

NOTES ON DISTRIBUTIONS AND THEIR FOURIER TRANSFORMS

1. DISTRIBUTIONS

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

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

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

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

Example 1.2. Let $h : \mathbb{R} \rightarrow \mathbb{C}$ be the map $h(x) = \mathbf{1}_{(0, \infty)} \exp(-\frac{1}{x})$. This is a smooth map with support $[0, \infty)$. Thus $\varphi : \mathbb{R} \rightarrow \mathbb{C}$ 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 $\lim_{n \rightarrow \infty} \phi_{1/n} \phi$ 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 $f : D(U) \rightarrow \mathbb{C}$ that is *continuous* in the sense that if $\phi_1, \phi_2, \dots \in D(U)$ converge to $\phi \in D(U)$ then $f(\phi_1), f(\phi_2), \dots$ 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 $\|\phi_j\|_\infty \rightarrow 0$ but $\phi_j'(0) = 1$, then if δ' coincides with a measure μ , then we have $-1 = \delta'(\phi_j) = \int \phi_j d\mu \rightarrow 0$, 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

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

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 $\phi(x_1, \dots, x_{n-1}, \pm R) = 0$ to get that this integral is

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

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 $-\mu_f(\partial_n) = \partial_n \mu_f$. \square

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 $H : \mathbb{R} \rightarrow \mathbb{R}$ 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

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

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 $f_1, f_2, \dots \in D'(U)$ converges to $f \in D'(U)$ (in $D'(U)$) if $f_i(\varphi) \rightarrow f(\varphi)$ for all $\varphi \in D(U)$

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

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

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 $f_n \rightarrow \delta$ in $D'(\mathbb{R})$.

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

Proof. First we note that

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

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 established. Now choose any $1 \leq j \leq n$ and use the product rule to see that

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

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

Definition 1.12 (Multiplying a distribution by a smooth function). If $f \in D(U)$ is a distribution 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 $\phi_n \rightarrow \phi$ in $D(U)$ then $\psi \phi_n \rightarrow \psi \phi$ in $D(U)$ as well, and so $\psi f(\phi_n) \rightarrow \psi f(\phi)$ 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 multiplication 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 $v_1, v_2, v \in V$) that is complete.

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

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

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 $\phi_1, \phi_2, \dots \in C_0^\infty(K)$ converge to $\phi \in C_0^\infty(K)$ if and only if for all $\alpha \in \mathbb{Z}_{\geq 0}^n$ we have that

$$\partial^\alpha \phi_i \rightarrow \phi$$

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 multiplication 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 $f : D(U) \rightarrow \mathbb{C}$ is a distribution (in $D'(U)$) if and only if for all compact subsets $K \subset U$ there exists $c > 0$ and $k \in \mathbb{Z}_{\geq 0}$ such that

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

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 $|f(\psi_k)| = 1$ but we have

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

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

Theorem 2.3 (Uniform boundedness). Let V be a Frechet space and suppose that \mathcal{F} is a set of continuous linear functions $f : V \rightarrow \mathbb{C}$ 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 $0 \in U$ 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. \square

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 neighbourhood $\mathcal{U} \subset C_0^\infty(K)$ 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 $\|\phi\|_{C_0^\infty(K)} < R$ 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 $c = \frac{2}{R}$. \square

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

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

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

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

Proof. Define $f(\varphi) = \lim_{j \rightarrow \infty} f_j(\varphi)$. 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 $f_j(\varphi) \rightarrow f(\varphi)$ 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 $\varphi_j \rightarrow \varphi$ while the second converges to 0 as $f_j \rightarrow f$. \square

3. SUPPORT OF A DISTRIBUTION

If $U \subset V \subset \mathbb{R}^d$ are open sets, then there is a continuous (preserves limits) inclusion $D(U) \rightarrow D(V)$. This induces a restriction map $p_{U,V} : D'(V) \rightarrow D'(U)$ where $(p_{U,V}f)(\phi) = f(\phi)$ for $\phi \in D(U) \subset D(V)$ and $f \in D'(V)$. 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 neighbourhood $U_x \subset U$ 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 $\text{supp } \psi_i \subset U_i$. Thus $\phi = \sum_i \psi_i \phi$ and so $f(\phi) = \sum_i f(\psi_i \phi) = 0$. \square

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 $f \in D'(U)$ we define the support of f , denoted by $\text{supp } f$, to be the smallest closed set such that $f|_V = 0$ where $V = U \setminus \text{supp } f$.

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

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

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 $\phi : \mathbb{R} \rightarrow [0, 1]$ is smooth and has support in $[0, 1]$ and $\phi(x) > e^{-1/x}$ on $(0, \frac{1}{2})$. Thus

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

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 $\phi : \mathbb{R}^d \rightarrow [0, 1]$ 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 $\psi : \mathbb{R} \rightarrow [0, 1]$ 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. \square

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 $\psi_1, \dots, \psi_n : \mathbb{R}^d \rightarrow [0, 1]$ 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 $\phi_i : \mathbb{R}^d \rightarrow [0, 1]$ 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. \square

Theorem 3.6 (Gluing 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 Partition of Unity to find open sets $U_1, \dots, U_n \in \mathcal{U}$ that cover K and $\psi_i : \mathbb{R}^d \rightarrow [0, 1]$ 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, $\psi_{i,j} := \psi_i \psi'_j : \mathbb{R}^d \rightarrow [0, 1]$ 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 $\phi \psi_i \sum_j \psi'_j = \phi \psi_i$ 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 $U_1, \dots, U_n \in \mathcal{U}$ 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 $\phi_k \rightarrow \phi \in D(X)$ then there is a compact set $K \subset X$ containing all their supports. Thus $\psi_i \phi_k \rightarrow \psi_i \phi$ 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 $\psi : \mathbb{R}^d \rightarrow [0, 1]$ smooth that equals 1 on K and has support inside U (so $\phi\psi = \phi$ everywhere). \square

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 $\alpha \in \mathbb{Z}_{\geq 0}^d$ we have that

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

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 $\phi : \mathbb{R} \rightarrow \mathbb{R}$ 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 $\phi_n \rightarrow 0$ 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 $f : \mathcal{E}(X) \rightarrow \mathbb{C}$ 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 $f_j(\phi) \rightarrow f(\phi)$ 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

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

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 $d(k_n, c_n) \rightarrow 0$ and by taking a subsequence we assume $k_n \rightarrow k \in K$, 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 $\iota : \mathcal{E}'(X) \rightarrow D(X)$ 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 $\phi_j \rightarrow \phi$ 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 exhaustion by compact subsets of X (as defined in the lemma above). Now by partition of unity theorem, we can find a map $\phi_n : \mathbb{R}^d \rightarrow [0, 1]$ 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 $\phi_n \psi \rightarrow \psi$ 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

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

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 $\phi_n : \mathbb{R} \rightarrow [0, 1]$ 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 $\psi_n \rightarrow \psi$ in $\mathcal{E}(X)$ then their partial derivatives converge uniformly on the support of ϕ , thus $\psi_n\phi \rightarrow \psi\phi$ 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 $\psi \in D(U \setminus K)$ 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 $D'(X)$ 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 $K_1 \subset K_2 \subset \dots$ be an exhaustion of X by compact sets. Set $\phi_n = \phi_{K_n}$. We claim that $\phi_n \rightarrow 0$ 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 $\phi_n \rightarrow 0$ uniformly on each compact set of K , thus $1 = g(\phi_n) = f(\phi_n) \rightarrow 0$ 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 $\phi : \mathbb{R} \rightarrow \mathbb{R}$ is a non-zero compactly supported smooth function, then $\phi_n \rightarrow 0$ in $\mathcal{E}(X)$ and indeed $f(\phi_n) \rightarrow 0$. 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 exhaustion of X by 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 $f_n \rightarrow f$ in $D'(X)$. To justify the first claim, note that for $\phi \in D(U \setminus \text{supp } \psi_n)$ we have that

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

which shows that f_n has support a subset of $\text{supp } f$, 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 $f, g : \mathbb{R}^d \rightarrow \mathbb{C}$ 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 $\mathbb{R}^d \rightarrow \mathbb{R}$ 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

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

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

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

as $h \rightarrow 0$ 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

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

By the mean value theorem¹, we have that $\frac{1}{t}(\phi(x + te_1) - \phi(x)) = \partial_1 \phi(x + s_{t,x}e_1)$ for some $0 \leq s_{t,x} \leq t$. Thus by uniform continuity of $\partial_1 \phi$ we have that

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

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

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

as $t \rightarrow 0$. □

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

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

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

and thus we have uniform convergence. □

Corollary 5.4. The space $D(\mathbb{R}^d)$ is closed under convolution. Moreover, if $\phi_n \rightarrow \phi$ in $D(\mathbb{R}^d)$ then $\phi_n * \psi \rightarrow \phi * \psi$ 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 $x - y \in \text{supp } \phi$ and $y \in \text{supp } \psi$. Thus $x \in \text{supp}(\phi) + \text{supp}(\psi)$, 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 $\phi_n \rightarrow \phi$, 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 $\partial^\alpha \phi_n \rightarrow \partial^\alpha \phi$ thus we have uniform convergence $(\partial^\alpha \phi_n) * \psi \rightarrow (\partial^\alpha \phi) * \psi$ 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 $f, g : \mathbb{R}^d \rightarrow \mathbb{R}$ be uniformly continuous bounded L^1 functions and suppose g has compact support. Then

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

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

¹To apply it we have to assume ϕ is real valued, but it is easy to see that it is enough to consider this case.

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 $O(\epsilon) \rightarrow 0$ as $N \rightarrow \infty$. This demonstrates the uniform convergence. \square

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

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

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

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

Then

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

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 $\Delta(r) \rightarrow 0$ as $r \rightarrow 0$ 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

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

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

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 $g : \mathbb{R} \rightarrow \mathbb{R}$ 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 $f : \mathbb{R} \rightarrow [0, \infty)$ 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

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

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 $x \in K$. \square

6. TENSOR PRODUCTS OF DISTRIBUTIONS

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

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

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

$$\theta_n \rightarrow \theta,$$

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

$$\partial^\alpha \theta_n \rightarrow \partial^\alpha \theta$$

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

Proof. The proof is similar 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 neighbourhood of the support of $\theta \in D(X \times Y)$. It now follows from Proposition 5.6 that $\phi_\epsilon * \psi \rightarrow \psi$ uniformly for all $\psi \in D(X \times Y)$ and thus in particular

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

uniformly for all $\alpha \in \mathbb{Z}_{\geq 0}^d$. In particular, this means that $\phi_\epsilon * \theta \rightarrow \theta$ 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

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

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

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

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 $\text{supp } \theta$ is a compact subset of $X \times Y$. Now suppose that x_0 is in the support of $x \mapsto \theta(x, y)$. This means that there is a sequence $x_1, x_2, \dots \rightarrow x_0$ such that $\phi(x_i, y) > 0$ and $x_0 \in X$. So in particular $(x_i, y) \in \text{supp } \theta$ and $(x_i, y) \rightarrow (x_0, y)$, thus as the support is closed (by definition) we have that $(x_0, y) \in \text{supp } \theta$. Thus we have shown that the support of $x \mapsto \theta(x, y)$ is contained in the projection onto X of the set $\text{supp } (\theta) \cap (X \times \{y\})$, 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 $V\theta : Y \rightarrow \mathbb{C}$ 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 i th 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

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

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

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

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 $(x, y) \in \text{supp}(\theta)$ 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 compact 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 $\theta_1, \theta_2, \dots \rightarrow \theta$ 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 its 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 $i = 1, 2, \dots$ and $y \in Y$ we have that

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

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

Finally, if $\phi_1 \in D(X)$ and $\phi_2 \in D(Y)$ then

$$(u \otimes v)(\phi_1 \otimes \phi_2) = v(y \mapsto u(x \mapsto \phi_1(x)\phi_2(y))) = v(y \mapsto \phi_2(y)u(\phi_1)) = u(\phi_1)v(\phi_2)$$

and so this construction does indeed satisfy the definition of a tensor product. \square

7. CONVOLUTION OF DISTRIBUTIONS

7.1. Well defined convolutions of distributions.

Lemma 7.1. Let $v \in D'(\mathbb{R}^d)$ be a distribution, then for $x \in \mathbb{R}^d$ the map define the distribution $T_x v \in D'(\mathbb{R}^d)$ given by

$$T_x v(\phi) = v(y \mapsto \phi(x + y)).$$

Then for each $\phi \in D(\mathbb{R}^d)$ the mapping

$$F_\phi : x \mapsto T_x v(\phi)$$

is smooth. In particular $(\partial^\alpha F_\phi)(x) = v(y \mapsto \partial^\alpha \phi(x + y)) = F_{\partial^\alpha \phi}(x)$.

Proof. Letting e_i be a standard basis vector, we see that

$$\frac{1}{t}(T_{x+te_i} v(\phi) - T_x v(\phi)) = v\left(y \mapsto \frac{\phi(x + te_i + y) - \phi(x + y)}{t}\right).$$

Now by the mean value theorem we see that

$$\frac{\phi(x + te_i + y) - \phi(x + y)}{t} = (\partial_i \phi)(x + se_i + y)$$

for some $0 \leq s \leq t$. By uniform continuity of partial derivatives of ϕ , we see that

$$\frac{\phi(x + te_i + y) - \phi(x + y)}{t} \rightarrow (\partial_i \phi)(x + y) \text{ uniformly over } x, y \in \mathbb{R}^d \text{ as } t \rightarrow 0.$$

We can replace ϕ by any of its partial derivatives and obtain the same uniform convergence for its partial derivatives. Thus the convergence to $(\partial_i \phi)(x + y)$ occurs in $D(\mathbb{R}^d)$. Thus by the continuity of v we get that

$$\frac{1}{t}(T_{x+te_i} v(\phi) - T_x v(\phi)) \rightarrow v(y \mapsto (\partial_i \phi)(x + y)) \quad \text{as } t \rightarrow 0.$$

Thus we obtain the relation that

$$(\partial_i F_\phi)(x) = v(y \mapsto (\partial_i \phi)(x + y)).$$

But the right hand side is clearly equal to

$$T_x v(\partial_i \phi) = F_{\partial_i \phi}(x)$$

and so we obtain the smoothness of F and the stated formula for its partial derivatives by repeatedly applying this relation. \square

Example 7.2. We have

$$T_x \delta_a = \delta_{a+x}$$

where $\delta_a(\phi) = \phi(a)$ is the point mass at $a \in \mathbb{R}^d$.

We define $T_{-x}\phi(y) = \phi(x + y)$ for functions ϕ on \mathbb{R}^d and $x \in \mathbb{R}^d$.

Lemma 7.3. If $v \in D(\mathbb{R}^d)$ and $\phi_1, \phi_2 \dots \in D(\mathbb{R}^d)$ converge to $\phi \in D(\mathbb{R}^d)$ in $D(\mathbb{R}^d)$ then for each compact $K \subset \mathbb{R}^d$ we have that

$$T_x v(\phi_i) \rightarrow T_x v(\phi) \quad \text{uniformly for } x \in K.$$

Proof. Fix a compact set $K \subset \mathbb{R}^d$. By linearity, we only need to show the case $\phi = 0$. Now since ϕ_i converge in $D(\mathbb{R}^d)$ we may find compact set $K' \supset K$ such that each of the compactly supported functions $T_{-x}\phi_i(y)$ for $x \in K$ and $i \in \{1, 2, \dots\}$ have support inside K' . It follows that we may now apply Theorem 2.2 to the compact set K' and find a constant $c > 0$ and integer $k > 0$ such that for all $x \in K$ and $i \in \mathbb{N}$ we have that

$$v(T_{-x}\phi_i) \leq c \|T_{-x}(\phi_i)\|_{C_k}.$$

Now since the ϕ_i and all its partial derivatives converge uniformly to 0, we see that $T_{-x}\phi_i$ must also converge uniformly to 0 at the same rate for $x \in K$ (translation preserves maximum of absolute values of functions). Thus $v(T_{-x}\phi_i) \rightarrow 0$ uniformly for $x \in K$. \square

The previous two lemmata now show that the convolution of a compactly supported distribution and another distribution is well defined.

Definition 7.4. If $u \in \mathcal{E}'(\mathbb{R}^d)$ and $v \in D'(\mathbb{R}^d)$ are distributions with v compactly supported then we can define their convolution $u * v \in D'(\mathbb{R}^d)$ by the formula

$$(u * v)(\phi) = u(x \mapsto v(y \mapsto \phi(x + y))).$$

If instead we assume $u \in D'(\mathbb{R}^d)$ and $v \in \mathcal{E}'(\mathbb{R}^d)$ then the convolution is also a well defined distribution.

Proof of well definedness. More precisely, this is well defined in the case $u \in \mathcal{E}'(\mathbb{R}^d)$ and $v \in D'(\mathbb{R}^d)$ since by Lemma 7.1 we have that the inner term $x \mapsto v(y \mapsto \phi(x + y))$ is a smooth function and since $u \in \mathcal{E}'(\mathbb{R}^d)$ it is defined on the space $C^\infty(\mathbb{R}^d)$ of smooth functions (that are not necessarily compactly supported). Thus the expression $(u * v)(\phi)$ is well defined. Now, it defines a distribution (continuous) since if $\phi_i \rightarrow \phi$ in $D(\mathbb{R}^d)$ then by Lemma 7.3 we have that

$$x \mapsto v(y \mapsto \phi_i(x + y))$$

converge uniformly to

$$x \mapsto v(y \mapsto \phi(x + y))$$

on compact sets K and thus by the definition of convergence in $\mathcal{E}'(\mathbb{R}^d)$ we have that $(u * v)(\phi_i) \rightarrow (u * v)(\phi)$, thus $u * v$ is a distribution (continuous linear functional).

Now consider the other case $u \in D'(\mathbb{R}^d)$ and $v \in \mathcal{E}'(\mathbb{R}^d)$. Then the inner map

$$x \mapsto v(y \mapsto \phi(x + y))$$

has compact support since for large enough y , we have that v and $x \mapsto \phi(y + x)$ will have disjoint supports, hence the expression $(u * v)(\phi)$ is well defined in this case. Now we show continuity. Suppose $\phi_i \rightarrow 0$ in $D(\mathbb{R}^d)$. Then there is a compact set K containing the supports of all the ϕ_i . Thus each map $y \mapsto \phi_i(x + y)$ is supported on the compact set $K - x$. But for all large enough x , the support of v is disjoint from $K - x$. Thus for large enough x we have that

$$v(y \mapsto (\phi_i(x + y))) = 0 \text{ for all } i \in \mathbb{N}.$$

Therefore indeed the map

$$x \mapsto v(y \mapsto (\phi_i(x + y)))$$

are supported on a single compact set and converge in $D(\mathbb{R}^d)$ (by Lemma 7.3) to 0. Thus by continuity of u , we have that $(u * v)(\phi) \rightarrow 0$ and thus $u * v$ is continuous. \square

Example 7.5. If δ_a is the point mass at $a \in \mathbb{R}^d$ and $u \in D'(\mathbb{R}^d)$. Then

$$(\delta_a * u) = \delta_a(x \mapsto u(y \mapsto \phi(y + a))) = u(y \mapsto \phi(y + a)).$$

In particular $\delta_0 * u = u$. Now if $u(\phi) = \int f(x)\phi(x)dx$ for some locally integrable function f . Then

$$(\delta_a * u) = \int f(y)\phi(y + a)dy = \int f(y - a)\phi(y)dy.$$

Thus $\delta_a * u$ has a density function $f(y - a)$, so convolution with δ_a translates the density by a . Moreover, we have that

$$\delta_a * \delta_b(\phi) = \delta_a(x \mapsto \delta_b(y \mapsto \phi(x + y))) = \delta_b(y \mapsto \phi(a + y)) = \phi(a + b) = \delta_{a+b}(\phi).$$

Thus

$$\delta_a * \delta_b = \delta_{a+b}.$$

Example 7.6. If u is not compactly supported then this formula may not be well defined. For example if u, v are both the uniform measure on \mathbb{R} then $u * v(\phi) = \infty$ for any positive function ϕ , thus not a distribution. An even more pathological example arises if

$$u = \sum_{n \in \mathbb{Z}} (-1)^n \text{Lebesgue}_{[n, n+1)}$$

is a signed measure which is still a well defined distribution.

Now observe that if $f, g : \mathbb{R}^d \rightarrow \mathbb{C}$ are locally integrable functions then they may be defined naturally as distributions. We now show that their convolution as distributions is the same as their convolution as functions. To see this, let $\phi \in D(\mathbb{R}^d)$, then

$$\begin{aligned} (f * g)(\phi) &= \int_{\mathbb{R}^d} (f * g)(x)\phi(x)dx \\ &= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(y)g(x - y)\phi(x)dy \right) dx \\ &= \int_{\mathbb{R}^d} f(y) \left(\int_{\mathbb{R}^d} g(x - y)\phi(x)dx \right) dy \\ &= \int_{\mathbb{R}^d} f(y) \left(\int_{\mathbb{R}^d} g(x)\phi(x + y)dx \right) dy \\ &= f(y \mapsto g(x \mapsto \phi(x + y))) \end{aligned}$$

where our use of Fubini's theorem is justified by the fact that $f(y)g(x - y)$ is an L^1 map on $\mathbb{R}^d \times \mathbb{R}^d$ as justified in the proof of Proposition 5.1 and thus so is $f(y)g(x - y)\phi(x)$ as ϕ is bounded. Thus we see that the convolution as functions gives rise to the distribution obtained by taking the convolution as distributions, where we assume either f or g has compact support for the definition of convolution of distributions to be valid.

7.2. Convolution of a compactly supported smooth function and a distribution. Next, we attempt to compute $\psi * u$ where $\psi \in C_0^\infty(\mathbb{R})$ is a compactly supported smooth function. The idea behind the next lemma is that convolution of distributions is bilinear and so must preserve linear combinations, yet convolutions of functions are approximate linear combinations. We let $(S\psi)(x) = \psi(-x)$.

Lemma 7.7. Let $\psi \in C_0^\infty(\mathbb{R}^d)$ and let $u \in D'(\mathbb{R}^d)$. Then for all $\phi \in D(\mathbb{R}^d)$ we have that

$$(\psi * u)(\phi) = u((S\psi) * \phi).$$

Proof. We have by Proposition 5.5 that

$$(1) \quad (\psi * \phi)(x) = \lim_{N \rightarrow \infty} \frac{1}{N^d} \sum_{y \in \frac{1}{N^d} \mathbb{Z}^d} \psi(y) \phi(x - y)$$

where the limit is uniform for $x \in \mathbb{R}^d$. Observe also that for all N , the right hand side is support on the compact set $K = \text{supp}(\phi) + \text{supp}(\psi)$. Note that by compactness of the support, this sum is finite for each N . Applying Proposition 5.5 to any partial derivative of ϕ we also get

$$\begin{aligned} \partial^\alpha(\psi * \phi)(x) &= \psi * (\partial^\alpha \phi)(x) \\ &= \lim_{N \rightarrow \infty} \frac{1}{N^d} \sum_{y \in \frac{1}{N^d} \mathbb{Z}^d} \psi(y) (\partial^\alpha \phi)(x - y) \\ &= \lim_{N \rightarrow \infty} \partial^\alpha \left(\frac{1}{N^d} \sum_{y \in \frac{1}{N^d} \mathbb{Z}^d} \psi(y) \phi(x - y) \right) \quad \text{uniformly for } x \in K \end{aligned}$$

where of course in the final expression the partial derivative is with respect to x and $\psi(y)$ is just a constant. Thus we have shown that the convergence (1) is actually in $D(\mathbb{R}^d)$. Thus we may use the continuity of u to conclude that

$$u(\psi * \phi) = \lim_{N \rightarrow \infty} \frac{1}{N^d} \sum_{y \in \frac{1}{N^d} \mathbb{Z}^d} \psi(y) u(x \mapsto \phi(x - y)).$$

Now we can interpret the right hand side as a Riemann sum approximation for the functions $F(y) = \psi(y)u(x \mapsto \phi(x - y))$ which we know is smooth hence Riemann integrable. Note that the sum is finite for each N as ψ is compactly supported. We thus obtain that

$$u(\psi * \phi) = \int_{\mathbb{R}^d} \psi(y) u(x \mapsto \phi(x - y)) dy = \int_{\mathbb{R}^d} (S\psi)(y) u(x \mapsto \phi(x + y)) dy$$

where we applied the Lebesgue measure preserving map $y \mapsto -y$. But this right hand side is precisely the convolution product

$$(S\psi)(y \mapsto u(x \mapsto \phi(x + y))) = (S\psi) * u.$$

This is clearly equivalent to what we wanted to show since $SS\psi = \psi$. □

Next, we show that the convolution of a compactly supported smooth function and a distribution is a smooth function (more precisely, the distribution corresponding to a smooth function).

Theorem 7.8. Let $\psi \in C_0^\infty(\mathbb{R}^d)$ and let $u \in D'(\mathbb{R}^d)$. Then $\psi * u \in C^\infty(\mathbb{R}^d)$, where we embed $C^\infty(\mathbb{R}^d)$ into $D'(\mathbb{R})$ as usual. In particular,

$$(\psi * u)(x) = u(y \mapsto \psi(y - x)).$$

Proof. If $\phi \in D(\mathbb{R})$ then we have

$$\begin{aligned}
(\psi * u)(\phi) &= u(S\psi * \phi) \\
&= u(SS\phi * S\psi) \\
&= (S\phi * u)(S\psi) \\
&= \int_{\mathbb{R}^d} (S\phi)(x)u(y \mapsto S\psi(x+y))dx = \int_{\mathbb{R}^d} \phi(-x)u(y \mapsto \psi(-x-y))dx \\
&= \int_{\mathbb{R}^d} \phi(x)u(y \mapsto \psi(x-y))dx.
\end{aligned}$$

Thus the distribution $\psi * u$ has a density function $x \mapsto u(y \mapsto \psi(y-x))$ which we know is smooth. \square

7.3. Applications: approximating a distribution with a smooth function.

Lemma 7.9. If $u_n \rightarrow u$ in $\mathcal{E}'(\mathbb{R})$ then $u_n * v \rightarrow u * v$ in $D'(\mathbb{R})$ for $v \in D'(\mathbb{R})$.

Proof. For any $\phi \in D(\mathbb{R})$ we have that

$$(u_n * v)(\phi) = u_n(x \mapsto v(y \mapsto \phi(x+y))) \rightarrow u(x \mapsto v(y \mapsto \phi(x+y))) = (u * v)(\phi).$$

\square

We can now show that any distribution is a limit of compactly supported smooth functions.

Theorem 7.10. The space $C_0^\infty(\mathbb{R}^d)$ is a dense subspace of $D'(\mathbb{R}^d)$. Moreover, if $u \in D'(\mathbb{R}^d)$ has compact support, then there exists a compact set $K \supset \text{supp } u$ and $\theta_n \in C_0^\infty(\mathbb{R}^d)$ such that all θ_n have support inside K and $\theta_n \rightarrow u$ in $\mathcal{E}'(\mathbb{R}^d)$.

Proof. Let $u \in D'(\mathbb{R}^d)$ be a distribution. We can find a sequence of functions $\psi_n \in C_0^\infty(\mathbb{R}^d)$ such that $\psi_n \rightarrow \delta$ in $\mathcal{E}'(\mathbb{R}^d)$ and ψ_n all have support inside $[-1, 1]^d$. It now follows that

$$\psi_n * u \rightarrow \delta * u = u$$

in $D'(\mathbb{R}^d)$. However we know that $\psi_n * u \in C^\infty(\mathbb{R}^d)$ corresponds to the smooth function

$$\theta_n(x) = u(y \mapsto \psi_n(y-x)).$$

Now θ_n does not necessarily have compact support. Let $\chi_n : \mathbb{R}^d \rightarrow [0, 1]$ be a sequence of smooth functions with compact support such that $\chi_n(x) = 1$ for $\|x\| \leq n$. Thus $\chi_n \theta_n \in C_0^\infty(\mathbb{R}^d)$. Now suppose $\phi \in D'(\mathbb{R}^d)$, thus the support of ϕ is contained in a ball of radius n_0 for some positive integer n_0 . For $n \geq n_0$ we thus have that

$$(\chi_n \theta_n)(\phi) = \int_{\mathbb{R}^d} \chi_n(x)dx \theta_n(x)\phi(x)dx = \int_{\mathbb{R}^d} \theta_n(x)\phi(x)dx = \theta_n(\phi) \rightarrow u(\phi).$$

Thus $\chi_n \theta_n \rightarrow u$ in $D(\mathbb{R}^d)$.

Now suppose that u has compact support. Then if $\theta_n(x) \neq 0$ then we must have $\psi_n(y-x) \neq 0$ for some $y \in \text{supp } \phi$ and thus $y-x \in [-1, 1]^d$ and so $x \in K = [-1, 1]^d + \text{supp } u$. Hence all these maps θ_n have support inside the compact set $K \supset \text{supp } u$. We know that $\theta_n \rightarrow u$ in $D'(\mathbb{R}^d)$ but now let us show that the convergence occurs in $\mathcal{E}'(\mathbb{R}^d)$ as well. Thus suppose that $\varphi \in C^\infty(\mathbb{R}^d)$ is a smooth function (not necessarily

of compact support). Then there exists a large enough integer R such that K is inside the ball of radius R around 0. Now it follows that

$$\theta_n(\varphi) = \int_{\mathbb{R}^d} \theta_n(x)\varphi(x)dx = \int_{\mathbb{R}^d} \theta_n(x)\varphi(x)\chi_R(x)dx = \theta_n(\chi_R\varphi) \rightarrow u(\chi_R\varphi) = u(\varphi)$$

where in the last equality we used that the support of u is inside K . Thus indeed the convergence occurs in $\mathcal{E}'(\mathbb{R}^d)$. \square

7.4. Commutativity. We now that the convolution of two L^1 functions is commutative. We now address the question for distributions.

Lemma 7.11. Let $\psi \in C_0^\infty(\mathbb{R}^d)$ and $u \in D'(\mathbb{R}^d)$. Then $u * \psi = \psi * u$.

Proof. Let $\phi \in D(\mathbb{R}^d)$. We know that

$$(\psi * u)(\phi) = u(S\psi * \phi) = u\left(x \mapsto \int (S\psi)(y)\phi(x-y)dy\right).$$

But

$$\begin{aligned} (u * \psi)(\phi) &= u(x \mapsto \psi(y \mapsto (\phi(x+y)))) \\ &= u\left(x \mapsto \int_{\mathbb{R}^d} \psi(y)\phi(x+y)dy\right) \\ &= u\left(x \mapsto \int_{\mathbb{R}^d} (S\psi)(y)\phi(x-y)dy\right) \end{aligned}$$

where in the last line we made the substitution $y \mapsto -y$. Thus we have the desired equality. \square

8. FOURIER TRANSFORM

8.1. Definition and Riemann-Lebesgue Lemma.

Definition 8.1. If $u \in L^1(\mathbb{R})$ is integrable, then we define the Fourier transform $\mathcal{F}u : \mathbb{R}^d \rightarrow \mathbb{C}$ by

$$\mathcal{F}u(\omega) = \int_{\mathbb{R}^d} u(x)e^{-2\pi i\langle x, \omega \rangle} dx.$$

Note that this is well defined since $|e^{-2\pi it}| = 1$ for $t \in \mathbb{R}$, thus

$$|\mathcal{F}u(x)| \leq \int_{\mathbb{R}^d} |u(\omega)| dx = \|u\|_1.$$

Thus $\mathcal{F}u$ is a bounded function on \mathbb{R}^d .

Proposition 8.2. If $u \in L^1(\mathbb{R}^d)$ then the Fourier transform $\mathcal{F}u : \mathbb{R}^d \rightarrow \mathbb{C}$ is uniformly continuous bounded function on \mathbb{R}^d .

Proof. We see that

$$|\mathcal{F}u(\omega + \eta) - \mathcal{F}u(\omega)| = \left| \int_{\mathbb{R}^d} u(x)e^{-2\pi i\langle \omega, x \rangle} (e^{-2\pi i\langle \eta, x \rangle} - 1) dx \right| \leq \int_{\mathbb{R}^d} |u(x)(e^{-2\pi i\langle \eta, x \rangle} - 1)| dx.$$

As this upper bound is independent of ω , uniform continuity of $\mathcal{F}u$ will follow once we show this upper bound converges to zero as $\eta \rightarrow 0$. But this follows from the dominated convergence theorem as the integrand converges pointwise to 0 as $\eta \rightarrow 0$ and the integrand is bounded for all η by the integrable function $2|u|$. \square

Example 8.3. Let $\mathbb{1}_{[a,b]} \in L^1(\mathbb{R})$ be the indicator function of an interval where $a < b \in \mathbb{R}$. Then

$$\mathcal{F}\mathbb{1}_{[a,b]}(\omega) = \int_a^b e^{-2\pi i \omega x} dx = \frac{e^{-2\pi i \omega a} - e^{-2\pi i \omega b}}{2\pi i \omega}.$$

More generally, if $B = [a_1, b_1] \times \cdots \times [a_d, b_d]$ is a box in \mathbb{R}^d then the Fourier transform of its indicator function is

$$\mathcal{F}\mathbb{1}_B((\omega_1, \dots, \omega_d)) = \prod_{j=1}^d \frac{e^{-2\pi i \omega_j a_j} - e^{-2\pi i \omega_j b_j}}{2\pi i \omega_j}.$$

We now show that

$$\lim_{\omega \rightarrow \infty} \mathcal{F}\mathbb{1}_B(\omega) \rightarrow 0.$$

This is clear in one dimension ($d = 1$) as it is $O(\frac{1}{|\omega|})$. In higher dimensions, we have the upper bound

$$O\left(\frac{1}{\max_j |\omega_j|}\right)$$

which also converges to 0 as $\omega \rightarrow \infty$.

The decay at infinity for the Fourier transform given above actually holds more generally as follows.

Theorem 8.4 (Riemann-Lebesgue Lemma). Let $u \in L^1(\mathbb{R}^d)$. Then

$$\lim_{\omega \rightarrow \infty} \mathcal{F}u(\omega) = 0.$$

More precisely, for all $\epsilon > 0$ there exists an $R > 0$ such that if $\|\omega\| > R$ then $|\mathcal{F}u(\omega)| < \epsilon$.

Proof. We know this is true if u is the indicator function of a box, as seen in the example above. Thus, it is also true if u is a linear combination of such functions. Thus the theorem is true on a dense subset of $L^1(\mathbb{R}^d)$. Now suppose that $u_1, u_2 \in L^1(\mathbb{R}^d)$ are such functions that converge to $u \in L^1(\mathbb{R}^d)$. Then

$$\|\mathcal{F}u - \mathcal{F}u_j\|_\infty \leq \|u - u_j\|_1 \rightarrow 0.$$

Thus $\mathcal{F}u_j \rightarrow \mathcal{F}u$ uniformly. But it follows that if $g_j : \mathbb{R}^d \rightarrow \mathbb{C}$ are continuous functions that vanish at ∞ , then any uniform limit $g : \mathbb{R}^d \rightarrow \mathbb{C}$ of such functions also vanishes at infinity. To see this, suppose that $\epsilon > 0$, then there exists j_0 such that for $j \geq j_0$ we have that $|g_j(x) - g(x)| < \frac{\epsilon}{2}$ for all $x \in \mathbb{R}^d$. For any such fixed j , there exists R such that $|g_j(x)| < \frac{\epsilon}{2}$ for $|x| > R$. Thus

$$|g(x) - 0| \leq |g(x) - g_j(x)| + |g_j(x) - 0| < \epsilon$$

for all $|x| > R$. □

8.2. Fourier-Laplace transform. We defined the Fourier transform has a function on \mathbb{R}^d . We now consider whether, with $d = 1$, we can naturally extend the Fourier transform to \mathbb{C} .

Thus we define

$$\mathcal{F}u(z) = \int_{\mathbb{R}} u(x) e^{-2\pi i x z} dx$$

for $z \in \mathbb{C}$ if the integral absolutely converges. This need not be the case, for example if $z = i$ and

$$u(x) = \frac{1}{1 + x^2}$$

which is L^1 then the integral is

$$\int_{\mathbb{R}^d} \frac{e^{2\pi x}}{1 + x^2} dx = \infty.$$

Example 8.5. If $u(x) = e^{-\alpha x^2}$ for some $\alpha > 0$, then we have absolute convergence for any $z \in \mathbb{C}$ since the integrand is bounded by $e^{-\frac{1}{2}\alpha x^2}$ for sufficiently large x .

In fact, let us be a bit more general. Let μ be a signed Borel measure that is the difference of two finite positive Borel measures (for example, and probability measure or a measure with a density in $L^1(\mathbb{R})$). Then define the Fourier-Laplace transform as

$$\mathcal{F}\mu(z) = \int_{\mathbb{R}} e^{-2\pi i x z} d\mu(x)$$

for $z \in \mathbb{C}$ where this integral absolutely converges.

Proposition 8.6 (Fourier-Laplace is holomorphic). Suppose that μ is a difference of finite positive measures on \mathbb{R} . Let $U \subset \mathbb{C}$ be an open set such that for all $z \in U$ we have

$$\int_{\mathbb{R}} |e^{-2\pi i x z}| d\mu(x) < \infty.$$

Then

$$\mathcal{F}\mu(z) = \int_{\mathbb{R}} e^{-2\pi i x z} d\mu(x)$$

defines a holomorphic function of z on U , where the integral absolutely converges, with derivative

$$(\mathcal{F}\mu)'(z) = \int_{\mathbb{R}} -2\pi i x z e^{-2\pi i x z} d\mu(x)$$

where again the integral absolutely converges.

Proof. We wish to compute the limit as $w \in \mathbb{C}$ tends to 0 of

$$\frac{\mathcal{F}u(z+w) - \mathcal{F}u(z)}{w} = \int_{\mathbb{R}} e^{-2\pi i x z} \left(\frac{e^{-2\pi x w i} - 1}{w} \right) dx.$$

Now clearly as $w \rightarrow 0$ then the integrand converges to the correct value. Thus, by the Lebesgue dominated convergence theorem, it remains to establish that for sufficiently small w the integrand can be bounded by a single integrable function. From the mean value inequality, it follows that

$$\left| \frac{e^{-2\pi x w i} - 1}{w} \right| = \sup_{t \in [0,1]} \left| \frac{\partial}{\partial z} \Big|_{z=tw} e^{-2\pi i x z} \right| \leq 2\pi |x| e^{2\pi x w}.$$

Now as $z \in U$, we can find small enough $\epsilon > 0$ such that the ball of radius 2ϵ around z is a subset of U . Now there exists a constant $C_\epsilon > 0$ such that

$$|x| < C_\epsilon (e^{2\pi \epsilon x} + e^{-2\pi \epsilon x}) \text{ for all } x \in \mathbb{R}.$$

We can thus bound the integrand

$$\begin{aligned} \left| e^{-2\pi i x z} \left(\frac{e^{-2\pi x w i} - 1}{w} \right) \right| &\leq |e^{-2\pi i x z}| \cdot 2\pi |x| e^{2\pi x w} \\ &\leq 2\pi C_\epsilon (|e^{-2\pi i x z + 2\pi x w - 2\pi \epsilon x}| + |e^{-2\pi i x z + 2\pi x w + 2\pi \epsilon x}|). \end{aligned}$$

Now if we assume $|w| \leq \epsilon$, we get that these terms in the upper bound are of the form $2\pi C_\epsilon |e^{-2\pi i x(z+\delta)}|$ where $|\delta| \leq \epsilon + \epsilon \leq 2\epsilon$. Thus they are of the form $|e^{-2\pi i x u}|$ where $u \in U$. Thus by assumption that this integral absolutely converges for $z \in U$, so we have the desired upper bound for $|w| \leq \epsilon$ so the uniform convergence theorem applies. \square

Example 8.7. The Fourier-Laplace transform of a Gaussian $u(x) = Ce^{-\alpha x^2}$, where $C, \alpha > 0$, must be an entire function (holomorphic everywhere) on \mathbb{C} . Let us compute it. Assume for convenience that $u(x) = Ce^{-x^2}$ where $C > 0$ is chosen so that

$$\int_{\mathbb{R}} u(x) dx = 1.$$

Now

$$\begin{aligned} \mathcal{F}u(z) &= \int_{\mathbb{R}} C \exp(-(x^2 + 2\pi i x z)) \\ &= \int_{\mathbb{R}} C \exp(-((x + \pi i z)^2 + \pi^2 z^2)) \\ &= e^{-\pi^2 z^2} \int_{\mathbb{R}} C \exp(-(x + \pi i z)^2) dx. \end{aligned}$$

Since $\mathcal{F}u(z)$ is analytic on \mathbb{C} and the first factor never vanishes, we must have that the second factor is also analytic on \mathbb{C} . That is,

$$\int_{\mathbb{R}} C \exp(-(x + \pi i z)^2) dx$$

is an analytic function on \mathbb{C} . But we know that this function is constant for $z \in i\mathbb{R}$ by translation invariance of the integral in the real domain. Thus by holomorphicity on \mathbb{C} , this integral is constant over all $z \in \mathbb{C}$ and so this integral is equal to 1 (as we assumed it is at $z = 0$). Thus we get that

$$\mathcal{F}u(z) = e^{-\pi^2 z^2}$$

which is another Gaussian. Of course, it is well known $C = \frac{1}{\sqrt{\pi}}$.

8.3. Schwarz space of functions.

Definition 8.8. A function $f : \mathbb{R}^d \rightarrow \mathbb{C}$ is said to be *rapidly decreasing* if for all polynomials $P(x_1, \dots, x_d) \in \mathbb{R}[x_1, \dots, x_d]$ we have that

$$\lim_{x \rightarrow \infty} f(x)P(x) = 0.$$

We say that f is a Schwarz function if f and all partial derivatives of $\partial^\alpha f$, for $\alpha \in \mathbb{Z}_{\geq 0}^d$ are rapidly decreasing. We let $\mathcal{S}(\mathbb{R}^d)$ denote the space of all Schwarz functions on \mathbb{R}^d .

Note that to check rapidly decreasing it is enough to consider P being a monomial.

Lemma 8.9. Let $f : \mathbb{R}^d \rightarrow \mathbb{C}$ be a function that is bounded on every bounded set. Then the following are equivalent:

- (1) f is rapidly decreasing.
- (2) For all integers $N > 0$ exists a constant $C > 0$ such that

$$|f(x)| \leq C\|x\|^{-N} \quad \text{for all } \|x\| > 1.$$

- (3) For all integers $N > 0$ there exists a constant $C > 0$ such that

$$|f(x)| \leq C(\|x\| + 1)^{-N} \quad \text{for all } x \in \mathbb{R}^d.$$

Proof. (i) \implies (ii): If f is rapidly then in particular $\|x\|^{2N}|f(x)| \rightarrow 0$ thus there exists large enough $R > 1$ so that

$$\|x\|^N \leq \|x\|^{2N}|f(x)| \leq 1$$

for all $\|x\| \geq R$. By assumption on the boundedness of f on bounded sets, $|f(x)|/\|x\|^N$ is bounded on the annulus $1 \leq \|x\| \leq R$, so we just take $C > 1$ to be larger than this maximum. (ii) \implies (iii): Note that there is $R > 0$ such that for $\|x\| \geq R$ we have

$$(\|x\| + 1)^{-N} \geq \|x\|^{-(N+1)} \geq \frac{1}{C}|f(x)|$$

for some $C > 0$. One can choose a constant $C' > C$ so that we have the desired inequality $|f(x)| \leq C'(\|x\| + 1)^{-N}$ for $\|x\| \leq R$ by boundedness of f on this ball. (iii) \implies (i): Any monomial $x_{i_1} \dots x_{i_N}$ for degree N can be bounded by $(\|x\| + 1)^{-N}$. \square

Proposition 8.10. If $f, g \in \mathcal{S}(\mathbb{R}^d)$ then $f * g \in \mathcal{S}(\mathbb{R}^d)$. In particular, the convolution is everywhere defined. Moreover, for all $\alpha \in \mathbb{Z}_{\geq 0}^d$ we have that

$$\partial^\alpha(f * g) = (\partial^\alpha f * g).$$

Proof. Fix any polynomial $P : \mathbb{R}^d \rightarrow \mathbb{R}$ of degree N . Now observe that

$$|P(x)(f * g)(x)| \leq |P(x)| \int_{\mathbb{R}^d} |f(y)| |g(x - y)| dy \leq C|P(x)| \int_{\mathbb{R}^d} (1 + \|y\|)^{-M} (1 + \|x - y\|)^{-M} dy$$

for some constant $C > 0$ where $M = \max\{N + 1, d + 2\}$. It is thus enough to show that as $x \rightarrow 0$ this integral behaves like $O(\|x\|^{-N-1})$. Observe that for each fixed $x \in \mathbb{R}^d$ we have $\mathbb{R}^d = A_1 \cup A_2$ where

$$A_1 = \{y \in \mathbb{R}^d \mid \|y - x\| \geq \frac{\|x\|}{2}\}$$

and

$$A_2 = \{y \in \mathbb{R}^d \mid \|y\| \geq \frac{\|x\|}{2}\}.$$

Now observe that for $i = 1$ we have that

$$\begin{aligned} \int_{A_i} (1 + \|y\|)^{-M} (1 + \|x - y\|)^{-M} &\leq \int_{A_i} (1 + \|y\|)^{-M} \left(1 + \frac{\|x\|}{2}\right)^{-M} \\ &= \left(1 + \frac{\|x\|}{2}\right)^{-M} \int_{A_i} (1 + \|y\|)^{-M} \\ &= O(\|x\|^{-M}), \end{aligned}$$

which holds since $M \geq d + 2$ and so the integral in the last line converges. The same inequality holds for $i = 2$ as well by translation symmetry. Now since $M \geq N + 1$ we have shown that

$$|P(x)(f * g)(x)| = O(\|x\|^{-1}) \rightarrow 0.$$

Thus this completes the proof that $f * g$ is rapidly decreasing.

Now let e_j be the j -th basis vector in \mathbb{R}^d and fix $x \in \mathbb{R}^d$. For $t \in \mathbb{R} \setminus \{0\}$ we consider

$$\frac{(f * g)(x + te_j) - (f * g)(x)}{t} = \int_{\mathbb{R}^d} f(y) \left(\frac{g(x - y + te_j) - g(x - y)}{t} \right) dy.$$

By the mean value theorem (only applies to f, g real valued, but from bilinearity of convolution we can assume this) we know that

$$\frac{g(x - y + te_j) - g(x - y)}{t} = (\partial_j g)(x - y + s_{y,t}e)$$

for some $s_{y,t} \in [0, t]$. But $\partial_j g$ is a rapidly decreasing continuous function, thus we obtain that it is uniformly continuous. This means that

$$\frac{g(x - y + te_j) - g(x - y)}{t} \rightarrow (\partial_j g)(x - y) \quad \text{uniformly for as } t \rightarrow 0.$$

As f is integrable, this means that

$$\int_{\mathbb{R}^d} f(y) \left(\frac{g(x - y + te_j) - g(x - y)}{t} \right) dy \rightarrow \int_{\mathbb{R}^d} f(y) (\partial_j g)(x - y) dy = (f * (\partial_j g))(x) \quad \text{as } t \rightarrow 0.$$

Thus we have shown that

$$\partial_j(f * g)(x) = (f * \partial_j g)(x).$$

Since $\partial_j g$ is also a Schwarz function, we can proceed inductively to show that

$$\partial^\alpha(f * g) = (f * (\partial^\alpha g)).$$

and this proves the stated inequality (convolution is commutative). Finally, this equality shows that all partial derivatives of $f * g$ are convolutions of Schwarz functions, and hence rapidly decreasing. This means that indeed $f * g \in \mathcal{S}(\mathbb{R}^d)$. \square

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