B4.3 Distribution Theory MT20

Lecture 11: Distributions whose first derivatives are regular distributions

1. Distributions whose first derivative is a regular distribution (one-dimensional case)
2. Absolute continuity and the fundamental theorem of calculus revisited
3. Distributions whose first order partial derivatives are regular distributions (higher dimensional case)
4. Definition of Sobolev functions

The material corresponds to pp. 51-55 in the lecture notes and should be covered in Week 6.

Recall from previous lectures that a regular distribution on an open non-empty subset Ω of $\mathbb{R}^n$ is a distribution

$$\mathscr{D}(\Omega) \ni \phi \mapsto \int_{\Omega} f\phi \,\mathrm{d}x$$

where $f \in \mathsf{L}^1_{\mathrm{loc}}(\Omega)$. Because this distribution uniquely determines $f \in \mathsf{L}^1_{\mathrm{loc}}(\Omega)$ (by the fundamental lemma of the calculus of variations) we identify the distribution with f and we use the same symbol for both interpretations. In fact, the symbol 'f' then stands for three different objects: *the distribution*, *the local* L^1 *function*, and *any of its representatives*.

What is intended follows from the given context or must be explicitly mentioned.

Main theme of lecture:

Let $u \in \mathscr{D}'(\Omega)$, where Ω is a non-empty open subset of $\mathbb{R}^n$. What can we say about those u for which the distributional partial derivatives

$$\partial_1 u, \ldots, \partial_n u$$

are all regular distributions on Ω?

We will see that the answer depends on the dimension n. In the proofs we assume that $\Omega = \mathbb{R}^n$, but the results remain true in the general case.

Recall from previous lectures:

- if $u \in \mathscr{D}'(\Omega)$ and $\nabla u \in \mathrm{C}(\Omega)$, then $u \in \mathrm{C}^1(\Omega)$.
- if $u \in \mathscr{D}'(a,b)$ and $u' \in \mathrm{L}^1_{\mathrm{loc}}(a,b)$, then for some constants $x_0 \in (a,b)$, $c \in \mathbb{C}$ we have

$$u(x) = c + \int_{x_0}^{x} u'(t)\,\mathrm{d}t \quad \text{a.e.}$$

- if $f \in \mathrm{L}^1_{\mathrm{loc}}(a,b)$, then the function

$$F(x) = \int_{x_0}^{x} f(t)\,\mathrm{d}t \quad (x_0 \in (a,b))$$

is continuous and its distributional derivative $F' = f$.

Definition: A function $u\colon (a,b) \to \mathbb{C}$ is **absolutely continuous** if there exist a function $f \in \mathrm{L}^1(a,b)$ and constants $x_0 \in (a,b)$, $c \in \mathbb{C}$ such that

$$u(x) = c + \int_{x_0}^{x} f(t)\,\mathrm{d}t$$

holds for all $x \in (a,b)$. It is **locally absolutely continuous** when only $f \in \mathrm{L}^1_{\mathrm{loc}}(a,b)$ above.

Corollaries:

- A function $f\colon (a,b) \to \mathbb{C}$ is locally absolutely continuous *iff* it is continuous and its distributional derivative $f' \in \mathrm{L}^1_{\mathrm{loc}}(a,b)$.
- A distribution $u \in \mathscr{D}'(a,b)$ has derivative $u' \in \mathrm{L}^1_{\mathrm{loc}}(a,b)$ *iff* u is a regular distribution with a locally absolutely continuous representative.
- If $f\colon (a,b) \to \mathbb{C}$ is locally absolutely continuous, then the distributional derivative $f' \in \mathrm{L}^1_{\mathrm{loc}}(a,b)$ and for all c, $d \in (a,b)$ we have

$$f(d) - f(c) = \int_{c}^{d} f'(t)\,\mathrm{d}t.$$

Weak derivatives

When $f\colon (a,b) \to \mathbb{C}$ is locally absolutely continuous its distributional derivative is often also called its weak derivative. More generally, $u \in \mathsf{L}^1_{\mathrm{loc}}(a,b)$ is said to have a weak derivative if its distributional derivative $u' \in \mathsf{L}^1_{\mathrm{loc}}(a,b)$.

Similar terminology is used in higher dimensions: $u \in \mathsf{L}^1_{\mathrm{loc}}(\Omega)$ has a weak partial derivative with respect to x_j if the distributional partial derivative $\partial_j u \in \mathsf{L}^1_{\mathrm{loc}}(\Omega)$. Note: This terminology is not universal and sometimes *weak derivative* is understood in a wider sense.

Example The function $f(x) = |x|$, $x \in \mathbb{R}$, has the weak derivative $f'(x) = x/|x|$.

The Heaviside function $H\colon \mathbb{R} \to \mathbb{R}$ has no weak derivative since $H' = \delta_0 \notin \mathsf{L}^1_{\mathrm{loc}}(\mathbb{R})$.

Example Assume $u \in \mathscr{D}'(\mathbb{R})$ and $u' \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R})$. Then u is a regular distribution with a locally absolutely continuous representative. We claim that

$$\frac{\tau_h u - u}{h} \to u' \text{ in } \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}) \text{ as } h \to 0,$$

that is, for each $a < b$,

$$\int_a^b \left| \frac{u(x+h) - u(x)}{h} - u'(x) \right| \mathrm{d}x \to 0 \text{ as } h \to 0.$$

Fix $a < b$ and assume that u is the locally absolutely continuous representative, so that for each $x \in (a,b)$ and $h \neq 0$,

$$\frac{u(x+h) - u(x)}{h} = \frac{1}{h} \int_x^{x+h} u'(t) \, \mathrm{d}t.$$

If u had been C^1, then the above difference quotient would converge locally uniformly in $x \in \mathbb{R}$ to $u'(x)$ as $h \to 0$, and the claim would in particular follow. In order to deal with the general case we use mollification.

Let $\left(\rho_\varepsilon\right)_{\varepsilon>0}$ be the standard mollifier on $\mathbb{R}$ and denote $u_\varepsilon = \rho_\varepsilon * u$. Then

$$\frac{(\tau_h - I)u}{h} = \frac{(\tau_h - I)u_\varepsilon}{h} + \frac{(\tau_h - I)(u - u_\varepsilon)}{h}$$

and so subtracting u', integrating over $x \in (a, b)$ and using the triangle inequality we get

$$\begin{aligned}
\int_a^b \left| \frac{(\tau_h - I)u(x)}{h} - u'(x) \right| \mathrm{d}x \quad &\leq \quad \int_a^b \left| \frac{(\tau_h - I)u_\varepsilon(x)}{h} - u_\varepsilon'(x) \right| \mathrm{d}x \\
&\qquad + \int_a^b \left| u_\varepsilon'(x) - u'(x) \right| \mathrm{d}x \\
&\qquad + \int_a^b \left| \frac{(\tau_h - I)(u - u_\varepsilon)(x)}{h} \right| \mathrm{d}x \\
&=: \quad I + II + III,
\end{aligned}$$

say.

Estimates for I and II: Let $\tau > 0$.

According to Proposition 2.7 in the lecture notes we can find $\varepsilon_\tau > 0$ so

$$\int_{a-1}^{b+1} \left| u' - \rho_\varepsilon * u' \right| \mathrm{d}x < \tau$$

for $\varepsilon \in (0, \varepsilon_\tau]$.

Since $u_\varepsilon' = \rho_\varepsilon * u'$ it follows that

$$II = \int_a^b \left| u_\varepsilon'(x) - u'(x) \right| \mathrm{d}x < \tau$$

for such ε and all $h \neq 0$.

Fix $\varepsilon = \varepsilon_\tau$. For this fixed ε we have since $u_\varepsilon \in \mathrm{C}^\infty(\mathbb{R})$ that for some $h_\tau \in (0, 1)$,

$$I = \int_a^b \left| \frac{(\tau_h - I)u_\varepsilon(x)}{h} - u_\varepsilon'(x) \right| \mathrm{d}x < \tau$$

holds for all $0 < |h| < h_\tau$.

In order to **estimate** *III* we use that for locally absolutely continuous v we have the fundamental theorem of calculus:

$$\frac{v(x+h)-v(x)}{h} = \frac{1}{h}\int_x^{x+h} v'(t)\,\mathrm{d}t$$

holds for all $x \in \mathbb{R}$ and $h \in \mathbb{R}\setminus\{0\}$, where v' is the distributional derivative. With $v = u - u_\varepsilon$ this yields:

$$\begin{aligned} III &= \int_a^b \left| \frac{1}{h}\int_x^{x+h} \bigl(u'(t) - u_\varepsilon'(t)\bigr)\,\mathrm{d}t \right| \mathrm{d}x \\ &\leq \int_a^b \frac{1}{|h|}\int_{x-|h|}^{x+|h|} \bigl|u'(t) - u_\varepsilon'(t)\bigr|\,\mathrm{d}t\,\mathrm{d}x \end{aligned}$$

We use Tonelli's theorem to swap the integration order:

$$III \leq \int_a^b \frac{1}{|h|} \int_{x-|h|}^{x+|h|} \left| u'(t) - u'_\varepsilon(t) \right| \mathrm{d}t \, \mathrm{d}x$$

$$= \int_{a-|h|}^{b+|h|} \int_a^b \frac{1}{|h|} \mathbf{1}_{(x-|h|,x+|h|)}(t) \left| u'(t) - u'_\varepsilon(t) \right| \mathrm{d}x \, \mathrm{d}t$$

$$= \int_{a-|h|}^{b+|h|} \int_a^b \frac{1}{|h|} \mathbf{1}_{(t-|h|,t+|h|)}(x) \left| u'(t) - u'_\varepsilon(t) \right| \mathrm{d}x \, \mathrm{d}t$$

$$\leq 2 \int_{a-1}^{b+1} \left| u'(t) - u'_\varepsilon(t) \right| \mathrm{d}t$$

since $0 < |h| < h_\tau$ and $h_\tau < 1$. By our choise of ε this is less than 2τ for all $0 < |h| < h_\tau$. This concludes the proof.

Remarks: It follows from the example that there exists a null sequence $h_j \to 0$ so

$$\frac{u(x + h_j) - u(x)}{h_j} \to u'(x) \text{ as } j \to \infty \tag{1}$$

pointwise in almost all $x \in \mathbb{R}$.

According to **Lebesgue's differentiation theorem** we have

$$\frac{1}{h} \int_x^{x+h} u'(t)\,\mathrm{d}t \to u'(x) \text{ as } h \to 0$$

pointwise in almost every $x \in \mathbb{R}$. It therefore follows that in fact (1) holds for the full limit $h \to 0$ pointwise outside a null set. Consequently, *a locally absolutely continuous function is differentiable almost everywhere in the usual sense and its almost everywhere defined usual derivative is a representative for its distributional derivative.*

What happens in higher dimensions?

Example Let

$$u(x,y) = \frac{1}{\sqrt{x^2+y^2}} \quad \text{for } (x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$$

Then $u \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^2)$ and one can show (as in Example 5.22 in lecture notes) that

$$\nabla u = -\frac{(x,y)}{\left(x^2+y^2\right)^{\frac{3}{2}}} \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^2)$$

Note that u *has no continuous representative*! This is different from the one-dimensional case where we saw that distributions whose first derivative was regular had a locally absolutely continuous representative. In the above example we started with a regular distribution–what if u is a distribution whose first order partial distributional derivatives $\partial_j u$ are all regular, then what can we say about u?

Theorem Let Ω be a non-empty open subset of $\mathbb{R}^n$ where we assume the dimension $n > 1$. Suppose $u \in \mathscr{D}'(\Omega)$ and that

$$\partial_j u \in \mathrm{L}^1_{\mathrm{loc}}(\Omega) \text{ for each } 1 \leq j \leq n.$$

Then $u \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$.

Remark It can be shown that $u \in \mathrm{L}^{\frac{n}{n-1}}_{\mathrm{loc}}(\Omega)$ and that u admits a representative (denoted again by) u whose restrictions to almost all lines parallel to the coordinate axes are locally absolutely continuous. The latter means that for each $1 \leq j \leq n$ and for $\mathscr{L}^{n-1}$ almost all $x' \in \mathbb{R}^{n-1} \times \{0\} \subset \mathbb{R}^n$ the function $t \mapsto u\big(x' + te_j\big)$ is locally absolutely continuous on $\{t \in \mathbb{R} : x' + te_j \in \Omega\}$. The partial derivatives $\partial u / \partial x_j$ therefore exist in the usual sense $\mathscr{L}^n$ almost everywhere and coincide with the distributional partial derivatives. Such functions are called local ACL functions (abbreviation for *absolutely continuous on lines*).
[The contents of this remark isn't examinable.]

Proof of theorem: We only give the proof in the case $\Omega = \mathbb{R}^n$ and use mollification. We also make use of the following result:

Fischer's Completeness Theorem: Let $p \in [1, \infty]$ and A be a measurable subset of $\mathbb{R}^n$. Then $\mathrm{L}^p(A)$ is a complete space.
(In the terminology of the course Functional Analysis 1 it is therefore a *Banach space*.)

Let $\left(\rho_\varepsilon\right)_{\varepsilon>0}$ be the standard mollifier on $\mathbb{R}^n$ and put $u_\varepsilon = \rho_\varepsilon * u$. Then $u_\varepsilon \in \mathrm{C}^\infty(\mathbb{R}^n)$ and

$$\nabla u_\varepsilon = \rho_\varepsilon * \nabla u.$$

Fix ε', $\varepsilon'' > 0$ and put $v = u_{\varepsilon'} - u_{\varepsilon''}$. Then $v \in \mathrm{C}^\infty(\mathbb{R}^n)$ and for x, $y \in \mathbb{R}^n$ the fundamental theorem of calculus yields

$$v(x) = v(y) + \int_0^1 \nabla v\big((1-t)y + tx\big) \cdot (x-y)\,\mathrm{d}t.$$

Multiply by $\rho(y)$ and integrate over $y \in \mathbb{R}^n$:

$$v(x) = \langle v, \rho\rangle + \int_{\mathbb{R}^n}\int_0^1 \nabla v\big((1-t)y + tx\big) \cdot (x-y)\,\mathrm{d}t\rho(y)\,\mathrm{d}y,$$

hence

$$\big|v(x)\big| \leq \big|\langle v, \rho\rangle\big| + \int_{\mathbb{R}^n}\int_0^1 \big|\nabla v\big((1-t)y + tx\big) \cdot (x-y)\big|\,\mathrm{d}t\rho(y)\,\mathrm{d}y.$$

Fix $R > 1$ and write $B_R = B_R(0)$. We now integrate over $x \in B_R$, estimate the right-hand side using Cauchy-Schwartz' inequality and that $\mathrm{supp}(\rho) = \overline{B_1} \subset B_R$:

$$\int_{B_R} |v(x)| \,\mathrm{d}x \le (I) + (II),$$

where

$$(I) := \big|\langle v, \rho\rangle\big| \mathscr{L}^n(B_R)$$

and

$$(II) := 2R \max \rho \int_{B_R}\int_{B_R}\int_0^1 \big|\nabla v\big((1-t)y + tx\big)\big| \,\mathrm{d}t \,\mathrm{d}y \,\mathrm{d}x.$$

Recall that $v = u_{\varepsilon'} - u_{\varepsilon''}$ with $u_\varepsilon = \rho_\varepsilon * u$, so

$$(I) \to 0 \text{ as } \varepsilon', \varepsilon'' \searrow 0.$$

We claim the same is true for (II).

Put $c = 2R \max \rho$. Rewrite (II) and use Tonelli's theorem to swap integration orders:

$$\begin{aligned} (II) &= c \int_0^{\frac{1}{2}} \int_{B_R} \int_{B_R} \big| \nabla v\big((1-t)y + tx\big)\big| \, \mathrm{d}y \, \mathrm{d}x \, \mathrm{d}t \\ &\quad + c \int_{\frac{1}{2}}^1 \int_{B_R} \int_{B_R} \big| \nabla v\big((1-t)y + tx\big)\big| \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}t \\ &=: cII_i + cII_{ii} \end{aligned}$$

In order to estimate the two multiple integrals II_i and II_{ii} on the right-hand side we use substitutions in the inner integrals. For II_i we substitute in the inner y-integral for each $t \in (0, \frac{1}{2})$ and $x \in B_R$:

$$\left\{ \begin{array}{l} y' = (1-t)y + tx \\ \mathrm{d}y = (1-t)^{-n} \mathrm{d}y' \leq 2^n \mathrm{d}y' \\ y' \in (1-t)B_R + tx \subset B_R. \end{array} \right.$$

Hereby

$$\begin{aligned} II_i &\leq 2^n \int_0^{\frac{1}{2}} \int_{B_R} \int_{B_R} \left|\nabla v(y')\right| \mathrm{d}y' \,\mathrm{d}x \,\mathrm{d}t \\ &= 2^{n-1} \mathscr{L}^n(B_R) \int_{B_R} \left|\nabla v(y')\right| \mathrm{d}y'. \end{aligned}$$

The estimate for II_{ii} is similar. We substitute in the inner x-integral for each $t \in (\frac{1}{2}, 1)$ and $y \in B_R$:

$$\begin{cases} x' = (1-t)y + tx \\ \mathrm{d}x = t^{-n}\mathrm{d}x' \leq 2^n \mathrm{d}x' \\ x' \in (1-t)y + tB_R \subset B_R, \end{cases}$$

whereby

$$II_{ii} \leq 2^{n-1} \mathscr{L}^n(B_R) \int_{B_R} \left|\nabla v(x')\right| \mathrm{d}x'.$$

We combine the obtained bounds and have in terms of the new constant $c_1 = 2^{n+1} R \mathscr{L}^n(B_R) \max \rho$:

$$(II) \leq c_1 \int_{B_R} \big| \nabla v(x) \big| \,\mathrm{d}x.$$

Recall that $v = u_{\varepsilon'} - u_{\varepsilon''}$ and $u_\varepsilon = \rho_\varepsilon * u$, so

$$\nabla v = \rho_{\varepsilon'} * \nabla u - \rho_{\varepsilon''} * \nabla u.$$

By virtue of Proposition 2.7 from the lecture notes (applied to each of $\mathbf{1}_{B_{R+1}} \partial_j u \in \mathrm{L}^1(\mathbb{R}^n)$) we have that $\rho_\varepsilon * \nabla u \to \nabla u$ in $\mathrm{L}^1(B_R)$ as $\varepsilon \searrow 0$, and therefore that

$$(II) \to 0 \text{ as } \varepsilon' \,,\, \varepsilon'' \searrow 0.$$

Consequently we have the Cauchy property:

$$\int_{B_R} \big| u_{\varepsilon'} - u_{\varepsilon''} \big| \,\mathrm{d}x \to 0 \text{ as } \varepsilon' \,,\, \varepsilon'' \searrow 0.$$

It follows by completeness of $\mathrm{L}^1(B_R)$ that there exists $w_R \in \mathrm{L}^1(B_R)$ so $u_\varepsilon \to w_R$ in $\mathrm{L}^1(B_R)$ as $\varepsilon \searrow 0$. Now this is true for any $R > 1$, so corresponding to any pair $1 < r < R$ we find $w_r \in \mathrm{L}^1(B_r)$, $w_R \in \mathrm{L}^1(B_R)$ with

$$\begin{cases} (\rho_\varepsilon * u)|_{B_r} \to w_r \text{ in } \mathrm{L}^1(B_r) & \text{as } \varepsilon \searrow 0 \\ (\rho_\varepsilon * u)|_{B_R} \to w_R \text{ in } \mathrm{L}^1(B_R) & \text{as } \varepsilon \searrow 0. \end{cases}$$

It follows that $w_r = w_R|_{B_r}$ almost everywhere, and so we may consistently define $w \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n)$ by $w|_{B_r} = w_r$ for $r > 1$. Because also $\rho_\varepsilon * u \to u$ in $\mathscr{D}'(\mathbb{R}^n)$ as $\varepsilon \searrow 0$ we conclude that

$$\langle u, \phi \rangle = \lim_{\varepsilon \searrow 0} \int_{\mathbb{R}^n} (\rho_\varepsilon * u)\phi \, \mathrm{d}x = \int_{\mathbb{R}^n} w\phi \, \mathrm{d}x$$

holds for all $\phi \in \mathscr{D}(\mathbb{R}^n)$ finishing the proof. □

Definition of Sobolev functions: (Sergei Lvovich Sobolev, 1908-1989)

Let Ω be a non-empty open subset of $\mathbb{R}^n$, $m \in \mathbb{N}$ and $p \in [1, \infty]$.

A $\mathrm{W}^{m,p}$ *Sobolev function* on Ω is any $u \in \mathrm{L}^p(\Omega)$ for which $\partial^\alpha u \in \mathrm{L}^p(\Omega)$ for each multi-index $\alpha \in \mathbb{N}_0^n$ of length $|\alpha| \leq m$.

The set of all $\mathrm{W}^{m,p}$ Sobolev functions on Ω is denoted by $\mathrm{W}^{m,p}(\Omega)$ and is called a *Sobolev space*. It is not difficult to see that it is a vector subspace of $\mathrm{L}^p(\Omega)$ and that it is normed by

$$\|u\|_{\mathrm{W}^{m,p}} = \begin{cases} \left(\sum_{|\alpha| \leq m} \|\partial^\alpha u\|_p^p\right)^{\frac{1}{p}} & \text{if } p < \infty \\ \max_{|\alpha| \leq m} \|\partial^\alpha u\|_\infty & \text{if } p = \infty. \end{cases}$$

It can be shown that $\mathrm{W}^{m,p}(\Omega)$ hereby is complete.