Partial Differential Equations I

lecture by

Prof. Dr. Felix Finster

during the summer semester 2013

revision and layout in LyX by

Andreas Völklein

Last changed: April 26, 2013

ATTENTION

This script does *not* replace the lecture.

Therefore it is recommended *strongly* to attend the lecture.

Links

The text of the "GNU Free Documentation License" can also be read on the following site:

`https://www.gnu.org/licenses/fdl-1.3.en.html`

A transparent copy of the recent version of this document can be downloaded from:

`https://github.com/andiv/PDE1`

Literature

Elliptic and parabolic partial differential equations:

– Jürgen Jost: *Partial Differential Equations*; Springer, 2007
ISBN: 978-0-387-49318-3; doi: 10.1007/978-0-387-49319-0
(good book, but not all details, small errors)

– Lawrence C. Evans: *Partial Differential Equations*; American Mathematical Society, 2010; ISBN: 978-0-8218-4974-3
(part of the lecture follows this book, lots of details)

– David Gilborg, Neil S. Trudinger: *Elliptic Partial Differential Equations of second order*; Springer, 2001; ISBN: 3-540-41160-7
(classic textbook, complete treatment)

Hyperbolic partial differential equations (for the lecture "Partial Differential Equations II"):

– Fritz John: *Partial Differential Equations*; Springer, 1999
ISBN: 0-387-90609-6

– Michael E. Taylor: *Partial Differential Equations I - III*; Springer, 1997
ISBN: 0-387-94653-5, 0-387-94651-9, 0-387-94652-7
(nice detailed text books)

– Joel Smoller: *Shock waves and reaction-diffusion equations*; Springer, 1994
ISBN: 3-540-94259-9
(nicely presented, good motivations, covers most of the material)

– Friedrich Sauvigny: Partial Differential Equations I-II; Springer, 2012
ISBN: 978-1-4471-2981-3, 978-1-4471-2984-4;
doi: 10.1007/978-1-4471-2981-3, 10.1007/3-540-27540-1

– and many more …

Contents

1 A Brief Introduction

An ordinary differential equation (ODE) can be written as:

$$\frac{\mathrm{d}}{\mathrm{d}t} u(t) = \dot{u}(t) = v(t, u)$$
$$u : I \subseteq \mathbb{R} \to \mathbb{R}^N$$

This equation involves only derivatives with respect to *one* variable t.

$$\frac{\partial}{\partial x} f(x, y) + \frac{\partial}{\partial y} f(x, y) = 0$$

This is an example for a partial differential equation.

1.1 Definition (Partial Differential Equation)

A *partial differential equation* (PDE) is a (scalar) equation, which involves partial derivatives of an unknown function $u : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$. We always assume that $\Omega \subseteq \mathbb{R}^n$ is open.
More generally, a *system of partial differential equations* is a system of equations involving partial derivatives of a function $u : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^N$.
Similarly one can define partial differential equations on manifolds.
For ordinary differential equations we considered the initial-value problem:

$$\dot{u}(t) = v(t, u) \qquad\qquad u(t_0) = u_0$$

For partial differential equations one considers
- the initial-value problem and
- the boundary-value problem.

1.2 Examples

1. Cauchy-Riemann equations: Let

   $$f : \Omega \overset{\text{open}}{\subseteq} \mathbb{C} \to \mathbb{C}$$

   be holomorphic.

   $$f = a + \mathbf{i}b \qquad a := \mathrm{Re}(f) \qquad b := \mathrm{Im}(f)$$

   $$\frac{\partial a}{\partial x} = \frac{\partial b}{\partial y} \qquad\qquad \frac{\partial b}{\partial x} = -\frac{\partial a}{\partial y}$$

This is a system of two partial differential equations.

$$u := \begin{pmatrix} a \\ b \end{pmatrix} \qquad u : \Omega \subseteq \mathbb{C} = \mathbb{R}^2 \to \mathbb{R}^2$$

$$\frac{\partial^2 a}{\partial x^2} + \frac{\partial^2 a}{\partial y^2} = \frac{\partial b}{\partial x \partial y} - \frac{\partial b}{\partial y \partial x} = 0$$

$$\Rightarrow \qquad \underbrace{\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)}_{=:\Delta} a = 0 \qquad \Delta b = 0$$

This is the *Laplace equation* with the *Laplace operator* Δ. Solutions of the Laplace equation are called *harmonic functions.*

2. Let (M, g) be a Riemannian manifold. Here exists the Laplace-Beltrami operator $\Delta\backslash$.
   - In the special case $M = \mathbb{R}^n$ we have:

     $$\begin{aligned} \Delta &= \frac{\partial^2}{\partial x_1^2} + \ldots + \frac{\partial^2}{\partial x_n^2} \\ \Delta\varphi &= 0 \end{aligned}$$

   - With the Riemannian metric g_{ij} we can define:

     $$\Delta\backslash\varphi = g^{ij}\nabla_i\nabla_j\varphi = \operatorname{div}(\operatorname{grad}(\varphi)) = \frac{1}{\sqrt{\det(g)}}\partial_j\left(\sqrt{\det(g)}g^{jk}\partial_k\varphi\right)$$

   This gives an elliptic equation.
3. Newton's gravitational law: Let $\varrho(x)$ be the mass density and $\varphi(x)$ the Newtonian potential.

   $$\Delta\varphi = \underbrace{-4\pi\varrho}_{\text{inhomogeneity}}$$

   Such an inhomogeneous Laplace equation is usually referred to as Poisson equation and it is elliptic.
4. Heat flow equation (Wärmeleitungsgleichung): Let $\varphi(t, x)$ be the temperature at time $t \in \mathbb{R}$ and position $x \in \mathbb{R}^n$.

   $$\partial_t\varphi(t, x) = \Delta\varphi(t, x)$$

   This is a parabolic equation.
5. The Schrödinger equation is a parabolic equation:

   $$\mathbf{i}\hbar\partial_t\psi(t, x) = \left(-\frac{\hbar^2}{2m}\Delta + V\right)\psi(t, x)$$

   Additionally to the the heat flow equation there is the potential V, but more important there is a factor of $\mathbf{i}$ in front of the partial derivative. The time-independent Schrödinger equation is:

   $$E\psi(x) = \left(-\frac{\hbar^2}{2m}\Delta + V\right)\psi(x)$$

   This is similar to the Poisson equation and also elliptic.

6. The wave equation

$$\left(\partial_t^2 - \Delta_x\right) \psi\left(t, x\right) = 0$$

is hyperbolic. We will consider it in the lecture "Partial Differential Equations II".

7. Maxwell's equations: $E\left(t,x\right)$ is the electric field and $B\left(t,x\right)$ the magnetic field.

$$\begin{aligned} \operatorname{div}(E) &= 4\pi\varrho && \text{Gauss law} \\ \operatorname{rot}(E) &= -\partial_t B && \text{Maxwell} \\ \operatorname{div}(B) &= 0 && \\ \operatorname{rot}(B) &= 4\pi j - \partial_t E && \text{Faraday} \end{aligned}$$

This is a system of 8 partial differential equation.

8. Einstein's field equation:

$$R_{ij} - \frac{1}{2} R g_{ij} = 4\pi\kappa T_{ij}$$

This is a geometric partial differential equation. R_{ij} is the Ricci curvature, R the scalar curvature and T_{ij} the energy-momentum tensor. It is a system of 10 partial differential equations.

9. Equations of relativistic quantum mechanics:

$$\left(-\partial_t^2 + \Delta\right) \psi = m^2 \psi$$

This is the Klein-Gordon equation with the mass m.

$$\mathbf{i}\gamma^j \partial_j \psi = m\psi$$

This is the Dirac equation, a system of 4 complex-valued or 8 real-valued partial differential equations for a particle with spin $\frac{1}{2}$.

10. Water waves can be described by the Korteweg-de Vries equation:

$$\partial_t u + u \partial_x u + \partial_x^3 u = 0$$

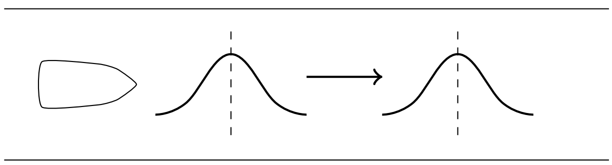

Figure 1.1: Solitons (discovered by John Russel in 1834): When the ship suddenly stops, the water flows on along the channel. This wave moves with a constant speed and its shape stays the same.

11. Shock waves: Burger's equation

$$\partial_t u + u \partial_x u = 0$$

is hyperbolic.

12. Turbulence can be described by the incompressible Navier-Stokes equations for the velocity $v : \Omega \subseteq \mathbb{R}^3 \to \mathbb{R}^3$.

$$\begin{aligned} \operatorname{div}(v) &= 0 \\ \rho \partial_t v^j + \rho v^i \partial_i v^j - \eta \Delta v^j &= -\partial_j P \end{aligned}$$

Here ϱ is the gas density, P the pressure and η the viscosity.

1.3 Classification

I) The *order* of a partial differential equation is the highest order of the derivatives in it.

$$\begin{aligned} \Delta u &= f && \text{second order} \\ \partial_t \varphi &= \Delta \varphi && \text{second order} \\ \partial_t u + u\partial_x u + \partial_x^3 u &= 0 && \text{third order} \end{aligned}$$

II) Algebraic classification:

a) *Linear* equations: The unknown function u and its derivatives appear only linearly.

$$\begin{aligned} \partial_t u &= u && \text{linear} \\ \partial_t u + u\partial_x u &= 0 && \text{non-linear} \end{aligned}$$

b) Linear *homogeneous* equations: If u is a solution, then λu for $\lambda \in \mathbb{R}$ is also a solution.

$$\begin{aligned} \Delta u &= 0 && \text{linear homogeneous} \\ \Delta u &= \varrho && \text{linear inhomogeneous} \\ \Delta\backslash u &= 0 && \text{linear homogeneous} \end{aligned}$$

c) Linear with *constant coefficients*:

$$\begin{aligned} \Delta u &= \rho && \text{linear with constant coefficients} \\ \Delta\backslash u &= \varrho && \text{in general non-constant coefficients} \end{aligned}$$

III) Classification by type: elliptic, parabolic, hyperbolic
Here we only consider scalar second order equations with $x \in \Omega \subseteq \mathbb{R}^n$.

$$\begin{gathered} F\left(x, u, \mathrm{D}u, \mathrm{D}^2 u\right) = 0 \\ F : \Omega \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^{n\times n} \to \mathbb{R} \end{gathered}$$

$$A_{ij} := \frac{\partial F\left(x, u, p_i, p_{ij}\right)}{\partial p_{ij}}$$

is a symmetric $n \times n$ matrix.

- If A is positive definite, the equation is called *elliptic*.
- If A has $n-1$ positive and one negative eigenvalue, the equation is called *hyperbolic*.
- If A has $n-1$ positive eigenvalues and a non-trivial kernel, the equation is called *parabolic*.
- If all eigenvalues are negative or $n-1$ are negative, then we replace F by $-F$.
  All other case of *mixed type* are difficult and we do not consider them in this lecture.

1.4 Examples

Consider the Poisson equation:

$$\Delta u = \varrho$$

$$
\begin{aligned}
F\left(x, u, \mathrm{D}u, \mathrm{D}^2 u\right) &= -\varrho(x) + \delta^{ij}\partial_{ij}u \\
F\left(x, u, p_i, p_{ij}\right) &= -\varrho(x) + \delta^{ij}p_{ij} \\
A_{ij} = \frac{\partial F\left(x, u, p_i, p_{ij}\right)}{\partial p_{ij}} &= \delta_{ij}
\end{aligned}
$$

So we have $A = \mathbb{1}$ and thus the equation is elliptic.
Now consider the inhomogeneous wave equation:

$$\left(\partial_t^2 - \Delta\right)\phi\left(t, x_1, x_2, x_3\right) = \varrho$$

$$F\left(x, u, p_i, p_{ij}\right) = \varrho(x) + \eta^{ij}p_{ij} \qquad \eta = \operatorname{diag}(-1, 1, 1, 1)$$

So $A = \eta$ has one negative and three positive eigenvalues which means that the equation is hyperbolic.

$$\partial_t \phi = \Delta\phi$$

$$
\begin{aligned}
F\left(x, u, \mathrm{D}\phi, \mathrm{D}^2\phi\right) &= -\partial_0\phi + \sum_{i,j=1}^{3}\delta^{ij}\partial_{ij}\phi \\
A_{ij} &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}
\end{aligned}
$$

Therefore the equation is parabolic.

2 Distributions and Fourier Transform

Motivation

We want to solve partial differential equations with constant coefficients in $\Omega = \mathbb{R}^n$, for example:

$$\left(-\partial_t^2 + \Delta\right)\phi = 0$$

Now we make a "plane wave ansatz" with $t, \omega \in \mathbb{R}$ and $k, x \in \mathbb{R}^{n-1}$:

$$\begin{aligned} \phi(t,x) &= e^{-\mathbf{i}\omega t + \mathbf{i}\langle k,x\rangle} \\ \partial_t \phi(t,x) &= -\mathbf{i}\omega\phi(t,x) \\ \partial_j \phi(t,x) &= \mathbf{i}k_j\phi(t,x) \end{aligned}$$

This gives an algebraic equation:

$$\begin{aligned} \left(-(-\mathbf{i}\omega)^2 + (\mathbf{i}k)^2\right)\phi &= 0 \\ \Leftrightarrow \qquad \omega^2 &= k^2 \end{aligned}$$

We also want to differentiate non-smooth functions, e.g.:

$$\Delta_{\mathbb{R}^3}\frac{1}{|x|} = -4\pi\delta(x)$$

$\delta(x)$ is called Dirac δ-distribution.

2.1 The Schwartz Space and Distributions

Laurent Schwartz was the first to investigate distributions systematically. He was awarded the fields medal for his research.

2.1.1 Definition (Multi-Index)

For $\mathbb{R}^n$ we denote indices by $i, j, k \in \{1, \ldots, n\}$. We call $\alpha = (i_1, \ldots, i_k)$ with $i_l \in \{1, \ldots, n\}$ a *multi-index*. $|\alpha| := k$ is called the *order* or *absolute value* of the multi-index.
With this we can write differentials of order k as

$$\mathrm{D}^\alpha := \frac{\partial}{\partial x^{i_1}} \cdots \frac{\partial}{\partial x^{i_k}} \tag{2.1}$$

and homogeneous polynomials of degree k in the components of a vector $x = \left(x^1, \ldots, x^n\right)$ as:

$$x^\alpha := x^{i_1} \cdots x^{i_k} \tag{2.2}$$

For $f \in C^\infty(\mathbb{R}^n)$ and $r, s \in \mathbb{N}$ we define the *Schwartz norm*:

$$\|f\|_{r,s} := \sum_{\alpha, |\alpha| \le r} \sum_{\beta, |\beta| \le s} \sup_{x \in \mathbb{R}^n} \left| x^\alpha \mathrm{D}^\beta f(x) \right| \tag{2.3}$$

For example for $r = 0 = s$ we have:

$$\|f\|_{0,0} = \sup_{x \in \mathbb{R}^n} |f(x)| = \|f\|_{C^0}$$

2.1.2 Definition (Schwartz Space)

The *Schwartz space* $\mathcal{S}(\mathbb{R}^n)$ is the vector space of all $f \in C^\infty(\mathbb{R}^n)$ for which all Schwartz norms are finite, i.e. for all $r, s \in \mathbb{N}$ holds:

$$\|f\|_{r,s} < \infty$$

This space is an infinite-dimensional vector space.
On a normed space $(E, \|.\|)$, the topology is given by the open sets.
$\Omega \subseteq E$ is defined as *open* if holds:

$$\underset{x \in \Omega}{\forall} \; \underset{\varepsilon > 0}{\exists} \; : \; B_\varepsilon(x) \subseteq \Omega$$

A subset $\Omega \subseteq \mathcal{S}(\mathbb{R}^n)$ is called *open* if for every $f \in \Omega$ there is a $\varepsilon > 0$ and $r, s \in \mathbb{N}$ such that holds:

$$\left\{ g \in \mathcal{S} \middle| \, \|g - f\|_{r,s} < \varepsilon \right\} \subseteq \Omega \tag{2.4}$$

Note: This topology is fine, because it involves many open sets, since the condition for open only involves the statement “there are $r, s \in \mathbb{N}$”.
Convergence $f_n \to f$ in $\mathcal{S}$ means that *every* open neighborhood U of f contains almost all f_n. For a finer topology, the condition for a sequence to converge is stronger.

2.1.3 Theorem (Criterion for Convergence)

Convergence $f_n \to f$ in $\mathcal{S}$ is equivalent to the convergence $\|f_n - f\|_{r,s} \to 0$ for all $r, s \in \mathbb{N}$.

Proof

“$\Rightarrow$”: Suppose that $f_n \to f$ converges. By definition of the convergence, every open neighborhood of f contains almost all f_n. For all $r, s \in \mathbb{N}$ the sets $U_\varepsilon^{r,s} := \left\{ g \middle| \, \|g - f\|_{r,s} < \varepsilon \right\}$ are open by definition. So the inequality

$$\|f_n - f\|_{r,s} < \varepsilon$$

holds for almost all f_n and thus converges $\|f_n - f\|_{r,s} \to 0$.
“$\Leftarrow$”: Assume that $\|f_n - f\|_{r,s} \to 0$ converges for all $r, s \in \mathbb{N}$. Let A be an open neighborhood of f. This means by definition that there exist $r, s \in \mathbb{N}$ and $\varepsilon \in \mathbb{R}_{>0}$ with $U_\varepsilon^{r,s} \subseteq A$. For this (r, s) we know that $\|f_n - f\|_{r,s} \to 0$ converges. Hence there exists a $N \in \mathbb{N}$ such that $\|f_n - f\|_{r,s} < \varepsilon$ holds for all $n \in \mathbb{N}_{>N}$, in other words $f_n \in U_\varepsilon^{r,s} \subseteq A$. So $f_n \to f$ converges in $\mathcal{S}$. $\square_{2.1.3}$

A vector space with a topology generated by a family of norms or semi-norms is a *uniform space* and is called *topological vector space.*

2.1.4 Definition (Tempered Distribution)

Let $\mathcal{S}^*(\mathbb{R}^n)$ be the dual space of $\mathcal{S}(\mathbb{R}^n)$. It is called the space of *tempered distributions* (temperierte Distributionen).

In linear algebra for a finite-dimensional vector space V, the dual space $V^* = L(V, \mathbb{R})$ is the space of linear functionals. V^* is again a vector space with $\dim(V^*) = \dim(V)$.
Here $\mathcal{S}(\mathbb{R}^n)$ is an infinite-dimensional vector space with a topology. $\mathcal{S}^*(\mathbb{R}^n)$ is the space of all *continuous* linear functionals.

In a Banach space $(E, \|.\|)$ holds: A linear functional $A : E \to \mathbb{R}$ is continuous if and only if A is bounded, i.e. $|Au| \le c\,\|U\|$ for all $u \in E$.

2.1.5 Lemma (Criterion for Continuity)

A linear functional $T : \mathcal{S}(\mathbb{R}^n) \to \mathbb{R}$ is continuous if and only if there are $r, s \in \mathbb{N}$ and a $c \in \mathbb{R}_{>0}$ such that holds:

$$|Tf| \le c\,\|f\|_{r,s} \tag{2.5}$$

Proof

"$\Leftarrow$": Assume that (2.5) holds for some $r, s \in \mathbb{N}$. We want to show that T is continuous. To this end, let $f_n \to f$ be a convergent series in $\mathcal{S}$. Our task is to show that $Tf_n \to Tf$ converges.
The convergence $f_n \to f$ implies $\|f_n - f\|_{r',s'} \to 0$ for all $r', s' \in \mathbb{N}$ and thus in particular for r, s satisfying the inequality (2.5). By linearity follows:

$$|Tf_n - Tf| = |T(f_n - f)| \le c\,\|f_n - f\|_{r,s} \xrightarrow{n\to\infty} 0$$

So T maps convergent sequences to convergent sequences and is thus continuous.

"$\Rightarrow$": Assume that T is continuous. Then the preimage of open sets is open, in particular $T^{-1}(B_1(0)) \subseteq \mathcal{S}(\mathbb{R}^n)$ is open. So there exist $r, s \in \mathbb{N}$ and a $\varepsilon \in \mathbb{R}_{>0}$ such that holds:

$$T^{-1}(B_1(0)) \supseteq U_\varepsilon^{r,s} := \left\{ g \middle| \|g\|_{r,s} < \varepsilon \right\}$$

This implies:

$$\|g\|_{r,s} < \varepsilon \quad \Rightarrow \quad g \in T^{-1}(B_1(0))$$

Now $g \in T^{-1}(B_1(0))$ means $|Tg| < 1$. For any $f \in \mathcal{S}(\mathbb{R}^n)$ apply this to $g = \frac{f}{\lambda}$ with $\lambda \in \mathbb{R}_{>0}$.

$$\begin{aligned} \frac{1}{\lambda}\|f\|_{r,s} < \varepsilon \quad &\Rightarrow \quad \frac{1}{\lambda}|Tf| < 1 \qquad / \cdot \lambda \\ \|f\|_{r,s} < \lambda\varepsilon \quad &\Rightarrow \quad |Tf| < \lambda \end{aligned}$$

Now choose $\lambda = \frac{2}{\varepsilon}\|f\|_{r,s}$, so the left side holds, which implies:

$$|Tf| < \frac{2}{\varepsilon}\|f\|_{r,s} \qquad \underset{f\in\mathcal{S}(\mathbb{R}^n)}{\forall}$$

$\square_{2.1.5}$

2.1.6 Example (δ-Distribution)

a) Consider the following functional:

$$\begin{aligned} \delta : \mathcal{S}(\mathbb{R}) &\to \mathbb{R} \\ f &\mapsto f(0) \end{aligned}$$

This is obviously linear.

$$|\delta(f)| = |f(0)| \leq \sup_{\mathbb{R}} |f| = \|f\|_{0,0}$$

Hence δ is continuous, which means that $\delta \in \mathcal{S}^*(\mathbb{R})$ is a tempered distribution.
A convenient *notation* with $f \in \mathcal{S}(\mathbb{R})$ is:

$$\delta(f) = \int_{\mathbb{R}} f(x)\,\delta(x)\,\mathrm{d}x$$

b) In higher dimension $n \in \mathbb{N}$ we define:

$$\begin{aligned} \delta : \mathcal{S}(\mathbb{R}^n) &\to \mathbb{R} \\ f &\mapsto f(0) \end{aligned}$$

Again holds $|\delta(f)| \leq \|f\|_{0,0}$. The physicists' notation for this is:

$$\begin{aligned} \delta(f) &= \int_{\mathbb{R}^n} f(x)\,\delta^{(n)}(x)\,\mathrm{d}^n x \\ \delta^{(n)}(x) &= \delta\left(x^1\right)\cdots\delta\left(x^n\right) \end{aligned}$$

Remark

δ can also be introduced as a *measure* on $\mathbb{R}^n$, the *Dirac measure.* For $A \subseteq \mathbb{R}^n$ define:

$$\delta(x) = \begin{cases} 1 & \text{if } 0 \in A \\ 0 & \text{otherwise} \end{cases}$$

Then for $f \in C^0(\mathbb{R}^n)$ the expression

$$\int_{\mathbb{R}^n} f(x)\,\mathrm{d}\delta(x) = f(0)$$

makes mathematical sense as an integral.
This is useful because convergence theorems and so on from measure theory are available. The problem is, that this does not work for every distribution and thus is not general enough for most purposes, e.g. the derivative $\delta'(x)$ is a distribution, but cannot be written as a measure.

2.1.7 Example (Integral Operator)

a) Consider $g \in C^\infty(\mathbb{R}^n)$ with at most polynomial growth, i.e. there are $c \in \mathbb{R}_{>0}$ and $r \in \mathbb{N}$ such that holds:

$$|g(x)| \leq c\,(1 + |x|^r)$$

Now define:

$$T_g : \mathcal{S}(\mathbb{R}^n) \to \mathbb{R}$$
$$f \mapsto \int_{\mathbb{R}^n} g(x) f(x) \,\mathrm{d}^n x$$

The integral here is just the Lebesgue integral and it exists:
For $f \in \mathcal{S}(\mathbb{R}^n)$ holds $\|f\|_{r,s} < \infty$ for all $r, s \in \mathbb{N}$.

$$\sup_{\mathbb{R}} |f| + \sup_{x\in\mathbb{R}} \left(|x|^{\tilde{r}} \cdot |f(x)|\right) \leq \|f\|_{\tilde{r},0} < \infty$$
$$\Rightarrow \quad \sup_{x\in\mathbb{R}} \left(\left(1 + |x|^{\tilde{r}}\right) |f(x)|\right) \leq \|f\|_{\tilde{r},0}$$
$$\Rightarrow \quad |f(x)| \leq \frac{\|f\|_{\tilde{r},0}}{1 + |x|^{\tilde{r}}} \qquad \forall_{\tilde{r}\in\mathbb{N}}$$

So we get:

$$T_g f = \int g(x) f(x) \,\mathrm{d}^n x \leq \int c\,(1 + |x|^r) \frac{\|f\|_{\tilde{r},0}}{\left(1 + |x|^{\tilde{r}}\right)} \mathrm{d}^n x =$$
$$\overset{\text{polar coordinates}}{\underset{\rho:=|x|}{=}} c\,\|f\|_{\tilde{r},0} \underbrace{\mu\left(S^{n-1}\right)}_{\text{volume of unit sphere}} \int_0^\infty \rho^{n-1} \frac{1 + \rho^r}{1 + \rho^{\tilde{r}}} \mathrm{d}\rho \overset{\tilde{r}>r+n}{<} \infty$$

This is finite if and only if the integrand decays faster than ρ^{-1}, i.e. $n - 1 + r - \tilde{r} < -1$ and thus $\tilde{r} > n + r$. Since $\tilde{r} \in \mathbb{N}$ is arbitrary, the integral exists.
Continuity: The previous estimate implies with $\tilde{r} = n + r + 1$:

$$|T_g f| \leq C(g, n) \|f\|_{\tilde{r},0}$$

Thus $T_g \in \mathcal{S}^*(\mathbb{R}^n)$ is a tempered distribution.

b) Chose $g(x) = e^x$ and define:

$$T_g f := \int_{-\infty}^{\infty} f(x)\, e^x \mathrm{d}x$$

This is *not* a well-defined tempered distribution. Namely, choose:

$$f(x) = \frac{1}{\cosh\left(\frac{x}{2}\right)}$$

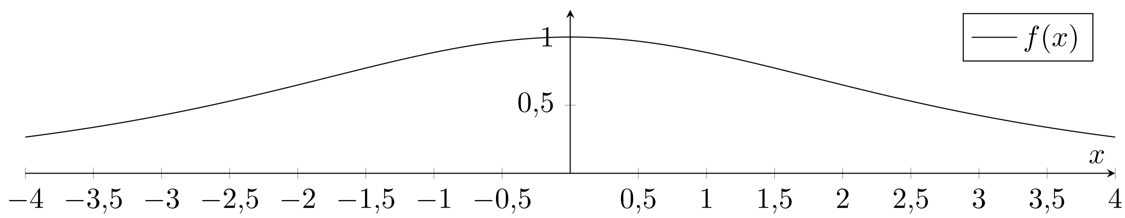

Figure 2.1: $f(x)$ decays rapidly.

$f(x)$ and all its derivatives decay rapidly (exponentially fast $\sim e^{-\frac{x}{2}}$) at $\pm\infty$, so $f \in \mathcal{S}$ is a Schwartz function. But $T_g f$ diverges:

$$T_g f = \int_{-\infty}^{\infty} \frac{e^x}{\cosh\left(\frac{x}{2}\right)} \mathrm{d}x = +\infty$$

2.1.8 Remark (Schwartz Functions as Distributions)

The mapping

$$\begin{aligned} T : \mathcal{S}(\mathbb{R}^n) &\to \mathcal{S}^*(\mathbb{R}^n) \\ f &\mapsto T_f \end{aligned}$$

is injective.

Proof

If T was injective, there were $f_1, f_2 \in \mathcal{S}(\mathbb{R}^n)$ with $T_{f_1} = T_{f_2}$ and $f_1 \neq f_2$. By linearity this would imply $T_g = 0$ with $g = f_1 - f_2 \neq 0$ and we could choose a $y \in \mathbb{R}^n$ with $g(y) \neq 0$. By continuity follows $g > 0$ or $g < 0$ in a neighborhood U of y. Now choose a test function h with $\mathrm{supp}(h) \subseteq U$ and $h \geq 0$. Then follows the contradiction:

$$0 = T_g(h) = \int_{\mathbb{R}^n} g(x)\, h(x)\, \mathrm{d}^n x \neq 0$$

$\square_{2.1.8}$

Thus we can regard distributions as "generalized functions". Namely we identify a function g with T_g. (Later on many people often do not distinguish between g and T_g.)

Operations on Schwartz Functions and Distributions

- $\mathcal{S}^*(\mathbb{R}^n)$ is a vector space with addition $T+S$ and scalar multiplication $\alpha \cdot f$ for distributions $T, S \in \mathcal{S}^*(\mathbb{R}^n)$ and $\alpha \in \mathbb{R}$.
- Multiplication of a distribution by a Schwartz function is defined for $T \in \mathcal{S}^*(\mathbb{R}^n)$ and $f, g \in \mathcal{S}(\mathbb{R}^n)$ as:

  $$(fT)(g) := T(f \cdot g)$$

  This is well defined, because $f \cdot g$ is again a Schwartz function and for $h \in \mathcal{S}(\mathbb{R}^n)$ holds:

  $$\begin{aligned} (fT_h)(g) &\overset{\text{definition of } fT_h}{=} T_h(f \cdot g) \overset{\text{definition of } T_h}{=} \int_{\mathbb{R}^n} h(x)\, (f \cdot g)(x)\, \mathrm{d}^n x = \\ &= \int_{\mathbb{R}^n} (f \cdot h)(x)\, g(x)\, \mathrm{d}^n x = T_{fh}(g) \end{aligned}$$

  So this definition is extends the multiplication of Schwartz functions to distributions. But we still have to show, that this operation gives a continuous functional.

2.1.9 Definition (regular/singular distribution)

A tempered distribution T is called *regular* distribution if there is a $g \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ (locally integrable, i.e. integrable on every compact interval) with $T = T_g$. Otherwise, T is called *singular*.
For example $\delta(x)$ is singular.

2.1.10 Lemma (Multiplication of a Distribution by a Schwartz Function)

Let $f \in \mathcal{S}(\mathbb{R}^n)$ be a Schwartz function and $T \in \mathcal{S}^*(\mathbb{R}^n)$ a distribution. Then fT is a *continuous* linear functional on $\mathcal{S}(\mathbb{R}^n)$, in other words $fT \in \mathcal{S}^*(\mathbb{R}^n)$ is also a distribution.

Proof

According to Lemma 2.1.5, our task is to show that there are $r, s \in \mathbb{N}$ and a $C \in \mathbb{R}_{>0}$ with:

$$|(fT)(g)| \leq C \|g\|_{r,s}$$

Since T is continuous, there exist $r, s \in \mathbb{N}$ and a $\tilde{C} \in \mathbb{R}_{>0}$ with:

$$|T(fg)| \leq \tilde{C} \|fg\|_{r,s}$$

Thus it remains to show that there is a $c(r,s) \in \mathbb{R}_{>0}$ such that for all $f, g \in \mathcal{S}(\mathbb{R}^n)$ holds:

$$\|fg\|_{r,s} \leq c(r,s) \|f\|_{r,s} \cdot \|g\|_{r,s}$$

This inequality can be proven by induction in s.
Induction basis $s = 0$:

$$\begin{aligned}\|fg\|_{r,0} &= \sum_{|\alpha| \leq r} \sup_{x \in \mathbb{R}^n} \left(|x^\alpha f(x) g(x)|\right) \leq \\ &\leq \sup_{y \in \mathbb{R}^n} |g(y)| \sum_{|\alpha| \leq r} \sup_{x \in \mathbb{R}^n} \left(|x^\alpha f(x)|\right) = \\ &= \|g\|_{0,0} \cdot \|f\|_{r,0} \leq \|g\|_{r,0} \cdot \|f\|_{r,0}\end{aligned}$$

Induction step $s \rightsquigarrow s+1$: Assume that the statement holds for a $s \in \mathbb{N}$ for all $r \in \mathbb{N}$. Let β be a multi-index with $|\beta| = s+1$, i.e. $\beta = (i_1, \ldots, i_{s+1})$. Now set:

$$\hat{\beta} := (i_1, \ldots, i_s) \qquad j := i_{s+1} \qquad \mathrm{D}^\beta = \mathrm{D}^{\hat{\beta}} \frac{\partial}{\partial x^j}$$

It holds:

$$\mathrm{D}^\beta(fg) = \mathrm{D}^{\hat{\beta}} \frac{\partial}{\partial x^j}(fg) = \mathrm{D}^{\hat{\beta}}\left(\left(\frac{\partial}{\partial x^j} f\right) g + f\left(\frac{\partial}{\partial x^j} g\right)\right)$$

$$\begin{aligned}\|fg\|_{r,s+1} &= \|f \cdot g\|_{r,s} + \sum_{\substack{|\alpha| \leq r \\ |\beta| = s+1}} \sup_{x \in \mathbb{R}} \left|x^\alpha \mathrm{D}^\beta (f \cdot g)(x)\right| = \\ &= \|fg\|_{r,s} + \sum_{\substack{|\alpha| \leq r \\ |\hat{\beta}| = s}} \sum_{j=1}^n \sup_{x \in \mathbb{R}} \left|x^\alpha \mathrm{D}^{\hat{\beta}}\left(\left(\frac{\partial}{\partial x^j} f(x)\right) g(x) + f(x)\left(\frac{\partial}{\partial x^j} g(x)\right)\right)\right| \leq \\ &\underset{\text{hypothesis}}{\overset{\text{induction}}{\leq}} c(r,s)\|f\|_{r,s}\|g\|_{r,s} + \sum_{j=1}^n c(r,s)\left(\left\|\frac{\partial}{\partial x^j} f\right\|_{r,s} \|g\|_{r,s} + \|f\|_{r,s}\left\|\frac{\partial}{\partial x^j} g\right\|_{r,s}\right) \leq \\ &\leq c(r,s)\|f\|_{r,s}\|g\|_{r,s} + n \cdot c(r,s)\left(\|f\|_{r,s+1}\|g\|_{r,s} + \|f\|_{r,s}\|g\|_{r,s+1}\right) \leq \\ &\leq \underbrace{(2n+1)\, c(r,s)}_{=c(r,s+1)} \|f\|_{r,s+1}\|g\|_{r,s+1}\end{aligned}$$

$\square_{2.1.10}$

2.1.11 Example (Derivative of the δ-Distribution)

We make a formal computation:

$$\int_{\mathbb{R}} \delta'(x) f(x)\,\mathrm{d}x = \int_{\mathbb{R}} \left(\frac{\mathrm{d}}{\mathrm{d}x}\delta(x)\right) f(x)\,\mathrm{d}x \overset{\text{integration}}{\underset{\text{by parts}}{=}} -\int_{\mathbb{R}} \delta(x) f'(x)\,\mathrm{d}x = -f'(0)$$

This motivates us to *define*:

$$\begin{aligned}\delta' : \mathcal{S}(\mathbb{R}) &\to \mathbb{R}\\ f &\mapsto -f'(0)\end{aligned}$$

This is obviously linear and it is continuous, because for all $f \in \mathcal{S}(\mathbb{R})$ holds:

$$\left|\delta'(f)\right| = \left|f'(0)\right| \leq \|f\|_{0,1}$$

Hence we have $\delta' \in \mathcal{S}^*(\mathbb{R})$.

Remark

δ' cannot be introduced as a measure, in contrast to δ:

$$\begin{aligned}\delta(\Omega) &= \begin{cases}1 & \text{if } 0 \in \Omega\\ 0 & \text{otherwise}\end{cases}\\ \delta'(\Omega) &=?\end{aligned}$$

2.1.12 Definition (Distributional Derivative, Convolution)

- For a tempered distribution $T \in \mathcal{S}^*(\mathbb{R}^n)$ we define the *distributional derivative* $\mathrm{D}^\alpha T$ by:

  $$(\mathrm{D}^\alpha T)(f) := (-1)^{|\alpha|}\, T(\mathrm{D}^\alpha f)$$

  $\mathrm{D}^\alpha T$ is a distribution, since it is a continuous functional:

  $$|(\mathrm{D}^\alpha T)(f)| = |T(\mathrm{D}^\alpha f)| \overset{T \text{ continuous}}{\leq} C\,\|\mathrm{D}^\alpha f\|_{r,s} \leq C\,\|f\|_{r,s+|\alpha|}$$

  So we have a mapping $\mathrm{D}^\alpha : \mathcal{S}^*(\mathbb{R}^n) \to \mathcal{S}^*(\mathbb{R}^n)$.
- The convolution (Faltung) for $f, g \in \mathcal{S}(\mathbb{R}^n)$ is defined as:

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

2.1.13 Lemma (Commutativity and Associativity of the Convolution)

The convolution is commutative and associative, i.e. for $f, g, h \in \mathcal{S}(\mathbb{R}^n)$ holds:

$$f * g = g * f \qquad\qquad f * (g * h) = (f * g) * h$$

Proof

$$(f*g)(x) = \int_{\mathbb{R}^n} f(x-y)\, g(y)\, \mathrm{d}^n y \underset{\mathrm{d}^n z = \mathrm{d}^n y}{\overset{z:=x-y}{=}} \int_{\mathbb{R}^n} f(z)\, g(x-z)\, \mathrm{d}^n z = (g*f)(x)$$

Associativity follows analogously using Fubini's theorem:

$$\begin{aligned}
f*(g*h)(x) &= \int_{\mathbb{R}^n} f(x-y)\,(g*h)(y)\, \mathrm{d}^n y = \\
&= \int_{\mathbb{R}^n} f(x-y) \left(\int_{\mathbb{R}^n} g(y-z)\, h(z)\, \mathrm{d}^n z \right) \mathrm{d}^n y = \\
&\underset{\mathrm{d}^n \tilde{y} = \mathrm{d} y}{\overset{\tilde{y}:=y-z}{=}} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x-\tilde{y}-z)\, g(\tilde{y})\, h(z)\, \mathrm{d}^n \tilde{y} \mathrm{d}^n z = \\
&= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} f(x-z-\tilde{y})\, g(\tilde{y})\, \mathrm{d}^n \tilde{y} \right) h(z)\, \mathrm{d}^n z = \\
&= \int_{\mathbb{R}^n} (f*g)(x-z)\, h(z)\, \mathrm{d}^n z = \\
&= (f*g)*h(x)
\end{aligned}$$

$\square_{2.1.13}$

2.1.14 Proposition (Convolution is Continuous)

The convolution $*: \mathcal{S} \times \mathcal{S} \to \mathcal{S}$ is continuous.

Proof

$$\|f*g\|_{r,s} = \sum_{\substack{|\alpha| \le r \\ |\beta| \le s}} \sup_{x \in \mathbb{R}^n} \left| x^\alpha \mathrm{D}_x^\beta \int_{\mathbb{R}^n} f(x-y)\, g(y)\, \mathrm{d}^n y \right|$$

The derivative may be commuted with the integral:

$$\begin{aligned}
\frac{\partial}{\partial x_j} (f*g)(x) &= \lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon} \left((f*g)(x+\varepsilon e_j) - (f*g)(x) \right) = \\
&= \lim_{\varepsilon \searrow 0} \int_{\mathbb{R}^n} \frac{f(x+\varepsilon e_j - y) - f(x-y)}{\varepsilon} g(y)\, \mathrm{d}^n y
\end{aligned}$$

Using the estimate

$$\begin{aligned}
\frac{f(x+\varepsilon e_j - y) - f(x-y)}{\varepsilon} &= \frac{1}{\varepsilon} \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\tau} \left(f(x+\varepsilon\tau e_j - y) \right) \mathrm{d}\tau = \\
&= \frac{1}{\varepsilon} \int_0^1 (\partial_j f)(x+\varepsilon\tau e_j - y)\, \varepsilon \mathrm{d}\tau \\
\Rightarrow \qquad \left| \frac{f(x+\varepsilon e_j - y) - f(x-y)}{\varepsilon} \right| &\le \sup_{\mathbb{R}^n} |\partial_j f| \le \|f\|_{0,1}
\end{aligned}$$

we get:

$$\left| \frac{f(x+\varepsilon e_j - y) - f(x-y)}{\varepsilon} g(y) \right| \le \|f\|_{0,1} \cdot |g(y)| \in L^1(\mathbb{R}^n)$$

Thus the dominated convergence theorem implies:

$$\frac{\partial}{\partial x_j}(f*g)(x)=\int_{\mathbb{R}^n}\frac{\partial}{\partial x^j}f(x-y)\,g(y)\,\mathrm{d}^n y$$

By induction follows:

$$x^\alpha \mathrm{D}^\beta \int_{\mathbb{R}} f(x-y)\,g(y)\,\mathrm{d}^n y = x^\alpha \int_{\mathbb{R}^n}\left(\mathrm{D}^\beta f\right)(x-y)\,g(y)\,\mathrm{d}^n y$$

Now we treat the x^α:

$$x^\alpha = ((x-y)+y)^\alpha = \sum_{\gamma,\delta \text{ with } \alpha=\gamma+\delta} c_{\gamma\delta}\,(x-y)^\gamma \cdot y^\delta$$

Here holds $|\gamma|,|\delta|\leq|\alpha|$. Now we can estimate:

$$\left|(x-y)^\gamma\,\mathrm{D}^\beta f(x)\right| \leq \|f\|_{r,s}$$

Hence we get:

$$\begin{aligned}\|f*g\|_{r,s} &\leq c(r,s)\,\|f\|_{r,s}\sum_\delta\int_{\mathbb{R}^n}\left|y^\delta g(y)\right|\cdot\frac{(1+|y|)^{n+1}}{(1+|y|)^{n+1}}\mathrm{d}^n y \leq \\ &\leq c(r,s)\,\|f\|_{r,s}\,\|g\|_{n+1+r,0}\underbrace{\int_{\mathbb{R}^n}\frac{\mathrm{d}^n y}{(1+|y|)^{n+1}}}_{<\infty}\end{aligned}$$

$\square_{2.1.14}$

How can we define the convolution of $T\in S^*(\mathbb{R}^n)$ with $f\in S(\mathbb{R}^n)$? $f*T_g$ should be equal to T_{f*g}. For $h\in\mathcal{S}(\mathbb{R}^n)$ holds:

$$\begin{aligned}T_{f*g}(h) &= \int_{\mathbb{R}^n}(f*g)(x)\cdot h(x)\,\mathrm{d}^n x = \\ &= \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} f(x-y)\,g(y)\,\mathrm{d}^n y\right)\cdot h(x)\,\mathrm{d}^n x\end{aligned}$$

By Fubini we may interchange the order of integration to get:

$$\begin{aligned}T_{f*g}(h) &= \int_{\mathbb{R}^n} g(y)\left(\int_{\mathbb{R}^n} f(x-y)\cdot h(x)\,\mathrm{d}^n x\right)\mathrm{d}^n y = \\ &\overset{\tilde f(z):=f(-z)}{=} \int_{\mathbb{R}^n} g(y)\left(\int_{\mathbb{R}^n}\tilde f(y-x)\cdot h(x)\,\mathrm{d}^n x\right)\mathrm{d}^n y = \\ &= \int_{\mathbb{R}^n} g(y)\left(\tilde f*h\right)(y)\,\mathrm{d}^n y = T_g\left(\tilde f*h\right)\end{aligned}$$

So for a distribution $T\in S^*(\mathbb{R}^n)$ we define the *convolution* as:

$$\begin{aligned}*:\mathcal{S}(\mathbb{R}^n)\times\mathcal{S}^*(\mathbb{R}^n)&\to\mathcal{S}^*(\mathbb{R}^n)\\ (f*T)(h) &:= T\left(\tilde f*h\right)\end{aligned}$$

For $S,T\in\mathcal{S}^*(\mathbb{R}^n)$, $S*T$ and $S\cdot T$ are ill-defined in general. For example $\delta(x)\cdot\delta(x)$ makes no sense, as well as T_f*T_f for $f=1$.

2.2 The Fourier Transform

First consider the Fourier transform on $\mathcal{S}(\mathbb{R}^n)$ and later on $\mathcal{S}^*(\mathbb{R}^n)$.

2.2.1 Definition (Fourier Transform)

Define linear functionals $\mathcal{F}$ and $\overline{\mathcal{F}}$ on $\mathcal{S}$:

$$\begin{aligned}(\mathcal{F}f)(p) &:= \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-\mathbf{i}px} f(x)\,\mathrm{d}^n x\\ \left(\overline{\mathcal{F}}f\right)(x) &:= \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{\mathbf{i}px} f(p)\,\mathrm{d}^n p\end{aligned}$$

The integrals are well-defined and finite, because f has suitable decay properties at infinity. An alternative convention, which is not convenient here, because it has less symmetry, is:

$$\begin{aligned}(\mathcal{F}f)(p) &:= \int_{\mathbb{R}^n} e^{-\mathbf{i}px} f(x)\,\mathrm{d}^n x\\ \left(\overline{\mathcal{F}}f\right)(x) &:= \frac{1}{(2\pi)^{n}}\int_{\mathbb{R}^n} e^{\mathbf{i}px} f(p)\,\mathrm{d}^n p\end{aligned}$$

2.2.2 Proposition (Fourier Transform)

$\mathcal{F}$ and $\overline{\mathcal{F}}$ are well-defined linear operators from $\mathcal{S}(\mathbb{R}^n)$ to $\mathcal{S}(\mathbb{R}^n)$.

Proof

The linearity is clear. We still have to show, that all norms $\|\mathcal{F}f\|_{r,s}$ are finite. First consider the norm $\|.\|_{0,0}$:

$$\begin{aligned}|(\mathcal{F}f)(p)| &\leq \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} |f(x)|\cdot\frac{(1+|x|)^{n+1}}{(1+|x|)^{n+1}}\mathrm{d}^n x \leq\\ &\leq \frac{1}{(2\pi)^{\frac{n}{2}}}\|f\|_{n+1,0}\underbrace{\int_{\mathbb{R}^n}\frac{\mathrm{d}^n x}{(1+|x|)^{n+1}}}_{<\infty} \leq c\,\|f\|_{n+1,0}\end{aligned}$$

Now we consider $\left|p^\alpha \mathrm{D}^\beta(\mathcal{F}f)(p)\right|$.

$$\begin{aligned}\frac{\partial}{\partial p^j}(\mathcal{F}f)(p) &= \frac{\partial}{\partial p^j}\int e^{-\mathbf{i}px} f(x)\,\mathrm{d}^n x = \ldots = \int\left(\frac{\partial}{\partial p^j}e^{-\mathbf{i}px}\right) f(x)\,\mathrm{d}^n x =\\ &= \int\left(-\mathbf{i}x^j\right)e^{-\mathbf{i}px} f(x)\,\mathrm{d}^n x\end{aligned}$$

That the derivative and the integral can be interchanged is shown as follows:

$$\frac{\partial}{\partial p^j}\int e^{-\mathbf{i}px} f(x)\,\mathrm{d}^n x = \lim_{\varepsilon\searrow 0}\int\frac{e^{-\mathbf{i}(p+\varepsilon e_j)x} - e^{-\mathbf{i}px}}{\varepsilon} f(x)\,\mathrm{d}^n x$$

$$\frac{e^{-\mathbf{i}(p+\varepsilon e_j)x} - e^{-\mathbf{i}px}}{\varepsilon} = \frac{1}{\varepsilon}\int_0^1\frac{\mathrm{d}}{\mathrm{d}\tau}e^{-\mathbf{i}(p+\varepsilon\tau e_j)x}\mathrm{d}\tau = -\mathbf{i}e_j x\int_0^1 e^{-\mathbf{i}(p+\varepsilon\tau e_j)x}\mathrm{d}\tau$$

$$\Rightarrow \qquad \left|\frac{e^{-\mathbf{i}(p+\varepsilon e_j)x} - e^{-\mathbf{i}px}}{\varepsilon}\right| \leq \|x\| \cdot \int_0^1 \underbrace{\left|e^{-\mathbf{i}(p+\varepsilon\tau e_j)x}\right|}_{=1} \mathrm{d}\tau = \|x\|$$

$$\left|\frac{e^{-\mathbf{i}(p+\varepsilon e_j)x} - e^{-\mathbf{i}px}}{\varepsilon} f(x)\right| \leq \|x\| \cdot |f(x)| \leq \frac{\|x\|}{1+\|x\|^{n+2}} \|f\|_{n+2,0} =: h(x) \in L^1(\mathbb{R}^n)$$

This allows us to apply the dominated convergence theorem to take the limit $\varepsilon \to 0$ inside the integral. Iteration of this process gives:

$$\begin{aligned}
\mathrm{D}^\beta (\mathcal{F}f)(p) &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}} (-\mathbf{i})^{|\beta|} x^\beta f(x) e^{-\mathbf{i}px} \mathrm{d}^n x \\
p^\alpha \mathrm{D}^\beta (\mathcal{F}f)(p) &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}} (-\mathbf{i})^{|\beta|} x^\beta f(x) p^\alpha e^{-\mathbf{i}px} \mathrm{d}^n x \\
p^j e^{-\mathbf{i}px} &= \mathbf{i} \frac{\partial}{\partial x^j} e^{-\mathbf{i}px} \\
\Rightarrow \qquad p^\alpha e^{-\mathbf{i}px} &= \mathbf{i}^{|\alpha|} \mathrm{D}_x^\alpha e^{-\mathbf{i}px} \\
p^\alpha \mathrm{D}^\beta (\mathcal{F}f)(p) &= \frac{(-\mathbf{i})^{|\beta|} \mathbf{i}^{|\alpha|}}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \underbrace{\left(x^\beta f(x)\right)}_{\text{rapid decay}} \left(\mathrm{D}_x^\alpha e^{-\mathbf{i}px}\right) \mathrm{d}^n x = \\
&\overset{\text{integration}}{\underset{\text{by parts}}{=}} \frac{(-\mathbf{i})^{|\beta|} \mathbf{i}^{|\alpha|}}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} (-1)^{|\alpha|} \mathrm{D}_x^\alpha \left(x^\beta f(x)\right) e^{-\mathbf{i}px} \mathrm{d}^n x = \\
&= \frac{(-\mathbf{i})^{|\alpha|+|\beta|}}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \mathrm{D}_x^\alpha \left(x^\beta f(x)\right) e^{-\mathbf{i}px} \mathrm{d}^n x
\end{aligned}$$

From the computation we did earlier in the proof we know:

$$|\mathrm{D}_x^\alpha (\mathcal{F}f)(p)| \leq C \left\|\mathrm{D}_x^\alpha \left(x^\beta f\right)\right\|_{n+1,0} \leq \tilde{C}(\alpha,\beta) \|f\|_{|\beta|+n+1,|\alpha|}$$

$$\|\mathcal{F}f\|_{r,s} \leq \tilde{c}(s,r,n) \|f\|_{s+n+1,r}$$

Therefore $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is well-defined. The same follows analogously for $\overline{\mathcal{F}}$. $\square_{2.2.2}$

So we have the following correspondence:

$$\begin{aligned}
-\mathbf{i}x^j &\leftrightarrow \frac{\partial}{\partial p^j} \\
-\mathbf{i}\frac{\partial}{\partial x^j} &\leftrightarrow p^j
\end{aligned}$$

Here the derivatives always act on f or $\mathcal{F}f$ and not on $e^{-\mathbf{i}px}$.

$$x^\alpha \mathrm{D}^\beta f \leftrightarrow \mathbf{i}^{|\alpha|+|\beta|} \mathrm{D}_p^\alpha p^\beta (\mathcal{F}f)(p) = \mathbf{i}^{|\alpha|+|\beta|} p^\beta \mathrm{D}^\alpha (\mathcal{F}f)(p) + \text{lower order terms}$$

Suppose we had worked with $\|f\|_{0,k} = |f|_{C^k}$ as family of norms. Then the norms of the Fourier transform of a function with finite norms would not necessarily be finite.

2.2.3 Theorem (Plancherel, convergence generating factor)

$\mathcal{F}$ and $\overline{\mathcal{F}}$ are inverse to each other:

$$\overline{\mathcal{F}}\mathcal{F} = \mathcal{F}\overline{\mathcal{F}} = \mathbb{1} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$$

Proof

$$\left(\mathcal{F}\overline{\mathcal{F}}f\right)(p) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{-\mathbf{i}px} \underbrace{\left(\int_{\mathbb{R}^n} e^{\mathbf{i}qx} f(q)\,\mathrm{d}^n q\right)}_{(\overline{\mathcal{F}}f)(x)} \mathrm{d}^n x \overset{?}{=} f(p)$$

$$\overset{?}{=} \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} f(q)\left(\int_{\mathbb{R}^n} e^{-\mathbf{i}(p-q)x}\mathrm{d}^n x\right)\mathrm{d}^n q$$

The problem here is, that $e^{-\mathbf{i}(p-q)x}$ does not decay at infinity, so the integral is not well-defined. Instead we have to introduce a *convergence generating factor* $e^{-\varepsilon x^2}$ and, after integrating, calculate the limes $\varepsilon \to 0$.

$$\left(\mathcal{F}\overline{\mathcal{F}}f\right)(p) = \frac{1}{(2\pi)^n}\lim_{\varepsilon\searrow 0}\int_{\mathbb{R}^n} e^{-\mathbf{i}px}e^{-\varepsilon x^2}\left(\int_{\mathbb{R}^n} e^{\mathbf{i}qx} f(q)\,\mathrm{d}^n q\right)\mathrm{d}^n x =$$

$$\overset{\text{Fubini}}{=} \frac{1}{(2\pi)^n}\lim_{\varepsilon\searrow 0}\int_{\mathbb{R}^n} f(q)\left(\int_{\mathbb{R}^n} e^{-\mathbf{i}(p-q)x}e^{-\varepsilon x^2}\mathrm{d}^n x\right)\mathrm{d}^n q$$

The resulting Gaussian integral can be computed in closed form. In one dimension it is:

$$\int_{\mathbb{R}} e^{-\mathbf{i}\lambda x}e^{-\varepsilon x^2}\mathrm{d}x = \int_{\mathbb{R}} e^{-\varepsilon\left(x+\frac{\mathbf{i}\lambda}{2\varepsilon}\right)^2-\frac{\lambda^2}{4\varepsilon}}\mathrm{d}x = e^{-\frac{\lambda^2}{4\varepsilon}}\int_{\mathbb{R}} e^{-\varepsilon\left(x+\frac{\mathbf{i}\lambda}{2\varepsilon}\right)^2}\mathrm{d}x =$$

$$\underset{\mathrm{d}z=\sqrt{\varepsilon}\mathrm{d}x}{\overset{z=\sqrt{\varepsilon}\left(x+\frac{\mathbf{i}\lambda}{2\varepsilon}\right)}{=}} e^{-\frac{\lambda^2}{4\varepsilon}}\int_{\mathbb{R}+\frac{\mathbf{i}\lambda}{2\sqrt{\varepsilon}}} e^{-z^2}\frac{\mathrm{d}z}{\sqrt{\varepsilon}} =$$

$$\underset{\text{deformation}}{\overset{\text{contour}}{=}} \frac{e^{-\frac{\lambda^2}{4\varepsilon}}}{\sqrt{\varepsilon}}\underbrace{\int_{\mathbb{R}} e^{-z^2}\mathrm{d}z}_{=\sqrt{\pi}} = \sqrt{\frac{\pi}{\varepsilon}}e^{-\frac{\lambda^2}{4\varepsilon}}$$

So we get:

$$\left(\mathcal{F}\overline{\mathcal{F}}f\right)(p) = \frac{1}{(2\pi)^n}\lim_{\varepsilon\searrow 0}\int_{\mathbb{R}^n} f(q)\left(\frac{\pi}{\varepsilon}\right)^{\frac{n}{2}} e^{-\frac{(p-q)^2}{4\varepsilon}}\mathrm{d}^n q$$

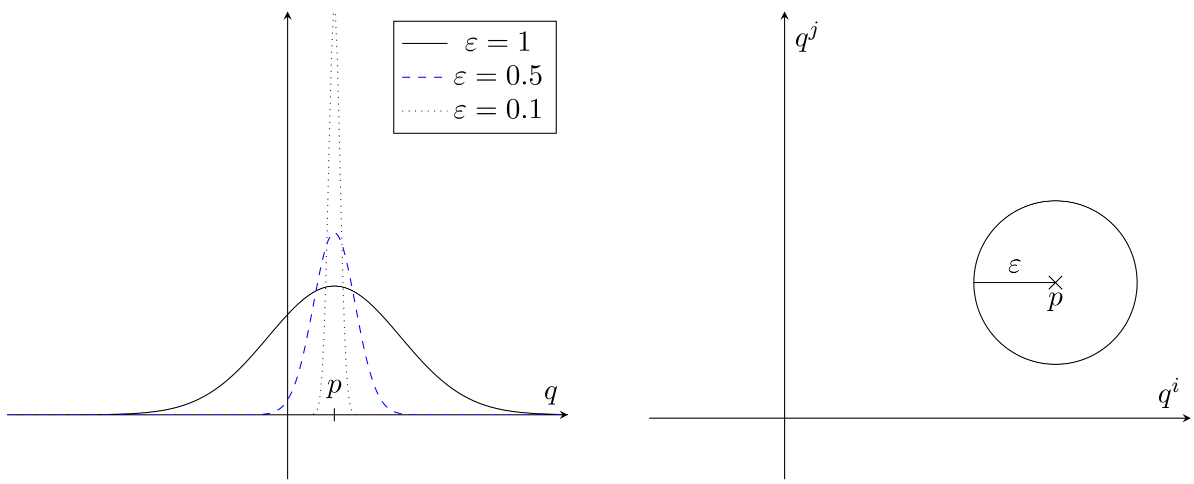

Figure 2.2: The Gaussian gets very narrow and very high as ε decreases.

Estimate the integral as follows:

$$\left(\mathcal{F}\overline{\mathcal{F}}f\right)(p) = \lim_{\varepsilon\searrow 0}\frac{1}{(4\pi\varepsilon)^{\frac{n}{2}}}\left(\int_{\mathbb{R}^n} f(p)\, e^{-\frac{(p-q)^2}{4\varepsilon}}\,\mathrm{d}^n q + \int_{\mathbb{R}^n}\left(f(q)-f(p)\right)e^{-\frac{(p-q)^2}{4\varepsilon}}\,\mathrm{d}^n q\right)$$

The first integral gives:

$$\int_{\mathbb{R}^n} e^{-\frac{(p-q)^2}{4\varepsilon}}\,\mathrm{d}^n q \underset{\mathrm{d}^n z = \frac{\mathrm{d}^n q}{(4\varepsilon)^{\frac{n}{2}}}}{\overset{z=\frac{p-q}{2\sqrt{\varepsilon}}}{=}} (4\varepsilon)^{\frac{n}{2}}\underbrace{\int_{\mathbb{R}^n} e^{-z^2}\,\mathrm{d}^n q}_{=\pi^{\frac{n}{2}}} = (4\pi\varepsilon)^{\frac{n}{2}}$$

So we get:

$$\left(\mathcal{F}\overline{\mathcal{F}}f\right)(p) = f(p) + \lim_{\varepsilon\searrow 0}\frac{1}{(4\pi\varepsilon)^{\frac{n}{2}}}\int_{\mathbb{R}^n}\left(f(q)-f(p)\right)e^{-\frac{(p-q)^2}{4\varepsilon}}\,\mathrm{d}^n q$$

It remains to show that the second summand goes to zero for $\varepsilon \to 0$. We use the following scaling argument:

$$\begin{aligned}\frac{1}{\varepsilon^{\frac{n}{2}}}\int_{\mathbb{R}^n}\left(f(q)-f(p)\right)e^{-\frac{(p-q)^2}{4\varepsilon}}\,\mathrm{d}^4 q &\underset{\mathrm{d}^n q = \varepsilon^{\frac{n}{2}}\mathrm{d}^n u}{\overset{u=\frac{p-q}{\sqrt{\varepsilon}}}{=}} \frac{1}{\varepsilon^{\frac{n}{2}}}\int_{\mathbb{R}^n}\left(f\left(p-\sqrt{\varepsilon}u\right)-f(p)\right)e^{-\frac{u^2}{4}}\varepsilon^{\frac{n}{2}}\,\mathrm{d}^n u = \\ &= \int_{\mathbb{R}^n}\underbrace{\left(f\left(p-\sqrt{\varepsilon}u\right)-f(p)\right)}_{\xrightarrow{\varepsilon\searrow 0} 0 \text{ pointwise}}e^{-\frac{u^2}{4}}\,\mathrm{d}^n u\end{aligned}$$

For the integrand holds:

$$\underbrace{\left(f\left(p-\sqrt{\varepsilon}u\right)-f(p)\right)}_{\xrightarrow{\varepsilon\searrow 0} 0 \text{ pointwise}}e^{-\frac{u^2}{4}} \le \|f\|_{0,0}\,e^{-\frac{u^2}{4}} \in L^1\left(\mathbb{R}^n\right)$$

So the dominated convergence theorem can be applied to get:

$$\left(\mathcal{F}\overline{\mathcal{F}}f\right)(p) = f(p) + \lim_{\varepsilon\searrow 0}\frac{1}{(4\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n}\underbrace{\lim_{\varepsilon\searrow 0}\left(f\left(p-\sqrt{\varepsilon}u\right)-f(p)\right)}_{=0}e^{-\frac{u^2}{4}}\,\mathrm{d}^n u = f(p)$$

$\overline{\mathcal{F}}\mathcal{F} = \mathbb{1}$ follows analogously. $\square_{2.2.3}$

We want to generalize the Fourier transform to $\mathcal{S}^*\left(\mathbb{R}^n\right)$. We begin with the case T_g with $g \in \mathcal{S}\left(\mathbb{R}^n\right)$. We want:

$$\mathcal{F}\left(T_g\right) = T_{\mathcal{F}g}$$

$$T_{\mathcal{F}g}(f) = \int_{\mathbb{R}^n}(\mathcal{F}g)(p)\, f(p)\,\mathrm{d}^n p = \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} e^{-\mathrm{i}px} g(x)\,\mathrm{d}^n x\right) f(p)\,\mathrm{d}^n p$$

Fubini's theorem allows us to interchange the orders of integration:

$$T_{\mathcal{F}g}(f) = \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} g(x)\underbrace{\left(\int_{\mathbb{R}^n} e^{-\mathrm{i}px} f(p)\,\mathrm{d}^n p\right)}_{=\mathcal{F}f(x)}\mathrm{d}^n x = \int_{\mathbb{R}^n} g(x)\,(\mathcal{F}f)(x)\,\mathrm{d}^n x = \mathrm{T}_g(\mathcal{F}f)$$

Since we want $\left(\mathcal{F}T_g\right)(f) = T_{\mathcal{F}g} = T_g(\mathcal{F}f)$, this motivates the following general definition:

2.2.4 Definition (Fourier transform of distributions)

$\mathcal{F}, \overline{\mathcal{F}} : \mathcal{S}^*(\mathbb{R}^n) \to \mathcal{S}^*(\mathbb{R}^n)$ are defined by their action on a test function $f \in \mathcal{S}(\mathbb{R}^n)$:

$$(\mathcal{F}T)(f) := T(\mathcal{F}f)$$
$$\left(\overline{\mathcal{F}}T\right)(f) := T\left(\overline{\mathcal{F}}f\right)$$

It holds:

$$|(\mathcal{F}T)(f)| = |T(\mathcal{F}f)| \underset{\Rightarrow \exists\, r,s\in\mathbb{N}, c\in\mathbb{R}_{>0}}{\overset{T\in\mathcal{S}^*(\mathbb{R}^n)}{\leq}} c\,\|\mathcal{F}f\|_{r,s} \leq \tilde{c}\,\|f\|_{s+n+1,r}$$

Thus $\mathcal{F}T$ is indeed a tempered distribution.

2.2.5 Theorem (Plancherel for distributions)

Plancherel's theorem holds on $\mathcal{S}^*(\mathbb{R}^n)$ as well:

$$\mathcal{F}\overline{\mathcal{F}} = \overline{\mathcal{F}}\mathcal{F} = \mathbb{1}_{\mathcal{S}^*(\mathbb{R}^n)}$$

Proof

$$\left(\mathcal{F}\overline{\mathcal{F}}T\right)(f) \overset{\text{Definition 2.2.4}}{=} \left(\overline{\mathcal{F}}T\right)(\mathcal{F}f) = T\left(\overline{\mathcal{F}}\mathcal{F}f\right) \overset{\text{Plancherel 2.2.3}}{=} T(f)$$

Since this holds for all $f \in \mathcal{S}(\mathbb{R}^n)$ and all $T \in \mathcal{S}^*(\mathbb{R}^n)$ it follows:

$$\mathcal{F}\overline{\mathcal{F}} = \mathbb{1}_{\mathcal{S}^*(\mathbb{R}^n)}$$

The same follows for $\overline{\mathcal{F}}\mathcal{F}$. $\square_{2.2.5}$

2.2.6 Example

1. Fourier transform of δ-Distribution:

$$\begin{aligned}(\mathcal{F}\delta)(f) &= \delta(\mathcal{F}f) = (\mathcal{F}f)(0) = \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\mathrm{i}px} f(x)\,\mathrm{d}^n x \bigg|_{p=0} = \int_{\mathbb{R}^n} \frac{1}{(2\pi)^{\frac{n}{2}}} f(x)\,\mathrm{d}^n x = T_{(2\pi)^{-\frac{n}{2}}}(f)\end{aligned}$$

   This means:

$$\mathcal{F}\delta = T_{(2\pi)^{-\frac{n}{2}}}$$

   Or, in a more computational manner, one can write this as:

$$\mathcal{F}\delta = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\mathrm{i}px} \delta(x)\,\mathrm{d}^n x = \frac{1}{(2\pi)^{\frac{n}{2}}}$$

   This is not satisfying from a mathematical point of view, because one does not know, that the usual formula for the Fourier transform also works for distributions.

2. Consider T_g with $g(p) = e^{\mathbf{i}py}$ for a given $y \in \mathbb{R}^n$.

$$\begin{aligned}(\mathcal{F}T_g)(f) = T_g(\mathcal{F}f) = T_g\left(\frac{1}{(2\pi)^{\frac{n}{2}}}\int e^{-\mathbf{i}px} f(x)\,\mathrm{d}^n x\right) = \\ = \frac{1}{(2\pi)^{\frac{n}{2}}}\int e^{\mathbf{i}py}\left(\int e^{-\mathbf{i}px} f(x)\,\mathrm{d}^n x\right)\mathrm{d}^n p = \\ = (2\pi)^{\frac{n}{2}}\left(\overline{\mathcal{F}}\mathcal{F}\right) f(y) = (2\pi)^{\frac{n}{2}} f(y)\end{aligned}$$

So we get:

$$\begin{aligned}\mathcal{F}T_g f = (2\pi)^{\frac{n}{2}} f(y) \\ (\mathcal{F}T_g)(x) = (2\pi)^{\frac{n}{2}} \delta^{(n)}(x-y)\end{aligned}$$

Formally one can write:

$$(\mathcal{F}g)(x) = \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-\mathbf{i}px} e^{\mathbf{i}py}\mathrm{d}^n p = \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-\mathbf{i}p(x-y)}\mathrm{d}^n p$$

This is ill-defined, but physicists use the formal relation:

$$\int_{\mathbb{R}^n} e^{-\mathbf{i}p(x-y)}\mathrm{d}^n p = (2\pi)^n \delta^{(n)}(x-y)$$

Appendix

Acknowledgements

My special thanks goes to Professor Finster, who gave this lecture and allowed me to publish this script of the lecture.

I would also like to thank all those, who found errors by careful reading and told me of them.

Andreas Völklein