v: 2019-04-01

REAL ANALYSIS LECTURE NOTES

RASUL SHAFIKOV

10. Structure theorems and convolution of distributions

10.1. **Structure theorems.** We introduce the following property ($\star$).

Definition 10.1. *A linear functional $f : \mathcal{D}(\Omega) \to \mathbb{R}$ satisfies condition ($\star$) if for every compact subset K in Ω there exists $C = C(K) > 0$ and a positive integer $k = k(K)$ such that*

$$|\langle f, \phi\rangle| \le C||\phi||_{C^k(K)},\ \forall \phi \in \mathcal{D}(\Omega),\ \operatorname{supp}\phi \subset K. \tag{1}$$

We have the following characterization of distributions.

Theorem 10.2. *A linear functional f on the space $\mathcal{D}(\Omega)$ is a distribution if and only if it satisfies condition ($\star$).*

Proof. If f satisfies ($\star$) then f is clearly continuous, and so $f \in \mathcal{D}'(\Omega)$. To prove the converse, assume that $f \in \mathcal{D}'(\Omega)$. Arguing by contradiction, suppose that f does not satisfy ($\star$). Then there exists a compact K in Ω such that for every C and k the inequality (1) fails for some $\phi \in \mathcal{D}(\Omega)$ with $\operatorname{supp}\phi \subset K$. In particular, we can set $C = k = j$ and take a function $\phi_j \in \mathcal{D}(\Omega)$ with $\operatorname{supp}\phi \subset K$ such that

$$|\langle f, \phi_j\rangle| > j \parallel \phi_j \parallel_{C^j(K)}, \quad j = 0, 1, 2, \dots. \tag{2}$$

By linearity of expressions on both sides, this inequality still holds if we replace ϕ_j by the function $\psi_j = \frac{\phi_j}{\langle f,\phi_j\rangle}$. Then

$$1/j > \parallel \psi_j \parallel_{C^j(K)}, \quad j = 0, 1, 2, \dots.$$

Fix a positive integer k. Then for $j \ge k$ we have

$$\parallel \psi_j \parallel_{C^k(K)} \le \parallel \psi_j \parallel_{C^j(K)} < j^{-1} \to 0, \quad \text{as} \quad j \to \infty.$$

Therefore, the sequence (ψ_j) converges to 0 in $\mathcal{D}(\Omega)$ but $\langle f, \psi_j\rangle = 1$. This contradiction proves the theorem. □

Let Ω be a domain in $\mathbb{R}^n$, $f \in \mathcal{D}'(\Omega)$ and $k \ge 0$ be an integer. We say that a distribution f has *the order of singularity* $\le k$ if there exists a constant $C = C(\Omega, f) > 0$ such that for every $\varphi \in \mathcal{D}(\Omega)$ we have

$$|\langle f, \varphi\rangle| \le C \parallel \varphi \parallel_{C^k(\Omega)}.$$

Thus, f satisfies condition (1) with the same k for every compact K in Ω, i.e., k can be chosen independently of K. We say that the order of singularity of f *is equal to* k if this estimate does not hold for some $k' < k$.

Example 10.1. If T_f is a regular distribution defined by a function $f \in L^1(\Omega)$. Then its order of singularity is 0. ⋄

Example 10.2. The order of singularity of $\delta^{(k)}(x)$ is equal to k. ⋄

The following property of distributions is often used.

Theorem 10.3. *Let Ω' be a domain in $\mathbb{R}^n$ and Ω be a bounded subdomain such that $\overline{\Omega} \subset \Omega'$. Then for every distribution $f \in \mathcal{D}'(\Omega')$ its restriction to Ω is a distribution of finite order of singularity.*

Thus, the theorem claims that there exist an integer $k \geq 0$ (depending on f and Ω) and a constant $C = C(\Omega, f) > 0$ such that for every $\varphi \in \mathcal{D}(\Omega)$ we have

$$|\langle f, \varphi \rangle| \leq C \parallel \varphi \parallel_{C^k(\Omega)}$$

The proof is similar to the previous one.

Proof. Arguing by contradiction, suppose that there exists a sequence $\varphi_m \in \mathcal{D}(\Omega)$ such that

$$|\langle f, \varphi_m \rangle| > m \parallel \varphi_m \parallel_{C^m(\Omega)}$$

for every $m = 1, 2, \ldots$. Set $\psi_m = \alpha_m \varphi_m$, where α_m is a real number. Then by linearity we still have

$$|\langle f, \psi_m \rangle| > m \parallel \psi_m \parallel_{C^m(\Omega)} .$$

Let $\alpha_m = (\parallel \varphi_m \parallel_{C^m(\Omega)})^{-1}/m$. Then

$$|\langle f, \psi_m \rangle| > m \parallel \psi_m \parallel_{C^m(\Omega)} = 1. \tag{3}$$

On the other hand, $\parallel \psi_m \parallel_{C^m(\Omega)} = 1/m$ for every m. Then for every β, such that $|\beta| \leq m$, we have

$$\parallel D^\beta \psi_m \parallel_{C(\Omega)} \leq 1/m.$$

Thus, the sequence (ψ_m) converges to 0 together with all partial derivatives of all orders and the supports of ψ^m are contained in the compact $\overline{\Omega}$ in Ω'. Then $\psi^m \longrightarrow 0$ in $\mathcal{D}(\Omega')$ and $\langle f, \psi_m \rangle \longrightarrow 0$. This contradicts (3). □

The following is a consequence of Theorem 10.3.

Proposition 10.4. *Let $f \in \mathcal{D}'(\mathbb{R}^n)$ satisfy* $\operatorname{supp} f = \{0\}$. *Then there exist an integer $k \geq 0$ and constants C_α such that*

$$f = \sum_{|\alpha| \leq k} C_\alpha D^\alpha \delta(x).$$

Proof. Let a function $\eta \in \mathcal{D}(\mathbb{R}^n)$ be equal to 1 in a neighbourhood of 0 and vanishes outside $B(0,1) = \{|x| < 1\}$. Consider a function $\varphi \in \mathcal{D}(\mathbb{R}^n)$. Let Ω be a domain in $\mathbb{R}^n$ containing $\operatorname{supp} \varphi \cup B(0,2)$. Applying Theorem 10.3 to f in Ω we conclude that there exist an integer $k \geq 0$ (depending on Ω) and a constant $C = C(\Omega, f) > 0$ such that for every $\phi \in \mathcal{D}(\Omega)$ we have

$$|\langle f, \phi \rangle| \leq C \parallel \phi \parallel_{C^k(\Omega)} . \tag{4}$$

Set $h(x) = \phi(x) - \sum_{|\alpha| \leq k} (D^\alpha \phi(0) x^\alpha)/\alpha!$ and

$$\psi_s(x) = h(x)\eta(sx).$$

Then for every integer $s \geq 1$ we have

$$\langle f(x), \psi_1 \rangle = \langle f, \psi_s \rangle. \tag{5}$$

Indeed, $\langle f, \psi_1 \rangle - \langle f, \psi_s \rangle = \langle f, (\psi_1 - \psi_s)h \rangle = 0$ since $(\psi_1 - \psi_s)h = 0$ in a neighborhood of 0 and $\operatorname{supp} f = \{0\}$. Since $\operatorname{supp} \psi_s \subset \Omega$ for every $s \geq 1$, we obtain that $\psi_s \in \mathcal{D}(\Omega)$ and by (4),

$$|\langle f, \psi_s \rangle| \leq C \parallel \psi_s \parallel_{C^k(\Omega)}, \quad s \geq 1.$$

It follows easily from the definition of ψ_s that $\| \psi_s \|_{C^k(\Omega)} \longrightarrow 0$ as $s \longrightarrow \infty$. But then (5) implies that $\langle f(x), \psi_1 \rangle = 0$. Therefore,

$$\langle f, \varphi\eta \rangle = \sum_{|\alpha|\le k} (\langle f, x^\alpha \eta \rangle / \alpha!) D^\alpha \varphi(0) = \sum_{|\alpha|\le k} C_\alpha \langle D^\alpha \delta(x), \varphi \rangle,$$

where $C_\alpha = (\langle f, x^\alpha \eta \rangle / \alpha!)$ are independent of ϕ. □

Example 10.3. Let a function $f \in L^1_{loc}(\mathbb{R}^n \setminus \{0\})$ satisfy the following condition: there exists a constant $C > 0$ and an integer $m > 0$ such that

$$|f(x)| \le \frac{C}{|x|^m}, \quad \forall x \in \{x \in \mathbb{R}^n : |x| \le 1\}. \tag{6}$$

We will show that f admits an extension past the origin as a distribution, i.e., there exists $\tilde{f} \in \mathcal{D}'(\mathbb{R}^n)$ such that

$$\langle \tilde{f}, \varphi \rangle = \int_{\mathbb{R}^n} f(x)\varphi(x)dx$$

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

First of all let us recall the general Taylor formula: let $a \in \mathbb{R}^n$ and let ψ be a smooth function $\in C^\infty$ in a neighbourhood of a. Then for every integer $k \ge 0$ there exists a neighbourhood U of a such that for $x \in U$ we have

$$\psi(x) = \sum_{0\le|\alpha|\le k} \frac{1}{\alpha!} D^\alpha \psi(a)(x-a)^\alpha + \int_0^1 (1-t)^k \sum_{|\alpha|=k+1} \frac{k+1}{\alpha!} D^\alpha \psi(tx + (1-t)a)(x-a)^\alpha dt.$$

As usual we use here the notation $\alpha! = \alpha_1!...\alpha_n!$ and $x^\alpha = x_1^{\alpha_1}...x_n^{\alpha_n}$. We define the distribution $\tilde{f}$ by the formula

$$\langle \tilde{f}, \varphi \rangle = I_1 + I_2,$$

where

$$I_1 = \int_{|x|\ge 1} f(x)\varphi(x)dx,$$

$$I_2 = \int_{|x|\le 1} f(x) \left(\varphi(x) - \sum_{|\alpha|\le m-1} \frac{1}{\alpha!} D^\alpha \varphi(0) x^\alpha \right) dx.$$

Using the Taylor formula and condition (6) we obtain

$$|I_2| \le C \sum_{|\alpha|=m} \sup_{\mathbb{R}} |D^\alpha \varphi|$$

Using the condition $\operatorname{supp} \varphi \subset \{x : |x| \le M\}$ we also obtain

$$|I_1| \le \int_{1\le|x|\le M} |f(x)\varphi(x)|dx \le C' \sup_{\mathbb{R}^n} |\varphi|.$$

From this and Theorem 10.2 we conclude that $\tilde{f}$ is a well-defined distribution in $\mathcal{D}'(\mathbb{R}^n)$. ⋄

Finally, consider an example of a distribution of infinite order.

Example 10.4. Consider a linear functional on $\mathcal{D}(\mathbb{R})$ defined by

$$\langle f, \varphi \rangle = \sum_{n=0}^{\infty} \varphi^{(n)}(n).$$

It follows by Theorem 8.5 that $f \in \mathcal{D}'(\mathbb{R})$. We leave to the reader to prove that f is not of a finite order, arguing by contradiction. $\diamond$

10.2. **Regularization and convolution with test-functions.** Recall that a convolution $f * g$ of two functions $L^2(\mathbb{R}^n)$ is defined by

$$f * g(x) = \int_{\mathbb{R}^n} f(y)g(x-y)dy = \int_{\mathbb{R}^n} f(x-y)g(y)dy. \tag{7}$$

This makes natural the following general definition.

Definition 10.5. *A convolution of a distribution $f \in \mathcal{D}'(\mathbb{R}^n)$ and a test function $\varphi \in \mathcal{D}(\mathbb{R}^n)$ is defined by*

$$f * \varphi(x) = \langle f(y), \varphi(x-y) \rangle. \tag{8}$$

Note the following: for every $x \in \mathbb{R}^n$ the function $y \mapsto \varphi(x-y)$ is a test-function; on the right-hand side of (8) we apply distribution f to this function which is stressed by the notation $f(y)$. Thus, $f * \varphi$ is defined as a usual function on $\mathbb{R}^n$.

Proposition 10.6. *We have $f * \phi \in C^\infty(\mathbb{R}^n)$ and*

$$D^\alpha(f * \phi) = f * D^\alpha \phi = (D^\alpha f) * \phi. \tag{9}$$

Proof. The regularity of $f * \varphi$ and the first equality of (9) follow from Theorem 8.7. Let us prove the second equality in (9). We have

$$\left((\frac{\partial}{\partial x_j} f) * \varphi \right)(x) = \langle \frac{\partial}{\partial y_j} f(y), \varphi(x-y) \rangle = -\langle f(y), \frac{\partial}{\partial y_j}(\varphi(x-y)) \rangle$$
$$= \langle f(y), \left(\frac{\partial}{\partial x_j} \varphi \right)(x-y)) \rangle = f * \frac{\partial}{\partial x_j} \varphi.$$

The rest of the proof is done by induction. $\square$

A very important special case arises if we take the bump function ω_ε as φ in the definition of convolution. This leads to

Definition 10.7. *The convolution $f_\varepsilon := f * \omega_\varepsilon$ is called the regularization of a distribution f.*

Proposition 10.8. *We have*

(i) $f_\varepsilon \in C^\infty(\mathbb{R}^n)$.
(ii) $(D^\alpha f)_\varepsilon = D^\alpha(f_\varepsilon)$.
(iii) *If $f \in C(\mathbb{R}^n)$ then $f_\varepsilon \longrightarrow f, \varepsilon \longrightarrow 0+$ in $C(\Omega)$ for every bounded subset Ω of $\mathbb{R}^n$.*
(iv) *If $\varphi \in \mathcal{D}(\mathbb{R}^n)$ then $\varphi_\varepsilon \in \mathcal{D}(\mathbb{R}^n)$ and $\varphi_\varepsilon \longrightarrow \varphi$ in $\mathcal{D}(\mathbb{R}^n)$ as $\varepsilon \longrightarrow 0+$.*
(v) *if $f \in \mathcal{D}'(\mathbb{R}^n)$ then $f_\varepsilon \longrightarrow f$ in $\mathcal{D}'(\mathbb{R}^n)$ as $\varepsilon \longrightarrow 0+$.*

Proof. Parts (i) and (ii) follow from Proposition 10.6. Part (iii) is established in Proposition 7.6, so it remains to show (iv) and (v). If $\varphi \in \mathcal{D}(\mathbb{R}^n)$, then its support is compact, say, $\varphi(x) = 0$ when $|x| \geq A$ for some $A > 0$. Then the formula for the regularization of functions shows that $\varphi_\varepsilon(x) = 0$ for $|x| \geq A + \varepsilon$ so $\operatorname{supp} \varphi_\varepsilon \subset K = \{x \in \mathbb{R}^n : |x| \leq A + 1\}$ for $\varepsilon < 1$. It follows now from (iii) and

(ii) that φ_ε converges to φ as $\varepsilon \longrightarrow 0+$ uniformly on K together with all partial derivatives of any order. Hence $\varphi_\varepsilon \longrightarrow \varphi, \varepsilon \longrightarrow 0+$ in $\mathcal{D}(\mathbb{R}^n)$ and we obtain (iv).

To prove (v), we view $f_\varepsilon = \langle f(y), \omega_\varepsilon(x-y)\rangle$ as a distribution acting on every $\psi \in \mathcal{D}(\mathbb{R}^n)$ by

$$\langle f_\varepsilon, \psi\rangle = \int_{\mathbb{R}^n} \langle f(y), \omega_\varepsilon(x-y)\rangle \psi(x)dx.$$

It follows from Theorem 8.7 that

$$\int_{\mathbb{R}^n} \langle f(y), \omega_\varepsilon(x-y)\rangle \psi(x)dx = \langle f(y), \int_{\mathbb{R}^n} \omega_\varepsilon(x-y)\psi(x)dx\rangle = \langle f, \psi_\varepsilon\rangle. \tag{10}$$

To see this we consider for simplicity of notation the case $n = 1$. Let

$$F(t) = \int_{-\infty}^{t} \langle f(y), \omega_\varepsilon(x-y)\rangle \psi(x)dx,$$

and

$$G(t) = \langle f(y), \int_{-\infty}^{t} \omega_\varepsilon(x-y)\psi(x)dx\rangle.$$

Then by Theorem 8.7, $F'(x) = G'(x) = \langle f(y), \omega_\varepsilon(x-y)\rangle\psi(x)$. Since $F(-\infty) = G(-\infty) = 0$, we have $F = G$ for all t and we pass to the limit as $t \to +\infty$. This proves (10). Then by (iv),

$$\langle f_\varepsilon, \psi\rangle = \langle f, \psi_\varepsilon\rangle \longrightarrow \langle f, \psi\rangle \quad \text{as} \quad \varepsilon \longrightarrow 0+.$$

This concludes the proof. □

As an application we give another proof of Corollary 9.4: if $f \in \mathcal{D}'(\mathbb{R}^n)$ are such that $\frac{\partial}{\partial x_j} f = 0$ in $\mathcal{D}'(\mathbb{R}^n)$ for $j = 1, ..., n$, then $f = const$.

Proof of Corollary 9.4. For every $\varepsilon > 0$ we have $0 = (\frac{\partial}{\partial x_j} f)_\varepsilon = \frac{\partial}{\partial x_j}(f_\varepsilon)$. Since f_ε is a usual function of class C^∞, we conclude that $f_\varepsilon = C(\varepsilon)$. Then,

$$\langle f, \varphi\rangle = \lim_{\varepsilon \longrightarrow 0+} \langle f_\varepsilon, \varphi\rangle = \lim_{\varepsilon \longrightarrow 0+} C(\varepsilon) \int \varphi(x)dx \quad \text{for any} \quad \varphi \in \mathcal{D}(\mathbb{R}^n).$$

In particular, set $\varphi = \omega(x)$ so that $\int \omega(x)dx = 1$. We obtain that $C = \lim_{\varepsilon \to 0+} C(\varepsilon)$ exists, and $f = C$. □

10.3. **Convolution of distributions.** Let f and g be functions in $L^2(\mathbb{R}^n)$. For a moment assume that $f * g$ is in $L^1_{loc}(\mathbb{R}^n)$. Then it defines a regular distribution acting on $\varphi \in \mathcal{D}(\mathbb{R}^n)$ by

$$\langle f * g(x), \varphi(x)\rangle = \int_{\mathbb{R}^n} f * g(x)\varphi(x)dx = \int \left(\int f(y)g(x-y)dy\right) \varphi(x)dx.$$

By Fubini's theorem,

$$\begin{aligned}
&\int \left(\int f(y)g(x-y)dy\right)\varphi(x)dx = \int f(y)\left(\int g(x-y)\varphi(x)dx\right)dy \\
&= \int f(y)\left(\int g(t)\varphi(t+y)dt\right)dy = \int f(y)\langle g(t), \varphi(t+y)\rangle dy \\
&= \langle f(y), \langle g(t), \varphi(t+y)\rangle\rangle.
\end{aligned}$$

Thus,

$$\langle f * g(x), \varphi(x)\rangle = \langle f(y), \langle g(t), \varphi(t+y)\rangle\rangle\,, \ \varphi \in \mathcal{D}(\mathbb{R}^n). \tag{11}$$

Therefore, in the general case of arbitrary distributions $f, g \in \mathcal{D}'(\mathbb{R}^n)$ it is natural to take equality (11) as a *definition* of the convolution $f * g$. However, the right-hand side of equality (11) is not defined for arbitrary distributions f and g since the function $y \mapsto \langle g(t), \varphi(t+y)\rangle$ is just of class $C^\infty(\mathbb{R}^n)$ and in general need not have compact support. The support is clearly compact if the distribution g itself has compact support. So in this case the convolution is well-defined. Similarly, if f has compact support then it acts on any function from $C^\infty(\mathbb{R}^n)$ as previously discussed, so the right-hand of equality (11) is also well-defined. We summarize this in the following.

Proposition 10.9. *The convolution $f * g$ of two distributions $f, g \in \mathcal{D}'(\mathbb{R}^n)$ is a distribution correctly defined by the equality (11) if at least one of the distributions f and g has compact support.*

Example 10.5. For any distribution $f \in \mathcal{D}'(\mathbb{R}^n)$ we have

$$\langle f * \delta, \varphi\rangle = \langle f(y), \langle \delta(t), \varphi(t+y)\rangle\rangle = \langle f(y), \varphi(y)\rangle,$$

that is $f * \delta = f$. Furthermore

$$\langle \delta * f, \varphi\rangle = \langle \delta(y), \langle f(t), \varphi(t+y)\rangle\rangle = \langle f(t), \varphi(t)\rangle,$$

so that $\delta * f = f$. We obtain the following fundamental identity

$$f * \delta = \delta * f = f$$

for any $f \in \mathcal{D}'(\mathbb{R}^n)$. $\diamond$

We conclude this section by some algebraic properties of convolution.

(1) *The map $(f, g) \mapsto f * g$ is bilinear.* This is obvious.
(2) *We have*

$$D^\alpha(f * g) = f * D^\alpha g = D^\alpha f * g. \tag{12}$$

For the proof we consider $\varphi \in \mathcal{D}(\mathbb{R}^n)$. Then

$$\begin{aligned}\langle D^\alpha(f * g), \varphi\rangle &= (-1)^{|\alpha|}\langle f * g, D^\alpha\varphi\rangle = \langle f(y), \langle g(t), (-1)^{|\alpha|}(D^\alpha\varphi)(t+y)\rangle\rangle \\ &= \langle f(y), \langle D^\alpha g(t), \varphi(t+y)\rangle\rangle = \langle f * D^\alpha g, \varphi\rangle,\end{aligned}$$

which proves the first equality of (12). For the second, we observe that $\langle D^\alpha g(t), \varphi(t+y)\rangle = (D^\alpha g) * \varphi(-y)$; hence it follows from (9) that

$$\langle D^\alpha g(t), \varphi(t+y)\rangle = (-1)^{|\alpha|} D^\alpha \langle g(t), \varphi(t+y)\rangle.$$

Therefore,

$$\begin{aligned}\langle f(y), \langle D^\alpha g(t), \varphi(t+y)\rangle\rangle &= (-1)^{|\alpha|}\langle f(y), D^\alpha\langle g(t), \varphi(t+y)\rangle\rangle \\ &= \langle D^\alpha f(y), \langle g(t), \varphi(t+y)\rangle\rangle = D^\alpha f * g.\end{aligned}$$

A simpler proof can be given using Theorem 8.7:

$$\begin{aligned}(-1)^{|\alpha|}\langle f(y), \langle g(t), (D^\alpha\varphi)(t+y)\rangle\rangle &= (-1)^{|\alpha|}\langle f(y), D^\alpha\langle g(t), \varphi(t+y)\rangle\rangle \\ &= \langle D^\alpha f(y), \langle g(t), \varphi(t+y)\rangle\rangle.\end{aligned}$$