Distribution Theory and Applications — Summary

Lucas Riedstra

November 23, 2020

Contents

0 Most important definitions

Spaces of test functions a function $f \in X \to \mathbb{C}$ is in:

1. $\mathcal{D}(X)$ if f is smooth and $\operatorname{supp} f \subseteq X$ is compact;
2. $\mathcal{S}(\mathbb{R}^n)$ if f is smooth and $\|f\|_{\alpha,\beta} := \sup_{x\in\mathbb{R}^n} \left|x^\alpha (D^\beta f)(x)\right| < \infty$ for all α, β;
3. $\mathcal{E}(X)$ if f is smooth.

Note $\mathcal{D}(\mathbb{R}^n) \subseteq \mathcal{S}(\mathbb{R}^n) \subseteq \mathcal{E}(\mathbb{R}^n)$. We have the following modes of convergence in these spaces:

1. $\varphi_m \to 0$ in $\mathcal{D}(X)$ if there exists a compact $K \subseteq X$ with $\operatorname{supp} \varphi_m \subseteq K$ for all m, and $\partial^\alpha \varphi_m \to 0$ uniformly for each α;
2. $\varphi_m \to 0$ in $\mathcal{S}(\mathbb{R}^n)$ if $\|\varphi_m\|_{\alpha,\beta} \to 0$ for all α, β;
3. $\varphi_m \to 0$ in $\mathcal{E}(X)$ if $\partial^\alpha \varphi_m \to 0$ uniformly on compact subsets of X for all α.

Spaces of distributions The continuous linear maps from $\mathcal{D}(X)$, $\mathcal{S}(X)$, and $\mathcal{E}(X)$ to $\mathbb{C}$ are called *distributions*, *tempered distributions*, and *compactly supported distributions* respectively, and these spaces are denoted $\mathcal{D}'(X)$, $\mathcal{S}'(\mathbb{R}^n)$, $\mathcal{E}'(X)$, employed with weak-$*$ convergence. Note that $\mathcal{E}'(\mathbb{R}^n) \subseteq \mathcal{S}'(\mathbb{R}^n) \subseteq \mathcal{D}'(\mathbb{R}^n)$. We have the following characterisations:

1. $u \in \mathcal{D}'(X)$ iff for every compact $K \subseteq X$ there exist non-negative C, N such that for all $\varphi \in \mathcal{D}(X)$ with $\operatorname{supp}(\varphi) \subseteq K$ we have
$$|\langle u, \varphi\rangle| \leqslant C \sum_{|\alpha|\leqslant N} \sup_x |\partial^\alpha \varphi(x)|.$$
2. $u \in \mathcal{S}'(\mathbb{R}^n)$ iff there exist constants C, N such that for all $\varphi \in \mathcal{S}(\mathbb{R}^n)$ we have
$$|\langle u, \varphi\rangle| \leqslant C \sum_{|\alpha|,|\beta|\leqslant N} \|\varphi\|_{\alpha,\beta}.$$
3. $u \in \mathcal{E}'(X)$ iff there exists a compact $K \subseteq X$ and non-negative C, N such that
$$|\langle u, \varphi\rangle| \leqslant C \sum_{|\alpha|\leqslant N} \sup_{x\in K} |\partial^\alpha \varphi(x)|.$$

Basic operations We define the following basic operations:

1. if f is smooth, then $\langle fu, \varphi\rangle := \langle u, f\varphi\rangle$ (for Schwarz functions, we must have that $f\varphi \in \mathcal{S}(\mathbb{R}^n)$ for all $\varphi \in \mathcal{S}(\mathbb{R}^n)$);
2. For a distribution u: $\langle \partial^\alpha u, \varphi\rangle := (-1)^{|\alpha|}\langle u, \partial^\alpha \varphi\rangle$;
3. For a test function φ we define $(\tau_h\varphi)(x) = \varphi(x-h)$, and for a distribution u we then define $\langle \tau_h u, \varphi\rangle := \langle u, \tau_{-h}\varphi\rangle$.
4. For a test function φ we define $\mathcal{R}[\varphi](x) = \check{\varphi}(x) := \varphi(-x)$, and for a distribution u we then define $\langle \mathcal{R}[u], \varphi\rangle = \langle \check{u}, \varphi\rangle := \langle u, \check{\varphi}\rangle$.

Convolution

1. For $u \in C^\infty(X)$, $\varphi \in \mathcal{D}(X)$, we define
$$(u * \varphi)(x) := \int_{\mathbb{R}^n} u(x-y)\varphi(y)\,\mathrm{d}y = \int_{\mathbb{R}^n} u(y)\varphi(x-y)\,\mathrm{d}y = \langle u, \tau_x\check{\varphi}\rangle.$$

2. For $u \in \mathcal{D}'(\mathbb{R}^n)$, $\varphi \in \mathcal{D}(\mathbb{R}^n)$, (or $u \in \mathcal{E}'(\mathbb{R}^n)$, $\varphi \in \mathcal{E}(\mathbb{R}^n)$, or $u \in \mathcal{S}'(\mathbb{R}^n)$, $\varphi \in \mathcal{S}(\mathbb{R}^n)$), we define

$$(u * \varphi)(x) := \langle u, \tau_x \check{\varphi} \rangle.$$

It can be shown that $u * \varphi$ is smooth, and that $\langle u, \varphi \rangle = (u * \check{\varphi})(0)$.

3. For $u \in \mathcal{D}'(\mathbb{R}^n), v \in \mathcal{E}'(\mathbb{R}^n)$ or $u \in \mathcal{E}'(\mathbb{R}^n), v \in \mathcal{D}'(\mathbb{R}^n)$, define $u * v \in \mathcal{D}'(\mathbb{R}^n)$ by the property

$$(u * v) * \varphi = u * (v * \varphi) \qquad \forall \varphi \in \mathcal{D}(\mathbb{R}^n).$$

Fourier transform

1. For $f \in L^1(\mathbb{R}^n)$, define the Fourier transform by

$$\mathcal{F}[f](\lambda) = \hat{f}(\lambda) := \int_{\mathbb{R}^n} e^{-i\lambda \cdot x} f(x) \, \mathrm{d}x \, .$$

It is known that $\mathcal{F}$ is a continuous bijection from $\mathcal{S}(\mathbb{R}^n)$ to itself with inverse

$$\mathcal{F}^{-1}[\hat{f}](x) := \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i\lambda \cdot x} \hat{f}(\lambda) \, \mathrm{d}\lambda \, .$$

Note that we can write $\mathcal{F}^{-1} = \frac{1}{(2\pi)^n} \mathcal{R} \mathcal{F} = \frac{1}{(2\pi)^n} \mathcal{F} \mathcal{R}$.

2. For $u \in \mathcal{S}'(\mathbb{R}^n)$, we define the Fourier transform of u by $\langle \mathcal{F}[u], \varphi \rangle = \langle \hat{u}, \varphi \rangle := \langle u, \hat{\varphi} \rangle$. It is known that $\mathcal{F}$ extends to a continuous bijection from $\mathcal{S}'(\mathbb{R}^n)$ to itself, with inverse $\mathcal{F}^{-1} = (2\pi)^{-n} \mathcal{R} \mathcal{F} = (2\pi)^{-n} \mathcal{F} \mathcal{R}$.

Sobolev space We define $\langle \lambda \rangle := \sqrt{1 + \|\lambda\|^2}$ for $\lambda \in \mathbb{R}^n$, and note that $\langle \lambda \rangle \sim \|\lambda\|$ for large λ.

For $s \in \mathbb{R}$, we define the *Sobolev space* $H^s(\mathbb{R}^n)$ as the set of tempered distributions $u \in \mathcal{S}'(\mathbb{R}^n)$ for which $\hat{u}$ can be identified with a measurable function $\hat{u}(\lambda)$ such that

$$\int_{\mathbb{R}^n} \langle \lambda \rangle^s \hat{u}(\lambda) \, \mathrm{d}\lambda < \infty.$$

1 Distributions

1.1 Test functions and distributions

Definition 1.1. Let $X \subseteq \mathbb{R}^n$ be open, then we define the set of *test functions* on X as

$$\mathcal{D}(X) := C_0^\infty(X) = \{f : X \to \mathbb{C} \mid f \text{ is smooth with compact support}\}.$$

Definition 1.2. Let $(\varphi_m) \subseteq \mathcal{D}(X)$. We say that $(\varphi_m) \to 0$ in $\mathcal{D}(X)$ if

1. there exists a compact $K \subseteq X$ such that $\operatorname{supp} \varphi_m \subseteq K$ for all m;
2. $\partial^\alpha \varphi_m \to 0$ uniformly for each multi-index α.

Note that, for any $\varphi, \psi \in \mathcal{D}(X)$ and any multi-index α we have

$$\int_X \varphi \cdot \partial^\alpha \psi \, \mathrm{d}x = (-1)^{|\alpha|} \int_X \psi \cdot \partial^\alpha \varphi \, \mathrm{d}x \,,$$

which follows from partial integration and the fact that all boundary terms vanish since φ and ψ have compact support.

Also, by Taylor's theorem, for any $\varphi \in \mathcal{D}(X)$, $x, h \in X$ and $N \in \mathbb{N}$ we have

$$\varphi(x+h) = \sum_{|\alpha| \leqslant N} \frac{h^\alpha}{\alpha!} \partial^\alpha \varphi(x) + R_N(x,h) \quad \text{where } R_N(x,h) = o(|h|^N) \text{ uniformly in } x.$$

Definition 1.3. A *distribution* on X is a linear map $u : \mathcal{D}(X) \to \mathbb{C}$ if for every compact set $K \subseteq X$ there exist constants C, N such that for all $\varphi \in \mathcal{D}(X)$ with $\operatorname{supp} \varphi \subseteq K$ we have

$$|u(\varphi)| \leqslant C \sum_{|\alpha| \leqslant N} \sup |\partial^\alpha \varphi|. \tag{1}$$

Condition 1 is called the *seminorm condition*. If, in the seminorm condition, the same N can be used for every compact set $K \subseteq X$, then the least such N is called the *order* of u, written $\operatorname{ord}(u)$.

The set of all distributions in X is denoted $\mathcal{D}'(X)$.

If $u \in \mathcal{D}'(X)$ and $\varphi \in \mathcal{D}(X)$, then instead of $u(\varphi)$ we usually write $\langle u, \varphi \rangle$.

Recap 1.4. A function $f : X \to \mathbb{C}$ is called *locally integrable* if $\int_K |f| \, \mathrm{d}x < \infty$ for all compact $K \subseteq X$.

The set of locally integrable functions on X is denoted $L^1_{\mathrm{loc}}(X)$.

Example 1.5. Let $M \in \mathbb{N}$ and let $f_\alpha \in L^1_{\mathrm{loc}}(X)$ for all $|\alpha| \leqslant M$. Define the linear map $T : \mathcal{D}(X) \to \mathbb{C}$ by

$$\langle T, \varphi \rangle := \sum_{|\alpha| \leqslant M} \int_X f_\alpha \cdot \partial^\alpha \varphi \, \mathrm{d}x \,.$$

It is trivial that T is linear, and we verify that T is a distribution as follows: take $\varphi \in \mathcal{D}(X)$ with $\operatorname{supp} \varphi \subseteq K$. Then we have

$$\begin{aligned}
|\langle T, \varphi \rangle| &\leqslant \sum_{|\alpha| \leqslant M} \int_K |f_\alpha| \cdot |\partial^\alpha \varphi| \, \mathrm{d}x \\
&\leqslant \sum_{|\alpha| \leqslant M} \sup |\partial^\alpha \varphi| \cdot \int_K |f_\alpha| \, \mathrm{d}x \\
&\leqslant \left(\max_\alpha \int_K |f_\alpha| \, \mathrm{d}x \right) \sum_{|\alpha| \leqslant M} \sup |\partial^\alpha \varphi|.
\end{aligned}$$

Therefore, the seminorm condition is satisfied with $N = M$. From this, it also follows that $\operatorname{ord}(T) \leqslant M$.

A special case of the previous example is the case $M = 0$: in this case the distribution simply becomes

$$\langle \tau_f, \varphi \rangle = \int_X f\varphi \,\mathrm{d}x \,.$$

Henceforth we will abuse notation: if $f \in L^1_{\mathrm{loc}}(X)$, then we will write f instead of τ_f, i.e., $\langle f, \varphi \rangle = \int_X f\varphi \,\mathrm{d}x$.

Lemma 1.6 (Sequential continuity)**.** *Let $u\colon \mathcal{D}(X) \to \mathbb{C}$ be a linear map. Then u is a distribution if and only if, for every sequence $(\varphi_m) \subseteq \mathcal{D}(X)$ with $\varphi_m \to 0$ as in definition 1.2, we have $\langle u, \varphi_m \rangle \to 0$.*

Proof. '$\implies$' If u is a distribution and $(\varphi_m) \to 0$, then $\operatorname{supp}\varphi_m \subseteq K$ for some compact K, and by eq. (1) there exist C, N such that

$$|\langle u, \varphi_m \rangle| \leqslant C \sum_{|\alpha| \leqslant N} \sup |\partial^\alpha \varphi_m| \to 0.$$

'$\impliedby$' Suppose there is a compact set K such that eq. (1) is not valid for any C, N. Let $m \in \mathbb{N}$ and $C = N = m$, then there is some φ_m with $\operatorname{supp}(\varphi_m) \subseteq K$, and

$$|\langle u, \varphi_m \rangle| > m \sum_{|\alpha| \leqslant m} \sup |\partial^\alpha \varphi_m|.$$

By dividing φ_m by $\langle u, \varphi_m \rangle \neq 0$, we may assume w.l.o.g. that $\langle u, \varphi_m \rangle = 1$. We now have a sequence (φ_m) such that

$$\frac{1}{m} > \sum_{|\alpha| \leqslant m} \sup |\partial^\alpha \varphi_m| \implies |\partial^\alpha \varphi_m| < \frac{1}{m} \quad \text{for } |\alpha| \leqslant m \implies \partial^\alpha \varphi_m \to 0 \text{ uniformly for all } \alpha.$$

Since each φ_m also satisfies $\operatorname{supp}\varphi_m \subseteq K$, by definition 1.2 we have that $\varphi_m \to 0$, but also $\langle u, \varphi_m \rangle \to 1$, a contradiction. □

1.2 Limits in the distribution space

Definition 1.7. We say that a sequence $(u_m) \subseteq \mathcal{D}'(X)$ converges to $u \in \mathcal{D}'(X)$ and write $u_m \to u$ if

$$\langle u_m, \varphi \rangle \to \langle u, \varphi \rangle \quad \text{for all } \varphi \in \mathcal{D}(X).$$

The following theorem is non-examinable but interesting:

Theorem 1.8. *Let (u_m) be a sequence in $\mathcal{D}'(X)$ such that $\lim_{m\to\infty} \langle u_m, \varphi \rangle$ exists for all $\varphi \in \mathcal{D}(X)$. Then the map $\langle u, \varphi \rangle := \lim_{m\to\infty} \langle u_m, \varphi \rangle$ is a distribution in X.*

Proof. This is a direct application of the uniform boundedness principle. □

Example 1.9. Let $X = \mathbb{R}$ and consider the sequence of functions $u_m \in L^1_{\mathrm{loc}}(\mathbb{R})$ defined by $u_m(x) = \sin(mx)$. Then, for all $\varphi \in \mathcal{D}(\mathbb{R})$, we have

$$\langle u_m, \varphi \rangle = \int_{\mathbb{R}} \sin(mx)\varphi(x) \,\mathrm{d}x = \frac{1}{m} \int_{\mathbb{R}} \cos(mx)\varphi'(x) \,\mathrm{d}x \leqslant \frac{1}{m} \int \left|\varphi'(x)\right| \mathrm{d}x \to 0.$$

Therefore, it holds that $u_m \to 0$ in $\mathcal{D}'(\mathbb{R})$. With our abuse of notation we write this as $\lim_{m\to\infty} \sin(mx) = 0$ in $\mathcal{D}'(\mathbb{R})$.

1.3 Basic operations

1.3.1 Differentiation and multiplication by smooth functions

For $u \in C^\infty(X)$ and $\varphi \in \mathcal{D}(X)$, we have noted that

$$\langle \partial^\alpha u, \varphi \rangle = \int_X \partial^\alpha u \cdot \varphi \, \mathrm{d}x = (-1)^{|\alpha|} \int_X u \cdot \partial^\alpha \varphi \, \mathrm{d}x = \langle u, (-1)^{|\alpha|} \partial^\alpha \varphi \rangle.$$

Since the RHS makes sense for any distribution u, we define

Definition 1.10. For $f \in C^\infty(X)$, $u \in \mathcal{D}'(X)$, we define $\partial^\alpha(fu)$ by

$$\langle \partial^\alpha(fu), \varphi \rangle := \langle u, (-1)^{|\alpha|} f \cdot \partial^\alpha \varphi \rangle$$

Remark. This definition outlines a more general pattern when working with distributions: first we take some well-defined operator on the collection of smooth maps, then we rewrite it to a form that is sensible for any distribution, and then we *define* that new form as the operator on distributions. This process is called *extending the definition by duality.*

Example 1.11. Let $u = \delta_x$, then we have

$$\langle \partial^\alpha \delta_x, \varphi \rangle = \langle \delta_x, (-1)^{|\alpha|} \partial^\alpha \varphi \rangle = (-1)^{|\alpha|} \partial^\alpha \varphi(x)$$

Furthermore, consider the Heaviside function $H(x) = \mathbb{1}_{x \geqslant 0}$. We have

$$\langle H', \varphi \rangle = -\langle H, \varphi' \rangle = -\int_{\mathbb{R}} H(x) \varphi'(x) \, \mathrm{d}x = -\int_0^\infty \varphi'(x) \, \mathrm{d}x = \varphi(0) = \langle \delta_0, \varphi \rangle,$$

so we write $H' = \delta_0$ *in the distributional sense.*

Lemma 1.12. *Suppose $u' \in \mathcal{D}'(\mathbb{R})$ satisfies $u' = 0$. Then u is constant (i.e., $\langle u, \varphi \rangle = \langle c, \varphi \rangle = c \int_{\mathbb{R}} \varphi \, \mathrm{d}x$ for some c).*

Proof. Fix any $\vartheta \in \mathcal{D}(\mathbb{R})$ with $\langle 1, \vartheta \rangle = 1$. Let $\varphi \in \mathcal{D}(\mathbb{R})$ and define

$$\varphi_A := \varphi - \langle 1, \varphi \rangle \vartheta, \quad \varphi_B := \langle 1, \varphi \rangle \vartheta \quad \text{such that } \varphi = \varphi_A + \varphi_B.$$

Note that $\langle 1, \varphi_A \rangle = \langle 1, \varphi \rangle - \langle 1, \varphi \rangle \langle 1, \vartheta \rangle = 0$.

We claim that the function $\Phi_A(x) := \int_{-\infty}^x \varphi_A(y) \, \mathrm{d}y$ has compact support: since $\operatorname{supp} \varphi_A \subseteq [a, b]$ for some $a, b \in \mathbb{R}$, clearly $\Phi_A(x) = 0$ for $x < a$, while for $x > b$ we have $\Phi_A(x) = \langle 1, \varphi_A \rangle = 0$. Obviously, it holds that $\Phi_A' = \varphi_a$. Now we compute

$$\langle u, \varphi \rangle = \langle u, \varphi_A \rangle + \langle u, \varphi_B \rangle = \langle u, \Phi_A' \rangle + \langle 1, \varphi \rangle \langle u, \vartheta \rangle = -\langle u', \Phi_A \rangle + c \langle 1, \varphi \rangle = \langle c, \varphi \rangle.$$

Since φ was chosen arbitrarily this shows that u is constant. □

1.3.2 Reflection and translation

For $\varphi \in \mathcal{D}(\mathbb{R}^n), h \in \mathbb{R}^n$, define the *translation of φ by h* by

$$(\tau_h \varphi)(x) := \varphi(x - h),$$

and the *reflection of* φ by $\check{\varphi}(x) := \varphi(-x)$.

Extending the definitions of translation and reflection by duality yields the following:

Definition 1.13. For $u \in \mathcal{D}'(\mathbb{R}^n)$, $h \in \mathbb{R}^n$, $\varphi \in \mathcal{D}(\mathbb{R}^n)$, define

$$\langle \tau_h u, \varphi \rangle := \langle u, \tau_{-h} \varphi \rangle \quad \text{and} \quad \langle \check{u}, \varphi \rangle := \langle u, \check{\varphi} \rangle.$$

Lemma 1.14. *For $u \in \mathcal{D}'(\mathbb{R}^n)$, define $V_h \in \mathcal{D}'(\mathbb{R}^n)$ for $0 \neq h \in \mathbb{R}^n$ by*

$$V_h := \frac{\tau_{-h}u - u}{\|h\|}$$

If $(h_j) \subseteq \mathbb{R}^n$ is a sequence for which $\lim_{j\to\infty} \frac{h_j}{\|h_j\|} = m \in S^{n-1}$, then $V_{h_j} \to \sum_i m_i \frac{\partial u}{\partial x_i}$ in $\mathcal{D}'(\mathbb{R}^n)$.

Proof. By definition, we can write $\langle V_h, \varphi\rangle = \frac{1}{\|h\|}\langle u, \tau_h\varphi - \varphi\rangle$. Now Taylor's theorem tells us that

$$(\tau_h\varphi - \varphi)(x) = \varphi(x-h) - \varphi(x) = -\sum_i h_i \frac{\partial \varphi}{\partial x_i} + R(x,h),$$

where $R(x,h) = o(\|h\|)$ in $D(\mathbb{R}^n)$ (exercise sheet 1, question 2).

By sequential continuity, we have

$$\lim_{j\to\infty}\langle V_{h_j}, \varphi\rangle = \langle u, -\sum_i m_i \frac{\partial\varphi}{\partial x_i}\rangle = \langle \sum_i m_i \frac{\partial u}{\partial x_i}, \varphi\rangle,$$

which shows that $V_{h_j} \to \sum_i m_i \frac{\partial u}{\partial x_i}$ in $\mathcal{D}'(\mathbb{R}^n)$. $\square$

1.3.3 Convolution

For $\varphi \in \mathcal{D}(\mathbb{R}^n)$, note that $(\tau_x\check{\varphi})(y) = \check{\varphi}(y-x) = \varphi(x-y)$.

Definition 1.15. For $u \in C^\infty(\mathbb{R}^n), \varphi \in \mathcal{D}(\mathbb{R}^n)$, we define the *convolution* $u * \varphi \colon \mathbb{R}^n \to \mathbb{C}$ as

$$(u*\varphi)(x) := \int_{\mathbb{R}^n} u(x-y)\varphi(y)\,\mathrm{d}y = \int_{\mathbb{R}^n} u(y)\varphi(x-y)\,\mathrm{d}y = \langle u, \tau_x\check{\varphi}\rangle.$$

Since the RHS makes sense for any $u \in \mathcal{D}'(\mathbb{R}^n)$, we extend the definition this way: for $u \in \mathcal{D}'(\mathbb{R}^n), \varphi \in \mathcal{D}(\mathbb{R}^n)$, we define the convolution $u * \varphi$ as

$$(u*\varphi)(x) := \langle u, \tau_x\check{\varphi}\rangle.$$

Lemma 1.16. *Let $\varphi \in C^\infty(\mathbb{R}^n \times \mathbb{R}^m)$ and define $\Phi_x(y) := \varphi(x,y)$. Suppose for any $x \in \mathbb{R}^n$ there exists a neighbourhood $N(x)$ and a compact $K \subseteq \mathbb{R}^m$ such that $\varphi(x,y)$ for all $x \in N(x), y \notin K$.*

Then $x \mapsto \langle u, \Phi_x\rangle$ is differentiable with

$$\partial_x^\alpha\langle u, \Phi_x\rangle = \langle u, \partial_x^\alpha \Phi_x\rangle$$

for any $u \in \mathcal{D}'(\mathbb{R}^m)$.

Proof. Fix $x \in \mathbb{R}^n$, then by Taylor's formula we have

$$\Phi_{x+h}(y) = \Phi_x(y) + \sum_{i=1}^n h_i \frac{\partial\varphi}{\partial x_i} + R(x,y,h),$$

where $\partial_y^\alpha R(x,y,h) = o(\|h\|)$, uniformly in y, for any multi-index α. Furthermore, by assumption there exists a compact K such that for h small enough, $\operatorname{supp} R(x,\cdot,h) \subseteq K$. Therefore, $R(x,\cdot,h)$ is a test function for h small enough.

Combining the previous two facts shows that $R(x,\cdot,h) = o(\|h\|)$ in $\mathcal{D}(\mathbb{R}^m)$ as $h \to 0$.

Let $u \in \mathcal{D}'(\mathbb{R}^m)$, then we find by sequential continuity that $\langle u, R(x,\cdot,h)\rangle$ is also $o(\|h\|)$, and therefore

$$\langle u, \Phi_{x+h}\rangle = \langle u, \Phi_x\rangle + \sum_{i=1}^n h_i\langle u, \frac{\partial\Phi_x}{\partial x_i}\rangle + o(\|h\|).$$

This shows that $x \mapsto \langle u, \Phi_x\rangle$ is differentiable with

$$\frac{\partial}{\partial x_i}\langle u, \Phi_x\rangle = \langle u, \frac{\partial\Phi_x}{\partial x_i}\rangle.$$

From this the result follows. $\square$

Corollary 1.17. *If $u \in \mathcal{D}'(\mathbb{R}^n)$, $\varphi \in \mathcal{D}(\mathbb{R}^n)$, then $u * \varphi$ is differentiable with $\partial^\alpha(u * \varphi) = u * \partial^\alpha \varphi$.*

Proof. Apply the previous lemma with $\Phi_x(y) := \varphi(x - y)$. □

Due to the previous corollary, we often call $u * \varphi$ a *regularisation* of u.

Convention. If $u \in \mathcal{D}'(X)$ and $\varphi \in \mathcal{D}(X)$, then instead of $\langle u, \varphi \rangle$ we also write $\langle u(t), \varphi(t) \rangle$ (or with any other dummy variable) when the variable used for φ is not directly clear.

1.4 Density of test functions in distribution space

Lemma 1.18. *If $u \in \mathcal{D}'(\mathbb{R}^n)$ and $\varphi, \psi \in \mathcal{D}(\mathbb{R}^n)$, then*

$$(u * \varphi) * \psi = u * (\varphi * \psi).$$

Proof. Fix $x \in \mathbb{R}^n$. Now we write

$$\begin{aligned}((u * \varphi) * \psi)(x) &= \int_{\mathbb{R}^n} \langle u(z), (\tau_y \check{\varphi})(z) \rangle \cdot \psi(x - y) \, \mathrm{d}y \\ &= \int_{\mathbb{R}^n} \langle u(z), (\tau_{x-y} \check{\varphi})(z) \rangle \cdot \psi(y) \, \mathrm{d}y \\ &= \int_{\mathbb{R}^n} \langle u(z), \psi(y)(\tau_{x-y} \check{\varphi})(z) \rangle \, \mathrm{d}y \,.\end{aligned}$$

We would like to interchange integral and application of u, and we will have to justify this using Riemann sums:

$$\begin{aligned}\int_{\mathbb{R}^n} \langle u(z), \psi(y)(\tau_{x-y} \check{\varphi})(z) \rangle \, \mathrm{d}y &= \lim_{\varepsilon \downarrow 0} \sum_{m \in \mathbb{Z}^n} \langle u(z), \psi(\varepsilon m) \varphi(x - z - \varepsilon m) \rangle \varepsilon^n \\ &\overset{*}{=} \lim_{\varepsilon \downarrow 0} \langle u(z), \sum_{m \in \mathbb{Z}^n} \psi(\varepsilon m) \varphi(x - z - \varepsilon m) \varepsilon^n \rangle \\ &\overset{**}{=} \left\langle u(z), \int_{\mathbb{R}^n} \psi(y) \varphi(x - z - y) \, \mathrm{d}y \right\rangle \\ &= \langle u(z), (\varphi * \psi)(x - z) \rangle = \langle u(z), (\tau_x \widecheck{\varphi * \psi})(z) \rangle = (u * (\varphi * \psi))(x).\end{aligned}$$

Here, $*$ is by the fact that the sum is finite since ψ has compact support, while $**$ is by sequential continuity of u and the fact that the Riemann sum converges to the convolution integral *in the space of test functions* (non-examinable fact). □

We will use the following trick many times:

Proposition 1.19. *For any $u \in \mathcal{D}'(\mathbb{R}^n), \varphi \in \mathcal{D}(\mathbb{R}^n)$ we have $\langle u, \varphi \rangle = (u * \check{\varphi})(0)$.*

Proof. We have $(u * \check{\varphi})(0) = \langle u, \tau_0 \varphi \rangle = \langle u, \varphi \rangle$. □

For example, from this trick it follows that if $u * \varphi = 0$ for all φ, then $u = 0$.

Theorem 1.20. *If $u \in \mathcal{D}'(\mathbb{R}^n)$, there exists a sequence $(\varphi_k) \subseteq \mathcal{D}(\mathbb{R}^n)$ such that $\varphi_k \to u$ in $\mathcal{D}'(\mathbb{R}^n)$.*

Proof. Fix $\psi \in \mathcal{D}(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} \psi \, \mathrm{d}x = 1$, and set $\psi_k(x) := k^n \psi(kx)$. Note that $\int_{\mathbb{R}^n} \psi_k \, \mathrm{d}x = 1$.

Now, fix any $\chi \in \mathcal{D}(\mathbb{R}^n)$ with $\chi \equiv 1$ on $\{\|x\| < 1\}$ and $\chi \equiv 0$ on $\{\|x\| < 2\}$. Define $\chi_k(x) := \chi(x/k)$, so that $\lim_{k \to \infty} \chi_k(x) = 1$ for all x. We will set

$$\varphi_k(x) := \chi_k(x)(u * \psi_k)(x).$$

Clearly we have $\varphi_k \in \mathcal{D}(\mathbb{R}^n)$ since each χ_k has compact support.

Now, take any $\vartheta \in \mathcal{D}(\mathbb{R}^n)$, then we have

$$\begin{aligned}\langle \varphi_k, \vartheta \rangle = \langle \chi_k(u * \psi_k), \vartheta \rangle = \langle u * \psi_k, \chi_k \vartheta \rangle &= \Big[(u * \psi_k) * \widetilde{\chi_k \vartheta}\Big](0) \\ &= \Big[u * \Big(\psi_k * \widetilde{\chi_k \vartheta}\Big)\Big](0).\end{aligned}$$

Now we compute $\psi_k * \widetilde{\chi_k \vartheta}$: note that

$$\begin{aligned}(\psi_k * \widetilde{\chi_k \vartheta})(x) &= \int_{\mathbb{R}^n} k^n \psi(k(x-y)) \chi\Big(-\frac{y}{k}\Big) \vartheta(-y)\,\mathrm{d}y \\ &= \int_{\mathbb{R}^n} \psi(y) \chi\Big(\frac{y}{k^2} - \frac{x}{k}\Big) \vartheta(\frac{y}{k} - x)\,\mathrm{d}y \\ &= \vartheta(-x) + R_k(-x) = (\widetilde{\vartheta + R_k})(x)\end{aligned}$$

where $R_k(x) = \int_{\mathbb{R}^n} \psi(y)\big[\chi\big(\frac{y}{k^2} + \frac{x}{k}\big)\vartheta\big(\frac{y}{k} + x\big) - \vartheta(x)\big]\,\mathrm{d}y$.

So

$$\langle \varphi_k, \vartheta \rangle = (u * (\widetilde{\vartheta + R_k}))(0) = (u * \check{\vartheta})(0) + (u * \check{R_k})(0) = \langle u, \vartheta \rangle + \langle u, R_k \rangle.$$

We must now only prove that $R_k \to 0$ in $\mathcal{D}(\mathbb{R}^n)$, and then by sequential continuity it follows that $\varphi_k \to u$ in $\mathcal{D}'(\mathbb{R}^n)$. □

2 Distributions with compact support

Definition 2.1. Let $Y \subseteq X$ be open and $u \in \mathcal{D}'(X)$. We say that u *vanishes* on Y if $\langle u, \varphi \rangle = 0$ for all $\varphi \in \mathcal{D}(Y)$.

Definition 2.2. For $u \in \mathcal{D}'(X)$, we define the *support* of u as

$$\operatorname{supp} u := X \backslash \bigcup \{Y \subseteq X \mid Y \text{ open}, u \text{ vanishes on } Y\}.$$

For example, the support of δ_x is simply $\{x\}$.

2.1 Test functions and distributions

We will now consider a bigger space of test functions, which yields a smaller space of distributions.

Definition 2.3. We define $\mathcal{E}(X)$ as the space of smooth functions $\varphi \colon X \to \mathbb{C}$. We say that a sequence $(\varphi_m) \subseteq \mathcal{E}(X)$ converges to 0 if $\partial^\alpha \varphi \to 0$ uniformly on compact subsets of X for every multi-index α.

Definition 2.4. We define $\mathcal{E}'(X)$ as the space of linear maps $u \colon \mathcal{E}(X) \to \mathbb{C}$ for which there exists a compact $K \subseteq X$ and nonnegative constants C, N such that

$$|\langle u, \varphi \rangle| \leqslant C \sum_{\alpha \leqslant N} \sup_K |\partial^\alpha \varphi| \tag{2}$$

for all $\varphi \in \mathcal{E}(X)$.

Lemma 2.5 (Sequential continuity)**.** *Let $u \colon \mathcal{E}(X) \to \mathbb{C}$ be a linear map. Then $u \in \mathcal{E}'(X)$ if and only if, for every sequence $(\varphi_m) \subseteq \mathcal{E}(X)$ with $\varphi_m \to 0$, we have $\langle u, \varphi_m \rangle \to 0$.*

Proof. TODO: □

Lemma 2.6. *If $u \in \mathcal{E}'(X)$, then $u\restriction_{\mathcal{D}(X)}$ defines an element of $\mathcal{D}'(X)$ with compact support and finite order.*

Conversely, for each $u \in \mathcal{D}'(X)$ with compact support there exists a unique extension $\tilde{u} \in \mathcal{E}'(X)$ with $\operatorname{supp}(\tilde{u}) = \operatorname{supp}(u)$ and $\tilde{u}\restriction_{\mathcal{D}(X)} = u$.

Proof. Let $u \in \mathcal{E}'(X)$, so that there exists a compact $K \subseteq X$ with $|\langle u, \varphi \rangle| \leqslant C \sum_{|\alpha| \leqslant N} \sup_K |\partial^\alpha \varphi|$. Now, for any compact $K' \subseteq X$ and any φ with $\operatorname{supp} \varphi \subseteq K'$, eq. (1) is clearly satisfied, and we can use the same N for all compact K', so clearly $u\restriction_{\mathcal{D}(X)}$ is an element of $\mathcal{D}'(X)$ with finite order. Finally, suppose φ is supported in $X \backslash K$, then it is clear that $\langle u, \varphi \rangle = 0$, which proves that $\operatorname{supp} u \subseteq K$ and therefore that u has compact support.

Now suppose $u \in \mathcal{D}'(X)$ has compact support, let $\rho \in \mathcal{D}(X)$ be 1 in a neighbourhood of $\operatorname{supp} u$, and define $\tilde{u} \in \mathcal{E}'(X)$ by

$$\langle \tilde{u}, \varphi \rangle := \langle u, \rho\varphi \rangle \quad \forall \varphi \in \mathcal{E}(X).$$

Clearly $\tilde{u}$ is an element of $\mathcal{E}'(X)$ since $\operatorname{supp}(\rho\varphi) \subseteq \operatorname{supp} \rho$ and

$$|\langle \tilde{u}, \varphi \rangle| = |\langle u, \rho\varphi \rangle| \leqslant C \sum_{|\alpha| \leqslant N} \sup_{\operatorname{supp}(\rho)} |\partial^\alpha (\rho\varphi)| \overset{\star}{\leqslant} C' \sum_{|\alpha| \leqslant N} \sup_{\operatorname{supp} \rho} |\partial^\alpha \varphi|,$$

where $\star$ follows from the Leibniz rule. It is also clear that $\operatorname{supp} \tilde{u} = \operatorname{supp} u$.

Finally we will show uniqueness: suppose $\tilde{v}$ is an extension of u with $\operatorname{supp} \tilde{v} = \operatorname{supp} u$, and write any $\varphi \in \mathcal{E}(X)$ as $\varphi = \rho\varphi + (1-\rho)\varphi = \varphi_0 + \varphi_1$. Then since $\varphi_0 \in \mathcal{D}(X)$ and φ_1 vanishes on a neighbourhood of $\operatorname{supp} u$, we find

$$\langle \tilde{v}, \varphi \rangle = \langle \tilde{v}, \varphi_0 + \varphi_1 \rangle = \langle \tilde{v}, \varphi_0 \rangle = \langle u, \varphi_0 \rangle,$$

which is independent of choice of extension. □

2.2 Convolution between distributions

Definition 2.7. Define for $u \in \mathcal{E}'(\mathbb{R}^n), \varphi \in \mathcal{E}(\mathbb{R}^n)$ the *convolution*

$$(u * \varphi)(x) = \langle u, \tau_x \check{\varphi} \rangle.$$

This convolution satisfies the same properties as the convolution between $\mathcal{D}'(\mathbb{R}^n)$ and $\mathcal{D}(\mathbb{R}^n)$. Also, if $\varphi \in \mathcal{D}(\mathbb{R}^n)$, then $u * \varphi \in \mathcal{D}(\mathbb{R}^n)$.

Definition 2.8. Let $u, v \in \mathcal{D}'(\mathbb{R}^n)$, at least one of which has compact support, define $u * v \colon \mathcal{D}(\mathbb{R}^n) \to \mathcal{E}(\mathbb{R}^n)$ by

$$(u * v) * \varphi = u * (v * \varphi) \quad \forall \varphi \in \mathcal{D}(\mathbb{R}^n).$$

We will show (example sheet 2, question 1) that $u * v$ is uniquely defined and gives rise to an element of $\mathcal{D}'(\mathbb{R}^n)$ via $\langle u * v, \varphi \rangle := [(u * v) * \check{\varphi}](0)$.

Lemma 2.9. *Given $u, v \in \mathcal{D}'(\mathbb{R}^n)$, at least one of which has compact support, we have $u * v = v * u$.*

Proof. First we note that $(u * \varphi) * \psi = u * (\varphi * \psi)$ holds if u has compact support and at least one of φ, ψ has compact support.

Fix $\varphi, \psi \in \mathcal{D}(\mathbb{R}^n)$, we see from our earlier shown properties that

$$(u * v) * (\varphi * \psi) = u * (v * (\varphi * \psi)) = u * ((v * \varphi) * \psi) = u * (\psi * (v * \varphi)) = (u * \psi) * (v * \varphi).$$

If we interchange u and v in the above, that is equivalent to interchanging φ and ψ, which we know must yield the same result. This shows $u * v$ and $v * u$ agree on $\varphi * \psi$ for all $\varphi, \psi \in \mathcal{D}(\mathbb{R}^n)$. Defining $E = u * v - v * u$, we find that $0 = E * (\varphi * \psi) = (E * \varphi) * \psi$ for all $\varphi, \psi \in \mathcal{D}(\mathbb{R}^n)$, so $E * \varphi = 0$ for all $\varphi \in \mathcal{D}(\mathbb{R}^n)$, so $E = 0$. $\square$

3 Tempered distributions and Fourier analysis

3.1 Functions of rapid decay

Definition 3.1. For any $f\colon \mathbb{R}^n \to \mathbb{C}$ and multi-indices α, β we define $\|f\|_{\alpha,\beta} := \sup_{x\in\mathbb{R}^n} \left|x^\alpha \partial^\beta \varphi\right|$.
We define the *Schwartz space*

$$\mathcal{S}(\mathbb{R}^n) := \left\{ f \in C^\infty(\mathbb{R}^n) : \|f\|_{\alpha,\beta} < \infty \text{ for all } \alpha, \beta \right\}.$$

We say that a sequence $(\varphi_n) \subseteq \mathcal{S}(\mathbb{R}^n)$ converges to 0 if $\|\varphi_m\|_{\alpha,\beta} \to 0$ for every α, β.

Example 3.2. The function $x \mapsto \exp\left(-\|x\|^2\right)$ lies in $\mathcal{S}(\mathbb{R}^n)$.

Proposition 3.3. *For all n we have that $\mathcal{S}(\mathbb{R}^n) \subseteq L^1(\mathbb{R}^n)$.*

Proof. Let $\varphi \in \mathcal{S}(\mathbb{R}^n)$, then for all $N \in \mathbb{N}$ we have

$$\int_{\mathbb{R}^n} |\varphi(x)| \,\mathrm{d}x = \int_{\mathbb{R}^n} (1 + \|x\|)^{-N}(1 + \|x\|)^N |\varphi(x)| \,\mathrm{d}x \overset{?}{\leqslant} C \sum_{|\alpha|\leqslant N} \|\varphi\|_{\alpha,0} \int_{\mathbb{R}^n} (1 + \|x\|)^{-N} \,\mathrm{d}x\,.$$

Since $\int_{\mathbb{R}^n} (1 + \|x\|)^{-N} \,\mathrm{d}x$ is finite for N large enough (??), this proves the claim. $\square$

Definition 3.4. A linear map $u\colon \mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$ is called a *tempered distribtion* if there exists constants C, N such that

$$|\langle u, \varphi\rangle| \leqslant C \sum_{|\alpha|,|\beta|\leqslant N} \|\varphi\|_{\alpha,\beta} \quad \text{for all } \varphi \in \mathcal{S}(\mathbb{R}^n).$$

The space of tempered distributions is denoted by $\mathcal{S}'(\mathbb{R}^n)$.

This definition is equivalent to sequential continuity.

3.2 The Fourier transform on Schwartz functions

Convention. We write $D := -i\partial$ and $D^\alpha = (-i)^{|\alpha|}\partial^\alpha$.

Definition 3.5. For $f \in L^1(\mathbb{R}^n)$, define the *Fourier transform* of f by

$$[\mathcal{F}(f)](\lambda) = \hat{f}(\lambda) := \int_{\mathbb{R}^n} e^{-i\lambda\cdot x} f(x) \,\mathrm{d}x \quad \text{where } \lambda \in \mathbb{R}^n.$$

Lemma 3.6. *If $f \in L^1(\mathbb{R}^n)$, then $\hat{f}$ is continuous.*

Proof. If $\lambda_m \to \lambda \in \mathbb{R}^n$, then by the dominated convergence theorem we have

$$\hat{f}(\lambda_m) = \int_{\mathbb{R}^n} e^{-i\lambda_m\cdot x} f(x) \,\mathrm{d}x \overset{\mathrm{DCT}}{=} \int_{\mathbb{R}^n} e^{-i\lambda\cdot x} f(x) \,\mathrm{d}x = \hat{f}(\lambda),$$

where we were able to use the dominated convergence theorem since the integrand is absolutely bounded by $|f|$ and $f \in L^1$. $\square$

It turns out that this idea generalises: the faster the function f decays, the smoother the Fourier transform $\hat{f}$ is.

Lemma 3.7. *For $\varphi \in \mathcal{S}(\mathbb{R}^n)$, we have $\mathcal{F}[D_x^\alpha \varphi](\lambda) = \lambda^\alpha \hat{\varphi}(\lambda)$ and $\mathcal{F}[x^\beta \varphi](\lambda) = (-1)^{|\beta|} D_\lambda^\beta \hat{\varphi}(\lambda)$.*

Proof. Since $\left|x^\alpha D^\beta \varphi\right| \to 0$ as $\|x\| \to \infty$, we have using integration by parts

$$\begin{aligned}\mathcal{F}[D^\alpha_\lambda \varphi](\lambda) &= \int_{\mathbb{R}^n} e^{-i\lambda\cdot x} D^\alpha_x \varphi(x)\,\mathrm{d}x\\ &= (-1)^{|\alpha|}\int_{\mathbb{R}^n} D^\alpha_x(e^{-i\lambda\cdot x})\varphi(x)\,\mathrm{d}x\\ &= \lambda^\alpha \hat{\varphi}(\lambda).\end{aligned}$$

For the second part of the lemma, by differentiation under the integral sign we get

$$\begin{aligned}\mathcal{F}[x^\beta \varphi](\lambda) &= \int_{\mathbb{R}^n} e^{-i\lambda\cdot x} x^\beta \varphi(x)\,\mathrm{d}x\\ &= \int_{\mathbb{R}^n} ((-D_\lambda)^\beta e^{-i\lambda\cdot x})\varphi(x)\,\mathrm{d}x\\ &= (-1)^{|\beta|} D^\beta_\lambda \hat{\varphi}(\lambda).\end{aligned}$$

□

We define the *inverse Fourier transform* by

$$\mathcal{F}^{-1}[\hat{f}](x) \coloneqq \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{i\lambda\cdot x}\hat{f}(\lambda)\,\mathrm{d}\lambda\,.$$

We will now show that on $\mathcal{S}(\mathbb{R}^n)$, the inverse Fourier transform is indeed an inverse:

Theorem 3.8. *The Fourier transform $\mathcal{F}$ defines a continuous isomorphism from $\mathcal{S}(\mathbb{R}^n)$ to itself.*

Proof. Let $\varphi \in \mathcal{S}(\mathbb{R}^n)$. First, we show that $\hat{\varphi} \in \mathcal{S}(\mathbb{R}^n)$: by the previous lemma we have for multi-indices α, β that

$$\begin{aligned}\left|\lambda^\alpha(-D_\lambda)^\beta \hat{\varphi}(\lambda)\right| = \left|\lambda^\alpha \mathcal{F}[x^\beta\varphi](\lambda)\right| = \left|\mathcal{F}[D^\alpha_x(x^\beta\varphi)](\lambda)\right| &= \left|\int_{\mathbb{R}^n} e^{-i\lambda\cdot x} D^\alpha(x^\beta\varphi)\,\mathrm{d}x\right|\\ &\leqslant \int_{\mathbb{R}^n}\left|D^\alpha(x^\beta\varphi)\right|\mathrm{d}x\,,\end{aligned}\tag{3}$$

which is finite since $D^\alpha(x^\beta\varphi)$ is also a Schwartz function and therefore integrable.

From the previous lemma we also infer that $\hat{\varphi}$ is smooth, so indeed we have $\hat{\varphi} \in \mathcal{S}(\mathbb{R}^n)$. From eq. (3) it is also easily seen that if $\varphi_m \to 0$ in $\mathcal{S}(\mathbb{R}^n)$, then $\hat{\varphi}_m \to 0$ in $\mathcal{S}(\mathbb{R}^n)$ also, which shows that $\mathcal{F}$ is continuous.

To prove surjectivity and injectivity, we will show that $\mathcal{F}^{-1}[\mathcal{F}[f]](x) = f(x)$ (???). Indeed we have

$$\begin{aligned}\mathcal{F}^{-1}[\mathcal{F}[\varphi]](x) &= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{i\lambda\cdot(x-y)}\varphi(y)\,\mathrm{d}y\,\mathrm{d}\lambda\\ &= \frac{1}{(2\pi)^n}\lim_{\varepsilon\downarrow 0}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{i\lambda\cdot(x-y)-\varepsilon\|\lambda\|^2}\varphi(y)\,\mathrm{d}y\,\mathrm{d}\lambda\\ &\overset{\star}{=} \frac{1}{(2\pi)^n}\lim_{\varepsilon\downarrow 0}\int_{\mathbb{R}^n}\varphi(y)\int_{\mathbb{R}^n} e^{i\lambda\cdot(x-y)-\varepsilon\|\lambda\|^2}\,\mathrm{d}\lambda\,\mathrm{d}y\,,\end{aligned}$$

where $\star$ follows from Fubini's theorem.

Now we note that

$$\int_{\mathbb{R}^n} e^{i\lambda\cdot(x-y)-\varepsilon\|\lambda\|^2}\,\mathrm{d}\lambda = \prod_{i=1}^n\int_{\mathbb{R}} e^{i\lambda_j(x_j-y_j)-\varepsilon\lambda_j^2}\,\mathrm{d}\lambda \overset{\star\star}{=} \prod_{i=1}^n\left(\frac{\pi}{e}\right)^{1/2} e^{-\frac{(x_i-y_i)^2}{4\varepsilon}} = \left(\frac{\pi}{\varepsilon}\right)^{n/2} e^{-\frac{\|x-y\|^2}{4\varepsilon}}\,.$$

To explain ⋆⋆, TODO: .
and plugging that into the above yields

$$\begin{aligned}\mathcal{F}^{-1}[\mathcal{F}[\varphi]](x) &= \lim_{\varepsilon\downarrow 0} 2^{-n}(\pi\varepsilon)^{-n/2}\int_{\mathbb{R}^n}\varphi(y)e^{-\|x-y\|^2/(4\varepsilon)}\,\mathrm{d}y \\ &\overset{\star\star\star}{=} \pi^{-n/2}\lim_{\varepsilon\downarrow 0}\int_{\mathbb{R}^n}\varphi(x-2\sqrt{\varepsilon}y')e^{-\|y'\|^2}\,\mathrm{d}y \\ &\overset{\text{DCT}}{=} \pi^{-n/2}\varphi(x)\int_{\mathbb{R}^n}e^{-\|y'\|^2}\,\mathrm{d}y = \varphi(x),\end{aligned}$$

where $\star\star\star$ follows from the substitution $x-y=2\sqrt{\varepsilon}y'$.

Finally, continuity of $\mathcal{F}^{-1}$ is easily shown with an argument analogous to that for continuity of $\mathcal{F}$ (????). □

3.3 The Fourier transform on tempered distributions

Lemma 3.9. *For $\varphi, \psi \in \mathcal{S}(\mathbb{R}^n)$, we have $\int_{\mathbb{R}^n}\varphi(x)\hat{\psi}(x)\,\mathrm{d}x = \int_{\mathbb{R}^n}\hat{\varphi}(x)\psi(x)\,\mathrm{d}x$.*

Proof. This follows from Fubini's theorem:

$$\int_{\mathbb{R}^n}\varphi(x)\hat{\psi}(x)\,\mathrm{d}x = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\varphi(x)\psi(\lambda)e^{-i\lambda\cdot x}\,\mathrm{d}\lambda\,\mathrm{d}x = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\psi(\lambda)\varphi(x)e^{-i\lambda\cdot x}\,\mathrm{d}x\,\mathrm{d}\lambda = \psi(\lambda)\psi(\lambda)\hat{\varphi}(\lambda)\,\mathrm{d}\lambda\,.$$

□

The above result can be rewritten as $\langle\hat{\varphi},\psi\rangle = \langle\varphi,\hat{\psi}\rangle$, which motivates the definition of the Fourier transform for tempered distributions:

Definition 3.10. For $u \in \mathcal{S}'(\mathbb{R}^n)$, we define its Fourier transform by

$$\langle\hat{u},\varphi\rangle := \langle u,\hat{\varphi}\rangle \quad \text{for all } \varphi \in \mathcal{S}(\mathbb{R}^n).$$

Using sequential continuity and theorem 3.8, it is easily seen that $\hat{u}$ is indeed an element of $\mathcal{S}'(\mathbb{R}^n)$.

Example 3.11. It is easily checked that $\delta_0 \in \mathcal{S}'(\mathbb{R}^n)$, and we can compute

$$\langle\hat{\delta_0},\varphi\rangle = \langle\delta_0,\hat{\varphi}\rangle = \hat{\varphi}(0) = \int_{\mathbb{R}^n}\varphi(x)\,\mathrm{d}x = \langle 1,\varphi\rangle,$$

so we can write $\hat{\delta_0} = 1$. Analogously, by the Fourier inversion theorem we have

$$\langle\hat{1},\varphi\rangle = \langle 1,\hat{\varphi}\rangle = \int_{\mathbb{R}^n}\hat{\varphi}(\lambda)\,\mathrm{d}\lambda = (2\pi)^n\varphi(0) = \langle(2\pi)^n\delta_0,\varphi\rangle,$$

so we write $\hat{1} = (2\pi)^n\delta_0$.

We can easily generalise lemma 3.7 to the Fourier transform on $\mathcal{S}'(\mathbb{R}^n)$, so

$$\mathcal{F}[D^\alpha u] = \lambda^\alpha\hat{u}, \quad \mathcal{F}[x^\beta u] = (-D^\beta)\hat{u}.$$

Theorem 3.12. *The Fourier transform $\mathcal{F}$ extends to a continuous isomorphism on $\mathcal{S}'(\mathbb{R}^n)$.*

Proof. We claim that $\check{u} = (2\pi)^{-n}\mathcal{F}[\mathcal{F}[u]]$. To check this, note that by the Fourier inversion theorem we have for $\varphi \in \mathcal{S}(\mathbb{R}^n)$ that

$$\check{\varphi}(x) = \varphi(-x) = (2\pi)^{-n}\int_{\mathbb{R}^n}e^{-i\lambda\cdot x}\hat{\varphi}(\lambda)\,\mathrm{d}\lambda = (2\pi)^{-n}\mathcal{F}[\mathcal{F}[\varphi]](x),$$

and therefore

$$\langle \check{u}, \varphi \rangle = \langle u, \check{\varphi} \rangle = \langle u, (2\pi)^{-n} \mathcal{F}[\mathcal{F}[\varphi]] \rangle = \langle (2\pi)^{-n} \mathcal{F}[\mathcal{F}[u]], \varphi \rangle.$$

This shows that $\mathcal{F}$ is bijective (since $\mathcal{F} \circ \mathcal{F}$ is bijective). For continuity of $\mathcal{F}$ and its inverse: using theorem 3.8, we find

$$\begin{aligned}
& u_m \to 0 \text{ in } \mathcal{S}'(\mathbb{R}^n) \\
\iff & \langle u_m, \varphi \rangle \to 0 \ \forall \varphi \in \mathcal{S}(\mathbb{R}^n) \\
\iff & \langle u_m, \hat{\varphi} \rangle \to 0 \ \forall \varphi \in \mathcal{S}(\mathbb{R}^n) \\
\iff & \langle \hat{u}_m, \varphi \rangle \to 0 \ \forall \varphi \in \mathcal{S}(\mathbb{R}^n) \\
\iff & \hat{u}_m \to 0 \text{ in } \mathcal{S}'(\mathbb{R}^n).
\end{aligned}$$

□

3.4 Sobolev spaces

Convention. We write $\langle \lambda \rangle := (1 + \|\lambda\|^2)^{1/2}$ for $\lambda \in \mathbb{R}^n$. Note that $\langle \lambda \rangle \sim 1$ as $\|\lambda\| \to 0$ and $\langle \lambda \rangle \to \|\lambda\|$ as $\|\lambda\| \to \infty$.

Definition 3.13. For $s \in \mathbb{R}$, define the *Sobolev space* $H^s(\mathbb{R}^n)$ to be the set of tempered distributions $u \in \mathcal{S}'(\mathbb{R}^n)$ for which $\hat{u}$ can be identified with a measurable function $\hat{u} \colon \mathbb{R}^n \to \mathbb{C}$ such that $\langle \lambda \rangle^s \hat{u} \in L^2(\mathbb{R}^n)$.

For $X \subseteq \mathbb{R}^n$ open, we define the *localised Sobolev space* $H^s_{\text{loc}}(X)$ by setting

$$u \in H^s_{\text{loc}}(X) \iff \varphi u \in H^s(\mathbb{R}^n) \text{ for all } \varphi \in \mathcal{D}(X).$$

Lemma 3.14 (Sobolev lemma)**.** *If $u \in H^s(\mathbb{R}^n)$ with $s > \frac{n}{2}$, then u is continuous.*

Proof. We will show that $\hat{u}$ is integrable. By Cauchy-Schwarz, we have

$$\begin{aligned}
\int_{\mathbb{R}^n} |\hat{u}(\lambda)| \, \mathrm{d}\lambda &= \int_{\mathbb{R}^n} \langle \lambda \rangle^{-s} \langle \lambda \rangle^s |\hat{u}(\lambda)| \, \mathrm{d}\lambda \\
&\leqslant \left(\int_{\mathbb{R}^n} \langle \lambda \rangle^{-2s} \, \mathrm{d}\lambda \right)^{1/2} \|u\|_{H^s} \\
&= C \|u\|_{H^s} \left(\int_0^\infty r^{n-1} (1 + r^2)^{-s} \, \mathrm{d}r \right)^{1/2},
\end{aligned}$$

where the last line follows from using polar coordinates and C is the area of the $(n-1)$-sphere.

Writing $s = \frac{n}{2} + \varepsilon$, we find that the integrand $r^{n-1}(1+r^2)^{-s}$ is of order $O(r^{-1-2\varepsilon})$ as $r \to \infty$, and therefore the integral is finite, so indeed we have $\hat{u} \in L^1(\mathbb{R}^n)$.

By applying theorem 3.8 to a test function, we can show that $u = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{i\lambda \cdot x} \hat{u}(\lambda) \, \mathrm{d}\lambda$, which is continuous by the dominated convergence theorem. □

Corollary 3.15. *If $u \in H^s(\mathbb{R}^n)$ for every $s > n/2$, then $u \in C^\infty(\mathbb{R}^n)$.*

4 Applications of Fourier transform

4.1 Elliptic regularity

Recall that $D = -i\partial$. If p is an N-th order polynomial, then $p(D)$ is called an N-th order differential operator.

Definition 4.1. For an N-th order differential operator $p(D) = \sum_{|\alpha|\leqslant N} c_\alpha D^\alpha$, define its *principal symbol* $\sigma_p(\lambda)$ by

$$\sigma_p(\lambda) := \sum_{|\alpha|=N} c_\alpha \lambda^\alpha \qquad (\lambda \in \mathbb{R}^n).$$

The operator $p(D)$ is called *elliptic* if $\sigma_p(\lambda) \neq 0$ for $\lambda \neq 0$.

Lemma 4.2. *If $p(D)$ is an N-th order elliptic partial differential operator, then there exist $R > 0$ such that, $C > 0$ such that*

$$|p(\lambda)| \geqslant C\langle\lambda\rangle^N \qquad \text{if } \|\lambda\| > R.$$

Proof. Let $C_0 > 0$ be the minimum of $|\sigma_p|$ on S^{n-1}, then for $\lambda \neq 0$ we have

$$|\sigma_p(\lambda)| = \left|\sum_{|\alpha|=N} c_\alpha \lambda^\alpha\right| = \|\lambda\|^N |\sigma_p(\lambda/\|\lambda\|)| \geqslant \|\lambda\|^N C_0.$$

By the triangle inequality we find

$$|p(\lambda)| \geqslant |\sigma_p(\lambda)| - |\sigma_p(\lambda) - p(\lambda)| \geqslant \left[C_0 - \left|\frac{p(\lambda) - \sigma_p(\lambda)}{\|\lambda\|^N}\right|\right] \|\lambda\|^N$$

Since $p - \sigma_p$ is a polynomial of order $N-1$, we can choose R sufficiently large s.t. $\left|\frac{p(\lambda)-\sigma_p(\lambda)}{\|\lambda\|^N}\right| < C_0/2$. Since $\langle\lambda\rangle \sim \|\lambda\|$ for λ large enough, we find that there exists C such that

$$|p(\lambda)| \geqslant \frac{C_0}{2}\|\lambda\|^N \geqslant C\langle\lambda\rangle^N$$

for $\|\lambda\| > R$. $\square$

We will try to prove the *elliptic regularity theorem*:

Theorem 4.3 (Elliptic regularity)**.** *Suppose $p(D)$ is an N-th order elliptic partial differential operator and $u \in \mathcal{D}'(X)$ satisfies $p(D)u \in H^s_{\text{loc}}(X)$, then $u \in H^{s+N}_{\text{loc}}(X)$.*

Corollary 4.4. *If $p(D)$ is N-th order elliptic and $p(D)u \in C^\infty(X)$, then $u \in C^\infty(X)$.*

We will first prove an "easy version" of theorem 4.3 using a *parametrix*:

Definition 4.5. If $p(D)$ is an N-th order differential operator, then $E \in D'(\mathbb{R}^n)$ is called a *parametrix* for $p(D)$ if

$$p(D)E = \delta_0 + \omega \qquad \text{for some } \omega \in \mathcal{E}(\mathbb{R}^n).$$

Lemma 4.6. *Every elliptic partial differential operator $p(D)$ has a parametrix which is smooth on $\mathbb{R}^n\backslash\{0\}$.*

Proof. Since $p(D)$ is elliptic, we can choose $R > 0$, $C > 0$ such that $|p(\lambda)| \geqslant C\langle\lambda\rangle^N$ for $\|\lambda\| > R$.

Fix some $\chi_R \in \mathcal{D}(\mathbb{R}^n)$ such that $\chi_R = 1$ on $\|\lambda\| \leqslant R$ and $\chi_R = 0$ on $\|\lambda\| > R + 1$, and define

$$\hat{E}(\lambda) := \frac{1 - \chi_R(\lambda)}{p(\lambda)}.$$

Then $\tilde{E}$ is smooth and for λ sufficiently large we have $|\hat{E}(\lambda)| \lesssim \langle\lambda\rangle^{-N}$ since χ_R vanishes for large λ, which implies $\hat{E} \in \mathcal{S}'(\mathbb{R}^n)$. Therefore, $p(\lambda)\hat{E} = 1 - \chi_R(\lambda)$ is also a tempered distribution and we can take its inverse Fourier transform $p(D)E = \delta_0 + \omega$ for some $\omega \in \mathcal{S}(\mathbb{R}^n)$, which shows that E is a parametrix.

To prove that E is smooth on $\mathbb{R}^n\backslash\{0\}$, consider for $\|\lambda\| > R+1$

$$\left|\mathcal{F}[D^\beta(x^\alpha E)]\right| = \left|\lambda^\beta D^\alpha \hat{E}\right| = \left|\lambda^\beta D^\alpha\left(\frac{1}{p(\lambda)}\right)\right| \overset{\star}{\lesssim} \|\lambda\|^{|\beta|-|\alpha|-N},$$

where $\star$ can be shown with an induction argument. For each β, we can simply choose $|\alpha|$ large enough such that $\mathcal{F}[D^\beta(x^\alpha E)] \in L^1(\mathbb{R}^n)$, and therefore $D^\beta(x^\alpha E)$ is continuous for $|\alpha|$ large enough. Since β was randomly chosen, E will be smooth outside the origin. $\square$

We will now consider the proof of theorem 4.3 in the special case that u and $f := p(D)u$ have compact support.

Proof. Let E be a parametrix for P, then we have

$$u = \delta_0 * u = [p(D)E - \omega] * u = p(D)E * u - \omega * u = E * f - \omega * u.$$

Since u has compact support, $\omega * u$ will be a Schwartz function, and it can be shown that

$$|\mathcal{F}[E * f](\lambda)| = \left|\hat{E}(\lambda)\hat{f}(\lambda)\right| \lesssim \langle\lambda\rangle^{-N}\left|\hat{f}(\lambda)\right|,$$

which shows that $f \in H^s(\mathbb{R}^n) \implies u \in H^{s+N}(\mathbb{R}^n)$. $\square$

To prove theorem 4.3 in general, we will need some facts which are proved on the second example sheet:

1. If $s > t$ then $H^s(\mathbb{R}^n) \subseteq H^t(\mathbb{R}^n)$;
2. If $\varphi \in \mathcal{S}(\mathbb{R}^n)$, $u \in H^s(\mathbb{R}^n)$, then $\varphi u \in H^s(\mathbb{R}^n)$;
3. If $u \in \mathcal{E}'(\mathbb{R}^n)$, then $u \in H^t(\mathbb{R}^n)$ for some $t \in \mathbb{R}$;
4. If $u \in H^s(\mathbb{R}^n)$, then $D^\alpha u \in H^{s-|\alpha|}(\mathbb{R}^n)$.

Now we prove the theorem:

Proof. Fix $\varphi \in \mathcal{D}(X)$, we wish to prove that $\varphi u \in H^{s+N}(\mathbb{R}^n)$ given that $p(D)u \in H^s_{\text{loc}}(X)$. Choosing $M \in \mathbb{N}$, we introduce a collection $\{\psi_0, \ldots, \psi_M\} \subseteq \mathcal{D}(X)$ such that

$$\operatorname{supp}(\varphi) \subseteq \operatorname{supp}(\psi_M) \subseteq \cdots \subseteq \operatorname{supp}(\psi_0), \quad \psi_{i-1} = 1 \text{ on } \operatorname{supp}\psi_i, \quad \psi_M = 1 \text{ on } \operatorname{supp}\varphi.$$

Consider $\psi_0 u \in \mathcal{E}'(\mathbb{R}^n)$. Then there exists $t \in \mathbb{R}$ for which $\varphi_0 u \in H^t(\mathbb{R}^n)$. We compute

$$p(D)(\psi_1 u) = \psi_1 p(D)u + [p(D), \psi_1](u) = \psi_1 f + [p(D), \psi_1](\psi_0 u),$$

where the last equality follows from the fact that $\psi_0 u \equiv u$ on $\operatorname{supp}\psi_1$. Now note that $[p(D), \psi_1]$ is an order $N-1$ differential operator. So we have $\psi_1 f \in H^s(\mathbb{R}^n)$ and $[p(D), \psi_1](\psi_0 u) \in H^{t-N+1}(\mathbb{R}^n)$. Setting $\tilde{A}_1 := \min(s, t-N+1)$ we find that $p(D)(\psi_1 u) \in H^{\tilde{A}_1}(\mathbb{R}^n)$.

Since $|p(\lambda)| \gtrsim \langle\lambda\rangle^N$, we find that

$$\begin{aligned} p(D)(\psi_1 u) \in H^{\tilde{A}_1}(\mathbb{R}^n) &\implies \int_{\mathbb{R}^n} \langle\lambda\rangle^{2\tilde{A}_1} |p(\lambda)\mathcal{F}[\psi_1 u](\lambda)|^2 \,\mathrm{d}\lambda \\ &\implies \int_{\mathbb{R}^n} \langle\lambda\rangle^{2\tilde{A}_1+2N} |\mathcal{F}[\psi_1 u](\lambda)|^2 \,\mathrm{d}\lambda \\ &\implies \psi_1 u \in H^{\tilde{A}_1+N}(\mathbb{R}^n). \end{aligned}$$

Define $A_1 := \tilde{A}_1 + N = \min\{s+N, t+1\}$, then we have shown that $\psi_1 u \in H^{A_1}(\mathbb{R}^n)$. By carrying on inductively, we can show that $\psi_M u \in H^{A_M}(\mathbb{R}^n)$ where $A_M = \min\{s+N, t+M\}$. By choosing M large enough we conclude $\psi_M u \in H^{s+N}(\mathbb{R}^n)$, and since $\psi_M = 1$ on $\operatorname{supp}\varphi$, this also shows that $\varphi u \in H^{s+N}(\mathbb{R}^n)$. Since φ was randomly chosen, it follows that $u \in H^{s+N}_{\mathrm{loc}}(X)$. □

4.2 Fundamental solutions

Definition 4.7. Let $p(D)$ be a partial differential operator, then $E \in \mathcal{D}'(\mathbb{R}^n)$ is called a *fundamental solution* for $p(D)$ if $p(D)E = \delta_0$.

Example 4.8. Let $z = x_1 + ix_2 \in \mathbb{C}$ and define the Cauchy-Riemann operator as $\frac{\partial}{\partial \bar{z}} := \frac{1}{2}\left(\frac{\partial}{\partial x_1} + i\frac{\partial}{\partial x_2}\right)$. It can be shown that $E := \frac{1}{\pi z}$ is a fundamental solution of this equation.

Example 4.9. Let $p(D) = \frac{\partial}{\partial t} - \Delta x$ be the heat operator (where $\Delta = \frac{\partial^2}{\partial x^2}_1 + \cdots + \frac{\partial^2}{\partial x_n^2}$) with coordinates $(t, x) \in \mathbb{R} \times \mathbb{R}^n$. Then it can be shown that

$$E := \begin{cases} (4\pi t)^{-n/2} \exp\left(-\frac{\|x\|}{4t}\right), & t > 0, \\ 0, & t \leqslant 0, \end{cases}$$

is a fundamental solution.

Furthermore, if f has compact support, then $u = E * f$ solves $p(D)u = f$, since in this case

$$p(D)(E * f) = (p(D)E * f) = \delta_0 * f = f.$$

As a guess to construct fundamental solutions, we can use the Fourier transform: we have

$$\begin{aligned} p(D)E = \delta_0 &\implies p(\lambda)\hat{E} = 1 \implies \hat{E} = \frac{1}{p(\lambda)} \\ &\implies \langle E, \varphi\rangle = \langle E, \frac{1}{(2\pi)^n}\mathcal{F}[\widetilde{\mathcal{F}[\varphi]}]\rangle = \frac{1}{(2\pi)^n}\langle \hat{E}, \widetilde{\mathcal{F}[\varphi]}\rangle = \frac{1}{(2\pi)^n}\int \frac{\hat{\varphi}(-\lambda)}{p(\lambda)}\,\mathrm{d}\lambda\,. \end{aligned}$$

Indeed, one can check that this E "works", but the problem is that we have no guarantee that $E \in \mathcal{D}'(\mathbb{R}^n)$, since $p(\lambda)$ may cause problems at its roots. To circumvent this, we have to use a construction called *Hörmander's staircase*. For this, we will first need a lemma. For $x \in \mathbb{R}^n$, we will write $x = (x', x_n)$ with $x' \in \mathbb{R}^{n-1}$.

Lemma 4.10. *For each $\varphi \in \mathcal{D}(\mathbb{R}^n)$ and $\lambda' \in \mathbb{R}^{n-1}$, the function $z \mapsto \hat{\varphi}(\lambda', z)$ is analytic in $z \in \mathbb{C}$. Furthermore, for each $m \in \mathbb{N}_0$ there exists constants $c_m, \delta > 0$ (independent of λ') such that*

$$\left|\hat{\varphi}(\lambda', z)\right| \leqslant c_m(1 + |z|)^{-m} e^{\delta|\operatorname{Im} z|}.$$

Proof. By definition of the Fourier transform and Fubini's theorem, we have

$$\hat{\varphi}(\lambda', z) = \int e^{-i\lambda'\cdot x'} \int e^{-izx_n}\varphi(x', x)\,\mathrm{d}x_n\,\mathrm{d}x'\,.$$

It is easily seen that this function is smooth in z and satisfies the Cauchy-Riemann equations, which means it is analytic.

Integrating by parts we find

$$\begin{aligned} \left|z^m\hat{\varphi}(\lambda', z)\right| &= \left|\int e^{-i\lambda'\cdot x'}\int\left(i\frac{\partial}{\partial x_n}\right)^m e^{-izx_n}\varphi(x', x_n)\,\mathrm{d}x_n\,\mathrm{d}x'\right| \\ &= \left|\int e^{-i\lambda'\cdot x'}\int e^{-izx_n}\left(\frac{\partial^m}{\partial x_n^m}\varphi(x', x_m)\right)\mathrm{d}x_m\,\mathrm{d}x'\right| \\ &\leqslant \iint \left|e^{-izx_n}\right| \cdot \left|\frac{\partial^m}{\partial x_n^m}\varphi(x', x_n)\right|\mathrm{d}x_n\,\mathrm{d}x' \\ &\leqslant c_m e^{\delta|\operatorname{Im} z|}, \end{aligned}$$

where δ is chosen such that $\varphi(x', x_n) = 0$ if $|x_n| > \delta$. □

Now, we can prove the main theorem of this section, which *almost* gives an explicit construction for a fundamental solution:

Theorem 4.11. *Every nonzero constant-coefficient partial differential operator has a fundamental solution.*

Proof. By rotating our coordinate axes, we can assume p takes the form

$$p(\lambda', \lambda_n) = \lambda_n^M + \sum_{m=1}^{M-1} a_m(\lambda')\lambda_n^m,$$

(??) (i.e., we simply write p as a polynomial in λ_n). Fix $\mu' \in \mathbb{R}^{n-1}$, then we can write

$$p(\mu', \lambda_n) = \prod_{i=1}^{M} (\lambda_n - \tau_i(\mu')),$$

where the τ_i are the roots of the polynomial $\lambda_n \mapsto p(\mu', \lambda_n)$. Now, by the pigeonhole principle, there exists a horizontal line $\operatorname{Im}\lambda_n = c(\mu')$ in the region $|\operatorname{Im}\lambda_n| \leqslant M+1$ such that

$$\left|\lambda_n - \tau_i(\mu')\right| > 1 \quad \text{on } \operatorname{Im}\lambda_n = c(\mu') \quad (i = 1, \ldots, m)$$

Therefore, on $\operatorname{Im}(\lambda_n) = c(\mu')$ we have $|p(\lambda', \lambda_n)| \gtrsim 1$.

Since roots of a polynomial vary continuously with its coefficients, we can use the same horizontal line $\operatorname{Im}\lambda_n = c(\mu')$ for all λ' in a (small) neighbourhood $N(\mu')$ of μ'. We can cover all of $\mathbb{R}^{n-1}$ with such neighbourhoods, and by the Heine-Borel theorem, we can extract a locally finite subcover $N_1 = N(\mu_1'), N_2 = N(\mu_2'), \ldots$. Furthermore, we can modify these neighbourhoods so that they are disjoint by defining

$$\Delta_i = N_i \backslash \left(\bigcup_{j=1}^{i-1} \overline{N_j} \right).$$

The Δ_i are all open, disjoint, and satisfy $\mathbb{R}^{n-1} = \cup_i \overline{\Delta_i}$. Now we define

$$\langle E, \varphi \rangle := \frac{1}{(2\pi)^n} \sum_{i=1}^{\infty} \int_{\Delta_i} \int_{\operatorname{Im}\lambda_n = c_i} \frac{\hat{\varphi}(-\lambda' - \lambda_n)}{p(\lambda', \lambda_n)} \, \mathrm{d}\lambda_n \, \mathrm{d}\lambda' .$$

In ES3, it is shown that $E \in \mathcal{D}'(\mathbb{R}^n)$. Furthermore, we have

$$\begin{aligned}
\langle p(D)E, \varphi \rangle = \langle E, p(-D)\varphi \rangle &= \frac{1}{(2\pi)^n} \sum_{i=1}^{\infty} \int_{\Delta_i} \int_{\operatorname{Im}\lambda_n = c_i} \frac{p(\lambda', \lambda_n)}{p(\lambda', \lambda_n)} \hat{\varphi}(-\lambda' - \lambda_n) \, \mathrm{d}\lambda_n \, \mathrm{d}\lambda' \\
&= \frac{1}{(2\pi)^n} \sum_{i=1}^{\infty} \int_{\Delta_i} \int_{\operatorname{Im}\lambda_n = c_i} \hat{\varphi}(-\lambda' - \lambda_n) \, \mathrm{d}\lambda_n \, \mathrm{d}\lambda' \\
&\overset{\star}{=} \frac{1}{(2\pi)^n} \sum_{i=1}^{\infty} \int_{\Delta_i} \int_{\operatorname{Im}\lambda_n = 0} \hat{\varphi}(-\lambda' - \lambda_n) \, \mathrm{d}\lambda_n \, \mathrm{d}\lambda' = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \hat{\varphi}(-\lambda) \, \mathrm{d}\lambda = \varphi(0).
\end{aligned}$$

by the Fourier inversion theorem. Here, $\star$ follows from the Cauchy's theorem and the previous lemma ($\hat{\varphi}$ decays rapidly in the horizontal direction, so taking a contour integral over a rectangle and letting the vertical side go to infinity shows that the integral over $\operatorname{Im}\lambda_n = c_i$ equals the integral over $\operatorname{Im}\lambda_n = 0$). We conclude that $p(D)E = \delta_0$. □

Note that the only nonconstructive part of the theorem is the extraction of a locally finite subcover of the neighbourhoods $N(\mu')$.

4.3 Structure theorem for distributions of compact support

In this section, we will prove that every $u \in \mathcal{E}'(X)$ can be written as a finite sum $u = \sum_\alpha \partial^\alpha f_\alpha$ where the f_α are continuous. The theorem generalises to $u \in \mathcal{D}'(X)$ (the sum can then be infinite, but locally finite), but we will not prove this, since it requires the use of partitions of unity.

We start with a lemma:

Lemma 4.12. *For $u \in \mathcal{E}'(\mathbb{R}^n)$, the Fourier transform $\hat{u} \in \mathcal{S}'(\mathbb{R}^n)$ can be identified with the smooth (real-analytic) function $\lambda \mapsto \langle u(x), e^{-i\lambda\cdot x}\rangle$, which we will denote $\hat{u}(\lambda)$.*

Proof. We will first prove the density of $\mathcal{D}(\mathbb{R}^n)$ in $\mathcal{S}(\mathbb{R}^n)$. Fix $\varphi \in \mathcal{S}(\mathbb{R}^n)$ and $\chi \in \mathcal{D}(\mathbb{R}^n)$ with $\chi = 1$ on $\|x\| \leqslant 1$ and $\chi = 0$ on $\|x\| > 2$. Define $\varphi_m \in \mathcal{D}(\mathbb{R}^n)$ by $\varphi_m(x) := \varphi(x)\chi(x/m)$. We will show that $\varphi_m \to \varphi \in \mathcal{S}(\mathbb{R}^n)$.

For each pair of multi-indices α, β, we have

$$\begin{aligned}\|\varphi - \varphi_m\|_{\alpha,\beta} &= \left\|x^\alpha D^\beta(\varphi - \varphi_m)\right\|_\infty = \left\|x^\alpha D^\beta(\varphi\cdot\{1-\chi(x/m)\})\right\| \\ &= \left\|x^\alpha \sum_{\gamma\leqslant\beta}\binom{\beta}{\gamma}(D^\gamma\varphi)(x)\cdot D^{\beta-\gamma}(1-\chi(x/m))\right\|.\end{aligned}$$

For $\gamma \neq \beta$, the derivative $D^\gamma\varphi$ is bounded uniformly while the derivative $D^{\beta-\gamma}(1-\chi(x/m))$ will converge uniformly to 0 since it will have at least one factor $1/m$. For $\gamma = \beta$, we have

$$\left\|x^\alpha(1-\chi(x/m))D^\beta\varphi\right\|_\infty \leqslant \sup_{\|x\|>M}\left\|x^\alpha D^\beta\varphi\right\| \to 0,$$

since $D^\beta\varphi$ decays rapidly. We conclude that $\|\varphi - \varphi_m\|_{\alpha,\beta} \to 0$, and therefore that $\varphi_m \to \varphi$ in $\mathcal{S}(\mathbb{R}^n)$.

Now, by a Riemann sum argument (like the one we have used in lemma 1.18) we have

$$\langle \hat{u}, \varphi_m\rangle = \langle u, \hat{\varphi}_m\rangle = \left\langle u(x), \int e^{-i\lambda\cdot x}\varphi_m(\lambda)\,\mathrm{d}\lambda\right\rangle \overset{\star}{=} \int\langle u(x), e^{-i\lambda\cdot x}\rangle\varphi_m(\lambda)\,\mathrm{d}\lambda\,,$$

where $\star$ is the Riemann sum argument (here, we need that φ_m has compact support). Now, since $u \in \mathcal{E}'(\mathbb{R}^n)$, there exists a compact K and constants $C', N > 0$ such that

$$\left|\langle u(x), e^{-i\lambda\cdot x}\rangle\right| \leqslant C' \sum_{|\alpha|\leqslant N}\sup_K\left|D_x e^{-i\lambda\cdot x}\right| \leqslant C\langle\lambda\rangle^N,$$

so $\lambda \mapsto \langle u(x), e^{-i\lambda\cdot x}\rangle$ is polynomially bounded, and therefore we can use the dominated convergence theorem to conclude

$$\langle \hat{u}, \varphi\rangle = \lim_{n\to\infty}\langle \hat{u}, \varphi_m\rangle = \lim_{n\to\infty}\int\langle u(x), e^{-i\lambda\cdot x}\rangle\varphi_m(\lambda)\,\mathrm{d}\lambda = \int\langle u(x), e^{-i\lambda\cdot x}\rangle\varphi(\lambda)\,\mathrm{d}\lambda\,,$$

which proves that $\hat{u}$ can be identified with the function $\hat{u}(\lambda) = \langle u(x), e^{-i\lambda\cdot x}\rangle$. □

Furthermore, it is clear that for $u \in \mathcal{E}'(\mathbb{R}^n)$, we have

$$|\hat{u}(\lambda)| \leqslant C\sum_{|\alpha|\leqslant N}\sup_K\left|\partial_x^\alpha e^{-i\lambda x}\right| \lesssim \langle\lambda\rangle^N. \tag{4}$$

Theorem 4.13 (Structure theorem)**.** *For $u \in \mathcal{E}'(X)$, there exists a finite collection $\{f_\alpha\} \subseteq C(X)$ with $\operatorname{supp}(f_\alpha) \subseteq X$ such that $u = \sum_\alpha \partial^\alpha f_\alpha$ in $\mathcal{E}'(X)$.*

Proof. Fix $\rho \in \mathcal{D}(X)$ with $\rho = 1$ on a neighbourhood of u, then we can extend u to $\mathcal{E}'(\mathbb{R}^n)$ by setting $\langle u, \varphi\rangle := \langle u, \rho\varphi\rangle$ (note that $\rho\varphi \in \mathcal{D}(X)$ for all $\varphi \in \mathcal{E}(\mathbb{R}^n)$). Since $\rho\varphi \in \mathcal{S}(\mathbb{R}^n)$, we know there exist $\psi \in \mathcal{S}(\mathbb{R}^n)$ such that

$$\rho\varphi = \mathcal{F}[\mathcal{F}[\psi]] = (2\pi)^n \check{\psi},$$

and therefore

$$\langle u, \rho\varphi\rangle = \langle u, \mathcal{F}[\mathcal{F}[\psi]]\rangle = \langle \hat{u}, \hat{\psi}\rangle.$$

Using the Laplacian $\Delta = \sum_i \partial^i \partial^i$, we can write for any $m \in \mathbb{N}$

$$\hat{\psi} = \langle\lambda\rangle^{-2M} \mathcal{F}\big[(1-\Delta)^M \psi\big],$$

since $\mathcal{F}[(1-\Delta)^m \psi] = (1 + \|\lambda\|^2)^m \hat{\psi} = \langle\lambda\rangle^{2M} \hat{\psi}$.

Plugging this back into our previous equations, we have

$$\langle \hat{u}, \hat{\psi}\rangle = \langle \hat{u}, \langle\lambda\rangle^{-2M} \mathcal{F}[(1-\Delta)^M \psi]\rangle = \big\langle \mathcal{F}[\hat{u}\langle\lambda\rangle^{-2M}], (1-\Delta)^M \psi\big\rangle. \tag{5}$$

Now, by eq. (4), we have $|\hat{u}(\lambda)| \lesssim \langle\lambda\rangle^N$, so we can choose M large enough such that $\hat{u}(\lambda) \cdot \langle\lambda\rangle^{-2M} \in L^1(\mathbb{R}^n)$, and by the dominated convergence theorem, the function

$$f(x) = \frac{1}{(2\pi)^n} \int e^{-i\lambda\cdot x} \langle\lambda\rangle^{-2M} \hat{u}(\lambda)\, \mathrm{d}\lambda$$

is continuous, and it is easily checked that $(2\pi)^n \check{f} = \mathcal{F}[\langle\lambda\rangle^{2M} \hat{u}(\lambda)]$.

Using the fact that $(2\pi)^n \check{\psi} = \rho\varphi$, and going back to eq. (5) we see

$$\langle u, \rho\varphi\rangle = \langle (2\pi)^n \check{f}, (1-\Delta)^M \psi\rangle = \langle \check{f}, (1-\Delta)^M \widetilde{(\rho\varphi)}\rangle = \langle f, (1-\Delta)^M (\rho\varphi)\rangle,$$

where the last step follows from the fact that the Laplacian is reflection invariant.

Expanding the derivatives using the Leibniz rule yields

$$(1-\Delta)^M (\rho\varphi) = \sum_\alpha (-1)^{|\alpha|} \rho_\alpha \partial^\alpha \varphi$$

for suitable $\rho_\alpha \in \mathcal{D}(X)$, and therefore we have

$$\langle u, \varphi\rangle = \sum_\alpha \langle f, (-1)^{|\alpha|} \rho_\alpha \partial^\alpha \varphi\rangle = \Big\langle \sum_\alpha \partial^\alpha(\rho_\alpha f), \varphi \Big\rangle,$$

so $u = \sum_\alpha \partial^\alpha(\rho_\alpha f) = \sum_\alpha \partial^\alpha f_\alpha \in \mathcal{E}'(\mathbb{R}^n)$, where f_α is continuous and $\operatorname{supp}(f_\alpha) = \operatorname{supp}(\rho_\alpha f) \subseteq X$. $\square$

There also exist nonconstructive proofs for the previous theorem using Hahn-Banach.

4.4 Paley-Wiener-Schwartz theorem

We have seen that if $u \in \mathcal{E}'(\mathbb{R}^n)$, then $\hat{u}$ is equivalent to the real-analytic function $\lambda \mapsto \langle u(x), e^{-i\lambda\cdot x}\rangle$ and $|\hat{u}(\lambda)| \lesssim \langle\lambda\rangle^N$ for some $N \geqslant 0$. We consider the complex extension of $\hat{u}$, and we call $\hat{u}(z)$ the *Fourier-Laplace transform* of u. Note that

$$\frac{\partial \hat{u}}{\partial \bar{z}_i} = \langle u, \frac{\partial}{\partial \bar{z}_i} e^{-iz\cdot x}\rangle = 0 \quad (i = 1, \ldots, n),$$

so $\hat{u}(z)$ is complex-analytic in each component z_i.

We can also estimate the size of $\hat{u}(z)$:

Lemma 4.14. *If $u \in \mathcal{E}'(\mathbb{R}^n)$ and $\operatorname{supp}(u) \subseteq \overline{B_\delta}$, then there exist nonnegative constant C, N such that*

$$\|\hat{u}(z)\| \leqslant C(1+\|z\|)^N e^{\delta|\operatorname{Im} z|} \qquad \forall z \in \mathbb{C}^n.$$

Proof. Let $\psi \in C^\infty(\mathbb{R})$ be such that $\psi(\tau) = 1$ for $\tau \geqslant -\frac{1}{2}$ and $\psi(\tau) = 0$ for $\tau \leqslant -1$. Introduce for $\varepsilon > 0$

$$\varphi_\varepsilon(x) := \psi(\varepsilon(\delta - \|x\|)),$$

then we have $\varphi_\varepsilon \in \mathcal{D}(\mathbb{R}^n)$ with $\varphi_\varepsilon(x) = 0$ for $\|x\| \geqslant \delta + \varepsilon^{-1}$, and $\varphi_\varepsilon(x) = 1$ for $\|x\| \leqslant \delta + \frac{1}{2}\varepsilon^{-1}$. In particular, every φ_ε is 1 on a neighbourhood of $\operatorname{supp}(u) \subseteq \overline{B_\delta}$.

Therefore, we have

$$\hat{u}(z) = \langle u(x), e^{-iz\cdot x}\rangle = \langle u(x), \varphi_\varepsilon(x) e^{-iz\cdot x}\rangle,$$

and since $u \in \mathcal{E}'(\mathbb{R}^n)$, by the seminorm condition we have nonnegative C, N such that

$$\hat{u}(z) \leqslant C \sum_{|\alpha| \leqslant N} \sup \left|\partial_x^\alpha \left(\varphi_\varepsilon(x) e^{-iz\cdot x}\right)\right|.$$

By definition we have $\left|\partial_x^\beta \varphi_\varepsilon\right| \lesssim \varepsilon^{|\beta|}$ while $\left|\partial_x^\gamma e^{-iz\cdot x}\right| \lesssim \|z\|^{|\gamma|} e^{(\delta+\varepsilon^{-1})|\operatorname{Im} z|}$ on $\operatorname{supp} \varphi_\varepsilon$.

Applying Leibniz, we obtain

$$|\hat{u}(z)| \leqslant C \sum_{|\beta|+|\gamma| \leqslant N} \|z\|^{|\beta|} e^{|\gamma|} e^{(\delta+\varepsilon^{-1}) \operatorname{Im} z},$$

and this holds for all $\varepsilon > 0$, so we can plug in $\varepsilon = \|z\|$ and obtain the result (since $\operatorname{Im} z/\|z\|$ is bounded). □

So if $u \in \mathcal{E}'(\mathbb{R}^n)$ with $\operatorname{supp}(u) \subseteq \overline{B_\delta}$, then $\hat{u}(z)$ is complex analytic and obeys $|\hat{u}(z)| \leqslant (1 + \|z\|)^N e^{\delta|\operatorname{Im} z|}$. The PWS theorem addresses the converse: if a complex analytic function obeys the estimate we just saw, is it the fourier transform of some distribution?

Theorem 4.15. *(a) If $u \in \mathcal{E}'(\mathbb{R}^n)$ with $\operatorname{ord}(u) = N$ and $\operatorname{supp} u \subseteq \overline{B_\delta}$, then $\hat{u}(z)$ is entire and*

$$|\hat{u}(z)| \lesssim (1+\|z\|)^N e^{\delta|\operatorname{Im} z|}. \tag{6}$$

Conversely, if $U(z)$ is entire and satisfies eq. (6) for some N, then $U = \hat{u}$ for some $u \in \mathcal{E}'(\mathbb{R}^n)$ with $\operatorname{supp}(u) \subseteq \overline{B_\delta}$.

(b) If $u \in \mathcal{D}(\mathbb{R}^n)$ and $\operatorname{supp}(u) \subseteq \overline{B_\delta}$, then for $M = 0, 1, 2, \ldots$ we have

$$|\hat{u}(z)| \lesssim_M (1+\|z\|)^{-M} e^{\delta|\operatorname{Im} z|}. \tag{7}$$

Conversely, if $U(z)$ is entire and satisfies eq. (7), then $U = \hat{u}$ for some $u \in \mathcal{D}(\mathbb{R}^n)$ with $\operatorname{supp}(u) \subseteq \overline{B_\delta}$.

Proof. TODO: write this out (lecture 12) □

4.5 Oscillatory integrals

We will study integrals of the form $\int_{\mathbb{R}^n} e^{i\Phi(x,\vartheta)} a(x,\vartheta)\, d\vartheta$, where Φ is called the *phase function* and a the *amplitude* of the signal. We wil use oscillations from the $e^{i\Phi}$-term to control the growth, while $a(\cdot, \vartheta)$ is allowed to grow modestly with ϑ.

Lemma 4.16 (Riemann-Lebesgue)**.** *If $f \in L^1(\mathbb{R})$, then $\left|\hat{f}(\lambda)\right| \to 0$ when $|\lambda| \to \infty$.*

Proof. Assume $f \in L^1$ is continuous, then setting $x' = x + \pi/\lambda$, we have

$$\begin{aligned}\hat{f}(\lambda) &= \int e^{-i\lambda\cdot x} f(x)\,\mathrm{d}x = \frac{1}{2}\int \left(e^{-i\lambda\cdot x} f(x)\,\mathrm{d}x + e^{-i\lambda\cdot(x+\pi/\lambda)} f(x+\pi/\lambda)\right)\mathrm{d}x \\ &= \frac{1}{2}\int e^{-i\lambda\cdot x}[f(x) - f(x+\pi/\lambda)]\,\mathrm{d}x\end{aligned}$$

Now let $\varepsilon > 0$ and choose $R = R(\varepsilon)$ such that $\int_{\|x\|>R} |f(x) - f(x+\pi/\lambda)|\,\mathrm{d}x < \frac{\varepsilon}{2}$. By the dominated convergence theorem, we can also choose $\lambda = \lambda(\varepsilon, R)$ large enough such that

$$\int_{\|x\|<R} \left|e^{-i\lambda x}[f(x) - f(x+\pi/\lambda)]\right|\mathrm{d}x < \frac{\varepsilon}{2}.$$

So for f we have $\left|\hat{f}(\lambda)\right| < \frac{\varepsilon}{2}$ for $|\lambda|$ large enough, which proves the result for continuous functions.

Now we use that continuous functions are dense in L^1, so for any $f \in L^1$, pick $g \in C(\mathbb{R}) \cap L^1$ such that $\|f - g\|_{L^1} < \frac{\varepsilon}{2}$, then it is easily checked that $\left|\hat{f}(\lambda)\right| \leqslant \|f - g\|_{L^1} + |\hat{g}(\lambda)| < \varepsilon$ for λ sufficiently large, which proves the claim. □

Now suppose $\varphi \in \mathcal{D}(\mathbb{R})$ and $\Phi \in C^\infty(\mathbb{R})$ with Φ' nowhere 0. Consider

$$I(\lambda) := \int e^{i\lambda\Phi(\vartheta)}\varphi(\vartheta)\,\mathrm{d}\vartheta\,.$$

Note that $|\Phi'(\vartheta)| \gtrsim 1$ for $\vartheta \in \operatorname{supp}\varphi$ (since $\operatorname{supp}\varphi$ is compact), and we can write

$$I(\lambda) = \int \frac{1}{i\lambda}\frac{\varphi(\vartheta)}{\Phi'(\vartheta)}\frac{\mathrm{d}}{\mathrm{d}\vartheta}\left(e^{i\lambda\Phi(\vartheta)}\right)\mathrm{d}\vartheta\,,$$

and with repeated integration by parts we find that

$$|I(\lambda)| \leqslant \langle\lambda\rangle^{-N} \quad \text{for any } N \geqslant 0.$$

A natural question is what happens if $\Phi'(\vartheta) = 0$ somewhere.

Lemma 4.17 (Stationary phase)**.** *Suppose $\Phi \in C^\infty(\mathbb{R})$ with $\Phi' \neq 0$ on $\mathbb{R}\backslash\{0\}$ and $\Phi(0) = \Phi'(0) = 0$, $\Phi''(0) \neq 0$. Then, for $\chi \in \mathcal{D}(\mathbb{R})$, we have*

$$\int e^{i\lambda\Phi(\vartheta)}\chi(\vartheta)\,\mathrm{d}\vartheta \lesssim \langle\lambda\rangle^{-1/2}.$$

Proof. TODO: write this (lecture 13, a terrible proof with lots of differentiation and garbage). □