v: 2019-03-05

REAL ANALYSIS LECTURE NOTES

RASUL SHAFIKOV

8. Basic theory of distributions

8.1. Definition and examples of distributions.

Definition 8.1. *A linear continuous functional f on the space $\mathcal{D}(\Omega)$ is called a distribution. The linear space of all distributions is denoted by $\mathcal{D}'(\Omega)$.*

Recall that continuity here means that for every sequence (φ_j) of test-functions converging to φ in $\mathcal{D}(\Omega)$ we have $\lim_{j\to\infty} f(\varphi_j) = f(\varphi)$. By the linearity of f this is equivalent to the continuity at the zero vector: f is continuous if and only if $\lim_{j\to\infty} f(\varphi_j) = 0$ for every sequence $\varphi_j \longrightarrow 0$ in $\mathcal{D}(\Omega)$. We will often use the notation $f(\varphi) = \langle f, \varphi\rangle$ which has certain advantages. We now consider some examples.

Example 8.1. Recall that $L^1_{loc}(\Omega)$ the space of locally Lebesgue-integrable functions on Ω. Let f be in $L^1_{loc}(\Omega)$, i.e., $\int_X |f(x)|dx < \infty$ for every compact measurable subset $X \subset \Omega$. Then f defines a distribution $T_f \in \mathcal{D}'(\Omega)$ acting on a test-function $\varphi \in \mathcal{D}(\Omega)$ by

$$\langle T_f, \varphi\rangle = \int_\Omega f(x)\varphi(x)dx. \tag{1}$$

By the linearity of the integral, T_f is a linear functional on $\mathcal{D}(\Omega)$. The continuity of T_f follows from the definition of the topology on $\mathcal{D}(\Omega)$ and the Lebesgue Dominated Convergence Theorem. $\diamond$

Consider the map

$$L : L^1_{loc}(\Omega) \longrightarrow \mathcal{D}'(\Omega),$$

given by

$$L : f \mapsto T_f.$$

By linearity, the injectivity of L is equivalent to the fact that $L^{-1}(0) = \{0\}$. The latter is a consequence of the following.

Proposition 8.2. *$T_f = 0$ in $\mathcal{D}'(\Omega)$ if and only if $f \in L^1_{loc}(\Omega)$ vanishes almost everywhere in Ω.*

Proof. We only show that if $T_f = 0$ in $\mathcal{D}'(\Omega)$ then f vanishes almost everywhere in Ω; the converse is obvious. Let p be an arbitrary point of Ω; fix an $r > 0$ such that the closed ball $\overline{B}(p, r)$ is contained in Ω. Let η be a function of class $C^\infty(\mathbb{R}^n)$ such that $\eta = 1$ on $B(p, r/2)$ and $\operatorname{supp}\eta \subset B(p, r)$. Then $\eta f \in L^1(\mathbb{R}^n)$ and $T_{\eta f}$ vanishes in $\mathcal{D}'(\mathbb{R}^n)$. On the other hand, for every ε and every x the function $y \mapsto \eta(y)\omega_\varepsilon(x-y)$ is in $\mathcal{D}(\mathbb{R}^n)$ so $(\eta f)_\varepsilon(x) = \int_{\mathbb{R}^n} f(y)\eta(y)\omega_\varepsilon(x-y)dy = 0$ for every $x \in \mathbb{R}^n$. By (iii) of Proposition 7.6, $\| \eta f - (\eta f)_\varepsilon \|_{L^1} \longrightarrow 0$, $\varepsilon \longrightarrow 0$. Hence ηf represents 0 in $L^1(\mathbb{R}^n)$ and so vanishes almost everywhere. Therefore, f vanishes almost everywhere on $B(p, r/2)$. Since p is an arbitrary point, the general statement follows. $\square$

Thus, every "usual" function f of class $L^1_{loc}(\Omega)$ can be identified with a distribution T_f. In what follows we often drop the T and write $\langle f, \varphi\rangle$ instead of $\langle T_f, \varphi\rangle$ viewing usual functions as distributions. Such distributions (defined by (1)) are called *regular*. However, the class of distributions is much larger, and so the space of distributions $\mathcal{D}'(\Omega)$ is a far-reaching generalization of the notion of a function defined point-wise. Distributions which are not regular, are called *singular*. The following example confirms their existence.

Example 8.2. Consider the distribution $\delta(x) \in \mathcal{D}'(\mathbb{R}^n)$ (Dirac delta function) defined by

$$\langle \delta(x), \varphi \rangle = \varphi(0)$$

for $\varphi \in \mathcal{D}(\mathbb{R}^n)$. Suppose that there exists a function $f \in L^1_{loc}(\Omega)$ such that $\delta = T_f$. For every $\varepsilon > 0$ consider the function $\psi_\varepsilon \in \mathcal{D}(\Omega)$ defined by $\psi_\varepsilon(x) = e^{\frac{-\varepsilon^2}{\varepsilon^2 - |x|^2}}$ for $|x| < \varepsilon$ and $\psi_\varepsilon(x) = 0$ for $|x| \geq \varepsilon$. Then $\langle \delta, \psi_\varepsilon \rangle = \psi_\varepsilon(0) = e^{-1}$. On the other hand,

$$\langle T_f, \psi_\varepsilon \rangle = \int f(x)\psi_\varepsilon(x)dx,$$

and by the Lebesgue convergence theorem the last integral tends to 0 as $\varepsilon \to 0$: a contradiction. Thus, the δ-function is a singular distribution. Similarly, for every $a \in \mathbb{R}^n$ one can define the translated delta function δ_a:

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

◇

Example 8.3. Another interesting and typical example of a singular distribution is given as follows:

$$\langle \mathcal{P}\frac{1}{x}, \phi \rangle = v.p. \int_{\mathbb{R}} \frac{\phi(x)}{x} dx := \lim_{\varepsilon \longrightarrow 0+} \left(\int_{-\infty}^{-\varepsilon} \frac{\phi(x)}{x} dx + \int_{\varepsilon}^{+\infty} \frac{\phi(x)}{x} dx \right),$$

where v.p. stands for *valeur principale* in the sense of Cauchy of the integral. First we show that $\mathcal{P}\frac{1}{x}$ is well-defined. Indeed, let $\phi \in \mathcal{D}(\mathbb{R})$ be arbitrary with $\operatorname{supp} \phi \subset [-A, A]$ for some $A > 0$. Then, using the Mean Value Theorem for ϕ on the intervals $[-x, 0]$ and $[0, x]$, we obtain

$$\begin{aligned}
\left| \langle \mathcal{P}\frac{1}{x}, \phi \rangle \right| &= \lim_{\varepsilon \longrightarrow 0+} \left| \int_{-A}^{-\varepsilon} \frac{\phi(x)}{x} dx + \int_{\varepsilon}^{A} \frac{\phi(x)}{x} dx \right| \\
&= \lim_{\varepsilon \longrightarrow 0+} \left| \int_{-A}^{-\varepsilon} \frac{\phi(0) - x\phi'(\xi(x))}{x} dx + \int_{\varepsilon}^{A} \frac{\phi(0) + x\phi'(\tilde{\xi}(x))}{x} dx \right| \\
&\leq \int_{-A}^{A} |\phi'(\xi(x))| dx + \int_{-A}^{A} |\phi'(\tilde{\xi}(x))| dx \leq 2A \sup_{[-A,A]} |\phi'| < \infty.
\end{aligned}$$

Clearly, $\mathcal{P}\frac{1}{x}$ is linear. Let us now show that it is continuous on $\mathcal{D}(\mathbb{R})$. Consider a sequence (ϕ_j) converging to 0 in $\mathcal{D}(\mathbb{R})$. This means that there exists $A > 0$ such that $\phi_j(x) = 0$ for every j and every $|x| \geq A$. Then, applying the Mean Value Theorem, we have

$$\begin{aligned}
\left| \langle \mathcal{P}\frac{1}{x}, \phi_j \rangle \right| &= \left| v.p. \int_{\mathbb{R}} \frac{\phi_j(x)}{x} dx \right| = \left| v.p. \int_{-A}^{A} \frac{\phi_j(0) + x\phi_j'(\xi(x))}{x} dx \right| \\
&\leq \int_{-A}^{A} |\phi_j'(\xi(x))| dx \leq 2A \sup_{[-A,A]} |\phi_j'| \longrightarrow 0, j \longrightarrow 0.
\end{aligned}$$

8.2. **Convergence of distributions.** Now we define a topology on the space of distributions. For applications it is sufficient to use the standard notion of weak* convergence.

Definition 8.3. *A sequence of distributions (f_j) converges to a distribution f in $\mathcal{D}'(\Omega)$ if for every $\varphi \in \mathcal{D}(\Omega)$ one has $\lim_{j \longrightarrow \infty} \langle f_j, \varphi \rangle = \langle f, \varphi \rangle$.*

The following simple example of convergence is very important.

Proposition 8.4. *$\omega_\varepsilon \longrightarrow \delta$ in $\mathcal{D}'(\mathbb{R}^n)$ as $\varepsilon \longrightarrow 0+$.*

Proof. Given $\phi \in \mathcal{D}(\mathbb{R}^n)$ we need to show that

$$\lim_{\varepsilon \longrightarrow 0+} \int \omega_\varepsilon(x)\phi(x)dx = \phi(0).$$

For every $\tau > 0$ there exists an $\varepsilon_0 > 0$ such that $|\phi(x) - \phi(0)| < \tau$ when $|x| < \varepsilon_0$. Using the properties of the bump-function we obtain

$$\left| \int_{\mathbb{R}^n} \omega_\varepsilon(x)\phi(x)dx - \phi(0) \right| \leq \int_{|x| \leq \varepsilon} \omega_\varepsilon(x)|\phi(x) - \phi(0)|dx < \tau.$$

□

Another fundamental property of the space $\mathcal{D}'(\Omega)$ is its completeness.

Theorem 8.5. *Let (f_j) be a sequence in $\mathcal{D}'(\Omega)$ such that for every $\varphi \in \mathcal{D}(\Omega)$ the sequence $(\langle f_j, \varphi \rangle)$ converges in $\mathbb{R}$. Consider the map $f : \mathcal{D}(\Omega) \longrightarrow \mathbb{R}$ defined by*

$$\langle f, \varphi \rangle := \lim_{j \longrightarrow \infty} \langle f_j, \varphi \rangle, \;\; \varphi \in \mathcal{D}(\Omega).$$

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

Proof. The linearity of f is obvious so we just need to establish its continuity. Let $\varphi_k \longrightarrow 0$ as $k \longrightarrow \infty$ in $\mathcal{D}(\Omega)$. Arguing by contradiction suppose that $\langle f, \varphi_k \rangle$ does not converge to 0. Passing to a subsequence we may assume that there exists an $\varepsilon > 0$ such that $|\langle f, \varphi_k \rangle| \geq 2\varepsilon$ for all k. Since $\langle f, \varphi_k \rangle = \lim_{j \longrightarrow \infty} \langle f_j, \varphi_k \rangle$, for every k there exists j_k such that $|\langle f_{j_k}, \varphi_k \rangle| \geq \varepsilon$. However, this contradicts the following statement:

Lemma 8.6. *Let (f_k) be a sequence in $\mathcal{D}'(\Omega)$ satisfying assumptions of Theorem 8.5 and $\varphi_k \longrightarrow 0$ in $\mathcal{D}(\Omega)$. Then $\langle f_k, \varphi_k \rangle \longrightarrow 0$, as $k \longrightarrow \infty$.*

Thus, in order to complete the proof of the theorem it remains to prove the lemma.

Proof of Lemma 8.6. Suppose on the contrary that the statement of the lemma is false. Passing to a subsequence we may assume that $|\langle f_k, \varphi_k \rangle| \geq C > 0$. Since $\varphi_k \longrightarrow 0$ in $\mathcal{D}(\Omega)$, we have:

(a) $\varphi_k = 0$ for all k outside a compact subset $K \subset \Omega$.

(b) For every α the sequence $D^\alpha \varphi_k$ converges uniformly to 0.

Passing to a subsequence we can assume that any $k = 0, 1, 2, \ldots,$

$$|D^\alpha \varphi_k(x)| \leq 1/4^k, \;\; |\alpha| \leq k.$$

Set $\psi_k = 2^k \varphi_k$; then

$$|D^\alpha \psi_k(x)| \leq 1/2^k, |\alpha| \leq k. \tag{2}$$

Furthermore, $\psi_k \longrightarrow 0$ in $\mathcal{D}(\mathbb{R}^n)$ and every series of type $\sum_s \psi_{k_s}(x)$ converges in $\mathcal{D}(\Omega)$. On the other hand,

$$|\langle f_k, \psi_k \rangle| = 2^k |\langle f_k, \varphi_k \rangle| \geq 2^k C \longrightarrow \infty \;\text{ as } k \longrightarrow \infty. \tag{3}$$

To reach a contradiction, we construct by induction suitable subsequences (f_{k_s}) and (ψ_{k_s}) that satisfy inequalities (7) and (8) below. Choose f_{k_1} and ψ_{k_1} such that $|\langle f_{k_1}, \psi_{k_1} \rangle| \geq 2$. This is always possible in view of (3). Suppose that f_{k_j}, ψ_{k_j}, $j = 1, ..., s-1$, are already constructed. We wish to find f_{k_s}, ψ_{k_s}. Since $\psi_k \longrightarrow 0$, $k \longrightarrow \infty$ in $\mathcal{D}(\Omega)$, we have $\lim_{k \to \infty} \langle f_{k_j}, \psi_k \rangle \longrightarrow 0$, for any $j = 1, ..., s-1$, and so there exists N such that for $k \geq N$

$$|\langle f_{k_j}, \psi_k \rangle| \leq 1/2^{s-j}, \;\; j = 1, ..., s-1. \tag{4}$$

Moreover, since

$$\lim_{k\to\infty}\langle f_k,\psi_{k_j}\rangle=\langle f,\psi_{k_j}\rangle \quad j=1,...,s-1,$$

there exists $N_1\geq N$ such that for all $k\geq N_1$

$$|\langle f_k,\psi_{k_j}\rangle|\leq|\langle f,\psi_{k_j}\rangle|+1,\quad j=1,...,s-1. \tag{5}$$

Finally, in view of (3) we fix $k_s\geq N_1$ such that

$$|\langle f_{k_s},\psi_{k_s}\rangle|\geq\sum_{j=1}^{s-1}|\langle f,\psi_{k_j}\rangle|+2s. \tag{6}$$

Now it follows from (4), (5), (6) that the functions f_{k_s} and ψ_{k_s} satisfy

$$|\langle f_{k_j},\psi_{k_s}\rangle|\leq 1/2^{s-j}, j=1,...,s-1, \tag{7}$$

$$|\langle f_{k_s},\psi_{k_s}\rangle|\geq\sum_{j=1}^{s-1}|\langle f_{k_s},\psi_{k_j}\rangle|+s+1. \tag{8}$$

This gives the inductive construction of the required subsequences. Set

$$\psi(x)=\sum_{s=1}^{\infty}\psi_{k_s}(x).$$

By (2) this series converges in $\mathcal{D}(\Omega)$. Its sum $\psi\in\mathcal{D}(\Omega)$ satisfies

$$\langle f_{k_s},\psi\rangle=\langle f_{k_s},\psi_{k_s}\rangle+\sum_{j=1,j\neq s}^{\infty}\langle f_{k_s},\psi_{k_j}\rangle.$$

Therefore, keeping in mind (7), (8) we obtain

$$\langle f_{k_s},\psi\rangle\geq|\langle f_{k_s},\psi_{k_s}\rangle|-\sum_{j=1}^{s-1}|\langle f_{k_s},\psi_{k_j}\rangle|-\sum_{j=s+1}^{\infty}|\langle f_{k_s},\psi_{k_j}\rangle|$$

$$\geq s+1-\sum_{j=s+1}^{\infty}1/2^{j-s}=s,$$

that is, $\langle f_{k_s},\psi\rangle\longrightarrow\infty$ as $s\longrightarrow\infty$. This contradicts the condition $\langle f_k,\psi\rangle\longrightarrow\langle f,\psi\rangle$, which completes the proof. □

8.3. **Multiplication of distributions.** The product of two functions of class $L^1_{loc}(\mathbb{R})$ in general is not in this class (consider, for instance, $f(x)=|x|^{-1/2}$ and f^2). This example shows that it is impossible in general to define in a natural way even the product of regular distributions. In fact, one can show that it is impossible to define a multiplication of two distributions which satisfies the standard algebraic properties (commutativity, associativity,...). However, one can define the product of a distribution $f\in\mathcal{D}'(\Omega)$ and a function $a\in C^\infty(\Omega)$.

First, consider the case when $f\in L^1_{loc}(\Omega)$, i.e., f is a regular distribution. Then the distribution corresponding to the usual product af acts on a test-function φ by

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

In the case of an arbitrary distribution f we take the right-hand side of this equality to be the definition of the distribution af, i.e., we *set*

$$\langle af, \varphi \rangle := \langle f, a\varphi \rangle, \quad \varphi \in \mathcal{D}(\Omega).$$

Observe two immediate properties of the algebraic operation of multiplication of a distribution by a smooth function a:

(a) Linearity: for every $f, g \in \mathcal{D}'(\Omega)$ and real λ, μ we have

$$a(\lambda f + \mu g) = \lambda(af) + \mu(ag).$$

(b) Continuity: if $f_j \longrightarrow f$ in $\mathcal{D}'(\Omega)$ then $af_j \longrightarrow af$ in $\mathcal{D}'(\Omega)$.

Example 8.4. $a(x)\delta(x) = a(0)\delta(x)$, since

$$\langle a\delta, \phi \rangle = \langle \delta, a\phi \rangle = a(0)\phi(0) = a(0)\langle \delta, \phi \rangle.$$

$\diamond$

Example 8.5. $x\mathcal{P}\frac{1}{x} = 1$. Indeed, for any $\phi \in \mathcal{D}(\Omega)$, we have

$$\langle x\mathcal{P}\frac{1}{x}, \phi \rangle = \langle \mathcal{P}\frac{1}{x}, x\phi \rangle = v.p. \int_{\mathbb{R}} \frac{x\phi(x)}{x} dx = \int_{\mathbb{R}} \phi(x) dx = \langle 1, \phi \rangle.$$

$\diamond$

8.4. **Composition with linear maps.** Let f be a function of class $L^1_{loc}(\mathbb{R}^n)$ and let $u : x \mapsto Ax+b$ be a bijective affine map of $\mathbb{R}^n$, i.e., $\det A \neq 0$. Given $\phi \in \mathcal{D}(\Omega)$ consider

$$\begin{aligned}\langle T_{f\circ u}, \phi \rangle &= \int f(Ay+b)\phi(y)dy = |\det A|^{-1} \int f(x)\phi(A^{-1}(x-b))dx \\ &= |\det A|^{-1}\langle T_f, \phi(A^{-1}(x-b))\rangle.\end{aligned}$$

For an arbitrary $f \in \mathcal{D}'(\Omega)$ we take the last equality as a definition of the distribution $f \circ u = f(Ay+b)$, that is,

$$\langle f(Ay+b), \phi \rangle := |\det A|^{-1}\langle f, \phi(A^{-1}(x-b))\rangle.$$

The distribution $f(x+b)$ is called the *translation* of a distribution f by a vector b. In particular,

$$\langle \delta(y-a), \phi \rangle = \langle \delta, \phi(x+a) \rangle = \phi(a).$$

Recall that we also denoted above this distribution by δ_a.

8.5. **Dependence on a parameter.** The continuity of distributions implies their "good" behaviour under an action on test-functions depending on a real parameter. We will use this property and its variations.

Theorem 8.7. *Let X and Y be domains in $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively and $\varphi \in C^\infty(X \times Y)$. Suppose that there exists a compact subset $K \subset X$ such that $\varphi(x,y) = 0$ for every (x,y) with $x \notin K$. Then for every $f \in \mathcal{D}'(X)$ the function*

$$F : Y \ni y \mapsto \langle f(x), \varphi(x,y) \rangle$$

is of class $C^\infty(Y)$ and

$$D^\alpha_y \langle f(x), \varphi(x,y) \rangle = \langle f(x), D^\alpha_y \varphi(x,y) \rangle.$$

Proof. (a) Let us show that F is a continuous function. Let $y^k \in \mathbb{R}^m$ be a sequence converging to $y \in Y$. We can assume that the points y^k are in a fixed closed ball $B \subset Y$. Then

$$\| D_x^\beta \varphi(x, y^k) - D_x^\beta \varphi(x, y) \|_{C(X)} \leq \| \nabla D_x^\beta \varphi(x, y) \|_{C(K \times B)} |y^k - y|.$$

Since the supports of all functions $x \mapsto D_x^\beta \varphi(x, y^k)$ are contained in K, the sequence $\varphi(x, y^k)$ converges to $\varphi(x, y)$ as $k \longrightarrow \infty$ in $\mathcal{D}(X)$ and $F(y^k) \longrightarrow F(y), k \longrightarrow \infty$, by continuity of f.

(b) Next we study the partial derivatives of F. For the element $\vec{e}_j$, $j = 1, ..., m$, of the standard basis of $\mathbb{R}^m$, and every fixed $y \in Y$ we have

$$\frac{\varphi(x, y + t\vec{e}_j) - \varphi(x, y)}{t} \longrightarrow \frac{\partial}{\partial y_j}(\varphi(x, y)),\ t \longrightarrow 0,$$

in $\mathcal{D}(X)$. Therefore,

$$\frac{1}{t}(F(y + t\vec{e}_j) - F(y)) = \langle f(x), \frac{1}{t}(\varphi(x, y + t\vec{e}_j) - \varphi(x, y))\rangle \longrightarrow \langle f(x), \frac{\partial}{\partial y_j}\varphi(x, y)\rangle.$$

Hence, the partial derivative of F in y_j exists and

$$\frac{\partial}{\partial y_j}\langle f(x), \varphi(x, y)\rangle = \langle f(x), \frac{\partial}{\partial y_j}\varphi(x, y)\rangle.$$

Part (a) shows that the partial derivative $\frac{\partial}{\partial y_j}F$ is continuous. Proceeding by induction, we obtain that $F \in C^\infty(Y)$ and satisfies the derivation rule stated in the theorem. $\square$