NOTES ON DISTRIBUTIONS AND THEIR FOURIER TRANSFORMS

1. Distributions

Definition 1.1. Let $U \subset \mathbb{R}^n$ be an open set. We define

$$D(U) = \{\varphi : \mathbb{R}^n \to \mathbb{C} \mid \varphi \text{ is smooth, compactly supported and } \operatorname{supp}(\varphi) \subset U\}$$

to be the set of *test functions* on U. Given $\varphi_1, \varphi_2, \ldots \in D(U)$ and $\varphi \in D(U)$, then we say that $\varphi = \lim_{n\to\infty} \varphi_n$ if there exists a compact set $K \subset U$ such that φ and all φ_n have support inside K and for all $\alpha \in \mathbb{Z}^n_{\geq 0}$ we have that

$$\partial^\alpha \varphi_n \to \partial^\alpha \varphi \text{ uniformly on K.}$$

Example 1.2. Let $h : \mathbb{R} \to \mathbb{C}$ be the map $h(x) = \mathbb{1}_{(0,\infty)} \exp(-\frac{1}{x})$. This is a smooth map with support $[0, \infty)$. Thus $\varphi : \mathbb{R} \to \mathbb{C}$ given by $\varphi(x) = h(x)h(1-x)$ is a smooth map with support $[0, 1]$. Thus $\varphi \in D((-\epsilon, 1+\epsilon))$ for all $\epsilon > 0$ (but not for $\epsilon = 0$). Now let $\phi_t \in D(\mathbb{R})$ be given by $\phi_t(x) = \phi(x+t)$, then clearly $\lim_{n\to\infty} \phi_{1/n}\phi$ in $D(\mathbb{R})$ but the sequence ϕ_n, $n \in \mathbb{Z}$, does not converge (because the union of the supports is unbounded, hence not compact).

Note that $D(U)$ is closed under partial differentiation, and partial differentiation is continuous (preserves limits).

Definition 1.3. A *distribution* on $U \subset \mathbb{R}^n$ is a linear functional $f : D(U) \to \mathbb{C}$ that is *continuous* in the sense that if $\phi_1, \phi_2, \ldots \in D(U)$ converge to $\phi \in D(U)$ then $f(\phi_1), f(\phi_1), \ldots$ converges to $f(\phi)$. We let $D'(U)$ denote the space of distributions on U.

Example 1.4. Any measure μ on $\mathbb{R}^n$ that is finite on compact sets is a distribution in $D(\mathbb{R}^n)$, e.g., $\phi \mapsto \int \phi d\mu$. Consider the distribution $\delta' \in D'(\mathbb{R})$ given by $\delta'(\phi) = -\phi'(0)$. This distribution cannot arise from a measure as can be seen as follows. Choose $\phi_j \in D(\mathbb{R})$ supported on $[-1, 1]$ such that $\|\phi_j\|_\infty \to 0$ but $\phi_j'(0) = 1$, then if δ' coincides with a measure μ, then we have $-1 = \delta'(\phi_j) = \int \phi_j d\mu \to 0$, a contradiction.

The following definition describes why we called the example above δ'.

Definition 1.5. Let $f \in D'(U)$ where $U \subset \mathbb{R}^n$. Let $\partial_j \phi$ denote the j-th partial derivative of a smooth map ϕ. We can define $\partial_j f \in D'(U)$ by

$$\partial_j f(\phi) = -f(\partial_j \phi) \quad \text{for all } \phi \in D(U).$$

We now explain the minus sign in the definition.

Proposition 1.6. Let f be a continuously differentiable function on $\mathbb{R}^n$ (not necessarily compactly supported). This defines a distribution μ_f on $\mathbb{R}^n$ via $\mu_f(\phi) = \int \phi(x) f(x) d^n x$ where $d^n x$ is the lebesgue measure on $\mathbb{R}^n$. Then

$$\partial_j \mu_f = \mu_{\partial_j f}.$$

Proof. For convenience, suppose $j = n$. Then for any $\phi \in D(U)$ we have

$$\mu_{\partial_n f}(\phi) = \int \partial_n f(x)\phi(x)d^n x$$

Now by Fubini's theorem (the integrand has compact support) we can write this integral as

$$\int \left(\int_{-R}^{R} \partial_n f(x_1, \ldots, x_{n-1}, t)\phi(x_1, \ldots, x_{n-1}, t)dt \right) d^{n-1}(x_1, \ldots, x_{n-1})$$

where $R > 0$ is chosen large enough so that $\phi = 0$ outside of $[-R, R]^n$. Finally, we apply integration by parts to the inner integral and use $\phi(x_1, \ldots, x_{n-1}, \pm R) = 0$ to get that this integral is

$$\int \left(\int_{-R}^{R} -f(x_1, \ldots, x_{n-1}, t)\partial_n \phi(x_1, \ldots, x_{n-1}, t)dt \right) d^{n-1}(x_1, \ldots, x_{n-1})$$

which equals $-\int f(x)\partial_n \phi(x)d^n x$ since $\partial_n \phi = 0$ outside of $[-R, R]^n$. But this is precisely $-\mu_f(\partial_n) = \partial_n \mu_f$. ☐

Thus we have extended the notion of differentiation to distributions, which include also non-differentiable but locally integrable functions via the embedding $f \mapsto \mu_f$ in the proposition above. We now identify μ_f and f as is standard practice.

Example 1.7. Let $H(x) = \mathbb{1}_{[0,\infty)}(x)$. Then $H : \mathbb{R} \to \mathbb{R}$ is discontinuous at 0 thus not differentiable in the classical. Yet it has a distribution derivative as follows $H' = \delta$ where $\delta(\phi) = \phi(0)$ is the Dirac delta distribution (which is the probability measure supported at a single point 0). To see this note that for any smooth $\phi \in D(\mathbb{R})$ supported on $[-R, R]$ we have that

$$-\int H(x)\phi'(x) = -\int_0^R \phi'(x)dx = -\phi(R) + \phi(0) = \phi(0) = \delta(\phi).$$

Example 1.8. Consider a ball that bounces off a wall. Its position can be modelled as $x(t) = t$ for $t < 0$ and $x(t) = -t$ for $t \geq 0$ (the wall is located at $x = 0$ and it hits it at $t = 0$). Its velocity is $x'(t) = 1$ for $t < 0$ and $x'(t) = -1$ for $t > 0$ and $x'(0)$ is undefined. What is its acceleration? It is 0 for all $t \neq 0$, but what is it at $t = 0$? As a distribution the acceleration $x''(t)$ is 2δ, which makes sense as all the impact happens at $t = 0$. Of course, in real life maybe $x''(t)$ is continuous and the impact happens on some very small time scale $[-\epsilon, \epsilon]$ as the ball is squashed and unsquashed, but nonetheless $\int_{-\epsilon}^{\epsilon} x''(t)dt = 2$ still holds.

Definition 1.9. (Convergence of Distributions) We say that a sequence of distributions $f_1, f_2, \ldots \in D'(U)$ converges to $f \in D'(U)$ (in $D'(U)$) if $f_i(\varphi) \to f(\varphi)$ for all $\varphi \in D(U)$

Example 1.10. Let $f : \mathbb{R} \to \mathbb{C}$ be an integrable function with $\int_{-\infty}^{\infty} f(x)dx = 1$. Let $f_n(x) = nf(nx)$. Thus if $\varphi \in D(\mathbb{R})$ then by making the substibution $u = nx$ we get

$$\int_{\mathbb{R}} f_n(x)\varphi(x)dx = \int_{\mathbb{R}} \frac{du}{dx} f(nx)\varphi(x)dx = \int_{\mathbb{R}} f(u)\varphi(\frac{u}{n})du \to \varphi(0) \quad \text{as } n \to \infty$$

where we used the dominated convergence theorem (the integrand is bounded by the integrable function $\|\varphi\|_\infty f$ and converges to $f(u)\varphi(0)$ pointwise).Thus $f_n \to \delta$ in $D'(\mathbb{R})$.

Lemma 1.11. Suppose that $\phi_n, \psi_n, \phi, \psi : U \to \mathbb{R}$ are smooth functions, where $U \subset \mathbb{R}^d$ is open, such that for some compact $K \subset U$ and for all $\alpha \in \mathbb{Z}^d_{\geq 0}$ we have that $\partial^\alpha \phi_n \to \partial^\alpha \phi$ and $\partial^\alpha \psi_n \to \partial^\alpha \psi$ uniformly on K. Then for all α, we have that $\partial^\alpha(\phi_n \psi_n) \to \partial^\alpha(\phi_n \psi_n)$ uniformly on K.

Proof. First we note that

$$|\phi_n\psi_n - \phi\psi| = |\phi_n(\psi_n - \psi) + \psi(\phi_n - \phi)| \le |\phi_n||\psi_n - \phi_n| + |\psi||\phi_n - \phi|$$

converges to 0 uniformly on K. We now prove by induction on k that all order n derivatives of $\phi_n\psi_n$ converge to the corresponding derivatives of $\phi\psi$ (the induction hypothesis is on any such ϕ, ψ and not just for this specific ones). The base case $k = 0$ has now been establishes. Now choose any $1 \le j \le n$ and use the product rule to see that

$$\partial_j(\phi_n\psi_n) = \partial_j(\phi_n)\psi_n + \phi_n(\partial_j\phi_n) \to \partial_j(\phi)\psi + \phi(\partial_j\phi) = \partial_j(\phi\psi)$$

uniformly on K where we applied this $k = 0$ case. Now by induction hypothesis for $|\alpha| = k$ we have $\partial_\alpha\partial_j(\phi_n)\psi_n \to \partial_\alpha\partial_j(\phi)\psi$ uniformly on K and likewise for the second term. Thus $\partial_\alpha\partial_j(\phi_n\psi_n) \to \partial_\alpha(\phi\psi)$, and this completes the induction step. $\square$

Definition 1.12 (Multiplying a distribution by a smooth function)**.** If $f \in D(U)$ is a distributiona and $\psi \in C^\infty(U)$ is any smooth function (not necessarily of compact support in U) then we can define $\psi f \in D(U)$ by

$$\psi f(\phi) = f(\psi\phi).$$

Note that ψf is indeed a distribution as the above lemma shows that if $\phi_n \to \phi$ in $D(U)$ then $\psi\phi_n \to \psi\phi$ in $D(U)$ as well, and so $\psi f(\phi_n) \to \psi f(\phi)$ by the continuity of f.

2. Test functions as a Frechet space

Definition 2.1. A Frechet space is a topological vector space (addition and scalar multiplcation is continuous, the field is either $\mathbb{R}$ or $\mathbb{C}$ which has the usual topology) whose topology comes from an invariant metric d (i.e., $d(v_1 + v, v_2 + v) = d(v_1, v_2)$ for all $v_1, v2, v \in V$) that is complete.

For $K \subset \mathbb{R}^n$ compact we define the norm

$$\|\phi\|_{C^k} = \sup_{x\in K, |\alpha|\le k} |\partial^\alpha\phi|(x).$$

Note that $C_0^k(K)$ is a Banach space and hence a Frechet space with respect to this norm. We define $C_0^\infty(K)$ to be the smooth functions with support inside K and for $\phi \in C_0^\infty(K)$ we define

$$\|\phi\|_{C_0^\infty(K)} = \sum_{k=0}^{\infty} 2^{-k} \min\{1, \|\phi\|_{C^k}\}$$

and we note that $d(\phi_1, \phi_2) = \|\phi_1 - \phi_2\|_{C_0^\infty(K)}$ is an invariant metric that is complete. Moreover, a sequence of test functions $\phi_1, \phi_2, \ldots \in C_0^\infty(K)$ converge to $\phi \in C_0^\infty(K)$ if and only if for all $\alpha \in \mathbb{Z}^n_{\ge 0}$ we have that

$$\partial^\alpha\phi_i \to \phi$$

uniformly on K. In other words, a sequence of test functions in $D(U)$ converges if they all have support inside the same compact subset $K \subset U$ and they converge in $C_0^\infty(K)$ with respect to this metric. This in particular verifies that $C_0^\infty(K)$ is a topological vector space with respect to this metric (the continuity of addition and scalar multiplcation inherits from the same properties of the norms $\|\cdot\|_{C_k}$).

Theorem 2.2 (Theorem 3.8 of [1])**.** Let $U \subset \mathbb{R}^d$ be open. A linear functional $f : D(U) \to \mathbb{C}$ is a distribution (in $D'(U)$) if and only if for all compact subsets $K \subset U$ there exists $c > 0$ and $k \in \mathbb{Z}_{\ge 0}$ such that

$$|f(\phi)| \le c\|\phi\|_{C^k} \quad \text{for all } \phi \in C_0^\infty(K).$$

Proof. Easy to see that any functional satisfying this property is a distribution. To see the converse, suppose that this conditional fails for some compact set $K \subset U$. Then for each positive integer $c = k$ we have

$$|f(\phi_k)| > k\|\phi_k\|_{C^k}$$

for some $\phi_k \in C_0^\infty(K)$. Let $\psi_k = \frac{1}{|f(\phi_k)|}\phi_k$. Thus $|f(\psi_n)| = 1$ but we have

$$|\psi_n|_{C^k} \leq |\psi_n|_{C^n} < \frac{1}{n}$$

for all $n \geq k$ so $\psi_n \to 0$ on $C_0^\infty(K)$, which shows that f is not continuous, i.e., not a distribution. □

Theorem 2.3 (Uniform boundedness)**.** Let V be a Frechet space and suppose that $\mathcal{F}$ is a set of continuous linear functions $f : V \to \mathbb{C}$ such that $\{f(x) \mid f \in \mathcal{F}\}$ is bounded in $\mathbb{C}$ for all $x \in V$. Then there is an open set $U \subset V$ with $0 \in V$ such that $|f(u)| \leq 1$ for all $f \in \mathcal{F}$ and $u \in U$.

Proof. Let

$$U_n = \{x \in V \mid |f(x)| > n \text{ for some } f \in \mathcal{F}\}$$

. Now U_n is an open set. For each $x \in V$, we have that there exists n such that $|f(x)| \leq n$ for all $f \in \mathcal{F}$, which means that $x \notin U_n$. Consequently

$$\emptyset = \bigcap_{n=1}^{\infty} U_n.$$

Thus not all U_n can be dense by Baire's theorem. As some U_n is dense, we have a non-empty open set V such that $V \cap U_n = \emptyset$. Choosing $v_0 \in V$, we have that if $u \in V - v_0$ then $u = v - v_0$ for some $v \in V$ and so

$$|f(u)| = |f(v) - f(v_0)| \leq |f(v)| + |f(v_0)| \leq 2n.$$

Thus we may set $U = \frac{1}{2n}(V - v_0)$, which is open by definition of topological vector space. □

Theorem 2.4 (Lemma 5.4 in [1], no proof given there)**.** Let f_j be a sequence of distributions in $D'(U)$, where $U \subset \mathbb{R}^d$ is open such that $f_j(\phi)$ is bounded for all $\phi \in D(U)$. Then for all compact $K \subset U$ there exists a constant $c > 0$ and $k \in \mathbb{Z}_{\geq 0}$ such that

$$\|f_j(\phi)\| \leq c\|\phi\|_{C_0^k(K)} \quad \text{for all } j \in \mathbb{N} \text{ and } \phi \in C_0^\infty(K).$$

Proof. We apply the uniform boundedness principle above. This implies that there is an open neighbhourhood $\mathcal{U} \subset C_0^\infty(K)$ such that $f_j(u) \leq 1$ for all $u \in \mathcal{U}$ and $j \in \mathbb{N}$. So there exists an R such that if $\|\phi\|_{C_0^\infty(K)} < R$ then $f_j(\phi) \leq 1$. Now take k large enough so that

$$\sum_{i=k}^{\infty} 2^{-i} < \frac{R}{2}.$$

This means that if $\|\phi\|_{C_0^k(K)} < \frac{R}{2}$ then $\|\phi\|_{C_0^k(K)} < R$ and so $f_j(\phi) < 1$. As $\|\cdot\|_{C_0^k(K)}$ is a norm on $C_0^\infty(K)$, we have completed the proof with $c = \frac{2}{R}$. □

Theorem 2.5. Let $U \subset \mathbb{R}^d$ be an open set and suppose that $f_1, f_2, \ldots \in D(U)$ is a sequence of distributions such that $\lim_{j\to\infty} f_j(\varphi)$ exists in $\mathbb{C}$ for all $\varphi \in D(U)$.

(1) Then there exists a a distribution $f \in D(U)$ such that

$$f = \lim_{j\to\infty} f_j.$$

(2) If $\varphi, \varphi_j \in D(U)$ are such that $\lim_{j\to\infty} \varphi_j = \varphi$ then $f_j(\varphi_j)$ converges to $f(\varphi)$.

Proof. Define $f(\varphi) = \lim_{j\to\infty} f_j(\varphi)$. It remains to show that this defines a distribution (is continuous). Let K be a compact set. Applying the uniform boundedness principle we have a constant $c > 0$ and $k \in \mathbb{Z}_{\geq 0}$ such that

$$|f_j(\varphi)| \leq c\|\varphi\|_{C_0^k(K)} \quad \text{for all } j \in \mathbb{N}, \varphi \in C_0^\infty(K).$$

Thus as $f_j(\varphi) \to f(\varphi)$ we have that

$$|f(\varphi)| \leq c\|\varphi\|_{C_0^k(K)} \quad \text{for all } \varphi \in C_0^\infty(K).$$

This implies the continuity of f, thus $f \in D(U)$. Now suppose that $\varphi_j \in C_0^\infty(K)$ converge to $\varphi \in C_0^\infty(K)$. Thus

$$|f_j(\varphi_j) - f(\varphi)| \leq |f_j(\varphi_j - \varphi)| + |f_j(\varphi) - f(\varphi)| \leq c\|\varphi_j - \varphi\|_{C^k} + |f_j(\varphi) - f(\varphi)|$$

and the first term converges to 0 as $\varphi_j \to \varphi$ while the second converges to 0 as $f_j \to f$. □

3. Support of a distribution

If $U \subset V \subset \mathbb{R}^d$ are open sets, then there is a continuous (preserves limits) inclusion $D(U) \to D(V)$. This induces a restriction map $p_{U,V} : D'(V) \to D'(U)$ where $(p_{U,V}f)(\phi) = f(\phi)$ for $\phi \in D(U) \subset D(V)$ and $f \in D'(U)$. Note that this is continuous (preserves limits of distributions). We also use the notation $f|_U = p_{U,V}f$.

Lemma 3.1. Suppose that U is an open set, $f \in D'(U)$ and suppose that for each $x \in U$ there exists an open neighbhourhood $U_x \subset U$ of x such that $p_{U_x,U}f = 0$. Then $f = 0$.

Proof. We take $\phi \in D(U)$, thus there is a compact set K such that $K \subset u$ and ϕ is supported on K. Now by compactness, we can find a finite cover of $U_1, \dots U_n$ of K such that f restricts to 0 on each $U_i \subset U$. Choose U_i such that the closure of U_i is in U. By partition of unity theorem, we may choose $\psi_1, \dots, \psi_n \in D(U)$ such that $\sum_{i=1}^n \psi_i(x) = 1$ for all $x \in K$ and $\operatorname{supp} \psi_i \subset U_i$. Thus $\phi = \phi \sum_i \psi_i$ and so $f(\phi) = \sum_i f(\phi\psi_i) = 0$. □

This lemma shows that if $f|_V = 0$ for all $V \in \mathcal{V}$, where $\mathcal{V}$ is a collection of open sets, then setting $V_{\max} = \cup_{V\in\mathcal{V}} V$ we have that $f|_{V_{\max}} = 0$. Thus the following definition is well defined.

Definition 3.2. If $U \subset \mathbb{R}^d$ is open and $f \in D'(X)$ we define the support of f, denoted by $\operatorname{supp} f$, to the smallest closed set such that $f|V = 0$ where $V = X \setminus \operatorname{supp} f$.

Example 3.3. Consider the distribution $f \in D'(\mathbb{R}_{>0})$ given by

$$f(\phi) = \int_0^\infty e^{1/x^2}\phi(x)dt.$$

It is indeed a distribution since e^{-1/x^2} is integrable on each compact subset of $\mathbb{R}_{>0}$, but on any other compact subset containing 0. This distribution is not the restriction of any distribution $g \in D(\mathbb{R})$. To see this, suppose that it was. Now let $\phi_n = \phi(x + \frac{1}{n})$ where $\phi : \mathbb{R} \to [0, 1]$ is smooth and has support in $[0, 1]$ and $\phi(x) > e^{-1/x}$ on $(0, \frac{1}{2})$. Thus

$$g(\phi) = \lim_{n\to\infty} g(\phi_n) = \lim_{n\to\infty} f(\phi_n) = \infty,$$

a contradiction (distributions have finite values in $\mathbb{C}$).

We justify the partition of unity used above.

Proposition 3.4. Let $B(a,r) \subset B(a,r') \subset \mathbb{R}^d$ are open balls. There is a smooth function $\phi : \mathbb{R}^d \to [0,1]$ that is 1 on $B(a,r)$ and 0 outside $B(a,r')$.

Proof sketch. We just need to prove this for $d = 1$ and then build such a radial function. We already saw that we have a compactly supported $\psi : \mathbb{R} \to [0,1]$ supported on $[0,\epsilon]$ where $0 < \epsilon < \frac{1}{2}$. Now let $\psi_2(x) = \int_{-\infty}^{x} \psi(t)dt$. We see that $\psi_2(x)$ is constant for $x > \epsilon$ and is zero on $x < 0$. Consequently $\psi_3(x) = \psi_2(x)\psi_2(1-x)$ has values in $[0,1]$, is compactly supported and is constant on the interval $(\epsilon, 1-\epsilon)$. We can now translate and scale ψ_3 appropriately. □

Proposition 3.5 (Partition of unity). Let $K \subset \mathbb{R}^d$ be a compact set and suppose that $U \supset K$ is open. Suppose that $\mathcal{U}$ is a collection of open subsets of U that covers K. Then there exist smooth functions $\psi_1, \dots, \psi_n : \mathbb{R}^d \to [0,1]$ such that

$$\psi := \sum_{i=1}^{n} \psi_i$$

satisfies that $\psi(x) = 1$ for $x \in K$ and ψ_i has support inside some element of $\mathcal{U}$.

Proof. By compactness, we may find finitely many balls $B(a_1, r_1), \dots, B(a_n, r_n)$ that cover K such that $B(a_i, 2r_i)$ is a subset of some element of $\mathcal{U}$ (and thus $B(a_i, 2r_i)$ are subsets of U). Now apply the previous construction to find some smooth $\phi_i : \mathbb{R}^d \to [0,1]$ that equals 1 on $B(a_i, r_i)$ and has support inside $B(a_i, 2r_i)$. Now let $\psi_1 = \phi_1$ and for $1 < i \leq n$ define $\psi_i = \phi_i \prod_{j<i}(1-\phi_j)$. Observe that ψ_i has support inside the support of ϕ_i, thus inside some element of $\mathcal{U}$, as required. Moreover, by induction we have that

$$\sum_{i=1}^{j} \psi_i = 1 - \prod_{i=1}^{j}(1-\phi_i).$$

In particular for $j = n$ this means that by setting $\psi = \sum_{i=1}^{n} \psi_i$ we have that $\psi(x) = 1$ for $x \in B(a_i, r_i)$, and thus for all $x \in K$. Moreover, if $\phi(x) = 0$ then $\psi_i(x) = 0$ and thus ψ_i has support inside some element of $\mathcal{U}$, as required. □

Theorem 3.6 (Gluding distributions). Suppose that $X \subset \mathbb{R}^d$ is an open set and suppose that $\mathcal{U}$ is a collection of open subsets of X that cover X. Suppose that for each $U \in \mathcal{U}$ there is a distribution $f_U \in D'(U)$ such that these f_U are compatible in the sense that $f_U|_{U\cap V} = f_V|_{U\cap V}$ are the same distributions on $D'(U \cap V)$. Then there is a unique distribution $f \in D'(X)$ such that $f|U = f_U$ for all $U \in \mathcal{U}$.

Proof. We construct f as follows (show that it is well defined later): For each $\phi \in D(X)$, choose a compact set $K \subset X$ containing the support of ϕ. Now we may apply Parition of Unity to find open sets $U_1, \dots, U_n \in \mathcal{U}$ that cover K and $\psi_i : \mathbb{R}^d \to [0,1]$ with support inside U_i such that $\psi := \sum_{i=1}^{n} \psi_i$ satisfies that $\psi(x) = 1$ for all $x \in K$. We now define

$$f(\phi) = \sum_{i=1}^{n} f_{U_i}(\phi\psi_i).$$

Note that this shows uniqueness since $\phi = \sum_{i=1}^{n} \phi\psi_i$ on $\mathbb{R}^d$.

We now show that f is well defined (does not depend on the choice of K or the choice of the U_i or the choice of ψ_i). To see this, suppose that K', U_j' and ψ_j' are such other choices. Then we make a common refinement and show it assigns the same value to our $f(\phi)$ as follows. Let $K'' = K \cap K'$, it clearly contains the support of ϕ and is compact. Now the sets U_i and U_j cover K''. Thus the sets $V_{i,j} = U_i \cap U_j$ cover K''. Moreover, $\psi_{i,j} := \psi_i \psi'_j : \mathbb{R}^d \to [0,1]$ has support inside $V_{i,j}$ and

$$\sum_{i,j} \psi_{i,j} = \left(\sum_i \psi_i\right)\left(\sum_j \psi'_j\right)$$

and thus equals 1 on K''. So this common refinement is a new partition of unity. But now

$$\sum_{i,j} f_{U_i}|_{V_{i,j}}(\psi_{i,j}\phi) = \sum_{i,j} f_{U_i}(\psi_{i,j}\phi) = \sum_i f_{U_i}(\phi\psi_i \sum_j \psi'_j) = \sum_i f_{U_i}(\phi\psi_i)$$

where we used that $\phi\psi_i \sum_j \psi'_j = \phi_i\psi_i$ since $\sum_j \psi'_j(x) = 1$ for all $x \in K'$ and thus all x in the support of ϕ. This completes the proof of well definedness since by assumption,

$$f_{U_i}|_{V_{i,j}}(\psi_{i,j}\phi) = f_{U'_j}|_{V_{i,j}}(\psi_{i,j}\phi)$$

and so

$$\sum_i f_{U_i}(\phi\psi_i) = \sum_j f_{U'_j}(\phi\psi'_j)$$

by the same calculation as above. Suppose that $\phi, \phi' \in D(X)$. Thus to compute $f(\phi + \phi')$ we may choose a compact set $K \subset U$ that contains the support of ϕ and ϕ'. Now choose $U_1, \ldots, U_n \in \mathcal{U}$ that cover K, thus by definition

$$f(\phi_1 + \phi_2) = \sum_i f_{U_i}(\psi_i(\phi_1 + \phi_2)) = \sum_i f_{U_i}(\psi_i\phi_1) + \sum_i f_{U_i}(\psi_i\phi_2) = f(\phi_1) + f(\phi_2)$$

where the ψ_i are chosen as in the construction. Linearity of f now easily follows. We now show the continuity of f. If $\phi_k \to \phi \in D(X)$ then there is a compact set $K \subset X$ containing all their supports. Thus $\psi_i\phi_k \to \psi\phi$ and the continuity of each f_{U_i} gives continuity of f. Finally, it remains to show that $f|_U = f_U$ for all $U \in \mathcal{U}$. Thus suppose that $\phi \in D(U)$ and choose a compact set $K \subset U$ such that ϕ has support inside K. As U already covers K, by definition we have that

$$f|_U(\phi) = f(\phi) = f_U(\psi\phi) = f_U(\phi)$$

for some $\psi : \mathbb{R}^d \to [0,1]$ smooth that equals 1 on K and has suppose inside U (so $\phi\psi = \psi$ everywhere). □

4. Distributions with compact supports

If $X \subset \mathbb{R}^d$ is an open set, we let $\mathcal{E}(X) = C^\infty(X)$ denote the set of all smooth functions on X. We have $D(X) \subset \mathcal{E}(X)$ and the inclusion may be strict, for example constant non-zero functions are not in $D(\mathbb{R})$ but are in $C(\mathbb{R})$.

Definition 4.1. We say that $\phi_j \in \mathcal{E}(X)$ converges to $\phi \in \mathcal{E}(X)$ if for all compact sets $K \subset X$ and $\alpha \in \mathbb{Z}^d_{\geq 0}$ we have that

$$\partial^\alpha \phi_j \to \partial^\alpha \phi \quad \text{uniformly on } K.$$

Example 4.2. In $\mathcal{E}(\mathbb{R})$, we have that $f_n(x) = \frac{1}{n}x$ converge to the 0 function. However the convergence is not uniform on the whole of $\mathbb{R}$ itself, but is on every bounded (hence every compact) set. If $\phi : \mathbb{R} \to \mathbb{R}$ is a smooth compactly supported non-zero function then $\phi \in D(X) \subset \mathcal{E}(X)$ and $\phi_n(x) = \phi(x - n)$ are also in $\mathcal{E}(X)$. Note that $\phi_n \to 0$ in $\mathcal{E}(X)$ but not in $D(X)$ (as there is no compact subset containing all supports of the ϕ_n).

Definition 4.3. If $X \subset \mathbb{R}^d$ is open, we let $\mathcal{E}'(X)$ denote the space of linear maps $f : \mathcal{E}(X) \to \mathbb{C}$ that satisfy the property that if $\phi_j \in \mathcal{E}(X)$ converges to $\phi \in \mathcal{E}(X)$. We say $f_j \in \mathcal{E}'(X)$ converges to $f \in \mathcal{E}'(X)$ if $f_j(\phi) \to f(\phi)$ for all $\phi \in \mathcal{E}(X)$.

We now aim to classify $\mathcal{E}'(X)$ and find a natural embedding into $D'(X)$. First observe that given $f \in \mathcal{E}'(X)$, if we restrict this function to $D(X)$, then we get an element of $\mathcal{D}'(X)$. Thus we have a map

$$\iota : \mathcal{E}'(X) \to D(X).$$

Lemma 4.4. If $X \subset \mathbb{R}^d$ is open, then it has an *exhaustion by compact sets*, which we define to be a sequence $K_1 \subset K_2 \subset \dots$ of compact sets such that

$$X = \bigcup_{n=1}^{\infty} K_n$$

and such that any compact sets $K \subset X$ satisfies that $K \subset K_n$ for some n.

Proof. If $X = \mathbb{R}^d$, just take a compact ball of radius n. Otherwise, let $C = \mathbb{R}^d \setminus X$. Note that C is closed. Now let $K_n \subset X$ be those points in ball of radius n around 0 with distance at most $\frac{1}{n}$ to C. Clearly $X = \bigcup_{n=1}^{\infty} K_n$ since any point not in C must have positive distance to C as C is closed. Now let $K \subset X$ be any open set. Then $K \cap C = \emptyset$ and it follows that K has positive distance to C, as otherwise there is $k_n \in K$ and $c_n \in C$ such that $d(k_n, c_n) \to 0$ and by taking a subsequence we assume $k_n \to k \in K$, thus $d(k, C) = 0$, contradicting the disjointness of K and C. It follows that $K \subset K_n$ for large enough n as K is bounded and has positive distance to C.

□

Lemma 4.5. The map $\iota : \mathcal{E}'(X) \to D(X)$ is well defined, continuous (preserves limits) and injective.

Proof. To see well defined, observe that $D(X) \subset \mathcal{E}(X)$ this embedding is continuous, i.e., if $\phi_j \to \phi$ in $D(X)$ then their partial derivatives converge uniformly on all compact sets K by definition. Thus if $f \in \mathcal{E}'(X)$ then indeed $f|_{D(X)} \in \mathcal{E}(X)$. Continuity is obvious as the definition is the same.

Now we show injectivity, thus we wish to show that this linear map ι has a trivial kernel. Thus suppose that $f \in \mathcal{E}'(X)$ satisfies that $f(\phi) = 0$ for all $\phi \in D(X)$. We must now show that $f(\psi) = 0$ for all $\psi \in \mathcal{E}(X)$. To see this, first write

$$X = \bigcup_{n=1}^{\infty} K_n$$

where $K_1 \subset K_2 \subset \dots$ is an exhuastion by compact subsets of X (as defined in the lemma above). Now by partition of unity theorem, we can find a map $\phi_n : \mathbb{R}^d \to [0, 1]$ with compact support inside X that is equal to 1 on K_n. Thus $\phi_n \psi \in D(X)$ as it has compact support in X. Hence $f(\phi_n \psi) = 0$ by assumption. Now $\phi_n \psi \to \psi$ in $\mathcal{E}(X)$ since for any compact set K, we have that $K \subset K_n$ for large enough n and so $\phi_n \psi = \psi$ on K. Thus by continuity of f, we have that

$$f(\phi) = \lim_{n \to \infty} f(\phi_n \psi) = 0,$$

as desired. □

The next example shows that this embedding is not surjective.

Example 4.6. Consider the distribution $f \in D'(\mathbb{R})$ given by $f(\phi) = \int_{\mathbb{R}} \phi(x)dx$. We claim that $f \neq \iota g$ for some $g \in \mathcal{E}'(X)$. Again let $\phi_n : \mathbb{R} \to [0,1]$ be a compactly supported function equal to 1 on $[-n, n]$. Then $g(\phi_n) = f(\phi_n) > 2n$ does not converge to any value. But ϕ_n converges in $\mathcal{E}(X)$ to the constant function 1 thus $g(\phi_n)$ should converge, a contradiction.

Lemma 4.7. Let $f \in D'(X)$ be a distribution with support K such that $K \subset X$. Then f extends to an element of $\mathcal{E}'(X)$ (is in the image of the map ι).

Proof. Let $\psi \in \mathcal{E}(X)$. Let $\phi \in D(X)$ be an element such that $\phi = 1$ on some open set $U \supset K$ such that $\overline{U} \subset X$ and $\overline{U}$ is compact (exists as $K \subset X$ and we use partition of unity). Now define $g(\psi) = f(\psi\phi)$. We claim that this defines $g \in \mathcal{E}'(X)$. To show that it is well defined, we see that $\psi\phi \in D(X)$ thus $f(\psi\phi)$ makes sense (no need to show independence on ϕ, just consider it fixed throughout). Continuity of g is clear since if $\psi_n \to \psi$ in $\mathcal{E}(X)$ then their partial derivatives convege uniformly on the support of ϕ, thus $\psi_n\phi \to \psi_n\phi$ in $D(X)$. Now it remains to show that g agrees with f on $D(X)$. Thus assume $\psi \in D(X)$ already. Now observe that $\psi - \psi\phi = 0$ on $\overline{U}$ and thus $X \in D(U \setminus K)$ and by definition of support, we have that $f|_{X \setminus K} = 0$ and so $f(\psi - \psi\phi) = 0$, thus $g(\psi) = f(\psi)$ as desired. $\square$

We now show that the converse is true, i.e., that $\mathcal{E}'(X)$ coincides with those distributions in $\mathcal{D}'(X)$ whose support is a compact subset of K.

Theorem 4.8. A distribution $f \in D'(X)$ is the restriction of some $g \in \mathcal{E}'(X)$ if and only if the support of f is some compact subset $K \subset X$.

Proof. Suppose that $g \in \mathcal{E}'(X)$, we wish to show that it has compact support K for some $K \subset X$ as an element $g \in D'(X)$. Suppose that it does not, thus for any compact set $K \subset X$ we have that $g|_{X \setminus K}$ is not the zero distribution on $X \setminus K$. Thus there exists $\phi_K \in D(X \setminus K)$ such that $g(\phi_K) = 1$. Now let $K_1 \subset K_2 \subset \ldots$ be an exhaustion of X by compact sets. Set $\phi_n = \phi_{K_n}$.We claim that $\phi_n \to 0$ in $\mathcal{E}(X)$ (not necessarily in $D(X)$ though). To see this, let $K \subset X$ be any compact set. Then $K \subset K_n$ for large enough n, but then $\phi_n = 0$ on such K_n as $\phi_n \in D(X \setminus K_n)$. Thus indeed $\phi_n \to 0$ uniformly on each compact set of K, thus $1 = g(\phi_n) = f(\phi_n) \to 0$ by continuity of f on $\mathcal{E}(X)$. A contradiction. $\square$

Example 4.9. The distribution $f(\phi) = \int_0^1 \phi(x)dx$ is a compactly supported distribution on $\mathbb{R}$, thus it is defined and continuous on all of $C^\infty(\mathbb{R})$. Note that if $\phi_n(x) = \phi(x+n)$ where $\phi : \mathbb{R} \to \mathbb{R}$ is a non-zero compactly supported smooth function, then $\phi_n \to 0$ in $\mathcal{E}(X)$ and indeed $f(\phi_n) \to 0$. However ϕ_n does not converge to 0 in $D(X)$.

Example 4.10. Consider the distribution on $(0,1)$ given by

$$f(\phi) = \int_0^1 e^{1/x^2} \phi(x)dx.$$

It is continuous and well defined since e^{1/x^2} is integrable on each compact subset K in $(0,1)$. However, $f \notin \mathcal{E}'((0,1))$ (not the restriction of any element in $\mathcal{E}'((0,1))$). Indeed, it is not supported on any compact subset K of $(0,1)$. This is despite the support being the bounded subset $(0,1)$ that is closed in $(0,1)$ but not in $\mathbb{R}$. One can directly see that there is not continuous extension of f to $\mathcal{E}((0,1))$ by considering a sequence of elements $\phi_n \in D((0,1))$ that converge monotonically in $\mathcal{E}(X)$ to the constant function 1 on $(0,1)$ and seeing a lack of convergence ($f(\phi_n)$ diverges to ∞).

Theorem 4.11. For any open $X \subset \mathbb{R}^d$, we have that $\mathcal{E}'(X)$ is a dense subset of $D'(X)$.

Proof. Let $K_1 \subset K_2 \subset \dots$ be an exhuastion of X be compact subsets. Choose a $\psi_n \in D(X)$ such that $\psi_n = 1$ on K_n. Now let $f \in D(X)$ be any distribution and define $f_n = \psi_n f \in D(X)$ by

$$f_n(\phi) = f_n(\phi\psi_n).$$

We claim that $f_n \in \mathcal{E}'(X)$ and $f_n \to f$ in $D'(X)$. To justify the first claim, note that for $\phi \in D(U \setminus \operatorname{supp} \psi_n)$ we have that

$$f_n(\phi) = f(\psi_n\phi) = f(0) = 0$$

which shows that f_n has support a subset of $\operatorname{supp} f$, thus has compact support. Now for any $\phi \in D(X)$, there is a compact set $K \subset X$ such that $\phi(x) = 0$ for $x \notin K$. Also, for n large enough we have $K \subset K_n$, thus $\phi = \phi\psi_n$ and so $f_n(\phi) = f(\phi)$ for large enough n. $\square$

5. Convolution of Functions

Proposition 5.1. Let $f, g : \mathbb{R}^d \to \mathbb{C}$ be L^1 functions. Then the function

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

is almost an everywhere well defined function $\mathbb{R}^d \to \mathbb{R}$ that is L^1.

Proof. By Tonelli's theorem (Fubini's theorem for positive functions that are not necessarily L^1) we have that

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} |f(x-y)||g(y)|dxdy = \left(\int_{\mathbb{R}^d} |f(x-y)|dx\right)\left(\int_{\mathbb{R}^d} |g(y)|dy\right) < \infty$$

thus we have that

$$\int_{\mathbb{R}^d} |f(x-y)||g(y)|dy < \infty$$

for almost all x. Thus $(f * g)(x)$ is well defined for almost all x. It also follows from Fubini's theorem that $(f * g)$ is Lebesgue measurable (almost Borel). Finally, the inequality above also shows it is in $L^1(\mathbb{R}^d)$ and in fact the L^1 norm bounded by the product of the L^1 norms. $\square$

Convolution is commutative (use translation invariance) and bilinear. It also preserves properties like continuity and differentiability as follows.

Proposition 5.2. Suppose that $\phi \in C_0(\mathbb{R}^d)$ and $g \in L^1(\mathbb{R}^d)$ and ϕ is n times continuously differentiable (where $n \in \mathbb{Z}_{\geq 0}$). Then $\phi * g$ is n times continuously differentiable, everywhere defined and

$$\partial^\alpha(\phi * g) = (\partial^\alpha\phi) * g$$

whenever $|\alpha| \leq n$.

Proof. We just focus on the $n = 1$ and $n = 0$ cases and the rest follows from induction. For the $n = 0$ case we just have to show that $\phi * g$ is continuous if ϕ is continuous and $g \in L^1(\mathbb{R}^d)$. As ϕ is compactly supported, it is uniformly continuous and so we have that

$$\phi(x+h) - \phi(x) \to 0 \text{ uniformly on } \mathbb{R}^d \text{ as } h \to 0.$$

It now follows that, for each fixed $x \in \mathbb{R}^d$, we have that

$$(\phi * g)(x+h) - \phi * g(x) = \int_{\mathbb{R}^d} (\phi * (x+h-y) - \phi * (x-y))g(y) \to 0$$

as $h \to 0$ by the dominated convergence theorem. Now assuming that ϕ is once continuously differentiable and letting $\partial_1 = \frac{\partial}{\partial x_1}$ we must show that

$$\lim_{t\to 0} \frac{(\phi * g)(x+te_1) - (\phi * g)(x)}{t} = ((\partial_1 \phi) * g)(x).$$

By the mean value theorem, we have that $\frac{1}{t}((\phi)(x+te_1) - (\phi)(x)) = \partial_1\phi(x + s_{t,x}e_1)$ for some $0 \le s_{t,x} \le t$. Thus by uniform continuity of $\partial_1\phi$ we have that

$$\frac{1}{t}((\phi)(x+te_1) - (\phi)(x)) \to \partial_1(\phi)(x)$$

uniformly for $x \in \mathbb{R}^d$ as $t \to 0$. We now by the dominated convergence theorem that

$$\begin{aligned}
\frac{(\phi * g)(x+te_1) - (\phi * g)(x)}{t} &= \int_{\mathbb{R}^d} \frac{\phi(x-y+te_1) - \phi(x-y)}{t} g(y)dy \\
&\to \int_{\mathbb{R}^d} (\partial_1\phi(x-y))g(y)dy \\
&= ((\partial_1\phi) * g)(x)
\end{aligned}$$

as $t \to 0$. □

Proposition 5.3. Suppose that $f_n, f, g \in L^1(\mathbb{R}^d)$ are such that $\|f_n - f\|_\infty \to 0$. Then $|(f_n * g) - (f * g)|_\infty \to 0$.

Proof. For each $x \in \mathbb{R}^d$ we have

$$|(f_n * g)(x) - (f * g)(x)| = |\int_{\mathbb{R}^d} (f_n(x-y) - f(x-y))g(y)| \le |f - f_n|_\infty \int_{\mathbb{R}^d} |g(y)|dy$$

and thus we have uniform convergence. □

Corollary 5.4. The space $D(\mathbb{R}^d)$ is closed under convolution. Moreover, if $\phi_n \to \phi$ in $D(\mathbb{R}^d)$ then $\phi_n * \psi \to \phi * \psi$ in $D(\mathbb{R}^d)$ for all $\psi \in D(\mathbb{R})$.

Proof. If

$$0 \neq (\phi * \psi)(x) = \int \phi(x-y)\psi(y)dy$$

then there exists y such that $x - y \in \operatorname{supp}\phi$ and $y \in \operatorname{supp}\psi$. Thus $x \in \operatorname{supp}(\phi) + \operatorname{supp}(\psi)$, which is a compact set. Thus $\phi * \psi$ is compactly supported. By the proposition above is smooth, thus $D(\mathbb{R}^d)$ is indeed closed under convolution. Now if $\phi_n \to \phi$, then there is some compact set K containing the supports of all ϕ_n, ϕ and ψ and we have uniform convergence of all partial derivatives $\partial^\alpha \phi_n \to \partial^\alpha \phi$ thus we have uniform convergence $(\partial^\alpha \phi_n) * \psi \to (\partial^\alpha \phi) * \psi$ and all these functions are supported on a single compact set $K + K$. Finally, using the identity $\partial^\alpha(\phi * \psi) = (\partial^\alpha \phi) * \psi$ the proof is complete. □

We now show that a convolution can be approximated by an average of translations. This is useful for establishing certain density results.

Proposition 5.5. Let $f, g : \mathbb{R}^d \to \mathbb{R}$ be uniformly continuous bounded L^1 functions and suppose g has compact support. Then

$$(f * g)(x) = \lim_{N\to\infty} \frac{1}{N^d} \sum_{y \in \frac{1}{N^d}\mathbb{Z}^d} f(x-y)g(y)$$

where the limit is uniform on $x \in \mathbb{R}^d$. In particular, $(f * g)$ is a uniform limit of finite linear combinations of transates of f.

Proof. Let for $y \in \frac{1}{N}\mathbb{Z}^d$, let

$$Q_{y,N} = \left[y_1, y_1 + \frac{1}{N}\right) \times \cdots \times \left[y_d, y_d + \frac{1}{N}\right)$$

be the cube of side-length $\frac{1}{N}$ whose leftmost bottom corner is y. Fix $\epsilon > 0$. By uniform continuity and boundedness of f and g, we may find a large enough N_0 such that if for $N > N_0$ we have that

$$|f(x-y)g(y) - f(x-u)g(u)| < \epsilon \quad \text{for all } u \in Q_{y,N}, y \in \frac{1}{N}\mathbb{Z}^d.$$

Now it follows that

$$\begin{aligned}(f * g)(x) &= \sum_{y \in \frac{1}{N}\mathbb{Z}^d} \int_{Q_{y,N}} f(x-u)g(u)du \\ &= \sum_{y \in \frac{1}{N}\mathbb{Z}^d} \frac{1}{N^d}(f(x-y)g(y) + E_{y,N})\end{aligned}$$

where $|E_{y,N}| < \epsilon$. However, notice that $E_{y,N} = 0$ if $Q_{y,N} \cap \text{supp}(g) = \emptyset$. But there are $O(N^d)$ such y since the support of g is compact (if the support is inside $[-B, B)^d$ for some positive integer B, then there are $(2BN)^d$ such y). Thus the total error is $O(\epsilon) \to 0$ as $N_0 \to 0$. This demonstrates the uniform convergence. □

Proposition 5.6 (Approximate identity)**.** Let $f : \mathbb{R}^d \to [0, \infty)$ be an L_1 function such that

$$\int_{\mathbb{R}^d} f(x)dx = 1$$

and suppose that $g : \mathbb{R}^d \to \mathbb{R}$ is a compactly supported continuous function. Let

$$f_\epsilon(x) = \frac{1}{\epsilon^d} f(\epsilon x).$$

Then

$$f_\epsilon * g \to g \quad \text{uniformly on } \mathbb{R}^d.$$

Proof. For each $r > 0$, we let

$$\Delta(r) = \sup_{x \in \mathbb{R}^d} \{|g(x+h) - g(x)| \mid \|h\| \leq r\}$$

and we observe that $\Delta(r) \to 0$ as $r \to 0$ by uniform continuity (g is compactly supported).

$$\begin{aligned}(f_\epsilon * g)(x) = \int_{\mathbb{R}^d} f_\epsilon(y)g(x-y)dy &= \int_{|x|<r} f_\epsilon(y)g(x-y)dy + O(\|g\|_\infty \int_{|x|>r} f_\epsilon(y)dy) \\ &= \int_{|x|<r} f_\epsilon(y)g(x)dy + O(\Delta(r) \int_{|x|<r} f_\epsilon(y)dy) + O(\|g\|_\infty \int_{|x|>r} f_\epsilon(y)dy)\end{aligned}$$

Now observe that since

$$\int_{\mathbb{R}^d} f_\epsilon(x) = 1$$

we have that for each $r > 0$ fixed this quantity converges, as $\epsilon > 0$, to

$$g(x) + O(\Delta(r)).$$

But by uniform continuity, we can choose r small enough so that this error term is arbitrarily small and thus get the desired uniform continuity as the the implicit $O()$ constant does not depend on x. □

The previous two propositions have same nice applications to approximating a function by a nice class of functions. For instance, we now recover the Weierstrass approximation theorem for polynomials.

Theorem 5.7 (Weierstrass approximation)**.** Let $g : \mathbb{R} \to \mathbb{R}$ be a compactly supported continuous function. Then for any compact set K and $\epsilon > 0$, there exists a polynomial $P(x) \in \mathbb{R}[x]$ such that $\sup_{x\in K} |P(x) - g(x)| < \epsilon$.

Proof. We let $f : \mathbb{R} \to [0, \infty)$ be an entire function (a function given by a power series with an infinite radius of convergence) that is bounded, uniformly continuous and such that

$$\int f(x)dx = 1.$$

For example, we can take

$$f(x) = Ce^{-x^2} = C\sum_{n=0}^{\infty} \frac{(-1)^n}{n!} x^{2n}$$

for some approximately chosen constant C. Thus, for $\epsilon > 0$, we obtain by Proposition 5.6 that for sufficiently small $\delta > 0$ we have

$$|g(x) - (f_\delta * g)(x)| < \frac{\epsilon}{2} \quad \text{for all } x \in \mathbb{R}.$$

But now we use Proposition 5.5 to show that

$$|(f_\delta * g)(x) - h(x)| < \frac{\delta}{2} \quad \text{for all } x \in \mathbb{R}$$

where

$$h(x) = \lim_{N\to\infty} \frac{1}{N^d} \sum_{y\in \frac{1}{N^d}\mathbb{Z}^d} f_\delta(x - y)g(y).$$

Thus as the sum defining $h(x)$ is finite, we see that $h(x)$ is a finite linear combination of entire functions $x \mapsto f_\delta(x - y)$. Thus $\|g - h\|_\infty < \epsilon$. Finally, we finish the proof by using the uniform convergence of the power series for $f_\delta(x)$ for $x \in K$. □

6. Tensor Products of Distributions

Given functions $\phi : X \to \mathbb{C}$ and $\psi : Y \to \mathbb{C}$, we define their tensor product

$$\phi \otimes \psi : X \times Y \to \mathbb{C}$$

to be

$$(\phi \otimes \psi)(x, y) = \phi(x)\psi(y).$$

If $X, Y \subset \mathbb{R}^d$ are open sets, then we let

$$D(X) \otimes D(Y) \subset D(X \times Y)$$

denote the set of linear combinations of tensor product

$$\{\phi \otimes \psi \mid \phi \in D(X), \psi \in D(Y)\}.$$

Proposition 6.1. Let $X, Y \subset \mathbb{R}^d$ be open sets. Then space $D(X) \otimes D(Y)$ is dense in $D(X \times Y)$, that is, for any $\theta \in D(X \times Y)$ we can find a sequence $\theta_n \in D(X) \otimes D(Y)$ such that

$$\theta_n \to \theta,$$

that is for each $\alpha \in \mathbb{Z}^d_{\geq 0}$ we have that

$$\partial^\alpha \theta_n \to \partial^\alpha \theta$$

uniformly on some compact set K that contains the support of all the θ_n and θ.

Proof. The proof is similair to that of Weierstrass theorem given above. First, we let $\phi_1 \in D(\mathbb{R}^d)$ and $\phi_2 \in D(\mathbb{R}^d)$ be functions whose integral equals 1. Now let $\phi = \phi_1 \otimes \phi_2$. Now define

$$\phi_\epsilon(z) = \frac{1}{\epsilon}\phi(\epsilon z).$$

Observe that for small enough $\epsilon > 0$ we have that

$$\phi_\epsilon * \theta \in D(X \times Y)$$

since the support of $\phi_\epsilon * \theta$ is an arbitrarily small neighbhourhood of the support of $\theta \in D(X \times Y)$. It now follows from Proposition 5.6 that $\phi_\epsilon * \psi \to \psi$ uniformly for all $\psi \in D(X \times Y)$ and thus in particular

$$\partial^\alpha(\phi_\epsilon * \psi) = \phi_\epsilon * (\partial^\alpha \psi) \to \partial^\alpha \psi$$

uniformly for all $\alpha \in \mathbb{Z}^d_{\geq 0}$. In particular, this means that $\phi_\epsilon * \theta \to \theta$ in $D(X \times Y)$. It now suffices to show that each

$$\phi_\epsilon * \theta$$

is the limit of some elements in $D(X) \otimes D(Y)$. But by Proposition 5.5 we have the uniform limit

$$(\partial^\alpha \phi_\epsilon * \theta)(x) = \lim_{N\to\infty} \frac{1}{N^d} \sum_{y \in \frac{1}{N}(\mathbb{Z}^d)} (\partial^\alpha \phi_\epsilon(x-y))\theta(y).$$

Thus we have written $\phi_\epsilon * \theta$ as a limit in $D(X \times Y)$ of linear combination of translates of ϕ_ϵ, which are clearly in $D(X) \otimes D(Y)$. $\square$

Theorem 6.2 (Tensor product of distributions)**.** Let $u \in D'(X)$ and $v \in D'(Y)$ be two distributions, where $X, Y \subset \mathbb{R}^d$ are open sets. Then there exists a unique distributions $u \otimes v \in D(X \times Y)$, called the tensor product of u and v, such that

$$(u \otimes v)(\phi \otimes \psi) = u(\phi) \cdot v(\psi) \qquad \text{for all } \phi \in D(X), \psi \in D(Y).$$

Moreover, the tensor product is given by the well defined formula

$$(u \otimes v)(\theta) = u(y \mapsto (v(x \mapsto \theta(x,y)))).$$

Proof. Note that uniqueness follows from the fact that $D(X) \otimes D(Y)$ is dense in $D(X \times Y)$, thus any distribution on $D(X \times Y)$ is determined uniquely by its restriction to $D(X) \otimes D(Y)$. Now we show existence by showing that the claimed formula gives a well defined distribution. Firstly, to show that $v(x \mapsto \theta(x,y))$ is well defined we must show that $x \mapsto \theta(x,y)$ is smooth (and it is since θ is) and it is an element of $D(X)$, i.e., its support is a compact subset of X. Now since θ is an element of $D(X \times Y)$ we have that $\operatorname{supp}\theta$ is a compact subset of $X \times Y$. Now suppose that x_0 is in the support of $x \mapsto \theta(x,y)$. This this means that there is a sequence $x_1, x_2, \ldots \to x_0$ such that $\phi(x_i, y) > 0$ and $x_0 \in X$. So in particular $(x_i, y) \in \operatorname{supp}\theta$ and $(x_i, y) \to (x_0, y)$, thus as the support is closed (by definition) we have that $(x_0, y) \in \operatorname{supp}\theta$. Thus we have shown that the support of $x \mapsto \theta(x,y)$ is contained in the projection onto X of the set $\operatorname{supp}(\theta) \cap (X \times \{y\})$, and thus is compact subset of $X \times Y$. Thus $x \mapsto \theta(x,y)$ is an element of $D(X)$, so the expression $v(x \mapsto \theta(x,y))$ is well defined.

Next, we must show that the map $V\theta : Y \to \mathbb{C}$ defined by $y \mapsto v(x \mapsto \theta(x,y))$ is an element of $D(Y)$. We thus need to show it is smooth and has compact support inside Y. For smoothness, we let $\partial_i = \frac{\partial}{\partial y_i}$ denote partial differentiation with respect to the ith coordinate in Y.

Claim: Let $\theta_y(x) = \theta(x, y)$, let e_i be the i-th basis vector in $\mathbb{R}^d$ and let

$$\theta_{y,t} = \frac{\theta_{y+te_i}(x) - \theta_y(x)}{t}.$$

Then $\theta_{y,t} \in D(X)$ and it converges in $D(X)$ to the map

$$(\partial_i \theta_y)(x) := (\partial_i \theta)(x, y).$$

Proof of Claim: For sufficiently small t we have that $y + te_i \in Y$ as Y is open, $\theta_{y+te_i} \in D(X)$ as we saw previously. Thus $\theta_{y,t} \in D(X)$ for sufficiently small t. But now by the mean value theorem we have that

$$\frac{\theta_{y+te_i}(x) - \theta_y(x)}{t} = \partial_i \theta(x, y + s_{t,x} e_i)$$

for some $s_{t,x} \in [0, t]$. Thus by uniform continuity of θ we have that

$$\frac{\theta_{y+te_i}(x) - \theta_y(x)}{t} \to (\partial_i \theta)(x, y) \text{ uniformly for } x \in X.$$

The same argument applies if we replace θ_y with any higher order partial derivative (with respect to the x coordinates). Which shows the convergence in $D(X)$ (uniform convergence of all partial derivatives). QED of claim.

But now the claim implies that

$$\frac{V\theta(y + te_i) - V\theta(y)}{t} = v(\theta_{y,t}) \to v((\partial_i \theta)_{y,t}) = V(\partial_i \theta)$$

thus $V\theta$ is differentiable and hence smooth by applying inductively the argument to $V(\partial_i \theta)$. To show that $V\theta \in D(Y)$ it now remains to show that the support of $V\theta$ is a compact subset of X. Thus suppose that $V(\theta)(y) \neq 0$. Then clearly $x \mapsto \theta(x, y)$ cannot be the zero function and thus $(x, y) \in \mathrm{supp}(\theta)$ for some $x \in X$. It now follows by projection onto Y that y is contained in the projection onto Y of the support of θ, which must be a comapct set. Thus $V\theta$ is indeed in $D(Y)$. This now completes the proof that the expression

$$(u \otimes v)(\theta) = u(y \mapsto (v(x \mapsto \theta(x, y))))$$

is well defined (gives a well defined number on the right hand side).

Now we must show that $(u \otimes v)$ is a distribution (continuous linear functional). Linearity is clear. For continuity, suppose now that $\theta_1, \theta_2, \ldots \to \theta$ in $D(X \times Y)$. Then the map

$$x \mapsto \theta_i(x, y)$$

converge in $D(X)$ to

$$x \mapsto \theta(x, y)$$

and thus continuity of v implies that the maps

$$V(\theta_i) : y \mapsto (v(x \mapsto \theta_i(x, y)))$$

converge to the map $V\theta$ pointwisely. We must do better: we must show that this convergence is uniform for θ and also for all it partial derivatives. As $\partial^\alpha V(\theta) = V(\partial^\alpha \theta)$ as shown above, it is enough to show the uniform convergence of θ and it will follow for partial derivatives. By linearity, let us assume that $\theta = 0$ is the zero function. Now let $K \subset X$ be a compact set that contains all the supports of the $(\theta_i)_y$ (there is a single compact subset of $X \times Y$ containing all the supports of the θ_i, so we just take K to be the projection of that). Now by Theorem 2.2 we have a constant $c > 0$ and integer $k > 0$ such that for all $i = 1, 2, \ldots$ and $y \in Y$ we have that

$$|v((\theta_i)_y)| \leq c \|(\theta_i)_y\|_{C_k}.$$

Thus since the $\theta_i \to \theta = 0$ in $D(X \times Y)$ we have for each $\alpha \in \mathbb{Z}^d_{\geq 0}$ that $\partial^\alpha(\theta_i)_y \to \partial^\alpha \theta_y$ uniformly accross all y. Thus $v((\theta_i)_y)$ converges to 0 uniformly in y. □

7. Convolution of Distributions

Definition 7.1. Let

8. Fourier Transform

References

[1] Duistermaat, J. J.; Kolk, J. A. C. *Distributions. Theory and applications.* Translated from the Dutch by J. P. van Braam Houckgeest. Cornerstones. Birkhäuser Boston, Inc., Boston, MA, 2010. xvi+445 pp. ISBN: 978-0-8176-4672-1