B4.3 Distribution Theory and Fourier Analysis: An Introduction*

Jan Kristensen

Michaelmas Term 2018

1 Lecture 1

1.1 Why Distributions?

There are many reasons to study distributions, but most of them are only really appreciated after the fact. Many physical quantities are naturally *not* defined pointwise. For instance, being able to measure temperature at a given point in space and time is an idealization – see the discussion in R.S. Strichartz's *A Guide to Distribution Theory and Fourier Transforms*, §1. Similarly, in the theory of Lebesgue integration as discussed in the Part A Integration course you encountered L^p functions and also they are not really defined uniquely everywhere, but only almost everywhere. In fact they are strictly speaking not even functions, but equivalence classes of functions under the equivalence relation *equal almost everywhere*. Nonetheless, for $f \in L^p(\mathbb{R}^n)$ and each measurable subset $A \subset \mathbb{R}^n$ the integral

$$\int_A f(x)\,\mathrm{d}x$$

is well-defined and does not depend on the representative used to calculate the integral. Note that if we know that f is continuous, then the integrals $\int_A f(x)\,\mathrm{d}x$ for measurable subsets A of $\mathbb{R}^n$ with, say $\mathcal{L}^n(A) < \infty$, determine $f(x)$ uniquely for *all* $x \in \mathbb{R}^n$. Specifically, we have

$$\frac{1}{\mathcal{L}^n(B_r(x_0))}\int_{B_r(x_0)} f(x)\,\mathrm{d}x \to f(x_0)$$

as $r \to 0^+$ for all $x_0 \in \mathbb{R}^n$. For a general L^p function f, knowing the values of

$$\langle f, \mathbf{1}_A\rangle := \int_A f(x)\,\mathrm{d}x$$

*Notes typed by Grigalius Taujanskas. Thanks to Edward Hart for helpful comments.

for all measurable $A \subset \mathbb{R}^n$ with $\mathcal{L}^n(A) < \infty$ determines $f(x)$ uniquely almost everywhere (and so uniquely as an L^p function). Here

$$\mathbf{1}_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A \end{cases}$$

acts as a *test function*, or *measurement* of f. It turns out that taking very nice test functions here is a good idea that allows us to extend aspects of differential calculus to L^p functions and beyond. This leads to the theory of distributions. But why should we bother?

1.1.1 One Dimensional Wave Equation

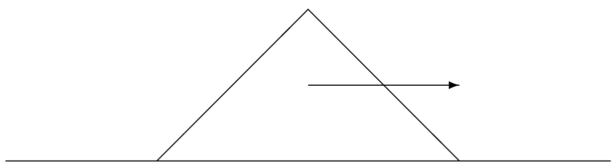
The equation

$$\frac{\partial^2 u}{\partial t^2}(x,t) = k^2 \frac{\partial^2 u}{\partial x^2}(x,t) \tag{1}$$

can be used to model a vibrating string. A function given by

$$u(x,t) = f(x - kt),$$

where f is a function of one variable, represents a travelling wave with shape $f(x)$ moving to the right with velocity k. When f is twice differentiable, one can check that u is a solution to (1). However, there is no physical reason for the shape of the travelling wave to be twice differentiable. For instance, the triangular profile



moving with speed k to the right is perfectly fine! We do not want to throw away physically meaningful solutions because of technicalities. Looking at the example above, one could think that if we accepted as solutions to differential equations any function that satisfies the differential equation except for some points (finitely many, say), where it fails to be differentiable, then all would be fine. But this is too simplistic and does not work, as the next example shows.

1.1.2 Laplace's Equation

On $\mathbb{R}^n$ Laplace's equation is

$$\Delta u := \frac{\partial^2 u}{\partial x_1^2} + \cdots + \frac{\partial^2 u}{\partial x_n^2} = 0. \tag{2}$$

For $n = 2$ or $n = 3$ a solution to the above equation has the physical interpretation of an electric potential in a region with no external charges. From physical experience we know that such potentials should be smooth. However, as you may have seen last year,

$$u = G_2(x_1, x_2) := \log\left(x_1^2 + x_2^2\right),$$

and

$$u = G_3(x_1, x_2, x_3) := \left(x_1^2 + x_2^2 + x_3^2\right)^{-\frac{1}{2}}$$

are solutions in $\mathbb{R}^2 \setminus \{(0,0)\}$ and in $\mathbb{R}^3 \setminus \{(0,0,0)\}$ respectively. Clearly neither can be extended to the origin in a smooth manner, and so these should not be considered as solutions on the full space.

Distribution theory allows us, among many other things, to distinguish between the case of the one dimensional wave equation (1) and Laplace's equation (2). Indeed, the standing wave satisfies the one dimensional wave equation in the sense of distributions for any continuous profile f, while

$$\Delta G_2 = c_2 \delta_0$$

and

$$\Delta G_3 = c_3 \delta_0$$

as distributions, where δ_0 is Dirac's delta function, a distribution.

We must spend some time developing the notion of a *test function* before we can define the notion of a *distribution*. It is worth mentioning that there is *not* just one class of test functions. In this course we will define two classes of test functions, the compactly supported ones and the Schwartz ones. But there are in fact many others, and, depending on the context, they can sometimes be very useful. To each class of test functions there corresponds a class of distributions. The principle to keep in mind here is that the nicer (smoother) the test functions are, the wilder (rougher) the corresponding distributions are allowed to be.

1.2 Test Functions I

Let Ω be a non-empty open subset of $\mathbb{R}^n$. We denote by

$$C(\Omega) := \{u \colon \Omega \to \mathbb{R} \,:\, u \text{ is continuous}\}$$

and sometimes say that such functions are C^0 functions, and write $C^0(\Omega) = C(\Omega)$. We will use the same notation for complex-valued functions, which will be clear from context or will be stated explicitly. Similarly, for $k \in \mathbb{N}$ we define

$$C^k(\Omega) := \{u : \Omega \to \mathbb{R} \,:\, u \text{ is } k \text{ times continuously differentiable}\}.$$

That is, $u \in C^k(\Omega)$ (or u is a C^k function) if and only if u and all its partial derivatives up to and including order k are continuous on Ω.

Example 1.1. The function u is C^1 if and only if u, $\frac{\partial u}{\partial x_1}, \ldots, \frac{\partial u}{\partial x_n}$ are all continuous on Ω. Note in particular that we require, for example, $\frac{\partial u}{\partial x_1}$ to be jointly continuous in $x = (x_1, x_2, \ldots, x_n)$.

Lemma 1.2. *If $u \colon \Omega \to \mathbb{R}$ is C^2, then*

$$\frac{\partial^2 u}{\partial x_j \partial x_k} = \frac{\partial^2 u}{\partial x_k \partial x_j}$$

on Ω. Hence the order in which we take (two) partial derivatives is unimportant for C^2 functions.

Proof. Let $\{e_j\}_{j=1}^n$ be the standard basis for $\mathbb{R}^n$ and denote by

$$\triangle_h u(x) := u(x+h) - u(x)$$

for x, $x+h \in \Omega$. Observe that $\triangle_{se_j}\triangle_{re_k}u = \triangle_{re_k}\triangle_{se_j}u$ for s, $r \in \mathbb{R}$. Because u is C^2, applying the Fundamental Theorem of Calculus twice we have

$$\frac{1}{sr}\triangle_{se_j}\triangle_{re_k}u(x) = \int_0^1\int_0^1 \frac{\partial^2 u}{\partial x_j \partial x_k}(x + \sigma s e_j + \rho r e_k)\,\mathrm{d}\sigma\,\mathrm{d}\rho,$$

and hence

$$\left|\frac{1}{sr}\triangle_{se_j}\triangle_{re_k}u(x) - \frac{\partial^2 u}{\partial x_j \partial x_k}(x)\right| \to 0$$

as $(r, s) \to (0, 0)$. □

We can extend this result to C^k functions for $k \geqslant 2$ by induction, and so for such functions we do not have to worry about the order in which we partially differentiate. When there are many independent variables we shall often rely on multi-index notation.

1.2.1 Multi-index Notation

A multi-index α is an n-tuple of non-negative integers, $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$. The length (or *order*) of α is

$$|\alpha| := \alpha_1 + \cdots + \alpha_n.$$

If α, $\beta \in \mathbb{N}_0^n$, then also $\alpha + \beta \in \mathbb{N}_0^n$. When $u \in C^k(\Omega)$ and $|\alpha| \leqslant k$, we write

$$D^\alpha u(x) := \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}}(x)$$

and by convention set $D^0 u(x) := u(x)$.

Example 1.3. For $\alpha = (1, 2)$, $\beta = (0, 2)$ and $u \in C^3(\Omega)$, where $\Omega \subset \mathbb{R}^2$,

$$D^\alpha u = \frac{\partial^3 u}{\partial x_1 \partial x_2^2}, \qquad D^\beta u = \frac{\partial^2 u}{\partial x_2^2}.$$

When $u \in C^5(\Omega)$,

$$D^{\alpha+\beta} u = \frac{\partial^5 u}{\partial x_1 \partial x_2^4}.$$

Note that by lemma 1.2, $D^{\alpha+\beta}u = D^\alpha(D^\beta u) = D^\beta(D^\alpha u) = D^{\beta+\alpha}u$. In a sense, lemma 1.2 justifies using multi-index notation for partial derivatives.

Note that $C^k(\Omega)$ is a descending sequence of sets, $C^{k+1}(\Omega) \subsetneq C^k(\Omega)$. We define

$$C^\infty(\Omega) := \bigcap_{k=0}^{\infty} C^k(\Omega),$$

the class of infinitely differentiable functions on Ω. Under the natural pointwise definitions of addition, multiplication by scalars, and multiplication, these classes form commutative rings with unity and vector spaces (over $\mathbb{R}$ or $\mathbb{C}$).

1.2.2 Support of a Continuous Function

For $u \in C(\Omega)$ we define the support of u as

$$\operatorname{supp}(u) := \Omega \cap \overline{\{x \in \Omega \ : \ u(x) \neq 0\}},$$

that is, the closure of the set $\{u \neq 0\}$ relative to Ω. As such, $\operatorname{supp}(u)$ is closed in Ω, but need not be closed in $\mathbb{R}^n$.

Example 1.4. Define $u_1 \colon \mathbb{R} \to \mathbb{R}$ by

$$u_1(x) = \begin{cases} 1 - |x|, & |x| < 1, \\ 0, & |x| \geqslant 1. \end{cases}$$

Then $\operatorname{supp}(u_1) = [-1, 1]$.

If instead we consider the restriction of u_1 to $\Omega = (-1, 1)$, that is $u_2(x) = 1 - |x|$, $x \in (-1, 1)$, then $\operatorname{supp}(u_2) = (-1, 1)$.

One sees that the support of a function u depends on the ambient set Ω, and we could instead write $\operatorname{supp}_\Omega(u)$. However, for our purposes it will suffice to write $\operatorname{supp}(u)$, where Ω will be understood from context.

We shall be particularly interested in having compact support. Recall that a set K is compact if and only if any open cover of K admits a finite subcover. Also recall that in $\mathbb{R}^n$, the Heine–Borel theorem tells us that K is compact if and only if K is closed and bounded. Note that $K \subset \Omega$ is compact in Ω if and only if K is compact in $\mathbb{R}^n$. Also when $K \subset \Omega$ is compact, then $\operatorname{dist}(K, \partial\Omega) := \inf_{x \in K,\, y \in \partial\Omega} |x - y| > 0$.

Definition 1.5. Let Ω be an non-empty open subset of $\mathbb{R}^n$. Then

$$\mathcal{D}(\Omega) := \{u \in C^\infty(\Omega) \ : \ \operatorname{supp}(u) \text{ is compact}\}$$

is the class of smooth compactly supported test functions.

Remark 1.6. Note that for $u \in \mathcal{D}(\Omega)$, u not identically zero, we have $\operatorname{dist}(\operatorname{supp}(u), \partial\Omega) > 0$.

We also write

$$C_c^k(\Omega) := \{u \in C^k(\Omega) \ : \ \operatorname{supp}(u) \text{ is compact}\}$$

for $k \in \mathbb{N}_0 \cup \{\infty\}$. So in fact $\mathcal{D}(\Omega) = C_c^\infty(\Omega)$. As before, we can define ring operations in the standard way, making $C_c^k(\Omega)$ and $\mathcal{D}(\Omega)$ into commutative rings (*without* unity) and vector spaces (over $\mathbb{R}$ or $\mathbb{C}$).

We have defined a *test function* to be any smooth and compactly supported function, but so far we have seen no example. Actually, are there any smooth compactly supported test functions other than the trivial one $\phi \equiv 0$?

2 Lecture 2

2.1 Bump Functions

Lemma 2.1. *The function*

$$\phi(x) = \begin{cases} \exp\left(\frac{1}{x^2-1}\right), & |x| < 1, \\ 0, & |x| \geqslant 1 \end{cases}$$

is in $C^\infty(\mathbb{R})$ *with* $\operatorname{supp}(\phi) = [-1, 1]$. *In particular,* $\phi \in \mathcal{D}(\mathbb{R})$.

Sketch of proof. One shows by induction on k that

$$\phi^{(k)}(x) = \frac{p_k(x)}{(x^2-1)^{2k}} \exp\left(\frac{1}{x^2-1}\right)$$

for $|x| < 1$, where $p_k(x)$ is a polynomial of order k. One deduces that $\phi \in C^k(\mathbb{R})$. The rest is left as an exercise. $\square$

Example 2.2. Let $B_r(x_0) := \{x \in \mathbb{R}^n \ : \ |x - x_0| < r\}$ and put

$$\varphi(x) := \phi\left(\frac{|x - x_0|^2}{r^2}\right), \quad x \in \mathbb{R}^n.$$

By the Chain Rule, we see that $\varphi \in C^\infty(\mathbb{R}^n)$. Clearly, $\operatorname{supp}(\varphi) = \overline{B_r(x_0)}$, and so $\varphi \in \mathcal{D}(\mathbb{R}^n)$.

We can do much more using ϕ as a building block. We shall use the operation of *convolution*; recall that when $f, g \in L^1(\mathbb{R}^n)$, then

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

is well-defined for almost every $x \in \mathbb{R}^n$, and $f * g \in L^1(\mathbb{R}^n)$. Furthermore, $f * g = g * f$ almost everywhere.

2.1.1 The Standard Mollifier in $\mathbb{R}^n$

Put

$$c_n := \int_{\mathbb{R}^n} \phi\left(|x|^2\right)\, \mathrm{d}x = \int_0^\infty \phi\left(r^2\right) r^{n-1}\, \mathrm{d}r\, \omega_{n-1} > 0,$$

where ω_{n-1} is the surface area of $\mathbb{S}^{n-1}$. It can be shown that

$$\omega_{n-1} = n\mathcal{L}^n\left(B_1(0)\right) = \frac{2\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)}.$$

We only need to know that $0 < \omega_{n-1} < \infty$, the actual value is not important here. Put

$$\rho(x) := \frac{1}{c_n}\phi\left(|x|^2\right), \quad x \in \mathbb{R}^n.$$

Then $\rho \in \mathcal{D}(\mathbb{R}^n)$, $\rho \geqslant 0$, $\operatorname{supp}(\rho) = \overline{B_1(0)}$ and

$$\int_{\mathbb{R}^n} \rho(x)\, \mathrm{d}x = 1.$$

For each $\varepsilon > 0$ we put

$$\rho_\varepsilon(x) := \varepsilon^{-n}\rho\left(\frac{x}{\varepsilon}\right), \quad x \in \mathbb{R}^n.$$

Then $\rho_\varepsilon \in \mathcal{D}(\mathbb{R}^n)$, $\rho_\varepsilon \geqslant 0$, $\operatorname{supp}\left(\rho_\varepsilon\right) = \overline{B_\varepsilon(0)}$ and

$$\int_{\mathbb{R}^n} \rho_\varepsilon(x)\, \mathrm{d}x = 1.$$

Definition 2.3. We call the family of functions $\left(\rho_\varepsilon\right)_{\varepsilon>0}$ the *standard mollifier* on $\mathbb{R}^n$.

Proposition 2.4. *Let $1 \leqslant p < \infty$ and $u \in L^p(\Omega)$. Define u to be zero outside Ω. Then*

(i) $\rho_\varepsilon * u \in C^\infty(\Omega)$,

(ii) $\|\rho_\varepsilon * u\|_p \leqslant \|u\|_p$, *and*

(iii) $\|\rho_\varepsilon * u - u\|_p \to 0$ *as* $\varepsilon \to 0^+$.

We require the following auxiliary result for the proof.

Lemma 2.5. *Let $1 \leqslant p \leqslant \infty$, $\varphi \in \mathcal{D}(\Omega)$, and $u \in L^p(\Omega)$. Define u to be zero outside of Ω. Then $\varphi * u \in C^1(\Omega)$ and for each $1 \leqslant j \leqslant n$,*

$$\frac{\partial(\varphi * u)}{\partial x_j} = \left(\frac{\partial \varphi}{\partial x_j}\right) * u.$$

Proof. This is a simple application of the Dominated Convergence Theorem. We omit the details. $\square$

Proof of Proposition 2.4. Part *(i)* follows by applying Lemma 2.5 inductively. For part *(ii)*, we use Hölder's inequality. Let

$$\frac{1}{p} + \frac{1}{q} = 1$$

and write for each x and almost every y,

$$|\rho_\varepsilon(x-y)u(y)| = \rho_\varepsilon(x-y)^{\frac{1}{q}}\rho_\varepsilon(x-y)^{\frac{1}{p}}|u(y)|.$$

Integrating over $y \in \mathbb{R}^n$ and using Hölder's inequality,

$$\begin{aligned}\int_{\mathbb{R}^n} |\rho_\varepsilon(x-y)u(y)|\,\mathrm{d}y &\leqslant \left(\int_{\mathbb{R}^n} \rho_\varepsilon(x-y)\,\mathrm{d}y\right)^{\frac{1}{q}} \left(\int_{\mathbb{R}^n} \rho_\varepsilon(x-y)|u(y)|^p\,\mathrm{d}y\right)^{\frac{1}{p}} \\ &= \left(\int_{\mathbb{R}^n} \rho_\varepsilon(x-y)|u(y)|^p\,\mathrm{d}y\right)^{\frac{1}{p}}.\end{aligned}$$

Integrating over $x \in \Omega$,

$$\begin{aligned}\int_\Omega |(\rho_\varepsilon * u)(x)|^p\,\mathrm{d}x &\leqslant \int_\Omega \int_{\mathbb{R}^n} \rho_\varepsilon(x-y)|u(y)|^p\,\mathrm{d}y\,\mathrm{d}x \\ &\overset{\dagger}{=} \int_{\mathbb{R}^n} |u(y)|^p \int_\Omega \rho_\varepsilon(x-y)\,\mathrm{d}x\,\mathrm{d}y \\ &\leqslant \int_{\mathbb{R}^n} |u(y)|^p \int_{\mathbb{R}^n} \rho_\varepsilon(x-y)\,\mathrm{d}x\,\mathrm{d}y = \|u\|_p^p,\end{aligned}$$

where in $\dagger$ we used Fubini–Tonelli. For *(iii)*, let $\tau > 0$ and take $v \in C_c(\Omega)$ such that $\|u-v\|_p \leqslant \tau$ (that this is possible follows from the way we defined the integral in Part A Integration–you might want to write out the details of this). By uniform continuity of v, we can find an $\varepsilon_0 > 0$ such that

$$\|\rho_\varepsilon * v - v\|_\infty < \tau$$

for $0 < \varepsilon \leqslant \varepsilon_0$. Using Minkowski's inequality, we have for $0 < \varepsilon \leqslant \varepsilon_0$

$$\begin{aligned}\|\rho_\varepsilon * u - u\|_p &\leqslant \|\rho_\varepsilon * (u-v)\|_p + \|\rho_\varepsilon * v - v\|_p + \|v-u\|_p \\ &\overset{\textit{(ii)}}{\leqslant} 2\|v-u\|_p + \|\rho_\varepsilon * v - v\|_p \\ &< 2\tau + \|\rho_\varepsilon * v - v\|_\infty \mathcal{L}^n\left(\overline{B_\varepsilon(\operatorname{supp}(v))}\right)^{\frac{1}{p}} \\ &< \left(2 + \mathcal{L}^n\left(\overline{B_\varepsilon(\operatorname{supp}(v))}\right)^{\frac{1}{p}}\right)\tau.\end{aligned}$$

□

We are now ready to prove two useful technical results.

2.1.2 Cut-off Functions and Partitions of Unity

Theorem 2.6. *Let K be a compact subset of Ω. There exists $\phi \in \mathcal{D}(\Omega)$ such that $0 \leqslant \phi \leqslant 1$ and $\phi \equiv 1$ on K. We refer to ϕ as a* cut-off function *between K and $\mathbb{R}^n \setminus \Omega$.*

Proof. Put $d := \operatorname{dist}(K, \partial\Omega) > 0$ and fix $\delta \in \left(0, \frac{d}{4}\right]$. Put $\tilde{K} = \overline{B_{2\delta}(K)}$. Recall that by definition,

$$\tilde{K} = \{x \in \mathbb{R}^n \ : \ \operatorname{dist}(x, K) \leqslant 2\delta\}.$$

Let $(\rho_\varepsilon)_{\varepsilon>0}$ be the standard mollifier and put $\phi := \rho_\delta * \mathbf{1}_{\tilde{K}}$. Then $\phi \in C^\infty(\mathbb{R}^n)$, $\operatorname{supp}(\phi) \subset \overline{B_\delta(\tilde{K})} = \overline{B_{3\delta}(K)}$, and since $\delta \leqslant \frac{d}{4}$, then $\operatorname{supp}(\phi) \subset \Omega$ and hence $\phi \in \mathcal{D}(\Omega)$. Next, $0 \leqslant \phi \leqslant 1$, and for $x \in K$ we have $\overline{B_\delta(x)} \subset \tilde{K}$, so

$$\phi(x) = \int_{\mathbb{R}^n} \rho_\delta(x-y)\mathbf{1}_{\tilde{K}}(y)\,\mathrm{d}y = \int_{\mathbb{R}^n} \rho_\delta(x-y)\,\mathrm{d}y = 1.$$

Note that $\rho_\delta(x-y)$ is supported in $\overline{B_\delta(x)}$. □

Remark 2.7. For a multi-index α we have

$$\begin{aligned}|D^\alpha\phi(x)| &= \left|\int_{\mathbb{R}^n} \delta^{-|\alpha|}(D^\alpha\rho)_\delta(x-y)\mathbf{1}_{\tilde{K}}(y)\,\mathrm{d}y\right| \\ &\leqslant \delta^{-|\alpha|}\int_{\mathbb{R}^n} |(D^\alpha\rho)_\delta(x-y)|\,\mathrm{d}y \\ &= \delta^{-|\alpha|}\|D^\alpha\rho\|_{L^1},\end{aligned}$$

hence

$$|D^\alpha\phi| \leqslant c_\alpha d^{-|\alpha|},$$

where $c_\alpha = 4^{|\alpha|}\|D^\alpha\rho\|_{L^1}$, a constant independent of d.

Theorem 2.8. *Let $\Omega = \bigcup_{j=1}^m \Omega_j$, where $\Omega_1, \ldots, \Omega_m$ are open, non-empty, potentially overlapping sets. For $K \subset \Omega$ compact there exist $\phi_1, \ldots, \phi_m \in \mathcal{D}(\Omega)$ satisfying $\operatorname{supp}(\phi_j) \subset \Omega_j$, $0 \leqslant \phi_j \leqslant 1$, and*

$$\sum_{j=1}^m \phi_j = 1$$

on K. We refer to $\phi_1, \ldots, \phi_m$ as a partition of unity *on K subordinate to the cover $\Omega_1, \ldots, \Omega_m$.*

Proof. (*Not examinable*) Let $x \in K \cap \Omega_j$. Because Ω_j is open, we can find $r_j(x) > 0$ such that $\overline{B_{r_j(x)}(x)} \subset \Omega_j$. The set

$$\left\{B_{r_j(x)}(x) \ : \ x \in K,\ 1 \leqslant j \leqslant m\right\}$$

is an open cover of K, so by compactness it admits a finite subcover, say

$$\left\{B_s := B_{r_{j_s}(x_s)}(x_s),\ 1 \leqslant s \leqslant N\right\}.$$

Put $J_j := \{s \ : \ j_s(x_s) = j\}$, so that

$$\bigcup_{s\in J_j} \overline{B_s} \subset \Omega_j.$$

Now $K_j = K \cap \left(\bigcup_{s\in J_j} \overline{B_s}\right)$ is compact, $K_j \subset \Omega_j$ and $K = \bigcup_{j=1}^m K_j$. We now apply Theorem 2.6 to each K_j, Ω_j to find corresponding cut-off functions $\psi_j \in \mathcal{D}(\Omega_j)$ satisfying $0 \leqslant \psi_j \leqslant 1$ and $\psi_j \equiv 1$ on K_j. We extend ψ_j to $\Omega \setminus \Omega_j$ by zero and, denoting this extension again by ψ_j, have $\psi_j \in \mathcal{D}(\Omega)$. Now define $\phi_1 := \psi_1$, $\phi_2 := \psi_2(1-\psi_1)$, $\dots$, $\phi_m := \psi_m \prod_{j=1}^{m-1}(1-\psi_j)$. By repeated use of the Leibniz rule we see that $\phi_1, \dots, \phi_m \in C^\infty(\Omega)$. Clearly $\mathrm{supp}(\phi_j) \subset \Omega_j$, and $0 \leqslant \phi_j \leqslant 1$. Finally, on K we have

$$\sum_{j=1}^m \phi_j - 1 = -\prod_{j=1}^m (1-\psi_j) = 0.$$

□

The proof of Theorem 2.8 is not examinable, but the result is.

3 Lecture 3

3.1 Convergence of Sequences in $\mathcal{D}(\Omega)$

Before defining distributions corresponding to smooth compactly supported test functions we must first discuss a notion of convergence in $\mathcal{D}(\Omega)$.

Definition 3.1. Let $\{\phi_j\}_j$ be a sequence in $\mathcal{D}(\Omega)$ and $\phi \in \mathcal{D}(\Omega)$. We say

$$\phi_j \to \phi \quad \text{in } \mathcal{D}(\Omega)$$

if and only if there exists a compact set $K \subset \Omega$ such that $\mathrm{supp}(\phi)$, $\mathrm{supp}(\phi_j) \subset K$ for all j, and for each multi-index α

$$\sup_K |D^\alpha(\phi_j - \phi)| \to 0.$$

In words, $\phi_j \to \phi$ in $\mathcal{D}(\Omega)$ if and only if all supports are contained in a fixed compact set in Ω, and we have uniform convergence of $\phi_j - \phi$ together with all partial derivatives to the zero function.

Remark 3.2. Convergence in $\mathcal{D}(\Omega)$ is a strong requirement. The requirement of all supports being contained in a fixed compact set is needed to ensure that $\phi(x-j)$ does not converge to zero in $\mathcal{D}(\mathbb{R})$.

Remark 3.3. It is possible to define a topology $\mathcal{T}$ on $\mathcal{D}(\Omega)$ in such a way that $\phi_j \to \phi$ in $\mathcal{D}(\Omega)$ corresponds to $\phi_j \to \phi$ in the topological space $(\mathcal{D}(\Omega), \mathcal{T})$. We shall not pursue this point of view in these notes, however, apart from noting that it can be shown that the topology $\mathcal{T}$ is not metrizable.

3.2 Distributions Corresponding to $\mathcal{D}(\Omega)$

Definition 3.4. A functional $u\colon \mathcal{D}(\Omega) \to \mathbb{R}$ (or $\mathbb{C}$) is a *distribution* on Ω if and only if

(i) u is linear,

$$u(\phi + t\psi) = u(\phi) + tu(\psi)$$

for ϕ, $\psi \in \mathcal{D}(\Omega)$, $t \in \mathbb{R}$ (or $\mathbb{C}$), and

(ii) u is continuous in the sense that $u(\phi_j) \to u(\phi)$ whenever $\phi_j \to \phi$ in $\mathcal{D}(\Omega)$.

The set of all distributions on Ω is denoted by $\mathcal{D}'(\Omega)$.

Remark 3.5. Firstly, because of linearity, the continuity condition (ii) holds if and only if it holds at $\phi = 0$. Indeed, if $u(\phi_j) \to 0$ whenever $\phi_j \to 0$ in $\mathcal{D}(\Omega)$ and $\psi_j \to \psi$ in $\mathcal{D}(\Omega)$, then we take $\phi_j = \psi_j - \psi$ and note that $\phi_j \to 0$ in $\mathcal{D}(\Omega)$. Then by assumption, $u(\phi_j) \to 0$. But u is linear, so $u(\phi_j) = u(\psi_j) - u(\psi)$ and so $u(\psi_j) \to u(\psi)$.

Secondly, when $u : \mathcal{D}(\Omega) \to \mathbb{R}$ is linear (and defined everywhere on $\mathcal{D}(\Omega)$), then chances are that u is continuous in the sense defined above and thus is a distribution on Ω. Indeed, the only counterexamples I know are constructed by use of the Axiom of Choice in the form of existence of a Hamel basis for $\mathcal{D}(\Omega)$.

Notation. When $u \in \mathcal{D}'(\Omega)$ and $\phi \in \mathcal{D}(\Omega)$, we often write $\langle u, \phi\rangle$ instead of $u(\phi)$.

Example 3.6. If $f \in L^p(\Omega)$, $p \in [1, \infty]$, then

$$\langle T_f, \phi\rangle = \int_\Omega f(x)\phi(x)\,\mathrm{d}x, \quad \phi \in \mathcal{D}(\Omega)$$

defines a distribution on Ω. Linearity follows from linearity of the integral, and continuity follows from the Dominated Convergence Theorem. Note that since each $\phi \in \mathcal{D}(\Omega)$ has compact support in Ω and since we defined convergence in $\mathcal{D}(\Omega)$ by requiring all supports to be in a fixed compact set in Ω, the above distribution T_f would also be well-defined if f was only locally in L^p.

3.3 Local Lebesgue Spaces

Definition 3.7. For $p \in [1, \infty]$ we write $f \in L^p_{\mathrm{loc}}(\Omega)$ and say f is locally L^p on Ω if and only if for each compact set $K \subset \Omega$ we have $f|_K \in L^p(K)$. Specifically, we require $\int_K |f|^p \,\mathrm{d}x < \infty$ when $p < \infty$ and $\operatorname{ess\,sup}_K |f| < \infty$ when $p = \infty$.

Example 3.8. The function $x^{-1} \notin L^1(0, \infty)$, but $x^{-1} \in L^1_{\mathrm{loc}}(0, \infty)$ and, in fact, $x^{-1} \in L^p_{\mathrm{loc}}(0, \infty)$ for all $p \in [1, \infty]$. Note that Ω determines what *local* means. For example, $x^{-1} \in L^1_{\mathrm{loc}}(0, \infty)$, but $x^{-1} \notin L^1_{\mathrm{loc}}(-1, 1)$.

Example 3.9. Summarizing a previous discussion, each $f \in L^p_{\mathrm{loc}}(\Omega)$, $p \in [1, \infty]$, gives rise to a distribution on Ω via

$$\langle T_f, \phi \rangle = \int_\Omega f(x)\phi(x)\,\mathrm{d}x$$

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

Example 3.10 (Dirac's delta function at $x_0 \in \Omega$)*.* The map

$$\phi \mapsto \langle \delta_{x_0}, \phi \rangle := \phi(x_0)$$

for $\phi \in \mathcal{D}(\Omega)$ is clearly a distribution on Ω. Furthermore, so is $\phi \mapsto (D^\alpha \phi)(x_0)$ for any multi-index α.

While the continuity condition (ii) in Definition 3.4 often is not an issue, it is nonetheless useful to reformulate it using linearity as follows.

Theorem 3.11. *A linear functional $u \colon \mathcal{D}(\Omega) \to \mathbb{R}$ (or $\mathbb{C}$) is a distribution if and only if for every compact set $K \subset \Omega$ there exist constants $c = c(K) > 0$ and $m = m(K) \in \mathbb{N}_0$ such that*

$$|\langle u, \phi \rangle| \leqslant c \sum_{|\alpha| \leqslant m} \sup_K |D^\alpha \phi| \tag{3}$$

for all $\phi \in \mathcal{D}(K) := \{\phi \in \mathcal{D}(\Omega) \ : \ \operatorname{supp}(\phi) \subset K\}$.

Proof. If $\phi_j \to 0$ in $\mathcal{D}(\Omega)$, then for some compact set $K \subset \Omega$ we have $\phi_j \in \mathcal{D}(K)$ for all j. Then by assumption we can find $c = c(K) > 0$ and $m = m(K) \in \mathbb{N}_0$ such that (3) holds. But then

$$|\langle u, \phi_j \rangle| \leqslant c \sum_{|\alpha| \leqslant m} \sup_K |D^\alpha \phi_j| \to 0.$$

For the converse, we argue by contradiction. Assume there exists $u \in \mathcal{D}(\Omega)$ and a compact set $K \subset \Omega$ such that (3) is violated for all choices of c and m. In particular, for $c = m = j$ we can find $\phi_j \in \mathcal{D}(K)$ with

$$|\langle u, \phi_j \rangle| > j \sum_{|\alpha| \leqslant j} \sup_K |D^\alpha \phi_j|.$$

Put $\lambda_j = \langle u, \phi_j \rangle$. Then $|\lambda_j| > 0$, $\psi_j := \frac{\phi_j}{\lambda_j} \in \mathcal{D}(K)$, $\langle u, \psi_j \rangle = 1$, and

$$1 > j \sum_{|\alpha| \leqslant j} \sup_K |D^\alpha \psi_j|.$$

Thus $|D^\alpha \psi_j| < j^{-1}$ on Ω for $j \geqslant |\alpha|$, and in particular $\psi_j \to 0$ in $\mathcal{D}(\Omega)$. But $\langle u, \psi_j \rangle = 1$, which does not converge to zero. $\square$

4 Lecture 4

4.1 The Fundamental Lemma of the Calculus of Variations

When $f \in L^p_{\mathrm{loc}}(\Omega)$, then

$$\langle T_f, \phi \rangle = \int_\Omega f(x)\phi(x)\,\mathrm{d}x,$$

$\phi \in \mathcal{D}(\Omega)$, defines a distribution on Ω. It is natural to ask if the distribution T_f determines f, that is, if for $f, g \in L^p_{\mathrm{loc}}(\Omega)$ we have $T_f = T_g$, must it be the case that $f = g$ almost everywhere? The answer is affirmative and relies on the following.

Lemma 4.1 (The Fundamental Lemma of the Calculus of Variations)**.** *If $f \in L^1_{loc}(\Omega)$ and*

$$\int_\Omega f(x)\phi(x)\,\mathrm{d}x = 0$$

for all $\phi \in \mathcal{D}(\Omega)$, then $f = 0$ almost everywhere.

Proof. Let $\mathcal{O}$ be a non-empty open subset of Ω such that $\overline{\mathcal{O}}$ is compact and $\overline{\mathcal{O}} \subset \Omega$. In this case we write $\mathcal{O} \Subset \Omega$.

Put $g = f\mathbf{1}_\mathcal{O}$ and extend g to $\mathbb{R}^n \setminus \Omega$ by zero. Because $\overline{\mathcal{O}} \subset \Omega$ is compact, we have $g \in L^1(\mathbb{R}^n)$. For the standard mollifier $(\rho_\varepsilon)_{\varepsilon>0}$ we know, by proposition 2.4, that $\|\rho_\varepsilon * g - g\|_1 \to 0$ as $\varepsilon \to 0^+$. Now note that for $x \in \Omega$,

$$\begin{aligned}(\rho_\varepsilon * g)(x) &= \int_{\mathbb{R}^n} \rho_\varepsilon(x-y)g(y)\,\mathrm{d}y \\ &= \int_\mathcal{O} \rho_\varepsilon(x-y)f(y)\,\mathrm{d}y.\end{aligned}$$

If we take $x \in \mathcal{O}$ and $\varepsilon \in (0, \operatorname{dist}(x, \partial\mathcal{O}))$, then, denoting $\phi^x(y) := \rho_\varepsilon(x-y)$ for $y \in \Omega$, we have $\phi^x \in C^\infty(\Omega)$ and $\operatorname{supp}(\phi^x) = \overline{B_\varepsilon(x)} \subset \mathcal{O} \subset \Omega$, so $\phi^x \in \mathcal{D}(\Omega)$. By assumption,

$$0 = \int_\Omega f(y)\phi^x(y)\,\mathrm{d}y = \int_\mathcal{O} f(y)\rho_\varepsilon(x-y)\,\mathrm{d}y = (\rho_\varepsilon * g)(x).$$

It follows that $(\rho_\varepsilon * g)(x) \to 0$ as $\varepsilon \to 0^+$ pointwise in $x \in \mathcal{O}$. From Fatou's Lemma, we therefore get

$$\begin{aligned}\int_\mathcal{O} |g|\,\mathrm{d}x &\leqslant \liminf_{\varepsilon\to 0^+} \int_\mathcal{O} |\rho_\varepsilon * g - g|\,\mathrm{d}x \\ &\leqslant \lim_{\varepsilon\to 0^+} \int_\Omega |\rho_\varepsilon * g - g|\,\mathrm{d}x = 0.\end{aligned}$$

Thus $f = g = 0$ almost everywhere in $\mathcal{O}$ and since $\mathcal{O} \Subset \Omega$ was arbitrary, we conclude that $f = 0$ almost everywhere. $\square$

Notation. When $f \in L^p_{\mathrm{loc}}(\Omega)$, we shall also use f to denote the distribution T_f. This is of course an abuse of notation, but it is convenient and should not cause too much trouble. We shall often refer to distributions that correspond to an $L^p_{\mathrm{loc}}(\Omega)$ function as a *regular distribution* on Ω.

Definition 4.2. Let $u \in \mathcal{D}'(\Omega)$. If there exists an $m \in \mathbb{N}_0$ with the property that for all compact subsets $K \subset \Omega$ there exists a constant $c = c_K > 0$ such that

$$|\langle u, \phi\rangle| \leqslant c \sum_{|\alpha| \leqslant m} \sup_K |D^\alpha \phi|$$

for all $\phi \in \mathcal{D}(K)$, then we say u has *order at most* m.

We say u *has order* m if and only if u has order at most m, but not order at most $m-1$. In particular, u has order 0 if and only if it has order at most 0.

We say u has order infinity if and only if u does not have order at most m for any $m \in \mathbb{N}_0$.

Note that by Theorem 3.11, any distribution has locally finite order.

Example 4.3. Let $f \in L^1_{\mathrm{loc}}(\Omega)$. Then the corresponding distribution has order 0. Indeed, if $K \subset \Omega$ is compact and $\varphi \in \mathcal{D}(K)$, then

$$\begin{aligned} |\langle f, \varphi\rangle| &= \left| \int_\Omega f(x)\varphi(x)\,\mathrm{d}x \right| \\ &\leqslant \int_K |f||\varphi|\,\mathrm{d}x \\ &\leqslant \sup_K |\varphi| \int_K |f|\,\mathrm{d}x, \end{aligned}$$

and so $|\langle f, \varphi\rangle| \leqslant c \sup_K |\varphi|$.

Example 4.4. Let $x_0 \in \Omega$ and $\alpha \in \mathbb{N}_0^n$. Define

$$\langle T, \varphi\rangle := (D^\alpha \varphi)(x_0)$$

for $\varphi \in \mathcal{D}(\Omega)$. Then $T \in \mathcal{D}'(\Omega)$ as T is clearly linear and for compact $K \subset \Omega$ and $\varphi \in \mathcal{D}(K)$,

$$|\langle T, \varphi\rangle| = |(D^\alpha \varphi)(x_0)| \leqslant \sup_K |D^\alpha \varphi|.$$

It also shows that T has order at most $|\alpha|$. If $\alpha = 0$ so that $T = \delta_{x_0}$, we see that T has order 0. Assume $|\alpha| > 0$. We shall prove that T has order $|\alpha|$. Suppose, for contradiction, that T has order at most $|\alpha| - 1$. Take $r \in (0, \operatorname{dist}(x_0, \partial\Omega))$ and put $K = \overline{B_r(x_0)}$. Then $K \subset \Omega$ is compact. By assumption, we can then find $c = c_K > 0$ such that

$$|\langle T, \varphi\rangle| = |(D^\alpha \varphi)(x_0)| \leqslant c \sum_{|\beta| \leqslant |\alpha|-1} \sup_K |D^\beta \varphi| \tag{4}$$

for all $\varphi \in \mathcal{D}(K)$. Take $\psi \in \mathcal{D}(B_1(0))$ with $\psi(0) = 1$ and define, for $\varepsilon \in (0, r)$,

$$\varphi(x) := \frac{(x-x_0)^\alpha}{\alpha!}\psi\left(\frac{x-x_0}{\varepsilon}\right)$$

for $x \in \Omega$. Note that φ is C^∞ and $\operatorname{supp}(\varphi) \subset \overline{B_\varepsilon(x_0)} \subset K$, so that $\varphi \in \mathcal{D}(K)$. Also,

$$D^\beta\left(\frac{(x-x_0)^\alpha}{\alpha!}\right)\Bigg|_{x=x_0} = \begin{cases} 1 \text{ if } \beta = \alpha \\ 0 \text{ if } \beta \neq \alpha \end{cases},$$

so that $D^\alpha\varphi(x_0) = 1$. If $\beta \in \mathbb{N}_0^n$, $\beta \leqslant \alpha$, and $|\beta| < |\alpha|$, then for $x \in B_\varepsilon(x_0)$ we get from the generalized Leibniz rule (see below)

$$\begin{aligned}
|D^\beta\varphi(x_0)| &\leqslant \sum_{\gamma\leqslant\beta}\binom{\beta}{\gamma}\left|D_x^\gamma\left(\frac{(x-x_0)^\alpha}{\alpha!}\right)\right|\left|D_x^{\beta-\gamma}\psi\left(\frac{x-x_0}{\varepsilon}\right)\right| \\
&\leqslant \sum_{\gamma\leqslant\beta}\binom{\beta}{\gamma}c_\gamma|x-x_0|^{|\alpha|-|\gamma|}\sup|D^{\beta-\gamma}\psi|\varepsilon^{|\gamma|-|\beta|} \\
&\leqslant \sum_{\gamma\leqslant\beta}\binom{\beta}{\gamma}c_\gamma\sup|D^{\beta-\gamma}\psi|\varepsilon^{|\alpha|-|\gamma|+|\gamma|-|\beta|} \\
&= c_{\psi,\beta}\varepsilon^{|\alpha|-|\beta|},
\end{aligned}$$

where

$$c_{\psi,\beta} := \sum_{\gamma\leqslant\beta}\binom{\beta}{\gamma}c_\gamma\sup|D^{\beta-\gamma}\psi|.$$

When $\varepsilon < 1$ we have, since $|\alpha| - |\beta| \geqslant 1$, that $\varepsilon^{|\alpha|-|\beta|} \leqslant \varepsilon$, and hence in combination with (4) we get

$$1 \leqslant c\sum_{|\beta|\leqslant|\alpha|-1}c_{\psi,\beta}\varepsilon^{|\alpha|-|\beta|} \leqslant c\sum_{|\beta|\leqslant|\alpha|-1}c_{\psi,\beta}\varepsilon =: \tilde{c}\varepsilon.$$

Here $\tilde{c} > 0$ is a constant (in particular, independent of ε) whose value is unimportant. We have shown that $1 \leqslant \tilde{c}\varepsilon$ holds for all $\varepsilon \in (0, \min(1, r))$. This is a contradiction for $\varepsilon < \tilde{c}^{-1}$.

Theorem 4.5 (Generalized Leibniz Rule)**.** *Let f, $g \in C^k(\Omega)$. Then $fg \in C^k(\Omega)$ and for $\alpha \in \mathbb{N}_0^n$, $|\alpha| \leqslant k$, we have*

$$D^\alpha(fg) = \sum_{\beta\leqslant\alpha}\binom{\alpha}{\beta}D^\beta f D^{\alpha-\beta}g,$$

where $\beta \leqslant \alpha$ means $\beta_i \leqslant \alpha_i$ for all $i = 1, \ldots n$,

$$\binom{\alpha}{\beta} := \frac{\alpha!}{\beta!(\alpha-\beta)!},$$

and $\alpha! = \alpha_1!\alpha_2!\ldots\alpha_n!$.

Proof. This can be proven by induction on $|\alpha|$. We omit the details. $\square$

A generalization of the above example is as follows. Let x_j, where $j \in J$ is countable or finite, be distinct points in Ω so that the set $\{x_j \,|\, j \in J\}$ has no limit points in Ω (that is, any limit points are on $\partial\Omega$). For any set of multi-indices $\alpha_j \in \mathbb{N}_0^n$ put

$$\langle T, \varphi\rangle := \sum_{j\in J} (D^{\alpha_j}\varphi)(x_j)$$

for $\varphi \in \mathcal{D}(\Omega)$. Then $T \in \mathcal{D}'(\Omega)$ and the order of T is $\sup_{j\in J} |\alpha_j|$.

5 Lecture 5

Definition 5.1. Let $\{u_j\}_j$ be a sequence in $\mathcal{D}'(\Omega)$ and let $u \in \mathcal{D}'(\Omega)$. We say u_j converges to u in the sense of distributions on Ω and write

$$u_j \longrightarrow u \text{ in } \mathcal{D}'(\Omega)$$

if and only if

$$\langle u_j, \varphi\rangle \longrightarrow \langle u, \varphi\rangle$$

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

Remark 5.2. As with convergence in $\mathcal{D}(\Omega)$, one can define a topology $\mathcal{T}$ on $\mathcal{D}'(\Omega)$ so that $u_j \to u$ in $\mathcal{D}'(\Omega)$ corresponds to $u_j \to u$ in the topological space $(\mathcal{D}'(\Omega), \mathcal{T})$. We will not need it and shall not pursue this here. We note, however, that the topology $\mathcal{T}$ is not metrizable.

Convergence in $\mathcal{D}'(\Omega)$ is very weak.

Example 5.3. Let $p \in [1, \infty]$ and $f_j, f \in L^p(\Omega)$. If $f_j \to f$ in $L^p(\Omega)$, then $f_j \to f$ in $\mathcal{D}'(\Omega)$. This can be proven using the Dominated Convergence Theorem. The converse, however, is false:

(i) Let $f_j(x) = \sin(jx)$, $x \in (0, 1)$. Then $f_j \to 0$ in $\mathcal{D}'(0, 1)$, but $f_j \not\to 0$ in $L^p(0, 1)$ for any $p \in [1, \infty]$.

(ii) Let $g_j(x) = g(jx)$, $x \in (0, 1)$, where g is T-periodic and on $(0, T]$ is given by $g = -117\mathbf{1}_{(0,\frac{T}{2}]} + 117\mathbf{1}_{(\frac{T}{2},T]}$. On Sheet 2 you will be asked to prove that $g_j \to 0$ in $\mathcal{D}'(0, 1)$, but $\|g_j\|_1 = 117 \not\to 0$.

Example 5.4. Let $v \in C_c(\mathbb{R}^n)$ and for $x_0 \in \Omega$ and $\varepsilon > 0$ put

$$v_\varepsilon(x) := \varepsilon^{-n} v\left(\frac{x - x_0}{\varepsilon}\right),$$

$x \in \Omega$. Then $v_\varepsilon \in \mathcal{D}'(\Omega)$ and $v_\varepsilon \to \int_{\mathbb{R}^n} v \, \mathrm{d}x \, \delta_{x_0}$ in $\mathcal{D}'(\Omega)$ as $\varepsilon \to 0^+$. Indeed, it is clear that $v_\varepsilon \in \mathcal{D}'(\Omega)$ for all $\varepsilon > 0$, and for $\varphi \in \mathcal{D}(\Omega)$ we have

$$\begin{aligned} \langle v_\varepsilon, \varphi \rangle &= \int_\Omega \varepsilon^{-n} v \left(\frac{x - x_0}{\varepsilon} \right) \varphi(x) \, \mathrm{d}x \\ &= \int_{\mathbb{R}^n} v(y) \varphi(x_0 + \varepsilon y) \, \mathrm{d}y \\ &\underset{\varepsilon \to 0^+}{\longrightarrow} \int_{\mathbb{R}^n} v(y) \varphi(x_0) \, \mathrm{d}y = \int_{\mathbb{R}^n} v \, \mathrm{d}y \, \langle \delta_{x_0}, \varphi \rangle, \end{aligned}$$

where in the second line we made the change of variables $y = \varepsilon^{-1}(x - x_0)$. In particular, note that if $(\rho_\varepsilon)_{\varepsilon > 0}$ is the standard mollifier, then $\rho_\varepsilon \to \delta_0$ in $\mathcal{D}'(\mathbb{R}^n)$ as $\varepsilon \to 0^+$.

Note that for $x \in \mathbb{R}^n$, $\phi^x(y) := \rho_\varepsilon(x - y)$, $y \in \mathbb{R}^n$, is C^∞ and has support $\operatorname{supp}(\phi^x) = \overline{B_\varepsilon(x)}$, so in particular $\phi^x \in \mathcal{D}(\Omega)$. Now recall that for $f \in L^p(\mathbb{R}^n)$ we defined

$$\begin{aligned} (\rho_\varepsilon * f)(x) &= \int_{\mathbb{R}^n} \rho_\varepsilon(x - y) f(y) \, \mathrm{d}y \\ &= \int_{\mathbb{R}^n} \phi^x(y) f(y) \, \mathrm{d}y = \langle f, \phi^x \rangle. \end{aligned}$$

This in fact all makes perfect sense for any $f \in L^p_{\mathrm{loc}}(\mathbb{R}^n)$, but we can go much further.

Definition 5.5 (Mollification of distributions)**.** Let $u \in \mathcal{D}'(\mathbb{R}^n)$ and $(\rho_\varepsilon)_{\varepsilon > 0}$ be the standard mollifier. Then $\rho_\varepsilon * u \in \mathcal{D}'(\mathbb{R}^n)$ is defined by

$$\langle \rho_\varepsilon * u, \varphi \rangle = \langle u, \tilde{\rho}_\varepsilon * \varphi \rangle,$$

$\varphi \in \mathcal{D}(\mathbb{R}^n)$, where $\tilde{\rho}_\varepsilon(x) = \rho_\varepsilon(-x)$, which is equal to $\rho_\varepsilon(x)$ in the case of the standard mollifier since it is an even function.

Remark 5.6. $\rho_\varepsilon * \varphi \in \mathcal{D}(\mathbb{R}^n)$ since clearly $\rho_\varepsilon * \varphi$ is C^∞ and $\operatorname{supp}(\rho_\varepsilon * \varphi) \subset \overline{B_{2\varepsilon}(\operatorname{supp}(\varphi))}$.

From Lemma 2.5, we get

$$D^\alpha(\rho_\varepsilon * \varphi) = \rho_\varepsilon * D^\alpha \varphi$$

by applying it $|\alpha|$ times, and it is easy to check that

$$\sup |D^\alpha(\rho_\varepsilon * \varphi)| \leqslant \sup |D^\alpha \varphi|,$$

so $\rho_\varepsilon * u$ will satisfy the same bounds as u (uniformly in $\varepsilon > 0$). But for each fixed $\varepsilon > 0$, $\rho_\varepsilon * u$ is much better, as the following theorem shows.

Theorem 5.7. *If $u \in \mathcal{D}'(\mathbb{R}^n)$, then $\rho_\varepsilon * u \in C^\infty(\mathbb{R}^n)$ and $\rho_\varepsilon * u \to u$ in $\mathcal{D}'(\mathbb{R}^n)$.*

Sketch of proof. For $\varepsilon > 0$ fixed we consider

$$\phi^x(y) := \rho_\varepsilon(x-y)$$

for $y \in \mathbb{R}^n$. Since $\phi^x \to \phi^{x_0}$ in $\mathcal{D}(\mathbb{R}^n)$ as $x \to x_0$, we have by the continuity condition

$$\langle u,\, \phi^x\rangle \longrightarrow \langle u,\, \phi^{x_0}\rangle$$

as $x \to x_0$. Thus $\rho_\varepsilon * u \in C^0(\mathbb{R}^n)$. Next, for a direction $1 \leqslant j \leqslant n$ and an increment $h \neq 0$ consider

$$\frac{1}{h}\left((\rho_\varepsilon * u)(x+he_j) - (\rho_\varepsilon * u)(x)\right) = \left\langle u,\, \frac{1}{h}\left(\phi^{x+he_j} - \phi^x\right)\right\rangle.$$

We can easily check that

$$\frac{1}{h}\left(\phi^{x+he_j} - \phi^x\right) \underset{h\to 0}{\longrightarrow} \psi_j^x$$

in $\mathcal{D}(\mathbb{R}^n)$, where $\psi_j^x(y) = (D_j\rho_\varepsilon)(x-y)$, $y \in \mathbb{R}^n$. Hence $\rho_\varepsilon * u$ admits partial derivatives $D_j(\rho_\varepsilon * u)$, and we see that, as above, they are continuous. Thus $\rho_\varepsilon * u$ is C^1, and an induction argument along these lines gives that $\rho_\varepsilon * u$ is C^∞. For the approximation, let $\varphi \in \mathcal{D}(\mathbb{R}^n)$. By definition,

$$\langle \rho_\varepsilon * u,\, \varphi\rangle = \langle u,\, \rho_\varepsilon * \varphi\rangle,$$

and it's easy to check that $\rho_\varepsilon * \varphi \underset{\varepsilon\to 0^+}{\longrightarrow} \varphi$ in $\mathcal{D}(\mathbb{R}^n)$, consequently $\rho_\varepsilon * u \underset{\varepsilon\to 0^+}{\longrightarrow} u$ in $\mathcal{D}'(\mathbb{R}^n)$. $\square$

We can refine the result even further.

Theorem 5.8. *If $u \in \mathcal{D}'(\Omega)$, then there exists a sequence u_j in $\mathcal{D}(\Omega)$ such that $u_j \to u$ in $\mathcal{D}'(\Omega)$.*

Therefore any distribution on Ω can be approximated in the sense of distributions by smooth compactly supported test functions.

Strategy. Prove results for test functions and then try to extend them to distributions by the above approximation result. We shall not prove Theorem 5.8 here, but we shall return to it later.

How should we differentiate $u \in \mathcal{D}'(\Omega)$? Take $u_j \in \mathcal{D}(\Omega)$ such that $u_j \to u$ in $\mathcal{D}'(\Omega)$. Then

$$\begin{aligned}
\langle D_k u_j,\, \varphi\rangle &= \int_\Omega (D_k u_j)\varphi \,\mathrm{d}x \\
&= \int_{\mathbb{R}^n} (D_k u_j)\varphi \,\mathrm{d}x \\
&\overset{\text{Fubini}}{=} \int_{\mathbb{R}^{n-1}}\int_{\mathbb{R}} (D_k u_j)\varphi \,\mathrm{d}x_k\,\mathrm{d}x^1 \\
&\overset{\text{parts}}{=} -\int_{\mathbb{R}^{n-1}}\int_{\mathbb{R}} u_j D_k\varphi \,\mathrm{d}x_k\,\mathrm{d}x^1 \\
&\overset{\text{Fubini}}{=} \langle u_j,\, -D_k\varphi\rangle \underset{j\to\infty}{\longrightarrow} \langle u, -D_k\varphi\rangle,
\end{aligned}$$

$\varphi \in \mathcal{D}(\Omega)$. Hence $\langle D_k u,\, \varphi\rangle := \langle u,\, -D_k\varphi\rangle$ seems reasonable.

6 Lecture 6

We outline a principle that often allows us to extend well-known operations on test functions to corresponding operations on distributions. Let T be an operation on test functions, that is $T : \mathcal{D}(\Omega) \to \mathcal{D}(\Omega)$ is a linear map. Suppose there exists a linear map $S : \mathcal{D}(\Omega) \to \mathcal{D}(\Omega)$ satisfying

$$\int_\Omega T(\varphi)\psi \, \mathrm{d}x = \int_\Omega \varphi S(\psi) \, \mathrm{d}x$$

for all φ, $\psi \in \mathcal{D}(\Omega)$. We call this an *adjoint identity.* If S is continuous in the sense that $S(\psi_j) \to S(\psi)$ in $\mathcal{D}(\Omega)$ whenever $\psi_j \to \psi$ in $\mathcal{D}(\Omega)$, then we can extend T to distributions u by the rule

$$\langle \bar{T}(u), \psi \rangle := \langle u, S(\psi) \rangle,$$

$\psi \in \mathcal{D}(\Omega)$. Because S is linear and continuous, it follows that $\bar{T}(u) \in \mathcal{D}'(\Omega)$, and in fact $\bar{T} : \mathcal{D}'(\Omega) \to \mathcal{D}'(\Omega)$ is linear and continuous,

$$\bar{T}(u + \lambda v) = \bar{T}(u) + \lambda \bar{T}(v)$$

for u, $v \in \mathcal{D}'(\Omega)$, $\lambda \in \mathbb{R}$ (or $\mathbb{C}$), since for $\psi \in \mathcal{D}(\Omega)$

$$\langle \bar{T}(u + \lambda v), \psi \rangle = \langle u + \lambda v, S(\psi) \rangle = \langle u, S(\psi) \rangle + \lambda \langle v, S(\psi) \rangle = \langle \bar{T}(u), \psi \rangle + \lambda \langle \bar{T}(v), \psi \rangle,$$

and if $u_j \to u$ in $\mathcal{D}'(\Omega)$, then for $\psi \in \mathcal{D}(\Omega)$

$$\langle \bar{T}(u_j), \psi \rangle = \langle u_j, S(\psi) \rangle \longrightarrow \langle u, S(\psi) \rangle = \bar{T}(u), \psi \rangle,$$

hence $\bar{T}(u_j) \to \bar{T}(u)$.

Example 6.1. 1. *(Differentiation).* $T = \frac{\mathrm{d}}{\mathrm{d}x} = \mathrm{D}$ on $\mathcal{D}(\mathbb{R})$. For φ, $\psi \in \mathcal{D}(\mathbb{R})$ we have by integration by parts

$$\int_\mathbb{R} \varphi' \psi \, \mathrm{d}x = [\varphi\psi]_{-\infty}^{+\infty} - \int_{-\infty}^{\infty} \varphi\psi' \, \mathrm{d}x = \int_\mathbb{R} \varphi(-\psi') \, \mathrm{d}x,$$

hence we have an adjoint identity with $S = -\mathrm{D}$. Clearly, $S : \mathcal{D}(\mathbb{R}) \to \mathcal{D}(\mathbb{R})$ is linear and continuous, so we may extend to distributions $u \in \mathcal{D}'(\mathbb{R})$ by

$$\langle \bar{\mathrm{D}}u, \psi \rangle = \langle u, -\mathrm{D}\psi \rangle,$$

$\psi \in \mathcal{D}(\mathbb{R})$. To check consistency, suppose $u \in C^1(\mathbb{R})$ and consider also u as an element of $\mathcal{D}'(\mathbb{R})$. We would like to know the relation between the distributional derivative $\bar{\mathrm{D}}u$ defined above and the usual derivative $\mathrm{D}u$. We have

$$\langle \bar{D}u, \psi \rangle = \langle u, -\mathrm{D}\psi \rangle = \int_{-\infty}^{\infty} u(-\mathrm{D}\psi) \, \mathrm{d}x = -[u\psi]_{-\infty}^{+\infty} + \int_{-\infty}^{\infty} \psi \mathrm{D}u, \mathrm{d}x = \langle \mathrm{D}u, \psi \rangle$$

for all $\psi \in \mathcal{D}(\mathbb{R})$, and so $\bar{\mathrm{D}}u = \mathrm{D}u$. In the following we shall therefore not distinguish between the distributional and the classical derivatives and simply denote both by $\mathrm{D}u$ or $\frac{\mathrm{d}u}{\mathrm{d}x}$ when they exist.

2. *(Multiplication by smooth functions).* For $f \in C^\infty(\mathbb{R})$ define $T(\varphi) := f\varphi$ for $\varphi \in \mathcal{D}(\mathbb{R})$. Clearly $T : \mathcal{D}(\mathbb{R}) \to \mathcal{D}(\mathbb{R})$ is linear and $S = T$ yields an adjoint identity:

$$\int_{\mathbb{R}} f\varphi\psi \,\mathrm{d}x = \int_{\mathbb{R}} \varphi f\psi \,\mathrm{d}x$$

for φ, $\psi \in \mathcal{D}(\mathbb{R})$. It is clear that $S : \mathcal{D}(\mathbb{R}) \to \mathcal{D}(\mathbb{R})$ is linear and continuous (checked by Leibniz), so we may extend T to distributions by the rule

$$\langle fu,\, \psi\rangle := \langle u,\, f\psi\rangle$$

for $u \in \mathcal{D}'(\mathbb{R})$, $\psi \in \mathcal{D}(\mathbb{R})$. Clearly we have consistency here: when $u \in L^1_{\mathrm{loc}}(\mathbb{R})$, then $fu \in L^1_{\mathrm{loc}}(\mathbb{R})$ and fu can be identified with the above distribution.

3. Many other useful operations admit extensions to distributions.

   Translation. $T = \tau_h$ defined by $\tau_h\varphi(x) = \varphi(x + h)$ yields adjoint identity with $S = \tau_{-h}$. Thus for $u \in \mathcal{D}'(\mathbb{R})$, $\tau_h u \in \mathcal{D}'(\mathbb{R})$ is defined by the rule

$$\langle \tau_h u,\, \psi\rangle := \langle u,\, \tau_{-h}\psi\rangle$$

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

   Dilation. $T = d_r$ defined by $d_r\varphi(x) = \varphi(rx)$, $r > 0$, yields the adjoint identity with $S = \frac{1}{r}d_{\frac{1}{r}}$. Thus for $u \in \mathcal{D}'(\mathbb{R})$, $d_r u \in \mathcal{D}'(\mathbb{R})$ is defined by the rule

$$\langle d_r u,\, \psi\rangle := \left\langle u,\, \frac{1}{r}d_{\frac{1}{r}}\psi\right\rangle$$

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

   Reflection through the origin. $(T\varphi)(x) = \tilde{\varphi}(x) = \varphi(-x)$ admits the adjoint identity with $S = T$. Thus for $u \in \mathcal{D}'(\mathbb{R})$, $\tilde{u} \in \mathcal{D}'(\mathbb{R})$ is defined by the rule

$$\langle \tilde{u},\, \psi\rangle := \langle u,\, \tilde{\psi}\rangle$$

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

   Convoltion with a test function. For $v \in \mathcal{D}(\mathbb{R})$, $T\varphi = v * \varphi$ admits an adjoint identity with $S\psi = \tilde{v} * \psi$. Indeed, by Fubini,

$$\begin{aligned}
\int_{-\infty}^{\infty} (v * \varphi)(x)\psi(x)\,\mathrm{d}x &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} v(x-y)\varphi(y)\,\mathrm{d}y\psi(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} v(x-y)\psi(x)\,\mathrm{d}x\varphi(y)\,\mathrm{d}y \\
&= \int_{-\infty}^{\infty} (\tilde{v} * \psi)(y)\varphi(y)\,\mathrm{d}y.
\end{aligned}$$

   Thus for $u \in \mathcal{D}'(\mathbb{R})$, $v * u \in \mathcal{D}'(\mathbb{R})$ is defined by the rule

$$\langle v * u,\, \psi\rangle := \langle u,\, \tilde{v} * \psi\rangle$$

   for $\psi \in \mathcal{D}(\mathbb{R})$. On Sheet 2 you will be asked to prove that $v * u \in C^\infty(\mathbb{R})$.

Definition 6.2. Let Ω be a non-empty open subset of $\mathbb{R}^n$. Let $u \in \mathcal{D}'(\Omega)$ and $j \in \{1, \dots, n\}$. The j-th partial derivative of u, $D_j u$ or $\frac{\partial u}{\partial x_j}$, in the sense of distributions is defined by the rule

$$\langle D_j u,\, \varphi\rangle := \langle u,\, -D_j \varphi\rangle$$

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

Note that D_j fits into the adjoint identity scheme with $T = D_j$ and $S = -D_j$, and so is well-defined. Also note that D_j is continuous in the sense that if $u_k \to u$ in $\mathcal{D}'(\Omega)$, then $D_j u_k \to D_j u$ in $\mathcal{D}'(\Omega)$. As in the one-dimensional case, when $u \in C^1(\Omega)$ the distributional and classical partial derivatives $D_1 u, \dots, D_n u$ coincide. Moreover, note that since for $\varphi \in \mathcal{D}(\Omega)$ we have

$$\frac{\partial^2 \varphi}{\partial x_j \partial x_k} = \frac{\partial^2 \varphi}{\partial x_k \partial x_j},$$

we also have $D_j D_k u = D_k D_j u$ for $u \in \mathcal{D}'(\Omega)$. We can therefore use multi-index notation for distributional derivatives. For $u \in \mathcal{D}'(\Omega)$ and $\alpha \in \mathbb{N}^n$ we have

$$\langle D^\alpha u,\, \varphi\rangle = (-1)^{|\alpha|}\langle u,\, D^\alpha \varphi\rangle$$

for $\varphi \in \mathcal{D}(\Omega)$, where we recall that $\alpha = (\alpha_1, \dots, \alpha_n)$, $|\alpha| = \alpha_1 + \dots + \alpha_n$, and

$$D^\alpha \varphi = \frac{\partial^{|\alpha|}\varphi}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}.$$

Definition 6.3. Let u be a distribution and f be a smooth function. Then the product fu in the sense of distributions is defined by the rule

$$\langle fu,\, \varphi\rangle := \langle u,\, f\varphi\rangle$$

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

This definition also fits into the adjoint identity scheme with $T = fx = S$ and so is well-defined. It is clearly consistent, as in the one-dimensional case.

Example 6.4. The *Heaviside function* is the function

$$H(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geqslant 0. \end{cases}$$

Note that the value of $H(x)$ at $x = 0$ is not particularly important and is sometimes taken to be 0 instead (or in some other contexts even $\frac{1}{2}$). Clearly $H \in L^1_{\mathrm{loc}}(\mathbb{R})$, so $H \in \mathcal{D}'(\mathbb{R})$ and we have $H' = \delta_0$. Indeed, since for $\varphi \in \mathcal{D}(\mathbb{R})$

$$\begin{aligned}\langle H', \varphi\rangle &= \langle H, -\varphi'\rangle \\ &= \int_{-\infty}^{\infty} H(x)(-\varphi'(x))\,\mathrm{d}x \\ &= -\int_0^{\infty} \varphi'(x)\,\mathrm{d}x \\ &\overset{\text{FTC}}{=} -[\varphi(x)]_{x=0}^{x=\infty} \\ &= \varphi(0) = \langle \delta_0, \varphi\rangle.\end{aligned}$$

Note also that for $m \in \mathbb{N}$

$$\left\langle \frac{\mathrm{d}^m}{\mathrm{d}x^m}\delta_0, \varphi\right\rangle = \langle \delta_0, (-1)^m \varphi^{(m)}\rangle = (-1)^m \varphi^{(m)}(0).$$

A slight extension of the above formula for H' is obtained by differentiation of a piecewise C^1 function

$$h(x) = \begin{cases} f(x) & x < 0 \\ g(x) & x \geqslant 0, \end{cases}$$

where $f, g \in C^1(\mathbb{R})$. This will be addressed on Sheet 2.

Example 6.5. We can define $\triangle_h = \frac{\tau_h - \mathbf{1}}{h}$ for $h \neq 0$ on distributions $u \in \mathcal{D}'(\mathbb{R})$ by the adjoint identity scheme: for $\varphi \in \mathcal{D}(\mathbb{R})$ put

$$\langle \triangle_h u, \varphi\rangle = \left\langle u, \frac{\tau_{-h} - \mathbf{1}}{h}\varphi\right\rangle,$$

where $\left(\frac{\tau_{-h}-\mathbf{1}}{h}\varphi\right)(x) = \frac{\varphi(x-h)-\varphi(x)}{h}$. If $u \in C^1(\mathbb{R})$, then clearly

$$\triangle_h u(x) = \frac{u(x+h) - u(x)}{h} \underset{h\to 0}{\longrightarrow} u'(x)$$

locally uniformly in x. What happens when $u \in \mathcal{D}'(\mathbb{R})$? One may check that

$$\frac{\tau_{-h} - \mathbf{1}}{h}\varphi \underset{h\to 0}{\longrightarrow} -\varphi'$$

in $\mathcal{D}(\mathbb{R})$ and therefore $\triangle_h u \underset{h\to 0}{\longrightarrow} u'$ in $\mathcal{D}'(\mathbb{R})$.

Theorem 6.6 (Leibniz Rule)**.** *If $u \in \mathcal{D}'(\Omega)$, $f \in C^\infty(\Omega)$, and $j \in \{1, \ldots, n\}$, then*

$$D_j(fu) = (D_j f)u + f D_j u$$

in $\mathcal{D}'(\Omega)$. In fact, the Generalized Leibniz Rule also holds for distributions: for a multi-index $\alpha \in \mathbb{N}^n$,

$$D^\alpha = \sum_{\beta \leqslant \alpha} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} D^\beta f D^{\alpha-\beta} u.$$

Proof. We only prove the basic case, the general case can be proved by induction, or simply by using the formula for test functions. First note that $D_j(fu)$, $(D_jf)u + fD_ju \in \mathcal{D}'(\Omega)$ and that for $\varphi \in \mathcal{D}(\Omega)$:

$$\langle D_j(fu),\, \varphi\rangle = \langle fu,\, -D_j\varphi\rangle = \langle u,\, -fD_j\varphi\rangle,$$

$$\begin{aligned}
\langle (D_jf)u + fD_ju,\, \varphi\rangle &= \langle (D_jf)u,\, \varphi\rangle + \langle fD_ju,\, \varphi\rangle \\
&= \langle u,\, (D_jf)\varphi\rangle + \langle D_ju,\, f\varphi\rangle \\
&= \langle u,\, (D_jf)\varphi\rangle + \langle u,\, -D_j(f\varphi)\rangle \\
&= \langle u,\, (D_jf)\varphi - D_j(f\varphi)\rangle \\
&= \langle u,\, -fD_j\varphi\rangle,
\end{aligned}$$

and we are done. □

7 Lecture 7

Theorem 7.1. *Let Ω be a nonempty connected open subset of $\mathbb{R}^n$. If $u \in \mathcal{D}'(\Omega)$ and*

$$D_ju = 0$$

for $j = 1, 2, \ldots, n$, then u is constant in the sense that there exists $c \in \mathbb{R}$ (or $\mathbb{C}$) such that

$$\langle u,\, \varphi\rangle = c\int_\Omega \varphi(x)\,\mathrm{d}x$$

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

Proof. We only give details for the case $n = 1$ and $\Omega = \mathbb{R}$. The general case can be done along similar lines (not examinable). Suppose $u \in \mathcal{D}'(\mathbb{R})$ satisfies $u' = 0$, that is

$$0 = \langle u',\, \varphi\rangle = -\langle u,\, \varphi'\rangle$$

for all $\varphi \in \mathcal{D}'(\mathbb{R})$. Fix $\rho \in \mathcal{D}(\mathbb{R})$ with $\int_\mathbb{R} \rho\,\mathrm{d}x = 1$ (the standard mollifier kernel on $\mathbb{R}$ will do). For $\varphi \in \mathcal{D}(\mathbb{R})$ we put

$$c_\varphi := \int_\mathbb{R} \varphi(x)\,\mathrm{d}x$$

and

$$\psi(x) := \int_{-\infty}^x (\varphi(t) - c_\varphi\rho(t))\,\mathrm{d}t$$

for $x \in \mathbb{R}$. Take $a, b \in \mathbb{R}$ with $a < b$ and $\varphi(t) = 0 = \rho(t)$ for $t \leqslant a$ and $t \geqslant b$. Then $\psi(x) = 0$ for $x \leqslant a$, and for $x \geqslant b$,

$$\psi(x) = \int_{-\infty}^x (\varphi(t) - c_\varphi\rho(t))\,\mathrm{d}t = \int_{-\infty}^\infty (\varphi(t) - c_\varphi\rho(t))\,\mathrm{d}t = 0.$$

By the FTC, ψ is C^1 with $\psi'(x) = \varphi(x) - c_\varphi \rho(x)$ and hence ψ is C^∞. Since also $\operatorname{supp}(\psi) \subset [a, b]$, $\psi \in \mathcal{D}(\mathbb{R})$. Now

$$\begin{aligned}\langle u, \varphi\rangle &= \langle u, \psi' + c_\varphi \rho\rangle \\ &= \langle u, \psi'\rangle + c_\varphi \langle u, \rho\rangle \\ &= \langle -u', \psi\rangle + c_\varphi \langle u, \rho\rangle \\ &= c_\varphi \langle u, \rho\rangle,\end{aligned}$$

so

$$\langle u, \varphi\rangle = \langle u, \rho\rangle \int_{\mathbb{R}} \varphi(x)\,\mathrm{d}x = c \int_{\mathbb{R}} \varphi(x)\,\mathrm{d}x,$$

where we denoted $c = \langle u, \rho\rangle$. □

Remark 7.2. Recall that any $u \in L^1_{\mathrm{loc}}(\Omega)$ is uniquely determined by the corresponding distribution

$$\langle u, \varphi\rangle = \int_\Omega u\varphi\,\mathrm{d}x,$$

$\varphi \in \mathcal{D}(\Omega)$, and we do not distinguish between u as an L^1_{loc} function and u as a distribution in our notation. In particular note that $C^k(\Omega) \subset L^1_{\mathrm{loc}}(\Omega)$, and that a distribution $u \in \mathcal{D}'(\Omega)$ is a C^k function precisely when

$$\langle u, \varphi\rangle = \int_\Omega f(x)\varphi(x)\,\mathrm{d}x$$

for some $f \in C^k(\Omega)$. Now following the above convention we write $u = f$. However, keep in mind that for $u \in L^1_{\mathrm{loc}}(\Omega)$ and $f \in C^k(\Omega)$ we have from

$$\int_\Omega u(x)\varphi(x)\,\mathrm{d}x = \int_\Omega f(x)\varphi(x)\,\mathrm{d}x$$

only that $u = f$ a.e. in Ω.

Theorem 7.3. *Let Ω be a nonempty open subset of $\mathbb{R}^n$ and $u \in \mathcal{D}'(\Omega)$. If $D_j u \in C(\Omega)$ for $j = 1, 2, \ldots, n$, then $u \in C^1(\Omega)$.*

Proof. We only consider the case $n = 1$ and $\Omega = \mathbb{R}$. Let $(\rho_\varepsilon)_{\varepsilon>0}$ be the standard mollifier and put $u_\varepsilon = \rho_\varepsilon * u$. Then by Theorem 5.7 we have $u_\varepsilon \in C^\infty(\mathbb{R})$ and $u_\varepsilon \longrightarrow u$ in $\mathcal{D}'(\mathbb{R})$. By the FTC,

$$u_\varepsilon(x) - u_\varepsilon(y) = \int_y^x u_\varepsilon'(t)\,\mathrm{d}t$$

for all $x, y \in \mathbb{R}$. By considering difference quotients as in Example 6.5, we find that $u_\varepsilon' = \rho_\varepsilon * u'$. Here the distributional derivative u' is a continuous function, so $u_\varepsilon' \to u'$ locally uniformly on $\mathbb{R}$. Now $u_\varepsilon(x) = u_\varepsilon(y) + \int_y^x u_\varepsilon'(t)\,\mathrm{d}t$, and multiplying by $\rho(y)$ and then integrating over $y \in \mathbb{R}$ gives

$$u_\varepsilon(x) = \int_{-\infty}^{\infty} u_\varepsilon(y)\rho(y)\,\mathrm{d}y + \int_{-\infty}^{\infty} \rho(y) \int_y^x u_\varepsilon'(y)\,\mathrm{d}t\,\mathrm{d}y.$$

Multiply by $\varphi(x) \in \mathcal{D}(\mathbb{R})$ and integrate over $x \in \mathbb{R}$:

$$\begin{aligned}
\langle u_\varepsilon, \varphi \rangle &= \int_{-\infty}^{\infty} u_\varepsilon(x)\varphi(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} u_\varepsilon(y)\rho(y)\,\mathrm{d}y \int_{-\infty}^{\infty} \varphi(x)\,\mathrm{d}x + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \rho(y) \int_y^x u_\varepsilon'(t)\,\mathrm{d}t\,\mathrm{d}y\,\varphi(x)\,\mathrm{d}x.
\end{aligned}$$

Taking $\varepsilon \to 0$ we get

$$\begin{aligned}
\langle u, \varphi \rangle &= \langle u, \varphi \rangle \int_{-\infty}^{\infty} \varphi(x)\,\mathrm{d}x + \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \rho(y) \int_y^x u'(t)\,\mathrm{d}t\,\mathrm{d}y \right) \varphi(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} \left(\langle u, \rho \rangle + \int_{-\infty}^{\infty} \rho(y) \int_y^x u'(t)\,\mathrm{d}t\,\mathrm{d}y \right) \varphi(x)\,\mathrm{d}x.
\end{aligned}$$

Notice that the first term in parentheses is a constant, while the second is a C^1 function in x (by the FTC). Therefore $u \in C^1(\mathbb{R})$. □

What happens if $u \in \mathcal{D}'(\Omega)$ and $D_j u \in L^p(\Omega)$ for $j = 1, 2, \ldots, n$? We shall return to this by the end of the course. Meanwhile we record a definition that is related to this question (and its answer):

Definition 7.4. Let Ω be a nonempty open subset of $\mathbb{R}^n$, $m \in \mathbb{N}$, and $p \in [1, \infty]$. Then any $u \in L^p(\Omega)$ such that $D^\alpha u \in L^p(\Omega)$ for all $|\alpha| \leqslant m$ is called a $W^{m,p}$ *Sobolev function* and the set of all such functions is denoted $W^{m,p}(\Omega)$. Thus

$$W^{m,p}(\Omega) := \{ u \in L^p(\Omega) : D^\alpha u \in L^p(\Omega) \text{ for all } |\alpha| \leqslant m \}.$$

It is not difficult to check that $W^{m,p}(\Omega)$ is a vector space under the usual definitions of addition and scalar multiplication. It is called a *Sobolev space*, and is equipped with the norm

$$\|u\|_{W^{m,p}} := \begin{cases} \left(\displaystyle\sum_{|\alpha| \leqslant m} \|D^\alpha u\|_p^p \right)^{1/p} & \text{if } p \in [1, \infty) \\ \displaystyle\max_{|\alpha| \leqslant m} \|D^\alpha u\|_\infty & \text{if } p = \infty. \end{cases}$$

This is a norm in the same sense that $\|\cdot\|_p$ is a norm on $L^p(\Omega)$: one identifies functions that agree a.e.

8 Lecture 8

Definition 8.1. Let $f \in L^1(\mathbb{R}^n)$. Then the *Fourier transform* of f is

$$\mathcal{F}(f)(\xi) = \hat{f}(\xi) := \int_{\mathbb{R}^n} f(x) \mathrm{e}^{-ix\cdot\xi}\,\mathrm{d}x, \quad \xi \in \mathbb{R}^n.$$

Remark 8.2. Note that this is well-defined for each $\xi \in \mathbb{R}^n$ since

$$|f(x)\mathrm{e}^{-ix\cdot\xi}| = |f(x)|$$

and $f \in L^1(\mathbb{R}^n)$. Here $x \cdot \xi = x_1\xi_1 + \cdots + x_n\xi_n$ is the usual dot product in $\mathbb{R}^n$. Observe that

- $|\hat{f}(\xi)| \leqslant \|f\|_1$ for all $\xi \in \mathbb{R}^n$, and if $f \geqslant 0$, then $\hat{f}(0) = \|f\|_1$,
- $\hat{f} \in C(\mathbb{R}^n)$ (by the DCT),
- $\hat{f}(\xi) \to 0$ as $|\xi| \to \infty$ (by the Riemann–Lebesgue lemma which we shall prove later in the course).

The precise range $\{\hat{f} : f \in L^1(\mathbb{R}^n)\}$ is not so easy to describe in terms not involving the Fourier transform; however, it is strictly smaller than

$$C_0(\mathbb{R}^n) := \left\{ g \in C(\mathbb{R}^n) : \lim_{|x|\to\infty} g(x) = 0 \right\}.$$

One reason that we are interested in the Fourier transform here is its ability to transform partial derivatives to an algebraic operation.

Lemma 8.3 (Differentiation Rule)**.** *Let $f \in L^1(\mathbb{R}^n)$ and assume $D_j f \in L^1(\mathbb{R}^n)$ for some $j \in \{1, \ldots, n\}$. Then*

$$\widehat{D_j f}(\xi) = i\xi_j \hat{f}(\xi).$$

Note that here $D_j f$ is the distributional derivative of f.

Proof. Let $\phi \in \mathcal{D}(\mathbb{R}^n)$ be such that $\phi(x) = 1$ for $|x| \leqslant 1$ (exercise: think about how to construct such a cut-off function). We then calculate

$$\begin{aligned}
\widehat{D_j f}(\xi) &= \int_{\mathbb{R}^n} D_j f(x)\mathrm{e}^{-ix\cdot\xi}\,\mathrm{d}x \\
&\overset{\text{DCT}}{=} \lim_{r\to\infty} \int_{\mathbb{R}^n} D_j f(x)\mathrm{e}^{-ix\cdot\xi}\phi\left(\frac{x}{r}\right)\mathrm{d}x \\
&= \lim_{r\to\infty} \left\langle D_j f,\, \mathrm{e}^{-i(\cdot)\cdot\xi}\phi\left(\frac{\cdot}{r}\right)\right\rangle \\
&= \lim_{r\to\infty} \left\langle f,\, i\xi_j\mathrm{e}^{-i(\cdot)\cdot\xi}\phi\left(\frac{\cdot}{r}\right) - \mathrm{e}^{-i(\cdot)\cdot\xi}(D_j\phi)\left(\frac{\cdot}{r}\right)\frac{1}{r}\right\rangle \\
&= \lim_{r\to\infty} \int_{\mathbb{R}^n} f(x)\left(i\xi_j\mathrm{e}^{-ix\cdot\xi}\phi\left(\frac{x}{r}\right) - \mathrm{e}^{-ix\cdot\xi}(D_j\phi)\left(\frac{x}{r}\right)\frac{1}{r}\right)\mathrm{d}x \\
&\overset{\text{DCT}}{=} \int_{\mathbb{R}^n} f(x) i\xi_j \mathrm{e}^{-ix\cdot\xi}\,\mathrm{d}x = i\xi_j\hat{f}(\xi).
\end{aligned}$$

□

Example 8.4. Let $f = \mathbf{1}_{(-1,1)}$. Clearly $f \in L^1(\mathbb{R})$, and

$$\hat{f}(\xi) = \begin{cases} \dfrac{2\sin\xi}{\xi} & \text{for } \xi \neq 0 \\ 2 & \text{for } \xi = 0. \end{cases}$$

Note that $\hat{f} \notin L^1(\mathbb{R})$.

Example 8.5. Let $\rho \in \mathcal{D}(\mathbb{R})$ the the standard mollifier kernel on $\mathbb{R}$ (that is, ρ is an even function satisfying $0 \leqslant \rho \leqslant 1$, $\operatorname{supp}(\rho) = [-1, 1]$, and $\int \rho = 1$). Then

$$\hat{\rho}(\xi) = 2\int_0^1 \rho(x)\cos(x \cdot \xi)\,\mathrm{d}x.$$

It is not hard to check that $\hat{\rho} \in C^\infty(\mathbb{R})$, but $\operatorname{supp}(\hat{\rho})$ is not compact. Therefore $\hat{\rho} \notin \mathcal{D}(\mathbb{R})$. We shall return to this point when discussing the uncertainty principle later in the course. However,

$$\hat{\rho}(\xi) \to 0$$

as $|\xi| \to \infty$, and in fact, for any $k, m \in \mathbb{N}_0$ we have

$$|\xi|^k D^m \hat{\rho}(\xi) \longrightarrow 0$$

as $|\xi| \to \infty$.

We would like to extend the Fourier transform to distributions, and to that end we seek an adjoint identity. The above example with $\hat{\rho} \notin \mathcal{D}(\mathbb{R})$ shows that we will have to define a new class of test functions. We shall return to that shortly, but first we need a few lemmas.

Lemma 8.6 (The Product Rule)**.** *Let f, $g \in L^1(\mathbb{R}^n)$. Then*

$$\int_{\mathbb{R}^n} f(x)\hat{g}(x)\,\mathrm{d}x = \int_{\mathbb{R}^n} \hat{f}(x)g(x)\,\mathrm{d}x.$$

Note that both sides are well-defined since the Fourier transform of an L^1 function is bounded and continuous. We thus have an adjoint identity with $S = \mathcal{F} = T$, but there is an issue with the domain that we shall address later.

Proof. This is an easy application of Fubini:

$$\begin{aligned} \int_{\mathbb{R}^n} f(x)\hat{g}(x)\,\mathrm{d}x &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x)g(y)\mathrm{e}^{-ix\cdot y}\,\mathrm{d}y\,\mathrm{d}x \\ &\overset{\text{Fubini}}{=} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x)g(y)\mathrm{e}^{-ix\cdot y}\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_{\mathbb{R}^n} \hat{f}(y)g(y)\,\mathrm{d}y \end{aligned}$$

□

Before addressing the issues with the domain and the appropriate class of test functions, let us investigate the properties of the Fourier transform on L^1 functions a little more. It will turn out to be a useful source of insight.

Lemma 8.7 (Translation Rules)**.** *Let $f \in L^1(\mathbb{R}^n)$. Then*

$$\mathcal{F}(\tau_h f)(\xi) = \mathrm{e}^{i\xi\cdot h}\hat{f}(\xi)$$

and

$$\mathcal{F}(\mathrm{e}^{-ix\cdot h}f(x))(\xi) = \tau_h\hat{f}(\xi)$$

for any $h \in \mathbb{R}^n$.

Proof. We simply calculate

$$\mathcal{F}(\tau_h f)(x) = \int_{\mathbb{R}^n} f(x+h)\mathrm{e}^{-ix\cdot\xi}\,\mathrm{d}x \overset{y=x+h}{=} \int_{\mathbb{R}^n} f(y)\mathrm{e}^{-i(y-h)\cdot\xi}\,\mathrm{d}y = \mathrm{e}^{ih\cdot\xi}\hat{f}(\xi),$$

and

$$\int_{\mathbb{R}^n} \mathrm{e}^{-ix\cdot h}f(x)\mathrm{e}^{-ix\cdot\xi}\,\mathrm{d}x = \int_{\mathbb{R}^n} f(x)\mathrm{e}^{-ix\cdot(\xi+h)}\,\mathrm{d}x = \hat{f}(\xi+h) = \tau_h\hat{f}(\xi).$$

□

Lemma 8.8 (Dilation Rules)**.** *Let $f \in L^1(\mathbb{R}^n)$ and denote*

$$(d_r f)(x) = f(rx)$$

for $r > 0$. Then

$$\mathcal{F}(d_r f)(\xi) = r^{-n}\hat{f}(r^{-1}\xi) = r^{-n}(d_{\frac{1}{r}}\hat{f})(\xi)$$

and

$$(d_r\hat{f})(\xi) = \mathcal{F}(r^{-n}d_{\frac{1}{r}}f)(\xi).$$

Proof. The proof is a simple calculation as in the previous lemma,

$$\begin{aligned}
\mathcal{F}(d_r f)(\xi) &= \int_{\mathbb{R}^n} f(rx)\mathrm{e}^{-ix\cdot\xi}\,\mathrm{d}x \\
&\underset{\mathrm{d}y=r^n\mathrm{d}x}{\overset{y=rx}{=}} \int_{\mathbb{R}^n} f(y)\mathrm{e}^{-i\frac{y}{r}\cdot\xi}r^{-n}\,\mathrm{d}y \\
&= r^{-n}\hat{f}\left(\frac{\xi}{r}\right) \\
&= r^{-n}(d_{\frac{1}{r}}\hat{f})(\xi),
\end{aligned}$$

and

$$\begin{aligned}
(d_r \hat{f})(\xi) &= \hat{f}(r\xi) \\
&= \int_{\mathbb{R}^n} f(x) \mathrm{e}^{-ix\cdot r\xi} \,\mathrm{d}x \\
&\overset{y=rx}{\underset{\mathrm{d}y=r^n\mathrm{d}x}{=}} \int_{\mathbb{R}^n} f\left(\frac{y}{r}\right) \mathrm{e}^{-i\frac{y}{r}\cdot r\xi} r^{-n} \,\mathrm{d}y \\
&= r^{-n} \int_{\mathbb{R}^n} (d_{\frac{1}{r}} f)(y) \mathrm{e}^{-iy\cdot\xi} \,\mathrm{d}y \\
&= r^{-n} \widehat{(d_{\frac{1}{r}} f)}(\xi).
\end{aligned}$$

□

Lemma 8.9 (Convolution Rule)**.** *Let $f,\, g \in L^1(\mathbb{R}^n)$. Then $f * g \in L^1(\mathbb{R}^n$ and*

$$\mathcal{F}(f * g)(\xi) = \hat{f}(\xi)\hat{g}(\xi).$$

Proof. By Fubini,

$$\begin{aligned}
\mathcal{F}(f * g)(\xi) &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x-y) g(y) \,\mathrm{d}y \,\mathrm{e}^{-ix\cdot\xi} \,\mathrm{d}x \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x-y) \mathrm{e}^{-i(x-y)\cdot\xi} \,\mathrm{d}x \, g(y) \mathrm{e}^{-iy\cdot\xi} \,\mathrm{d}y \\
&\overset{z=x-y}{\underset{\mathrm{d}z=\mathrm{d}x}{=}} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(z) \mathrm{e}^{-iz\cdot\xi} \,\mathrm{d}z \, g(y) \mathrm{e}^{-iy\cdot\xi} \,\mathrm{d}y \\
&= \hat{f}(\xi)\hat{g}(\xi).
\end{aligned}$$

□

Lemma 8.10 (Reverse Differentiation Rule)**.** *Let $f \in L^1(\mathbb{R}^n)$ and assume $x_j f(x) \in L^1(\mathbb{R}^n)$ for some $j \in \{1, \ldots, n\}$. Then the distributional partial derivative $D_j \hat{f}$ is a continuous function,*

$$(D_j \hat{f})(\xi) = \mathcal{F}(-ix_j f(x))(\xi).$$

In fact, $D_j \hat{f}$ exists classically.

Proof. Let us start with the latter statement. Fix $\xi \in \mathbb{R}^n$ and $h \in \mathbb{R} \setminus \{0\}$ and consider the following difference quotient. We have

$$\begin{aligned}
\triangle_{he_j} \hat{f}(\xi) &:= \frac{\hat{f}(\xi + he_j) - \hat{f}(\xi)}{h} \\
&= \int_{\mathbb{R}^n} f(x) \triangle_{he_j} \mathrm{e}^{-ix\cdot(\cdot)}(\xi) \,\mathrm{d}x \\
&\underset{h\to 0}{\overset{\mathrm{DCT}}{\longrightarrow}} \int_{\mathbb{R}^n} -ix_j f(x) \mathrm{e}^{-ix\cdot\xi} \,\mathrm{d}x \\
&= \mathcal{F}(-ix_j f(x))(\xi),
\end{aligned}$$

so the partial derivative $D_j\hat{f}$ exists classically at ξ. Moreover, since the application $\xi \mapsto \mathcal{F}(-ix_jf(x))(\xi)$ is continuous, so is $D_j\hat{f}$. This is also the distributional derivative since as in Example 6.5, we have $\triangle_{he_j}\hat{f} \to D_j\hat{f}$ in $\mathcal{D}'(\mathbb{R}^n)$ as $h \to 0$. More precisely, one has that

$$\langle \triangle_{he_j}\hat{f}, \varphi\rangle \underset{h\to 0}{\longrightarrow} \langle D_j\hat{f}, \varphi\rangle$$

for every $\varphi \in \mathcal{D}(\mathbb{R}^n)$. But the difference quotient $\triangle_{he_j}\hat{f}(\xi)$ converges locally uniformly in ξ to the classical derivative too, so $D_j\hat{f}$ can be understood in either sense. If unconvinced, the reader is invited to write out the precise details of the last part of the argument as an exercise. $\square$

9 Lecture 9

We start with the observation that the differentiation rules for the Fourier transform on L^1 admit generalizations to higher order derivatives. We formalize this using the following notation. Recall that for a multi-index $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ and $x = (x_1, \dots, x_n \in \mathbb{R}^n$, $D = (D_1, \dots, D_n)$, we denoted by

$$x^\alpha := x_1^{\alpha_1} \dots x_n^{\alpha_n}$$

and

$$D^\alpha := \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}.$$

We can use this notation to write out polynomials in n variables: if $p(x)$ is a polynomial of degree at most k, then

$$p(x) = \sum_{|\alpha|\leqslant k} c_\alpha x^\alpha,$$

where $c_\alpha \in \mathbb{R}$ and we sum over all multi-indices $\alpha \in \mathbb{N}_0^n$ of length $|\alpha| \leqslant k$. Corresponding to the polynomial $p(x)$ is a *linear partial differential operator*

$$p(D) := \sum_{|\alpha|\leqslant k} c_\alpha D^\alpha.$$

If $c_\alpha \neq 0$ for some $\alpha \in \mathbb{N}_0^n$ with $|\alpha| = k$, then we say $p(D)$ has order k. Sometimes we also write $p(iD)$ or $p(-iD)$, the notation being self explanatory:

$$p(iD) = \sum_{|\alpha|\leqslant k} c_\alpha (iD)^\alpha = \sum_{|\alpha|\leqslant k} c_\alpha i^{|\alpha|} D^\alpha,$$

and so on.

Corollary 9.1 (Generalized Differentiation Rules)**.** *Let $p(x) \in \mathbb{C}[x]$ be a polynomial in n variables.*

1. *If $f \in L^1(\mathbb{R}^n)$ and $p(d)f \in L^1(\mathbb{R}^n)$, then*

$$\widehat{p(D)f}(\xi) = p(i\xi)\hat{f}(\xi).$$

2. *If $f \in L^1(\mathbb{R}^n)$ and $pf \in L^1(\mathbb{R}^n)$, then*

$$(p(iD)\hat{f})(\xi) = \mathcal{F}(pf)(\xi).$$

Note that in 1. $p(D)f$ is understood distributionally.

We are now ready to address the domain issues in the adjoint identity for the Fourier transform (Lemma 8.6).

Definition 9.2. A function $f : \mathbb{R}^n \to \mathbb{R}$ (or into $\mathbb{C}$) is said to be *rapidly decreasing* if and only if for every $m \in \mathbb{N}$ there exist $r_m, c_m > 0$ such that

$$|f(x)| \leqslant c_m |x|^{-m}$$

for all $|x| \geqslant r_m$.

Remark 9.3. A continuous function f is rapidly decreasing if and only if for any polynomial $p(x)$ the function $x \mapsto p(x)f(x)$ is bounded on $\mathbb{R}^n$:

$$\sup_{x\in\mathbb{R}^n} |p(x)f(x)| < \infty.$$

Clearly this bound will depend on the polynomial $p(x)$. Exercise: prove this.

Example 9.4.

- $\frac{1}{1+x^{2m}}$ is *not* rapidly decreasing for any $m \in \mathbb{N}$,
- e^{-x^2} is rapidly decreasing,
- $\mathrm{e}^{-|x|}$ is rapidly decreasing.

Definition 9.5 (The Schwartz Space $\mathcal{S}(\mathbb{R}^n)$)**.** We say that φ is a *Schwartz test function on* $\mathbb{R}^n$ and write $\varphi \in \mathcal{S}(\mathbb{R}^n)$ if and only if $\varphi \in C^\infty(\mathbb{R}^n)$ and for all $\alpha \in \mathbb{N}_0^n$, $D^\alpha\varphi$ is rapidly decreasing.

Example 9.6.

- $\mathrm{e}^{-|x|^2} \in \mathcal{S}(\mathbb{R}^n) \setminus \mathcal{D}(\mathbb{R}^n)$,
- $\mathrm{e}^{-|x|} \notin \mathcal{S}(\mathbb{R}^n)$ because it is not differentiable at zero,
- $\hat{\rho}(\xi) \in \mathcal{S}(\mathbb{R}^n) \setminus \mathcal{D}(\mathbb{R}^n)$, where ρ is the standard mollifier kernel on $\mathbb{R}^n$.

The following lemma collects a few of the properties of the class of Schwartz test functions.

Lemma 9.7.

(i) $\mathcal{S}(\mathbb{R}^n)$ *is a vector space (over* $\mathbb{R}$ *or* $\mathbb{C}$*),*

(ii) If $p(x)$ *is a polynomial and* $\varphi \in \mathcal{S}(\mathbb{R}^n)$*, then* $p\varphi \in \mathcal{S}(\mathbb{R}^n)$*,*

(iii) If $p(x)$ *is a polynomial and* $\varphi \in \mathcal{S}(\mathbb{R}^n)$*, then* $p(D)\varphi \in \mathcal{S}(\mathbb{R}^n)$*,*

(iv) $\mathcal{D}(\mathbb{R}^n) \subsetneq \mathcal{S}(\mathbb{R}^n) \subsetneq L^1(\mathbb{R}^n)$*.*

Proof.

(i) If φ, $\psi \in \mathcal{S}(\mathbb{R}^n)$ and $\lambda \in \mathbb{R}$ (or $\mathbb{C}$), then certainly $\varphi + \lambda\psi \in C^\infty(\mathbb{R}^n)$ because $C^\infty(\mathbb{R}^n)$ is a vector space. Now for a polynomial p on $\mathbb{R}^n$ we have

$$\sup_x |p(x)(\varphi + \lambda\psi)(x)| \leqslant \sup_x |p(x)\varphi(x)| + |\lambda| \sup_x |p(x)\psi(x)| < \infty.$$

Thus $\varphi + \lambda\psi$ is rapidly decreasing. All derivatives $D^\alpha(\varphi + \lambda\psi)$ are similarly seen to be rapidly decreasing for all $\alpha \in \mathbb{N}_0^n$. Consequently $\varphi + \lambda\psi \in \mathcal{S}(\mathbb{R}^n)$.

(ii) Clearly $p\varphi \in C^\infty(\mathbb{R}^n)$. Fix $\alpha \in \mathbb{N}_0^n$ and a polynomial q on $\mathbb{R}^n$. By the Leibniz Rule

$$D^\alpha(p\varphi) = \sum_{\beta \leqslant \alpha} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} D^\beta p D^{\alpha-\beta}\varphi,$$

and so

$$\begin{aligned} \sup_x |qD^\alpha(p\varphi)| &\leqslant \sup_x \sum_{\beta \leqslant \alpha} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} |qD^\beta p D^{\alpha-\beta}\varphi| \\ &\leqslant \sum_{\beta \leqslant \alpha} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \sup_x |qD^\beta p D^{\alpha-\beta}\varphi| \\ &< \infty \end{aligned}$$

since $qD^\beta p$ is a polynomial on $\mathbb{R}^n$.

(iii) For each $\alpha \in \mathbb{N}_0^n$ we have $D^\alpha\varphi \in \mathcal{S}(\mathbb{R}^n)$, and so $p(D)\varphi \in \mathcal{S}(\mathbb{R}^n)$ follows from (i).

(iv) We have already seen that $\mathcal{D}(\mathbb{R}^n) \subsetneq \mathcal{S}(\mathbb{R}^n)$. Let $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Then in particular φ is continuous and rapidly decreasing, so for $m = n + 2$ we may find $r = r_m$ and $c = c_m > 0$ such that

$$|\varphi(x)| \leqslant c|x|^{-n-2}$$

for $|x| \geqslant r$. By continuity, $M \coloneqq \sup_{|x| \leqslant r} |\varphi(x)| < \infty$, and so

$$\int_{\mathbb{R}^n} |\varphi(x)| \,\mathrm{d}x \leqslant M\mathcal{L}^n(B_r(0)) + \int_r^\infty c\omega_{n-1} \frac{\mathrm{d}t}{t^2} < \infty.$$

$\square$

10 Lecture 10

Theorem 10.1. *$\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is a linear map.*

Proof. By part (iv) of Lemma 9.7, $\mathcal{S}(\mathbb{R}^n) \subset L^1(\mathbb{R}^n)$ and so $\hat{\varphi}$ is well-defined for $\varphi \in \mathcal{S}(\mathbb{R}^n)$. We thus only need to check that $\hat{\varphi} \in \mathcal{S}(\mathbb{R}^n)$ whenever $\varphi \in \mathcal{S}(\mathbb{R}^n)$.

Recall that if $\varphi \in \mathcal{S}(\mathbb{R}^n) \subset L^1(\mathbb{R}^n)$, then $\hat{\varphi}$ is continuous and

$$|\hat{\varphi}(\xi)| = \left| \int_{\mathbb{R}^n} \varphi(x) \mathrm{e}^{-ix\cdot\xi} \, \mathrm{d}x \right| \leqslant \|\varphi\|_1 \tag{5}$$

for all ξ. Fix $\varphi \in \mathcal{S}(\mathbb{R}^n)$.

Step 1: $\hat{\varphi}$ is rapidly decreasing.

Since $\hat{\varphi}$ is continuous it suffices to show that $p\hat{\varphi}$ is bounded whenever p is a polynomial in n variables. So let p be such a polynomial, and recall from parts (iii) and (iv) of Lemma 9.7 that

$$p(-iD) \in \mathcal{S}(\mathbb{R}^n) \subset L^1(\mathbb{R}^n).$$

Hence by Corollary 9.1

$$p(\xi)\hat{\varphi}(\xi) = \widehat{p(-iD)\varphi}(\xi),$$

and so $\sup_\xi |p(\xi)\hat{\varphi}(\xi)| < \infty$ by (5).

Step 2: Let $\alpha \in \mathbb{N}_0^n$. Then $D^\alpha\hat{\varphi}$ is rapidly decreasing.

Again, $D^\alpha\hat{\varphi}$ is continuous so we only need to show that $\sup_\xi |p(\xi)D^\alpha\hat{\varphi}(\xi)| < \infty$ whenever p is a polynomial in n variables. By part (ii) of Lemma 9.7, $\psi(x) := (-ix)^\alpha\varphi(x) \in \mathcal{S}(\mathbb{R}^n)$, ad so by parts (iii) and (iv)

$$p(-iD)\psi \in \mathcal{S}(\mathbb{R}^n) \subset L^1(\mathbb{R}^n).$$

Then Corollary 9.1 yields

$$p(\xi)D^\alpha\hat{\varphi}(\xi) = p(\xi)\hat{\psi}(\xi) = \widehat{p(-iD)\psi}(\xi),$$

which is bounded by (5). $\square$

Remark 10.2. We record the following principle that is implicit in the above proof.

(a) Let $m \in \mathbb{N}_0$. If $f \in W^{m,1}(\mathbb{R}^n)$, then $\sup_\xi (1 + |\xi|)^m |\hat{f}(\xi)| < \infty$,

(b) Let $m \in \mathbb{N}$, $m \geqslant n+1$. If $f \in L^1(\mathbb{R}^n)$ and $\sup_x (1+|x|)^m |f(x)| < \infty$, then $\hat{f} \in C^{m-n-1}(\mathbb{R}^n)$ and $\sup_\xi |D^\alpha \hat{f}(\xi)| < \infty$ for $|\alpha| \leqslant m - n - 1$.

There is clearly a gap of $n+1$ derivatives between (a) and (b), but in the proof of Theorem 10.1 this did not matter because the definition of a Schwartz function involves C^∞ smoothness and rapid decrease.

We could now proceed to define the Fourier transform for *certain* distributions using the adjoint identity scheme: the product rule holds in particular for φ, $\psi \in \mathcal{S}(\mathbb{R}^n)$, namely

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

However, we shall first show that the Fourier transform is bijective on the Schwartz space.

Theorem 10.3 (Fourier Inversion Theorem on $\mathcal{S}(\mathbb{R}^n)$)**.** *The Fourier transform $\mathcal{F}\colon \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is bijective with inverse given by*

$$(\mathcal{F}^{-1}\psi)(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} \psi(\xi)\mathrm{e}^{ix\cdot\xi}\,\mathrm{d}\xi.$$

Proof. Let $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Can we recover φ from $\hat{\varphi}$? Now $\hat{\varphi} \in \mathcal{S}(\mathbb{R}^n) \subset L^1(\mathbb{R}^n)$ so we may consider

$$\begin{aligned}\mathcal{F}^2(\varphi)(x) &= \int_{\mathbb{R}^n} \hat{\varphi}(\xi)\mathrm{e}^{-ix\cdot\xi}\,\mathrm{d}\xi \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \varphi(y)\mathrm{e}^{-i(x+y)\cdot\xi}\,\mathrm{d}y\,\mathrm{d}\xi.\end{aligned}$$

Observe that $|\varphi(y)\mathrm{e}^{-i(x+y)\cdot\xi}| = |\varphi(y)|$ is *not* in general integrable over $(y,\xi) \in \mathbb{R}^n \times \mathbb{R}^n$, so we cannot use Fubini to swap the order of integration. Instead we use a trick (in reality it's a tool from the theory of 'summability methods'). This is based on the following.

Lemma 10.4 (Auxiliary Lemma)**.** *For $t > 0$ let $G_t(x) := \mathrm{e}^{-t|x|^2}$, $x \in \mathbb{R}^n$. Then $\hat{G}_t(\xi) = \left(\frac{\pi}{t}\right)^{n/2}\mathrm{e}^{-\frac{|\xi|^2}{4t}}$, $\xi \in \mathbb{R}^n$, and the family $(\hat{G}_t)_{t>0}$ satisfies*

(1) $\int_{\mathbb{R}^n} \hat{G}_t(\xi)\,\mathrm{d}\xi = (2\pi)^n$ *for* $t > 0$,

(2) $\hat{G}_t(\xi) \geqslant 0$ *for* $\xi \in \mathbb{R}^n$, $t > 0$, *and*

(3) $\lim_{t\to 0^+} \hat{G}_t(\xi) = 0$ *uniformly in* $|\xi| > \varepsilon$ *for each fixed* $\varepsilon > 0$.

As a consequence of points (1)–(3), the family

$$\left((2\pi)^{-n}\hat{G}_t\right)_{t>0} = \left(\frac{1}{(4\pi t)^{n/2}}\mathrm{e}^{-\frac{|\xi|^2}{4t}}\right)_{t>0}$$

is an *approximate identity*, and for $f \in \mathcal{S}(\mathbb{R}^n)$ we have that

$$(\hat{G}_t * f)(x) \underset{t\to 0^+}{\longrightarrow} (2\pi)^n f(x)$$

uniformly in $x \in \mathbb{R}^n$. You might recognize the function $(2\pi)^{-n}\hat{G}_t$ as the heat kernel from the Part A Differential Equations course. We postpone the proof of the Auxiliary Lemma for now. Note that

$$\mathcal{F}^2(\varphi)(x) = \lim_{t\to 0^+} \int_{\mathbb{R}^n} \hat{\varphi}(\xi)\mathrm{e}^{-ix\cdot\xi}G_t(\xi)\,\mathrm{d}\xi.$$

Now for $t > 0$ we have

$$\begin{aligned}
\int_{\mathbb{R}^n} \hat{\varphi}(\xi)\mathrm{e}^{-ix\cdot\xi} G_t(\xi)\,\mathrm{d}\xi &= \int_{\mathbb{R}^n} \hat{\varphi}(\xi)\underbrace{\mathrm{e}^{-ix\cdot\xi - t|\xi|^2}}_{\in L^1}\mathrm{d}\xi \\
\text{product rule} \to &= \int_{\mathbb{R}^n} \varphi(\xi)\mathcal{F}_{y\to\xi}\left(\mathrm{e}^{-ix\cdot y - t|y|^2}\right)\mathrm{d}\xi \\
\text{translation rule} \to &= \int_{\mathbb{R}^n} \varphi(\xi)\hat{G}_t(\xi + x)\,\mathrm{d}\xi \\
\begin{cases} \eta = -\xi \\ \mathrm{d}\eta = \mathrm{d}\xi \end{cases} \to &= \int_{\mathbb{R}^n} \varphi(-\eta)\hat{G}_t(x - \eta)\,\mathrm{d}\eta \\
&= (\hat{G}_t * \tilde{\varphi})(x).
\end{aligned}$$

From the Auxiliary Lemma it follows that

$$(\hat{G}_t * \tilde{\varphi})(x) \underset{t\to 0^+}{\longrightarrow} (2\pi)^n \tilde{\varphi}(x) = (2\pi)^n \varphi(-x)$$

uniformly in $x \in \mathbb{R}^n$. Collecting the pieces we conclude that

$$\mathcal{F}^2(\varphi)(x) = (2\pi)^n \varphi(-x),$$

or equivalently that

$$\varphi(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} \hat{\varphi}(\xi)\mathrm{e}^{ix\cdot\xi}\,\mathrm{d}\xi,$$

$x \in \mathbb{R}^n$. It follows easily that $\mathcal{F}$ is bijective. It thus remains to prove the Auxiliary Lemma. □

11 Lecture 11

Proof of the Auxiliary Lemma. We express (1)-(3) by saying that $((2\pi)^{-n}\hat{G}_t)_{t>0}$ is an approximate identity and for $f \in \mathcal{S}(\mathbb{R}^n)$ we have

$$(\hat{G}_t * f)(x) \underset{t\to 0^+}{\longrightarrow} (2\pi)^n f(x)$$

uniformly in x.

Remark 11.1. In addition to (3) we record that also

(3') $\displaystyle\int_{|\xi|>\varepsilon} \hat{G}_t(\xi)\,\mathrm{d}\xi \underset{t\to 0^+}{\longrightarrow} 0$ for each fixed $\varepsilon > 0$.

Note that $G_t = d_{\sqrt{t}} G_1$, so from the dilation rule $\hat{G}_t = t^{-n/2} d_{t^{-1/2}} \hat{G}_1$. We can therefore focus on the computation for $t = 1$. Put $G := G_1$, and note further that

$$G(x) = \mathrm{e}^{-|x|^2} = \prod_{j=1}^n \mathrm{e}^{-x_j^2},$$

hence by Fubini

$$\hat{G}(\xi) = \prod_{j=1}^{n} \mathcal{F}_{x_j \to \xi_j} \left(\mathrm{e}^{-x_j^2} \right) (\xi_j).$$

We may therefore further assume that $n = 1$ and $G(x) = \mathrm{e}^{-|x|^2}$, $x \in \mathbb{R}$. In this case we compute

$$\hat{G}(\xi) = \int_{-\infty}^{\infty} \mathrm{e}^{-x^2 - ix\xi} \, \mathrm{d}x = \mathrm{e}^{-\frac{1}{4}\xi^2} \int_{-\infty}^{\infty} \mathrm{e}^{-(x+\frac{i}{2}\xi)^2} \, \mathrm{d}x.$$

Put

$$F(\xi) = \int_{-\infty}^{\infty} \mathrm{e}^{-(x+\frac{i}{2}\xi)^2} \, \mathrm{d}x$$

for $\xi \in \mathbb{R}$, and observe that by the DCT that F is C^1 with

$$\begin{aligned} F'(\xi) &= \int_{-\infty}^{\infty} \mathrm{e}^{-(x+\frac{i}{2}\xi)^2} \left(-2 \left(x + \frac{i}{2}\xi \right) \frac{i}{2} \right) \mathrm{d}x \\ &= \frac{i}{2} \int_{-\infty}^{\infty} \frac{\mathrm{d}}{\mathrm{d}x} \left(\mathrm{e}^{-(x+\frac{i}{2}\xi)^2} \right) \mathrm{d}x \\ &\overset{\text{FTC}}{=} \frac{i}{2} \left[\mathrm{e}^{-(x+\frac{i}{2}\xi)^2} \right]_{-\infty}^{\infty} = 0. \end{aligned}$$

F is therefore constant: $F(\xi) = F(0) = \int_{-\infty}^{\infty} \mathrm{e}^{-x^2} \, \mathrm{d}x = \sqrt{\pi}$. Thus

$$\hat{G}(\xi) = \sqrt{\pi} \mathrm{e}^{-\frac{\xi^2}{4}},$$

as required.

We return to the general case and the proof of (1)-(3). Point (1) follows from the above calculation upon performing a substitution, and (2) is clear. For (3) we note that when $|\xi| > \delta$,

$$|\hat{G}_t(\xi)| = \hat{G}_t(\xi) \leqslant \left(\frac{\pi}{t} \right)^{n/2} \mathrm{e}^{-\frac{\delta^2}{4t}}.$$

For $s > 0$ one has $\mathrm{e}^s > \frac{s^n}{n!}$, so $\mathrm{e}^{-s} < \frac{n!}{s^n}$, and so

$$\left(\frac{\pi}{t} \right)^{n/2} \mathrm{e}^{-\frac{\delta^2}{4t}} < \frac{\pi^{n/2} n! 4^n}{\delta^{2n}} t^{n/2} \underset{t \to 0^+}{\longrightarrow} 0,$$

from which (3) follows. Point (3') is a minor variation:

$$\int_{|\xi| > \delta} \hat{G}_t(\xi) \, \mathrm{d}\xi \underset{\mathrm{d}\eta = (2\sqrt{t})^{-n} \mathrm{d}\xi}{\overset{\eta = \frac{\xi}{2\sqrt{t}}}{=}} 2^n \pi^{n/2} \int_{|\eta| > (2\sqrt{t})^{-1}\delta} \mathrm{e}^{-|\eta|^2} \, \mathrm{d}\eta \underset{t \to 0^+}{\longrightarrow} 0.$$

Finally, for $f \in \mathcal{S}(\mathbb{R}^n)$ we have that f is in particular uniformly continuous, so given $\varepsilon > 0$ we can find $\delta > 0$ such that $|f(y) - f(x)| < \varepsilon$ whenever $|x - y| \leqslant \delta$. As f is also bounded we get

$$\begin{aligned}|(\hat{G}_t * f)(x) - f(x)| &\overset{(1)}{\leqslant} \int_{\mathbb{R}^n} \hat{G}_t(x-y)|f(y) - f(x)| \, \mathrm{d}y \\ &\leqslant \int_{|x-y|>\delta} \hat{G}_t(x-y) \, \mathrm{d}y \cdot 2\|f\|_\infty + \int_{\overline{B_\delta(x)}} \hat{G}_t(x-y) \, \mathrm{d}y \cdot \varepsilon,\end{aligned}$$

and hence from (1) and (3')

$$\limsup_{t \to 0^+} |(\hat{G}_t * f)(x) - f(x)| \leqslant \varepsilon.$$

The choices of δ and t above only depend on ε, so convergence is uniform in $x \in \mathbb{R}^n$. One can show that the conclusion can also be obtained only from (1)-(3); the use of (3') is not necessary. $\square$

Remark 11.2. When $(K_t)_{t>0}$ is an approximate identity, it is not difficult to show that

$$\|K_t * f - f\|_1 \underset{t \to 0^+}{\longrightarrow} 0$$

whenever $f \in L^1(\mathbb{R}^n)$.

11.1 Recap on the Fourier Transform

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

$$\mathcal{F}(f)(\xi) = \hat{f}(\xi) := \int_{\mathbb{R}^n} f(x) \mathrm{e}^{-ix \cdot \xi} \, \mathrm{d}x,$$

for $x \in \mathbb{R}^n$. Here $\mathcal{F}$ maps $L^1(\mathbb{R}^n)$ into $C_0(\mathbb{R}^n)$, but one can show that it is not onto. The situation is better when we consider the Fourier transform on $\mathcal{S}(\mathbb{R}^n)$; then

$$\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \longrightarrow \mathcal{S}(\mathbb{R}^n)$$

is bijective with

$$\mathcal{F}^{-1}(g)(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} g(\xi) \mathrm{e}^{ix \cdot \xi} \, \mathrm{d}\xi,$$

$x \in \mathbb{R}^n$. Note that we may write this as $\mathcal{F}^{-1} = (2\pi)^{-n} \tilde{\mathcal{F}}$, where the order in which we perform the operations $g \mapsto \mathcal{F}(g)$ and $g \mapsto \tilde{g}$ is unimportant.

We return to the task of defining the Fourier transform on distributions using the adjoint identity:

$$\int_{\mathbb{R}^n} \varphi(x) \hat{\psi}(x) \, \mathrm{d}x = \int_{\mathbb{R}^n} \hat{\varphi}(x) \psi(x) \, \mathrm{d}x$$

for $\varphi, \psi \in \mathcal{S}(\mathbb{R}^n)$. Observe that the distribution should be defined on $\mathcal{S}(\mathbb{R}^n)$ rather than on $\mathcal{D}(\mathbb{R}^n)$, and since $\mathcal{D}(\mathbb{R}^n) \subsetneq \mathcal{S}(\mathbb{R}^n)$, it is likely that we will have to exclude some distributions. As with $\mathcal{D}'(\mathbb{R}^n)$ we start with a notion of convergence on $\mathcal{S}(\mathbb{R}^n)$. In connection with this we recall the following characterization of $\mathcal{S}(\mathbb{R}^n)$:

$$\varphi \in \mathcal{S}(\mathbb{R}^n) \iff \left(\varphi \in C^\infty(\mathbb{R}^n) \text{ and } \forall \alpha, \beta \in \mathbb{N}_0^n \;\; S_{\alpha,\beta}(\varphi) := \sup_{x \in \mathbb{R}^n} |x^\alpha D^\beta \varphi(x)| < \infty\right).$$

We may also replace the condition for $S_{\alpha,\beta}(\varphi)$ by the following apparently stronger, but in fact equivalent, condition

$$\mathsf{S}_{p,q}(\varphi) := \sup_{x \in \mathbb{R}^n} |p(x)(q(D)\varphi)(x)| < \infty$$

for all polynomials p and q on $\mathbb{R}^n$. For $k, l \in \mathbb{N}_0$ put

$$\bar{S}_{k,l}(\varphi) := \max_{|\alpha| \leqslant k,\, |\beta| \leqslant l} S_{\alpha,\beta}(\varphi).$$

Note that all three of $S_{\alpha,\beta}$, $\mathsf{S}_{p,q}$, and $\bar{S}_{k,l}$ are semi-norms. Observe that

$$\mathsf{S}_{p,q}(\varphi) \leqslant c\bar{S}_{\deg p,\, \deg q}(\varphi)$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^n)$, where c is a constant that only depends on the polynomials p and q and whose precise value is not important here.

Lemma 11.3. *Let p be a polynomial on $\mathbb{R}^n$. Then there exists a constant $c = c(p)$ such that for all $\alpha, \beta \in \mathbb{N}_0^n$ and $\varphi \in \mathcal{S}(\mathbb{R}^n)$,*

$$\begin{aligned} S_{\alpha,\beta}(p\varphi) &\leqslant c\bar{S}_{|\alpha|+\deg p, |\beta|}(\varphi), \\ S_{\alpha,\beta}(p(D)\varphi) &\leqslant c\bar{S}_{|\alpha|, |\beta|+\deg p}(\varphi). \end{aligned}$$

The proof is a simple but somewhat tedious application of the Leibniz rule and we omit the details here.

Definition 11.4. Let $\varphi_j, \varphi \in \mathcal{S}(\mathbb{R}^n)$. Then we say φ_j converges to φ in the sense of the Schwartz test functions, and write $\varphi_j \to \varphi$ in $\mathcal{S}(\mathbb{R}^n)$, if and only if

$$S_{\alpha,\beta}(\varphi - \varphi_j) \longrightarrow 0$$

as $j \to \infty$ for all $\alpha, \beta \in \mathbb{N}_0^n$. This can also be stated in terms of $\mathsf{S}_{p,q}$ or $\bar{S}_{k,l}$.

Remark 11.5 (A metric on $\mathcal{S}(\mathbb{R}^n)$). Define for $\varphi, \psi \in \mathcal{S}(\mathbb{R}^n)$

$$d(\varphi, \psi) := \sum_{k,l \in \mathbb{N}_0} 2^{-k-l} \frac{\bar{S}_{k,l}(\varphi - \psi)}{1 + \bar{S}_{k,l}(\varphi - \psi)}.$$

Then d is a metric on $\mathcal{S}(\mathbb{R}^n)$, and we have $\varphi_j \to \varphi$ in $\mathcal{S}(\mathbb{R}^n)$ if and only if $d(\varphi_j, \varphi) \to 0$. Note that d is translation invariant,

$$d(\varphi + \eta, \psi + \eta) = d(\varphi, \psi),$$

and it can be shown that $(\mathcal{S}(\mathbb{R}^n), d)$ is complete (such a space is called a Fréchet space).

We may compare this to the notion of convergence in $\mathcal{D}(\mathbb{R}^n)$. If $\varphi_j, \varphi \in \mathcal{D}(\mathbb{R}^n)$ and $\varphi_j \to \varphi$ in $\mathcal{D}(\mathbb{R}^n)$, then $\varphi_j, \varphi \in \mathcal{S}(\mathbb{R}^n)$ and $\varphi_j \to \varphi$ in $\varphi_j \to \varphi$ in $\mathcal{S}(\mathbb{R}^n)$. The converse, however, is clearly false.

Lemma 11.6. *Let p be a polynomial on $\mathbb{R}^n$. Then the maps $\varphi \mapsto p\varphi$, $\varphi \mapsto p(D)\varphi$, and $\varphi \mapsto \hat{\varphi}$ are continuous with respect to the convergence on $\mathcal{S}(\mathbb{R}^n)$.*

Proof. The first two follow immediately from Lemma 11.3. For the Fourier transform, note that for all $\alpha, \beta \in \mathbb{N}_0^n$ $\xi^\alpha D^\beta \hat{\varphi}(\xi) = (-1)^{|\alpha|+|\beta|} \mathcal{F}_{x\to\xi}\left(D^\alpha(x^\beta \varphi(x))\right)$, so

$$\begin{aligned} S_{\alpha,\beta}(\hat{\varphi}) &= \sup_\xi \left| \mathcal{F}_{x\to\xi}\left(D^\alpha(x^\beta\varphi(x))\right)\right| \\ &\leqslant \int_{\mathbb{R}^n} |D^\alpha(x^\beta\varphi(x))| \,\mathrm{d}x \\ &= \int_{\mathbb{R}^n} \frac{1}{1+|x|^{2n}}(1+|x|^{2n})|D^\alpha(x^\beta\varphi(x))| \,\mathrm{d}x \\ &\leqslant \underbrace{\int_{\mathbb{R}^n} \frac{\mathrm{d}x}{1+|x|^{2n}}}_{=:c} \sup_x \left|(1+|x|^{2n})D^\alpha(x^\beta\varphi(x))\right|. \end{aligned}$$

By convexity of the map $t \mapsto t^n$, we have

$$|x|^{2n} = (x_1^2 + \cdots + x_n^2)^n \leqslant n^{n-1}(x_1^{2n} + \cdots + x_n^{2n})$$

and thus

$$\begin{aligned} \sup_x \left|(1+|x|^{2n})\mathrm{D}^\alpha(x^\beta\varphi(x))\right| &\leqslant \sup_x \left| \left(1 + n^{n-1}\sum_{j=1}^n x_j^{2n}\right) \mathrm{D}^\alpha(x^\beta\varphi)\right| \\ &\leqslant \sup_x \left|\mathrm{D}^\alpha(x^\beta\varphi)\right| + n^{n-1}\sum_{j=1}^n \sup_x \left|x_j^{2n}\mathrm{D}^\alpha(x^\beta\varphi)\right| \\ &\overset{\text{Lemma 11.3}}{\leqslant} c_\beta \bar{S}_{|\beta|,|\alpha|}(\varphi) + n^n c_\beta \bar{S}_{2n+|\beta|,|\alpha|}(\varphi) \\ &\leqslant C\bar{S}_{2n+|\beta|,|\alpha|}(\varphi), \end{aligned}$$

where $C = (1+n^n)c_\beta$. We have thus shown that

$$S_{\alpha,\beta}(\hat{\varphi}) \leqslant cC\bar{S}_{2n+|\beta|,|\alpha|}(\varphi),$$

and hence $\mathcal{F}$ is continuous. □

Definition 11.7 (Tempered Distributions)**.** A functional $u : \mathcal{S}(\mathbb{R}^n) \to \mathbb{R}$ (or into $\mathbb{C}$) is a tempered distribution if and only if u is linear and u is continuous on $\mathcal{S}(\mathbb{R}^n)$ in the sense that

$$\langle u,\, \varphi_j\rangle \longrightarrow \langle u,\, \varphi\rangle$$

whenever $\varphi_j \to \varphi$ in $\mathcal{S}(\mathbb{R}^n)$. The set of all tempered distributions is denoted by $\mathcal{S}'(\mathbb{R}^n)$. It is clearly a vector space with the usual definitions of vector space operations.

Remark 11.8. Since $\mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$ and $\mathcal{D}$-convergence implies $\mathcal{S}$-convergence, it follows that $\mathcal{S}'(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n)$. The space $\mathcal{D}'(\mathbb{R}^n)$ is genuinely larger than $\mathcal{S}'(\mathbb{R}^n)$, since for example $u = \mathrm{e}^{x^2} \in L^1_{\mathrm{loc}}(\mathbb{R}) \subset \mathcal{D}'(\mathbb{R})$ by the rule

$$\langle u, \varphi \rangle = \int_{-\infty}^{\infty} \mathrm{e}^{x^2} \varphi(x) \, \mathrm{d}x.$$

However, $u \notin \mathcal{S}'(\mathbb{R})$ as for instance $\varphi = \mathrm{e}^{-x^2} \in \mathcal{S}(\mathbb{R})$ but $u\varphi = \mathbf{1}_{\mathbb{R}} \notin L^1(\mathbb{R})$.

Definition 11.9 (Convergence of Tempered Distributions)**.** For a sequence (u_j) in $\mathcal{S}'(\mathbb{R}^n)$ and $u \in \mathcal{S}'(\mathbb{R}^n)$ we write

$$u_j \longrightarrow u \;\text{ in } \mathcal{S}'(\mathbb{R}^n)$$

if and only if $\langle u_j, \varphi \rangle \to \langle u, \varphi \rangle$ for each fixed $\varphi \in \mathcal{S}(\mathbb{R}^n)$.

Remark 11.10. This is stronger than convergence in $\mathcal{D}'(\mathbb{R}^n)$.

Using adjoint identities and lemma 11.6 we get the following.

Definition 11.11. For $\alpha \in \mathbb{N}_0^n$, a polynomial $in \mathbb{C}[x]$ and $u \in \mathcal{S}'(\mathbb{R}^n)$ we define the tempered distributions $D^\alpha u$, pu, and $\hat{u}$ by the rules

$$\langle D^\alpha u, \varphi \rangle := (-1)^{|\alpha|} \langle u, D^\alpha \varphi \rangle,$$

$$\langle pu, \varphi \rangle := \langle u, p\varphi \rangle,$$
$$\langle \hat{u}, \varphi \rangle := \langle u, \hat{\varphi} \rangle.$$

We also define the translation $\tau_h u$, dilation $d_r u$, and $\tilde{u}$ as on $\mathcal{D}'(\mathbb{R}^n)$. All of the above operations are linear and continuous in the sense of $\mathcal{S}'(\mathbb{R}^n)$, and the rules for the Fourier Transform on $\mathcal{S}(\mathbb{R}^n)$ also hold on $\mathcal{S}'(\mathbb{R}^n)$.

Theorem 11.12 (Fourier Inversion Formula on $\mathcal{S}'(\mathbb{R}^n)$)**.** *The map $\mathcal{F} \colon \mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$ is a linear bijection with inverse $\mathcal{F}^{-1} = (2\pi)^{-n} \tilde{\mathcal{F}}$.*

Proof. By chasing definitions and using the Fourier Inversion Formula on $\mathcal{S}(\mathbb{R}^n)$, for $u \in \mathcal{S}'(\mathbb{R}^n)$ and $\varphi \in \mathcal{S}(\mathbb{R}^n)$ we have

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

□

Example 11.13. Let $u \in L^p(\mathbb{R}^n)$, $p \in [1, \infty]$. Then $u \in \mathcal{S}'(\mathbb{R}^n)$ by the rule

$$\langle u, \varphi \rangle = \int_{\mathbb{R}^n} u\varphi \, \mathrm{d}x$$

for $\varphi \in \mathcal{S}(\mathbb{R}^n)$. We show that it is well-defined and continuous by noting that

$$\mathcal{S}(\mathbb{R}^n) \subset (L^1 \cap L^\infty)(\mathbb{R}^n) \subset L^q(\mathbb{R}^n)$$

for $\frac{1}{p} + \frac{1}{q} = 1$, so by Hölder's inequality

$$\int_{\mathbb{R}^n} |u\varphi| \, \mathrm{d}x \leqslant \|u\|_p \|\varphi\|_q.$$

If $1 \leqslant q < \infty$, then

$$\begin{aligned}
\|\varphi\|_q^q &= \int_{\mathbb{R}^n} \frac{1}{1 + |x|^{2nq}} (1 + |x|^{2nq}) |\varphi|^q \, \mathrm{d}x \\
&\leqslant \int_{\mathbb{R}^n} \frac{\mathrm{d}x}{1 + |x|^{2nq}} \left(\sup_x (1 + |x|^{2nq}) |\varphi(x)|^q \right) \\
&\leqslant \int_{\mathbb{R}^n} \frac{\mathrm{d}x}{1 + |x|^{2nq}} \left((\sup_x |\varphi|)^q + (\sup_x |x|^{2n} |\varphi(x)|)^q \right) \\
&\leqslant \int_{\mathbb{R}^n} \frac{\mathrm{d}x}{1 + |x|^{2nq}} \left(S_{0,0}(\varphi)^q + S_{|x|^{2n},0}(\varphi)^q \right) < \infty.
\end{aligned}$$

If $q = \infty$, then $\|\varphi\|_\infty = S_{0,0}(\varphi)$. It follows that $\langle u, \varphi \rangle$ is well-defined and continuous: if $\varphi_j \to \varphi$ in $\mathcal{S}(\mathbb{R}^n)$, then $S_{0,0}(\varphi_j - \varphi) \to 0$, $S_{|x|^{2n},0}(\varphi_j - \varphi) \to 0$ and hence $\langle u, \varphi_j - \varphi \rangle \to 0$. This implies continuity by linearity of u.

The example with $u = \mathrm{e}^{x^2}$ shows that we cannot expect to make sense of general L^p_{loc} functions as tempered distributions: to be a tempered distribution, a function cannot grow too quickly at infinity. This is admittedly quite vague, but it has to be. Indeed, consider $u = \cos(\mathrm{e}^x)$. This is bounded, so $u \in \mathcal{S}'(\mathbb{R})$ by the above. Thus also $u' \in \mathcal{S}'(\mathbb{R})$, but it is easy to see that $u' = -\sin(\mathrm{e}^x)\mathrm{e}^x$. So is exponential growth allowed or not?

Finally, let us record that the Dirac delta function is a tempered distribution:

$$\langle \delta_{x_0}, \varphi \rangle = \varphi(x_0)$$

for $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Note that $\delta_{x_0} = \tau_{-x_0}\delta_0$ and $\hat{\delta}_0 = 1$.

Proposition 11.14. *Let $u : \mathcal{S}(\mathbb{R}^n) \to \mathbb{R}$ (or into $\mathbb{C}$) be linear. Then $u \in \mathcal{S}'(\mathbb{R}^n)$ if and only if there exists constants $c > 0$, $k, l \in \mathbb{N}_0$ such that*

$$|\langle u, \varphi \rangle| \leqslant c\bar{S}_{k,l}(\varphi) \tag{6}$$

holds for all $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Recall that we denoted by $\bar{S}_{k,l}(\varphi) = \max_{|\alpha| \leqslant k,\, |\beta| \leqslant l} S_{\alpha,\beta}(\varphi)$ and $S_{\alpha,\beta}(\varphi) = \sup_x |x^\alpha D^\beta \varphi(x)|$.

Remark 11.15. Tempered distributions have finite order.

Proof. The 'if' part is clear. To prove the 'only if' statement, assume u is φ-continuous but that (6) fails for all $c = k = l = j \in \mathbb{N}$: there exist $\varphi_j \in \mathcal{S}(\mathbb{R}^n)$ such that

$$|\langle u,\, \varphi_j\rangle| > j\bar{S}_{j,j}(\varphi_j).$$

Clearly $\varphi_j \neq 0$, so $\bar{S}_{j,j}(\varphi_j) > 0$ and we may define

$$\psi_j = \frac{\varphi_j}{j\bar{S}_{j,j}(\varphi_j)} \in \mathcal{S}(\mathbb{R}^n).$$

For $\alpha, \beta \in \mathbb{N}_0^n$ we have for $|\alpha|, |\beta| \leqslant j$ that $S_{\alpha,\beta}(\psi_j) \leqslant j^{-1} \to 0$, hence $\psi_j \to 0$ in $\mathcal{S}(\mathbb{R}^n)$ and so $\langle u,\, \psi_j\rangle \to 0$. But this contradicts $|\langle u,\, \psi_j\rangle| > 1$. ☐

12 Lecture 12

Definition 12.1. A function $f\colon \mathbb{R}^n \to \mathbb{R}$ (or into $\mathbb{C}$) is of *polynomial growth* if and only if there exist constants $c > 0$, $m \in \mathbb{N}_0$ such that

$$|f(x)| \leqslant c(1 + |x|)^m$$

for all $x \in \mathbb{R}^n$.

Remark 12.2. Obviously f is of polynomial growth if and only if there exists a polynomial p on $\mathbb{R}^n$ such that $|f(x)| \leqslant |p(x)|$ for all x.

Lemma 12.3. *If $u \in L^\infty_{\mathrm{loc}}(\mathbb{R}^n)$ is of polynomial growth, then*

$$\langle u,\, \varphi\rangle = \int_{\mathbb{R}^n} \varphi(x)u(x)\,\mathrm{d}x$$

for $\varphi \in \mathcal{S}(\mathbb{R}^n)$ is a tempered distribution.

Proof. Let $p \in \mathbb{R}[x]$ be chosen so that

$$|u(x)| \leqslant |p(x)|$$

for almost every x. Then for $\varphi \in \mathcal{S}(\mathbb{R}^n)$

$$\begin{aligned}
|\langle u,\, \varphi\rangle| &\leqslant \int_{\mathbb{R}^n} |p(x)\varphi(x)|\,\mathrm{d}x \\
&= \int_{\mathbb{R}^n} \frac{1}{1 + |x|^{2n}} \left|(1 + |x|^{2n})p(x)\varphi(x)\right|\,\mathrm{d}x \\
&\leqslant \int_{\mathbb{R}^n} \frac{\mathrm{d}x}{1 + |x|^{2n}} \sup_x \left|(1 + |x|^{2n})p(x)\varphi(x)\right| \\
&= cS_{P,0}(\varphi),
\end{aligned}$$

where $c := \int_{\mathbb{R}^n} \frac{\mathrm{d}x}{1+|x|^{2n}}$, $P(x) = (1 + |x|^{2n})p(x)$. ☐

Definition 12.4. A function $a \in C^\infty(\mathbb{R}^n)$ is said to be of *moderate growth* if and only if a and all partial derivatives $D^\alpha a$, $\alpha \in \mathbb{N}_0^n$, have polynomial growth: $\forall \alpha \in \mathbb{N}_0^n \ \exists p_\alpha \in \mathbb{R}[x]$ such that

$$|D^\alpha a(x)| \leqslant |p_\alpha(x)|$$

for all $x \in \mathbb{R}^n$.

Example 12.5. Polynomials have moderate growth.

Lemma 12.6. *Let $a \in C^\infty(\mathbb{R}^n)$ be of moderate growth (a is a moderate C^∞ function) and $u \in \mathcal{S}'(\mathbb{R}^n)$. Then*

$$\langle au, \varphi \rangle := \langle u, a\varphi \rangle,$$

$\varphi \in \mathcal{S}(\mathbb{R}^n)$, defines a tempered distribution au. Furthermore, the map

$$\mathcal{S}'(\mathbb{R}^n) \ni u \longmapsto au \in \mathcal{S}'(\mathbb{R}^n)$$

is linear and continuous.

Proof. Fix $\alpha, \beta \in \mathbb{N}_0^n$. Then for $\varphi \in \mathcal{S}(\mathbb{R}^n)$ we compute using the Leibniz rule

$$x^\alpha D^\beta(a\varphi) = \sum_{\gamma \leqslant \beta} \left(\begin{array}{c} \beta \\ \gamma \end{array} \right) (D^\gamma a) x^\alpha (D^{\beta-\gamma}\varphi).$$

For each $\gamma \leqslant \beta$, $D^\gamma a$ has polynomial growth so we can find constants $c_\gamma > 0$, $m_\gamma \in \mathbb{N}_0$ such that

$$|D^\gamma a(x)| \leqslant c_\gamma (1 + |x|)^{m_\gamma}$$

for all $x \in \mathbb{R}^n$. Here

$$(1 + |x|)^{m_\gamma} \leqslant (1 + |x|)^{2m_\gamma} \leqslant 2^{2m_\gamma - 1}(1 + |x|^{2m_\gamma})$$

and

$$|x|^{2m_\gamma} = (x_1^2 + \cdots + x_n^2)^{m_\gamma} \leqslant n^{m_\gamma - 1}(x_1^{2m_\gamma} + \cdots + x_n^{2m_\gamma}),$$

and so

$$|D^\gamma a(x)| \leqslant 2^{2m_\gamma - 1} c_\gamma + 2^{2m_\gamma - 1} n^{m_\gamma - 1} c_\gamma \sum_{j=1}^n x_j^{2m_\gamma}.$$

Put $\bar{m} = \max_{\gamma \leqslant \beta} m_\gamma$, $\bar{c} = 2^{2m-1} n^m \max_{\gamma \leqslant \beta} c_\gamma$. Then

$$|D^\gamma a(x)| \leqslant \bar{c} \left(1 + \sum_{j=1}^n x_j^{2\bar{m}} \right)$$

and so

$$\begin{aligned}
S_{\alpha,\beta}(a\varphi) &\leqslant \bar{c}\sum_{\gamma\leqslant\beta}\begin{pmatrix}\beta\\ \gamma\end{pmatrix}\sup_x\left|\left(1+\sum_{j=1}^{n}x_j^{2\bar{m}}\right)x^\alpha D^{\beta-\gamma}\varphi\right| \\
&\leqslant \bar{c}\sum_{\gamma\leqslant\beta}\begin{pmatrix}\beta\\ \gamma\end{pmatrix}\left(S_{\alpha,\beta-\gamma}(\varphi)+n\bar{S}_{|\alpha|+2\bar{m},|\beta|}(\varphi)\right) \\
&\leqslant \bar{c}\sum_{\gamma\leqslant\beta}\begin{pmatrix}\beta\\ \gamma\end{pmatrix}(n+1)\bar{S}_{|\alpha|+2\bar{m},|\beta|}(\varphi) \\
&\leqslant c\bar{S}_{|\alpha|+2\bar{m},|\beta|}(\varphi).
\end{aligned}$$

It follows that $\varphi\mapsto a\varphi$ is $\mathcal{S}$-continuous and hence $au\in\mathcal{S}'(\mathbb{R}^n)$ since the linearity of au is clear once we know it is well-defined. Next, $u\mapsto au$ is clearly linear and $\mathcal{S}'$-continuous: the latter follows by definition-chasing. Indeed, when $u_j\to u$ in $\mathcal{S}'(\mathbb{R}^n)$, then

$$\langle u_j,\, a\varphi\rangle \longrightarrow \langle u,\, a\varphi\rangle$$

for each $\varphi\in\mathcal{S}(\mathbb{R}^n)$, that is

$$\langle au_j,\, \varphi\rangle \longrightarrow \langle au,\, \varphi\rangle$$

for each $\varphi\in\mathcal{S}(\mathbb{R}^n)$. □

Lemma 12.7. *If $u\in\mathcal{S}'(\mathbb{R}^n)$ and $\theta\in\mathcal{S}(\mathbb{R}^n)$, then $u*\theta$ can be defined the adjoint identity scheme as*

$$\langle u*\theta,\, \varphi\rangle = \langle u,\, \tilde{\theta}*\varphi\rangle$$

for all $\varphi\in\mathcal{S}(\mathbb{R}^n)$. Furthermore, $u*\theta$ is a C^∞ function of moderate growth and is given by

$$(u*\theta)(x) = \langle u,\, \theta(x-\cdot)\rangle$$

for $x\in\mathbb{R}^n$.

Remark 12.8. We have not emphasized it so far, but since the convolution product is commutative on $\mathcal{S}(\mathbb{R}^n)$, $\varphi*\psi=\psi*\varphi$ for $\varphi,\,\psi\in\mathcal{S}(\mathbb{R}^n)$, we also have $u*\theta=\theta*u$. The proof is easy (exercise).

Definition 12.9. We can define convolutions of $u,\,v\in\mathcal{S}'(\mathbb{R}^n)$ provided $\hat{v}$ is a C^∞ function of moderate growth:

$$u*v := \mathcal{F}^{-1}(\hat{u}\hat{v}).$$

This is a good definition by virtue of the Fourier Inversion Formula on $\mathcal{S}'(\mathbb{R}^n)$ and lemma 12.6.

The various rules for the Fourier transform continue to hold for tempered distributions.

Theorem 12.10 (Convolution Rule)**.** *Let $u,\,v\in\mathcal{S}'(\mathbb{R}^n)$ and assume v is a C^∞ function of moderate growth. Then $uv\in\mathcal{S}'(\mathbb{R}^n)$ and $\widehat{uv}=(2\pi)^{-n}\hat{u}*\hat{v}$.*

Proof. By the Inversion Formula we define

$$\hat{u} * \hat{v} := \mathcal{F}^{-1}\left(\hat{\hat{u}}\hat{\hat{v}}\right),$$

where $\hat{\hat{u}} = (2\pi)^n\tilde{u}$, $\hat{\hat{v}} = (2\pi)^n\tilde{v}$, and the latter is clearly a C^∞ function of moderate growth. Hence $\hat{\hat{u}}\hat{\hat{v}} \in \mathcal{S}'(\mathbb{R}^n)$ by lemma 12.6, and using $\tilde{u}\tilde{v} = \widetilde{uv}$ we get

$$\hat{u} * \hat{v} = \mathcal{F}^{-1}\left((2\pi)^{2n}\widetilde{uv}\right) = \tilde{\mathcal{F}}\left((2\pi)^{2n}\widetilde{uv}\right) = (2\pi)^n\widehat{uv}.$$

□

Theorem 12.11 (Plancherel's Theorem)**.** *The Fourier transform $\mathcal{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is bijective, and $(2\pi)^{-n/2}\mathcal{F}$ is* unitary *(isometric and onto). That is, $\mathcal{F}(L^2) = L^2$ and*

$$\|\hat{f}\|_{L^2} = (2\pi)^{n/2}\|f\|_{L^2}$$

for $f \in L^2(\mathbb{R}^n)$, and more generally

$$\int_{\mathbb{R}^n} f(x)\overline{g(x)}\,\mathrm{d}x = (2\pi)^{-n}\int_{\mathbb{R}^n} \hat{f}(\xi)\overline{\hat{g}(\xi)}\,\mathrm{d}\xi.$$

for f, $g \in L^2(\mathbb{R}^n)$.

Proof. We start by observing that for φ, $\psi \in \mathcal{S}(\mathbb{R}^n)$

$$\int_{\mathbb{R}^n} \varphi\bar{\psi}\,\mathrm{d}x = (2\pi)^{-n}\int_{\mathbb{R}^n} \hat{\varphi}\bar{\hat{\psi}}\,\mathrm{d}\xi,$$

and in particular for $\varphi = \psi$

$$\int_{\mathbb{R}^n} |\varphi|^2\,\mathrm{d}x = (2\pi)^{-n}\int_{\mathbb{R}^n} |\hat{\varphi}|^2\,\mathrm{d}\xi. \tag{7}$$

This follows from the Product Rule and the Inversion Formula on $\mathcal{S}(\mathbb{R}^n)$: clearly $\bar{\psi} \in \mathcal{S}(\mathbb{R}^n)$, so $\mathcal{F}^{-1}(\bar{\psi}) \in \mathcal{S}(\mathbb{R}^n)$ and so

$$\int \varphi\bar{\psi}\,\mathrm{d}x = \int \varphi\mathcal{F}(\mathcal{F}^{-1}\bar{\psi})\,\mathrm{d}x = \int \hat{\varphi}\mathcal{F}^{-1}(\bar{\psi})\,\mathrm{d}x.$$

Now

$$\mathcal{F}^{-1}(\bar{\psi})(x) = (2\pi)^{-n}\int \bar{\psi}(y)\mathrm{e}^{ix\cdot y}\,\mathrm{d}y = (2\pi)^{-n}\overline{\int \psi(y)\mathrm{e}^{-ix\cdot y}\,\mathrm{d}y} = (2\pi)^{-n}\overline{\hat{\psi}(x)}.$$

If now $f \in L^2(\mathbb{R}^n)$ we know that there exist $f_j \in \mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$ so $\|f - f_j\|_{L^2} \to 0$. Clearly this means in particular that $f_j \to f$ in $\mathcal{S}'(\mathbb{R}^n)$, and thus by $\mathcal{S}'$-continuity of the Fourier transform, $\hat{f}_j \to \hat{f}$ in $\mathcal{S}'(\mathbb{R}^n)$. By (7) we see that

$$\int_{\mathbb{R}^n} |\hat{f}_j - \hat{f}_k|^2\,\mathrm{d}\xi = (2\pi)^n\int_{\mathbb{R}^n} |f_j - f_k|^2\,\mathrm{d}x,$$

so $(\hat{f}_j)$ is Cauchy. It is thus convergent in L^2 (by Riesz–Fischer), $\hat{f}_j \to g$ in $L^2(\mathbb{R}^n)$ for some $g \in L^2(\mathbb{R}^n)$. Clearly then $\hat{f}_j \to g$ in $\mathcal{S}'(\mathbb{R}^n)$ too, and so $g = \hat{f}$. □

What happens on the other L^p spaces?

Theorem 12.12 (Hausdorff–Young)**.** *For $p \in (1,2)$ and $\frac{1}{p} + \frac{1}{q} = 1$ we have for $f \in L^p(\mathbb{R}^n)$ that $\hat{f} \in L^q(\mathbb{R}^n)$ with*

$$\|\hat{f}\|_{L^q} \leqslant (2\pi)^{n/q} \|f\|_{L^p}.$$

Remark 12.13. For $p > 2$ the image $\mathcal{F}(L^p(\mathbb{R}^n))$ contains tempered distributions of positive orders.

13 Lecture 13

Recall that a partial differential operator (PDO) with constant coefficients can be written as

$$P(D) = \sum_{|\alpha| \leqslant m} a_\alpha D^\alpha$$

for $a_\alpha \in \mathbb{C}$, and if $a_\alpha \neq 0$ for some α with $|\alpha| = m$, we say $P(D)$ is of order m.

Definition 13.1. A *fundamental solution* for $P(D)$ is any $E \in \mathcal{S}'(\mathbb{R}^n)$ such that $P(D)E = \delta_0$.

Example 13.2. Recall from Problem Sheet 1 that $E(x) = \frac{x_+^m}{m!}$ satisfies

$$\frac{\mathrm{d}^{m+1}}{\mathrm{d}x^{m+1}} E = \delta_0$$

in $\mathcal{D}'(\mathbb{R})$, and from Problem Sheet 2 that $E(x) = \frac{1}{2\pi} \log |x|$ satisfies

$$\Delta E = \delta_0$$

in $\mathcal{D}'(\mathbb{R}^2)$. It is not difficult to check that in both cases the distribution E is tempered and so is a fundamental solution.

We refer to the polynomial

$$p(\xi) = \sum_{|\alpha| \leqslant m} a_\alpha \xi^\alpha$$

as the *symbol* of the PDO $P(D)$, and, provided $P(D)$ has order m, we call

$$\sum_{|\alpha| = m} a_\alpha \xi^\alpha$$

the *principal symbol* of $P(D)$.

Definition 13.3. A PDO
$$P(D) = \sum_{|\alpha| \leqslant m} a_\alpha D^\alpha$$
of order m is called *elliptic* if and only if the principal symbol satisfies
$$\sum_{|\alpha|=m} a_\alpha \xi^\alpha \neq 0$$
for all $\xi \in \mathbb{R}^n \setminus \{0\}$.

The most prominent examples of elliptic PDOs are the Laplacian on $\mathbb{R}^n$,
$$\Delta = D_1^2 + \cdots + D_n^2,$$
and the Cauchy–Riemann operators on the plane,
$$\frac{\partial}{\partial \bar{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right), \qquad \frac{\partial}{\partial z} = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right).$$
They are of orders 2 and 1 respectively, and their principal symbols are
$$-|\xi|^2, \quad -\frac{1}{2}(\xi_1 + i\xi_2), \quad -\frac{1}{2}(\xi_1 - i\xi_2),$$
respectively. Clearly the condition for ellipticity is satisfied by each of them.

Example 13.4. Recall that on $\mathbb{R}^2 \simeq \mathbb{C}$ we have
$$\Delta = 4\frac{\partial^2}{\partial z \partial \bar{z}},$$
and $\frac{1}{2\pi} \log|z|$ is a fundamental solution for Δ:
$$\begin{aligned}
\delta_0 &= \Delta\left(\frac{1}{2\pi}\log|z|\right) \\
&= \frac{2}{\pi}\frac{\partial^2}{\partial \bar{z}\partial z}(\log|z|) \\
&= \frac{1}{\pi}\frac{\partial}{\partial \bar{z}}\left(\frac{\partial}{\partial z}\log(z\bar{z})\right) \\
&= \frac{1}{\pi}\frac{\partial}{\partial \bar{z}}\left(\frac{1}{z\bar{z}}\bar{z}\right) \\
&= \frac{\partial}{\partial \bar{z}}\left(\frac{1}{\pi z}\right),
\end{aligned}$$
so $(\pi z)^{-1}$ is a fundamental solution for $\frac{\partial}{\partial \bar{z}}$.

Theorem 13.5. *Assume E is a fundamental solution for a PDO $P(D)$. Then for $f \in \mathcal{S}(\mathbb{R}^n)$ the general solution to*

$$P(D)u = f$$

in $\mathcal{S}'(\mathbb{R}^n)$ is given by

$$u = E * f + h,$$

where $h \in \mathcal{S}'(\mathbb{R}^n) \cap \ker P(D)$.

Remark 13.6. We can allow any $f \in \mathcal{S}'(\mathbb{R}^n)$ on the right-hand side for which we can define $E * f$ as a tempered distribution (for instance, if $\hat{f}$ is a moderate C^∞ function, but also much more generally; we have not defined convolutions in full generality).

Proof. We simply chase definitions: $E * f \in \mathcal{S}'(\mathbb{R}^n)$ and

$$P(D)(E * f) = (P(D)E) * f = \delta_0 * f = f,$$

so $E * f$ is a particular solution and hence the general solution is as stated. □

Theorem 13.7. *Let $n \geqslant 3$. Then*

$$E(x) = -\frac{1}{(n-2)\omega_{n-1}} |x|^{2-n},$$

$x \in \mathbb{R}^n \setminus \{0\}$, is a fundamental solution for Δ.

Remark 13.8. Note that E is C^∞ away from zero, $E \in L^1_{\mathrm{loc}}(\mathbb{R}^n) \cap \mathcal{S}'(\mathbb{R}^n)$, and that

$$D_j E = \frac{1}{\omega_{n-1}} x_j |x|^{-n} \in L^1_{\mathrm{loc}}(\mathbb{R}^n).$$

The constant ω_{n-1} is the surface area of $\mathbb{S}^{n-1}$ in $\mathbb{R}^n$. One can show that $\omega_{n-1} = n\mathcal{L}^n(B_1(0))$, and

$$\mathcal{L}^n(B_1(0)) = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2} + 1\right)}$$

for all $n \in \mathbb{N}$. In particular, we record the values for $n = 2$ and 3:

$$\omega_1 = 2\mathcal{L}^2(B_1(0)) = 2\pi, \qquad \omega_2 = 3\mathcal{L}^3(B_1(0)) = 4\pi.$$

The calculation uses $\Gamma(x+1) = x\Gamma(x)$ and $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. Note also that

- in $n = 1$ the Laplacian $\frac{\mathrm{d}^2}{\mathrm{d}x^2}$ has fundamental solution x^+,
- in $n = 2$ the Laplacian $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y}$ has fundamental solution $\frac{1}{2\pi} \log \sqrt{x^2 + y^2}$. This is called the *logarithmic potential*;

- in $n = 3$ the Laplacian $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ has fundamental solution
$$-\frac{1}{4\pi}\frac{1}{\sqrt{x^2+y^2+z^2}}.$$
This is called the *Newtonian potential*.

Proof. Fourier transforming $\Delta E = \delta_0$ we get
$$1 = \hat{\delta}_0 = \widehat{\Delta E} = -|\xi|^2 \hat{E}(\xi).$$
This is not enough to deduce that $\hat{E}(\xi) = -\frac{1}{|\xi|^2}$, only that
$$\hat{E}(\xi) = -\frac{1}{|\xi|^2} + \hat{T}$$
for some $T \in \mathcal{S}'(\mathbb{R}^n)$ satisfies $\Delta T = 0$. This means that $-|\xi|^2\hat{T} = 0$ in $\mathcal{S}'(\mathbb{R}^n)$. Hence if $\varphi \in \mathcal{S}(\mathbb{R}^n)$ and $0 \notin \operatorname{supp}\varphi$, then
$$\psi(\xi) = -\frac{\varphi(\xi)}{|\xi|^2}$$
for $\xi \neq 0$ and $\psi(0) = 0$ belongs to $\mathcal{S}(\mathbb{R}^n)$, and so
$$\langle \hat{T}, \varphi \rangle = \langle -|\xi|^2 \hat{T}, \psi \rangle = 0.$$
We express this by $\operatorname{supp}\hat{T} = \{0\}$, that is $\hat{T}$ has support $\{0\}$. We discuss below how this implies that $\hat{T} \in \operatorname{span}\{D^\alpha \delta_0 \,:\, \alpha \in \mathbb{N}_0^n\}$, and hence that $T \in \operatorname{span}\{(2\pi)^{-n}(ix)^\alpha \,:\, \alpha \in \mathbb{N}_0^n\} = \mathbb{C}[x]$. Since also $\Delta T = 0$, we see that T must be a harmonic polynomial. Note that implicit in this is the Liouville-type result saying that if $T \in \mathcal{S}'(\mathbb{R}^n)$ is harmonic, then T is a polynomial. Now we return to the quest for fundamental solutions:
$$\hat{E}(\xi) = -\frac{1}{|\xi|^2} + \hat{T}(\xi).$$
We only need one, so consider $\hat{E} = -\frac{1}{|\xi|^2}$. The result then follows from the following.

Lemma 13.9 (Auxiliary Lemma)**.** *Let $\alpha \in (-n, 0)$ and put $f(x) = |x|^\alpha$. Then $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n) \cap \mathcal{S}'(\mathbb{R}^n)$ and $\hat{f}(\xi) = c(n,\alpha)|\xi|^{-n-\alpha}$, where*
$$c(n,\alpha) = 2^{\alpha+n}\pi^{\frac{n}{2}} \frac{\Gamma\left(\frac{n+\alpha}{2}\right)}{\Gamma\left(-\frac{\alpha}{2}\right)}$$
and $-n < -n-\alpha < 0$.

Proof. We start with the observation that for $x \neq 0$

$$|x|^\alpha \Gamma\left(-\frac{\alpha}{2}\right) = |x|^\alpha \int_0^\infty t^{-\frac{\alpha}{2}-1} \mathrm{e}^{-t}\, \mathrm{d}t = \int_0^\infty s^{-\frac{\alpha}{2}-1} \mathrm{e}^{-s|x|^2}\, \mathrm{d}s,$$

where we made the substitution $t = s|x|^2$, and hence

$$|x|^\alpha = \frac{1}{\Gamma\left(-\frac{\alpha}{2}\right)} \int_0^\infty s^{-\frac{\alpha}{2}-1} \mathrm{e}^{-s|x|^2}\, \mathrm{d}s.$$

Note that

$$\frac{1}{\Gamma\left(-\frac{\alpha}{2}\right)} \int_0^j s^{-\frac{\alpha}{2}-1} \mathrm{e}^{-s|x|^2}\, \mathrm{d}s \underset{j\to\infty}{\longrightarrow} |x|^\alpha$$

in $\mathcal{S}'(\mathbb{R}^n)$ and Riemann sums for the integrals

$$\frac{1}{\Gamma\left(-\frac{\alpha}{2}\right)} \int_0^j s^{-\frac{\alpha}{2}-1} \mathrm{e}^{-s|x|^2}\, \mathrm{d}s.$$

for j fixed converge as mesh size tends to zero in the $\mathcal{S}'(\mathbb{R}^n)$ sense. Consequently we get by $\mathcal{S}'$-continuity and linearity of $\mathcal{F}$ that

$$\begin{aligned}
\mathcal{F}_{x\to\xi}(|x|^\alpha) &= \frac{1}{\Gamma\left(-\frac{\alpha}{2}\right)} \int_0^\infty s^{-\frac{\alpha}{2}-1} \mathcal{F}_{x\to\xi}(\mathrm{e}^{-s|x|^2})\, \mathrm{d}s \\
&= \frac{1}{\Gamma\left(-\frac{\alpha}{2}\right)} \int_0^\infty s^{-\frac{\alpha}{2}-1} \left(\frac{\pi}{s}\right)^{\frac{n}{2}} \mathrm{e}^{-\frac{|\xi|^2}{4s}}\, \mathrm{d}s \\
&= \frac{\pi^{\frac{n}{2}}}{\Gamma\left(-\frac{\alpha}{2}\right)} \left(\frac{|\xi|}{2}\right)^{-n-\alpha} \int_0^\infty t^{\frac{n+\alpha}{2}-1} \mathrm{e}^{-t}\, \mathrm{d}t \\
&= 2^{n+\alpha} \pi^{\frac{n}{2}} \frac{\Gamma\left(\frac{n+\alpha}{2}\right)}{\Gamma\left(-\frac{\alpha}{2}\right)} |\xi|^{-n-\alpha}.
\end{aligned}$$

□

□

13.1 Localization of Distributions

Theorem 13.10. *If $u \in \mathcal{D}'(\Omega)$ and for each $x \in \Omega$ there exists $r_x > 0$ such that $\langle u, \varphi\rangle = 0$ for all $\varphi \in \mathcal{D}(\Omega \cap B_{r_x}(x))$, then $u = 0$.*

Remark 13.11. For an open subset $\omega \subset \Omega$ we define the *restriction* of $u \in \mathcal{D}'(\Omega)$ to ω, denoted $u|_\omega$, by

$$\langle u|_\omega, \varphi\rangle := \langle u, \varphi\rangle$$

for all $\varphi \in \mathcal{D}(\omega)$. Clearly $u|_\omega \in \mathcal{D}'(\omega)$, and the assumption in the above theorem is that $u|_{\Omega\cap B_{r_x}(x)} = 0$.

Proof. (*Not examinable*) Let $\varphi \in \mathcal{D}(\Omega)$. Since we clearly have

$$\operatorname{supp}\varphi \subset \bigcup \{\Omega \cap B_{r_x}(x) \,:\, x \in \Omega\},$$

the compactness of $\operatorname{supp}\varphi$ means that we can find a finite subcover, say

$$\operatorname{supp}\varphi \subset \big(\Omega \cap B_{r_{x_1}}(x_1)\big) \cup \cdots \cup \big(\Omega \cap B_{r_{x_m}}(x_m)\big).$$

Using Theorem 2.8 we find a partition of unity $\phi_1, \ldots, \phi_m \in \mathcal{D}(\Omega)$ with $\operatorname{supp}\phi_j \subset \Omega \cap B_{r_{x_j}}(x_j)$, $0 \leqslant \phi_j \leqslant 1$, and $\sum_{j=1}^m \phi_j = 1$ on $\operatorname{supp}\varphi$. Thus

$$\langle u,\, \varphi\rangle = \left\langle u,\, \sum_{j=1}^m \varphi\phi_j \right\rangle = \sum_{j=1}^m \langle u,\, \varphi\phi_j\rangle = 0.$$

□

Definition 13.12. For $u \in \mathcal{D}'(\Omega)$ the *support* of u is

$$\operatorname{supp} u := \{x \in \Omega \,:\, u|_{\Omega\cap B_r(x)} \neq 0 \text{ for all } r > 0\}.$$

Thus $x \in \Omega \setminus \operatorname{supp} u$ if and only if there exists $r > 0$ such that $u|_{\Omega\cap B_r(x)} = 0$. Consequently the above theorem means in particular that $\Omega \setminus \operatorname{supp} u$ must be the largest open subset of Ω on which u vanishes. It also follows from this that the support of u is a relatively closed subset of Ω.

Theorem 13.13. *Let $u \in \mathcal{D}'(\Omega)$ and $x_0 \in \Omega$. If $\operatorname{supp} u = \{x_0\}$, then $u \in \operatorname{span}\{D^\alpha \delta_{x_0} \,:\, \alpha \in \mathbb{N}_0^n\}$.*

The proof is omitted. □

Example 13.14. If $u \in C(\Omega)$, its support is by definition

$$\operatorname{supp} u = \Omega \cap \overline{\{x \in \Omega \,:\, u(x) \neq 0\}}.$$

We can also obviously consider the corresponding distribution

$$\langle u,\, \varphi\rangle = \int_\Omega u\varphi \,\mathrm{d}x$$

for $\varphi \in \mathcal{D}(\Omega)$. Its support as a distribution coincides with the above definition of the support for a continuous function and we may therefore use the same notation for both: In order to justify this let D be the support of the distribution. If $\varphi \in \mathcal{D}(\Omega)$ and $\operatorname{supp}\varphi \subset \Omega \setminus \operatorname{supp} u$, then

$$\langle u,\, \varphi\rangle = \int_\Omega u\varphi \,\mathrm{d}x = 0,$$

so the open set $\Omega \setminus \operatorname{supp} u$ must be contained in the largest open set on which u vanishes, that is $\Omega \setminus \operatorname{supp} u \subset \Omega \setminus D$, i.e. $D \subset \operatorname{supp} u$. If $x_0 \in \operatorname{supp} u \setminus D$, then we find $\varphi \in \mathcal{D}(\Omega)$ supported near x_0 and in $\Omega \setminus D$ such that $\langle u,\, \varphi\rangle \neq 0$. □

14 Lecture 14

Theorem 14.1 (Weyl's Lemma)**.** *Assume $u \in \mathcal{D}'(\Omega)$ and $\Delta u = 0$ in $\mathcal{D}'(\Omega)$. Then $u \in C^\infty(\Omega)$ and u is harmonic.*

Corollary 14.2. *Let $\Omega \subset \mathbb{C}$ be open and assume $f \in \mathcal{D}'(\Omega)$ satisfies*

$$\frac{\partial f}{\partial \bar{z}} = 0$$

in $\mathcal{D}'(\Omega)$. Then f is holomorphic.

Proof. This is clear since we obviously also have that

$$\Delta f = 4\frac{\partial}{\partial z}\left(\frac{\partial}{\partial \bar{z}} f\right) = 0$$

in $\mathcal{D}'(\Omega)$. Weyl's Lemma then implies that f is C^∞, in which case distributional and classical derivatives coincide. Thus f satisfies the usual Cauchy–Riemann equations and is a holomorphic function. $\square$

Proof of Weyl's Lemma. Let $(\rho_\varepsilon)_{\varepsilon>0}$ be the standard mollifier. Fix $\Omega' \Subset \Omega$ and put $\varepsilon_0 = \operatorname{dist}(\Omega', \partial\Omega)$. For each $x \in \Omega'$ and $\varepsilon \in (0, \varepsilon_0)$ the function

$$y \longmapsto \rho_\varepsilon(x-y)$$

belongs to $\mathcal{D}(\Omega)$ and so we may consider $\langle u, \rho_\varepsilon(x - \cdot)\rangle$. We assert that it is independent of $\varepsilon \in (0, \varepsilon_0)$. To prove it we calculate $\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\rho_\varepsilon(x-y)$ for $x, y \in \mathbb{R}^n$. Recall that

$$\rho_\varepsilon(x-y) = \varepsilon^{-n}\rho\left(\frac{x-y}{\varepsilon}\right)$$

and that $\rho(x) = \theta(|x|^2)$ (since ρ was a smooth radial function), where $\theta \in C^\infty(\mathbb{R})$ satisfies $\theta(t) = 0$ for $t \geqslant 1$. Now calculate

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left(\varepsilon^{-n}\rho\left(\frac{x-y}{\varepsilon}\right)\right) &= -n\varepsilon^{-n-1}\rho\left(\frac{x-y}{\varepsilon}\right) - \varepsilon^{-n}\nabla\rho\left(\frac{x-y}{\varepsilon}\right)\cdot\frac{x-y}{\varepsilon^2}\\ &= -\frac{1}{\varepsilon^{n+1}}\left(n\rho\left(\frac{x-y}{\varepsilon}\right) + \nabla\rho\left(\frac{x-y}{\varepsilon}\right)\cdot\frac{x-y}{\varepsilon}\right).\end{aligned}$$

Put $K(x) = -n\rho(x) - \nabla\rho(x)\cdot x$ so that

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left(\varepsilon^{-n}\rho\left(\frac{x-y}{\varepsilon}\right)\right) = \frac{1}{\varepsilon^{n+1}}K\left(\frac{x-y}{\varepsilon}\right).$$

We now use that $\rho(x) = \theta(|x|^2)$. Hereby

$$K(x) = -\operatorname{div}\left(\rho(x)x\right) = -\operatorname{div}\left(\theta(|x|^2)x\right)$$

and if we set

$$\Theta(t) = \frac{1}{2}\int_t^\infty \theta(s)\,\mathrm{d}s,$$

then $\Theta \in C^\infty(\mathbb{R})$ with $\Theta(t) = 0$ for $t \geqslant 1$, and $\Theta'(t) = -\frac{1}{2}\theta(t)$. Consequently

$$-\theta(|x|^2)x = \nabla\left(\Theta(|x|^2)\right),$$

and so $K(x) = \operatorname{div}\nabla\left(\Theta(|x|^2)\right) = (\Delta\Phi)(x)$, where $\Phi(x) = \Theta(|x|^2)$. Observe that $\Phi \in \mathcal{D}(\overline{B_1(0)})$, and

$$\begin{aligned}-\frac{1}{\varepsilon^{n+1}}\left(n\rho\left(\frac{x-y}{\varepsilon}\right) + \nabla\rho\left(\frac{x-y}{\varepsilon}\right)\cdot\frac{x-y}{\varepsilon}\right) &= \frac{1}{\varepsilon^{n+1}}\Delta_y\left(\Phi\left(\frac{x-y}{\varepsilon}\right)\right)\\ &= \Delta_y\left(\varepsilon^{1-n}\Phi\left(\frac{x-y}{\varepsilon}\right)\right).\end{aligned}$$

Here $y \mapsto \varepsilon^{1-n}\Phi\left(\frac{x-y}{\varepsilon}\right)$ is supported in $\overline{B_\varepsilon(x)} \subset \Omega$, and so by assumption

$$\left\langle u,\, \Delta_y\left(\varepsilon^{1-n}\Phi\left(\frac{x-y}{\varepsilon}\right)\right)\right\rangle = 0.$$

Now by considering difference quotients we see that

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\langle u,\, \rho_\varepsilon(x-\cdot)\rangle = \left\langle u,\, \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\rho_\varepsilon(x-\cdot)\right\rangle.$$

Indeed, for ε, $\varepsilon' > 0$ we have

$$\begin{aligned}\frac{\rho_{\varepsilon+\varepsilon'}(x-y) - \rho_\varepsilon(x-y)}{\varepsilon'} &\overset{\text{FTC}}{=} \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\rho_{\varepsilon+t\varepsilon'}(x-y)\,\mathrm{d}t\\ &\underset{\varepsilon'\to 0}{\longrightarrow} \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=\varepsilon}\rho_s(x-y)\end{aligned}$$

in $\mathcal{D}'(\Omega)$ with respect to y, provided $x \in \Omega'$ and $0 < \varepsilon < \varepsilon_0$ (since we may differentiate both sides with respect to y). But then $\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\langle u,\, \rho_\varepsilon(x-\cdot)\rangle = 0$, and so $\langle u,\, \rho_\varepsilon(x-\cdot)\rangle = \langle u,\, \rho_{\varepsilon_1}(x-\cdot)\rangle$ for all $\varepsilon \in (0,\varepsilon_0)$, where $\varepsilon_1 \in (0,\varepsilon_0)$. Now let $\varphi \in \mathcal{D}(\Omega')$. Then, by the usual trick when convolving distributions with test functions,

$$\begin{aligned}\int_{\Omega'}\langle u,\, \rho_\varepsilon(x-\cdot)\rangle\varphi(x)\,\mathrm{d}x &= \left\langle u,\, \int_{\Omega'}\rho_\varepsilon(x-\cdot)\varphi(x)\,\mathrm{d}x\right\rangle\\ &= \langle u,\, \rho_\varepsilon * \varphi\rangle,\end{aligned}$$

and so for $\varepsilon \in (0,\varepsilon_1)$ we have

$$\langle u,\, \rho_\varepsilon * \varphi\rangle = \int_{\Omega'}\langle u,\, \rho_{\varepsilon_1}(x-\cdot)\rangle\varphi(x)\,\mathrm{d}x.$$

Hence, as $\rho_\varepsilon * \varphi \to \varphi$ in $\mathcal{D}(\Omega)$ as $\varepsilon \to 0^+$, we get

$$\langle u, \varphi\rangle = \int_{\Omega'} \langle u, \rho_{\varepsilon_1}(x - \cdot)\rangle \varphi(x)\,\mathrm{d}x.$$

Consequently $u|_{\Omega'} \in C^\infty(\Omega')$, and since Ω' was arbitrary, we are done. $\square$

Remark 14.3. The above proof is inspired by the mean value property that is known to characterize harmonic functions in the following sense. Let $h \in C(\Omega)$. Then h is harmonic in the usual sense ($h \in C^2(\Omega)$ and $\Delta h = 0$) if and only if for all balls $B_r(x_0) \Subset \Omega$ we have

$$h(x_0) = \frac{1}{\omega_{n-1} r^{n-1}} \int_{\partial B_r(x_0)} h(x)\,\mathrm{d}S_x.$$

Using polar coordinates we see that when h is harmonic, then for $B_r(x_0) \Subset \Omega$

$$h(x_0) = (\rho_r * h)(x_0).$$

15 Lecture 15

In this lecture we use the Plancherel Theorem to obtain estimates and regularity results for distributional solutions to the Poisson equation.

We are interested in solving the Poisson equation $\Delta u = f$ in $\mathbb{R}^n$ in the context of tempered distributions. If E denotes the fundamental solution for Δ in $\mathbb{R}^n$ found in Example 13.2 (for $n = 2$) and in Theorem 13.7 (for $n \geqslant 3$), then the general solution in $\mathcal{S}'(\mathbb{R}^n)$ is $E * f + h$, where h is any harmonic polynomial on $\mathbb{R}^n$. This follows from Theorem 13.5 *provided* we can make sense of $E * f$ as a tempered distribution.

Example 15.1. If $f \in \mathcal{S}'(\mathbb{R}^n)$ has compact support, then $\hat{f}$ is a moderate C^∞ function and so $\hat{E}\hat{f}$ is well-defined as a tempered distribution by Lemma 12.6. We can then define

$$E * f := \mathcal{F}^{-1}\left(\hat{E}\hat{f}\right).$$

This is consistent and a good definition (extending the Convolution Rule).

Theorem 15.2 (An L^2 identity for the Laplacian)**.** *Let $f \in L^2(\mathbb{R}^n)$ and assume $E * f$ is well-defined as a tempered distribution. Then the general solution to Poisson's equation $\Delta v = f$ in $\mathcal{S}'(\mathbb{R}^n)$ is $v = E * f + h$, where h is any harmonic polynomial on $\mathbb{R}^n$. Furthermore, if $u = E * f$, then $D_j D_k u \in L^2(\mathbb{R}^n)$ for $1 \leqslant j, k \leqslant n$, and*

$$\sum_{j,k=1}^{n} \int_{\mathbb{R}^n} |D_j D_k u|^2\,\mathrm{d}x = \int_{\mathbb{R}^n} |f|^2\,\mathrm{d}x. \tag{8}$$

Remark 15.3. The symmetric $n \times n$ matrix

$$D^2 u = (D_j D_k u)$$

is called the *Hessian matrix* of u. When u is a distribution with the property that the second order partial derivatives $D_j D_k u$ are *regular distributions* (i.e. they are L^1_{loc} functions), then

$$|D^2 u|^2 = \sum_{j,k=1}^{n} |D_j D_k u|^2.$$

The right-hand side serves to define the left-hand side, and we record that for an $n \times n$ matrix $A = (a_{ij}) \in M_{n\times n}(\mathbb{C})$

$$|A| := \sqrt{\operatorname{tr}(A^\top \bar{A})} = \left(\sum_{j,k=1}^{n} |a_{jk}|^2 \right)^{\frac{1}{2}}.$$

This is the standard norm on $M_{n\times n}(\mathbb{C})$, sometimes called the Frobenius norm or the Hilbert–Schmidt norm. In terms of the Hessian matrix we may rewrite (8) as $\|D^2 u\|_{L^2} = \|f\|_{L^2}$.

Proof. By the Differentiation Rule,

$$\widehat{D_j D_k u} = \frac{\xi_j \xi_k}{|\xi|^2} \hat{f}.$$

Since $\left| \frac{\xi_j \xi_k}{|\xi|^2} \right| \leqslant 1$ and $\hat{f} \in L^2(\mathbb{R}^n)$ by Plancherel, also $\widehat{D_j D_k u} \in L^2(\mathbb{R}^n)$ and thus $D_j D_k u \in L^2(\mathbb{R}^n)$, by applying Plancherel once again. Note that

$$|\xi|^4 = \left(\sum_{j=1}^{n} \xi_j^2 \right)^2 = \sum_{j,k=1}^{n} \xi_j^2 \xi_k^2$$

and therefore by Plancherel's Formula

$$\begin{aligned}
\int_{\mathbb{R}^n} |D^2 u|^2 \,\mathrm{d}x &= \sum_{j,k=1}^{n} \int_{\mathbb{R}^n} |D_j D_k u|^2 \,\mathrm{d}x \\
&= \sum_{j,k=1}^{n} \int_{\mathbb{R}^n} \left| \mathcal{F}^{-1}_{\xi\to x} \left(\frac{\xi_j \xi_k}{|\xi|^2} \hat{f} \right) \right|^2 \mathrm{d}x \\
&= \sum_{j,k=1}^{n} (2\pi)^{-n} \int_{\mathbb{R}^n} \left| \frac{\xi_j \xi_k}{|\xi|^2} \hat{f} \right|^2 \mathrm{d}\xi \\
&= (2\pi)^{-n} \int_{\mathbb{R}^n} \left(\frac{1}{|\xi|^4} \sum_{j,k=1}^{n} \xi_j^2 \xi_k^2 \right) |\hat{f}|^2 \,\mathrm{d}\xi \\
&= (2\pi)^{-n} \int_{\mathbb{R}^n} |\hat{f}|^2 \,\mathrm{d}\xi \\
&= \int_{\mathbb{R}^n} |f|^2 \,\mathrm{d}x.
\end{aligned}$$

□

Remark 15.4 (Calderon–Zygmund L^p estimate for the Laplacian). Let $p \in (1, \infty)$. Then there exists a constant $c = c(p, n)$ such that for $f \in L^p(\mathbb{R}^n)$ (with compact support, say)

$$\int_{\mathbb{R}^n} |D^2(E * f)|^p \,\mathrm{d}x \leqslant c \int_{\mathbb{R}^n} |f|^p \,\mathrm{d}x.$$

This fails for $p = 1$ and for $p = \infty$. The Sobolev Embedding Theorem then implies that $E * f \in W^{2,p}_{\mathrm{loc}}$ (see Theorem 16.1 for $p = 2$).

Proposition 15.5 (Localization)**.** *Let Ω be an open subset of $\mathbb{R}^n$ and suppose that $u \in \mathcal{D}'(\Omega)$ satisfies $\Delta u = f$ in $\mathcal{D}'(\Omega)$. If $f \in L^2_{\mathrm{loc}}(\Omega)$, then $u \in W^{2,2}_{\mathrm{loc}}(\Omega)$. Recall that*

$$W^{k,2}_{\mathrm{loc}}(\Omega) = \left\{ v \in L^2_{\mathrm{loc}}(\Omega) \,:\, D^\alpha v \in L^2_{\mathrm{loc}}(\Omega) \ \forall |\alpha| \leqslant k \right\}.$$

Proof. Fix $\Omega' \Subset \Omega$ and take a cut-of function $\theta \in \mathcal{D}(\Omega)$ between Ω' and $\partial\Omega$: $\mathbf{1}_{\Omega'} \leqslant \theta \leqslant \mathbf{1}_{\Omega}$. Now $\theta u \in \mathcal{S}'(\mathbb{R}^n)$ in view of the bound u must satisfy on the compact set $\operatorname{supp}\theta \subset \Omega$. Using Leibniz we calculate

$$\begin{aligned}
\Delta(\theta u) &= (\Delta\theta)u + 2\sum_{j=1}^{n} D_j\theta D_j u + \theta\Delta u \\
&= f_1 + f_2,
\end{aligned}$$

say, where $f_1 = (\Delta\theta)u + 2\nabla\theta\cdot\nabla u$ and $f_2 = \theta f$. Observe that $f_1, f_2 \in \mathcal{S}'(\mathbb{R}^n)$ both have compact supports contained in $\operatorname{supp}\theta$. By linearity we then have $\theta u = u_1 + u_2$, where $u_i \in \mathcal{S}'(\mathbb{R}^n)$

satisfies $\Delta u_i = f_i$ in $\mathcal{S}'(\mathbb{R}^n)$. Now $f_2 \in L^2(\mathbb{R}^n)$ and so we may apply Theorem 15.2 which, together with the Sobolev Embedding stated in Theorem 16.1, gives that $u_2 \in W^{2,2}_{\mathrm{loc}}(\mathbb{R}^n)$. For u_1 we observe that $u_1|_{\Omega'} \in \mathcal{D}'(\Omega')$ satisfies

$$\Delta(u_1|_{\Omega'}) = f_1|_{\Omega'} = 0$$

in $\mathcal{D}'(\Omega')$ since $\theta \equiv 1$ on Ω' so that $\Delta\theta = 0 = D_j\theta$ for all $1 \leqslant j \leqslant n$. But then Weyl's Lemma (Theorem 14.1) implies that $u_1|_{\Omega'}$ is C^∞ and harmonic, and so

$$u|_{\Omega'} = (\theta u)|_{\Omega'} = u_1|_{\Omega'} + u_2|_{\Omega'} \in W^{2,2}_{\mathrm{loc}}(\Omega').$$

Because $\Omega' \Subset \Omega$ was arbitrary, the proof is complete. □

Remark 15.6. The above can also be done for other elliptic PDEs. For instance for $P(D)u = f$ in $\mathcal{D}'(\Omega)$, where

$$P(D) = \sum_{|\alpha|=2} a_\alpha D^\alpha,$$

$a_\alpha \in \mathbb{R}$, and

$$P(\xi) = \sum_{|\alpha|=2} a_\alpha \xi^\alpha \neq 0$$

for all $\xi \in \mathbb{R}^n \setminus \{0\}$. For more along these lines, see the course C4.3 Functional Analytic Methods for PDEs.

16 Lecture 16

Theorem 16.1 (L^2 Sobolev Inequality)**.** *Let $u \in \mathcal{S}'(\mathbb{R}^n)$ and assume that for some $m \in \mathbb{N}$ $D^\alpha u \in L^2(\mathbb{R}^n)$ for all $\alpha \in \mathbb{N}_0^n$ with $|\alpha| = m$. Then*

1. *if $2m > n$, then $u \in C(\mathbb{R}^n)$ (u need not be bounded),*
2. *if $2m = n$, then $u \in L^p_{\mathrm{loc}}(\mathbb{R}^n)$ for all $p < \infty$ (but not in general for $p = \infty$),*
3. *if $2m < n$, then there exists a constant $c = c(m,n) > 0$ such that*

$$\inf_{k \in \mathbb{C}_{m-1}[x]} \left(\int_{\mathbb{R}^n} |u - k|^{\frac{2n}{n-2m}} \,\mathrm{d}x \right)^{\frac{n-2m}{2n}} \leqslant c \left(\int_{\mathbb{R}^n} \sum_{|\alpha|=m} |D^\alpha u|^2 \,\mathrm{d}x \right)^{\frac{1}{2}},$$

where $\mathbb{C}_{m-1}[x]$ denotes polynomials over $\mathbb{C}$ of degree at most $m-1$.

Proof. We omit the proof here. Instead we prove a slightly weaker variant of Theorem 16.1 (1). □

Theorem 16.2 (An L^2 Sobolev Embedding)**.** *If $m \in \mathbb{N}$, $m > \frac{n}{2}$, and $u \in W^{m,2}(\mathbb{R}^n)$, then $u \in C_0(\mathbb{R}^n)$*[1]*.*

Proof. Since $u \in \mathcal{S}'(\mathbb{R}^n)$ we can express $u \in W^{m,2}(\mathbb{R}^n)$ equivalently by use of Plancherel as

$$\widehat{D^\alpha u}(\xi) = (-i\xi)^\alpha \hat{u}(\xi) \in L^2(\mathbb{R}^n)$$

for $|\alpha| \leqslant m$. Indeed,

$$\sum_{|\alpha|\leqslant m} \int_{\mathbb{R}^n} |D^\alpha u|^2 \,\mathrm{d}x = \sum_{|\alpha|\leqslant m} (2\pi)^{-n} \int_{\mathbb{R}^n} |\xi^\alpha \hat{u}|^2 \,\mathrm{d}\xi = (2\pi)^{-n} \int_{\mathbb{R}^n} \left(\sum_{|\alpha|\leqslant m} |\xi^\alpha|^2 \right) |\hat{u}|^2 \,\mathrm{d}\xi.$$

Observe that for some positive constants $c_1 = c_1(m,n)$, $c_2 = c_2(m,n)$

$$c_1(1+|\xi|^2)^m \leqslant \sum_{|\alpha|\leqslant m} |\xi^\alpha|^2 \leqslant c_2(1+|\xi|^2)m$$

for all $\xi \in \mathbb{R}^n$. Thus $u \in W^{m,2}(\mathbb{R}^n)$ if and only if $(1+|\xi|^2)^{\frac{m}{2}}\hat{u} \in L^2(\mathbb{R}^n)$[2]. Now $g := (1+|\xi|^2)^{\frac{m}{2}}\hat{u} \in L^2(\mathbb{R}^n)$ and $2m > n$, so

$$|\hat{u}| = (1+|\xi|^2)^{-\frac{m}{2}}|g| \leqslant \frac{1}{2}(1+|\xi|^2)^{-m} + \frac{1}{2}|g|^2 \in L^1(\mathbb{R}^n)$$

and hence $u \in C_0(\mathbb{R}^n)$ by the Fourier Inversion Formula in $\mathcal{S}'(\mathbb{R}^n)$:

$$u = \mathcal{F}^{-1}(\hat{u}) = (2\pi)^{-n}\tilde{\mathcal{F}}(\hat{u}).$$

□

Theorem 16.3 (Qualitative Version of the Uncertainty Principle)**.** *If $u \in C_c(\mathbb{R}^n)$ and $\hat{u} \in C_c(\mathbb{R}^n)$, then $u = 0$.*

Proof. Put

$$f(z) = \int_{\mathbb{R}^n} u(x)\mathrm{e}^{-ix\cdot z} \,\mathrm{d}x$$

for $z \in \mathbb{C}^n$. This is well-defined because $u \in C_c(\mathbb{R}^n)$, and using the DCT we see that $u \in C^1(\mathbb{R}^n)$, with

$$\frac{\partial}{\partial \bar{z}_k} f(z) = \int_{\mathbb{R}^n} u(x) \frac{\partial}{\partial \bar{z}_k}(\mathrm{e}^{-ix\cdot z}) \,\mathrm{d}x = 0$$

for each $k = 1,\dots,n$. But then $z_k \mapsto f(z)$ is holomorphic (for fixed $z_1,\dots,z_{k-1},z_{k+1},\dots,z_n$). Now observe that for $\xi \in \mathbb{R}^n$ we have $f(\xi) = \hat{u}(\xi)$, and so if we fix $\xi_2,\dots,\xi_n \in \mathbb{R}$ and denote by $f_1(z_1) = f(z_1,\xi_2,\dots,\xi_n)$, then f_1 is entire and $f_1(\xi_1) = \hat{u}(\xi)$. Because $\hat{u}$ has compact support, it follows that f_1 must vanish on a half-line, so the Identity Theorem for holomorphic functions implies that $f_1 \equiv 0$. But then $\hat{u}(\xi) = 0$, and since $\xi_2,\dots,\xi_n \in \mathbb{R}$ were arbitrary, we have shown that $\hat{u} \equiv 0$. By the Fourier Inversion Formula in $\mathcal{S}'(\mathbb{R}^n)$ it follows that $u = 0$. □

[1] Recall that this means that there exists a representative of u belonging to $C_0(\mathbb{R}^n)$.

[2] Note that this gives us a way to define the Sobolev spaces $W^{m,2}$ in terms of the Fourier transform.