Theory of Distributions

Summer Term 2019

Michael Kunzinger

michael.kunzinger@univie.ac.at
Universität Wien
Fakultät für Mathematik
Oskar-Morgenstern-Platz 1
A-1090 Wien

Preface

These are lecture notes for a one semester course on the theory of distributions. Throughout we emphasize the applicability of distributions in concrete problems from analysis. Particular emphasis is laid on the calculation of fundamental solutions to linear partial differential equations. My main sources are the books [OW15] and [FJ98], as well as the lecture notes [HS09]. I am greatly indebted to Norbert Ortner, Peter Wagner, Günther Hörmann and Roland Steinbauer for allowing me to shamelessly copy material from [OW15] and [HS09], and for many helpful discussions. I also want to thank Norbert Ortner, Peter Wagner, Roman Popovych, and Eduard Nigsch for numerous helpful remarks and corrections.

Michael Kunzinger, summer term 2019

Contents

Chapter 1

Test functions and distributions

1.1 Basic definitions

Throughout these notes, Ω will denote a non-empty open subset of $\mathbb{R}^n$ and we write $\partial_i \equiv \frac{\partial}{\partial x^i}$ for the i-th partial derivative. We use Greek letters to denote multi-indices: for $\alpha, \beta \in \mathbb{N}_0^n$, $x \in \mathbb{R}^n$ we write

$$x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}, \quad |\alpha| := \alpha_1 + \cdots + \alpha_n, \quad \alpha! := \alpha_1! \cdots \alpha_n!$$
$$\alpha \geq \beta :\Leftrightarrow \forall i : \alpha_i \geq \beta_i$$
$$\binom{\alpha}{\beta} := \frac{\alpha!}{\beta!(\alpha - \beta!)} \quad \text{for } \alpha \geq \beta, \quad \partial^\alpha := \partial_1^{\alpha_1} \ldots \partial_n^{\alpha_n} = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}}.$$

Partial differential operators of order at most m with constant coefficients will be written in the form

$$P(\partial) = P(\partial_1, \ldots, \partial_n) = \sum_{|\alpha| \leq m} a_\alpha \partial^\alpha,$$

with $a_\alpha \in \mathbb{C}$, $m \in \mathbb{N}_0$.

1.1.1 Definition. *The $\mathbb{C}$-vector space of smooth functions on Ω is*

$$\mathcal{E}(\Omega) = C^\infty(\Omega) = \{\varphi : \Omega \to \mathbb{C} \mid \varphi \text{ is } C^\infty\}.$$

The space of test functions *on Ω is defined as*

$$\mathcal{D}(\Omega) = \{\varphi \in \mathcal{E}(\Omega) \mid \operatorname{supp}(\varphi) \Subset \Omega\}.$$

Here, $\operatorname{supp}(\varphi) = \overline{\{x \in \Omega \mid \varphi(x) \neq 0\}}$ and $\Subset$ means 'is a compact subset of'. We will sometimes write $\mathcal{E}$ or $\mathcal{D}$ instead of $\mathcal{E}(\Omega)$ and $\mathcal{D}(\Omega)$, respectively.

Both $\mathcal{E}(\Omega)$ and $\mathcal{D}(\Omega) \subseteq \mathcal{E}(\Omega)$ are locally convex spaces and there is an elaborate theory of the topological and functional analytic properties of these spaces, cf. e.g., [Hor66]. However, for the purpose of this lecture course we can and will confine ourselves to sequential convergence:

1.1.2 Definition.

(i) *A sequence* (φ_k) *in* $\mathcal{E}(\Omega)$ *converges to* $\varphi \in \mathcal{E}(\Omega)$ *if it converges uniformly on compact subsets of* Ω *in all derivatives, i.e.,*

$$\forall K \Subset \Omega\ \forall \alpha \in \mathbb{N}_0^n\ \forall \varepsilon > 0\ \exists N \in \mathbb{N}:\ \forall k \geq N:\ \|\partial^\alpha \varphi_k - \partial^\alpha \varphi\|_{\infty,K} \leq \varepsilon.$$

(ii) *A sequence* (φ_k) *in* $\mathcal{D}(\Omega)$ *converges to* $\varphi \in \mathcal{D}(\Omega)$ *if it converges in* $\mathcal{E}(\Omega)$ *and if, in addition, all* φ_k *are supported in some fixed* $K \Subset \Omega$*:*

$$\varphi_k \to \varphi \quad \textit{in } \mathcal{E}(\Omega) \textit{ and } \exists K \Subset \Omega : \forall k \in \mathbb{N} : \operatorname{supp}(\varphi_k) \subseteq K.$$

With the usual operations of addition, multiplication and multiplication by scalars, both $\mathcal{E}(\Omega)$ and $\mathcal{D}(\Omega)$ become topological $\mathbb{C}$-algebras.

1.1.3 Lemma. *Let* $f\colon \mathbb{R} \to \mathbb{R}$*,*

$$f(x) := \begin{cases} 0 & x \leq 0 \\ e^{-\frac{1}{x}} & x > 0 \end{cases}$$

Then f *is smooth.*

Proof. By induction we obtain that

$$f^{(n)}(x) = \begin{cases} 0 & x \leq 0 \\ e^{-\frac{1}{x}} P_n(\frac{1}{x}) & x > 0 \end{cases}$$

where P_n is a polynomial. Hence $\lim_{x\nearrow 0} f^{(n)}(x) = \lim_{x\searrow 0} f^{(n)}(x) = 0$ for all n. □

For $x_0 \in \Omega$ and $r > 0$ such that the closed ball $\overline{B_r(x_0)}$ is contained in Ω, it follows that the function $\varphi(x) := f(1-|x-x_0|^2/r^2)$ lies in $\mathcal{D}(\Omega)$, showing that $\mathcal{D}(\Omega) \neq \{0\}$. The following is an important topological property of $\mathcal{D}(\Omega)$:

1.1.4 Theorem. $\mathcal{D}(\Omega)$ *is sequentially complete.*

Proof. Let (φ_l) be a Cauchy sequence in $\mathcal{D}(\Omega)$ and let $K \Subset \Omega$ be such that $\operatorname{supp}(\varphi_l) \subseteq K$ for all $l \in \mathbb{N}$. Then for each $\alpha \in \mathbb{N}_0^n$ $(\partial^\alpha \varphi_l)$ is an $\|\,\|_\infty$-Cauchy sequence of continuous functions, hence has a continuous limit. More precisely, we have

$$\forall \alpha \in \mathbb{N}_0^n\ \exists \psi_\alpha \in C(\Omega) \text{ with } \partial^\alpha \varphi_l \to \psi_\alpha \text{ uniformly on } K.$$

We now claim that $\varphi_l \to \psi_0$ in $\mathcal{D}(\Omega)$. Indeed, $\operatorname{supp}(\psi_0) \subseteq K$ is clear, and for any $\alpha \in \mathbb{N}_0^n$, $1 \leq j \leq n$, $\beta = \alpha + e_j$ (with e_j the j-th unit vector) we have

$$\psi_\alpha = \lim_{l\to\infty} \partial^\alpha \varphi_l = \lim_{l\to\infty} \int_{-\infty}^{x_j} \partial^\beta \varphi_l(x_1, \ldots, s, \ldots, x_n)ds = \int_{-\infty}^{x_j} \psi_\beta(x_1, \ldots, s, \ldots, x_n)ds.$$

Hence $\partial_j \psi_\alpha = \psi_\beta$ and since α, β, and j were arbitrary we have $\psi_\alpha = \partial^\alpha \psi_0$ for all $\alpha \in \mathbb{N}_0^n$. But this implies that $\partial^\alpha \varphi_l \to \partial^\alpha \psi_0$ uniformly on K for all $\alpha \in \mathbb{N}_0^n$, and we are done. □

1.1.5 Proposition. $\mathcal{E}(\Omega)$ *is sequentially complete.*

Proof. Let (φ_l) be a Cauchy sequence in $\mathcal{E}(\Omega)$ and let $K \Subset \Omega$ and $\alpha \in \mathbb{N}_0$. Then $(\partial^\alpha \varphi_l)$ is a $\|\,\|_{\infty,K}$-Cauchy sequence, hence converges in $C(K)$. The limits must coincide on any non-trivial intersection of two such compact sets, giving a function $\psi_\alpha \in C(\Omega)$. As in the previous proof it then follows that $\varphi_l \to \psi_0$ in $\mathcal{E}(\Omega)$ (only the lower limit of integration is no longer $-\infty$ but some appropriate value a_j such that the line connecting $(x_1, \ldots, a_j, \ldots, x_n)$ and $(x_1, \ldots, x_j, \ldots, x_n)$ lies in Ω). □

We may now introduce the space of distributions on Ω as the set of sequentially continuous linear functionals on the space of test functions:

1.1.6 Definition.

(i) The space of distributions on Ω *is defined as*

$$\mathcal{D}'(\Omega) := \{T : \mathcal{D}(\Omega) \to \mathbb{C} \mid T \text{ linear and } T(\varphi_k) \to 0 \text{ if } \varphi_k \to 0 \text{ in } \mathcal{D}(\Omega)\}.$$

(ii) Let T_k, $T \in \mathcal{D}'(\Omega)$. *Then* $T_k \to T$ *if, for each* $\varphi \in \mathcal{D}(\Omega)$, $\langle T_k, \varphi\rangle \to \langle T, \varphi\rangle$. *Analogously, if* T_λ *is a family of distributions with* $\lambda \in \mathbb{C}^N$ *for some* N, *then* $\lim_{\lambda\to\lambda_0} T_\lambda = T$ *means that* $\langle T_\lambda, \varphi\rangle \to \langle T_{\lambda_0}, \varphi\rangle$ *for each* $\varphi \in \mathcal{D}(\Omega)$.

As already done in (ii) above, the action of a distribution $T \in \mathcal{D}'(\Omega)$ on a test function $\varphi \in \mathcal{D}(\Omega)$ is often denoted by $\langle T, \varphi\rangle$, emphasizing the bilinearity of the operation. If a test function depends on an additional variable, $\varphi = \varphi(x, y)$ then it is often useful to emphasize the variable which T acts upon, e.g., $T \in \mathcal{D}'(\mathbb{R}^n_x)$, $\langle T(x), \varphi(x, y)\rangle$. Moreover, given $K \Subset \Omega$, we write $\mathcal{D}(K)$ for the subspace of $\mathcal{D}(\Omega)$ consisting of those functions whose support is contained in K.

The following is a useful characterization of distributions by a *seminorm-estimate*:

1.1.7 Theorem. *Let* $T : \mathcal{D}(\Omega) \to \mathbb{C}$ *be linear. Then* $T \in \mathcal{D}'(\Omega) \iff \forall K \Subset \Omega$ $\exists C > 0$ $\exists m \in \mathbb{N}_0$:

$$|\langle T, \varphi\rangle| \le C \sum_{|\alpha|\le m} \|\partial^\alpha \varphi\|_{\infty,K} \qquad \forall \varphi \in \mathcal{D}(K). \tag{1.1.1}$$

Proof. ($\Leftarrow$): Let $\varphi_k \to 0$ in $\mathcal{D}(\Omega)$ with $K \supseteq \operatorname{supp}(\varphi_k)$ for all k. Choose $C > 0$ and $m \in \mathbb{N}_0$ according to (1.1.1). Then we have

$$|\langle T, \varphi_k\rangle| \le C \sum_{|\alpha|\le m} \|\partial^\alpha \varphi_k\|_{\infty,K} \to 0 \qquad (k \to \infty),$$

hence $T \in \mathcal{D}'(\Omega)$.

($\Rightarrow$): By contradiction: Assume that we have $T \in \mathcal{D}'(\Omega)$ but $\exists K \Subset \Omega$ $\forall m \in \mathbb{N}_0$ there is some $\varphi_m \in \mathcal{D}(K)$ such that

$$|\langle T, \varphi_m\rangle| > m \cdot \sum_{|\alpha|\le m} \|\partial^\alpha \varphi_m\|_{\infty,K}.$$

(Note that necessarily $\varphi_m \neq 0$ and thus $0 < \|\varphi_m\|_{\infty,K} \le \sum_{|\alpha|\le m} \|\partial^\alpha \varphi_m\|_{\infty,K}$.)

Now for $x \in \Omega$ define

$$\psi_m(x) := \frac{\varphi_m(x)}{m \cdot \sum_{|\alpha|\le m} \|\partial^\alpha \varphi_m\|_{\infty,K}}.$$

Then $\psi_m \in \mathcal{D}(K)$ and for any $\beta \in \mathbb{N}_0^n$ with $m \ge |\beta|$ we have

$$\|\partial^\beta \psi_m\|_{\infty,K} \le \sum_{|\gamma|\le m} \|\partial^\gamma \psi_m\|_{\infty,K} = \sum_{|\gamma|\le m} \frac{\|\partial^\gamma \varphi_m\|_{\infty,K}}{m \cdot \sum\limits_{|\alpha|\le m} \|\partial^\alpha \varphi_m\|_{\infty,K}} = \frac{1}{m},$$

so $\psi_m \to 0$ in $\mathcal{D}(\Omega)$ (as $m \to \infty$).

On the other hand, we obtain by construction

$$|\langle T, \psi_m\rangle| = \frac{|\langle T, \varphi_m\rangle|}{m \cdot \sum\limits_{|\alpha|\le m} \|\partial^\alpha \varphi_m\|_{\infty,K}} > 1,$$

and therefore $|\langle T, \psi_m\rangle| \not\to 0$ in $\mathbb{C}$, a contradiction. □

An important topological property of the space $\mathcal{D}'(\Omega)$ is the following:

1.1.8 Theorem. *The space $\mathcal{D}'(\Omega)$ is sequentially complete, i.e., if (T_k) is a Cauchy sequence in $\mathcal{D}'(\Omega)$, then $\lim_{k\to\infty}\langle T_k, \varphi\rangle$ exists for each $\varphi \in \mathcal{D}(\Omega)$, and*

$$\langle T, \varphi\rangle := \lim_{k\to\infty} \langle T_k, \varphi\rangle$$

defines a distribution $T \in \mathcal{D}'(\Omega)$.

Proof. Since $(\langle T_k, \varphi\rangle)$ is a Cauchy sequence for each $\varphi \in \mathcal{D}(\Omega)$, it converges due to the completeness of $\mathbb{C}$. Also, the resulting map $T : \mathcal{D}(\Omega) \to \mathbb{C}$ is clearly linear, so it only remains to show that it is sequentially continuous. Let us assume, to the contrary, that there exists some sequence $\varphi_j \to 0$ in $\mathcal{D}(\Omega)$ such that $\langle T, \varphi_j\rangle \not\to 0$. We will use a *gliding hump* technique to derive a contradiction. First, for some $c > 0$ we have $|\langle T, \varphi_j\rangle| > c$ for infinitely many j. Extracting a subsequence and multiplying by suitable numbers $\frac{2}{c}e^{i\theta}$ we obtain a new sequence, again denoted by the same letters, such that $\varphi_j \to 0$ in $\mathcal{D}(\Omega)$ and $\langle T, \varphi_j\rangle \geq 2$ (in particular: $\in \mathbb{R}$) for all $j \in \mathbb{N}$. Since $\varphi_j \to 0$ uniformly in all derivatives, we may extract another subsequence (ψ_j) such that

$$\begin{aligned}
\|\psi_1\|_\infty &\leq \frac{1}{2} \\
\|\partial^\alpha\psi_2\|_\infty &\leq \frac{1}{2^2} \quad \forall\alpha \text{ with } |\alpha| \leq 1 \\
\vdots \quad &\quad \vdots \quad \vdots \\
\|\partial^\alpha\psi_j\|_\infty &\leq \frac{1}{2^j} \quad \forall\alpha \text{ with } |\alpha| \leq j-1 \\
\vdots \quad &\quad \vdots \quad \vdots
\end{aligned} \tag{1.1.2}$$

Being a subsequence of (φ_j), $\psi_j \to 0$ in $\mathcal{D}(\Omega)$. As $T_k \in \mathcal{D}'(\Omega)$ for each k, this implies:

$$\lim_{j\to\infty} \langle T_k, \psi_j\rangle = 0 \quad \forall k. \tag{1.1.3}$$

Also, by definition of T we have

$$\lim_{k\to\infty} \langle T_k, \psi_j\rangle = \langle T, \psi_j\rangle \geq 2 \quad \forall j. \tag{1.1.4}$$

Next we use (1.1.3) and (1.1.4) to iteratively select a subsequence (ω_j) of (ψ_j), as well as a subsequence (S_k) of (T_k) as follows:

$$\begin{array}{rl}
\omega_1 \colon\ \omega_1 = \psi_1 & S_1 \colon\ \mathrm{Re}\langle S_1, \omega_1\rangle > 1 \\
\omega_2 \colon\ |\langle S_1, \omega_2\rangle| < 2^{-2} & S_2 \colon\ \mathrm{Re}\langle S_2, \omega_1\rangle > 1,\ \mathrm{Re}\langle S_2, \omega_2\rangle > 1 \\
\omega_3 \colon\ |\langle S_1, \omega_3\rangle|,\ |\langle S_2, \omega_3\rangle| < 2^{-3} & S_3 \colon\ \mathrm{Re}\langle S_3, \omega_j\rangle > 1 \quad (j \leq 3) \\
\vdots & \vdots \\
\omega_m \colon\ |\langle S_k, \omega_m\rangle| < 2^{-m} \quad (k < m) & S_m \colon\ \mathrm{Re}\langle S_m, \omega_j\rangle > 1 \quad (j \leq m) \\
\vdots & \vdots
\end{array}$$

Now set $\varphi := \sum_{j=1}^\infty \omega_j$. Then due to (1.1.2), this series converges uniformly in every derivative. Furthermore, all the supports of the ω_j lie in one fixed compact subset of Ω because (ω_j) is a subsequence of (φ_j), which converges to 0 in $\mathcal{D}(\Omega)$.

Consequently, φ is an element of $\mathcal{D}(\Omega)$ by Theorem 1.1.4. By definition of T we have $\langle T, \varphi\rangle = \lim_{k\to\infty}\langle S_k, \varphi\rangle$. On the other hand, the above construction implies

$$\operatorname{Re}\langle S_k, \varphi\rangle = \sum_{j=1}^{\infty} \operatorname{Re}\langle S_k, \omega_j\rangle > \underbrace{1+1+\cdots+1}_{=k} - \sum_{j=k+1}^{\infty} \frac{1}{2^j} \to \infty,$$

giving the desired contradiction. □

To conclude this section we state a result that we will only prove much later (Section 4.3). Of course we will carefully avoid any circular arguments when using it in the meantime.

1.1.9 Theorem. *$\mathcal{D}(\Omega)$ is sequentially dense in $\mathcal{D}'(\Omega)$, i.e., for any $T \in \mathcal{D}'(\Omega)$ there exists a sequence (u_j) in $\mathcal{D}(\Omega)$ such that, for any $\varphi \in \mathcal{D}(\Omega)$,*

$$\langle u_j, \varphi\rangle := \int_\Omega u_j(x)\varphi(x)\,dx \to \langle T, \varphi\rangle \qquad (j \to \infty).$$

Concerning the action of u_j on φ used here, see Theorem 1.2.3 below.

1.2 Examples of distributions

Distributions are simultaneous generalizations of locally integrable functions and of Radon measures. Let us first examine the space $L^1_{\mathrm{loc}}(\Omega)$ as a subspace of $\mathcal{D}'(\Omega)$.

1.2.1 Definition. *A measurable function $f : \Omega \to \mathbb{C}$ is called* locally integrable *if $\int_K |f(x)|\,dx < \infty$ for all $K \Subset \Omega$. The space of all (equivalence classes of) locally integrable functions on Ω is denoted by $L^1_{\mathrm{loc}}(\Omega)$.*

1.2.2 Example. Every continuous function is locally integrable, so $C(\Omega) \subseteq L^1_{\mathrm{loc}}(\Omega)$. The function $\frac{1}{x}$ is locally integrable on $\mathbb{R}\setminus\{0\}$, but not on $\mathbb{R}$, while $\frac{1}{\sqrt{|x|}} \in L^1_{\mathrm{loc}}(\mathbb{R})$.

1.2.3 Theorem.

(i) If $f \in L^1_{\mathrm{loc}}(\Omega)$, then the associated linear functional

$$T_f : \mathcal{D}(\Omega) \to \mathbb{C}, \quad \langle T_f, \varphi\rangle := \int_\Omega f(x)\varphi(x)\,dx$$

is a distribution, $T_f \in \mathcal{D}'(\Omega)$.

(ii) The map

$$L^1_{\mathrm{loc}}(\Omega) \to \mathcal{D}'(\Omega), \quad f \mapsto T_f$$

is linear and injective. In this sense we may consider $L^1_{\mathrm{loc}}(\Omega)$ as a linear subspace of $\mathcal{D}'(\Omega)$, and we will often write f instead of T_f.

Proof. (i) Clearly $T_f : \mathcal{D}(\Omega) \to \mathbb{C}$ is linear. Moreover, it is continuous: let $\varphi_k \to \varphi$ in $\mathcal{D}(\Omega)$. Then

$$\left|\int_\Omega f(x)(\varphi_k(x) - \varphi(x))\,dx\right| \le \|\varphi_k - \varphi\|_{\infty,K} \int_K |f(x)|\,dx \to 0 \ (k \to \infty)$$

where $K \Subset \Omega$ is chosen such that $\operatorname{supp}(\varphi_k) \subseteq K$ for all k (and hence $\operatorname{supp}(\varphi) \subseteq K$).

(ii) Since $f \mapsto T_f$ is linear, it suffices to show that $T_f = 0$ entails $f = 0$ a.e. on Ω. To this end, note that $T_f = 0$ implies

$$\mathcal{F}(f\varphi)(\xi) = \int_{\mathbb{R}^n} e^{-i\xi x} f(x)\varphi(x)\,dx = 0$$

for each $\varphi \in \mathcal{D}(\Omega)$ and each $\xi \in \mathbb{R}^n$. Thus injectivity of the Fourier transform (see Proposition 1.2.4 below) implies $f(x) = 0$ a.e., as claimed. □

1.2.4 Proposition. *Denote by $C_{\mathrm{b}}(\mathbb{R}^n)$ the space of continuous and bounded functions $\mathbb{R}^n \to \mathbb{C}$. Then the Fourier transform*

$$\mathcal{F} : L^1(\mathbb{R}^n) \to C_{\mathrm{b}}(\mathbb{R}^n), \quad f \mapsto \left(\xi \mapsto \int_{\mathbb{R}^n} e^{-i\xi x} f(x)\,dx\right)$$

is well defined, linear, continuous and injective.

Proof. We follow [OW15], based on [New74].

$\mathcal{F}(f)$ is continuous by dominated convergence, and bounded because $\|\mathcal{F}(f)\|_\infty \leq \|f\|_1$. The latter estimate also implies continuity of $\mathcal{F}$.

To show injectivity, we first look at the case $n = 1$ and suppose that $\mathcal{F}(f) = 0$ for some $f \in L^1(\mathbb{R})$. Then we define a function $g : \mathbb{C} \to \mathbb{C}$ by

$$g(z) := \begin{cases} \int_{-\infty}^0 f(x)e^{-ixz}\,dx & \mathrm{Im}(z) \geq 0, \\ -\int_0^\infty f(x)e^{-ixz}\,dx & \mathrm{Im}(z) \leq 0. \end{cases}$$

Then g is well defined on $\mathbb{R}$ because $\mathcal{F}(f) = 0$. Also, it is continuous on $\mathbb{C}$ and analytic on $\mathbb{C} \setminus \mathbb{R}$. By continuity and Cauchy's integral theorem, the integral of g over any triangle that lies either in $\{\mathrm{Im}(z) \geq 0\}$ or in $\{\mathrm{Im}(z) \leq 0\}$ vanishes. An arbitrary triangle in $\mathbb{C}$ can be decomposed into finitely many such triangles, so indeed $\int_\Delta g(z)\,dz = 0$ for *any* triangle Δ in $\mathbb{C}$. Therefore, Morera's theorem shows that g is an entire function. Since $\|g\|_\infty \leq \|f\|_1$ (as can be seen by writing $z = a+ib$ and taking into account the cases $b \geq 0$ resp. $b \leq 0$ in the definition of g), Liouville's theorem implies that g is constant. Moreover, by dominated convergence we have $\lim_{y\to\infty} g(iy) = 0$, so in fact g has to vanish identically. Therefore,

$$0 = g(0) = \int_{-\infty}^0 f(x)\,dx.$$

Replacing f by $x \mapsto f(x+\xi)$ we in fact obtain $\int_{-\infty}^\xi f(x)\,dx = 0$ for any $\xi \in \mathbb{R}$. The derivative of $\xi \mapsto \int_{-\infty}^\xi f(x)\,dx$ equals $f(\xi)$ a.e., so the claim follows in this case.

To obtain the result for general n we use induction: By Fubini's theorem, if $f \in L^1(\mathbb{R}^n)$, then

$$f_{x_1} : \mathbb{R}^{n-1} \to \mathbb{C}, \quad x' \mapsto f(x_1, x')$$

belongs to $L^1(\mathbb{R}^{n-1})$ for all $x_1 \in \mathbb{R} \setminus N$, where N is a Lebesgue null-set. Also, again by Fubini, for any $\xi' \in \mathbb{R}^{n-1}$, the map

$$g_{\xi'}(x_1) := \mathcal{F}f_{x_1}(\xi') = \int_{\mathbb{R}^{n-1}} f(x_1, x')e^{-i\xi' x'}\,dx'$$

is integrable and $\mathcal{F}g_{\xi'}(\xi_1) = \mathcal{F}f(\xi) = 0$ for any $\xi_1 \in \mathbb{R}$. By what was shown above in the case $n = 1$ it follows that $g_{\xi'} = 0$ almost everywhere. Another application of Fubini's theorem gives

$$0 = \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}\setminus N} |g_{\xi'}(x_1)|\,dx_1 d\xi' = \int_{\mathbb{R}\setminus N} \int_{\mathbb{R}^{n-1}} |g_{\xi'}(x_1)|\,d\xi' dx_1,$$

and so $\int_{\mathbb{R}^{n-1}} |g_{\xi'}(x_1)|\, d\xi' = 0$ for $x_1 \in \mathbb{R} \setminus N_1$, where N_1 is another null-set that contains N. Since $\xi' \mapsto g_{\xi'}(x_1)$ is continuous, we obtain that $\mathcal{F}(f_{x_1})(\xi') = g_{\xi'}(x_1) = 0$ for all $\xi' \in \mathbb{R}^{n-1}$ and all $x_1 \in \mathbb{R} \setminus N_1$. By induction hypothesis, we have

$$0 = \|f_{x_1}\|_{L^1(\mathbb{R}^{n-1})} = \int_{\mathbb{R}^{n-1}} |f(x_1, x')|\, dx'$$

for all $x_1 \in \mathbb{R} \setminus N_1$. But then in fact

$$\|f\|_1 = \int_{\mathbb{R}} \int_{\mathbb{R}^{n-1}} |f(x_1, x')|\, dx' dx_1 = \int_{\mathbb{R}\setminus N_1} \int_{\mathbb{R}^{n-1}} |f(x_1, x')|\, dx' dx_1 = 0,$$

and thereby $f = 0$ a.e., as claimed. □

1.2.5 Remark.

(i) Since, by the above, we may consider $L^1_{\mathrm{loc}}(\Omega)$ as a subset of $\mathcal{D}'(\Omega)$, the same is true for $L^p_{\mathrm{loc}}(\Omega)$. In fact, for $1 \leq p < \infty$ and $K \Subset \Omega$, by Hölder's inequality $\|f\|_{1,K} = \|f \cdot 1\|_{1,K} \leq \|f\|_{p,K} \|1\|_{q,K}$, where q is such that $1/p + 1/q = 1$. Since $\|1\|_{q,K} = \lambda(K)^{1/q}$, this gives the claim. For $p = \infty$, it is immediate that $L^\infty_{\mathrm{loc}}(\Omega) \subseteq L^1_{\mathrm{loc}}(\Omega)$. In particular, $L^p_{\mathrm{c}}(\Omega)$, the space of locally p-integrable functions of compact support is contained in $\mathcal{D}'(\Omega)$ as well.

(ii) Any Radon measure is a distribution. Indeed, the space $\mathcal{M}(\Omega)$ of Radon measures on Ω consists of all continuous linear functionals on the space $C_{\mathrm{c}}(\Omega)$. Here, a sequence (φ_k) in $C_{\mathrm{c}}(\Omega)$ is convergent if all the supports of the φ_k are contained in a single compact subset of Ω and if the φ_k converge uniformly. From this it is immediate that $\mathcal{M}(\Omega) \subseteq \mathcal{D}'(\Omega)$.

Elements of $L^1_{\mathrm{loc}}(\Omega)$ are often also called *regular* distributions. Not surprisingly, these do not exhaust $\mathcal{D}'(\Omega)$:

1.2.6 Example. (i) An important example of a distribution that is not regular is the *Dirac measure*: For any $a \in \Omega$, set

$$\delta_a : \mathcal{D}(\Omega) \to \mathbb{C}, \quad \delta_a(\varphi) = \varphi(a),$$

and set $\delta := \delta_0 \in \mathcal{D}'(\mathbb{R}^n)$. Then obviously $\delta_a \in \mathcal{M}(\Omega) \subseteq \mathcal{D}'(\Omega)$. To see that indeed $\delta_a \notin L^1_{\mathrm{loc}}(\Omega)$, let f be as in Lemma 1.1.3 and set $\rho_k(x) := f(1 - k|x - a|^2)$ $(k \in \mathbb{N})$. Then for any $g \in L^1_{\mathrm{loc}}(\Omega)$, $\langle g, \rho_k \rangle \to 0$ as $k \to \infty$. However, $\langle \delta_a, \rho_k \rangle = f(1) \neq 0$ for all k.

On the other hand, δ is the weak limit of functions from $L^1(\mathbb{R}^n)$: in fact, fix any $\rho \in L^1(\mathbb{R}^n)$ with $\int \rho(x)\, dx = 1$ and set

$$\rho_\varepsilon(x) := \varepsilon^{-n} \rho\left(\frac{x}{\varepsilon}\right). \tag{1.2.1}$$

Then $\lim_{\varepsilon \to 0+} \rho_\varepsilon = \delta$ because for any $\varphi \in \mathcal{D}(\mathbb{R}^n)$ we have:

$$\begin{aligned} \langle \rho_\varepsilon, \varphi \rangle = \varepsilon^{-n} \int_{\mathbb{R}^n} \rho(x/\varepsilon)\varphi(x)\, dx &= \int_{\mathbb{R}^n} \rho(y)\varphi(\varepsilon y)\, dy \\ \to \int_{\mathbb{R}^n} \rho(y)\varphi(0)\, dy &= \varphi(0) = \langle \delta, \varphi \rangle \end{aligned}$$

by dominated convergence.

(ii) For a general $f \in L^1(\mathbb{R}^n)$ with $f \geq 0$ and $f \neq 0$, we may set $\rho := \frac{1}{\|f\|_1} f$ to obtain $\rho_\varepsilon \to \delta$ by (i). In this way we can derive several important examples of so-called *delta-nets*:

First, denote by Y the Heaviside function: $Y(x) = 1$ for $x > 0$, $Y(x) = 0$ for $x \le 0$. Then $f(x) := Y(1 - |x|) = \chi_{B_1(0)}$ is the characteristic function of the unit ball $B_1(0)$ in $\mathbb{R}^n$. Therefore, we have $\|f\|_1 = \mathrm{vol}(B_1(0)) = \pi^{n/2}/\Gamma(n/2+1)$, and so

$$\lim_{\varepsilon\to 0+} C\varepsilon^{-n} Y(\varepsilon - |x|) = \delta \quad \text{with } C = \pi^{-n/2}\Gamma(n/2+1) \tag{1.2.2}$$

Next, let $f(x) := \frac{1}{(1+|x|^2)^{\frac{n+1}{2}}}$. Then $\|f\|_1 = \frac{\pi^{\frac{n+1}{2}}}{\Gamma(\frac{n+1}{2})}$, so setting $\rho := \frac{1}{\|f\|_1} f$ we get

$$\lim_{\varepsilon\to 0+} \frac{C\varepsilon}{(|x|^2+\varepsilon^2)^{\frac{n+1}{2}}} = \delta \quad \text{with } C = \pi^{-\frac{n+1}{2}}\Gamma(\frac{n+1}{2}). \tag{1.2.3}$$

Finally, for $f(x) := e^{-|x|^2}$ and this time using $\rho_{\sqrt{\varepsilon}}$ we obtain

$$\lim_{\varepsilon\to 0+} Ce^{-|x|^2/\varepsilon} = \delta \quad \text{with } C = \pi^{-\frac{n}{2}}. \tag{1.2.4}$$

1.2.7 Example. Let us now look at examples of distributions that do not correspond to measures (cf. Remark 1.2.5). Let σ denote the surface measure on S^{n-1} and let $f \in L^1(S^{n-1})$ be such that f satisfies the *mean value zero condition*

$$\int_{S^{n-1}} f(\omega)d\sigma(\omega) = 0. \tag{1.2.5}$$

Then we define the *principal value* (*valeur principale*) $\mathrm{vp}(|x|^{-n}f(\frac{x}{|x|}))$ by the limit

$$\mathrm{vp}(|x|^{-n}f(\frac{x}{|x|})) := \lim_{\varepsilon\to 0+} Y(|x|-\varepsilon)|x|^{-n}f(\frac{x}{|x|}),$$

i.e.,

$$\langle \mathrm{vp}(|x|^{-n}f(\frac{x}{|x|})), \varphi\rangle := \lim_{\varepsilon\to 0+} \int_{|x|\ge\varepsilon} \frac{\varphi(x)}{|x|^n} f\big(\frac{x}{|x|}\big)\,dx \quad (\varphi \in \mathcal{D}).$$

To see that this limit actually exists, we first note that, by (1.2.5),

$$\int_{\varepsilon\le|x|\le R} \frac{1}{|x|^n} f\big(\frac{x}{|x|}\big)\,dx = \int_\varepsilon^R \frac{1}{r}\int_{S^{n-1}} f(\omega)d\sigma(\omega)dr = 0,$$

so that, if $\operatorname{supp}\varphi \subseteq B_R(0)$,

$$\begin{aligned}
\langle \mathrm{vp}(|x|^{-n}f(\frac{x}{|x|})), \varphi\rangle &= \lim_{\varepsilon\to 0+} \int_{\varepsilon\le|x|\le R} \frac{\varphi(x)}{|x|^n} f\Big(\frac{x}{|x|}\Big)\,dx \\
&= \lim_{\varepsilon\to 0+} \int_{\varepsilon\le|x|\le R} \frac{\varphi(x)-\varphi(0)}{|x|^n} f\Big(\frac{x}{|x|}\Big)\,dx \qquad (1.2.6)\\
&= \int_{|x|\le R} \frac{\varphi(x)-\varphi(0)}{|x|^n} f\Big(\frac{x}{|x|}\Big)\,dx.
\end{aligned}$$

Here, the last integral converges due to $|\varphi(x) - \varphi(0)| \le \|\nabla\varphi\|_\infty \cdot |x|$. This estimate also shows that $\langle \mathrm{vp}(|x|^{-n}f(\frac{x}{|x|})), \varphi_k\rangle \to 0$ if $\varphi_k \to 0$ in $\mathcal{D}$, i.e., $\mathrm{vp}(|x|^{-n}f(\frac{x}{|x|}))$ is indeed a distribution. The latter fact also follows from Theorem 1.1.8.

Important examples of principal value distributions are the following:

$$\mathrm{vp}\Big(\frac{1}{x}\Big) := \mathrm{vp}(|x|^{-1}\mathrm{sign}(x/|x|)), \quad \langle \mathrm{vp}\Big(\frac{1}{x}\Big), \varphi\rangle = \int_{-R}^R \frac{\varphi(x)-\varphi(0)}{x}\,dx,$$

for $\operatorname{supp}\varphi \subseteq [-R,R]$, as well as

$$T_{jk} = \operatorname{vp}\left(\frac{\delta_{jk}|x^2| - nx_jx_k}{|x|^{n+2}}\right) \in \mathcal{D}'(\mathbb{R}^n) \quad (1 \leq j,k \leq n),$$

which arise as the second derivatives of the kernel of the Newton potential, i.e., of $|x|^{2-n}$ for $n \neq 2$, resp. of $\log(|x|)$ for $n = 2$. Here, $f_{jk}(\omega) = \delta_{jk} - n\omega_j\omega_k$ satisfies (1.2.5), because $\int_{S^{n-1}} \omega_j^2 \, d\sigma(\omega)$ has the same value for each j by symmetry, so for $j = k$ we get

$$n \int_{S^{n-1}} \omega_j^2 \, d\sigma(\omega) = \int_{S^{n-1}} \sum_{j=1}^{n} \omega_j^2 \, d\sigma(\omega) = \int_{S^{n-1}} 1 \, d\sigma(\omega) = |S^{n-1}|.$$

For $j \neq k$, note that the map $(\omega_1, \ldots, \omega_n) \mapsto (\omega_1, \ldots, -\omega_j, \ldots, \omega_n)$ transforms $\int_{S^{n-1}} \omega_j\omega_k \, d\sigma(\omega)$ into its own negative, so the integral must vanish.

Finally, that principal value distributions are not measures follows from the considerations above which relied on properties of $\nabla\varphi$, hence require derivatives of the test function (as opposed to the case of Radon measures).

1.2.8 Example. A complex approximation of $\operatorname{vp}(\frac{1}{x})$ is given by Sokhotski's formula:

$$\lim_{\varepsilon \to 0+} \frac{1}{x \pm i\varepsilon} = \operatorname{vp}\Big(\frac{1}{x}\Big) \mp i\pi\delta. \tag{1.2.7}$$

To see this note, first, that by (1.2.3) we have

$$\lim_{\varepsilon \to 0+} \frac{1}{x \pm i\varepsilon} = \lim_{\varepsilon \to 0+} \frac{x \mp i\varepsilon}{x^2 + \varepsilon^2} = \lim_{\varepsilon \to 0+} \frac{x}{x^2 + \varepsilon^2} \mp i\pi\delta.$$

Second, for $\operatorname{supp}(\varphi) \subseteq [-R,R]$ and $\varepsilon \to 0+$ we have

$$\left\langle \frac{x}{x^2+\varepsilon^2}, \varphi \right\rangle = \int_{-R}^{R} \frac{(\varphi(x) - \varphi(0))x}{x^2+\varepsilon^2} \, dx \to \int_{-R}^{R} \frac{\varphi(x) - \varphi(0)}{x} \, dx = \left\langle \operatorname{vp}\Big(\frac{1}{x}\Big), \varphi \right\rangle,$$

which gives the claim. From (1.2.7) we directly obtain *Heisenberg's formula*

$$\lim_{\varepsilon \to 0+} \left(\frac{1}{x + i\varepsilon} - \frac{1}{x - i\varepsilon} \right) = -2\pi i\delta.$$

The two cases of (1.2.7) correspond to one another via complex conjugation in the following sense:

1.2.9 Definition. *For $T \in \mathcal{D}'(\Omega)$, the* complex conjugate $\bar{T}$, *the* real part $\operatorname{Re}(T)$, *and the* imaginary part $\operatorname{Im}(T)$, *are defined by*

$$\langle \bar{T}, \varphi \rangle := \overline{\langle T, \bar{\varphi} \rangle}, \ \operatorname{Re}(T) := \frac{1}{2}(T + \bar{T}), \ \operatorname{Im}(T) := \frac{1}{2i}(T - \bar{T}),$$

respectively.

1.3 Distributions of finite order

1.3.1 Definition. *A distribution $T \in \mathcal{D}'(\Omega)$ is said to be* of finite order, *if in the seminorm-estimate* (1.1.1), *the integer m may be chosen uniformly for all K, i.e.*

$$\exists m \in \mathbb{N}_0 \, \forall K \Subset \Omega \, \exists C > 0 : \qquad |\langle T, \varphi \rangle| \leq C \sum_{|\alpha| \leq m} \|\partial^\alpha \varphi\|_{\infty,K} \quad (\varphi \in \mathcal{D}(K)). \tag{1.3.1}$$

The minimal $m \in \mathbb{N}_0$ satisfying the above is then called the order of the distribution. *The space of all distributions of order less or equal to m is denoted by $\mathcal{D}'^m(\Omega)$.*

The subspace of all distributions of finite order is denoted by $\mathcal{D}'_F(\Omega)$,

$$\mathcal{D}'_F(\Omega) = \bigcup_{m\in\mathbb{N}_0} \mathcal{D}'^m(\Omega).$$

1.3.2 Example.

(i) Any regular distribution is of order 0.

(ii) δ_a is of order 0.

(iii) Let $|\alpha| = m$ and define $T \in \mathcal{D}'(\mathbb{R}^n)$ by $\langle T, \varphi\rangle := \partial^\alpha\varphi(0)$. Then T is of order m.

(iv) There exist distributions that are *not of finite order.* E.g., consider $T \in \mathcal{D}'(\mathbb{R})$ defined by

$$\langle T, \varphi\rangle := \sum_{k=0}^{\infty} \varphi^{(k)}(k).$$

If $K \Subset \mathbb{R}$ then we have to choose $m \geq \sup\{k \in \mathbb{N}_0 \mid k \in K\}$ to ensure (1.3.1). There can be no m such that (1.3.1) holds with this fixed m and for all compact subsets K. The farther outward $\operatorname{supp}(\varphi)$ reaches the higher the derivatives that have to be taken into account.

For a distribution of order m, the continuity condition (1.1.1) involves only derivatives up to order m of the test functions. Thus, if we define the space $\mathcal{D}^m(\Omega)$ as the set of all m-times continuously differentiable functions with compact support contained in Ω, we may expect that $\mathcal{D}'^m(\Omega)$ will be the (sequential) dual of $\mathcal{D}^m(\Omega)$. Here, sequential convergence in $\mathcal{D}^m(\Omega)$ is defined analogously to $\mathcal{D}(\Omega)$, only taking into account derivatives up to order m. $\mathcal{D}'^m(\Omega)$ then consists of all sequentially continuous linear forms on $\mathcal{D}^m(\Omega)$.

A slimmed-down version of the proof of Theorem 1.1.4 gives:

1.3.3 Proposition. *$\mathcal{D}^m(\Omega)$ is sequentially complete.*

The technical problem with the above identification is that given $T \in \mathcal{D}'^m(\Omega)$ we have to extend it to a linear functional on $\mathcal{D}^m(\Omega)$, which is strictly larger than $\mathcal{D}(\Omega)$. The precise result is Theorem 1.3.6 below.

We note that, in particular, distributions of order 0 define continuous linear forms on $C_c(\Omega)$. Therefore they can be identified with complex *Radon measures* on Ω.

We are going to need the following result on approximation of functions of finite differentiability, which is also of independent interest.

1.3.4 Theorem. *Let $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and let $\rho \in \mathcal{D}(\mathbb{R}^n)$, $\operatorname{supp}(\rho) \subseteq \overline{B_1(0)}$, and $\int \rho(x)\,dx = 1$ (such a ρ is called a* mollifier*). For $\varepsilon \in (0, 1]$ we define*

$$f_\varepsilon(x) := \int_{\mathbb{R}^n} f(y)\rho_\varepsilon(x-y)\,dy = \frac{1}{\varepsilon^n}\int_{\mathbb{R}^n} f(y)\rho(\frac{x-y}{\varepsilon})\,dy. \tag{1.3.2}$$

($f_\varepsilon = f * \rho_\varepsilon$, the convolution of f and ρ_ε).
Then

(i) $f_\varepsilon \in \mathcal{E}(\mathbb{R}^n)$ with $\operatorname{supp}(f_\varepsilon) \subseteq \{x \in \mathbb{R}^n \mid d(x, \operatorname{supp}(f)) \leq \varepsilon\}$; (where $d(x_0, A) := \inf_{x\in A}|x - x_0|$ for any $x_0 \in \mathbb{R}^n$ and $A \subseteq \mathbb{R}^n$)

(ii) If f is compactly supported then $f_\varepsilon \in \mathcal{D}(\mathbb{R}^n)$.

(iii) *If $m < \infty$ and $f \in C^m(\mathbb{R}^n)$, then $f_\varepsilon \to f$ in $C^m(\mathbb{R}^n)$ (as $\varepsilon \to 0$);*

(iv) *If $f \in \mathcal{E}(\mathbb{R}^n)$ then $f_\varepsilon \to f$ in $\mathcal{E}(\mathbb{R}^n)$ (as $\varepsilon \to 0$).*

(v) *If $f \in \mathcal{D}^m(\mathbb{R}^n)$ then $f_\varepsilon \to f$ in $\mathcal{D}^m(\mathbb{R}^n)$ (as $\varepsilon \to 0$).*

(vi) *If $f \in \mathcal{D}(\mathbb{R}^n)$ then $f_\varepsilon \to f$ in $\mathcal{D}(\mathbb{R}^n)$ (as $\varepsilon \to 0$).*

In (iii), convergence in $C^m(\mathbb{R}^n)$ is defined analogously to that in $\mathcal{E}(\mathbb{R}^n)$, only taking into account derivatives up to order m.

Proof. (i) and (ii): By dominated convergence we may pull derivatives of f_ε into the integral, so smoothness of f_ε follows from that of ρ_ε. Noting that $\operatorname{supp}(\rho_\varepsilon) \subseteq \overline{B_\varepsilon(0)}$ and changing integration variables from y to $y' = x - y$ we may write

$$f_\varepsilon(x) = \int_{\mathbb{R}^n} f(y)\rho_\varepsilon(x-y)\,dy = \int_{\overline{B_\varepsilon(0)}} f(x-y')\rho_\varepsilon(y')\,dy'.$$

If $x \in \mathbb{R}^n$ with $d(x, \operatorname{supp}(f)) > \varepsilon$ then $f(x - y') = 0$ for all y' in the integration domain, thus $f_\varepsilon(x) = 0$. Therefore $\operatorname{supp}(f_\varepsilon) \subseteq \{x \in \mathbb{R}^n \mid d(x, \operatorname{supp}(f)) \leq \varepsilon\}$.
Concerning the remaining claims, we first prove uniform convergence on any $K \Subset \mathbb{R}^n$ of $f_\varepsilon \to f$ ($\varepsilon \to 0$).
The change of variables $y \mapsto z = (x-y)/\varepsilon$ (hence $dz = dy/\varepsilon^n$) yields

$$f_\varepsilon(x) = \frac{1}{\varepsilon^n}\int_{\mathbb{R}^n} f(y)\rho\Big(\frac{x-y}{\varepsilon}\Big)\,dy = \int_{\mathbb{R}^n} f(x-\varepsilon z)\rho(z)\,dz.$$

Therefore, by uniform continuity of f on compact sets, we have for any $x \in K$:

$$\begin{aligned}
|f_\varepsilon(x) - f(x)| &\overset{[\int\rho=1]}{=} \Big|\int_{\mathbb{R}^n} f(x-\varepsilon z)\rho(z)\,dz - \int_{\mathbb{R}^n} f(x)\rho(z)\,dz\Big| \\
&\overset{[\operatorname{supp}(\rho)\subseteq\overline{B_1(0)}]}{\leq} \int_{\overline{B_1(0)}} |f(x-\varepsilon z) - f(x)|\,|\rho(z)|\,dz \\
&\leq \underbrace{\Big(\int |\rho(z)|\,dz\Big)}_{\text{constant}} \cdot \sup_{|y|\leq\varepsilon, x\in K} |f(x-y) - f(x)| \quad \to 0 \qquad (\varepsilon \to 0).
\end{aligned}$$

If $|\alpha| \leq m$ then the same game can be played with $\partial^\alpha f_\varepsilon(x) = \int \partial^\alpha f(x-\varepsilon z)\rho(z)\,dz$ to show uniform convergence $\partial^\alpha f_\varepsilon \to \partial^\alpha f$ on K. Thus we obtain, in particular, uniform convergence of all derivatives up to order m on compact subsets of $\mathbb{R}^n$. From this and what we already know about $\operatorname{supp}(f_\varepsilon)$, (iii)-(vi) follow. □

1.3.5 Remark. (i) Given $f \in \mathcal{D}^k(\Omega)$ and constructing f_ε as in the above theorem, we obtain for $\varepsilon < d(\operatorname{supp}(f), \Omega^c)$ that $\operatorname{supp}(f_\varepsilon) \subseteq \Omega$, hence $f_\varepsilon \in \mathcal{D}(\Omega)$ and $f_\varepsilon \to f$ in $\mathcal{D}^k(\Omega)$. (Here, $d(A,B) = \inf_{x\in A, y\in B} |x-y|$ for subsets $A, B \subseteq \mathbb{R}^n$.)
(ii) As a special case of the result in (i) we conclude that $\mathcal{D}(\Omega)$ is dense in $C_c(\Omega)$.

Based on these preparations we can now prove:

1.3.6 Theorem.

(i) *Every $T \in \mathcal{D}'^m(\Omega)$ can uniquely be extended to a continuous linear form on $\mathcal{D}^m(\Omega)$.*

(ii) Conversely, if T is a continuous linear form on $\mathcal{D}^m(\Omega)$ then $T|_{\mathcal{D}(\Omega)} \in \mathcal{D}'^m(\Omega)$.

Proof. (i) Let $T \in \mathcal{D}'^m(\Omega)$, then we have: $\forall K \Subset \Omega\ \exists C > 0$ (depending on K) with

$$|\langle T, \psi\rangle| \leq C \sum_{|\alpha|\leq m} \|\partial^\alpha \psi\|_{\infty,K} \qquad \forall \psi \in \mathcal{D}(K). \tag{1.3.3}$$

Let $\varphi \in \mathcal{D}^m(\Omega)$. By Remark 1.3.5 there is a sequence (φ_l) in $\mathcal{D}(\Omega)$ such that $\varphi_l \to \varphi$ in $\mathcal{D}^m(\Omega)$ (as $l \to \infty$). That is, there exists $K_0 \Subset \Omega$ with $\operatorname{supp}(\varphi) \subseteq K_0$ and $\operatorname{supp}(\varphi_l) \subseteq K_0$ for all l such that for all α with $|\alpha| \leq m$ we have $\partial^\alpha \varphi_l \to \partial^\alpha \varphi$ uniformly on K_0. In particular, for $|\alpha| \leq m$ we obtain a Cauchy sequence $(\partial^\alpha \varphi_l)$ with respect to the L^∞-norm on K_0.

Choosing $C_0 > 0$ according to (1.3.3) we obtain

$$|\langle T, \varphi_k - \varphi_l\rangle| \leq C_0 \sum_{|\alpha|\leq m} \|\partial^\alpha \varphi_k - \partial^\alpha \varphi_l\|_{\infty,K_0},$$

which implies that $(\langle T, \varphi_l\rangle)$ is a Cauchy sequence in $\mathbb{C}$, hence possesses a limit $\overline{T}(\varphi) := \lim\langle T, \varphi_l\rangle$. By a standard sequence mixing argument we see that the value $\overline{T}(\varphi)$ is independent of the approximating sequence (φ_l). Linearity with respect to φ is clear, hence we obtain a linear form $\overline{T}$ on $\mathcal{D}^m(\Omega)$. Moreover,

$$|\langle \overline{T}, \varphi\rangle| = \lim |\langle T, \varphi_l\rangle| \leq \lim C_0 \sum_{|\alpha|\leq m} \|\partial^\alpha \varphi_l\|_{\infty,K_0} = C_0 \sum_{|\alpha|\leq m} \|\partial^\alpha \varphi\|_{\infty,K_0}$$

shows (sequential) continuity of $\overline{T}$. That $\overline{T}|_{\mathcal{D}(\Omega)} = T$ follows by choosing constant approximating sequences $\varphi_l \equiv \varphi$ for $\varphi \in \mathcal{D}(\Omega)$, and uniqueness of $\overline{T}$ as an extension of T follows from the density of $\mathcal{D}(\Omega)$ in $\mathcal{D}^m(\Omega)$ (observed in Remark 1.3.5).

(ii) is clear since sequential continuity of T implies the same for $T|_{\mathcal{D}(\Omega)}$: Convergent sequences in $\mathcal{D}(\Omega)$ converge also in $\mathcal{D}^m(\Omega)$ and have the same limit. Thus $T|_{\mathcal{D}(\Omega)} \in \mathcal{D}'(\Omega)$. Since T is sequentially continuous on $\mathcal{D}^m(\Omega)$, it follows as in the proof of Theorem 1.1.7 that it satisfies a seminorm estimate of the form (1.3.3), hence it is of order m. $\square$

Chapter 2

Multiplication, support, composition

2.1 Multiplication by smooth functions

In Definition 1.2.9 we have already seen an example of extending classical operations to distributions via *transposition*, i.e., roughly, by having the test function carry the burden. More precisely, the evaluation of the transformed distribution AT on a test function φ is defined as $\langle T, A^t\varphi\rangle$, where A^t is a transposed operator constructed in such a way that, for regular distributions T_f, $AT_f = T_{Af}$, i.e.,

$$\langle T_{Af}, \varphi\rangle = \langle T_f, A^t\varphi\rangle \quad (\varphi \in \mathcal{D}(\Omega)).$$

As a model case, let us first consider the multiplication of a distribution by a smooth function:

2.1.1 Definition. *For $T \in \mathcal{D}'(\Omega)$ and $g \in \mathcal{E}(\Omega)$, the* multiplication gT *is defined by*

$$\langle gT, \varphi\rangle := \langle T, g\varphi\rangle \quad (\varphi \in \mathcal{D}(\Omega)).$$

Since $\varphi_k \to \varphi$ in $\mathcal{D}(\Omega)$ implies $g\varphi_k \to g\varphi$, the linear functional gT so defined is indeed a distribution. This is the first example of a general property of transposition in the above sense:

2.1.2 Proposition. *Let $L : \mathcal{D}(\Omega_2) \to \mathcal{D}(\Omega_1)$ be linear and sequentially continuous. Then the transpose $L^t : \mathcal{D}'(\Omega_1) \to \mathcal{D}'(\Omega_2)$, $\langle L^tT, \varphi\rangle := \langle T, L\varphi\rangle$ is linear and sequentially continuous.*

Proof. To see that $L^tT \in \mathcal{D}'(\Omega_2)$ for any $T \in \mathcal{D}'(\Omega_1)$, let $\varphi_l \to 0$ in $\mathcal{D}(\Omega_2)$. Then $L\varphi_l \to 0$ in $\mathcal{D}(\Omega_1)$ by the sequential continuity of L. Hence

$$\langle L^tT, \varphi_l\rangle = \langle T, L\varphi_l\rangle \to 0.$$

Linearity is immediate from the definition.
As for continuity of L^t, let $T_k \to T$ in $\mathcal{D}'(\Omega_1)$. Then for any $\varphi \in \mathcal{D}(\Omega_2)$

$$\langle L^tT_k, \varphi\rangle = \langle T_k, L\varphi\rangle \stackrel{(k\to\infty)}{\longrightarrow} \langle T, L\varphi\rangle = \langle L^tT, \varphi\rangle,$$

hence $L^tT_k \to L^tT$ in $\mathcal{D}'(\Omega_2)$. □

In the case of multiplication, L simply corresponds to the map $\varphi \mapsto g\varphi$. Consistency with the classical multiplication in the case of regular distributions follows from

$$\langle T_{gf}, \varphi\rangle = \int_\Omega g(x)f(x)\varphi(x)\,dx = \langle T_f, g\varphi\rangle.$$

2.2 Localization and support

Let $\Omega' \subseteq \Omega$ open, then there is a natural embedding $L : \mathcal{D}(\Omega') \hookrightarrow \mathcal{D}(\Omega)$ by extending test functions on Ω' by 0 to $\Omega \setminus \Omega'$. The restriction of distributions on Ω to Ω' is then defined as the transpose of this L:

2.2.1 Definition. *For $\Omega' \subseteq \Omega$, the* restriction *(or* localization*) of any $T \in \mathcal{D}'(\Omega)$ to Ω' is the element of $\mathcal{D}'(\Omega')$ given by $T|_{\Omega'} := L^t T$, i.e.,*

$$\langle T|_{\Omega'}, \varphi\rangle := \langle T, L\varphi\rangle \quad (\varphi \in \mathcal{D}(\Omega')).$$

In addition, we define the support *of $T \in \mathcal{D}'(\Omega)$ by*

$$\mathrm{supp}\,(T) := \Omega \setminus \bigcup\{\Omega' \subseteq \Omega \mid \Omega' \textit{ open},\ T|_{\Omega'} = 0\}.$$

For $f \in C(\Omega)$ it is easy to see that $\mathrm{supp}\,(T_f) = \mathrm{supp}\,(f) = \overline{\{x \in \Omega \mid f(x) \neq 0\}}$ (where the closure is to be taken in the subset topology of Ω). On the other hand, if $f \in L^1_{\mathrm{loc}}(\Omega)$, then $\Omega\setminus\mathrm{supp}\,(T_f)$ is the largest open set U such that $\int_U |f(x)|\,dx = 0$.

In order to prove some fundamental properties of localization we need the following basic version of partitions of unity.

2.2.2 Proposition. *Let $K \Subset \Omega$ and let $\Omega_i \subseteq \Omega$ be open $(1 \leq i \leq m)$ such that $K \subseteq \bigcup_{i=1}^m \Omega_i$. Then there exist functions $\psi_i \in \mathcal{D}(\Omega_i)$, $(1 \leq i \leq m)$ with $0 \leq \psi_i \leq 1$, $\sum_{i=1}^m \psi_i \leq 1$ on Ω and $\sum_{i=1}^m \psi_i = 1$ on a neighborhood of K.*

Proof. Consider first the case where $m = 1$. For $x \in \mathbb{R}^n$ and $\emptyset \neq A \subseteq \mathbb{R}^n$, the map $x \mapsto d(x, A)$ is continuous. Indeed, by the triangle inequality, $|d(x, A) - d(y, A)| \leq |x - y|$. Therefore, for any $\varepsilon > 0$, $K_\varepsilon := \{x \in \mathbb{R}^n \mid d(x, K) < \varepsilon\}$ is an open neighborhood of K. Pick $\varepsilon > 0$ such that $4\varepsilon < d(K, \mathbb{R}^n \setminus \Omega_1)$ and set

$$f(x) := \begin{cases} 1 - \frac{1}{\varepsilon} d(x, K_{2\varepsilon}) & x \in K_{3\varepsilon} \\ 0 & x \notin K_{3\varepsilon} \end{cases}$$

Then f is continuous and $\mathrm{supp}\,(f) \subseteq \overline{K_{3\varepsilon}}$. Set $\psi := f_\varepsilon$, as in (1.3.2). Then $\mathrm{supp}\,(\psi) \subseteq \overline{K_{4\varepsilon}} \subseteq \Omega_1$, $0 \leq \psi \leq 1$, and $\psi \equiv 1$ on K_ε, since for $x \in K_\varepsilon$

$$\psi(x) = \int_{|x-y|<\varepsilon} f(y)\rho_\varepsilon(x - y)\,dy = \int 1 \cdot \rho_\varepsilon(x - y)\,dy = 1.$$

Turning now to the case $m > 1$, note first that for each $x \in K$ there exists some $r(x) > 0$ and some $i \in \{1, \dots, m\}$ such that $\overline{B_{r(x)}(x)} \subseteq \Omega_i$. These $B_{r(x)}$ form an open cover of K from which we may therefore extract a finite subcover $B_1, \dots, B_N$. For $i \in \{1, \dots, m\}$, let $K_i := \bigcup_{\overline{B_j} \subseteq \Omega_i} \overline{B_j}$. Then $K_i \Subset \Omega_i$ and $K \subseteq \bigcup_{i=1}^m K_i$. By what we have shown in the first case, for each $i \in \{1, \dots, m\}$ there exists some $\varphi_i \in \mathcal{D}(\Omega_i)$, $0 \leq \varphi_i \leq 1$ and $\varphi_i \equiv 1$ on a neighborhood U_i of K_i.

Now set $\psi_1 := \varphi_1$, $\psi_2 := \varphi_2 \cdot (1 - \varphi_1)$, $\psi_m := \varphi_m(1 - \varphi_1)\cdots(1 - \varphi_{m-1})$. Then $\mathrm{supp}\,(\psi_i) \subseteq \Omega_i$, $0 \leq \psi_i \leq 1$, and

$$\sum_{i=1}^m \psi_i = 1 - (1 - \varphi_1)\cdots(1 - \varphi_m),$$

as is readily verified by induction. Thus $\sum_{i=1}^m \psi_i \le 1$ and $\sum_{i=1}^m \psi_i = 1$ on the neighborhood $\bigcup_{i=1}^m U_i$ of K. □

2.2.3 Corollary. *Let $K \Subset \Omega$. Then there exists a cut-off (or plateau-) function for K, i.e., $\chi \in \mathcal{D}(\Omega)$ such that $\chi \equiv 1$ on a neighborhood of K.*

Proof. This is the special case $m = 1$ of Theorem 2.2.2. □

2.2.4 Proposition. *Let $T \in \mathcal{D}'(\Omega)$ and $\varphi \in \mathcal{D}(\Omega)$ with* $\operatorname{supp}(T) \cap \operatorname{supp}(\varphi) = \emptyset$. *Then $\langle T, \varphi \rangle = 0$.*

Proof. Let $K := \operatorname{supp}(\varphi)$. Since $\operatorname{supp}(T) \cap K = \emptyset$, for any $x \in K$ there exists a neighborhood U_x of x in Ω such that $T|_{U_x} = 0$.

Choose a finite subcovering $(U_{x_i})_{i=1}^m$ and a subordinated partition of unity $(\psi_i)_{i=1}^m$ as in Proposition 2.2.2, that is, with $\psi_j \in \mathcal{D}(U_{x_j})$ and $\sum_{i=1}^m \psi_i = 1$ on a neighborhood of K. Then

$$\langle T, \varphi \rangle = \langle T, \underbrace{\sum_{i=1}^m \varphi\, \psi_i}_{=\varphi} \rangle = \sum_{i=1}^m \langle T, \underbrace{\varphi\, \psi_i}_{\in \mathcal{D}(U_{x_i})} \rangle = 0.$$

□

2.2.5 Corollary. *Let $T \in \mathcal{D}'(\Omega)$. If any $x \in \Omega$ has a neighborhood $U_x \subseteq \Omega$ such that $T|_{U_x} = 0$, then $T = 0$ in $\mathcal{D}'(\Omega)$.*

Proof. It follows from the definition of the support of T and our assumption that $\operatorname{supp}(T) = \emptyset$. Proposition 2.2.4 then implies $\langle T, \varphi \rangle = 0$ for all $\varphi \in \mathcal{D}(\Omega)$, so $T = 0$. □

In fact the space of distributions forms a *fine sheaf* (of C^∞-modules), which basically amounts to the following result:

2.2.6 Theorem. *Let I be a set and $(\Omega_i)_{i\in I}$ be an open covering of Ω. For every $i \in I$ let $T_i \in \mathcal{D}'(\Omega_i)$ such that the following holds:*

$$T_i|_{\Omega_i \cap \Omega_j} = T_j|_{\Omega_i \cap \Omega_j} \qquad \forall i, j \in I \text{ with } \Omega_i \cap \Omega_j \neq \emptyset. \tag{2.2.1}$$

Then $\exists!\, T \in \mathcal{D}'(\Omega)$ with $T|_{\Omega_j} = T_j$ for all $j \in I$.

(A collection of distributions $T_i \in \mathcal{D}'(\Omega_i)$ satisfying (2.2.1) *is called a* coherent family.*)*

Proof. *Uniqueness:* Let $T_1, T_2 \in \mathcal{D}'(\Omega)$ and suppose that $\forall i \in I$: $T_1|_{\Omega_i} = T_i = T_2|_{\Omega_i}$. Set $T := T_1 - T_2$, then $T|_{\Omega_i} = 0$ for all $i \in I$, so Corollary 2.2.5 implies $T = 0$, thus $T_1 = T_2$.

Existence: Let $K \Subset \Omega$. Since $K \subseteq \bigcup_{i\in I} \Omega_i$ we may pick a finite subcovering: $\exists i_1, \dots, i_m \in I$ such that $K \subseteq \bigcup_{l=1}^m \Omega_{i_l}$. According to Proposition 2.2.2 we can find a subordinate partition of unity, i.e. $\psi_l \in \mathcal{D}(\Omega_{i_l})$ $(l = 1 \dots, m)$ with $\sum_{l=1}^m \psi_l = 1$ in a neighborhood of K. For every compact subset K of Ω we choose a corresponding partition of unity.

Now we define the action of T on $\varphi \in \mathcal{D}(\Omega)$ as follows: Let $K := \operatorname{supp}(\varphi)$ and let $\psi_1, \dots, \psi_m$ be the partition of unity chosen above. Then we set

$$\langle T, \varphi \rangle := \sum_{l=1}^m \langle T_{i_l}, \varphi\, \psi_l \rangle. \tag{2.2.2}$$

We have to show that

(a) the value of $\langle T, \varphi \rangle$ is well defined by (2.2.2) (i.e. it depends only on $(T_i)_{i \in I}$ and φ),

(b) $T \in \mathcal{D}'(\Omega)$, and

(c) $T|_{\Omega_i} = T_i$ for all $i \in I$.

(a) Let K' be a compact subset with $K' \supseteq \operatorname{supp}(\varphi)$ and suppose that $\Omega_{r_1}, \ldots, \Omega_{r_p}$ is a corresponding subcovering with subordinate partition of unity $\psi'_1, \ldots, \psi'_p$. Then we have

$$\sum_{k=1}^{p} \langle T_{r_k}, \varphi\, \psi'_k \rangle = \sum_{k=1}^{p} \sum_{l=1}^{m} \langle T_{r_k}, \overbrace{\varphi\, \psi'_k \psi_l}^{\in \mathcal{D}(\Omega_{i_l} \cap \Omega_{r_k})} \rangle \underset{(2.2.1)}{=} \sum_{l=1}^{m} \sum_{k=1}^{p} \langle T_{i_l}, \varphi\, \psi_l \psi'_k \rangle = \sum_{l=1}^{m} \langle T_{i_l}, \varphi\, \psi_l \rangle.$$

(b) We prove the seminorm estimate (1.1.1). If $K \Subset \Omega$ we have (with subcovering and partition of unity as chosen above) the action on any $\varphi \in \mathcal{D}(K)$ given by (2.2.2). Using (1.1.1) for every $T_{i_1}, \ldots, T_{i_m}$ (with compact set $\operatorname{supp}(\psi_{i_l}) \cap K$, and $C > 0$ and order N uniformly over $l = 1, \ldots, m$) we obtain the estimate

$$\begin{aligned} |\langle T, \varphi \rangle| \le \sum_{l=1}^{m} |\langle T_{i_l}, \varphi \psi_l \rangle| &\le C \sum_{l=1}^{m} \sum_{|\alpha| \le N} \|\partial^\alpha (\varphi \psi_l)\|_{\infty, K} \\ {\scriptstyle [\text{Leibniz rule: } \partial^\alpha(\varphi\psi_l) = \sum_{\beta \le \alpha} \binom{\alpha}{\beta} \partial^\beta \varphi \partial^{\alpha-\beta} \psi_l]} \quad &\le C' \sum_{|\alpha| \le N} \|\partial^\alpha \varphi\|_{\infty, K}, \end{aligned}$$

where C' depends only on K (via ψ_{i_l}, $l = 1, \ldots, m$).

(c) Let $\varphi \in \mathcal{D}(\Omega_i)$ and $K := \operatorname{supp}(\varphi)$. Use Corollary 2.2.3 to construct a cut-off function $\chi \in \mathcal{D}(\Omega_i)$ over K, i.e. $\chi = 1$ in a neighborhood of K. Then the single set Ω_i provides a finite covering of K and χ is a partition of unity subordinate to it. Consequently, $\varphi = \varphi\, \chi$ and (2.2.2) yields

$$\langle T, \varphi \rangle = \langle T_i, \varphi\, \chi \rangle = \langle T_i, \varphi \rangle.$$

□

2.3 Distributions with compact support

Now that we have developed a notion of support for distributions, we may single out from $\mathcal{D}'(\Omega)$ the subspace of those distributions that have compact support. As it will turn out, this subspace can be identified with the dual space $\mathcal{E}'(\Omega)$ of $\mathcal{E}(\Omega)$, which we now introduce:

2.3.1 Definition. *We denote the space of sequentially continuous linear functionals on $\mathcal{E}(\Omega)$ by $\mathcal{E}'(\Omega)$. Thus if $T : \mathcal{E}(\Omega) \to \mathbb{C}$ is linear, then $T \in \mathcal{E}'(\Omega)$ if and only if, whenever $\varphi_k \to \varphi$ in $\mathcal{E}(\Omega)$, then $\langle T, \varphi_k \rangle \to \langle T, \varphi \rangle$.*

Our first aim is to derive a characterization of continuity for linear functionals $\mathcal{E}(\Omega) \to \mathbb{C}$ via a seminorm estimate similar to Theorem 1.1.7. To prove it, we need a technical tool called *compact exhaustion* of an open set Ω. By this we mean a sequence of compact sets (K_m) such that $K_m \subseteq K_{m+1}^\circ$ for each m and $\Omega = \bigcup_{m \in \mathbb{N}} K_m$. Such an exhaustion always exists: e.g., one may set $K_m := \{x \in \Omega \mid |x| \le m \text{ and } d(x, \partial\Omega) \ge 1/m\}$.

2.3.2 Theorem. *Let $T : \mathcal{E}(\Omega) \to \mathbb{C}$ be linear. Then $T \in \mathcal{E}'(\Omega) \iff \exists K \Subset \Omega$ $\exists C > 0$ $\exists m \in \mathbb{N}_0$:*

$$|\langle T, \varphi\rangle| \le C \sum_{|\alpha|\le m} \|\partial^\alpha \varphi\|_{\infty,K} \qquad \forall \varphi \in \mathcal{E}(\Omega). \tag{2.3.1}$$

Proof. $\Leftarrow$: Let $\varphi_k \to 0$ in $\mathcal{E}(\Omega)$. By assumption $\exists K, C, m$ as in (2.3.1), hence

$$|\langle T, \varphi_k\rangle| \le C \sum_{|\alpha|\le m} \|\partial^\alpha \varphi_k\|_{\infty,K} \;\to\; 0 \quad (k \to \infty).$$

$\Rightarrow$: By contradiction: Suppose that $T \in \mathcal{E}'(\Omega)$ but that (2.3.1) is violated and let (K_m) be a compact exhaustion of Ω. Then setting $C = m$ it follows that, for each $m \in \mathbb{N}$, there exists some $\varphi_m \in \mathcal{E}(\Omega)$ such that

$$|\langle T, \varphi_m\rangle| > m \sum_{|\alpha|\le m} \|\partial^\alpha \varphi_m\|_{\infty,K_m}. \tag{2.3.2}$$

Without loss of generality we may assume that $K_1^\circ \neq \emptyset$ and we may pick some $0 \neq \psi \in \mathcal{D}(K_1^\circ)$. Then for each $m \in \mathbb{N}$, replacing (if necessary) φ_m by $\varphi_m + c_m\psi$ for some sufficiently small $c_m > 0$, the inequality (2.3.2) remains intact and the new φ_m additionally satisfies $\|\varphi_m\|_{\infty,K_m} \neq 0$. Therefore, $0 < \|\varphi_m\|_{\infty,K_m} \le \sum_{|\alpha|\le m}\|\partial^\alpha \varphi_m\|_{\infty,K_m}$, and so we may define

$$\psi_m := \frac{\varphi_m}{m\sum_{|\alpha|\le m}\|\partial^\alpha \varphi_m\|_{\infty,K_m}}$$

to obtain an element of $\mathcal{E}(\Omega)$ for each $m \in \mathbb{N}$.
Now given $K \Subset \Omega$ and $\beta \in \mathbb{N}_0^n$, let $m \in \mathbb{N}$ be such that $m \ge |\beta|$ and $K \subseteq K_m$. Then

$$\|\partial^\beta \psi_m\|_{\infty,K} \le \sum_{|\gamma|\le m} \|\partial^\gamma \psi_m\|_{\infty,K_m} = \sum_{|\gamma|\le m} \frac{\|\partial^\gamma \varphi_m\|_{\infty,K_m}}{m \cdot \sum\limits_{|\alpha|\le m} \|\partial^\alpha \varphi_m\|_{\infty,K_m}} = \frac{1}{m},$$

hence $\psi_m \to 0$ in $\mathcal{E}(\Omega)$ (as $m \to \infty$). By construction, however,

$$|\langle T, \psi_m\rangle| = \frac{|\langle T, \varphi_m\rangle|}{m \cdot \sum\limits_{|\alpha|\le m} \|\partial^\alpha \varphi_m\|_{\infty,K_m}} > 1,$$

and therefore $|\langle T, \psi_m\rangle| \not\to 0$ in $\mathbb{C}$, a contradiction. $\square$

We have $\mathcal{D}(\Omega) \subseteq \mathcal{E}(\Omega)$, and this imbedding is sequentially continuous: $\varphi_k \to 0$ in $\mathcal{D}(\Omega)$ implies $\varphi_k \to 0$ in $\mathcal{E}(\Omega)$. Moreover:

2.3.3 Proposition. *$\mathcal{D}(\Omega)$ is sequentially dense in $\mathcal{E}(\Omega)$.*

Proof. Let $\varphi \in \mathcal{E}(\Omega)$ and pick a compact exhaustion (K_m) ($m \in \mathbb{N}$) of Ω. For each $m \in \mathbb{N}$, let χ_m be a cut-off function for K_m (cf. Corollary 2.2.3), and set $\varphi_m := \varphi \cdot \chi_m \in \mathcal{D}(\Omega)$. For any $K \Subset \Omega$ there exists some $m \in \mathbb{N}$ such that $K \subseteq K_m$. Then $\varphi_k = \varphi$ on a neighborhood of K for each $k \ge m$, so $\varphi_k \to \varphi$ in $\mathcal{E}(\Omega)$. $\square$

2.3.4 Remark. Using the previous results, we may now consider $\mathcal{E}'(\Omega)$ as a subspace of $\mathcal{D}'(\Omega)$:

(i) If $T \in \mathcal{E}'(\Omega)$ then $T|_{\mathcal{D}(\Omega)} \in \mathcal{D}'(\Omega)$. Indeed, we have

$$\varphi_k \to 0 \text{ in } \mathcal{D}(\Omega) \Rightarrow \varphi_k \to 0 \text{ in } \mathcal{E}(\Omega) \Rightarrow \langle T, \varphi_k\rangle \to 0.$$

Moreover, $T \mapsto T|_{\mathcal{D}(\Omega)}$ is an embedding of $\mathcal{E}'(\Omega)$ into $\mathcal{D}'(\Omega)$ (i.e., injective) because $\mathcal{D}(\Omega)$ is dense in $\mathcal{E}(\Omega)$ by Proposition 2.3.3.

(ii) If $T \in \mathcal{E}'(\Omega)$ then $T|_{\mathcal{D}(\Omega)}$ has compact support. Indeed, let $K \Subset \Omega$ be as in (2.3.1). Then for any $\varphi \in \mathcal{D}(\Omega)$, if $\operatorname{supp}(\varphi) \cap K = \emptyset$, then $\langle T, \varphi \rangle = 0$. Therefore $\operatorname{supp}(T|_{\mathcal{D}(\Omega)}) \subseteq K$.

In fact, the elements of $\mathcal{E}'(\Omega)$ correspond *precisely* to the compactly supported distributions in $\mathcal{D}'(\Omega)$:

2.3.5 Theorem. *Let $T \in \mathcal{D}'(\Omega)$ and suppose that $\operatorname{supp}(T)$ is compact. Then there exists a unique $\widetilde{T} \in \mathcal{E}'(\Omega)$ with $\widetilde{T}|_{\mathcal{D}(\Omega)} = T$.*

Proof. *Uniqueness* follows from the density of $\mathcal{D}(\Omega)$ in $\mathcal{E}(\Omega)$ (Proposition 2.3.3).

Existence: Let $\rho \in \mathcal{D}(\Omega)$ with $\rho = 1$ on a neighborhood of $\operatorname{supp}(T)$ (Corollary 2.2.3) and define $\widetilde{T}$ by

$$\langle \widetilde{T}, \varphi \rangle := \langle T, \rho\varphi \rangle \qquad (\varphi \in \mathcal{E}(\Omega)).$$

Then $\widetilde{T} : \mathcal{E}(\Omega) \to \mathbb{C}$ is linear and for any $\varphi \in \mathcal{D}(\Omega)$ we have

$$\langle \widetilde{T}, \varphi \rangle = \langle T, \rho\varphi \rangle = \langle T, \varphi \rangle + \underbrace{\langle T, (\rho - 1)\varphi \rangle}_{\substack{=0 \text{ by 2.2.4, since} \\ \operatorname{supp}((\rho-1)\varphi) \cap \operatorname{supp}(T) = \emptyset}} = \langle T, \varphi \rangle,$$

thus $\widetilde{T}\,|_{\mathcal{D}(\Omega)} = T$.

It remains to show that $\widetilde{T} \in \mathcal{E}'(\Omega)$. Let $K := \operatorname{supp}(\rho)$. Then for every $\varphi \in \mathcal{E}(\Omega)$ we have $\operatorname{supp}(\rho\varphi) \subseteq K$ and therefore $\rho\varphi \in \mathcal{D}(K)$. Thanks to (1.1.1) we can find $C > 0$ and $m \in \mathbb{N}_0$ such that

$$|\langle \widetilde{T}, \varphi \rangle| = |\langle T, \rho\varphi \rangle| \leq C \sum_{|\alpha| \leq m} \|\partial^\alpha(\rho\varphi)\|_{\infty, K} \qquad \forall \varphi \in \mathcal{E}(\Omega).$$

Applying the Leibniz rule to the terms $\partial^\alpha(\rho\varphi)$ we obtain the estimate (2.3.1) (as in the proof of Theorem 2.2.6, part (b)). □

Based on this result we shall henceforth consider $\mathcal{E}'(\Omega)$ as the subspace of $\mathcal{D}'(\Omega)$ consisting of those distributions that have compact support.

2.3.6 Remark. We observe that the definition of $\widetilde{T}$ in the proof of Theorem 2.3.5 does not depend on the choice of ρ. In fact, let $\chi \in \mathcal{D}(\Omega)$ also be a cut-off over (a neighborhood of) $\operatorname{supp}(T)$, then $\operatorname{supp}(\rho - \chi) \cap \operatorname{supp}(T) = \emptyset$ and Proposition 2.2.4 yields

$$\langle T, \rho\varphi \rangle - \langle T, \chi\varphi \rangle = \langle T, (\rho - \chi)\varphi \rangle = 0.$$

2.3.7 Corollary. *Any compactly supported distribution is of finite order: $\mathcal{E}'(\Omega) \subseteq \mathcal{D}'_F(\Omega)$.*

Proof. The seminorm estimate (2.3.1) in particular implies (1.1.1) with one fixed N. □

2.3.8 Theorem. *$\mathcal{E}'(\Omega)$ is sequentially dense in $\mathcal{D}'(\Omega)$.*

Proof. Let $T \in \mathcal{D}'(\Omega)$ and let (K_m) be a compact exhaustion of Ω. For each m, let $\chi_m \in \mathcal{D}(\Omega)$ be a cut-off function for K_m (Corollary 2.2.3) and set $T_m := \chi_m \cdot T$. Then $T_m \in \mathcal{E}'(\Omega)$ and we show that $T_m \to T$ in $\mathcal{D}'(\Omega)$. To this end, let $\varphi \in \mathcal{D}(\Omega)$ and let $K := \operatorname{supp}(\varphi)$. Then there exists some $m \in \mathbb{N}$ such that $K \subseteq K_m$, and so for each $k \geq m$ we have $\varphi = \chi_k \cdot \varphi$. Consequently,

$$\langle T_k, \varphi \rangle = \langle T, \chi_k\varphi \rangle = \langle T, \varphi \rangle$$

for these k. □

2.3.9 Examples.

(i) For $1 \le p \le \infty$, $L^p_c(\Omega) \subseteq \mathcal{E}'(\Omega)$.

(ii) Any element of $T \in \mathcal{E}'(\Omega)$ can be extended by 0 to a distribution in $\widetilde{T} \in \mathcal{E}'(\mathbb{R}^n)$:

$$\widetilde{T} : \mathcal{E}(\mathbb{R}^n) \to \mathbb{C}, \quad \langle \tilde{T}, \varphi \rangle := \langle T, \rho\varphi \rangle,$$

where ρ is as in the proof of Theorem 2.3.5.

On the other hand, for general distributions in $\mathcal{D}'(\Omega)$, an extension to an element of $\mathcal{D}'(\mathbb{R}^n)$ may, but need not be possible in general:

2.3.10 Examples.

(i) We first consider a situation where extension is possible. If $f \in L^1(S^{n-1})$ does not satisfy the mean-value zero condition (1.2.5), then $\lim_{\varepsilon \to 0+} Y(|x| - \varepsilon)|x|^{-n} f(\frac{x}{|x|})$ does not exist in $\mathcal{D}'(\mathbb{R}^n)$. Nevertheless, we may extend

$$S = |x|^{-n} f\Big(\frac{x}{|x|}\Big) \in L^1_{\mathrm{loc}}(\mathbb{R}^n \setminus \{0\}) \subseteq \mathcal{D}'(\mathbb{R}^n \setminus \{0\})$$

to a distribution $T \in \mathcal{D}'(\mathbb{R}^n)$ by setting

$$\langle T, \varphi \rangle := \int_{|x| \le 1} \frac{\varphi(x) - \varphi(0)}{|x|^n} f\Big(\frac{x}{|x|}\Big)\, dx + \int_{|x|>1} \frac{\varphi(x)}{|x|^n} f\Big(\frac{x}{|x|}\Big)\, dx.$$

That T is a distribution follows from the argument given after (1.2.6). Also, $T|_{\mathbb{R}^n \setminus \{0\}} = S$. Note, however, that there is no 'canonical' extension of S to $\mathbb{R}^n$. Changing 1 to R in the domains of integration results in an additional term proportional to $\varphi(0) = \langle \delta, \varphi \rangle$. Indeed, any extension of S to $\mathbb{R}^n$ can only be unique up to a distribution supported in $\{0\}$.

As in Example 1.2.7, T may be represented as a limit in $\mathcal{D}'(\mathbb{R}^n)$:

$$\begin{aligned}\langle T, \varphi \rangle &= \lim_{\varepsilon \to 0+} \left[\int_{\varepsilon \le |x| \le 1} \frac{\varphi(x) - \varphi(0)}{|x|^n} f\Big(\frac{x}{|x|}\Big)\, dx + \int_{|x|>1} \frac{\varphi(x)}{|x|^n} f\Big(\frac{x}{|x|}\Big)\, dx \right] \\ &= \lim_{\varepsilon \to 0+} \left[\langle Y(|x| - \varepsilon)|x|^{-n} f\Big(\frac{x}{|x|}\Big), \varphi \rangle - \varphi(0) \int_\varepsilon^1 \frac{dr}{r} \int_{S^{n-1}} f(\omega)\, d\sigma(\omega) \right],\end{aligned}$$

i.e.,

$$T = \lim_{\varepsilon \to 0+} \left[Y(|x| - \varepsilon)|x|^{-n} f\Big(\frac{x}{|x|}\Big) + \delta \log(\varepsilon) \int_{S^{n-1}} f(\omega)\, d\sigma(\omega) \right]. \tag{2.3.3}$$

Note that, if $\int_{S^{n-1}} f(\omega)\, d\sigma(\omega) = 0$, then $T = \mathrm{vp}(|x|^{-n} f\big(\frac{x}{|x|}\big))$.

In particular, for $n = 1$ and $f(\omega) := Y(\omega)$ we call the resulting distribution x_+^{-1}, since it is a distributional extension of $|x|^{-1} Y(\frac{x}{|x|}) = x^{-1} Y(x)$. By the above, we have

$$x_+^{-1} = \lim_{\varepsilon \to 0+} [Y(x - \varepsilon)x^{-1} + \log(\varepsilon)\delta]. \tag{2.3.4}$$

Note, however, that $x_+^{-1} \notin L^1_{\mathrm{loc}}(\mathbb{R})$ because $x^{-1}Y(x)$ is not locally integrable on $\mathbb{R}$.

(ii) Next we look at an example where extension is not possible. Let $S(x) := Y(x)e^{1/x} \in L^1_{\text{loc}}(\mathbb{R} \setminus \{0\})$. We claim that there does not exist any $T \in \mathcal{D}'(\mathbb{R})$ with $T|_{\mathbb{R}\setminus\{0\}} = S$. In fact, suppose there was such a T and let $\varphi \in \mathcal{D}(\mathbb{R})$ with $\operatorname{supp}(\varphi) \subseteq (0, \infty)$, $\varphi \geq 0$, and $\varphi(x) \geq 1$ for all $x \in [1, 2]$. Now set $\varphi_k(x) := e^{-k}\varphi(k^2 x)$. Then $\varphi_k \to 0$ in $\mathcal{D}(\mathbb{R})$ (but *not* in $\mathcal{D}(\mathbb{R} \setminus \{0\})$ because the supports of φ_k are not uniformly bounded in $\mathbb{R} \setminus \{0\}$). However,

$$\begin{aligned}\langle T, \varphi_k\rangle = \langle S, \varphi_k\rangle = e^{-k} \int_0^\infty e^{1/x}\varphi(k^2 x)\, dx &\geq e^{-k} \int_{k^{-2}}^{2k^{-2}} e^{1/x}\, dx \\ &\geq \frac{1}{k^2} \exp(-k + \frac{1}{2}k^2) \to \infty,\end{aligned}$$

a contradiction.

2.4 Composition with diffeomorphisms and submersions

Our first aim is to define the composition of a distribution with a diffeomorphism. As usual, we will insist on compatibility with the classical definition in the case of a regular distribution. Thus, let $h : \Omega_1 \to \Omega_2$ be a diffeomorphism of open subsets of $\mathbb{R}^n$ and let $f \in L^1_{\text{loc}}(\Omega_2)$. Then $f \circ h \in L^1_{\text{loc}}(\Omega_1)$, and for $\varphi \in \mathcal{D}(\Omega_1)$ we have

$$\begin{aligned}\langle T_{f\circ h}, \varphi\rangle &= \int_{\Omega_1} f(h(x))\varphi(x)\, dx = \int_{\Omega_2} f(y)\varphi(h^{-1}(y))|\det(Dh^{-1})(y)|\, dy \\ &= \left\langle T_f, \frac{\varphi}{|\det(Dh)|} \circ h^{-1} \right\rangle.\end{aligned} \tag{2.4.1}$$

We conclude that the only possible general definition is the following:

2.4.1 Definition. *Let $h : \Omega_1 \to \Omega_2$ be a diffeomorphism of open subsets of $\mathbb{R}^n$ and let $T \in \mathcal{D}'(\Omega_2)$. Then the composition $T \circ h \equiv h^*T \in \mathcal{D}'(\Omega_1)$ of T with h, also called the* pullback *of T under h, is defined by*

$$\langle T \circ h, \varphi\rangle := \left\langle T, \frac{\varphi}{|\det(Dh)|} \circ h^{-1} \right\rangle \quad (\varphi \in \mathcal{D}(\Omega_1)).$$

Note that $T \circ h$ is the transpose of the map $L : \mathcal{D}(\Omega_1) \to \mathcal{D}(\Omega_2)$, $L(\varphi) = \frac{\varphi}{|\det(Dh)|} \circ h^{-1}$, which clearly is sequentially continuous. Thus $T \circ h$ is indeed a distribution by Proposition 2.1.2.

2.4.2 Example. If $a \in \Omega_1$ and $b = h(a) \in \Omega_2$, then

$$\delta_b \circ h = |\det Dh(a)|^{-1}\delta_a.$$

In particular, if $A \in \mathrm{GL}(n, \mathbb{R})$, then $\delta \circ A = |\det A|^{-1}\delta$.

2.4.3 Definition. *Let $\Omega \subseteq \mathbb{R}^n$ be an open cone (i.e., $c \cdot \Omega \subseteq \Omega$ for each $c > 0$). Then $T \in \mathcal{D}'(\Omega)$ is called* homogeneous of degree $\lambda \in \mathbb{C}$, *if*

$$\forall c > 0 :\ T(cx) \equiv T \circ cI_n = c^\lambda T.$$

Explicitly, since $\langle T \circ cI_n, \varphi\rangle = c^{-n}\langle T(x), \varphi\big(\frac{x}{c}\big)\rangle$, a distribution $T \in \mathcal{D}'(\Omega)$ is homogeneous of degree λ if and only if

$$\langle T, \varphi\rangle = c^{n+\lambda}\langle T(x), \varphi(cx)\rangle, \quad c > 0,\ \varphi \in \mathcal{D}(\Omega). \tag{2.4.2}$$

2.4.4 Examples.

(i) By what we have seen in Example 2.4.2, δ is homogeneous of degree $-n$, i.e., $\delta(cx) = c^{-n}\delta$.

(ii) If $f \in L^1(S^{n-1})$ satisfies the mean-value zero condition (1.2.5), then $T = \mathrm{vp}(|x|^{-n} f\big(\frac{x}{|x|}\big)) \in \mathcal{D}'(\mathbb{R}^n)$ is homogeneous of degree $-n$. Indeed,

$$\begin{aligned}\langle T \circ (cI_n), \varphi\rangle &= \langle T, \varphi(x/c)c^{-n}\rangle = c^{-n} \lim_{\varepsilon\to 0+} \int_{|x|\geq\varepsilon} \frac{1}{|x|^n} \varphi\Big(\frac{x}{c}\Big) f\Big(\frac{x}{|x|}\Big)\, dx \\ &= \lim_{\varepsilon\to 0+} \int_{|y|\geq \varepsilon/c} \frac{\varphi(y)}{c^n |y|^n} f\Big(\frac{y}{|y|}\Big)\, dy = c^{-n}\langle T, \varphi\rangle.\end{aligned}$$

(iii) Again let $f \in L^1(S^{n-1})$, but this time suppose that (1.2.5) is violated, so $\int_{S^{n-1}} f(\omega)\, d\omega \neq 0$. Then $|x|^{-n} f\big(\frac{x}{|x|}\big) \in L^1_{\mathrm{loc}}(\mathbb{R}^n \setminus \{0\}) \subseteq \mathcal{D}'(\mathbb{R}^n \setminus \{0\})$ is still homogeneous of degree $-n$, as follows exactly as in (ii). However, the extension constructed in Example 2.3.10 (i) is *not* homogeneous in $\mathcal{D}'(\mathbb{R}^n)$. In fact, if $c > 0$ then using (2.4.1) and Example 2.4.4 (i), from (2.3.3) we get

$$\begin{aligned}T(cx) &= \lim_{\varepsilon\to 0+} \Big[c^{-n} Y(c|x| - \varepsilon)|x|^{-n} f\Big(\frac{x}{|x|}\Big) + c^{-n}\log(\varepsilon) \int_{S^{n-1}} f(\omega)\, d\sigma(\omega)\, \delta \Big] \\ &= c^{-n} \lim_{\varepsilon\to 0+} \Big[Y\Big(|x| - \frac{\varepsilon}{c}\Big)|x|^{-n} f\Big(\frac{x}{|x|}\Big) + \log(\varepsilon/c) \int_{S^{n-1}} f(\omega)\, d\sigma(\omega)\, \delta \\ &\qquad + \log(c) \int_{S^{n-1}} f(\omega)\, d\sigma(\omega)\, \delta \Big].\end{aligned}$$

Therefore, again by (2.3.3),

$$T(cx) = c^{-n} T + c^{-n} \log(c) \int_{S^{n-1}} f(\omega)\, d\sigma(\omega)\, \delta. \tag{2.4.3}$$

2.4.5 Example. Here we look at the composition of principal value distributions with linear maps.

Note first that if $T \in \mathcal{D}'(\mathbb{R}^n)$ is homogeneous of degree λ then so is $T \circ A$ for any $A \in \mathrm{GL}(n, \mathbb{R})$. Indeed, for any $c > 0$ we have

$$(T \circ A) \circ (cI_n) = (T \circ (cI_n)) \circ A = c^\lambda (T \circ A).$$

As we have seen in Example 2.4.4 (ii), if $f \in L^1(S^{n-1})$ satisfies the mean-value zero condition (1.2.5), then $T = \mathrm{vp}(|x|^{-n} f\big(\frac{x}{|x|}\big)) \in \mathcal{D}'(\mathbb{R}^n)$ is homogeneous of degree $-n$, so the same is true for $T \circ A$. As was shown in Example 2.3.10 (i), on $\mathbb{R}^n \setminus \{0\}$ the distribution T coincides with the L^1_{loc}-function $|x|^{-n} f\big(\frac{x}{|x|}\big)$. Therefore,

$$T \circ A|_{\mathbb{R}^n\setminus\{0\}} = |Ax|^{-n} f\Big(\frac{Ax}{|Ax|}\Big) \in L^1_{\mathrm{loc}}(\mathbb{R}^n \setminus \{0\}),$$

so $T \circ A$ coincides with $S := \mathrm{vp}(|Ax|^{-n} f\Big(\frac{Ax}{|Ax|}\Big))$ outside the origin. To determine the difference $T \circ A - S$ in $\mathcal{D}'(\mathbb{R}^n)$, let $\varphi \in \mathcal{D}(\mathbb{R}^n)$. Then

$$
\begin{aligned}
&\langle T \circ A - S, \varphi \rangle \\
&\quad = \lim_{\varepsilon \to 0+} \Big\langle \Big(Y(|x| - \varepsilon)|x|^{-n} f\Big(\frac{x}{|x|}\Big)\Big) \circ A - Y(|x| - \varepsilon)|Ax|^{-n} f\Big(\frac{Ax}{|Ax|}\Big), \varphi \Big\rangle \\
&\quad = \lim_{\varepsilon \to 0+} \int_{\mathbb{R}^n} |Ax|^{-n} f\Big(\frac{Ax}{|Ax|}\Big)(Y(|Ax| - \varepsilon) - Y(|x| - \varepsilon))\varphi(x)\, dx \\
&\quad = \lim_{\varepsilon \to 0+} |\det A|^{-1} \int_{\mathbb{R}^n} |y|^{-n} f\Big(\frac{y}{|y|}\Big)(Y(|y| - \varepsilon) - Y(|A^{-1}y| - \varepsilon))\varphi(A^{-1}y)\, dy \\
&\quad = \lim_{\varepsilon \to 0+} |\det A|^{-1} \times \\
&\qquad \times \int_{S^{n-1}} f(\omega) \int_0^\infty (Y(r - \varepsilon) - Y(r|A^{-1}\omega| - \varepsilon))\varphi(rA^{-1}\omega) \frac{dr}{r}\, d\sigma(\omega).
\end{aligned}
$$

Here,

$$
\begin{aligned}
\int_0^\infty (Y(r - \varepsilon) - Y(r|A^{-1}\omega| - \varepsilon))\varphi(rA^{-1}\omega)\frac{dr}{r} &= \int_\varepsilon^{\varepsilon/|A^{-1}\omega|} \varphi(rA^{-1}\omega)\frac{dr}{r} \\
&= \int_\varepsilon^{\varepsilon/|A^{-1}\omega|} \varphi(0)\frac{dr}{r} + O(\varepsilon) = -\varphi(0)\log|A^{-1}\omega| + O(\varepsilon),
\end{aligned}
$$

by Taylor expansion. Letting $\varepsilon \to 0+$, we arrive at

$$
T \circ A = \mathrm{vp}\Big(|Ax|^{-n} f\Big(\frac{Ax}{|Ax|}\Big)\Big) - \frac{\delta}{|\det A|} \int_{S^{n-1}} f(\omega) \log|A^{-1}\omega|\, d\sigma(\omega).
$$

For $A = cI$ this reduces to (2.4.3), as it should.

Next we turn to the problem of composing distributions with maps that are more general than diffeomorphisms. As it turns out, this is indeed possible for *submersions*, i.e., maps whose differential is surjective at each point. For $h : \Omega \to \mathbb{R}$ this means that $\nabla h(x) = (\partial_1 h(x), \dots, \partial_n h(x)) \neq 0$ for all $x \in \Omega$. In this case, we want to define the pullback under h, $h^* : \mathcal{D}'(\mathbb{R}) \to \mathcal{D}'(\Omega)$ as the unique continuous extension of $h^* : \mathcal{D}(\mathbb{R}) \to \mathcal{D}(\Omega)$, $f \mapsto f \circ h$.
To see how to go about this, let $f \in \mathcal{D}(\mathbb{R})$, $\varphi \in \mathcal{D}(\Omega)$. Then

$$
\begin{aligned}
\langle T_{h^*f}, \varphi \rangle &= \langle T_{f\circ h}, \varphi \rangle = \int_\Omega f(h(x))\varphi(x)\, dx = -\int_\Omega \Big(\int_{h(x)}^\infty f'(s)\, ds\Big)\varphi(x)\, dx \\
&= -\int_{\{(x,s)\in\Omega\times\mathbb{R} | h(x)<s\}} f'(s)\varphi(x)\, dxds = -\int_{-\infty}^\infty f'(s)\Big(\int_{\{x|h(x)<s\}} \varphi(x)\, dx\Big) ds \\
&= \int_{-\infty}^\infty f(s)\frac{d}{ds}\Big(\int_{\{x|h(x)<s\}} \varphi(x)\, dx\Big) dt =: \int_{-\infty}^\infty f(s)\varphi_h(s)\, ds = \langle f, \varphi_h \rangle,
\end{aligned}
$$

where

$$
\varphi_h(s) := \frac{d}{ds}\int_\Omega Y(s - h(x))\varphi(x)\, dx = \frac{d}{ds}\int_{\{x|h(x)<s\}} \varphi(x)\, dx. \tag{2.4.4}
$$

This motivates:

2.4.6 Definition. *Let $h : \Omega \to \mathbb{R}$ be a smooth submersion, i.e., $\nabla h(x) \neq 0$ for each $x \in \Omega$. Moreover, let $T \in \mathcal{D}'(\mathbb{R})$. Then the pullback h^*T (also called the composition $T \circ h$) of T with h is defined by*

$$
\langle h^*T, \varphi \rangle = \langle T, \varphi_h \rangle \equiv \Big\langle T(s), \frac{d}{ds}\int_\Omega Y(s - h(x))\varphi(x)\, dx \Big\rangle.
$$

The following result shows, among others, that indeed h^*T is a distribution on Ω.

2.4.7 Proposition. *The map $h^* : \mathcal{D}'(\mathbb{R}) \to \mathcal{D}'(\Omega)$, $T \mapsto h^*T = T \circ h$ is well defined and is the unique sequentially continuous extension of the map $h^* : \mathcal{D}(\mathbb{R}) \to \mathcal{D}(\Omega)$, $f \mapsto f \circ h$. Moreover, if $T = T_f$ for some $f \in L^1_{\mathrm{loc}}(\mathbb{R})$, then $T \circ h = T_{f\circ h}$.*

Proof. Let $y \in \Omega$. Then $\nabla h(y) \neq 0$, and without loss of generality we may assume that $\partial_n h(y) \neq 0$. Let $F_y := (x_1, \ldots, x_n) \mapsto \xi = (\xi_1, \ldots, \xi_n) := (x_1, \ldots, x_{n-1}, h(x))$. Then

$$DF_y(y) = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 \\ * & * & * & \ldots & * & \partial_n h(y) \end{pmatrix},$$

so by the inverse function theorem there exists a neighborhood U_y of y such that $F_y : U_y \to F_y(U_y)$ is a diffeomorphism. Let $G_y := F_y^{-1}$. Given $\varphi \in \mathcal{D}(U_y)$, define φ_h by (2.4.4). Substituting $x = G_y(\xi)$ we get

$$\begin{aligned} \varphi_h(s) &= \frac{d}{ds} \int_{\xi_n < s} \varphi(G_y(\xi)) |\det DG_y(\xi)|\, d\xi \\ &= \int \varphi \circ G_y(\xi', s) |\det DG_y(\xi', s)|\, d\xi', \end{aligned} \tag{2.4.5}$$

where $\xi' = (\xi_1, \ldots, \xi_{n-1})$. Note that the integrand in the last integral is an element of $\mathcal{D}(\mathbb{R}^n)$. Consequently, $\varphi_h \in \mathcal{D}(\mathbb{R})$.

Now given any $\varphi \in \mathcal{D}(\Omega)$, we may cover $\operatorname{supp}(\varphi)$ by finitely many U_{y_j} $(1 \le j \le N)$ as above. For this covering, we choose a subordinate partition of unity $\{\psi_j \mid 1 \le j \le N\}$ as in Proposition 2.2.2. Then $\varphi = \sum_{j=1}^N \psi_j \varphi$, and so

$$\varphi_h(s) = \frac{d}{ds} \int_{\{x \mid h(x) < s\}} \varphi(x)\, dx = \sum_{j=1}^N \frac{d}{ds} \int_{\{x \mid h(x) < s\}} (\psi_j \varphi)(x)\, dx = \sum_{j=1}^N (\psi_j \varphi)_h(s),$$

so $\varphi_h \in \mathcal{D}(\mathbb{R})$.

The map $\varphi \mapsto \varphi_h$ is obviously linear. It is also sequentially continuous, being a finite sum of compositions of the following operations:

$$\varphi \mapsto \varphi\psi_j \mapsto (\varphi\psi_j)|_{U_{y_j}} \circ G_{y_j} \mapsto ((\varphi\psi_j)|_{U_{y_j}} \circ G_{y_j}) |\det DG_{y_j}|,$$

with G_{y_j} defined analogously as above, for $\partial_i F_{y_j} \neq 0$, followed by $I_i : \mathcal{D}(\Omega) \to \mathcal{D}(\mathbb{R})$,

$$I_i(\psi)(t) = \int_{\mathbb{R}^{n-1}} \psi(\xi_1, \ldots, \xi_{i-1}, t, \xi_{i+1}, \ldots, \xi_n)\, d\xi_1 \ldots d\xi_{i-1}\xi_{i+1} \ldots d\xi_n.$$

It follows that $h^*T \in \mathcal{D}'(\Omega)$. Moreover, by Proposition 2.1.2, the map $h^* : T \mapsto h^*T$, being the adjoint of $\varphi \mapsto \varphi_h$, is sequentially continuous: $\mathcal{D}'(\mathbb{R}) \to \mathcal{D}'(\Omega)$. We have already seen above that $h^*|_{\mathcal{D}(\mathbb{R})} = f \mapsto f \circ h$, so the uniqueness claim follows from the fact that $\mathcal{D}(\mathbb{R})$ is dense in $\mathcal{D}'(\mathbb{R})$ (by Theorem 1.1.9).

Finally, let $f \in L^1_{\mathrm{loc}}(\Omega)$. To show that $\langle h^*T_f, \varphi \rangle = \langle T_{f\circ h}, \varphi \rangle$, using a partition of unity on a neighborhood of $\operatorname{supp}(\varphi)$ as above, we may reduce to the case $\Omega = U_y$. Then by (2.4.5),

$$\begin{aligned} \langle h^*T_f, \varphi \rangle &= \int f(s)\varphi_h(s)\, ds = \int \int f(s) \varphi \circ G_y(\xi', s) |\det DG_y(\xi', s)|\, d\xi' ds \\ &= \int_\Omega f(h(x))\varphi(x)\, dx = \langle T_{f\circ h}, \varphi \rangle, \end{aligned}$$

where the penultimate equality follows by substituting $x = G_y(\xi', s)$. □

2.4.8 Example. (*Single layer distributions*) An important example of the pullback of a distribution under a submersion is given by $T \circ h$, for $T = \delta_a \in \mathcal{D}'(\mathbb{R})$, $h : \Omega \to \mathbb{R}$ smooth, $a \in h(\Omega)$, where h is a submersion in a neighborhood of $M = h^{-1}(a)$. Note that M is a smooth submanifold of $\mathbb{R}^n$ (a hypersurface), and it can naturally be equipped with the restriction g of the standard Euclidean metric $\sum_{i=1}^n dx_i \otimes dx_i$ of $\mathbb{R}^n$ to M. The metric g gives rise to a surface measure $d\sigma$ by setting

$$\int_M \varphi(x)\, d\sigma(x) = \int_U \varphi(x(u)) \sqrt{\det(g_{jk}(u))}\, du.$$

Here, $\varphi \in C(M)$ is such that $\operatorname{supp} \varphi$ lies in a coordinate patch parametrized by $x = x(u)$, $u \in U \subseteq \mathbb{R}^{n-1}$, and $g = \sum_{j,k=1}^{n-1} g_{jk}(u) du_j \otimes du_k$.

Now for $f : M \to \mathbb{C}$ locally integrable with respect to $d\sigma$, the *single layer distribution* $S_M(f) \in \mathcal{D}'(\Omega)$ with density f is defined by

$$\langle S_M(f), \varphi \rangle := \int_M f(x)\varphi(x)\, d\sigma(x) \qquad (\varphi \in \mathcal{D}(\Omega)). \tag{2.4.6}$$

Note that $S_M(f)$ is a well-defined distribution in $\mathcal{D}'(\Omega)$ by Theorem 1.1.7, because $M \cap \operatorname{supp} \varphi$ is compact, and $|\langle S_M(f), \varphi \rangle| \le \|\varphi\|_{\infty,K} \|f\|_{L^1(K\cap M)}$ for $\varphi \in \mathcal{D}(K)$ (where the L^1-norm is with respect to $d\sigma$).

We now want to calculate $\delta_a \circ h$ and express it as a single layer distribution. To get an intuition of what to expect, note first that since $\Gamma(3/2) = \frac{1}{2}\sqrt{\pi}$, by (1.2.2) we have

$$\delta_a = \lim_{\varepsilon \to 0+} \frac{1}{2\varepsilon} Y(\varepsilon - |s - a|)$$

in $\mathcal{D}'(\mathbb{R}_s)$. By Proposition 2.4.7 we therefore obtain

$$\delta_a \circ h = \lim_{\varepsilon \to 0+} \frac{1}{2\varepsilon} Y(\varepsilon - |h(x) - a|) \tag{2.4.7}$$

in $\mathcal{D}'(\Omega)$. Heuristically, this means that $\delta_a \circ h$ is the limit of constant mass densities on the layers $\{x \in \Omega \mid a - \varepsilon < h(x) < a + \varepsilon\}$. Any such layer has the approximate width $\frac{2\varepsilon}{|\nabla h(x)|}$: In fact, for $x_0 \in M$ and x close to x_0,

$$h(x) \approx h(x_0) + \nabla h(x_0)(x - x_0) = a + \nabla h(x_0)(x - x_0),$$

so $|x - x_0| < \frac{\varepsilon}{|\nabla h(x_0)|}$. Combining this with (2.4.7), we therefore conjecture that

$$\langle \delta_a \circ h, \varphi \rangle = \int_M \frac{\varphi(x)}{|\nabla h(x)|}\, d\sigma(x) \qquad (\varphi \in \mathcal{D}(\Omega)). \tag{2.4.8}$$

We now want to prove (2.4.8) formally. To this end, as in the proof of Proposition 2.4.7, using a partition of unity we may reduce to the case where $F_y := (x_1, \dots, x_n) \mapsto (x_1, \dots, x_{n-1}, h(x))$ is a diffeomorphism on all of Ω. This implies that $x_1, \dots, x_{n-1}$ are local coordinates for M: in fact, $M = h^{-1}(a)$, so a chart for M is then given by

$$M \ni x \mapsto F_y(x) = (x_1, \dots, x_{n-1}, h(x)) = (x_1, \dots, x_{n-1}, a) \mapsto (x_1, \dots, x_{n-1}).$$

Now h is constant on M, so

$$0 = dh = \sum_{j=1}^n \partial_j h\, dx_j \Rightarrow dx_n = -\sum_{j=1}^{n-1} \frac{\partial_j h}{\partial_n h} dx_j$$

on M. Hence the metric g on M can be written in the form

$$g = \Big(\sum_{l=1}^{n} dx_l \otimes dx_l\Big)\Big|_M = \sum_{j,k=1}^{n-1} (\delta_{jk} + \frac{\partial_j h \cdot \partial_k h}{(\partial_n h)^2}) dx_j \otimes dx_k =: \sum_{j,k=1}^{n-1} g_{jk} dx_j \otimes dx_k.$$

In other words,

$$(g_{jk})_{j,k=1}^{n-1} = I_{n-1} + vv^\top \quad \text{with} \;\; v = \frac{1}{\partial_n h}\begin{pmatrix} \partial_1 h \\ \vdots \\ \partial_{n-1} h \end{pmatrix}.$$

From this it follows that $\det((g_{jk})_{j,k=1}^{n-1}) = 1 + |v|^2$: in fact, extending v to a basis of $\mathbb{R}^{n-1}$ by vectors w_i with $w_i \perp v$ for $i = 1, \dots, n-2$, the map $I_{n-1} + vv^\top : \mathbb{R}^{n-1} \to \mathbb{R}^{n-1}$ has the matrix

$$\begin{pmatrix} 1+|v|^2 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Consequently,

$$\det((g_{jk})_{j,k=1}^{n-1}) = 1 + |v|^2 = \frac{|\nabla h|^2}{(\partial_n h)^2}.$$

As in the proof of Proposition 2.4.7 we may calculate, for any $\varphi \in \mathcal{D}(\Omega)$:

$$\langle \delta_a \circ h, \varphi \rangle = \langle \delta_a, \varphi_h \rangle = \varphi_h(a).$$

Now by (2.4.5), we have, setting $\xi' := (\xi_1, \dots, \xi_{n-1})$,

$$\begin{aligned} \varphi_h(a) &= \int \frac{\varphi(G_y(\xi', a))}{|(\det DF_y)(G_y(\xi', a))|} \, d\xi' = \int \frac{\varphi(G_y(\xi', a))}{|(\partial_n h)(G_y(\xi', a))|} \, d\xi' \\ &= \int \frac{\varphi(G_y(\xi', a))}{|\nabla h|} \cdot \sqrt{\det(g_{jk})} \, d\xi' = \int_M \frac{\varphi(x)}{|\nabla h(x)|} \, d\sigma(x), \end{aligned}$$

concluding the proof of (2.4.8). In short-hand form,

$$\delta_a \circ h = \delta(h(x) - a) = S_M(|\nabla h|^{-1}). \tag{2.4.9}$$

We also note that $\operatorname{supp}(\delta_a \circ h) = M$.

2.4.9 Examples. Special cases of single layer distributions:

(i) For $n = 1$, a function $h : \mathbb{R} \supseteq \Omega \to \mathbb{R}$ is submersive if and only if $h'(x) \neq 0$ for each $x \in M = h^{-1}(a)$, which now must be a discrete set in Ω (i.e., has no accumulation points). Indeed, if x_0 were an accumulation point of M, then there would be $x_k \in M$, $x_k \neq x_0$, $x_k \to x$. But then $h'(x_0) = \lim_{k\to\infty} \frac{h(x_0)-h(x_k)}{x_0-x_k} = 0$, a contradiction. Using a partition of unity, we may reduce the calculation of $\langle h^*\delta_a, \varphi \rangle$ to the case where in fact $M = h^{-1}(a) = \{x_0\}$ consists only of a single point. In the notation from Proposition 2.4.7, we have $F_y = x \mapsto h(x)$, and $G_y = F_y^{-1} = h^{-1}$ (which exists since $h' \neq 0$ everywhere). By (2.4.5) we have

$$\varphi_h(s) = \frac{d}{ds} \int_{\xi < s} \varphi(h^{-1}(\xi)) |(h^{-1})'(\xi)| \, d\xi = \frac{\varphi(h^{-1}(s))}{|h'(h^{-1}(s))|}, \tag{2.4.10}$$

so

$$\langle h^*\delta_a, \varphi \rangle = \varphi_h(a) = \frac{\varphi(h^{-1}(a))}{|h'(h^{-1}(a))|} = \frac{\varphi(x_0)}{|h'(x_0)|} = \frac{1}{|h'(x_0)|} \langle \delta_{x_0}, \varphi \rangle$$

in this case. In general, we have to sum over all $x \in h^{-1}(a)$, so we arrive at

$$h^*\delta_a = \delta_a \circ h := \sum_{x \in h^{-1}(a)} \frac{1}{|h'(x)|}\delta_x \in \mathcal{D}'(\Omega).$$

Concrete examples are:

$$\begin{aligned}
\delta_a(x^2) = \delta(x^2 - a) &= \frac{1}{2\sqrt{a}}(\delta_{\sqrt{a}} + \delta_{-\sqrt{a}}) \in \mathcal{D}'(\mathbb{R} \setminus \{0\}) \qquad (a > 0) \\
\delta \circ \sin = \delta(\sin x) &= \sum_{k \in \mathbb{Z}} \delta_{k\pi} \in \mathcal{D}'(\mathbb{R}) \\
\delta(\sin(1/x)) &= \frac{1}{\pi^2} \sum_{k \in \mathbb{Z}\setminus\{0\}} \frac{1}{k^2}\delta_{1/(k\pi)} \in \mathcal{D}'(\mathbb{R} \setminus \{0\}).
\end{aligned}$$

(ii) Let h be a positive definite quadratic form, $h(x) = x^\top Cx$ with $C = C^\top \in \mathrm{GL}(n, \mathbb{R})$ positive definite. Via diagonalization and multiplication by suitable diagonal matrices it follows that there exists some $A \in \mathrm{GL}(n, \mathbb{R})$ with $A^\top CA = I_n$, so $h(Ay) = y^\top A^\top CAy = |y|^2$. Therefore,

$$\begin{aligned}
\langle \delta(x^\top Cx - 1), \varphi\rangle &= \langle \delta_1 \circ h, \varphi\rangle = \langle \delta_1 \circ h \circ A \circ A^{-1}, \varphi\rangle \\
&= |\det A| \langle \delta_1(|y|^2), \varphi \circ A\rangle = \frac{|\det A|}{2} \int_{S^{n-1}} \varphi(A\omega) d\sigma(\omega),
\end{aligned}$$

where we used Definition 2.4.1, (2.4.9), and the fact that

$$|\nabla(|y|^2)| = 2|y| = 2$$

for $y = \omega \in S^{n-1}$. Note also that $|\det A| = (\det C)^{-1/2}$.

As a concrete example, let $h(x) := \sum_{i=1}^3 \frac{x_i^2}{a_i^2}$, $x \in \mathbb{R}^3$, $a_i \in (0, \infty)$ for $i = 1, 2, 3$, so the level sets of h are ellipsoids. In this case, $A = \mathrm{diag}(a_1, a_2, a_3)$, so that

$$\begin{aligned}
\langle \delta_1 \circ h, \varphi\rangle &= \left\langle \delta\Big(1 - \sum_{i=1}^3 \frac{x_i^2}{a_i^2}\Big), \varphi \right\rangle \\
&= \frac{a_1 a_2 a_3}{2} \int_0^\pi \int_0^{2\pi} \varphi(a_1 \cos\varphi \sin\theta, a_2 \sin\varphi \sin\theta, a_3 \cos\theta) \sin\theta d\varphi d\theta.
\end{aligned}$$

2.4.10 Remark. It follows from (2.4.10) that in case $h : \Omega \to \Omega'$ is a diffeomorphism between open subsets of $\mathbb{R}$ and $T \in \mathcal{D}'(\Omega')$, Definitions 2.4.1 and 2.4.6 for $T \circ h$ coincide.

Chapter 3

Differentiation

3.1 Definition and basic properties

A main motivation for the introduction of spaces of distributions is the wish to assign partial derivatives to functions that are merely continuous. In this section we shall see that it is indeed possible to define such operations for arbitrary distributions while preserving 'backwards compatibility' with respect to subspaces of $\mathcal{D}'$ consisting of functions that are differentiable in the classical sense.

We want to define a notion of differentiation in $\mathcal{D}'$ that is compatible with the classical derivative of, say, a C^1-function. More precisely, let $f \in C^1(\Omega) \subseteq L^1_{\text{loc}}(\Omega) \subseteq \mathcal{D}'(\Omega)$. We wish to achieve that $\partial_j^{\text{new}}(T_f) = T_{\partial_j f}$ holds, which requires the following diagram to be commutative:

$$\begin{tikzcd} C^1 \arrow[r, hook] \arrow[d, "\partial_j"'] & \mathcal{D}' \arrow[d, "\partial_j^{\text{new}}"] \\ C^0 \arrow[r, hook] & \mathcal{D}' \end{tikzcd}$$

In terms of the action on a test function $\varphi \in \mathcal{D}(\Omega)$, this means that

$$\langle \partial_j^{\text{new}}(T_f), \varphi \rangle \stackrel{!}{=} \langle T_{\partial_j f}, \varphi \rangle = \int \partial_j f(x)\varphi(x)\,dx = -\int f(x)\partial_j\varphi(x)\,dx = -\langle T_f, \partial_j\varphi \rangle,$$

where we employed integration by parts. In this way, we are naturally led to the following

3.1.1 Definition. *Let $T \in \mathcal{D}'(\Omega)$ and $1 \le j \le n$. We define the* partial derivative *$\partial_j T$ of T by*

$$\langle \partial_j T, \varphi \rangle := -\langle T, \partial_j \varphi \rangle \qquad \forall \varphi \in \mathcal{D}(\Omega).$$

If $n = 1$ we denote the derivative by T' instead of $\partial_1 T$. By iteration, we obtain the general formula

$$\langle \partial^\alpha T, \varphi \rangle = (-1)^{|\alpha|}\langle T, \partial^\alpha \varphi \rangle \qquad (\alpha \in \mathbb{N}_0^n).$$

Let us now examine some basic properties of differentiation on the space of distributions:

3.1.2 Proposition. *The map $\partial^\alpha : \mathcal{D}'(\Omega) \to \mathcal{D}'(\Omega)$ is well defined, linear, and sequentially continuous. If $f \in C^m(\Omega)$ and $|\alpha| \le m$, then $\partial^\alpha T_f = T_{\partial^\alpha f}$. Moreover,*

$$\partial_j T = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}(T(x_1, \ldots, x_j + \varepsilon, \ldots, x_n) - T(x)),$$

where $T(x_1, \dots, x_j + \varepsilon, \dots, x_n)$ is defined as $T \circ h_\varepsilon$ with h the diffeomorphism $x \mapsto (x_1, \dots, x_j + \varepsilon, \dots, x_n)$.

Proof. The first claims follow from Proposition 2.1.2, together with the fact that the map

$$\mathcal{D}(\Omega) \to \mathcal{D}(\Omega), \qquad \varphi \mapsto (-1)^{|\alpha|}\partial^\alpha \varphi$$

clearly is linear and sequentially continuous. Compatibility with classical derivatives on $C^m(\Omega)$ is built into our definition. For the final claim, it will suffice to consider $j = 1$. Since $\det Dh = 1$, Definition 2.4.1 implies that

$$\frac{1}{\varepsilon}\langle T(x_1 + \varepsilon, x') - T(x), \varphi\rangle = \left\langle T, \frac{\varphi(x_1 - \varepsilon, x') - \varphi(x)}{\varepsilon} \right\rangle$$

(where $x' = (x_2, \dots, x_n)$). Now

$$\lim_{\varepsilon \to 0} \frac{\varphi(x_1 - \varepsilon, x') - \varphi(x)}{\varepsilon} = -\lim_{\varepsilon \to 0} \int_0^1 \partial_1 \varphi(x_1 - \varepsilon t, x')\, dt = -\partial_1 \varphi(x),$$

and this limit even holds in $\mathcal{D}(\Omega)$. □

3.1.3 Remark. Completely analogously to the last part of the above proof, by transposition, we obtain the following result on the *directional derivative* of a distribution $T \in \mathcal{D}'(\Omega)$ in the direction of $v \in \mathbb{R}^n$:

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon}(T(x + \varepsilon v) - T(x)) = \sum_{j=1}^n v_j \cdot \partial_j T. \tag{3.1.1}$$

3.1.4 Examples.
(i) For Y the Heaviside function on $\mathbb{R}$ and $\varphi \in \mathcal{D}(\mathbb{R})$ we have

$$\langle Y', \varphi\rangle = -\langle Y, \varphi'\rangle = -\int_0^\infty \varphi'(x)\, dx = -0 + \varphi(0) = \varphi(0) = \langle \delta, \varphi\rangle.$$

Therefore in $\mathcal{D}'(\mathbb{R})$ we obtain the important result

$$Y' = \delta.$$

(ii) *Derivative of a jump.* Let $f \in C^\infty(\mathbb{R})$. Then $f \cdot Y \in L^1_{\mathrm{loc}}(\mathbb{R}) \subseteq \mathcal{D}'(\mathbb{R})$ and we have

$$\begin{aligned}
\langle (T_{fY})', \varphi\rangle &= -\langle T_{fY}, \varphi'\rangle = -\int_0^\infty f(x)\varphi'(x)\, dx \\
&= f(0)\varphi(0) + \int_0^\infty f'(x)\varphi(x)\, dx = f(0)\langle \delta, \varphi\rangle + \int_{-\infty}^\infty Y(x) f'(x)\varphi(x)\, dx \\
&= \langle f(0)\delta + T_{f'Y}, \varphi\rangle.
\end{aligned}$$

Thus we obtain the first instance of a *jump formula*

$$(fY)' := (T_{fY})' = f(0)\delta + f'Y \tag{3.1.2}$$

Further important properties are collected in the following result:

3.1.5 Proposition. *Let $T \in \mathcal{D}'(\Omega)$. Then:*

(i) For $\alpha, \beta \in \mathbb{N}_0^n$, $\partial^\alpha \partial^\beta T = \partial^\beta \partial^\alpha T = \partial^{\alpha+\beta} T$.

(ii) For $\alpha \in \mathbb{N}_0^n$, $f \in C^\infty(\Omega)$, $\partial^\alpha(f \cdot T) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \partial^\beta f \cdot \partial^{\alpha-\beta} T$ (Leibniz rule).

(iii) If $T \in \mathcal{D}'(\mathbb{R})$ and $h : \Omega \to \mathbb{R}$ is a submersion, then $\partial_j(T \circ h) = \frac{\partial h}{\partial x_j} \cdot (T' \circ h)$ (chain rule).

Proof. (i) follows directly from the definition.
(ii) By induction, it suffices to prove the case $\partial^\alpha = \partial_j$. Here we have

$$\begin{aligned}\langle \partial_j(fT), \varphi \rangle = -\langle fT, \partial_j \varphi \rangle = -\langle T, f\, \partial_j \varphi \rangle &= -\langle T, \partial_j(f\varphi) - (\partial_j f)\varphi \rangle \\ &= \langle \partial_j T, f\varphi \rangle + \langle (\partial_j f) T, \varphi \rangle = \langle f\, \partial_j T + (\partial_j f)\, T, \varphi \rangle.\end{aligned}$$

(iii) If $T = T_f$ for some $f \in \mathcal{D}(\mathbb{R})$, then this follows from the classical chain rule. Since $\mathcal{D}(\mathbb{R})$ is sequentially dense in $\mathcal{D}'(\mathbb{R})$ by Proposition 1.1.9, the result follows from Proposition 2.4.7. □

Finally, we note the following useful fact on the differentiation of homogeneous distributions:

3.1.6 Lemma. *Let $T \in \mathcal{D}'(\Omega)$ be homogeneous of degree λ, and let $\alpha \in \mathbb{N}_0^n$. Then $\partial^\alpha T$ is homogeneous of degree $\lambda - |\alpha|$.*

Proof. We use (2.4.2):

$$\begin{aligned}\langle \partial^\alpha T, \varphi \rangle &= (-1)^{|\alpha|} \langle T, \partial^\alpha \varphi \rangle = c^{n+\lambda} (-1)^{|\alpha|} \langle T(x), (\partial^\alpha \varphi)(cx) \rangle \\ &= c^{n+\lambda} (-1)^{|\alpha|} \langle T(x), c^{-|\alpha|} \partial^\alpha (\varphi(cx)) \rangle = c^{n+\lambda-|\alpha|} \langle \partial^\alpha T(x), \varphi(cx) \rangle.\end{aligned}$$

□

3.2 Examples and applications

In this section we look at several more elaborate examples of differentiation of distributions, as well as at first applications to differential equations.

3.2.1 Example. *Differentiation of sequences and series*
The rather innocuous statement on the sequential continuity of differentiation in $\mathcal{D}'(\Omega)$ (Proposition 3.1.2) is a first indication on the differences between the rules of calculation in classical functions versus distributions. Indeed, it implies that any sequence (or, which amounts to the same, series) converging in $\mathcal{D}'$ can be differentiated term-wise, in stark contrast to classical analysis.
As a concrete example, let $f_k \in C^\infty(\mathbb{R}) \subseteq \mathcal{D}'(\mathbb{R})$ ($k \in \mathbb{N}$) be given by

$$f_k(x) = \frac{1}{\sqrt{k}} \sin(kx).$$

Since $\|f_k\|_{L^\infty(\mathbb{R})} = 1/\sqrt{k}$, $f_k \to 0$ uniformly, hence also $f_k \to 0$ in $\mathcal{D}'(\mathbb{R})$.
The derivatives are $f_k'(x) = \sqrt{k}\, \cos(kx)$, thus (f_k') does not even converge pointwise to 0 [if $\cos(kx) \to 0$ ($k \to \infty$), then $1 = \lim(1 + \cos(2kx)) = \lim 2\cos^2(kx) = 0$]. Nevertheless, we know that $f_k' \to 0$ in $\mathcal{D}'(\Omega)$ by the continuity of the distributional derivative.

3.2.2 Example. Consider the Cauchy principal value $\mathrm{vp}(\frac{1}{x})$ from Example 1.2.7. We claim that

$$(\log|x|)' = \mathrm{vp}(\frac{1}{x}). \qquad (3.2.1)$$

Here $\log|x|$ is to be understood as the regular distribution $\varphi \mapsto \int \log|x|\,\varphi(x)\,dx$, and the derivative is in the $\mathcal{D}'$-sense. Note that $\log|.| \in L^1_{\mathrm{loc}}(\mathbb{R})$, since

$$\int_0^1 |\log(x)|\,dx = -(x\log(x) - x)|_0^1 = 1.$$

We can compare the above distributional formula to the classical statement that for $x > 0$ we have $\log'(x) = 1/x$, hence, when $x < 0$, also $(\log(|x|))' = (\log(-x))' = -1/(-x) = 1/x$.

To prove (3.2.1), we evaluate on test functions

$$\begin{aligned}
\langle (\log|x|)', \varphi \rangle &= -\langle \log|x|, \varphi' \rangle = -\int \log|x|\,\varphi'(x)\,dx \\
&= -\lim_{\varepsilon\to 0+} \int_{|x|>\varepsilon} \log|x|\,\varphi'(x)\,dx \\
&\overset{\substack{\text{[int. by}\\ \text{parts]}\\ \downarrow}}{=} \lim_{\varepsilon\to 0+} \left(\int_{|x|>\varepsilon} \frac{\varphi(x)}{x}\,dx - \big(\varphi(x)\log|x|\big)\Big|_{-\varepsilon}^{\varepsilon} \right) \\
&= \lim_{\varepsilon\to 0+} \int_{|x|>\varepsilon} \frac{\varphi(x)}{x}\,dx - \underbrace{\lim_{\varepsilon\to 0+} \log(\varepsilon)\,\underbrace{\big(\varphi(\varepsilon) - \varphi(-\varepsilon)\big)}_{=0+2\varepsilon\varphi'(0)+O(\varepsilon^2)}}_{=0} \\
&= \lim_{\varepsilon\to 0+} \int_{|x|>\varepsilon} \frac{\varphi(x)}{x}\,dx = \left\langle \mathrm{vp}\Big(\frac{1}{x}\Big), \varphi \right\rangle.
\end{aligned}$$

3.2.3 Examples. Whereas point charges in electrodynamics can be described by Dirac measures, hence can be modelled within measure theory, the same is no longer true for dipoles and double layer potentials, whose description requires distribution theory:

(i) *Dipoles:* Let $\omega \in S^{n-1}$, $l \in \mathbb{R}$, and $a \in \mathbb{R}^n$. A *dipole* is a limiting case of the configuration where two point charges of equal but opposite strength are separated along a vector $\varepsilon\omega \in S^{n-1}$, the limit being taken in such a way that the distance ε of the two point charges goes to zero. The *dipole moment* of strength l (corresponding to the charge) can then be calculated, according to (3.1.1), as the distribution

$$\begin{aligned}
\lim_{\varepsilon\to 0} \frac{l}{\varepsilon}(\delta_{a+\varepsilon\omega} - \delta_a) = \lim_{\varepsilon\to 0} \frac{l}{\varepsilon}(\delta_a(x - \varepsilon\omega) - \delta_a) &= -l\sum_{j=1}^n \omega_j \partial_j \delta_a \\
&= -l\omega^\top \cdot \nabla\delta_a \in \mathcal{D}'(\mathbb{R}^n).
\end{aligned}$$

For $\varphi \in \mathcal{D}(\mathbb{R}^n)$ we have

$$\langle -l\omega^\top \cdot \nabla\delta_a, \varphi \rangle = l\omega^\top \cdot \nabla\varphi(a).$$

(ii) *Double layer distributions:* Let M be a C^∞-hypersurface in $\mathbb{R}^n$, given as a level set $M = h^{-1}(a)$ of a smooth function $h : \Omega \to \mathbb{R}$ that we suppose to be a submersion near M (cf. Example 2.4.8). Then we may orient M by the unit normal $\nu := \frac{\nabla h}{|\nabla h|}$. For $f : M \to \mathbb{C}$ locally integrable, in Example 2.4.8 we defined the single layer distribution $S_M(f) \in \mathcal{D}'(\Omega)$. Similarly, we now define the *double layer distribution* $D_M(f) \in \mathcal{D}'(\Omega)$ with density f by

$$\langle D_M(f), \varphi \rangle := -\int_M f(x) \cdot \partial_\nu \varphi(x)\,d\sigma(x), \qquad \varphi \in \mathcal{D}(\Omega). \tag{3.2.2}$$

Here, $\partial_\nu \varphi(x) = \nu(x)^\top \cdot \nabla\varphi(x) = \sum_{j=1}^n \nu_j(x)\partial_j\varphi(x)$, $x \in M$, is the normal derivative of φ on M. According to (i), we may view $D_M(f)$ as a collection of dipoles in direction ν, spread out across M with density $-f$. Then analogously to the equation $g \cdot \delta' = g(0)\delta' - g'(0)\delta$, we have:

$$g \cdot D_M(f) = D_M(f \cdot g|_M) - S_M(f \cdot (\partial_\nu g)|_M), \qquad g \in \mathcal{E}(\Omega). \tag{3.2.3}$$

To prove this, we evaluate on a test function $\varphi \in \mathcal{D}(\Omega)$:

$$\begin{aligned}\langle g \cdot D_M(f), \varphi\rangle = -\int_M f(x)\partial_\nu(g\varphi)(x)\, d\sigma(x) = -\int_M f(x)g(x)\partial_\nu\varphi(x)\, d\sigma(x) \\ - \int_M f(x)\varphi(x)\partial_\nu g(x)\, d\sigma(x) = \langle D_M(f \cdot g|_M) - S_M(f \cdot (\partial_\nu g)|_M), \varphi\rangle.\end{aligned}$$

Our next aim is to express $\delta_a' \circ h$ in terms of single and double layer distributions. Using the chain rule from Proposition 3.1.5 (iii), together with (2.4.8), we calculate for $\varphi \in \mathcal{D}(\Omega)$:

$$\begin{aligned}\langle \delta_a' \circ h, \varphi\rangle &= \sum_{j=1}^n \left\langle \frac{\partial_j h}{|\nabla h|^2}\partial_j(\delta_a \circ h), \varphi\right\rangle = -\sum_{j=1}^n \left\langle \delta_a \circ h, \partial_j\left(\frac{\varphi\partial_j h}{|\nabla h|^2}\right)\right\rangle \\ &= -\sum_{j=1}^n \int_M \partial_j\left(\frac{\varphi\partial_j h}{|\nabla h|^2}\right)\frac{d\sigma}{|\nabla h|} \\ &= -\int_M \partial_\nu\varphi \frac{d\sigma}{|\nabla h|^2} - \int_M \varphi \sum_{j=1}^n \partial_j\left(\frac{\partial_j h}{|\nabla h|^2}\right)\frac{d\sigma}{|\nabla h|},\end{aligned}$$

so that

$$\delta_a' \circ h = D_M(|\nabla h|^{-2}) - S_M\Big(|\nabla h|^{-1} \cdot \nabla^\top\Big(\frac{\nu}{|\nabla h|}\Big)\Big), \tag{3.2.4}$$

where $\nabla^\top$ denotes the divergence operator.

(iii) Let us illustrate the previous point in the case that M is a sphere. Thus let $h(x) = |x|$, $h : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$, and let $a > 0$. Then $M = h^{-1}(a) = aS^{n-1}$, and $|\nabla h| = 1$. Therefore, (2.4.9) gives

$$\begin{aligned}\langle \delta_a \circ h, \varphi\rangle = \langle S_M(1), \varphi\rangle = \int_M \varphi(x)\, d\sigma(x) = \int_{aS^{n-1}} \varphi(x)\, d\sigma(x) \\ = a^{n-1}\int_{S^{n-1}} \varphi(a\omega)\, d\sigma(\omega)\end{aligned}$$

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

Next, as can be easily checked, $\nabla^\top\left(\frac{\nu}{|\nabla h|}\right) = \sum_{j=1}^n \partial_j\left(\frac{\partial_j h}{|\nabla h|^2}\right) = \frac{n-1}{|x|}$, so (3.2.4) gives

$$\delta_a' \circ h = \delta'(|x| - a) = D_M(1) - \frac{n-1}{a}S_M(1).$$

Applied to a test function, this gives

$$\begin{aligned}\langle \delta_a' \circ h, \varphi\rangle &= -a^{n-1}\int_{S^{n-1}} \omega^\top \cdot \nabla\varphi(a\omega)\, d\sigma(\omega) - (n-1)a^{n-2}\int_{S^{n-1}} \varphi(a\omega)\, d\sigma(\omega) \\ &= -\frac{\partial}{\partial a}\int_{S^{n-1}} a^{n-1}\varphi(a\omega)\, d\sigma(\omega) = -\frac{\partial}{\partial a}\langle \delta_a \circ h, \varphi\rangle.\end{aligned}$$

Note that this is consistent with Proposition 3.1.2 because

$$\delta_a' = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}(\delta_a(x + \varepsilon) - \delta_a) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}(\delta_{a-\varepsilon} - \delta_a) = -\frac{\partial}{\partial a}\delta_a$$

in $\mathcal{D}'(\mathbb{R})$.

Next we introduce a central notion for the application of distribution theory to linear partial differential equations:

3.2.4 Definition. *Let $P(x,\partial) = \sum_{|\alpha|\le m} a_\alpha(x)\partial^\alpha$ be a linear differential operator with coefficients $a_\alpha \in C^\infty(\Omega)$. A distribution $E \in \mathcal{D}'(\Omega)$ is called a* fundamental solution *of $P(x,\partial)$ at $\xi \in \Omega$ if $P(x,\partial)E = \delta_\xi$ holds in $\mathcal{D}'(\Omega)$. If $P(\partial) = \sum_{|\alpha|\le m} a_\alpha \partial^\alpha$ has constant coefficients, then $E \in \mathcal{D}'(\Omega)$ is called a fundamental (or elementary) solution of $P(\partial)$ if $P(\partial)E = \delta$ holds in $\mathcal{D}'(\mathbb{R}^n)$.*

As we shall see in Section 4.6, knowledge of a fundamental solution often allows one to solve the corresponding partial differential equation for general right hand sides. Anticipating some constructions from later chapters, the reason for this is that to solve an equation of the form

$$P(\partial)T = S$$

for a constant coefficient linear PDO $P(\partial)$ and a nice enough right hand side, knowledge of a fundamental solution E allows one to simply set $T := S * E$ (convolution), to obtain a solution via

$$P(\partial)(T) = P(\partial)(E) * S = \delta * S = S. \tag{3.2.5}$$

3.2.5 Example. We know from Example 3.1.4 that $\frac{d}{dx}Y = \delta$ in $\mathcal{D}'(\mathbb{R})$. This means that the Heaviside function is a fundamental solution of the differential operator $\frac{d}{dx}$ on $\mathbb{R}$.

More generally, let $f : \mathbb{R} \to \mathbb{C}$ be continuously differentiable outside a discrete set D (i.e., $D \cap K$ is finite for each $K \Subset \mathbb{R}$). Further, suppose that f has left and right limits in all points of D, so $f \in L^\infty_{\mathrm{loc}}(\mathbb{R})$. Finally, assume that f', which is defined on $\mathbb{R} \setminus D$, is locally integrable on $\mathbb{R}$. Then we have the following *jump formula*:

$$(T_f)' = T_{f'} + \sum_{a\in D} s(f,a)\delta_a, \tag{3.2.6}$$

where $s(f,a) = \lim_{\varepsilon\to 0+}[f(a+\varepsilon) - f(a-\varepsilon)]$ is the jump of f at a (cf. the special case (3.1.2)). To prove this formula, suppose first that f only has a single jump, i.e., $D = \{a\}$. Then

$$\begin{aligned}
\langle T_f', \varphi\rangle &= -\int_{-\infty}^{\infty} f(x)\varphi'(x)\,dx = -\int_{-\infty}^{a} f(x)\varphi'(x)\,dx - \int_{a}^{\infty} f(x)\varphi'(x)\,dx \\
&= -(f\varphi)\Big|_{-\infty}^{a} + \int_{-\infty}^{a} f'(x)\varphi(x)\,dx - (f\varphi)\Big|_{a}^{\infty} + \int_{a}^{\infty} f'(x)\varphi(x)\,dx \\
&= \lim_{\varepsilon\to 0+}[f(a+\varepsilon) - f(a-\varepsilon)]\varphi(a) + \int_{-\infty}^{\infty} f'(x)\varphi(x)\,dx \\
&= \langle T_{f'} + s(f,a)\delta_a, \varphi\rangle.
\end{aligned}$$

In general, $D \cap \operatorname{supp}\varphi$ is a finite set $\{a_1,\dots,a_k\}$ (with $a_1 < a_s < \dots < a_k$). Then

$$\langle T_f', \varphi\rangle = -\int_{-\infty}^{a_1} f(x)\varphi'(x)\,dx - \sum_{j=1}^{k-1}\int_{a_j}^{a_{j+1}} f(x)\varphi'(x)\,dx - \int_{a_k}^{\infty} f(x)\varphi'(x)\,dx$$

and applying the above calculation to each term gives (3.2.6).

By induction, one obtains the following generalization of (3.2.6) to higher order derivatives: Let $f : \mathbb{R} \to \mathbb{C}$ be m times continuously differentiable outside the discrete set D, and such that $f^{(k)}$ has limits from the left and the right, for $0 \leq k < m$. Finally, suppose that $f^{(m)}$, which is defined on $\mathbb{R} \setminus D$, is locally integrable on $\mathbb{R}$. Then

$$\frac{d^m}{dx^m} T_f = T_{f^{(m)}} + \sum_{k=0}^{m-1} \sum_{a \in D} s(f^{(m-k-1)}, a) \delta_a^{(k)}. \tag{3.2.7}$$

As a first concrete application of (3.2.6), let $f(x) := Y(x)e^{\lambda x}$, $\lambda \in \mathbb{C}$. Then

$$T_f' = \lambda Y \cdot e^{\lambda x} + \delta = \lambda f + \delta \quad \Rightarrow \quad \left(\frac{d}{dx} - \lambda\right) f = \delta.$$

In other words, $f(x) := Y(x)e^{\lambda x}$ is a fundamental solution of the differential operator $\frac{d}{dx} - \lambda$.

For a general constant-coefficient ordinary differential operator $P(\frac{d}{dx}) = \sum_{j=0}^m a_j \frac{d^j}{dx^j}$, we may factorize the *characteristic polynomial* $\sum_{j=0}^m a_j \xi^j$ into linear factors, and therefore an analogous factorization applies to P itself. Thus the following result shows how to calculate a fundamental solution for an arbitrary such operator $P(\frac{d}{dx})$.

3.2.6 Theorem. *Let $m \in \mathbb{N}$, $\alpha \in \mathbb{N}_0^m$, and let $\lambda_1, \ldots, \lambda_m \in \mathbb{C}$ be pairwise different. Then the ordinary differential operator*

$$P_{\lambda,\alpha}\Big(\frac{d}{dx}\Big) = \prod_{j=1}^m \Big(\frac{d}{dx} - \lambda_j\Big)^{\alpha_j+1}$$

has as fundamental solution the L^∞_{loc}-function $E_{\lambda,\alpha}$ given by

$$\begin{aligned} E_{\lambda,\alpha}(x) &= \frac{Y(x)}{\alpha!}\Big(\frac{\partial}{\partial\lambda}\Big)^{\alpha}\Big(\sum_{j=1}^m e^{\lambda_j x} \prod_{k\neq j}(\lambda_j - \lambda_k)^{-1}\Big) \\ &= Y(x) \sum_{j=1}^m \frac{1}{\alpha_j!}\Big(\frac{\partial}{\partial\lambda_j}\Big)^{\alpha_j}\Big(e^{\lambda_j x} \prod_{k\neq j}(\lambda_j - \lambda_k)^{-\alpha_k-1}\Big). \end{aligned} \tag{3.2.8}$$

$E_{\lambda,\alpha}$ is the only fundamental solution of $P_{\lambda,\alpha}(\frac{d}{dx})$ with support in $[0, \infty)$.

Proof. We first consider the special case $\alpha = 0$, making an ansatz for $E := E_{\lambda,0}$ in the form

$$E = Y \cdot \sum_{j=1}^m a_j e^{\lambda_j x}, \quad a_j \in \mathbb{C}.$$

Then obviously $P_{\lambda,0}(\frac{d}{dx})E|_{\mathbb{R}\setminus\{0\}} = 0$. We want to pick the coefficients a_j in such a way that $E \in C^{m-2}$ and that $E^{(m-1)}$ has a jump of height one at 0, i.e.,

$$\begin{aligned} &0 = s(E, 0) = \sum_{j=1}^m a_j,\ 0 = s(E', 0) = \sum_{j=1}^m \lambda_j a_j,\ \ldots, \\ &\qquad \ldots,\ 0 = s(E^{(m-2)}, 0) = \sum_{j=1}^m \lambda_j^{m-2} a_j,\ 1 = s(E^{(m-1)}, 0) = \sum_{j=1}^m \lambda_j^{m-1} a_j. \end{aligned}$$

The leading term of $P_{\lambda,0}(\frac{d}{dx})$ is $\frac{d^m}{dx^m}$, so by (3.2.7) we obtain

$$P_{\lambda,0}(\frac{d}{dx})E = s(E^{(m-1)}, 0)\delta + T_{P_{\lambda,0}(\frac{d}{dx})E} = 1 \cdot \delta + 0 = \delta.$$

Finally, the above choice of coefficients is possible because the a_j are the solution to the following Vandermonde system of linear equations:

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_m \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_1^{m-1} & \lambda_2^{m-1} & \dots & \lambda_m^{m-1} \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

As is well known from linear algebra, the determinant of this matrix is $\prod_{j>k}(\lambda_j - \lambda_k)$, so Cramer's rule implies that

$$a_j = \prod_{k\neq j}(\lambda_j - \lambda_k)^{-1} = P'_{\lambda,0}(\lambda_j)^{-1}.$$

This proves (3.2.8) for $\alpha = 0$.

Next we note that the maps

$$\{\lambda \in \mathbb{C}^m \mid \lambda_1, \dots, \lambda_m \text{ pairwise different}\} \to \mathbb{C}: \ \lambda \mapsto \langle E_{\lambda,0}, \varphi\rangle$$

are smooth (in fact, even holomorphic) for each $\varphi \in \mathcal{D}(\mathbb{R})$. The first representation in (3.2.8) shows that the α-th derivative of $E_{\lambda,0}$ with respect to λ is the distribution $\alpha! E_{\lambda,\alpha}$. To check that $E_{\lambda,\alpha}$ is a fundamental solution of $P_{\lambda,\alpha}(\frac{d}{dx})$ we therefore use induction with respect to $|\alpha|$. The base case $\alpha = 0$ was already shown above. Suppose, therefore, that $\varphi \in \mathcal{D}(\mathbb{R})$, $\alpha = (\alpha_1, \alpha') \in \mathbb{N}_0^m$, and $P_{\lambda,\alpha}(\frac{d}{dx})E_{\lambda,\alpha} = \delta$ (obviously it suffices to consider the case where the first component of α is raised by one). Then

$$\begin{aligned} 0 = \frac{\partial}{\partial\lambda_1}\varphi(0) &= \frac{\partial}{\partial\lambda_1}\left\langle P_{\lambda,\alpha}\Big(\frac{d}{dx}\Big)E_{\lambda,\alpha}, \varphi\right\rangle = \frac{\partial}{\partial\lambda_1}\left\langle E_{\lambda,\alpha}, P_{\lambda,\alpha}\Big(-\frac{d}{dx}\Big)\varphi\right\rangle \\ &= -(1+\alpha_1)\left\langle E_{\lambda,\alpha}, P_{\lambda,(\alpha_1-1,\alpha')}\Big(-\frac{d}{dx}\Big)\varphi\right\rangle + \left\langle \frac{\partial E_{\lambda,\alpha}}{\partial\lambda_1}, P_{\lambda,\alpha}\Big(-\frac{d}{dx}\Big)\varphi\right\rangle. \end{aligned}$$

Now given $\psi \in \mathcal{D}(\mathbb{R})$, insert $\varphi := \big(-\frac{d}{dx} - \lambda_1\big)\psi$ here, to obtain

$$\left\langle \frac{\partial E_{\lambda,\alpha}}{\partial\lambda_1}, P_{\lambda,(\alpha_1+1,\alpha')}\Big(-\frac{d}{dx}\Big)\psi\right\rangle = (1+\alpha_1)\left\langle E_{\lambda,\alpha}, P_{\lambda,\alpha}\Big(-\frac{d}{dx}\Big)\psi\right\rangle = (1+\alpha_1)\psi(0).$$

Consequently, $\frac{1}{1+\alpha_1}\frac{\partial E_{\lambda,\alpha}}{\partial\lambda_1}$ is a fundamental solution of $P_{\lambda,(\alpha_1+1,\alpha')}(\frac{d}{dx})$, concluding the induction step.

Finally, suppose that $\tilde{E}_{\lambda,\alpha}$ is another fundamental solution of $P_{\lambda,\alpha}(\frac{d}{dx})$ with support contained in $[0,\infty)$. Then the difference $T := E_{\lambda,\alpha} - \tilde{E}_{\lambda,\alpha}$ is a solution of the homogeneous linear differential equation $P_{\lambda,\alpha}(\frac{d}{dx})T = 0$. As we shall see in Theorem 3.4.4 below, such an equation only has classical solutions. Thus T must be a polynomial, but since its support is supposed to be contained in $[0,\infty)$ it must in fact vanish. □

3.2.7 Examples. (i) Let $P(\frac{d}{dx}) = P_{\lambda,0}(\frac{d}{dx}) = \prod_{j=1}^m(\frac{d}{dx} - \lambda_j)$ with $\lambda_j \in \mathbb{C}$ pairwise different. Then Theorem 3.2.6 gives the fundamental solution

$$E = E_{\lambda,0}(x) = Y(x)\sum_{j=1}^m e^{\lambda_j x}\prod_{k\neq j}(\lambda_j - \lambda_k)^{-1}.$$

In particular, for $\lambda_1 = -\lambda_2 = i\omega$, $\omega \in \mathbb{C}\backslash\{0\}$ this shows that a fundamental solution of the operator $\frac{d^2}{dx^2} + \omega^2$ is given by

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

(ii) Let $P(\frac{d}{dx}) = (\frac{d}{dx} - \lambda)^{r+1}$. Then $m = 1$ and $\alpha = (r, 0, \dots, 0)$, so

$$E = \frac{1}{r!} Y(x) \Big(\frac{\partial}{\partial\lambda}\Big)^r e^{\lambda x} = \frac{1}{r!} Y(x) x^r e^{\lambda x}.$$

3.3 The multi-dimensional jump-formula

In this section we seek to generalize the distributional jump-formula (3.2.6) to jumps along hypersurfaces in several dimensions. We shall see that this generalization has numerous important applications, in particular for the determination of fundamental solutions.

Let $M \subseteq \Omega$ be a closed oriented C^∞-hypersurface, and let $f : \Omega \setminus M \to \mathbb{R}$ be C^∞. (The results in this section in fact hold for suitable finite differentiability as well, as can be seen transferring the arguments given below to the setting of distributions of finite order). We assume that f has (in general different) boundary values from both sides of M, and that the partial derivatives $\partial_j f$, defined on $\Omega \setminus M$, are locally integrable when viewed as functions on Ω. Then we define the *jump vector field* $s(f)$ of f along M by

$$s(f) : M \to \mathbb{R}^n, \quad x \mapsto \nu(x) \cdot \lim_{\varepsilon \to 0} (f(x + \varepsilon\nu(x)) - f(x - \varepsilon\nu(x))),$$

where $\nu(x)$ is a unit normal of M at $x \in M$. Note that $s(f)(x)$ is independent of the choice of $\nu(x)$.

We assume that the gradient ∇f of f is an element of $L^1_{\mathrm{loc}}(\Omega)^n \subseteq \mathcal{D}'(\Omega)^n$ and that $s(f) \in L^1_{\mathrm{loc}}(M, d\sigma)$.

3.3.1 Theorem. *Under the above assumptions, the following* distributional jump formula *holds:*

$$\nabla T_f = T_{\nabla f} + S_M(s(f)),$$

where, in analogy to (2.4.6), $S_M(s(f)) \in \mathcal{D}'(\Omega)^n$ *is defined by*

$$\langle S_M(s(f)), \varphi\rangle := \int_M s(f)(x)\varphi(x)\, d\sigma(x) \in \mathbb{R}^n \qquad (\varphi \in \mathcal{D}(\Omega)).$$

Proof. Using a partition of unity, we may reduce the proof to the case where $M = h^{-1}(0)$ for some smooth function $h : \Omega \to \mathbb{R}$ that is a submersion on M. Since M is oriented there is a global choice of smooth unit normal vector ν and we may choose the sign of h such that $\nu = \frac{\nabla h}{|\nabla h|}$. Recall that the Gauss divergence theorem implies, for any $X \in \mathfrak{X}(\Omega)$:

$$\int_\Omega \nabla \cdot X\, dx = \int_{\partial\Omega} X \cdot \nu(x)\, d\sigma(x),$$

or, for the components of X,

$$\int_\Omega \partial_j X_j\, dx = \int_{\partial\Omega} X_j \cdot \nu_j(x)\, d\sigma(x).$$

In the present setting, if X is compactly supported, then the only boundary term that contributes is the integral over M, and $\Omega \setminus M = \{h < 0\} \cup \{h > 0\}$. Thus we obtain, for any $\varphi \in \mathcal{D}(\Omega)$ and any $j \in \{1, \ldots, n\}$

$$\begin{aligned}
\langle \partial_j T_f, \varphi \rangle &= -\int_\Omega f \partial_j \varphi dx = -\int_{h(x)<0} f \partial_j \varphi dx - \int_{h(x)>0} f \partial_j \varphi dx \\
&= \int_{\Omega \setminus M} \partial_j f \varphi \, dx - \int_M \varphi(x) \frac{\partial_j h}{|\nabla h|} \lim_{\varepsilon \to 0+} (f(x - \varepsilon \nu(x))) \, d\sigma(x) \\
&\qquad\qquad + \int_M \varphi(x) \frac{\partial_j h}{|\nabla h|} \lim_{\varepsilon \to 0+} (f(x + \varepsilon \nu(x))) \, d\sigma(x) \\
&= \langle T_{\partial_j f}, \varphi \rangle + \langle S_M(s(f)_j), \varphi \rangle.
\end{aligned}$$

Collecting all components, this concludes the proof. □

3.3.2 Corollary. *Let $f, h \in C^\infty(\Omega)$, let h be a submersion on $M = h^{-1}(0)$, and set $\nu(x) := \frac{\nabla h(x)}{|\nabla h(x)|}$. Then*

$$\nabla(Y(h) \cdot f) = Y(h) \cdot \nabla f + S_M(\nu \cdot f) \tag{3.3.1}$$

Proof. $Y(h) \cdot f$ satisfies the assumptions of Theorem 3.3.1, jumping from 0 to f at M. The gradient of this function is 0 on one side of M and ∇f on the other, so the first term in the jump formula from Theorem 3.3.1 gives $Y(h) \cdot \nabla f$. For the second, note that

$$s(f)(x) = \frac{\nabla h(x)}{|\nabla h(x)|} \lim_{\varepsilon \to 0+} (f(x + \varepsilon \nabla h(x)) - 0) = f(x)\nu(x).$$

Alternatively, we may use the product and chain rules from Proposition 3.1.5 to obtain $\nabla(Y(h) \cdot f) = Y(h) \cdot \nabla f + f \cdot \nabla h \cdot \delta \circ h$. The claim then follows from (2.4.9) because

$$\begin{aligned}
\langle f \cdot \nabla h \cdot \delta \circ h, \varphi \rangle &= \langle \delta \circ h, f \cdot \nabla h \cdot \varphi \rangle = \langle S_M(|\nabla h|^{-1}), f \cdot \nabla h \cdot \varphi \rangle \\
&= \int_M f(x) \frac{\nabla h(x)}{|\nabla h(x)|} \varphi(x) \, d\sigma(x) = \langle S_M(\nu \cdot f), \varphi \rangle.
\end{aligned}$$

□

Applying (3.3.1) componentwise, we obtain the following jump formula for a smooth vector field $v \in \mathfrak{X}(\Omega)$:

$$\nabla^\top (Y(h) \cdot v) = Y(h) \cdot \nabla^\top v + S_M(\nu^\top \cdot v).$$

3.3.3 Proposition. *Let $M \subseteq \Omega$ be a closed oriented C^∞-hypersurface (so $\dim(M) = n - 1$) with unit normal ν. Then*

$$\nabla S_M(1) = D_M(\nu) - S_M(\nu \cdot \nabla^\top \nu), \tag{3.3.2}$$

where $\nabla^\top \nu$ is defined by arbitrarily extending ν to a smooth vector field in a neighborhood of M.

Proof. Any point in M possesses an open neighborhood U in Ω such that $M \cap U = h^{-1}(0)$ for some smooth function h with $\nabla h \neq 0$ in U and such that $\nu = \nabla h / |\nabla h|$ on U. For $\varphi \in \mathcal{D}(U)$ and fixed $j \in \{1, \ldots, n\}$ we define the smooth vector field w by

$$w_k(x) := \nabla^\top (\nu \varphi) \delta_{jk} - \partial_j (\nu_k \varphi) \qquad (1 \leq k \leq n).$$

Then

$$\operatorname{div}(w) = \sum_k \partial_k \Big(\sum_m \partial_m(\nu_m\varphi)\delta_{jk} \Big) - \sum_k \partial_k\partial_j(\nu_k\varphi)$$
$$= \sum_m \partial_j\partial_m(\nu_m\varphi) - \sum_k \partial_k\partial_j(\nu_k\varphi) = 0.$$

Note that, since ν is a unit vector field, $\nu^\top \cdot \nu = 1$, so $0 = \partial_j(\nu^\top \cdot \nu) = 2\nu^\top \partial_j \nu$. Therefore, the Gauss theorem implies

$$0 = \int_{\{x\in U: h(x)<0\}} \operatorname{div}(w)\, dx = \int_{M\cap U} \nu^\top w\, d\sigma = \int_{M\cap U} [\nu_j \nabla^\top(\nu\varphi) - \nu^\top \partial_j(\nu\varphi)]\, d\sigma$$
$$= \int_{M\cap U} [\nu_j \partial_\nu \varphi + \nu_j\varphi\nabla^\top\nu - \varphi \underbrace{\nu^\top\partial_j\nu}_{=0} - \partial_j\varphi]\, d\sigma$$
$$= \langle -D_M(\nu_j) + S_M(\nu_j\nabla^\top\nu) + \partial_j S_M(1), \varphi\rangle,$$

which gives the claim. The result for Ω instead of U then follows as usual by using a partition of unity. □

3.3.4 Remark. To relate the above result to the geometry of hypersurfaces, let us recall some basic facts from differential geometry (cf., e.g., [Kun08, Bae10]). The map $x \mapsto \nu(x)$, assigning to each point x the unit normal at x is called the Gauss map. Its differential $\widetilde{W}_x := D_x\nu : T_xM \to T_xM$ is called the Weingarten map or shape operator. We first note that indeed $\widetilde{W}$ takes values in T_xM: We have $T_xM = \nu(x)^\perp$. Now denoting the standard scalar product on $\mathbb{R}^n$ by $(\ ,\ )$, let $w \in T_xM$ and let $c : \mathbb{R} \to M$ be a smooth curve with $c(0) = x$ and $c'(0) = w$. Then since $(\nu(c(t)), \nu(c(t))) \equiv 1$ we get

$$0 = \frac{d}{dt}\Big|_{t=0} (\nu(c(t)), \nu(c(t)) = 2(D_x\nu(w), \nu(x)),$$

so indeed $\widetilde{W}_x(w) \in T_xM$. Note that for any $w \in \mathbb{R}^n$ we have

$$(w^\top\nabla)\nu = (\sum_{i=1}^n w_i\partial_i)\nu = \begin{pmatrix} \sum_{i=1}^n w_i\partial_i\nu_1 \\ \vdots \\ \sum_{i=1}^n w_i\partial_i\nu_n \end{pmatrix}$$

and

$$\nabla\nu^\top = \begin{pmatrix} \partial_1 \\ \vdots \\ \partial_n \end{pmatrix} \cdot (\nu_1, \dots, \nu_n) = \begin{pmatrix} \partial_1\nu_1 & \dots & \partial_1\nu_n \\ \vdots & \vdots & \vdots \\ \partial_n\nu_1 & \dots & \partial_n\nu_n \end{pmatrix},$$

so

$$(\nabla\nu^\top)^\top w = \begin{pmatrix} \partial_1\nu_1 & \dots & \partial_n\nu_1 \\ \vdots & \vdots & \vdots \\ \partial_1\nu_n & \dots & \partial_n\nu_n \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n w_i\partial_i\nu_1 \\ \vdots \\ \sum_{i=1}^n w_i\partial_i\nu_n \end{pmatrix} = (w^\top\nabla)\nu. \quad (3.3.3)$$

Since, by the above, $(\nabla\nu^\top)^\top$ is the Jacobian of ν, it follows that

$$\widetilde{W}_x : T_xM \to T_xM : w \mapsto (w^\top\nabla)\nu = (\nabla\nu^\top)^\top w.$$

One can show that in fact $\widetilde{W}_x$ is a symmetric map, i.e.,

$$(\widetilde{W}_x u, w) = (u, \widetilde{W}_x w)$$

for all $u, w \in T_xM$. For a proof we refer to [Kun08, Prop. 3.1.7].

To extend $\widetilde{W}_x$ to a map (again called Weingarten map) on all of $\mathbb{R}^n$ we make use of the orthogonal decomposition $\mathbb{R}^n = T_xM \oplus \langle \nu(x) \rangle$ and add a zero component in the orthogonal direction. Thus, denoting the orthogonal projection onto T_xM by π_x, we set

$$\begin{aligned} W_x : \mathbb{R}^n = T_xM \oplus \langle \nu(x) \rangle &\mapsto T_xM \oplus \langle \nu(x) \rangle = \mathbb{R}^n \\ w &\mapsto \widetilde{W}_x(\pi_x(w)) \end{aligned}$$

Since in this orthogonal decomposition we have only added a trivial component in the ν-direction, W_x is also symmetric:

$$\begin{aligned} (W_x(w), u) &= (W_x(w), \pi_x(u)) = (\widetilde{W}_x(\pi_x(w)), \pi_x(u)) \\ &= (\pi_x(w), \widetilde{W}_x(\pi_x(u))) = (w, W_x(u)). \end{aligned}$$

Due to

$$(\nu\nu^\top)w = \begin{pmatrix} \nu_1\nu_1 & \dots & \nu_1\nu_n \\ \vdots & \vdots & \vdots \\ \nu_n\nu_1 & \dots & \nu_n\nu_n \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} \nu_1 \cdot \sum_{i=1}^n \nu_i w_i \\ \vdots \\ \nu_n \cdot \sum_{i=1}^n \nu_i w_i \end{pmatrix} = (w, \nu)\nu,$$

we see that $\pi_x(w) = w - (w, \nu(x))\nu(x) = (I_n - \nu(x)\nu(x)^\top)w$, so (keeping in mind that $W = W^\top$)

$$W = \widetilde{W} \circ \pi = (\nabla\nu^\top)^\top \cdot (I_n - \nu\nu^\top) = (I_n - \nu\nu^\top) \cdot \nabla\nu^\top, \tag{3.3.4}$$

Finally, let us calculate the trace $\operatorname{tr}(W_x)$ of W_x. For this, let $(e_1, \dots, e_{n-1})$ be an orthonormal basis in T_xM of eigenvectors of $\widetilde{W}_x$ (which exists since $\widetilde{W}_x$ is symmetric). Then $(e_1, \dots, e_{n-1}, \nu_x)$ is an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of W_x, hence

$$\operatorname{tr}(W_x) = \sum_{k=1}^{n-1} (W_x e_i, e_i) + \underbrace{(W_x\nu, \nu)}_{=0} = \sum_{k=1}^{n-1} (\widetilde{W}_x e_i, e_i) = \operatorname{tr}(\widetilde{W}_x).$$

Here, $\operatorname{tr}(\widetilde{W}_x)$ equals the trace of the $n \times n$ matrix $(\nabla\nu^\top)^\top$: The matrix of the linear map $(\nabla\nu^\top)^\top : \mathbb{R}^n \to \mathbb{R}^n$ with respect to the onb $(e_1, \dots, e_{n-1}, \nu_x)$ is

$$\begin{pmatrix} (\widetilde{W}_x e_1, e_1) & 0 & \dots & 0 & * \\ 0 & (\widetilde{W}_x e_2, e_2) & \dots & 0 & * \\ \vdots & \vdots & \dots & (\widetilde{W}_x e_{n-1}, e_{n-1}) & * \\ 0 & 0 & \dots & 0 & ((\nabla\nu^\top)^\top \nu, \nu) \end{pmatrix}$$

In this expression,

$$((\nabla\nu^\top)^\top \nu, \nu) \underset{(3.3.3)}{=} ((\nu^\top\nabla)\nu, \nu) = (\partial_\nu \nu, \nu) = \frac{1}{2}\partial_\nu \underbrace{(\nu, \nu)}_{=1} = 0.$$

It follows that $\operatorname{tr}(\widetilde{W}_x)$ equals the trace of the operator $(\nabla\nu^\top)^\top$, which (using the standard basis) due to (3.3.3) is given by $\nabla^\top\nu$. Summing up,

$$\operatorname{tr}(W_x) = \operatorname{tr}(\widetilde{W}_x) = \nabla^\top\nu(x) = \sum_{k=1}^{n} \partial_k \nu_k(x). \tag{3.3.5}$$

(By the Theorem of Rodriguez (cf. [Kun08, Th. 3.1.11]), the eigenvalues of $\widetilde{W}_x$ are the principal curvatures of M in x, hence this trace is $(n-1)$ times the mean curvature of M.)

Returning to the interpretation of Proposition 3.3.3, by (3.3.5) we have

$$\nabla S_M(1) = D_M(\nu) - S_M(\nu \cdot \mathrm{tr}(W)), \tag{3.3.6}$$

which has to be read component-wise: $\partial_j S_M(1) = D_M(\nu_j) - S_M(\nu_j \mathrm{tr}(W))$ ($j = 1, \ldots, n$).

Our next aim is to follow up on the jump formula for vector fields (3.3.1) and calculate the corresponding second derivative:

3.3.5 Lemma. *Let $f, h \in C^\infty(\Omega)$, with h a submersion on $M = h^{-1}(0)$, and let ν and W be as above. Then*

$$\begin{aligned}\nabla\nabla^\top(Y(h)\cdot f) =& Y(h)\cdot\nabla\nabla^\top f + D_M(f\nu\nu^\top) + S_M\big(\nu\cdot\nabla^\top f \\ &+ \nabla f\cdot\nu^\top - \partial_\nu(f)\nu\nu^\top + f\cdot(W - \nu\nu^\top\mathrm{tr}(W))\big).\end{aligned} \tag{3.3.7}$$

Proof. Note that, in the above claim, for any $T \in \mathcal{D}'(\Omega)$, by $\nabla\nabla^\top T$ we denote the Hessian matrix

$$\nabla\nabla^\top T = \begin{pmatrix}\partial_1\\ \vdots \\ \partial_n\end{pmatrix}\cdot(\partial_1 T, \ldots, \partial_n T) = \begin{pmatrix}\partial_1^2 T & \ldots & \partial_1\partial_n T\\ \vdots & \vdots & \vdots\\ \partial_n\partial_1 T & \ldots & \partial_n^2 T\end{pmatrix}.$$

Applying (3.3.1) twice, as well as the product rule, we calculate

$$\begin{aligned}\nabla\nabla^\top(Y(h)\cdot f) &= \nabla[Y(h)\cdot\nabla^\top f + f\nu^\top S_M(1)]\\ &= Y(h)\cdot\nabla\nabla^\top f + \nu\cdot\nabla^\top f\cdot S_M(1)\\ &+ \nabla f\cdot\nu^\top\cdot S_M(1) + f\cdot\nabla(\nu^\top)\cdot S_M(1) + f\cdot\nabla S_M(1)\cdot\nu^\top.\end{aligned}$$

If we insert here for $\nabla S_M(1)$ from (3.3.6), we obtain

$$\begin{aligned}\nabla\nabla^\top(Y(h)\cdot f) &= Y(h)\cdot\nabla\nabla^\top f + f\cdot D_M(\nu)\cdot\nu^\top\\ &+ (\nu\cdot\nabla^\top f + \nabla f\cdot\nu^\top + f\cdot\nabla(\nu^\top) - f\cdot\mathrm{tr}(W)\cdot\nu\nu^\top)S_M(1).\end{aligned}$$

Also, from (3.2.3) it follows that, for any $1 \le i, j \le n$,

$$f\cdot\nu_j D_M(\nu_i) = D_M(f\nu_i\nu_j) - S_M(\nu_i\partial_\nu(f\nu_j)).$$

Written in matrix form, this means

$$f\cdot D_M(\nu)\cdot\nu^\top = D_M(f\nu\nu^\top) - S_M(\nu\cdot\partial_\nu(f\nu^\top)).$$

Here, $\partial_\nu(f\nu^\top) = (\partial_\nu f)\nu^\top + f\partial_\nu\nu^\top$, where $\partial_\nu\nu^\top = \nu^\top\nabla\nu^\top$. Finally, since (3.3.4) shows that

$$W = \nabla\nu^\top - \nu\nu^\top\nabla\nu^\top,$$

the claim follows. □

3.3.6 Example. Let $h(x) := |x| - R$, so $M = h^{-1}(0) = R\cdot S^{n-1}$ ($R > 0$). Then

$$\nu(x) = \frac{x}{|x|}, \qquad \mathrm{tr}(W) = \nabla^\top\nu = \frac{n-1}{R},$$

and

$$\nabla\nu^\top = \frac{|x|^2 I_n - xx^\top}{|x|^3},$$

implying that in this situation $\nu\nu^\top\nabla\nu^\top = 0$, hence

$$W = \nabla\nu^\top - \nu\nu^\top\nabla\nu^\top = \nabla\nu^\top = \frac{|x|^2 I_n - xx^\top}{|x|^3}.$$

Therefore, (3.3.7) in this case reads

$$\begin{aligned}\nabla\nabla^\top(Y(|x|-R)\cdot f) &= Y(|x|-R)\cdot\nabla\nabla^\top f + D_{R\cdot S^{n-1}}\Big(\frac{fxx^\top}{R^2}\Big)\\ &+ S_{R\cdot S^{n-1}}\Big(\frac{x}{R}\cdot\nabla^T f + \nabla f\cdot\frac{x^\top}{R} - (x^\top\nabla f)\cdot\frac{xx^\top}{R^3} + f\cdot\frac{R^2 I_n - nxx^\top}{R^3}\Big)\end{aligned}$$

If, in particular, f is rotationally symmetric, i.e., $f(x) = g(|x|)$ for some smooth function g on $\mathbb{R}\setminus\{0\}$, we obtain

$$\begin{aligned}&\nabla\nabla^\top(Y(|x|-R)\cdot g(|x|))\\ &=Y(|x|-R)\Big(\frac{|x|^2 I_n - xx^\top}{|x|^3}g'(|x|) + \frac{xx^\top}{|x|^2}g''(|x|)\Big)\\ &+\frac{g(R)}{R^2}D_{R\cdot S^{n-1}}(xx^\top) + \Big(\frac{g'(R)}{R^2} - \frac{ng(R)}{R^3}\Big)S_{R\cdot S^{n-1}}(xx^\top)\\ &+\frac{g(R)}{R}I_n S_{R\cdot S^{n-1}}(1).\end{aligned}\tag{3.3.8}$$

3.4 Distributions with support in a single point

Our next aim is to characterize all distributions whose support consists of a single point. Clearly any linear combination of derivatives of δ_{x_0} has this property. We will show that the converse of this observation is true as well. For this, we need the following auxiliary result:

3.4.1 Lemma. *Let $x_0 \in \Omega$ and let $T \in \mathcal{D}'(\Omega)$ with* $\operatorname{supp}(T) = \{x_0\}$. *Then $\exists m \in \mathbb{N}_0$ such that*

$$\langle T,\varphi\rangle = 0 \quad \forall\varphi\in\mathcal{D}(\Omega) \text{ with } \partial^\alpha\varphi(x_0) = 0 \quad \forall\ |\alpha|\le m.$$

Proof. Without loss of generality we may suppose that $x_0 = 0$ and $\Omega = \mathbb{R}^n$.
Let $\psi \in \mathcal{D}(\mathbb{R}^n)$ with $\psi = 1$ on $B_{1/2}(0)$ and $\psi = 0$ on $\mathbb{R}^n\setminus B_1(0)$. If $\varepsilon\in(0,1]$ then we have for any $\varphi\in\mathcal{D}(\mathbb{R}^n)$ that

$$\varphi(x) - \varphi(x)\,\psi\Big(\frac{x}{\varepsilon}\Big) = 0 \qquad \text{when } x \in B_{\varepsilon/2}(0).$$

Therefore $\operatorname{supp}(T)\cap\operatorname{supp}(\varphi - \varphi\,\psi(\frac{\cdot}{\varepsilon})) = \emptyset$ and so by Proposition 2.2.4

$$\langle T,\varphi\rangle = \Big\langle T, \varphi\,\psi\Big(\frac{\cdot}{\varepsilon}\Big)\Big\rangle. \tag{3.4.1}$$

Now for each $\varepsilon\in(0,1]$ we have $\operatorname{supp}(\varphi\,\psi(\frac{\cdot}{\varepsilon})) \subseteq \operatorname{supp}(\psi) \subseteq \overline{B_1(0)} =: K$. Hence (1.1.1), together with (3.4.1) implies that $\exists m\in\mathbb{N}_0\ \exists C>0$ such that

$$|\langle T,\varphi\rangle| \le C\sum_{|\alpha|\le m}\|\partial^\alpha\Big(\varphi\,\psi\Big(\frac{\cdot}{\varepsilon}\Big)\Big)\|_{\infty,K} \qquad \forall\varphi\in\mathcal{D}(\Omega). \tag{3.4.2}$$

Suppose now that $\partial^\alpha\varphi(0) = 0$ for all $|\alpha| \le m$, and let $|\beta| \le m$. Then we have by Taylor's theorem

$$\partial^\beta\varphi(x) = \sum_{|\gamma|\le m-|\beta|} \frac{x^\gamma}{\gamma!}\partial^{\beta+\gamma}\varphi(0) \qquad \text{[=0 by hypothesis]}$$
$$+ (m-|\beta|+1) \sum_{|\gamma|=m-|\beta|+1} \frac{x^\gamma}{\gamma!} \cdot \int_0^1 (1-t)^{m-|\beta|}(\partial^{\beta+\gamma}\varphi)(tx)\,dt.$$

Hence by the compactness of $\operatorname{supp}(\varphi)$ we obtain the estimate

$$|\partial^\beta\varphi(x)| \le \overbrace{(m-|\beta|+1)\int_0^1 (1-t)^{m-|\beta|}\,dt}^{=1} \cdot \sum_{|\gamma|=m-|\beta|+1} \frac{|x|^{|\gamma|}}{\gamma!}\|(\partial^{\beta+\gamma}\varphi)\|_{\infty,\operatorname{supp}(\varphi)}$$
$$= \underbrace{\left(\sum_{|\gamma|=m-|\beta|+1} \frac{\|(\partial^{\beta+\gamma}\varphi)\|_{\infty,\operatorname{supp}(\varphi)}}{\gamma!} \right)}_{=:C(m,\beta,\varphi)} \cdot |x|^{m-|\beta|+1},$$

which in turn gives

$$|\partial^\beta\varphi(x)| \le C(m,\beta,\varphi)\,\varepsilon^{m-|\beta|+1} \qquad \text{for } |x|\le\varepsilon. \tag{3.4.3}$$

To prepare for the application of (3.4.3) to the estimate (3.4.2), we apply the Leibniz rule and use the fact that $\operatorname{supp}(\psi(\frac{\cdot}{\varepsilon})) \subseteq \overline{B_\varepsilon(0)}$ to deduce

$$\|\partial^\alpha\left(\varphi\,\psi\left(\frac{\cdot}{\varepsilon}\right)\right)\|_{\infty,K} \le \sum_{\beta\le\alpha}\binom{\alpha}{\beta}\|\partial^\beta\varphi(.)(\partial^{\alpha-\beta}\psi)\left(\frac{\cdot}{\varepsilon}\right)\varepsilon^{-|\alpha-\beta|}\|_{\infty,\overline{B_\varepsilon(0)}}$$
$$\underset{\substack{\uparrow\\(3.4.3)}}{\le} \sum_{\beta\le\alpha}\binom{\alpha}{\beta}C(m,\beta,\varphi)\varepsilon^{m-|\beta|+1}C\varepsilon^{-(|\alpha|-|\beta|)} = O(\varepsilon^{m+1-|\alpha|}).$$

Inserting these upper bounds into (3.4.2) yields

$$|\langle T,\varphi\rangle| = O\big(\sum_{|\alpha|\le m}\varepsilon^{m+1-|\alpha|}\big) = O(\varepsilon).$$

Since $0<\varepsilon\le 1$ was arbitrary we obtain that $\langle T,\varphi\rangle = 0$. □

Based on this we can now prove:

3.4.2 Theorem. *Let $x_0\in\Omega$ and $T\in\mathcal{D}'(\Omega)$ with* $\operatorname{supp}(T)=\{x_0\}$. *Then $\exists m\in\mathbb{N}_0$ and $c_\alpha\in\mathbb{C}$ ($|\alpha|\le m$), such that*

$$\langle T,\varphi\rangle = \sum_{|\alpha|\le m} c_\alpha\partial^\alpha\varphi(x_0) \qquad \forall\varphi\in\mathcal{D}(\Omega).$$

In other words, $T=\sum_{|\alpha|\le m}(-1)^{|\alpha|}c_\alpha\partial^\alpha\delta_{x_0}$.

Proof. Again, without loss of generality, let $x_0=0$, $\Omega=\mathbb{R}^n$.
Let ψ, K, and m be as in the proof of Lemma 3.4.1. Let $\varphi\in\mathcal{D}(\Omega)$ be arbitrary.

We have by Taylor's theorem

$$\varphi(x) = \sum_{|\gamma|\le m} \frac{x^\gamma}{\gamma!}\partial^\gamma\varphi(0) \;+\; \overbrace{(m+1)\sum_{|\gamma|=m+1} \frac{x^\gamma}{\gamma!}\cdot\int_0^1 (1-t)^m(\partial^\gamma\varphi)(tx)\,dt}^{=:R_\varphi(x)}$$
$$= \sum_{|\gamma|\le m} \frac{x^\gamma\psi(x)}{\gamma!}\,\partial^\gamma\varphi(0) \;+\; \underbrace{\big(1-\psi(x)\big)\sum_{|\gamma|\le m}\frac{x^\gamma}{\gamma!}\partial^\gamma\varphi(0) \;+R_\varphi(x)}_{=:\widetilde{\varphi}(x)},$$

where $\widetilde{\varphi} \in \mathcal{D}(\mathbb{R}^n)$ satisfies $\partial^\alpha\widetilde{\varphi}(0) = 0$ when $|\alpha| \le m$ (due to the polynomial factors in R_φ and the fact that $\psi = 1$ on a neighborhood of 0). Thus Lemma 3.4.1 gives $\langle T, \widetilde{\varphi}\rangle = 0$ and therefore

$$\langle T, \varphi\rangle = \sum_{|\gamma|\le m} \frac{1}{\gamma!}\,\langle T, x^\gamma\psi\rangle\,\partial^\gamma\varphi(0).$$

Setting $c_\gamma = \langle T, x^\gamma\psi\rangle/\gamma!$ yields the claim. □

3.4.3 Example. As an application of Theorem 3.4.2, let us determine all $T \in \mathcal{D}'(\mathbb{R})$ with $x^m \cdot T = 0$ for some $m \in \mathbb{N}$. First we observe that we necessarily have $\operatorname{supp} T \subseteq \{0\}$, so Theorem 3.4.2 implies that

$$T = \sum_{j=0}^{k} a_j\delta^{(j)}$$

for some $k \in \mathbb{N}_0$ and $a_j \in \mathbb{C}$, $j = 1, \ldots, k$, $a_k \neq 0$. Now if $k \ge m$, then for $\varphi(x) := x^{k-m}\chi(x)$, $\chi \in \mathcal{D}(\mathbb{R})$, $\chi \equiv 1$ near 0 we get

$$0 = \langle x^m \cdot T, \varphi\rangle = \langle T, x^k\chi\rangle = \sum_{j=0}^{k} a_j(-1)^j(x^k\chi)^{(j)}(0) = (-1)^k k! a_k \neq 0,$$

a contradiction. Conversely, $x^m \cdot \delta^{(j)} = 0$ for $j = 0, \ldots, m-1$, so

$$\{T \in \mathcal{D}'(\mathbb{R}) \mid x^m \cdot T = 0\} = \Big\{\sum_{j=0}^{m-1} a_j\delta^{(j)} \mid a_j \in \mathbb{C}\Big\}.$$

An essential property of smooth functions on intervals is that their derivative vanishes if and only if they are constant. An analogous result is true for distributions:

3.4.4 Theorem. *Let $a < b \in \mathbb{R}$ and $\Omega = (a, b) \subseteq \mathbb{R}$. Then*

(i) $\forall T \in \mathcal{D}'(\Omega)$: $T' = 0 \Leftrightarrow T \in \mathbb{C} \subseteq \mathcal{D}'(\Omega)$.

(ii) $\forall m \in \mathbb{N}\ \forall T \in \mathcal{D}'(\Omega)$: $T^{(m)} = 0 \Leftrightarrow T = \sum_{j=0}^{m-1} c_j x^j$, $c_j \in \mathbb{C}$.

(iii) The differential operator

$$\frac{d}{dx} : \mathcal{D}'(\Omega) \to \mathcal{D}'(\Omega)$$

has a sequentially continuous linear right-inverse $R : \mathcal{D}'(\Omega) \to \mathcal{D}'(\Omega)$, i.e., $R(T)' = T$ for all $T \in \mathcal{D}(\Omega)$. Thus $R(T)$ is a primitive *of T.*

(iv) *Any linear constant coefficient ordinary differential operator $P\big(\frac{d}{dx}\big)$ has only classical solutions:*

$$T \in \mathcal{D}'(\Omega),\ P\Big(\frac{d}{dx}\Big)T = 0 \Rightarrow T \in C^\infty(\Omega).$$

Proof. (i) Consider the subset $H := \{\varphi' \mid \varphi \in \mathcal{D}(\Omega)\}$ of $\mathcal{D}(\Omega)$. Then H is the following hyperplane:

$$H = \{\psi \in \mathcal{D}(\Omega) \mid \int_a^b \psi(x)\,dx = \langle 1, \psi\rangle = 0\} = \ker(1).$$

Indeed, $\int_a^b \varphi'(x)\,dx = 0$ is clear. Conversely, if $\int_a^b \psi(x)\,dx = 0$, then set $\varphi(x) := \int_a^x \psi(t)\,dt$ to obtain $\varphi' = \psi$ and $\varphi \in \mathcal{D}(\Omega)$.
Now fix any $\chi \in \mathcal{D}(\Omega)$ with $\langle 1, \chi\rangle = 1$ and define

$$\mathrm{pr} : \mathcal{D}(\Omega) \to H, \quad \mathrm{pr}(\varphi) := \varphi - \langle 1, \varphi\rangle \cdot \chi.$$

Then since T vanishes on H by assumption,

$$\langle T, \varphi\rangle = \langle T, \mathrm{pr}(\varphi) + \langle 1, \varphi\rangle \cdot \chi\rangle = \langle 1, \varphi\rangle\langle T, \chi\rangle = \langle\langle T, \chi\rangle, \varphi\rangle,$$

i.e., $T = \langle T, \chi\rangle \in \mathbb{C}$.
(ii) follows from (i) by induction.
(iii) For χ and pr as in (i), set

$$R : \mathcal{D}'(\Omega) \to \mathcal{D}'(\Omega) : \quad \langle R(T), \varphi\rangle := \left\langle T, -\int_a^x (\mathrm{pr}\varphi)(t)\,dt\right\rangle.$$

Note that $\varphi \mapsto -\int_a^x (\mathrm{pr}\varphi)(t)\,dt$ is sequentially continuous, so indeed $R(T) \in \mathcal{D}'(\Omega)$. Also,

$$\begin{aligned}\langle R(T)', \varphi\rangle &= -\langle R(T), \varphi'\rangle = \left\langle T, \int_a^x (\mathrm{pr}\varphi')(t)\,dt\right\rangle \\ &= \left\langle T, \int_a^x \varphi'(t)\,dt - \underbrace{\langle 1, \varphi'\rangle}_{=0} \int_a^x \chi(t)\,dt\right\rangle = \langle T, \varphi\rangle.\end{aligned}$$

(iv) Any such operator factorizes into a composition of terms of the form $P\big(\frac{d}{dx}\big) = \frac{d}{dx} - \lambda$. Using induction it therefore suffices to verify the claim for this P. Here,

$$0 = \Big(\frac{d}{dx} - \lambda\Big)T = e^{\lambda x}\frac{d}{dx}(e^{-\lambda x}T),$$

so (i) gives $e^{-\lambda x}T = c \in \mathbb{C}$, i.e., $T = ce^{\lambda x} \in C^\infty(\Omega)$. $\square$

3.5 Hypoellipticity and fundamental solutions

We now want to apply the results of the previous section to verify fundamental solutions of some important differential operators.

3.5.1 Definition.

(i) The singular support *of a distribution* $T \in \mathcal{D}'(\Omega)$ *is the complement of the largest open set* $U \subseteq \Omega$ *such that* $T|_U \in C^\infty(U)$.

(ii) A differential operator $P(\partial) = \sum_{|\alpha|\le m} a_\alpha \partial^\alpha$, $a_\alpha \in \mathbb{C}$, *is called* hypoelliptic *if*

$$\operatorname{singsupp}(P(\partial)T) = \operatorname{singsupp}(T)$$

for each $T \in \mathcal{D}'(\mathbb{R}^n)$.

Note that the open set U from Definition 3.5.1 does indeed exist and equals the union of all open subsets of Ω restricted to which T is smooth.

3.5.2 Examples. (i) $\operatorname{singsupp} Y^{(m)} = \{0\}$ for each $m \in \mathbb{N}_0$.

(ii) Any fundamental solution of a hypoelliptic operator must have singular support $\{0\}$. As we shall see later (Theorem 6.2.5), the converse is true as well: if a differential operator has such a fundamental solution it must be hypoelliptic. Thus by Theorem 3.2.6, any constant coefficient ordinary differential operator is hypoelliptic.

Let us now use the jump formula from Lemma 3.3.5 to verify fundamental solutions.

3.5.3 Example. *Fundamental solution of the Laplace operator*
The Laplace operator on $\mathbb{R}^n$

$$\Delta_n = \sum_{i=1}^n \partial_i^2$$

has the following fundamental solution:

$$E = \begin{cases} \frac{1}{2\pi}\log|x| & \text{if } n = 2 \\ c_n|x|^{2-n} & \text{if } n \neq 2 \end{cases} \qquad c_n = \frac{\Gamma\left(\frac{n}{2}\right)}{(2-n)2\pi^{n/2}}.$$

E is locally integrable on $\mathbb{R}^n$. Since E is smooth outside the origin, Δ_n is hypoelliptic. Also, it can easily be verified that $\Delta_n E = 0$ on $\mathbb{R}^n \setminus \{0\}$. We first note that $Y(|x|-R)\cdot E \to E$ as $R \to 0+$ in $\mathcal{D}'(\mathbb{R}^n)$, as can immediately be verified by application to a test function. Since differentiation is sequentially continuous (by Proposition 3.1.2), this implies that

$$\Delta_n(E) = \lim_{R\to 0+} \Delta_n(Y(|x|-R)\cdot E),$$

and this is what we will use to verify that indeed $\Delta_n(E) = \delta$. Next we observe that $\Delta_n = \operatorname{tr}(\nabla\nabla^\top)$. So we have to determine $\nabla\nabla^\top E$, which we will do using Lemma 3.3.5. Since in all of the above cases we can write E in the form $E = g(|x|)$, we may use (3.3.8) to calculate $\operatorname{tr}(\nabla\nabla^\top(Y(|x|-R)\cdot E))$. Noting that $\operatorname{tr}(xx^\top) = |x|^2$ we get

$$\begin{aligned}\Delta_n(Y(|x|-R)\cdot E) &= \operatorname{tr}(\nabla\nabla^\top(Y(|x|-R)\cdot E)) \\ &= Y(|x|-R)\Big(\frac{n-1}{|x|}g'(|x|) + g''(|x|)\Big) \\ &+ g(R)D_{R\cdot S^{n-1}}(1) + \Big(g'(R) - \cancel{\frac{n}{R}g(R)} + \cancel{\frac{n}{R}g(R)}\Big)S_{R\cdot S^{n-1}}(1).\end{aligned}$$

As can easily be checked,

$$\frac{n-1}{|x|}g'(|x|) + g''(|x|) = 0$$

in both cases of the definition of E, so we are left with

$$\Delta_n(Y(|x|-R)\cdot E) = g(R)D_{R\cdot S^{n-1}}(1) + g'(R)S_{R\cdot S^{n-1}}(1).$$

Now for any $\varphi \in \mathcal{D}(\mathbb{R}^n)$,

$$\begin{aligned}\langle g(R)D_{R\cdot S^{n-1}}(1), \varphi\rangle &= -\int_{R\cdot S^{n-1}} g(R)\partial_\nu \varphi(y)\, d\sigma(y) \\ &\underset{y=R\omega}{=} -g(R)R^{n-1}\int_{S^{n-1}} \omega^\top \nabla\varphi(R\omega)\, d\sigma(\omega) \to 0 \ (R\to 0+).\end{aligned}$$

Analogously,

$$\langle g'(R)S_{R\cdot S^{n-1}}(1), \varphi\rangle = g'(R)R^{n-1}\int_{S^{n-1}} \varphi(R\omega)\, d\sigma(\omega).$$

For $n \neq 2$, observe that $g'(R) = (2-n)c_n R^{1-n}$ and that $(2-n)c_n \int_{S^{n-1}} d\sigma = 1$. Therefore,

$$g'(R)R^{n-1}\int_{S^{n-1}} \varphi(R\omega)\, d\sigma(\omega) - \varphi(0) = (2-n)c_n \int_{S^{n-1}} (\varphi(R\omega)-\varphi(0))\, d\sigma(\omega) \to 0$$

as $R \to 0+$, and analogously for $n = 2$. Altogether, we arrive at

$$\Delta_n E = \lim_{R\to 0+} \Delta_n(Y(|x|-R)\cdot E) = \delta,$$

as claimed.

3.5.4 Example. *Fundamental solution of the Cauchy–Riemann operator*
We claim that the Cauchy–Riemann operator

$$P(\partial) = \partial_1 + i\partial_2$$

on $\mathbb{R}^2$ has the fundamental solution

$$E = \frac{1}{2\pi(x_1+ix_2)} = \frac{1}{2\pi z}.$$

To show this, noting that $E \in L^1_{\text{loc}}(\mathbb{R}^2)$, we proceed analogously to the previous example. Let $h(x) := |x| - R$, so $M = h^{-1}(0) = R\cdot S^1$, and

$$\nu(x) = \frac{\nabla h(x)}{|\nabla h(x)|} = \frac{1}{R}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \qquad (x \in R\cdot S^1).$$

By the jump formula (3.3.1) we have

$$\begin{aligned}\partial_1(Y(|x|-R)\cdot E) &= Y(h)\partial_1 E + S_M(\nu_1\cdot E)\\ \partial_2(Y(|x|-R)\cdot E) &= Y(h)\partial_2 E + S_M(\nu_2\cdot E),\end{aligned}$$

so

$$(\partial_1+i\partial_2)(Y(|x|-R)\cdot E) = Y(h)(\underbrace{\partial_1 E + i\partial_2 E}_{=0}) + S_M((\nu_1+i\nu_2)E) = \frac{1}{2\pi R}S_{RS^1}(1).$$

Moreover,

$$\left\langle \frac{1}{2\pi R}S_{RS^1}(1), \varphi\right\rangle - \varphi(0) = \frac{1}{2\pi}\int_{S^1} (\varphi(R\omega)-\varphi(0))\, d\sigma(\omega) \to 0 \quad (R\to 0+).$$

Consequently,

$$(\partial_1+i\partial_2)E = \lim_{R\to 0+} (\partial_1+i\partial_2)(Y(|x|-R)\cdot E) = \delta,$$

as claimed.

3.5.5 Example. *Fundamental solution of the heat operator*
Here we show that the heat operator

$$P(\partial) = \partial_t - \Delta_n$$

possesses as a fundamental solution the function

$$E(t,x) = \frac{Y(t)}{(4\pi t)^{n/2}} e^{-\frac{|x|^2}{4t}} \in L^1_{\mathrm{loc}}(\mathbb{R}^{n+1}_{t,x}). \tag{3.5.1}$$

This time, for any $\varepsilon > 0$, we set $h(t,x) := t - \varepsilon$, and $h^{-1}(0) =: M_\varepsilon = \{(t,x) \in \mathbb{R}^{n+1} \mid t = \varepsilon\}$. Then $\nu(t,x) = (1,0,\dots,0)^\top$ is constant. We have

$$\Delta_n(Y(t-\varepsilon)\cdot E) = Y(t-\varepsilon)\Delta_n E.$$

Also, the t-component of equation (3.3.1) gives

$$\partial_t(Y(t-\varepsilon)\cdot E) = Y(t-\varepsilon)\partial_t E + S_{M_\varepsilon}(E).$$

Since $\partial_t E - \Delta_n E = 0$ for $t \neq 0$, we obtain

$$(\partial_t - \Delta_n)(Y(t-\varepsilon)\cdot E) = S_{M_\varepsilon}(E).$$

Consequently,

$$(\partial_t - \Delta_n)(E) = \lim_{\varepsilon\to 0+} (\partial_t - \Delta_n)(Y(t-\varepsilon)\cdot E) = \lim_{\varepsilon\to 0+} S_{M_\varepsilon}\Big(\frac{e^{-\frac{|x|^2}{4\varepsilon}}}{(4\pi\varepsilon)^{n/2}}\Big).$$

Finally, for any $\varphi \in \mathcal{D}(\mathbb{R}^{n+1})$,

$$\begin{aligned}
\Big\langle S_{M_\varepsilon}\Big(\frac{e^{-\frac{|x|^2}{4\varepsilon}}}{(4\pi\varepsilon)^{n/2}}\Big), \varphi\Big\rangle &= (4\pi\varepsilon)^{-n/2} \int_{\mathbb{R}^n} \varphi(\varepsilon,x) e^{-\frac{|x|^2}{4\varepsilon}}\, dx \\
&\underset{x=2\varepsilon^{1/2}y}{=} \pi^{-n/2} \int_{\mathbb{R}^n} \varphi(\varepsilon, 2\sqrt{\varepsilon}y) e^{-|y|^2}\, dy \\
&\to \pi^{-n/2}\varphi(0) \int_{\mathbb{R}^n} e^{-|y|^2}\, dy = \varphi(0) = \langle \delta, \varphi\rangle
\end{aligned}$$

as $\varepsilon \to 0+$ by dominated convergence.
An alternative proof, not based on jump formulas, goes as follows:

$$\begin{aligned}
\langle (\partial_t - \Delta_n)E, \varphi\rangle &= -\int E(t,x)(\partial_t\varphi + \Delta_n\varphi)\, dxdt \\
&= -\lim_{\varepsilon\to 0+} \int_{t>\varepsilon} E(t,x)(\partial_t\varphi + \Delta_n\varphi)\, dxdt.
\end{aligned}$$

Here,

$$\begin{aligned}
\int_\varepsilon^\infty E(t,x)\partial_t\varphi\, dt &= -E(\varepsilon,x)\varphi(\varepsilon,x) - \int_\varepsilon^\infty \partial_t E(t,x)\varphi(t,x)\, dt \\
\int E(t,x)\Delta_n\varphi\, dx &= \int \Delta_n E(t,x)\varphi(t,x)\, dx.
\end{aligned}$$

Thus since $(\partial_t - \Delta_n)E = 0$ on $\mathbb{R}\setminus\{0\}\times\mathbb{R}^n$,

$$\langle (\partial_t - \Delta_n)E, \varphi\rangle = \lim_{\varepsilon\to 0+} \int E(\varepsilon,x)\varphi(\varepsilon,x)\, dx,$$

which, as shown above, equals $\langle \delta, \varphi\rangle$.

To conclude this section, we want to determine a *fundamental solution of the wave operator on* $\mathbb{R}^4$.

The wave operator (or d'Alembert operator) on $\mathbb{R}^4$ with coordinates $(x_0, \dots, x_3) \equiv (x_0, x')$ has the form

$$P(\partial) = \Box = -\partial_0^2 + \partial_1^2 + \partial_2^2 + \partial_3^2 = -\partial_0^2 + \Delta_3.$$

Let $h : \mathbb{R}^4 \setminus \{0\} \to \mathbb{R}$, $h(x_0, x') := x_0^2 - |x'|^2$. Then $\nabla h \neq 0$ on $\mathbb{R}^4 \setminus \{0\}$, so h is a submersion. Let $\varphi \in \mathcal{D}(\mathbb{R}^4 \setminus \{0\})$. We first calculate (cf. Definition 2.4.6) $\langle h^*\delta, \varphi\rangle = \varphi_h(0)$. Noting that $h(x) < t$ means that $-\sqrt{t + |x'|^2} < x_0 < \sqrt{t + |x'|^2}$, we have

$$\begin{aligned}
\varphi_h(0) &= \partial_s|_{s=0} \int_{\{h(x)<s\}} \varphi(x)\,dx = \partial_s|_{s=0} \int\int_{-\sqrt{s+|x'|^2}}^{\sqrt{s+|x'|^2}} \varphi(x_0, x')\,dx_0 dx' \\
&= \frac{1}{2}\int \frac{\varphi(|x'|, x')}{|x'|}\,dx' + \frac{1}{2}\int \frac{\varphi(-|x'|, x')}{|x'|}\,dx' \qquad (3.5.2)\\
&=: \langle \delta_+(h), \varphi\rangle + \langle \delta_-(h), \varphi\rangle.
\end{aligned}$$

Thus

$$h^*\delta = \delta_+(h) + \delta_-(h). \tag{3.5.3}$$

By symmetry, it will suffice to analyze $\delta_+(h)$. Since $|x'|^{-1} \in L^1_{\mathrm{loc}}(\mathbb{R}^3)$ (as can be seen using polar coordinates), it follows that $\delta_+(h) \in \mathcal{D}'^0(\mathbb{R}^4)$. Moreover, $\mathrm{supp}\,(\delta_+(h)) \subseteq \Gamma^+$, where

$$\Gamma^+ = \{x \in \mathbb{R}^4 \mid x_0 = |x'|\} \tag{3.5.4}$$

is the *forward light cone*. As a first step we show:

3.5.6 Lemma. *On* $\mathbb{R}^4 \setminus \{0\}$, $\Box(h^*\delta) = 0$.

Proof. Since h is a submersion, by Proposition 3.1.5 (iii) we may use the chain rule to calculate the derivatives:

$$\begin{aligned}
\Box(h^*\delta) &= \Box(\delta\circ h) = -\partial_0^2(\delta \circ h) + \sum_{i=1}^3 \partial_i^2(\delta\circ h)\\
&= -\partial_0(\delta'(x_0^2 - |x'|^2)2x_0) - \sum_{i=1}^3 \partial_i(\delta'(x_0^2 - |x'|^2)2x_i)\\
&= -\delta''(x_0^2 - |x'|^2)4x_0^2 - 2\delta'(x_0^2 - |x'|^2) + \sum_{i=1}^3 \delta''(x_0^2 - |x'|^2)4x_i^2 - 6\delta'(x_0^2 - |x'|^2)\\
&= -4\delta''(x_0^2 - |x'|^2)\cdot\Big(x_0^2 - \sum_{i=1}^3 x_i^2\Big) - 8\delta'(x_0^2 - |x'|^2).
\end{aligned}$$

Now let $\varphi \in \mathcal{D}(\mathbb{R}^4 \setminus \{0\})$. Then

$$\begin{aligned}
\langle \Box(h^*\delta), \varphi\rangle &= -4\langle \delta''(x_0^2 - |x'|^2), (x_0^2 - |x'|^2)\cdot\varphi\rangle - 8\langle \delta'(x_0^2 - |x'|^2), \varphi\rangle\\
&\underset{2.4.6}{=} \langle \delta, -4[(x_0^2 - |x'|^2)\varphi]_h'' + 8\varphi_h'\rangle =: \langle \delta, -4\psi_h'' + 8\varphi_h'\rangle.
\end{aligned}$$

Here, by (2.4.4),

$$\begin{aligned}
\varphi_h(s) &= \partial_s \int_{\{x|h(x)<s\}} \varphi(x)\,dx = \partial_s \int\int_{-(|x'|^2+s)^{1/2}}^{(|x'|^2+s)^{1/2}} \varphi(x_0, x')\,dx_0 dx'\\
&= \int \frac{\varphi((|x'|^2+s)^{1/2}, x')}{2(|x'|^2+s)^{1/2}}\,dx' + \int \frac{\varphi(-(|x'|^2+s)^{1/2}, x')}{2(|x'|^2+s)^{1/2}}\,dx'
\end{aligned}$$

and

$$\psi_h(s) = [(x_0^2 - |x'|^2) \cdot \varphi]_h(s) = \partial_s \int \int_{-(|x'|^2+s)^{1/2}}^{(|x'|^2+s)^{1/2}} (x_0^2 - |x'|^2)\varphi(x_0, x')\, dx_0 dx'$$
$$= \int s \cdot \frac{\varphi((|x'|^2+s)^{1/2}, x')}{2(|x'|^2+s)^{1/2}}\, dx' + \int s \cdot \frac{\varphi(-(|x'|^2+s)^{1/2}, x')}{2(|x'|^2+s)^{1/2}}\, dx' = s\varphi_h(s),$$

so

$$\psi_h'(s) = \varphi_h(s) + s\varphi_h'(s), \quad \psi_h''(s) = 2\varphi_h'(s) + s\varphi_h''(s).$$

Thus, finally,

$$\langle \Box(h^*\delta), \varphi\rangle = \langle \delta(s), \cancel{-8\varphi_h'(s)} - 4s\varphi_h''(s) + \cancel{8\varphi_h'(s)}\rangle = 0.$$

□

Since $\operatorname{supp}(\delta_+ h) \cap \operatorname{supp}(\delta_- h) \subseteq \Gamma^+ \cap \Gamma^- = \{0\}$, Lemma 3.5.6, together with (3.5.3) implies that both $\Box(\delta_+(h))$ and $\Box(\delta_-(h))$ vanish on $\mathbb{R}^4 \setminus \{0\}$. Thus $\operatorname{supp}(\Box(\delta_+(h)))$ $\subseteq \{0\}$, and Theorem 3.4.2 implies the existence of some $m \in \mathbb{N}_0$ and $c_\alpha \in \mathbb{C}$ ($|\alpha| \le m$) such that

$$\Box(\delta_+(h)) = \sum_{|\alpha|\le m} c_\alpha \partial^\alpha \delta. \tag{3.5.5}$$

Next we note that $\delta_+(h)$ is homogeneous of degree -2: Indeed, by (3.5.2) we have

$$\langle \delta_+(h), \varphi\rangle = \frac{1}{2}\int \frac{\varphi(|x'|, x')}{|x'|}\, dx' \underset{y'=x'/c}{=} \frac{1}{2}\int \frac{\varphi(c|y'|, cy')}{c|y'|} c^3\, dy'$$
$$= c^2 \langle \delta_+(h)(x), \varphi(cx)\rangle = c^{4-2} \langle \delta_+(h)(x), \varphi(cx)\rangle,$$

so the claim follows from (2.4.2). By Lemma 3.1.6, therefore, $\Box(\delta_+(h))$ is homogeneous of degree -4. Also, from Example 2.4.4 (i) and Lemma 3.1.6, $\partial^\alpha\delta$ is homogeneous of degree $(-4-|\alpha|)$. This implies that in (3.5.5) we must have $c_\alpha = 0$ for all $|\alpha| > 0$, i.e.,

$$\Box(\delta_+(h)) = c\delta \tag{3.5.6}$$

and it remains to determine this c.

To this end, we first note that since $\operatorname{supp}(\delta_+(h)) \subseteq \Gamma^+$, $\langle \delta_+(h), \varphi\rangle$ is well defined for any $\varphi \in C^\infty(\mathbb{R}^4)$ with $\operatorname{supp}(\varphi) \cap \Gamma^+$ compact. Taking $\rho \in \mathcal{D}(\mathbb{R})$, $\varphi(x_0, x') := \rho(x_0)$ is such a φ. Hence

$$c\rho(0) = \langle c\delta, \varphi\rangle \underset{(3.5.6)}{=} \langle \delta_+(h), \Box(\varphi)\rangle = -\langle \delta_+(h)(x_0, x'), \rho''(x_0)\rangle$$
$$\underset{(3.5.2)}{=} -\frac{1}{2}\int \frac{\rho''(|x'|)}{|x'|}\, dx' = -\frac{1}{2}4\pi \int_0^\infty \rho''(r) r\, dr = -2\pi\rho(0),$$

resulting in $c = -2\pi$. Analogous considerations apply to $\delta_-(h)$, so that altogether we have proved:

3.5.7 Theorem. *Fundamental solutions of the wave operator $\Box$ on $\mathbb{R}^4$ are given by*

$$\langle E^+, \varphi\rangle = -\frac{1}{4\pi}\int \frac{\varphi(|x'|, x')}{|x'|}\, dx' \quad (\varphi \in \mathcal{D}(\mathbb{R}^4)) \tag{3.5.7}$$

and

$$\langle E^-, \varphi\rangle = -\frac{1}{4\pi}\int \frac{\varphi(-|x'|, x')}{|x'|}\, dx' \quad (\varphi \in \mathcal{D}(\mathbb{R}^4)) \tag{3.5.8}$$

Furthermore, $\operatorname{supp} E^+ = \Gamma^+$*, and* $\operatorname{supp} E^- = \Gamma^-$.

3.6 Distribution-valued functions

3.6.1 Definition.

(i) Let X be a metric space. A map $f : X \to \mathcal{D}'(\Omega)$ is called continuous *if for each $\varphi \in \mathcal{D}(\Omega)$ the map*

$$\langle f, \varphi\rangle : X \to \mathbb{C}, \quad \lambda \mapsto \langle f(\lambda), \varphi\rangle$$

is continuous (i.e., f is continuous with respect to the weak topology). The vector space of all continuous functions $f : X \to \mathcal{D}'(\Omega)$ is denoted by $C(X, \mathcal{D}'(\Omega))$.

(ii) If $\emptyset \neq U \subseteq \mathbb{R}^l$ is open and $m \in \mathbb{N}_0 \cup \{\infty\}$, then $f : U \to \mathcal{D}'(\Omega)$ is called m times continuously differentiable *if this holds in the weak sense, i.e.,*

$$\langle f, \varphi\rangle : U \to \mathbb{C}, \quad \lambda \mapsto \langle f(\lambda), \varphi\rangle \in C^m(U)$$

for each $\varphi \in \mathcal{D}(\Omega)$. The vector space of all such mappings is denoted by $C^m(U, \mathcal{D}'(\Omega))$.

(iii) Analogously, $C^m([0,\infty), \mathcal{D}'(\Omega))$ consists of all $f \in C^m((0,\infty), \mathcal{D}'(\Omega))$ such that $f^{(j)}(t)$ converges in $\mathcal{D}'(\Omega)$ as $t \to 0$ for each $j = 0, \dots, m$.

As a first application of these notions we obtain the following result on fundamental solutions:

3.6.2 Proposition. *Let $P_1(\partial), \dots, P_l(\partial)$ be linear differential operators with constant coefficients and let $U \subseteq \mathbb{R}^l$ be open, $m \in \mathbb{N}$, and $E \in C^m(U, \mathcal{D}'(\mathbb{R}^n))$ such that $E(\lambda)$ is a fundamental solution of $\prod_{j=1}^l (P_j(\partial) - \lambda_j)$. Then $\frac{1}{\alpha!}\partial_\lambda^\alpha E(\lambda)$ is a fundamental solution of $\prod_{j=1}^l (P_j(\partial) - \lambda_j)^{\alpha_j+1}$ for $\lambda \in U$ and $\alpha \in \mathbb{N}_0^l$ with $|\alpha| \leq m$.*

Proof. This follows exactly as in the proof of Theorem 3.2.6, which treats the case $P_j(\partial) = \frac{d}{dx}$. □

3.6.3 Definition. *Let $\emptyset \neq U$ be open in $\mathbb{C}^l$.*

(i) A map $f : U \to \mathcal{D}'(\Omega)$ is called holomorphic *(or* analytic*) if*

$$\langle f, \varphi\rangle : U \to \mathbb{C}, \quad \lambda \mapsto \langle f(\lambda), \varphi\rangle$$

is holomorphic for each $\varphi \in \mathcal{D}(\Omega)$.

(ii) If $l = 1$ then f is called meromorphic *if f is defined and holomorphic in $U \setminus D$ for some discrete set $D \subseteq U$, and if for each $\lambda_0 \in D$ there exists some $k \in \mathbb{N}_0$ such that $(\lambda - \lambda_0)^k f(\lambda)$ can be continued holomorphically to λ_0 (in the weak sense). Also the* residue *is defined in the weak sense:*

$$\langle \mathrm{Res}_{\lambda=\lambda_0} f(\lambda), \varphi\rangle := \mathrm{Res}_{\lambda=\lambda_0} \langle f(\lambda), \varphi\rangle,$$

for $\varphi \in \mathcal{D}(\Omega)$.

Let us now look more closely at some basic properties of distributions depending on a real parameter. To this end, we first consider test functions depending on such a parameter.

3.6.4 Proposition. *Let $n_1, n_2 \in \mathbb{N}$ and $\Omega_j \subseteq \mathbb{R}^{n_j}$ ($j = 1, 2$) be open subsets. Assume that $\varphi \in \mathcal{C}^\infty(\Omega_1 \times \Omega_2)$ satisfies the following:*

$$\forall y' \in \Omega_2 \; \exists \text{ neighborhood } U(y') \text{ of } y' \text{ in } \Omega_2 \; \exists K(y') \Subset \Omega_1 : \qquad \operatorname{supp}\big(\varphi(.,y)\big) \subseteq K(y') \quad \forall y \in U(y').$$

(i.e., the support of the map $x \mapsto \varphi(x,y)$ is contained in $K(y')$).
Then for any $T \in \mathcal{D}'(\Omega_1)$ we have

$$y \mapsto \langle T(x), \varphi(x,y)\rangle := \langle T, \varphi(.,y)\rangle \in \mathcal{C}^\infty(\Omega_2)$$

and for all $\alpha \in \mathbb{N}_0^{n_2}$

$$\partial^\alpha \langle T, \varphi(.,y)\rangle = \langle T, \partial_y^\alpha \varphi(.,y)\rangle. \tag{3.6.1}$$

3.6.5 Remark.

(i) Note that for a regular distribution $f \in L^1_{\text{loc}}(\Omega_1)$, (3.6.1) reads

$$\partial_y^\alpha \int_{\Omega_1} f(x)\varphi(x,y)\,dx = \int_{\Omega_1} f(x)\partial_y^\alpha \varphi(x,y)\,dx,$$

hence it includes a variant of the classical theorem on "differentiation under the integral".

(ii) To be prepared for the proof we recall a basic estimate, which is a consequence of the mean value theorem (cf. [Hör09, 18.18, equation (18.13)], or [For05, p. 6, Corollar zu Satz 5]): If $f \in C^1(\Omega)$ and the line segment $\overline{xy}$ joining $x, y \in \Omega$ lies entirely in Ω, then we have

$$|f(x) - f(y)| \leq \|Df\|_{L^\infty(\overline{xy})}\, |x-y|. \tag{3.6.2}$$

Proof of Proposition 3.6.4. By hypothesis we have for any $y \in \Omega_2$ that $x \mapsto \varphi(x,y)$ belongs to $\mathcal{D}(\Omega_1)$. Thus we may define

$$\Psi(y) := \langle T(x), \varphi(x,y)\rangle \qquad (y \in \Omega_2). \tag{3.6.3}$$

Let $y' \in \Omega_2$ and choose $U(y')$ and $K(y')$ as in the hypothesis. Let $r > 0$ be such that $B_r(y') \subseteq U(y')$.

- Ψ is continuous: For any $h \in \mathbb{R}^{n_2}$ with $|h| < r$ set

  $$\varphi_h(x,y') := \varphi(x, y'+h) - \varphi(x,y').$$

  Then we may write

  $$\Psi(y'+h) - \Psi(y') = \langle T(x), \varphi_h(x,y')\rangle$$

  and conclude that it suffices to show $\varphi_h(.,y') \to 0$ in $\mathcal{D}(\Omega_1)$ as $h \to 0$.

  From the hypothesis we have $\operatorname{supp}(\varphi_h(.,y')) \subseteq K(y') \Subset \Omega_1$ for all h with $|h| < r$. Furthermore, if $\beta \in \mathbb{N}_0^{n_1}$ we may apply (3.6.2) to the function $y \mapsto \partial_x^\beta \varphi(x,y)$ and obtain

  $$\begin{aligned}|\partial_x^\beta \varphi_h(x,y')| &= |\partial_x^\beta \varphi(x,y'+h) - \partial_x^\beta \varphi(x,y')| \\ &\leq \|D_y \partial_x^\beta \varphi\|_{L^\infty(K(y') \times \overline{B_r(y')})} \cdot |h| \;\to\; 0 \qquad (h \to 0).\end{aligned}$$

- Ψ is continuously differentiable: Let e_j denote the jth standard basis vector in $\mathbb{R}^{n_2}$ $(1 \leq j \leq n_2)$ and define for $0 < \varepsilon < r$

$$\chi_\varepsilon(x, y') := \frac{\varphi(x, y' + \varepsilon e_j) - \varphi(x, y')}{\varepsilon} - \partial_{y_j}\varphi(x, y') \qquad (x \in \Omega_1).$$

  By (3.6.3) we obtain

$$\frac{\Psi(y' + \varepsilon e_j) - \Psi(y')}{\varepsilon} - \langle T(x), \partial_{y_j}\varphi(x, y')\rangle = \langle T(x), \chi_\varepsilon(x, y')\rangle$$

  and thus recognize that it suffices to prove $\chi_\varepsilon(., y') \to 0$ in $\mathcal{D}(\Omega_1)$ as $\varepsilon \to 0$, since we know that $y' \mapsto \langle T(x), \partial_{y_j}\varphi(x, y')\rangle$ is continuous (by an application of the first part of this proof to $\partial_{y_j}\varphi$ in place of φ).

  From the hypothesis we get $\operatorname{supp}(\chi_\varepsilon(., y')) \subseteq K(y') \Subset \Omega_1$ for all $\varepsilon \in]0, r[$. Furthermore, if $\beta \in \mathbb{N}_0^{n_1}$ we may apply the mean value theorem[1] to the function $\varepsilon \mapsto \partial_x^\beta \varphi(x, y' + \varepsilon e_j)$ and obtain with some $\varepsilon_1 \in [0, \varepsilon]$

$$\partial_x^\beta \chi_\varepsilon(x, y') = \partial_{y_j}\partial_x^\beta \varphi(x, y' + \varepsilon_1 e_j) - \partial_{y_j}\partial_x^\beta \varphi(x, y').$$

  Hence another application of (3.6.2) to the function $y \mapsto \partial_{y_j}\partial_x^\beta\varphi(x, y)$ now gives

$$|\partial_x^\beta \chi_\varepsilon(x, y')| \leq C(\beta, \varphi)\, \varepsilon_1 \;\to\; 0 \qquad (0 \leq \varepsilon_1 \leq \varepsilon \to 0),$$

  where the constant $C(\beta, \varphi)$ equals the maximum of $|D_y(\partial_{y_j}\partial_x^\beta\varphi)|$ on $K(y') \times \overline{B_r(y')}$.

  In particular, we obtain the special case of (3.6.1) when $\alpha = e_j$, i.e.,

$$\partial_j \Psi(y') = \langle T(x), \partial_{y_j}\varphi(x, y')\rangle.$$

- $\Psi \in \mathcal{C}^\infty(\Omega_2)$: Proceeding inductively we obtain that $\partial^\alpha \Psi$ is continuously differentiable and thereby that (3.6.1) holds for all $\alpha \in \mathbb{N}_0^{n_2}$.

$\square$

3.6.6 Corollary.

(i) If $T \in \mathcal{D}'(\Omega_1)$ and $\varphi \in \mathcal{D}(\Omega_1 \times \Omega_2)$, then the function $y \mapsto \langle T, \varphi(., y)\rangle$ belongs to $\mathcal{D}(\Omega_2)$ and (3.6.1) *holds.*

(ii) If $T \in \mathcal{E}'(\Omega_1)$ and $\varphi \in \mathcal{C}^\infty(\Omega_1 \times \Omega_2)$, then the function $y \mapsto \langle T, \varphi(., y)\rangle$ belongs to $\mathcal{C}^\infty(\Omega_2)$ and (3.6.1) *holds.*

Proof. (i) The hypothesis of Proposition 3.6.4 is satisfied with $U(y') = \Omega_2$ and $K(y') = \pi_1(\operatorname{supp}(\varphi))$, where π_1 denotes the projection $\Omega_1 \times \Omega_2 \to \Omega_1$, $(x, y) \mapsto x$.

(ii) As in the proof of Theorem 2.3.5 we may choose a suitable cut-off ρ over a neighborhood of $\operatorname{supp}(T)$. Then the function Ψ defined by the action of T on $\varphi(., y)$ is given by

$$\Psi(y) = \langle T, \rho(.)\varphi(., y)\rangle.$$

Hence the assumptions of Proposition 3.6.4 hold upon taking $U(y') = \Omega_2$ and $K(y') = \operatorname{supp}(\rho)$. $\square$

[1] Note that by splitting into real- and imaginary part we may w.l.o.g. assume that φ is real-valued, so the mean value theorem applies.

3.6.7 Example. Let $\psi \in \mathcal{D}(\mathbb{R}^n)$ and define the function $\varphi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ by $\varphi(x, y) := \psi(x + y)$. Then $\varphi \in \mathcal{C}^\infty(\mathbb{R}^n \times \mathbb{R}^n)$ and satisfies the hypothesis of Proposition 3.6.4, e.g. with $U(y') := B_1(y')$ and $K(y') := \operatorname{supp}(\psi) - \overline{B_1(y')} = \{z - y \mid z \in \operatorname{supp}(\psi), y \in \overline{B_1(y')}\}$.

Thus, for any $T \in \mathcal{D}'(\mathbb{R}^n)$ the map $y \mapsto \langle T(x), \psi(x + y)\rangle$ is smooth $\mathbb{R}^n \to \mathbb{C}$ and

$$\begin{aligned}
\langle \partial_j T, \psi\rangle = -\langle T, \partial_j \psi\rangle = -\langle T, \partial_{y_j} \varphi(.,0)\rangle &\overset{\overset{(3.6.1)}{\downarrow}}{=} -\partial_j \langle T, \varphi(., y)\rangle \,|_{y=0} \\
&= -\lim_{\varepsilon \to 0} \frac{\langle T(x), \psi(x - \varepsilon e_j)\rangle - \langle T(x), \psi(x)\rangle}{-\varepsilon} \\
&= \lim_{\varepsilon \to 0} \frac{\langle T(x), \psi(x - \varepsilon e_j)\rangle - \langle T(x), \psi(x)\rangle}{\varepsilon} \\
&= \lim_{\varepsilon \to 0} \langle \frac{T(x + \varepsilon e_j) - T(x)}{\varepsilon}, \psi(x)\rangle,
\end{aligned}$$

confirming Proposition 3.1.2.

Chapter 4

Tensor product and convolution

4.1 Tensor product of distributions

To begin with, let us first look at the tensor product of functions. Thus let $n_1, n_2 \in \mathbb{N}$ and $\Omega_j \subseteq \mathbb{R}^{n_j}$ $(j = 1, 2)$ be open subsets. For functions $f \in \mathcal{C}^\infty(\Omega_1)$ and $g \in \mathcal{C}^\infty(\Omega_2)$ we define the *tensor product* $f \otimes g \in \mathcal{C}^\infty(\Omega_1 \times \Omega_2)$ by

$$f \otimes g\,(x, y) := f(x)\,g(y) \qquad (x \in \Omega_1, y \in \Omega_2).$$

We may consider $f \otimes g$ as a regular distribution on $\Omega_1 \times \Omega_2$ whose action on a test function $\Phi \in \mathcal{D}(\Omega_1 \times \Omega_2)$ is defined by

$$\begin{aligned}\langle f \otimes g, \Phi\rangle &= \int_{\Omega_1 \times \Omega_2} f(x)g(y)\Phi(x, y)\,d(x, y)\\ &= \int_{\Omega_2} g(y)\Big(\int_{\Omega_1} f(x)\Phi(x, y)\,dx\Big)\,dy = \langle g(y), \langle f(x), \Phi(x, y)\rangle\rangle.\end{aligned}$$

In particular, if $\Phi(x, y) = \varphi(x)\psi(y)$, i.e. $\Phi = \varphi \otimes \psi$, with $\varphi \in \mathcal{D}(\Omega_1)$ and $\psi \in \mathcal{D}(\Omega_2)$, then we obtain

$$\langle f \otimes g, \varphi \otimes \psi\rangle = \langle f, \varphi\rangle\langle g, \psi\rangle. \tag{4.1.1}$$

Our aim is to extend the tensor product to distributions in such a way that the analogue of (4.1.1) holds for all test functions φ, ψ and determines the distributional tensor product uniquely.

The first step will be to show that the linear combinations of all elements of the form $\varphi \otimes \psi$ are dense in $\mathcal{D}(\Omega_1 \times \Omega_2)$.

4.1.1 Lemma. *Let M denote the subspace of $\mathcal{D}(\Omega_1 \times \Omega_2)$ defined by the linear span of the set*

$$M_0 := \{\varphi \otimes \psi \mid \varphi \in \mathcal{D}(\Omega_1), \psi \in \mathcal{D}(\Omega_2)\}.$$

Then M is dense in $\mathcal{D}(\Omega_1 \times \Omega_2)$.

Proof. It suffices to consider sums of elements in M_0 to generate all of M, since scalar factors can always be subsumed into one of the functions. Let $\Phi \in \mathcal{D}(\Omega_1 \times \Omega_2)$. We will show that there exist sequences (φ_j) in $\mathcal{D}(\Omega_1)$ and (ψ_j) in $\mathcal{D}(\Omega_2)$ such that

$$\sum_{j=0}^{m} \varphi_j \otimes \psi_j \;\to\; \Phi \qquad \text{in } \mathcal{D}(\Omega_1 \times \Omega_2) \text{ as } m \to \infty.$$

We first cover the support of Φ by finitely many products $U_j \times V_j$, where $U_j \subseteq \Omega_1$ and $V_j \subseteq \Omega_2$ are cubes in $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$, respectively. Then using a partition of unity $\chi_1, \ldots, \chi_N$ subordinate to this cover and considering $\chi_i \cdot \Phi$ instead of Φ, we can reduce the claim to the case where Ω_i is itself such a cube ($i = 1, 2$). Moreover, after an appropriate translation and scaling we may even assume that $\operatorname{supp}(\Phi) \subseteq (0,1)^{n_1+n_2}$ and that $\Omega_l = (0,1)^{n_l}$ ($l = 1, 2$).

Setting $n = n_1 + n_2$ and $I := (0,1)^n$ we now claim that for any $\Phi \in \mathcal{D}(I)$ we can find n sequences $(\mu_{j,1})_{j\in\mathbb{N}_0}, \ldots, (\mu_{j,n})_{j\in\mathbb{N}_0}$ in $\mathcal{D}((0,1))$ such that putting

$$\Phi_m(x_1, \ldots, x_n) := \sum_{j=0}^{m} \mu_{j,1}(x_1) \cdots \mu_{j,n}(x_n) \qquad ((x_1, \ldots, x_n) \in I)$$

we obtain

$$\Phi_m \to \Phi \quad \text{in } \mathcal{D}(I). \tag{4.1.2}$$

Assuming the claim to be true for the moment, we first show how it implies the statement of the lemma: we simply set

$$\begin{aligned} \varphi_j(x_1, \ldots, x_{n_1}) &:= \mu_{j,1}(x_1) \cdots \mu_{j,n_1}(x_{n_1}), \\ \psi_j(y_1, \ldots, y_{n_2}) &:= \mu_{j,n_1+1}(y_1) \cdots \mu_{j,n_1+n_2}(y_{n_2}), \end{aligned}$$

then $\sum_{j=0}^{m} \varphi_j \otimes \psi_j = \Phi_m$ and (4.1.2) completes the proof of the lemma.

Returning now to the proof of the claim, we first extend Φ periodically to $\mathbb{R}^n$. This extension is in $\mathcal{C}^\infty(\mathbb{R}^n)$ since $\operatorname{supp}(\Phi)$ has positive distance from the boundary ∂I. Thus we set $\Phi(x + k) = \Phi(x)$ for all $k \in \mathbb{Z}^n$. Now we may expand Φ into a Fourier series

$$\Phi(x) = \sum_{k\in\mathbb{Z}^n} c_k e^{2\pi i (k,x)} \qquad (x \in I),$$

where the Fourier coefficients are given by $c_k = \int_I \Phi(x) e^{-2\pi i (k,x)}\, dx$ ($k \in \mathbb{Z}^n$).

By smoothness of Φ we have convergence of the (partial sums of the) Fourier series to Φ in $\mathcal{C}^\infty(\mathbb{R}^n)$, that is, uniformly on $\mathbb{R}^n$ in all derivatives. (Via several integrations by parts it is routine to deduce the following: $\forall l \in \mathbb{N}\ \exists \gamma_l > 0$ such that $|c_k| \leq \gamma_l (1 + |k|^2)^{-l}$; thus we obtain uniform and absolute convergence of every derivative of the Fourier series, hence convergence to some function in $\mathcal{C}^\infty(I)$; since by abstract Hilbert space theory the Fourier series converges to Φ in $L^2(I)$, the $\mathcal{C}^\infty$-limit of the series must also be Φ.)

Since $\operatorname{supp}(\Phi)$ is compact in I we can find $\eta > 0$ such that $\operatorname{supp}(\Phi) \subseteq [2\eta, 1 - 2\eta]^n$. Let $\rho \in \mathcal{D}((0,1))$ with $\rho = 1$ on $(\eta, 1 - \eta)$ and define $\Psi_N \in \mathcal{D}(I)$ for $N \in \mathbb{N}_0$ by

$$\Psi_N(x_1, \ldots, x_n) := \sum_{\substack{(k_1,\ldots,k_n)\in\mathbb{Z}^n \\ |k_1|,\ldots,|k_n|\leq N}} c_k \prod_{l=l}^{n} \rho(x_l) e^{2\pi i k_l x_l}.$$

Clearly $\operatorname{supp}(\Psi_N) \subset \operatorname{supp}(\rho)^n$ for all N, $\operatorname{supp}(\Phi) \subseteq \operatorname{supp}(\rho)^n$, and $\Psi_N\ |_{\operatorname{supp}(\Phi)}$ agrees with the corresponding partial sum of the Fourier series. Hence by Leibniz' rule we obtain that $\Psi_N \to \Phi$ in $\mathcal{D}(I)$. Finally, since the series $(\Psi_N)_{N\in\mathbb{N}_0}$ converges uniformly absolutely (for all derivatives) it may be brought into the form as claimed by a standard relabeling procedure. (A few details on the relabeling: First, for any $k = (k_1, \ldots, k_n) \in \mathbb{Z}^n$ we put $\widetilde{\mu}_1^k(x_1) := c_k\, \rho(x_1) e^{2\pi i k_1 x_1}$ and $\widetilde{\mu}_l^k(x_l) := \rho(x_l) e^{2\pi i k_l x_l}$ ($l = 2, \ldots n$), so that $\Psi_N(x_1, \ldots, x_n) = \sum_{k\in\mathbb{Z}^n, \|k\|_\infty \leq N} \widetilde{\mu}_1^k(x_1) \cdots \widetilde{\mu}_n^k(x_n)$; second, choose a bijection $\beta : \mathbb{N}_0 \to \mathbb{Z}^n$ and define $\mu_{j,l} := \widetilde{\mu}_l^{\beta(j)}$; then the partial sums $\Phi_m(x_1, \ldots, x_n) := \sum_{j=0}^{m} \mu_{j,1}(x_1) \cdots \mu_{j,n}(x_n)$ are re-arrangements of the original series; by uniform absolute convergence of the original series (for every derivative) we obtain also $\Phi_m \to \Phi$ in $\mathcal{D}(I)$.) $\square$

Based on these preparations we can now establish the existence of a unique tensor product of distributions extending the classical operation:

4.1.2 Theorem. *Let $S \in \mathcal{D}'(\Omega_1)$ and $T \in \mathcal{D}'(\Omega_2)$. There exists a unique distribution $S \otimes T \in \mathcal{D}'(\Omega_1 \times \Omega_2)$, called the* tensor product *of S and T, such that*

$$\langle S \otimes T, \varphi \otimes \psi \rangle = \langle S, \varphi \rangle \langle T, \psi \rangle \qquad \forall \varphi \in \mathcal{D}(\Omega_1), \forall \psi \in \mathcal{D}(\Omega_2). \tag{4.1.3}$$

Proof. *Uniqueness:* By (4.1.3) the linear form $S \otimes T$ is determined on the subspace $M \subseteq \mathcal{D}(\Omega_1 \times \Omega_2)$ generated by splitting tensors of the form $\varphi \otimes \psi$ (M as in Lemma 4.1.1). Indeed, if $\chi = \sum_{j=1}^m \varphi_j \otimes \psi_j$ (with $\varphi_j \in \mathcal{D}(\Omega_1)$, $\psi_j \in \mathcal{D}(\Omega_2)$), then by linearity and (4.1.3)

$$\langle S \otimes T, \chi \rangle = \sum_{j=1}^{m} \langle S \otimes T, \varphi_j \otimes \psi_j \rangle = \sum_{j=1}^{m} \langle S, \varphi_j \rangle \langle T, \psi_j \rangle. \tag{4.1.4}$$

By assumption, $S \otimes T$ is continuous on $\mathcal{D}(\Omega_1 \times \Omega_2)$. Therefore uniqueness of $S \otimes T$ follows since M is dense by Lemma 4.1.1.

Take any $\chi \in \mathcal{D}(\Omega_1 \times \Omega_2)$. By Corollary 3.6.6 (i) the function $y \mapsto \langle S(x), \chi(x,y) \rangle$ belongs to $\mathcal{D}(\Omega_2)$, hence we may define a linear form $S \otimes T$ on $\mathcal{D}(\Omega_1 \times \Omega_2)$ by

$$\langle S \otimes T, \chi \rangle := \langle T(y), \langle S(x), \chi(x,y) \rangle \rangle \qquad \forall \chi \in \mathcal{D}(\Omega_1 \times \Omega_2). \tag{4.1.5}$$

On the subspace M this definition reproduces (4.1.4), in particular (4.1.3) holds. It remains to show that $S \otimes T$ is continuous.

Let $K \Subset \Omega_1 \times \Omega_2$ and denote by $K_i \Subset \Omega_i$ the projection of K to Ω_i ($i = 1, 2$). Let $\chi \in \mathcal{D}(K)$ and define $g \in \mathcal{D}(K_2)$ by

$$g(y) := \langle S, \chi(.,y) \rangle \qquad (y \in \Omega_2).$$

Recall that (3.6.1) gives $\partial^\beta g(y) = \langle S, \partial_y^\beta \chi(.,y) \rangle$.

The continuity condition (1.1.1) applied to T provides m and C (depending on K_2 only, not on g or χ) such that

$$|\langle T, g \rangle| \leq C \sum_{|\beta| \leq m} \|\partial^\beta g\|_{\infty, K_2}. \tag{4.1.6}$$

Since $\operatorname{supp}(\chi(.,y)) \subseteq K_1$ we may employ (1.1.1) for S to obtain N and C' (depending on K_1 only, but not on χ) such that

$$|\partial^\beta g(y)| = |\langle S, \partial_y^\beta \chi(.,y) \rangle| \leq C' \sum_{|\alpha| \leq N} \|\partial_x^\alpha \partial_y^\beta \chi(.,y)\|_{\infty, K_1}. \tag{4.1.7}$$

Combining (4.1.6) and (4.1.7) yields an estimate of the form (1.1.1) for $S \otimes T$ (with constant CC' and maximal order of derivatives $m + N$). $\square$

The main properties of the tensor product are collected in the following result.

4.1.3 Theorem. *Let $S \in \mathcal{D}'(\Omega_1)$ and $T \in \mathcal{D}'(\Omega_2)$. The tensor product $S \otimes T \in \mathcal{D}'(\Omega_1 \times \Omega_2)$ satisfies the following "Fubini-like" relation for all $\Phi \in \mathcal{D}(\Omega_1 \times \Omega_2)$:*

$$\langle S \otimes T, \Phi \rangle = \langle T(y), \langle S(x), \Phi(x,y) \rangle \rangle = \langle S(x), \langle T(y), \Phi(x,y) \rangle \rangle. \tag{4.1.8}$$

Moreover,

(i) $\mathrm{supp}\,(S \otimes T) = \mathrm{supp}\,(S) \times \mathrm{supp}\,(T)$.

(ii) $\partial_x^\alpha \partial_y^\beta (S \otimes T) = \partial_x^\alpha S \otimes \partial_y^\beta T$.

(iii) $\otimes : \mathcal{D}'(\Omega_1) \times \mathcal{D}'(\Omega_2) \to \mathcal{D}'(\Omega_1 \times \Omega_2)$ *is bilinear and sequentially continuous in each factor.*

Proof. We have used the equation $\langle S \otimes T, \Phi \rangle = \langle T(y), \langle S(x), \Phi(x,y) \rangle \rangle$ already to prove existence of the tensor product.

If we consider the functional $R\colon \Phi \mapsto \langle S(x), \langle T(y), \Phi(x,y) \rangle \rangle$, then it is easy to see that it satisfies (4.1.3) and continuity of R follows exactly as in the proof of Theorem 4.1.2. Moreover,

$$\langle S \otimes T, \varphi \otimes \psi \rangle = \langle S, \varphi \rangle \langle T, \psi \rangle = \langle R, \varphi \otimes \psi \rangle,$$

so by uniqueness we necessarily have $R = S \otimes T$ and, consequently, (4.1.8) holds.

(i) We first show $\mathrm{supp}\,(S) \times \mathrm{supp}\,(T) \subseteq \mathrm{supp}\,(S \otimes T)$.

Let $(x,y) \in \mathrm{supp}\,(S) \times \mathrm{supp}\,(T)$ and let W be a neighborhood of (x,y) in $\Omega_1 \times \Omega_2$. We may find a neighborhood U_x of x (in Ω_1) and a neighborhood V_y of y in Ω_2 with $U_x \times V_y \subseteq W$. Now

$$\left.\begin{array}{lcl} x \in \mathrm{supp}\,(S) & \Longrightarrow & \exists \varphi \in \mathcal{D}(U_x) : \langle S, \varphi \rangle \neq 0 \\ y \in \mathrm{supp}\,(T) & \Longrightarrow & \exists \psi \in \mathcal{D}(V_y) : \langle T, \psi \rangle \neq 0 \end{array}\right\} \ \mathrm{supp}\,(\varphi \otimes \psi) \subseteq U_x \times V_y \subseteq W$$

and

$$\langle S \otimes T, \varphi \otimes \psi \rangle = \langle S, \varphi \rangle \langle T, \psi \rangle \neq 0,$$

so $(x,y) \in \mathrm{supp}\,(S \otimes T)$.

To show the reverse inclusion suppose $(x,y) \in \big(\Omega_1 \times \Omega_2\big) \setminus \big(\mathrm{supp}\,(S) \times \mathrm{supp}\,(T)\big)$. We may assume w.l.o.g. that $x \notin \mathrm{supp}\,(S)$. Then there exists some neighborhood U_x of x (in Ω_1) such that $\overline{U_x} \cap \mathrm{supp}\,(S) = \emptyset$.

Let $\chi \in \mathcal{D}(\Omega_1 \times \Omega_2)$ with $\mathrm{supp}\,(\chi) \subseteq U_x \times \Omega_2$. Then we obtain $\forall y' \in \Omega_2$

$$\{x' \in \Omega_1 \mid \chi(x',y') \neq 0\} \subseteq \pi_1(\mathrm{supp}\,(\chi)) \subseteq U_x, \text{ i.e., } \mathrm{supp}\,(\chi(.,y')) \subseteq U_x.$$

Hence Proposition 2.2.4 implies that $\langle S(.), \chi(.,y') \rangle = 0$ for all $y' \in \Omega_2$. Therefore

$$\langle S \otimes T, \chi \rangle = \langle T(y'), \langle S(x'), \chi(x',y') \rangle \rangle = 0.$$

Since χ was an arbitrary element of $\mathcal{D}(U_x \times \Omega_2)$ we conclude that $(x,y) \notin \mathrm{supp}\,(S \otimes T)$.

(ii) By a direct calculation of the action on any $\chi \in \mathcal{D}(\Omega_1 \times \Omega_2)$ we get:

$$\begin{aligned} \langle \partial_x^\alpha \partial_y^\beta (S \otimes T), \chi \rangle &= (-1)^{|\alpha|+|\beta|} \langle S \otimes T, \partial_x^\alpha \partial_y^\beta \chi \rangle \\ &= (-1)^{|\alpha|+|\beta|} \langle S(x), \langle T(y), \partial_y^\beta \partial_x^\alpha \chi(x,y) \rangle \rangle \\ &= (-1)^{|\alpha|} \langle S(x), \langle \partial_y^\beta T(y), \partial_x^\alpha \chi(x,y) \rangle \rangle \\ &\underset{(3.6.1)}{=} (-1)^{|\alpha|} \langle S(x), \partial_x^\alpha \langle \partial_y^\beta T(y), \chi(x,y) \rangle \rangle \\ &= \langle \partial_x^\alpha S(x), \langle \partial_y^\beta T(y), \chi(x,y) \rangle \rangle = \langle \partial_x^\alpha S \otimes \partial_y^\beta T, \chi \rangle. \end{aligned}$$

(iii) Both bilinearity and separate sequential continuity follow immediately from (4.1.8). □

4.1.4 Remark. Using the theory of locally convex spaces one can even show that the tensor product is *jointly* sequentially continuous in both factors.

4.1.5 Example. Let $\Omega_1 = \Omega_2 = \mathbb{R}$ and $S = T = \delta$. We have

$$\langle \delta \otimes \delta, \chi \rangle = \langle \delta(x), \langle \delta(y), \chi(x,y) \rangle \rangle = \langle \delta(x), \chi(x,0) \rangle = \chi(0,0) = \langle \delta(x,y), \chi(x,y) \rangle,$$

i.e.,

$$\delta(x) \otimes \delta(y) = \delta(x,y).$$

Moreover,

$$\partial_x \partial_y (Y(x) \otimes Y(y)) = \delta(x) \otimes \delta(y) = \delta(x,y).$$

4.1.6 Remark. Tensor products of any finite number of distributional factors are constructed in a similar way and the properties are analogous. For example, we obtain on $\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R}$ (n times) by a calculation as above

$$\delta(x) = \delta(x_1, \ldots, x_n) = \delta(x_1) \otimes \cdots \otimes \delta(x_n)$$

and

$$\partial_1 \cdots \partial_n (Y(x_1) \otimes \cdots Y(x_n)) = \delta(x_1, \ldots, x_n).$$

The following result characterizes distributions that do not depend on one of the variables.

4.1.7 Theorem. *Let $T \in \mathcal{D}'(\mathbb{R}^n)$. Then we have:*

$$\partial_n T = 0 \quad \Longleftrightarrow \quad \exists S \in \mathcal{D}'(\mathbb{R}^{n-1}) : \; T(x) = S(x') \otimes 1(x_n),$$

with the notation $x' = (x_1, \ldots, x_{n-1})$ and $1(x_n)$ for the constant function $x_n \mapsto 1$. Note that in this case the action of T on a test function $\varphi \in \mathcal{D}(\mathbb{R}^n)$ is thus given by

$$\langle T, \varphi \rangle = \langle 1(x_n), \langle S(x'), \varphi(x', x_n) \rangle \rangle = \int_{\mathbb{R}} \langle S, \varphi(., t) \rangle \, dt. \tag{4.1.9}$$

Proof. ($\Leftarrow$) is immediate from Theorem 4.1.3 (ii).
($\Rightarrow$) Let $\chi \in \mathcal{D}(\mathbb{R})$ with $\int \chi = 1$ and define the linear functional[1] $S \colon \mathcal{D}(\mathbb{R}^{n-1}) \to \mathbb{C}$ by

$$\langle S, \psi \rangle := \langle T(x', x_n), \psi(x') \otimes \chi(x_n) \rangle \qquad (\psi \in \mathcal{D}(\mathbb{R}^{n-1})). \tag{4.1.10}$$

Continuity of S follows from the observation that $\psi_k \to 0$ in $\mathcal{D}(\mathbb{R}^{n-1})$ ($k \to \infty$) implies $\psi_k \otimes \chi \to 0$ in $\mathcal{D}(\mathbb{R}^n)$, hence $S \in \mathcal{D}'(\mathbb{R}^{n-1})$.
Now let $\varphi \in \mathcal{D}(\mathbb{R}^n)$, then we calculate

$$\begin{aligned} \langle S \otimes 1, \varphi \rangle = \langle S(x'), \langle 1(t), \varphi(x', t) \rangle \rangle &= \langle S(x'), \int \varphi(x', t) \, dt \rangle \\ &\underset{\substack{\uparrow \\ (4.1.10)}}{=} \langle T(x', x_n), (\int \varphi(x', t) \, dt) \otimes \chi(x_n) \rangle. \end{aligned}$$

Hence we may write

$$\langle T - S \otimes 1, \varphi \rangle = \langle T(x', x_n), \underbrace{\varphi(x', x_n) - (\int \varphi(x', t) \, dt) \otimes \chi(x_n)}_{=: \Phi(x', x_n)} \rangle.$$

[1] It is not difficult to guess S by making the ansatz $T = S \otimes 1$ and considering the action of T on a tensor product: $\langle T, \psi \otimes \chi \rangle = \langle S(x'), \langle 1(x_n), \psi(x') \, \chi(x_n) \rangle \rangle = \langle S(x'), \langle 1, \chi \rangle \, \psi(x') \rangle = \langle 1, \chi \rangle \langle S, \psi \rangle = (\int \chi) \, \langle S, \psi \rangle$.

Observe that for every $x' \in \mathbb{R}^{n-1}$ we have

$$\int \Phi(x', x_n)\, dx_n = \int \varphi(x', x_n)\, dx_n - \left(\int \varphi(x', t)\, dt\right) \cdot \underbrace{\int \chi(x_n)\, dx_n}_{=1} = 0.$$

Therefore $\Psi(x', x_n) := \int_{-\infty}^{x_n} \Phi(x', s)\, ds$ defines a function $\Psi \in \mathcal{D}(\mathbb{R}^n)$ with the property $\partial_n \Psi = \Phi$.

Therefore we obtain finally

$$\langle T - S \otimes 1, \varphi \rangle = \langle T, \Phi \rangle = \langle T, \partial_n \Psi \rangle = -\langle \partial_n T, \Psi \rangle \underset{\substack{\uparrow \\ \partial_n T = 0}}{=} 0,$$

so $T = S \otimes 1$. □

4.2 Convolution

For functions $f \in C_c(\mathbb{R}^n)$ and $g \in C(\mathbb{R}^n)$ the *convolution* $f * g \in C(\mathbb{R}^n)$ is defined by

$$f * g(x) := \int f(y) g(x - y)\, dy = \int f(x - y) g(y)\, dy \qquad (x \in \mathbb{R}^n).$$

We may consider $f * g$ as a regular distribution on $\mathbb{R}^n$ and calculate its action on a test function as follows:

$$\begin{aligned} \langle f * g, \varphi \rangle &= \int f * g(z) \varphi(z)\, dz = \int \int f(z - y) g(y) \varphi(z)\, dy\, dz \qquad \text{[Fubini]} \\ &= \int \int f(z - y) g(y) \varphi(z)\, dz\, dy \underset{\substack{\uparrow \\ \text{[inner integral} \\ x = z - y]}}{=} \int \int f(x) g(y) \varphi(x + y)\, dx\, dy \\ &\underset{\substack{\uparrow \\ \text{[Fubini]}}}{=} \int_{\mathbb{R}^{2n}} f(x) g(y) \varphi(x + y)\, d(x, y). \end{aligned}$$

This suggests to generalize the convolution to distributions $S \in \mathcal{E}'(\mathbb{R}^n)$ and $T \in \mathcal{D}'(\mathbb{R}^n)$ by a formula like

$$\langle S * T, \varphi \rangle := \langle S(x) \otimes T(y), \varphi(x + y) \rangle.$$

However, the status of the right-hand side of this equation has to be clarified, which will be achieved by an appropriate cut-off to adjust the support properties of the function $(x, y) \mapsto \varphi(x + y)$. (Note that this function will not be compactly supported, unless $\varphi = 0$: if $\varphi(z_0) \neq 0$, then for every $x \in \mathbb{R}^n$ and $y := z_0 - x$ we have $\varphi(x + y) \neq 0$.)

We begin by introducing some general vector space operations on subsets of $\mathbb{R}^n$. For $A, B \subseteq \mathbb{R}^n$, we set

$$\begin{aligned} -A &:= \{-x \mid x \in A\}, \\ A + B &:= \{x + y \mid x \in A, y \in B\}, \text{ and} \\ A - B &:= \{x - y \mid x \in A, y \in B\}. \end{aligned}$$

4.2.1 Remark.

(i) A compact and B closed $\implies$ $A \pm B$ is closed (Also: A, B compact $\Rightarrow$ $A \pm B$ compact.)

(ii) $\overline{A}$ compact $\implies$ $\overline{A} \pm \overline{B} = \overline{A \pm B}$.

If $\rho \in \mathcal{D}(\mathbb{R}^n)$ and $\varphi \in \mathcal{C}^\infty(\mathbb{R}^n)$, then

$$\operatorname{supp}\big(\rho(x)\varphi(x+y)\big) \subseteq \operatorname{supp}(\rho) \times \big(\operatorname{supp}(\varphi) - \operatorname{supp}(\rho)\big). \tag{4.2.1}$$

(Proof: $\rho(x)\varphi(x+y) \neq 0 \quad \Rightarrow \quad x \in \operatorname{supp}(\rho)$ and $x + y \in \operatorname{supp}(\varphi)$.)

If in addition $\operatorname{supp}(\varphi)$ is compact, then $\operatorname{supp}(\rho(x)\varphi(x+y))$ is compact in $\mathbb{R}^n \times \mathbb{R}^n$.

4.2.2 Theorem. *Let $S \in \mathcal{E}'(\mathbb{R}^n)$, $T \in \mathcal{D}'(\mathbb{R}^n)$. Choose a cut-off function $\rho \in \mathcal{D}(\mathbb{R}^n)$ with $\rho = 1$ on a neighborhood of* $\operatorname{supp}(S)$. *We define the* convolution $S * T$ *of S and T by setting*

$$\langle S * T, \varphi\rangle := \langle S(x) \otimes T(y), \rho(x)\varphi(x+y)\rangle \qquad \forall \varphi \in \mathcal{D}(\mathbb{R}^n). \tag{4.2.2}$$

Then

(i) The value of $\langle S * T, \varphi\rangle$ is independent of the choice of the cut-off ρ, i.e. $S * T$ is well defined.

(ii) Equation (4.2.2) *defines a distribution on $\mathbb{R}^n$, i.e. $S * T \in \mathcal{D}'(\mathbb{R}^n)$.*

Proof. (i) If $\sigma \in \mathcal{D}(\mathbb{R}^n)$ is also a cut-off over $\operatorname{supp}(S)$, then there is a neighborhood U of $\operatorname{supp}(S)$ such that $(\rho(x) - \sigma(x))\varphi(x+y) = 0$ when $(x, y) \in U \times \mathbb{R}^n$, which in turn is a neighborhood of $\operatorname{supp}(S \otimes T) = \operatorname{supp}(S) \times \operatorname{supp}(T)$. Hence Proposition 2.2.4 implies

$$\langle S(x) \otimes T(y), (\rho(x) - \sigma(x))\varphi(x+y)\rangle = 0.$$

(ii) Linearity of $S * T$ is obvious. To show the continuity condition (1.1.1), let $K \Subset \mathbb{R}^n$ and $\varphi \in \mathcal{D}(K)$ arbitrary. By (4.2.1) we have

$$\operatorname{supp}\big(\rho(x)\varphi(x+y)\big) \subseteq \operatorname{supp}(\rho) \times \big(K - \operatorname{supp}(\rho)\big) =: K',$$

and K' is compact in $\mathbb{R}^n \times \mathbb{R}^n$. The corresponding seminorm estimate (1.1.1) for $S \otimes T$ on K' then implies an estimate of the form (1.1.1) on K for $S * T$. □

4.2.3 Corollary. *Let $S \in \mathcal{E}'(\mathbb{R}^n)$, $T \in \mathcal{D}'(\mathbb{R}^n)$. Then we have for each $\varphi \in \mathcal{D}(\mathbb{R}^n)$*

$$\langle S * T, \varphi\rangle = \langle T(y), \langle S(x), \varphi(x+y)\rangle\rangle = \langle S(x), \langle T(y), \varphi(x+y)\rangle\rangle. \tag{4.2.3}$$

Moreover, we see that the roles of S and T may be interchanged in this formula. In this sense we have commutativity $S * T = T * S$.

Proof. Choosing a cut-off ρ as above we have by (4.2.2) and (4.1.8)

$$\begin{aligned}\langle S * T, \varphi\rangle &= \langle S(x) \otimes T(y), \rho(x)\varphi(x+y)\rangle \\ &= \langle S(x), \rho(x)\langle T(y), \varphi(x+y)\rangle\rangle = \langle T(y), \langle S(x), \rho(x)\varphi(x+y)\rangle\rangle.\end{aligned}$$

As follows from Remark 2.3.6, we may drop reference to the cut-off ρ in the action of an $\mathcal{E}'$-distribution, so we obtain (4.2.3). □

In the following result, for any $h \in \mathbb{R}^n$ and any $T \in \mathcal{D}'(\mathbb{R}^n)$, by τ_h we denote the translation operation

$$\langle \tau_h T, \varphi\rangle = \langle T, \tau_{-h}\varphi\rangle = \langle T(x), \varphi(x+h)\rangle. \tag{4.2.4}$$

4.2.4 Proposition. *Let $S, T \in \mathcal{D}'(\mathbb{R}^n)$, at least one of the two with compact support. Then*

(i) $\mathrm{supp}\,(S * T) \subseteq \mathrm{supp}\,(S) + \mathrm{supp}\,(T)$.

(ii) $j = 1, \ldots, n$: $\quad \partial_j(S * T) = (\partial_j S) * T = S * (\partial_j T)$.

(iii) $\forall h \in \mathbb{R}^n$: $\quad \tau_h(S * T) = (\tau_h S) * T = S * \tau_h T$.

Furthermore, $\delta = \delta_0$ plays the role of a neutral element for convolution

(iv) $\forall T \in \mathcal{D}'(\mathbb{R}^n)$: $T * \delta = \delta * T = T$.

Proof. (i) W.l.o.g. we may assume that $\mathrm{supp}\,(T)$ is compact and ρ is a suitable cut-off. Let $z \in \mathbb{R}^n \backslash (\mathrm{supp}\,(S) + \mathrm{supp}\,(T))$. As noted in Remark 4.2.1, $\mathrm{supp}\,(S) + \mathrm{supp}\,(T)$ is closed, hence there is an open neighborhood U of z such that $U \cap (\mathrm{supp}\,(S) + \mathrm{supp}\,(T)) = \emptyset$. Let $\varphi \in \mathcal{D}(U)$, then $(x, y) \in \mathrm{supp}\,((x', y') \mapsto \varphi(x' + y'))$ implies $x + y \in \mathrm{supp}\,(\varphi) \Subset U$. Hence

$$\mathrm{supp}\,(\varphi(x+y)) \cap \underbrace{(\mathrm{supp}\,(S) \times \mathrm{supp}\,(T))}_{=\mathrm{supp}\,(S \otimes T)} = \emptyset$$

and therefore

$$\langle S * T, \varphi \rangle = \langle S(x) \otimes T(x), \rho(x)\varphi(x+y) \rangle = 0.$$

We conclude that $z \notin \mathrm{supp}\,(S * T)$.

(ii) We calculate the action on a test function φ applying (4.2.3)

$$\begin{aligned} \langle (\partial_j S) * T, \varphi \rangle &= \langle T(y), \langle \partial_j S(x), \varphi(x+y) \rangle \rangle = -\langle T(y), \langle S(x), \partial_j \varphi(x+y) \rangle \rangle \\ &= -\langle S * T, \partial_j \varphi \rangle = \langle \partial_j (S * T), \varphi \rangle \end{aligned}$$

and by commutativity also $\partial_j(S * T) = \partial_j(T * S) = (\partial_j T) * S = S * \partial_j T$.

(iii) As in (ii) by use of (4.2.3)

$$\begin{aligned} \langle (\tau_h S) * T, \varphi \rangle &= \langle T(y), \langle \tau_h S(x), \varphi(x+y) \rangle \rangle = \langle T(y), \langle S(x), \tau_{-h} \varphi(x+y) \rangle \rangle \\ &= \langle S * T, \tau_{-h} \varphi \rangle = \langle \tau_h (S * T), \varphi \rangle \end{aligned}$$

and again by commutativity also $\tau_h(S * T) = \tau_h(T * S) = (\tau_h T) * S = S * \tau_h T$.

(iv) The action on a test function φ gives

$$\langle \delta * T, \varphi \rangle = \langle T(y), \langle \delta(x), \varphi(x+y) \rangle \rangle = \langle T(y), \varphi(0 + y) \rangle = \langle T, \varphi \rangle,$$

and again by commutativity also $T * \delta = \delta * T = T$. □

The following result establishes sequential continuity properties of the convolution of distributions:

4.2.5 Theorem. *Suppose that either*

(i) $S \in \mathcal{E}'(\mathbb{R}^n)$ *and* $T, T_m \in \mathcal{D}'(\mathbb{R}^n)$ $(m \in \mathbb{N})$ *with* $T_m \to T$ *in* $\mathcal{D}'(\mathbb{R}^n)$ $(m \to \infty)$

or

(ii) $S \in \mathcal{D}'(\mathbb{R}^n)$ *and* $T, T_m \in \mathcal{E}'(\mathbb{R}^n)$ $(m \in \mathbb{N})$ *satisfy the following:* $\exists K \Subset \mathbb{R}^n$ *such that* $\mathrm{supp}\,(T) \subseteq K$, $\mathrm{supp}\,(T_m) \subseteq K$ *holds* $\forall m \in \mathbb{N}$ *and* $T_m \to T$ *in* $\mathcal{D}'(\mathbb{R}^n)$ $(m \to \infty)$.[2]

[2] It would suffice to assume $T \in \mathcal{D}'(\mathbb{R}^n)$ without a support condition, since then $\mathrm{supp}\,(T) \subseteq K$ follows from the convergence $T_m \to T$.

Then $S * T_m \to S * T$ *in* $\mathcal{D}'(\mathbb{R}^n)$ $(m \to \infty)$.

Proof. (i) If $\varphi \in \mathcal{D}(\mathbb{R}^n)$, then by Corollary 3.6.6 (ii) the function $\varphi_S \colon y \mapsto \langle S(x), \varphi(x+y)\rangle$ is smooth. Moreover, φ_S vanishes when $y \notin \operatorname{supp}(\varphi) - \operatorname{supp}(S)$, since this implies $\operatorname{supp}(S) \cap \operatorname{supp}(\varphi(.+y)) = \emptyset$ and Proposition 2.2.4 yields $\varphi_S(y) = \langle S(x), \varphi(x+y)\rangle = 0$. Thus φ_S is a test function on $\mathbb{R}^n$ and we obtain

$$\langle S{*}T_m, \varphi\rangle \underset{(4.2.3)}{=} \langle T_m(y), \langle S(x), \varphi(x{+}y)\rangle\rangle \overset{m\to\infty}{\to} \langle T(y), \langle S(x), \varphi(x{+}y)\rangle\rangle = \langle S{*}T, \varphi\rangle.$$

(ii) Let $\rho \in \mathcal{D}(\mathbb{R}^n)$ be a cut-off over some neighborhood of K. Recall that the action of any $R \in \mathcal{E}'(\mathbb{R}^n)$ with $\operatorname{supp}(R) \subseteq K$ on a function $\psi \in \mathcal{C}^\infty(\mathbb{R}^n)$ was obtained by $\langle R, \rho\psi\rangle$. Furthermore, the function $y \mapsto \langle S(x), \varphi(x+y)\rangle$ is smooth by Proposition 3.6.4, so by (4.2.3) again we have as $m \to \infty$

$$\begin{aligned}
\langle S * T_m, \varphi\rangle &= \langle T_m(y), \langle S(x), \varphi(x+y)\rangle\rangle \\
&= \langle T_m(y), \rho(y)\langle S(x), \varphi(x+y)\rangle\rangle \ \to \ \langle T(y), \rho(y)\langle S(x), \varphi(x+y)\rangle\rangle \\
&= \langle T(y), \langle S(x), \varphi(x+y)\rangle\rangle = \langle S * T, \varphi\rangle.
\end{aligned}$$

$\square$

4.3 Regularization

In the preceding section we have developed a theory of convolution as a map $\mathcal{E}' \times \mathcal{D}' \to \mathcal{D}'$. Now we will change the point of view by restricting the $\mathcal{E}'$-factor to $\mathcal{C}^\infty$-functions, that is we consider the convolution $\mathcal{D} * \mathcal{D}'$. As we will see, this provides a process of *regularizing* (smoothing) a given distribution. More precisely, if $\rho \in \mathcal{D}$ is a mollifier, then for any $T \in \mathcal{D}'$ we obtain a net of smooth functions $T * \rho_\varepsilon$ $(\varepsilon > 0)$ with the property $T * \rho_\varepsilon \to T$ in $\mathcal{D}'$ as $\varepsilon \to 0$. Recall that we already used this technique in Theorem 1.3.4 to approximate C^k-functions by $\mathcal{C}^\infty$-functions.

To get an intuitive idea why convolution has a smoothing effect, we consider $f \in C(\mathbb{R}) \subseteq \mathcal{D}'(\mathbb{R})$. Let $\rho \in \mathcal{D}(\mathbb{R})$ with $\rho \geq 0$, $\operatorname{supp}(\rho) \subseteq [-1,1]$, and $\rho(x) = 1$ when $|x| \leq 1/2$. If $\rho_\varepsilon(z) := \rho(z/\varepsilon)/\varepsilon$, then $\operatorname{supp}(\rho_\varepsilon)(x - .) \subseteq [x-\varepsilon, x+\varepsilon]$, $\rho_\varepsilon(x-y) = 1$ when $|x-y| \leq \varepsilon/2$, and we obtain

$$f * \rho_\varepsilon(x) = \int_{-\infty}^{\infty} f(y)\rho_\varepsilon(x-y)\,dy \approx \frac{1}{\varepsilon}\int_{x-\varepsilon/2}^{x+\varepsilon/2} f(y)\,dy =: M_\varepsilon(f)(x).$$

Here, $M_\varepsilon(f)(x)$ is the "mean value of f near x" and we easily deduce that $M_\varepsilon(f)(x) \to f(x)$ $(\varepsilon \to 0)$. Moreover, as noted in the proof of Theorem 1.3.4 the functions $f * \rho_\varepsilon$ $(\varepsilon > 0)$ are smooth.

4.3.1 Theorem. *Let* $\rho \in \mathcal{C}^\infty(\mathbb{R}^n)$ *and* $T \in \mathcal{D}'(\mathbb{R}^n)$ *such that*

(i) $\operatorname{supp}(\rho)$ *is compact, or*

(ii) $\operatorname{supp}(T)$ *is compact.*

Then

$$T * \rho(x) = \langle T(y), \rho(x-y)\rangle \qquad (x \in \mathbb{R}^n) \tag{4.3.1}$$

and $T * \rho \in \mathcal{C}^\infty(\mathbb{R}^n)$.

Proof. Suppose first that (i) holds. Let $\varphi \in \mathcal{D}(\mathbb{R}^n)$. We may regard φ as a regular element in $\mathcal{E}'(\mathbb{R}^n)$ and choose a cut-off $\sigma \in \mathcal{D}(\mathbb{R}^n)$ over a neighborhood of $\operatorname{supp}(\varphi)$. Then $(x,y) \mapsto \sigma(x)\rho(x-y)$ is in $\mathcal{D}(\mathbb{R}^n \times \mathbb{R}^n)$.

Recall that the function $x \mapsto \langle T(y), \rho(x-y)\rangle$ is smooth due to Corollary 3.6.6 (ii). We may thus calculate the action of this function on φ by

$$\begin{aligned}
\int \langle T(y), \rho(x-y)\rangle\, \varphi(x)\, dx &= \int \varphi(x)\sigma(x)\langle T(y), \rho(x-y)\rangle\, dx \\
&= \langle \varphi(x), \langle T(y), \sigma(x)\rho(x-y)\rangle\rangle = \langle \varphi(x) \otimes T(y), \sigma(x)\rho(x-y)\rangle \\
&= \langle T(y), \langle \varphi(x), \sigma(x)\rho(x-y)\rangle\rangle \\
&= \langle T(y), \langle \varphi(x), \rho(x-y)\rangle\rangle = \langle T(y), \langle \varphi, \tau_y\rho\rangle\rangle \\
&= \langle T(y), \langle \tau_{-y}\varphi, \rho\rangle\rangle = \langle T(y), \langle \rho, \tau_{-y}\varphi\rangle\rangle \\
&= \langle T(y), \langle \rho(x), \varphi(x+y)\rangle\rangle = \langle T * \rho, \varphi\rangle.
\end{aligned}$$

Since φ was arbitrary, we obtain that $T * \rho$ is a regular distribution and is given by (4.3.1).

Now suppose that (ii) holds. Then we pick a cut-off $\chi \in \mathcal{D}(\mathbb{R}^n)$ over a neighborhood of $\operatorname{supp}(T)$ and note that the right-hand side in (4.3.1) actually means $\langle T(y), \chi(y)\rho(x-y)\rangle$. A calculation similar to the above then shows

$$\int \langle T(y), \chi(y)\rho(x-y)\rangle\, \varphi(x)\, dx = \langle T * \rho, \varphi\rangle.$$

□

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

Proof. Let $T \in \mathcal{D}'(\mathbb{R}^n)$. We have to show that there exists a sequence (T_m) in $\mathcal{D}(\mathbb{R}^n)$ with $T_m \to T$ in $\mathcal{D}'(\mathbb{R}^n)$ as $m \to \infty$.

Let $\rho \in \mathcal{D}(\mathbb{R}^n)$ be a mollifier, i.e. $\operatorname{supp}(\rho) \subseteq \overline{B_1(0)}$ and $\int \rho = 1$, and set

$$\rho_m(x) = m^n \rho(mx) \qquad (x \in \mathbb{R}^n, m \in \mathbb{N}).$$

(This corresponds to ρ_ε when $\varepsilon = 1/m$ as used in the proof of Theorem 1.3.4.)

By Example 1.2.6 we have $\rho_m \to \delta$ in $\mathcal{D}'(\mathbb{R}^n)$. Since $\operatorname{supp}(\rho_m) \subseteq \operatorname{supp}(\rho) \subseteq \overline{B_1(0)}$ holds for all m, Theorem 4.2.5, case (ii), implies

$$\widetilde{T_m} := T * \rho_m \to T * \delta = T \qquad (m \to \infty).$$

Case (i) of Theorem 4.3.1 ensures that $\widetilde{T_m}$ belongs to $\mathcal{C}^\infty(\mathbb{R}^n)$. Thus, it remains to adjust the supports for our approximating sequence. To achieve this we take $\chi \in \mathcal{D}(\mathbb{R}^n)$ with $\chi = 1$ on $B_1(0)$ and set

$$T_m(x) := \chi\Big(\frac{x}{m}\Big)\, \widetilde{T_m}(x) = \chi\Big(\frac{x}{m}\Big) \cdot (\rho_m * T)(x) \qquad (x \in \mathbb{R}^n, m \in \mathbb{N}).$$

(Note that $T_m = \widetilde{T_m}$ on $B_m(0)$ and $\operatorname{supp}(T_m) \subseteq m \cdot \operatorname{supp}(\chi)$.)

Let $\varphi \in \mathcal{D}(\mathbb{R}^n)$ arbitrary. Then we have for m sufficiently large that

$$\langle T_m, \varphi\rangle = \int T_m(x)\varphi(x)\, dx = \int \widetilde{T_m}(x)\varphi(x)\, dx = \langle \widetilde{T_m}, \varphi\rangle.$$

Hence also $T_m \to T$ in $\mathcal{D}'(\mathbb{R}^n)$. □

This allows us to finally give a proof of Theorem 1.1.9, stating that $\mathcal{D}(\Omega)$ is dense in $\mathcal{D}'(\Omega)$ for any open set $\emptyset \neq \Omega \subseteq \mathbb{R}^n$:

Proof of Theorem 1.1.9. Let (K_m) be a compact exhaustion of Ω (as introduced after Definition 2.3.1), and pick $\chi_m \in \mathcal{D}(K_{m+1})$ such that $\chi_m \equiv 1$ in a neighborhood of K_m. Let $T \in \mathcal{D}'(\Omega)$ and set $T_m := \chi_m T$. Then $T_m \in \mathcal{E}'(\Omega) \subseteq \mathcal{E}'(\mathbb{R}^n)$. Let $\rho \in \mathcal{D}(\mathbb{R}^n)$ be a mollifier as in the proof of Theorem 4.3.2. Then it follows from Proposition 4.2.4 (i) that for each m there exists some $\varepsilon_m > 0$, $\varepsilon_m \searrow 0$ ($m \to \infty$) such that, setting $\rho_m(x) := \varepsilon_m^{-n}\rho(x/\varepsilon_m)$ we have $\operatorname{supp}(T_m * \rho_m) \subseteq \Omega$ (hence $T_m * \rho_m \in \mathcal{D}(\Omega)$).

Our aim now is to show that $T_m * \rho_m \to T$ in $\mathcal{D}'(\Omega)$. Thus let $\varphi \in \mathcal{D}(\Omega)$. Then for some $l' \in \mathbb{N}$ we have $\operatorname{supp}(\varphi) \subseteq K_{l'}$, so

$$\langle T, \varphi\rangle = \langle T, \chi_m \varphi\rangle = \langle T_m, \varphi\rangle \quad \forall m \geq l'. \tag{4.3.2}$$

Furthermore, there exists some $k_0 \in \mathbb{N}$ such that for all $k \geq k_0$ we have

$$L_k := \operatorname{supp}(x \mapsto \rho_k(y)\varphi(x+y)) \subseteq \operatorname{supp}(\varphi) + \overline{B_{\varepsilon_k}(0)} \Subset \Omega.$$

As $L_k \supseteq L_{k+1}$ for all k, there exists some $l'' > k_0$ such that $L_k \subseteq K_m$ for all $k, m \geq l''$. Now set $l := \max(l', l'')$. Then for each $m \geq l$ we obtain

$$\operatorname{supp}(x \mapsto \int \rho_m(y)\varphi(x+y)\,dy) \subseteq L_m \subseteq K_l,$$

so

$$\begin{aligned}
\langle T_m * \rho_m, \varphi\rangle &= \left\langle T_m(x), \int \rho_m(y)\varphi(x+y)\,dy \right\rangle \\
&= \left\langle T_m(x), \chi_l(x)\cdot \int \rho_m(y)\varphi(x+y)\,dy \right\rangle \\
&= \left\langle T_l(x), \int \rho_m(y)\varphi(x+y)\,dy \right\rangle = \langle T_l * \rho_m, \varphi\rangle.
\end{aligned}$$

Now $\rho_m \to \delta$ in $\mathcal{D}'(\mathbb{R}^n)$ and $\operatorname{supp}(\rho_m) \subseteq \operatorname{supp}(\rho_1)$ for all m. Therefore Theorem 4.2.5 (ii) implies

$$\langle T_m * \rho_m, \varphi\rangle = \langle T_l * \rho_m, \varphi\rangle \overset{m\to\infty}{\longrightarrow} \langle T_l, \varphi\rangle = \langle T, \varphi\rangle$$

by (4.3.2). □

4.3.3 Theorem. *Let $L\colon \mathcal{D}(\mathbb{R}^n) \to \mathcal{E}(\mathbb{R}^n)$ be a linear map. Then the following statements are equivalent:*

(i) L is continuous and $\forall h \in \mathbb{R}^n$: $\tau_h \circ L = L \circ \tau_h$, i.e. L commutes with translations.

(ii) $\exists!\, T \in \mathcal{D}'(\mathbb{R}^n)$: $L\varphi = T * \varphi$ holds for all $\varphi \in \mathcal{D}(\mathbb{R}^n)$.

Proof. (ii) $\Rightarrow$ (i): Linearity of L is clear and that convolution commutes with translations follows from Proposition 4.2.4 (iii). Smoothness of $T * \varphi$ is ensured by case (i) in Theorem 4.3.1.

It remains to prove that $\varphi_j \to 0$ in $\mathcal{D}(\mathbb{R}^n)$ implies $T * \varphi_j \to 0$ in $\mathcal{E}(\mathbb{R}^n)$. Since $\partial^\alpha(T * \varphi_j) = T * (\partial^\alpha \varphi_j)$ by Proposition 4.2.4 (ii), it suffices to show the following: $\forall K \Subset \mathbb{R}^n$ we have $T * \varphi_j \to 0$ uniformly on K.

Let $K \Subset \mathbb{R}^n$ be arbitrary and let $K_0 \Subset \mathbb{R}^n$ such that $\operatorname{supp}(\varphi_j) \subseteq K_0$ for all j. Then by (4.3.1) and (1.1.1) applied to T we can find $m \in \mathbb{N}_0$ and $C > 0$ such that for all $x \in K$

$$|T * \varphi_j(x)| = |\langle T, \varphi_j(x - .)\rangle| \leq C \sum_{|\alpha| \leq m} \|\partial^\alpha \varphi_j\|_{\infty, K-K_0},$$

hence

$$\|T * \varphi_j\|_{\infty,K} \leq C \sum_{|\alpha|\leq m} \|\partial^\alpha \varphi_j\|_{\infty,K-K_0} \to 0 \qquad (j \to \infty).$$

(i) $\Rightarrow$ (ii): Uniqueness of T follows by considering $\forall \varphi \in \mathcal{D}(\mathbb{R}^n)$

$$\langle T(y), \varphi(-y)\rangle = \langle T, \varphi(0 - .)\rangle = T * \varphi(0) = (L\varphi)(0),$$

since this equation determines the action of T.
To show existence we make the ansatz

$$\langle T, \varphi\rangle := L\,(R\varphi)(0) = L\check{\varphi}(0) \qquad \forall \varphi \in \mathcal{D}(\mathbb{R}^n),$$

where we have put $R\varphi(y) := \check{\varphi}(y) = \varphi(-y)$. (Note that $RR\varphi = \varphi$.) By continuity and linearity of L the corresponding properties for T follow, that is, $T \in \mathcal{D}'(\mathbb{R}^n)$. (We have $T = (\text{evaluation at } 0) \circ L \circ R$, which is a composition of linear continuous maps.) Finally, if $\varphi \in \mathcal{D}(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$ are arbitrary, then commutativity with translations implies

$$\begin{aligned} L\varphi(x) &= \tau_{-x}(L\varphi)(0) = L(\tau_{-x}\varphi)(0) = \langle T, R(\tau_{-x}\varphi)\rangle = \langle T, R(\varphi(. + x))\rangle \\ &= \langle T(y), \varphi(-y + x)\rangle = T * \varphi(x). \end{aligned}$$

□

4.3.4 Examples. (i) Let $h \in \mathbb{R}^n$ and $\varphi \in \mathcal{D}(\mathbb{R}^n)$, then

$$\delta_h * \varphi(x) = \langle \delta_h, \varphi(x - .)\rangle = \varphi(x - h) = \tau_h\varphi(x),$$

thus translation τ_h corresponds to convolution with δ_h.
(ii) Let $P(\partial) = \sum_{|\alpha|\leq m} a_\alpha \partial^\alpha$ be a PDO with constant coefficients $a_\alpha \in \mathbb{C}$. Clearly, $P(\partial)$ defines a translation invariant map $\mathcal{D}(\mathbb{R}^n) \to \mathcal{D}(\mathbb{R}^n) \subseteq \mathcal{E}(\mathbb{R}^n)$. We have for any $\varphi \in \mathcal{D}(\mathbb{R}^n)$

$$\begin{aligned} P(\partial)\varphi = \sum_{|\alpha|\leq m} a_\alpha\, \partial^\alpha \varphi &= \sum_{|\alpha|\leq m} a_\alpha\, \partial^\alpha (\delta * \varphi) \\ &= \sum_{|\alpha|\leq m} a_\alpha\, (\partial^\alpha \delta) * \varphi = \big(\sum_{|\alpha|\leq m} a_\alpha\, \partial^\alpha \delta\big) * \varphi = T * \varphi, \end{aligned}$$

where $T := \sum_{|\alpha|\leq m} a_\alpha\, \partial^\alpha \delta = P(\partial)\delta$.

4.4 Non-compact supports

So far, we have defined the convolution for $S \in \mathcal{E}'$ and $T \in \mathcal{D}'$ by its action on a test function φ as

$$\langle S * T, \varphi\rangle = \langle S(x) \otimes T(y), \varphi(x + y)\rangle.$$

Inspecting the right-hand side of this equation, we realize that all that is required for the above formula to work is to have the function

$$\operatorname{supp}(S) \times \operatorname{supp}(T) \to \mathbb{C}, \quad (x, y) \mapsto \varphi(x + y)$$

compactly supported. This in turn would be guaranteed (for all φ), if the map $\operatorname{supp}(S) \times \operatorname{supp}(T) \to \mathbb{R}^n$, $(x, y) \mapsto x + y$ has the property that inverse images of compact subsets of $\mathbb{R}^n$ are compact in $\operatorname{supp}(S) \times \operatorname{supp}(T)$.

4.4.1 Definition. *Let X and Y be locally compact topological spaces and $f\colon X \to Y$ be continuous. Then f is said to be* proper, *if for every compact subset $K \subseteq Y$ the inverse image $f^{-1}(K) \subseteq X$ is compact.*

4.4.2 Lemma. *Let $A \subseteq \mathbb{R}^n$ be closed and $f\colon A \to \mathbb{R}^m$ be continuous. Then A is locally compact*[3]*. Furthermore, f is proper if and only if the following holds:*

$$\forall \eta > 0\, \exists \gamma > 0\, \forall x \in A: \ |f(x)| \leq \eta \ \Rightarrow \ |x| \leq \gamma.$$

Proof. Let $x \in A$ and $K(x)$ be a compact neighborhood of x in $\mathbb{R}^n$. Then $U(x) = A \cap K(x)$ is a compact neighborhood of x in A. Hence A is locally compact (since the Hausdorff property of A is clear).

If f is proper, then $f^{-1}(\overline{B_\eta(0)})$ is compact in A, hence (also compact and) bounded in $\mathbb{R}^n$. Therefore we can find $\gamma > 0$ such that $f^{-1}(\overline{B_\eta(0)}) \subseteq \overline{B_\gamma(0)}$, which means that $|f(x)| \leq \eta$ implies $|x| \leq \gamma$.

Conversely, suppose that for any $\eta > 0$ we can find $\gamma > 0$ with the above property, i.e., $f^{-1}(\overline{B_\eta(0)}) \subseteq \overline{B_\gamma(0)}$. Let $K \Subset \mathbb{R}^m$. By continuity the set $f^{-1}(K)$ is closed in A, thus also closed in $\mathbb{R}^n$ (since A is closed in $\mathbb{R}^n$). Choose $\eta > 0$ so that $K \subseteq \overline{B_\eta(0)}$. There is $\gamma > 0$ such that $f^{-1}(K) \subseteq f^{-1}(\overline{B_\eta(0)}) \subseteq \overline{B_\gamma(0)}$. Hence $f^{-1}(K)$ is also bounded. In summary, $f^{-1}(K)$ is compact. □

Suppose $S, T \in \mathcal{D}'(\mathbb{R}^n)$ are such that the map $\operatorname{supp}(S) \times \operatorname{supp}(T) \to \mathbb{R}^n, (x, y) \mapsto x + y$ is proper. If $\eta > 0$, then there exists $\gamma > 0$ such that the following holds for every $(x, y) \in \operatorname{supp}(S) \times \operatorname{supp}(T)$:

$$|x + y| \leq \eta \quad \Longrightarrow \quad \max(|x|, |y|) \leq \gamma. \tag{4.4.1}$$

Let $\rho, \chi \in \mathcal{D}(\mathbb{R}^n)$ with $\rho = 1$, $\chi = 1$ on a neighborhood of $\overline{B_\gamma(0)}$.

We claim that the restriction of the distribution $(\chi S) * (\rho T)$ to $B_\eta(0)$ is independent of the choice of ρ and χ. By commutativity of the convolution it suffices to show this for χ.

Thus let $\chi_1 \in \mathcal{D}(\mathbb{R}^n)$ also have the property that $\chi_1 = 1$ on a neighborhood of $\overline{B_\gamma(0)}$. Then $\operatorname{supp}((\chi_1 - \chi)S) \cap \overline{B_\gamma(0)} = \emptyset$ and we will show that also

$$\operatorname{supp}(((\chi_1 - \chi)S) * (\rho T)) \cap \overline{B_\eta(0)} = \emptyset \tag{4.4.2}$$

holds. Indeed, let $z \in \mathbb{R}^n$ satisfy $|z| \leq \eta$ and

$$z \in \operatorname{supp}(((\chi_1 - \chi)S) * (\rho T)) \subseteq \operatorname{supp}((\chi_1 - \chi)S) + \operatorname{supp}(\rho T) \subseteq \operatorname{supp}(S) + \operatorname{supp}(T).$$

Then $z = x + y$ with $x \in \operatorname{supp}((\chi_1 - \chi)S)$ and $y \in \operatorname{supp}(\rho T)$ and (4.4.1) implies that $x, y \in \overline{B_\gamma(0)}$. In particular $x \in \operatorname{supp}((\chi_1 - \chi)S) \cap \overline{B_\gamma(0)} = \emptyset$, a contradiction.

By (4.4.2) we have $((\chi_1 - \chi)S) * (\rho T) \mid_{B_\eta(0)} = 0$ and therefore

$$(\chi_1 S) * (\rho T) = (\chi S) * (\rho T) + ((\chi_1 - \chi)S) * (\rho T) = (\chi S) * (\rho T) \quad \text{on } B_\eta(0).$$

Thus we are led to the following way of defining the convolution $S * T$ when neither S nor T need to be compactly supported.

4.4.3 Definition. *Let $S, T \in \mathcal{D}'(\mathbb{R}^n)$ such that the map*

$$\operatorname{supp}(S) \times \operatorname{supp}(T) \to \mathbb{R}^n, (x, y) \mapsto x + y \quad \textit{is proper.}$$

[3] In general (topological) subspaces of a locally compact space may fail to be locally compact. (E.g. $\mathbb{Q}$ with the inherited euclidean topology of $\mathbb{R}$ is not locally compact; see also examples with sine curves in $\mathbb{R}^2$ as in [SJ95, No. 118,1]).

Then we define the convolution $S * T \in \mathcal{D}'(\mathbb{R}^n)$ as follows: For any $\eta > 0$ and $\varphi \in \mathcal{D}(\mathbb{R}^n)$ with $\operatorname{supp}(\varphi) \subseteq B_\eta(0)$ *we set*

$$\langle S * T, \varphi\rangle := \langle (\chi S) * (\rho T), \varphi\rangle, \tag{4.4.3}$$

where the cut-off functions χ and ρ are as above.

4.4.4 Remark.

(i) If $S \in \mathcal{E}'(\mathbb{R}^n)$ and $T \in \mathcal{D}'(\mathbb{R}^n)$, then the convolution according to Definition 4.4.3 coincides with $S * T$ as constructed in Theorem 4.2.2.

Proof. By compactness, $\operatorname{supp}(S)$ is bounded, say, $\operatorname{supp}(S) \subseteq \overline{B_R(0)}$. Hence properness of the map $\operatorname{supp}(S) \times \operatorname{supp}(T) \to \mathbb{R}^n, (x, y) \mapsto x+y$ follows, since $x \in \operatorname{supp}(S)$, $y \in \operatorname{supp}(T)$ and $|x + y| \leq \eta$ implies $|y| \leq \eta + R$. Thus we may put $\gamma := \eta + R$ to satisfy (4.4.1). With cut-off functions χ and ρ as above, we now obtain $\chi S = S$. Furthermore, if $\varphi \in \mathcal{D}(B_\eta(0))$ then $\rho T = T$ on the set $\{y \mid \exists x \in \operatorname{supp}(S) : x + y \in \operatorname{supp}(\varphi)\}$, hence $\langle S * (\rho T), \varphi\rangle = \langle S(x), \langle \rho(y)T(y), \varphi(x + y)\rangle\rangle = \langle S(x), \langle T(y), \varphi(x + y)\rangle\rangle = \langle S * T, \varphi\rangle$. □

(ii) Relations analogous to those stated in Proposition 4.2.4 also hold for the convolution defined in Definition 4.4.3. In particular, we have again the formulae

$$\langle S * T, \varphi\rangle = \langle S(x), \langle T(y), \varphi(x + y)\rangle\rangle = \langle T(y), \langle S(x), \varphi(x + y)\rangle\rangle,$$
$$\operatorname{supp}(S * T) \subseteq \operatorname{supp}(S) + \operatorname{supp}(T),$$
$$\partial_j(S * T) = (\partial_j S) * T = S * (\partial_j T) \text{ and } \tau_h(S * T) = (\tau_h S) * T = u * (\tau_h T),$$

and separate sequential continuity of $(S, T) \mapsto S * T$.

The proofs are easy adaptations of those in Proposition 4.2.4, based on (4.4.3). Alternatively, cf. [Hor66, Chapter 4.9] for an equivalent approach and more detailed proofs. (Equivalence of the approaches follows from Exercise 2 in the same Section of that book.)

(iii) Similarly, convolution of finitely many distributions $T_1, \ldots, T_m \in \mathcal{D}'(\mathbb{R}^n)$ can be defined under the condition that

$$\operatorname{supp}(T_1) \times \ldots \times \operatorname{supp}(T_m) \to \mathbb{R}^n,\ (x^{(1)}, \ldots, x^{(m)}) \mapsto x^{(1)} + \ldots + x^{(m)} \quad \text{is proper.}$$

In this case, we have also *associativity* of the convolution, in particular, if T_1, T_2, T_3 satisfy the above properness condition, then

$$T_1 * T_2 * T_3 = (T_1 * T_2) * T_3 = T_1 * (T_2 * T_3).$$

(Cf. [FJ98, Section 5.3].)

Warning: Associativity may fail if the properness condition is violated even in cases where both convolutions $(T_1 * T_2) * T_3$ and $T_1 * (T_2 * T_3)$ do exist. For example, on $\mathbb{R}$ we have

$$(1 * \delta') * Y = (1' * \delta) * Y = 0 * Y = 0, \text{ whereas}$$
$$1 * (\delta' * Y) = 1 * (\delta * Y') = 1 * (\delta * \delta) = 1 * \delta = 1.$$

4.4.5 Example. $\mathcal{D}'_+(\mathbb{R}) := \{T \in \mathcal{D}'(\mathbb{R}) \mid \exists a \in \mathbb{R} : \operatorname{supp}(T) \subseteq [a, \infty[\,\}$ is a *convolution algebra*,, i.e., $\mathcal{D}'_+(\mathbb{R})$ is a vector subspace such that convolution is a bilinear map $*\colon \mathcal{D}'_+(\mathbb{R}) \times \mathcal{D}'_+(\mathbb{R}) \to \mathcal{D}'_+(\mathbb{R})$ and $(\mathcal{D}'_+(\mathbb{R}), +, *)$ forms a ring.
(In addition, we have $\delta_0 \in \mathcal{D}'_+(\mathbb{R})$ as an identity with respect to convolution and commutativity of $*$.)

If $T_1, \ldots, T_m \in \mathcal{D}'_+(\mathbb{R})$, then the properness condition in (iii) above holds, since we may first choose a common lower bound for the supports and then boundedness of the sum forces boundedness of each summand. Thus we obtain convolvability and associativity. That $T_1 * T_2$ again belongs to $\mathcal{D}'_+(\mathbb{R})$ follows from the relation $\operatorname{supp}(T_1 * T_2) \subseteq \operatorname{supp}(T_1) + \operatorname{supp}(T_2)$. Finally, bilinearity is immediate from the definition.

Similarly, one can show that $\mathcal{D}'_-(\mathbb{R}) := \{T \in \mathcal{D}'(\mathbb{R}) \mid \exists a \in \mathbb{R} : \operatorname{supp}(T) \subseteq]-\infty, a]\}$ is a convolution algebra.

4.4.6 Remark. *Alternative description of primitives (or antiderivatives)*: Let $a, b \in \mathbb{R}$ with $a < b$ and $\rho \in \mathcal{C}^\infty(\mathbb{R})$ be such that $\rho = 0$ when $x < a$ and $\rho = 1$ when $x > b$.

For any $T \in \mathcal{D}'(\mathbb{R})$ put

$$T_- := (1 - \rho)T \quad \text{and} \quad T_+ := \rho T.$$

Then $T = T_- + T_+$ with $T_- \in \mathcal{D}'_-(\mathbb{R})$ and $T_+ \in \mathcal{D}'_+(\mathbb{R})$. Since also $(Y - 1) \in \mathcal{D}'_-(\mathbb{R})$ and $Y \in \mathcal{D}'_+(\mathbb{R})$, we may define

$$S := (Y - 1) * T_- + Y * T_+ \in \mathcal{D}'(\mathbb{R})$$

and obtain

$$S' = ((Y-1)*T_-)' + (Y*T_+)' = (Y-1)'*T_- + Y'*T_+ = \delta*T_- + \delta*T_+ = T_- + T_+ = T.$$

In particular, for any $R \in \mathcal{D}'_+(\mathbb{R})$ the distribution $Y*R \in \mathcal{D}'_+(\mathbb{R})$ is an antiderivative.

4.5 The local structure of distributions

In the introduction to [Hör90], Lars Hoermander states that *"In differential calculus one encounters immediately the unpleasant fact that not every function is differentiable. The purpose of distribution theory is to remedy this flaw; indeed, the space of distributions is essentially the smallest extension of the space of continuous functions where differentiability is always well defined."*

While we already know that any continuous function has derivatives of arbitrary order, we have not yet verified the second part of Hoermander's statement, on $\mathcal{D}'$ being the *smallest* extension of the space of continuous functions. This is the purpose of the current section.

4.5.1 Examples. (i) As one of the first examples of differentiation we had $Y' = \delta$ in $\mathcal{D}'(\mathbb{R})$. We see that in a sense the "primitive function" of δ is thus more regular than δ itself. In fact, Y is a regular distribution, since $Y \in L^\infty_{\text{loc}}(\mathbb{R}) \subseteq L^1_{\text{loc}}(\mathbb{R})$. Let us look at a "primitive function" of Y, namely the *kink function*

$$x_+ := xY(x) \qquad (x \in \mathbb{R}).$$

Indeed we have by the Leibniz rule (Proposition 3.1.5) $x'_+ = (xY(x))' = Y(x) + x\,\delta(x) = Y(x)$. We observe that x_+ is even continuous[4] and that $x''_+ = \delta$.

Successively defining primitive functions with value 0 at $x = 0$ we obtain with the functions $x_+^{k-1} \in C^{k-2}(\mathbb{R})$ the relations

$$\left(\frac{x_+^{k-1}}{(k-1)!}\right)^{(k)} = \delta \qquad (k = 2, 3, \ldots).$$

[4]Since any other antiderivative differs from x_+ by a constant, we deduce continuity of any primitive function of Y.

This follows easily by induction.
(ii) *The multidimensional case*: We use coordinates $x = (x_1, \ldots, x_n)$ in $\mathbb{R}^n$ and put

$$E_k(x) := \frac{(x_1)_+^{k-1}(x_2)_+^{k-1} \cdots (x_n)_+^{k-1}}{((k-1)!)^n}. \tag{4.5.1}$$

Then $E_k \in C^{k-2}(\mathbb{R}^n)$ and we have

$$(\partial_1\partial_2 \cdots \partial_n)^k E_k = \delta \qquad (k = 2, 3, \ldots). \tag{4.5.2}$$

In the terminology of Section 3.5, we may restate (4.5.2) as follows: *E_k is a fundamental solution for the partial differential operator $(\partial_1 \cdots \partial_n)^k$.* Furthermore, Y is a fundamental solution for $\frac{d}{dx}$, x_+ is a fundamental solution for $(\frac{d}{dx})^2$ etc.

The following theorem shows that indeed distributions are a *minimal* extension of the space of continuous functions:

4.5.2 Theorem. *Let $T \in \mathcal{D}'(\mathbb{R}^n)$ and $\Omega \subseteq \mathbb{R}^n$ be open and bounded. Then there exists $f \in C(\mathbb{R}^n)$ and $\alpha \in \mathbb{N}_0^n$ such that*

$$T\mid_\Omega = \partial^\alpha(f\mid_\Omega).$$

Thus, locally every distribution is the (distributional) derivative of a continuous function.

Proof. The boundedness of Ω allows us to choose $\psi \in \mathcal{D}(\mathbb{R}^n)$ with $\psi = 1$ on Ω. We set $\tilde{T} = \psi T$, then $T\mid_\Omega = \tilde{T}\mid_\Omega$ and $\tilde{T} \in \mathcal{E}'(\mathbb{R}^n)$. By Corollary 2.3.7, $\tilde{T}$ is of finite order N, say. We have

$$\tilde{T} = \delta * \tilde{T} \underset{(4.5.2)}{=} (\partial_1 \cdots \partial_n)^{N+2}\, E_{N+2} * \tilde{T},$$

hence it suffices to show that $E_{N+2} * \tilde{T}$ is continuous.
Let $\rho \in \mathcal{D}(\mathbb{R}^n)$ be a mollifier and $\rho_\varepsilon(x) = \rho(x/\varepsilon)/\varepsilon^n$ ($x \in \mathbb{R}^n$, $0 < \varepsilon \leq 1$). Consider

$$f_\varepsilon := \big(E_{N+2} * \tilde{T}\big) * \rho_\varepsilon,$$

which is in $C^\infty(\mathbb{R}^n)$ by Theorem 4.3.1. Since $\tilde{T}$ and ρ_ε both have compact support, we may use associativity and commutativity of the convolution and obtain

$$f_\varepsilon(x) = \tilde{T} * \underbrace{(E_{N+2} * \rho_\varepsilon)}_{\in\mathcal{C}^\infty \text{ Th. 4.3.1}}(x) \underset{(4.3.1)}{=} \langle \tilde{T}(y), (E_{N+2} * \rho_\varepsilon)(x-y)\rangle.$$

Recall from Theorem 1.3.4 (ii) that $E_{N+2} * \rho_\varepsilon \to E_{N+2}$ in $C^N(\mathbb{R}^n)$ as $\varepsilon \to 0$. In particular, $(E_{N+2} * \rho_\varepsilon)$ is a Cauchy net in $C^N(\mathbb{R}^n)$.
Since $\tilde{T} \in \mathcal{E}'(\mathbb{R}^n)$ and is of order N we have the seminorm estimate (2.3.1) with derivative order N and some $C > 0$ and a fixed compact set $K \Subset \mathbb{R}^n$. Applying this to $f_\varepsilon - f_\eta$ ($0 < \varepsilon, \eta \leq 1$) we obtain for any compact subset $L \Subset \mathbb{R}^n$ and arbitrary $x \in L$

$$\begin{aligned}
|f_\varepsilon(x) - f_\eta(x)| &= |\langle \tilde{T}(y), (E_{N+2} * \rho_\varepsilon - E_{N+2} * \rho_\eta)(x-y)\rangle| \\
&\leq C \sum_{|\alpha|\leq N} \|\partial^\alpha(E_{N+2} * \rho_\varepsilon - E_{N+2} * \rho_\eta)(x - .)\|_{\infty,K} \\
&\leq C \sum_{|\alpha|\leq N} \|\partial^\alpha(E_{N+2} * \rho_\varepsilon - E_{N+2} * \rho_\eta)\|_{\infty,L-K}.
\end{aligned}$$

Upon taking the supremum over $x \in L$ we deduce that (f_ε) is a Cauchy net in $C(\mathbb{R}^n)$, thus converges uniformly on compact sets to some function $f \in C(\mathbb{R}^n)$.

On the other hand, by separate sequential continuity of convolution we also obtain the convergence

$$f_\varepsilon = E_{N+2} * (\tilde{T} * \rho_\varepsilon) \to E_{N+2} * (\tilde{T} * \delta) = E_{N+2} * \tilde{T} \qquad (\varepsilon \to 0).$$

Therefore the equality $E_{N+2} * \tilde{T} = f \in C(\mathbb{R}^n)$ must hold, completing the proof. □

For compactly supported distributions we even obtain a global result:

4.5.3 Theorem. *Let $T \in \mathcal{E}'(\mathbb{R}^n)$ and U be an open neighborhood of* $\operatorname{supp}(T)$. *Then we can find $m \in \mathbb{N}$ and functions $f_\beta \in C(\mathbb{R}^n)$ ($|\beta| \leq m$) with* $\operatorname{supp}(f_\beta) \Subset U$ *such that*

$$T = \sum_{|\beta| \leq m} \partial^\beta f_\beta.$$

In other words, every compactly supported distribution can be (globally) represented by a finite sum of (distributional) derivatives of continuous functions.

Proof. Choose $\Omega \subseteq \mathbb{R}^n$ open and bounded such that $\operatorname{supp}(T) \subseteq \Omega \subseteq \overline{\Omega} \Subset U$. The local structure theorem provides us with a function $f \in C(\mathbb{R}^n)$ such that $T \mid_\Omega = \partial^\alpha (f \mid_\Omega)$.

Let $\chi \in \mathcal{D}(\Omega)$ with $\chi = 1$ on a neighborhood of $\operatorname{supp}(T)$. Then we have for any $\varphi \in \mathcal{C}^\infty(\mathbb{R}^n)$

$$\begin{aligned}
\langle T, \varphi \rangle &= \langle T, \chi\varphi \rangle = \langle \partial^\alpha f, \chi\varphi \rangle = (-1)^{|\alpha|} \langle f, \partial^\alpha(\chi\varphi) \rangle \qquad \text{[Leibniz' rule]} \\
&= \sum_{\beta \leq \alpha} (-1)^{|\alpha|} \binom{\alpha}{\beta} \underbrace{\langle f, \partial^{\alpha-\beta}\chi\, \partial^\beta \varphi \rangle}_{(-1)^{|\beta|} \langle \partial^\beta (f\, \partial^{\alpha-\beta}\chi), \varphi \rangle} = \sum_{\beta \leq \alpha} (-1)^{|\alpha|+|\beta|} \binom{\alpha}{\beta} \langle \partial^\beta (f\, \partial^{\alpha-\beta}\chi), \varphi \rangle \\
&= \sum_{\beta \leq \alpha} \langle \partial^\beta \Big(\underbrace{(-1)^{|\alpha|+|\beta|} \binom{\alpha}{\beta} f\, \partial^{\alpha-\beta}\chi}_{=: f_\beta} \Big), \varphi \rangle = \langle \sum_{\beta \leq \alpha} \partial^\beta f_\beta, \varphi \rangle.
\end{aligned}$$

□

4.6 Fundamental solutions and convolution

Now that we have convolution as a tool at our disposal, let us revisit the concept of fundamental solution. Recall from the discussion preceding (3.2.5) that in order to solve

$$P(\partial)T = S$$

for a constant coefficient linear PDO $P(\partial)$, knowledge of a fundamental solution E allows one to obtain a solution by convolution: $T := S * E$

$$P(\partial)(T) = P(\partial)(E) * S = \delta * S = S.$$

This requires that E and S be convolvable, i.e., we need to ensure the existence of the convolution defining T (e.g., by requiring S to have compact support).

Let us illustrate the implications of these observations in two cases of differential operators whose fundamental solutions we have already calculated in Section 3.5, beginning with the wave operator on $\mathbb{R}^4$.

4.6.1 Theorem. *Let $S \in \mathcal{D}'(\mathbb{R}^4)$ with* $\operatorname{supp}(S) \subseteq \{x \mid x_0 \geq 0\}$. *Then the wave equation*

$$\Box T = (-\partial_0^2 + \partial_1^2 + \partial_2^2 + \partial_3^2)T = S \tag{4.6.1}$$

has a unique solution $T \in \mathcal{D}'(\mathbb{R}^4)$ with $\operatorname{supp}(T) \subseteq \{x \mid x_0 \geq 0\}$. *This solution is given by $T = S * E^+$. It satisfies* $\operatorname{supp}(T) \subseteq \operatorname{supp}(S) + \Gamma^+$.

Proof. We first have to show that $T = E^+ * S$ is well defined. According to Definition 4.4.3 this can be done by verifying that

$$+\colon \Gamma^+ \times \operatorname{supp}(S) \to \mathbb{R}^4$$

is proper. For this we use Lemma 4.4.2. So let $(|x'|, x') \in \Gamma^+$, $y \in \operatorname{supp}(S)$, and $|(|x'|, x')| + (y_0, y')| \leq \eta$. Then $|x'| + y_0 \leq \eta$, so $|x'| \leq \eta$ and $y_0 \leq \eta$, and

$$|y'| = |y' + x' - x'| \leq |y' + x'| + |x'| \leq 2\eta,$$

giving the claim. So indeed $T = E^+ * S$ is well defined and

$$\Box(T) = \Box(E^+) * S = \delta * S = S.$$

Finally, to obtain uniqueness, suppose that $\tilde{T} \in \mathcal{D}'(\mathbb{R}^4)$ is another solution with support in $\{x \mid x_0 \geq 0\}$, and set $R := T - \tilde{T}$. Then since $\operatorname{supp}(R) \subseteq \{x \mid x_0 \geq 0\}$, the convolutions in the following calculation are well defined:

$$R = \delta * R = \Box(E^+) * R = E^+ * \Box(R) = 0.$$

$\Box$

4.6.2 Corollary. *E^+ is the unique fundamental solution of $\Box$ with support in $\{x \mid x_0 \geq 0\}$.*

Proof. Set $S = \delta$ in Theorem 4.6.1. $\Box$

Theorem 4.6.1 remains valid under the assumption $\operatorname{supp}(S) \subseteq \{x \mid x_0 \geq c\}$ for some $c \in \mathbb{R}$. Also, analogous results hold for E^-.

Now let us compute the solution $E^+ * S$ more explicitly. Given $\varphi \in \mathcal{D}(\mathbb{R}^4)$, using Remark 4.4.4 (ii) we have

$$\begin{aligned}
\langle E^+ * S, \varphi \rangle &= \langle S(y), \langle E^+(x), \varphi(x+y) \rangle \rangle \\
&= -\frac{1}{4\pi} \left\langle S(y), \int \frac{\varphi(|x'| + y_0, x' + y')}{|x'|}\, dx' \right\rangle \\
&\underset{x'+y' \to x'}{=} -\frac{1}{4\pi} \left\langle S(y), \int \frac{\varphi(|x' - y'| + y_0, x')}{|x' - y'|}\, dx' \right\rangle
\end{aligned} \tag{4.6.2}$$

In particular, for $S = T_f$ with $f \in C^\infty(\mathbb{R}^4)$ we obtain

$$\begin{aligned}
\langle E^+ * S, \varphi \rangle &= -\frac{1}{4\pi} \int \frac{\varphi(|x' - y'| + y_0, x')}{|x' - y'|} f(y)\, dx' dy' dy_0 \\
&\underset{|x'-y'|+y_0 \to x_0}{=} -\frac{1}{4\pi} \int \varphi(x) \frac{f(x_0 - |x' - y'|, y')}{|x' - y'|}\, dx dy'.
\end{aligned}$$

Thus for $f \in C^\infty(\mathbb{R}^4)$:

$$E^+ * f(x) = -\frac{1}{4\pi} \int \frac{f(x_0 - |x' - y'|, y')}{|x' - y'|}\, dy' \ \ldots \textit{retarded potential.}$$

Analogously,

$$E^- * f(x) = -\frac{1}{4\pi}\int \frac{f(x_0+|x'-y'|,y')}{|x'-y'|}\,dy' \ldots \textit{advanced potential.}$$

Using this, (4.6.2) gives:

$$\langle E^+ * f,\varphi\rangle = \langle f, E^- * \varphi\rangle \qquad (\varphi\in\mathcal{D}(\mathbb{R}^4)).$$

Next we turn to the heat operator $P=\partial_t-\Delta_n$, this time considering the corresponding initial value problem

$$\begin{aligned} Pv &= 0 \quad \text{ on } \mathbb{R}_+\times\mathbb{R}^n \\ v(0,x) &= f \quad v\in C^2(\mathbb{R}_+\times\mathbb{R}^n)\cap C^0(\overline{\mathbb{R}}_+\times\mathbb{R}^n). \end{aligned} \tag{4.6.3}$$

Here,

$$C^0(\overline{\mathbb{R}}_+\times\mathbb{R}^n) = \{v\in C^0(\mathbb{R}_+\times\mathbb{R}^n)\mid \exists\bar v\in C^0(\mathbb{R}\times\mathbb{R}^n): v=\bar v|_{\mathbb{R}_+\times\mathbb{R}^n}\}.$$

Equivalently, $C^0(\overline{\mathbb{R}}_+\times\mathbb{R}^n)$ consists of those $v\in C^0(\mathbb{R}_+\times\mathbb{R}^n)$ that can be extended continuously to the boundary (a $\bar v$ as above can then be defined by $\bar v(t,x)=v(-t,x)$) for $t<0$.

Since distributions in general do not possess restrictions to boundaries, we first need to find a reformulation of the initial value problem (4.6.3) in a purely distributional way. The procedure we shall use can also serve as a blueprint for other such problems.

Let us begin by supposing that v is a solution to (4.6.3) and define

$$v^{\mathrm{c}}(t,x) := \begin{cases} v(t,x) & t>0 \\ 0 & t\le 0. \end{cases} \tag{4.6.4}$$

Thus v^{c} is the *cutoff* of an extension of v to $\mathbb{R}\times\mathbb{R}^n$ at the boundary $\{0\}\times\mathbb{R}^n$ of $\mathbb{R}^+\times\mathbb{R}^n$. Then using integration by parts and (4.6.4),

$$\begin{aligned}\langle Pv^{\mathrm{c}},\varphi\rangle = \langle v^{\mathrm{c}},P^\dagger\varphi\rangle &= -\int_{\{t>0\}} v(\partial_t+\Delta)\varphi\,dxdt \\ &= \int_{\{t>0\}}\underbrace{(\partial_t v-\Delta v)}_{=0}\varphi\,dxdt + \int v(0,x)\varphi(0,x)\,dx = \int f(x)\varphi(0,x)\,dx\end{aligned}$$

for any $\varphi\in\mathcal{D}(\mathbb{R}\times\mathbb{R}^n)$. Thus in $\mathcal{D}'(\mathbb{R}\times\mathbb{R}^n)$ we have

$$Pv^{\mathrm{c}} = \delta(t)\otimes f(x). \tag{4.6.5}$$

This motivates us to view (4.6.5), together with $\operatorname{supp}(v^{\mathrm{c}})\subseteq\overline{\mathbb{R}}_+\times\mathbb{R}^n$ as the distributional equivalent of the initial value problem (4.6.3). So we are looking for some $T\in\mathcal{D}'(\mathbb{R}\times\mathbb{R}^n)$ such that

$$P(T) = \delta(t)\otimes f(x), \tag{4.6.6}$$

where (we can even allow) $f\in\mathcal{D}'(\mathbb{R}^n)$, with $\operatorname{supp}(T)\subseteq\overline{\mathbb{R}}_+\times\mathbb{R}^n$. To simplify things we assume $f\in\mathcal{E}'(\mathbb{R}^n)$, so that $\delta(t)\otimes f(x)$ has compact support. Then

$$T := E(t,x) * (\delta(t)\otimes f(x)) \tag{4.6.7}$$

with E the fundamental solution from (3.5.1) exists and satisfies (4.6.6). Also, the support of E is $\overline{\mathbb{R}}_+\times\mathbb{R}^n$, and $\operatorname{supp}(\delta(t)\otimes f(x))$ is contained in $\{0\}\times\mathbb{R}^n$, so

$\text{supp}\,(T) \subseteq \overline{\mathbb{R}}_+ \times \mathbb{R}^n$ by Proposition 4.2.4 (i), verifying that T is indeed a solution to the distributional initial value problem.

We close this section by showing the compatibility of the distributional solution concept with the classical one. To this end we will establish that the solution T from (4.6.7) is in fact smooth on $\mathbb{R}_+ \times \mathbb{R}^n$ and that it converges to f in $\mathcal{D}'(\mathbb{R}^n)$ as $t \to 0+$. So let $\varphi \in \mathcal{D}(\mathbb{R} \times \mathbb{R}^n)$. Then

$$\begin{aligned}\langle T, \varphi \rangle &= \langle E(t,x) \otimes \delta(s) \otimes f(y), \varphi(s+t, x+y) \rangle \\ &= \left\langle f(y), \int E(t,x)\varphi(t, x+y)\, dtdx \right\rangle \\ &= \left\langle f(y), \int E(t, x-y)\varphi(t,x)\, dtdx \right\rangle.\end{aligned}$$

Recall from (3.5.1) that E is C^∞ on $\mathbb{R}_+ \times \mathbb{R}^n$. Therefore, if $\varphi \in \mathcal{D}(\mathbb{R}_+ \times \mathbb{R}^n)$ and if $\sigma \in \mathcal{D}(\mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^n)$ is a cutoff-function that is 1 on a neighborhood of $\text{supp}\,(\varphi) \times \text{supp}\,(f)$, we have

$$\begin{aligned}\left\langle f(y), \int E(t, x-y)\varphi(t,x)\, dtdx \right\rangle &= \langle f(y) \otimes \varphi(t,x), \sigma(t,x,y)E(t,x-y) \rangle \\ &= \int \langle f(y), E(t, x-y) \rangle \varphi(t,x)\, dtdx.\end{aligned}$$

It follows that

$$T = \langle f(y), E(t, x-y) \rangle \qquad t > 0. \tag{4.6.8}$$

By Corollary 3.6.6, this shows that T is C^∞ on $t > 0$. This also implies that T satisfies $P(T) = 0$ on $t > 0$ in the classical sense. To determine the limiting behavior, let $\varphi \in \mathcal{D}(\mathbb{R}^n)$. Then for $t > 0$ we get

$$\begin{aligned}\int E(t,x)\varphi(x)\, dx &= \frac{1}{(4\pi t)^{n/2}} \int \varphi(x) e^{-\frac{|x|^2}{4t}}\, dx \\ &= \pi^{-\frac{n}{2}} \int \varphi(2yt^{\frac{1}{2}}) e^{-|y|^2}\, dy \to \varphi(0)\end{aligned}$$

as $t \to 0+$ by dominated convergence. This means that

$$\lim_{t \to 0+} E(t,x) = \delta(x)$$

in $\mathcal{D}'(\mathbb{R}^n)$. Combining this with (4.6.8), Theorems 4.3.1 and 4.2.5 imply that $T \to f$ in $\mathcal{D}'(\mathbb{R}^n)$ as $t \to 0+$.

Chapter 5

Temperate distributions and Fourier transform

5.1 The classical Fourier transform

Recall that $L^1(\mathbb{R}^n)$ is the vector space of equivalence classes of Lebesgue integrable functions f on $\mathbb{R}^n$, i.e. $\int_{\mathbb{R}^n} |f(x)|\, dx < \infty$ as Lebesgue integral, modulo the relation of 'being equal (Lebesgue) almost everywhere'. Following traditional abuse of notion and notation we typically work with elements of L^1 as if they were functions, thus, strictly speaking, mixing up a representative with its equivalence class.

The *Fourier transform* of a function $f \in L^1(\mathbb{R}^n)$ is defined to be the function $\mathcal{F}(f)\colon \mathbb{R}^n \to \mathbb{C}$, given by ($x\xi$ denoting the standard inner product of x and ξ on $\mathbb{R}^n$)

$$\mathcal{F}(f)(\xi) = \widehat{f}(\xi) := \int_{\mathbb{R}^n} f(x) e^{-ix\xi}\, dx \qquad (\xi \in \mathbb{R}^n). \tag{5.1.1}$$

(For every ξ the value of the integral is finite, since $|f(x)e^{-ix\xi}| = |f(x)|$ is L-integrable; furthermore, $\mathcal{F}(f)(\xi)$ does not depend on the L^1-representative, since altering f on a set of Lebesgue measure zero does not change the value of the integral.)

The following result lists some basic properties of the classical Fourier transform.

5.1.1 Theorem.

(i) For every $f \in L^1(\mathbb{R}^n)$ the Fourier transform $\widehat{f}\colon \mathbb{R}^n \to \mathbb{C}$ is continuous and satisfies

$$|\widehat{f}(\xi)| \leq \|f\|_1 \qquad \forall \xi \in \mathbb{R}^n. \tag{5.1.2}$$

(ii) If $f, g \in L^1(\mathbb{R}^n)$, then

$$\int f(x)\, \widehat{g}(x)\, dx = \int \widehat{f}(\xi)\, g(\xi)\, d\xi. \tag{5.1.3}$$

(iii) If $f, g \in L^1(\mathbb{R}^n)$, then $x \mapsto f * g(x) = \int f(y)g(x-y)dy$ defines an element $f * g \in L^1(\mathbb{R}^n)$ and we have

$$\widehat{(f * g)} = \widehat{f} \cdot \widehat{g}. \tag{5.1.4}$$

Proof. (i) As remarked immediately after the definition, $\widehat{f}(\xi)$ is well defined and finite. Moreover, the triangle inequality for integrals yields

$$|\widehat{f}(\xi)| \leq \int |f(x)e^{-ix\xi}|\, dx = \int |f(x)|\, dx = \|f\|_1.$$

If $\xi_k \to \xi$ as $k \to \infty$, then $f(x)e^{-ix\xi_k} \to f(x)e^{-ix\xi}$ pointwise and $|f(x)e^{-ix\xi_k}| \leq |f(x)|$ provides an L^1-bound uniformly for all k. Thus dominated convergence implies $\widehat{f}(\xi_k) \to \widehat{f}(\xi)$ $(k \to \infty)$, hence continuity of $\widehat{f}$.

(ii) By (i) we have that $|\widehat{g}| \leq \|g\|_1$, hence $\widehat{g}$ is bounded and $f\,\widehat{g} \in L^1(\mathbb{R}^n)$. Furthermore,

$$\begin{aligned}\int f(x)\,\widehat{g}(x)\, dx &= \int f(x) \int g(\xi)\, e^{-ix\xi}\, d\xi\, dx = \int g(\xi) \int f(x)e^{-ix\xi}\, dx\, d\xi \\ &= \int g(\xi)\,\widehat{f}(\xi)\, d\xi.\end{aligned}$$

(iii) Observe that $(x, y) \mapsto f(y)g(x-y)$ is L-measurable and

$$\begin{aligned}\int_{\mathbb{R}^n\times\mathbb{R}^n} |f(y)g(x-y)|\, d(x,y) &= \int |f(y)| \int |g(x-y)|\, dx\, dy \\ &= \int |f(y)|\|g\|_1\, dy = \|g\|_1\|f\|_1 < \infty.\end{aligned}$$

Hence $x \mapsto \int f(y)g(x-y)\, dy = f * g(x)$ defines an integrable function on $\mathbb{R}^n$. We determine its Fourier transform as follows

$$\begin{aligned}\widehat{(f*g)}(\xi) &= \int e^{-ix\xi}\, (f*g)(x)\, dx = \int e^{-ix\xi} \int f(y)g(x-y)dy\, dx \\ &= \int f(y) \int g(x-y)e^{-ix\xi}\, dx\, dy \underset{[z=x-y]}{\overset{}{\underset{\uparrow}{=}}} \int f(y) \int g(z)e^{-i(z+y)\xi}\, dz\, dy \\ &= \int f(y)e^{-iy\xi} \underbrace{\int g(z)e^{-iz\xi}\, dz}_{=\widehat{g}(\xi)}\, dy = \widehat{f}(\xi)\,\widehat{g}(\xi).\end{aligned}$$

□

Our plan is to extend the Fourier transform to distributions by the standard procedure of transposition, i.e., letting $\mathcal{F}$ act on test functions. It turns out, however, that $\mathcal{D}(\mathbb{R}^n)$ is not an appropriate space of test functions for this operation:

If $f \in L^1(\mathbb{R}^n) \subseteq \mathcal{D}'(\mathbb{R}^n)$ and $\varphi \in \mathcal{D}(\mathbb{R}^n) \subseteq L^1(\mathbb{R}^n)$, then (5.1.3) gives

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

Thus it is tempting to try defining the Fourier transform of any $T \in \mathcal{D}'(\mathbb{R}^n)$ by our standard duality trick in the form

$$\langle \widehat{T}, \varphi \rangle := \langle T, \widehat{\varphi} \rangle.$$

However, if $\varphi \in \mathcal{D}(\mathbb{R}^n)$ then $\widehat{\varphi}$ cannot have compact support unless $\varphi = 0$. For simplicity, let us give details in the one-dimensional case: Let $R > 0$ such that $\operatorname{supp}(\varphi) \subseteq [-R, R]$. Using the power series expansion of the exponential function and its uniform convergence on the compact set $[-R, R]$ we may write

$$\widehat{\varphi}(\xi) = \int_{-R}^{R} \sum_{k=0}^{\infty} \frac{(-ix\xi)^k}{k!} \varphi(x)\, dx = \sum_{k=0}^{\infty} \big(\underbrace{\frac{(-i)^k}{k!} \int_{-R}^{R} \varphi(x)x^k\, dx}_{=:a_k} \big) \cdot \xi^k = \sum_{k=0}^{\infty} a_k \xi^k,$$

where $|a_k| \leq 2R\|\varphi\|_\infty R^k/k!$. This shows that $\widehat{\varphi}(\xi)$ is represented by a power series with infinite radius of convergence, thus is a real analytic function (that can be extended to a holomorphic function on all of $\mathbb{C}$). If $\operatorname{supp}(\widehat{\varphi})$ is compact, then $\widehat{\varphi}$ vanishes on a set with accumulation points, hence $\widehat{\varphi} = 0$ (everywhere). (As we will see below, this implies $\varphi = 0$.)

We conclude that $\mathcal{F}(\mathcal{D}) \not\subseteq \mathcal{D}$ and ask the question, whether there is a function space $\mathcal{Y}$ on $\mathbb{R}^n$ with $\mathcal{D} \subseteq \mathcal{Y} \subseteq L^1 \cap \mathcal{E}$ such that $\mathcal{F}(\mathcal{Y}) \subseteq \mathcal{Y}$.

A further natural requirement will be that $\mathcal{Y}$ should be invariant under differentiation. Observe that then we further obtain that $\forall \varphi \in \mathcal{Y}$

$$\widehat{\partial_j \varphi}(\xi) = \int e^{-ix\xi} \partial_j \varphi(x)\, dx = -\int (-i\xi_j) e^{-ix\xi} \varphi(x)\, dx = i\xi_j \widehat{\varphi}(\xi)$$

should belong to $\mathcal{Y}$. By induction we deduce that also multiplication by polynomials should leave $\mathcal{Y}$ invariant. Furthermore, a calculation similar to the above shows $(x_j\varphi)\widehat{\ } = i\partial_j \widehat{\varphi}$ etc. Thus, we are led to the additional condition that also

$$x^\alpha \partial^\beta \mathcal{Y} \subseteq \mathcal{Y} \qquad \forall \alpha, \beta \in \mathbb{N}_0^n$$

should hold. As we shall see in the following section, an appropriate function space is given by considering smooth functions φ such that $x^\alpha \partial^\beta \varphi(x)$ is bounded (for all α, β).

5.2 The space of rapidly decreasing functions

To avoid extra factors of the form $(-i)^{|\alpha|}$ from popping up in many calculations, it is very common to introduce the operator

$$D_j := \frac{1}{i}\partial_j \qquad (j = 1, \ldots, n)$$

and $D = (D_1, \ldots, D_n)$. Note that we then have $D_j(e^{ix\xi}) = \xi_j e^{ix\xi}$ and furthermore, $p(D)(e^{ix\xi}) = p(\xi)e^{ix\xi}$, if p is any polynomial function on $\mathbb{R}^n$.

5.2.1 Definition. *Let $\varphi \in \mathcal{C}^\infty(\mathbb{R}^n)$.*

(i) The function φ is said to be rapidly decreasing *if it satisfies the following semi-norm condition*

$$\forall \alpha, \beta \in \mathbb{N}_0^n : \quad q_{\alpha,\beta}(\varphi) := \sup_{x \in \mathbb{R}^n} |x^\alpha D^\beta \varphi(x)| < \infty. \tag{5.2.1}$$

(ii) The vector space of all rapidly decreasing functions on $\mathbb{R}^n$ is denoted by $\mathcal{S}(\mathbb{R}^n)$.

(iii) Let (φ_m) be a sequence in $\mathcal{S}(\mathbb{R}^n)$. We define convergence of (φ_m) to φ in $\mathcal{S}(\mathbb{R}^n)$ (as $m \to \infty$), denoted also by $\varphi_m \stackrel{\mathcal{S}}{\to} \varphi$, by the property

$$\forall \alpha, \beta \in \mathbb{N}_0^n : \quad q_{\alpha,\beta}(\varphi_m - \varphi) \to 0 \quad (m \to \infty).$$

(Similarly for nets like $(\varphi_\varepsilon)_{0<\varepsilon\leq 1}$.)

5.2.2 Remark.

(i) Let $\varphi \in \mathcal{C}^\infty(\mathbb{R}^n)$, then condition (5.2.1) is equivalent to the following statement

$$\forall \gamma \in \mathbb{N}_0^n\, \forall l \in \mathbb{N}_0\, \exists C > 0 : \quad |D^\gamma \varphi(x)| \leq \frac{C}{(1+|x|)^l} \qquad \forall x \in \mathbb{R}^n. \tag{5.2.2}$$

It is obvious that (5.2.1) is a consequence of (5.2.2). On the other hand, (5.2.1) implies[1]

$$\forall \gamma \in \mathbb{N}_0^n \, \forall k \in \mathbb{N}_0 \, \exists C > 0 : \quad \underbrace{\sup_{x\in\mathbb{R}^n} |D^\gamma \varphi(x)| + \sup_{x\in\mathbb{R}^n} |\,|x|^{2k} D^\gamma \varphi(x)|}_{\geq \sup_{x\in\mathbb{R}^n} |\,(1+|x|^{2k}) D^\gamma \varphi(x)|} \leq C,$$

which in turn gives (5.2.2) upon noting that $1/(1+|x|^{2k}) \leq C_{k,l}/(1+|x|)^l$ when $2k \geq l$ with an appropriate constant $C_{k,l}$.

(ii) We clearly have $\mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n) \subset \mathcal{E}(\mathbb{R}^n)$.

(iii) An explicit example of a function $\varphi \in \mathcal{S}(\mathbb{R}^n) \setminus \mathcal{D}(\mathbb{R}^n)$ is $\varphi(x) = e^{-c|x|^2}$ with $\mathrm{Re}(c) > 0$.

(iv) Convergence in $\mathcal{S}(\mathbb{R}^n)$ (and, in fact, also the topology of $\mathcal{S}(\mathbb{R}^n)$) is equivalently described by the *increasing* sequence of semi-norms

$$Q_k(\varphi) := \sum_{|\alpha|,|\beta|\leq k} q_{\alpha,\beta}(\varphi) \qquad (k \in \mathbb{N}_0).$$

('Increasing' since $k \leq k'$ implies $Q_k(\varphi) \leq Q_{k'}(\varphi)$.)
Moreover, we claim that convergence (and also the topology) in $\mathcal{S}(\mathbb{R}^n)$ can also be described by the metric $d\colon \mathcal{S}(\mathbb{R}^n) \times \mathcal{S}(\mathbb{R}^n) \to \mathbb{R}$, defined by

$$d(\varphi,\psi) := \sum_{k=0}^{\infty} 2^{-k} \frac{Q_k(\varphi-\psi)}{1+Q_k(\varphi-\psi)}.$$

(Regarding the abstract theory in the background, this stems from the general fact that a locally convex vector space is metrizable if and only if its topology is generated by a countable number of semi-norms. The construction of the metric is as in [Hor66, Chapter 2.6, Proposition 2].)
We comment on the proof of the above claim:

- In showing that d indeed defines a metric the only nontrivial part is the triangle inequality $d(\varphi,\psi) \leq d(\varphi,\rho) + d(\rho,\psi)$. Use that the function $f\colon [0,\infty[\to [0,\infty[$, $f(x) = x/(1+x)$, is increasing and that $Q_k(\varphi-\psi) \leq Q_k(\varphi-\rho) + Q_k(\rho-\psi)$. Finally, in every summand (as $k = 0,1,2,\ldots$) use the following simple estimate valid for any $a,b \geq 0$: $\frac{a+b}{1+a+b} = \frac{a}{1+a+b} + \frac{b}{1+a+b} \leq \frac{a}{1+a} + \frac{b}{1+b}$.
- That convergence with respect to the metric d implies $\mathcal{S}$-convergence as defined above is clear. Conversely, assume that (φ_m) is a sequence converging to φ in $\mathcal{S}$. We have to show that $d(\varphi_m,\varphi) \to 0$ as $m \to \infty$.

  Let $\varepsilon > 0$. Chose $N \in \mathbb{N}$ so that $\varepsilon > 1/2^{N-1}$. There exists $m_0 \in \mathbb{N}$ such that $Q_N(\varphi_m - \varphi) < \varepsilon/4$ holds for all $m \geq m_0$. Thus we obtain for any $m \geq m_0$

$$\begin{aligned} d(\varphi_m,\varphi) &= \sum_{k=0}^{N} 2^{-k} \overbrace{\frac{Q_k(\varphi_m-\varphi)}{1+Q_k(\varphi_m-\varphi)}}^{\leq Q_N(\varphi_m-\varphi)} + \sum_{k=N+1}^{\infty} 2^{-k} \overbrace{\frac{Q_k(\varphi_m-\varphi)}{1+Q_k(\varphi_m-\varphi)}}^{\leq 1} \\ &\leq Q_N(\varphi_m-\varphi) \sum_{k=0}^{N} 2^{-k} + \sum_{k=N+1}^{\infty} 2^{-k} = Q_N(\varphi_m-\varphi)\cdot 2(1-2^{-N-1}) + \frac{1}{2^{N+1}}\cdot 2 \\ &< \frac{\varepsilon}{4}\cdot 2(1-0) + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

[1](putting $\beta = \gamma$ and $\alpha = 0, 2e_1, \ldots 2e_n, 4e_1, \ldots, 4e_n, \ldots, 2ke_1, \ldots, 2ke_n$ successively)

In particular, since $\mathcal{S}(\mathbb{R}^n)$ is a metric space, we need not distinguish between continuity and sequential continuity for maps defined on $\mathcal{S}(\mathbb{R}^n)$.

5.2.3 Theorem. $\mathcal{S}(\mathbb{R}^n)$ *is complete (as a metric space).*

Proof. Suppose (φ_j) is a Cauchy sequence in $\mathcal{S}(\mathbb{R}^n)$. Recall that $C_{\mathrm{b}}(\mathbb{R}^n) := C(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ equipped with the norm $\|\ \|_\infty$ is a Banach space. For every $\alpha, \beta \in \mathbb{N}_0^n$ we obtain that $(x^\alpha D^\beta \varphi_j)$ is a Cauchy sequence in C_{b}, thus converges to some $\varphi_{\alpha,\beta} \in C_{\mathrm{b}}$.

Put $\varphi := \varphi_{0,0}$. As in the proof of Theorem 1.1.4 it follows that $\varphi \in \mathcal{C}^\infty$ and that $D^\beta \varphi = \varphi_{0,\beta}$ for all $\beta \in \mathbb{N}_0^n$.

Moreover, since $\varphi_{\alpha,\beta} = C_{\mathrm{b}}\text{-}\lim x^\alpha D^\beta \varphi_j = \text{pointwise-}\lim x^\alpha D^\beta \varphi_j = x^\alpha \varphi_{0,\beta}$ we deduce that also $x^\alpha D^\beta \varphi = x^\alpha \varphi_{0,\beta} = \varphi_{\alpha,\beta}$.

If $N \in \mathbb{N}_0$ is arbitrary, but fixed, then we have for any $\alpha, \beta \in \mathbb{N}_0^n$ that $\|x^\alpha D^\beta \varphi\|_\infty \leq \|\varphi_{\alpha,\beta} - x^\alpha D^\beta \varphi_N\|_\infty + \|x^\alpha D^\beta \varphi_N\|_\infty < \infty$, hence $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Finally, the Cauchy sequence property of $(x^\alpha D^\beta \varphi_j)$ provides for any $\varepsilon > 0$ an index m_0 such that

$$\|x^\alpha D^\beta \varphi - x^\alpha D^\beta \varphi_l\|_\infty = \lim_{j\to\infty} \|x^\alpha D^\beta \varphi_j - x^\alpha D^\beta \varphi_l\|_\infty \leq \varepsilon \qquad (l \geq m_0).$$

Therefore $\varphi_l \to \varphi$ in $\mathcal{S}(\mathbb{R}^n)$ as $l \to \infty$. □

5.2.4 Definition. *The space of slowly increasing smooth functions is defined by*

$$\mathcal{O}_M(\mathbb{R}^n) := \{ f \in \mathcal{C}^\infty(\mathbb{R}^n) \mid \forall \alpha \in \mathbb{N}_0^n\, \exists N \in \mathbb{N}_0\, \exists C > 0\, \forall x \in \mathbb{R}^n : \\ |\partial^\alpha f(x)| \leq C(1+|x|)^N \ \}.$$

Clearly, polynomials belong to $\mathcal{O}_M(\mathbb{R}^n)$.

The following result clarifies the relation between $\mathcal{S}(\mathbb{R}^n)$ and other function spaces we have been considering before.

5.2.5 Theorem.

(i) Let $P(x,D)$ be a partial differential operator with coefficients in $\mathcal{O}_M(\mathbb{R}^n)$, i.e.,

$$P(x,D) = \sum_{|\gamma|\leq m} a_\gamma(x) D^\gamma \qquad (a_\gamma \in \mathcal{O}_M(\mathbb{R}^n)).$$

Then $P(x,D)\colon \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is linear and continuous.

(ii) $\mathcal{D}(\mathbb{R}^n) \subseteq \mathcal{S}(\mathbb{R}^n)$ with continuous embedding.

(iii) $\mathcal{D}(\mathbb{R}^n)$ is dense in $\mathcal{S}(\mathbb{R}^n)$.

(iv) $\mathcal{S}(\mathbb{R}^n) \subseteq L^1(\mathbb{R}^n)$ with continuous embedding.

Proof. (i): Linearity is clear. To show continuity we prove that $\varphi_j \to 0$ in $\mathcal{S}$ implies $P(x,D)\varphi_j \to 0$ in $\mathcal{S}$ (as $j \to \infty$). For any $\alpha, \beta \in \mathbb{N}_0^n$ we have $q_{\alpha,\beta}(P(x,D)\varphi_j) \leq \sum_{|\gamma|\leq m} q_{\alpha,\beta}(a_\gamma D^\gamma \varphi_j)$. Upon application of the Leibniz rule it only remains to estimate a linear combination of terms of the form $(\sigma \leq \beta)$

$$|x^\alpha|\, |D^{\beta-\sigma} a_\gamma(x)|\, |D^\sigma \varphi_j(x)| \leq |x^\alpha|\, C(1+|x|)^N |D^\sigma \varphi_j(x)| \leq \tilde{C} Q_k(\varphi_j),$$

for k sufficiently large, and this upper bound tends to 0 as $j \to \infty$.

(ii): Clearly $\varphi_j \to 0$ in $\mathcal{D}$ implies $\varphi_j \to 0$ in $\mathcal{S}$.

(iii): Choose a cut-off function $\rho \in \mathcal{D}(\mathbb{R}^n)$ with $\rho(x) = 1$ when $|x| \le 1$.

Let $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Define $\varphi_j(x) := \varphi(x)\rho(x/j)$ ($j \in \mathbb{N}$), then $\varphi_j \in \mathcal{D}(\mathbb{R}^n)$ and $\varphi(x) - \varphi_j(x) = 0$ when $|x| \le j$. We show that $\varphi_j \to \varphi$ in $\mathcal{S}$. In fact,

$$q_{\alpha,\beta}(\varphi - \varphi_j) = q_{\alpha,\beta}(\varphi(x)(1 - \rho(x/j))) \le C_{\alpha,\beta} \sup_{|\gamma| \le |\beta|} \sup_{|x| \ge j} |x^\alpha D^\gamma \varphi(x)| \to 0$$

as $j \to \infty$ because $|x^\alpha D^\gamma \varphi(x)| \le |x|^{-2} \sum_{|\delta| \le |\alpha|+2} q_{\delta,\gamma}(\varphi)$ for $|x| \ge j$.

(iv): For any $\varphi \in \mathcal{S}(\mathbb{R}^n)$ and N such that $(1 + |x|^2)^{-N} \in L^1(\mathbb{R}^n)$,

$$(1 + |x|^2)^N |\varphi(x)| \le C \sum_{|\alpha| \le 2N} q_{\alpha,0}(\varphi).$$

Therefore, $\|\varphi\|_{L^1} \le \tilde{C} \sum_{|\alpha| \le 2N} q_{\alpha,0}(\varphi)$, with $\tilde{C} = C\|(1 + |x|^2)^{-N}\|_{L^1}$. □

Thanks to Theorem 5.2.5 (iv) the Fourier transform is defined on $\mathcal{S} \subseteq L^1$ and Theorem 5.1.1(i) gives $\mathcal{F}(\mathcal{S}) \subseteq C_{\mathrm{b}}$. We will show that, in fact, Fourier transform is an isomorphism of $\mathcal{S}$. We split this task into several steps, starting by proving the following *exchange formulae*:

5.2.6 Lemma. *For any $\varphi \in \mathcal{S}(\mathbb{R}^n)$ and $\alpha \in \mathbb{N}_0^n$ we have*

$$(D^\alpha \varphi)\widehat{\ }(\xi) = \xi^\alpha \widehat{\varphi}(\xi) \tag{5.2.3}$$

and

$$(x^\alpha \varphi)\widehat{\ }(\xi) = (-1)^{|\alpha|} D^\alpha \widehat{\varphi}(\xi), \tag{5.2.4}$$

in particular, $\widehat{\varphi} \in \mathcal{C}^\infty(\mathbb{R}^n)$.

Proof. By Theorem 5.2.5 the functions $D^\alpha \varphi$ and $x^\alpha \varphi$ belong to $\mathcal{S}(\mathbb{R}^n) \subseteq L^1(\mathbb{R}^n)$, hence we may take their Fourier transforms and relieve the calculations in Section 5.1 of their informal status: First, we apply integration by parts and find

$$(D_j \varphi)\widehat{\ }(\xi) = \int D_j \varphi(x) e^{-ix\xi}\, dx = -\int \varphi(x)(-\xi_j) e^{-ix\xi}\, dx = \xi_j \widehat{\varphi}(\xi)$$

and (5.2.3) follows by induction. Second, standard theorems on differentiation of the parameter in the integral imply that $\widehat{\varphi}$ is continuously differentiable and

$$-D_j \widehat{\varphi}(\xi) = -D_{\xi_j} \int \varphi(x) e^{-ix\xi}\, dx = \int \varphi(x)\, x_j e^{-ix\xi}\, dx = (x_j \varphi)\widehat{\ }(\xi).$$

Equation (5.2.4) and smoothness of $\widehat{\varphi}$ then follow by induction. □

5.2.7 Lemma. *$\mathcal{F}(\mathcal{S}(\mathbb{R}^n)) \subseteq \mathcal{S}(\mathbb{R}^n)$ and $\varphi \mapsto \widehat{\varphi}$ is continuous $\mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$.*

Proof. Let $\varphi \in \mathcal{S}(\mathbb{R}^n)$. We know from the previous lemma that $\widehat{\varphi} \in \mathcal{C}^\infty(\mathbb{R}^n)$ and that the exchange formulae hold. To show that $\widehat{\varphi}$ belongs to $\mathcal{S}(\mathbb{R}^n)$ we have to establish an upper bound for $q_{\alpha,\beta}(\widehat{\varphi}) = \sup_{\xi \in \mathbb{R}^n} |\xi^\alpha D^\beta \widehat{\varphi}(\xi)|$, where $\alpha, \beta \in \mathbb{N}_0^n$ are arbitrary. Repeated application of the exchange formulae and the basic L^∞-L^1-estimate 5.1.1 (i) give

$$|\xi^\alpha D^\beta \widehat{\varphi}(\xi)| = |\xi^\alpha \mathcal{F}(x^\beta \varphi)(\xi)| = |\mathcal{F}(D^\alpha(x^\beta \varphi))(\xi)| \le \int |D^\alpha(x^\beta \varphi(x))|\, dx.$$

Hence $q_{\alpha,\beta}(\hat{\varphi}) \le \|D^\alpha(x^\beta \varphi(x))\|_{L^1}$, which as in the proof of Theorem 5.2.5 can be estimated some $C \sum_{|\gamma|,|\delta| \le N} q_{\gamma,\delta}(\varphi)$ for suitable N. This proves that $\widehat{\varphi} \in \mathcal{S}(\mathbb{R}^n)$ and also shows continuity of the Fourier transform as a linear operator on $\mathcal{S}(\mathbb{R}^n)$. □

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

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

Proof. In dimension $n = 1$ we note that $g(x) = e^{-x^2/2}$ satisfies the following linear first-order ordinary differential equation

$$g'(x) + xg(x) = 0 \tag{5.2.5}$$

with initial value $g(0) = 1$. (Recall that any solution to (5.2.5) is of the form $f(x) = ce^{-x^2/2}$, where $c = f(0)$.) Applying Fourier transform, (5.2.5) and the exchange formulae yield

$$i\xi\,\widehat{g}(\xi) + i\widehat{g}'(\xi) = 0,$$

thus $\widehat{g}$ also solves the differential equation (5.2.5). Therefore we must have $\widehat{g}(\xi) = c\exp(-\xi^2/2)$ and $c = \widehat{g}(0) = \int \exp(-\xi^2/2)d\xi = \sqrt{2\pi}$.
In dimension $n > 1$ we then calculate directly using Fubini's theorem

$$\begin{aligned}\big(e^{-|x|^2/2}\big)\widehat{\ }(\xi) &= \prod_{k=1}^{n} \int e^{-ix_k\xi_k} e^{-x_k^2/2}\,dx_k = \prod_{k=1}^{n} \widehat{g}(\xi_k)\\ &= \prod_{k=1}^{n} (2\pi)^{1/2} e^{-\xi_k^2/2} = (2\pi)^{n/2} e^{-|\xi|^2/2}.\end{aligned}$$

□

5.2.9 Lemma. *For any* $\varphi \in \mathcal{S}(\mathbb{R}^n)$ *the* Fourier inversion formula

$$\varphi(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \widehat{\varphi}(\xi) e^{ix\xi}\,d\xi \qquad (x \in \mathbb{R}^n) \tag{5.2.6}$$

holds.

Proof. For any $a \neq 0$ and $f \in \mathcal{S}(\mathbb{R}^n)$:

$$(x \mapsto f(ax))\widehat{\ }(\xi) = \int f(ax)e^{-ix\xi}\,dx = \int f(y)e^{-iy\xi/a}\,\frac{dy}{|a|^n} = \frac{1}{|a|^n}\,\widehat{f}\Big(\frac{1}{a}\xi\Big). \tag{5.2.7}$$

Recall that (5.1.3) gives for any $\varphi, \psi \in \mathcal{S}(\mathbb{R}^n)$

$$\int \widehat{\varphi}\,\psi = \int \varphi\,\widehat{\psi}. \tag{5.2.8}$$

Now let $\varphi \in \mathcal{S}(\mathbb{R}^n)$ be arbitrary and put $\psi(\xi) = e^{-|\varepsilon\xi|^2/2}$, where $\varepsilon > 0$. Then (5.2.8) together with (5.2.7) and Lemma 5.2.8 imply

$$\underbrace{\int \widehat{\varphi}(\xi)e^{-|\varepsilon\xi|^2/2}\,d\xi}_{=:l_\varepsilon} = \frac{(2\pi)^{n/2}}{\varepsilon^n} \int \varphi(x)e^{-|x|^2/(2\varepsilon^2)}\,dx \overset{[z=x/\varepsilon]}{=} \underbrace{(2\pi)^{n/2} \int \varphi(\varepsilon z)e^{-|z|^2/2}\,dz}_{=:r_\varepsilon}.$$

As $\varepsilon \to 0$ we have by dominated convergence

$$l_\varepsilon \to \int \widehat{\varphi}(\xi)\,d\xi$$

and that

$$r_\varepsilon \to (2\pi)^{n/2}\varphi(0) \underbrace{\int e^{-|z|^2/2}\,dz}_{=(2\pi)^{n/2}[\text{by } 5.2.8]} = (2\pi)^n\varphi(0),$$

hence

$$\varphi(0) = (2\pi)^{-n} \int \widehat{\varphi}(\xi)\, d\xi. \tag{5.2.9}$$

From this we will obtain the result by translation. To this end, we first note that for any $h \in \mathbb{R}^n$ and $f \in \mathcal{S}(\mathbb{R}^n)$

$$(\tau_{-h} f)\widehat{\ }(\xi) = \int f(x+h) e^{-ix\xi}\, dx = \int f(y) e^{-i(y-h)\xi}\, dx = e^{ih\xi}\, \widehat{f}(\xi). \tag{5.2.10}$$

(Translation of f corresponds to modulation of $\widehat{f}$.)

Thus we finally arrive at

$$\varphi(x) = (\tau_{-x}\varphi)(0) \underset{(5.2.9)}{=} (2\pi)^{-n} \int \widehat{(\tau_{-x}\varphi)}(\xi)\, d\xi = (2\pi)^{-n} \int e^{ix\xi} \widehat{\varphi}(\xi)\, d\xi.$$

□

5.2.10 Theorem. *The Fourier transform $\mathcal{F}\colon \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is linear and continuous with continuous inverse $\mathcal{F}^{-1}$ given by*

$$\mathcal{F}^{-1}\psi(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} \psi(\xi) e^{ix\xi}\, d\xi \qquad (\psi \in \mathcal{S}(\mathbb{R}^n)).$$

Hence we have for all $\varphi \in \mathcal{S}(\mathbb{R}^n)$

$$\widehat{\widehat{\varphi}} = (2\pi)^n \check{\varphi} \tag{5.2.11}$$

(recall that $\check{\varphi}(x) = \varphi(-x)$ and $\check{\check{\varphi}} = \varphi$).

Proof. By Lemma 5.2.9 the formula for $\mathcal{F}^{-1}$ gives a left-inverse of $\mathcal{F}$ on $\mathcal{S}(\mathbb{R}^n)$, i.e., $\mathcal{F}^{-1} \circ \mathcal{F} = \mathrm{id}_{\mathcal{S}}$. Hence $\mathcal{F}$ is injective.

To prove surjectivity of $\mathcal{F}$ let $\varphi \in \mathcal{S}(\mathbb{R}^n)$ be arbitrary. Then (5.2.6) yields

$$\varphi(x) = \frac{1}{(2\pi)^n} \int \widehat{\varphi}(\xi) e^{ix\xi}\, d\xi = \int \big((2\pi)^{-n} \widehat{\varphi}(-\xi)\big) e^{-ix\xi}\, d\xi = \mathcal{F}((2\pi)^{-n} \check{\widehat{\varphi}}).$$

Moreover, noting that (5.2.7) with $a = -1$ implies $\mathcal{F}(\check{\psi}) = (\mathcal{F}\psi)\check{\ }$ for any $\psi \in \mathcal{S}(\mathbb{R}^n)$, the above equation means

$$\check{\varphi} = \mathcal{F}((2\pi)^{-n} \check{\check{\widehat{\varphi}}}) = (2\pi)^{-n} \mathcal{F}(\widehat{\varphi}) = (2\pi)^{-n} \widehat{\widehat{\varphi}}.$$

Continuity of $\mathcal{F}$ has been shown in Lemma 5.2.7 above and that of $\mathcal{F}^{-1}$ follows in the same way (the integral formulae are completely analogous). □

5.3 Temperate distributions

5.3.1 Definition. *A* temperate distribution *(also:* tempered *distribution) on $\mathbb{R}^n$ is a continuous linear functional $T\colon \mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$, i.e., $\varphi_k \to 0$ in $\mathcal{S} \Longrightarrow \langle T, \varphi_k \rangle \to 0$ in $\mathbb{C}$. The space of temperate distributions on $\mathbb{R}^n$ is denoted by $\mathcal{S}'(\mathbb{R}^n)$.*

As noted in 5.2.2(iv), for maps defined on the metrizable space $\mathcal{S}(\mathbb{R}^n)$ continuity is equivalent to sequential continuity. Also, as in the cases of $\mathcal{D}'$ and $\mathcal{E}'$ there is an "analytic" characterization of continuity of linear functions on $\mathcal{S}$ in terms of seminorm estimates:

5.3.2 Theorem. *Let $T\colon \mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$ be linear. Then $T \in \mathcal{S}'(\mathbb{R}^n)$ if and only if the following holds: $\exists C > 0\ \exists N \in \mathbb{N}_0$ such that $\forall \varphi \in \mathcal{S}(\mathbb{R}^n)$*

$$|\langle T, \varphi\rangle| \leq C\, Q_N(\varphi) = C \sum_{|\alpha|,|\beta| \leq N} q_{\alpha,\beta}(\varphi) = C \sum_{|\alpha|,|\beta|\leq N} \|x^\alpha D^\beta \varphi\|_\infty. \tag{5.3.1}$$

Proof. Clearly, (5.3.1) and $\varphi_k \to 0$ in $\mathcal{S}$ imply $\langle u, \varphi_k\rangle \to 0$. Conversely, suppose T is continuous but (5.3.1) does not hold: Then $\forall N \in \mathbb{N}\ \exists \varphi_N \in \mathcal{S}(\mathbb{R}^n)$ such that

$$|\langle T, \varphi_N\rangle| > N Q_N(\varphi_N).$$

Then $\varphi_N \neq 0$ and $\psi_N := \varphi_N/(N\, Q_N(\varphi_N))$ ($N \in \mathbb{N}$) defines a sequence in $\mathcal{S}$ with $q_{\alpha,\beta}(\psi_N) \leq 1/N$ when $N \geq \max(|\alpha|, |\beta|)$, but $|\langle T, \psi_N\rangle| \geq 1$, a contradiction. □

5.3.3 Remark. Since $\mathcal{D}(\mathbb{R}^n) \subseteq \mathcal{S}(\mathbb{R}^n)$ with continuous dense embedding (cf. Theorem 5.2.5(ii),(iii)) we have for any $T \in \mathcal{S}'(\mathbb{R}^n)$ that

$$T|_{\mathcal{D}(\mathbb{R}^n)} \in \mathcal{D}'(\mathbb{R}^n)$$

and that the map $T \mapsto T\,|_{\mathcal{D}(\mathbb{R}^n)}$ is injective $\mathcal{S}'(\mathbb{R}^n) \to \mathcal{D}'(\mathbb{R}^n)$. Thus we may consider $\mathcal{S}'(\mathbb{R}^n)$ as a subspace of $\mathcal{D}'(\mathbb{R}^n)$. The latter point of view can serve in alternatively to *define* $\mathcal{S}'$ to consist of those distributions in $\mathcal{D}'$ which can be extended to continuous linear forms on $\mathcal{S} \supseteq \mathcal{D}$ (e.g., see [FJ98, Definition 8.3.1]).

5.3.4 Theorem. *Any $T \in \mathcal{S}'(\mathbb{R}^n)$ is of finite order, i.e., $\mathcal{S}'(\mathbb{R}^n) \subseteq \mathcal{D}'_F(\mathbb{R}^n)$.*

Proof. This is immediate from the seminorm-estimate (5.3.1). □

In particular, we may consider operations defined originally on $\mathcal{D}'$ (differentiation, multiplication etc.) and study under what conditions these leave $\mathcal{S}'$ invariant, thus defining corresponding operations on $\mathcal{S}'$.

5.3.5 Proposition. *Let $T \in \mathcal{S}'(\mathbb{R}^n)$. Then we have*

(i) $\forall \alpha \in \mathbb{N}_0$*:* $\quad \partial^\alpha T \in \mathcal{S}'(\mathbb{R}^n)$

(ii) $\forall f \in \mathcal{O}_M(\mathbb{R}^n)$*:* $\quad f \cdot T \in \mathcal{S}'(\mathbb{R}^n)$

(iii) Let $P(x, D)$ be a partial differential operator with coefficients in $\mathcal{O}_M(\mathbb{R}^n)$, then $P(x, D)\colon \mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$ is linear and sequentially continuous.

Proof. (i) and (ii) are immediate from Theorem 5.2.5(i).

(iii) follows by direct inspection from from (i) and (ii); alternatively, one may use the general property that adjoints of (sequentially) continuous linear maps are weak*-sequentially continuous; cf. Proposition 2.1.2, where this property was shown explicitly for $\mathcal{D}'$ and a transfer to $\mathcal{S}'$ is easy. □

5.3.6 Remark.

(i) We have $\mathcal{E}'(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n) \subset \mathcal{D}'_F(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n)$.

That $\mathcal{E}' \subseteq \mathcal{S}'$ follows from the discussion in Remark 5.3.3, since $\mathcal{E}' \subseteq \mathcal{D}'$ and $\varphi_k \to 0$ in $\mathcal{S}$ implies $\varphi_k \to 0$ in $\mathcal{E}$.

The inclusion $\mathcal{S}' \subseteq \mathcal{D}'_F$ follows from the fact that N occurring in the estimate (5.3.1) is valid with global L^∞-norms.

Each of the above inclusions is indeed strict: For example, $1 \in \mathcal{S}'(\mathbb{R}^n)\setminus\mathcal{E}'(\mathbb{R}^n)$, since $|\langle 1, \varphi\rangle| \le \int |\varphi| \le \int_{\mathbb{R}^n} (1+|x|)^{-(n-1)}dx \cdot Q_{n+1}(\varphi)$, but $\operatorname{supp}(1) = \mathbb{R}^n$ is not compact.

Second, the function $u(x) = e^{x^2}$ defines a regular distribution in $u \in \mathcal{D}'^0(\mathbb{R})$, but u is not defined on all of $\mathcal{S}(\mathbb{R})$, since $\varphi(x) = e^{-x^2}$ yields $\langle u, \varphi\rangle = \int 1dx = \infty$ (alternatively, any approximating sequence $\mathcal{D}(\mathbb{R}) \ni \varphi_j \to \varphi$ in $\mathcal{S}(\mathbb{R})$ yields $(\langle u, \varphi_j\rangle)$ unbounded).

(ii) Recall that for any $1 \le p < \infty$ the vector space $L^p(\mathbb{R}^n)$ is defined analogously to L^1, only changing the integrability condition to $\|f\|_p := (\int_{\mathbb{R}^n} |f(x)|^p\, dx)^{1/p} < \infty$. Furthermore, $L^\infty(\mathbb{R}^n)$ consists of (classes of) essentially bounded L-measurable functions f on $\mathbb{R}^n$ with norm $\|f\|_\infty = \operatorname{ess\,sup}_{x\in\mathbb{R}^n} |f(x)|$ ($= \inf\{M \in [0,\infty[\mid |f| \le M$ almost everywhere$\}$).
$(L^p(\mathbb{R}^n), \|.\|_p)$ is a Banach space for every $1 \le p \le \infty$.

We have
$$L^p(\mathbb{R}^n) \subseteq \mathcal{S}'(\mathbb{R}^n).$$

Proof: By Hölder's inequality ([Fol99, 6.2 and Theorem 6.8.a]), if $f \in L^p(\mathbb{R}^n)$ and $\varphi \in \mathcal{S}(\mathbb{R}^n)$, then
$$|\langle f, \varphi\rangle| \le \int |f\varphi| \le \|f\|_p \|\varphi\|_q \qquad (\frac{1}{p} + \frac{1}{q} = 1).$$

The standard $\mathcal{S}$-estimate $|\varphi(x)| \le Q_l(\varphi)/(1+|x|)^l$, valid for every $l \in \mathbb{N}_0$, gives
$$\|\varphi\|_q^q = \int |\varphi(x)|^q\, dx \le Q_l(\varphi)^q \int \frac{dx}{(1+|x|)^{lq}}$$
and thus shows continuity of $\varphi \mapsto \langle f, \varphi\rangle$ upon choosing l sufficiently large to ensure $lq > n$. □

(iii) Let $f \in C(\mathbb{R}^n)$ be of *polynomial growth*, i.e., $\exists C, M \ge 0$:
$$|f(x)| \le C(1+|x|)^M \qquad \forall x \in \mathbb{R}^n.$$

Then $f \in \mathcal{S}'(\mathbb{R}^n)$, since we for any $\varphi \in \mathcal{S}(\mathbb{R}^n)$
$$\begin{aligned}|\langle f, \varphi\rangle| &\le \int |f(x)||\varphi(x)| \le \int C(1+|x|)^M \frac{Q_l(\varphi)}{(1+|x|)^l}\, dx \\ &= CQ_l(\varphi) \underbrace{\int \frac{dx}{(1+|x|)^{l-M}}}_{=:C_l},\end{aligned}$$

where C_l is finite, if $l > M + n$.

Note that we automatically obtain that also $\partial^\alpha f \in \mathcal{S}'(\mathbb{R}^n)$ due to Proposition 5.3.5(i).

5.3.7 Definition. *Let (T_j) be a sequence in $\mathcal{S}'(\mathbb{R}^n)$ and $T \in \mathcal{S}'(\mathbb{R}^n)$. We say that (T_j)* converges *to T in $\mathcal{S}'(\mathbb{R}^n)$, denoted $T_j \to T$ $(j \to \infty)$, if $\forall \varphi \in \mathcal{S}(\mathbb{R}^n)$: $\langle T_j, \varphi\rangle \to \langle T, \varphi\rangle$ $(j \to \infty)$. (Similarly for nets $(T_\varepsilon)_{\varepsilon\in\,]0,1]}$ etc.)*

5.3.8 Theorem. *$\mathcal{S}'(\mathbb{R}^n)$ is sequentially complete.*

Proof. Using the norms $\varphi \mapsto q_{\alpha,\beta}(\varphi)$ instead of $\varphi \mapsto \|\varphi\|_\infty$, this follows exactly as in the proof of Theorem 1.1.8. □

5.3.9 Theorem. *$\mathcal{D}(\mathbb{R}^n)$ is sequentially dense in $\mathcal{S}'(\mathbb{R}^n)$.*

Proof. Let $\omega \in \mathcal{D}(\mathbb{R}^n)$, $\omega \equiv 1$ on $B_1(0)$, and let $\rho \in \mathcal{D}(\mathbb{R}^n)$, $\int \rho = 1$, $\operatorname{supp}(\rho) \subseteq B_1(0)$ and $\rho = \check{\rho}$ for all x. Given any $T \in S'(\mathbb{R}^n)$ and $k \in \mathbb{N}$, $\omega_k \cdot (T * \rho_{1/k}) \in \mathcal{D}(\mathbb{R}^n)$, where $\omega_k(x) := \omega(x/k)$ and $\rho_{1/k}(x) = k^n \rho(kx)$. Therefore it suffices to show that $\omega_k \cdot (T * \rho_{1/k}) \to T$ in $\mathcal{S}'(\mathbb{R}^n)$, i.e., that for any $\psi \in \mathcal{S}(\mathbb{R}^n)$ we have

$$\langle \omega_k \cdot (T * \rho_{1/k}) - T, \psi \rangle = \langle T, \rho_{1/k} * (\omega_k \psi) - \psi \rangle \to 0 \quad (k \to \infty).$$

So we have reduced the proof to showing that $\rho_{1/k} * (\omega_k \psi) - \psi \to 0$ in $\mathcal{S}(\mathbb{R}^n)$. Now

$$\begin{aligned}
x^\alpha \rho_{1/k} * \partial^\beta(\omega_k \psi)(x) - x^\alpha \partial^\beta \psi(x) &= x^\alpha \int [\partial^\beta(\omega_k \psi)\Big(x - \frac{y}{k}\Big) - \partial^\beta \psi(x)]\rho(y)\, dy \\
= x^\alpha \int [\partial^\beta(\omega_k \psi)\Big(x - \frac{y}{k}\Big) &- \partial^\beta(\omega_k \psi)(x)]\rho(y)\, dy \\
&+ x^\alpha \int [\partial^\beta(\omega_k \psi)(x) - \partial^\beta \psi(x)]\rho(y)\, dy =: A + B.
\end{aligned}$$

Here,

$$\begin{aligned}
A = x^\alpha \int \int_0^1 \frac{d}{d\sigma}\Big(\partial^\beta(\omega_k \psi)\Big(x - \frac{\sigma y}{k}\Big)\Big)\, d\sigma \rho(y)\, dy \\
= -x^\alpha \int \int_0^1 \sum_{j=1}^n \partial_j \partial^\beta(\omega_k \psi)\Big(x - \frac{\sigma y}{k}\Big) \frac{y_j}{k}\, d\sigma \rho(y)\, dy,
\end{aligned}$$

so since $|y| \le 1$ in the integrand,

$$|A| \le \frac{C}{k} \sup_{|\gamma| \le |\beta|+1} \sup_{|z| \le 1} |x^\alpha \partial^\gamma \psi(x - z)| \tag{5.3.2}$$

(with C a generic constant, here and below). Moreover,

$$\begin{aligned}
|x^\alpha \partial^\gamma \psi(x - z)| &\le |x|^{|\alpha|} \frac{1}{(1 + |x - z|)^{|\alpha|}} \underbrace{((1 + |x - z|)^{|\alpha|})|\partial^\gamma \psi(x - z)|}_{\le C} \\
&\le C \frac{(1 + |x - z| + |z|)^{|\alpha|}}{(1 + |x - z|)^{|\alpha|}} = C\Big(1 + \underbrace{\frac{|z|}{1 + |x - z|}}_{\le |z| \le 1}\Big)^{|\alpha|} \le C 2^{|\alpha|},
\end{aligned}$$

so $|A| \le \frac{C}{k} \to 0$ $(k \to \infty)$. Finally, $\partial^\beta(\omega_k \psi)(x) - \partial^\beta \psi(x) = 0$ on $B_k(0)$, so also

$$|B| \le \sup_{|x| > k} C |x^\alpha \partial^\beta \psi| \le C \sup_{|x| > k} \frac{1}{x_1} q_{\alpha + e_1, \beta}(\psi) \to 0 \quad (k \to \infty).$$

□

5.3.10 Corollary. *$L^p(\mathbb{R}^n)$ is dense in $\mathcal{S}'(\mathbb{R}^n)$ for all $1 \le p \le \infty$*

Proof. This is clear since $\mathcal{D}(\mathbb{R}^n) \subseteq L^p(\mathbb{R}^n)$ $(1 \le p \le \infty)$. □

To conclude this section we mention the following global structure theorem for temperate distributions (see [FJ98, Th. 8.3.1] for a proof similar to that of Theorem 4.5.2):

5.3.11 Theorem. *Any temperate distribution is a derivative of some polynomially bounded continuous function.*

5.4 Fourier transform of temperate distributions

We have now collected sufficient information to construct an appropriate extension of Fourier transformation to distribution theory. If $u \in L^1(\mathbb{R}^n)$ we may consider it as an element of $\mathcal{S}'(\mathbb{R}^n)$ (by Remark 5.3.6(ii)) and Theorem 5.1.1 (i) gives $\widehat{u} \in C_{\mathrm{b}}(\mathbb{R}^n) \subseteq L^\infty(\mathbb{R}^n)$, so $\widehat{u} \in \mathcal{S}'(\mathbb{R}^n)$ by 5.3.6(ii) again.

Thus, since for any $\varphi \in \mathcal{S}(\mathbb{R}^n) \subseteq L^1(\mathbb{R}^n)$ also $\widehat{\varphi} \in \mathcal{S}(\mathbb{R}^n) \subseteq L^1(\mathbb{R}^n)$, we obtain

$$\langle \widehat{u}, \varphi \rangle = \int \widehat{u}(\xi)\varphi(\xi)\, d\xi \underset{\substack{\uparrow \\ [5.1.1(ii)]}}{=} \int u(x)\widehat{\varphi}(x)\, dx = \langle u, \widehat{\varphi} \rangle. \tag{5.4.1}$$

Observe that the right-most term can be extended to the general case $T \in \mathcal{S}'(\mathbb{R}^n)$: Since $\varphi \mapsto \widehat{\varphi}$ is a (continuous) isomorphism on $\mathcal{S}(\mathbb{R}^n)$, the map $\varphi \mapsto \langle T, \widehat{\varphi} \rangle$ defines an element in $\mathcal{S}'(\mathbb{R}^n)$.

5.4.1 Definition. *If $T \in \mathcal{S}'(\mathbb{R}^n)$, then the* Fourier transform *$\widehat{T}$, or $\mathcal{F}T$, is defined by*

$$\langle \widehat{T}, \varphi \rangle := \langle T, \widehat{\varphi} \rangle \qquad \forall \varphi \in \mathcal{S}(\mathbb{R}^n). \tag{5.4.2}$$

5.4.2 Theorem. *The Fourier transform $\mathcal{F}\colon \mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$ is linear and bijective, $\mathcal{F}$ as well as $\mathcal{F}^{-1}$ are sequentially continuous. We have again the formulae*

$$\widehat{\widehat{T}} = (2\pi)^n\, \check{T} \tag{5.4.3}$$

(where $\langle \check{T}, \varphi \rangle := \langle T, \check{\varphi} \rangle$) as well as $\mathcal{F}^{-1}T = (2\pi)^{-n}(\widehat{T})\check{}$.

Moreover, if $u \in L^1(\mathbb{R}^n)$, then $\widehat{u}$ according to (5.4.2) *coincides with its classical Fourier transform (as L^1-function).*

Proof. Compatibility of the distributional with the classical Fourier transform on functions $u \in L^1(\mathbb{R}^n)$ follows from (5.4.1).

Speaking in abstract terms, it follows that $\mathcal{F}$ is an isomorphism on $\mathcal{S}'(\mathbb{R}^n)$ since it is the adjoint of an isomorphism on $\mathcal{S}(\mathbb{R}^n)$. The latter statement includes also the weak*-continuity of $\mathcal{F}$ and $\mathcal{F}^{-1}$. However, we give an independent proof for our case here.

Linearity of $\mathcal{F}$ is clear, as is the sequential continuity of $\mathcal{F}$ from (5.4.2).

To prove injectivity, assume that $\mathcal{F}T = 0$. Then we have for every $\varphi \in \mathcal{S}(\mathbb{R}^n)$ that $0 = \langle \mathcal{F}T, \mathcal{F}^{-1}\varphi \rangle = \langle T, \varphi \rangle$, hence $T = 0$.

To show surjectivity we will first derive (5.4.3). Let $T \in \mathcal{S}'(\mathbb{R}^n)$ and $\varphi \in \mathcal{S}(\mathbb{R}^n)$, then

$$\langle \widehat{\widehat{T}}, \varphi \rangle = \langle \widehat{T}, \widehat{\varphi} \rangle = \langle T, \widehat{\widehat{\varphi}} \rangle \underset{\substack{\uparrow \\ [(5.2.11)]}}{=} (2\pi)^n \langle T, \check{\varphi} \rangle = (2\pi)^n \langle \check{T}, \varphi \rangle, \tag{5.4.4}$$

which (since $\check{}$ and $\mathcal{F}$ commute on $\mathcal{S}$, hence also on $\mathcal{S}'$), replacing T by $\check{T}$ in (5.4.4) implies $T = \mathcal{F}((2\pi)^{-n}\widehat{T}\check{})$. In particular, this yields surjectivity of $\mathcal{F}$ and the stated formula for the inverse. □

Next we list a number of properties of the Fourier transform on $\mathcal{S}'$, which follow directly from the corresponding formulae on $\mathcal{S}$ and the definition as adjoint of the Fourier transform on $\mathcal{S}$.

5.4.3 Proposition. *For any $S, T \in \mathcal{S}'(\mathbb{R}^n)$ we have*

(i) $\forall \alpha \in \mathbb{N}_0^n$: $(D^\alpha T)\widehat{} = \xi^\alpha\, \widehat{T}$, $D^\alpha(\mathcal{F}^{-1}S)(\xi) = \mathcal{F}^{-1}(\xi^\alpha S)$

(ii) $\forall \alpha \in \mathbb{N}_0^n$: $(x^\alpha T)\widehat{\ } = (-1)^{|\alpha|} D^\alpha \widehat{T}$, $x^\alpha \mathcal{F}^{-1}(S) = (-1)^{|\alpha|} \mathcal{F}^{-1}(D^\alpha S)$.

(iii) $\forall h \in \mathbb{R}^n$: $(\tau_h T)\widehat{\ } = e^{-i\xi h}\,\widehat{T}$,

(iv) $\forall h \in \mathbb{R}^n$: $(e^{ixh}T)\widehat{\ } = \tau_h \widehat{T}$,

(v) $(\check{T})\widehat{\ } = (\widehat{T})\check{\ }$.

Proof. Applying the definition of the action of T on any $\varphi \in \mathcal{S}(\mathbb{R}^n)$ we obtain (i) and (ii) from Lemma 5.2.6 (the formulas for S follow by setting $S := \mathcal{F}(T)$), (iii) and (iv) from (5.2.10) and a similar direct calculation showing

$$(e^{ixh}\varphi)\widehat{\ } = \tau_h \widehat{\varphi}$$

(alternatively, use (5.2.11)), and (v) from (5.2.7) with $a = -1$. Finally, we note that all identities claimed above also follow by continuous extension from the fact that they hold on the dense subspace $\mathcal{S}(\mathbb{R}^n)$ of $\mathcal{S}'(\mathbb{R}^n)$. □

5.4.4 Examples. (i) We directly calculate for arbitrary $\varphi \in \mathcal{S}(\mathbb{R}^n)$

$$\langle \widehat{\delta}, \varphi \rangle = \langle \delta, \widehat{\varphi} \rangle = \widehat{\varphi}(0) = \int \underbrace{e^{-i0x}}_{=1} \varphi(x)\, dx = \langle 1, \varphi \rangle,$$

therefore we obtain

$$\widehat{\delta} = 1. \tag{5.4.5}$$

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

$$\langle \delta, \varphi \rangle = \langle \delta, \check{\varphi} \rangle = (2\pi)^{-n} \langle \widehat{\delta}, \widehat{\varphi} \rangle = (2\pi)^{-n} \langle 1, \widehat{\varphi} \rangle = \langle (2\pi)^{-n}\, \widehat{1}, \varphi \rangle,$$

hence

$$\widehat{1} = (2\pi)^n\, \delta. \tag{5.4.6}$$

(ii) Let $e_h(x) := e^{ixh}$ $(h \in \mathbb{R}^n)$, then $e_h \in \mathcal{O}_M(\mathbb{R}^n) \subseteq \mathcal{S}'(\mathbb{R}^n)$ and

$$\mathcal{F}e_h = \mathcal{F}(e^{ihx} \cdot 1) = \tau_h \widehat{1} = (2\pi)^n\, \tau_h \delta = (2\pi)^n\, \delta_h.$$

(iii) The Fourier transform of the Heaviside function $Y \in L^\infty(\mathbb{R}) \subseteq \mathcal{S}'(\mathbb{R})$:
Since $Y' = \delta$ we have $i\xi\, \widehat{Y} = \widehat{\delta} = 1$ and hence $\widehat{Y}(\xi) = -i/\xi$ when $\xi \neq 0$.
Note that on $\mathbb{R} \setminus \{0\}$ the function $\xi \mapsto 1/\xi$ coincides, as a distribution, with the principal value $\mathrm{vp}(1/\xi)$. Furthermore, since by Example 1.2.7 $\xi \cdot \mathrm{vp}(1/\xi) = 1$,

$$\xi \left(\widehat{Y}(\xi) + i\mathrm{vp}(1/\xi)\right) = 0.$$

By Example 3.4.3, $\xi T(\xi) = 0$ implies $T = c\delta$ with a complex constant c. Hence it remains to determine the constant c in the equation $\widehat{Y} + i\,\mathrm{vp}(1/\xi) = c\delta$.
Note that $\check{\delta} = \delta$, $\check{Y} = 1 - Y$, and $\mathrm{vp}(1/\xi)\check{\ } = -\mathrm{vp}(1/\xi)$ and recall that Fourier transform commutes with reflection $\check{\ }$. Thus we calculate

$$\begin{aligned} c\delta = c\check{\delta} = \widehat{\check{Y}} + i\,\mathrm{vp}(1/\xi)\check{\ } = (1-Y)\widehat{\ } - i\,\mathrm{vp}(1/\xi) = \widehat{1} - \overbrace{(\widehat{Y} + i\,\mathrm{vp}(1/\xi))}^{=c\delta} \\ = 2\pi\,\delta - c\delta = (2\pi - c)\delta, \end{aligned}$$

hence $c = \pi$ and we arrive at

$$\widehat{Y} = \pi\,\delta - i\,\mathrm{vp}(\frac{1}{\xi}) \tag{5.4.7}$$

(iv) Since $\widehat{\widehat{Y}} = (2\pi)\check{Y} = (2\pi)(1 - Y(x))$ we may use the result of (iii) to deduce a formula for $\mathcal{F}(\mathrm{vp}(1/x))$ as follows

$$i\,\mathrm{vp}(1/x)\widehat{\ } = \pi\,\widehat{\delta} - \widehat{\widehat{Y}} = \pi\,1 - 2\pi\,\check{Y} = \pi - 2\pi(1 - Y) = \pi\,\underbrace{(2Y - 1)}_{=\mathrm{sgn}(\xi)}.$$

Therefore we obtain

$$\mathrm{vp}(1/x)\widehat{\ }(\xi) = -i\pi\,\mathrm{sgn}(\xi).$$

5.4.5 Remark. Let $P(D)$ be a linear partial differential operator with constant coefficients, i.e.

$$P(D) = \sum_{|\alpha|\leq m} c_\alpha\, D^\alpha \qquad (c_\alpha \in \mathbb{C}).$$

If $T, f \in \mathcal{S}'(\mathbb{R}^n)$, then the exchange formulae from Proposition 5.4.3 (i),(ii) applied to each term in $P(D)$ give

$$P(D)T = f \quad \Longleftrightarrow \quad P(\xi)\,\widehat{T} = \widehat{f}.$$

Thus, the action of $P(D)$ is translated into multiplication with the polynomial $P(\xi)$. In certain cases, this trick allows (a more or less) explicit representation of solutions. Moreover, the above equivalence provides important additional information in theoretical investigations regarding regularity and solvability questions, see Chapter 6.

5.5 Fourier transform on $\mathcal{E}'$ and the convolution theorem

Recall from Theorem 5.1.1 (iii) and (5.1.4) that we have for any $f, g \in L^1(\mathbb{R}^n)$

$$\widehat{(f * g)} = \widehat{f} \cdot \widehat{g},$$

where the product on the right-hand side means the usual (pointwise) multiplication of continuous functions. In the current section we will prove the analogous result for the convolution product, if $f \in \mathcal{S}'$ and $g \in \mathcal{E}'$. Then $\widehat{f} \in \mathcal{S}'$ and we have to clarify the meaning of $\widehat{f} \cdot \widehat{g}$ in a preparatory result on the Fourier transforms of distributions in $\mathcal{E}'$.

5.5.1 Theorem. *If $T \in \mathcal{E}'(\mathbb{R}^n)$, then $\widehat{T} \in \mathcal{O}_M(\mathbb{R}^n) \subseteq \mathcal{C}^\infty(\mathbb{R}^n)$ and we have*

$$\widehat{T}(\xi) = \langle T(x), e^{-ix\xi}\rangle \qquad \forall \xi \in \mathbb{R}^n. \tag{5.5.1}$$

Moreover, $\widehat{T}$ can be extended to an entire (holomorphic) function on $\mathbb{C}^n$.

Proof. Smoothness of the function $h\colon \xi \mapsto \langle T(x), e^{-ix\xi}\rangle$ follows from Corollary 3.6.6 (ii), and in particular any derivative $D^\alpha h$ is given by

$$D^\alpha h(\xi) = \langle T(x), D^\alpha_\xi(e^{-ix\xi})\rangle = \langle T(x), (-x)^\alpha e^{-ix\xi}\rangle.$$

The $\mathcal{O}_M$-estimates for h follow directly from the seminorm estimate (2.3.1) for T, which provide a compact neighborhood K of $\mathrm{supp}\,(T)$, a constant $C > 0$, and an order of derivative $N \in \mathbb{N}_0$ such that

$$|D^\alpha h(\xi)| = |\langle T(x), (-x)^\alpha e^{-ix\xi}\rangle| \leq C \sum_{|\beta|\leq N} \sup_{x\in K} \underbrace{|\partial^\beta_x (x^\alpha e^{-ix\xi})|}_{\leq C_{\beta,K}|\xi|^{|\beta|}} \leq C'(1 + |\xi|)^N,$$

where $C_{\beta,K}$ and C' denote appropriate constants.
Let $\varphi \in \mathcal{D}(\mathbb{R}^n)$ arbitrary, then $T \otimes \varphi \in \mathcal{E}'(\mathbb{R}^n \times \mathbb{R}^n)$ and

$$\langle T \otimes \varphi(x,\xi), e^{-ix\xi} \rangle = \langle T(x), \underbrace{\langle \varphi(\xi), e^{-ix\xi} \rangle}_{=\widehat{\varphi}(x)} \rangle = \langle T, \widehat{\varphi} \rangle = \langle \widehat{T}, \varphi \rangle.$$

On the other hand,

$$\langle T \otimes \varphi(x,\xi), e^{-ix\xi} \rangle = \langle \varphi(\xi), \underbrace{\langle T(x), e^{-ix\xi} \rangle}_{=h(\xi)} \rangle = \int_{\mathbb{R}^n} h(\xi)\varphi(\xi)\, d\xi,$$

hence $\widehat{T} = h$ and therefore (5.5.1) holds.
Finally, again by Corollary 3.6.6(ii) we obtain smoothness of the function

$$\mathbb{R}^n + i\mathbb{R}^n = \mathbb{C}^n \ni \zeta = \xi + i\eta \;\mapsto\; \langle T(x), e^{-ix\zeta} \rangle = \langle T(x), e^{-ix\xi + x\eta} \rangle \in \mathbb{C}$$

and, since $2\partial_{\overline{\zeta_j}}(e^{-ix\xi+x\eta}) := (\partial_{\xi_j} + i\partial_{\eta_j})(e^{-ix\xi+x\eta}) = 0$, we also obtain

$$\partial_{\overline{\zeta_j}} \widehat{T}(\zeta) = \langle T(x), \partial_{\overline{\zeta_j}}(e^{-ix\xi+x\eta}) \rangle = 0.$$

Thus the Cauchy–Riemann equations are satisfied in each complex variable, which means holomorphicity of $\widehat{T}$ as a function on $\mathbb{C}^n$ (cf. [For84, p. 261]). □

5.5.2 Lemma. *If $S, T \in \mathcal{E}'(\mathbb{R}^n)$, then $\widehat{(S * T)} = \widehat{S} \cdot \widehat{T}$.*

Proof. Since $\operatorname{supp}(S*T) \subseteq \operatorname{supp}(S) + \operatorname{supp}(T)$ we have $S*T \in \mathcal{E}'(\mathbb{R}^n)$. Therefore (5.5.1) gives

$$\begin{aligned}\widehat{(S*T)}(\xi) = \langle S*T(z), e^{-iz\xi} \rangle &= \langle S \otimes T(x,y), e^{-i\xi(x+y)} \rangle \\ &= \langle S(x), e^{-ix\xi} \langle T(y), e^{-iy\xi} \rangle \rangle = \widehat{S}(\xi)\, \widehat{T}(\xi).\end{aligned}$$

□

5.5.3 Theorem. (Convolution Theorem) *Let $S \in \mathcal{S}'(\mathbb{R}^n)$ and $T \in \mathcal{E}'(\mathbb{R}^n)$. Then $S * T$ belongs to $\mathcal{S}'(\mathbb{R}^n)$ and we have*

$$\widehat{(S*T)} = \widehat{S} \cdot \widehat{T}. \tag{5.5.2}$$

Proof. Theorem 5.5.1 implies that $\widehat{T} \in \mathcal{O}_M(\mathbb{R}^n)$, hence $\widehat{T}\,\widehat{S} \in \mathcal{S}'(\mathbb{R}^n)$ by Proposition 5.3.5 (ii). Since the Fourier transform is an isomorphism on $\mathcal{S}'(\mathbb{R}^n)$,

$$\exists!\, R \in \mathcal{S}'(\mathbb{R}^n): \quad \widehat{R} = \widehat{T}\,\widehat{S}.$$

We determine R by its action on any test function $\varphi \in \mathcal{D}(\mathbb{R}^n)$, upon noting that $\check{\varphi} = (2\pi)^{-n}\,\widehat{\widehat{\varphi}}$, as follows:

$$\begin{aligned}\langle R, \check{\varphi} \rangle &= (2\pi)^{-n} \langle R, \widehat{\widehat{\varphi}} \rangle = (2\pi)^{-n} \langle \widehat{R}, \widehat{\varphi} \rangle = (2\pi)^{-n} \langle \widehat{T}\,\widehat{S}, \widehat{\varphi} \rangle \\ &= (2\pi)^{-n} \langle \widehat{S}, \widehat{T}\,\widehat{\varphi} \rangle \underset{\substack{\uparrow \\ [5.5.2]}}{=} (2\pi)^{-n} \langle \widehat{S}, \widehat{T*\varphi} \rangle = (2\pi)^{-n} \langle \widehat{\widehat{S}}, T*\varphi \rangle \\ &= \langle \check{S}, T*\varphi \rangle \underset{\substack{\uparrow \\ (4.3.1)}}{=} S*(T*\varphi)(0) = (S*T)*\varphi(0) \underset{\substack{\uparrow \\ (4.3.1)}}{=} \langle S*T, \check{\varphi} \rangle.\end{aligned}$$

Since $\mathcal{D} \subseteq \mathcal{S}$ is dense we obtain $R = S * T$. □

Chapter 6

Fourier Analysis and Partial Differential Equations

In this final chapter we consider further applications of distribution theory to linear partial differential equations. The key tools we are going to employ are the space of temperate distributions and the Fourier transform.

6.1 The Malgrange–Ehrenpreis Theorem

In Section 3.5 we derived fundamental solutions of several important linear PDEs. A key question determining the usefulness of distribution theory as a tool for solving such equations is whether *any* linear partial differential equation with constant coefficients possesses a fundamental solution in the space of distributions. It was one of the great early breakthroughs of the theory when in 1954/55, B. Malgrange and L. Ehrenpreis independently were able to answer this question affirmatively. Both proofs were non-constructive, relying on the Hahn–Banach theorem. Subsequently, a number of authors were concerned with finding explicit general formulae for such fundamental solutions. For a brief historical account and many references we refer to [OW15]. Below we present the most elegant proof of an explicit representation known to date, due to Peter Wagner, [Wag09].

6.1.1 Theorem. *(Malgrange–Ehrenpreis) Let $P(\xi) = \sum_{|\alpha|\le m} c_\alpha \xi^\alpha$ be a non-identically vanishing polynomial with complex coefficients of degree m. Let $P_m(\xi) := \sum_{|\alpha|=m} c_\alpha \xi^\alpha$ and pick $\eta \in \mathbb{R}^n$ with $P_m(\eta) \neq 0$. Let $\lambda_0, \dots, \lambda_m$ be pairwise different and set*

$$a_j := \prod_{k=0, k\neq j}^{m} (\lambda_j - \lambda_k)^{-1}.$$

Then

$$E = \frac{1}{\overline{P_m(2\eta)}} \sum_{j=0}^{m} a_j e^{\lambda_j \eta x} \mathcal{F}_\xi^{-1}\left(\frac{\overline{P(i\xi + \lambda_j \eta)}}{P(i\xi + \lambda_j \eta)}\right) \tag{6.1.1}$$

is a fundamental solution of $P(\partial)$, i.e., $P(\partial)E = \delta$.

Proof. We first note that, for any fixed $\lambda \in \mathbb{R}$, the set $N := \{\xi \in \mathbb{R}^n \mid P(i\xi + \lambda\eta) = 0\}$ has Lebesgue measure zero. Indeed, after a linear change of variables we may assume that $P_m(1, 0, \dots, 0) \neq 0$. Then any set $N_{\xi'} := \{\xi_1 \in \mathbb{R} \mid P(i(\xi_1, \xi') + \lambda\eta) =$

$0\}$ is finite, and so by Fubini's theorem we have

$$\lambda(N) = \int_N d\xi = \int_{\mathbb{R}^{n-1}} \int_{N_{\xi'}} d\xi_1 \, d\xi' = 0.$$

Consequently,

$$S(\xi) := \frac{\overline{P(i\xi + \lambda\eta)}}{P(i\xi + \lambda\eta)} \in L^\infty(\mathbb{R}^n) \subseteq \mathcal{S}'(\mathbb{R}^n),$$

showing that (6.1.1) gives a well-defined distribution.
Next, observe that for any $T \in \mathcal{D}'(\mathbb{R}^n)$ and any $l \in \{1, \dots, n\}$ we have

$$\partial_l(e^{\zeta x} T) = e^{\zeta x}(\partial_l + \zeta_l)T,$$

and iterating this we get $\partial^\alpha(e^{\zeta x}T) = e^{\zeta x}(\partial + \zeta)^\alpha T$ for any α, so

$$P(\partial)(e^{\zeta x}T) = e^{\zeta x} P(\partial + \zeta)T. \tag{6.1.2}$$

Using (6.1.2), together with Proposition 5.4.3, for any $S \in \mathcal{S}'(\mathbb{R}^n)$ and any $\zeta \in \mathbb{C}^n$ we get

$$P(\partial)(e^{\zeta x}\mathcal{F}^{-1}S) = e^{\zeta x}P(\partial + \zeta)\mathcal{F}^{-1}S = e^{\zeta x}\mathcal{F}_\xi^{-1}(P(i\xi + \zeta)S)$$

Inserting $S = \frac{\overline{P(i\xi+\lambda\eta)}}{P(i\xi+\lambda\eta)}$ and $\zeta := \lambda\eta$ for any $\lambda \in \mathbb{R}$, this gives

$$P(\partial)\left(e^{\lambda\eta x}\mathcal{F}^{-1}\left(\frac{\overline{P(i\xi + \lambda\eta)}}{P(i\xi + \lambda\eta)}\right)\right) = e^{\lambda\eta x}\mathcal{F}_\xi^{-1}\big(\overline{P(i\xi + \lambda\eta)}\big).$$

Again by Proposition 5.4.3,

$$\begin{aligned}\mathcal{F}_\xi^{-1}\big(\overline{P(i\xi + \lambda\eta)}\big) &= \mathcal{F}_\xi^{-1}\big(\overline{P}(-i\xi + \lambda\eta)\big) = \mathcal{F}_\xi^{-1}\big(\overline{P}(-i\xi + \lambda\eta) \cdot 1\big) \\ &= \overline{P}(-\partial + \lambda\eta)\mathcal{F}^{-1}(1) = \overline{P}(-\partial + \lambda\eta)\delta.\end{aligned}$$

Also, analogous to (6.1.2), we have

$$(-\partial_j + 2\lambda\eta_j)(e^{\lambda\eta x}T) = -\lambda\eta_j e^{\lambda\eta x}T - e^{\lambda\eta x}\partial_j T + 2\lambda\eta_j e^{\lambda\eta x}T = e^{\lambda\eta x}(-\partial_j + \lambda\eta_j)T,$$

and iterating this gives $\overline{P}(-\partial + 2\lambda\eta)(e^{\lambda\eta x}T) = e^{\lambda\eta x}\overline{P}(-\partial + \lambda\eta)T$. Putting all of the above together, we arrive at

$$\begin{aligned}P(\partial)\left(e^{\lambda\eta x}\mathcal{F}^{-1}\left(\frac{\overline{P(i\xi + \lambda\eta)}}{P(i\xi + \lambda\eta)}\right)\right) &= e^{\lambda\eta x}\overline{P}(-\partial + \lambda\eta)\delta = \overline{P}(-\partial + 2\lambda\eta)(\underbrace{e^{\lambda\eta x}\delta}_{=\delta}) \\ &= \overline{P}(-\partial + 2\lambda\eta)\delta = \lambda^m\overline{P_m(2\eta)}\delta + \sum_{k=0}^{m-1}\lambda^k T_k,\end{aligned}$$

for certain distributions $T_k \in \mathcal{E}'(\mathbb{R}^n)$. Therefore,

$$P(\partial)E = \frac{1}{\overline{P_m(2\eta)}}\sum_{j=0}^{m} a_j\big[\lambda_j^m\overline{P_m(2\eta)}\delta + \sum_{k=0}^{m-1}\lambda_j^k T_k\big]$$

Finally, as we know from the Vandermonde-calculation in the proof of Theorem 3.2.6, the a_j are chosen such that they satisfy the system of linear equations

$$\sum_{j=0}^{m} a_j\lambda_j^k = \begin{cases} 0 & k = 0, \dots, m-1 \\ 1 & k = m \end{cases},$$

so indeed $P(\partial)E = \delta$, as claimed. $\square$

6.2 The elliptic regularity theorem

Recall from Definition 3.5.1 that the singular support of a distribution $T \in \mathcal{D}'(\Omega)$ is the complement of the largest open set $U \subseteq \Omega$ such that $T|_U \in C^\infty(U)$.

6.2.1 Lemma. *Let $S \in \mathcal{E}'(\mathbb{R}^n)$ and $T \in \mathcal{D}'(\mathbb{R}^n)$. Then*

$$\operatorname{singsupp} S * T \subseteq \operatorname{singsupp} S + \operatorname{singsupp} T. \tag{6.2.1}$$

Proof. Choose $\rho, \psi \in C^\infty(\mathbb{R}^n)$ such that $\rho \equiv 1$ in a neighborhood of $\operatorname{singsupp} S$ and $\psi \equiv 1$ in a neighborhood of $\operatorname{singsupp} T$. Then

$$\begin{aligned} S*T &= (\rho S + (1-\rho)S) * (\psi T + (1-\psi)T) \\ &= (\rho S) * (\psi T) + (\rho S) * ((1-\psi)T) + ((1-\rho)S) * T + ((1-\rho)S) * ((1-\psi)T). \end{aligned}$$

Here, the last three terms are all smooth functions, so

$$\operatorname{singsupp} S * T \subseteq \operatorname{singsupp} (\rho S) * (\psi T) \subseteq \operatorname{supp} (\rho S) * (\psi T) \subseteq \operatorname{supp} \rho + \operatorname{supp} \psi.$$

Since this holds for all ρ and ψ as above, the claim follows. □

6.2.2 Definition. *Let $P(D) = \sum_{|\alpha| \le m} a_\alpha D^\alpha$ be a linear partial differential operator with constant coefficients. Then $P(\xi) := \sum_{|\alpha| \le m} a_\alpha \xi^\alpha$ is called the* symbol *of P and the homogeneous polynomial $\sigma_P(\xi) := \sum_{|\alpha|=m} a_\alpha \xi^\alpha := P_m(\xi)$ is called the* principal symbol *of P. The PDO $P_m(D)$ is called the* principal part *of P. P is called* elliptic *if $\sigma_P(\xi) \neq 0$ for each $\xi \neq 0$.*

6.2.3 Example. Both the Laplace operator Δ and the Cauchy–Riemann operator $\partial_1 + i\partial_2$ are elliptic.

The following important result states that solutions T to elliptic equations $P(\partial)T = S$ are as regular as the right hand side S:

6.2.4 Theorem. *Let P be an elliptic PDO with constant coefficients. Then for all $T \in \mathcal{D}'(\Omega)$ we have*

$$\operatorname{singsupp} T = \operatorname{singsupp} P(T). \tag{6.2.2}$$

In other words, any elliptic operator is hypoelliptic.

Proof. Since P is elliptic, $c := \min_{\xi \in S^{n-1}} |\sigma_P(\xi)| > 0$. Then for all $\xi \in \mathbb{R}^n$ we have $|\sigma_P(\xi)| \ge c|\xi|^m$, and there are constants $c_1, \dots, c_m$ such that

$$|P(\xi)| \ge c|\xi|^m - c_1|\xi|^{m-1} - \dots - c_m.$$

Consequently, for each $a > 0$ there exists some $t > 0$ such that $|P(\xi)| \ge a$ for all $|\xi| \ge t$.
Pick $\chi \in \mathcal{D}(\mathbb{R}^n)$ such that $0 \le \chi \le 1$ and $\chi \equiv 1$ on $|\xi| \le t$. Then the function $\frac{1-\chi}{P}$ is globally bounded by $\frac{1}{a}$, hence is an element of $L^\infty(\mathbb{R}^n) \subseteq \mathcal{S}'(\mathbb{R}^n)$. Fixing such a and t, set $E := \mathcal{F}^{-1}(\frac{1-\chi}{P})$, then

$$\hat{E}(\xi) = \frac{1-\chi(\xi)}{P(\xi)} \tag{6.2.3}$$

and we claim that $E|_{\mathbb{R}^n \setminus \{0\}}$ is smooth. To see this, note first that for all α, β

$$(x^\beta D^\alpha E)\hat{} = (-1)^{|\beta|} D^\beta (\xi^\alpha \hat{E}).$$

Now for $|\xi|$ large we have $|D^\beta(\xi^\alpha \hat{E})| \le \left|\frac{Q(\xi)}{P^{2^{|\beta|}}}\right|$, where Q is a polynomial of degree $m2^{|\beta|} + |\alpha| - |\beta| - m$ (as is easily checked by induction). Therefore,

$$|D^\beta(\xi^\alpha \hat{E})| = O(|\xi|^{|\alpha|-|\beta|-m}) \qquad (|\xi| \to \infty).$$

Setting $|\beta| := |\alpha| + n - m + 1$ we get that $(x^\beta D^\alpha E)\hat{} \in L^1(\mathbb{R}^n)$ and the proof of Theorem 5.1.1 shows that $x^\beta D^\alpha E \in C(\mathbb{R}^n)$. Given any $x_0 \neq 0$, there exists such a β with $x_0^\beta \neq 0$ and so $D^\alpha E$ is smooth in a neighborhood of x_0, which establishes our claim. In particular, $\operatorname{singsupp} E \subseteq \{0\}$.

By (6.2.3), $(P(D)E)\hat{}(\xi) = P(\xi)\hat{E}(\xi) = 1 - \chi(\xi)$, so

$$P(D)E = \delta - \mathcal{F}^{-1}\chi =: \delta - \rho, \tag{6.2.4}$$

where $\rho \in C^\infty(\mathbb{R}^n)$, so $\operatorname{singsupp} E = \{0\}$.

Now let Ω' be open and relatively compact in Ω and choose some $\psi \in \mathcal{D}(\Omega)$ with $\psi \equiv 1$ on Ω'. Then (6.2.4) gives

$$\psi T = \delta * (\psi T) = E * (P(\psi T)) + \rho * (\psi T).$$

Here, $\rho * (\psi T) \in C^\infty$ and $\operatorname{singsupp}(E * P(\psi T)) \subseteq \operatorname{singsupp} P(\psi T)$ by Lemma 6.2.1. It follows that $\operatorname{singsupp}(\psi T) \subseteq \operatorname{singsupp} P(\psi T)$. But the converse inclusion is obvious, so

$$\operatorname{singsupp}(\psi T) = \operatorname{singsupp} P(\psi T).$$

Since Ω' as above was arbitrary, $\operatorname{singsupp}(T) = \operatorname{singsupp}(PT)$. □

A distribution E as constructed in the previous proof, i.e., such that $P(D)E = \delta + \rho$, with ρ smooth is called a C^∞-parametrix of P. Next, we prove the characterization of hypoellipticity referred to in Example 3.5.2:

6.2.5 Theorem. *Let $P(D)$ be a linear PDO with constant coefficients. Then the following are equivalent:*

(i) $P(D)$ is hypoelliptic.

(ii) There exists a fundamental solution E of $P(D)$ with $\operatorname{singsupp} E = \{0\}$.

Proof. (i)⇒(ii): By the Malgrange–Ehrenpreis Theorem 6.1.1 there exists a fundamental solution $E \in \mathcal{D}'(\mathbb{R}^n)$ for $P(D)$. Hence $P(D)E = \delta$. It follows that $\operatorname{singsupp} E = \operatorname{singsupp} \delta = \{0\}$.

(ii)⇒(i): If such an E exists, then (6.2.4) holds with $\rho = 0$. From here, it follows exactly as in the proof of Theorem 6.2.4 that $\operatorname{singsupp} PT = \operatorname{singsupp} T$ for any T. □

6.3 The Paley–Wiener–Schwartz theorem

For a function f on $\mathbb{R}^n$ the *Laplace transform* of f is defined by

$$p \mapsto \int e^{-px} f(x)\,dx \qquad (p \in \mathbb{C}^n).$$

Setting $p = i\zeta$ we obtain the *Fourier–Laplace transform*

$$\zeta \mapsto \int e^{-i\zeta x} f(x)\,dx, \tag{6.3.1}$$

which formally equals the Fourier transform but with a complex 'dual' variable ζ and reduces to the Fourier transform if $\zeta = \xi \in \mathbb{R}^n$. If f is a bounded measurable function with compact support then (6.3.1) defines an analytic function on $\mathbb{C}^n$. We are going to extend these notions to tempered distributions and bring in some complex variable techniques. In order to properly study this transform, we first collect some basic facts on analytic functions of several complex variables.

Let $f \in C^1(X)$ with $X \subseteq \mathbb{C}^n$, then

$$df = \sum_{j=1}^{n} \frac{\partial f}{\partial z_j}\, dz_j + \sum_{j=1}^{n} \frac{\partial f}{\partial \bar{z}_j}\, d\bar{z}_j,$$

where we have used the notation

$$\begin{aligned} z &= (z_1, \ldots, z_n) = (x_1 + iy_1, \ldots, x_n + iy_n) \\ \bar{z} &= (\bar{z}_1, \ldots, \bar{z}_n) = (x_1 - iy_1, \ldots, x_n - iy_n) \\ \frac{\partial}{\partial z_j} &= \frac{1}{2}\big(\partial_{x_j} - i\partial_{y_j}\big), \quad \frac{\partial}{\partial \bar{z}_j} = \frac{1}{2}\big(\partial_{x_j} + i\partial_{y_j}\big). \end{aligned}$$

The function f is called *analytic* if the Cauchy–Riemann differential equations, i.e., $\partial f/\partial \bar{z}_j = 0$ hold for all $1 \leq j \leq n$.

For $w \in \mathbb{C}^n$ and $r = (r_1, \ldots, r_n) \in (\mathbb{R}_+)^n$ we call the set

$$D(w, r) := \{z : \ |z_j - w_j| < r_j \quad (1 \leq j \leq n)\}$$

a *polydisc* of radius r around w and we clearly have $D(w, r) = D_1 \times \cdots \times D_n$, with $D_k = D(w_k, r_k)$. A repeated application of the one-dimensional Cauchy formula gives for any analytic function f on X and any $z \in D(w, r)$, a polydisc in X,

$$f(z) = \left(\frac{1}{2\pi i}\right)^n \int\limits_{\partial D_1} \cdots \int\limits_{\partial D_n} \frac{f(\zeta_1, \ldots, \zeta_n)}{(\zeta_1 - z_1) \cdots (\zeta_n - z_n)}\, d\zeta_1 \ldots d\zeta_n. \tag{6.3.2}$$

Since differentiation under the integral is permitted we see that $f \in \mathcal{C}^\infty(D(w, r))$. Moreover, $z_k \mapsto \frac{1}{\zeta_k - z_k}$ is analytic on D_k, so the derivatives $\partial^\alpha f$ of f satisfy the Cauchy–Riemann equations, hence are analytic in $D(w, r)$ as well. By the fact that X can be covered by polydiscs we have that any analytic f on X is actually smooth on X with all its derivatives again analytic on X.

We may now proceed as in the one-dimensional case to see that any f that is analytic on X has a power series expansion around any point $w \in X$. To make this more explicit, note first that on D_k

$$\frac{1}{\zeta_k - z_k} = \frac{1}{\zeta_k - w_k} \frac{1}{1 - \frac{z_k - w_k}{\zeta_k - w_k}} = \frac{1}{\zeta_k - w_k} \sum_{l=0}^{\infty} \Big(\frac{z_k - w_k}{\zeta_k - w_k}\Big)^l.$$

Therefore the series

$$\sum_{|\alpha| \geq 0} \frac{(z - w)^\alpha}{(\zeta_1 - w_1) \cdots (\zeta_n - w_n)(\zeta - w)^\alpha} = \frac{1}{(\zeta_1 - z_1) \cdots (\zeta_n - z_n)}$$

converges uniformly and absolutely on any compact subset of $D(w, r)$. Hence we may integrate in (6.3.2) term-wise to obtain

$$\begin{aligned}
f(z) = & \Big(\frac{1}{2\pi i}\Big)^n \sum_{|\alpha|\geq 0} (z-w)^\alpha \cdot \\
& \cdot \int\limits_{\partial D_1} \cdots \int\limits_{\partial D_n} \frac{f(\zeta_1,\dots,\zeta_n)}{(\zeta_1-w_1)\cdots(\zeta_n-w_n)(\zeta-w)^\alpha}\, d\zeta_1\cdots d\zeta_n \\
& = \sum_{|\alpha|\geq 0} \frac{(z-w)^\alpha}{\alpha!}\, \partial^\alpha f(w),
\end{aligned} \tag{6.3.3}$$

with the convergence being absolute and uniform on compact subsets of $D(w,r)$. For the last equality we have used

$$\partial^\alpha f(w) = \Big(\frac{1}{2\pi i}\Big)^n \alpha! \int\limits_{\partial D_1} \cdots \int\limits_{\partial D_n} \frac{f(\zeta)}{(\zeta_1-w_1)\cdots(\zeta_n-w_n)(\zeta-w)^\alpha}\, d\zeta_1\dots d\zeta_n,$$

which again is a consequence of (6.3.2).

The only fact about these basic properties of analytic functions we are going to use in the sequel is uniqueness of the analytic extension, i.e., the following statement.

Let $X \subseteq \mathbb{C}^n$ be open and connected. If f is analytic on X and there is a point $w \in X$ with $\partial^\alpha f(w) = 0$ for all α, then f=0 on X.

To prove this assertion set $Y := \{z \in X : \ \partial^\alpha f(z) = 0 \ \forall \alpha\}$, which is closed as the intersection of a family of closed sets. But by (6.3.3) each point in Y has a polydisc-shaped neighbourhood contained in Y, so Y is also open. By connectedness of X we have that $Y = X$ or $Y = \emptyset$. The latter is impossible since $w \in Y$ and we are done.

Finally, recall from Theorem 5.5.1 that for any $T \in \mathcal{E}'(\mathbb{R}^n)$ the Fourier transform, which is given by

$$\widehat{T}(\xi) = \langle T(x), e^{-ix\xi}\rangle \qquad (\xi \in \mathbb{R}^n),$$

can be extended to a holomorphic function on $\mathbb{C}^n$. This leads the way to the following definition.

6.3.1 Definition. *Let $T \in \mathcal{E}'(\mathbb{R}^n)$. Then we call the function*

$$\hat{T}(\zeta) := \langle T(x), e^{-ix\zeta}\rangle \qquad (\zeta \in \mathbb{C}^n) \tag{6.3.4}$$

the Fourier–Laplace transform *of T.*

6.3.2 Remark.

(i) As already indicated above, Theorem 5.5.1 tells us that the Fourier–Laplace transform of any $T \in \mathcal{E}'(\mathbb{R}^n)$ is an entire function.

(ii) In case $T = u \in \mathcal{D}(\mathbb{R}^n)(= \mathcal{E}'(\mathbb{R}^n) \cap \mathcal{C}^\infty(\mathbb{R}^n))$ equation (6.3.4) obviously reduces to

$$\hat{T}(\zeta) = \int\limits_{\mathbb{R}^n} u(x) e^{-ix\zeta}\, dx \tag{6.3.5}$$

and in case $\zeta = \xi \in \mathbb{R}^n$ (6.3.4) gives back the Fourier transform.

The following result establishes a connection between the support of a distribution and fall-off properties of its Fourier–Laplace transform. This fact has far-reaching consequences in microlocal analysis, cf. Section 6.4 below.

6.3.3 Proposition.

(i) Let $T \in \mathcal{E}'(\mathbb{R}^n)$ *with* $\operatorname{supp}(T) \subseteq \overline{B_a(0)}$*. Then there exist* C*,* $N > 0$ *such that*

$$|\hat{T}(\zeta)| \leq C(1+|\zeta|)^N \, e^{a|\mathrm{Im}\zeta|} \qquad (\zeta \in \mathbb{C}^n). \tag{6.3.6}$$

(ii) Let $T \in \mathcal{D}(\mathbb{R}^n)$ *with* $\operatorname{supp}(T) \subseteq \overline{B_a(0)}$*. Then for any* $m \geq 0$ *there exist* $C_m > 0$ *such that*

$$|\hat{T}(\zeta)| \leq C_m(1+|\zeta|)^{-m} \, e^{a|\mathrm{Im}\zeta|} \qquad (\zeta \in \mathbb{C}^n). \tag{6.3.7}$$

Proof. (i) Let $\varphi \in \mathcal{C}^\infty(\mathbb{R})$ such that $\varphi(t) \equiv 0$ for $t \leq -1$ and $\varphi(t) \equiv 1$ for $t \geq -1/2$ and set

$$\varphi_\zeta(x) := \varphi(|\zeta|^2(a^2 - |x|^2)) \qquad (\zeta \in \mathbb{C}^n).$$

We then have $\varphi_\zeta \equiv 1$ for $\zeta = 0$, while for $\zeta \neq 0$ we find

$$\varphi_\zeta(x) \equiv 0 \text{ for } |x|^2 \geq a^2 + |\zeta|^{-2} \text{ and } \varphi_\zeta(x) \equiv 1 \text{ for } |x|^2 \leq a^2 + \frac{1}{2}|\zeta|^{-2},$$

hence, in particular, $\varphi_\zeta \in \mathcal{D}(\mathbb{R}^n)$ for $\zeta \neq 0$. By the support condition on T we may write

$$\hat{T}(\zeta) = \langle T(x), \varphi_\zeta(x) e^{-ix\zeta} \rangle.$$

Now since $\hat{T}$ is smooth, it is bounded on $|\zeta| \leq 1$. To obtain the estimate (6.3.6) also for $|\zeta| \geq 1$, we note that in this region we have $\operatorname{supp}(\varphi_\zeta) \subseteq \{|x|^2 \leq a^2 + 1\} =: K$. So we may use (1.1.1) to obtain the existence of N, C with

$$|\hat{T}(\zeta)| \leq C \sum_{|\alpha| \leq N} \|D_x^\alpha(\varphi_\zeta(x) e^{-ix\zeta})\|_{L^\infty(K)},$$

hence the claim follows from the Leibnitz rule since on K we have for $|\zeta| \geq 1$:

$$|D_x^\beta \varphi_\zeta(x)| \quad \leq \quad C_\beta |\zeta|^{2|\beta|}$$

and, since $|x|^2 \leq a^2 + |\zeta|^{-2}$ implies $|x| \leq \sqrt{a^2 + |\zeta|^{-2}} \leq a + |\zeta|^{-1}$:

$$|D_x^\gamma e^{-ix\zeta}| \quad = \quad |\zeta|^{|\gamma|} e^{|\mathrm{Im}\zeta||x|} \leq |\zeta|^{|\gamma|} e^{|\mathrm{Im}\zeta|(a + |\zeta|^{-1})} \leq |\zeta|^{|\gamma|} e^{a|\mathrm{Im}\zeta|} e^1.$$

(ii) Let now $T \in \mathcal{D}(\mathbb{R}^n)$, then by (6.3.5) $\hat{T}(\zeta) = \int e^{-ix\zeta} T(x) dx$ and we may use integration by parts to obtain

$$\zeta^\alpha \hat{T}(\zeta) = \int e^{-ix\zeta} D^\alpha T(x) \, dx \qquad \forall \alpha \in \mathbb{N}_0^n.$$

So for all α

$$\begin{aligned} |\zeta^\alpha \hat{T}(\zeta)| \quad &\leq \quad \|D^\alpha T\|_{L^\infty(\mathbb{R}^n)} \sup_{x \in \operatorname{supp}(T)} e^{|x||\mathrm{Im}\zeta|} \int\limits_{|x| \leq a} dx \\ &\leq \quad C \|D^\alpha T\|_{L^\infty(\mathbb{R}^n)} e^{a|\mathrm{Im}\zeta|}, \end{aligned}$$

which gives the claim. □

The following theorem shows that the above estimates are actually characterizing.

6.3.4 Theorem. *(Paley–Wiener–Schwartz) Let* $a > 0$ *and let* $f : \mathbb{C}^n \to \mathbb{C}$ *be analytic.*

(i) f is the Fourier–Laplace transform of some $T \in \mathcal{E}'(\mathbb{R}^n)$ with $\mathrm{supp}\,(T) \subseteq \overline{B_a(0)}$ *if and only if*

$$\exists C, N > 0 : \; |f(\zeta)| \le C(1+|\zeta|)^N e^{a|\mathrm{Im}\zeta|} \qquad (\zeta \in \mathbb{C}^n). \tag{6.3.8}$$

(ii) f is the Fourier–Laplace transform of some $T \in \mathcal{D}(\mathbb{R}^n)$ with $\mathrm{supp}\,(T) \subseteq \overline{B_a(0)}$ *if and only if*

$$\forall m \ge 0 \; \exists C_m > 0 : \; |f(\zeta)| \le C_m(1+|\zeta|)^{-m} e^{a|\mathrm{Im}\zeta|} \qquad (\zeta \in \mathbb{C}^n). \tag{6.3.9}$$

Proof. (i),(ii) $\Rightarrow$: In both cases this is just Proposition 6.3.3.
(ii) $\Leftarrow$: To begin with we set $m = n+1$ and $\zeta = \xi \in \mathbb{R}^n$ in (6.3.9). Then $\xi \mapsto f(\xi)$ is in $L^1(\mathbb{R}^n)$ and by the same reasoning as in 5.1.1(i) we find that

$$T(x) := (2\pi)^{-n} \int e^{ix\xi} f(\xi)\, d\xi \tag{6.3.10}$$

is continuous.
Now setting $m = |\alpha| + n + 1$ in (6.3.9) we find that also $\xi \mapsto \xi^\alpha f(\xi)$ is in $L^1(\mathbb{R}^n)$ for all $|\alpha| \le m$ and we may differentiate under the integral in (6.3.10). So we have that $T \in \mathcal{C}^\infty(\mathbb{R}^n)$.
We claim that

$$\mathrm{supp}\,(T) \subseteq \{|x| \le a\}. \tag{6.3.11}$$

Since each of the functions $\zeta_j \mapsto f(\zeta)$ is analytic we may use Cauchy's theorem in each of the variables ζ_j $(1 \le j \le n)$ to shift the integration with respect to ξ_j in (6.3.10) into the complex domain, namely (given any $\eta_j \in \mathbb{R}$) from $\mathbb{R} \subseteq \mathbb{C}$ to $\mathbb{R} + i\eta_j$. By (6.3.9) the integrals parallel to the imaginary axis vanish in the limit, and we may replace (6.3.10) by

$$T(x) = (2\pi)^{-n} \int\limits_{\mathrm{Im}\zeta = \eta} e^{ix\zeta} f(\zeta)\, d\zeta \overset{\substack{\zeta=\xi+i\eta \\ \downarrow}}{=} (2\pi)^{-n} \int\limits_{\mathbb{R}^n} e^{ix(\xi+i\eta)} f(\xi+i\eta)\, d\xi,$$

where $\eta \in \mathbb{R}^n$ is arbitrary. Now again setting $m = n+1$ in (6.3.9) we obtain

$$|T(x)| \le (2\pi)^{-n} C_{n+1} e^{a|\eta| - x\eta} \int\limits_{\mathbb{R}^n} \underbrace{(1+|\xi+i\eta|)^{-n-1}}_{\text{integrable}} d\xi \le C e^{a|\eta|-x\eta}.$$

Next, we set $\eta = tx/|x|$ $(t > 0)$ and obtain $|T(x)| \le Ce^{(a-|x|)t}$. By taking the limit $t \to \infty$ we see that $T(x)$=0 if $|x| > a$, which establishes (6.3.11).
Knowing that $T \in \mathcal{D}(\mathbb{R}^n)$ we may apply the Fourier inversion formula in $\mathcal{S}$ (Lemma 5.2.9) to obtain $\hat{T}(\xi) = f(\xi)$ for all $\xi \in \mathbb{R}^n$. Moreover, we know from Theorem 5.5.1 that $\hat{T}$ extends to an analytic function on $\mathbb{C}^n$, so by uniqueness of analytic continuation we obtain $\hat{T} = f$ on $\mathbb{C}^n$.
(i) $\Leftarrow$: By Remark 5.3.6 (iii) and (6.3.8) we have that $f|_{\mathbb{R}^n} \in \mathcal{S}'(\mathbb{R}^n)$. So by Theorem 5.4.2 $f|_{\mathbb{R}^n}$ is the Fourier transform of some $T \in \mathcal{S}'(\mathbb{R}^n)$, i.e., $\hat{T} = f|_{\mathbb{R}^n}$.
Let now ρ be a mollifier and set $T_\varepsilon := \rho_\varepsilon * T$. Then by the convolution theorem 5.5.3 and formula (5.2.7) we find

$$\widehat{T_\varepsilon}(\xi) = \widehat{\rho_\varepsilon}(\xi)\hat{T}(\xi) = \hat{\rho}(\varepsilon\xi) f(\xi).$$

Next we combine 6.3.3 (ii) for $\hat{\rho}(\varepsilon.)$ with (6.3.8) to obtain that $\widehat{T_\varepsilon}$ extends to an analytic function on $\mathbb{C}^n$ such that for all $m \in \mathbb{N}_0$ the estimate

$$|\widehat{T_\varepsilon}(\zeta)| \le C_m(1+\varepsilon|\zeta|)^{-m}\, C(1+|\zeta|)^N e^{(a+\varepsilon)|\mathrm{Im}\zeta|}$$

holds. Upon replacing m by $m+N$ and noting that $\frac{(1+|\zeta|)^N}{(1+\varepsilon|\zeta|)^{m+N}} \le \frac{1}{\varepsilon^{m+N}} \frac{1}{(1+|\zeta|)^m}$ we see that T_ε satisfies (6.3.9) with $a+\varepsilon$ replacing a. So by (ii) we find some $S \in \mathcal{D}(\mathbb{R}^n)$ with $\hat{S} = \widehat{T_\varepsilon}$ and $\operatorname{supp}(S) \subseteq \{|x| \le a+\varepsilon\}$. Hence $S = T_\varepsilon$ and we obtain $\operatorname{supp}(T_\varepsilon) \subseteq \{|x| \le a+\varepsilon\}$.

Next we show that actually $\operatorname{supp}(T) \subseteq \{|x| \le a\}$. If $x_0 \notin \overline{B_a(0)}$ then there exists $\varepsilon_0 > 0$ and a neighborhood V of x_0 such that $|y| > a + \varepsilon_0$ for all $y \in V$. So by the above $T_\varepsilon = 0$ on V for all $\varepsilon < \varepsilon_0$ and we have for all $\varphi \in \mathcal{D}(V)$

$$\langle T, \varphi \rangle = \lim_{\varepsilon \to 0} \langle T_\varepsilon, \varphi \rangle = 0.$$

So $T|_V = 0$, hence $x_0 \notin \operatorname{supp}(T)$.

Finally we proceed as in (ii): Again Theorem 5.5.1 says that $\hat{T}$ extends analytically to $\mathbb{C}^n$ and by the fact that $\hat{T}|_{\mathbb{R}^n} = f|_{\mathbb{R}^n}$ and by uniqueness of analytic continuation we have $\hat{T} = f$ on $\mathbb{C}^n$, concluding the proof. □

6.4 Spectral analysis of singularities, wave front sets

Up to now we have only rather crude tools at hand to study the structure of the singularities of a given distribution. In particular, the singular support gives a measure of the size of the set where a distribution is not smooth. Using it one can also obtain certain insights into the regularity of various operations on distributions. For the case of convolution, recall from (6.2.1) that

$$\operatorname{singsupp}(S * T) \subseteq \operatorname{singsupp}(S) + \operatorname{singsupp}(T) \qquad (S \in \mathcal{E}'(\mathbb{R}^n),\ T \in \mathcal{D}'(\mathbb{R}^n)).$$

A versatile and powerful tool for a closer study of the structure of the singularity set of a distribution is provided by the Fourier transformation. It allows one to distinguish regular and irregular frequency directions. We first introduce some terminology:

6.4.1 Definition. *A set $\Gamma \subseteq \mathbb{R}^n \setminus \{0\}$ is called a* cone *if for every $\xi \in \Gamma$ and every $\lambda > 0$ we have $\lambda\xi \in \Gamma$. A cone Γ is called a conic neighborhood of $\xi_0 \in \mathbb{R}^n$ if Γ is open and $\xi_0 \in \Gamma$.*

The following abbreviation will be useful for studying asymptotic behavior:

$$\langle \xi \rangle := (1 + |\xi|^2)^{1/2} \qquad (\xi \in \mathbb{R}^n). \tag{6.4.1}$$

6.4.2 Definition. *The set of regular frequency directions of $T \in \mathcal{E}'(\mathbb{R}^n)$ is*

$$\begin{aligned}\Gamma_\infty(T) := \{\xi_0 \in \mathbb{R}^n \setminus \{0\} \mid \exists \text{ conic nbhd } V_0 \text{ of } \xi_0 :\\ \forall N \in \mathbb{N}\ \exists C_N : \forall \xi \in V_0\ |\hat{T}(\xi)| \le C_N \langle \xi \rangle^{-N}\}.\end{aligned}$$

The set of irregular frequency directions of T is

$$\Sigma(T) := \mathbb{R}^n \setminus (\Gamma_\infty(T) \cup \{0\}) = \mathbb{R}^n \setminus \Gamma_\infty(T) \cap \mathbb{R}^n \setminus \{0\} \subseteq \mathbb{R}^n \setminus \{0\}.$$

Since $\Gamma_\infty(T)$ is an open cone, $\Sigma(T)$ is closed in $\mathbb{R}^n \setminus \{0\}$. Moreover, $\Sigma(T)$ is a cone: let $\xi_0 \in \Sigma(T)$ and let $\lambda > 0$. Then $\xi_0 \notin \Gamma_\infty(T)$ obviously implies $\lambda\xi_0 \notin \Gamma_\infty(T)$.

The following result shows that $\Sigma(T)$ encodes the regularity of T:

6.4.3 Lemma. *Let $T \in \mathcal{E}'(\mathbb{R}^n)$. TFAE:*

(i) $T \in \mathcal{D}(\mathbb{R}^n)$.

(ii) $\Sigma(T) = \emptyset$.

Proof. (i)$\Rightarrow$(ii): If $T \in \mathcal{D}(\mathbb{R}^n)$, then (6.3.7) shows that $\Gamma_\infty(T) = \mathbb{R}^n \setminus \{0\}$, so $\Sigma(T) = \emptyset$.
(ii)$\Rightarrow$(i): By assumption, $\Gamma_\infty(T) = \mathbb{R}^n \setminus \{0\} \supseteq S^{n-1}$. Thus any $\xi_0 \in S^{n-1}$ has a conic neighborhood V_{ξ_0} in which $\hat{T}$ is rapidly decreasing. Since S^{n-1} is compact, it is covered by finitely many such neighborhoods, so for any $N \in \mathbb{N}$ we can find some $C_N > 0$ with

$$|\hat{T}(\xi)| \leq C_N \langle \xi \rangle^{-N} \qquad (\xi \in \mathbb{R}^n).$$

Choosing $N > n$ it follows that, in particular, $\hat{T} \in L^1(\mathbb{R}^n)$, so

$$T(x) = \mathcal{F}^{-1}(\hat{T})(x) = (2\pi)^{-n} \int e^{ix\xi} \hat{T}(\xi)\, d\xi.$$

In particular, T is continuous. Moreover,

$$\partial^\alpha T(x) = (2\pi)^{-n} i^{|\alpha|} \int \xi^\alpha e^{ix\xi} \hat{T}(\xi)\, d\xi$$

is continuous for any α since the integrand is in $L^1(\mathbb{R}^n)$. Thus, T is smooth. □

In order to extend the above notions to distributions with not necessarily compact support, we use localization. We first need the following preparatory result:

6.4.4 Lemma. *Let $T \in \mathcal{E}'(\mathbb{R}^n)$ and let $\varphi \in \mathcal{D}(\mathbb{R}^n)$. Then*

$$\Sigma(\varphi T) \subseteq \Sigma(T).$$

Proof. Let $0 \neq \xi_0$, $\xi_0 \notin \Sigma(T)$. Then there exists a conic neighborhood Γ of ξ_0 and for each $N \in \mathbb{N}$ there is some $C_N > 0$ such that

$$|\hat{T}(\xi)| \leq C_N \langle \xi \rangle^{-N} \qquad \forall \xi \in \Gamma.$$

Let Γ' be a conic neighborhood of ξ_0 with $\Gamma' \cap S^{n-1} \subseteq \overline{\Gamma'} \cap S^{n-1} \Subset \Gamma \cap S^{n-1}$. It then suffices to show that $(\varphi T)\hat{}$ is rapidly decreasing in Γ' (as this will imply that $\xi_0 \notin \Sigma(\varphi T)$).
Let $\psi \in \mathcal{D}(\mathbb{R}^n)$. Then

$$\begin{aligned}
\langle (\varphi T)\hat{}, \psi \rangle &= \langle \varphi T, \hat{\psi} \rangle = \langle T, \varphi \hat{\psi} \rangle \underset{(5.2.11)}{=} (2\pi)^{-n} \langle T, (\widehat{\hat{\varphi}})\check{}\, \hat{\psi} \rangle = (2\pi)^{-n} \langle T, ((\hat{\varphi})\check{})\hat{}\, \hat{\psi} \rangle \\
&\underset{(5.5.3)}{=} (2\pi)^{-n} \langle T, ((\hat{\varphi})\check{} * \psi)\hat{} \rangle = (2\pi)^{-n} \langle \hat{T}, (\hat{\varphi})\check{} * \psi) \rangle \\
&= \langle (2\pi)^{-n} \hat{T} * \hat{\varphi}, \psi \rangle,
\end{aligned}$$

so $(\varphi T)\hat{} = (2\pi)^{-n} \hat{T} * \hat{\varphi}$.
Since $T \in \mathcal{E}'(\mathbb{R}^n)$, Theorem 6.3.4 implies that $\hat{T}$ is analytic and that there exist C, $M > 0$ such that

$$|\hat{T}(\xi)| \leq C(1 + |\xi|)^M \qquad (\xi \in \mathbb{R}^n). \tag{6.4.2}$$

Also, $\hat{\varphi} \in \mathcal{S}(\mathbb{R}^n)$, so

$$(\varphi T)\hat{}(\xi) = (2\pi)^{-n} \hat{T} * \hat{\varphi}(\xi) = (2\pi)^{-n} \int \hat{\varphi}(\eta) \hat{T}(\xi - \eta)\, d\eta.$$

Since $\overline{\Gamma'} \cap S^{n-1} \Subset \Gamma$, there exists some $\varepsilon \in (0,1)$ such that

$$B_\varepsilon(\Gamma' \cap S^{n-1}) := \{\xi \in S^{n-1} \mid \exists \eta \in \Gamma' \cap S^{n-1} : |\xi - \eta| < \varepsilon\} \subseteq \Gamma, \tag{6.4.3}$$

We now claim that, whenever $\xi \in \Gamma'$, $\eta \in \mathbb{R}^n$, and $|\eta| < \varepsilon|\xi|$, it follows that $\xi - \eta \in \Gamma$. To see this, we distinguish two cases: First, if $|\xi| = 1$, then $|\eta| < \varepsilon$ and the claim follows directly from (6.4.3). Second, if $0 \neq \xi$ is arbitrary, then by the first case

$$\left|\frac{\eta}{|\xi|}\right| < \varepsilon,\ \frac{\xi}{|\xi|} \in \Gamma' \cap S^{n-1} \Rightarrow \frac{\xi}{|\xi|} - \frac{\eta}{|\xi|} \in \Gamma \Rightarrow \xi - \eta \in \Gamma$$

since Γ is a cone.

After these preparations we are now ready to show that indeed $(\varphi T)\hat{}$ is rapidly decreasing in Γ'. Let $\xi \in \Gamma'$. Then

$$(\varphi T)\hat{}(\xi) = (2\pi)^{-n}\Big[\int_{|\eta|>\varepsilon|\xi|} \hat{\varphi}(\eta)\hat{T}(\xi - \eta)\, d\eta + \int_{|\eta|<\varepsilon|\xi|} \hat{\varphi}(\eta)\hat{T}(\xi-\eta)\, d\eta\Big],$$

and we will show that both summands are rapidly decreasing in $\xi \in \Gamma'$.

Since $\hat{\varphi} \in \mathcal{S}(\mathbb{R}^n)$, for any $N \in \mathbb{N}$ there exists some $C_N > 0$ such that

$$|\hat{\varphi}(\eta)| \leq C_N \langle \eta \rangle^{-N}. \tag{6.4.4}$$

Combining this with (6.4.2) we obtain

$$\begin{aligned}
&\Big|\int_{|\eta|>\varepsilon|\xi|} \hat{\varphi}(\eta)\hat{T}(\xi-\eta)\, d\eta\Big| \\
&\qquad \leq C_N \sup_{|\eta|>\varepsilon|\xi|} \langle \eta \rangle^{n+1-N} \cdot C \cdot \int_{|\eta|>\varepsilon|\xi|} \langle \eta \rangle^{-(n+1)} \underbrace{(1 + |\xi - \eta|)^M}_{\leq \langle \xi \rangle^M \langle \eta \rangle^M} d\eta \\
&\qquad \leq \tilde{C}_N \sup_{|\eta|>\varepsilon|\xi|} \langle \eta \rangle^{n+1-N+M} \langle \xi \rangle^M \int_{|\eta|>\varepsilon|\xi|} \langle \eta \rangle^{-(n+1)}\, d\eta \\
&\qquad \leq C(\varepsilon, N)\langle \xi \rangle^{n+1-N+2M},
\end{aligned}$$

for $N > n + 1 + 2M$, implying rapid decrease in ξ. For the second term we note that, as we saw above, $|\eta| < \varepsilon|\xi|$ and $\xi \in \Gamma'$ imply $\xi - \eta \in \Gamma$, where $\hat{T}$ is rapidly decreasing. Combining this with (6.4.4), for any N we get

$$\Big|\int_{|\eta|<\varepsilon|\xi|} \hat{\varphi}(\eta)\hat{T}(\xi-\eta)\, d\eta\Big| \leq \tilde{C}_N \int_{|\eta|<\varepsilon|\xi|} \langle \eta \rangle^{-(n+1)} \langle \xi - \eta \rangle^{-N}\, d\eta.$$

Here, $|\xi - \eta| \geq |\xi| - |\eta| \geq (1 - \varepsilon)|\xi|$, so

$$\begin{aligned}
&\tilde{C}_N \int_{|\eta|<\varepsilon|\xi|} \langle \eta \rangle^{-(n+1)} \langle \xi - \eta \rangle^{-N}\, d\eta \\
&\qquad \leq \tilde{C}_N (1-\varepsilon)^{-N} \langle \xi \rangle^{-N} \int_{\mathbb{R}^n} \langle \eta \rangle^{-(n+1)}\, d\eta \leq C(\varepsilon, N)\langle \xi \rangle^{-N},
\end{aligned}$$

showing that also in this case we obtain rapid decrease in ξ. □

The previous result motivates the idea of *micro-localization*:

6.4.5 Definition. *Let $T \in \mathcal{D}'(\Omega)$, and let $x \in \Omega$. The cone of irregular frequency directions of T at x is*

$$\Sigma_x(T) := \bigcap \{\Sigma(\varphi \cdot T) \mid \varphi \in \mathcal{D}(\Omega),\ \varphi(x) \neq 0\}. \tag{6.4.5}$$

Note that, being an intersection of closed cones in $\mathbb{R}^n \setminus \{0\}$, $\Sigma_x(T)$ itself is such a cone.

6.4.6 Lemma. *Let $T \in \mathcal{D}'(\Omega)$ and let $I \ni \iota \mapsto \varphi_\iota$ be a net in $\mathcal{D}(\Omega)$ such that $\varphi_\iota(x) \neq 0$ for all ι and such that $\operatorname{supp}\varphi_\iota \to \{x\}$ (i.e., for any neighborhood U of x there exists some ι_0 such that $\operatorname{supp}(\varphi_\iota) \subseteq U$ for all $\iota \geq \iota_0$). Then*

(i) $\Sigma(\varphi_\iota T) \to \Sigma_x(T)$ *(as closed cones).*

(ii) $\bigcap_{\iota\in I} \Sigma(\varphi_\iota T) = \Sigma_x(T)$.

Proof. (i) Let V be an open cone containing $\Sigma_x(T) = \bigcap\{\Sigma(\varphi{\cdot}T) \mid \varphi \in \mathcal{D}(\Omega),\ \varphi(x) \neq 0\}$. Then

$$\bigcap_{\varphi}(\Sigma(\varphi T) \cap S^{n-1}) \subseteq V \Rightarrow \emptyset = \bigcap_{\varphi}(\Sigma(\varphi T) \cap S^{n-1} \cap (\mathbb{R}^n \setminus V)).$$

The sets in this last intersection are compact, so there exist $\varphi_1, \dots, \varphi_j \in \mathcal{D}(\Omega)$ with $\varphi_i(x) \neq 0$ for $1 \leq i \leq j$ such that $\emptyset = \bigcap_{i=1}^{j}(\Sigma(\varphi_i T) \cap S^{n-1} \cap (\mathbb{R}^n \setminus V))$. Since V is a cone we conclude that

$$\bigcap_{i=1}^{j}(\Sigma(\varphi_i T) \cap S^{n-1}) \subseteq V \Rightarrow \bigcap_{i=1}^{j}(\Sigma(\varphi_i T)) \subseteq V.$$

By assumption there exists some $\iota_0 \in I$ such that for each $\iota \geq \iota_0$ we have

$$\operatorname{supp}\varphi_\iota \subseteq \bigcap_{i=1}^{j}\{y \in \Omega \mid \varphi_i(y) \neq 0\}$$

(which is an open neighborhood of x). In particular, $\prod_{i=1}^{j}\varphi_i \neq 0$ on $\operatorname{supp}\varphi_\iota$, so $\psi_\iota := \varphi_\iota / \prod_{i=1}^{j}\varphi_i \in \mathcal{D}(\Omega)$, and we can write

$$\varphi_\iota = \varphi_1 \cdots \varphi_j \psi_\iota$$

for each $\iota \geq \iota_0$. For such ι, Lemma 6.4.4 implies

$$\Sigma(\varphi_\iota T) = \Sigma(\varphi_1 \cdots \varphi_j \psi_\iota T) \subseteq \Sigma(\varphi_1 \cdots \varphi_j T) \subseteq \bigcap_{i=1}^{j}\Sigma(\varphi_i T) \subseteq V,$$

giving the claim.
(ii) Any closed cone Γ is the intersection of all open cones containing it: indeed, it suffices to observe that

$$\Gamma \cap S^{n-1} = \bigcap_{\varepsilon>0} B_\varepsilon(\Gamma \cap S^{n-1}).$$

By (i), any open cone containing $\Sigma_x(T)$ also contains $\bigcap_{\iota\in I}\Sigma(\varphi_\iota T)$, so (6.4.5) implies

$$\bigcap_{\iota\in I}\Sigma(\varphi_\iota T) \subseteq \Sigma_x(T) \subseteq \bigcap_{\iota\in I}\Sigma(\varphi_\iota T).$$

□

6.4.7 Proposition. *Let $T \in \mathcal{D}'(\Omega)$ and let $x \in \Omega$. TFAE:*

(i) $x \notin \operatorname{singsupp}(T)$.

(ii) There exists some $\varphi \in \mathcal{D}(\Omega)$ with $\varphi(x) \neq 0$ such that $\varphi T \in C^\infty(\Omega)$.

(iii) $\Sigma_x(T) = \emptyset$.

Proof. (i)$\Leftrightarrow$(ii) is clear.
(ii)$\Rightarrow$(iii): For φ as in (ii), Lemma 6.4.3 implies that $\Sigma(\varphi T) = \emptyset$. Thus

$$\Sigma_x(T) = \bigcap \{\Sigma(\varphi \cdot T) \mid \varphi \in \mathcal{D}(\Omega),\, \varphi(x) \neq 0\} = \emptyset.$$

(iii)$\Rightarrow$(ii): As in the proof of Lemma 6.4.6 (i) it follows that there exist $\varphi_1, \dots, \varphi_j \in \mathcal{D}(\Omega)$ with $\varphi_i(x) \neq 0$ for all $i = 1, \dots, j$ and $\bigcap_{i=1}^j \Sigma(\varphi_i T) = \emptyset$. Now choose some $\varphi \in \mathcal{D}(\Omega)$, $\varphi(x) \neq 0$, such that $\operatorname{supp} \varphi \subseteq \bigcap_{i=1}^j \{y \in \Omega \mid \varphi_i(y) \neq 0\}$. Then, with a ψ as in the proof of Lemma 6.4.6,

$$\Sigma(\varphi T) = \Sigma(\varphi_1 \cdots \varphi_j \psi T) \subseteq \bigcap_{i=1}^{j} \Sigma(\varphi_i T) = \emptyset.$$

Lemma 6.4.3 now implies that $\varphi T \in \mathcal{D}(\Omega)$, so (ii) follows. □

6.4.8 Definition. *Let $T \in \mathcal{D}'(\Omega)$. The wave front set of T is*

$$\mathrm{WF}(T) := \{(x, \xi) \in \Omega \times (\mathbb{R}^n \setminus \{0\}) \mid \xi \in \Sigma_x(T)\}.$$

Some fundamental properties of wave front sets are collected in the following result:

6.4.9 Theorem. *Let $T \in \mathcal{D}'(\Omega)$, then:*

(i) $\mathrm{WF}(T)$ *is closed in $\Omega \times (\mathbb{R}^n \setminus \{0\})$ and conic in the second component, i.e.: $(x, \xi) \in \mathrm{WF}(T)$, $\lambda > 0 \Rightarrow (x, \lambda\xi) \in \mathrm{WF}(T)$.*

(ii) $\mathrm{pr}_1(\mathrm{WF}(T)) = \operatorname{singsupp}(T)$.

(iii) If $T \in \mathcal{E}'(\mathbb{R}^n)$, then $\mathrm{pr}_2(\mathrm{WF}(T)) = \Sigma(T)$.

Proof. (i) Let $(y, \eta) \in (\Omega \times \mathbb{R}^n \setminus \{0\}) \setminus \mathrm{WF}(T)$. Then

$$0 \neq \eta \notin \Sigma_y(T) = \bigcap \{\Sigma(\varphi \cdot T) \mid \varphi \in \mathcal{D}(\Omega),\, \varphi(y) \neq 0\}.$$

Hence there exists some $\varphi \in \mathcal{D}(\Omega)$ with $\varphi(y) \neq 0$ and $\eta \notin \Sigma(\varphi T)$, i.e., $\eta \in \Gamma_\infty(\varphi T)$. Pick an open neighborhood U of y contained in the interior of $\operatorname{supp} \varphi$, as well as a conic neighborhood V of η contained in $\Gamma_\infty(\varphi T)$. Then

$$(y, \eta) \in U \times V \subseteq (\Omega \times (\mathbb{R}^n \setminus \{0\})) \setminus \mathrm{WF}(T).$$

$\mathrm{WF}(T)$ is conic in the second component since each $\Sigma_x(T)$ is a cone.
(ii) By Proposition 6.4.7,

$$\mathrm{pr}_1(\mathrm{WF}(T)) = \{x \in \Omega \mid \Sigma_x(T) \neq \emptyset\} = \operatorname{singsupp}(T).$$

(iii) Let $W := \mathrm{pr}_2(\mathrm{WF}(T))$. Then if $\xi \in W$, for a suitable x and any $\varphi \in \mathcal{D}(\Omega)$ with $\varphi(x) \neq 0$ we have $\xi \in \Sigma_x(T) \subseteq \Sigma(\varphi T) \subseteq \Sigma(T)$. Thus $W \subseteq \Sigma(T)$. Moreover, W is a cone.
We claim that

$$W \cap S^{n-1} = \mathrm{pr}_2(\mathrm{WF}(T) \cap (\operatorname{supp}(T) \times S^{n-1})).$$

Indeed, if $\xi \in W \cap S^{n-1}$ then there exists some $x \in \Omega$ such that $\xi \in \Sigma_x(T)$, and $|\xi| = 1$. In particular, $x \in \operatorname{singsupp}(T) \subseteq \operatorname{supp}(T)$. Since $(x, \xi) \in \mathrm{WF}(T)$, it follows that $\xi \in \mathrm{pr}_2(\mathrm{WF}(T) \cap (\operatorname{supp}(T) \times S^{n-1}))$. If, conversely, ξ is contained in the right hand side, then there exists some $x \in \operatorname{supp}(T)$ such that $(x, \xi) \in \mathrm{WF}(T)$ and $|\xi| = 1$.

In particular, $W \cap S^{n-1}$ is compact and W is a cone, so W is a closed cone.

Now let V be a conic neighborhood of W. Then V also is a conic neighborhood of $\Sigma_x(T)$, for each $x \in \Omega$. Hence for any $x \in \Omega$ there exists an open neighborhood U_x of x with $\Sigma(\varphi T) \subseteq V$ for every $\varphi \in \mathcal{D}(U_x)$ with $\varphi(x) \neq 0$ (see Lemma 6.4.6 (i)). Since $\operatorname{supp}(T)$ is compact, there exists a finite covering of $\operatorname{supp}(T)$ by such $U_{x_1}, \ldots, U_{x_j}$. Now pick such $\varphi_i \in \mathcal{D}(U_{x_i})$ with $\sum_{i=1}^{j} \varphi_i = 1$ near $\operatorname{supp}(T)$. Then

$$\Sigma(T) = \Sigma\Big(\sum_{i=1}^{j} \varphi_i T\Big) \subseteq \bigcup_{i=1}^{j} \Sigma(\varphi_i T) \subseteq V.$$

Again using that W is the intersection of all open cones containing it, we conclude that $\Sigma(T) \subseteq W$. Altogether, it follows that $W = \Sigma(T)$. □

6.4.10 Examples.

(i) $\mathrm{WF}(\delta) = \{0\} \times (\mathbb{R}^n \setminus \{0\})$.

Indeed, using Theorem 6.4.9 we have

$$\begin{aligned} &\mathrm{pr}_1(\mathrm{WF}(\delta)) = \operatorname{singsupp}(\delta) = \{0\} \\ &\mathrm{pr}_2(\mathrm{WF}(\delta)) = \Sigma(\delta) = \mathbb{R}^n \setminus (\{0\} \cup \Gamma_\infty(\delta)). \end{aligned}$$

Here, $\Gamma_\infty(\delta)$ is the set of all $\xi_0 \in \mathbb{R}^n$ such that $\hat{\delta} = 1$ is rapidly decreasing in a conic neighborhood of ξ_0, hence is empty. Consequently, $\Sigma(\delta) = \mathbb{R}^n \setminus \{0\}$ and our claim follows.

(ii) Here we consider an entire class of examples, namely that of Euclidean measures on subspaces of $\mathbb{R}^n$. Let V be a linear subspace of $\mathbb{R}^n$, dS the surface measure on V and let $u_0 \in C^\infty(V)$. Then

$$\langle u, \varphi \rangle := \int_V \varphi(x) u_0(x)\, dS(x) \qquad (\varphi \in \mathcal{D}(\mathbb{R}^n))$$

is a Radon measure, hence a distribution of order 0 on $\mathbb{R}^n$. We claim that

$$\mathrm{WF}(u) = \operatorname{supp}(u) \times V^\perp \setminus \{0\}. \tag{6.4.6}$$

To see this, let $\varphi \in \mathcal{D}(\mathbb{R}^n)$. Then $\varphi \cdot u \in \mathcal{E}'(\mathbb{R}^n)$, and decomposing any $\xi \in \mathbb{R}^n$ as $\xi = \xi' + \xi'' \in V \oplus V^\perp$ we have by Theorem 5.5.1

$$(\varphi u)\hat{}\,(\xi) = \int_V e^{-ix\xi} \varphi(x) u_0(x)\, dS(x) = \int_V e^{-ix\xi'} \varphi(x) u_0(x)\, dS(x). \tag{6.4.7}$$

If $x \notin \operatorname{supp}(u) \supseteq \operatorname{singsupp}(u)$, then $(x, \xi) \notin \mathrm{WF}(u)$ for all ξ by Theorem 6.4.9 (ii). So let $x \in \operatorname{supp}(u)$ and let Γ be a closed cone in $\mathbb{R}^n \setminus \{0\}$. We distinguish two cases:

a) $\Gamma \cap V^\perp = \emptyset$.

We first show that under this assumption we have

$$\exists c > 0\ \forall \xi \in \Gamma: \ |\xi| \leq c|\xi'|. \tag{6.4.8}$$

Replacing $\xi \in \Gamma$ by $\xi/|\xi|$ and noting that $(\xi/|\xi|)' = \xi'/|\xi|$ we see that it is sufficient to prove (6.4.8) for any $\xi \in \Gamma \cap S^{n-1}$. So suppose that this property is not satisfied, then there exists a sequence ξ_j in $\Gamma \cap S^{n-1}$ with $1 = |\xi_j| > j|\xi_j'|$, implying that $\xi_j' \to 0$.

We have $\xi_j = \xi_j' + \xi_j''$, and since S^{n-1} is compact we may extract subsequences (denoted by the same letters) such that $\xi_j \to \xi_0 \in S^{n-1} \cap \Gamma$, $\xi_j' \to 0$, and $\xi_j'' \to \xi_0''$ (noting that V, $V^\perp$ are closed). Hence $\xi_0 = \xi_0'' \in \Gamma \cap V^\perp$, a contradiction.

Now let $\varphi \in \mathcal{D}(\mathbb{R}^n)$ with $\varphi(x) \neq 0$. Using integration by parts in (6.4.7) we obtain

$$|(\varphi u)\hat{}(\xi)| \le C_N(1+|\xi'|)^{-N} \underset{(6.4.8)}{\le} \tilde{C}_N(1+|\xi|)^{-N}$$

for all $\xi \in \Gamma$. Thus we have shown that $(\varphi u)\hat{}$ is rapidly decreasing in any closed cone that does not intersect $V^\perp$.

b) $\Gamma \cap V^\perp \neq \emptyset$.

Then there exists some $\xi \in \Gamma \cap V^\perp$, and so also $\lambda\xi \in \Gamma \cap V^\perp$ for each $\lambda > 0$. Let $\varphi \in \mathcal{D}(\mathbb{R}^n)$ with $\varphi(x) \neq 0$. Then by (6.4.7),

$$(\varphi u)\hat{}(\lambda\xi) = \int_V e^{-ix\cdot 0}\varphi(x)u_0(x)dS(x) = \langle u, \varphi\rangle =: c_\varphi \neq 0$$

for a suitable φ (since $x \in \operatorname{supp}(u)$). Consequently, $(\varphi u)\hat{}$ is not rapidly decreasing in the direction of ξ, implying that $(x,\xi) \in \mathrm{WF}(u)$ for each $\xi \in \Gamma \cap V^\perp$.

This class of examples underlies Hörmander's motivation for the name 'wave front set': By the Huygens principle a wave always propagates in a direction perpendicular to the tangent space of the wave front.

(iii) By Theorem 6.4.9, any wave front set is closed in $\Omega \times \mathbb{R}^n \setminus \{0\}$ and conic in its second component. One may ask, conversely, which subsets of $\Omega \times \mathbb{R}^n \setminus \{0\}$ can be realized as wave front sets of distributions on Ω. Interestingly, the answer is that *all* such sets are wave front sets of distributions. Indeed, given such a set S, pick a set

$$\{(x_k, \xi_k) \in S \mid k \in \mathbb{N},\ |\xi_k| = 1\}$$

that is dense in $(\mathbb{R}^n \times S^{n-1}) \cap S$, and let $\varphi \in \mathcal{D}(\Omega)$ such that $\int \varphi(x)\,dx = 1$. Then

$$T(x) := \sum_{k=1}^{\infty} \frac{1}{k^2}\varphi(k(x-x_k))e^{ik^3\xi_k} \qquad (x \in \mathbb{R}^n).$$

is even continuous and $\mathrm{WF}(T) = S$. For a proof we refer to [Hör90, 8.14].

Next we address the question of how standard operations on distributions influence the wave front set.

6.4.11 Proposition. *Let $T \in \mathcal{D}'(\Omega)$.*

(i) If $f \in C^\infty(\Omega)$, then

$$\mathrm{WF}(fT) \subseteq \mathrm{WF}(T). \tag{6.4.9}$$

(ii) Let P be a linear PDO with C^∞ coefficients. Then

$$\mathrm{WF}(P(T)) \subseteq \mathrm{WF}(T). \tag{6.4.10}$$

Proof. (i) It suffices to show that, for any $x \in \Omega$ we have $\Sigma_x(fT) \subseteq \Sigma_x(T)$. This, in turn, will follow once we verify that, for any $\varphi \in \mathcal{D}(\Omega)$, $\Sigma(\varphi fT) \subseteq \Sigma(\varphi T)$. Now pick $\psi \in \mathcal{D}(\Omega)$ such that $\psi \equiv 1$ on a neighborhood of $\operatorname{supp}(\varphi)$, so that $\psi\varphi = \varphi$. Then by Lemma 6.4.4,

$$\Sigma(\varphi fT) = \Sigma((\psi f)\varphi T) \subseteq \Sigma(\varphi T).$$

(ii) By (i) it suffices to show the claim for $P = D^\alpha$. Pick nets φ_ι, χ_ι in $\mathcal{D}(\Omega)$, $\varphi_\iota \equiv 1$ near x, $\chi_\iota \equiv 1$ near $\operatorname{supp}\varphi_\iota$, $\operatorname{supp}(\chi_\iota) \to \{x\}$. Then by Lemma 6.4.4,

$$\Sigma_x(D^\alpha T) \subseteq \Sigma(\varphi_\iota D^\alpha T) = \Sigma(\varphi_\iota D^\alpha(\chi_\iota T)) \subseteq \Sigma(D^\alpha(\chi_\iota T))$$

Now for any $S \in \mathcal{E}'(\Omega)$, $(D^\alpha S)\hat{}(\xi) = \xi^\alpha \hat{S}(\xi)$, so $\Sigma(D^\alpha S) \subseteq \Sigma(S)$. Applying this to the above, we get

$$\Sigma_x(D^\alpha T) \subseteq \Sigma(\chi_\iota T),$$

which converges to $\Sigma_x(T)$ by Proposition 6.4.6 (i). □

6.4.12 Remark. More can be said about the transformation of wave front sets by elementary operations on distributions. For the proofs of the following results we refer to [FJ98, Hör90].

(i) Let $S \in \mathcal{D}'(\mathbb{R}^n)$, $T \in \mathcal{D}'(\mathbb{R}^k)$. Then

$$\begin{aligned}\mathrm{WF}(S \otimes T) \subseteq (\mathrm{WF}(S) \times \mathrm{WF}(T)) \cup (\mathrm{WF}(S) \times (\operatorname{supp}(T) \times \{0\}))\\ \cup ((\operatorname{supp}(S) \times \{0\}) \times \mathrm{WF}(T)).\end{aligned}$$

(ii) Let $f : \Omega \to \Omega'$ be a diffeomorphism, and let $T \in \mathcal{D}'(\Omega')$. Then

$$\mathrm{WF}(f^*T) = \{(x, (Df(x))^\top \cdot \eta) \mid (f(x), \eta) \in \mathrm{WF}(T)\}.$$

This means that under coordinate transformations the wave front set transforms like a subset of the cotangent bundle. Based on this observation, one can define the wave front set of a distribution on a manifold M as a subset of T^*M.

(iii) Suppose that $S, T \in \mathcal{D}'(\mathbb{R}^n)$ are such that $\mathrm{WF}(S) + \mathrm{WF}(T)$ (with $+$ the usual sum of subsets, cf. Section 4.2) does not contain any 0-direction. Then one can form $\hat{S} * \hat{T}$ and use this to define the product of S and T by

$$S \cdot T := \mathcal{F}^{-1}(\hat{S} * \hat{T})$$

(which is obviously compatible with the product of smooth functions). Then

$$\mathrm{WF}(S \cdot T) \subseteq \mathrm{WF}(S) \cup \mathrm{WF}(T) \cup (\mathrm{WF}(S) + \mathrm{WF}(T)).$$

(iv) The main application of microlocal analysis is the study of regularity properties of partial differential equations. If P is a PDO of order m with C^∞ coefficients and principal part P_m, the *characteristic set* of P is defined as

$$\mathrm{Char}(P) := \{(x, \xi) \in \mathbb{R}^n \times (\mathbb{R}^n \setminus \{0\}) \mid P_m(x, \xi) = 0\}.$$

Then for any $T \in \mathcal{D}'(\mathbb{R}^n)$, we have the following fundamental result on the *propagation of singularities*

$$\mathrm{WF}(T) \subseteq \mathrm{Char}(P) \cup \mathrm{WF}(P(T))$$

Thus, if T is a solution to the PDE $P(T) = f$, then the singularities of T are contained in the set $\mathrm{Char}(P) \cup \mathrm{WF}(f)$. Note that, if P is elliptic then $\mathrm{Char}(P) = \emptyset$. In this case, due to Proposition 6.4.11 we obtain that $\mathrm{WF}(T) = \mathrm{WF}(P(T))$, i.e., the singularity set is exactly preserved. In particular, by applying pr_1 we confirm the elliptic regularity Theorem 6.2.4, i.e., $\operatorname{singsupp}(P(T)) = \operatorname{singsupp}(T)$.

6.5 Fourier transform on L^2, Sobolev spaces

A standard result (from analysis or measure theory) states that $C_c(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ when $1 \le p < \infty$ (see, e.g., [Fol99, Proposition 7.9]). Using the regularization techniques already seen in Theorem 4.3.1 we may in turn approximate C_c-functions uniformly by test functions in $\mathcal{D}$ and thus conclude in summary that

$$\mathcal{D}(\mathbb{R}^n) \text{ (as well as } \mathcal{S}(\mathbb{R}^n)\text{) is dense in } L^p(\mathbb{R}^n) \quad (1 \le p < \infty). \tag{6.5.1}$$

(Cf. also [Fol99, Proposition 8.17])

6.5.1 Theorem. *(Plancherel) If $f \in L^2(\mathbb{R}^n)$ then the ($\mathcal{S}'$-)Fourier transform $\widehat{f}$ is also in $L^2(\mathbb{R}^n)$. Moreover, Parseval's formula* (5.1.3)*, i.e.,*

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

is valid for all $f, g \in L^2(\mathbb{R}^n)$ and we have

$$\|\widehat{f}\|_{L^2} = (2\pi)^{n/2}\,\|f\|_{L^2}. \tag{6.5.2}$$

Proof. *Step 1:* Let $f, g \in \mathcal{S}(\mathbb{R}^n)$.

Then $\widehat{f}, \widehat{g} \in \mathcal{S}(\mathbb{R}^n)$ and (5.1.3) holds. If we set $g = \overline{\widehat{f}}$ then

$$g(x) = \overline{\widehat{f}(x)} = \int \overline{f(y)} e^{ixy} dy = \widehat{\overline{f}}(-x) = (\widehat{\overline{f}})\check{}\,(x),$$

hence $\widehat{g} = (2\pi)^n\,\overline{f}$ and (5.1.3) implies (6.5.2) in this case.

Step 2: Let $f \in L^2(\mathbb{R}^n)$ and $g \in \mathcal{S}(\mathbb{R}^n)$.

We have $\langle \widehat{f}, g\rangle = \langle f, \widehat{g}\rangle = \int f(x)\,\widehat{g}(x)dx$ and the Cauchy-Schwarz inequality gives

$$|\langle \widehat{f}, g\rangle| = |\langle f, \widehat{g}\rangle| \le \|f\|_{L^2}\|\widehat{g}\|_{L^2} \underset{\substack{\uparrow\\ \text{[Step 1]}}}{=} (2\pi)^{n/2}\,\|f\|_{L^2}\|g\|_{L^2}.$$

Since $\mathcal{S}$ is dense in L^2, the above inequality shows that the linear functional $g \mapsto \langle \widehat{f}, g\rangle$ on $\mathcal{S}(\mathbb{R}^n)$ has a unique continuous extension to $L^2(\mathbb{R}^n)$, which we denote again by $\widehat{f}$. In view of the Fréchet-Riesz theorem ([Wer05, Theorem V.3.6]) there exists a unique $v \in L^2(\mathbb{R}^n)$ such that we have

$$\forall \varphi \in L^2: \quad \langle \widehat{f}, \varphi\rangle = (\varphi, v)_{L^2} = \int \varphi(x)\,\overline{v}(x)\,dx.$$

If $\varphi \in \mathcal{S}$ we obtain $\langle \widehat{f}, \varphi\rangle = \langle \overline{v}, \varphi\rangle$, thus $\overline{v} = \widehat{f}$ holds in $\mathcal{S}'$ and therefore $\widehat{f} \in L^2$ and Parseval's formula $\langle \widehat{f}, g\rangle = \langle f, \widehat{g}\rangle$ is valid with $f, g \in L^2$. Now (6.5.2) follows exactly as in Step 1. □

6.5.2 Corollary. *The linear map $(2\pi)^{-n/2}\,\mathcal{F}\,|_{L^2(\mathbb{R}^n)}$ defines a unitary operator on $L^2(\mathbb{R}^n)$.*

Proof. Replacing f by $\widehat{f}$ and g by $\check{\overline{g}}$ in Parseval's formula yields

$$(\widehat{f}, \widehat{g})_{L^2} = \int \widehat{f}(\xi)\,\overline{\widehat{g}}(\xi)\,d\xi = \int \widehat{f}(\xi)\,\widehat{\check{\overline{g}}}(\xi)\,d\xi \underset{\substack{\uparrow\\ \text{[Parseval]}}}{=} \int \widehat{\widehat{f}}(x)\,\check{\overline{g}}(x)\,dx$$

$$= (2\pi)^n \int \check{f}(x)\,\check{\overline{g}}(x)\,dx = (2\pi)^n\,(f, g)_{L^2},$$

thus the linear map $f \mapsto (2\pi)^{-n/2}\,\widehat{f}$ is an isometry on L^2. Since $f = (2\pi)^{-n/2}\,\mathcal{F}(h)$, where $h := (2\pi)^{-n/2}\,\widehat{f}\,\check{}\, \in L^2$, the map $(2\pi)^{-n/2}\,\mathcal{F}\,|_{L^2(\mathbb{R}^n)}$ is also surjective (as operator on L^2), hence it is unitary. $\square$

Plancherel's Theorem 6.5.1 implies that for $u \in \mathcal{S}'(\mathbb{R}^n)$ we have

$$u \in L^2(\mathbb{R}^n) \iff \hat{u} \in L^2(\mathbb{R}^n).$$

Moreover, by the exchange formulas (Proposition 5.4.3(i),(ii)) we know that differentiation of u amounts to multiplication of $\hat{u}$ with polynomials and vice versa. In this way derivatives of u are linked to growth at infinity of $\hat{u}$. The definition of Sobolev spaces is based on this observation and allows one to measure smoothness of u in terms of L^2-estimates of its Fourier transform. We start by introducing some notation.

Let $s \in \mathbb{R}$ and $\xi \in \mathbb{R}^n$. In the following we shall write

$$\lambda \equiv \lambda(\xi) := (1 + |\xi|^2)^{\frac{1}{2}} (= \langle \xi \rangle) \text{ and hence } \lambda^s \equiv \lambda^s(\xi) = (1 + |\xi|^2)^{\frac{s}{2}}.$$

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

$$H^s(\mathbb{R}^n) := \{u \in \mathcal{S}'(\mathbb{R}^n) : \ \lambda^s \hat{u} \in L^2(\mathbb{R}^n)\}.$$

6.5.4 Remark.

(i) Note that if $u \in H^s(\mathbb{R}^n)$ then by definition $\hat{u}$ is a function.

(ii) From Plancherel's theorem 6.5.1 it follows that $H^0(\mathbb{R}^n) = L^2(\mathbb{R}^n)$.

6.5.5 Proposition.

(i) The spaces $H^s(\mathbb{R}^n)$ are Hilbert spaces with scalar product

$$\langle u|v\rangle_s := (2\pi)^{-n} \int \lambda^{2s}(\xi)\hat{u}(\xi)\bar{\hat{v}}(\xi)\, d\xi \tag{6.5.3}$$

and (associated) norm

$$\|u\|_{H^s}^2 = (2\pi)^{-n} \int \lambda^{2s}(\xi)|\hat{u}(\xi)|^2 d\xi. \tag{6.5.4}$$

(ii) For all $s \in \mathbb{R}$ we have that $\mathcal{S}(\mathbb{R}^n) \subseteq H^s(\mathbb{R}^n)$ is dense.

Note: The factor $(2\pi)^{-n}$ in (6.5.3) was introduced to have

$$\|u\|_{L^2} = \|u\|_{H^0},$$

cf. (6.5.2) in Plancherel's theorem. Moreover, we have

$$\|u\|_{H^s} = (2\pi)^{-\frac{n}{2}}\|\lambda^s \hat{u}\|_{L^2}.$$

Proof. (i) The scalar product exists by the Cauchy-Schwarz inequality and the definition of H^s. Indeed we have

$$\int |\lambda^s \hat{u}\ \lambda^s \bar{\hat{v}}| \le \|\lambda^s \hat{u}\|_{L^2}\|\lambda^s \hat{v}\|_{L^2}.$$

Moreover, sesquilinearity and non-negativity is clear. To show positive definiteness assume that $\langle u|u\rangle_s = 0$. Then we have

$$\begin{aligned}\int \lambda^{2s}|\hat{u}|^2 = 0 \quad &\Longrightarrow \quad \hat{u}(\xi) = 0 \text{ a.e.} \Longrightarrow \hat{u} = 0 \in L^2(\mathbb{R}^n)\\ &\underset{\substack{\uparrow\\ 5.3.6(\text{ii})}}{\Longrightarrow} \quad \hat{u} = 0 \in \mathcal{S}'(\mathbb{R}^n) \Longrightarrow u = 0 \in H^s(\mathbb{R}^n).\end{aligned}$$

To show completeness of H^s, let $(u_j)_j$ be a Cauchy sequence in $H^s(\mathbb{R}^n)$. Then $\lambda^s \hat{u}_j$ is a Cauchy sequence in $L^2(\mathbb{R}^n)$, hence is convergent. Let $v \in L^2(\mathbb{R}^n)$ be its limit and set $u := \mathcal{F}^{-1}(v(\xi)\lambda(\xi)^{-s})$. Then $\hat{u}\lambda^s = v \in L^2(\mathbb{R}^n)$, so $u \in H^s(\mathbb{R}^n)$, and

$$\begin{aligned}(2\pi)^{n/2}\|u_j - u\|_{H^s} &= \Big[\int \lambda^{2s}(\xi)|\hat{u}_j - \hat{u}|^2(\xi)\, d\xi\Big]^{\frac{1}{2}} = \Big[\int \big|\lambda^s(\xi)\hat{u}_j(\xi) - v(\xi)\big|^2 d\xi\Big]^{\frac{1}{2}}\\ &= \|\lambda^s\hat{u}_j - v\|_{L^2} \to 0 \qquad (j \to \infty).\end{aligned}$$

(ii) Clearly, $\mathcal{S}(\mathbb{R}^n) \subseteq H^s(\mathbb{R}^n)$ for all s: By Theorems 5.2.5(i) and 5.2.10, $u \in \mathcal{S}(\mathbb{R}^n)$ implies $\lambda^s\hat{u} \in \mathcal{S}(\mathbb{R}^n) \subseteq L^2(\mathbb{R}^n)$.
To show denseness, let $u \in H^s(\mathbb{R}^n)$. By (6.5.1) there exists $(\varphi_j)_j \in \mathcal{D}(\mathbb{R}^n)$ with

$$\varphi_j \to \lambda^s\hat{u} \text{ in } L^2(\mathbb{R}^n). \tag{6.5.5}$$

Now set $\psi_j := \mathcal{F}^{-1}(\underbrace{\lambda^{-s}\varphi_j}_{\in\mathcal{D}\subseteq\mathcal{S}}) \in \mathcal{S}(\mathbb{R}^n)$. Then we obtain

$$\begin{aligned}\|u - \psi_j\|_{H^s} &= (2\pi)^{-\frac{n}{2}}\left(\int \lambda^{2s}(\xi)|\hat{u}(\xi) - \lambda^{-s}\varphi_j(\xi)|^2 d\xi\right)^{\frac{1}{2}}\\ &= (2\pi)^{-\frac{n}{2}}\left(\int |\lambda^s(\xi)\hat{u}(\xi) - \varphi_j(\xi)|^2 d\xi\right)^{\frac{1}{2}} \to 0,\end{aligned}$$

where convergence is due to (6.5.5). □

6.5.6 Example. We have

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

Indeed,

$$\begin{aligned}\delta \in H^{-s}(\mathbb{R}^n) \quad &\iff \quad \lambda^{-s}\hat{\delta} \in L^2(\mathbb{R}^n) \overset{\substack{\hat{\delta}=1\\ \downarrow}}{\iff} \lambda^{-s}(\xi) \le C(1+|\xi|)^{-s} \in L^2(\mathbb{R}^n)\\ &\iff \quad \int\limits_{R^n} \frac{d\xi}{(1+|\xi|)^{2s}} < \infty \iff 2s > n.\end{aligned}$$

6.5.7 Proposition.

(i) For $s \ge t$ we have $H^s(\mathbb{R}^n) \hookrightarrow H^t(\mathbb{R}^n)$ continuously.

(ii) Let $P(D)$ be a linear PDO with constant coefficients of order m, then

$$P(D):\ H^s(\mathbb{R}^n) \to H^{s-m}(\mathbb{R}^n)$$

is continuous.

Proof. (i) Let $s \geq t$, then

$$\begin{aligned}
\|u\|_{H^t} &= (2\pi)^{-\frac{n}{2}} \left(\int |\lambda^t(\xi)\hat{u}(\xi)|^2 d\xi \right)^{\frac{1}{2}} \\
&= (2\pi)^{-\frac{n}{2}} \Big(\int |\underbrace{(1+|\xi|^2)^{\frac{t-s}{2}}}_{\leq 1} \lambda^s(\xi)\, \hat{u}(\xi)|^2 d\xi \Big)^{\frac{1}{2}} \leq \|u\|_{H^s}.
\end{aligned}$$

(ii) We prove the statement for $P(D) = D^\alpha$, the general case follows analogously. Let $u \in H^s(\mathbb{R}^n)$, then by the exchange formula (Prop. 5.4.3(i)) we have

$$\begin{aligned}
\lambda^{s-m}(\xi)\, |\widehat{D^\alpha u}| &\leq (1+|\xi|^2)^{\frac{s-m}{2}}\, |\xi|^m\, |\hat{u}(\xi)| \\
&\leq (1+|\xi|^2)^{\frac{s-m}{2}}\, (1+|\xi|^2)^{\frac{m}{2}}\, |\hat{u}(\xi)| \\
&= (1+|\xi|^2)^{\frac{s}{2}}\, |\hat{u}(\xi)| = \lambda^s(\xi)\, |\hat{u}(\xi)|,
\end{aligned}$$

so $D^\alpha u \in H^{s-m}(\mathbb{R}^n)$ and $\|D^\alpha u\|_{H^{s-m}} \leq \|u\|_{H^s}$. □

6.5.8 Remark. Due to Proposition 6.5.7(i) it makes sense to introduce the spaces

$$H^\infty(\mathbb{R}^n) := \bigcap_{s\in\mathbb{R}} H^s(\mathbb{R}^n) \text{ and } H^{-\infty}(\mathbb{R}^n) := \bigcup_{s\in\mathbb{R}} H^s(\mathbb{R}^n).$$

We immediately see that we have the inclusions

$$\mathcal{S}(\mathbb{R}^n) \subseteq H^\infty(\mathbb{R}^n) \subseteq H^{-\infty}(\mathbb{R}^n) \subseteq \mathcal{S}'(\mathbb{R}^n).$$

In fact all these inclusions are strict: $(1+|x|^2)^{-n} \in H^\infty(\mathbb{R}^n)$ by Theorem 6.5.13 below but not in $\mathcal{S}(\mathbb{R}^n)$ and $1 \in \mathcal{S}'(\mathbb{R}^n) \setminus H^{-\infty}(\mathbb{R}^n)$.

6.5.9 Remark. Let $\varphi, \psi \in \mathcal{S}(\mathbb{R}^n)$ and regard φ as a regular $\mathcal{S}'$-distribution. Then we have by Corollary 6.5.2

$$\begin{aligned}
\langle \varphi, \psi \rangle = \int \varphi(x)\psi(x)\, dx &= (\psi, \bar{\varphi})_{L^2} = (2\pi)^{-n}(\hat{\psi}, \hat{\bar{\varphi}})_{L^2} \\
\underset{\hat{\bar{\bar{\varphi}}}(\xi)=\hat{\varphi}(-\xi)}{\overset{\uparrow}{=}} &(2\pi)^{-n} \int \hat{\psi}(\xi)\hat{\varphi}(-\xi)\, d\xi = (2\pi)^{-n} \int \lambda^{-s}(\xi)\hat{\psi}(\xi)\, \lambda^s(\xi)\hat{\varphi}(-\xi)\, d\xi,
\end{aligned}$$

hence by the Cauchy-Schwarz inequality

$$|\langle \varphi, \psi \rangle| \leq (2\pi)^{-n} \int |\hat{\psi}(\xi)\hat{\varphi}(-\xi)| d\xi \leq \|\psi\|_{H^{-s}} \|\varphi\|_{H^s}. \tag{6.5.6}$$

Since $\mathcal{S}(\mathbb{R}^n)$ is dense in $H^s(\mathbb{R}^n)$ for all s (Proposition 6.5.5(ii)) we may extend the map

$$(\varphi, \psi) \mapsto \langle \varphi, \psi \rangle$$

uniquely to a continuous bilinear map $H^{-s}(\mathbb{R}^n) \times H^s(\mathbb{R}^n) \to \mathbb{C}$ which we write as

$$(u, v) \mapsto \langle u, v \rangle_{H^{-s}, H^s} := (2\pi)^{-n} \int \hat{u}(\xi)\hat{v}(-\xi)\, d\xi. \tag{6.5.7}$$

Note that (6.5.6) gives

$$|\langle u, v \rangle_{H^{-s}, H^s}| \leq \|u\|_{H^{-s}} \|v\|_{H^s}. \tag{6.5.8}$$

6.5.10 Theorem. *The bilinear form $\langle\ ,\ \rangle_{H^{-s},H^s}$ of (6.5.7) induces an isometric isomorphism*

$$H^{-s}(\mathbb{R}^n) \to \big(H^s(\mathbb{R}^n)\big)' \quad \textit{(the topological dual of } H^s\textit{)}.$$

In other words, $H^{-s}(\mathbb{R}^n)$-distributions are precisely the continuous linear forms on $H^s(\mathbb{R}^n)$.

Proof. By (6.5.8), for any $u \in H^{-s}(\mathbb{R}^n)$, the map $\Phi_u : v \mapsto \langle u, v\rangle_{H^{-s},H^s}$ is a continuous linear form on $H^s(\mathbb{R}^n)$ with $\|\Phi_u\| \le \|u\|_{H^{-s}}$. Now set

$$v_0 := \mathcal{F}^{-1}(\lambda^{-2s}(\xi)\hat{u}(-\xi)).$$

Then since $u \in H^{-s}(\mathbb{R}^n)$, $\lambda^s \hat{v}_0 = \lambda^{-s}\hat{u}(-\xi) \in L^2(\mathbb{R}^n)$, meaning that $v_0 \in H^s(\mathbb{R}^n)$. Furthermore, by (6.5.7),

$$\langle u, v_0\rangle_{H^{-s},H^s} = (2\pi)^{-n} \int |\hat{u}(\xi)|^2 \lambda^{-2s}(\xi)\, d\xi = \|u\|^2_{H^{-s}}.$$

We have

$$\|v_0\|_{H^s} = (2\pi)^{-n/2} \Big[\int \lambda^{2s}(\xi)\lambda^{-4s}(\xi)|\hat{u}(\xi)|^2\, d\xi\Big]^{1/2} = \|u\|_{H^{-s}}.$$

Thus $\tilde{v}_0 := v_0/\|v_0\|_{H^s}$ satisfies $\|\tilde{v}_0\|_{H^s} = 1$ and $\Phi_u(\tilde{v}_0) = \|u\|_{H^{-s}}$. This shows that $\|\Phi_u\| = \|u\|_{H^{-s}}$, so that

$$\begin{gathered} H^{-s}(\mathbb{R}^n) \to H^s(\mathbb{R}^n)' \\ u \mapsto \Phi_u \end{gathered}$$

is an isometry. It therefore only remains to show that this map is surjective. So let $u' \in H^s(\mathbb{R}^n)'$. Then since $H^s(\mathbb{R}^n)$ is a Hilbert space, by the Riesz-Fréchet theorem there exists some $w \in H^s(\mathbb{R}^n)$ with

$$u'(v) = \langle v|w\rangle_s = (2\pi)^{-n} \int \lambda^{2s}(\xi)\hat{v}(\xi)\overline{\hat{w}(\xi)}\, d\xi.$$

Now set $u := \mathcal{F}^{-1}(\overline{\hat{w}(-\xi)}\lambda^{2s}(\xi))$. Then $u \in H^{-s}(\mathbb{R}^n)$ and $u'(v) = \langle u, v\rangle_{H^{-s},H^s} = \Phi_u(v)$ for all $v \in H^s(\mathbb{R}^n)$, i.e., $u' = \Phi_u$. □

6.5.11 Remark. Do not be confused by the fact that (as is the case for any Hilbert space) $(H^s)'$ is also isometrically isomorphic to H^s itself. This isomorphism is induced by the mapping $\langle\ |\ \rangle_s$ rather than $\langle\ ,\ \rangle_{H^{-s},H^s}$.

Composing these two mappings we obtain an isometric isomorphism from $H^s(\mathbb{R}^n)$ to $H^{-s}(\mathbb{R}^n)$ which is essentially given by the (Pseudo-differential) operator $\lambda^{2s}(D)$ defined via $\mathcal{F}(\lambda^{2s}(D)u) := \lambda^{2s}\hat{u}$.

Our next task is to show that Sobolev spaces consist of functions whose derivatives belong to L^2. An overall understanding of this statement is best reached via the use of Pseudo-differential operators. Since this is beyond the focus of the present course we will restrict ourselves to the case of the spaces $H^m(\mathbb{R}^n)$ with $m \in \mathbb{N}_0$. We start with a little technical Lemma, which, however is easily proved also in the general case $s \in \mathbb{R}$.

6.5.12 Lemma. *For all $s \in \mathbb{R}$ we have*

$$u \in H^{s+1}(\mathbb{R}^n) \iff u, D_1 u, \ldots, D_n u \in H^s(\mathbb{R}^n)$$

and in this case the norms satisfy the equality

$$\|u\|^2_{H^{s+1}} = \|u\|^2_{H^s} + \sum_{j=1}^n \|D_j u\|^2_{H^s}.$$

Proof. We have $\lambda^2(\xi) = 1 + |\xi|^2 = 1 + \sum_j \xi_j^2$ and so again by the exchange formula Proposition 5.4.3 (i)

$$|\lambda^{s+1}\hat{u}|^2 = \lambda^2|\lambda^s\hat{u}|^2 = |\lambda^s\hat{u}|^2 + \sum_{j=1}^n |\lambda^s\xi_j\hat{u}|^2 = |\lambda^s\hat{u}|^2 + \sum_{j=1}^n |\lambda^s\widehat{D_j u}|^2.$$

$\square$

This leads to the following characterization of H^m for $m \in \mathbb{N}_0$:

6.5.13 Theorem. *Let $m \in \mathbb{N}_0$, then we have*

$$H^m(\mathbb{R}^n) = \{u \in \mathcal{S}'(\mathbb{R}^n) : \ D^\alpha u \in L^2(\mathbb{R}^n) \text{ for all } \alpha \le m\}.$$

Furthermore, $H^m(\mathbb{R}^n)$ is the completion of $\mathcal{D}(\mathbb{R}^n)$ w.r.t. the norm

$$\|\varphi\|^{(m)} := \Big(\int \sum_{|\alpha|\le m} |D^\alpha\varphi(x)|^2 dx\Big)^{\frac{1}{2}}.$$

Proof. We prove the first assertion by induction. The case $m = 0$ is clear from 6.5.4(ii) (resp. Plancherel's Theorem). The inductive step is due to Lemma 6.5.12, since

$$u \in H^{m+1}(\mathbb{R}^n) \overset{\substack{6.5.12\\ \downarrow}}{\Longleftrightarrow} u, D_j u \in H^m(\mathbb{R}^n)\ \forall 1 \le j \le n \overset{\substack{\text{Ind. hyp.}\\ \downarrow}}{\Longleftrightarrow} D^\alpha u \in L^2(\mathbb{R}^n)\ \forall |\alpha| \le m+1.$$

We now show that the completion of $(\mathcal{D}(\mathbb{R}^n), \|\ \|^{(m)})$ is $H^m(\mathbb{R}^n)$.

$\boxed{\subseteq}$ Let $(\varphi_j)_j$ be a Cauchy sequence in $\mathcal{D}(\mathbb{R}^n)$ w.r.t. $\|\ \|^{(m)}$. Then $(D^\alpha\varphi_j)_j$ is a Cauchy sequence in $L^2(\mathbb{R}^n)$ for all $|\alpha| \le m$ and so there exist $u_\alpha \in L^2(\mathbb{R}^n)$ with

$$D^\alpha\varphi_j \longrightarrow u_\alpha \text{ in } L^2(\mathbb{R}^n)$$

and we claim $\varphi_j \to u_0$ w.r.t. $\|\ \|^{(m)}$. Hence we have to show that $\|D^\alpha\varphi_j - D^\alpha u_0\|_{L^2} \to 0\ \forall|\alpha| \le m$. To do so it suffices to show $D^\alpha u_0 = u_\alpha\ \forall|\alpha| \le m$ since then $u_0 \in H^m(\mathbb{R}^n)$ by (i). We have for all $\psi \in \mathcal{D}(\mathbb{R}^n)$ (where convergence in both cases is due to the Cauchy-Schwarz inequality)

$$\int \psi D^\alpha\varphi_j \to \int \psi u_\alpha \quad \text{and}$$

$$\int \psi D^\alpha\varphi_j = (-1)^{|\alpha|}\int \varphi_j D^\alpha\psi \to (-1)^{|\alpha|}\int u_0 D^\alpha\psi = \int D^\alpha u_0\psi.$$

So we obtain $\int(D^\alpha u_0 - u_\alpha)\psi = 0\ \forall\psi \in \mathcal{D}(\mathbb{R}^n)$ which establishes the claim due to denseness of $\mathcal{D}(\mathbb{R}^n)$ in $L^2(\mathbb{R}^n)$ (cf. Remark 6.5.1).

$\boxed{\supseteq}$ Let $u \in H^m(\mathbb{R}^n)$. Then by (i) $D^\alpha u \in L^2(\mathbb{R}^n)$ for all $|\alpha| \le m$. We proceed by smoothing u: Let ρ be a mollifier and set $u_\varepsilon := u * \rho_\varepsilon$. Then by the standard results on smoothing (see 4.2.4(ii) and e.g. [Fol99, Theorem 8.14]) we have

$$D^\alpha u_\varepsilon = D^\alpha(u * \rho_\varepsilon) = (D^\alpha u) * \rho_\varepsilon \text{ and } \|D^\alpha u_\varepsilon - D^\alpha u\|_{L^2} \to 0\ \forall|\alpha| \le m. \quad (6.5.9)$$

Let now $\varphi \in \mathcal{D}(\mathbb{R}^n)$, $0 \le \varphi \le 1$ and $\varphi \equiv 1$ on $\overline{B_1(0)}$ and set $g_\varepsilon(x) := \varphi(\varepsilon x)u_\varepsilon(x)$. Then $g_\varepsilon \in \mathcal{D}(\mathbb{R}^n)$ and we have

$$\begin{aligned} D^\alpha(g_\varepsilon - u) &= \varphi(\varepsilon.)\big(D^\alpha u_\varepsilon - D^\alpha u\big) + \overbrace{\big(\varphi(\varepsilon.) - 1\big)D^\alpha u}^{\to 0 \text{ in } L^2} \\ &\quad + \underbrace{D^\alpha\big(\varphi(\varepsilon.)u_\varepsilon\big) - \varphi(\varepsilon.)D^\alpha u_\varepsilon} \\ &\qquad = \sum_{0<\beta\le\alpha}\binom{\alpha}{\beta}\underbrace{\varepsilon^{|\beta|}}_{\to 0}\underbrace{(D^\beta\varphi)(\varepsilon.)}_{\|\ \|_{L^\infty}<\infty}\underbrace{D^{\alpha-\beta}u_\varepsilon}_{\|\ \|_{L^2}<\infty}. \end{aligned}$$

So by (6.5.9) $D^\alpha g_\varepsilon \to D^\alpha u$ in $L^2(\mathbb{R}^n)$ $\forall |\alpha| \le m$, hence $\|g_\varepsilon - u\|^{(m)} \to 0$. $\square$

6.5.14 Corollary. *Let $m \in \mathbb{N}_0$.*

(i) The norms $\| \ \|^{(m)}$ and $\| \ \|_{H^m}$ are equivalent on $H^m(\mathbb{R}^n)$.

(ii) Let $u \in H^{-m}(\mathbb{R}^n)$. Then there exists $f_\alpha \in L^2(\mathbb{R}^n)$ ($|\alpha| \le m$) such that

$$u = \sum_{|\alpha| \le m} D^\alpha f_\alpha.$$

Proof. (i) Let $(u_j)_j \in H^m(\mathbb{R}^n)$. Then we have

$$\begin{aligned} u_j \to 0 \quad \text{w.r.t. } \| \ \|_{H^s} &\iff \xi^\alpha \hat{u}_j \to 0 \text{ in } L^2 \text{ for all } |\alpha| \le m \\ &\iff \mathcal{F}(D^\alpha u_j) \to 0 \text{ in } L^2 \text{ for all } |\alpha| \le m \\ &\iff D^\alpha u_j \to 0 \text{ in } L^2 \text{ for all } |\alpha| \le m \\ &\iff u_j \to 0 \text{ w.r.t. } \| \ \|^{(m)}. \end{aligned}$$

(ii) Let $u \in H^{-m}(\mathbb{R}^n)$. Then $(1+|\xi|^2)^{-\frac{m}{2}}\hat{u} \in L^2(\mathbb{R}^n)$ and so $\hat{g}(\xi) := \hat{u}(\xi)\big(1 + \sum_{j=1}^n |\xi_j|^m\big)^{-1} \in L^2(\mathbb{R}^n)$ and finally

$$\hat{u}(\xi) = \hat{g}(\xi) + \sum_{j=1}^n |\xi_j|^m \hat{g}(\xi) = \hat{g}(\xi) + \sum_{j=1}^n \xi_j^m \underbrace{\frac{|\xi_j|^m}{\xi_j^m}\,\hat{g}(\xi)}_{\in L^2}.$$

Hence the assertion follows by applying $\mathcal{F}^{-1}$. $\square$

One of the most useful features of Sobolev spaces is also connected with the fact that Sobolev norms measure smoothness. Indeed if we suppose the Sobolev order, i.e., s in $H^s(\mathbb{R}^n)$ to be high enough as compared to the dimension n of the space, then the functions are actually continuous and vanish at infinity. This means in the context of the H^m-spaces: if one can prove the L^2-property of sufficiently many orders of derivatives one in fact gains regularity. To prove this statement we need two results from the classical theory of the Fourier transform, the Lemma of Riemann-Lebesgue and the Fourier inversion formula for L^1-functions.

6.5.15 Lemma. *(Classical Fourier inversion formula) Let $g \in L^1(\mathbb{R}^n)$. then*

$$(\mathcal{F}^{-1}g)(x) = (2\pi)^{-n} \int e^{ix\xi} g(\xi)\, d\xi,$$

which is a continuous and bounded function.

Proof. Since $g \in L^1(\mathbb{R}^n) \subseteq \mathcal{S}'(\mathbb{R}^n)$ (Remark 5.3.6 (ii)) we conclude from Theorem 5.4.2 that $\mathcal{F}^{-1}g =: f \in \mathcal{S}'(\mathbb{R}^n)$. So for $\varphi \in \mathcal{S}(\mathbb{R}^n)$ we obtain

$$\begin{aligned} \langle f, \check{\varphi} \rangle &\overset{\overset{(5.4.4)}{\downarrow}}{=} (2\pi)^{-n} \langle \hat{f}, \hat{\varphi} \rangle \overset{\overset{\hat{f}=g\in L^1}{\downarrow}}{=} (2\pi)^{-n} \int g(\xi)\hat{\varphi}(\xi)\, d\xi \\ &\overset{\overset{\varphi\in\mathcal{S}}{\downarrow}}{=} (2\pi)^{-n} \int g(\xi) \int \varphi(x) e^{-ix\xi}\, dx\, d\xi \\ &\underset{\underset{x\mapsto -x}{\uparrow}}{\overset{\overset{\text{Fubini}}{\downarrow}}{=}} (2\pi)^{-n} \int\int g(\xi) e^{ix\xi}\, d\xi\, \check{\varphi}(x)\, dx, \end{aligned}$$

which shows that $(\mathcal{F}^{-1}g)(x) = f(x) = (2\pi)^{-n} \int g(\xi) e^{ix\xi} d\xi$. $\square$

6.5.16 Lemma. *(Riemann-Lebesgue) If $f \in L^1(\mathbb{R}^n)$ then $\hat{f} \in C_0^0(\mathbb{R}^n)$ (where $C_0^0 = \{f \in C^0(\mathbb{R}^n) | \lim_{|x|\to\infty} f(x) = 0\}$ is the space of continuous functions vanishing at infinity).*

Proof. In view of Theorem 5.1.1 (i) we only have to show that $\hat{f}$ vanishes at infinity. This is elementary for f being the characteristic function of a rectangle. Indeed for $n = 1$ we have

$$\hat{f}(\xi) = \int_a^b e^{-ix\xi}\, dx = \frac{e^{-ia\xi} - e^{-ib\xi}}{i\xi} \to 0 \qquad (|\xi| \to \infty)$$

and the general case is analogous. The case of a general $f \in L^1(\mathbb{R}^n)$ follows since the characteristic functions of rectangles are a total set in $L^1(\mathbb{R}^n)$ (i.e., their finite linear combinations are dense). □

6.5.17 Theorem. *(Sobolev embedding theorem)*

(i) If $s > \frac{n}{2}$ then $H^s(\mathbb{R}^n) \subseteq C_0^0(\mathbb{R}^n)$.

(ii) If $s > k + \frac{n}{2}$ then $H^s(\mathbb{R}^n) \subseteq C_0^k(\mathbb{R}^n)$ $(= \{f \in C^k(\mathbb{R}^n) | \lim_{|x|\to\infty} \partial^\alpha f(x) = 0\ \forall |\alpha| \leq k\})$.

Proof. (i) Let $u \in H^s(\mathbb{R}^n)$ with $s > n/2$. Note that $\xi \mapsto (1 + |\xi|^2)^{-s} \in L^1(\mathbb{R}^n)$ and we set $f = \lambda^s \hat{u}$ which by definition is in $L^2(\mathbb{R}^n)$ with $\|f\|_{L^2} = (2\pi)^{n/2}\|u\|_{H^s}$. So we obtain using the Cauchy-Schwarz inequality

$$\|\hat{u}\|_{L^1} \leq \|f\|_{L^2}\Big(\int \underbrace{(1 + |\xi|^2)^{-s}}_{\in L^1} d\xi\Big)^{\frac{1}{2}} \leq C\|f\|_{L^2} \leq C\|u\|_{H^s}$$

hence $\hat{u} \in L^1(\mathbb{R}^n)$. So Lemma 6.5.15 tells us that $u = \mathcal{F}^{-1}\hat{u}$, where $\mathcal{F}^{-1}$ is the *classical* inverse Fourier transform. Finally by Lemma 6.5.16 (applied to $\mathcal{F}^{-1}$) u is in C_0^0.

(ii) Let $u \in H^s(\mathbb{R}^n)$ with $s > k + n/2$. Then by Proposition 6.5.7(i),(ii) $D^\alpha u \in H^{s-k}(\mathbb{R}^n)$ for all $|\alpha| \leq k$ and by (i) $D^\alpha u \in C_0^0(\mathbb{R}^n)$ for these α. □

6.5.18 Corollary. *If $u \in H^\infty(\mathbb{R}^n)$ then $u \in C_0^\infty(\mathbb{R}^n)$.*

6.5.19 Remark. One can show that

$$u \in H^s(\mathbb{R}^n),\ \varphi \in \mathcal{S}(\mathbb{R}^n) \implies \varphi u \in H^s(\mathbb{R}^n)$$

with the map $u \mapsto \varphi u$ being continuous on $H^s(\mathbb{R}^n)$. This result tells us that PDOs with $\mathcal{S}$-coefficients operate continuously on the scale of Sobolev spaces. A proof involves Young's inequality for $p = 1$ and $q = 2$ (hence $r = 2$) and *Petree's inequality* which can be proven by elementary means and reads

$$\left(\frac{1 + |\xi|^2}{1 + |\eta|^2}\right)^t \leq 2^{|t|}(1 + |\xi - \eta|^2)^{|t|}$$

for $t \in \mathbb{R}$ and $\xi, \eta \in \mathbb{R}^n$. For details see [FJ98, p. 125].

The theory of Sobolev spaces is vast and has many applications in the theory of PDE, see e.g. [Fol95, Chapter 6] for a start. One striking feature is *Rellich's theorem* which states that under certain conditions the embedding $H^s \hookrightarrow H^t$ $(s > t)$ is compact, hence from any bounded sequence in H^s one may extract an H^t-converging subsequence — an argument which is frequently used in existence proofs in PDE.

A standard reference on Sobolev spaces, with emphasis put on the L^p-based spaces of integer order, i.e.,

$$W^{m,p}(\Omega) := \{u \in L^p(\Omega) : \; D^\alpha u \in L^p(\Omega) \text{ for all } |\alpha| \leq m\}$$

for $m \in \mathbb{N}_0$ and $1 \leq p \leq \infty$ is [AF03].

Bibliography

[AF03] R. Adams and J.J.F. Fournier. *Sobolev Spaces.* Elsevier, Oxford, second edition, 2003.

[Bae10] Christian Baer. *Elementary differential geometry.* Cambridge University Press, Cambridge, 2010. Translated from the 2001 German original by P. Meerkamp.

[FJ98] G. Friedlander and M. Joshi. *Introduction to the theory of distributions.* Cambridge University Press, New York, second edition, 1998.

[Fol95] G. B. Folland. *Introduction to partial differential equations.* Princeton University Press, Princeton, New Jersey, second edition, 1995.

[Fol99] G. B. Folland. *Real Analysis.* John Wiley and Sons, New York, 1999.

[For84] O. Forster. *Analysis 3.* Vieweg Verlag, Wiesbaden, 1984. 3. Auflage.

[For05] O. Forster. *Analysis 2.* Vieweg Verlag, Wiesbaden, 2005. 6. Auflage.

[Hor66] J. Horváth. *Topological vector spaces and distributions.* Addison-Wesley, Reading, MA, 1966.

[Hör90] L. Hörmander. *The analysis of linear partial differential operators*, volume I. Springer-Verlag, second edition, 1990.

[Hör09] G. Hörmann. Analysis (Lecture notes, University of Vienna). available electronically at http://www.mat.univie.ac.at/gue/material.html, 2008-09.

[HS09] G. Hörmann and R. Steinbauer. Theory of distributions. University Lecture, 2009.

[Kun08] M. Kunzinger. Differential geometry 1. available electronically at https://www.mat.univie.ac.at/ mike/teaching/ss08/dg.pdf, 2008.

[New74] D. J. Newman. Fourier uniqueness via complex variables. *Amer. Math. Monthly*, 81:379–380, 1974.

[OW15] N. Ortner and P. Wagner. *Fundamental Solutions of Linear Partial Differential Operators – Theory and Practice.* Berlin, 2015.

[SJ95] L. A. Steen and J. A. Seebach Jr. *Counterexamples in topology.* Dover Publications Inc., Mineola, NY, 1995. Reprint of the second (1978) edition.

[Wag09] P. Wagner. A new constructive proof of the Malgrange-Ehrenpreis theorem. *Amer. Math. Monthly*, 116:457–462, 2009.

[Wer05] D. Werner. *Funktionalanalysis.* Springer-Verlag, Berlin, 2005. fünfte Auflage.

Index