Theory of function spaces

Willem van Zuijlen

22nd July 2020

Contents

Conventions and Notation

- $\mathbb{N} = \{1, 2, 3, \dots\}$, $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ and $\mathbb{N}_{-1} = \{-1, 0\} \cup \mathbb{N}$.
- d **will be a fixed element of** $\mathbb{N}$**.**
- For $x \in \mathbb{R}^d$ or $x \in \mathbb{C}^d$ we write $|x|$ for its euclidean norm $\sqrt{\sum_{i=1}^d |x_i|^2}$, $|x|_1 = \sum_{i=1}^d |x_i|$ and $|x|_\infty = \max_{i\in\{1,\dots,d\}} |x_i|$.
- For $\alpha \in \mathbb{N}_0^d$ we write $|\alpha| = \sum_{i=1}^d \alpha_i$ and D^α or ∂^α for the operation on (smooth) functions by

  $$\partial^\alpha f = \partial_1^{\alpha_1} \dots \partial_d^{\alpha_d} f,$$

  where ∂_i is the partial derivative with respect to the i-th coordinate.
- We write $\| \cdot \|_{L^p}$ for the norm on the L^p spaces. See Section A. As is common, we will not distinguish between an element in L^p with (one of) the function(s) it represents.
- For the inner product on L^2 we write $\langle \cdot, \cdot \rangle_{L^2}$ (as not to get confusion with the notation $\langle \cdot, \cdot \rangle$ for the pairing between distributions and test functions). So

  $$\langle f, g \rangle_{L^2} = \int f \overline{g}.$$

1 Spaces of differentiable functions and distributions

Let Ω be a nonempty open subset of $\mathbb{R}^d$. The underlying field $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$. We write $\mathcal{N}_x(\Omega)$ or just $\mathcal{N}_x$ for the set of neighbourhoods of x in Ω.

Definition 1.1. For a function $f : \Omega \to \mathbb{F}$ we define the *support* $\operatorname{supp} f$ to be the set

$$\operatorname{supp} f = \{x \in \Omega : \forall V \in \mathcal{N}_x \ \exists y \in V \ [\![f(y) \neq 0]\!]\}. \tag{1}$$

Or in words, it is the set of all x such that for all neighbourhoods V of x there exists an element y in that neighbourhood such that $f(y) \neq 0$. Observe that we also have the following equalities

$$\begin{aligned} \operatorname{supp} f &= \overline{\{x \in \Omega : f(x) \neq 0\}} \\ &= \Omega \setminus \{U \subset \Omega : U \text{ is open and } f = 0 \text{ on } U\}. \end{aligned} \tag{2}$$

If F is a set of functions $\Omega \to \mathbb{F}$, we write F_c for the subset of compactly supported functions in F, i.e., $F_c = \{f \in F : \operatorname{supp} f \text{ is compact}\}$.

Definition 1.2.
- We write $C(\Omega, \mathbb{F})$ or $C(\Omega)$ for the set of continuous functions $\Omega \to \mathbb{F}$. We will also write $C^0(\Omega) = C(\Omega)$ and

  $$\|\varphi\|_{C^0} = \sup_{x \in \Omega} |\varphi(x)| \qquad (\varphi \in C(\Omega)),$$

  observe that $\|\varphi\|_{C^0} = \|\varphi\|_{L^\infty}$ for $\varphi \in C(\Omega)$.

- For $k \in \mathbb{N}$ we write $C^k(\Omega, \mathbb{F})$ or $C^k(\Omega)$ for the k-times continuously differentiable functions $\Omega \to \mathbb{F}$ and $\| \cdot \| : C^k(\Omega) \to [0, \infty)$ for

  $$\|f\|_{C^k} = \sum_{\beta : |\beta| \leq k} \|\partial^\beta f\|_{L^\infty} \qquad (f \in C^k(\Omega, \mathbb{F})). \tag{3}$$

- We say that $f : \overline{\Omega} \to \mathbb{F}$ is *k-times continuously differentiable* if $f|_\Omega \in C^k(\Omega)$ and if $\partial^\beta f|_\Omega$ can be extended to a continuous function on $\overline{\Omega}$. We write $C^k(\overline{\Omega}, \mathbb{F})$ or $C^k(\overline{\Omega})$ for the set of such functions.

Definition 1.3. $\mathcal{D}(\Omega)$ is defined to be the vector space $C_c^\infty(\Omega)$. An element of $\mathcal{D}(\Omega)$ is called a *testfunction.*

A linear function $u : \mathcal{D}(\Omega) \to \mathbb{F}$, is called a *distribution* if for all compact sets $K \subset \Omega$, there exist $C > 0$ and $k \in \mathbb{N}_0$ such that

$$|u(\varphi)| \leq C \|\varphi\|_{C^k} \qquad (\varphi \in \mathcal{D}(\Omega), \ \operatorname{supp} \varphi \subset K). \tag{4}$$

If u is a distribution and $k \in \mathbb{N}_0$ is such that for all compact sets K there exists a $C > 0$ such that (4) holds, then u is said to be of *order k*.

1.4. Observe that if u and v are distributions (on Ω) and $\lambda, \mu \in \mathbb{F}$, then $w : \mathcal{D}(\Omega) \to \mathbb{F}$ defined by $w(\varphi) = \lambda u(\varphi) + \mu v(\varphi)$ is a distribution.

Definition 1.5. We define $\mathcal{D}'(\Omega)$ to be the vector space of distributions.

Before we consider the topologies we equip $\mathcal{D}$ and $\mathcal{D}'$ with, let us give some examples of distributions.

Example 1.6. Let f be a locally integrable function on Ω, also written $f \in L^1_{\text{loc}}(\Omega)$. Then $u_f : \mathcal{D}(\Omega) \to \mathbb{F}$ defined by

$$u_f(\varphi) = \int_\Omega f\varphi = \int_\Omega f(x)\varphi(x) \,\mathrm{d}x \tag{5}$$

is a distribution and is of order 0.

Definition 1.7. A *(positive) Radon measure* μ on Ω is a σ-additive function (a measure) on the Lebesgue measurable subsets of Ω (with values in $[0, \infty]$) such that $\mu(K) < \infty$ for all compact sets $K \subset \Omega$.

Example 1.8. If $\mathbb{F} = \mathbb{R}$ let μ be a Radon measure on Ω. Then $u_\mu : \mathcal{D}(\Omega) \to \mathbb{F}$ defined by

$$u_\mu(\varphi) = \int_\Omega \varphi \,\mathrm{d}\mu \tag{6}$$

is a distribution and is of order 0.

Observe that positive locally integrable functions give rise to Radon measures. But not vice versa, as the following example illustrates.

One important example to highlight here is the *Dirac-δ* measure, for which we write δ_0 (or δ_x if we center it at x). It is defined on measurable sets A by

$$\delta_0(A) = \begin{cases} 1 & 0 \in A, \\ 0 & 0 \notin A. \end{cases}$$

Therefore $\int \varphi \,\mathrm{d}\delta_0 = \varphi(0)$.

$\boxed{\textbf{Exercise}}$ **1.1.** Prove that every function in $L^p(\Omega)$ is locally integrable, where p is an element of $[1, \infty]$.

This implies that all element of function spaces like L^p or C^k represent distributions. But these examples only represent distributions of order 0, as the following theorem states.

Theorem 1.9. *A distribution u is of order 0 if and only if it is represented by the difference of two Radon measures, in the sense that there exist Radon measures $\mu_1, \mu_2, \mu_3, \mu_4$ such that $u = u_{\mu_1} - u_{\mu_2} + \mathrm{i}[u_{\mu_3} - u_{\mu_4}]$ (see (6)).*

Remark 1.10. Consider $d = 1$ and $\mathbb{F} = \mathbb{R}$. Observe that u_f with $f(x) = \sin(x)$ for $x \in \mathbb{R}$ is of order zero and so there exists two Radon measures μ_1, μ_2 such that $u = u_{\mu_1} - u_{\mu_2}$. Observe however that $\mu_1 - \mu_2$ is not a signed measure, that is, it is not a σ-additive function on the Lebesgue measurable subsets of $\mathbb{R}$ into $\mathbb{R}$ as "its" value on $\mathbb{R}$ is ill-defined. This means that the theorem in [7] is stated incorrectly (which states that every distribution of zero order is represented by a signed (or complex) Radon measure).

You are asked to do the proof of Theorem 1.9 yourself in Exercise1.3. One might want to use a partition of unity, which we will present in the following.

First we will prove some facts about the topology on Ω (inherited by $\mathbb{R}^d$). Moreover, we use this to prove Theorem 1.16, which shows that if u_μ or u_f is zero, then μ or f is zero. In other words, the function that maps Radon measures into the space of distributions $\mu \mapsto u_\mu$ and the function $L^1_{\text{loc}}(\Omega) \to \mathcal{D}'(\Omega)$ given by $f \mapsto u_f$ are injective.

1.11 (Notation). For $x \in \mathbb{R}^d$, $r > 0$ we write $B(x, r)$ for the (Euclidean) ball in $\mathbb{R}^d$ with center x and radius r:

$$B(x, r) = \{y \in \mathbb{R}^d : |x - y| \le r\}.$$

Theorem 1.12. *There exists an increasing sequence of compact sets $(K_n)_{n\in\mathbb{N}}$ such that $K_n \subset K_{n+1}^\circ$ and $K_n \subset \Omega$ for all $n \in \mathbb{N}$ and*

$$\Omega = \bigcup_{n\in\mathbb{N}} K_n.$$

Proof. Observe that if $\Omega = \mathbb{R}^d$, then we can take K_n to be the closure of the ball around 0 with radius n: $\overline{B(0, n)}$.

Let us first prove that Ω is the union of closed sets. Let $f : \Omega \to [0, \infty]$ be such that $f(x)$ is the distance from x to $\mathbb{R}^d \setminus \Omega$, i.e.,

$$f(x) = \inf\{|x - y| : y \in \mathbb{R}^d \setminus \Omega\} \qquad (x \in \Omega).$$

Then f is a continuous function and therefore $A_n = f^{-1}[\frac{1}{n}, \infty)$ is a closed subset of Ω, $A_n \subset A_{n+1}$ for all $n \in \mathbb{N}$ and $\Omega = \bigcup_{n\in\mathbb{N}} A_n$.

Now it is straightforward to check that $K_n = A_n \cap \overline{B(0, n)}$ satisfies the conditions. $\square$

The next lemma shows there exist many smooth functions.

Lemma 1.13. *Let K be a compact subset of $\mathbb{R}^d$ and U be an open subset of $\mathbb{R}^d$ such that $K \subset U$. There exists a C^∞ function $\varphi : \mathbb{R}^d \to [0, \infty)$ such that φ is strictly positive on K and is zero outside U.*

Proof. By a covering argument it is sufficient to prove the lemma for K being the set that consists one point. For this we consider the following function, $\psi_\varepsilon : \mathbb{R}^d \to [0, \infty)$ defined by

$$\psi_\varepsilon(x) = \begin{cases} e^{\frac{1}{\|x\|^2 - \varepsilon^2}} & \text{if } |x| < \varepsilon, \\ 0 & \text{if } |x| \ge \varepsilon. \end{cases}$$

One can prove that this function is C^∞ by using that $\lim_{t\to\infty} p(t)e^{-t} = 0$. $\square$

Before we turn to the partition of unity, we recall a definition and some theorems from topology.

Definition 1.14. Let E be a topological space. A collection of subsets of E, $\mathcal{U}$, is called a *covering* of E if $\bigcup \mathcal{U} = E$. It is called an *open covering* if each element in $\mathcal{U}$ is an open set. If $\mathcal{U}$ and $\mathcal{V}$ are covers of E, then $\mathcal{V}$ is called a *refinement* of $\mathcal{U}$ or *finer* than $\mathcal{U}$ if for each $V \in \mathcal{V}$ there exists a $U \in \mathcal{U}$ with $V \subset U$. A covering $\mathcal{U}$ is called *locally finite* if for all $x \in E$ there exists a neighbourhood V of x such that V intersects only finitely many elements of $\mathcal{U}$.

1.15 (Partition of unity). There exists a countable covering of Ω of open sets $(U_n)_{n\in\mathbb{N}}$ which closure $\overline{U_n}$ is a compact subset of Ω (by Theorem 1.12). We may and do assume this covering is locally finite, which in particular implies that for all $x \in X$ there exist at most finitely many n such that $x \in U_n$. Now we can find another cover $(V_n)_{n\in\mathbb{N}}$ such that $\overline{V}_n \subset U_n$ for all $n \in \mathbb{N}$ (and $\Omega = \bigcup_{n\in\mathbb{N}} V_n$). As $\overline{V}_n$ is compact, by Lemma 1.13 there exists a $\psi_n \in C^\infty(\Omega, [0,\infty))$ such that $\psi_n > 0$ on V_n and $\psi_n = 0$ outside U_n. Let us define $\Psi = \sum_{n\in\mathbb{N}} \psi_n$, which is finite and strictly positive everywhere. Then we can define $\chi_n := \frac{\psi_n}{\Psi}$ and we obtain

$$0 \le \chi_n(x) \le 1 \quad \text{and} \quad \sum_{n\in\mathbb{N}} \chi_n(x) = 1 \qquad (x \in \Omega).$$

$(\chi_n)_{n\in\mathbb{N}}$ is called a *partition of unity* subordinate to the covering $(U_n)_{n\in\mathbb{N}}$.

$\boxed{\textbf{Exercise}}$ **1.2.** Prove that for any compact set $K \subset \Omega$ there exists a testfunction χ such that $\chi = 1$ on K.

Theorem 1.16. *If μ is a Radon measure on Ω, then for all open sets $U \subset \Omega$*

$$\mu(U) = \sup\{\mu(K) : K \subset U, K \text{ is compact}\}. \tag{7}$$

Moreover,

$$\mu(U) = \sup\{\int \varphi \, \mathrm{d}\mu : \varphi \in C_c^\infty(\Omega, [0,1]), \operatorname{supp}\varphi \subset U\}. \tag{8}$$

Consequently, if $\int \varphi \, \mathrm{d}\mu = 0$ for all $\varphi \in \mathcal{D}(\Omega)$, then $\mu = 0$. Moreover, if $f \in L^1_{\mathrm{loc}}(\Omega)$ and $\int f\varphi = 0$ for all $\varphi \in \mathcal{D}(\Omega)$, then $f = 0$ (in L^1_{loc}).

Proof. (7) follows from Theorem 1.12. (8) follows by taking a partition of unity for an open set U: there exist $\chi_n \in C^\infty(\Omega, [0,1])$ such that $\sum_{n\in\mathbb{N}} \chi_n(x) = \mathbb{1}_U(x)$, then $\int \sum_{n=1}^N \chi_n \, \mathrm{d}\mu \uparrow \int \mathbb{1}_U \mu = \mu(U)$ as $N \uparrow \infty$ by Levi's monotone convergence theorem. □

$\boxed{\textbf{Exercise}}$ **1.3.** Prove Theorem 1.9. One can follow the following steps.

(a) Show that for any testfunction χ the distribution χu is represented by a complex Radon measure (which means $\mu = \mu_1 - \mu_2 + \mathrm{i}[\mu_3 - \mu_4]$ for finite (postive) Radon measures $\mu_1, \mu_2, \mu_3, \mu_4$) by using Riesz' representation theorem (Theorem H.10). You might also want to use the Hahn-Banach theorem (Theorem J.2).

(b) Use the partition of unity to prove Theorem 1.9.

Knowing about the existence of partitions of unity, we can prove that distributions are determined by their "local behaviour", in the sense of the following theorem.

Theorem 1.17. *If $u, v \in \mathcal{D}'(\Omega)$ are such that for all $x \in \Omega$ there exists an open neighbourhood U of x such that $u(\varphi) = v(\varphi)$ for all $\varphi \in \mathcal{D}(\Omega)$ with* $\operatorname{supp} \varphi \subset U$*, then $u = v$.*

Exercise **1.4.** Prove Theorem 1.17.

Of course there are also distributions of higher order. You are asked to give examples, after we define certain operations for distributions. First we make observations that hold for operations on distributions represented by functions.

1.18 (Notation). For a function $f : \Omega \to \mathbb{F}$ and $y \in \mathbb{R}^d$ we define the functions $\check{f} : -\Omega \to \mathbb{F}$ and $\mathcal{T}_y f : \Omega + y \to \mathbb{F}$ by

$$\check{f}(x) = f(-x), \qquad \mathcal{T}_y f(x) = f(x-y) \qquad (x \in \mathbb{R}^d). \tag{9}$$

1.19. Let $f \in L^1_{\mathrm{loc}}(\Omega)$. The following statement follow by applying the change of variables formulae and integration by parts. (For the notation u_f see (5).)

(a) $u_{\check{f}}$ is a distribution on $-\Omega$ and for $\varphi \in \mathcal{D}(-\Omega)$

$$u_{\check{f}}(\varphi) = \int_{-\Omega} f(-x)\varphi(x)\,\mathrm{d}x = \int_{\Omega} f(x)\varphi(-x)\,\mathrm{d}x = u_f(\check{\varphi}) \tag{10}$$

(b) $u_{\mathcal{T}_y f}$ is a distribution on $\Omega + y$ and for $\varphi \in \mathcal{D}(\Omega + y)$

$$u_{\mathcal{T}_y f}(\varphi) = \int_{\Omega+y} f(x-y)\varphi(x)\,\mathrm{d}x = \int_{\Omega} f(x)\varphi(x+y)\,\mathrm{d}x = u_f(\mathcal{T}_{-y}\varphi). \tag{11}$$

(c) Suppose $f \in C^k(\Omega)$ for some $k \in \mathbb{N}$. Let $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \le k$. Then u_{∂^α} is a distribution and for $\varphi \in \mathcal{D}(\Omega)$

$$u_{\partial^\alpha f}(\varphi) = \int_{\Omega} \partial^\alpha f(x)\varphi(x)\,\mathrm{d}x = (-1)^{|\alpha|}\int_{\Omega} f(x)\partial^\alpha\varphi(x)\,\mathrm{d}x = (-1)^{|\alpha|}u_f(\partial^\alpha\varphi). \tag{12}$$

(d) Let $\psi \in C^\infty(\Omega)$. Then $u_{\psi f}$ is a distribution and for $\varphi \in \mathcal{D}(\Omega)$

$$u_{\psi f}(\varphi) = \int_{\Omega} \psi(x) f(x)\varphi(x)\,\mathrm{d}x = u_f(\psi\varphi). \tag{13}$$

(e) Let $l : \mathbb{R}^d \to \mathbb{R}^d$ be linear and bijective. Then $f \circ l$ is locally integrable and $u_{f\circ l}$ is a distribution on $l^{-1}(\Omega)$ and for $\varphi \in \mathcal{D}(l^{-1}(\Omega))$

$$\begin{aligned} u_{f\circ l}(\varphi) &= \int_{l^{-1}(\Omega)} f \circ l(x)\varphi(x)\,\mathrm{d}x = \frac{1}{|\det l|}\int_{\Omega} f(x)\varphi \circ l^{-1}(x)\,\mathrm{d}x \\ &= \frac{1}{|\det l|} u_f(\varphi \circ l^{-1}). \end{aligned} \tag{14}$$

$\boxed{\textbf{Exercise}}$ **1.5.** Let $u \in \mathcal{D}'(\Omega)$. Check that if one defines $w(\varphi)$ to be equal to the right-hand side of (10) with "u" instead of "u_f", i.e., $w(\varphi) = u(\check{\varphi})$, that w is a distribution (on $-\Omega$). Do the same for (11), (12), (13) and (14).

The analogues operations for distributions generalise the previous relations.

Definition 1.20. Let $y \in \mathbb{R}^d$, $\alpha \in \mathbb{N}_0^d, \psi \in \mathcal{E}$ and $l : \mathbb{R}^d \to \mathbb{R}^d$ linear and bijective. For a distribution $u \in \mathcal{D}'$ we define

(a) $\check{u} \in \mathcal{D}'(-\Omega)$ by

$$\check{u}(\varphi) = u(\check{\varphi}) \qquad (\varphi \in \mathcal{D}(-\Omega)),$$

(b) $\mathcal{T}_y u \in \mathcal{D}'(\Omega + y)$ by

$$\mathcal{T}_y u(\varphi) = u(\mathcal{T}_{-y}\varphi) \qquad (\varphi \in \mathcal{D}(\Omega + y)),$$

(c) $\partial^\alpha u \in \mathcal{D}'(\Omega)$ by

$$\partial^\alpha u(\varphi) = (-1)^{|\alpha|} u(\partial^\alpha \varphi) \qquad (\varphi \in \mathcal{D}(\Omega)),$$

(d) $\psi u \in \mathcal{D}'(\Omega)$ by

$$\psi u(\varphi) = u(\psi\varphi) \qquad (\varphi \in \mathcal{D}(\Omega)),$$

(e) $u \circ l \in \mathcal{D}'(l(\Omega))$ by

$$u \circ l(\varphi) = \frac{1}{|\det l|} u(\varphi \circ l^{-1}) \qquad (\varphi \in \mathcal{D}(l(\Omega))).$$

1.21. Observe that all the above operations are "linear in u".

$\boxed{\textbf{Exercise}}$ **1.6.** Construct distributions that are of order k, for any $k \in \mathbb{N}$. Also construct a distribution that is not of any finite order.

$\boxed{\textbf{Exercise}}$ **1.7.** Let $d = 1$ and let $f : \mathbb{R} \to \mathbb{R}$ be the absolute value function: $f(x) = |x|$ for $x \in \mathbb{R}$. Show that the derivative $\mathrm{D}\, u_f$ can be represented by u_g for some locally integrable function g.

$\boxed{\textbf{Exercise}}$ **1.8.** In this exercise we consider dimension one and want to consider the function $x \mapsto \frac{1}{x}$ as a distribution. However, there is a problem of defining the integral of testing it against a testfunction and integrating around zero. Therefore we define the distribution differently.
(1) First prove that for all $\varphi \in \mathcal{D}(\mathbb{R})$ the limit

$$\lim_{\varepsilon \downarrow 0} \int_{\mathbb{R} \setminus [-\varepsilon, \varepsilon]} \frac{\varphi(x)}{x} \, \mathrm{d}x$$

exists and equals $-\int_{\mathbb{R}} \varphi'(x) \log|x| \, \mathrm{d}x$. For this check that $x \mapsto \log|x|$ is integrable around zero and conclude that it is locally integrable.
(2) Prove that $u : \mathcal{D} \to \mathbb{R}$ defined by

$$u(\varphi) := \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}\setminus[-\varepsilon,\varepsilon]} \frac{\varphi(x)}{x} \, \mathrm{d}x$$

is of order 1. Observe that by (1) $u = -\mathrm{D}\, u_f$, where $f(x) = \log|x|$.

Remark 1.22. In Theorem 1.12 we basically used the Heine-Borel theorem, as the closure of a ball is closed and bounded and therefore compact by this theorem. For general topological spaces, one has the following analogues to Theorem 1.12.

Theorem 1.23. *[6, (12.6.1)] Let E be a separable, locally compact, metrizable space and let $\mathfrak{B}$ be a basis of open sets in E. If $\mathcal{U}$ is an open covering of E, then there exists a countable locally finite open covering $(B_n)_{n\in\mathbb{N}}$ of E that is finer than $\mathcal{U}$ and such that $\overline{B_n}$ is compact and belong to $\mathfrak{B}$.*

Theorem 1.24. *[6, (12.6.2)] Let E be a metrizable space that possesses a countable locally finite open covering $(A_n)_{n\in\mathbb{N}}$. Then there exists a countable open covering $(B_n)_{n\in\mathbb{N}}$ such that $\overline{B_n} \subset A_n$ for all $n \in \mathbb{N}$.*

2 Topologies on the spaces of testfunctions and distributions

In this section we introduce the topologies that we equip $\mathcal{D}(\Omega)$ and $\mathcal{D}'(\Omega)$ with.

2.1. One could equip $\mathcal{D}(\Omega)$ with the locally convex topology generated by the seminorms $\|\cdot\|_{C^k}$ for $k \in \mathbb{N}_0$. With this topology the space $\mathcal{D}(\Omega)$ is metrizable but not complete. Therefore we consider a different topology on $\mathcal{D}(\Omega)$.

Theorem 2.2. *Every topological vector space with a topology generated by countably many seminorms is metrizable.*

$\boxed{\textbf{Exercise}}$ **2.1.** Prove that $\mathcal{D}(\Omega)$ equipped with the locally convex topology generated by the seminorm $\|\cdot\|_{C^k}$ is metrizable (i.e., prove (/look up a proof of) Theorem 2.2) but not complete.

Definition 2.3. [27, Chapter IV] Let $\mathcal{X}$ and $\mathcal{Y}$ be vector spaces over $\mathbb{F}$ and $\langle \cdot, \cdot \rangle$: $\mathcal{X} \times \mathcal{Y} \to \mathbb{F}$ be a bilinear form that satisfies the separation axioms:

$$\langle x, y \rangle = 0 \text{ for all } y \in \mathcal{Y} \quad \text{implies } x = 0,$$
$$\langle x, y \rangle = 0 \text{ for all } x \in \mathcal{X} \quad \text{implies } y = 0.$$

The *weak topology* $\sigma(\mathcal{X}, \mathcal{Y})$ on $\mathcal{X}$ is the coarest topology on $\mathcal{X}$ such that all maps $\langle \cdot, y \rangle$ with $y \in \mathcal{Y}$ are continuous. This topology is generated by the seminorms $x \mapsto |\langle x, y \rangle|$ for $y \in \mathcal{Y}$. Similarly one defines the weak topology $\sigma(\mathcal{Y}, \mathcal{X})$ on $\mathcal{Y}$.

Theorem 2.4. [27, Chapter IV, Theorem 1.2] *The dual of the topological space* $(\mathcal{Y}, \sigma(\mathcal{Y}, \mathcal{X}))$ *is* $\mathcal{X}$. *This means that if* $f : \mathcal{Y} \to \mathbb{F}$ *is continuous and linear, then there exists a unique* $x \in \mathcal{X}$ *such that* $f(y) = \langle x, y \rangle$.

Definition 2.5. We define $\langle \cdot, \cdot \rangle : \mathcal{D}'(\Omega) \times \mathcal{D}(\Omega) \to \mathbb{F}$ by

$$\langle u, \varphi \rangle = u(\varphi) \qquad ((u, \varphi) \in \mathcal{D}'(\Omega) \times \mathcal{D}(\Omega)). \tag{15}$$

We equip $\mathcal{D}(\Omega)$ with the weak topology $\sigma(\mathcal{D}(\Omega), \mathcal{D}'(\Omega))$ and $\mathcal{D}'(\Omega)$ with the weak topology $\sigma(\mathcal{D}(\Omega)', \mathcal{D}(\Omega))$, also called the *weak* topology*.

2.6. As is usual in the literature of topological vector spaces, one writes $\mathcal{X}'$ for the topological dual of $\mathcal{X}$. Theorem 2.4 justifies our notation $\mathcal{D}'$, in the sense that $\mathcal{D}'$ is indeed the topological dual of $\mathcal{D}$.

Convergence of sequences in $\mathcal{D}$ can be described very explicitly.

Theorem 2.7. *[7, Theorem on page 99] A sequence* $(\varphi_n)_{n\in\mathbb{N}}$ *converges to a* φ *in* $\mathcal{D}(\Omega)$ *if and only if* (a) *and* (b)*:*

(a) *There exists a compact set* $K \subset \Omega$ *such that the support of* φ_n *and* φ *lies within* K *for all* $n \in \mathbb{N}$.

(b) $\|\varphi_n - \varphi\|_{C^k} \to 0$ *for all* $k \in \mathbb{N}$.

$\boxed{\textbf{Exercise}}$ **2.2.** Prove the "if" part of Theorem 2.7: that (a) and (b) imply $\varphi_n \to \varphi$ in $\mathcal{D}(\Omega)$.

Proof of the "only if" part of Theorem 2.7. Suppose that $\varphi_n \to 0$ in $\mathcal{D}(\Omega)$. We deduce (a) and (b) by arguing by contradiction.

Suppose (a) is not satisfied. Then there exists a sequence $(x_k)_{k\in\mathbb{N}}$ such that no subsequence converges in Ω, and a subsequence $(\varphi_{n_k})_{k\in\mathbb{N}}$ such that $\varphi_{n_k}(x_k) \neq 0$ and $\varphi_{n_k}(x_j) = 0$ for $j > k$. Now let us define a measure with support being equal to the set of x_k's as follows. We let $\mu = \sum_{i\in\mathbb{N}} a_i \delta_{x_i}$, where the a_i's are chosen such that $\sum_{i=1}^{k} a_i \varphi_k(x_i) = 1$; this can always be done inductively. By assumption on the sequence $(x_k)_{k\in\mathbb{N}}$, this measure is a Radon measure, as any compact set $K \subset \Omega$ contains only finitely many x_k's. Therefore it defines a distribution. But $\int \varphi_k \, \mathrm{d}\mu = 1$ for all k, which contradicts the hypothesis that $\varphi_k \to 0$ in $\mathcal{D}(\Omega)$.

In order to show (b) we show that the following statement holds:

$$\text{If } \psi_k \to 0 \text{ in } \mathcal{D}(\Omega), \text{ then } (\psi_k)_{k\in\mathbb{N}} \text{ is uniformly bounded.} \tag{16}$$

In case (16) holds, then $(\partial^\alpha \varphi_k)_{k\in\mathbb{N}}$ is uniformly bounded for any choice of $\alpha \in \mathbb{N}_0$. This implies that these sequences also are uniformly Lipschitz and thus equicontinuous. Therefore, by applying the Arzela-Ascoli theorem (see Theorem G.1) it follows that those sequences converge uniformly on any compact set. As there is a compact set that contains the support of all the functions, this implies that $\|\partial^\alpha \varphi_k\|_{L^\infty} \to 0$ for all $\alpha \in \mathbb{N}_0^d$. This implies (b).

To prove the statement (16) let us assume that $\psi_k \to 0$ in $\mathcal{D}(\Omega)$ and that ψ_k is not uniformly bounded. Therefore, by possibly passing to a subsequence, we may assume that $\|\psi_k\|_{L^\infty} > 3^k$ for all $k \in \mathbb{N}$. Then we can find a sequence $(x_k)_{k\in\mathbb{N}}$ in Ω such that

$$|\psi_k(x_k)| = \|\psi_k\|_{L^\infty}.$$

As ψ_k converges pointwise to zero, we may and do assume –by possibly passing to a subsequence– that $\sum_{i=1}^{k-1} \psi_k(x_i) < 4^{-k}$. As we did before let us construct a Radon measure. We let $\mu = \sum_{i\in\mathbb{N}} 3^{-i}\delta_i$. Then

$$\int \psi_k \, \mathrm{d}\mu = \sum_{i=1}^{k-1} 3^{-i}\psi_k(x_i) + 3^{-k}\psi_k(x_k) + \sum_{i=k+1}^{\infty} 3^{-i}\psi_k(x_i).$$

By the assumptions we have $|\int \psi_k \, \mathrm{d}\mu| \geq -4^{-k} + 1 - 3^{-1} \geq \frac{5}{12}$. Therefore $\int \psi_k \, \mathrm{d}\mu$ does not converge to zero, which contradicts our hypothesis. $\square$

As we have observed in 2.6, $\mathcal{D}'$ is the topological dual of $\mathcal{D}$. This means that a linear function $u : \mathcal{D}(\Omega) \to \mathbb{F}$ is an element of $\mathcal{D}'(\Omega)$ if and only if it is continuous, which means that $u(\varphi_\iota) \to u(\varphi)$ for any net $(\varphi_\iota)_{\iota\in\mathbb{I}}$ with $\varphi_\iota \to \varphi$ in $\mathcal{D}$. The next theorem shows us that it is equivalent to consider only convergent sequences/ to show sequential continuity.

Theorem 2.8. *[7, Page 100] A linear function $u : \mathcal{D}(\Omega) \to \mathbb{F}$ is a distribution if and only if it is sequentially continuous, i.e., $\varphi_n \to \varphi$ implies $u(\varphi_n) \to u(\varphi)$ for all sequences $(\varphi_n)_{n\in\mathbb{N}}$ and φ in $\mathcal{D}(\Omega)$.*

Exercise 2.3. Prove Theorem 2.8.

2.9 ($\mathcal{D}$ is not metrizable). Let us show that $\mathcal{D}(\Omega)$ is not metrizable. We show that if there is a metric on $\mathcal{D}(\Omega)$, then it generates a different topology. Suppose d is a metric on $\mathcal{D}(\Omega)$, such that under the topology of d, $\mathcal{D}(\Omega)$ is a topological vector space. We can find a sequence of increasing compact sets $(K_n)_{n\in\mathbb{N}}$ who's union equals Ω. Let χ_n be a test function that equals 1 on K_n for all n. We can and do choose $\lambda_n \in \mathbb{R}$ such that $d(\lambda_n\chi_n, 0) \leq 2^{-n}$. Then $\lambda_n\chi_n$ converges to 0 but (a) of Theorem 2.7 is not satisfied, which means that $\lambda_n\chi_n$ converges in the topology generated by d but not in the weak topology $\sigma(\mathcal{D}, \mathcal{D}')$.

Definition 2.10. Let $\mathcal{X}$ and $\mathcal{Y}$ be as in Definition 2.3. We say that $\mathcal{X}$ equipped with the $\sigma(\mathcal{X}, \mathcal{Y})$ topology is *sequentially complete* if for every sequence $(x_n)_{n\in\mathbb{N}}$ it holds that if $(\langle x_n, y\rangle)_{n\in\mathbb{N}}$ is a Cauchy sequence for all $y \in \mathcal{Y}$, then there exists an x in $\mathcal{X}$ such that $x_n \to x$ in $\mathcal{X}$.

2.11. In alignment with Definition 2.10, we call $\mathcal{D}'$ *weak* sequentially complete* when it is sequentially complete (with respect to the $\sigma(\mathcal{D}', \mathcal{D})$ topology). This means that the following holds: if $(u_n)_{n\in\mathbb{N}}$ is a sequence such that $(\langle u_n, \varphi\rangle)_{n\in\mathbb{N}}$ is a Cauchy sequence, then there exists a u such that $u_n \to u$ in $\mathcal{D}'$.

2.12. For the proof of the next theorem we introduce the following notation. For a compact subset K of Ω we write $\mathcal{D}_K(\Omega)$ for the subset of $\mathcal{D}(\Omega)$ of functions with support in K. The topology on $\mathcal{D}_K(\Omega)$ is defined by the seminorms $\|\cdot\|_{C^k}$. Moreover, $\mathcal{D}_K(\Omega)$ is a complete metric space.

Theorem 2.13. *$\mathcal{D}'(\Omega)$ is weak* sequentially complete.*

Proof. Suppose that $(u_n)_{n\in\mathbb{N}}$ is a sequence such that $(\langle u_n,\varphi\rangle)_{n\in\mathbb{N}}$ is a Cauchy sequence for all $\varphi \in \mathcal{D}(\Omega)$. It will be clear what the limit should be: We define $u : \mathcal{D}(\Omega) \to \mathbb{F}$ such that $\langle u, \varphi\rangle = \lim_{n\to\infty}\langle u_n, \varphi\rangle$ for any $\varphi \in \mathcal{D}(\Omega)$. Clearly u is linear, so let us show that it is continuous. Let $K \subset \Omega$ be compact and $\mathcal{D}_K(\Omega)$ be as in 2.12. On this space define the function $|||\cdot||| : \mathcal{D}_K(\Omega) \to \mathbb{F}$ by

$$|||\varphi||| := \sup_{k\in\mathbb{N}} |u_k(\varphi)|.$$

This defines a seminorm as it is the supremum of a family of seminorms. This function is lower semicontinuous as it is the supremum of continuous functions. Therefore, the set $S_j := \{f \in \mathcal{D}_K(\Omega) : |||f||| \le j\}$ is closed and convex $\mathcal{D}_K(\Omega)$. As $\bigcup_{j\in\mathbb{N}} S_j = \mathcal{D}_K(\Omega)$, by Baire's Category theorem (Theorem I.1) there exists an $M \in \mathbb{N}$ such that S_M contains an open ball of $\mathcal{D}_K(\Omega)$. Let d_K denote the metric on $\mathcal{D}_K(\Omega)$ and let us write $B_K(\psi,\varepsilon)$ for the ball around ψ with radius $\varepsilon > 0$: $\{\varphi \in \mathcal{D}_K(\Omega) : d_K(\varphi,\psi) < \varepsilon)$. Suppose $\psi \in \mathcal{D}_K(\Omega)$ and $\varepsilon > 0$ are such that the ball $B_K(\psi,\varepsilon)$ is contained in S_M. As S_M is symmetric around zero, also $B_K(-\psi,\varepsilon)$ is contained in S_M. As S_M is convex, this implies that 0 is in the interior of S_M. Therefore we assume $\psi = 0$. This implies that for such $\varphi \in \mathcal{D}_K(\Omega)$

$$d_K(\varphi, 0) < \varepsilon \Longrightarrow |u_n(\varphi)| \le M \text{ for all } n \in \mathbb{N}.$$

Hence $|u(\varphi)| \le M$ for $\varphi \in B_K(0,\varepsilon)$. This implies that u is continuous on $\mathcal{D}_K(\Omega)$. As the topology on $\mathcal{D}_K(\Omega)$ is defined by the seminorms $\|\cdot\|_{C^k}$ this means that there exists a $C > 0$ and a $k \in \mathbb{N}$ such that

$$|u(\varphi)| \le C\|\varphi\|_{C^k(\Omega)} \qquad (\varphi \in \mathcal{D}_K(\Omega)).$$

This proves that u is a distribution. □

2.14. We equip the space of locally integrable functions on Ω with the topology defined by the seminorms $\|\cdot\|_{L^1,K}$ with $K \subset \Omega$ being compact, where

$$\|\varphi\|_{L^1,K} := \|\varphi \mathbb{1}_K\|_{L^1} = \int_K |\varphi| \qquad (\varphi \in \mathcal{D}(\Omega)).$$

Similarly, $L^p_{\mathrm{loc}}(\Omega)$ is equipped with the seminorms $\|\cdot\|_{L^p,K}$ with $K \subset \Omega$ being compact, defined by $\|\varphi\|_{L^1,K} := \|\varphi\mathbb{1}_K\|_{L^1}$.

Theorem 2.15. *As a function $C^k(\Omega) \to \mathcal{D}'(\Omega)$ or as a function $L^1_{\mathrm{loc}}(\Omega) \to \mathcal{D}'(\Omega)$, the map $f \mapsto u_f$ is continuous and injective.*

Proof. The injectivity follows from Theorem 1.16. We leave the proof of the continuity as an exercise (see Exercise 2.4). $\square$

$\boxed{\textbf{Exercise}}$ **2.4.** Prove the continuity of the functions in Theorem 2.15. Think about a topology on the space of Radon measures such that the function that maps a Radon measure μ to the distribution u_μ is continuous.

2.16. [7, Page 98] Let $\psi \in C^\infty(\Omega)$ and $\alpha \in \mathbb{N}_0^d$. Observe that the maps

$$\mathcal{D}(\Omega) \to \mathcal{D}(\Omega), \quad \varphi \mapsto \psi \partial^\alpha \varphi,$$
$$\mathcal{D}'(\Omega) \to \mathcal{D}'(\Omega), \quad u \mapsto \psi \partial^\alpha u,$$

are continuous.

2.17 (Convention/Notation). As is common in literature, and convenient, is to view elements of $C^k(\Omega)$ and $L^1_{\text{loc}}(\Omega)$ as distributions. That is, not to distinguish f from u_f. However, we still prefer not to write "$f(\varphi)$" for "$u_f(\varphi)$" so we will write "$\langle f, \varphi \rangle$" instead.

The notation "$\langle \cdot, \cdot \rangle$" is also commonly used for inner products and this might cause confusion. Indeed, say we take $f, g \in \mathcal{D}$ and as mentioned above, view f as the distribution u_f. Then $\langle f, g \rangle = \int fg$ which is not the same (at least not for general $\mathbb{C}$-valued functions) as $\int f\overline{g}$, which is the inner product between f and g, for which we also write $\langle f, g \rangle_{L^2}$.

2.18 (Restriction of a distribution to a smaller set). Suppose U is an open subset of Ω. Then there exists a linear injection

$$\iota : \mathcal{D}(U) \to \mathcal{D}(\Omega),$$

where $\iota(\varphi)(x) = \varphi(x)$ for $x \in U$ and $\iota(\varphi)(x) = 0$ for $x \in \Omega \setminus U$, and $\varphi \in \mathcal{D}(U)$. As any compact set in U is a compact set in Ω it follows that ι is continuous. Let now $\rho : \mathcal{D}'(\Omega) \to \mathcal{D}'(U)$ be defined by

$$\langle \rho(u), \varphi \rangle = \langle u, \iota(\varphi) \rangle \qquad (\varphi \in \mathcal{D}(U)).$$

Then ρ is also linear and continuous. So $\mathcal{D}(U)$ can be continuously embedded in $\mathcal{D}(\Omega)$ and $\mathcal{D}'(\Omega)$ can be continuously embedded in $\mathcal{D}'(U)$.

For this reason, we will view $\rho(u)$ as the restriction of u to $\mathcal{D}(U)$. Therefore, when $v \in \mathcal{D}'(U)$ we will say "$u = v$ on U" instead of "$\rho(u) = v$".

The following theorem is a kind of counterpart to Theorem 1.17.

Theorem 2.19. *[8, Theorem 7.4] Let $\mathcal{U}$ be a collection of open subsets of $\mathbb{R}^d$ with $\bigcup \mathcal{U} = \Omega$. Let u_U be a distribution on U for all $U \in \mathcal{U}$. Suppose that $u_U = u_V$ on $U \cap V$, in the sense that $u_U(\varphi) = u_V(\varphi)$ for all $\varphi \in \mathcal{D}(\mathbb{R}^d)$ with $\operatorname{supp} \varphi \subset U \cap V$. Then there exists a unique distribution u on Ω such that $u = u_U$ on $\mathcal{D}(U)$ for all $U \in \mathcal{U}$.*

Proof. By 2.18 we may assume that if $U \in \mathcal{U}$ and V is an open subset of U, that $V \in \mathcal{U}$ (as we can take u_V to be the "restriction" of u_U to $\mathcal{D}(V)$).

Let $(U_n)_{n\in\mathbb{N}}$ and $(V_n)_{n\in\mathbb{N}}$ be locally finite covers of Ω as in 1.15: U_n and V_n are open, $\overline{V_n} \subset U_n$ and $\overline{U_n}$ is a compact subset of Ω for all $n \in \mathbb{N}$. For all n we can therefore find finitely many elements in $\mathcal{U}$ that are subsets of U_n and cover V_n. In this way we can also obtain a locally finite covering of Ω that consists of elements in $\mathcal{U}$.

Therefore we may assume instead that U_n is in $\mathcal{U}$ for all $n \in \mathbb{N}$. Let $(\chi_n)_{n\in\mathbb{N}}$ be a partition of unity subordinate to $(U_n)_{n\in\mathbb{N}}$ (like in 1.15). As $(U_n)_{n\in\mathbb{N}}$ is a locally finite cover, this means for all $\varphi \in \mathcal{D}(\Omega)$ that $\chi_n\varphi$ is nonzero for finitely many n. Therefore we can define $u : \mathcal{D}(\Omega) \to \mathbb{F}$ by

$$u(\varphi) = \sum_{n\in\mathbb{N}} u_{U_n}(\chi_n\varphi) \qquad (\varphi \in \mathcal{D}(\Omega)).$$

By Theorem 1.17 it follows that $u = u_U$ for all $U \in \mathcal{U}$, and also the uniqueness follows (and so the definition of u does not depend on the choice of partition of unity).

It is left to check that u is a distribution. That it is a linear function on $\mathcal{D}(\Omega)$ is straightforward to check. For the continuity we use Theorem 2.8 to restrict to sequential continuity. By Theorem 2.7 we know that if $\varphi_n \to \varphi$ in $\mathcal{D}(\Omega)$ that there exists a compact set K that contains the support of all φ_n's. Therefore, there are only finitely many k such that $u_{U_k}(\chi_k\varphi_n)$ is nonzero for some n. That is, there exists a $L \in \mathbb{N}$ such that $u(\varphi_n) = \sum_{k=1}^{L} u_{U_k}(\chi_k\varphi_n)$ for all $n \in \mathbb{N}$. As for all k we have $\chi_k\varphi_n \to \chi_k\varphi$, we have $u_{U_k}(\chi_k\varphi_n) \to u_{U_k}(\chi_k\varphi)$. From this we conclude the continuity of u. $\square$

Remark 2.20 (Regarding the literature on the topologies of $\mathcal{D}$ and $\mathcal{D}'$).
In some books like for example [8], the topology on $\mathcal{D}$ and $\mathcal{D}'$ are not regarded, but only convergence of sequences. Now, after we have the knowledge that we can judge whether a linear map $\mathcal{D} \to \mathbb{F}$ is a distribution by considering whether it is sequentially continuous, we could say "we could forget about the other nets and consider only convergence of sequences". However, this is a posteriori knowledge. Moreover, the topology of $\mathcal{D}$ and $\mathcal{D}'$ is not determined by the convergence of sequences, in the sense that neither of those spaces is first countable. This can be shown as their dimension is not countable, and by using the Hahn-Banach theorem (Theorem J.2).

3 Convolutions

We still consider Ω to be an open subset of $\mathbb{R}^d$ and make some statements in terms of Ω. But regarding the convolution, we will only consider $\Omega = \mathbb{R}^d$, as we will write.

Definition 3.1. Let $f, g : \mathbb{R}^d \to \mathbb{F}$ be measurable functions. If $y \mapsto f(x-y)g(y)$ is integrable for all $x \in \mathbb{R}^d$, then we define the function $f * g : \mathbb{R}^d \to \mathbb{R}$ by

$$f * g(x) = \int f(x-y)g(y)\ \mathrm{d}y.$$

$f * g$ is called the *convolution* of f with g. We will say "$f * g$ exists" instead of "$y \mapsto f(x-y)g(y)$ is integrable for all $x \in \mathbb{R}^d$".

If f and/or g are defined only on a region in $\mathbb{R}^d$, then we will understand $f * g$ to be the convolution of the extended functions that are equal to zero outside their domain in the following sense. Suppose A and B are measurable subsets of $\mathbb{R}^d$ and $f : A \to \mathbb{F}$, $g : B \to \mathbb{F}$, then

$$f * g(x) = \int \overline{f}(x-y)\overline{g}(y) \, \mathrm{d}y,$$

where

$$\overline{f}(x) = \begin{cases} f(x) & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases} \qquad \overline{g}(x) = \begin{cases} g(x) & \text{if } x \in B, \\ 0 & \text{if } x \notin B. \end{cases}$$

3.2 (Commutativity of the convolution). Observe that if $y \mapsto f(x-y)g(y)$ is integrable for all $x \in \mathbb{R}^d$, then also $y \mapsto g(x-y)f(y)$ is integrable and $f * g = g * f$. Because of this commutativity we also call $f * g$ the convolution of f and g (instead of f with g).

3.3. Observe also that if both f and g are integrable, that $f * g$ is integrable and $\|f * g\|_{L^1} \leq \|f\|_{L^1}\|g\|_{L^1}$.

The following theorem will be used often later on. It generalises 3.3.

Theorem 3.4 (Young's inequality). *Let $p, q, r \in [1, \infty]$ be such that*

$$\tfrac{1}{p} + \tfrac{1}{q} = 1 + \tfrac{1}{r}.$$

For $f \in L^p(\mathbb{R}^d)$, $g \in L^q(\mathbb{R}^d)$ we have $f * g \in L^r(\mathbb{R}^d)$ and

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

Proof. If $\{p, q\} = \{1, \infty\}$ and $r = \infty$, then the inequality follows directly. If $p = q = r = 1$ too, as we already mentioned in 3.3. So we may assume $1 < r < \infty$. Then

$$|f(x-y)g(y)| = (|f(x-y)|^p|g(y)|^q)^{\frac{1}{r}}|f(x-y)|^{1-\frac{p}{r}}|g(y)|^{1-\frac{q}{r}}.$$

So by applying the Generalized Hölder inequality (see Theorem A.4), with

$$p_1 = r, \quad p_2 = \frac{p}{1-\frac{p}{r}}, \quad p_3 = \frac{q}{1-\frac{q}{r}},$$

(or just the Hölder inequality when either $p = r$ or $q = r$) we obtain

$$|f * g(x)| \leq \left(\int |f(x-y)|^p|g(y)|^q \, \mathrm{d}y\right)^{\frac{1}{r}} \|f\|_{L^p}^{1-\frac{p}{r}}\|g\|_{L^q}^{1-\frac{q}{r}},$$

and so

$$\begin{aligned} \|f * g\|_{L^r} &\leq \left(\int\int |f(x-y)|^p|g(y)|^q \, \mathrm{d}y \, \mathrm{d}x\right)^{\frac{1}{r}} \|f\|_{L_p}^{1-\frac{p}{r}}\|g\|_{L^q}^{1-\frac{q}{r}} \\ &\leq \|f\|_{L^p}\|g\|_{L^q}. \end{aligned}$$

$\square$

The following is a consequence of the Young's inequality.

Corollary 3.5. *Let $p, q, r \in [1,\infty]$ be such that*

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 2.$$

For $f \in L^p(\mathbb{R}^d)$, $g \in L^q(\mathbb{R}^d)$ and $h \in L^r(\mathbb{R}^d)$ we have that $(f * g)\check{h}$ is integrable and

$$\int (f * g)\check{h} = \int (f * h)\check{g} = \int (f * h)\check{\ } g = \int (\check{f} * \check{h}) g, \tag{17}$$

$$\|(f * g)\check{h}\|_{L^1} \le \|f\|_{L^p} \|g\|_{L^q} \|h\|_{L^r}. \tag{18}$$

$\boxed{\textbf{Exercise}}$ **3.1.** Prove Corollary 3.5.

Definition 3.6. We define the *essential support* of a measurable function $f : \Omega \to \mathbb{F}$ by

$$\operatorname{ess\,supp} f = \mathbb{R}^d \setminus \bigcup \{U \subset \mathbb{R}^d : U \text{ is open and } f = 0 \text{ almost everywhere on } U\}.$$

The above definition is similar to the description of the support of a function as in (2). Observe that also, similar to (1), the essential support is equal to those points for which each neighbourhood of that point, the function f is not equal to zero almost everywhere, which we describe differently by saying that the Lebesgue measure of the set $f^{-1}(\mathbb{F} \setminus \{0\}) \cap V$ is positive:

$$\operatorname{ess\,supp} f = \{x \in \Omega : \forall V \in \mathcal{N}_x [\![\int \mathbb{1}_{f^{-1}(\mathbb{F}\setminus\{0\})\cap V} > 0]\!]\}.$$

As the essential support of f is equal to the one of g if f and g are "essentially the same" in the sense that $f = g$ almost everywhere, one can make sense of the essential support for locally integrable functions (in the usual way by identifying an equivalence class with an element in it).

Of course, for a continuous function f we have

$$\operatorname{supp} f = \operatorname{ess\,sup} f.$$

We recall the following facts about summation of closed sets, Lemma 3.7 and Example 3.8.

Lemma 3.7. *Let $A, B \subset \mathbb{R}^d$ and A be compact and B closed. Then $A + B$ is closed.*

Proof. Let $(d_n)_{n\in\mathbb{N}}$ be a sequence in $A + B$ that converges to an element d in $\mathbb{R}^d$. We prove that $d \in A + B$. By definition, for each n there exist $a_n \in A$ and $b_n \in B$ such that $d_n = a_n + b_n$. As A is compact, $(a_n)_{n\in\mathbb{N}}$ has a convergent subsequence. Let us assume $(a_n)_{n\in\mathbb{N}}$ itself converges in A to an element a. Then $d_n - a_n \to d - a$ and as $d_n - a_n \in B$ for all n and B is closed, $d - a \in B$, which implies $d = a + d - a \in A + B$. $\square$

The assumption that A is not only closed, but also bounded (which together is the same as compact for subsets of $\mathbb{R}^d$) is essential as the following example illustrates.

Example 3.8. Let $A = \mathbb{N}$ and $B = \{-m + \frac{1}{m} : m \in \mathbb{N}, m \geq 2\}$. Then $A + B$ is not closed as $\frac{1}{m}$ is an element of $A + B$ for all $m \in 1 + \mathbb{N}$ but 0 is not.

Theorem 3.9. *For any two measurable functions f, g on $\mathbb{R}^d$ such that $f * g$ exists, we have*

$$\operatorname{supp} f * g \subset \overline{\operatorname{ess\,supp} f + \operatorname{ess\,supp} g}.$$

Proof. Let $x \notin \overline{\operatorname{ess\,supp} f + \operatorname{ess\,supp} g}$, which means that there exists an open neighbourhood V of x such that $V \cap \operatorname{ess\,supp} f + \operatorname{ess\,supp} g = \emptyset$. This means that for all $z \in V$ we have $z - \operatorname{ess\,supp} f \cap \operatorname{ess\,supp} g = \emptyset$, which in turn implies

$$z - \operatorname{ess\,supp} f \subset \bigcup \{U \subset \mathbb{R}^d : U \text{ is open and } g = 0 \text{ almost everywhere on } U\}.$$

Hence for all $z \in V$ the function $y \mapsto f(y)g(z - y)$ is almost everywhere equal to zero and thus $f * g(z) = 0$. Therefore x is in the complement of $\operatorname{supp} f * g$ (see (2)). $\square$

Remark 3.10. In case we view f and g as in Theorem 3.9 not as functions but as equivalence classes of functions (up to equivalence with respect to begin almost everywhere equal) then we will also view $f*g$ as such an equivalence class and also write $\operatorname{ess\,supp} f*g$ instead of $\operatorname{supp} f * g$.

Example 3.11 ($\operatorname{ess\,supp} f + \operatorname{ess\,supp} g \subsetneq \operatorname{supp} f * g = \overline{\operatorname{ess\,supp} f + \operatorname{ess\,supp} g}$)**.**
We adapt Example 3.8 to obtain two measurable functions f and g which are not almost everywhere equal to zero. We define the sets $A, B \subset \mathbb{R}$ by

$$A = \bigcup_{n=2}^{\infty} \Big[n, n + \frac{1}{n}\Big], \qquad B = \bigcup_{m=2}^{\infty} \Big[-m + \frac{1}{m}, -m + \frac{2}{m}\Big].$$

We define $f, g : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = |x|^{-2} \mathbb{1}_A(x), \qquad g(x) = |x|^{-2} \mathbb{1}_B(x) \qquad (x \in \mathbb{R}).$$

Then f and g are integrable functions and so $f * g$ exists (and is integrable). Moreover, $\operatorname{supp} f = A$, $\operatorname{supp} g = B$,

$$A + B = \bigcup_{n,m=2}^{\infty} \Big[n - m + \frac{1}{m}, n - m + \frac{2}{m} + \frac{1}{n}\Big].$$

As in Example 3.8, the set $A + B$ is not closed as 0 is not in $A + B$ but $\frac{1}{m}$ is for all $m \in 1 + \mathbb{N}$. For each $n, m \in 1 + \mathbb{N}$ and $z \in (n - m + \frac{1}{m}, n - m + \frac{2}{m} + \frac{1}{n})$ we can show that $f * g(z) \neq 0$, so that as the support of a function is closed, $\overline{A + B} \subset \operatorname{supp} f * g$. And thus in this case $A + B \subsetneq \operatorname{supp} f * g = \overline{A + B}$

3.12. (a) Let $h \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ and f and g be bounded measurable functions with compact support. Then $f * g$ exists and is bounded with compact support. Therefore $(f * g)\check{h}$ is integrable and (17) holds.

(b) Observe that in the language of distributions, we can rewrite (17) as

$$\langle f * g, \check{h} \rangle = \langle f * h, \check{g} \rangle = \langle (f * h)\check{}, g \rangle = \langle \check{f} * \check{h}, g \rangle,$$

and with the inner product notation

$$\langle f * g, h \rangle_{L^2} = \langle \check{f} * \overline{h}, \overline{g} \rangle_{L^2}.$$

Now let us turn to the definition of the convolution of a distribution with a testfunction. The following observation shows how the definition of the convolution for functions should be extended to distributions.

3.13. Let $\varphi \in \mathcal{D}(\mathbb{R}^d)$ and let $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$. For $x \in \mathbb{R}^d$

$$f * \varphi(x) = \int f(y)\varphi(x-y)\ \mathrm{d}y = \int f(y)\mathcal{T}_x\check{\varphi}(y)\ \mathrm{d}y = u_f(\mathcal{T}_x\check{\varphi}).$$

This lets us naturally generalise the notion of convolution between distributions and testfunctions:

Definition 3.14. Let $u \in \mathcal{D}'(\mathbb{R}^d)$ and $\varphi \in \mathcal{D}(\mathbb{R}^d)$. We define the *convolution* of u with φ to be the function $\mathbb{R}^d \to \mathbb{F}$ defined by

$$u * \varphi(x) = u(\mathcal{T}_x\check{\varphi}) \qquad (x \in \mathbb{R}^d).$$

3.15. Observe that for a distribution u and a testfunction φ the following identities hold.

$$(u * \varphi)\check{} = \check{u} * \check{\varphi} \qquad u(\varphi) = u * \check{\varphi}(0).$$

Now we turn to the differentiability of the convolution $u * \varphi$.

3.16. As translation and differentiation commute for functions, the same is valid for distributions: For $u \in \mathcal{D}'(\Omega)$, $y \in \mathbb{R}^d$, $\alpha \in \mathbb{N}_0^d$ we have

$$\mathcal{T}_y\partial^\alpha u = \partial^\alpha\mathcal{T}_y u.$$

3.17 (Convergence of difference quotients in $\mathcal{D}$ and $\mathcal{D}'$). Let e_i be the basis vector in $\mathbb{R}^d$ in the i-th direction. We write "∂_i" for "∂^{e_i}". Let $\varphi \in \mathcal{D}(\Omega)$ and $i \in \{1, \ldots, d\}$. Then

$$\left(\frac{\mathcal{T}_{he_i} - \mathcal{T}_0}{h}\right)\varphi(x) = \frac{\varphi(x - he_i) - \varphi(x)}{h}. \tag{19}$$

Let us write ψ_h for the function on $\mathbb{R}^d$ for which $\psi_h(x)$ equals (19) (for some h one might have to interpret the right-hand side as the value of the extended φ on $\mathbb{R}^d$ being zero outside Ω). Observe that there exists a compact set that contains the support of ψ_h for all $h \in \mathbb{R}$ such that $0 < |h| \le 1$. We also have that $\psi_h(x) \to \partial_i \varphi(x)$ for all $x \in \Omega$ and that the set $\{\psi_h : h \in \mathbb{R}, 0 < |h| \le 1\}$ is uniformly bounded by the mean value theorem as for all such h and all $x \in \Omega$ we have that $\psi_h(x) = \partial_i \varphi(x + \theta e_i)$ for some $\theta \in \mathbb{R}$ with $|\theta| \le 1$. Similarly, the set

$$\left\{ \left(\frac{\mathcal{T}_{he_i} - \mathcal{T}_0}{h} \right) \partial^\alpha \varphi : h \in \mathbb{R}, 0 < |h| \le 1 \right\}$$

is uniformly bounded. Therefore, by an application of the Arzela-Ascoli theorem (see Theorem G.1, we apply it in the same spirit as in the proof of Theorem 2.7) we deduce that $\psi_h \to \partial_i \varphi$ in $\mathcal{D}(\Omega)$ as $h \to 0$.

Consequently, we have for any $u \in \mathcal{D}'(\Omega)$

$$\left(\frac{\mathcal{T}_{he_i} - \mathcal{T}_0}{h} \right) u \xrightarrow{h \to 0} \partial_i u \quad \text{in } \mathcal{D}'(\Omega).$$

Instead of the distribution u we could have taken a translation of u by x, $\mathcal{T}_x u$, and conclude that $x \mapsto u(\mathcal{T}_x \varphi)$ is differentiable for any $\varphi \in \mathcal{D}(\Omega)$. Moreover, because of the identity

$$\begin{aligned} \frac{u(\mathcal{T}_{x+he_i}\varphi) - u(\mathcal{T}_x \varphi)}{h} &= \left[\mathcal{T}_{-x} \left(\frac{\mathcal{T}_{-he_i} - \mathcal{T}_0}{h} \right) u \right](\varphi) \\ &= u \left(\mathcal{T}_x \left(\frac{\mathcal{T}_{he_i} - \mathcal{T}_0}{h} \right) \varphi \right), \end{aligned} \tag{20}$$

we have $\partial_i (u * \varphi) = u * (\partial_i \varphi)$. Of course, additionally one can continue and iterate the above for derivatives. Then we obtain the following.

Lemma 3.18. *For all distributions u and testfunctions φ the convolution $u * \varphi$ is an element of $C^\infty(\mathbb{R}^d)$, moreover, for $\alpha \in \mathbb{N}_0^d$*

$$\partial^\alpha (u * \varphi) = u * (\partial^\alpha \varphi) = (\partial^\alpha u) * \varphi. \tag{21}$$

The statement of Theorem 3.9, which states that the support of the convolution of two functions is included in the closure of the sum of the supports, extends to distributions (see Theorem 3.21). For this we need to extend the definition of the support to distributions.

Definition 3.19. For a distribution u on Ω we define the *support* $\operatorname{supp} u$ to be the set of $x \in \Omega$ such that for every neighbourhood U of x there exists a $\varphi \in \mathcal{D}(\Omega)$ with support in U and $u(\varphi) \neq 0$:

$$\operatorname{supp} u = \{ x \in \Omega : \forall U \in \mathcal{N}_x \ \exists \varphi \in \mathcal{D}(\Omega) \ [\![\operatorname{supp} \varphi \subset U, \ u(\varphi) \neq 0]\!] \}.$$

Observe that

$$\operatorname{supp} u = \Omega \setminus \{ x \in \Omega : \exists U \in \mathcal{N}_x \ \forall \varphi \in \mathcal{D}(\Omega) \ [\![\operatorname{supp} \varphi \subset U \Rightarrow u(\varphi) = 0]\!] \}.$$

3.20. For a continuous function f the support of f equals the support of u_f, and for a locally integrable function g the essential support of g equals the support of u_g:

$$\begin{aligned} \operatorname{supp} u_f &= \operatorname{supp} f \qquad (f \in C(\Omega)), \\ \operatorname{supp} u_g &= \operatorname{ess\,sup} g \qquad (f \in L^1_{\text{loc}}(\Omega). \end{aligned}$$

Observe moreover that for $\alpha \in \mathbb{N}_0^d$ and $\psi \in C^\infty(\Omega)$

$$\operatorname{supp} \partial^\alpha u \subset \operatorname{supp} u, \qquad \operatorname{supp} \psi u \subset \operatorname{supp} \psi \cap \operatorname{supp} u.$$

Exercise **3.2.** Show that if $\varphi \in \mathcal{D}(\Omega)$, $u \in \mathcal{D}'(\Omega)$ and $\operatorname{supp} \varphi \subset \Omega \setminus \operatorname{supp} u$, then $u(\varphi) = 0$.

Theorem 3.21. *Let u be a distribution and φ be a testfunction. Then* $\operatorname{supp} u * \varphi \subset \operatorname{supp} u + \operatorname{supp} \varphi$.

Exercise **3.3.** Prove Theorem 3.21.

3.22. Observe that Lemma 3.18 also implies that $f*g$ is infinitely differentiable if $f \in L^1_{\text{loc}}$ and $g \in C_c^\infty$ (as $f * g = u_f * g$). The same is true if not g is compactly supported, but f is, in the sence that $f \in L^1$ with compact support and $g \in C^\infty$. This will also be a consequence of Lemma 6.3 in which we prove that if u is a distribution with compact support and φ is smooth –but not necessarily compactly supported– that the convolution exists and is a smooth function.

Observe that this, together with Theorem 3.9 implies that if φ, ψ are testfunctions, then so is $\varphi * \psi$. As $u * \varphi$ for a distribution u and testfunction φ is in $C^\infty(\mathbb{R}^d)$, it defines in particular another distribution. Therefore one can take the convolution of $u * \varphi$ with another testfunction ψ, and also take the convolution of u with the testfunction $\varphi * \psi$. Theorem 3.24 tells us these convolutions are equal. Before we prove an auxiliary lemma that considers an approximation of $\varphi * \psi$.

Lemma 3.23. *Let $\varphi, \psi \in \mathcal{D}(\mathbb{R}^d)$. For any $\varepsilon > 0$ we define $S_\varepsilon \in \mathcal{D}(\mathbb{R}^d)$ as follows. We take finitely many disjoint measurable sets $(F_i)_{i \in I}$ that cover* $\operatorname{supp} \psi$ *and are of diameter at most ε (and so I is assumed to be a finite index set). For every i we choose a $y_i \in F_i$ and define*

$$S_\varepsilon(x) = \sum_{i \in I} (\int_{F_i} \psi) \mathcal{T}_{y_i} \varphi(x).$$

Then $S_\varepsilon \to \varphi * \psi$ in $\mathcal{D}(\mathbb{R}^d)$.

Proof. Observe that S_ε, being a finite linear combination of testfunctions is indeed a testfunction. Moreover, there exists a compact set K such that $\operatorname{supp} S_\varepsilon \subset K$ for all $\varepsilon > 0$, namely $K = \operatorname{supp} \psi + \operatorname{supp} \varphi$, $S_\varepsilon \to \varphi * \psi$ pointwise, and S_ε is uniformly bounded by $\|\varphi\|_{L^\infty} \|\psi\|_{L^1}$. Similarly, $\partial^\alpha S_\varepsilon$ is of the same form but for "$\partial^\alpha \varphi$" instead of "φ". Therefore by an application of the Arzela-Ascoli theorem (see Theorem G.1) we obtain that $S_\varepsilon \to \varphi * \psi$ in $\mathcal{D}(\mathbb{R}^d)$. $\square$

$\boxed{\textbf{Exercise}}$ **3.4.** Why is $\operatorname{supp}\psi + \operatorname{supp}\varphi$ compact for two testfunctions ψ and φ?

Theorem 3.24. *Let u be a distribution and φ, ψ be testfunctions. Then*

$$(u * \varphi) * \psi = u * (\varphi * \psi). \tag{22}$$

Proof. First observe that with S_ε as in Lemma 3.23

$$u * (\varphi * \psi)(x) = u(\mathcal{T}_x(\varphi * \psi)\check{}) = \lim_{\varepsilon \downarrow 0} u(\mathcal{T}_x \check{S}_\varepsilon)$$

This follows by applying Lemma 3.23 as follows

$$\begin{aligned} u(\mathcal{T}_x \check{S}_\varepsilon) &= u\left(\sum_{i \in I} (\int_{F_i} \psi) \mathcal{T}_{x-y_i}\varphi\right) = \sum_{i \in I} (\int_{F_i} \psi) \mathcal{T}_{y_i} u(\mathcal{T}_x \varphi) \\ &= \sum_{i \in I} (\int_{F_i} \psi)(u * \varphi)(x - y_i). \end{aligned}$$

This in turn is an approximation for $\psi * (u * \varphi)$ by Lemma 3.23 and so by taking a limit we obtain (22). □

3.25. As a direct consequence of Theorem 3.24 we have for $u \in \mathcal{D}'(\mathbb{R}^d)$ and $\varphi, \psi \in \mathcal{D}(\mathbb{R}^d)$

$$\langle u * \varphi, \psi \rangle = (u * \varphi) * \check{\psi}(0) = u * (\varphi * \check{\psi})(0) = \langle u, \check{\varphi} * \psi \rangle.$$

Compare this with 3.12 (a).

The next question that arises is: "Can one take a convolution between distributions?". The answer is not completely yes, in the sense that one can take a convolution if one of the distributions has compact support (think about taking the convolution of the constant function equal to 1 everywhere with itself). We will turn to distributions with compact in the Section 5. But first we focus on the approximation of distributions by smooth functions formed by mollification.

$\boxed{\textbf{Exercise}}$ **3.5.** Let $y \in \mathbb{R}^d$ and $\varphi \in \mathcal{D}(\mathbb{R}^d)$. Calculate $\delta_y * \varphi$.

$\boxed{\textbf{Exercise}}$ **3.6.** For each of the following cases, find $u \in \mathcal{D}'(\mathbb{R}^d)$ and $\varphi \in \mathcal{D}(\mathbb{R}^d)$ such that:

(a) $u * \varphi(x) = 0$ for all $x \in \mathbb{R}^d$,

(b) $u * \varphi(x) = 1$ for all $x \in \mathbb{R}^d$,

(c) $u * \varphi(x) = x$ for all $x \in \mathbb{R}^d$,

(d) $u * \varphi(x) = \sin(x)$ for all $x \in \mathbb{R}^d$.

$\boxed{\textbf{Exercise}}$ **3.7.** Consider the distribution on $\mathbb{R}$ given by $h = \mathbb{1}_{[0,\infty)}$, also called the Heaviside function. For $\varphi \in \mathcal{D}(\mathbb{R})$ calculate $h * \varphi'$, where φ' denotes the derivative of φ. Calculate the distributional derivative h' of h. Show that $(h * \varphi)' = h * \varphi' = h' * \varphi$.

4 Mollifiers

Now we turn to mollifiers and mollifications of locally integrable functions. To prove their convergence, we recall Lebesgue's differentiation theorem, which states that almost all points in $\mathbb{R}^d$ are Lebesgue points for any locally integrable function f.

Theorem 4.1 (Lebesgue's differentiation theorem). *[17, Theorem 2.3.4] For all $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ almost every point in $\mathbb{R}^d$ is a Lebesgue point, i.e., for almost all points x,*

$$\varepsilon^{-d} \int_{B(x,\varepsilon)} |f(y) - f(x)| \, \mathrm{d}y \xrightarrow{\varepsilon \downarrow 0} 0. \tag{23}$$

4.2 (Notation). For any closed set $A \subset \mathbb{R}^d$ we write A_ε for those points in $\mathbb{R}^d$ that are at most at ε distance from A, so that

$$A_\varepsilon = A + \overline{B(0,\varepsilon)} = \{y \in \mathbb{R}^d : \inf_{x \in A} |x - y| \le \varepsilon\}.$$

For a $f \in L^p_{\mathrm{loc}}(\mathbb{R}^d)$ and a compact set $K \subset \mathbb{R}^d$ we will also write $\|f\|_{L^p(K)}$ for $\|f \mathbb{1}_K\|_{L^p}$.

Theorem 4.3. *Let $f \in L^p_{\mathrm{loc}}(\mathbb{R}^d)$ for some $p \in [1,\infty)$ and $\psi \in C_c(\mathbb{R}^d)$. For $\varepsilon > 0$ we write ψ_ε for the function defined by $\psi_\varepsilon(x) = \varepsilon^{-d}\psi(\varepsilon^{-1}x)$. Then the following statements hold.*

(a) *$f * \psi_\varepsilon(x) \xrightarrow{\varepsilon \downarrow 0} (\int \psi) f(x)$ for all Lebesgue points $x \in \mathbb{R}^d$ of f.*

(b) *If f is continuous on an open set $U \subset \mathbb{R}^d$, then $f * \psi_\varepsilon \to (\int \psi) f$ uniformly on all compact subset of U.*

(c) *If $f * \psi_\varepsilon \to (\int \psi) f$ in $L^p_{\mathrm{loc}}(\mathbb{R}^d)$.*

Proof. As $\int \psi_\varepsilon = \int \psi$ for all $\varepsilon > 0$, we have

$$f * \psi_\varepsilon(x) - (\int \psi) f(x) = \int \psi_\varepsilon(x - y)\Big(f(y) - f(x)\Big) \, \mathrm{d}y.$$

As we can find an ε such that $\operatorname{supp} \psi_\varepsilon \subset B(0,1)$, we may without loss of generality assume that $\operatorname{supp} \psi \subset B(0,1)$. Then

$$|f * \psi_\varepsilon(x) - (\int \psi) f(x)| \le \|\psi\|_{L^\infty} \varepsilon^{-d} \int_{B(x,\varepsilon)} |f(y) - f(x)| \, \mathrm{d}y. \tag{24}$$

From this (a) follows. Suppose f is continuous on an open set U and $K \subset U$ is compact. Let $\delta > 0$ be such that $K_\delta \subset U$. As f is uniformly continuous on K_δ, the convergence in (23) is valid uniformly for $x \in K$. Hence (b) also follows from (24).

Let us turn to the proof of (c). Let $K \subset \mathbb{R}^d$ be compact. We will show

$$\|f * \psi_\varepsilon - (\int \psi) f\|_{L^p(K)} \to 0.$$

But first we observe that for all $h \in L^p_{\mathrm{loc}}$ we have

$$|h * \psi_\varepsilon(x)| \leq \int |h(y)\psi_\varepsilon(x-y)| \,\mathrm{d}y \leq \int |(h\mathbb{1}_{K_\varepsilon})(y)\psi_\varepsilon(x-y)| \,\mathrm{d}y \qquad (x \in K),$$

so that with Young's inequality and as $\|\psi_\varepsilon\|_{L^1} = \|\psi\|_{L^1}$

$$\|(h * \psi_\varepsilon)\mathbb{1}_K\|_{L^p} \leq \|\psi\|_{L^1}\|h\mathbb{1}_{K_\varepsilon}\|_{L^p} = \|\psi\|_{L^1}\|h\|_{L^p(K_\varepsilon)}. \tag{25}$$

Let $\delta > 0$. Take a function g that is continuous on K_1 and equals 0 outside K_1 such that

$$\|f - g\|_{L^p(K_1)} < \delta.$$

Then, as $|f * \psi_\varepsilon - (\int \psi)f| \leq |f * \psi_\varepsilon - g * \psi_\varepsilon| + |g * \psi_\varepsilon - (\int \psi)g| + |(\int \psi)g - (\int \psi)f|$, we obtain for $\varepsilon \in (0,1)$ by using (25)

$$\begin{aligned}\|f * \psi_\varepsilon - (\textstyle\int \psi)f\|_{L^p(K)} &\leq \|(f-g) * \psi_\varepsilon\|_{L^p(K)} + \|g * \psi_\varepsilon - (\textstyle\int \psi)g\|_{L^p(K)} \\ &\quad + \|\psi\|_{L^1}\|g - f\|_{L^p(K)} \\ &\leq 2\delta\|\psi\|_{L^1} + (\int \mathbb{1}_K)\|g * \psi_\varepsilon - (\textstyle\int \psi)g\|_{L^\infty(K)}.\end{aligned}$$

As by (b) $\|g * \psi_\varepsilon - (\int \psi)g\|_{L^\infty(K)} \xrightarrow{\varepsilon \downarrow 0} 0$ (take for example $U = K_1^\circ$, so that g is continuous on U and $K \subset U$) this implies (c). $\square$

Definition 4.4. Let ψ be a testfunction such that $\operatorname{supp}\psi \subset B(0,1)$ and $\int \psi = 1$ (the existence is guaranteed by Lemma 1.13). Such a function is called a *mollifier*. For $\varepsilon > 0$ we define ψ_ε to be the function on $\mathbb{R}^d$ defined by

$$\psi_\varepsilon(x) = \varepsilon^{-d}\psi(\tfrac{x}{\varepsilon}) \qquad (x \in \mathbb{R}^d).$$

Then $\operatorname{supp}\psi_\varepsilon \subset B(0,\varepsilon)$ and $\int \psi_\varepsilon = 1$. For a distribution u we call u_ε defined by

$$u_\varepsilon := u * \psi_\varepsilon \tag{26}$$

a *mollification* of u (with respect to ψ of order ε).

By Lemma 3.18 we know that u_ε is a smooth function. For a function f in L^p_{loc} we also know that $f_\varepsilon \to f$ in L^p_{loc}, by Theorem 4.3. So in particular,

$$\int f_\varepsilon \varphi \to \int f\varphi \qquad (\varphi \in \mathcal{D}(\mathbb{R}^d)),$$

which implies that $f_\varepsilon \to f$ in $\mathcal{D}'(\mathbb{R}^d)$. This "extends" to any distribution, see the following theorem. This theorem follows by Theorem 4.3.

Theorem 4.5. *Let ψ be a mollifier and u a distribution. Then $u_\varepsilon \in C^\infty(\mathbb{R}^d)$,*

$$\operatorname{supp} u_\varepsilon \subset (\operatorname{supp} u)_\varepsilon,$$

and $u_\varepsilon \xrightarrow{\varepsilon \downarrow 0} u$ in $\mathcal{D}'(\mathbb{R}^d)$.

$\boxed{\textbf{Exercise}}$ **4.1.** Prove Theorem 4.5.

By using the previous theorem, we can actually find $u_\varepsilon \in C_c^\infty(\mathbb{R}^d)$ such that $u_\varepsilon \to u$ for any $u \in \mathcal{D}'(\mathbb{R}^d)$:

Theorem 4.6. *$\mathcal{D}(\mathbb{R}^d)$ is dense in $\mathcal{D}'(\mathbb{R}^d)$.*

$\boxed{\textbf{Exercise}}$ **4.2.** Prove Theorem 4.6.

Remark 4.7. By choosing a mollifier ψ (for example the one of Lemma 1.13) which is supported on $B(0,1)$ we can define another mollifier $\tilde{\psi}$ that is supported in $B(-\frac{1}{2}, \frac{1}{4}) \cup B(\frac{1}{2}, \frac{1}{4})$ as follows:

$$\tilde{\psi} = \tfrac{1}{2}(\mathcal{T}_{-\frac{1}{2}}\psi_{\frac{1}{4}} + \mathcal{T}_{\frac{1}{2}}\psi_{\frac{1}{4}}).$$

As $\delta * \tilde{\psi} = \tilde{\psi}$, which is zero around zero and thus the inclusion in Theorem 4.5 does not need to be an inclusion.

5 Compactly supported distributions

Similar to the spaces $\mathcal{D}$ and $\mathcal{D}'$ we introduce the space $\mathcal{E}$ that consists of all smooth functions and its dual $\mathcal{E}'$. We will see that $\mathcal{E}'$ corresponds to the distributions with compact support.

Definition 5.1. We define $\mathcal{E}(\Omega)$ to be the set $C^\infty(\Omega)$ equipped with the topology generated by the seminorms $\|\cdot\|_{C^k,K}$ with $K \subset \Omega$ compact and $k \in \mathbb{N}_0$ given by

$$\|f\|_{C^k,K} = \|f|_K\|_{C^k(K)} = \sum_{\beta \in \mathbb{N}_0^d : |\beta| \le k} \sup_{x \in K} |\partial^\beta f(x)|.$$

We write $\mathcal{E}'(\Omega)$ for the space of continuous linear functions $u : \mathcal{E}(\Omega) \to \mathbb{F}$. This means (see for example [4, Theorem IV.3.1]) that $u \in \mathcal{E}'(\Omega)$ if and only if there exists a compact set K, a $k \in \mathbb{N}_0$ and a $C > 0$ such that

$$|u(\varphi)| \le C\|\varphi\|_{C^k,K} \qquad (\varphi \in \mathcal{E}(\Omega)). \tag{27}$$

We equip $\mathcal{E}'(\Omega)$ with the weak* topology $\sigma(\mathcal{E}'(\Omega), \mathcal{E}(\Omega))$.

5.2. Observe that if $u \in \mathcal{E}'(\Omega)$ and $K \subset \Omega$ is compact, $k \in \mathbb{N}_0$ and $C > 0$ are such that (27) holds, then the following holds. If $\varphi \in \mathcal{E}(\Omega)$ and $\operatorname{supp}\varphi \cap K = \emptyset$ then $u(\varphi) = 0$. Hence $\operatorname{supp} u \subset K$ and so an element of $\mathcal{E}'(\Omega)$ defines a distribution with compact support. We will prove that a distribution with compact support can be extended to an element of $\mathcal{E}'(\Omega)$ in 5.6.

Let us recall the Leibniz differentiation rule.

5.3 (Leibniz' rule). If $k \in \mathbb{N}_0$, $f, g \in C^k(\Omega)$ and $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \le k$, then

$$\partial^\alpha (fg) = \sum_{\substack{\beta \in \mathbb{N}_0^d \\ \beta \le \alpha}} \binom{\alpha}{\beta} (\partial^\beta f)(\partial^{\alpha-\beta} g), \tag{28}$$

where $\beta \le \alpha$ means $\beta_i \le \alpha_i$ for all $i \in \{1, \dots, d\}$ and with $\alpha! = \prod_{i=1}^d \alpha_i!$,

$$\binom{\alpha}{\beta} = \frac{\alpha!}{(\alpha-\beta)!\beta!} = \prod_{i=1}^d \binom{\alpha_i}{\beta_i}.$$

As a consequence, for $x \in \Omega$

$$\begin{aligned}\sum_{\substack{\alpha \in \mathbb{N}_0^d \\ |\alpha| \le k}} |\partial^\alpha (fg)(x)| &\le \sum_{\substack{\alpha \in \mathbb{N}_0^d \\ |\alpha| \le k}} \sum_{\substack{\beta,\gamma \in \mathbb{N}_0^d \\ \beta+\gamma=\alpha}} \binom{\alpha}{\beta} |\partial^\beta f(x)||\partial^\gamma g(x)| \\ &\le \left(\sum_{\substack{\alpha \in \mathbb{N}_0^d \\ |\alpha| \le k}} \sum_{\substack{\beta \in \mathbb{N}_0^d \\ \beta \le \alpha}} \binom{\alpha}{\beta} \right) \sum_{\substack{\beta \in \mathbb{N}_0^d \\ |\beta| \le k}} \sum_{\substack{\gamma \in \mathbb{N}_0^d \\ |\gamma| \le k}} |\partial^\beta f(x)||\partial^\gamma g(x)|. \end{aligned} \tag{29}$$

Hence for $C = \sum_{\alpha \in \mathbb{N}_0^d : |\alpha| \le k} \sum_{\beta \le \alpha} \binom{\alpha}{\beta}$

$$\|fg\|_{C^k} \le C \|f\|_{C^k} \|g\|_{C^k} \qquad (f, g \in C^k(\Omega)). \tag{30}$$

5.4. Recall 2.16. Now with Leibniz' rule and with the topology on $\mathcal{E}(\Omega)$ we conclude for example also that the map $\mathcal{E}(\Omega) \times \mathcal{D}(\Omega) \to \mathcal{D}(\Omega)$ given by $(\psi, \varphi) \mapsto \psi \partial^\alpha \varphi$ is continuous for all $\alpha \in \mathbb{N}_0^d$.

5.5. Observe that (28) extends to the product of a distribution with a smooth function. That is, if $u \in \mathcal{D}(\Omega)$ and $\psi \in C^\infty(\Omega)$, then

$$\partial^\alpha (\psi u) = \sum_{\substack{\beta \in \mathbb{N}_0^d \\ \beta \le \alpha}} \binom{\alpha}{\beta} (\partial^\beta \psi)(\partial^{\alpha-\beta} u),$$

5.6. Let u be a distribution on Ω with compact support K. We have already seen in Exercise 3.2 that if $\varphi \in \mathcal{D}(\Omega)$ and $\operatorname{supp} \varphi \subset \Omega \setminus K$, then $u(\varphi) = 0$. Let $\varepsilon > 0$ be such that $K_\varepsilon \subset \Omega$ and let $\chi \in \mathcal{D}(\Omega)$ be equal to 1 on K_ε (see 1.15). As $\operatorname{supp}(\varphi - \chi\varphi) \subset \Omega \setminus K_\varepsilon^\circ \subset \Omega \setminus K$, we have $u(\varphi) = u(\chi\varphi)$. Let $K_0 = \operatorname{supp} \chi$. As u is a distribution, there exist $C_1 > 0$ and $k \in \mathbb{N}_0$ such that $|u(\varphi)| \le C_1 \|\varphi\|_{C^k}$ for all $\varphi \in \mathcal{D}(\Omega)$ with $\operatorname{supp} \varphi \subset K_0$. This implies for all $\varphi \in \mathcal{D}(\Omega)$

$$|u(\varphi)| = |u(\chi\varphi)| \le C_1 \|\chi\varphi\|_{C^k}.$$

By Leibniz' rule we therefore have with $C' = C_1 C\|\chi\|_{C^k}$, where $C > 0$ is such that (30) holds,

$$|u(\varphi)| \le C'\|\varphi\|_{C^k,K_0} \qquad (\varphi \in \mathcal{D}(\Omega)). \tag{31}$$

Therefore u extends to an element of $\mathcal{E}'(\Omega)$ (for example by defining the extension v by $v(\varphi) = u(\chi\varphi)$ for $\varphi \in \mathcal{E}$). This will be used to prove Theorem 5.7.

$\boxed{\textbf{Exercise}}$ **5.1.** Show that if $u \in \mathcal{E}'(\Omega)$ and $u \neq 0$, that there exists a $\varphi \in \mathcal{D}(\Omega)$ such that $u(\varphi) \neq 0$.

Theorem 5.7. *The inclusion map $\mathcal{D}(\Omega) \to \mathcal{E}(\Omega)$ is sequentially continuous; $\mathcal{D}(\Omega)$ is dense in $\mathcal{E}(\Omega)$; and, the map $\iota : \mathcal{E}'(\Omega) \to \mathcal{D}'(\Omega)$ defined by $\iota(u) = u|_{\mathcal{D}(\Omega)}$ is continuous and injective and its image is the set of compactly supported distributions.*

Proof. That $\mathcal{D}(\Omega) \to \mathcal{E}(\Omega)$ is sequentially continuous follows from Theorem 2.7. The statement about ι follows from 5.2 and 5.6. We show that $\mathcal{D}(\Omega)$ is dense in $\mathcal{E}(\Omega)$. Let $\psi \in \mathcal{E}(\Omega)$ and let $(\chi_n)_{n\in\mathbb{N}}$ be a partition of unity. Let $K \subset \Omega$ be compact and $k \in \mathbb{N}$. By the properties of a partition of unity, there exists an n such that $\sum_{m=1}^n \chi_m = 1$ on K_δ for some $\delta > 0$ such that $K_\delta \subset \Omega$ (this you might have even proved in Exercise 1.2). Therefore

$$\left\| \psi - \sum_{m=1}^n \chi_m \psi \right\|_{C^k,K} = 0.$$

Hence $\sum_{m=1}^n \chi_m \psi \to \psi$ in $\mathcal{E}$. ☐

In 5.6 the inequality (31) holds for K_0 which is larger than K. The next exercise proves that it might be that (31) does not hold for $K_0 = \operatorname{supp} u$.

$\boxed{\textbf{Exercise}}$ **5.2.** [8, Exercise 8.3] Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of distinct elements and x be in $\mathbb{R}$ such that $x_n \to x$.

(a) Show that there exists a sequence $(a_n)_{n\in\mathbb{N}}$ in $(0,\infty)$ such that

$$\sum_{n\in\mathbb{N}} a_n = \infty, \qquad \sum_{n\in\mathbb{N}} a_n |x_n - x| < \infty.$$

(b) Prove that u defined by

$$u(\varphi) = \sum_{n\in\mathbb{N}} a_n(\varphi(x_n) - \varphi(x)) \qquad (\varphi \in \mathcal{D})$$

defines a distribution of order ≤ 1. Prove that the support of u is the compact set $\{x_n : n \in \mathbb{N}\} \cup \{x\}$.

(c) Show that for all $n \in \mathbb{N}$ there exists a $\varphi_n \in \mathcal{D}$ such that $\varphi_n = 1$ on a neighbourhood of x_i for all $i \in \{1, \dots, n\}$ and $\varphi_n = 0$ on a neighbourhood of x_j for all $j > n$ and $\varphi_n = 0$ on a neighbourhood of x. Prove that for all $k \in \mathbb{N}$

$$\|\varphi_n\|_{C^k, \operatorname{supp} u} = 1, \qquad u(\varphi_n) = \sum_{i=1}^{n} a_i.$$

(d) Conclude that for $K = \operatorname{supp} u$, (27) does not hold for any $k \in \mathbb{N}$.

The following example illustrates that these maps are not homeomorphisms on their image.

Example 5.8. Let $\Omega = \mathbb{R}$.

- Let ϕ be an element of $\mathcal{D}$ such that $\int \phi = 1$. Define $\phi_n = \mathcal{T}_n \phi$, i.e., $\phi_n(x) = \phi(x-n)$ for $x \in \mathbb{R}$. Then $\phi_n \to 0$ in $\mathcal{E}$. As for all compact sets K there exists an N such that $\operatorname{supp} \phi_n \cap K = \emptyset$ for all $n \geq N$. However, for u the distribution corresponding to the Lebesgue measure, or equivalently to the constant function $\mathbb{1}$, we have $u(\phi_n) = 1$ for all n, whence $(\phi_n)_{n \in \mathbb{N}}$ does not converge in $\mathcal{D}$.
- δ_n is an element of $\mathcal{D}'$ and of $\mathcal{E}'$ for all $n \in \mathbb{N}$. We have $\delta_n \to 0$ in $\mathcal{D}'$ but not in $\mathcal{E}'$, as we have $\delta_n(\mathbb{1}) = 1$ for all n.

5.9. So $\mathcal{E}'$ does not have the same topology as $\iota(\mathcal{E}')$. However, $(\mathcal{E}', \sigma(\mathcal{E}', \mathcal{D}))$ is homeomorphic to $\iota(\mathcal{E}')$.

So the relative topology of $\mathcal{D}$ as a subspace of $\mathcal{E}$ is different from the topology on $\mathcal{D}$, namely $\sigma(\mathcal{D}, \mathcal{D}')$.

5.10. (31) also implies that every distribution with compact support is of finite order. We will show in Theorem 5.13 that if the derivatives of a testfunction up to that order are zero on the support of the distribution, that the distribution evaluated in the testfunction equals zero.

5.11. Let φ be a testfunction on Ω and $a \in \Omega$. Suppose that $\partial^\alpha \varphi(a) = 0$ for all $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \leq k$. Let $\varepsilon > 0$ be such that $B(a, \varepsilon) \subset \Omega$. By Taylor's formula (see Theorem C.7) we know that φ equals a function ψ (use (169) with $l = k+1$) and that there exists a $C > 0$ (namely $C = \sum_{\alpha \in \mathbb{N}_0^d : |\alpha| = k+1} \|\partial^\alpha \varphi\|_\infty$) such that

$$|\psi(x)| \leq C|x-a|^{k+1} \leq C\varepsilon^{k+1} \qquad (x \in B(a, \varepsilon)).$$

By a repetition of the above argument for the derivatives of φ, we obtain a $C > 0$ such that for all $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \leq k$

$$|\partial^\alpha \varphi(x)| \leq C\varepsilon^{k+1-|\alpha|} \qquad (x \in B(a, \varepsilon)).$$

5.12. In 5.6 we have already used that for any compact set $K \subset \Omega$ and $\varepsilon > 0$ such that $K_\varepsilon \subset \Omega$, we can find a $\chi \in \mathcal{D}(\Omega)$ that equals 1 on K_ε by using 1.15.

As we have the tools of mollifiers, we can construct such a χ in such a way that the support is within $K_{3\varepsilon}$. Moreover, we can construct them for any closed set as follows.

Let F be a closed set. For $\varepsilon > 0$ let $\chi_\varepsilon = \mathbb{1}_{F_{2\varepsilon}} * \psi_\varepsilon$ for a positive mollifier ψ. Then $\operatorname{supp} \chi_\varepsilon \subset F_{3\varepsilon}$ and $\chi_\varepsilon = 1$ on F_ε.

Theorem 5.13. *[7, Theorem on p.102] Let $u \in \mathcal{D}'(\Omega)$ be a distribution of order k. Suppose that φ is a testfunction for which*

$$\partial^\alpha \varphi = 0 \text{ on } \operatorname{supp} u \text{ for all } \alpha \in \mathbb{N}_0 \text{ with } |\alpha| \leq k.$$

Then $u(\varphi) = 0$.

Proof. For convenience we assume $\Omega = \mathbb{R}^d$ for this proof. Instead one can interpret all the sets and functions appearing to be the restrictions to Ω. Let $F = \operatorname{supp} u$. For $\varepsilon > 0$ let $\chi_\varepsilon = \mathbb{1}_{F_{2\varepsilon}} * \psi_\varepsilon$ for a positive mollifier ψ as in 5.6, so that $\operatorname{supp} \chi_\varepsilon \subset F_{3\varepsilon}$ and $\chi_\varepsilon = 1$ on F_ε. Hence, as we have seen in 5.6, $u(\varphi) = u(\chi_\varepsilon \varphi)$ and $\operatorname{supp}(\chi_\varepsilon \varphi) \subset F_{3\varepsilon}$ and $\chi_\varepsilon \varphi = 0$ on F. Let $K = \operatorname{supp} \varphi$ and let $C > 0$ and $k \in \mathbb{N}_0$ be such that (4) holds, which implies that

$$|u(\varphi)| = |u(\chi_\varepsilon \varphi)| \leq C \|\chi_\varepsilon \varphi\|_{C^k} \qquad (\varepsilon > 0).$$

We show that $\|\chi_\varepsilon \varphi\|_{C^k} \xrightarrow{\varepsilon \downarrow 0} 0$. By Young's inequality we have

$$\|\chi_\varepsilon\|_{L^\infty} \leq \|\mathbb{1}_{F_{2\varepsilon}}\|_{L^\infty} \|\psi_\varepsilon\|_{L^1} = \|\psi\|_{L^1} = 1,$$

and moreover, with $C_1 = \sum_{\alpha \in \mathbb{N}_0^d : |\alpha| \leq k} \|\partial^\alpha \psi\|_{L^1}$, because $\psi_\varepsilon(x) = \varepsilon^{-d} \psi(\frac{x}{\varepsilon})$,

$$\|\partial^\alpha \chi_\varepsilon\|_{L^\infty} \leq \|\mathbb{1}_{F_{2\varepsilon}}\|_{L^\infty} \|\partial^\alpha \psi_\varepsilon\|_{L^1} \leq C_1 \varepsilon^{-|\alpha|}.$$

By 5.11 there exists a $C_2 > 0$ such that for all $x \in F_{3\varepsilon}$ and $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \leq k$

$$|\partial^\alpha \varphi(x)| \leq C_2 \varepsilon^{k+1-|\alpha|}.$$

Therefore by Leibniz' rule 5.3 we obtain for all $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \leq k$ and $x \in F_{4\varepsilon}$

$$\begin{aligned}
|\partial^\alpha (\chi_\varepsilon \varphi)(x)| &\leq \sum_{\substack{\beta \in \mathbb{N}_0^d \\ \beta \leq \alpha}} \binom{\alpha}{\beta} |\partial^\beta \chi_\varepsilon(x)| |\partial^{\alpha-\beta} \varphi(x)| \\
&\leq \sum_{\substack{\beta \in \mathbb{N}_0^d \\ \beta \leq \alpha}} \binom{\alpha}{\beta} C_1 \varepsilon^{-|\beta|} C_2 \varepsilon^{k+1-|\alpha-\beta|} \leq C_1 C_2 \Big(\sum_{\substack{\beta \in \mathbb{N}_0^d \\ \beta \leq \alpha}} \binom{\alpha}{\beta} \Big) \varepsilon.
\end{aligned}$$

Therefore $\|\chi_\varepsilon \varphi\|_{C^k} \xrightarrow{\varepsilon \downarrow 0} 0$ and thus $u(\varphi) = 0$. $\square$

Corollary 5.14. *[7, Corollary on p.103] If u is a distribution supported by $\{x\}$, then $u = \sum_{\alpha \in \mathbb{N}_0^d : |\alpha| \le k} c_\alpha \partial^\alpha \delta_x$ for some $k \in \mathbb{N}_0$ and $c_\alpha \in \mathbb{R}$.*

Moreover, $c_\alpha = \langle \iota^{-1}(u), \boldsymbol{x}^\alpha \rangle$, where $\boldsymbol{x} : x \mapsto x$ (and thus $\boldsymbol{x}^\alpha : x \mapsto x^\alpha$) and with ι as in Theorem 5.7.

Proof. By taking a translation of the distribution, we may as well assume that $x = 0$. Let $\varepsilon > 0$ be such that $B(0, \varepsilon) \subset \Omega$. By Taylor's formula (see Theorem C.7) $\varphi = P + \psi$ on $B(0, \varepsilon)$, for a polynomial P of order k given by

$$P(y) = \sum_{\substack{\alpha \in \mathbb{N}_0^d \\ |\alpha| \le k}} \frac{1}{\alpha!} \partial^\alpha \varphi(0) y^\alpha,$$

and ψ satisfying $\partial^\alpha \psi(0) = 0$ for all $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \le k$. Let χ be a testfunction that equals 1 on $B(0, \frac{\varepsilon}{2})$ and has support within $B(0, \varepsilon)$. Then $u(\varphi) = u(\chi\varphi) = u(P\chi)$ by the previous theorem. And thus,

$$u(\varphi) = u(P\chi) = \sum_{\substack{\alpha \in \mathbb{N}_0^d \\ |\alpha| \le k}} \frac{1}{\alpha!} \partial^\alpha \varphi(0) u(\boldsymbol{x}^\alpha \chi).$$

□

Theorem 5.15. *$\mathcal{E}'(\Omega)$ is weak* sequentially complete.*

Proof. One can follow the lines in the argument as in the proof Theorem 2.13 as follows: With d being the metric on $\mathcal{E}$, one replaces "$\mathcal{D}_K$" and "d_K" by "$\mathcal{E}$" and "d" and follows the same lines. □

5.16 ($\mathcal{E}$ is metrizable but $\mathcal{E}'$ is not). Theorem 1.12 implies that the topology of $\mathcal{E}(\Omega)$ is generated by a countable number seminorms, so that $\mathcal{E}(\Omega)$ is metrizable, see for example [4, Proposition IV.2.1].

$\mathcal{E}'(\Omega)$ is not metrizable, as we will show. We show that any metric on $\mathcal{E}'(\Omega)$ generates a different topology. Suppose d is a metric on $\mathcal{E}'(\Omega)$, such that under the topology generated by d the space is a topological vector space. We mimic the idea in 2.9. Let $(x_n)_{n \in \mathbb{N}}$ be a sequence in Ω such that no subsequence of it converges in Ω, i.e., for each compact set K there are finitely many elements of the sequence in K. For all $n \in \mathbb{N}$ let $\lambda_n > 0$ be such that $d(\lambda_n \delta_{x_n}, 0) < \frac{1}{n}$. Then $\lambda_n \delta_{x_n} \to 0$. But there exists a smooth function $\psi \in \mathcal{E}$ such that $\psi(n) = \frac{1}{\lambda_n}$ (use the partition of unity 1.15), so that $\lambda_n \delta_{x_n}(\psi)$ does not converge to 0. This means that $\lambda_n \delta_{x_n}$ does not converge to 0 in $\mathcal{E}'(\Omega)$.

6 Convolutions of distributions

In this section we consider only $\Omega = \mathbb{R}^d$ and write '$\mathcal{E}$' and '$\mathcal{D}$' instead of '$\mathcal{E}(\mathbb{R}^d)$' and '$\mathcal{D}(\mathbb{R}^d)$'.

6.1. In Section 3 we have defined the convolution between distributions and testfunctions. For $u \in \mathcal{D}'$ and $\varphi \in \mathcal{D}$ we have seen that the convolution $u*\varphi$ is a smooth function, or say an element of $\mathcal{E}$, and its support is included in the sum $\operatorname{supp} u + \operatorname{supp} \varphi$. As $\operatorname{supp} \varphi$ is compact, the sum is compact as soon as u has compact support. And so $u * \varphi$ is an element of $\mathcal{D}$ if u has compact support. Let us formally define the convolution between an element of $\mathcal{E}'$ and an element of $\mathcal{E}$.

Definition 6.2. Let $u \in \mathcal{E}'$ and $\varphi \in \mathcal{E}$. We define the *convolution* of u with φ to be the function $\mathbb{R}^d \to \mathbb{F}$ defined by

$$u * \varphi(x) = u(\mathcal{T}_x \check{\varphi}) \qquad (x \in \mathbb{R}^d).$$

The arguments of 3.17 extend to $u \in \mathcal{E}'$ and $\varphi \in \mathcal{E}$ (as the topology of $\mathcal{E}$ allows us to consider compact sets only), and we obtain the following.

Lemma 6.3. *For all $u \in \mathcal{E}'$ and $\varphi \in \mathcal{E}$ we have $u * \varphi \in \mathcal{E}$ and for for $\alpha \in \mathbb{N}_0^d$*

$$\partial^\alpha (u * \varphi) = u * (\partial^\alpha \varphi) = (\partial^\alpha u) * \varphi. \tag{32}$$

So by Lemma 3.18 and Lemma 6.3 we have $u * \varphi \in \mathcal{E}$ if either $u \in \mathcal{D}'$ and $\varphi \in \mathcal{D}$ or $u \in \mathcal{E}'$ and $\varphi \in \mathcal{E}$. By Theorem 3.21 we have $u * \varphi \in \mathcal{D}$ if $u \in \mathcal{E}'$ and $\varphi \in \mathcal{E}$. We can show that convolution with a distribution is sequentially continuous:

Lemma 6.4. *Let $u \in \mathcal{D}'$ and $v \in \mathcal{E}'$.*

(a) *The function $\mathcal{D} \to \mathcal{E}$ given by $\varphi \mapsto u * \varphi$ is sequentially continuous.*

(b) *The function $\mathcal{D} \to \mathcal{D}$ given by $\varphi \mapsto v * \varphi$ is sequentially continuous.*

(c) *The function $\mathcal{E} \to \mathcal{E}$ given by $\psi \mapsto v * \psi$ is continuous.*

Proof. (a) and (b) are left as an exercise, see Exercise 6.1. Let us only mention that (b) follows from (c) by Theorem 2.7.

Let K be a compact set, $C > 0$ and $k \in \mathbb{N}$ be such that

$$|v(\psi)| \le C \|\psi\|_{C^k, K} \qquad (\psi \in \mathcal{E}).$$

Let M be an arbitrary compact subset of $\mathbb{R}^d$ and let $m \in \mathbb{N}$. It is sufficient to show that there exists a compact set L, an $l \in \mathbb{N}$ and a $C' > 0$ such that

$$\|v * \psi\|_{C^m, M} \le C' \|\psi\|_{C^l, L} \qquad (\psi \in \mathcal{E}).$$

We have by Lemma 6.3, with $E = \sum_{\beta \in \mathbb{N}_0^d : |\beta| \le m}$

$$\begin{aligned}
\|v * \psi\|_{C^m, M} &= \sum_{\beta \in \mathbb{N}_0^d : |\beta| \le m} \sup_{x \in M} |\partial^\beta (v * \varphi)(x)| = \sum_{\beta \in \mathbb{N}_0^d : |\beta| \le m} \sup_{x \in M} |v(\mathcal{T}_x \partial^\beta \check{\varphi})| \\
&\le C \sum_{\beta \in \mathbb{N}_0^d : |\beta| \le m} \sup_{x \in M} \|\mathcal{T}_x \partial^\beta \check{\varphi})\|_{C^k, K} = CE \sup_{x \in M} \|\mathcal{T}_x \partial^\beta \varphi\|_{C^{k+m}, K} \\
&= CE \|\partial^\beta \check{\varphi})\|_{C^{k+m}, K+M},
\end{aligned}$$

We choose $L = K + M$, $l = k + m$ and $C' = CE$. L is compact as $K \times M$ is compact in $\mathbb{R}^d \times \mathbb{R}^d$, as addition is a continuous function $\mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ and as the image of a compact set under a continuous function is compact. $\square$

6.5. Lemma 6.4 allows us to compose the convolution with u and the convolution with v and obtain a sequentially continuous linear map $\mathcal{D} \to \mathcal{E}$ defined by $u * (v * \varphi)$. Moreover, for

$$(u, \varphi) \in (\mathcal{D}' \times \mathcal{D}) \cup (\mathcal{E}' \times \mathcal{E}) \cup (\mathcal{E}' \times \mathcal{D})$$

and $a \in \mathbb{R}^d$ we have

$$\mathcal{T}_a(u * \varphi) = (\mathcal{T}_a u) * \varphi = u * (\mathcal{T}_a \varphi).$$

In particular, this means that the map $\varphi \mapsto u * \varphi$ commutes with translation.

Therefore also the above mentioned composition, the map $\varphi \mapsto u * (v * \varphi)$ commutes with translation. Theorem 6.6 tells us that there exists a unique distribution w such that $w * \varphi = u * (v * \varphi)$.

We will show that with the definition of $u * v$ in Definition 6.9 we have $w = u * v$.

Exercise **6.1.** Prove Lemma 6.4 (a) and (b).

Theorem 6.6. [7, Theorem on page 121] *Let A be a linear map $\mathcal{D} \to \mathcal{E}$ which commutes with translation, i.e., $\mathcal{T}_a(A\varphi) - A(\mathcal{T}_a\varphi)$ for all $a \in \mathbb{R}^d$, and which is sequentially continuous, then there exists a unique distribution u such that $A\varphi = u * \varphi$ for all $\varphi \in \mathcal{D}$.*

Theorem 6.7. [7, Corollary on page 122] *Every linear map $\mathcal{E} \to \mathcal{E}$ which is sequentially continuous and commutes with translation is of the form $A\varphi = u * \varphi$ for some uniquely determined distribution u with compact support.*

6.8. Remember (17) in Corollary 3.5, which tells us that for integrable f and g and a testfunction φ we have (by viewing $f * g$ as a distribution and $\langle u, v \rangle = \int uv$)

$$\langle g * f, h \rangle = \langle f * g, h \rangle = \langle g, \check{f} * h \rangle.$$

This shows that the definition of $u * v$ as given below extends this relation.

Definition 6.9. For $u \in \mathcal{D}'$ and $v \in \mathcal{E}'$ we define $u * v$ to be the distribution given by

$$u * v(\varphi) = u(\check{v} * \varphi) \qquad (\varphi \in \mathcal{D}).$$

Moreover, we define $v * u$ to be the distribution

$$v * u(\varphi) = v(\check{u} * \varphi) \qquad (\varphi \in \mathcal{D}).$$

6.10. Observe that for $(u, v) \in (\mathcal{D}' \times \mathcal{E}') \cup (\mathcal{E}' \times \mathcal{D}')$, $\varphi \in \mathcal{D}$ and $x \in \mathbb{R}^d$ we have

$$(u * v) * \varphi(x) = u * v(\mathcal{T}_x \check{\varphi}) = u(\check{v} * \mathcal{T}_x \check{\varphi}) = u(\mathcal{T}_x (v * \varphi)\check{}) = u * (v * \varphi)(x).$$

We will now show that $u * v = v * u$. As

$$u * v(\varphi) = (u * v) * \check{\varphi}(0) = u * (v * \check{\varphi})(0),$$

it is a consequence of the identity (34) in Theorem 6.12. We will need to extend the associativity in Theorem 3.24 first.

Theorem 6.11. *Let $v \in \mathcal{E}'$, $\varphi \in \mathcal{D}$ and $\eta \in \mathcal{E}$. Then*

$$v * (\varphi * \eta) = (v * \varphi) * \eta = (v * \eta) * \varphi \tag{33}$$

Proof. If $\eta \in \mathcal{D}$, then this follows directly from Theorem 3.24. By Lemma 6.4 the functions $\mathcal{E} \mapsto \mathcal{E}$ given by $\eta \mapsto v * (\varphi * \eta)$, $\eta \mapsto (v * \varphi) * \eta$ and $(v * \eta) * \varphi$ are sequentially continuous (for the last, observe that the inclusion $\mathcal{D} \to \mathcal{E}'$ is sequentially continuous). Therefore, as $\mathcal{D}$ is dense in $\mathcal{E}$ (see Theorem 5.7), we obtain (33) as a consequence Theorem 6.11 by a limiting argument. □

Theorem 6.12. *Let $u \in \mathcal{D}'$, $v \in \mathcal{E}'$ and $\varphi \in \mathcal{D}$. Then*

$$u * (v * \varphi) = v * (u * \varphi). \tag{34}$$

Consequently, $u * v = v * u$.

Proof. Let also $\psi \in \mathcal{D}$. Then by Theorem 3.24, Theorem 6.11 and by the commutativity of convolution of functions (see 3.2)

$$\begin{aligned}
(u * (v * \varphi)) * \psi &= u * ((v * \varphi) * \psi) = u * (\psi * (v * \varphi)) = (u * \psi) * (v * \varphi) \\
&= (v * \varphi) * (u * \psi) = v * (\varphi * (u * \psi)) = v * ((u * \psi) * \varphi) \\
&= v * (u * (\psi * \varphi)) = v * (u * (\varphi * \psi)) = v * ((u * \varphi) * \psi) \\
&= v * ((u * \varphi) * \psi) = (v * (u * \varphi)) * \psi.
\end{aligned}$$

By taking ψ a mollifier, from the above identity by a limiting argument one obtains (34) (using Theorem 4.5). □

6.13. We will from now on write $\varphi * u$ for the function $u * \varphi$ for $u \in \mathcal{D}'(\mathbb{R}^d)$ and $\varphi \in \mathcal{D}(\mathbb{R}^d)$.

Theorem 6.14. *For $u \in \mathcal{D}'$ and $v \in \mathcal{E}'$*

$$\operatorname{supp} u * v \subset \operatorname{supp} u + \operatorname{supp} v.$$

Proof. Let $x \in \operatorname{supp} u * v$. For all $\varepsilon > 0$ there exists a $\varphi \in \mathcal{D}$ supported in $B(x, \varepsilon)$ such that $u * v(\varphi) \neq 0$, i.e., $u(\check{v} * \varphi) \neq 0$. Therefore $\operatorname{supp} u \cap (\operatorname{supp} \check{v} * \varphi) \neq \emptyset$. Let y be in this intersection. By Theorem 3.21 we know that there exists a $z \in \operatorname{supp} v$ and $w \in \operatorname{supp} \varphi$ such that $y = -z + w$. Then $w = y + z \in \operatorname{supp} u + \operatorname{supp} v$ and $|x - w| < \varepsilon$. As we can find such w for each ε and $\operatorname{supp} u + \operatorname{supp} v$ is closed, we conclude that $x \in \operatorname{supp} u + \operatorname{supp} v$. □

Remark 6.15. One can also define the convolution of two distributions, where instead of assuming that one of the two has compact support the map $\Sigma : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$, $\Sigma(x, y) = x + y$ is proper on $\operatorname{supp} u \times \operatorname{supp} v$, meaning that $\Sigma^{-1}(K) \cap \operatorname{supp} u \times \operatorname{supp} v$ is a compact subset of $\mathbb{R}^d \times \mathbb{R}^d$ for all compact sets $K \subset \mathbb{R}^d$. The details can be found for example in [8, Section 11].

7 Fundamental solutions of PDEs

Definition 7.1. We call an map $P : \mathcal{D}'(\Omega) \to \mathcal{D}'(\Omega)$ a *linear partial differential operator* in $\mathbb{R}^d$ with constant coefficients if there exists an $m \in \mathbb{N}$ and $c_\alpha \in \mathbb{F}$ for $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \le m$ such that

$$P = \sum_{\substack{\alpha \in \mathbb{N}_0^d \\ |\alpha| \le m}} c_\alpha \partial^\alpha.$$

Often, the following notation is also used. When we take $p : \mathbb{R}^d \to \mathbb{F}$ the polynomial

$$p(x) = \sum_{\substack{\alpha \in \mathbb{N}_0^d \\ |\alpha| \le m}} c_\alpha x^\alpha,$$

then it is common to write '$p(\partial)$' for 'P', so that one interpret $p(\partial)$ as the formal polynomial evaluated in ∂. One also uses 'D' instead of '∂' in literature, so that one writes '$p(\mathrm{D})$' for 'P'.

One says a distribution E is called a *fundamental solution* of P if $PE = \delta$ (the Dirac measure at zero).

Fundamental solutions can help to find distributional solutions to partial differential equations of the form $Pu = v$ as the following theorem illustrates.

Theorem 7.2. *[8, Theorem 12.2] Let P be a linear partial differential operator with constant coefficients and E a fundamental solution of P. For all $v \in \mathcal{E}'(\mathbb{R}^d)$ we have*

$$P(E * v) = v = E * (Pv).$$

Proof. This follows by the fact that $\partial^\alpha(E * v) = (\partial^\alpha E) * v = E * (\partial^\alpha v)$. $\square$

7.3. Observe that if E is a fundamental solution of a linear partial differential operator with constant coefficients P, and if $u \in \mathcal{D}'(\mathbb{R}^d)$ satisfies $Pu = 0$, then $E + u$ is also a fundamental solution of P.

Let us consider the example where we consider P to be the Laplacian Δ (which equals $\sum_{i=1}^d \partial_i^2$).

Example 7.4. Let E be the function on $\mathbb{R}^d$ (for $d \ge 2$) defined by $E(0) = 0$ and

$$E(x) = \begin{cases} \frac{1}{(2-d)V_d} |x|^{2-d} & d \ne 2, \\ \frac{1}{2\pi} \log |x| & d = 2, \end{cases} \tag{35}$$

where V_d is the $n-1$ dimensional volume of the sphere $\{x \in \mathbb{R}^d : |x| = 1\}$ (observe that $2\pi = V_2$). Then E is the fundamental solution of Δ (see Exercise 7.1).

$\boxed{\textbf{Exercise}}$ **7.1.** (a) [8, Problem 4.5] For $i \in \{1, \dots, d\}$ let v_i be the function on $\mathbb{R}^d$ defined by $v_i(0) = 0$ and

$$v_i(x) = \frac{x_i}{|x|^d} \qquad (x \in \mathbb{R}^d \setminus \{0\}).$$

Prove that v_i is locally integrable on $\mathbb{R}^d$ and that in $\mathcal{D}'$

$$\sum_{i=1}^{d} \partial_i v_i = V_d \delta,$$

where V_d is the $d-1$ dimensional volume of the sphere $\{x \in \mathbb{R}^d : |x| = 1\}$. (Hint: Observe that $\langle \partial_i v_i, \varphi \rangle = -\lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d \setminus B(0,\varepsilon)} v_i \partial_i \varphi$ and apply integration by parts (see Theorem E.1).)

(b) [8, Problem 4.6] Prove that E as in (35) is locally integrable on $\mathbb{R}^d$ and that E is the fundamental solution of Δ, i.e., $\Delta E = \delta$ (first you might want to prove that $\partial_i E = c v_i$ for some $c \in \mathbb{R}$).

7.5. With E being the fundamental solution to Δ as defined in (35), we conclude that for $v \in \mathcal{E}'(\mathbb{R}^d)$ we have a solution to the *Poisson equation*

$$\Delta u = v,$$

given by $E * v$.

Definition 7.6. A function $f \in C^2(\Omega)$ is called *harmonic*, or an *harmonic function* if $\Delta f = 0$. A distribution $u \in \mathcal{D}'(\Omega)$ is called *harmonic* if $\Delta u = 0$.

$\boxed{\textbf{Exercise}}$ **7.2.** For $\mathbb{F} = \mathbb{C}$ and $d \geq 2$, check that for all $k \in \mathbb{N}_0$ the polynomial $x \mapsto (x_1 + \mathrm{i} x_2)^k$ is harmonic.

For $d = 1$, observe that for $f \in C^2(\Omega)$, $\Delta f = 0$ if and only if $f(x) = a + bx$ for some $a, b \in \mathbb{F}$.

As is mentioned in 7.3 for any harmonic distribution u we have that $E + u$ is a fundamental solution of Δ. We will prove that any harmonic distribution is actually (represented by) a harmonic function in $C^\infty(\mathbb{R}^d)$. This statement is called Weyl's theorem. We prove a generalisation, for which we introduce the singular support, which indicates "where a distribution is smooth".

Definition 7.7 (Singular support). For a distribution u in $\mathcal{D}'(\Omega)$ we define the *singular support* as those points at which there exists no neighbourhood of that point on which the distribution is (represented by) a smooth function, for which we write $\operatorname{sing\,supp} u$, so that

$$\begin{aligned} \operatorname{sing\,supp} u &= \{x \in \Omega : \forall U \in \mathcal{N}_x \; [\![u|_U \notin C^\infty(U)]\!]\} \\ &= \Omega \setminus \{x \in \Omega : \exists U \in \mathcal{N}_x \; [\![u|_U \in C^\infty(U)]\!]\}, \end{aligned}$$

where $u|_U$ is written for the element $\mathcal{D}'(U)$ given by $\rho(u)$, with ρ as in 2.18.

Observe that $\operatorname{sing\,supp} u \subset \operatorname{supp} u$, which basically means that "where u equals zero it is smooth". The singular support satisfies the same rule as the support does for convolutions:

Lemma 7.8. *[8, Theorem 11.16] Let $u \in \mathcal{D}'(\mathbb{R}^d)$ and $v \in \mathcal{E}'(\mathbb{R}^d)$. Then*

$$\operatorname{sing\,supp} u * v \subset \operatorname{sing\,supp} u + \operatorname{sing\,supp} v. \tag{36}$$

Proof. Let us write A for $\operatorname{sing\,supp} u$ and B for $\operatorname{sing\,supp} v$. Let $\delta > 0$ (be such that $B_\delta \subset \Omega$) and $\chi \in \mathcal{D}(\Omega)$ be such that χ is equal to 1 on $A_{\frac{\delta}{2}}$ and 0 outside A_δ (see 5.12). Then $u_2 := (1 - \chi_A)u$ is (represented by) a smooth function and so $u = u_1 + u_2$ for $u_1 = \chi_A u$, and $\operatorname{supp} u_1 \subset A_\delta$. Similarly, we can write $v = v_1 + v_2$, where $\operatorname{supp} v_1 \subset B_\delta$ and v_2 is (represented by) a smooth function. Then

$$u * v = u_1 * v_1 + u_1 * v_2 + u_2 * v_1 + u_2 * v_2.$$

The last three terms are smooth (by Lemma 3.18 and Lemma 6.3), and the support of $u_1 * v_1$ is included in $A_\delta + B_\delta$ (Theorem 6.14), which in turn is included in $(A + B)_{2\delta}$. Therefore

$$\operatorname{sing\,supp} u * v \subset (A + B)_\delta.$$

As δ is chosen arbitrarily and the set $A+B$ is closed (see Lemma 3.7), we have $\bigcap_{\delta>0}(A+B)_\delta = A + B$ and conclude (36). □

We will now consider a generalisation of fundamental solutions, in the sense that we consider distributions that are "a fundamental solution modulo a smooth function".

Definition 7.9. Let P be a linear partial differential operator with constant coefficients. A distribution E is called a *parametrix* of P if there exists a $\psi \in \mathcal{E}(\mathbb{R}^d)$ such that $PE = \delta + \psi$.

Theorem 7.10. *[8, Theorem 12.4] Let P be a linear partial differential operator with constant coefficients. Suppose E is a parametrix of P with $\operatorname{sing\,supp} E = \{0\}$. Then for all open $\Omega \subset \mathbb{R}^d$*

$$\operatorname{sing\,supp} u = \operatorname{sing\,supp} Pu \qquad (u \in \mathcal{D}'(\Omega)). \tag{37}$$

Proof. Similar to 3.20 we have $\operatorname{sing\,supp} Pu \subset \operatorname{sing\,supp} u$, which basically means that 'Pu is smooth where u is'.

First suppose that u has compact support, so that we may assume $u \in \mathcal{E}'(\Omega)$. Let $\psi \in \mathcal{E}(\mathbb{R}^d)$ be such that $PE = \delta + \psi$. Then

$$E * (Pu) = (PE) * u = (\delta + \psi) * u = u + \psi * u.$$

Therefore $\operatorname{sing\,supp} u = \operatorname{sing\,supp} E * (Pu)$ as $\psi * u \in \mathcal{E}(\mathbb{R}^d)$ by Lemma 6.3. Therefore, by Lemma 7.8

$$\operatorname{sing\,supp} u \subset \operatorname{sing\,supp} E + \operatorname{sing\,supp} Pu = \operatorname{sing\,supp} Pu,$$

as $\operatorname{sing\,supp} E = \{0\}$.

Let $x \in \Omega \setminus \operatorname{sing\,supp} Pu$. Let $\chi \in \mathcal{D}(\Omega)$ be equal to 1 on an open neighbourhood U of x. Then $P(\chi u) = Pu$ on U (by which we mean that the restrictions to $\mathcal{D}(U)$ as in 2.18 are the same). Therefore $x \in \Omega \setminus \operatorname{sing\,supp} P(\chi u)$. As χu has compact support we have $\operatorname{sing\,supp} P(\chi u) = \operatorname{sing\,supp}(\chi u)$ and so χu is smooth on a neighbourhood of x, and so is u as $\chi u = u$ on U. So $x \in \Omega \setminus \operatorname{sing\,supp} u$ from which we conclude (37). □

7.11. Let P and E are as in Theorem 7.10. This theorem tells us that the solution u to $Pu = v$ for a $v \in \mathcal{E}'(\mathbb{R}^d)$ is smooth where v is, in the sense that if U is open and $v|_{\mathcal{D}(U)}$ is smooth, then $u|_U$ is smooth. Therefore, in particular we obtain Weyl's theorem as a consequence.

Theorem 7.12 (Weyl's Theorem). [7, Page 127]
Every harmonic distribution is (represented by) a smooth harmonic function.

Example 7.13. For $t > 0$ we define the function $h_t : \mathbb{R}^d \to \mathbb{R}$ by

$$h_t(x) = (4\pi t)^{-\frac{d}{2}} e^{-\frac{1}{4t}|x|^2} \qquad (x \in \mathbb{R}^d). \tag{38}$$

Then (see Exerise 7.3) it solves the *heat equation* on $(0, \infty) \times \mathbb{R}^d$:

$$\partial_t h_t(x) = \Delta_x h_t(x) \qquad ((t, x) \in (0, \infty) \times \mathbb{R}^d), \tag{39}$$

where Δ_x denotes the Laplacian acting on the x variable(s) only, i.e., $\Delta_x = \sum_{i=1}^d \partial_{x_i}^2$.

Exercise **7.3.** Show that (39) is satisfied for h_t as in (38).

7.14. Observe that

$$\int_{\mathbb{R}^d} h_t(x) \, \mathrm{d}x = \left(\int_{\mathbb{R}} (4\pi t)^{-\frac{1}{2}} e^{-\frac{1}{4t}s^2} \, \mathrm{d}s \right)^d = 1,$$

which follows by the fact that

$$\int_{\mathbb{R}} e^{-x^2} \, \mathrm{d}x = \sqrt{\pi}.$$

This can be proved using polar coordinates:

$$\begin{aligned} \left(\int_{\mathbb{R}} e^{-x^2} \, \mathrm{d}x \right)^2 &= \int_{\mathbb{R}} \int_{\mathbb{R}} e^{-(x^2+y^2)} \, \mathrm{d}x \, \mathrm{d}y = 2\pi \int_0^\infty r e^{-r^2} \, \mathrm{d}r \\ &= 2\pi \int_0^\infty \frac{1}{2} e^{-s} \, \mathrm{d}s = \pi. \end{aligned}$$

From this we can show that

$$\langle h_t, \varphi \rangle \xrightarrow{t \downarrow 0} \varphi(0) \qquad (\varphi \in C_{\mathrm{b}}(\mathbb{R}^d)). \tag{40}$$

Indeed,

$$\langle h_t, \varphi \rangle - \varphi(0) = \int_{\mathbb{R}^d} h_t(x)(\varphi(x) - \varphi(0)) \, \mathrm{d}x.$$

By a substitution $y = \frac{x}{\sqrt{t}}$ we have

$$\int_{\mathbb{R}^d} h_t(x)(\varphi(x) - \varphi(0)) \, \mathrm{d}x = \int_{\mathbb{R}^d} h_1(y)(\varphi(\sqrt{t}y) - \varphi(0)) \, \mathrm{d}y.$$

So that by the Lebesgue dominated convergence theorem we indeed obtain (40).

$\boxed{\textbf{Exercise}}$ **7.4.** Calculate the limit in $\mathcal{D}'(\mathbb{R}^d)$ of $\partial_t h_t$ as $t \downarrow 0$.

Example 7.15. Define $E : \mathbb{R}^{d+1} \to \mathbb{R}$ by

$$E(t,x) = \begin{cases} h_t(x) & (t,x) \in (0,\infty) \times \mathbb{R}^d, \\ 0 & (t,x) \in (-\infty, 0] \times \mathbb{R}^d. \end{cases}$$

Then (see Exercise 7.5) E is a fundamental solution of $\partial_t - \Delta_x$ (one also says, E is a fundamental solution of the heat equation).

Definition 7.16. The *gamma function* is the function $\Gamma : (0,\infty) \to (0,\infty)$ given by

$$\Gamma(s) = \int_0^\infty t^{s-1} e^{-t} \, \mathrm{d}t \qquad (s \in (0,\infty)).$$

It is sometimes also defined on the complex plane for those numbers for which the real part is strictly positive. By partial integration it follows that $\Gamma(s+1) = s\Gamma(s)$. Therefore $\Gamma(n) = (n-1)!$. Moreover, $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

$\boxed{\textbf{Exercise}}$ **7.5.** (a) Calculate $\int_0^\infty h_t(x) \, \mathrm{d}t$ for $x \neq 0$ (in terms of the gamma function).

(b) Show that $\lim_{t\downarrow 0} \int_{\mathbb{R}^d} h_t(x)\varphi(t,x) \, \mathrm{d}x = \varphi(0)$ for any $\varphi \in \mathcal{D}(\mathbb{R}^{d+1})$.

(c) Show that E is locally integrable.

(d) Calculate $\operatorname{sing\,supp} E$.

(e) Calculate $\operatorname{supp}(\partial_t - \Delta_x)E$.

(f) Estimate the order of $(\partial_t - \Delta_x)E$.

(g) Show that E is a fundamental solution of $\partial_t - \Delta_x$ (Hint: Observe that $\langle (\partial_t - \Delta)E, \varphi \rangle = \lim_{T\uparrow\infty, s\downarrow 0} - \int_s^T \int_{\mathbb{R}^d} h_t(x)(\partial_t + \Delta_x)\varphi(t,x) \, \mathrm{d}x \, \mathrm{d}t$ and apply integration by parts.)

(h) Conclude that if $v \in \mathcal{E}'(\mathbb{R}^{d+1})$ is smooth on an open set U, then so is the solution u of $(\partial_t - \Delta)u = v$.

Remark 7.17. In [8, Section 12] one finds references for the proof of the statement that every linear partial differential operator with constant coefficients, of which at least one coefficient is nonzero, has a fundamental solution.

8 Sobolev spaces

In Theorem 2.15 we have seen that L^1_{loc} is continuously embedded in $\mathcal{D}'$. This implies that L^p is continuously embedded in $\mathcal{D}'$, as L^p is continuously embedded in L^1_{loc}: This follows by Hölder's inequality, which implies that for all $p, q \in [1, \infty]$ such that $\frac{1}{p} + \frac{1}{q} = 1$ one has $\|f\|_{L^1,K} \leq \|f\|_{L^p} \|1_K\|_{L^q}$ for $f \in L^p$ and any compact set K.

In this section we will consider Sobolev spaces as subspaces of $\mathcal{D}'$. These spaces are subsets of L^p for which not only the function itself, but also its derivatives (in the distributional sense) up to a certain order are all included in L^p.

Definition 8.1. Let $p \in [1, \infty]$ and $k \in \mathbb{N}_0$. We define the *Sobolev space* of order k and integrability p, denoted $W^{k,p}(\Omega)$, by

$$W^{k,p}(\Omega) = \{u \in \mathcal{D}'(\Omega) : \partial^\beta u \in L^p(\Omega) \text{ for all } \beta \in \mathbb{N}_0^d \text{ with } |\beta| \leq k\}.$$

In some literature, for example in [9], the definition of a Sobolev space looks a bit different and does not use the language of distributions. In that case, the ∂^β is interpreted as the weak derivative and "$u \in \mathcal{D}'(\Omega)$" is replaced by "$u \in L^p(\Omega)$" and the part "$\partial^\beta \in L^p(\Omega)$" instead reads somehow like "$\partial^\beta u$ exists (as a weak derivative) and is in $L^p(\Omega)$". Let us give the definition of such weak derivatives.

Definition 8.2. Let $u, v \in L^1_{\mathrm{loc}}(\Omega)$ and $\alpha \in \mathbb{N}_0^d$. v is called the α-th *weak partial derivative* of u if $v = \partial^\alpha u$ in the distributional sense, i.e., if

$$\int v\varphi = \int u \cdot (-1)^{|\alpha|} \partial^\alpha \varphi \qquad (\varphi \in C_c^\infty(\Omega)).$$

Lemma 8.3. *Let $u \in L^1_{\mathrm{loc}}(\Omega)$. If u has an α-th weak partial derivative, then it is unique.*

Proof. This is a consequence of Theorem 1.16. □

$\boxed{\textbf{Exercise}}$ **8.1.** Consider $\Omega = (0, 2)$, $u, v \in L^1_{\mathrm{loc}}(\Omega)$ given by

$$u(x) = \begin{cases} x & x \in (0, 1], \\ 1 & x \in (1, 2), \end{cases} \qquad v(x) = \begin{cases} x & x \in (0, 1], \\ 2 & x \in (1, 2). \end{cases}$$

(a) Show that u has a weak derivative that is in L^p, so that $u \in W^{1,p}(\Omega)$ for all $p \in [1, \infty]$.

(b) Show that v has no weak derivative, but calculate its distributional derivative.

(c) Give an example of an element $u \in W^{1,p}(0, 2)$ such that the function v defined on $\mathbb{R}$ by $v(x) = u(x)$ for $x \in (0, 2)$ and $v(x) = 0$ for other x, is not in $W^{1,p}(\mathbb{R})$.

Definition 8.4. We equip the Sobolev space $W^{k,p}(\Omega)$ for $p \in [1, \infty)$ with the norm

$$\|u\|_{W^{k,p}} = \Big(\sum_{\beta \in \mathbb{N}_0^d, |\beta| \leq k} \|\partial^\beta u\|_{L^p}^p \Big)^{\frac{1}{p}},$$

and for $p = \infty$ with the norm

$$\|u\|_{W^{k,\infty}} = \max_{\beta \in \mathbb{N}_0^d, |\beta| \le k} \|\partial^\beta u\|_{L^\infty}.$$

$\boxed{\textbf{Exercise}}$ **8.2.** Verify that $\|\cdot\|_{W^{k,p}}$ indeed defines a norm on $W^{k,p}(\Omega)$.

Definition 8.5. Let $\mathfrak{X}$ be a normed space and $\|\cdot\|_1, \|\cdot\|_2 : X \to [0,\infty)$ be norms on $\mathfrak{X}$. They are said to be *equivalent* if they define the same topology.

Two norms $\|\cdot\|_1, \|\cdot\|_2$ are equivalent if and only if (see for example [4, Proposition III.1.5]) there exists $c, C > 0$ such that

$$c\|f\|_1 \le \|f\|_2 \le C\|f\|_1 \qquad (f \in \mathfrak{X}).$$

8.6. As any two norms on $\mathbb{R}^d$ (as it is a finite dimensional normed space, see [4, Theorem III.3.1]) are equivalent (observe that $x \mapsto \sum_{i=1}^d |x_i|$, $x \mapsto (\sum_{i=1}^d |x_i|^p)^{\frac{1}{p}}$ and $x \mapsto \max_{i=1}^d |x_i|$ are norms on $\mathbb{R}^d$), the following functions are norms that are equivalent to $\|\cdot\|_{W^{k,p}}$ for any $k \in \mathbb{N}_0$ and $p \in [1,\infty]$

$$u \mapsto \max_{\beta \in \mathbb{N}_0^d, |\beta| \le k} \|\partial^\beta u\|_{L^p}, \qquad u \mapsto \sum_{\beta \in \mathbb{N}_0^d, |\beta| \le k} \|\partial^\beta u\|_{L^p}.$$

The order of the Sobolev space determines the 'regularity' in the same way that the k of C^k does; as also here taking a derivative ∂^α decreases the order by $|\alpha|$:

Theorem 8.7. [9, p.247, Theorem 1(i),(iii)] Let $u \in W^{k,p}(\Omega)$. Then

(a) $\partial^\alpha u \in W^{k-|\alpha|,p}$ for all $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \le k$ and $\partial^\beta(\partial^\alpha u) = \partial^{\alpha+\beta} u$ for all $\alpha, \beta \in \mathbb{N}_0^d$ with $|\alpha| + |\beta| \le k$.

(b) If U is an open subset of Ω, then $u|_U \in W^{k,p}(U)$.

Proof. We leave this for the reader to verify. □

Observe that $W^{0,p}(\Omega) = L^p(\Omega)$, so that the Sobolev space of 0-th order is a Banach space. This extends to any order:

Theorem 8.8. *[9, p.249, Theorem 2] For all $p \in [1,\infty]$ and $k \in \mathbb{N}_0$, $W^{k,p}(\Omega)$ is a Banach space.*

Proof. Suppose that $(u_n)_{n \in \mathbb{N}}$ is a Cauchy sequence in $W^{k,p}(\Omega)$. Then $(\partial^\alpha u_n)_{n \in \mathbb{N}}$ is a Cauchy sequence in $L^p(\Omega)$ for all $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \le k$. As $L^p(\Omega)$ is a Banach space, there exist $u^{(\alpha)} \in L^p(\Omega)$ such that $\partial^\alpha u_n \to u^{(\alpha)}$ in L^p for all such α.

Let us write u for $u^{(0)}$. We are finished by showing that $\partial^\alpha u = u^{(\alpha)}$ for all such α, as this implies $u_n \to u$ in $W^{k,p}(\Omega)$. This follows by testing against a testfunction φ, using; if $f_n \to f$ in L^p, then $\int f_n \varphi \to \int f \varphi$ (which follows by Hölder's inequality):

$$\langle \partial^\alpha u, \varphi \rangle = \int u \cdot (-1)^{|\alpha|} \partial^\alpha \varphi = \lim_{n \to \infty} \int u_n \cdot (-1)^{|\alpha|} \partial^\alpha \varphi = \lim_{n \to \infty} \int \partial^\alpha u_n \cdot \varphi = \langle u^{(\alpha)}, \varphi \rangle.$$

As this holds for all $\varphi \in C_c^\infty(\Omega)$, we have $\partial^\alpha u = u^{(\alpha)}$ (by Theorem 1.16). □

Definition 8.9. Let $k \in \mathbb{N}_0$ and $p \in [1, \infty]$.

(a) We write $W_0^{k,p}(\Omega)$ for the closure of $C_c^\infty(\Omega)$ in $W^{k,p}(\Omega)$.

(b) We write $H^k(\Omega) = W^{k,2}(\Omega)$, $\|\cdot\|_{H^k} = \|\cdot\|_{W^{k,2}}$ and $H_0^k(\Omega) = W_0^{k,2}(\Omega)$.

Remark 8.10. One interprets $W_0^{k,p}(\Omega)$ as the subspace of $W^{k,p}(\Omega)$ of elements that vanish at the boundary of Ω, in symbols; $u = 0$ on $\partial\Omega$.

Similar to Theorem 8.8, in which we showed that $W^{k,p}$ is a Banach space by using that L^p is a Banach space, one can show that H^k is a Hilbert space because L^2 is:

Theorem 8.11. *Let $k \in \mathbb{N}_0$. $\langle \cdot, \cdot \rangle_{H^k} : H^k(\Omega) \times H^k(\Omega) \to \mathbb{F}$ defined by*

$$\langle u, v \rangle_{H^k} = \sum_{\alpha \in \mathbb{N}_0^d : |\alpha| \le k} \langle \partial^\alpha u, \partial^\alpha v \rangle_{L^2},$$

is an inner product on $H^k(\Omega)$, so that $H^k(\Omega)$ (and $H_0^k(\Omega)$) equipped with this inner product is a Hilbert space.

Proof. We leave it for the reader to check that $\langle \cdot, \cdot \rangle_{H^k}$ defines an inner product. The rest follows from Theorem 8.8 and because $\langle u, u \rangle_{H^k} = \|u\|_{H^k}^2$ for $u \in H^k(\Omega)$. □

There is a lot of theory on Sobolev spaces, which we will not treat here. Sobolev spaces play a central role in the theory of partial differential equations, and we still want to show one application of the theory. One classical reference for PDE theory, which contains a whole section on Sobolev spaces is [9] (see Section 5). There are different estimates that are useful, of which we present one important example; the Poincaré inequality.

Theorem 8.12. *[18, Theorem 12.17] Let Ω be a bounded open subset of $\mathbb{R}^d$. There for all $p \in [1, \infty)$ exists a $C > 0$ such that*

$$\|u\|_{L^p} \le C \|\nabla u\|_{L^p} \qquad (u \in W_0^{1,p}(\Omega)).$$

8.13. Consider the context of Theorem 8.12. As for functions $f : \mathbb{R}^d \to \mathbb{R}^d$ we interpret $\|f\|_{L^p} = (\int |f|^p)^{\frac{1}{p}}$, with $|f|$ being the composition of the euclidean norm with the function f, we have

$$\|\nabla u\|_{L^p} = (\int |\nabla u|^p)^{\frac{1}{p}} = (\int (\sum_{i=1}^d |\partial_i u|^2)^{\frac{p}{2}})^{\frac{1}{p}}.$$

Observe that for $p = 2$

$$\|\nabla u\|_{L^2} = (\int |\nabla u|^p)^{\frac{1}{p}} = (\int \sum_{i=1}^d |\partial_i u|^2)^{\frac{1}{2}}.$$

But view of 8.6, for all $p \in [1, \infty]$ there exists a $M > 0$ such that

$$\|\nabla u\|_{L^p} \le M (\int \sum_{i=1}^d |\partial_i u|^p)^{\frac{1}{p}}.$$

In this way one can reformulate Theorem 8.12 for the desired form of the norm on the right-hand side.

9 Solutions of elliptic PDEs in Sobolev spaces

In this section we will show the existence of solutions to elliptic equations. The notion of solution will be defined in the language of Sobolev spaces. Let $\mathbb{F} = \mathbb{R}$ and Ω be a bounded open subset of $\mathbb{R}^d$. We will consider the following Dirichlet boundary problem

$$\begin{cases} Lu = f & \text{on } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases} \tag{41}$$

where $f : \Omega \to \mathbb{R}$ is given, $\partial\Omega$ denotes the boundary of Ω and L is the following second-order partial differential operator (with variable coefficients), with $\mathfrak{a}_{i,j}, \mathfrak{b}_i, \mathfrak{c} : \Omega \to \mathbb{F}$,

$$Lu(x) = -\sum_{i,j=1}^{d} \partial_i(\mathfrak{a}_{ij}(x)\partial_j u)(x) + \sum_{i=1}^{d} \mathfrak{b}_i(x)\partial_i u(x) + \mathfrak{c}(x)u(x) \qquad (x \in \Omega). \tag{42}$$

This problem is called a *Dirichlet boundary problem* because of the condition that $u = 0$ on ∂U; this zero-boundary condition is called a *Dirichlet boundary condition.* As the operator L is defined by (42), one says that the PDE $Lu = f$ is in *divergence form.* Observe that with $\tilde{\mathfrak{b}}_i = \mathfrak{b}_i - \sum_{j=1}^{d} \partial_j \mathfrak{a}_{ij}$ we have $Lu = \tilde{L}u$, with

$$\tilde{L}u(x) = -\sum_{i,j=1}^{d} \mathfrak{a}_{ij}(x)\partial_{ij}u(x) + \sum_{i=1}^{d} \tilde{\mathfrak{b}}_i(x)\partial_i u(x) + \mathfrak{c}(x)u(x) \qquad (x \in \Omega).$$

The PDE $\tilde{L}u = f$ is said to be of *nondivergence form.*

9.1 (Assumption). We will assume the following symmetry for the operator L, namely that $\mathfrak{a}_{ij} = \mathfrak{a}_{ji}$ for all i and j. Moreover, we assume that $\mathfrak{a}_{i,j}, \mathfrak{b}_i$ and $\mathfrak{c}$ are in $L^\infty(\Omega)$ for all i and j, $f \in L^2(\Omega)$ and that the operator is assumed to be elliptic.

Definition 9.2. The partial differential operator L is called *elliptic* if there exists a $\theta > 0$ such that

$$\sum_{i,j=1}^{d} \mathfrak{a}_{ij}(x)y_i y_j \geq \theta|y|^2 \qquad (x \in \Omega, y \in \mathbb{R}^d). \tag{43}$$

Observe that $-\Delta$ is an elliptic operator.

We will consider a bilinear form associated to L. This bilinear form arises by integration by parts, or by the interpretation of Lu as a distribution, the bilinear form equals the function $(u, v) \mapsto \langle Lu, v\rangle$. Namely:

Definition 9.3. (a) We define the *bilinear form associated to* L by $B : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ by

$$B(u, v) = \int_\Omega \sum_{i,j=1}^{d} \mathfrak{a}_{ij}(\partial_i u)(\partial_i v) + \sum_{i=1}^{d} \mathfrak{b}_i(\partial_i u)v + \mathfrak{c}uv.$$

(b) A $u \in H_0^1(\Omega)$ is called a *weak solution* to the Dirichlet boundary problem (41) if

$$B(u,v) = \langle f, v\rangle_{L^2} \qquad (v \in H_0^1(\Omega)).$$

Remark 9.4. As one sees, the Dirichlet boundary condition of (41) has been put on the space in which one considers the solution to be, namely in $H_0^1(\Omega)$, for which already had the interpretation that $u = 0$ on $\partial\Omega$ for $u \in H_0^1(\Omega)$ as is mentioned in Remark 8.10.

We can use tools from functional analysis to prove that under certain conditions there exists a weak solution of the Dirichlet boundary problem (41). Let us first recall the Riesz-Fréchet theorem.

Theorem 9.5 (Riesz-Fréchet). *[5, Theorem 13.15] Let H be a Hilbert space over $\mathbb{F}$ with inner product $\langle\cdot,\cdot\rangle$. If $f : H \to \mathbb{F}$ is a bounded linear functional, then there exists a unique $a \in H$ such that*

$$f(x) = \langle a, x\rangle \qquad (x \in H).$$

Theorem 9.6 (Lax-Milgram). *[9, 6.2.1, Theorem 1] Let H be a Hilbert space over $\mathbb{R}$, with inner product $\langle\cdot,\cdot\rangle$ and norm $\|\cdot\|$. Let $B : H \times H \to \mathbb{R}$ be a bilinear map. Suppose there exist $c, C > 0$ such that*

$$|B(u,v)| \le C\|u\|\|v\| \qquad (u, v \in H), \tag{44}$$

$$c\|u\|^2 \le B(u,u) \qquad (u \in H). \tag{45}$$

Let $g : H \to \mathbb{R}$ be a bounded linear functional. Then there exists a unique $u \in H$ such that

$$B(u,v) = g(v) \qquad (v \in H).$$

Proof. If $B(u,v) = B(v,u)$, then B defines another inner product on H and so the theorem follows directly by the Riesz-Fréchet theorem.

As for $u \in H$ the map $v \mapsto B(u,v)$ is a bounded linear functional, the Riesz-Fréchet theorem implies that there exists an element in H, for which we write $A(u)$, such that

$$B(u,v) = \langle A(u), v\rangle \qquad (v \in H).$$

We will show that A is a bounded linear bijection (it is actually even a homeomorphism). By a straightforward calculation one checks that A is linear. Moreover,

$$\|Au\|^2 = \langle Au, Au\rangle = B(u, Au) \le C\|u\|\|Au\| \qquad (u \in H).$$

Therefore $\|Au\| \le C\|u\|$ for $u \in H$, so that A is bounded.

Let us first show that A is injective and its range is closed in H. This follows from (45), as

$$c\|u\|^2 \le B(u,u) = \langle Au, u\rangle \le \|Au\|\|u\|.$$

Now, let us prove that $A(H)$, the range of A, equals H. As $A(H)$ is closed we have $A(H)+A(H)^\perp = H$ (where $A(H)^\perp$ are those elements that are orthogonal to $A(H)$), so it is sufficient to show that $A(H)^\perp = \{0\}$. Let $w \in A(H)^\perp$. Then $0 = \langle Aw, w\rangle = B(w,w) \ge c\|w\|^2$. So $w = 0$.

By Riesz-Fréchet theorem, there exists a unique $w \in H$ such that $g(v) = \langle w, v\rangle$ for all $v \in H$. Therefore, with $u = A^{-1}w$, we have $B(u,v) = \langle w, v\rangle = g(v)$ for all $v \in H$. □

Let us verify the assumptions of the Lax-Milgram theorem for B as in Definition 9.3 (a).

Theorem 9.7. *Let B be as in Definition 9.3* (a) *under Assumption 9.1. There exists a $\gamma \ge 0$ and $c, C > 0$ such that*

$$|B(u,v)| \le C\|u\|_{H^1}\|v\|_{H^1} \qquad (u,v \in H_0^1(\Omega)), \tag{46}$$

$$c\|u\|_{H^1}^2 \le B(u,u) + \gamma\|u\|_{L^2}^2 \qquad (u \in H_0^1(\Omega)). \tag{47}$$

Proof. (46) we obtain as for $u, v \in H_0^1(\Omega)$

$$\begin{aligned} |B(u,v)| &\le \sum_{i,j=1}^d \|\mathfrak{a}_{ij}\|_{L^\infty}\int_\Omega |\partial_i u||\partial_j v| + \sum_{i=1}^d \|\mathfrak{b}_i\|_{L^\infty}\int_\Omega |\partial_i u||v| + \|\mathfrak{c}\|_{L^\infty}\int_\Omega |\partial_i u||\partial_j v| \\ &\le \left(\sum_{i,j=1}^d \|\mathfrak{a}_{ij}\|_{L^\infty} + \sum_{i=1}^d \|\mathfrak{b}_i\|_{L^\infty} + \|\mathfrak{c}\|_{L^\infty}\right)\|u\|_{H^1}\|v\|_{H^1}. \end{aligned}$$

On the other hand, for $\theta > 0$ as in (43) we have

$$\begin{aligned} \theta\sum_{i=1}^d \int_\Omega |\partial_i u|^2 &\le \int_\Omega \sum_{i,j=1}^d \mathfrak{a}_{ij}(\partial_i u)(\partial_j u) \\ &= B(u,u) - \int_\Omega \sum_{i=1}^d \mathfrak{b}_i u\partial_i u - \int_\Omega \mathfrak{c}u^2 \\ &\le B(u,u) + \sum_{i=1}^d \|\mathfrak{b}_i\|_{L^\infty}\int_\Omega |\partial_i u||u| + \|\mathfrak{c}\|_{L^\infty}\int_\Omega u^2. \end{aligned}$$

As $ab \le \varepsilon a^2 + \frac{1}{4\varepsilon}b^2$ for any $a, b \in \mathbb{R}$ and $\varepsilon > 0$ we have

$$\int_\Omega |\partial_i u||u| \le \varepsilon\int_\Omega |\partial_i u|^2 + \frac{1}{4\varepsilon}\int_\Omega |u|^2.$$

Now take ε small enough such that $\varepsilon\|\mathfrak{b}_i\|_{L^\infty} \le \frac{\theta}{2}$. Then by the Poincaré inequality (see Theorem 8.12 (and 8.13)) we have $\beta\|u\|_{H^1}^2 \le \sum_{i=1}^2 \int_\Omega |\partial_i u|^2$ and thus

$$\beta\frac{\theta}{2}\|u\|_{H^1}^2 \le \frac{\theta}{2}\sum_{i=1}^d \int_\Omega |\partial_i u|^2 \le B(u,u) + (\|\mathfrak{c}\|_{L^\infty} + \frac{1}{4\varepsilon})\int_\Omega u^2.$$

□

Now we can prove that under certain conditions (41) has a weak solution.

Theorem 9.8. [9, 6.2.2 Theorem 3] There exists a $\gamma \geq 0$ such that for all $\beta \geq \gamma$ and $f \in L^2(\Omega)$ there exists a unique weak solution $u \in H_0^1(\Omega)$ of the Dirichlet boundary problem

$$\begin{cases} Lu + \beta u = f & \text{on } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{48}$$

Proof. We apply the Lax-Milgram theorem to B_β, the bilinear operator corresponding to the elliptic operator L_β given by $L_\beta u = Lu + \beta u$:

$$B_\beta(u, v) = B(u, v) + \beta \langle u, v \rangle_{L^2} \qquad (u, v \in H_0^1(\Omega)).$$

Observe that for $f \in L^2(\Omega)$ the map $g : H_0^1(\Omega) \to \mathbb{R}$ given by $g(v) = \langle f, v \rangle_{L^2}$ is bounded and linear, because $\|v\|_{L^2}^2 \leq \|v\|_{H^1}^2$. So the Lax-Milgram theorem implies the existence of a $u \in H_0^1(\Omega)$ such that $B_\beta(u, v) = \langle f, v \rangle_{L^2}$ for all $v \in H_0^1(\Omega)$, which means that u is a weak solution to (48). ☐

There are more theorems on weak solutions of elliptic Dirichlet boundary problems, see [9, Section 6.2] (for example see the Fredholm alternative). Also, one can show that the solutions have a certain regularity that depends on the regularity of the coefficients $\mathfrak{a}_{i,j}, \mathfrak{b}_i, \mathfrak{c}$, see [9, Section 6.3].

Exercise 9.1. Show that one can choose $\gamma = 0$ in Theorem 9.7 and Theorem 9.8 in case $\mathfrak{b}_i = 0$ for all i and $\mathfrak{c} = 0$.

10 The Schwartz space and tempered distributions

We introduce the Schwartz space in this section, which is the space of smooth functions that quite rapidly decay at infinity. This space is suitable for the Fourier transform, as the Fourier transform maps the Schwarz into itself (let us mention that in $\mathcal{E}$ there are functions which do not have a Fourier transform as for this a function needs to be integrable, and on the other hand, the only smooth function with compact support that has a Fourier transform with a compact support is the zero function). We will turn to that later and first discuss here the topological properties of the Schwartz space and its dual, the space of tempered distributions. As our underlying space we consider $\mathbb{R}^d$ (only). For this reason we can leave out the part "$(\mathbb{R}^d)$" in the notation of function spaces or spaces of distributions.

Definition 10.1. The *Schwartz space* $\mathcal{S}$ (or $\mathcal{S}(\mathbb{R}^d)$) is the space of functions $\varphi \in C^\infty$ such that $\|\varphi\|_{k,\mathcal{S}} < \infty$ for all $k \in \mathbb{N}_0$, where $\|\cdot\|_{k,\mathcal{S}}$ is defined by

$$\|\varphi\|_{k,\mathcal{S}} := \sum_{\alpha:|\alpha|\leq k} \sup_{x \in \mathbb{R}^d} (1 + |x|)^k |\partial^\alpha \varphi(x)| \qquad (\varphi \in C^\infty). \tag{49}$$

$\|\varphi\|_{k,\mathcal{S}} < \infty$ is a seminorm on $\mathcal{S}$ for all $k \in \mathbb{N}_0$. A function in the Schwartz space will also be called a *Schwartz function.* The space $\mathcal{S}$ is equipped with the topology generated by the seminorms $\|\cdot\|_{k,\mathcal{S}}$.

We write $\mathcal{S}'$ (or $\mathcal{S}'(\mathbb{R}^d)$) for space of continuous linear maps $\mathcal{S} \to \mathbb{C}$. This means that $u \in \mathcal{S}'$ if and only if u is linear and there exists a $k \in \mathbb{N}_0$ and a $C > 0$ such that

$$|u(\varphi)| \le C\|\varphi\|_{k,\mathcal{S}} \qquad (\varphi \in \mathcal{S}).$$

An element of $\mathcal{S}'(\mathbb{R}^d)$ will also be called a *tempered distribution.*

$\mathcal{S}'$ is equipped with the $\sigma(\mathcal{S}', \mathcal{S})$ topology.

Observe that a smooth function φ is in the Schwartz space if the function and all its derivatives are decaying faster than any polynomial.

$\boxed{\textbf{Exercise}}$ **10.1.** Let $\varphi \in C^\infty$. Show that $\varphi \in \mathcal{S}$ if and only if $\lim_{|x|\to\infty} P(x)\varphi(x) = 0$ for all polynomials P.

10.2. There are different choices of seminorms that one can take, which generate the same topology. The seminorm as in (50) is the same as in [2]. In [7] instead the following seminorms are used

$$\varphi \mapsto \sum_{\alpha:|\alpha|\le k} \sup_{x\in\mathbb{R}^d} |(1+|x|^2)^k \partial^\alpha \varphi(x)| \tag{50}$$

is used. Basically because of the following inequality, the topologies generated are equivalent.

$$1 + |x|^2 \le (1+|x|)^2 \le 2(1+|x|^2) \qquad (x \in \mathbb{R}^d).$$

$\boxed{\textbf{Exercise}}$ **10.2.** Convince yourself of the statement in 10.2.

It will be clear that all compactly supported smooth functions are Schwartz functions. We will give a central example of a Schwartz function that is not compactly supported.

Definition 10.3. A function $f : \mathbb{R}^d \to \mathbb{R}$ is called a *Gaussian function* if there exist $a, b \in \mathbb{R}$, $a > 0$, $y \in \mathbb{R}^d$ such that

$$f(x) = be^{-a|x-y|^2}.$$

Example 10.4. An example of a C^∞ function without compact support that is a Schwartz function, is a Gaussian function. Indeed, for f as in Definition 10.3 for $\alpha \in \mathbb{N}_0^d$ with $|\alpha| = k$ one has $|\partial^\alpha f(x)| \le |b|(2a)^k|x-y|^k e^{-a|x-y|^2}$, so that because $(1+|x+y|)^k \le (1+|y|)^k(1+|x|)^k$,

$$\|f\|_{k,\mathcal{S}} \le |b|(2a(1+|y|))^k \sup_{x\in\mathbb{R}^d} (1+|x|)^{2k} e^{-a|x|^2} < \infty. \tag{51}$$

10.5 (Notation). For $\lambda \in \mathbb{R}$ we write $l_\lambda : \mathbb{R}^d \to \mathbb{R}^d$ for the linear function that multiplies the vector by the scalar λ. It is of course bijective in case $\lambda \neq 0$. For a function $f : \mathbb{R}^d \to \mathbb{F}$ we will also write "$l_\lambda f$" for the composition $f \circ l_\lambda$, which is the function $x \mapsto f(\lambda x)$. For a distribution u we also write "$l_\lambda u$" instead of "$u \circ l_\lambda$".

10.6 (Some norm estimates and convergence facts).

(a)

$$\|\varphi\|_{C^k} \leq \|\varphi\|_{k,\mathcal{S}} \qquad (\varphi \in \mathcal{S}).$$

(b) Let $K \subset \mathbb{R}^d$ be compact and $M = \sup_{x \in K} |x|$, $k \in \mathbb{N}$. Then

$$\|\varphi\|_{k,\mathcal{S}} \leq (1+M)^k \|\varphi\|_{C^k} \qquad (\varphi \in \mathcal{D}, \operatorname{supp} \varphi \subset K).$$

(c) By 5.3 for all $k \in \mathbb{N}_0$ there exists a $C > 0$ such that

$$\|f\varphi\|_{k,\mathcal{S}} \leq C \|f\|_{C^k} \|\varphi\|_{k,\mathcal{S}} \qquad (f \in C^k, \varphi \in \mathcal{S}).$$

(d) For any function χ in C^k we have $\partial^\alpha(l_\lambda \chi)(x) = \lambda^{|\alpha|}(\partial^\alpha \chi)(\lambda x)$ one has $\|l_\lambda \chi\|_{C^k} \leq \|\chi\|_{C^k}$ for $\lambda \in [-1,1]$. Hence

$$\|(l_\lambda \chi) f\|_{k,\mathcal{S}} \leq \|\chi\|_{C^k} \|f\|_{k,\mathcal{S}} \qquad (f \in \mathcal{S}, \chi \in C^k, \lambda \in [-1,1]).$$

(e) Let $\chi \in \mathcal{D}$ have values in $[0,1]$, $\chi = 1$ on the unit ball. Then (see Exercise 10.3)

- $(l_\lambda \chi) f \to f$ in L^p as $\lambda \downarrow 0$, for $f \in L^p$.
- $(l_\lambda \chi) f \to f$ in $\mathcal{S}$ as $\lambda \downarrow 0$, for $f \in \mathcal{S}$.

(f) Let $(\chi_n)_{n \in \mathbb{N}}$ be a partition of unity, i.e., $\chi_n \in C_c^\infty(\mathbb{R}^d, [0,1])$ and $\sum_{n \in \mathbb{N}} \eta_n(x) = 1$ for all $x \in \mathbb{R}^d$. Suppose that $\sup_{N \in \mathbb{N}} \| \sum_{n=1}^N \chi_n \|_{C^k} < \infty$ for all $k \in \mathbb{N}_0$. Similar to (e) (see Exercise 10.3) we have

- $(\sum_{n=1}^N \chi_n) f \to f$ in L^p as $N \to \infty$, for $f \in L^p$.
- $(\sum_{n=1}^N \chi_n) f \to f$ in $\mathcal{S}$ as $N \to \infty$, for $f \in \mathcal{S}$.

Exercise 10.3. Prove 10.6(e) and (f). Hint: Observe that

$$\sup_{x \in \mathbb{R}^d : |x| > \frac{1}{\lambda}} (1+|x|)^k |g(x)| \leq \sup_{x \in \mathbb{R}^d} \frac{(1+|x|)^{k+1}}{1 + \frac{1}{\lambda}} |g(x)|.$$

Bonus: In (f), is the condition that $\sup_{N \in \mathbb{N}} \| \sum_{n=1}^N \chi_n \|_{C^k} < \infty$ for all $k \in \mathbb{N}_0$ necessary?

10.7 (Multiplication). For all $k \in \mathbb{N}$ there exists a $C > 0$ such that (for example by 10.6)

$$\|fg\|_{k,\mathcal{S}} \leq C\|f\|_{k,\mathcal{S}}\|g\|_{k,\mathcal{S}} \qquad (f, g \in \mathcal{S}),$$

so that multiplication of functions as a map $\mathcal{S} \times \mathcal{S} \to \mathcal{S}$ is continuous.

Because of this, if $u \in \mathcal{S}'$ and $f \in \mathcal{S}$, then fu is a tempered distribution, defined by $\langle fu, \varphi\rangle = \langle u, f\varphi\rangle$.

Before we turn to the topological properties of $\mathcal{S}$, let us recall the following definition.

Definition 10.8. For $k \in \mathbb{N}_0 \cup \{\infty\}$ we write C_{b}^k for the subset of C^k that consists of those functions that are bounded and for which their derivatives up to order k are bounded, which means $f \in C_{\mathrm{b}}^k$ if and only if $f \in C^k$ and $\partial^\alpha f \in C_{\mathrm{b}}$ for all $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \leq k$.

10.9. Observe/remember that C_b^k is complete under the norm $\|\cdot\|_{C^k}$ for $k \in \mathbb{N}_0$. Moreover, observe that $\mathcal{S} \subset C_b^k$ for all $k \in \mathbb{N}_0$ and

$$\|f\|_{C_b^k} \leq \|f\|_{k,\mathcal{S}}.$$

Theorem 10.10. *The space $\mathcal{S}$ (equipped with the seminorms $\|\cdot\|_{k,\mathcal{S}}$) is a complete separable metrizable space and $\mathcal{D}$ is dense in $\mathcal{S}$.*

Proof. As $\mathcal{S}$ is equipped with a countable number of seminorms, it metrizable, see for example [4, Proposition IV.2.1]. The completeness follows easily from the fact that $\mathcal{S}$ is continuously embedded in the complete space C_b^k for all k, see 10.9. That $\mathcal{D}$ is dense in $\mathcal{S}$ follows from 10.6(e). Let us prove the separability. For this, let ψ be a mollifier. We know that $\varphi * \psi_\varepsilon \to \varphi$ in $\|\cdot\|_{C^k}$-norm for all $k \in \mathbb{N}_0$ and $\varphi \in \mathcal{D}$ by Theorem 4.3 (b) (and Lemma 3.18). Because the support of $\varphi * \psi_\varepsilon$ is included a compact set for all $\varepsilon \in (0,1)$, the convergence $\varphi * \psi_\varepsilon \to \varphi$ also holds in $\mathcal{S}$, i.e., with respect to the seminorms $\|\cdot\|_{k,\mathcal{S}}$. Hence $D = \{\varphi * \psi_\varepsilon : \varepsilon \in (0,1) \cap \mathbb{Q}, \varphi \in \mathcal{D}\}$ is dense in $\mathcal{S}$ (but not yet countable). By Lemma 3.23 we know that for each ε and φ there exist a sequence of Riemann type sums that converge to $\varphi * \psi_\varepsilon$, in other words, there exists a sequence $(\rho_k)_{k\in\mathbb{N}}$ in

$$R = \Big\{ \sum_{i=1}^n a_i \mathcal{T}_{y_i} \psi_\varepsilon : n \in \mathbb{N}, a_1, \ldots, a_n \in \mathbb{Q}, y_1, \ldots, y_n \in \mathbb{Q}^d, \varepsilon \in (0,1) \cap \mathbb{Q} \Big\}$$

such that $\rho_k \xrightarrow{k\to\infty} \varphi * \psi_\varepsilon$ in $\mathcal{S}$. Hence R is dense in D (in the topology of $\mathcal{S}$) and D is dense in $\mathcal{S}$, so that the countable set R is dense in $\mathcal{S}$. $\square$

10.11. By 10.6(b) every tempered distribution restricted to $\mathcal{D}$ defines a distribution.

Observe also that for all $k \in \mathbb{N}_0$ and compact set $K \subset \mathbb{R}^d$

$$\|\varphi\|_{C^k,K} \leq \|\varphi\|_{k,\mathcal{S}} \qquad (\varphi \in \mathcal{S}).$$

Therefore every $u \in \mathcal{E}'$ restricted to $\mathcal{S}$ defines a tempered distribution.

Theorem 10.12. *The following inclusion maps are continuous*

$$\mathcal{D} \to (\mathcal{S}, \sigma(\mathcal{S}, \mathcal{S}')), \qquad (\mathcal{S}, \sigma(\mathcal{S}, \mathcal{S}')) \to (\mathcal{E}, \sigma(\mathcal{E}, \mathcal{E}')), \qquad \mathcal{S} \to \mathcal{E},$$

The inclusion map $\mathcal{D} \to \mathcal{S}$ is sequentially continuous.

The map $\iota : \mathcal{S}' \to \mathcal{D}'$ defined by $\iota(u) = u|_{\mathcal{D}}$ is continuous and injective.

The map $\iota : \mathcal{E}' \to \mathcal{S}'$ defined by $\iota(u) = u|_{\mathcal{S}}$ is continuous and injective.

Proof. For the sequential continuity use Theorem 2.7. For the other continuity, see 10.11. That the ι maps are injective follows as $\mathcal{D}$ is dense in $\mathcal{S}$; similarly $\mathcal{D}$ and thus $\mathcal{S}$ is dense in $\mathcal{E}$. □

In Example 10.15 we show that none of the embeddings in Theorem 10.12 is a homeomorphism on its image. But first we state two elementary facts on integrability of $x \mapsto (1+|x|)^\alpha$ and integrability of $k \mapsto (1+|k|)^\alpha$, from which we conclude integrability properties of Schwartz functions.

Lemma 10.13. *The functions $\mathbb{R}^d \to \mathbb{R}$, $x \mapsto (1+|x|)^\alpha$ and $x \mapsto (1+|x|^2)^{\frac{\alpha}{2}}$ are integrable if and only if $\alpha < -d$.*

Proof. As $1+|x|^2 \le (1+|x|)^2 \le 2(1+|x|^2)$, is it sufficient to only consider $x \mapsto (1+|x|)^\alpha$. Integrating the function on $B(0,1)$ always gives a finite integral. It will be clear that $\alpha < 0$ is required. By changing to spherical coordinates and observing that $(2r)^\alpha \le (1+r)^\alpha \le r^\alpha$ for $\alpha < 0$ and $r \ge 1$, we see that $(1+|x|)^\alpha$ is integrable if and only if $\int_1^\infty r^{d-1+\alpha}\, \mathrm{d}r$ is finite. The latter is of course the case if and only if $\alpha < -d$. □

Lemma 10.14. *The functions $\mathbb{Z}^d \to \mathbb{R}$, $k \mapsto (1+|k|)^\alpha$ and $k \mapsto (1+|k|^2)^{\frac{\alpha}{2}}$ are summable if and only if $\alpha < -d$.*

Proof. We consider the function $k \mapsto (1+|k|^2)^{\frac{\alpha}{2}}$ only (this is sufficient by the inequality $1+|k|^2 \le (1+|k|)^2 \le 2(1+|k|^2)$). We write $\lfloor x \rfloor = (\lfloor x_1 \rfloor, \dots, \lfloor x_d \rfloor)$ for $x \in \mathbb{R}^d$ where $\lfloor x_1 \rfloor$ is the largest integer that is smaller or equal to x_1. Then $\sum_{k \in \mathbb{Z}^d} (1+|k|^2)^{\frac{\alpha}{2}} = \int_{\mathbb{R}^d} (1+|\lfloor x \rfloor|^2)^{\frac{\alpha}{2}}\, \mathrm{d}x$. Note that $|x - \lfloor x \rfloor| \le \sqrt{d}$. Therefore, if $|x| \ge 2\sqrt{d}$ we have

$$\tfrac{1}{2}|x| \le |x| - \sqrt{d} \le |\lfloor x \rfloor| \le |x| + \sqrt{d} \le \tfrac{3}{2}|x|.$$

Hence, $\frac{1}{4}(1+|x|^2) \le (1+|\lfloor x \rfloor|^2) \le \frac{9}{4}(1+|x|^2)$ for those x and so the statement follows by Lemma 10.13. □

$\boxed{\textbf{Exercise}}$ **10.4.** Prove that $\sum_{n \in \mathbb{N}} n\delta_n \in \mathcal{S}'(\mathbb{R})$. Moreover, if $p : \mathbb{R} \to \mathbb{F}$ is a polynomial, show that $\sum_{n \in \mathbb{N}} p(n)\delta_n$ is a tempered distribution. What about $p(x) = e^x$?

Example 10.15. We consider only one dimension, i.e., $d = 1$ for convenience.

(a) As in Example 5.8 let $\phi_n = \frac{1}{n}\mathcal{T}_n\phi$ for $n \in \mathbb{N}$, where $\phi \in \mathcal{D}$ with $\int \phi = 1$ and $\phi(0) = 1$. Then $\phi_n \to 0$ in $\mathcal{E}$, but $(\phi_n)_{n\in\mathbb{N}}$ does not converge in $\mathcal{S}$: Indeed, $u = \sum_{n\in\mathbb{N}} n\delta_n$ is a tempered distribution (see Exercise 10.4) and $u(\phi_n) = 1$ for all n.

(b) Let $f_n(x) = e^{-n(1+x^2)}$. By Example 10.4 $\lim_{n\to\infty} \|f_n\|_{k,\mathcal{S}} = 0$ for all $k \in \mathbb{N}_0$. Let $\phi \in \mathcal{D}$ be nonzero and define $\psi_n = f_n \mathcal{T}_n \phi$. Observe that $\psi_n \in \mathcal{D}$. By 10.6(c), $\|\psi_n\|_{k,\mathcal{S}} \le \|\phi\|_{C^k} \|f_n\|_{k,\mathcal{S}} \to 0$ as $n \to \infty$ for all $k \in \mathbb{N}_0$. Hence $\psi_n \to 0$ in $\mathcal{S}$. However, ψ_n does not converge in $\mathcal{D}$ by Theorem 2.7.

(c) $\delta_n \to 0$ in $\mathcal{S}'$ but not in $\mathcal{E}'$.

(d) Consider the element of $\mathcal{S}'$ given by $a_n \delta_n$, where $a_n = e^{n^2}$. Then $a_n \delta_n \to 0$ in $\mathcal{D}'$ but not in $\mathcal{S}'$ (and not in $\mathcal{E}'$), as for $\varphi(x) = e^{-x^2}$ we have $a_n \delta_n(\varphi) = 1$ for all n.

Remark 10.16. Observe that even if we equip $\mathcal{D}$ with the topology generated by the seminorms $(\|\cdot\|_{C^k})_{k\in\mathbb{N}_0}$ it is not continuously embedded in $\mathcal{S}$. This follows from Example 10.15 (a) as also $\|\phi_n\|_{C^k} \to 0$ for all $k \in \mathbb{N}_0$.

Theorem 10.17. *[7, Page 137] The space $\mathcal{S}'$ is weak* sequentially complete.*

Proof. The proof is very similar to the proof of Theorem 2.13 and uses the fact that $\mathcal{S}$ is a metric space: One replaces "$\mathcal{D}_K(\Omega)$" and "d_K" by "$\mathcal{S}$" and "d", where d is the metric on $\mathcal{S}$. $\square$

10.18 ($\mathcal{S}'$ is not metrizable). [7, Page 137] only mentions this, no arguments.

Lemma 10.19. *Let $p \in [1,\infty)$. For all $k \in \mathbb{N}$ such that $pk > d$ we have $(1+|\boldsymbol{x}|)^{-k} \in L^p$ and*

$$\|\cdot\|_{L^p} \le \|(1+|\boldsymbol{x}|)^{-k}\|_{L^p} \|\cdot\|_{k,\mathcal{S}}.$$

Hence $\mathcal{S}$ is continuously embedded in L^p. Moreover, $\mathcal{S}$ is dense in L^p. As

$$\|\cdot\|_{L^\infty} = \|\cdot\|_{0,\mathcal{S}},$$

$\mathcal{S}$ is also continuously embedded in L^∞.

Proof. Let $f \in \mathcal{S}$ and $p \in [1,\infty)$ (note that $\|\cdot\|_{L^\infty} \le \|\cdot\|_{0,\mathcal{S}}$). By definition of $\|\cdot\|_{k,\mathcal{S}}$ we have

$$|f(x)| \le \|f\|_{k,\mathcal{S}} (1+|x|)^{-k} \qquad (x \in \mathbb{R}^d),$$

whence by picking k large enough such that $pk > d$, by Lemma 10.13 it follows that $\|f\|_{L^p} \le C\|f\|_{k,\mathcal{S}}$ with C being the $\|\cdot\|_{L^p}$-norm of $x \mapsto (1+|x|)^{-k}$. The denseness follows by Lemma A.13. $\square$

10.20. Observe that L^1_{loc} is not a subset of $\mathcal{S}'$ as for example the function $e^{|\boldsymbol{x}|^2}$ is not in $\mathcal{S}'(\mathbb{R})$ (see Exercise 10.5). However, L^p is a subset of $\mathcal{S}'$ for all $p \in [1,\infty]$ and is continuously embedded, see Theorem 10.21.

$\boxed{\textbf{Exercise}}$ **10.5.** Verify that $x \mapsto e^{|x|^2}$ is not in $\mathcal{S}'$.

We present the analogue statement to Theorem 2.15 in Theorem 10.21, but first introduce the following notation.

Theorem 10.21. *Let $p \in [1,\infty]$. We have $L^p \subset \mathcal{S}'$, moreover the function $L^p \to \mathcal{S}'$, $f \mapsto u_f$ is continuous and injective.*

Proof. The injectivity follows from the injectivity of $L^p \to \mathcal{D}'$ and as $\mathcal{D}$ is dense in $\mathcal{S}$. Let $f \in L^p(\mathbb{R}^d)$. Then by Hölder's inequality, with $q \in [1,\infty]$ such that $\frac{1}{p} + \frac{1}{q} = 1$,

$$|u_f(\varphi)| = \left| \int_{\mathbb{R}^d} f\varphi \right| \le \|f\|_{L^p} \|\varphi\|_{L^q} \le C\|f\|_{L^p} \|\varphi\|_{k,\mathcal{S}},$$

where $k \in \mathbb{N}_0$ and $C > 0$ are as in Lemma 10.19. $\square$

Definition 10.22. A function $f : \mathbb{R}^d \to \mathbb{F}$ is said to be *of at most polynomial growth* if there there exists a $C > 0$ and a $k \in \mathbb{N}_0$ such that

$$|f(x)| \le C(1+|x|)^k \qquad (x \in \mathbb{R}^d). \tag{52}$$

We write C_{p}^∞ for the set of C^∞ functions f such that for all $\alpha \in \mathbb{N}_0^d$, the function $\partial^\alpha f$ that are of at most polynomial growth.

$\boxed{\textbf{Exercise}}$ **10.6.** Show that (52) holds if and only if there exists a polynomial $p : \mathbb{R} \to \mathbb{R}$ such that $|f(x)| \le p(|x|)$.

10.23. Observe that $f \in C_{\mathrm{p}}^\infty$ if and only if for all $m \in \mathbb{N}_0$ there exists an $k \in \mathbb{N}_0$ such that

$$\mathfrak{q}_k(f) := \sum_{\substack{\alpha \in \mathbb{N}_0^d \\ |\alpha| \le m}} \sup_{x \in \mathbb{R}^d} (1+|x|)^{-k} |\partial^\alpha f(x)| < \infty. \tag{53}$$

Let $k, m \in \mathbb{N}_0$. By Leibniz formula 5.3, for example (29), there exists a $C > 0$ such that for all $f \in C_{\mathrm{p}}^\infty$ and $\varphi \in \mathcal{S}$

$$\begin{aligned} &\sum_{\substack{\alpha \in \mathbb{N}_0^d \\ |\alpha| \le m}} \sup_{x \in \mathbb{R}^d} (1+|x|)^m |\partial^\alpha (f\varphi)(x)| \\ &\le C\Big(\sum_{\substack{\alpha \in \mathbb{N}_0^d \\ |\alpha| \le m}} \sup_{x \in \mathbb{R}^d} (1+|x|)^{-k} |\partial^\alpha f(x)| \Big) \Big(\sum_{\substack{\beta \in \mathbb{N}_0^d \\ |\beta| \le m}} \sup_{x \in \mathbb{R}^d} (1+|x|)^{m+k} |\partial^\beta \varphi)(x)| \Big), \end{aligned}$$

for all $f \in C_{\mathrm{p}}^\infty$ with $\mathfrak{q}_k(f) < \infty$ and $\varphi \in \mathcal{S}$,

$$\|f\varphi\|_{m,\mathcal{S}} \le C\mathfrak{q}_k(f) \|\varphi\|_{m+k,\mathcal{S}}. \tag{54}$$

Therefore

$$f \in C_{\mathrm{p}}^{\infty},\ \varphi \in \mathcal{S} \implies f\varphi \in \mathcal{S}, \tag{55}$$
$$f \in C_{\mathrm{p}}^{\infty},\ u \in \mathcal{S}' \implies fu \in \mathcal{S}'. \tag{56}$$

Moreover, if $f_n \in C_{\mathrm{p}}^{\infty}$ for all $n \in \mathbb{N}$ and $f \in C_{\mathrm{p}}^{\infty}$ and $\mathfrak{q}_k(f_n - f) \xrightarrow{n\to\infty} 0$ for some $k \in \mathbb{N}$, then

$$f_n\varphi \xrightarrow{n\to\infty} f\varphi \text{ in } \mathcal{S} \qquad (\varphi \in \mathcal{S}),$$
$$f_n u \xrightarrow{n\to\infty} fu \text{ in } \mathcal{S}' \qquad (u \in \mathcal{S}').$$

10.24. Coming back to 10.20: There are more functions that are representing tempered distributions than those in L^p spaces. For example, by (56), as $\mathbb{1} \in \mathcal{S}'$, we have

$$C_{\mathrm{p}}^{\infty} \subset \mathcal{S}'.$$

Moreover, locally integrable functions that are of at most polynomial growth are representing tempered distribution (as we have seen in Exercise 10.5, exponential growth is "too fast").

Exercise **10.7.** Verify 10.24.

10.25 (Summary of embeddings). If $\mathcal{X}$ and $\mathcal{Y}$ are two topological spaces, we write $X \hookrightarrow Y$ for "X is continuously embedded in Y", i.e., $\mathcal{X} \subset \mathcal{Y}$ and the inclusion map $X \to Y$ is continuous.

$\mathcal{S} \hookrightarrow L^p \hookrightarrow \mathcal{S}'$	Lemma 10.19 and Theorem 10.21
$\mathcal{D} = (\mathcal{D}, \sigma(\mathcal{D},\mathcal{D}')) \hookrightarrow (\mathcal{S}, \sigma(\mathcal{S},\mathcal{S}')) \hookrightarrow (\mathcal{E}, \sigma(\mathcal{E},\mathcal{E}'))$	Theorems 5.7 and 10.12
$\mathcal{D} \hookrightarrow \mathcal{S}$ (sequentially)	Theorem 10.12
$\mathcal{S} \hookrightarrow \mathcal{E}$	Theorem 10.12
$\mathcal{E}' \hookrightarrow \mathcal{S}' \hookrightarrow \mathcal{D}'$	Theorem 10.12

11 Fourier Transforms

Still we consider as our underlying space $\mathbb{R}^d$ and leave out the notation "$(\mathbb{R}^d)$" in funtion spaces or spaces of distributions. We also allow $\mathbb{F}$ to be either $\mathbb{R}$ or $\mathbb{C}$, and leave out the notation "$\mathbb{F}$" of the function spaces unless it matters. Let us first introduce some auxiliary lemmas.

Lemma 11.1. *Let $f \in L^1$. Then for all $a \in \mathbb{R}^d$*

$$\lim_{x\to a} \|\mathcal{T}_a f - \mathcal{T}_x f\|_{L^1} = 0.$$

Proof. If f is the indicator function of a rectangle in $\mathbb{R}^d$ (see Definition A.12), then it is easy to see that the above limit holds. Therefore it follows that for any finite linear combination of such indicator functions, the above limit holds. As the linear span of indicator functions of rectangles is dense in L^1, see Lemma A.13, by a 3ε argument one can finish the proof. $\square$

Lemma 11.2 (Lemma of Riemann-Lebesgue). *Let $g \in L^1(\mathbb{R})$. Then*

$$\left| \int_{\mathbb{R}} g(x) e^{-2\pi i a x} \, \mathrm{d}x \right| \leq \frac{1}{2} \| g - \mathcal{T}_{\frac{1}{2a}} g \|_{L^1} \qquad (a \in \mathbb{R}, a \neq 0).$$

Proof. Let $a \in \mathbb{R}$, $a \neq 0$. As $e^{\pi i} = -1$ we have

$$\int_{\mathbb{R}} g(x) e^{-2\pi i a x} \, \mathrm{d}x = \int_{\mathbb{R}} g(x - \frac{1}{2a}) e^{-2\pi i a (x - \frac{1}{2a})} \, \mathrm{d}x = - \int_{\mathbb{R}} \mathcal{T}_{\frac{1}{2a}} g(x) e^{-2\pi i a x} \, \mathrm{d}x.$$

Therefore

$$\int_{\mathbb{R}} g(x) e^{-2\pi i a x} \, \mathrm{d}x = \frac{1}{2} \int_{\mathbb{R}} [g(x) - \mathcal{T}_{\frac{1}{2a}} g(x)] e^{-2\pi i a x} \, \mathrm{d}x,$$

so that the desired inequality follows. $\square$

Definition 11.3 (Fourier transform of a function). Let $f \in \mathcal{L}^1$. The *Fourier transform* of f, $\hat{f} : \mathbb{R}^d \to \mathbb{C}$ is given by

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} e^{-2\pi i \langle x, \xi \rangle} f(x) \, \mathrm{d}x, \tag{57}$$

where $\langle x, \xi \rangle$ is the inner product on $\mathbb{R}^d$ (the notation $\langle \cdot, \cdot \rangle$ is of course also used as the pairing between distributions, but we assume there will be no confusing arising).

In case $g \in \mathcal{L}^1$ equals f almost everywhere, then $\hat{f} = \hat{g}$. This allows us to define the Fourier transform of an element of L^1 as the Fourier transform of one of its representatives in $\mathcal{L}^1$ and will use the formula (57) also for $f \in L^1$.

Exercise 11.1 (Voorbeeld 15.8). [23] Calculate the Fourier transform of the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = \max(1 - |x|, 0)$.

Exercise 11.2. Let $f \in L^1(\mathbb{R}^d)$ (for some $d \in \mathbb{N}$). Show that if $|\xi_n| \to \infty$, then $\hat{f}(\xi_n) \to 0$.

11.4 (Notation). Let $g : \mathbb{R}^d \to \mathbb{F}$ be a function. Suppose there exists an $L \in \mathbb{F}$ such that for all $\varepsilon > 0$ there exists an $M > 0$ such that $|g(x) - L| < \varepsilon$ for all $x \in \mathbb{R}^d$ with $|x| > M$. Then we will write "$\lim_{|x| \to \infty} g(x)$" for "$L$".

Theorem 11.5. *If $f \in L^1$, then $\hat{f} \in C_b(\mathbb{R}^d, \mathbb{C})$, $\lim_{|\xi| \to \infty} \hat{f}(\xi) = 0$ and*

$$\| \hat{f} \|_{L^\infty} \leq \| f \|_{L^1}.$$

Proof. For the convergence we refer to Exercise 11.2. The countinuity follows by Lebesgue's dominated convergence theorem. The bound on the norm is easy. $\square$

Definition 11.6 (Fourier transform as a function). We write $\mathcal{F}$ for the linear function $L^1 \to C_b(\mathbb{R}^d, \mathbb{C})$, $f \mapsto \hat{f}$ and call this map the *Fourier transform.*

Theorem 11.7. [23, Stelling 13.4] [30, Chapter 5 Proposition 1.8] [7, Page 142]
Let $f, g \in L^1$. Then $f\hat{g}$ and $\hat{f}g$ are integrable and

$$\int f\hat{g} = \int \hat{f}g. \tag{58}$$

Proof. The integrability follows by Theorem 11.5. The identity follows by Fubini's theorem. $\square$

Exercise **11.3.** Check that (58) holds.

11.8 (Notation). We will use the bold symbols $\boldsymbol{\xi}$ and $\boldsymbol{x}$ to denote the identity maps on $\mathbb{R}^d$, which means $\boldsymbol{\xi} : \xi \mapsto \xi$ and $\boldsymbol{x} : x \mapsto x$.

By substitution rules for integration we obtain the following.

Theorem 11.9. (a) *Let $f \in L^1$ and $y \in \mathbb{R}^d$. Then*

$$\mathcal{F}(\mathcal{T}_y f) = e^{-2\pi \mathrm{i}\langle \boldsymbol{\xi}, y\rangle} \hat{f}, \qquad \mathcal{T}_y \hat{f} = \mathcal{F}(e^{2\pi \mathrm{i}\langle \boldsymbol{x}, y\rangle} f). \tag{59}$$

(b) *Let $l : \mathbb{R}^d \to \mathbb{R}^d$ be linear and bijective. Then*

$$\mathcal{F}(f \circ l) = \frac{1}{|\det l|} \hat{f} \circ l_*,$$

where l_* is the transpose of l^{-1}, which means that $\langle l^{-1} y, \xi\rangle = \langle y, l_*(\xi)\rangle$ for all $x, \xi \in \mathbb{R}^d$.

In particular, for $\lambda \in \mathbb{R} \setminus \{0\}$ (for the notation see 10.5)

$$\mathcal{F}(l_\lambda f) = |\lambda|^{-d} l_{\frac{1}{\lambda}} \hat{f}.$$

Or, differently written $\mathcal{F}(f(\lambda \boldsymbol{x})) = |\lambda|^{-d} \hat{f}(\frac{\boldsymbol{\xi}}{\lambda})$.

Exercise **11.4.** Verify the statements of Theorem 11.9.

Theorem 11.10. [23, Stelling 13.5] *Let $g \in L^1(\mathbb{R})$.*

(a) *If $\boldsymbol{x}g \in L^1(\mathbb{R})$, then $\hat{g}$ is continuously differentiable and*

$$\hat{g}' = \mathcal{F}(-2\pi \mathrm{i} \boldsymbol{x} g). \tag{60}$$

(b) *If g is the indefinite integral of a function $h \in L^1(\mathbb{R})$ (which means that for any $a, b \in \mathbb{R}$ with $a < b$, $g(b) - g(a) = \int_a^b h$), then $\hat{h} = 2\pi \mathrm{i} \boldsymbol{\xi} \hat{g}$.*

In particular, if g is continuously differentiable and $g' \in L^1(\mathbb{R})$, then $\mathcal{F}(g') = 2\pi \mathrm{i} \boldsymbol{\xi} \hat{g}$.

Proof. (a) Let $a, b \in \mathbb{R}$, $a < b$. Then, by Theorem 11.7

$$\begin{aligned}
\int_a^b \mathcal{F}(-2\pi \mathrm{i} \boldsymbol{x} g) &= \int_{\mathbb{R}} \mathcal{F}(-2\pi \mathrm{i} \boldsymbol{x} g) \mathbb{1}_{[a,b]} \\
&= \int_{\mathbb{R}} -2\pi \mathrm{i} x g(x) \mathcal{F}(\mathbb{1}_{[a,b]})(x) \, \mathrm{d}x \\
&= \int_{\mathbb{R}} -2\pi \mathrm{i} x g(x) \frac{e^{-2\pi \mathrm{i} b x} - e^{-2\pi \mathrm{i} a x}}{-2\pi \mathrm{i} x} \, \mathrm{d}x \\
&= \int_{\mathbb{R}} g(x)(e^{-2\pi \mathrm{i} b x} - e^{-2\pi \mathrm{i} a x}) \, \mathrm{d}x = \hat{g}(b) - \hat{g}(a).
\end{aligned}$$

As the Fourier transform of an integrable function is continuous, we conclude that $\hat{g}$ is continuously differentiable with derivative given by (60).

(b) For $\xi \in \mathbb{R}$ we have by applying integration by parts

$$\begin{aligned}
\hat{h}(\xi) - 2\pi \mathrm{i} \xi \hat{g}(\xi) &= \lim_{N \to \infty} \int_{-N}^{N} h(x) e^{-2\pi \mathrm{i} \xi x} + g(x)(-2\pi \mathrm{i} \xi) e^{-2\pi \mathrm{i} \xi x} \, \mathrm{d}x \\
&= \lim_{N \to \infty} (g(N) e^{-2\pi \mathrm{i} N \xi} - g(-N) e^{2\pi \mathrm{i} N \xi}).
\end{aligned}$$

Therefore it suffices to show that $\lim_{|x| \to \infty} g(x) = 0$. As g is the indefinite integral of h, which means for example that $g(y) = g(0) + \int_0^y h(x) \, \mathrm{d}x$, both $\lim_{y \to \infty} g(y)$ and $\lim_{y \to -\infty} g(y)$ exist. By the integrability of g, these limits need to be equal to zero. $\square$

11.11 (Note to Theorem 11.10 (b)**).** If $g \in L^1(\mathbb{R})$ is not an indefinite integral of an integrable function, but almost everywhere differentiable and its derivative is equal (there where it exists) to an integrable function h, then $\hat{g}$ might not be equal to $\widehat{(2\pi \mathrm{i} \boldsymbol{\xi}) h}$, e.g., take $g = \mathbb{1}_{[a,b]}$. Indeed, g is almost everywhere differentiable with derivative 0 but its Fourier transform is given by

$$\mathcal{F}(\mathbb{1}_{[a,b]})(\xi) = \begin{cases} \frac{e^{-2\pi \mathrm{i} b \xi} - e^{-2\pi \mathrm{i} a \xi}}{-2\pi \mathrm{i} \xi} & \xi \neq 0, \\ b - a & \xi = 0. \end{cases}$$

11.12. Let us compute the Fourier transform of a Gaussian function in one dimension, with the help of Theorem 11.10.

Let $g : \mathbb{R} \to \mathbb{R}$ be the Gaussian function given by $g(x) = e^{-x^2}$. By Theorem 11.10

$$\frac{d}{d\xi} \hat{g}(\xi) = \mathcal{F}(-2\pi \mathrm{i} \boldsymbol{x} e^{-\boldsymbol{x}^2})(\xi) = \pi i \mathcal{F}\left(\mathrm{D}\, e^{-\boldsymbol{x}^2}\right)(\xi) = -2\pi^2 \xi \hat{g}(\xi). \tag{61}$$

As $\hat{g}(0) = \int_{\mathbb{R}} e^{-x^2} \, \mathrm{d}x = \sqrt{\pi}$ (see 7.14) we have $\hat{g}(\xi) = \sqrt{\pi} e^{-\pi^2 \xi^2}$, as this is the unique solution to the ordinary differential equation (61).

We can so to say 'apply' Theorem 11.10 to any of the directions in $\mathbb{R}^d$, to obtain the following.

Theorem 11.13. *Let $k \in \mathbb{N}_0$.*

(a) *If $\boldsymbol{x}^\beta f \in L^1$ for all $\beta \in \mathbb{N}_0^d$ with $|\beta| \le k$, then $\hat{f} \in C^k$ and*

$$\partial^\beta \hat{f} = \mathcal{F}((-2\pi \mathrm{i}\boldsymbol{x})^\beta f) \qquad (\beta \in \mathbb{N}_0^d, |\beta| \le k).$$

(b) *If $f \in L^1 \cap C^k$ and $\partial^\beta f \in L^1$ for all $\beta \in \mathbb{N}_0^d$ with $|\beta| \le k$, then $\boldsymbol{\xi}^\beta \hat{f}$ is continuous and bounded and*

$$(2\pi \mathrm{i}\boldsymbol{\xi})^\beta \hat{f} = \mathcal{F}(\partial^\beta f).$$

$\boxed{\textbf{Exercise}}$ **11.5.** Verify that Theorem 11.13 follows from Theorem 11.10.

11.14. Observe that if (b) holds for all k, which means that $f \in L^1 \cap C^\infty$ and $\partial^\beta f \in L^1$ for all $\beta \in \mathbb{N}_0^d$, then $\boldsymbol{\xi}^\beta \hat{f}$ is not only continuous and bounded but also integrable for all $\beta \in \mathbb{N}_0^d$: This follows from Lemma 10.13 as also $(1+|\boldsymbol{\xi}|)^{d+1}\boldsymbol{\xi}^\beta \hat{f}$ is bounded for all $\beta \in \mathbb{N}_0^d$.

As we have also seen that Schwartz functions are integrable, and also their derivatives as they are again Schwartz functions, Theorem 11.13 implies that $\hat{f}$ is a Schwartz function if f is a Schwartz function. The Fourier transform actually forms a bijection on the Schwartz functions, which follows from the inversion theorem, Theorem 11.16.

First we turn to Fourier transform of Gaussian functions.

Theorem 11.15. *Let $a > 0$, $y \in \mathbb{R}^d$ and $f : \mathbb{R}^d \to \mathbb{R}$ be the Gaussian function $f(x) = e^{-a|x-y|^2}$, then $f \in L^1$ and*

$$\hat{f}(\xi) = \left(\frac{\pi}{a}\right)^{\frac{d}{2}} e^{-\frac{\pi^2|\xi|^2}{a}} e^{-2\pi \mathrm{i}\langle y, \xi\rangle}. \tag{62}$$

$\boxed{\textbf{Exercise}}$ **11.6.** Prove Theorem 11.15.

Theorem 11.16. *The Fourier transform $\mathcal{F}$ forms a linear homeomorphism $\mathcal{S}(\mathbb{R}^d, \mathbb{C}) \to \mathcal{S}(\mathbb{R}^d, \mathbb{C})$ with*

$$f(x) = \mathcal{F}(\hat{f})(-x) = \int_{\mathbb{R}^d} \hat{f}(\xi) e^{2\pi \mathrm{i}\langle x, \xi\rangle} \,\mathrm{d}\xi \qquad (x \in \mathbb{R}^d). \tag{63}$$

Proof. Let us first prove that $\mathcal{F}$ is bijective by proving (63). Let $f \in \mathcal{S}$ and $x \in \mathbb{R}^d$. Let $g = \mathcal{T}_{-x} f$. Then $\hat{g} = \hat{f} e^{2\pi \mathrm{i}\langle x, \boldsymbol{\xi}\rangle}$ by Theorem 11.9. Therefore, it is sufficient to show (63) for $x = 0$, which means it is sufficient to show

$$f(0) = \int_{\mathbb{R}^d} \hat{f}(\xi) \,\mathrm{d}\xi.$$

Let h_t be as in Example 7.13, i.e.,

$$h_t(x) = (4\pi t)^{-\frac{d}{2}} e^{-\frac{1}{4t}|x|^2} \qquad (x \in \mathbb{R}^d).$$

Observe that by Theorem 11.15 we have $\hat{g}_t = h_t$ for (take $a = t\pi^2 t$)

$$g_t(x) = e^{-4\pi^2 t|x|^2} \qquad (x \in \mathbb{R}^d).$$

By Theorem 11.7

$$\int_{\mathbb{R}^d} f(x) h_t(x)\ \mathrm{d}x = \int_{\mathbb{R}^d} \hat{f}(\xi) g_t(\xi)\ \mathrm{d}\xi.$$

As f is continuous and bounded, the left-hand side converges to $f(0)$ as $t \downarrow 0$ by (40). As $\hat{f}$ is an element of $\mathcal{S}$ it is integrable, therefore by Lebesgue's dominated convergence theorem we have that the right-hand side converges to $\int_{\mathbb{R}^d} \hat{f}(\xi)\ \mathrm{d}\xi$ as $t \downarrow 0$, because $g_t(\xi) \uparrow 1$ as $t \downarrow 0$ for all $\xi \in \mathbb{R}^d$.

Let $m = d + 1$. As $C := \int_{\mathbb{R}^d} (1 + |x|)^{-m}\ \mathrm{d}x$ is finite by Lemma 10.13, we have

$$\|\hat{f}\|_{L^\infty} \leq \|f\|_{L^1} = \int_{\mathbb{R}^d} (1 + |x|)^m (1 + |x|)^{-m} |f(x)|\ \mathrm{d}x \leq C\|f\|_{m,\mathcal{S}} \qquad (f \in \mathcal{S}). \tag{64}$$

Let $k \in \mathbb{N}_0$. By Theorem 11.13 for $k \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \leq k$.

$$(1 + |\boldsymbol{\xi}|^2)^{\frac{k}{2}} \partial^\alpha \hat{f} = \mathcal{F}\Big(\Big(1 - \frac{\Delta}{4\pi^2}\Big)^{\frac{k}{2}} \big((2\pi \mathrm{i}\boldsymbol{x})^\alpha f\big)\Big).$$

Now if $f_n \xrightarrow{n\to\infty} 0$ in $\mathcal{S}$, then $(2\pi \mathrm{i}\boldsymbol{x})^\alpha f_n \xrightarrow{n\to\infty} 0$ in $\mathcal{S}$ for all $\alpha \in \mathbb{N}_0^d$. Hence for all $\alpha \in \mathbb{N}_0^d$

$$\Big\|\Big(1 - \frac{\Delta}{4\pi^2}\Big)^{\frac{k}{2}} \big((2\pi \mathrm{i}\boldsymbol{x})^\alpha f_n\big)\Big\|_{m,\mathcal{S}} \xrightarrow{n\to\infty} 0,$$

Therefore by (64)

$$\sum_{\substack{\alpha \in \mathbb{N}_0^d \\ |\alpha| \leq k}} \sup_{x \in \mathbb{R}^d} (1 + |\xi|^2)^{\frac{k}{2}} |\partial^\alpha \hat{f}_n(\xi)| \xrightarrow{n\to\infty} 0,$$

and thus $\|\hat{f}_n\|_{k,\mathcal{S}} \xrightarrow{n\to\infty} 0$ (remember 10.2). $\square$

11.17. With the $\check{}$ notation at hand we could instead of (63) also write

$$\mathcal{F}(\hat{f})(-x) = \hat{\hat{f}}(-x) = \check{\hat{\hat{f}}}(x) = \mathcal{F}(\hat{f})\check{}(x).$$

Actually, $\check{}$ commutes with the Fourier transform by Theorem 11.9, i.e.,

$$\check{\hat{f}} = \hat{\check{f}} \qquad (f \in L^1).$$

Be aware however, that in some literature the symbol $\check{}$ is used as a symbol for the Fourier inverse.

Actually, the previous theorem extends in the following way, in the sense that the Fourier transform is a bijection on a larger space.

Theorem 11.18. [23, Stelling 15.9] *Suppose that f is integrable and that $\hat{f}$ is too (so both f and $\hat{f}$ are in $\mathcal{L}^1$). Then*

$$f(x) = \mathcal{F}(\hat{f})(-x) = \int_{\mathbb{R}^d} \hat{f}(\xi) e^{2\pi \mathrm{i}\langle x,\xi\rangle} \, \mathrm{d}\xi \qquad \textit{for almost all } x \in \mathbb{R}^d. \tag{65}$$

Consequently, $\mathcal{F}$ also forms a bijection on $\{f \in L^1 : \hat{f} \in L^1\}$.

Proof. For all $\varphi \in \mathcal{S}$ we have by Theorem 11.7 and Theorem 11.16.

$$\int_{\mathbb{R}^d} \mathcal{F}(\hat{f})\varphi = \int_{\mathbb{R}^d} \hat{f}\hat{\varphi} = \int_{\mathbb{R}^d} f\mathcal{F}(\hat{\varphi}) = \int_{\mathbb{R}^d} f\check{\varphi} = \int_{\mathbb{R}^d} \check{f}\varphi.$$

Therefore, by Theorem 1.16, we have $\mathcal{F}(\hat{f}) = \check{f}$ almost everywhere. ☐

11.19. Observe that by Theorem 11.5 the set $\{f \in L^1 : \hat{f} \in L^1\}$ is included in C_b (where C_b is viewed as subset of L^∞).

11.20. With $\overline{f}$ being the complex conjugate of f, observe that for $f \in L^1(\mathbb{R}^d)$

$$\hat{\overline{f}} = \overline{\check{\hat{f}}}.$$

Therefore, as a consequence of Theorem 11.7 we have

$$\langle f, \hat{g}\rangle_{L^2} = \langle \hat{f}, \check{g}\rangle_{L^2}. \tag{66}$$

By the above observation and the Fourier inversion formula, we obtain the following identity, which is due to Parseval and Plancherel.

Theorem 11.21. [23, Stelling 19.7] [Parseval, Plancherel] *$\mathcal{F}$ extends to an isometric isomorphism from L^2 into L^2, so that in particular*

$$\|\hat{f}\|_{L^2} = \|f\|_{L^2} \qquad (f \in L^2). \tag{67}$$

Proof. As C_c is dense in L^2 (see Lemma A.13) and a subset of L^1 it is sufficient to show (67) for $f \in C_c$. Let $f \in C_c$. By Theorem 11.7 (see also 11.20) we have

$$\|f\|_{L^2}^2 = \langle f, f\rangle_{L^2} = \langle \check{f}, \mathcal{F}(\hat{f})\rangle_{L^2} = \langle \hat{f}, \hat{f}\rangle_{L^2} = \|\hat{f}\|_{L^2}^2.$$

☐

11.22. For $f \in L^1$ and $\varphi \in \mathcal{S}$ we have by Theorem 11.7

$$\int f\hat{\varphi} = \int \hat{f}\varphi.$$

So that with the notation of Example 1.6, we have

$$u_{\hat{f}}(\varphi) = u_f(\hat{\varphi}).$$

This relation lies at the basis for the definition of the Fourier transform of a tempered distribution, see the next definition.

11.23. Before we define the Fourier transform, let us first mention the following. As the Fourier transform of a real valued function is complex valued function (and in general not a real valued function), we have to be able to evaluate a tempered distribution in functions in $\mathcal{S}(\mathbb{R}^d,\mathbb{C})$. If $u \in \mathcal{S}'(\mathbb{R}^d,\mathbb{R})$, then u naturally extends to a tempered distribution in $\mathcal{S}'(\mathbb{R}^d,\mathbb{C})$ as follows. Every function $\varphi \in \mathcal{S}(\mathbb{R}^d,\mathbb{C})$ can be decomposed in a real $\Re\varphi$ and imaginary part $\Im\varphi$ in $\mathcal{S}(\mathbb{R}^d,\mathbb{R})$, so that $\varphi = \Re\varphi + \mathrm{i}\Im\varphi$. Therefore we can extend u to an element $\overline{u}$ in $\mathcal{S}'(\mathbb{R}^d,\mathbb{C})$ by defining

$$\overline{u}(\varphi) = u(\Re\varphi) + \mathrm{i}u(\Im\varphi).$$

Then $\overline{u} \in \mathcal{S}'(\mathbb{R}^d,\mathbb{C})$. We will from here on also write "u" for "$\overline{u}$".

Definition 11.24. Let u be a tempered distribution. We define the *Fourier transform* of u, $\hat{u}$ by

$$\hat{u}(\varphi) = u(\hat{\varphi}) \qquad (\varphi \in \mathcal{S}).$$

We will also write $\mathcal{F}$ for the map $\mathcal{S}' \to \mathcal{S}'$.

Example 11.25. The function $\mathbb{1}$ represents a Schwarz distribution, and so does δ_0. We calculate their Fourier transforms. For $\varphi \in \mathcal{S}$ we have

$$\begin{aligned}\langle \hat{\delta}_0, \varphi\rangle = \delta_0(\hat{\varphi}) = \hat{\varphi}(0) = \int \varphi = \langle \mathbb{1}, \varphi\rangle,\\ \langle \hat{\mathbb{1}}, \varphi\rangle = \int \hat{\varphi} = \varphi(0) = \langle \delta_0, \varphi\rangle,\end{aligned}$$

where we used the inversion formula in the second line. Hence

$$\hat{\delta}_0 = \mathbb{1}, \qquad \hat{\mathbb{1}} = \delta_0.$$

The following theorem is a consequence of Theorem 11.9, Theorem 11.13 and Theorem 11.16.

Theorem 11.26. *The Fourier transform $\mathcal{F} : u \mapsto \hat{u}$ forms a linear homeomorphism $\mathcal{S}'(\mathbb{R}^d,\mathbb{C}) \to \mathcal{S}'(\mathbb{R}^d,\mathbb{C})$. Moreover,*

$$u = \mathcal{F}(\hat{u})^{\vee} \qquad (u \in \mathcal{S}'(\mathbb{R}^d,\mathbb{C})),$$

and for $u \in \mathcal{S}'$, $\beta \in \mathbb{N}_0^d$, $y \in \mathbb{R}^d$, $l : \mathbb{R}^d \to \mathbb{R}^d$ a linear bijection and $\lambda \in \mathbb{R}$,

$$\mathcal{F}(\partial^\beta u) = (2\pi\mathrm{i}\boldsymbol{\xi})^\beta \hat{u}, \qquad \partial^\beta \hat{u} = \mathcal{F}((-2\pi\mathrm{i}\boldsymbol{x})^\beta u), \tag{68}$$

$$\mathcal{F}(\mathcal{T}_y u) = e^{-2\pi\mathrm{i}\langle \boldsymbol{\xi}, y\rangle}\hat{u}, \qquad \mathcal{T}_y \hat{u} = \mathcal{F}(e^{2\pi\mathrm{i}\langle \boldsymbol{x}, y\rangle} u), \tag{69}$$

$$\mathcal{F}(u \circ l) = \frac{1}{|\det l|}\hat{u} \circ l_*, \qquad \mathcal{F}(l_\lambda u) = \frac{1}{|\lambda|^d} l_{\frac{1}{\lambda}} \hat{u}, \tag{70}$$

where l_* is the transpose of l^{-1} as in Theorem 11.9 and similar to 10.5 "$l_\lambda u$" is written for "$u \circ l_\lambda$".

12 Convolution of tempered distributions

We first introduced the Fourier transform, as we will use this to prove statements about convolutions. The following theorem shows the key relation of the Fourier transform with convolution that we will use.

Theorem 12.1. *Let $f, g \in L^1$. Then $f * g \in L^1$ and*

$$\mathcal{F}(f * g) = \hat{f}\hat{g}.$$

Proof. By Young's inequality, Theorem 3.4 we have $f * g \in L^1$. Therefore, by Fubini's theorem, we have for $\xi \in \mathbb{R}^d$,

$$\begin{aligned}
\mathcal{F}(f * g)(\xi) &= \int_{\mathbb{R}^d} f * g(x) e^{-2\pi \mathrm{i}\langle x, \xi\rangle} \, \mathrm{d}x \\
&= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(y) g(x - y) \, \mathrm{d}y e^{-2\pi \mathrm{i}\langle x, \xi\rangle} \, \mathrm{d}x \\
&= \int_{\mathbb{R}^d} f(y) \int_{\mathbb{R}^d} g(x - y) e^{-2\pi \mathrm{i}\langle x, \xi\rangle} \, \mathrm{d}x \, \mathrm{d}y \\
&= \int_{\mathbb{R}^d} f(y) \mathcal{F}(\mathcal{T}_y g)(\xi) \, \mathrm{d}y \\
&= \int_{\mathbb{R}^d} f(y) \hat{g}(\xi) e^{-2\pi \mathrm{i}\langle y, \xi\rangle} \, \mathrm{d}y = \hat{f}(\xi)\hat{g}(\xi),
\end{aligned}$$

where we used Theorem 11.9. □

As a direct consequence, by Theorem 11.16 and as multiplication is a continuous operation on $\mathcal{S}$ (see 10.7):

Lemma 12.2. *Let $\varphi, \psi \in \mathcal{S}$. Then*

$$\mathcal{F}(\varphi * \psi) = \hat{\varphi}\hat{\psi}, \qquad \mathcal{F}(\varphi\psi) = \hat{\varphi} * \hat{\psi}. \tag{71}$$

Consequently, $\varphi * \psi \in \mathcal{S}$ and the function $\mathcal{S} \times \mathcal{S} \to \mathcal{S}$, $(f, g) \mapsto f * g$ is continuous.

Definition 12.3. Let $u \in \mathcal{S}'$ and $\varphi \in \mathcal{S}$. We define the *convolution* of u with φ to be the function $\mathbb{R}^d \to \mathbb{F}$ defined by

$$u * \varphi(x) = u(\mathcal{T}_x \check{\varphi}) \qquad (x \in \mathbb{R}^d).$$

Similar to Lemma 3.18 and Lemma 6.3 we have that the convolution between a Schwartz function and a tempered distribution is smooth, as we will see in Theorem 12.5. However, it need not be a Schwartz function as will be clear from the following exercise.

$\boxed{\textbf{Exercise}}$ **12.1.** Compute the convolution of the tempered distribution $\mathbb{1}$ with the Schwartz function $e^{-|x|^2}$.

Let us consider the convergence of difference quotients as we did in 3.17.

12.4 (Convergence of difference quotients in $\mathcal{S}$ and $\mathcal{S}'$). Observe that for $f \in \mathcal{S}$, $j \in \{1, \dots, d\}$ and $h \in \mathbb{R} \setminus \{0\}$

$$\mathcal{F}\Big(\Big(\frac{\mathcal{T}_{-he_j} - \mathcal{T}_0}{h} \Big) f \Big) = \frac{e^{2\pi \mathrm{i} \langle he_j, \boldsymbol{\xi} \rangle} - 1}{h} \hat{f}.$$

Therefore we have the following convergence in $\mathcal{S}$

$$\Big(\frac{\mathcal{T}_{-he_j} - \mathcal{T}_0}{h} \Big) f \xrightarrow{h \to 0} \partial_j f, \tag{72}$$

if and only if the following convergence holds in $\mathcal{S}$, where $\boldsymbol{\xi}_j = \langle \boldsymbol{\xi}, e_j \rangle$

$$\Big(\frac{e^{2\pi \mathrm{i} h \boldsymbol{\xi}_j} - 1}{h} \Big) \hat{f} \xrightarrow{h \to 0} 2\pi \mathrm{i} \boldsymbol{\xi}_j \hat{f}. \tag{73}$$

We simplify the notation for the moment "by substituting $t = 2\pi \xi_j$". For $h \in \mathbb{R} \setminus \{0\}$ let us write g_h for the function $\mathbb{R} \to \mathbb{C}$ given by

$$g_h(t) = \frac{e^{\mathrm{i}ht} - 1}{h} - \mathrm{i}t \qquad (t \in \mathbb{R}).$$

Then

$$\frac{\mathrm{d}^{\mathrm{n}}}{\mathrm{d}t^n} g_h(t) = \begin{cases} \mathrm{i}(e^{\mathrm{i}ht} - 1) & n = 1, \\ \mathrm{i}^n h^{n-1} e^{\mathrm{i}ht} & n \geq 2. \end{cases}$$

As $e^{\mathrm{i}ht} - 1 = \mathrm{i}t \int_0^h e^{\mathrm{i}rt} \,\mathrm{d}r$,

$$g_h(t) = \frac{e^{\mathrm{i}ht} - 1}{h} - \mathrm{i}t = \mathrm{i}t \frac{\int_0^h e^{\mathrm{i}rt} - 1 \,\mathrm{d}r}{h} = (\mathrm{i}t)^2 \frac{\int_0^h \int_0^r e^{\mathrm{i}ut} \,\mathrm{d}u \,\mathrm{d}r}{h},$$

and therefore we obtain

$$\Big| \frac{\mathrm{d}^{\mathrm{n}}}{\mathrm{d}t^n} g_h(t) \Big| \leq \begin{cases} |h||t|^2 & n = 0, \\ |h||t| & n = 1, \\ |h|^{n-1} & n \geq 2. \end{cases}$$

From this we obtain for $k \in \mathbb{N}_0$ and $h \in \mathbb{R}$ with $|h| \leq 1$, that for all $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \leq k$

$$\Big| \partial^\alpha \Big(\frac{e^{2\pi \mathrm{i} h \boldsymbol{\xi}_j} - 1}{h} - 2\pi \mathrm{i} \boldsymbol{\xi}_j \Big) \Big| \leq (2\pi)^k |h| (1 + |\boldsymbol{\xi}|)^2,$$

and thus by Leibniz rule 5.3

$$\Big\| \Big(\frac{e^{2\pi \mathrm{i} h \boldsymbol{\xi}_j} - 1}{h} - 2\pi \mathrm{i} \boldsymbol{\xi}_j \Big) \hat{f} \Big\|_{k, \mathcal{S}} \leq (2\pi)^k |h| \|\hat{f}\|_{k+2, \mathcal{S}}.$$

So that (73) and (72) hold.

Consequently, we have for any $u \in \mathcal{S}'$

$$\left(\frac{\mathcal{T}_{he_i} - \mathcal{T}_0}{h}\right) u \xrightarrow{h \to 0} \partial_i u \quad \text{in } \mathcal{S}'.$$

As in 3.17, the identity (20) is also valid for $u \in \mathcal{S}'$ and $\varphi \in \mathcal{S}$. Therefore $\partial_j(u * \varphi) = u * \partial_j \varphi$.

So the convolution between a tempered distribution and a Schwartz function is smooth. Moreover, it is of at most polynomial growth:

Theorem 12.5. [7, Theorem on page 151] [24, Theorem II.7.10] *Let $u \in \mathcal{S}'$ and $\varphi \in \mathcal{S}$. Then $u * \varphi$ is smooth and of at most polynomial growth, that is $u * \varphi \in C^\infty_{\mathrm{p}}$. For all $\alpha \in \mathbb{N}_0^d$*

$$\partial^\alpha (u * \varphi) = u * (\partial^\alpha \varphi) = (\partial^\alpha u) * \varphi. \tag{74}$$

Moreover, the map $\mathcal{S} \mapsto \mathcal{E}$, $\varphi \mapsto u * \varphi$ is continuous.

Proof. That the convolution is smooth is proven in 12.4. The fact that $u*\varphi$ is of at most polynomial growth basically follows by the inequality $(1 + |x + y|) \leq (1 + |x|)(1 + |y|)$ and together with the continuity this is left as an exercise. □

$\boxed{\textbf{Exercise}}$ **12.2.** Finish the proof of Theorem 12.5.

So convolution with a tempered distribution is a continuous operation and it commutes with translation. Like in Theorem 6.6 and Theorem 6.7 also each such operation is a convolution:

Theorem 12.6. [7, Theorem on page 151] *Let A be a linear map $\mathcal{S}(\mathbb{R}^d) \to \mathcal{E}(\mathbb{R}^d)$ which commutes with translation, i.e., $\mathcal{T}_h(A\varphi) - A(\mathcal{T}_h\varphi)$, and which is continuous, then there exists a unique tempered distribution $u \in \mathcal{S}'$ such that $A\varphi = u * \varphi$ for all $\varphi \in \mathcal{D}(\mathbb{R}^d)$.*

Theorem 12.7. [7, Theorem on page 151] *Let $u \in \mathcal{S}'$ and $\varphi \in \mathcal{S}$. Then in $\mathcal{S}'$*

$$\mathcal{F}(u * \varphi) = \hat{\varphi}\hat{u}, \qquad \mathcal{F}(\varphi u) = \hat{\varphi} * \hat{u}.$$

Proof. As $u*\varphi$ is of at most polynomial growth, it is a tempered distribution (see 10.24). As $\hat{\varphi}$ is a Schwartz function, also $\hat{\varphi}\hat{u}$ is tempered.

First let us consider $\varphi \in \mathcal{D}$. By 3.25 and using that $\check{\varphi} = \hat{\hat{\varphi}}$, for $\psi \in \mathcal{S}$ with $\hat{\psi} \in \mathcal{D}$

$$\begin{aligned} \langle \mathcal{F}(u * \varphi), \psi \rangle &= \langle u * \varphi, \hat{\psi} \rangle = \langle u, \check{\varphi} * \hat{\psi} \rangle = \langle u, \hat{\hat{\varphi}} * \hat{\psi} \rangle \\ &= \langle u, \mathcal{F}(\hat{\varphi}\psi) \rangle = \langle \hat{u}, \hat{\varphi}\psi \rangle = \langle \hat{\varphi}\hat{u}, \psi \rangle. \end{aligned}$$

As $\mathcal{D}$ is dense in $\mathcal{S}$ and the Fourier transformation is continuous, also $\{\psi \in \mathcal{S} : \hat{\psi} \in \mathcal{D}\}$ is dense in $\mathcal{S}$ and therefore from the above we obtain that $\mathcal{F}(u * \varphi) = \hat{\varphi}\hat{u}$ for all $u \in \mathcal{S}'$ and $\varphi \in \mathcal{S}$. Again by the density of $\mathcal{D}$ in $\mathcal{S}$ and by continuity of the Fourier transform and the map $\varphi \mapsto u * \varphi$, we conclude that $\mathcal{F}(u * \varphi) = \hat{\varphi}\hat{u}$ also for $\varphi \in \mathcal{S}$. The other identity then follows by the first (Exercise 12.3). □

$\boxed{\textbf{Exercise}}$ **12.3.** Prove that the identity $\mathcal{F}(\varphi u) = \hat{\varphi} * \hat{u}$ holds for all $u \in \mathcal{S}'$ and $\varphi \in \mathcal{S}$ by using that $\mathcal{F}(u * \varphi) = \hat{\varphi}\hat{u}$ holds for all $u \in \mathcal{S}'$ and $\varphi \in \mathcal{S}$.

Like in Theorem 3.24 we have the following associativity rule.

Theorem 12.8. [7, Theorem on page 151] *If $u \in \mathcal{S}'$ and $\varphi, \psi \in \mathcal{S}$ then*

$$u * (\varphi * \psi) = (u * \varphi) * \psi. \tag{75}$$

Proof. By Theorem 12.7 and Lemma 12.2 we have

$$\begin{aligned}\mathcal{F}(u * (\varphi * \psi)) &= \mathcal{F}(\varphi * \psi)\hat{u} = \hat{\varphi}\hat{\psi}\hat{u},\\ \mathcal{F}((u * \varphi) * \psi) &= \hat{\psi}\mathcal{F}(u * \varphi) = \hat{\psi}\hat{\varphi}\hat{u}.\end{aligned}$$

As the Fourier transform is injective on $\mathcal{S}'$ (by Theorem 11.26), we have (75). $\square$

Remark 12.9. One other way to prove Theorem 12.8 is by extending Theorem 3.24 by a limiting argument: the identity (75) holds in case φ and ψ are testfunctions, and every Schwartz function can be approximated in the topology of $\mathcal{S}$ by testfunctions, see for example 10.6 (e).

As a tempered distribution is a distribution, we can convolve it with a distribution with compact support in the sense of Definition 6.9. Remember that every compactly supported distribution is a tempered distribution (so that its Fourier transform is defined).

We will show that such convolution is a tempered distribution for which the Fourier transform equals the product of the Fourier transform of each of the distributions. But first, we will show that the Fourier transform of a compactly supported distribution is a smooth function.

Lemma 12.10. *Let $v \in \mathcal{E}'$. Then $\hat{v} \in C_{\mathrm{p}}^{\infty}$.*

Proof. Let $\chi \in C_c^\infty$ be equal to 1 on $(\operatorname{supp} v)_\delta$ for some $\delta > 0$, so that $v = \chi v$. Then

$$\hat{v} = \mathcal{F}(\chi v) = \hat{\chi} * \hat{v}.$$

As χ is a Schwartz function, so is $\hat{\chi}$. Therefore $\hat{v} \in C_{\mathrm{p}}^\infty$ by Theorem 12.5. $\square$

As a direct consequence:

Lemma 12.11. *Let $f \in \mathcal{S}'$, then*

$$\operatorname{supp} \hat{f} \textit{ is compact } \Longrightarrow f \in C_{\mathrm{p}}^\infty. \tag{76}$$

Regarding Lemma 12.10, we can actually characterize the Fourier transforms of compactly supported distributions explicitly. For this we recall the notion of analytic function or holomorphic function.

Definition 12.12. We say that a function $f : \mathbb{C}^d \to \mathbb{C}$ is *entire* or call it an *entire function* if it is holomorphic everywhere in $\mathbb{C}^d$, by which we mean that for all $i \in \{1, \dots, d\}$ and $x \in \mathbb{C}^d$ the following limit exists

$$\lim_{h \to 0} \frac{f(x + he_i) - f(x)}{h}.$$

The proof of the following theorem requires a little work, the proof can be found in Rudin's book on Functional Analysis for example.

Theorem 12.13 (Paley-Wiener). [24, Theorem 7.23]

(a) *If* $v \in \mathcal{E}'$, $R > 0$, $\operatorname{supp} v \subset B(0, R)$, *$v$ has order k and*

$$f(z) = \langle v, e^{-2\pi \mathrm{i} \langle z, \cdot \rangle} \rangle \qquad (z \in \mathbb{C}^d), \tag{77}$$

then f is entire, $f|_{\mathbb{R}^d} = \hat{v}$ *and there exists a* $C > 0$ *such that*

$$|f(z)| \leq C(1 + |z|)^k e^{R|\Im z|} \qquad (z \in \mathbb{C}^d). \tag{78}$$

(b) *Conversely, if f is an entire function on* $\mathbb{C}^d$ *which satisfies* (78) *for some* $k \in \mathbb{N}_0$ *and* $C > 0$*, then there exists a* $v \in \mathcal{E}'$ *with support in* $B(0, R)$ *such that* (77) *holds.*

(77) is also called the Fourier-Laplace transform of v. More on this topic see for example [10, Section 10].

Theorem 12.14. *If* $v \in \mathcal{E}'$ *and $\hat{v}$ has compact support, then* $v = 0$.

Proof. We know that $\hat{v}$ is entire, which implies that it is analytic, meaning that for each point $z_0 \in \mathbb{C}^d$ there exist $(a_\alpha)_{\alpha \in \mathbb{N}_0^d}$ such that $\sum_{\alpha \in \mathbb{N}_0^d} a_\alpha (z - z_0)^\alpha$ is convergent and equals $\hat{v}(z)$ for all $z \in \mathbb{C}^d$. Therefore, if it is zero on an open set, it is equal to zero everywhere (as the coefficients a_α are determined by derivatives by Taylor 's theorem). $\square$

12.15. Observe that as the Fourier transform of a Gaussian function is again a Gaussian function, and therefore its support is the whole $\mathbb{R}^d$, we conclude from the Paley-Wiener theorem that Gaussian functions cannot be extended to entire functions on $\mathbb{C}^d$ that satisfies (78). Indeed, if we consider the Gaussian function $f : \mathbb{R}^d \to \mathbb{R}$ given by $f(x) = e^{-|x|^2}$ for $x \in \mathbb{R}^d$, then it is the restriction to $\mathbb{R}^d$ of the function $g : \mathbb{C}^d \to \mathbb{C}$ given by

$$g(x) = \sum_{n=0}^{\infty} \frac{(-x \cdot x)^n}{n!} \qquad (x \in \mathbb{C}^d).$$

For $a \in \mathbb{R}$ we have $g(\mathrm{i}a) = e^{a^2}$ from which we see that (78) is not satisfied.

In Definition 6.9 we defined the convolution between distributions u and v, of which at least one has compact support, by

$$u * v(\varphi) = u(\check{v} * \varphi) \qquad (\varphi \in \mathcal{D}).$$

This extends to convolution between a tempered distribution and a distribution of compact support, see Definition 12.17. Let us first check that the map $\varphi \mapsto u(\check{v} * \varphi)$ is indeed a tempered distribution.

12.16. As $\hat{v}$ is a smooth function of at most polynomial growth for a distribution v with compact support, and multiplication with such functions is a continuous operation on the Schwartz space (see 10.23), the function $\varphi \mapsto \hat{v}\hat{\varphi}$ is continuous as function $\mathcal{S} \to \mathcal{S}$. Therefore, as $v * \varphi = \mathcal{F}^{-1}(\hat{v}\hat{\varphi})$, the function $\mathcal{S} \to \mathcal{S}$ given by convolving with v, $\varphi \mapsto v * \varphi$ is continuous. Therefore, if $u \in \mathcal{S}'$, the function $\mathcal{S} \to \mathbb{F}$ given by $\varphi \mapsto u(\check{v} * \varphi)$ is continuous and therefore is a tempered distribution.

For the definition of $v * u$, we have already showed in Theorem 12.5 that $\varphi \mapsto u * \varphi$ is continuous as function $\mathcal{S} \to \mathcal{E}$. Therefore $\varphi \mapsto v(\check{u} * \varphi)$ defines a tempered distribution as well.

Definition 12.17. For $u \in \mathcal{S}'$ and $v \in \mathcal{E}'$ we define $u * v$ to be the tempered distribution given by

$$u * v(\varphi) = u(\check{v} * \varphi) \qquad (\varphi \in \mathcal{S}).$$

Moreover, we define $v * u$ to be the tempered distribution

$$v * u(\varphi) = v(\check{u} * \varphi) \qquad (\varphi \in \mathcal{S}).$$

12.18. As in 6.10, if u and v are tempered distributions of which at least one has compact support and if φ is a Schwartz function, then

$$(u * v) * \varphi = u * (v * \varphi).$$

This follows by 6.10 as $\mathcal{D}$ is dense in $\mathcal{S}$ and as the convolution with a tempered distribution is a continuous operation on $\mathcal{S}$.

Theorem 12.19. *Let $u \in \mathcal{S}'$ and $v \in \mathcal{E}'$. Then $u * v = v * u$ and*

$$\mathcal{F}(u * v) = \hat{v}\hat{u}.$$

Proof. That $u * v = v * u$ follows by Theorem 6.12 because $\mathcal{D}$ is dense in $\mathcal{S}$.

For $\varphi \in \mathcal{S}$ we have

$$\langle \mathcal{F}(u * v), \varphi \rangle = \langle u * v, \hat{\varphi} \rangle = u * v * \check{\hat{\varphi}}(0) = \langle u, (v * \check{\hat{\varphi}})^{\vee} \rangle.$$

Now $v * \check{\hat{\varphi}} = \mathcal{F}^{-1}(\mathcal{F}(v * \check{\hat{\varphi}})) = \mathcal{F}^{-1}(\hat{v}\varphi)) = [\mathcal{F}(\hat{v}\varphi))]^{\vee}$, therefore

$$\langle u, (v * \check{\hat{\varphi}})^{\vee} \rangle = \langle u, \mathcal{F}(\hat{v}\varphi) \rangle = \langle \hat{u}, \hat{v}\varphi \rangle = \langle \hat{v}\hat{u}, \varphi \rangle.$$

$\square$

13 Fourier multipliers

We will now turn to the definition of a Fourier multiplier. The idea is that we multiply on the level of the Fourier transform. Formally, if σ is a function and u a tempered distribution we will define $\sigma(\mathrm{D})u$ by $\mathcal{F}^{-1}(\sigma\hat{u})$. We will consider different conditions for which this formula makes sense.

Definition 13.1 (Fourier multiplier). For $\sigma \in C_{\mathrm{p}}^{\infty}$ we define $\sigma(\mathrm{D}) : \mathcal{S}' \to \mathcal{S}'$ by

$$\sigma(\mathrm{D})u = \mathcal{F}^{-1}(\sigma\hat{u}) \qquad (u \in \mathcal{S}'),$$

and call the function or operation $\sigma(\mathrm{D})$ a *Fourier multiplier.*

Example 13.2. By Theorem 11.26 we have $\partial^\beta u = \sigma(\mathrm{D})u$ for $\sigma = (2\pi\mathrm{i}\boldsymbol{\xi})^\beta$ and $\mathcal{T}_y u = \sigma(\mathrm{D})u$ for $\sigma = e^{-2\pi\mathrm{i}\langle\boldsymbol{\xi},y\rangle}$, i.e.,

$$\partial^\beta u = (2\pi\mathrm{i}\boldsymbol{\xi})^\beta(\mathrm{D})u, \qquad \mathcal{T}_y u = e^{-2\pi\mathrm{i}\langle\boldsymbol{\xi},y\rangle}(\mathrm{D})u \qquad (\beta \in \mathbb{N}_0^d, y \in \mathbb{R}^d). \tag{79}$$

By the commutativity of multiplication, we obtain that also Fourier multipliers commute. Moreover, if $\mathcal{F}^{-1}(\sigma)$ is compactly supported, then the Fourier multiplier of σ equal convolution with $\mathcal{F}^{-1}(\sigma)$:

Lemma 13.3. *Let $\sigma, \tau \in C_{\mathrm{p}}^{\infty}$. Then*

$$\tau(\mathrm{D})\sigma(\mathrm{D})u = (\sigma\tau)(\mathrm{D})u = \sigma(\mathrm{D})\tau(\mathrm{D})u \qquad (u \in \mathcal{S}').$$

Consequently, Fourier multiplier commute with partial differential operators with constant coefficients and with translations. Moreover,

$$\sigma(\mathrm{D})(l_\lambda u) = [(\sigma \circ l_\lambda)(\mathrm{D})u] \circ l_\lambda \qquad (u \in \mathcal{S}', \lambda > 0). \tag{80}$$

If $\sigma \in \mathcal{S}$ or if $\sigma \in C_{\mathrm{p}}^{\infty}$ is such that $\hat{\sigma} \in \mathcal{E}'$, or equivalently, if σ can be extended to an entire function on $\mathbb{C}^d$, then

$$\sigma(\mathrm{D})u = \mathcal{F}^{-1}(\sigma\hat{u}) = \mathcal{F}^{-1}(\sigma) * u \qquad (u \in \mathcal{S}').$$

Exercise 13.1. Prove (80).

We extend the notation of a Fourier multiplier in case σ is only smooth on set that contains the support of $\hat{u}$.

Exercise 13.2. Let ψ be a mollifier function and F be a closed set. Show that $\psi_\varepsilon * \mathbb{1}_F$ is a smooth function and that all derivatives are bounded, i.e., $\psi_\varepsilon * \mathbb{1}_F \in C_{\mathrm{b}}^{\infty}$.

13.4. Let F be a closed set in $\mathbb{R}^d$ and suppose that $\sigma : \mathbb{R}^d \to \mathbb{F}$ is smooth and of at most polynomial growth on $F_{3\delta}$ for some $\delta > 0$ (the latter means that $\sigma\mathbb{1}_U$ is of at most polynomial growth).

Let $\chi, \psi \in C_{\mathrm{p}}^{\infty}$ be smooth functions $\mathbb{R}^d \to [0,1]$ that equal 1 on F_δ and 0 outside $F_{2\delta}$ (that such functions exist follows for example by 5.12, see for the properties Exercise 13.2). Then $\sigma\chi$ and $\sigma\psi$ are in C_{p}^{∞}.

Suppose that $u \in \mathcal{D}'$ and $\operatorname{supp} u \subset F$. Then we know that $\sigma\chi u = \sigma\psi u$ by Exercise 3.2. Therefore we can define the multiplication of u with σ to be equal to $\sigma\chi u$, as this is independent of the choice of χ.

We use this to define the Fourier multiplier of σ by being the Fourier multiplier of $\sigma\chi$:

Definition 13.5. Let $F \subset \mathbb{R}^d$ be closed and $\sigma : \mathbb{R}^d \to \mathbb{F}$ be smooth on $F_{3\delta}^{\circ}$ for some $\delta > 0$. Let χ be a smooth function that equals 1 on F_δ and equals 0 outside $F_{2\delta}$. We define the *Fourier multiplier*

$$\sigma(\mathrm{D}) : \{u \in \mathcal{S}' : \operatorname{supp} \hat{u} \subset F\} \to \{u \in \mathcal{S}' : \operatorname{supp} \hat{u} \subset F\}, \qquad \sigma(\mathrm{D})u := \mathcal{F}^{-1}(\sigma\chi\hat{u}),$$

so that $\sigma(\mathrm{D})u = (\sigma\chi)(\mathrm{D})u$ for all $u \in \mathcal{S}'$ with $\operatorname{supp} \hat{u} \subset F$.

If σ instead is a smooth function that is only defined on $F_{3\delta}^{\circ}$, we can define the Fourier multiplier in an analogues way as it does not matter how σ is defined outside $F_{3\delta}^{\circ}$.

We will show that the so-called Bessel potentials are examples of Fourier multipliers in 13.8. Moreover, we show that the fractional Laplacian is an example of a Fourier multiplier in the sense of Definition 13.5, in 13.9.

But first we give some other examples of Fourier multipliers, or that can be interpreted as Fourier multipliers.

13.6 (*Other Fourier multipliers). There are some other operations that we could consider to be Fourier multipliers.

(a) Let $u \in \mathcal{S}'$ and $v \in \mathcal{E}'$. We have seen that $\hat{v} \in C_p^{\infty}$ and therefore $\hat{v}u$ and thus $\mathcal{F}^{-1}(\hat{v}u)$ defines a tempered distribution (as we always wrote multiplication of distributions u by functions f as fu, we write $\hat{v}u$ and not $u\hat{v}$). We could also write $u(\mathrm{D})$ for the function $\mathcal{E}' \to \mathcal{S}'$ given by $u(\mathrm{D})v = \mathcal{F}^{-1}(\hat{v}u)$. Observe that as $\mathcal{F}^{-1}(u)$ is a tempered distribution, by Theorem 12.19 we have

$$\mathcal{F}^{-1}(\hat{v}u) = \mathcal{F}^{-1}(u) * v. \tag{81}$$

Example. Observe that convolution is a special case of such Fourier multiplier as $u * v = \mathcal{F}^{-1}(\hat{u}\hat{v}) = v(\mathrm{D})\hat{u}$.

(b) Consider $u \in L^p$ and $\sigma \in L^1$ with $\hat{\sigma} \in L^1$. By Young's inequality we know that the convolution of an L^1 function with a L^p function is again an L^p function, so that $\mathcal{F}^{-1}(\sigma) * u \in L^p$. As u also represents a tempered distribution, $\hat{u}$ is also a tempered distribution. We can also argue that taking the convolution with $\mathcal{F}^{-1}(\sigma)$ can be seen as a Fourier multiplier. However, a priori the multiplication of σ with $\hat{u}$ is not defined as a (tempered) distribution, as σ need not be smooth. However, let us argue that the identity still makes sense in this case. By Corollary 3.5, equation (17) and Theorem 12.1

$$\langle \mathcal{F}^{-1}(\sigma) * u, \varphi \rangle = \langle u, \mathcal{F}(\sigma) * \varphi \rangle = \langle u, \mathcal{F}(\sigma \mathcal{F}^{-1}(\varphi)) \rangle.$$

Therefore we could also define

$$\sigma(\mathrm{D}) : L^p \to L^p, \qquad \sigma(\mathrm{D})u := \mathcal{F}^{-1}(\sigma) * u.$$

13.7. Observe that the Fourier multiplier definition in 13.6 (a) does not allow –in general– for composition of Fourier multipliers, as $u(\mathrm{D})v$ may not be compactly supported. For example, if w would be also tempered, the composition of Fourier multipliers $w(\mathrm{D})u(\mathrm{D})v$ would equal be $\mathcal{F}^{-1}(\hat{v}uw)$, but then one would need to make sense of uw.

13.8 (Bessel potentials). Let us consider the following partial differential equation for a given $g \in \mathcal{S}'$:

$$(1 - \Delta)u = g.$$

We can write $(1-\Delta)$ as a Fourier multiplier (by for example (79)), namely $(1-\Delta) = \sigma(\mathrm{D})$, for

$$\sigma(\xi) = (1 + 4\pi^2|\xi|^2) \qquad (\xi \in \mathbb{R}^d).$$

As this function is strictly positive, we can divide by it: We define $\tau : \mathbb{R}^d \to \mathbb{R}$ by

$$\tau(\xi) = (1 + 4\pi^2|\xi|^2)^{-1} \qquad (\xi \in \mathbb{R}^d).$$

It is not too difficult to show that $\tau \in C_{\mathrm{p}}^{\infty}$. As $\tau\sigma = \mathbb{1}$,

$$u = \mathcal{F}^{-1}(\tau\sigma\hat{u}) = \tau(\mathrm{D})\sigma(\mathrm{D})u = \tau(\mathrm{D})(1 - \Delta)u = \tau(\mathrm{D})g.$$

So we could view $\tau(\mathrm{D})$ as the inverse of the operator $(1 - \Delta)$.

With the use of the Fourier multiplier, one defines the operator $(1 - \Delta)^s$ for $s \in \mathbb{R}$ by

$$(1 - \Delta)^s u = \sigma^s(\mathrm{D})u, \tag{82}$$

where $\sigma^s(\xi) = (\sigma(\xi))^s$ for $\xi \in \mathbb{R}^d$ (observe that with this notation $\sigma^{-s} = \tau^s$). Even though for $s < d$ the function τ^s is not integrable on $\mathbb{R}^d$, the Fourier inverse of it as a tempered distribution is represented by a function that is smooth on $\mathbb{R}^d \setminus \{0\}$. The function (on $\mathbb{R}^d \setminus \{0\}$) $\mathcal{F}^{-1}(\tau^s)$ is also called a *Bessel potential*. For more on Bessel potentials we refer to [31, Section 7.7], [9, Section 4.3] and [12, Section 6.1.2]. In the last reference, not the function $\mathcal{F}^{-1}(\tau^s)$ but the operator $(1 - \Delta)^{-s}$ is called a Bessel potential. We come back to Bessel potentials in 13.13 and 13.15.

13.9 (Fractional Laplacian). As σ and τ in 13.8 are in C_{p}^{∞}, the Fourier multipliers $\sigma(\mathrm{D})$ and $\tau(\mathrm{D})$ are defined on the whole of $\mathcal{S}'$. Let us give an example of a Fourier multiplier in the sense of Definition 13.5. For $s \in \mathbb{R}$ the function $\sigma^s : \mathbb{R}^d \setminus \{0\} \to \mathbb{R}$ given by

$$\sigma^s(\xi) = |2\pi\xi|^{2s}, \qquad (\xi \in \mathbb{R}^d \setminus \{0\})$$

is in $C_{\mathrm{p}}^{\infty}(\mathbb{R}^d \setminus \{0\})$. For $k \in \mathbb{N}$ that the operator Δ^k equals the Fourier multiplier $\sigma^{2k}(\mathrm{D})$ (where σ^{2k} can actually be viewed as a smooth function on the whole of $\mathbb{R}^d$). Let $s \in \mathbb{R}$. We define the *fractional Laplacian* for $u \in \mathcal{S}'$ with $\operatorname{supp} \hat{u} \subset \mathbb{R}^d \setminus B(0,\delta)$ for some $\delta > 0$ by

$$(-\Delta)^s u = \sigma^s(\mathrm{D}) u.$$

In the rest of this section we get back to Sobolev spaces and describe them in terms of Fourier transforms and Fourier multipliers.

13.10 (Sobolev spaces described by their Fourier transforms). In Theorem 8.11 we have seen that H^k, being the Sobolev space given by

$$H^k = W^{k,2} = \{u \in \mathcal{D}' : \partial^\beta u \in L^2(\Omega) \text{ for all } \beta \in \mathbb{N}_0^d \text{ with } |\beta| \le k\}.$$

is a Hilbert space with norm

$$\|u\|_{H^k} = \Big(\sum_{\alpha \in \mathbb{N}_0^d : |\alpha| \le k} \|\partial^\alpha u\|_{L^2}^2 \Big)^{\frac{1}{2}} \qquad (u, v \in H^k).$$

It turns out that H^k be described using Fourier transforms (see Exercise 13.3) , as follows

$$H^k = \{u \in \mathcal{S}' : (1 + |\boldsymbol{\xi}|)^k \hat{u} \in L^2\}, \tag{83}$$

moreover, the norm is equivalent to $u \mapsto \|(1 + |\boldsymbol{\xi}|)^k \hat{u}\|_{L^2}$, which means there exists a $C > 1$ such that

$$\frac{1}{C}\|u\|_{H^k} \le \|(1 + |\boldsymbol{\xi}|)^k \hat{u}\|_{L^2} \le C\|u\|_{H^k} \qquad (u \in H^k). \tag{84}$$

To prove (83) the Multinomial theorem might be beneficial.

Theorem 13.11 (Multinomial theorem). *For $x = (x_1, \dots, x_d) \in \mathbb{F}^d$ and $k \in \mathbb{N}$*

$$(x_1 + \cdots + x_d)^k = \sum_{\alpha \in \mathbb{N}_0^d : |\alpha| = k} \binom{k}{\alpha} x^\alpha, \tag{85}$$

where with $\alpha! = \alpha_1!\alpha_2! \cdots \alpha_d!$,

$$\binom{k}{\alpha} = \frac{k!}{\alpha!} = \frac{k!}{\alpha_1!\alpha_2! \cdots \alpha_d!}.$$

Proof. This follows by induction on the induction. For $d = 1$ the formula is trivial for all $k \in \mathbb{N}$. For $d = 2$ it is the usual binomial formula. Suppose (85) holds for a fixed $d \in \mathbb{N}$ and for any $k \in \mathbb{N}$. Then for $y = x_1 + \cdots + x_d$ and $z = (x_1, \dots, x_d)$ we have

$$\begin{aligned}(y + x_{d+1})^k &= \sum_{m \in \mathbb{N}_0 : m \leq k} \binom{k}{m} y^m x_{d+1}^{k-m} \\ &= \sum_{m \in \mathbb{N}_0 : m \leq k} \binom{k}{m} \sum_{\alpha \in \mathbb{N}_0^d : |\alpha| = k} \binom{m}{\alpha} z^\alpha x_{d+1}^{k-m},\end{aligned}$$

as for $\beta = (\alpha_1, \dots, \alpha_d, k - m)$ we have $|\beta| = k$ and

$$\binom{k}{m}\binom{m}{\alpha} = \frac{k!}{(k-m)!m!} \frac{m!}{\alpha_1! \cdots \alpha_d!} = \binom{k}{\beta},$$

it follows that (85) is valid also for for $d + 1$. □

$\boxed{\textbf{Exercise}}$ **13.3.** Show that (83) holds and show the existence of a $C > 1$ such that (84) holds.

The above equivalence of norms lets us extend the notation of H^k spaces to non-integer values of k, as follows.

Definition 13.12. For $s \in \mathbb{R} \setminus \mathbb{N}_0$ we define the *fractional Sobolev space* H^s by

$$H^s = \{u \in \mathcal{S}' : (1 + |\boldsymbol{\xi}|)^s \hat{u} \in L^2\}, \tag{86}$$

and define a norm on H^s by

$$\|u\|_{H^s} = \|(1 + |\boldsymbol{\xi}|)^s \hat{u}\|_{L^2} \qquad (u \in H^s).$$

13.13. As

$$\frac{1}{\sqrt{2}}(1 + |\xi|) \leq (1 + |\xi|^2)^{\frac{1}{2}} \leq (1 + |\xi|) \qquad (\xi \in \mathbb{R}^d),$$

also the norms $u \mapsto \|(1 + |\boldsymbol{\xi}|)^s \hat{u}\|_{L^2}$ and $u \mapsto \|(1 + |\boldsymbol{\xi}|^2)^{\frac{s}{2}} \hat{u}\|_{L^2}$ are equivalent. Therefore, by by Plancherels equality and by the definition of $(1 - \Delta)^s$ as in (82) we have that $\|\cdot\|_{H^s}$ is equivalent to $u \mapsto \|(1 - \Delta)^{\frac{s}{2}} u\|_{L^2}$.

Example 13.14. We have already seen that $\hat{\delta} = \mathbb{1}$. As $(1 + |\boldsymbol{\xi}|)^s$ is in L^2 if and only if $2s < -d$ by Lemma 10.13, it follows that $\delta \in H^s$ if and only if $s < -\frac{d}{2}$.

13.15 (Bessel potential spaces). We have only considered a generalisation for the Sobolev space $W^{k,p}$ for $p = 2$ but for any p one can actually define fractional Sobolev spaces, or also called *Bessel potential spaces.* In [32, Section 1.2] for example, is shown that (with the definition of $(1-\Delta)^{\frac{k}{2}}$ by (82))

$$W^{k,p} = \{f \in \mathcal{S}' : (1-\Delta)^{\frac{k}{2}} f \in L^p\}, \tag{87}$$

and that $\|\cdot\|_{W^{k,p}}$ is equivalent to

$$f \mapsto \|(1-\Delta)^{\frac{k}{2}} f\|_{L^p}. \tag{88}$$

Similar to Definition 13.12 one defines the *fractional Sobolev space* H_p^s for $s \in \mathbb{R}$ by replacing "k" in (87) by "s":

$$H_p^s = \{u \in \mathcal{S}' : (1-\Delta)^{\frac{s}{2}} u \in L^p\}, \tag{89}$$

and define a norm on H_p^s by

$$\|u\|_{H_p^s} = \|(1-\Delta)^{\frac{s}{2}} u\|_{L^p} \qquad (u \in \mathcal{S}').$$

Then $H_p^k = W^{k,p}$ and by Plancherel's identity it follows that $H_2^s = H^s$.

14 Bernstein and the Hörmander-Mikhlin inequalities

In this section we consider Fourier multipliers of tempered distributions of which their Fourier transform has compact support, either in a ball or annulus. By Lemma 12.11 we know that these tempered distributions are C_{p}^∞ functions. We will prove the Bernstein inequality and the Hörmander–Mikhlin inequality. In the next section we will define Besov spaces by the Littlewood–Paley decomposition. In 10.6 (f) we have seen that a tempered distribution u can be written as the sum over $\chi_n u$, where χ_n is a certain partition of unity. The Littlewood–Paley decomposition happens on the level of Fourier transform, one decomposes a tempered distribution u to be the sum over $\chi_n(\mathrm{D})u$ for a certain partition of unity χ_n of which each function is either supported in a ball or annulus.

Definition 14.1 (Annulus). An *annulus* in $\mathbb{R}^d$ is a set of the form $\{x \in \mathbb{R}^d : r \le |x| \le s\}$, for $s, r \in \mathbb{R}$ with $0 < r < s$. We will write

$$A(r,s) = \{x \in \mathbb{R}^d : r \le |x| \le s\},$$

and $A^\circ(r,s)$ for its interior $\{x \in \mathbb{R}^d : r < |x| < s\}$.

Before we turn to the Bernstein inequality, we prove how a tempered distribution with Fourier support in an annulus can be described as a convolution of a function with the k-th order derivatives of u.

Lemma 14.2. *Let $\mathcal{A}$ be an annulus and $\mathcal{B}$ be a ball around the origin in $\mathbb{R}^d$. Let $\chi \in C_c^\infty$ be equal to* 1 *on $\mathcal{B}_\delta$ for some $\delta > 0$. Let $\phi \in C_c^\infty$ be supported in an annulus and be equal to* 1 *on $\mathcal{A}_\delta$ for some $\delta > 0$. Let $k \in \mathbb{N}_0$.*

(a) *If u is a tempered distribution with* $\operatorname{supp} \hat{u} \subset \mathcal{B}$*, then for all* $\alpha \in \mathbb{N}_0^d$

$$\partial^\alpha u = h_\alpha * u,$$

where $h_\alpha = \partial^\alpha \mathcal{F}^{-1}(\chi)$*.*

(b) *If u is a tempered distribution with* $\operatorname{supp} \hat{u} \subset \mathcal{A}$*, then*

$$u = \sum_{\alpha \in \mathbb{N}_0^d : |\alpha| = k} g_\alpha * \partial^\alpha u,$$

where

$$g_\alpha = \binom{k}{\alpha} \mathcal{F}^{-1} \left((-2\pi \mathrm{i} \boldsymbol{\xi})^\alpha |2\pi \boldsymbol{\xi}|^{-2k} \phi \right).$$

(c) *There exists a $C > 0$ such that for all $r \in [1, \infty]$*

$$\|h_\alpha\|_{L^r}, \|g_\alpha\|_{L^r} \le C^{k+1} \qquad (\alpha \in \mathbb{N}_0^d, |\alpha| = k).$$

Proof. (a) follows from the fact that $\hat{u} = \chi \hat{u}$.

For (b), as $\hat{u}$ is supported on an annulus, we can divide (and multiply) by $|2\pi\boldsymbol{\xi}|^{2k}$. By the multinormial theorem (see Theorem 13.11, take $x_i = |\xi_i|^2 = (-\mathrm{i}\xi_i)(\mathrm{i}\xi_i)$):

$$|2\pi\xi|^{2k} = \sum_{\alpha \in \mathbb{N}_0^d \ : \ |\alpha| = k} \binom{k}{\alpha} (-2\pi \mathrm{i} \xi)^\alpha (2\pi \mathrm{i} \xi)^\alpha \qquad (\xi \in \mathbb{R}^d). \tag{90}$$

With this we have by Lemma 13.3 and as $(2\pi\mathrm{i}\boldsymbol{\xi})^\alpha(\mathrm{D}) = \partial^\alpha$,

$$\begin{aligned} u = \mathbb{1}(\mathrm{D})u &= \left(\frac{\sum_{\alpha \in \mathbb{N}_0^d : |\alpha| = k} \binom{k}{\alpha} (-2\pi \mathrm{i} \boldsymbol{\xi})^\alpha (2\pi i \boldsymbol{\xi})^\alpha}{|2\pi\boldsymbol{\xi}|^{2k}} \right) (\mathrm{D}) u \\ &= \left(\frac{\sum_{\alpha \in \mathbb{N}_0^d : |\alpha| = k} \binom{k}{\alpha} (-2\pi \mathrm{i} \boldsymbol{\xi})^\alpha}{|2\pi\boldsymbol{\xi}|^{2k}} \right) (\mathrm{D}) \ \partial^\alpha u. \end{aligned}$$

(c) By Corollary A.10, with $C_1 = 1 + \|(1 + |\boldsymbol{x}|^2)^{-d}\|_{L^1}$ (which if finite by Lemma 10.13), we have for all $f \in \mathcal{S}$

$$\|f\|_{L^r} \le \|f\|_{L^1} + \|f\|_{L^\infty} \le C_1 \|(1 + |\boldsymbol{x}|^2)^d f\|_{L^\infty}.$$

Let $\alpha \in \mathbb{N}_0^d$, $|\alpha| = k$. We first consider the bound for h_α. As $\|\hat{f}\|_{L^\infty} \le \|f\|_{L^1}$, by (68) we have

$$\|(1 + |\boldsymbol{x}|^2)^d h_\alpha\|_{L^\infty} \le \|(1 - \Delta)^d (2\pi \mathrm{i} \boldsymbol{\xi})^\alpha \chi\|_{L^1}.$$

By the Multinomial theorem (Theorem 13.11) we have

$$(1-\Delta)^d = \sum_{\beta\in\mathbb{N}_0^{d+1}:|\beta|=d} \binom{d}{\beta} \prod_{i=1}^{d} (-\partial_i^2)^{\beta_i}.$$

As $\sum_{\beta\in\mathbb{N}_0^{d+1}:|\beta|=d} \binom{d}{\beta} \le (d+1)^d$, for C_2 being the Lebesgue measure of the support of χ,

$$\|(1-\Delta)^d (2\pi\mathrm{i}\boldsymbol{\xi})^\alpha \chi\|_{L^1} \le C_2(d+1)^d \|(2\pi\mathrm{i}\boldsymbol{\xi})^\alpha\chi\|_{C^{2d}}$$

and by Leibniz formula 5.3 there exists a $C_3 > 0$ (only depending on d) such that

$$\|(2\pi\mathrm{i}\boldsymbol{\xi})^\alpha\chi\|_{C^{2d}} \le C_3 \|(2\pi\mathrm{i}\boldsymbol{\xi})^\alpha\|_{C^{2d}(\operatorname{supp}\chi)} \|\chi\|_{C^{2d}}.$$

As $|\alpha| = k$, by applying Leibniz formula k times, for $C_4 > 0$ given by

$$C_4 = \max_{i=1}^{d} \|2\pi\mathrm{i}\boldsymbol{\xi_i}\|_{C^{2d}(\operatorname{supp}\chi)},$$

we have

$$\|(2\pi\mathrm{i}\boldsymbol{\xi})^\alpha\|_{C^{2d}(\operatorname{supp}\chi)} \le C_3^k C_4^k.$$

Therefore by choosing $C > 0$ large enough (for example $C = C_1 + C_2(d+1)^d + C_3 + C_4$) we obtain the bound for h_α.

For g_α the factor $\binom{k}{\alpha}$ can be bounded by d^k. The rest is very similar to h_α: by following the lines above with $\chi = |2\pi\boldsymbol{\xi}|^{-2k}\phi$, we have

$$\|g_\alpha\|_{L^r} \le d^k C^{k+1} \||2\pi\boldsymbol{\xi}|^{-2k}\phi\|_{C^{2d}}.$$

By applying Leibniz formula again on the last term k times, we get another factor M^k for $M = \||2\pi\boldsymbol{\xi}|^{-2}\|_{C^{2d}(\operatorname{supp}\phi)}$. □

Now we will use the descriptions of u and $\partial^\alpha u$ by the convolutions in Lemma 14.2 together with the Young's inequality in Theorem 14.3.

Exercise **14.1.** Show that for $f \in L^p$ and $\lambda > 0$

$$\|l_\lambda f\|_{L^p} = \lambda^{-\frac{d}{p}} \|f\|_{L^p}, \tag{91}$$

where $l_\lambda f(x) = f(\lambda x)$ (as in 10.5).

Theorem 14.3 (Bernstein inequality). [2, Lemma 2.1] *Let $\mathcal{A}$ be an annulus and $\mathcal{B}$ be a ball around the origin in $\mathbb{R}^d$. There exists a $C > 0$ such that for all $k \in \mathbb{N}$ and $p, q \in [1,\infty]$ with $q \ge p$ and any $u \in L^p$ we have for all $\lambda > 0$*

$$\operatorname{supp}\hat{u} \subset \lambda\mathcal{B} \Longrightarrow \max_{\alpha\in\mathbb{N}_0^d:|\alpha|=k} \|\partial^\alpha u\|_{L^q} \le C^{k+1}\lambda^{k+d(\frac{1}{p}-\frac{1}{q})}\|u\|_{L^p}, \tag{92}$$

$$\operatorname{supp}\hat{u} \subset \lambda\mathcal{A} \Longrightarrow C^{-k-1}\lambda^k\|u\|_{L^p} \le \max_{\alpha\in\mathbb{N}_0^d:|\alpha|=k} \|\partial^\alpha u\|_{L^p} \le C^{k+1}\lambda^k\|u\|_{L^p}. \tag{93}$$

Proof. First we argue that we may restrict to the case $\lambda = 1$. Let $\mathcal{C}$ denote either $\mathcal{A}$ or $\mathcal{B}$. If $\operatorname{supp} \hat{u} \subset \lambda\mathcal{C}$, then $\operatorname{supp} \hat{v} \subset \mathcal{C}$ for $v = l_{\frac{1}{\lambda}} u$ as $\hat{v} = \lambda^d l_\lambda \hat{u}$. For the norms we have for $\alpha \in \mathbb{N}_0^d$ with $|\alpha| = k$

$$\|v\|_{L^p} = \lambda^{\frac{d}{p}} \|u\|_{L^p}, \qquad \|\partial^\alpha v\|_{L^q} = \lambda^{-k} \|l_{\frac{1}{\lambda}}(\partial^\alpha u)\|_{L^q} = \lambda^{-k+\frac{d}{q}} \|\partial^\alpha u\|_{L^q}.$$

Hence, we may indeed assume $\lambda = 1$.

• Assume that $\operatorname{supp} \hat{u} \subset \mathcal{B}$ and h_α be as in Lemma 14.2 (a), so that $\partial^\alpha u = h_\alpha * u$. Let $r \in [1, \infty]$ be such that $\frac{1}{q} + 1 = \frac{1}{r} + \frac{1}{p}$. By Young's inequality (Theorem 3.4) and Lemma 14.2 (c) there exists a $C > 0$ such that

$$\|\partial^\alpha u\|_{L^q} \le \|h_\alpha\|_{L^r} \|u\|_{L^p} \le C^{k+1} \|u\|_{L^p}.$$

• The upper bound in (93) follows immediately from (92). Let g_α be as in Lemma 14.2 (b). By Young's inequality

$$\|u\|_{L^p} \le \sum_{\alpha \in \mathbb{N}_0^d : |\alpha| = k} \|g_\alpha\|_{L^1} \|\partial^\alpha u\|_{L^p} \le \Big(\max_{\alpha \in \mathbb{N}_0^d : |\alpha| = k} \|\partial^\alpha u\|_{L^p} \Big) \Big(\sum_{\alpha \in \mathbb{N}_0^d : |\alpha| = k} \|g_\alpha\|_{L^1} \Big).$$

So that the lower bound follows by Lemma 14.2 (c). □

$\boxed{\textbf{Exercise}}$ **14.2.** Let $u \in C_{\mathrm{b}}^k$ (see Definition 10.8) for some $k \in \mathbb{N}$. Let $\rho \in C_c^\infty$ be supported in an annulus $\mathcal{A}$, so that $\rho_\lambda := l_{\frac{1}{\lambda}} \rho$ is supported in $\lambda\mathcal{A}$. Show that there exists a $C > 0$ such that

$$\|\rho_\lambda(\mathrm{D}) u\|_{L^\infty} \le C^{k+1} \lambda^{-k} \|u\|_{C^k}.$$

14.4 (Towards Besov spaces). As we see from Exercise 14.2, if a function has bounded derivatives of a certain order, this implies a decay on the L^∞ norm of the Fourier multiplier of a function ρ_λ that is supported in $\lambda\mathcal{A}$ as $\lambda \to \infty$.

One could also say that by multiplying the Fourier transform $\hat{u}$ by ρ_λ, one takes the frequencies of order λ. The bound then gives a control of the frequencies of this order. In the theory of Besov spaces, this control on the frequencies is the behind describing the regularity of a distribution. We will get back to this later. Observe that this agrees with the fractional Sobolev space H^s introduced in Definition 13.12, in which we also obtain the regularity s by describing a control on the frequencies of the distributions.

We will now turn to a lemma that describes the effect of certain Fourier multipliers on L^p norms distributions with support in annuli or balls (Lemma 14.8). This will later be used to describe the increase of decrease of regularity with respect to certain Fourier multipliers.

Let us still first introduce some notation and auxiliary facts:

14.5. Let us write $l^*_\lambda = \lambda^{-d} l_{\frac{1}{\lambda}}$ for $\lambda > 0$. The * notation agrees with the fact that l^*_λ is the adjoint of l_λ as an operator on L^2, i.e.,

$$\langle l_\lambda f, g\rangle_{L^2} = \langle f, l^*_\lambda g\rangle_{L^2} \qquad (f, g \in L^2).$$

By Theorem 11.26 we know that for a distribution $u \in \mathcal{S}'$,

$$\mathcal{F}(l_\lambda u) = l^*_\lambda \hat{u} \qquad \mathcal{F}(l^*_\lambda u) = l_\lambda \hat{u}. \tag{94}$$

Observe that by (91)

$$\|l^*_\lambda f\|_{L^1} = \|f\|_{L^1} \qquad (\lambda > 0, f \in L^1). \tag{95}$$

Definition 14.6 (Mikhlin norm). Let $m \in \mathbb{R}$ and $k = 2\lfloor 1 + \frac{d}{2}\rfloor$. For $\sigma \in C^k(\mathbb{R}^d \setminus \{0\})$ we define its *Mikhlin norm* of order $m \in \mathbb{R}$ by

$$\mathfrak{M}_m(\sigma) = \max_{\alpha \in \mathbb{N}_0^d : |\alpha| \le k} \sup_{x \in \mathbb{R}^d \setminus \{0\}} |x|^{|\alpha| - m} |\partial^\alpha \sigma(x)|.$$

Observe that $\mathfrak{M}_m(\sigma) < \infty$ if and only if there exists a $C > 0$ such that

$$|\partial^\alpha \sigma(x)| \le C|x|^{m - |\alpha|} \qquad (x \in \mathbb{R}^d \setminus \{0\}, \alpha \in \mathbb{N}_0^d, |\alpha| \le k). \tag{96}$$

In the case that the norm is finite one can of course take $C = \mathfrak{M}_m(\sigma)$ in (96).

14.7. Observe that for $m \le 0$ and for $\sigma \in \mathcal{S}$ the Mikhlin norm $\mathfrak{M}_m(\sigma)$ is finite as every derivative decays faster than polynomially.

Moreover, observe that we have the following scaling relation

$$\mathfrak{M}_m(l_\lambda \sigma) = \lambda^m \mathfrak{M}_m(\sigma) \qquad (\sigma \in C^k(\mathbb{R}^d \setminus \{0\})). \tag{97}$$

$\boxed{\textbf{Exercise}}$ **14.3.** Is the Mikhlin norm $\mathfrak{M}_m(\sigma)$ also finite for all $\sigma \in \mathcal{S}$ and $m > 0$?

Lemma 14.8 (Hörmander-Mikhlin inequality). [2, Lemma 2.2] *Let $m \in \mathbb{R}$ and $k = 2\lfloor 1 + \frac{d}{2}\rfloor$.*

(a) *Let $\mathcal{A}$ be an annulus in $\mathbb{R}^d$. There exists a $C > 0$ such that for all $p \in [1, \infty]$, $\lambda > 0$, all $\sigma \in C^\infty(\mathbb{R}^d \setminus \{0\})$ and all $u \in L^p$*

$$\operatorname{supp} \hat{u} \subset \lambda\mathcal{A} \implies \|\sigma(\mathrm{D})u\|_{L^p} \le C\mathfrak{M}_m(\sigma)\lambda^m \|u\|_{L^p}. \tag{98}$$

(b) *Let $\mathcal{B}$ be a ball around the origin. There exists a constant C such that for all $p \in [1, \infty]$, $\lambda > 0$, all $\sigma \in C^\infty(\mathbb{R}^d)$ and all $u \in L^p$*

$$\operatorname{supp} \hat{u} \subset \lambda\mathcal{B} \implies \|\sigma(\mathrm{D})u\|_{L^p} \le C\mathfrak{M}_m(\sigma)\lambda^m \|u\|_{L^p}. \tag{99}$$

Proof. The proof of (a) and (b) are very similar. Without loss of generality we may assume $\mathfrak{M}_m(\sigma) < \infty$. For (a) assume that $\phi \in C_c^\infty$ is supported in an annulus be such that $\phi = 1$ on $\mathcal{A}_\delta$ for some $\delta > 0$ and for (b) assume that ϕ is instead supported in a ball such that $\phi = 1$ on $\mathcal{B}_\delta$.

Let $\lambda > 0$. Let $u \in L^p$ be such that its Fourier transform is supported in $\lambda\mathcal{A}$. Then $\hat{u} = (l_{\frac{1}{\lambda}}\phi)\hat{u}$ and thus

$$\sigma(\mathrm{D})u = (\sigma l_{\frac{1}{\lambda}}\phi)(\mathrm{D})u = \mathcal{F}^{-1}(\sigma l_{\frac{1}{\lambda}}\phi) * u.$$

By Young's inequality we have $\|\sigma(\mathrm{D})u\|_{L^p} \le \|\mathcal{F}^{-1}(\sigma l_{\frac{1}{\lambda}}\phi)\|_{L^1}\|u\|_{L^p}$. As by (94)

$$\mathcal{F}^{-1}(\sigma l_{\frac{1}{\lambda}}\phi) = \mathcal{F}^{-1}(l_{\frac{1}{\lambda}}(\phi l_\lambda \sigma)) = l_{\frac{1}{\lambda}}^* \mathcal{F}^{-1}(\phi l_\lambda \sigma),$$

by (95) it suffices to show that there exists a $C > 0$ such that

$$\|\mathcal{F}^{-1}(\phi l_\lambda \sigma)\|_{L^1} \le C\mathfrak{M}_m(\sigma)\lambda^m.$$

Observe that $\frac{k}{2} = \lfloor 1 + \frac{d}{2} \rfloor$ is the smallest integer such that $(1 + |\boldsymbol{x}|^2)^{-\frac{k}{2}}$ is integrable. We multiply and divide by this function to estimate the L^1 norm by the integral of $(1 + |\boldsymbol{x}|^2)^{-\frac{k}{2}}$, for which we write M, and the supremum norm of the rest

$$\|\mathcal{F}^{-1}(\phi l_\lambda \sigma)\|_{L^1} \le M\|(1 + |\boldsymbol{x}|^2)^{\frac{k}{2}}\mathcal{F}^{-1}(\phi l_\lambda \sigma)\|_{L^\infty}.$$

With $c_{\alpha,\beta} \in \mathbb{R}$ being such that

$$(1 - (2\pi)^{-2}\Delta)^{\frac{k}{2}}(fg) = \sum_{\alpha,\beta \in \mathbb{N}_0^d : |\alpha| + |\beta| \le k} c_{\alpha,\beta}\partial^\alpha f \cdot \partial^\beta g,$$

we have by Theorem 11.5 and Theorem 11.26

$$\begin{aligned}
\|(1 + |\boldsymbol{x}|^2)^{\frac{k}{2}}\mathcal{F}^{-1}(\phi l_\lambda \sigma)\|_{L^\infty} &= \|\mathcal{F}^{-1}((1 - (2\pi)^{-2}\Delta)^{\frac{k}{2}}(\phi l_\lambda \sigma))\|_{L^\infty} \\
&\le \sum_{\alpha,\beta \in \mathbb{N}_0^d : |\alpha| + |\beta| \le k} |c_{\alpha,\beta}| \cdot \|\partial^\alpha \phi \cdot \partial^\beta l_\lambda \sigma\|_{L^1} \\
&\le \sum_{\alpha,\beta \in \mathbb{N}_0^d : |\alpha| + |\beta| \le k} |c_{\alpha,\beta}| \cdot \|\partial^\alpha \phi\|_{L^1} \sup_{\xi \in \operatorname{supp}\phi} |\partial^\beta l_\lambda \sigma(\xi)|.
\end{aligned}$$

We estimate the latter by

$$\begin{aligned}
\sup_{\xi \in \operatorname{supp}\phi} |\partial^\beta l_\lambda \sigma(\xi)| = \sup_{\xi \in \operatorname{supp}\phi} \lambda^{|\beta|}|\partial^\beta \sigma(\lambda\xi)| &\le \sup_{\xi \in \operatorname{supp}\phi} \lambda^{|\beta|}\mathfrak{M}_m(\sigma)|\lambda\xi|^{m-|\beta|} \\
&\le \lambda^m \mathfrak{M}_m(\sigma) \sup_{\xi \in \operatorname{supp}\phi} |\xi|^{m-|\beta|}.
\end{aligned}$$

Hence from the above estimates we conclude that there exists a C that only depends on d, m, ϕ (which only depends on $\mathcal{A}$ or $\mathcal{B}$) and k, such that (98) and (99) hold. $\square$

Remark 14.9. We called Lemma 14.8 the Hörmander-Mikhlin inequality as it is strongly related to what in literature is called the Hörmander–Mikhlin multiplier theorem, see [29] for example, or [19] (in Russian) or [16] for the work of Mikhlin and Hörmander. As unfortunately happens with names from languages with different alphabets, we also found instead of Mikhlin the names Michlin or Mihlin.

Their theorem dealt with the case $m = 0$. See for example also [17, Theorem 5.5.10] (which looks again a bit different). We decided to call the norm the Mikhlin norm as that seems to align with the literature and it seems that the Hörmander and Mikhlin statements are slightly different.

$\boxed{\textbf{Exercise}}$ **14.4.** The upper bounds in the Bernstein inequalities can also be proved using the Hörmander–Mikhlin inequalities, as follows. Prove that there exists an $M > 0$ such that for all $k \in \mathbb{N}$ and $\alpha \in \mathbb{N}_0^d$ with $|\alpha| = k$,

$$\mathfrak{M}_k((2\pi \mathrm{i}\boldsymbol{x})^\alpha) = M^k.$$

Conclude the upper bounds in (93) and (92) for $q = p$ from the Hörmander-Mikhlin inequality.

Let us regard the applicability of the Hörmander–Mikhlin inequality for the Fourier multipliers we have considered in 13.8 and 13.9, namely $(1 - \Delta)^s$ and $(-\Delta)^s$. For $\sigma = |2\pi\boldsymbol{x}|^s$ the Mikhlin norm of order s is finite (as we will see), but for $(1 + |\boldsymbol{x}|^2)^{\frac{s}{2}}$ it is not for s and m being strictly positive, as the function at zero equals 0 but $|\boldsymbol{x}|^m$ equals zero for $m > 0$. However, if we apply the Fourier multiplier only to those $u \in L^p$ that have the support of their Fourier transform bounded away from zero, we can still obtain a bound like (98). We state the exact statement in Lemma 14.12, after extending the notion of the Mikhlin norm to a seminorm that only considers the space $\mathbb{R}^d$ without a ball at the origin.

Definition 14.10 (Mikhlin seminorm). Let $m \in \mathbb{R}$ and $k = 2\lfloor 1 + \frac{d}{2} \rfloor$. For $\sigma \in C^k(\mathbb{R}^d \setminus \{0\})$ we define its *Mikhlin seminorm* of order $m \in \mathbb{R}$ on the complement of a ball of radius θ by

$$\mathfrak{M}_{m,\theta}(\sigma) = \max_{\alpha \in \mathbb{N}_0^d : |\alpha| \le k} \sup_{x \in \mathbb{R}^d \setminus B(0,\theta)} |x|^{|\alpha| - m} |\partial^\alpha \sigma(x)|.$$

14.11. As for the Mikhlin norm, $\mathfrak{M}_{m,\theta}(\sigma) < \infty$ if and only if there exists a $C > 0$ such that

$$|\partial^\alpha \sigma(x)| \le C|x|^{m - |\alpha|} \qquad (x \in \mathbb{R}^d \setminus B(0, \theta), \alpha \in \mathbb{N}_0^d, |\alpha| \le k). \tag{100}$$

Moreover, as σ is smooth on $\mathbb{R}^d \setminus \{0\}$, if $\mathfrak{M}_{m,\theta}(\sigma)$ is finite for some $\theta > 0$, then it is finite for all $\theta > 0$.

Lemma 14.12 (Hörmander-Mikhlin inequality 2). *Let $m \in \mathbb{R}$, $k = 2\lfloor 1 + \frac{d}{2} \rfloor$ and $\theta > 0$. Let $\mathcal{A}$ be an annulus in $\mathbb{R}^d$. There exist $C > 0$ and $a > 0$ such that for all*

$p \in [1,\infty]$, $\lambda > \theta$, all $\sigma \in C^\infty(\mathbb{R}^d \setminus \{0\})$ and all $u \in L^p$ which Fourier transform is supported in $\lambda\mathcal{A}$,

$$\|\sigma(\mathrm{D})u\|_{L^p} \le C\mathfrak{M}_{m,a\theta}(\sigma)\lambda^m\|u\|_{L^p}. \tag{101}$$

Proof. Without loss of generality, we may assume that there exists a $\theta > 0$ such that $\mathfrak{M}_{m,\theta}(\sigma) < \infty$ (so that this is actually finite for all θ). Let $r, s \in (0,\infty)$, $r < s$ be such that $\mathcal{A} = A(r,s)$. Let $\chi \in C_{\mathrm{p}}^\infty(\mathbb{R}^d)$ be equal to 1 on $\mathbb{R}^d \setminus B(0,\frac{\theta}{2}r)$ and 0 on $B(0,\frac{\theta}{4}r)$. Then $\sigma(\mathrm{D})u = (\sigma\chi)(\mathrm{D})u$ and by Leibniz formula there exists a $C > 0$ such that

$$\mathfrak{M}_m(\sigma\chi) \le \|\chi\|_{C^k}\mathfrak{M}_{m,\frac{\theta r}{4}}(\sigma).$$

Therefore (101) follows from Lemma 14.8 (a). □

14.13. For $\sigma \in C^k(\mathbb{R}^d)$ there exists a $\theta > 0$ such that $\mathfrak{M}_{m,\theta}(\sigma) < \infty$ if and only if

$$\max_{\alpha\in\mathbb{N}_0^d:|\alpha|\le k}\sup_{x\in\mathbb{R}^d}(1+|x|)^{|\alpha|-m}|\partial^\alpha\sigma(x)| < \infty,$$

(see Exercise 14.5) or equivalently, there exists a $C > 0$ such that for all $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \le k$

$$|\partial^\alpha\sigma(x)| \le C(1+|x|)^{m-|\alpha|} \qquad (x \in \mathbb{R}^d).$$

$\boxed{\textbf{Exercise}}$ **14.5.** Prove the statement in 14.13.

Lemma 14.14. *Let $l \in \mathbb{R}$. For all $x \neq 0$*

$$\partial^\alpha|x|^l = \begin{cases} \sum_{i=0}^n Q_{2i}^\alpha(x)|x|^{l-2(n+i)} & \text{if } |\alpha| = 2n \text{ for some } n \in \mathbb{N}_0, \\ \sum_{i=0}^n Q_{2i+1}^\alpha(x)|x|^{l-2(n+i+1)} & \text{if } |\alpha| = 2n+1 \text{ for some } n \in \mathbb{N}_0, \end{cases}$$

where $Q_k^\alpha(x) = \sum_{\beta:|\beta|=k} c_{k,\alpha,\beta}x^\beta$ for some $c_{k,\alpha,\beta} \in \mathbb{R}$. Consequently,

$$\mathfrak{M}_l(|\boldsymbol{x}|^l) < \infty.$$

Proof. First note that $\partial_{x_i}|x|^l = \partial_{x_i}(x_1^2 + \cdots + x_d^2)^{\frac{l}{2}} = l|x|^{l-2}x_i$ for all $l \in \mathbb{R}$. Moreover, for all multi-indices β we have

$$\partial_{x_i}x^\beta|x|^l = \begin{cases} (\beta_i - 1)x^{\beta-e_i}|x|^l + lx^{\beta+e_i}|x|^{l-2} & \text{if } \beta_i \ge 1, \\ lx^{\beta+e_i}|x|^{l-2} & \text{if } \beta_i = 0, \end{cases}$$

This argument can be used to proof the statement by induction. □

Lemma 14.15. *Let $m \in \mathbb{R}$ and $l \in \mathbb{N}$. Then*

$$\mathfrak{M}_{lm,1}((1+|\boldsymbol{x}|^l)^m) < \infty.$$

Proof. Let $\alpha \in \mathbb{N}_0^d$. We will use Theorem D.1. Let $g(y) = (1+y)^m$ for $y \in (0,\infty)$. Then for $k \in \mathbb{N}$

$$\mathrm{D}^k\, g(y) = \begin{cases} (1+y)^{m-k} & k \notin \{m+n : n \in \mathbb{N}\}, \\ 0 \le (1+y)^{m-k} & k \in \{m+n : n \in \mathbb{N}\}. \end{cases}$$

On the other hand, by Lemma 14.14 we have for all $\beta \in \mathbb{N}_0^d$ that there exists a $c_\beta \ge 0$ such that $\mathrm{D}^\beta\, |x|^l = c_\beta |x|^{l-|\beta|}$. Hence if $1 \le k \le |\alpha|$ and $b \in (\mathbb{N}^d)^k \setminus \{0\}$, $b_1 + \cdots + b_k = \alpha$ then

$$\left| \prod_{i=1}^k \mathrm{D}^{b_i}\, |x|^l \right| \le |x|^{lk-|\alpha|}$$

And so $\left| \mathrm{D}^k\, g(|x|^l) \prod_{i=1}^k \mathrm{D}^{b_i}\, |x|^l \right| \le (1+|x|^l)^{m-k} |x|^{lk-|\alpha|}$ for all k with $1 \le k \le |\alpha|$ and all x. Let $\theta > 0$. Then there exists a $C > 0$ such that for all k with $1 \le k \le |\alpha|$:

$$(1+|x|^l)^{m-k} \le C|x|^{lm-lk} \qquad (x : |x| > \theta).$$

Hence with Theorem D.1

$$|\partial^\alpha \sigma(x)| \le \sum_{k=1}^{|\alpha|} \left| \mathrm{D}^k\, g(|x|^l) \prod_{i=1}^k \partial^{b_i} |x|^l \right| \le |\alpha| C |x|^{lm-|\alpha|} \qquad (x : |x| > \theta).$$

Remark 14.16. In [2, Lemma 2.2], the σ is not assumed to be infinitely differentiable, but have k-th order derivatives. However, in that case one has to justify the formula $\sigma(\mathrm{D})u = \mathcal{F}^{-1}(\sigma\hat{u})$. For $u \in L^p$ -to me- it is not clear whether $\hat{u}$ (with compact support) is such that one can make sense of $\sigma\hat{u}$ as a tempered distribution. If $\hat{u}$ is given by a Radon measure (or of order 0), then $\sigma\hat{u}$ would be again a Radon measure and with compact support, therefore a tempered distribution. Observe that $\mathbb{1} \in L^\infty$ and that $\hat{\mathbb{1}}$ is not represented by a function but by δ_0.

15 Besov spaces defined by Littlewood–Paley decompositions

We write "$\mathbb{N}_{-1}$" for the set $\{-1, 0, 1, 2, \dots\}$. Next we introduce the notion of a dyadic partition of unity, which consists of one function that is supported in a ball and equals 1 on a smaller ball around zero and of functions that are supported in annuli which are scaled versions of each other.

Remember that a function $f : \mathbb{R}^d \to \mathbb{F}$ is called *radial* if that $f(x) = f(y)$ for all $x, y \in \mathbb{R}^d$ with $|x| = |y|$.

Definition 15.1. Let $\mathcal{B}$ be a ball around zero and $\mathcal{A}$ be an annulus. Let χ and ρ be C^∞ radial functions with values in $[0,1]$, χ supported in $\mathcal{B}$ and ρ supported in $\mathcal{A}$. We say that (χ, ρ) forms a *dyadic partition of unity* if for $\rho_{-1} := \chi$ and $\rho_j := \rho(2^{-j}\cdot) = l_{2^{-j}}\rho$ we have

$$\sum_{j\in\mathbb{N}_{-1}} \rho_j(\xi) = 1, \qquad \frac{1}{2} \le \sum_{j\in\mathbb{N}_{-1}} \rho_j(\xi)^2 \le 1 \qquad (\xi \in \mathbb{R}^d), \tag{102}$$

$$|i-j| \ge 2 \Longrightarrow \operatorname{supp}\rho_i \cap \operatorname{supp}\rho_j = \emptyset \qquad (i, j \in \mathbb{N}_0). \tag{103}$$

$(\rho_j)_{j\in\mathbb{N}_{-1}}$ will also be called a *dyadic partition of unity.*

15.2. Next, we show the existence of a dyadic partition of unity. For this we take the annulus $\mathcal{A} = A(\frac{3}{4}, \frac{8}{3})$ so that (103) follows directly from the fact that $\operatorname{supp}\rho_i \subset 2^i\mathcal{A}$.

Indeed, if $k \in \mathbb{N}_0$ and $2^k\mathcal{A} \cap \mathcal{A} \neq \emptyset$, then $\frac{8}{3}\cdot \ge 2^k\frac{3}{4}$, i.e.,

$$2^k \le \frac{2^5}{3^2} < 2^2, \text{ which implies } k \le 1.$$

Therefore

$$|i-j| \ge 2 \Longrightarrow 2^i\mathcal{A} \cap 2^j\mathcal{A} = \emptyset \qquad (i, j \in \mathbb{Z}). \tag{104}$$

Theorem 15.3. [2, Proposition 2.10] *There exist C^∞ radial functions χ and ρ such that (χ, ρ) forms a dyadic partition of unity, where χ has support in the ball $\mathcal{B} = B(0, \frac{4}{3})$ and ρ has support in the annulus $\mathcal{A} = A(\frac{3}{4}, \frac{8}{3})$. Moreover,*

$$\sum_{j\in\mathbb{Z}} \rho(2^{-j}x) = 1, \qquad \frac{1}{2} \le \sum_{j\in\mathbb{Z}} \rho(2^{-j}x)^2 \le 1 \qquad (x \in \mathbb{R}^d \setminus \{0\}). \tag{105}$$

Proof. Let $a \in (1, \frac{4}{3})$ and $\mathcal{C} = A(\frac{1}{a}, 2a)$. Then, as $(\frac{1}{a}, 2a) \supset [1, 2]$, we have

$$\bigcup_{j\in\mathbb{Z}} 2^j\mathcal{C} = \mathbb{R}^d \setminus \{0\}, \tag{106}$$

Let θ be a smooth radial function supported in $\mathcal{A}$ that equals 1 on $\mathcal{C}_\delta$ for some $\delta > 0$. By (104) for each $\xi \in \mathbb{R}^d \setminus \{0\}$ there exists an $\varepsilon > 0$ such that $\theta(2^{-j}\cdot)$ is nonzero on $B(\xi, \varepsilon)$ only for finitely many $j \in \mathbb{Z}$. Therefore the function $S : \mathbb{R}^d \to \mathbb{R}$ defined by

$$S(\xi) = \sum_{j\in\mathbb{Z}} \theta(2^{-j}\xi) \qquad (\xi \in \mathbb{R}^d),$$

is smooth. As $\theta(2^{-j}\cdot)$ is one on $2^j\mathcal{C}$ for all $j \in Z$, by (106) it follow that $S(\xi) > 0$ for $\xi \in \mathbb{R}^d \setminus \{0\}$. We define the functions $\chi, \rho : \mathbb{R}^d \to \mathbb{R}$ by

$$\rho(\xi) = \frac{\theta(\xi)}{S(\xi)}, \qquad \chi(\xi) = 1 - \sum_{j\in\mathbb{N}_0} \rho(2^{-j}\xi) \qquad (\xi \in \mathbb{R}^d).$$

Then ρ and χ are radial functions because both θ and S are. Moreover, they are smooth, as θ is smooth and S is smooth on $\mathbb{R}^d \setminus \{0\}$. For $\xi \in \mathbb{R}^d \setminus \{0\}$ we have $\sum_{j\in\mathbb{Z}} \rho(2^{-j}\xi) = \frac{S(\xi)}{S(\xi)} = 1$. As $\operatorname{supp}\theta(2^j\cdot) \subset B(0,\frac{4}{3})$ for $j \in \mathbb{N}$, it follows that

$$\operatorname{supp}\chi \subset B(0,\tfrac{4}{3}), \qquad \sum_{j\in\mathbb{N}_0} \rho(2^{-j}\xi) = 1 \qquad (\xi \in \mathbb{R}^d \setminus B(0,\tfrac{4}{3})),$$

and in particular $\sum_{j\in\mathbb{N}_{-1}} \rho_j = \mathbb{1}$ with the notation for ρ_j as in Definition 15.1.

We are left to show that $\frac{1}{2} \le \sum_{j\in\mathbb{Z}} \rho(2^{-j}\xi)^2$ for all $\xi \in \mathbb{R}^d\setminus\{0\}$ and $\frac{1}{2} \le \sum_{j\in\mathbb{N}_{-1}} \rho_j(\xi)^2$ for $\xi \in \mathbb{R}^d$. Let us write $\Sigma_{odd} = \sum_{j\in 2\mathbb{Z}+1} \rho(2^{-j}\cdot)$ and $\Sigma_{even} = \sum_{j\in 2\mathbb{Z}} \rho(2^{-j}\cdot)$. As the functions $\rho(2^{-j}\cdot)$ for j being odd have disjoint support by (104), we have $\Sigma_{odd}^2 = \sum_{j\in 2\mathbb{Z}+1} \rho(2^{-j}\cdot)^2$. Similarly, $\Sigma_{even}^2 = \sum_{j\in 2\mathbb{Z}} \rho(2^{-j}\cdot)^2$. Therefore, for $\xi \in \mathbb{R}^d \setminus \{0\}$,

$$1 = (\Sigma_{odd}(\xi) + \Sigma_{even}(\xi))^2 \le 2(\Sigma_{odd}^2(\xi) + \Sigma_{even}^2(\xi)) = 2\sum_{j\in\mathbb{Z}} \rho(2^{-j}\xi)^2.$$

$\square$

15.4. As we have seen in the proof, to form a dyadic partition of unity it is sufficient to consider only a function supported on an annulus with certain properties. In the sense that such function also can be said to "form a dyadic partition of unity". In other words, a dyadic partition of unity $(\rho_j)_{j\in\mathbb{N}_{-1}}$ is generated by ρ_0, as $\rho_j = \rho_0(2^{-j}\cdot)$ for $j \in \mathbb{N}_0$ and $\rho_{-1} = 1 - \sum_{j\in\mathbb{N}_0} \rho_0(2^{-j}\cdot)$.

Lemma 15.5. *Let $(\rho_j)_{j\in\mathbb{N}_{-1}}$ be a dyadic partition of unity. Let $\chi = \rho_{-1}$, $\rho = \rho_0$ and write $\Delta_j = \rho_j(\mathrm{D})$. Then*

$$f \in \mathcal{S} \Longrightarrow \Delta_j f \in \mathcal{S}, \quad u \in \mathcal{S}' \Longrightarrow \Delta_j u \in \mathcal{S}', \quad u \in L^p \Longrightarrow \Delta_j u \in L^p \quad (j \in \mathbb{N}_{-1}),$$

and

$$\chi(2^{-J}\cdot) = \sum_{j=-1}^{J-1} \rho_j \qquad (J \in \mathbb{N}_0), \tag{107}$$

$$\sum_{j\in\mathbb{N}_{-1}} \Delta_j f = f \quad \text{in } \mathcal{S} \qquad (f \in \mathcal{S}), \tag{108}$$

$$\sum_{j\in\mathbb{N}_{-1}} \Delta_j u = u \quad \text{in } \mathcal{S}' \qquad (u \in \mathcal{S}'), \tag{109}$$

$$\|\Delta_j f\|_{L^p} \le \|\mathcal{F}^{-1}(\rho)\|_{L^1} \|f\|_{L^p} \qquad (j \in \mathbb{N}_0, f \in L^p), \tag{110}$$

$$\left\| \sum_{j=-1}^{J} \Delta_j f \right\|_{L^p} \le \|\mathcal{F}^{-1}(\chi)\|_{L^1} \|f\|_{L^p} \qquad (J \in \mathbb{N}_{-1}, f \in L^p). \tag{111}$$

Exercise **15.1.** Prove Lemma 15.5. (Hint: Use 10.6 (e).)

15.6 (The notation "$\sum_{j\in\mathbb{N}_{-1}}$"). Let $u_j \in \mathcal{S}'$ for $j \in \mathbb{N}_{-1}$. We write

$$u = \sum_{j\in\mathbb{N}_{-1}} u_j$$

to denote that u is the limit of the series independent of its reordering, which means that for each bijection $q : \mathbb{N}_{-1} \to \mathbb{N}_{-1}$ we have

$$u = \lim_{J\to\infty} \sum_{j=1}^{J} u_{q(j)}. \tag{112}$$

This will however not be of any importance, so one may as well interpret it as $\sum_{j=1}^{\infty} u_j$.

(Hint: Use 10.6 (f), to prove the condition observe that

$$\sup_{N\in\mathbb{N}} \|\sum_{n=1}^{N} \rho_{q(n)}\|_{C^k} \le \sup_{N\in\mathbb{N}} \Big\| \sum_{\substack{j\in 2\mathbb{N}_0-1 \\ j\in\{q(n):n\in\{1,\dots,N\}\}}} \rho_j \Big\|_{C^k} + \sup_{N\in\mathbb{N}} \Big\| \sum_{\substack{j\in 2\mathbb{N}_0 \\ j\in\{q(n):n\in\{1,\dots,N\}\}}} \rho_j \Big\|_{C^k}$$
$$\le \sup_{j\in 2\mathbb{N}_0-1} \|\rho_j\|_{C^k} + \sup_{j\in 2\mathbb{N}_0} \|\rho_j\|_{C^k} \le 2 \sup_{j\in\mathbb{N}_{-1}} \|\rho_j\|_{C^k},$$

where we used that $\|f+g\|_{C^k} = \|f\|_{C^k} \vee \|g\|_{C^k}$ for f and g with disjoint support.)

In these lecture notes we interpret for example $\|u\|_{L^p}$ for $u \in \mathcal{S}'$ that is not represented by a L^p function to be equal to infinity.

Definition 15.7 (Besov Space). [2, Definition 2.68] Let $\alpha \in \mathbb{R}$ and $p, q \in [1,\infty]$. Let $(\rho_j)_{j\in\mathbb{N}_{-1}}$ be a dyadic partition of unity. We write $\rho = \rho_0$ and $\Delta_j = \rho_j(\mathrm{D})$. Δ_j is also called a *Littlewood-Paley block*. We define the *nonhomogeneous Besov space* $B^{\alpha}_{p,q}[\rho]$ to be the space of all tempered distributions u such that

$$\|u\|_{B^{\alpha}_{p,q}[\rho]} := \Big\| \Big(2^{j\alpha} \|\Delta_j u\|_{L^p} \Big)_{j\in\mathbb{N}_{-1}} \Big\|_{\ell^q} < \infty. \tag{113}$$

Here we wrote "$\|\cdot\|_{\ell^q}$" as an abbreviation for "$\|\cdot\|_{\ell^q(\mathbb{N}_{-1})}$". The parameter α can be interpreted as a "regularity parameter". See for example Exercise 15.2 for the implication that "being of a certain regularity" implies "being also of lower regularity".

We will drop the notation "$[\rho]$" later, as the space does not depend on the (choice of) dyadic partition of unity, this follows from Theorem 15.9: we mention this in 15.11.

Exercise 15.2. Let $p, q \in [1,\infty]$, $\alpha, \beta \in \mathbb{R}$. Show that $B^{\alpha}_{p,q}[\rho] \subset B^{\beta}_{p,q}[\rho]$ for $\beta < \alpha$ and $B^{\alpha}_{p,\infty}[\rho] \subset B^{\alpha-\varepsilon}_{p,q}[\rho]$ for $\varepsilon > 0$.

In the proof of Theorem 15.9 we use Young's inequality for ℓ^p spaces:

Theorem 15.8 (Young's inequality for ℓ^p spaces). *Let $p, q, r \in [1,\infty]$ be such that*

$$\tfrac{1}{p} + \tfrac{1}{q} = 1 + \tfrac{1}{r}.$$

For $f \in \ell^p(\mathbb{Z})$, $g \in \ell^q(\mathbb{Z})$ we have $f * g \in \ell^r(\mathbb{Z})$ and

$$\|f * g\|_{\ell^r} \le \|f\|_{\ell^p} \|g\|_{\ell^q}.$$

Proof. Follows in the same way as Theorem 3.4 but with applying Hölder's inequality to the sequence spaces $\ell^p(\mathbb{Z})$, which means the underlying measure space is $\mathbb{Z}$ equipped with the counting measure. $\square$

Theorem 15.9. *Let $\alpha \in \mathbb{R}$ and $p, q \in [1, \infty]$. Let $\mathcal{B}$ be a ball around zero and $\mathcal{A}$ be an annulus.*

(a) *There exist $C > 0$ such that for all dyadic partitions of unity $(\rho_j)_{j \in \mathbb{N}_{-1}}$ and $(\sigma_j)_{j \in \mathbb{N}_{-1}}$ with $\operatorname{supp} \rho_{-1}, \operatorname{supp} \sigma_{-1} \subset \mathcal{B}$ and $\operatorname{supp} \rho_0, \operatorname{supp} \sigma_0 \subset \mathcal{A}$, and for all $u \in \mathcal{S}'$*

$$\left\| \left(2^{j\alpha} \| \rho_j(\mathrm{D}) u \|_{L^p} \right)_{j \in \mathbb{N}_{-1}} \right\|_{\ell^q} \leq C \left\| \left(2^{j\alpha} \| \sigma_j(\mathrm{D}) u \|_{L^p} \right)_{j \in \mathbb{N}_{-1}} \right\|_{\ell^q}.$$

(b) *Then there exist $C > 0$ and $m \in \mathbb{N}_0$ such that for all sequences of smooth functions $(u_j)_{j \in \mathbb{N}_{-1}}$ with*

$$\operatorname{supp} \hat{u}_{-1} \subset \mathcal{B}, \quad \operatorname{supp} \hat{u}_j \subset 2^j \mathcal{A} \text{ for } j \geq 0, \qquad \left\| \left(2^{j\alpha} \| u_j \|_{L^p} \right)_{j \in \mathbb{N}_{-1}} \right\|_{\ell^q} < \infty,$$

$u := \sum_{j \in \mathbb{N}_{-1}} u_j$ exists in $\mathcal{S}'$,

$$|\langle u, \varphi \rangle| \leq C \left\| \left(2^{j\alpha} \| u_j \|_{L^p} \right)_{j \in \mathbb{N}_{-1}} \right\|_{\ell^q} \| \varphi \|_{m, \mathcal{S}} \qquad (\varphi \in \mathcal{S}), \tag{114}$$

and for all dyadic partitions of unity $(\rho_j)_{j \in \mathbb{N}_{-1}}$

$$\left\| \left(2^{j\alpha} \| \rho_j(\mathrm{D}) u \|_{L^p} \right)_{j \in \mathbb{N}_{-1}} \right\|_{\ell^q} \leq C \left\| \left(2^{j\alpha} \| u_j \|_{L^p} \right)_{j \in \mathbb{N}_{-1}} \right\|_{\ell^q}. \tag{115}$$

(c) *If $\alpha > 0$, then there exist $C > 0$ and $m \in \mathbb{N}_0$ such that for all sequences of smooth functions $(u_j)_{j \in \mathbb{N}_{-1}}$ with*

$$\operatorname{supp} \hat{u}_j \subset 2^j \mathcal{B} \text{ for all } j \in \mathbb{N}_{-1}, \qquad \left\| \left(2^{j\alpha} \| u_j \|_{L^p} \right)_{j \in \mathbb{N}_{-1}} \right\|_{\ell^q} < \infty. \tag{116}$$

one has that $u := \sum_{j \in \mathbb{N}_{-1}} u_j$ exists in $\mathcal{S}'$, (114) *holds and* (115) *holds for all dyadic partitions of unity $(\rho_j)_{j \in \mathbb{N}_{-1}}$.*

(d) *If $\alpha = 0$ and $q = 1$, then there exist $C > 0$ and $m \in \mathbb{N}_0$ such that for all sequences of smooth functions $(u_j)_{j \in \mathbb{N}_{-1}}$ with* (116) *one has that $u := \sum_{j \in \mathbb{N}_{-1}} u_j$ exists in $\mathcal{S}'$,* (114) *holds and for all dyadic partitions of unity $(\rho_j)_{j \in \mathbb{N}_{-1}}$*

$$\sup_{j \in \mathbb{N}_{-1}} \| \rho_j(\mathrm{D}) u \|_{L^p} \leq C \left\| \left(\| u_j \|_{L^p} \right)_{j \in \mathbb{N}_{-1}} \right\|_{\ell^1}. \tag{117}$$

Proof. (a) follows from (b).

Let $(u_j)_{j \in \mathbb{N}_{-1}}$ be as in (b).

• First we prove that for all bijections $q : \mathbb{N}_{-1} \to \mathbb{N}_{-1}$ the sum $\sum_{j=-1}^{J} u_{q(j)}$ converges in $\mathcal{S}'$ as $J \to \infty$ and prove (114). By Theorem 10.17 it suffices to prove that $\sum_{j=-1}^{\infty} |\langle u_j, \varphi\rangle| < \infty$ for all $\varphi \in \mathcal{S}$. Let $\varphi \in \mathcal{S}$. Let $k \in \mathbb{N}_0$ be such that $k > -\alpha$. For all $j \in \mathbb{N}_0$ we have by Lemma 14.2

$$u_j = 2^{-jk} \sum_{\beta\in\mathbb{N}_0^d:|\beta|=k} 2^{jd}(l_{2^j} g_\beta) * \partial^\beta u_j.$$

And thus

$$\langle u_j, \varphi\rangle = 2^{-jk}(-1)^k \sum_{\beta\in\mathbb{N}_0^d:|\beta|=k} \langle u_j, 2^{jd}(l_{2^j} \check{g}_\beta) * \partial^\beta \varphi\rangle.$$

By Hölder's and by Young's inequality, with $r \in [1,\infty]$ such that $\frac{1}{p} + \frac{1}{r} = 1$,

$$|\langle u_j, \varphi\rangle| \le 2^{-jk} \|u_j\|_{L^p} \sum_{\beta\in\mathbb{N}_0^d:|\beta|=k} \|2^{jd}(l_{2^j} \check{g}_\beta)\|_{L^1} \|\partial^\beta \varphi\|_{L^r}.$$

By Lemma 10.19 there is an $n \in \mathbb{N}$ and a $C_1 > 0$ such that

$$\|\partial^\beta \varphi\|_{L^r} \le C_1 \|\partial^\beta \varphi\|_{n,\mathcal{S}} \le C_1 \|\varphi\|_{n+k,\mathcal{S}}.$$

By (91) $\|2^{jd}(l_{2^j} \check{g}_\beta)\|_{L^1} = \|g_\beta\|_{L^1}$. Therefore with

$$C_2 = \sum_{\beta\in\mathbb{N}_0^d:|\beta|=k} \|g_\beta\|_{L^1},$$

we have for all $j \in \mathbb{N}_0$

$$\begin{aligned}|\langle u_j, \varphi\rangle| &\le C_1 C_2 2^{-jk} \|u_j\|_{L^p} \|\varphi\|_{n+k,\mathcal{S}} \\ &\le C_1 C_2 2^{-j(k+\alpha)} \left\| \left(2^{j\alpha} \|u_j\|_{L^p}\right)_{j\in\mathbb{N}_{-1}} \right\|_{\ell^q} \|\varphi\|_{n+k,\mathcal{S}}.\end{aligned}$$

We may assume that the above also holds for $j = -1$, as by a direct application of Hölder's inequality we have

$$\begin{aligned}|\langle u_{-1}, \varphi\rangle| &\le \|u_{-1}\|_{L^p} \|\varphi\|_{L^r} \le C_1 \|u_{-1}\|_{L^p} \|\varphi\|_{n,\mathcal{S}} \\ &\le C_1 2^{k+\alpha} \left\| \left(2^{j\alpha} \|u_j\|_{L^p}\right)_{j\in\mathbb{N}_{-1}} \right\|_{\ell^q} \|\varphi\|_{n+k,\mathcal{S}}.\end{aligned}$$

As $k + \alpha > 0$, there exists a $C > 0$ such that (114) holds with $m = n + k$.

• Let now $(\rho_j)_{j\in\mathbb{N}_{-1}}$ be a dyadic partition of unity. We prove (115). Let $\tilde{\mathcal{B}}$ be a ball around zero and $\tilde{\mathcal{A}}$ be an annulus such that

$$\operatorname{supp} \rho_{-1} \subset \tilde{\mathcal{B}}, \qquad \operatorname{supp} \rho_0 \subset \tilde{\mathcal{A}}. \tag{118}$$

Let $N \in \mathbb{N}$ be such that

$$2^j\mathcal{A} \cap \tilde{\mathcal{A}} = 2^j\mathcal{A} \cap \tilde{\mathcal{B}} = 2^j\tilde{\mathcal{A}} \cap \mathcal{A} = 2^j\tilde{\mathcal{A}} \cap \mathcal{B} = \emptyset \qquad (j \ge N). \tag{119}$$

We write $\Delta_j = \rho_j(\mathrm{D})$. Then $\Delta_j u_i = 0$ for all $i, j \in \mathbb{N}_{-1}$ with $|i-j| \ge N$. As $2^{(j-i)\alpha} \le 2^{|\alpha| N}$ for $i, j \in \mathbb{N}_{-1}$ with $|i-j| \le N$, by (110) and (111) (for $j = -1$) there exists a $C > 0$ such that for all $j \in \mathbb{N}_{-1}$

$$2^{j\alpha}\|\Delta_j u\|_{L^p} \le \sum_{i=(j-N)\vee -1}^{j+N} 2^{j\alpha}\|\Delta_j u_i\|_{L^p} \le C \sum_{i=(j-N)\vee -1}^{j+N} 2^{i\alpha}\|u_i\|_{L^p}$$

For $k \in \mathbb{Z}$ let $a_k = \mathbb{1}_{[-N,N]}(k)$ and $b_k = 2^{k\alpha}\|u_k\|_{L^p}$ for $k \in \mathbb{N}_{-1}$ and $b_k = 0$ otherwise. Write $a = (a_k)_{k\in\mathbb{Z}}$ and $b = (b_k)_{k\in\mathbb{Z}}$. Then

$$\sum_{i=(j-N)\vee -1}^{j+N} 2^{i\alpha}\|u_i\|_{L^p} = (a * b)_j \qquad (j \in \mathbb{Z})$$

Therefore by Young's inequality of Theorem 15.8,

$$\left\|(2^{j\alpha}\|\Delta_j u\|_{L^p})_{j\in\mathbb{N}_{-1}}\right\|_{\ell^q(\mathbb{N}_{-1})} \le C\|a * b\|_{\ell^q(\mathbb{Z})} \le C\|a\|_{\ell^1(\mathbb{Z})}\|b\|_{\ell^q(\mathbb{Z})},$$

as $\|b\|_{\ell^q(\mathbb{Z})} = \left\|\left(2^{j\alpha}\|u_j\|_{L^p}\right)_{j\in\mathbb{N}_{-1}}\right\|_{\ell^q(\mathbb{N}_{-1})}$ and $\|a\|_{\ell^1(\mathbb{Z})} = 2N+1$ this finished the proof for (b).

- Suppose that $\alpha > 0$ and $(u_j)_{j\in\mathbb{N}_{-1}}$ is as in (c). By Hölder's inequality we obtain

$$|\langle u_j, \varphi\rangle| \le \|u_j\|_{L^p}\|\varphi\|_{L^q} \le 2^{-j\alpha}\left\|\left(2^{j\alpha}\|u_j\|_{L^p}\right)_{j\in\mathbb{N}_{-1}}\right\|_{\ell^q}\|\varphi\|_{L^q}$$

which is summable as $\alpha > 0$. (114) can be obtained in the same way as above.

Let N again be such that $2^j\tilde{\mathcal{A}} \cap \mathcal{B} = \emptyset$ for $j \ge N$ (as in (119)). Then $\Delta_j u_i = 0$ for all $j \ge i + N$ and so by (110) and (111) there exists a $C_1 > 0$ such that

$$\begin{aligned} 2^{j\alpha}\|\Delta_j u\|_{L^p} &\le \sum_{i\in\mathbb{N}_{-1}:i>j-N} 2^{(j-i)\alpha}2^{i\alpha}\|\Delta_j u_i\|_{L^p} \\ &\le C_1 \sum_{i\in\mathbb{N}_{-1}:i>j-N} 2^{(j-i)\alpha}2^{i\alpha}\|u_i\|_{L^p} = C_1(a * b)(j), \end{aligned}$$

where $a_k = 2^{k\alpha}\mathbb{1}_{(-\infty,N)}(k)$ for $k \in \mathbb{Z}$ and $b_k = 2^{k\alpha}\|u_k\|_{L^p}$ for $k \in \mathbb{N}_{-1}$ and $b_k = 0$ otherwise. So that again with Young's inequality, we obtain the desired bound as

$$\|a\|_{\ell^1} = \sum_{k\in\mathbb{Z}:k<N} 2^{k\alpha} = \sum_{k\in\mathbb{N}_0} 2^{(N-1-k)\alpha} = \frac{2^{(N-1)\alpha}}{1-2^{-\alpha}} = \frac{2^{N\alpha}}{2^\alpha - 1}.$$

- (d) follows again by applying Hölder's inequality and the estimate $\|\Delta_j u\|_{L^p} \le \sum_{i\in\mathbb{N}_{-1}}\|\Delta_j u_i\|_{L^p}$.□

Remark 15.10. In general the inequality (115) only holds in this direction. E.g., let $u_0 = -u_1 \neq 0$ in L^p be supported in $\mathcal{A} \cap 2\mathcal{A}$ and $u_j = 0$ for $j \notin \{0, 1\}$. For this example the left-hand side of (115) is zero and the right-hand side is not.

15.11. Suppose ρ and σ form dyadic partitions of unity and that $\alpha \in \mathbb{R}$, $p, q \in [1, \infty]$. By Theorem 15.9 (a) it follows that there exists a $C > 0$ such that

$$\frac{1}{C}\|u\|_{B^\alpha_{p,q}[\sigma]} \leq \|u\|_{B^\alpha_{p,q}}[\rho] \leq C\|u\|_{B^\alpha_{p,q}[\sigma]} \qquad (u \in \mathcal{S}').$$

Therefore $B^\alpha_{p,q}[\rho] = B^\alpha_{p,q}[\sigma]$ and their norms are equivalent. For this reason we will write "$B^\alpha_{p,q}$" instead of "$B^\alpha_{p,q}[\rho]$" and "$\| \cdot \|_{B^\alpha_{p,q}}$" instead of "$\| \cdot \|_{B^\alpha_{p,q}[\rho]}$"; of course the norm depends on the choice of partition, but as our statements only consider estimates, the choice of partition is irrelevant for our purposes.

For the rest of this section we fix (χ and) ρ and also the annulus $\mathcal{A}$ and ball $\mathcal{B}$ such that $\operatorname{supp}\rho \subset \mathcal{A}$ and $\operatorname{supp}\chi \subset \mathcal{B}$.

In Theorem 15.15 we will show that Besov spaces are Banach spaces. Moreover, one could say that they are sequentially compactly embedded in $\mathcal{S}'$. In other words, every bounded sequence in a Besov space has a subsequence that converges in $\mathcal{S}'$ to an element of that Besov space. Moreover, the norm of the limit is bounded from above by the lim inf of the norm of the subsequence. This is similar to the statement in Fatou's lemma, in [2] they also call this the "Fatou property".

We will first prove this sequentially compact embedding for L^p and $\mathcal{M}$, after making the following observation.

15.12. Let $p \in (1, \infty]$. Then L^p is isometrically isomorphic to $(L^q)'$, the dual of L^q, for $q \in [1, \infty)$ being such that $\frac{1}{p} + \frac{1}{q} = 1$. Moreover,

$$\|v\|_{L^p} = \sup\{|\langle v, f\rangle| : f \in L^q, \|f\|_{L^q} \leq 1\} \qquad (v \in L^p).$$

Let $\mathcal{M}$ the space of signed ($\mathbb{F} = \mathbb{R}$) or complex ($\mathbb{F} = \mathbb{C}$) Radon measures (see Definition H.6 and H.8), is the dual of C_0, the space of continuous functions that vanish at infinity (see Definition H.9). Moreover,

$$\|\mu\|_{\mathcal{M}} = \sup\{|\langle \mu, f\rangle| : f \in C_0, \|f\|_{C^0} \leq 1\} \qquad (\mu \in \mathcal{M}).$$

Lemma 15.13. *Let $p \in (1, \infty]$. Let $\mathfrak{X}$ be either the Banach space L^p or $\mathcal{M}$. If $(u_n)_{n\in\mathbb{N}}$ is a sequence in $\mathfrak{X}$ that is bounded in the $\mathfrak{X}$ norm, then it has a subsequence $(u_{n_m})_{m\in\mathbb{N}}$ that converges in $\mathcal{S}'$ to an element u, which is also in $\mathfrak{X}$ and*

$$\|u\|_{\mathfrak{X}} \leq \liminf_{m\to\infty} \|u_{n_m}\|_{\mathfrak{X}}. \tag{120}$$

Proof. Let $\mathfrak{Y}$ be either $\mathfrak{Y} = L^q$ or $\mathfrak{Y} = C_0$ (see 15.12), so that $\mathfrak{X}$ is isometrically isomorphic to the dual of $\mathfrak{Y}$, $\mathfrak{Y}'$ and

$$\|u\|_{\mathfrak{X}} = \sup\{|\langle u, f\rangle| : f \in \mathfrak{Y}, \|f\|_{\mathfrak{Y}} \leq 1\} \qquad (u \in \mathfrak{X}). \tag{121}$$

In either case there exist $k \in \mathbb{N}$ and $C > 0$ such that $\|\cdot\|_{\mathfrak{Y}} \le C\|\cdot\|_{k,\mathcal{S}}$, see Lemma 10.19.

Let $(u_n)_{n\in\mathbb{N}}$ be a sequence that is bounded in $\mathfrak{X}$. Without loss of generality we assume $\|u_n\|_{\mathfrak{X}} \le 1$ for all $n \in \mathbb{N}$. Then

$$|\langle u_n, \varphi\rangle| \le \|u_n\|_{\mathfrak{X}}\|\varphi\|_{\mathfrak{Y}} \le C\|\varphi\|_{k,\mathcal{S}} \qquad (n \in \mathbb{N}, \varphi \in \mathcal{S}). \tag{122}$$

Therefore, for each $\varphi \in \mathcal{S}$ the sequence $(\langle u_n, \varphi\rangle)_{n\in\mathbb{N}}$ is bounded in $\mathbb{F}$ and hence has a convergent subsequence. Let D be a countable dense subset of $\mathcal{S}$ (see Theorem 10.10). We may assume that D is a $\mathbb{Q}$-linear space (first of all we may assume that $\mathbb{Q}D = D$, then we can take the countable union of the countable sets D, $D+D$, $D+D+D$, $\dots$). By a Cantor's diagonal method we find a subsequence $(u_{n_m})_{m\in\mathbb{N}}$ such that $\langle u_{n_m}, \varphi\rangle$ converges as $m \to \infty$ for all $\varphi \in D$. We define $u : D \to \mathbb{F}$ by

$$\langle u, \varphi\rangle = \lim_{m\to\infty} \langle u_{n_m}, \varphi\rangle \qquad (\varphi \in D).$$

As each u_n is linear, u is $\mathbb{Q}$-linear. By (122) we have

$$|\langle u, \varphi\rangle| \le \|\varphi\|_{\mathfrak{Y}} \le C\|\varphi\|_{k,\mathcal{S}},$$

so that u extends continuously on the whole of $\mathcal{S}$, as an element of $\mathcal{S}'$, and moreover, also extends to an element of $\mathfrak{Y}'$ and thus to $\mathfrak{X}$. As $\mathcal{S}$ is dense in $\mathfrak{Y}$ we may replace "$\mathfrak{Y}$" in (121) by "$\mathcal{S}$", and obtain

$$\begin{aligned}
\|u\|_{\mathfrak{X}} &= \sup\{\liminf_{m\to\infty} |\langle u_{n_m}, \varphi\rangle| : \varphi \in \mathcal{S}, \|\varphi\|_{\mathfrak{Y}} \le 1\} \\
&\le \sup\{\liminf_{m\to\infty} \|u_{n_m}\|_{\mathfrak{X}}\|\varphi\|_{\mathfrak{Y}} : \varphi \in \mathcal{S}, \|\varphi\|_{\mathfrak{Y}} \le 1\} \\
&\le \liminf_{m\to\infty} \|u_{n_m}\|_{\mathfrak{X}}.
\end{aligned}$$

□

$\boxed{\textbf{Exercise}}$ **15.3.** Show that the statement in Lemma 15.13 for $p = 1$ does not hold.

15.14. For $p = 1$, we still have the following: If $(u_n)_{n\in\mathbb{N}}$ is a sequence in L^1 that is bounded in the L^1 norm, and there exists a compact set K such that $\operatorname{supp} \hat{u}_n \subset K$ for all $n \in \mathbb{N}$, then there exists a subsequence $(u_{n_m})_{m\in\mathbb{N}}$ that converges in $\mathcal{S}'$ to an element u, which is also in L^1 and (120) holds for $\mathfrak{X} = L^1$.

First of all, that the limit in $\mathcal{S}'$ is actually in L^1 follows from the fact that it is a (signed or) complex Radon measure by Lemma 15.13 and because $\operatorname{supp} \hat{u} \subset K$, so that $u \in C_p^\infty$ (by Lemma 12.11). To obtain (120), it is sufficient to show that $\|f\|_{\mathcal{M}} = \|f\|_{L^1}$ for $f \in L^1$.

There exists a sequence of functions $(f_n)_{n\in\mathbb{N}}$ in $C_c(\mathbb{R}^d, [0, 1])$ such that $f_n(x) \to 1$ for those x such that $u(x) > 0$, $f_n(x) \to 0$ for those x such that $u(x) = 0$ and $f_n(x) \to -1$ for those x such that $u(x) < 0$. By Lebesgue's dominated convergence theorem

$$\langle u, f_n\rangle = \int u f_n \to \|u\|_{L^1}.$$

Theorem 15.15. *Let $\alpha \in \mathbb{R}$ and $p, q \in [1, \infty]$. The function $\| \cdot \|_{B^\alpha_{p,q}} : B^\alpha_{p,q} \to [0, \infty)$ defined as in (113) is a norm. $B^\alpha_{p,q}$ equipped with this norm is a Banach space that is continuously embedded in $\mathcal{S}'$. Moreover, if $(u_n)_{n\in\mathbb{N}}$ is a sequence in $B^\alpha_{p,q}$ that is bounded in the $B^\alpha_{p,q}$ norm, then it has a subsequence $(u_{n_m})_{m\in\mathbb{N}}$ that converges in $\mathcal{S}'$ to an element u, which is also in $B^\alpha_{p,q}$ and*

$$\|u\|_{B^\alpha_{p,q}} \le \liminf_{m\to\infty} \|u_{n_m}\|_{B^\alpha_{p,q}}.$$

Proof. By its definition it follows rather immediately that $\| \cdot \|_{B^\alpha_{p,q}}$ is a semi-norm. That it is a norm follows from the following: If $\|u\|_{B^\alpha_{p,q}} = 0$, then $\Delta_j u = 0$ and so $\rho_j \hat{u} = 0$ for all $j \in \mathbb{N}_{-1}$, whence $\operatorname{supp} \hat{u} = \emptyset$ and so $\hat{u} = 0$ and thus $u = 0$. That $B^\alpha_{p,q}$ is continuously embedded in $\mathcal{S}'$ follows from (114) in Theorem 15.9. We will prove that $B^\alpha_{p,q}$ is complete after proving the "Moreover" statement.

Let $(u_n)_{n\in\mathbb{N}}$ be a sequence that is bounded in $B^\alpha_{p,q}$. Without loss of generality we may assume that $\|u_n\|_{B^\alpha_{p,q}} \le 1$ for all $n \in \mathbb{N}$. Then

$$\|\Delta_j u_n\|_{L^p} \le 2^{-\alpha j} \qquad (n \in \mathbb{N}, j \in \mathbb{N}_{-1}).$$

By applying Lemma 15.13 to $(\Delta_j u_n)_{n\in\mathbb{N}}$ for each j, and applying Cantor's diagonal argument, we find a subsequence $(u_{n_m})_{m\in\mathbb{N}}$ of $(u_n)_{n\in\mathbb{N}}$ such that there exist $u_j \in \mathcal{S}'$ for all $j \in \mathbb{N}_{-1}$ such that

$$\Delta_j u_{n_m} \xrightarrow{m\to\infty} u_j, \qquad \|u_j\|_{L^p} \le \liminf_{m\to\infty} \|\Delta_j u_{n_m}\|_{L^p} \le 2^{-\alpha j} \qquad (j \in \mathbb{N}_{-1}).$$

As the support of the Fourier transform of $\Delta_j u_n$ is in the annulus $2^j\mathcal{A}$ (or ball $\mathcal{B}$), so is the support of $\hat{u}_j$ for $j \in \mathbb{N}_0$ (for $j = -1$).

By Theorem 15.9 (b) it follows that $u := \sum_{j\in\mathbb{N}_{-1}} u_j$ exists in $\mathcal{S}'$ and that there is a $C > 0$ such that

$$\begin{aligned}
\|u\|_{B^\alpha_{p,q}} &\le C \left\| \left(2^{j\alpha} \|u_j\|_{L^p} \right)_{j\in\mathbb{N}_{-1}} \right\|_{\ell^q} \\
&\le C \left\| \left(2^{j\alpha} \liminf_{m\to\infty} \|\Delta_j u_{n_m}\|_{L^p} \right)_{j\in\mathbb{N}_{-1}} \right\|_{\ell^q} \\
&\le C \liminf_{m\to\infty} \left\| \left(2^{j\alpha} \|\Delta_j u_{n_m}\|_{L^p} \right)_{j\in\mathbb{N}_{-1}} \right\|_{\ell^q} = C \liminf_{m\to\infty} \|u_{n_m}\|_{B^\alpha_{p,q}}.
\end{aligned}$$

To prove that $B^\alpha_{p,q}$ is complete, we assume that the sequence $(u_n)_{n\in\mathbb{N}}$ as above is also Cauchy. Let u be the limit of the subsequence as above. It suffices to show that $u_n \to u$ in $B^\alpha_{p,q}$. Let $\varepsilon > 0$ and $N \in \mathbb{N}$ be such that $m, k \ge N$ implies $\|u_k - u_m\|_{B^\alpha_{p,q}} < \varepsilon$. Let $k \ge N$. Apply the above limiting argument to the sequence $(u_n - u_k)_{n\in\mathbb{N}}$, so that for some sequence $(n_m)_{m\in\mathbb{N}}$ in $\mathbb{N}$

$$\|u - u_k\|_{B^\alpha_{p,q}} \le C \liminf_{m\to\infty} \|u_{n_m} - u_k\|_{B^\alpha_{p,q}} < C\varepsilon.$$

Therefore, $u_n \to u$ in $B^\alpha_{p,q}$. $\square$

For a negative regularity index α the Besov norm is equivalent to the function that replaces "Δ_j" in the norm by "$\sum_{i=-1}^{j} \Delta_i$":

Theorem 15.16. [2, Theorem 2.33] *Let $\alpha < 0$ and $p, q \in [1, \infty]$. For $u \in \mathcal{S}'$ we write $S_j u = \sum_{i=-1}^{j} \Delta_i u$ for $j \in \mathbb{N}_{-1}$. Then we have for $u \in \mathcal{S}'$*

$$u \in B^{\alpha}_{p,q} \iff \|(2^{j\alpha}\|S_j u\|_{L^p})_{j\in\mathbb{N}_{-1}}\|_{\ell^q} < \infty.$$

Moreover,

$$(1+2^{\alpha})^{-1}\|u\|_{B^{\alpha}_{p,q}} \le \|(2^{j\alpha}\|S_j u\|_{L^p})_{j\in\mathbb{N}_{-1}}\|_{\ell^q} \le (1-2^{\alpha})^{-1}\|u\|_{B^{\alpha}_{p,q}} \qquad (u \in \mathcal{S}'). \tag{123}$$

Proof. For the inequality on the left–hand side of (123):

$$2^{j\alpha}\|\Delta_j u\|_{L^p} \le 2^{j\alpha}(\|S_j u\|_{L^p} + 2^{\alpha}2^{(j-1)\alpha}\|S_{j-1}u\|_{L^p}).$$

Therefore

$$\|u\|_{B^{\alpha}_{p,q}} \le (1+2^{\alpha})\|(2^{j\alpha}\|S_j u\|_{L^p})_{j\in\mathbb{N}_{-1}}\|_{\ell^q}.$$

For the inequality on the right hand side of (123):

$$2^{j\alpha}\|S_j u\|_{L^p} \le 2^{j\alpha}\sum_{i=-1}^{j}\|\Delta_i u\|_{L^p} = \sum_{i=-1}^{j} 2^{(j-i)\alpha}2^{i\alpha}\|\Delta_i u\|_{L^p} = (a * b)(j),$$

where $a, b : \mathbb{Z} \to \mathbb{R}$ are given for $j \in \mathbb{Z}$ by

$$a_j = \begin{cases} 2^{j\alpha} & j \in \mathbb{N}_0, \\ 0 & j \le -1, \end{cases} \qquad b_j = \begin{cases} 2^{j\alpha}\|\Delta_j u\|_{L^p} & j \in \mathbb{N}_{-1}, \\ 0 & j \le -2. \end{cases}$$

Hence, by Young's inequality Theorem 15.8

$$\|(2^{j\alpha}\|S_j u\|_{L^p})_{j\in\mathbb{N}_{-1}}\|_{\ell^q} = \|a * b\|_{\ell^q} \le \|a\|_{\ell^1}\|b\|_{\ell^q} = \|a\|_{\ell^1}\|u\|_{B^{\alpha}_{p,q}}.$$

As $\alpha < 0$ we have $\|a\|_{\ell^1} = \sum_{j\in\mathbb{N}_0} 2^{j\alpha} = (1-2^{\alpha})^{-1}$. □

Example 15.17. We will consider in which Besov space the Dirac delta, δ_0, lies. Note that $\Delta_i\delta_0 = \mathcal{F}^{-1}(\rho_i)$ so that for $i \ge 0$ and $p \in [1, \infty)$ see (94) and Exercise 14.1

$$\begin{aligned} \|\Delta_i\delta_0\|_{L^\infty} &\le \|\rho_i\|_{L^1} = 2^{id}\|\rho\|_{L^1}, \\ \|\Delta_i\delta_0\|_{L^p} &= \|\mathcal{F}^{-1}(\rho_i)\|_{L^p} = 2^{-i(\frac{d}{p}-d)}\|\mathcal{F}^{-1}\rho\|_{L^p}. \end{aligned}$$

Therefore, $\delta_0 \in B^{-d(1-\frac{1}{p})}_{p,\infty}$ and $\delta_0 \in B^{-d(1-\frac{1}{p})-\varepsilon}_{p,q}$ for all $q \in [1, \infty)$ and $\varepsilon > 0$.

Exercise **15.4.** Show that for $\varepsilon > 0$ the function $z \mapsto \delta_z$ is continuous in $B^{-\varepsilon}_{1,\infty}$ but that it is not continuous in $B^{0}_{1,\infty}$. Hint: Use (but show) that for any $\varphi \in \mathcal{S} \setminus \{0\}$ and $z \in \mathbb{R}^d \setminus \{0\}$, there exists an $\kappa > 0$ such that

$$\limsup_{a\to\infty}\|\mathcal{T}_{az}\varphi - \varphi\|_{L^1} > \kappa.$$

16 Embeddings of Besov spaces and Sobolev spaces

The following lemma also justifies that one can view α as the regularity, as the regularity decreases by the number of derivatives one takes.

Lemma 16.1. *For all $\gamma \in \mathbb{N}_0^d$ there exists a $C > 0$ such that for all $\alpha, \beta \in \mathbb{R}$, $p_1, p_2, q_1, q_2 \in [1, \infty]$, with*

$$p_2 \geq p_1, \qquad q_2 \geq q_1 \qquad \beta \leq \alpha - d(\tfrac{1}{p_1} - \tfrac{1}{p_2}), \tag{124}$$

one has

$$\|\partial^\gamma u\|_{B^{\beta-|\gamma|}_{p_2,q_2}} \leq C \|u\|_{B^\alpha_{p_1,q_1}} \qquad (u \in \mathcal{S}'). \tag{125}$$

In particular, $B^\alpha_{p_1,q_1}$ is continuously embedded in $B^\beta_{p_2,q_2}$.

Proof. This follows by Bernstein's inequality, Lemma 14.3, as $\Delta_j \partial^\gamma = \partial^\gamma \Delta_j$ it implies that there exists a $C > 0$ such that

$$\|\Delta_j \partial^\gamma u\|_{L^{p_2}} \leq C 2^{j(|\gamma| + d(\frac{1}{p_1} - \frac{1}{p_2}))} \|\Delta_j u\|_{L^{p_1}}.$$

Therefore

$$\|\partial^\gamma u\|_{B^\beta_{p_2,q_2}} \leq C \|u\|_{B^{\beta+|\gamma|+d(\frac{1}{p_1}-\frac{1}{p_2})}_{p_1,q_2}} \qquad (u \in \mathcal{S}').$$

By monotonicity of the norm $\|\cdot\|_{\ell^q}$ in q (see A.6) and by monotonicity of the norm $\|\cdot\|_{B^\alpha_{p,q}}$ in α (see Exercise 15.2) we obtain (125). $\square$

16.2. Observe that the third condition in 124 can be rewritten as

$$\alpha - \frac{d}{p_1} \geq \beta - \frac{d}{p_2} \qquad \text{or} \qquad \alpha + \frac{d}{p_2} \geq \beta + \frac{d}{p_1}.$$

So given that u is an element of $B^\alpha_{p_1,q}$, one can obtain that u is also in a Besov space with a larger parameter than p_1 at the cost of a smaller regularity parameter than α.

An alternative presentation to (125) is

$$\|\partial^\gamma u\|_{B^{\alpha+\frac{d}{p_2}-|\gamma|}_{p_2,q_2}} \leq C \|u\|_{B^{\alpha+\frac{d}{p_1}}_{p_1,q_1}} \qquad (u \in B^{\alpha+\frac{d}{p_1}}_{p_1,q_1}).$$

On the other hand, observe that one can always "increase" the second parameter q, without the need to change the regularity parameter. The following lemma states that one can also decrease the second parameter by paying the littlest amount of regularity.

Lemma 16.3. *For all $q_1, q_2 \in [1, \infty]$ and $\varepsilon > 0$ there exists a $C > 0$ such that for all $\alpha \in \mathbb{R}$ and $p \in [1, \infty]$*

$$\|u\|_{B^{\alpha-\varepsilon}_{p,q_2}} \leq C \|u\|_{B^\alpha_{p,q_1}} \qquad (u \in \mathcal{S}'), \tag{126}$$

that is, B^α_{p,q_1} is continuously embedded in $B^{\alpha-\varepsilon}_{p,q_2}$.

Proof. If $q_1 \le q_2$, then this follows directly from Lemma 16.1 (even for $\varepsilon = 0$). Therefore we assume $q_1 > q_2$. The case $q_1 = \infty$ has already been treated in Exercise 15.2. Let $u \in B^{\alpha-\varepsilon}_{p,q_2}$ and $a_j := \|\Delta_j u\|_{L^p}$. Then by Hölder's inequality (observe that $\frac{q_2}{q_1} + \frac{q_1-q_2}{q_1} = 1$)

$$
\begin{aligned}
\|u\|_{B^{\alpha-\varepsilon}_{p,q_2}} &= \|(2^{j(\alpha-\varepsilon)}a_j)_{j\in\mathbb{N}_{-1}}\|_{\ell^{q_2}} = \left(\sum_{j\in\mathbb{N}_{-1}} 2^{-j\varepsilon q_2}(2^{j\alpha}a_j)^{q_2}\right)^{\frac{1}{q_2}} \\
&\le \left(\sum_{j\in\mathbb{N}_{-1}} 2^{-j\varepsilon\frac{q_1q_2}{q_1-q_2}}\right)^{\frac{q_1-q_2}{q_1q_2}} \left(\sum_{j\in\mathbb{N}_{-1}} (2^{j\alpha}a_j)^{q_1}\right)^{\frac{1}{q_1}} \\
&= \|(2^{-j\varepsilon})_{j\in\mathbb{N}_{-1}}\|_{\frac{q_1q_2}{q_1-q_2}} \|(2^{j\alpha}a_j)_{j\in\mathbb{N}_{-1}}\|_{\ell^{q_1}}.
\end{aligned}
$$

So that with $C = \|(2^{-j\varepsilon})_{j\in\mathbb{N}_{-1}}\|_{\frac{q_1q_2}{q_1-q_2}}$ we have (126). □

16.4. Let $\alpha \in \mathbb{R}$, $p, q \in [1,\infty]$. Let us show that C_c^∞ is a subset of $B^\alpha_{p,q}$ by using Bernstein's inequality. Let $k \in \mathbb{N}$ be such that $k > \alpha$. By Bernstein's inequality (Theorem 14.3) and (110) there exist $C_1, C_2 > 0$ such that for all $u \in \mathcal{S}'$ and $j \in \mathbb{N}_0$

$$
\begin{aligned}
\|\Delta_j u\|_{L^p} &\le C_1^{k+1} 2^{-kj} \max_{\beta\in\mathbb{N}_0^d:|\beta|=k} \|\partial^\beta \Delta_j u\|_{L^p} \\
&\le C_1^{k+1} C_2 2^{-kj} \max_{\beta\in\mathbb{N}_0^d:|\beta|=k} \|\partial^\beta u\|_{L^p}.
\end{aligned}
$$

As the L^p norm of $\Delta_{-1}u$ is also bounded a multiple of $\|u\|_{L^p}$ (see (111)), and as $(2^{(\alpha-k)j})_{j\in\mathbb{N}_{-1}}$ is in ℓ^q, we obtain that there exists a $C > 0$ such that for all $u \in \mathcal{S}'$

$$
\begin{aligned}
\|u\|_{B^\alpha_{p,q}} &\le C\left(\|u\|_{L^p} + \max_{\beta\in\mathbb{N}_0^d:|\beta|=k} \|\partial^\beta u\|_{L^p}\right) && (127) \\
&\le C|\operatorname{supp} u|^{\frac{1}{p}} \|u\|_{C^k}, && (128)
\end{aligned}
$$

where $|\operatorname{supp} u|$ is the Lebesgue measure of $\operatorname{supp} u$. By the above estimate, we in particular obtain that $\mathcal{D}$ is sequentially continuously embedded in $B^\alpha_{p,q}$: If $\varphi_n \to \varphi$ in $\mathcal{D}$ then $\varphi_n \to \varphi$ in $B^\alpha_{p,q}$. Moreover, as we can bound the right–hand side of (127) by the Sobolev norm, we have also obtained part of the following theorem (by observing that $k \ge \alpha$ is sufficient for $q = \infty$).

Theorem 16.5. *Let $\alpha, \beta \in \mathbb{R}$, $k \in \mathbb{N}_0$ and $p, q \in [1,\infty]$. If $\alpha < k < \beta$, then*

$$B^\beta_{p,q} \subset W^{k,p} \subset B^\alpha_{p,q},$$

and there exists a $C > 0$ such that

$$\frac{1}{C}\|\cdot\|_{B^\alpha_{p,q}} \le \|\cdot\|_{W^{k,p}} \le C\|\cdot\|_{B^\beta_{p,q}}. \tag{129}$$

Moreover,

$$B^k_{p,1} \subset W^{k,p} \subset B^k_{p,\infty},$$

and there exists a $C > 0$ such that

$$\frac{1}{C}\|\cdot\|_{B^k_{p,\infty}} \le \|\cdot\|_{W^{k,p}} \le C\|\cdot\|_{B^k_{p,1}}. \tag{130}$$

Proof. By 16.4 we obtain that $W_{k,p}$ is continuously embedded in $B^\alpha_{p,q}$ if $k > \alpha$, and where we may take $k = \alpha$ in case $\alpha \in \mathbb{N}_0$ and $q = \infty$.

Let us first consider $k = 0$ and show that $B^\beta_{p,q}$ is continuously embedded in L^p for $\beta > 0$ or if $(\beta, q) = (0, 1)$ (remember that $W^{0,p} = L^p$). By Hölder's inequality (for ℓ^p spaces: Corollary A.7), for $r \in [1,\infty]$ being such that $1 = \frac{1}{r} + \frac{1}{q}$,

$$\|u\|_{L^p} \le \sum_{i=-1}^{\infty} 2^{-\beta i} 2^{\beta i} \|\Delta_i u\|_{L^p} \le \|(2^{-\beta i})_{i\in\mathbb{N}_{-1}}\|_{\ell^r} \|(2^{\beta i}\|\Delta_i u\|_{L^p})_{i\in\mathbb{N}_{-1}}\|_{\ell^q}.$$

As $\beta > 0$ or $(\beta, q) = (0, 1)$ (and thus $r = \infty$), we have $M := \|(2^{-\beta i})_{i\in\mathbb{N}_{-1}}\|_{\ell^r} \in (0,\infty)$. For general $k \in \mathbb{N}_0$, by the above estimate and by Lemma 16.1 there exists a $L > 0$ such that for all $\gamma \in \mathbb{N}_0^d$ with $|\gamma| \le k$

$$\|\partial^\gamma u\|_{L^p} \le M\|\partial^\gamma u\|_{B^\beta_{p,q}} \le LM\|u\|_{B^{|\gamma|+\beta}_{p,q}} \le LM\|u\|_{B^{k+\beta}_{p,q}}.$$

From this we conclude (129) and (130). □

We have already seen that $\mathcal{D}$ is sequentially continuously embedded in $B^\alpha_{p,q}$. In Theorem 16.7 we will show that $\mathcal{D}$ is also dense in $B^\alpha_{p,q}$ in case p and q are both finite. For this we will use the following lemma.

Lemma 16.6. [2, Lemma 2.73] *Let $\alpha \in \mathbb{R}$ and $p, q \in [1,\infty]$. Suppose $q < \infty$. Then $\sum_{j=-1}^{J} \Delta_j u \to u$ in $B^\alpha_{p,q}$ as $J \to \infty$ for all $u \in B^\alpha_{p,q}$.*

Exercise **16.1.** Prove Lemma 16.6.

Theorem 16.7. [2, Proposition 2.74] *Let $\alpha \in \mathbb{R}$, $k \in \mathbb{N}_0$ and $p, q \in [1,\infty]$. Suppose $p < \infty$ and $q < \infty$. Then $\mathcal{D}$ is dense in $B^\alpha_{p,q}$ and sequentially continuously embedded and $\mathcal{D}$ is dense in $W^{k,p}$ and sequentially continuously embedded.*

Proof. That $\mathcal{D}$ is sequentially continuously embedded in $B^\alpha_{p,q}$ we have already seen in 16.4. That it is also sequentially continuously embedded in $W^{k,p}$ follows by Theorem 16.5.

Let $\varepsilon > 0$ and $u \in B^\alpha_{p,q}$. By Lemma 16.6 there exists a $J \in \mathbb{N}_{-1}$ such that for $u_J = \sum_{j=-1}^{J} \Delta_j u$ one has $\|u_J - u\|_{B^\alpha_{p,q}} < \varepsilon$. As the Fourier support of u_J is compact, u_J is smooth (see Lemma 12.11). Therefore $C^\infty \cap B^\alpha_{p,q}$ is dense in $B^\alpha_{p,q}$ (for all $\alpha \in \mathbb{R}$) and therefore by Theorem 16.5 $C^\infty \cap W^{k,p}$ is dense in $W^{k,p}$ (for all $k \in \mathbb{N}_0$).

Let $\theta \in C_c^\infty$ be equal to 1 on $B(0,1)$ and have support in $B(0,2)$. Write $\theta_R = \theta(\frac{1}{R}\cdot)$. As $\theta_R u_J \in C_c^\infty$ for $u \in C^\infty$ we show

$$\|(\theta_R - 1)u_J\|_{B_{p,q}^\alpha} \xrightarrow{R\to\infty} 0 \text{ for } u \in C^\infty \cap B_{p,q}^\alpha \text{ and } \alpha \in \mathbb{R}, \tag{131}$$

$$\|(\theta_R - 1)u_J\|_{W^{k,p}} \xrightarrow{R\to\infty} 0 \text{ for } u \in C^\infty \cap W^{k,p} \text{ and } k \in \mathbb{N}_0. \tag{132}$$

By Theorem 16.5 (131) follows from (132). Let $k \in \mathbb{N}_0$ and $u \in C^\infty \cap W^{k,p}$. Then by Leibniz rule (see 5.3) there exists a $C > 0$ such that

$$\|(\theta_R - 1)u_J\|_{W^{k,p}} \le C \max_{\beta\in\mathbb{N}_0^d:|\beta|\le k} \max_{\gamma\in\mathbb{N}_0^d:|\gamma|\le k} \|\partial^\gamma(\theta_R - 1)\partial^\beta u_J\|_{L^p}.$$

For all $\gamma \in \mathbb{N}_0^d$ the function $\partial^\gamma(\theta_R - 1)$ converges pointwise to zero as $R \to \infty$. As this function is uniformly bounded in R and $\partial^\beta u_J \in L^p$, which can be concluded from Bernstein's inequality, by Lebesgue's dominated convergence theorem we obtain (132).□

Let us show that $\mathcal{D}$ is not dense in case both p and q equal infinity.

16.8. We have seen that $C_c^\infty \subset B_{p,q}^\alpha$ for all $\alpha \in \mathbb{R}$, $p, q \in [1,\infty]$. Therefore, if $\varphi \in C_c^\infty$, then for all $\alpha \in \mathbb{R}$ and $p \in [1,\infty]$ there exists a $C > 0$ such that $\|\Delta_j\varphi\|_{L^p} \le C2^{-\alpha j}$ for all $j \in \mathbb{N}_{-1}$ and therefore

$$\lim_{j\to\infty} 2^{\alpha j}\|\Delta_j\varphi\|_{L^p} = 0. \tag{133}$$

In particular, if $q \in [1,\infty]$ and φ is in the closure of $\mathcal{D}$ in $B_{p,q}^\alpha$, then (133) holds.

Example 16.9 ($\mathcal{D}$ is not dense in $B_{\infty,\infty}^0$, $L^\infty \subsetneq B_{\infty,\infty}^0$). Let $d = 1$ and $(\rho_j)_{j\in\mathbb{N}_{-1}}$ be a dyadic partition of unity. Let $a > 0$ be such that $\rho_0(a) = 1$ and thus $\rho_i(a2^j)$ equals 1 if $i = j$ and zero otherwise. Let $v \in \mathcal{S}'$ be given by

$$v = \frac{1}{2}\sum_{n\in\mathbb{N}} \delta_{a2^n} + \delta_{-a2^n}.$$

Then $v = \mathcal{F}\hat{v}$ and $u := \hat{v}$ is given by

$$u = \sum_{n\in\mathbb{N}} \cos(2\pi a 2^n \cdot).$$

By assumption we have $\Delta_j u = \mathcal{F}^{-1}(\rho_j v) = \mathcal{F}^{-1}(\delta_{a2^j} + \delta_{-a2^j}) = \cos(2\pi a 2^j \cdot)$. Therefore $\|\Delta_j u\|_{L^\infty} = 1$ for all $j \in \mathbb{N}_{-1}$ and thus

$$\|u\|_{B_{\infty,\infty}^0} = 1.$$

Therefore u cannot be in the closure of $\mathcal{D}$, see 16.8.

Let us show that u is not locally integrable and therefore in particular not an element of L^∞. Let us consider the function $w : \mathbb{R} \to \mathbb{R}$ given by

$$w(x) = \sum_{n\in\mathbb{N}} 2^{-n}\cos(\pi 2^n x) \qquad (x \in \mathbb{R}).$$

Then in distributional sense, u is the derivative of $\frac{1}{2\pi a} l_{2a} w$. w is a Weierstrass function, as Hardy showed, see [15]. This means that w is a continuous function that is nowhere differentiable.

We will derive to a contradiction by assuming that u is locally integrable by the use of Lebesgue's differentiation theorem (Theorem 4.1). For notational convenience we assume $a = 1$ (otherwise replace "u" by "$l_{\frac{1}{2a}} u$" in the following). As we will show, w is the indefinite integral of u, and therefore the mentioned theorem implies that w is almost everywhere integrable, which clearly is a contradiction to the fact that it is nowhere differentiable.

We show that w is the indefinite integral of u, that is

$$\int_a^b u = w(b) - w(a) \qquad (a, b \in \mathbb{R}, a < b).$$

Let $a, b \in \mathbb{R}$ and $a < b$. Similar to the Heaviside function (see Exercise 3.7) in the distributional sense we have that $D\mathbb{1}_{[a,b]} = \delta_a - \delta_b$. Let ψ be a mollifier. Then $\psi_\varepsilon * \mathbb{1}_{[a,b]}$ converges almost everywhere to $\mathbb{1}_{[a,b]}$ by Theorem 4.3. As it is bounded, by Lebesgue's dominated convergence theorem we have

$$\begin{aligned}
\int_a^b u &= \lim_{\varepsilon \downarrow 0} \langle u, \psi_\varepsilon * \mathbb{1}_{[a,b]} \rangle = \lim_{\varepsilon \downarrow 0} \langle w', \psi_\varepsilon * \mathbb{1}_{[a,b]} \rangle \\
&= - \lim_{\varepsilon \downarrow 0} \langle w, \psi_\varepsilon * D\mathbb{1}_{[a,b]} \rangle = - \lim_{\varepsilon \downarrow 0} \langle w, \psi_\varepsilon * (\delta_a - \delta_b) \rangle \\
&= \lim_{\varepsilon \downarrow 0} \langle w, \mathcal{T}_b \psi_\varepsilon \rangle - \langle w, \mathcal{T}_a \psi_\varepsilon \rangle = \lim_{\varepsilon \downarrow 0} w * (\check{\psi})_\varepsilon(b) - w * (\check{\psi})_\varepsilon(a) \\
&= w(b) - w(a),
\end{aligned}$$

where for the last equality we used that w is continuous and Theorem 4.3 (b).

Due to the following lemma we easily show in Theorem 16.11 that $B^\alpha_{2,2} = H^\alpha$, where H^α is as in Definition 13.12.

Lemma 16.10. *For all $\alpha \in \mathbb{R}$ there exists a $C > 0$ such that*

$$2^{2\alpha j} \rho_j(\xi)^2 \le C \left(1 + |\xi|^2\right)^\alpha \qquad (j \in \mathbb{N}_{-1}, \xi \in \mathbb{R}^d). \tag{134}$$

Moreover, for all $\alpha \in \mathbb{R}$ there exist $c, C > 0$ such that

$$c \left(1 + |\xi|^2\right)^\alpha \le \sum_{j \in \mathbb{N}_{-1}} 2^{2\alpha j} \rho_j(\xi)^2 \le C \left(1 + |\xi|^2\right)^\alpha \qquad (\xi \in \mathbb{R}^d). \tag{135}$$

Proof. First we give the proof for $\alpha \ge 0$. Let $a > 0$ be such that $B(0, a) \cap \operatorname{supp} \rho = \emptyset$. Then for $j \in \mathbb{N}_0$ we have

$$\rho_j(\xi)^2 \le \left(\frac{1 + |\xi|^2}{a^2 2^{2j}} \right)^\alpha \qquad (\xi \in \mathbb{R}^d).$$

Hence (134) follows as ρ_{-1} is bounded.

This also implies the upper bound in (135) as $\rho_j(\xi)$ and $\rho_i(\xi)$ are only both nonzero if $|i-j| \le 1$ (by (103)).

By (102) and (103) it follows that

$$\bigcup_{j\in\mathbb{N}_{-1}} [\rho_j^2 \ge \frac{1}{4}] = \mathbb{R}^d,$$

where $[\rho_j^2 \ge \frac{1}{4}] = \{\xi \in \mathbb{R}^d : \rho_j^2(\xi) \ge \frac{1}{4}\}$. Let $b > 2$ be such that $[\rho_j^2 \ge \frac{1}{4}] \subset B(0, 2^j b)$ for all $j \in \mathbb{N}_{-1}$. Then for $\xi \in [\rho_j^2 \ge \frac{1}{4}]$ (note that as $b \ge 2$ one has $b2^j \ge 1$ for all $j \in \mathbb{N}_{-1}$)

$$\frac{1}{4}\left(\frac{1+|\xi|^2}{2b^2 2^{2j}}\right)^{\alpha} \le \frac{1}{4}\left(\frac{1+b^2 2^{2j}}{2b^2 2^{2j}}\right)^{\alpha} \le \frac{1}{4} \le \rho_j(\xi)^2,$$

so that the lower bound in (135) follows.

Now we give the proof for $\alpha < 0$. Let $b \ge 2$ be such that $\operatorname{supp} \rho_j \subset B(0, 2^j b)$ for all $j \in \mathbb{N}_{-1}$. As $\rho_j \le 1$ the bound in (134) follows as for $\xi \in \operatorname{supp} \rho_j \subset B(0, 2^j b)$ one has $1 + |\xi|^2 \le 2b^2 2^{2j}$. Let $a > 0$ be such that $[\rho_j^2 \ge \frac{1}{4}] \subset B(0, 2^j a)^c$ for all $j \in \mathbb{N}_0$. Then for $\xi \in [\rho_j^2 \ge \frac{1}{4}]$

$$\frac{1}{4}\left(\frac{a^2 2^{2j}}{1+|\xi|^2}\right)^{-\alpha} \le \frac{1}{4}\left(\frac{a^2 2^{2j}}{1+2^{2j}a^2}\right)^{-\alpha} \le \frac{1}{4} \le \rho_j(\xi)^2,$$

which implies the lower bound in (135). □

Theorem 16.11. *For all $\alpha \in \mathbb{R}$ we have*

$$B^{\alpha}_{2,2} = H^{\alpha},$$

with equivalent norms.

Proof. By the Plancherel formula (Theorem 11.21),

$$\begin{aligned}\|u\|^2_{B^{\alpha}_{2,2}} &= \sum_{j\in\mathbb{N}_{-1}} 2^{2\alpha j}\|\rho_j(\mathrm{D})u\|^2_{L^2} = \sum_{j\in\mathbb{N}_{-1}} 2^{2\alpha j}\|\rho_j \hat{u}\|^2_{L^2} \\ &= \int_{\mathbb{R}^d} \sum_{j\in\mathbb{N}_{-1}} 2^{2\alpha j}|\rho_j(\xi)|^2|\hat{u}(\xi)|^2 \,\mathrm{d}\xi.\end{aligned}$$

The rest follows from Lemma 16.10. □

16.12. In particular, Theorem 16.11 implies $L^2 = B^0_{2,2}$. However, there do not exist $s \in \mathbb{R}$, $p, q \in [1,\infty]$ such that $L^1 = B^s_{p,q}$, see Exercise 16.2

Exercise **16.2.** Show that there do not exist $s \in \mathbb{R}$, $p, q \in [1,\infty]$ such that $L^1 = B^s_{p,q}$. Hint: Use the property of Theorem 15.15 and Exercise 15.3.

Question **16.1.** Is each tempered distribution in a Besov space? That is, does the following equality hold?

$$\mathcal{S}' = \bigcup_{p,q\in[1,\infty]} \bigcup_{s\in\mathbb{R}} B^s_{p,q}.$$

17 *Besov spaces related to other spaces*

In this section we give an overview of other spaces and embeddings between those and Besov spaces.

Definition 17.1 (Hölder spaces). Let Ω be an open subset of $\mathbb{R}^d$ and $k \in \mathbb{N}_0$. We write also $C^{k,0}(\Omega)$ for $C^k(\Omega)$. Let $\alpha \in (0,1]$.

- A function $f : \Omega \to \mathbb{F}$ is *α-Hölder continuous* if there exists a $C > 0$ such that
$$|f(x) - f(y)| \le C|x-y|^\alpha \qquad (x, y \in \Omega). \tag{136}$$
- $C^{0,\alpha}(\Omega)$ is defined to be the space of α-Hölder continuous functions $\Omega \to \mathbb{F}$. The Hölder coefficient of a function f is given by
$$|f|_{C^{0,\alpha}(\Omega)} = \sup_{x,y\in\Omega: x\ne y} \frac{|f(x)-f(y)|}{|x-y|^\alpha}.$$
- $C^{k,\alpha}(\Omega)$ is defined to be the space of functions $\Omega \to \mathbb{F}$ that are k-times continuously differentiable for which their derivatives of order k are α-Hölder continuous.

We already defined $C_\mathrm{b}^k(\Omega)$ to be those elements of $C^k(\Omega)$ for which $\|\cdot\|_{C^k}$ is finite, similarly we define

$$\|f\|_{C^{k,\alpha}(\Omega)} = \|f\|_{C^{k,\alpha}(\Omega)} = \|f\|_{C^k(\Omega)} + \sum_{\beta\in\mathbb{N}_0^d:|\beta|=k} |\partial^\beta f|_{C^{0,\alpha}} \qquad (f \in C^{k,\alpha}(\Omega)), \tag{137}$$
$$C_\mathrm{b}^{k,\alpha}(\Omega) = \{f \in C^{k,\alpha}(\Omega) : \|f\|_{C^{k,\alpha}(\Omega)} < \infty\}. \tag{138}$$

17.2. For the rest of this section we consider $\Omega = \mathbb{R}^d$ and write "$C^{k,\alpha}$" instead of "$C^{k,\alpha}(\mathbb{R}^d)$".

Observe that $C^{0,1}$ consists of all the Lipschitz functions and that for $k \in \mathbb{N}$, $C_\mathrm{b}^{k+1} \subsetneq C_\mathrm{b}^{k,1}$.

For $s \in (0,\infty) \setminus \mathbb{N}$ it is also common in literature to write C^s for $C^{k,\alpha}$, where $k = \lfloor s \rfloor$ and $\alpha = s - \lfloor s \rfloor$.

$\boxed{\textbf{Exercise}}$ **17.1.** Can you classify the space of α-Hölder functions with $\alpha > 1$, that is, which functions f satisfy (136) for $\alpha > 1$.

In Definition 8.1 we introduced the Sobolev spaces $W^{k,p}$ for $k \in \mathbb{N}_0$ and $p \in [1,\infty]$. In Definition 13.12 and 13.15 we introduced the fractional Sobolev or Bessel–potential spaces H_p^s for $s \in \mathbb{R} \setminus \mathbb{N}_0$ and $p \in [1,\infty]$. We will now consider Slobodeckij spaces, $W^{s,p}$ with $s \in (0,\infty) \setminus \mathbb{N}$ as subspaces of $W^{k,p}$ with $k = \lfloor s \rfloor$ in a similar way as C^s or $C^{k,\alpha}$ is defined to be a subspace of C^k, where (with $\alpha = s - k$).

Definition 17.3 (Slobodeckij spaces). Let $p \in [1, \infty)$ and $s \in (0, \infty) \setminus \mathbb{N}$. Let $k \in \mathbb{Z}$ and $\alpha \in (0, 1)$ be given by

$$k = \lfloor s \rfloor, \qquad \alpha = s - k.$$

For $f \in W^{k,p}$ we define

$$\|f\|_{W^{s,p}} := \|f\|_{W^{k,p}} + \sum_{\alpha \in \mathbb{N}_0^d : |\alpha| = k} \left(\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{|\partial^\alpha f(x) - \partial^\alpha f(y)|^p}{|x - y|^{d + \alpha p}} \, \mathrm{d}x \, \mathrm{d}y \right)^{\frac{1}{p}}.$$

We define the *Slobodeckij space* $W^{s,p}$ by

$$W^{s,p} = \{ f \in W^{k,p} : \|f\|_{W^{s,p}} < \infty \}.$$

Definition 17.4 (Zygmund spaces). Let $s \in \mathbb{R}$. Let $k \in \mathbb{Z}$ and $\alpha \in (0, 1]$ be given by

$$k = \lceil s - 1 \rceil, \qquad \alpha = s - k, \tag{139}$$

in other words, k is such that $s - k \in (0, 1]$. We define $\| \cdot \|_{\mathcal{C}^s} : C^k \to \mathbb{R}$, by

$$\|f\|_{\mathcal{C}^s} = \|f\|_{C^k} + \sum_{\beta \in \mathbb{N}_0^d : |\beta| = k} \sup_{h \in \mathbb{R}^d \setminus \{0\}} \frac{\|(\mathcal{T}_h - 1)^2 \partial^\beta f\|_{C^0}}{|h|^\alpha},$$

and the *Zygmund space* $\mathcal{C}^s$ by

$$\mathcal{C}^s = \{ f \in C^k : \|f\|_{\mathcal{C}^s} < \infty \}.$$

Observe that

$$\begin{aligned} (\mathcal{T}_h - 1)^2 g(x) &= (\mathcal{T}_h - 1)(\mathcal{T}_h - 1) g(x) = (\mathcal{T}_h - 1) g(x - h) - (\mathcal{T}_h - 1) g(x) \\ &= g(x - 2h) - 2g(x - h) + g(x). \end{aligned}$$

Definition 17.5 (Besov–Lipschitz spaces). Let $s \in (0, \infty)$. Let $k \in \mathbb{Z}$ and $\alpha \in (0, 1]$ be as in (139). For $p, q \in [1, \infty)$ we define for $f \in W^{k,p}$

$$\begin{aligned} \|f\|_{\Lambda^s_{p,q}} &:= \|f\|_{W^{k,p}} + \sum_{\beta \in \mathbb{N}_0^d : |\beta| = k} \left(\int_{\mathbb{R}^d} \frac{\|(\mathcal{T}_h - 1)^2 \partial^\beta f\|_{L^p}^q}{|h|^{d + \alpha q}} \, \mathrm{d}h \right)^{\frac{1}{q}}, \\ \|f\|_{\Lambda^s_{p,\infty}} &:= \|f\|_{W^{k,p}} + \sum_{\beta \in \mathbb{N}_0^d : |\beta| = k} \sup_{h \in \mathbb{R}^d \setminus \{0\}} \frac{\|(\mathcal{T}_h - 1)^2 \partial^\alpha f\|_{L^p}^q}{|h|^\alpha}. \end{aligned}$$

For $q \in [1, \infty]$ we define the *Besov-Lipschitz space* $\Lambda^s_{p,q}$ to be the set of for which the above norm is finite:

$$\Lambda^s_{p,q} = \{ f \in W^{k,p} : \|f\|_{\Lambda^s_{p,q}} < \infty \}.$$

The Triebel–Lizorkin spaces are defined as the Besov spaces, but with the "L^p" and "ℓ^q" norm interchanged:

Definition 17.6 (Triebel–Lizorkin spaces). Let $(\rho_j)_{j\in\mathbb{N}_{-1}}$ be a dyadic parition of unity. Let $s \in \mathbb{R}$. For $p \in [1,\infty)$ and $q \in [1,\infty]$ we define

$$\|u\|_{F^s_{p,q}} := \|\|(2^{js}|\Delta_i u|)_{j\in\mathbb{N}_{-1}}\|_{\ell^q}\|_{L^p},$$

for example, for $q < \infty$ this means

$$\|u\|_{F^s_{p,q}} = \left[\int_{\mathbb{R}^d} \Big(\sum_{j\in\mathbb{N}_{-1}} 2^{qjs}|\Delta_i u(x)|^q \Big)^{\frac{p}{q}} \,\mathrm{d}x\right]^{\frac{1}{p}}.$$

We define the *Triebel–Lizorkin space* $F^s_{p,q}$ to be the set of tempered distributions for which the above norm is finite:

$$F^s_{p,q} = \{u \in \mathcal{S}' : \|u\|_{F^s_{p,q}} < \infty\}.$$

Remark 17.7. As for Besov spaces, the norm of $F^s_{p,q}$ depends on the choice of dyadic partition, but the space itself does not. This is shown in [33, Section 2.3.2].

17.8. Let us summarize for which parameters we have either continuous embeddings or equality between spaces with equivalent norms. Here, "$A \subset B$" means that the space A is continuously embedded in B, and "$A = B$" means that A and B are the same space with equivalent norms.

(a) [33, p.90, (9)] $C^s_{\mathrm{b}} = \mathcal{C}^s$ for $s \in (0,\infty)\setminus\mathbb{N}$ (C^s is as in 17.2).

(b) [33, p.90, (9)] $W^{s,p} = \Lambda^s_{p,p}$ for $s \in (0,\infty)\setminus\mathbb{N}$ and $p \in (1,\infty)$.

(c) [33, p.88] $H^s_p = F^s_{p,2}$ for $s \in \mathbb{R}$ and $p \in (1,\infty)$.

(d) [33, p.88] $H^k_p = W^{k,p}$ for $k \in \mathbb{N}$ and $p \in (1,\infty)$.

(e) [33, p.89] $B^0_{p,1} \subset L^p \subset B^0_{p,\infty}$ for $p \in [1,\infty)$.

(f) [33, p.89] $B^0_{\infty,1} \subset C^0_{\mathrm{b}} \subset B^0_{\infty,\infty}$.

(g) [33, p.90, p.113] $\Lambda^s_{p,q} = B^s_{p,q}$ for $s > 0$, $p \in [1,\infty)$ and $q \in [1,\infty]$.

(h) [33, p.90, p.113] $\mathcal{C}^s = B^s_{\infty,\infty}$ for $s > 0$.

(i) [33, p.47] $B^s_{p,\min\{p,q\}} \subset F^s_{p,q} \subset B^s_{p,\max\{p,q\}}$ for $s \in \mathbb{R}$, $p \in [1,\infty)$ and $q \in [1,\infty]$.

(j) [33, p.60] For $s_1, s_2 \in \mathbb{R}$, $p_1, p_2, q_1, q_2 \in [1,\infty]$:
$B^{s_1}_{p_1,q_1}(\mathbb{R}^d) = B^{s_2}_{p_2,q_2}(\mathbb{R}^d)$ if and only if $s_1 = s_2$ and $p_1 = p_2, q_1 = q_2$.
For $p_1, p_2 < \infty$:
$F^{s_1}_{p_1,q_1}(\mathbb{R}^d) = F^{s_2}_{p_2,q_2}(\mathbb{R}^d)$ if and only if $s_1 = s_2$ and $p_1 = p_2, q_1 = q_2$,
$F^{s_1}_{p_1,q_1}(\mathbb{R}^d) = B^{s_2}_{p_2,q_2}(\mathbb{R}^d)$ if and only if $s_1 = s_2$ and $p_1 = p_2 = q_1 = q_2$.

Observe that we can combine some of the above to obtain:

$$\begin{aligned}
&C_{\mathrm{b}}^{s} = \mathcal{C}^{s} = B^{s}_{\infty,\infty} \qquad (s \in (0,\infty) \setminus \mathbb{N}),\\
&W^{s,p} = \Lambda^{s}_{p,p} = B^{s}_{p,p} = F^{s}_{p,p} \qquad (s \in (0,\infty) \setminus \mathbb{N}, p \in (1,\infty)).\\
&H^{k}_{p} = W^{k,p} = F^{k}_{p,2} \qquad (k \in \mathbb{N}, p \in (1,\infty)).
\end{aligned}$$

17.9. In 16.12 we mentioned that there is no Besov space that is equal to L^1. We can generalise this as follows: For $r \in [1,2) \cup (2,\infty)$ there are no $s \in \mathbb{R}$, $p,q \in [1,\infty]$ such that $L^r = B^{s}_{p,q}$.

Exercise **17.2.** Let $r \in (1,\infty)$. Show that $B^{s}_{p,q} = L^r$ if and only if $p = q = r = 2$ and $s = 0$. Hint: $H^0_r = L^r$ (see 13.15).

Remark 17.10. The proof of $C^{\alpha} = B^{\alpha}_{\infty,\infty}$ for $\alpha \in (0,1)$ can also be found in [20, Lemma 8.6].

18 Fourier–multipliers on Besov spaces

We will use the Hörmander–Mikhlin inequality to show that under some conditions, Fourier multipliers map Besov spaces into other Besov space. We will only need the version of the Hörmander–Mikhlin inequality, Lemma 14.12. The following lemma shows that the condition of Lemma 14.12 can be described in a different way.

Lemma 18.1. *Let $\sigma \in C^{\infty}$ and $m \in \mathbb{R}$. Then there exists a $\theta > 0$ such that $\mathfrak{M}_{m,\theta}(\sigma) < \infty$ if and only if*

$$\max_{\alpha \in \mathbb{N}_0^d : |\alpha| \le k} \sup_{x \in \mathbb{R}^d} (1+|x|)^{|\alpha|-m} |\partial^{\alpha}\sigma(x)| < \infty, \tag{140}$$

i.e., if and only if there exists a $C > 0$ such that

$$|\partial^{\alpha}\sigma(x)| \le C(1+|x|)^{m-|\alpha|} \qquad (x \in \mathbb{R}^d, \alpha \in \mathbb{N}_0^d, |\alpha| \le k). \tag{141}$$

Proof. We have already seen that $\mathfrak{M}_{m,\theta}(\sigma) < \infty$ for all $\theta > 0$ (see 14.11). Hence we may take $\theta = 1$. We have $\mathfrak{M}_{m,1} < \infty$ if and only if there exists a $C > 0$ such that

$$|\partial^{\alpha}\sigma(x)| \le C|x|^{m-|\alpha|} \qquad (x \in \mathbb{R}^d, |x| \ge 1, \alpha \in \mathbb{N}_0^d, |\alpha| \le k). \tag{142}$$

As

$$\frac{1}{2}(1+|x|) \le |x| \le 1+|x| \qquad (x \in \mathbb{R}^d, |x| \ge 1),$$

(142) is equivalent to

$$|\partial^{\alpha}\sigma(x)| \le C(1+|x|)^{m-|\alpha|} \qquad (x \in \mathbb{R}^d, |x| \ge 1, \alpha \in \mathbb{N}_0^d, |\alpha| \le k). \tag{143}$$

Let us show that (143) is equivalent to (141).

As $1 \le 1+|x| \le 2$ for all $x \in B(0,1)$, there exists an $C_1 > 0$ such that

$$1 \le C_1(1+|x|)^{m-|\alpha|} \qquad (x \in B(0,1), \alpha \in \mathbb{N}_0^d, |\alpha| \le k).$$

As σ is smooth, its restriction to $B(0,1)$ is bounded in C^k-norm. Let $C_2 = \|\sigma\|_{C^k(B(0,1))}$. Then

$$|\partial^\alpha \sigma(x)| \le C_2 \le C_1 C_2 (1+|x|)^{m-|\alpha|} \qquad (x \in B(0,1), \alpha \in \mathbb{N}_0^d, |\alpha| \le k).$$

This shows that (143) is equivalent to (141). □

Now we show that if the condition in Lemma 18.1 is satisfied for σ that it forms a continuous map between Besov spaces.

Theorem 18.2. [2, Theorem 2.78] and [13, Lemma A.5] *Let $m, s \in \mathbb{R}$, $p, q \in [1,\infty]$. Let $\sigma \in C_{\mathrm{p}}^\infty$ be such that* (140) *holds. Then there exists a $C > 0$ such that*

$$\|\sigma(\mathrm{D})u\|_{B_{p,q}^{s-m}} \le C\|u\|_{B_{p,q}^s} \qquad (u \in B_{p,q}^s). \tag{144}$$

In other words, $\sigma(\mathrm{D})$ forms a continuous operator $B_{p,r}^s \to B_{p,r}^{s-m}$. Moreover, if $\mathcal{F}(\sigma) \in L^1$, then there exists a $C > 0$ such that

$$\|(l_\lambda \sigma)(\mathrm{D})u\|_{B_{p,q}^{s-m}} \le C(\lambda^m \vee 1)\|u\|_{B_{p,q}^s} \qquad (u \in B_{p,q}^s, \lambda > 0). \tag{145}$$

Proof. By Lemma 14.12, (97) and Lemma 18.1 there exists a $C > 0$ such that

$$\|(l_\lambda \sigma)(\mathrm{D})\Delta_j u\|_{L^p} \le C 2^{jm} \lambda^m \|\Delta_j u\|_{L^p} \qquad (j \in \mathbb{N}_0, u \in \mathcal{S}'). \tag{146}$$

Therefore, by Theorem 15.9, for (144) it is sufficient to show

$$\|\sigma(\mathrm{D})\Delta_{-1} u\|_{L^p} \le C\|\Delta_{-1} u\|_{L^p} \qquad (u \in \mathcal{S}'), \tag{147}$$

and for (145) it is sufficient to show

$$\|(l_\lambda \sigma)(\mathrm{D})\Delta_{-1} u\|_{L^p} \le C\|\Delta_{-1} u\|_{L^p} \qquad (u \in \mathcal{S}', \lambda > 0). \tag{148}$$

Let $\psi \in C_c^\infty$ be such that $\psi = 1$ on $\operatorname{supp} \rho_{-1}$. Then

$$\sigma(\mathrm{D})\Delta_{-1} u = \sigma(\mathrm{D})\psi(\mathrm{D})\Delta_{-1} u = (\sigma\psi)(\mathrm{D})\Delta_{-1} u = \mathcal{F}^{-1}(\sigma\psi) * u.$$

Hence, by applying Young's inequality we obtain (147) with $C = \|\mathcal{F}^{-1}(\sigma\psi)\|_{L^1}$, which is finite as $\sigma\psi \in C_c^\infty \subset \mathcal{S}$ and thus $\mathcal{F}^{-1}(\sigma\psi) \in \mathcal{S} \subset L^1$.

If $\mathcal{F}(\sigma) \in L^1$, then (148) holds with $C = \|\mathcal{F}^{-1}(\sigma)\|_{L^1}$ as this equals $\|\mathcal{F}^{-1}(\sigma)\|_{L^1}$ for all $\lambda > 0$. □

$\boxed{\textbf{Exercise}}$ **18.1.** Let $\beta \in \mathbb{R}$, $p, q \in [1, \infty]$ and $\gamma \in \mathbb{N}_0^d$. In Lemma 16.1 we have seen that there exists a $C > 0$ such that

$$\|\partial^\gamma u\|_{B_{p,q}^{\beta-|\gamma|}} \leq C\|u\|_{B_{p,q}^\beta} \qquad (u \in \mathcal{S}').$$

Show that this can also be derived from Theorem 18.2.

Example 18.3. Let us also apply the above to the Bessel–Potentials, $(1 - \Delta)^{\frac{s}{2}}$ for $s \in \mathbb{R}$. By Lemma 14.15 we have that for $\sigma(x) = (1 + |x|^2)^{\frac{s}{2}}$, $\mathfrak{M}_{s,1}(\sigma) < \infty$. Therefore, by Theorem 18.2 we have that $(1 - \Delta)^{\frac{s}{2}}$, being $\sigma(\mathrm{D})$, maps $B_{p,q}^\alpha$ continuously into $B_{p,q}^{\alpha-s}$. Of course, if we take $s = 2$, as one should expect, $(1 - \Delta)$ lowers the regularity by 2.

19 Paraproducts

In this section we consider the definition of a product of two distributions (for which this product makes sense). Let $u, v \in \mathcal{S}'$. As (see Lemma 15.5)

$$u = \sum_{i \in \mathbb{N}_{-1}} \Delta_i u, \quad v = \sum_{i \in \mathbb{N}_{-1}} \Delta_i v,$$

formally the product of u, v should equal

$$uv = \sum_{i,j \in \mathbb{N}_{-1}} \Delta_i u \Delta_j v.$$

This decomposition, that is the series on the right-hand side, will be further decomposed in terms of two 'paraproducts' and a 'resonance product' (see Definition 19.1).

$\boxed{\textbf{Question}}$ **19.1.** Does it hold that

$$\psi v = \lim_{J \to \infty} \sum_{i,j=-1}^{J} \Delta_i \psi \Delta_j v?$$

Definition 19.1. For $u \in \mathcal{S}'$ and $j \in \mathbb{N}_{-1}$ we write

$$S_j u = \sum_{i=-1}^{j} \Delta_i u \quad (j \in \mathbb{N}_{-1}), \qquad S_{-2}u = 0, \qquad S_{-3}u = 0.$$

Moreover, we will use the following -a priori formal- notations for $u, v \in \mathcal{S}'$:

$$u \varolessthan v = \sum_{j=-1}^{\infty} \sum_{i=-1}^{j-2} \Delta_i u \Delta_j v = \sum_{j=-1}^{\infty} S_{j-2} u \Delta_j v,$$

$$u \odot v = \sum_{j=-1}^{\infty} \sum_{i=j-1}^{j+1} \Delta_i u \Delta_j v.$$

If $u \varolessthan v$ exists in $\mathcal{S}'$, then it is called the *paraproduct* of v by u. If $u \odot v$ exists in $\mathcal{S}'$, then it is called the *resonance product* of u and v. We also write $u \varogreaterthan v := v \varolessthan u$.

Remark 19.2 (About notation). In many textbooks one writes "$T_u v$" for the paraproduct instead of "$u \varolessthan v$" (for example in [2]). In this sense one views T_u as an multiplying operator. Also "$\Pi(u,v)$" or "$R(u,v)$" is written for the resonance product. In the application to SPDEs in the authors of the paper [13] wrote "$u \prec v$" and "$u \circ v$" for the para- and resonance product, respectively. The latter notation changed in the SPDE literature, with some authors creating new symbols, for example "$<$" and "$=$" with a circle around it. In the latter case, "$\leq$" with a circle around it is then used for the sum of the paraproduct and the resonance product, for which the authors of [13] used "$\preceq$".

For this in our notation we could write $\mathrlap{\varodot}{\varolessthan}$. The following table presents the latex commands for the symbols used in these notes.

\varolessthan	$\varolessthan$
\varogreaterthan	$\varogreaterthan$
\varodot	$\varodot$
\mathrlap{\varodot}{\varolessthan}	$\mathrlap{\varodot}{\varolessthan}$

19.3 (Intuition behind the bound on paraproducts). In Theorem **??** we will bound the Besov norm of the paraproduct $u \varolessthan v$. The idea is as follows. Let us say that u is of regularity α if it is in some $B^{\alpha}_{p,q}$ space for some $p,q \in [1,\infty]$. It will turn out that we need some restrictions for the different parameters p,q for u and v but this we forget for the moment and concentrate on the regularity.

Suppose u is of regularity α and v of regularity β. Then the regularity of $u \varolessthan v$ for strictly positive α equals the regularity of β. The idea behind this is that if one multiplies a low frequency function with a high frequency function, the frequency of the product has a frequency equal to the highest frequency. For an illustration see Figure 1 and 2. In case the regularity of u is strictly negative, then the regularity of the product is the sum $\alpha+\beta$. So that the low frequencies of u still worsen the regularity of $u \varolessthan v$.

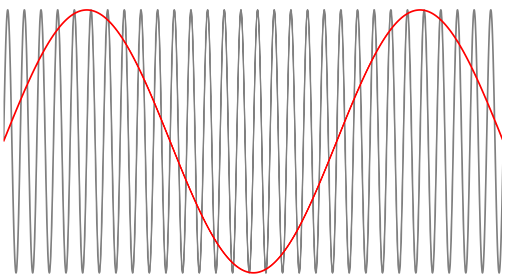

Figure 1: A function with high and low frequency.

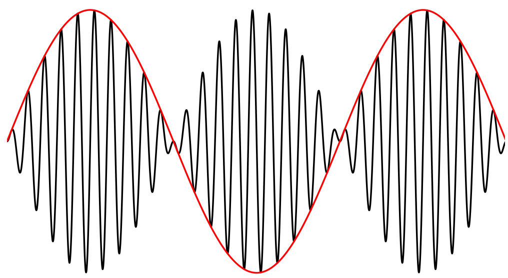

Figure 2: The product of the functions with high and low frequencies.

Theorem 19.4. [2, Theorem 2.82], [21, Lemma 2.1] *Let $p, p_1, p_2, q, q_1, q_2 \in [1,\infty]$ be such that $\frac{1}{p_1} + \frac{1}{p_2} \leq 1$ and*

$$\tfrac{1}{p} = \tfrac{1}{p_1} + \tfrac{1}{p_2}, \qquad \tfrac{1}{q} = \min\{1, \tfrac{1}{q_1} + \tfrac{1}{q_2}\},$$

For $C = \|\mathcal{F}^{-1}(\chi)\|_{L^1} > 0$ *and all* $\alpha \in \mathbb{R}$

$$\|u \otimes v\|_{B^{\alpha}_{p,q}} \le C\|u\|_{L^{p_1}}\|v\|_{B^{\alpha}_{p_2,q}} \qquad (u, v \in \mathcal{S}'). \tag{149}$$

For all $\alpha < 0$ *and* $\beta \in \mathbb{R}$

$$\|u \otimes v\|_{B^{\alpha+\beta}_{p,q}} \le (1 - 2^{\alpha})^{-1}\|u\|_{B^{\alpha}_{p_1,q_1}}\|v\|_{B^{\beta}_{p_2,q_2}} \qquad (u, v \in \mathcal{S}'). \tag{150}$$

Proof. Let $u, v \in \mathcal{S}'$. We will invoke Theorem 15.9 (b) for the functions $S_{j-2}u\Delta_j v = \sum_{i=-1}^{j-2} \Delta_i u \Delta_j v$ with $j \in \mathbb{N}_{-1}$. Observe that $S_{-3}u\Delta_{-1}v = 0$ and $S_{-2}u\Delta_0 v = 0$, and so their Fourier transform is trivially supported in any ball. Let us check that for $j \in \mathbb{N}_0$ the Fourier transform of $S_{j-2}u\Delta_j v$ is supported in $2^j\tilde{\mathcal{A}}$ for some annulus $\tilde{\mathcal{A}}$.

Let $a, b > 0$, $a < b$ be such that $\operatorname{supp}\rho = A(a, b)$. Then in particular $\operatorname{supp}\rho_{-1} \subset B(0, b)$ and so $\operatorname{supp}\rho_j \subset B(0, 2^j b)$. We write $\mathcal{A} = A(a, b)$ and $\mathcal{B} = \overline{B(0, b)}$.

As $\mathcal{F}(\Delta_i u \Delta_j v) = (\rho_i \hat{u}) * (\rho_j \hat{v})$, we have (see Theorem 3.9 and Lemma 3.7):

$$\operatorname{supp}\mathcal{F}(\Delta_i u \Delta_j v) \subset 2^j(2^{i-j}\mathcal{B} + \mathcal{A}) \qquad (i \in \mathbb{N}_{-1}, j \in \mathbb{N}_0). \tag{151}$$

For all $i \in \mathbb{N}_{-1}$ with $i \le j - 2$ we have $2^{i-j} \le 2^{-2}$ and thus $2^{i-j}\mathcal{B} \subset 2^{-2}\mathcal{B}$. As $\operatorname{supp}\rho_0 \cap \operatorname{supp}\rho_2 = \emptyset$ we have $\mathcal{A} \cap 2^2\mathcal{A} = \emptyset$ and thus $2^{-2}\mathcal{B} \cap \mathcal{A} = \emptyset$. Therefore $\tilde{\mathcal{A}} := \mathcal{A} + \overline{B(0, 2^{-2}b)}$ is an annulus and

$$\operatorname{supp}\mathcal{F}\left(\sum_{i=-1}^{j-2} \Delta_i u \Delta_j v\right) \subset 2^j\tilde{\mathcal{A}} \qquad (j \in \mathbb{N}_0).$$

By Hölder's inequality (Theorem A.3) and by (111),

$$\|S_{j-2}u\Delta_j v\|_{L^p} \le \|S_{j-2}u\|_{L^{p_1}}\|\Delta_j v\|_{L^{p_2}} \le \|\mathcal{F}^{-1}(\chi)\|_{L^1}\|u\|_{L^{p_1}}\|\Delta_j v\|_{L^{p_2}} \qquad (j \in \mathbb{N}_{-1}).$$

By this and Theorem 15.9 one obtains (149).

By Hölder's inequality (Corollary A.7), we get

$$\begin{aligned}
\left\|\left(2^{j(\alpha+\beta)}\|S_{j-2}u\Delta_j v\|_{L^p}\right)_{j\in\mathbb{N}_{-1}}\right\|_{\ell^q} &\le \left\|\left(2^{j\alpha}\|S_{j-2}u\|_{L^{p_1}}2^{j\beta}\|\Delta_j v\|_{L^{p_2}}\right)_{j\in\mathbb{N}_{-1}}\right\|_{\ell^q} \\
&\le \left\|\left(2^{j\alpha}\|S_{j-2}u\|_{L^{p_1}}\right)_{j\in\mathbb{N}_{-1}}\right\|_{\ell^{q_1}} \left\|\left(2^{j\beta}\|\Delta_j v\|_{L^{p_2}}\right)_{j\in\mathbb{N}_{-1}}\right\|_{\ell^{q_2}}.
\end{aligned}$$

By Theorem 15.16

$$\left\|\left(2^{j\alpha}\|S_{j-2}u\|_{L^{p_1}}\right)_{j\in\mathbb{N}_{-1}}\right\|_{\ell^{q_1}} = 2^{2\alpha}\left\|\left(2^{j\alpha}\|S_j u\|_{L^{p_1}}\right)_{j\in\mathbb{N}_{-1}}\right\|_{\ell^{q_1}} \le \frac{2^{2\alpha}}{1 - 2^{\alpha}}\|u\|_{B^{\alpha}_{p_1,q_1}},$$

as so we conclude (150) by Theorem 15.9. □

$\boxed{\textbf{Question}}$ **19.2.** Can we show that (150) does not hold in case $\alpha = 0$? For example, can we take $u = v$ equal to the example in Example 16.9. [[I plan to look at this, but maybe you want to provide me the answer before I possibly include this in the notes]]

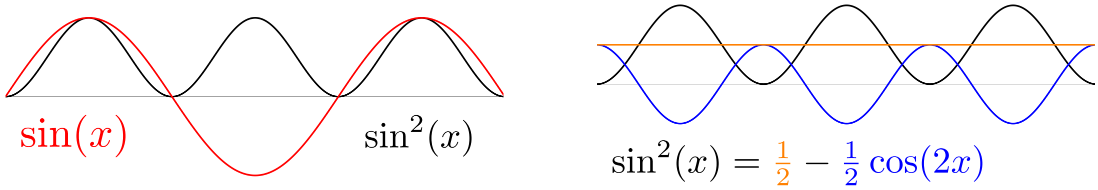

Figure 3: The sine function and its square and the decomposition of the square of the sine function in a low and high frequency function.

19.5 (Intuition behind the bound on the resonance product). In Theorem 19.6 we consider the resonance product of u and v. Let us use the language as in 19.3, with u of regularity α and v of regularity β. The block of $\Delta_i u$ has a frequency of the order $2^{-i\alpha}$ and $\Delta_i v$ of $2^{-i\beta}$, so that the product is of frequency $2^{-i(\alpha+\beta)}$. This already indicates that the regularity of the resonance product should be $\alpha+\beta$. The term "resonance" is used as one considers the outcome of two 'systems' that interact with the same or similar frequency, but also as this may 'strengthen' the outcome. Let us for example consider the product of two sine functions. As is illustrated in Figure 3 we see that the product can be decomposed in terms of a function with lower and one with higher frequencies. Therefore the frequencies of $\Delta_i u \Delta_i v$ range from zero frequencies up to the order $2i$. Hence the k-th Littlewood-Paley block $\Delta_k(u \odot v)$ possibly contains information of $\Delta_i u$ for i inbetween 0 and $2i$. In order to 'deal' with that it makes sense to impose the condition $\alpha+\beta>0$ in order to have some summability (such condition is also assumed in Theorem 15.9(c), on which the proof of Theorem 19.6 relies).

Theorem 19.6. *Let $\alpha, \beta \in \mathbb{R}$ and $\alpha+\beta \geq 0$. There exists a $C>0$ such that for all $p, p_1, p_2, q, q_1, q_2 \in [1,\infty]$ such that $\frac{1}{p_1}+\frac{1}{p_2} \leq 1$, with*

$$\tfrac{1}{p} = \tfrac{1}{p_1} + \tfrac{1}{p_2}, \qquad \tfrac{1}{q} = \min\{1, \tfrac{1}{q_1} + \tfrac{1}{q_2}\},$$

we have in case $\alpha+\beta>0$,

$$\|u \odot v\|_{B^{\alpha+\beta}_{p,q}} \leq C\|u\|_{B^{\alpha}_{p_1,q_1}} \|v\|_{B^{\beta}_{p_2,q_2}} \qquad (u, v \in \mathcal{S}'), \tag{152}$$

and in case $\alpha+\beta=0$ and $\{q_1,q_2\}=\{1,\infty\}$,

$$\|u \odot v\|_{B^{0}_{p,\infty}} \leq C\|u\|_{B^{\alpha}_{p_1,q_1}} \|v\|_{B^{\beta}_{p_2,q_2}} \qquad (u, v \in \mathcal{S}'). \tag{153}$$

Proof. Let $u \in B^{\alpha}_{p_1,q_1}, v \in B^{\beta}_{p_2,q_2}$. We define $\Delta_{-2}u := 0$ and

$$R_j := \left(\sum_{i=-1}^{1} \Delta_{j+i} u\right) \Delta_j v \qquad (j \in \mathbb{N}_{-1}),$$

so that $u \odot v = \sum_{j\in\mathbb{N}_{-1}} R_j$. Then

$$\operatorname{supp} \hat{R}_j \subset \operatorname{supp} \rho_{j+i}\hat{u} * \rho_j \hat{v} \subset \operatorname{supp} \rho_{j+i} + \operatorname{supp} \rho_j \subset 2^j \mathcal{B}.$$

for some ball $\mathcal{B}$ around the origin: Indeed, if $\mathcal{B}_1$ is a ball around the origin such that $\operatorname{supp} \rho_{-1} \cup \operatorname{supp} \rho_0 \subset \mathcal{B}_1$. Then $\rho_{j+i} + \operatorname{supp} \rho_j \subset (2^{j+i} + 2^j)\mathcal{B}_1 \subset 2^j 3\mathcal{B}_1$, so that we can take $\mathcal{B} = 3\mathcal{B}_1$. Observe that $\{q_1, q_2\} = \{1, \infty\}$ is equivalent to $q = 1$. Therefore, in case $\alpha + \beta > 0$, we can use Theorem 15.9(c) and in case $\alpha + \beta = 0$ and $\{q_1, q_2\} = \{1, \infty\}$ we can use Theorem 15.9 (d) to obtain that it is sufficient to show the existence of a $C > 0$ such that

$$\left\| (2^{j(\alpha+\beta)} \|R_j\|_{L^p})_{j\in\mathbb{N}_{-1}} \right\|_{\ell^q} \le C \|u\|_{B^\alpha_{p_1,q_1}} \|v\|_{B^\beta_{p_2,q_2}}$$

This follows by Hölder's inequality (both Theorem A.3 and Corollary A.7):

$$\begin{aligned}
\left\| (2^{j(\alpha+\beta)} \|R_j\|_{L^p})_{j\in\mathbb{N}_{-1}} \right\|_{\ell^q} &\le \left\| \left(2^{j(\alpha+\beta)} \Big\| \sum_{i=-1}^{1} \Delta_{j+i} u \Big\|_{L^{p_1}} \|\Delta_j v\|_{L^{p_2}} \right)_{j\in\mathbb{N}_{-1}} \right\|_{\ell^q} \\
&\le \left\| \left(2^{j\alpha} \sum_{i=-1}^{1} \big\|\Delta_{j+i} u\big\|_{L^{p_1}} \right)_{j\in\mathbb{N}_{-1}} \right\|_{\ell^{q_1}} \left\| \left(2^{j\beta} \|\Delta_j v\|_{L^{p_2}} \right)_{j\in\mathbb{N}_{-1}} \right\|_{\ell^{q_2}} \\
&\le (2^{-\alpha} + 1 + 2^{\alpha}) \left\| \left(2^{j\alpha} \big\|\Delta_j u\big\|_{L^{p_1}} \right)_{j\in\mathbb{N}_{-1}} \right\|_{\ell^{q_1}} \|v\|_{B^\beta_{p_2,q_2}} \qquad (154)
\end{aligned}$$

□

The following is basically the same as Corollary 2.1.35 of Jörg Martin's thesis. Such statement that combines the estimates on the paraproducts and resonance products is missing in [2].

Theorem 19.7. *Let $\alpha, \beta \in \mathbb{R} \setminus \{0\}$ be such that $\alpha + \beta > 0$, $\alpha \le \beta$. Let $\delta > 0$. There exists a $C > 0$ such that for all $p, p_1, p_2, q, q_1, q_2 \in [1, \infty]$ such that $\frac{1}{p_1} + \frac{1}{p_2} \le 1$, with*

$$\tfrac{1}{p} = \tfrac{1}{p_1} + \tfrac{1}{p_2}, \qquad \tfrac{1}{q} = \min\{1, \tfrac{1}{q_1} + \tfrac{1}{q_2}\}, \qquad (155)$$

we have

$$\|u \cdot v\|_{B^{\alpha-\delta}_{p,q}} \le C \|u\|_{B^\alpha_{p_1,q_1}} \|v\|_{B^\beta_{p_2,q_2}} \qquad (u, v \in \mathcal{S}'), \qquad (156)$$

and, for all $r \ge q_1$

$$\|u \cdot v\|_{B^\alpha_{p,r}} \le C \|u\|_{B^\alpha_{p_1,q_1}} \|v\|_{B^\beta_{p_2,q_2}}. \qquad (u, v \in \mathcal{S}'). \qquad (157)$$

Proof. Let $\delta > 0$ and $r \geq q_1$. Observe that $\beta > 0$ by assumption. By Theorem 19.6 and Theorem 19.4 there exists a $C_1 > 0$, by Theorem 16.5 there exists a $C_2 > 0$ and by Lemma 16.1 there exists a $C_3 > 0$ such that for all $u, v \in \mathcal{S}'$

$$\|u \odot v\|_{B^{\alpha+\beta}_{p,q}} \leq C_1 \|u\|_{B^{\alpha}_{p_1,q_1}} \|v\|_{B^{\beta}_{p_2,q_2}}, \tag{158}$$

$$\|u \varolessthan v\|_{B^{\alpha}_{p,q}} \leq \begin{cases} C_1 \|u\|_{L^{p_1}} \|v\|_{B^{\alpha}_{p_2,q}} \leq C_1 C_2 C_3 \|u\|_{B^{\alpha}_{p_1,q_1}} \|v\|_{B^{\alpha}_{p_2,q_2}} & \alpha > 0, \\ \|u \varolessthan v\|_{B^{\alpha+\beta}_{p,q}} \leq C_1 \|u\|_{B^{\alpha}_{p_1,q_1}} \|v\|_{B^{\beta}_{p_2,q_2}} & \alpha < 0, \end{cases} \tag{159}$$

$$\|u \varogreaterthan v\|_{B^{\alpha-\delta}_{p,q}} = \|v \varolessthan u\|_{B^{\alpha-\delta}_{p,q}} \leq C_1 \|v\|_{B^{-\delta}_{p_2,q_2}} \|u\|_{B^{\alpha}_{p_1,q_1}} \leq C_1 \|v\|_{B^{\beta}_{p_2,q_2}} \|u\|_{B^{\alpha}_{p_1,q_1}}, \tag{160}$$

$$\|u \varogreaterthan v\|_{B^{\alpha}_{p,r}} \leq C_1 \|v\|_{L^{p_2}} \|u\|_{B^{\alpha}_{p_1,r}} \leq C_1 C_2 C_3 \|v\|_{B^{\beta}_{p_2,q_2}} \|u\|_{B^{\alpha}_{p_1,q_1}}, \tag{161}$$

(156) follows then from (158), (159) and (160). For $r \geq q_1$ we also have $r \geq q$ and thus $\| \cdot \|_{B^{\alpha+\beta}_{p,r}} \leq C_3 \| \cdot \|_{B^{\alpha+\beta}_{p,q}}$, so that by combining (158), (159) and (161) we obtain (157). □

19.8 (The notation $\lesssim$). As we have seen in a couple of proofs, keeping track of which constant comes from which statement can became quite administrative. The benefit is that one actually sees where things come from and on which parameters they depend. However, when one reaches a higher number of constants, say 10, at the moment that C_{10} appears one can probably not tell the different constants apart any ways. For this reason often the notation "$\lesssim$" is used. The usage is as follows: For families $(a_i)_{i \in \mathbb{I}}, (b_i)_{i \in \mathbb{I}}$ in $\mathbb{R}$ for an index set $\mathbb{I}$, one write $a_i \lesssim b_i$ to denote the existence of a $C > 0$ such that $a_i \leq C b_i$ for all $i \in \mathbb{I}$. With this notation (161) could instead read

$$\|u \varogreaterthan v\|_{B^{\alpha}_{p,r}} \lesssim \|v\|_{L^{p_2}} \|u\|_{B^{\alpha}_{p_1,r}} \lesssim \|v\|_{B^{\beta}_{p_2,q_2}} \|u\|_{B^{\alpha}_{p_1,q_1}}.$$

Now, we however do have multiple parameters on the left and right-hand side: α, β and p, r, p_1, q_2, q_1. But the C_1, C_2 and C_3 introduced in that proof depend on α and β, which one now does not see in the notation. One way to overcome this is to write "$\lesssim_{\alpha,\beta}$". For families $(a_{i,\alpha})_{i \in \mathbb{I}, \alpha \in \mathbb{A}}, (b_{i,\alpha})_{i \in \mathbb{I}, \alpha \in \mathbb{A}}$ we write $a_{i,\alpha} \lesssim_{\alpha} b_{i,\alpha}$ to denote that for all $\alpha \in \mathbb{A}$ there exists a $C > 0$ such that $a_{i,\alpha} \leq C b_{i,\alpha}$ for all $i \in \mathbb{I}$. But of course it might be that there are many parameters that change the C, or in other words, which the C depends on. Another way how some authors overcome this is to write "$a_{i,\alpha} \lesssim b_{i,\alpha}$ uniformly in $i \in I$", which means the same as "there exists a $C > 0$ such that $a_{i,\alpha} \leq C b_{i,\alpha}$ for all $i \in \mathbb{I}$".

If we will use the notation "$\lesssim$", then it will be in proofs, without the dependence on the parameters. From the statement in the theorem or lemma (of the order of the "for all" and "there exists") it will be clear on which parameters the constant depends and on which not (as we fix the dimension for example, basically one should a priori assume that the constant depends on the dimension).

We turn now to a specific case of products between Besov spaces with $p = q = \infty$.

19.9 (Notation). In 17.8 we have mentioned that the Zygmund space (Definition 17.4) $\mathcal{C}^s$ for $s > 0$ equals $B^s_{\infty,\infty}$. It is common, and therefore we will follow this convention, to write $\mathcal{C}^s$ for $B^s_{\infty,\infty}$ for all $s \in R$. As the norms $\| \cdot \|_{\mathcal{C}^s}$ and $\| \cdot \|_{B^s_{\infty,\infty}}$ are equivalent for $s > 0$, and as we are interested in bounds of norms, as is mentioned in 15.11 for the dependence of the Besov norm on the choice of partition, we do not distinguish between the norms as we consider statements about estimates on those norms: In other words, when reading "$\| \cdot \|_{\mathcal{C}^s}$" one may as well read "$\| \cdot \|_{B^s_{\infty,\infty}}$".

19.10. By Leibniz rule (see 5.3) we have seen in (30) that the product of two C^k functions is again in C^k and that the product map

$$C^k \times C^k \to C^k, \quad (f,g) \mapsto fg,$$

is a bilinear continuous map. It follows that if $k, m \in \mathbb{N}_0$ that the product of a C^k function and a C^m function is a $C^{k\wedge m}$ function and

$$C^k \times C^m \to C^{k\wedge m}, \quad (f,g) \mapsto fg,$$

is a bilinear continuous map with

$$\|fg\|_{C^{k\wedge m}} \le \|f\|_{C^k} \|g\|_{C^m} \qquad (f \in C^k, g \in C^m).$$

The following theorem states something similar for the $\mathcal{C}^\alpha$ spaces.

By taking $p = p_1 = p_2 = q_1 = q_2 = r = \infty$ in (157) we obtain the following consequence of Theorem 19.7, which is widely used in the theory of SPDEs. See for example [14, Proposition 4.14] and [13, Lemma 2.1 and text below].

Corollary 19.11. *Let $\alpha, \beta \in \mathbb{R}$ and $\alpha + \beta > 0$. If $u \in \mathcal{C}^\alpha$ and $v \in \mathcal{C}^\beta$, then $uv = u \cdot v = u \varolessthan v + u \odot v + u \varogreaterthan v$ is an element of $\mathcal{C}^{\alpha\wedge\beta}$. Moreover, the map*

$$\mathcal{C}^\alpha \times \mathcal{C}^\beta \to \mathcal{C}^{\alpha\wedge\beta}, \quad (u,v) \mapsto uv,$$

is a bilinear continuous map and there exists a $C > 0$ such that

$$\|u \cdot v\|_{\mathcal{C}^{\alpha\wedge\beta}} \le C \|u\|_{\mathcal{C}^\alpha} \|v\|_{\mathcal{C}^\beta} \qquad (u, v \in \mathcal{S}'). \tag{162}$$

19.12. The product map in Corollary 19.11 agrees for $\alpha, \beta \in (0,\infty) \setminus \mathbb{N}$ with the map

$$C^\alpha \times C^\beta \mapsto C^{\alpha\wedge\beta}, \quad (f,g) \mapsto fg,$$

as for those α and β we have $C^\alpha = \mathcal{C}^\alpha$ (see 17.8).

The following corollary is another consequence of Theorem 19.4 and Theorem 19.6 and is left as an exercise:

Corollary 19.13. *Let $\alpha \in (0,\infty)$. There exists a $C > 0$ such that for all $p, q \in [1,\infty]$*

$$\|uv\|_{B^\alpha_{p,q}} \le C \left(\|u\|_{B^\alpha_{p,q}} \|v\|_{L^\infty} + \|u\|_{L^\infty} \|v\|_{B^\alpha_{p,q}} \right) \qquad (u, v \in \mathcal{S}').$$

Consequently, $L^\infty \cap B^\alpha_{p,q}$ is a Banach algebra under the norm $(C \vee 1)(\| \cdot \|_{L^\infty} + \| \cdot \|_{B^\alpha_{p,q}})$.

$\boxed{\textbf{Exercise}}$ **19.1.** Prove Corollary 19.13.

Another consequence is the following:

Theorem 19.14. *Let $\alpha, \beta \in \mathbb{R} \setminus \mathbb{N}$ and $\alpha + \beta > 0$. For $u \in H^\alpha$ and $v \in H^\beta$ we have $uv \in W^{\alpha+\beta,1}$. Moreover, the product map $H^\alpha \times H^\beta \to W^{\alpha+\beta,1}$, $(u,v) \mapsto uv$ is bilinear and continuous, moreover there exists a $C > 0$ such that*

$$\|uv\|_{W^{\alpha\wedge\beta,1}} \le C\|u\|_{H^\alpha}\|v\|_{H^\beta}.$$

$\boxed{\textbf{Exercise}}$ **19.2.** Prove Theorem 19.14. (Hint: 17.8)

Remark 19.15. In [25, Theorem 4.3.1] one can find that certain conditions are necessary for such product embeddings.

$\boxed{\textbf{Question}}$ **19.3.** In [26, Theorem 4.3.6] is stated a similar estimate on the product as in Theorem 19.7. How do they relate to each other, is the one a consequence of the other or do the describe different cases?

20 The heat kernel and heat equation

In Example 7.15 we have seen that $E : \mathbb{R}^{d+1} \to \mathbb{R}$ given by

$$E(t,x) = \begin{cases} h_t(x) & (t,x) \in (0,\infty) \times \mathbb{R}^d, \\ 0 & (t,x) \in (-\infty, 0] \times \mathbb{R}^d, \end{cases}$$

where

$$h_t(x) = (4\pi t)^{-\frac{d}{2}} e^{-\frac{1}{4t}|x|^2} \qquad ((t,x) \in (0,\infty) \times \mathbb{R}^d), \tag{163}$$

is the fundamental solution of the partial differential operator $\partial_t - \Delta_x$, also called *heat operator*. h_t is also called the *heat kernel* (at time t).

In this section we consider the heat equation and solutions described by the heat kernel. We write "Δ" for "Δ_x".

20.1 (Heat equation with initial condition). The following equation is called the *heat equation with initial condition* f (which is also called "heat equation")

$$\begin{cases} \partial_t u = \Delta u & \text{on } (0,\infty) \times \mathbb{R}^d, \\ u(0,\cdot) = f & \text{on } \mathbb{R}^d, \end{cases} \tag{164}$$

where $f \in \mathcal{S}'$. We have already seen that the function $(0,\infty) \times \mathbb{R}^d \to \mathbb{R}$, $(t,x) \mapsto h_t(x)$ satisfies the heat equation $(\partial_t - \Delta)h_t(x) = 0$ on $(0,\infty) \times \mathbb{R}^d$. In Lemma 20.3 we will show that $h_t * \varphi \to \varphi$ in $\mathcal{S}$ as $t \downarrow 0$ for $\varphi \in \mathcal{S}$. Because of these facts, u defined by

$$\begin{cases} u(t,x) = h_t * f(x) & (t > 0), \\ u(0,\cdot) = f, \end{cases} \tag{165}$$

is a solution to the heat equation with initial condition f such that $t \mapsto u(t,\cdot)$ is in $C([0,\infty),\mathcal{S}')$. Observe that u is smooth on $(0,\infty)\times\mathbb{R}^d$. If $f \in C_{\mathrm{b}}$, then u is also continuous as a function on $[0,\infty)\times\mathbb{R}^d$ by (40) and therefore u is a classical solution to this partial differential equation.

Observe that if $f=\delta_0$, then $u(t,x)=h_t(x)$ for $(t,x)\in(0,\infty)\times\mathbb{R}^d$. For this reason the heat kernel is also called the *fundamental solution to the heat equation.* Moreover, for $t>0$, as $h_t(x)=E(\cdot,x)*\delta_0(t)$ (for which we could write formally $\int_{\mathbb{R}} E(s,x)\delta_0(t-s)\,\mathrm{d}s$), we see that $h_t * f(x) = E*(\delta_0\times f)(t,x)$, where $\delta_0\times f$ is the distribution on $\mathbb{R}\times\mathbb{R}^d$ given by

$$\langle \delta_0\times f,\varphi\rangle = \langle f,\varphi(0,\cdot)\rangle \qquad (\varphi\in\mathcal{S}(\mathbb{R}\times\mathbb{R}^d)).$$

$\boxed{\textbf{Exercise}}$ **20.1.** Verify that u defined in (165) for $f\in C_{\mathrm{b}}$ solves the heat equation and is smooth on $(0,\infty)\times\mathbb{R}^d$ and continuous on $[0,\infty)\times\mathbb{R}^d$. (Hint: This follows from a couple of results from Section 7)

$\boxed{\textbf{Question}}$ **20.1.** It kind of would make sense to let $h_0=\delta_0$, also because of Lemma 20.2. Then we would have that $h(t,x)=h_t*f$ not only for $t>0$ but also for $t=0$. However, $E(0,\cdot)=0$. What would happen if we instead take $E'(t,\cdot)=E(t,\cdot)$ for $t\neq 0$ and $E'(0,\cdot)=\delta_0$?

For the proof of Lemma 20.3 we will use the following fact:

Lemma 20.2. *For all $k\in\mathbb{N}_0$ and $t>0$, $y\mapsto |y|^k h_t(y)$ is integrable.*

$\boxed{\textbf{Exercise}}$ **20.2.** Prove Lemma 20.2.

Lemma 20.3. *Let $\varphi\in\mathcal{S}$. Then $h_t*\varphi\to\varphi$ in $\mathcal{S}$ as $t\downarrow 0$.*

Proof. As derivatives of Schwartz functions are Schwartz functions, it suffices to show that for all $k\in\mathbb{N}_0$ and $\varphi\in\mathcal{S}$

$$\sup_{x\in\mathbb{R}^d}(1+|x|)^k|h_t*\varphi(x)-\varphi(x)| \xrightarrow{t\downarrow 0} 0.$$

Let $k\in\mathbb{N}_0$ and $\varphi\in\mathcal{S}$. For all $x\in\mathbb{R}^d$ we have

$$(1+|x|)^k|h_t*\varphi(x)-\varphi(x)| \le (1+|x|)^k\int_{\mathbb{R}^d} h_t(y)|\varphi(x-y)-\varphi(x)|\,\mathrm{d}y.$$

By (40) one can easily show (Exercise 20.1) that we supremum over x in $B(0,1)$ converges to zero. Therefore we may consider the complement, the x with $|x|>1$. We split the integral into two parts, the integral over $B(0,\frac{x}{2})$ and over its complement. Observe that for $y\in B(x,\frac{x}{2})$ we have $|y|\ge\frac12|x|$. Therefore, by Taylor's theorem (Theorem C.7) we have

$$\begin{aligned}\sup_{y\in B(0,\frac{x}{2})}|\varphi(x-y)-\varphi(x)| &\le \sup_{z\in B(x,\frac{x}{2})}\max_{\alpha\in\mathbb{N}_0^d:|\alpha|=1}|\partial^\alpha\varphi(z)||y|\\ &\le \|\varphi\|_{k+1,\mathcal{S}}(1+\frac12|x|)^{-k}|y|.\end{aligned}$$

As $x \mapsto (1+|x|)^k(1+\frac{1}{2}|x|)^{-k}$ is bounded, we see there exists a $C > 0$ such that

$$\int_{B(0,\frac{x}{2})} h_t(y)|\varphi(x-y)-\varphi(x)| \, \mathrm{d}y \le C \int_{\mathbb{R}^d} |y| h_t(y) \, \mathrm{d}y = C\sqrt{t} \int_{\mathbb{R}^d} |y| h_1(y) \, \mathrm{d}y,$$

which converges to zero as $t \downarrow 0$ (because is $y \mapsto |y|h_1(y)$ integrable by Lemma 20.2). On the complement we bound $|\varphi(x-y)-\varphi(x)|$ by M which is two times the supremum norm of φ and obtain

$$\int_{\mathbb{R}^d \setminus B(0,\frac{x}{2})} h_t(y)|\varphi(x-y)-\varphi(x)| \, \mathrm{d}y \le M \left|\frac{x}{2}\right|^{-k} \int_{\mathbb{R}^d} |y|^k h_t(y) \, \mathrm{d}y.$$

As $x \mapsto (1+|x|)^k|\frac{x}{2}|^{-k}$ is bounded uniformly in x for $|x| \ge 1$ and $\int_{\mathbb{R}^d} |y|^k h_t(y) \, \mathrm{d}y = \sqrt{t}^k \int_{\mathbb{R}^d} |y|^k h_1(y) \, \mathrm{d}y \xrightarrow{t\downarrow 0} 0$ (see Lemma 20.2) we conclude that $h_t * \varphi \to \varphi$ in $\mathcal{S}$ as $t \downarrow 0$.□

Remark 20.4 (Stochastic analogue). Another common definition of a heat kernel is

$$p_t(x) = (2\pi t)^{-\frac{d}{2}} e^{-\frac{1}{2t}|x|^2} \qquad ((t,x) \in (0,\infty) \times \mathbb{R}^d),$$

so that $p_t = h_{\frac{t}{2}}$. This is the common choice in stochastics, as this is also the density of a normal distributed random variable with variance t. In other words, it is the density of B_t, where $(B_t)_{t\ge 0}$ is a Brownian motion. So instead of our language, we could have used "p_t" instead of "h_t" if we also used "$\frac{1}{2}\Delta$" instead of "Δ". This also relates to the fact that $\frac{1}{2}\Delta$ is the generator of the Brownian motion. In this way, one can also represent the solution to the heat equation in the following stochastic way. If $\mathbb{E}_x$ is the expectation corresponding to the probability space in which $(B_t)_{t\ge 0}$ is a Brownian motion with $B_0 = x$, then $u(t,x) = \mathbb{E}_x[f(B_t)]$.

This is a special case of the Feynman–Kac formula, which describes solutions of parabolic partial differential equations in terms of diffusion process, which in turn satisfies a stochastic differential equation. For more details see for example [11, Section 6.5].

Exercise 20.3. Let $\lambda > 0$. Show that if u is given by (165) that the function v given by $v(t,x) := u(\lambda t, x)$ for $(t,x) \in (0,\infty) \times \mathbb{R}^d$ satisfies $\partial_t v = \lambda \Delta u$. From this one can verify the statement in Remark 20.4 about replacing "h_t" by "p_t" and "Δ" by "$\frac{1}{2}\Delta$".

20.5. For $t \ge 0$ we will now write $H_t : \mathcal{S}' \to \mathcal{S}'$ for the function given by

$$H_t f = \begin{cases} h_t * f & t > 0, \\ f & t = 0. \end{cases}$$

For $t > 0$, $H_t f \in C_{\mathrm{p}}^\infty$. For $t > 0$ we will consider in which Besov space the function $H_t f$ lies, when f is in the Besov space $B_{p,q}^s$.

H_t is a Fourier multiplier, by Theorem 11.15. Indeed, we have

$$h_t = \mathcal{F}^{-1}(g_t) = \hat{g}_t,$$

for

$$g_t(x) = e^{-4\pi^2 t|x|^2} \qquad (x \in \mathbb{R}^d),$$

and therefore for $g = g_1$

$$H_t = g_t(\mathrm{D}) = (l_{\sqrt{t}} g)(\mathrm{D}).$$

Therefore, we can apply Theorem 18.2. For $i \in \{1, \dots, d\}$ we have that $\partial_i g(x) = -8\pi^2 t x_i g(x)$. Therefore, inductively we obtain for $\alpha \in \mathbb{N}_0^d$ that $\partial^\alpha g = Pg$, where P is a polynomial of order $|\alpha|$, hence bounded in absolute value by a multiple of $(1+|x|)^{|\alpha|}$. Moreover, as $(1+|x|)^m g$ is bounded for all $m \in \mathbb{R}$, we have we obtain the following.

Theorem 20.6. *Let $s, m \in \mathbb{R}$, $p, q \in [1, \infty]$. There exists a $C > 0$ such that*

$$\|H_t f\|_{B^{s+m}_{p,q}} \le C(t^{-\frac{m}{2}} \vee 1)\|f\|_{B^s_{p,q}} \qquad (f \in B^s_{p,q}, t > 0).$$

So we have that $H_t f \in B^s_{p,q}$ for all $s \in \mathbb{R}$ if $f \in B^\alpha_{p,q}$ for some $\alpha \in \mathbb{R}$. A similar statement holds for Sobolev spaces:

$\boxed{\textbf{Exercise}}$ **20.4.** Let $p \in [1, \infty]$, $f \in L^p$ and $t > 0$. Show that $H_t f \in L^p$. Moreover, show that $H_t f \in W^{k,p}$ for all $k \in \mathbb{N}_0$.

20.7. For a normed space $\mathfrak{X}$, the space of continuous functions on $[0, T]$ with values in $\mathfrak{X}$, for which we write $C([0,T], \mathfrak{X})$, is equipped with the supremum norm

$$\|u\|_{C([0,T],\mathfrak{X})} = \sup_{s \in [0,T]} \|u(t)\|_{\mathfrak{X}} \qquad (u \in (\mathcal{S}')^{[0,T]}).$$

Equipped with this norm, $C([0,T], \mathfrak{X})$ is a Banach space.

20.8. If the initial condition f to the heat equation (164) is in a Besov space $B^\alpha_{p,q}$ for some $\alpha \in \mathbb{R}$, $p, q \in [1, \infty]$, then by Theorem 20.6 the solution u as in (165) satisfies

$$\|u\|_{C([0,T], B^\alpha_{p,q})} < \infty \qquad (T > 0).$$

Actually u is in $C([0,T], B^{\alpha-\varepsilon}_{p,q})$ for all $\varepsilon > 0$ (see for example Lemma 4.5 in the lecture notes of N. Perkowski on SPDEs, this considers $p = q = \infty$). Moreover, we have for $m \in \mathbb{R}$

$$\sup_{t \in (0,T]} t^{\frac{m}{2}} \|u(t)\|_{B^{\alpha+m}_{p,q}} < \infty. \tag{166}$$

20.9 (The heat equation with additive noise). Let ξ be a tempered distribution on $\mathbb{R} \times \mathbb{R}^d$ such that $\operatorname{supp} \xi \subset [0, \infty) \times \mathbb{R}^d$ and for all mollifier functions $\psi \in \mathcal{D}(\mathbb{R} \times \mathbb{R}^d)$ and $x \in \mathbb{R}^d$, $\langle \xi, \mathcal{T}_{(0,x)} \psi_\varepsilon \rangle \xrightarrow{\varepsilon \downarrow 0} 0$. In this way we regard ξ as a tempered distribution on $(0, \infty) \times \mathbb{R}^d$, which "is zero" on $\{0\} \times \mathbb{R}^d$. Let f be a tempered distribution on $\mathbb{R}^d$. Let us regard ξ as a noise term, often also called potential. We consider the *heat equation with additive potential/noise*:

$$\begin{cases} \partial_t u = \Delta u + \xi & \text{on } (0,\infty) \times \mathbb{R}^d, \\ u(0,\cdot) = f & \text{on } \mathbb{R}^d. \end{cases} \tag{167}$$

A solution is given by (at least if $E * \xi$ exists and is a function)

$$E * (\xi + \delta_0 \times f)(t,x) \qquad ((t,x) \in (0,\infty) \times \mathbb{R}^d),$$

which we can also view (at least formally, see Question 20.2) as a solution to $(\partial_t - \Delta)u = \xi + \delta_0 \times f$. Moreover, we can interpret convoluting with E to be the 'inverse' of $\partial_t - \Delta$.

A ξ as above is also called a *space–time noise*. If it is independent of its time variable, by which we mean that $\mathcal{T}_{(t,0)}\xi = \xi$, then we call ξ a *space noise*. If the last is the case, then $u : [0,\infty) \times \mathcal{S}'$ given by $u(0,\cdot) = f$ and

$$u(t,x) = \int_0^t H_s(\xi + \delta_0 \times f)(x) \,\mathrm{d}s = \int_0^t H_s \xi(x) \,\mathrm{d}s + H_t * f(x) \quad ((t,x) \in (0,\infty) \times \mathbb{R}^d),$$

is a solution that is smooth on $(0,\infty) \times \mathbb{R}^d$. Such u is also called a *mild solution* to (167).

Question **20.2.** How can we interpret $E * (\xi + \delta_0 \times f)(0,\cdot)$? Do we have $E * (\xi + \delta_0 \times f)(t,\cdot) \to f$ in $\mathcal{S}'$ as $t \downarrow 0$? Depending on the answers, can we interpret $E * (\xi + \delta_0 \times f)$ as a continuous function on $[0,\infty)$ with values in $\mathcal{S}'$? Is the condition on ξ the 'right' one, that it "is zero" on $\{0\} \times \mathbb{R}^d$.

20.10. In the course on SPDEs by N. Perkowski it is shown (in Theorem 4.6) that if $f \in \mathcal{C}^{\alpha+2}$ and $\xi \in C([0,T],\mathcal{C}^\alpha)$, then

$$\|u\|_{C([0,T],\mathcal{C}^{\alpha+2})} < \infty.$$

Moreover, by replacing $\mathcal{C}^\alpha$ by the closure of $\mathcal{D}$ in $\mathcal{C}^\alpha$ (which is the interpretation as in the lecture notes by N. Perkowski), lets call it $\mathcal{C}^\alpha_*$, one also has continuity: $u \in C([0,T],\mathcal{C}^{\alpha+2}_*)$.

20.11 (The heat equation with multiplicative noise). Let ξ and f be as in 20.9. We consider the *heat equation with multiplicative potential/noise*:

$$\begin{cases} \partial_t u = \Delta u + \xi u & \text{on } (0,\infty) \times \mathbb{R}^d, \\ u(0,\cdot) = f & \text{on } \mathbb{R}^d. \end{cases} \tag{168}$$

Here we are possibly dealing with a product that is ill–posed, as a priori u is a distribution and we cannot multiply two distributions in general. We can find a solution by finding a u that satisfies

$$u(t,x) = \int_0^t H_s(\xi u)(x) \,\mathrm{d}s + H_t * f(x) \quad ((t,x) \in (0,\infty) \times \mathbb{R}^d).$$

As for the heat equation with additive noise, such u is also called a *mild solution* to (168).

Such mild solution can be derived via a fixed point argument (by taking the right–hand side as the outcome of a map Φ acting on u and showing that this map has a fixed point). We consider a different equation which has a similar flavour as the heat equation in 20.14.

$\boxed{\textbf{Question}}$ **20.3.** What about uniqueness of distributional solutions? About uniqueness of mild solutions?

Remark 20.12 (The stochastic heat equation). In stochastics, ξ is often regarded as a random variable and (167) and (168) are called the *stochastic heat equation with additive noise* and *stochastic heat equation with multiplicative noise*, respectively. Also often "stochastic heat equation" is abbreviated by "SHE". The stochastic heat equation with multiplicative noise is also called the *parabolic Anderson model.* The interpretation is as follows. ξ being a random variable means it is a measurable map $\Omega \to \mathcal{S}'$, where Ω is the underlying space of a probability space. A (distributional) solution to for example (167) is then also a random map $\Omega \to (\mathcal{S}')^{[0,\infty)}$. For example one could say u is *almost surely a solution* to (167). This means that when we write "u_ω" for "$u(\omega)$", that for almost all ω (that means for all omega in a set of probability one), u_ω is the solution to

$$\begin{cases} \partial_t u_\omega = \Delta u_\omega + \xi_\omega & \text{on } (0,\infty) \times \mathbb{R}^d, \\ u_\omega(0,\cdot) = f_\omega & \text{on } \mathbb{R}^d, \end{cases}$$

where we have also written "ξ_ω" for "$\xi(\omega)$" and also allowed the initial condition f to be random.

In general, ξ is assumed to be white noise, either space–time white noise or space white noise. White noise is a totally uncorrelated noise, which informally means that the outcome of it at some point in space (and time) is independent from the outcome of a different point in space (and time). It can be shown that space–time white noise is almost surely (which means for almost all realisation) an element of $B^{-1-\frac{d}{2}-\varepsilon}_{\infty,\infty,w}$ and space white noise is almost surely in $B^{-\frac{d}{2}-\varepsilon}_{\infty,\infty,w}$ for all $\varepsilon > 0$; where the w is a "weight" and $B^s_{p,q,w}$ is a "weighted Besov space".

Instead of considering the heat equation with multiplicative noise, let us consider a different type of equation which we solve by a fixed point argument. First we recall Banach's fixed point theorem.

Theorem 20.13 (Banach's fixed point theorem). *Let $(\mathfrak{X}, d)$ be a complete metric space. Suppose that $\Phi : \mathfrak{X} \to \mathfrak{X}$ is a contraction, i.e., there exists a $c \in (0,1)$ such that*

$$d(\Phi(x), \Phi(y)) \le c d(x,y).$$

Then there exists a unique point x_* in $\mathfrak{X}$ such that $\Phi(x_*) = x_*$. Moreover, by defining $\Phi^1 = \Phi$ and $\Phi^k = \Phi \circ \Phi^{k-1}$ for $k \in \mathbb{N}$ with $k \ge 2$, we have for each $x \in \mathfrak{X}$ that

$$\lim_{k\to\infty} \Phi^k(x) = x_*.$$

20.14. Let $\xi, \psi \in \mathcal{S}'$ and consider the following partial differential equation

$$(1-\Delta)u = u\xi + \psi.$$

We define (at least formally)

$$\Phi(u) = (1-\Delta)^{-1}(u\xi + \psi),$$

and want to show that Φ has a fixed point under some conditions on ξ and ψ.

First of all, let us observe that by Theorem 18.2, see also Example 18.3, for all $p, q \in [1, \infty]$ and $\alpha \in \mathbb{R}$ there exists a $C_1 > 0$ such that

$$\|(1-\Delta)^{-1}w\|_{B^{\alpha}_{p,q}} \leq C_1 \|w\|_{B^{\alpha-2}_{p,q}} \qquad (w \in \mathcal{S}').$$

In order to have our Φ mapping $B^{\alpha}_{p,q}$ into itself, let us consider $\alpha = 1$, $p = q = 2$, $\psi \in H^{-1}$ and $\xi \in \mathcal{C}^{-1+\delta}$ for some $\delta > 0$ (where $H^s = B^s_{2,2}$ and $\mathcal{C}^s = B^s_{\infty,\infty}$). By Theorem 19.7there exists a $C_2 > 0$ such that

$$\|\xi u\|_{H^{-1}} \leq C_2 \|u\|_{H^1} \|\xi\|_{\mathcal{C}^{-1+\delta}}.$$

Therefore we have

$$\begin{aligned}\|\Phi(u)\|_{H^1} &\leq C_1 \|u\xi + \psi\|_{H^{-1}} \\ &\leq C_1 C_2 \|u\|_{H^1} \|\xi\|_{\mathcal{C}^{-1+\delta}} + C_2 \|\psi\|_{H^{-1}} \qquad (u \in H^1),\end{aligned}$$

which shows that Φ maps H^1 into itself, and

$$\|\Phi(u) - \Phi(v)\|_{H^1} \leq C_1 C_2 \|u - v\|_{H^1} \|\xi\|_{\mathcal{C}^{-1+\delta}} \qquad (u, v \in H^1).$$

Therefore, if $\|\xi\|_{\mathcal{C}^{-1+\delta}} < (C_1 C_2)^{-1+\delta}$, Φ is a contraction map on H^1 and hence possesses a unique fixed point. Alternatively, for each $\xi \in \mathcal{C}^{-1+\delta}$ and $\psi \in H^{-1}$ there exists an $a \in (0, \infty)$ such that the equation

$$(1-\Delta)u = au\xi + \psi$$

has a unique solution.

$\boxed{\textbf{Exercise}}$ **20.5.** We have shown that Φ is a contraction map on $B^{\alpha}_{p,q}$ in case

$$\xi \in B^{\beta}_{p_1,q_1}, \qquad \psi \in B^{\gamma}_{p_2,q_2},$$

and $\|\xi\|^{\beta}_{B_{p_1,q_1}}$ is small enough, with $p_1 = q_1 = \infty$, $p = q = p_2 = q_2$, $\beta = -1 + \delta$, $\gamma = -1$ and $\alpha = 1$.

Which other choices can be made for those parameters so that one still obtains a contraction map? For example, with $p_1 = q_1 = \infty$, $p = q = p_2 = q_2$ consider for which other α, β, γ the map is a contraction. And, for $p = p_1 = p_2 = q = q_1 = q_2 = \infty$, do we need $\delta > 0$ or does $\delta = 0$ suffices?

What can you say if instead one considers

$$(1-\Delta)^s u = u\xi + \psi,$$

for $s \in \mathbb{R}$?

A Preliminaries on L^p spaces

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $\mathbb{F}$ be either $\mathbb{R}$ or $\mathbb{C}$.

Definition A.1. We say that a subset A of X is an (μ-)*null set*, if there exists a $B \in \mathcal{A}$ with $A \subset B$ and $\mu(B) = 0$. We write A^c for the complement of A in X, so that $A^c = X \backslash A$.

Definition A.2. Let $p \in [1, \infty)$. $\mathcal{L}^p(\mu)$ is the space of measurable functions $f : X \to \mathbb{F}$ for which

$$\int |f(x)|^p \, \mathrm{d}\mu(x) < \infty.$$

We say that two measurable functions f and g are equivalent, written $f \sim g$ if there exists a null set $A \in \mathcal{A}$ such that $f = g$ on A^c. We write $L^p(\mu)$ for the space that consists of all equivalence classes in $\mathcal{L}^p(\mu)$, in formula $L^p(\mu) = \mathcal{L}^p(\mu)/\sim$ or when we define $[f]_\sim = \{g \in \mathcal{L}^p : g \sim f\}$ for $f \in \mathcal{L}^p$, then

$$L^p(\mu) = \{[f]_\sim : f \in \mathcal{L}^p(\mu)\}.$$

We define

$$\|f\|_{L^p} := \left(\int |f(x)|^p \, \mathrm{d}\mu \right)^{\frac{1}{p}}.$$

Similarly, we define $\mathcal{L}^\infty$ to be the space of measurable functions $f : X \to \mathbb{F}$ for which there exist a null set A such that f is bounded on A^c. In other words, those functions that are almost everywhere (abbreviated "a.e.") bounded. We define

$$\|f\|_{\mathcal{L}^\infty} = \inf\{M > 0 : |f| \le M \text{ a.e. }\}.$$

Similarly as for $p \in [1, \infty)$, we define

$$L^\infty(\mu) = \{[f]_\sim : f \in \mathcal{L}^\infty(\mu)\},$$

and write for $f \in L^\infty(\mu)$ and $g \in f$ (the following is independent of the choice of g)

$$\|f\|_{L^\infty} = \|g\|_{\mathcal{L}^\infty}.$$

We say that a function $f : \Omega \to \mathbb{F}$ is *locally integrable* if $f \mathbb{1}_K$ is an integrable function for all compact sets $K \subset \Omega$. We write $\mathcal{L}^1_{\text{loc}}(\Omega)$ for the space of all locally integrable functions and $L^1_{\text{loc}}(\Omega)$ for the space of their equivalence classes similarly as for $\mathcal{L}^p$ and L^p. Similarly, we write $L^p_{\text{loc}}(\Omega)$ for the space of functions that are locally in $L^p(\Omega)$.

But from now on we 'identify' functions f with their equivalence class $[f]_\sim$, and so use also consider elements of L^p as functions.

Theorem A.3 (Hölder's inequality). *[2, Theorem 1.1] Let $p, q, r \in [1, \infty]$ satisfy*

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{r}.$$

If $f \in L^p(\mu)$ and $g \in L^q(\mu)$, then $fg \in L^r(\mu)$ and

$$\|fg\|_{L^r} \le \|f\|_{L^p} \|g\|_{L^q}.$$

Theorem A.4 (Generalized Hölder inequality). *Let $n \in \mathbb{N}$ and $p_1, \ldots, p_n, r \in [1, \infty]$. Suppose*

$$\frac{1}{p_1} + \cdots + \frac{1}{p_n} = \frac{1}{r}.$$

For $i \in \{1, \ldots, n\}$ let $f_i \in L^{p_i}(\mu)$. Then $f_1 \cdots f_n \in L^r$ and

$$\|f_1 \cdots f_n\|_{L^r} \leq \|f_1\|_{L^{p_1}} \cdots \|f_n\|_{L^{p_n}}.$$

Proof. Let $q \in [1, \infty]$ be such that

$$\frac{1}{q} = \frac{1}{p_1} + \cdots + \frac{1}{p_{n-1}}.$$

Let $g = f_1 \cdots f_{n-1}$. If $g \in L^q$, then by the Hölder inequality, as $\frac{1}{q} + \frac{1}{p_n} = \frac{1}{r}$

$$\|g f_n\|_{L^r} \leq \|g\|_{L^q} \|f_n\|_{L^{p_n}}.$$

From this one can finish the proof by an induction argument. □

Lemma A.5. *We have $L^p(\mu) \subset L^1(\mu) + L^\infty(\mu)$ for all $p \in [1, \infty]$.*

Proof. Let $f \in L^p(\mu)$. Then $[|f| > 1]$ has finite measure. Define $f_1 := f \mathbb{1}_{[|f| \leq 1]}$ and $f_2 := f \mathbb{1}_{[|f|>1]}$. Then $f_1 \in L^\infty(\mu)$ and with Hölder's inequality we have

$$\|f_2\|_{L^1} \leq \|f\|_{L^p} \|\mathbb{1}_{[|f|>1]}\|_{L^q} < \infty,$$

for $q \geq 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$. □

A.6. [5, Exercise 5.17] Let $1 \leq p \leq r \leq \infty$. If $x \in \ell^p$, then $x \in \ell^r$ and $\|x\|_{\ell^r} \leq \|x\|_{\ell^p}$. Moreover, if $x \in \ell^p$, then $x \in c_0$.

Corollary A.7 (Hölder's inequality for ℓ^p spaces). *Let $p, q \in [1, \infty]$ and $r \in [1, \infty]$ be such that*

$$\min\{1, \frac{1}{p} + \frac{1}{q}\} = \frac{1}{r}.$$

If $f \in \ell^p$ and $g \in \ell^q$, then $fg \in \ell^r$ with

$$\|fg\|_{\ell^r} \leq \|f\|_{\ell^p} \|g\|_{\ell^q}.$$

Proof. Suppose that $\frac{1}{p} + \frac{1}{q} > 1$, in the other case we can apply Hölder's inequality immediately. Then both p and q are finite, and we can find $\tilde{p}$, $\tilde{q}$ with $p \leq \tilde{p} < \infty$, $q \leq \tilde{q} < \infty$ such that

$$\frac{1}{\tilde{p}} + \frac{1}{\tilde{q}} = 1,$$

Let $f \in \ell^p$ and $g \in \ell^q$. Then $f \in \ell^{\tilde{p}}$ and $g \in \ell^{\tilde{q}}$ and $\|fg\|_{\ell^1} \leq \|f\|_{\ell^{\tilde{p}}} \|g\|_{\ell^{\tilde{q}}} \leq \|f\|_{\ell^p} \|g\|_{\ell^q}$. □

Theorem A.8 (Log-convexity of L^p norms). *Let p, r be such that $1 \le p < r \le \infty$. Then $L^p(\mu) \cap L^r(\mu) \subset L^q(\mu)$ for all q with $p \le q \le r$ and with $\theta \in [0,1]$ such that*

$$\frac{1}{q} = \frac{\theta}{p} + \frac{1-\theta}{r}.$$

we have

$$\|f\|_{L^q} \le \|f\|_{L^p}^{\theta} \|f\|_{L^r}^{1-\theta} \qquad (f \in L^p \cap L^r).$$

Proof. As $1 = \frac{\theta q}{p} + \frac{(1-\theta)q}{r}$, we obtain by Hölder's inequality,

$$\|f\|_{L^q}^q = \int |f|^{\theta q} |f|^{(1-\theta)q} \le \||f|^{\theta q}\|_{L^{\frac{p}{\theta q}}} \||f|^{(1-\theta)q}\|_{L^{\frac{r}{(1-\theta)q}}} = \|f\|_{L^p}^{\theta q} \|f\|_{L^r}^{(1-\theta)q}.$$

Lemma A.9 (Young's inequality for products). *For $p, q > 0$ with $\frac{1}{p} + \frac{1}{q} = 1$,*

$$ab \le \frac{1}{p} a^p + \frac{1}{q} b^q \qquad (a, b \ge 0).$$

In an other formulation; if $\theta \in [0,1]$ then $a^\theta b^{1-\theta} \le \theta a + (1-\theta) b$ for all $a, b \ge 0$.

Proof. As the exponential function is convex, we have for p, q as above and $a, b \ge 0$,

$$ab = \exp\left(\frac{1}{p}\log a^p + \frac{1}{q}\log b^q\right) \le \frac{1}{p}\exp(\log a^p) + \frac{1}{q}\exp(\log b^q) = \frac{1}{p}a^p + \frac{1}{q}b^q.$$

Corollary A.10. *Let $p, q \in [1, \infty]$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. Then $L^1(\mu) \cap L^\infty(\mu) \subset L^p(\mu)$ and*

$$\|f\|_{L^p} \le \frac{1}{p}\|f\|_{L^1} + \frac{1}{q}\|f\|_{L^\infty} \qquad (f \in L^1 \cap L^\infty).$$

Proof. Note that $\theta = \frac{1}{p}$ is such that $\frac{1}{p} = \frac{\theta}{1} + \frac{(1-\theta)}{\infty}$. Apply Theorem A.8, to obtain $\|f\|_{L^p} \le \|f\|_{L^1}^{\frac{1}{p}} \|f\|_{L^\infty}^{\frac{1}{q}}$. Then apply Lemma A.9. $\square$

A.11 (Notation). Let $d \in \mathbb{N}$. We write $\mathcal{B}(\mathbb{R}^d)$ for the Borel-σ-algebra on $\mathbb{R}^d$. If μ is the Lebesgue measure on the measurable space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ and $p \in [1, \infty]$, then we write $L^p(\mathbb{R}^d)$ instead of $L^p(\mu)$.

Definition A.12. We call a set of the form $\prod_{i=1}^d [a_i, b_i]$, where $a_i, b_i \in \mathbb{R}$ and $a_i < b_i$, a *rectangle* (in $\mathbb{R}^d$).

Lemma A.13. *Let $p \in [1, \infty)$.*

(a) *$C_c(\mathbb{R}^d)$ is dense in $L^p(\mathbb{R}^d)$.*
(b) *$C_c^\infty(\mathbb{R}^d)$ is dense in $L^p(\mathbb{R}^d)$.*
(c) *Let $\mathcal{R}$ denote the set of rectangles in $\mathbb{R}^d$. The linear span of $\{\mathbb{1}_A : A \in \mathcal{R}\}$ is dense in $L^p(\mathbb{R}^d)$.*

Observe that ((b)) follows from ((a)) by the Stone-Weierstrass Theorem.

B Preliminaries on topological spaces (incomplete of course)

Definition B.1 (Neighbourhood of a point). We say that a set S is a *neighbourhood* of a point x, if there exists an open subset $U \subset S$, with $x \in U$.

C Taylor's formula

C.1 For one dimension

Let us first recall the fundamental theorem of calculus.

Theorem C.1. *[22, §15] Let $g : [a, b] \to \mathbb{R}$ be continuous. Then*

$$\frac{d}{dx} \int_a^x g(y) \, \mathrm{d}y = g(x).$$

The following is a direct consequence.

Corollary C.2. *If $f \in C^1[a, b]$, then*

$$f(x) = f(a) + \int_a^x f'(y) \, \mathrm{d}y.$$

C.3. If $f \in C^2$, then we have

$$f'(y) = f'(a) + \int_a^y f''(z) \, \mathrm{d}z,$$

and thus

$$\begin{aligned} f(x) &= f(a) + \int_a^x f'(y) \, \mathrm{d}y \\ &= f(a) + \int_a^x \left(f'(a) + \int_a^y f''(z) \, \mathrm{d}z \right) \mathrm{d}y \\ &= f(a) + (x - a) f'(a) + \int_a^x \int_a^y f''(z) \, \mathrm{d}z \, \mathrm{d}y. \end{aligned}$$

This can be iterated:
For $f \in C^k[a, b]$, we have

$$f(x) = \sum_{i=0}^{k-1} \frac{(x-a)^i}{i!} \mathrm{D}^i f(a) + R^k_{f,a}(x),$$

where by Fubini

$$\begin{aligned} R^k_{f,a}(x) &= \int_a^x \int_a^{y_1} \cdots \int_a^{y_{k-1}} \mathrm{D}^k f(y_k) \, \mathrm{d}y_k \, \mathrm{d}y_{k-1} \cdots \mathrm{d}y_1 \\ &= \int_{[a,b]^k} \mathbb{1}_{\{y: a \le y_k \le y_{k-1} \le \cdots \le y_1 \le x\}}(y) \, \mathrm{D}^k f(y_k) \, \mathrm{d}y. \\ &= \int_a^x \int_{y_k}^x \int_{y_{k-1}}^x \cdots \int_{y_2}^x \mathrm{d}y_1 \, \mathrm{d}y_2 \cdots \mathrm{d}y_{k-1} \, \mathrm{D}^k f(y_k) \, \mathrm{d}y_k \end{aligned}$$

By induction one can easily see that

$$\int_{y_k}^{x} \int_{y_{k-1}}^{x} \cdots \int_{y_2}^{x} \mathrm{d}y_1 \; \mathrm{d}y_2 \cdots \mathrm{d}y_{k-1} = \frac{(x-y_k)^{k-1}}{(k-1)!}.$$

So we have obtained the following.

Theorem C.4. *Let $f \in C^k[a,b]$, then*

$$\begin{aligned} f(x) &= \sum_{i=0}^{k-1} \frac{(x-a)^i}{i!} \mathrm{D}^i f(a) + \int_a^x \frac{(x-y)^{k-1}}{(k-1)!} \mathrm{D}^k f(y) \; \mathrm{d}y \\ &= \sum_{i=0}^{k} \frac{(x-a)^i}{i!} \mathrm{D}^i f(a) + \int_a^x \frac{(x-y)^{k-1}}{(k-1)!} [\mathrm{D}^k f(y) - \mathrm{D}^k f(a)] \; \mathrm{d}y. \end{aligned}$$

Let

$$\begin{aligned} L &= \max_{y \in [a,b]} | \mathrm{D}^k f(y)| \\ M &= \max_{y \in [a,b]} | \mathrm{D}^k f(y) - \mathrm{D}^k f(a)|. \end{aligned}$$

Then

$$\begin{aligned} \left| f(x) - \sum_{i=0}^{k-1} \frac{(x-a)^i}{i!} \mathrm{D}^i f(a) \right| &\le \frac{L}{k!}(x-a)^k, \\ \left| f(x) - \sum_{i=0}^{k} \frac{(x-a)^i}{i!} \mathrm{D}^i f(a) \right| &\le \frac{M}{k!}(x-a)^k. \end{aligned}$$

C.2 Taylor expansion in higher dimensions

Definition C.5. Let $f \in C^k(U, \mathbb{R}^p)$ for $U \subset \mathbb{R}^d$ open. Let $a \in U$. The *Taylor polynomial of order* of order k at the point a, written $T^k_{f,a}$, is given by

$$T^k_{f,a}(x) = \sum_{\alpha : |\alpha| \le k} \frac{1}{\alpha!} \partial^\alpha f(a)(x-a)^\alpha$$

The *remainder of order* k at the point a is given by $R^k_{f,a}(x) = f(x) - T^k_{f,a}(x)$.

Lemma C.6. *[8, Lemma 6.1] Let $f \in C^k(U, \mathbb{R}^d)$. Then for $l \in \{0, 1, \dots, k\}$ and $a, h \in \mathbb{R}^d$ and $t \in \mathbb{R}$ such that $a + th \in U$ we have*

$$\frac{1}{j!} \frac{d^j}{dt^j} f(a+th) = \sum_{\alpha : |\alpha| = j} \frac{h^\alpha}{\alpha!} \partial^\alpha f(a+th)$$

Theorem C.7 (Taylor's Formula). *[8, Theorem 6.2] Let $f \in C^k(U, \mathbb{R}^p)$ for $U \subset \mathbb{R}^d$ being an open ball. Let $a \in U$. For all $l \in \{1, \ldots, k\}$ and $x \in U$*

$$f(x) = T_{f,a}^{l-1}(x) + \sum_{\alpha:|\alpha|=l} \frac{(x-a)^\alpha}{\alpha!} \int_0^1 \frac{(1-s)^{l-1}}{(l-1)!} \partial^\alpha f(a + s(x-a)) \, \mathrm{d}s \tag{169}$$

$$= T_{f,a}^{l}(x) + \sum_{\alpha:|\alpha|=l} \frac{(x-a)^\alpha}{\alpha!} \int_0^1 \frac{(1-s)^{l-1}}{(l-1)!} \left[\partial^\alpha f(a + s(x-a)) - \partial^\alpha f(a)\right] \mathrm{d}s. \tag{170}$$

For $a, x \in U$ let us define

$$R_{f,a}^l(x) = \sum_{\alpha:|\alpha|=l} \frac{(x-a)^\alpha}{\alpha!} \int_0^1 \frac{(1-s)^{l-1}}{(l-1)!} \left(\partial^\alpha f(a + s(x-a)) - \partial^\alpha f(a)\right) \mathrm{d}s. \tag{171}$$

The map $U \times U \to \mathbb{R}$ given by $(a, x) \mapsto R_{f,a}^l(x)$ is in C^{k-l}, and for every compact $K \subset U$ and every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$|R_{f,a}^l(x)| \le \varepsilon |x-a|^l \qquad \text{for } x, a \in K \text{ and } |x-a| < \delta.$$

Moreover, for all $a \in U$ the map $R_{f,a}^l : U \to \mathbb{R}$ is in C^k and $\partial^\alpha R_{f,a}^l(a) = 0$ for all $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \le l$.

Proof. Let g be the one-dimensional function given by $g(t) = f(a + t(x-a))$. Then by Theorem C.4

$$g(t) = \sum_{i=0}^{l-1} \frac{t^i}{i!} \frac{d^i}{dt^i} g(0) + \int_0^t \frac{(t-s)^{l-1}}{(l-1)!} \frac{d^l}{ds^l} g(s) \, \mathrm{d}s.$$

So that with Lemma C.6 one obtains (169) and (170). □

D Multivariate chain rule for mixed derivatives

The following theorem is a special case of the Faà di Bruno formula (in which the codomain of g is allowed to be of higher dimensions than one).

Theorem D.1 (Chain rule for multi-index differentiation). *Let $\alpha \in \mathbb{N}_0^d$ and let $m = |\alpha|$. Assume $m \ge 1$. Let $U \subset \mathbb{R}$ be open, $g : U \to \mathbb{R}$ and $f : \mathbb{R}^d \to U$ both be C^m-functions. Then there exist $c_{m,a} \in \mathbb{R}$ for $a \in (\mathbb{N}_0^d)^k$ such that*

$$\partial^\alpha [g \circ f](x) = \sum_{k=1}^m \mathrm{D}^k g(f(x)) \sum_{\substack{a \in (\mathbb{N}_0^d \setminus \{0\})^k \\ a_1 + \cdots + a_k = \alpha}} c_{m,a} \prod_{i=1}^k \partial^{a_i} f(x)$$

Proof. We give a proof by induction. In case $m = 1$, then the formula follows by the chain rule for one derivative.

Let $m \in \mathbb{N}$ and assume that the formula holds for all $\beta \in \mathbb{N}_0^d$ with $|\beta| = m$. Let $\alpha \in \mathbb{N}_0^d$ be such that $|\alpha| = m+1$. Then we can find a $e \in \mathbb{N}_0^d$ and $|e| = 1$ and $\beta \in \mathbb{N}_0^d$ with $|\beta| = m$ such that $\alpha = e + \beta$.

For $k \in \{1, \dots, m\}$ and $b \in (\mathbb{N}_0^d \setminus \{0\})^k$ with $b_1 + \dots + b_k = \beta$ we have

$$\partial^e \left(\mathrm{D}^k g(f(x)) \right) = \mathrm{D}^{k+1} g(f(x)) \partial^e f(x),$$
$$\partial^e \left(\prod_{i=1}^k \partial^{b_i} f(x) \right) = \sum_{j=1}^k \partial^{b_j + e} f(x) \cdot \prod_{i \in \{1,\dots,k\} \setminus \{j\}} \partial^{b_i} f(x)$$

Hence

$$\begin{aligned}
&\partial^e \partial^\beta [g \circ f](x) \\
&= \sum_{k=1}^m \mathrm{D}^{k+1} g(f(x)) \sum_{\substack{b \in (\mathbb{N}_0^d \setminus \{0\})^k \\ b_1 + \dots + b_k = \beta}} \partial^e f(x) c_{m,b} \prod_{i=1}^k \partial^{b_i} f(x) \\
&\quad + \sum_{k=1}^m \mathrm{D}^k g(f(x)) \sum_{\substack{b \in (\mathbb{N}_0^d \setminus \{0\})^k \\ b_1 + \dots + b_k = \beta}} c_{m,b} \sum_{j=1}^k \partial^{b_j + e} f(x) \cdot \prod_{i \in \{1,\dots,k\} \setminus \{j\}} \partial^{b_i} f(x) \\
&= \mathrm{D}^{m+1} g(f(x)) \sum_{\substack{b \in (\mathbb{N}_0^d \setminus \{0\})^m \\ b_1 + \dots + b_m = \beta}} c_{m,b} \partial^e f(x) \prod_{i=1}^m \partial^{b_i} f(x) \\
&\quad + \sum_{k=2}^{m-1} \mathrm{D}^k g(f(x)) \Bigg[\sum_{\substack{b \in (\mathbb{N}_0^d \setminus \{0\})^{k-1} \\ b_1 + \dots + b_{k-1} = \beta}} c_{m,b} \partial^e f(x) \cdot \prod_{i=1}^{k-1} \partial^{b_i} f(x) \\
&\qquad + \sum_{\substack{b \in (\mathbb{N}_0^d \setminus \{0\})^k \\ b_1 + \dots + b_k = \beta}} c_{m,b} \sum_{j=1}^k \partial^{b_j + e} f(x) \cdot \prod_{i \in \{1,\dots,k\} \setminus \{j\}} \partial^{b_i} f(x) \Bigg] \\
&\quad + \mathrm{D}^1 g(f(x)) \partial^\alpha f(x).
\end{aligned}$$

Note that for $b \in (\mathbb{N}_0^d)^{k-1}$ with $b_1 + \cdots b_{k-1} = \beta$ one has (for example) $(b_1, \dots, b_k, e) \in (\mathbb{N}_0^d)^k$ and of course $\beta + e = \alpha$. Also for $b \in (\mathbb{N}_0^d)^k$ with $b_1 + \dots + b_k = \beta$ one has $(b_1, \dots, b_j + e, \dots, b_k) \in (\mathbb{N}_0^d)^k$ and $b_1 + \dots + b_j + e + \dots + b_k = \beta + e = \alpha$. From this we can conclude that there exists $c_{m+1,a}$ for $a \in (\mathbb{N}_0^d \setminus \{0\})^k$ with $a_1 + \dots + a_k = \alpha$ such that the chain rule holds for $m+1$. Why do the constants not depend on the choice of e and β.... consistency left to prove. □

E Integration by parts

Theorem E.1. *[9, Appendix C.2, Theorem 2] Let Ω be a bounded open set with C^1 boundary $\partial\Omega$. We write σ for the $d-1$ dimensional "surface" measure on $\partial\Omega$. For*

$f, g \in C(\overline{\Omega})$ and $i \in \{1, \ldots, d\}$ we have

$$\int_U f \partial_i g = -\int_U g \partial_i f + \int_{\partial U} f g \mathfrak{n}_i \, \mathrm{d}\sigma,$$

where $\mathfrak{n}(x)$ for $x \in \partial U$ is the outward pointing normal vector and $\mathfrak{n}_i$ its i-th component.

F The Stone-Weierstrass Theorem

Theorem F.1 (Stone-Weierstrass, algebra version). *[5, Theorem 6.14] Let K be a compact topological space. Let D be a linear subspace of $C(K)$ with:*

(a) *if $f, g \in D$, then $fg \in D$;*

(b) *if $\mathbb{F} = \mathbb{C}$: if $f \in D$, then $\overline{f} \in D$;*

(c) *if $x, y \in K$ and $x \neq y$, then there is an $f \in D$ with $f(x) = 0$ and $f(y) = 1$.*

Then D is dense in $C(K)$.

G The Arzela-Ascoli Theorem

Theorem G.1 (Arzela-Ascoli). *[7, Page 3] Let $\mathcal{X}$ be a compact metric space. Suppose F is an infinite collection of functions that $\mathcal{X} \to \mathbb{F}$ that is equicontinuous and uniformly bounded, i.e.,*

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall f \in F \; \forall x, y \in X \; [\![d(x, y) < \delta \Rightarrow |f(x) - f(y)| < \varepsilon]\!], \tag{172}$$

$$\exists M > 0 \; \forall x \in X \; \forall f \in R \; [\![|f(x)| \leq M]\!]. \tag{173}$$

Then there exists a sequence $(f_n)_{n \in \mathbb{N}}$ in F that uniformly converges to a function $f \in C(\mathcal{X})$.

H Riesz representation theorem

Definition H.1. Let $(X, \mathcal{A})$ be a measurable space. A *(positive) measure* on $\mathcal{A}$ is a countably additive function $\mu : \mathcal{A} \to [0, \infty]$ such that $\mu(\emptyset) = 0$. A *signed measure* is a countably additive function $\mu : \mathcal{A} \to \mathbb{R}$ such that $\mu(\emptyset) = 0$. A *complex measure* is a countably additive function $\mu : \mathcal{A} \to \mathbb{C}$ such that $\mu(\emptyset) = 0$.

Theorem H.2 (Hahn-Jordan Decomposition). *[4, Theorem C.1] Let μ be a signed measure on a measurable space $(X, \mathcal{A})$. Then there exist positive measures μ_1, μ_2, measurable sets $E_1, E_2 \in \mathcal{A}$ such that $E_1 \cap E_2 = \emptyset$ and $E_1 \cup E_2 = X$, $\mu_1(E_2) = 0$, $\mu_2(E_1) = 0$ and $\mu = \mu_1 - \mu_2$. These measures μ_1 and μ_2 are unique, and one writes also μ^+ for μ_1 and μ^- for μ_2. The sets E_1 and E_2 are unique up to $\mu_1 + \mu_2$–null sets.*

Consequently, if μ is a complex measure, then there exist positive measures $\mu_1, \mu_2, \mu_3, \mu_4$ such that

$$\mu = \mu_1 - \mu_2 + \mathrm{i}[\mu_3 - \mu_4]. \tag{174}$$

Definition H.3. If μ is a positive, signed or complex measure, we define its *total variation* $|\mu|$ to be the function $\mathcal{A} \to [0, \infty]$ given by

$$|\mu|(A) = \sup\{\sum_{i=1}^{n} |\mu(E_i)| : E_1, \ldots, E_n \text{ are pairwise disjoint and in } \mathcal{A} \text{ and } \bigcup_{i=1}^{n} E_i = A\}.$$

Theorem H.4. *[4, Proposition C.3] If μ is a positive measure, then $|\mu| = \mu$.*
If μ is a signed measure, then $|\mu|$ is a positive finite measure and $|\mu| = \mu^+ - \mu^-$.
If μ is a complex measure, then $|\mu|$ is a positive finite measure then $|\mu| \le \sum_{i=1}^{4} \mu_i$.

Definition H.5. We say that a positive measure μ is *inner regular* if

$$\mu(A) = \sup\{\mu(K) : K \subset A, K \text{ is compact}\},$$

and is *outer regular* if

$$\mu(A) = \inf\{\mu(U) : A \subset U, U \text{ is open}\}.$$

Definition H.6. Let X be a topological space. It is also considered to be a measurable space equipped with the Borel-σ-algebra. We define $\mathcal{M}(X, \mathbb{R})$ and $\mathcal{M}(X, \mathbb{C})$ to be the set of signed and complex measures on X, respectively, such that their total variation is inner and outer regular. We define $\|\cdot\|_{\mathcal{M}} : \mathcal{M}(X, \mathbb{F}) \to [0, \infty)$, which is called the *total variation norm*, by

$$\|\mu\|_{\mathcal{M}} = |\mu|(X) \qquad (\mu \in \mathcal{M}(X, \mathbb{F})).$$

Theorem H.7. *[4, Proposition C.12] $\|\cdot\|_{\mathcal{M}}$ is a norm on $\mathcal{M}(X, \mathbb{F})$.*

H.8. If $X = \mathbb{R}^d$, then every Radon measure is inner and outer regular.

Definition H.9. Let X be a locally compact space. We write $C_0(X, \mathbb{F})$ for the continuous functions $X \to \mathbb{F}$ that vanish at infinity: $f \in C(X, \mathbb{F})$ is in $C_0(X, \mathbb{F})$ if for all $\varepsilon > 0$ there exists a compact set K such that $|f| < \varepsilon$ on $X \setminus K$. $C_0(X, \mathbb{F})$ is equipped with the norm $\|\cdot\|_{C^0}$ (for which we sometimes also write $\|\cdot\|_{C_0}$).

Theorem H.10 (Riesz(-Markov-Kakutani) representation theorem). *[4, Theorem C.18] Let X be a locally compact space. For $\mu \in \mathcal{M}(X, \mathbb{F})$ define $\Psi_\mu : C_0(X, \mathbb{F}) \to \mathbb{F}$ by*

$$\Psi_\mu(f) = \int f \, \mathrm{d}\mu \qquad (f \in C_0(X, \mathbb{F})).$$

Then $\psi_\mu \in C_0(X, \mathbb{F})'$ and the map $\mathcal{M}(X, \mathbb{F}) \to C_0(X, \mathbb{F})'$, $\mu \mapsto \Psi_\mu$ is an isometric isomorphism.

H.11. Of course if X be a compact Hausdorff space, the $C_0(X, \mathbb{F}) = C(X, \mathbb{F})$.

I Baire's category theorem

Theorem I.1 (Baire's Theorem). *[5, Theorem 11.1] Let X be a complete nonempty metric space and let $U_1, U_2, \dots$ be dense open subsets of X. Then the intersection of those sets, $\bigcap_{n\in\mathbb{N}} U_n$ is dense in X.*

J Hahn-Banach Theorem

Definition J.1. Let X be a vector space and $q : X \to \mathbb{R}$. Then q is called a *sublinear function* if

(a) $q(x+y) \leq q(x) + q(y)$ for all $x, y \in X$,

(b) $q(\lambda x) = \lambda q(x)$ for all $x \in X$ and $\lambda \geq 0$.

Theorem J.2. *[4, Theorem III.6.2] Let X be a vector space over $\mathbb{R}$ and $q : X \to \mathbb{R}$ be a sublinear functional. Let M be a linear subspace of X. If $f : M \to \mathbb{R}$ is a linear function such that $f(x) \leq q(x)$ for all $x \in M$, then there is a linear function $F : X \to \mathbb{R}$ such that $G|_M = f$ and $F(x) \leq q(x)$ for all $x \in X$.*

References

[1] R. Allez and K. Chouk. The continuous anderson hamiltonian in dimension two. Preprint available at `https://arxiv.org/abs/1511.02718`.

[2] H. Bahouri, J.-Y. Chemin, and R. Danchin. *Fourier analysis and nonlinear partial differential equations*, volume 343 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Heidelberg, 2011.

[3] V. Bogachev. *Measure theory. Vol. I, II.* Springer-Verlag, Berlin, 2007.

[4] J. B. Conway. *A Course in Functional Analysis*, volume 96 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1990.

[5] B. de Pagter and A. van Rooij. *An Invitation to Functional Analysis*. Epsilon Uitgaven. Epsilon Uitgaven, 2013.

[6] J. Dieudonné. *Treatise on analysis. Vol. II.* Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1976. Enlarged and corrected printing, Translated by I. G. Macdonald, With a loose erratum, Pure and Applied Mathematics, 10-II.

[7] W. F. Donoghue, Jr. *Distributions and Fourier transforms*, volume 32 of *Pure and Applied Mathematics*. Academic Press, New York, 1969.

[8] J. J. Duistermaat and J. A. C. Kolk. *Distributions*. Cornerstones. Birkhäuser Boston, Inc., Boston, MA, 2010. Theory and applications, Translated from the Dutch by J. P. van Braam Houckgeest.

[9] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

[10] F. G. Friedlander. *Introduction to the theory of distributions*. Cambridge University Press, Cambridge, second edition, 1998. With additional material by M. Joshi.

[11] A. Friedman. *Stochastic differential equations and applications. Vol. 1.* Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1975. Probability and Mathematical Statistics, Vol. 28.

[12] L. Grafakos. *Modern Fourier analysis*, volume 250 of *Graduate Texts in Mathematics*. Springer, New York, third edition, 2014.

[13] M. Gubinelli, P. Imkeller, and N. Perkowski. Paracontrolled distributions and singular PDEs. *Forum Math. Pi*, 3:e6, 75, 2015.

[14] M. Hairer. A theory of regularity structures. *Invent. Math.*, 198(2):269–504, 2014.

[15] G. H. Hardy. Weierstrass's non-differentiable function. *Trans. Amer. Math. Soc.*, 17(3):301–325, 1916. Pdf available at ams.org.

[16] L. Hörmander. Estimates for translation invariant operators in L^p spaces. *Acta Math.*, 104:93–140, 1960.

[17] T. Hytönen, J. van Neerven, M. Veraar, and L. Weis. *Analysis in Banach spaces. Vol. I. Martingales and Littlewood-Paley theory*, volume 63 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]*. Springer, Cham, 2016.

[18] G. Leoni. *A first course in Sobolev spaces*, volume 181 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2017.

[19] S. G. Michlin. Fourier integrals and multiple singular integrals. *Vestnik Leningrad. Univ. Ser. Mat. Meh. Astr.*, 12(7):143–155, 1957.

[20] C. Muscalu and W. Schlag. *Classical and multilinear harmonic analysis. Vol. I*, volume 137 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2013.

[21] D. J. Prömel and M. Trabs. Rough differential equations driven by signals in Besov spaces. *J. Differential Equations*, 260(6):5202–5249, 2016.

[22] A. C. M. van Rooij and W. H. Schikhof. *A Second Course on Real Functions*. Cambridge University Press, Cambridge-New York, 1982.

[23] A. C. M. van Rooij. *Fouriertheorie*. Epsilon Uitgaven, Utrecht, 1988.

[24] W. Rudin. *Functional analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, second edition, 1991.

[25] T. Runst and W. Sickel. *Sobolev spaces of fractional order, Nemytskij operators, and nonlinear partial differential equations*, volume 3 of *De Gruyter Series in Nonlinear Analysis and Applications*. Walter de Gruyter & Co., Berlin, 1996.

[26] Y. Sawano. *Theory of Besov spaces*, volume 56 of *Developments in Mathematics*. Springer, Singapore, 2018.

[27] H. H. Schaefer. *Topological Vector Spaces*. Springer-Verlag, New York-Berlin, 1971. Third printing corrected, Graduate Texts in Mathematics, Vol. 3.

[28] H.-J. Schmeisser and H. Triebel. *Topics in Fourier analysis and function spaces*. A Wiley-Interscience Publication. John Wiley & Sons, Ltd., Chichester, 1987.

[29] E. Shamir. A remark on Mikhlin-Hörmander multipliers theorem. *J. Math. Anal. Appl.*, 16:104–107, 1966.

[30] E. M. Stein and R. Shakarchi. *Fourier analysis*, volume 1 of *Princeton Lectures in Analysis*. Princeton University Press, Princeton, NJ, 2003. An introduction.

[31] R. S. Strichartz. *A guide to distribution theory and Fourier transforms*. World Scientific Publishing Co., Inc., River Edge, NJ, 2003. Reprint of the 1994 original [CRC, Boca Raton; MR1276724 (95f:42001)].

[32] F. Trèves. *Topological vector spaces, distributions and kernels*. Dover Publications, Inc., Mineola, NY, 2006. Unabridged republication of the 1967 original.

[33] H. Triebel. *Theory of function spaces*, volume 78 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 1983.