NOTES ON DISTRIBUTIONS AND THEIR FOURIER TRANSFORMS

CONTENTS

2

Part 1. **Basic theory**

1. DISTRIBUTIONS

Recall that the *support* of a function $\phi : \mathbb{R}^n \to \mathbb{C}$ is the set

$$\operatorname{supp} \phi = \overline{\{x \in \mathbb{R}^n \mid \phi(x) \neq 0\}} \subset \mathbb{R}^n,$$

i.e., the closure of the set of points where ϕ does not vanish. We say that ϕ is *compactly supported* if the support of ϕ is compact.

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

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

to be the set of *test functions* on U. Given $\phi_1, \phi_2, \ldots \in D(U)$ and $\phi \in D(U)$, then we say that $\phi = \lim_{j \to \infty} \phi_j$ 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

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

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 $\phi : \mathbb{R} \to \mathbb{C}$ given by $\phi(x) = h(x)h(1-x)$ is a smooth map with support $[0, 1]$. Thus $\phi \in D((-\epsilon, 1+\epsilon))$ 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 $\lim_{j \to \infty} \phi_{1/j} = \phi$ in $D(\mathbb{R})$ but the sequence $\phi_1, \phi_2, \ldots$ does not converge because the union of the supports is unbounded, hence not contained in any compact set.

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 $\mu_f(\phi) = \int_{\mathbb{R}^n} \phi(x)f(x)d^n x$ 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

$$\mu_{\partial_n f}(\phi) = \int_{\mathbb{R}^n} \partial_n f(x)\phi(x)d^n x$$

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

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

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

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

which equals $-\int_{\mathbb{R}^n} f(x)\partial_n \phi(x)d^n x$ since $\partial_n \phi = 0$ outside of $[-R, R]^n$. But this is precisely $-\mu_f(\partial_n) = \partial_n \mu_f$. $\qquad \square$

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 $H : \mathbb{R} \to \mathbb{R}$ is discontinuous at 0 thus not differentiable in the classical sense. Yet it has a distribution derivative $H' = \delta$ 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 substitution $u = nx$ we get

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

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 $\alpha \in \mathbb{Z}_{\geq 0}^d$ 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 $\partial^\alpha(\phi_n \psi_n) \to \partial^\alpha(\phi\psi)$ uniformly on K.

Proof. The idea is to apply the product rule and then use the fact that uniform convergence is multiplicative (i.e., that multiplication is continuous with respect to the $\| \cdot \|_\infty$ norm on $C(K)$). We present the details using induction on the order of the partial derivatives.

First we note that

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

converges to 0 uniformly on K. We now prove by induction on k that all order k derivatives of $\phi_n \psi_n$ converge to the corresponding derivatives of $\phi\psi$. Note that we have already shown the $k = 0$ case. Thus suppose $\alpha \in \mathbb{Z}_{\geq 0}^d$ is an index with $k = |\alpha|$ such that

$$\partial^\alpha(\phi_n \psi_n) \to \partial^\alpha(\phi\psi)$$

uniformly on K for any ϕ, ψ satisfying the hypothesis of this Lemma (in particular, the partial derivatives of ϕ, ψ). We wish to show that for $1 \leq j \leq n$ we have that $\partial^\alpha \partial_j(\phi_n \psi_n) \to \partial^\alpha \partial_j(\phi\psi)$, as this will show the induction step. Now use the product rule to see that

$$\partial^\alpha \partial_j(\phi_n \psi_n) = \partial^\alpha(\partial_j(\phi_n)\psi_n) + \partial^\alpha(\phi_n(\partial_j \psi_n)).$$

Now by induction hypothesis for $|\alpha| = k$ applied to $\partial_j \phi_n$ and ψ_n we have that

$$\partial^\alpha(\partial_j(\phi_n)\psi_n) \to \partial^\alpha(\partial_j(\phi)\psi)$$

uniformly on K.

Likewise we get that the same statement when swapping ϕ_n with ψ_n and ϕ with ψ. 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 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 $\phi_n \to \phi$ 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 $v_1, v_2, v \in V$) that is complete.

We define $C_0^k(K)$ to be the C^k functions $\mathbb{R}^n \to \mathbb{C}$ with support inside K (where $k \in \mathbb{Z}_{\geq 0} \cup \{\infty\}$).

For $K \subset \mathbb{R}^n$ compact and $k \in \mathbb{Z}_{\geq 0}$ we define on $C_0^k(K)$ 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 for $k < \infty$. 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 $\alpha \in \mathbb{Z}_{\geq 0}^n$ 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 $\|\cdot\|_{C_k}$).

Remark 2.2. Warning!: Note that $\|\cdot\|_{C_0^\infty(K)}$ is not a norm because it fails the scaling property. In fact, it is bounded by 2. So we cannot say that $C_0^\infty(K)$ is a Banach space (TODO: really is not?).

Theorem 2.3 (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 $k \in \mathbb{Z}_{\geq 0}$ 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 $|f(\psi_n)| = 1$ but for all $n \geq k$ we have

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

So $\psi_n \to 0$ on $C_0^\infty(K)$, which shows that f is not a distribution. $\qquad\square$

Example 2.4. Let $f \in D(\mathbb{R})$ be the distribution $f(\phi) = \phi^{(k)}(0)$, i.e., kth derivative at 0. For any compact set K we have that

$$|f(\phi)| \leq \|\phi\|_{C^k}.$$

However, there is not constant c so that

$$|f(\phi)| \leq c\|\phi\|_{C^{k-1}} \quad \text{for all } \phi \in C_0^\infty(K)$$

if $0 \in K$. To see this, first choose $\phi_1 \in D(\mathbb{R})$ so that $\phi_1^{(k)}(0) = 1$ and support of ϕ_1 is $K = [-1, 1]$. Now set $\phi_\alpha(x) = \alpha^{1-k}\phi_1(\alpha x)$ for $\alpha \geq 1$. Then $\phi_\alpha^{(k)}(x) = \alpha\phi_1(\alpha x)$ and so

$$f(\phi_\alpha) = \alpha \to \infty \text{ as } \alpha \to \infty.$$

On the other hand, for $j \leq k - 1$ we have

$$\|\phi_\alpha^{(j)}\|_\infty = \alpha^{1-k+j}\|\phi_1^{(j)}\|_\infty \leq \|\phi_1^{(j)}\|_\infty$$

is bounded over all $\alpha \geq 1$. Note that all ϕ_α have support inside the support of ϕ_1 (as $\alpha \geq 1$) and hence are all in $C_0^\infty(K)$ for the same compact set K, which justifies this counterexample.

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. By compactness, we can find a finite cover of $U_1, \ldots U_n$ 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 $\psi_1, \ldots, \psi_n \in D(U)$ 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 $f(\phi) = \sum_i f(\phi\psi_i) = 0$. $\square$

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 $\operatorname{supp} f$, to be the smallest closed set such that $f|_V = 0$ where $V = U \setminus \operatorname{supp} f$.

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 is not on any open set 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 $B(a,r) \subset B(a,r') \subset \mathbb{R}^d$ be 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 non-zero smooth $\psi : \mathbb{R} \to [0,1]$ 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. $\square$

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 $\psi_1, \ldots, \psi_n : \mathbb{R}^d \to [0,1]$ 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 $B(a_1, r_1), \ldots, B(a_n, r_n)$ 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 $1 < i \le n$ 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 $x \in K$. $\qquad\square$

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 $U_1, \ldots, U_n \in \mathcal{U}$ 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

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

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 $\psi_i\phi_k \to \psi_i\phi$ (for each fixed i, as $k \to \infty$) 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 any chosen $\psi : \mathbb{R}^d \to [0,1]$ smooth that equals 1 on K and has support inside U (so $\phi\psi = \psi$ everywhere). $\qquad\square$

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[1] $D(X) \subset \mathcal{E}(X)$ 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

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

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 $\phi \in \mathcal{E}(X)$, then $f(\phi_j)$ converges to $f(\phi)$. 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 functional to $D(X)$, then we get an element of $\mathcal{D}'(X)$. Thus we have a map

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

---

[1]Technically, if $\phi \in D(X)$ then $\phi : \mathbb{R}^d \to \mathbb{C}$ by definition, we then restrict ϕ to X to get $\phi in \mathcal{E}(X)$.

Lemma 4.4. If $X \subset \mathbb{R}^d$ is open, then it has an *exhuastion 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 set $K \subset X$ 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 the closed ball of radius n around 0 with distance at least $\frac{1}{n}$ to C, i.e.,

$$K_n = \{x \in X \mid \|x\| \leq n \text{ and } d(x, C) \geq 1/n\}.$$

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 compact 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.

$\square$

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.

$\square$

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 any $g \in \mathcal{E}'(X)$. 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 compact 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 $\phi \in D(X)$ 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 $g(\psi) = f(\psi\phi)$ for $\psi \in \mathcal{E}(X)$. 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 $\psi - \psi\phi$ has support inside $X \setminus K$. 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. $\qquad\square$

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 X.

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 $f \in D'(X)$. Suppose that it does not, thus for any compact set $K \subset X$ we have that $f|_{X \setminus K}$ is not the zero distribution on $X \setminus K$. Thus there exists $\phi_K \in D(X \setminus K)$ such that $f(\phi_K) = 1$. 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 $1 = f(\phi_n) = g(\phi_n) \to 0$ by continuity of g on $\mathcal{E}(X)$. A contradiction. $\qquad\square$

Example 4.9. The distribution $f(\phi) = \int_0^1 \phi(x)dx$ is a compactly supported distribution on $\mathbb{R}$ (the support is $[0, 1]$), 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 $D(X)$.

Now let $\tilde{f} = f|_{(0,1)} \in D'((0,1))$ be the restriction of f to $(0,1)$. Then $\tilde{f} \notin \mathcal{E}'((0,1))$ as $\tilde{f}$ now has support $(0,1)$ in $(0,1)$, which is not compact. Thus the underlying set matters.

Example 4.10. Consider the distribution $f \in D'(X)$ on $X = (0,1)$ 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 $\mathcal{E}'(X)$). Indeed, it is not supported on any compact subset K of $X = (0,1)$. This is despite the support being the bounded subset $(0,1)$ that is closed in $X = (0,1)$, but is not closed in $\mathbb{R}$. One can directly see that there is no continuous extension of f to $\mathcal{E}(X)$ 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 $K_1 \subset K_2 \subset \ldots$ be an exhuastion 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 $f_n \to f$ 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 $\operatorname{supp} \psi_n$, 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. $\qquad\square$

The next technical lemma will be useful for studying convolutions.

Lemma 4.12. Let $v \in D'(\mathbb{R}^d)$ and $\psi : \mathbb{R}^d \to \mathbb{C}$ smooth such that

$$\operatorname{supp}(\psi) \cap \operatorname{supp}(v) = \emptyset.$$

Suppose that at least one of these conditions holds:

- The support of ψ is compact, i.e., $\psi \in D(\mathbb{R}^d)$.
- The support of v is compact, i.e, $v \in \mathcal{E}'(\mathbb{R}^d)$.

Then $v(\psi) = 0$.

Proof. Let $X = \mathbb{R}^d \setminus \operatorname{supp}(v)$. The first case where $\psi \in D(\mathbb{R}^d)$ and v does not necessarily have compact support follows straight from the definition of support of a distribution since we have $v|_X = 0$ and $\psi \in D(X)$ (as its support is inside X by the disjointness assumption) thus $v(\psi) = v|_X(\psi) = 0$.

Now we show the second case where ψ does not necessarily have compact support and $\psi \in \mathcal{E}(\mathbb{R}^d)$. Note that the same argument does not work, as the restriction $v|_X$ may not be in $\mathcal{E}'(X)$ (see Example 4.10). However, we do know that $v(\psi_n) = 0$ for any $\psi_n \in D(X)$, so it is sufficient to construct such ψ_n which converge to ψ in $\mathcal{E}(\mathbb{R}^d)$. To do this, let $K_1 \subset K_2 \subset \ldots$ be an exhaustion of X by compact sets and let $\chi_n \in D(X)$ be such that $\chi_n = 1$ on an open subset of X containing K_n. Let $\psi_n = \chi_n \psi \in D(X)$. We claim that $\psi_n \to \psi$ in $\mathcal{E}'(\mathbb{R}^d)$. To do this, let $K \subset \mathbb{R}^d$ be compact. We must show that $\partial^\alpha \psi_n \to \partial^\alpha \psi$ uniformly on K. However, notice that ψ_n and ψ are all supported inside $\operatorname{supp}(\psi) \subset X$ so we only need to show uniform convergence on $K' = K \cap \operatorname{supp}(\psi) \subset X$, which is compact (as K is compact, support is closed, hence K' is closed and bounded thus compact). Thus by definition of exhaustion by compact sets, there exists n_0 such that $K_{n_0} \supset K'$. However, for $n \geq n_0$ we have $\partial^\alpha \psi_n = \partial^\alpha \psi$ on K' and thus we have uniform convergence on K'. This completes the proof. $\qquad\square$

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 $\mathbb{R}^d \to \mathbb{C}$ that is L^1.

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

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

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. $\qquad\square$

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

$$|(\phi * g)(x + h) - \phi * g(x)| = |\int_{\mathbb{R}^d} (\phi * (x + h - y) - \phi * (x - y))g(y)|$$
$$\leq \sup_{u \in \mathbb{R}^d} |\phi(u + h) - \phi(u)| \|g\|_1 \to 0 \quad \text{as } h \to 0.$$

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[2], 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

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

uniformly for $x \in \mathbb{R}^d$ as $t \to 0$. This implies that

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

as $t \to 0$. $\qquad\square$

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

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

and thus we have uniform convergence. $\qquad\square$

---

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

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 $\phi * \psi$ 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. $\qquad\square$

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 over $x \in \mathbb{R}^d$. 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 for $N > N_0$ we have that

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

Now it follows that

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

$$= \sum_{y \in \frac{1}{N}\mathbb{Z}^d} \frac{1}{N^d}\left(f(x - y)g(y) + E_{y,N}\right)$$

where $|E_{y,N}| < \epsilon$. However, notice that $E_{y,N} = 0$ if $Q_{y,N} \cap \operatorname{supp}(g) = \emptyset$. But there are only $O(N^d)$ many y such that $Q_{y,N} \cap \operatorname{supp}(g) \neq \emptyset$ 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. $\qquad\square$

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

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

Then

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

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

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

and we observe that $\Delta(r) \to 0$ as $r \to 0$ by uniform continuity (g is compactly supported). Now fix $\epsilon > 0$ and choose r so that $\Delta(r) < \epsilon$. In the remainder of the proof, $O(a)$ will refer to a quantity such that $|O(a)| \le |a|$. For all $x \in \mathbb{R}^d$ we have

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

and thus

$$|(f_\delta * g)(x) - g(x)| \le |g(x)| \left| \int_{|y| < r} f_\delta(y) dy - 1 \right| + O\left(\epsilon \int_{|y| < r} f_\delta(y) dy\right) + O\left(\|g\|_\infty \int_{|y| > r} f_\delta(y) dy\right).$$

Now observe that since

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

we have that for each fixed $r > 0$ that

$$\lim_{\delta \to 0} \int_{|y| < r} f_\delta(y) dy = 1$$

and

$$\lim_{\delta \to 0} \int_{|y| > r} f_\delta(y) dy = 0.$$

This shows that

$$\lim_{\delta \to 0} (f_\delta * g)(x) = g(x) + O(\epsilon) \quad \text{uniformly for } x \in \mathbb{R}^d.$$

As ϵ is arbitrary, the proof is complete. $\qquad \square$

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 $g : \mathbb{R} \to \mathbb{R}$ 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

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

where

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

with N large enough.

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

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 products

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

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 $\alpha \in \mathbb{Z}_{\geq 0}^d$ 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 $\alpha \in \mathbb{Z}_{\geq 0}^d$. 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

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

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 $\mathrm{supp}\,\theta$ 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 $(x_i, y) \in \mathrm{supp}\,\theta$ and $(x_i, y) \to (x_0, y)$, thus as the support is closed (by definition) we have that $(x_0, y) \in \mathrm{supp}\,\theta$. Thus we have shown that the support of $x \mapsto \theta(x,y)$ is contained in the projection onto X of the set $\mathrm{supp}(\theta) \cap (X \times \{y\})$, 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 $(x, y) \in \text{supp}(\theta)$ 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 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 2.3 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 $\alpha \in \mathbb{Z}_{\geq 0}^d$ 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.

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. $\qquad\square$

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 $x \in \mathbb{R}^d$ 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

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

We can replace ϕ by any of its partial derivatives and obtain the same uniform convergence for its partial derivatives, namely, that

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

Thus the convergence to $(\partial_i\phi)(x + y)$ occurs in $D(\mathbb{R}^d)$. Thus by the continuity of v we get that

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

Thus we obtain the relation that

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

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. $\qquad\square$

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 $x \in \mathbb{R}^d$. Thus $F_\theta(x) = T_x v(\phi) = v(T_{-x}\phi)$. The above lemma says that $F_\phi \in \mathcal{E}(\mathbb{R}^d) = C^\infty(\mathbb{R}^d)$. We now show that F depends continuously on ϕ in the topology on $\mathcal{E}(\mathbb{R}^d)$.

Lemma 7.3. If $v \in D'(\mathbb{R}^d)$ and $\phi_1, \phi_2 \ldots \in D(\mathbb{R}^d)$ converge to $\phi \in D(\mathbb{R}^d)$ in $D(\mathbb{R}^d)$ then $F_{\phi_i} \to F_\phi$ in $\mathcal{E}(\mathbb{R}^d)$. That is, for each $\alpha \in \mathbb{Z}_{\geq 0}^d$ and each compact $K \subset \mathbb{R}^d$ we have that

$$\partial^\alpha F_{\phi_i}(x) \to \partial^\alpha F_\phi(x) \quad \text{uniformly for } x \in K.$$

Proof. Fix a compact set $K \subset \mathbb{R}^d$ and $\alpha \in \mathbb{Z}_{\geq 0}^d$. By linearity, we only need to show the case $\phi = 0$. Now since ϕ_i converge in $D(\mathbb{R}^d)$, there exists a compact set K_0 such that $\operatorname{supp}\phi_i \subset K_0$ for all ϕ_i. We may find a compact set $K' \supset K$ such that $\operatorname{supp} T_{-x}\phi_i \subset K'$ for $x \in K$ and $i \in \{1, 2, \ldots\}$. It follows that we may now apply Theorem 2.3 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

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

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

$$\sup_{x \in K} \|T_{-x}(\partial^\alpha\phi_i)\|_{C^k} \to 0,$$

as these norms are translation invariant. Thus $v(T_{-x}\partial^\alpha\phi_i) \to 0$ uniformly for $x \in K$. The proof is now complete by noticing that by Lemma 7.1 we have that

$$(\partial^\alpha F_{\phi_i})(x) = F_{\partial^\alpha\phi_i}(x) = v(T_{-x}\partial^\alpha\phi_i)$$

and thus the desired convergence to 0. $\square$

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 u 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 $x \mapsto v(y \mapsto \phi(x+y)) = F_\theta(x)$ 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 $\phi_i \to \phi$ in $D(\mathbb{R}^d)$ then by Lemma 7.3 we have that

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

converges in $\mathcal{E}(\mathbb{R}^d)$ to

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

and thus by the definition of convergence in $\mathcal{E}'(\mathbb{R}^d)$ we have that $(u * v)(\phi_i) \to (u * v)(\phi)$, 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 x, we have that v and $y \mapsto \phi(y + x)$ will have disjoint supports. This means that this inner map is an element of $D(\mathbb{R}^d)$ (it is smooth again by Lemma 7.1 and the fact that $v \in D'(\mathbb{R}^d)$ as $\mathcal{E}'(\mathbb{R}^d) \subset D'(\mathbb{R}^d)$) hence the expression $(u * v)(\phi)$ is well defined in this case. Now we show continuity. Suppose $\phi_i \to 0$ 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 maps

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

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

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

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

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

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

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 pathalogical example arises if

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

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

Now observe that if $f, g : \mathbb{R}^d \to \mathbb{C}$ are L^1 functions then they may be defined naturally as distributions. We now show that their convolution as distributions (assuming at least one has compact support) is the same as their convolution as functions. To see this, let $\phi \in D(\mathbb{R}^d)$, then

$$(f * g)(\phi) = \int_{\mathbb{R}^d} (f * g)(x)\phi(x)dx$$

$$= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(y)g(x-y)\phi(x)dy \right) dx$$

$$= \int_{\mathbb{R}^d} f(y) \left(\int_{\mathbb{R}^d} g(x-y)\phi(x)dx \right) dy$$

$$= \int_{\mathbb{R}^d} f(y) \left(\int_{\mathbb{R}^d} g(x)\phi(x+y)dx \right) dy$$

$$= f(y \mapsto g(x \mapsto \phi(x+y)))$$

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 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 $u \in D'(\mathbb{R}^d)$. The 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

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

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 and using Proposition 5.2 we also get

$$\partial^\alpha(\psi * \phi)(x) = \psi * (\partial^\alpha \phi)(x)$$

$$= \lim_{N \to \infty} \frac{1}{N^d} \sum_{y \in \frac{1}{N^d}\mathbb{Z}^d} \psi(y)(\partial^\alpha \phi)(x-y)$$

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

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

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

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 $SS\psi = \psi$. $\qquad\square$

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 $\phi \in D(\mathbb{R})$ thenby the previous Lemma 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.

$\qquad\square$

7.3. **Applications: approximating a distribution with a smooth function.** We now show that distributions are limits of smooth functions with compact support, i.e., $C_0^\infty(\mathbb{R}^d)$ is dense in $D'(\mathbb{R}^d)$.

Lemma 7.9. Let V be a normed vector space and suppose that $C \subset V$ is closed and $K \subset V$ is compact. Then $C + K$ is closed.

Proof. Suppose that $v \in V$ is in the closure of $C + K$. Thus these exists $c_i \in C$ and $k_i \in K$ so that $c_i + k_i \to v$. After passing to a subsequence, we assume that $k_i \to k$ for some $k \in K$. Thus $\|(c_i + k) - (c_i + k_i)\| \to 0$ and so $c_i + k \to v$. Thus $c_i \to v - k$. As C is closed, this means $v - k \in C$ and so $v \in C + k \subset C + K$ as required. $\qquad\square$

Proposition 7.10. Let $u, v \in D'(\mathbb{R}^d)$ be distributions with at least one of them of compact support. Then

$$\operatorname{supp}(u * v) \subset \operatorname{supp}(u) + \operatorname{supp}(v).$$

Proof. We must show that the open (by previous lemma) set $U = \mathbb{R}^d \setminus (\text{supp}(u) + \text{supp}(v))$ satisfies

$$U \subset \mathbb{R}^d \setminus \text{supp}(u * v).$$

So we must show that if $\phi \in D(\mathbb{R}^d)$ has support inside U then $(u * v)(\phi) = 0$. Suppose for contradiction that

$$0 \neq (u * v)(\phi) = u(x \mapsto v(y \mapsto \phi(x + y))).$$

It now follows by Lemma 4.12 that the support of the inner map $x \mapsto v(y \mapsto \phi(x + y))$ has non-trivial intersection with $\text{supp}(u)$. Thus there exists $x_0 \in \text{supp}(u)$ so that x_0 is in the support of this inner map. Thus there exists $x_1, x_2, \ldots$ converging to x_0 such that $v(y \mapsto \phi(x_i + y)) \neq 0$. Applying Lemma 4.12 now to v we get that $\text{supp}(v)$ has non-trivial intersection with the support of $y \mapsto \phi(x_i + y)$. Thus there exists $y_i \in \text{supp}(v)$ such that $x_i + y_i \in \text{supp}(\phi)$. By compactness of $\text{supp}(\phi)$, we get that $x_i + y_i \to z_0 \in \text{supp}(\phi)$ after passing to a subsequence. Thus $y_i \to y_0 := z_0 - x_0$ with $y_0 \in \text{supp}(v)$ as support is closed by definition. In summary, $x_0 + y_0 = z_0$ with $x_0 \in \text{supp}(u)$, $y_0 \in \text{supp}(v)$ and $z_0 \in \text{supp}(\phi) \subset U$. Thus $U \cap (\text{supp}(u) + \text{supp}(v))$ contains z_0 and is thus non-empty, contradicting the definition of U given above. $\qquad\square$

Lemma 7.11. If $u_n \to u$ in $\mathcal{E}'(\mathbb{R}^d)$ then $u_n * v \to u * v$ in $D'(\mathbb{R}^d)$ for $v \in D'(\mathbb{R}^d)$.

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

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

$$\square$$

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

Theorem 7.12. 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 $\theta_n \to u$ 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 $\psi_n \to \delta$ in $\mathcal{E}'(\mathbb{R}^d)$ and ψ_n all have support inside $[-1, 1]^d$. It now follows that

$$\psi_n * u \to \delta * u = u$$

in $D'(\mathbb{R}^d)$. However we know that $\theta_n = \psi_n * u \in C^\infty(\mathbb{R}^d)$ 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 $\chi_n : \mathbb{R}^d \to [0, 1]$ 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

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

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

Now suppose that u has compact support. By Proposition 7.10, θ_n have support inside the compact set $K = [-1, 1]^d + \text{supp}\, u \supset \text{supp}\, u$. We know that $\theta_n \to u$ 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

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

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

7.4. **Commutativity and associativity.** We know that the convolution of two L^1 functions is commutative and associative. We now address the question for distributions. Using Proposition 7.10 , it now makes sense to talk about associativity.

Theorem 7.13. (Associativity of convolution) Suppose that $u, v, w \in D'(\mathbb{R}^d)$ are such that at least two of them have compact support. Then

$$(u * v) * w = u * (v * w).$$

Proof. Note that by the Proposition 7.10, both expressions are well defined (i.e., for the first expression: if both u, v have compact support then $u * v$ has compact support and so $(u * v) * w$ makes sense. Otherwise, w must have compact support and so $(u * v) * w$ still makes sense). Now we compute

$$\begin{aligned}
((u * v) * w)(\phi) &= (u * v)(x \mapsto w(y \mapsto \phi(x + y))) \\
&= u(x' \mapsto v(y' \mapsto w(y \mapsto \phi(x' + y' + y)))).
\end{aligned}$$

We perform a similar calculation

$$\begin{aligned}
(u * (v * w))(\phi) &= u(x \mapsto (v * w)(y \mapsto \phi(x + y))) \\
&= u(x \mapsto v(x' \mapsto w(y' \mapsto \phi(x + x' + y'))))
\end{aligned}$$

and so we see that these two expressions agree. $\qquad\square$

We first prove commutativity for a special case.

Lemma 7.14. 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. $\qquad\square$

Now by regularization we can deduce commutativity in general.

Theorem 7.15 (Commutativity)**.** Suppose that $u, v \in D'(\mathbb{R}^d)$ and v has compact support. Then

$$u * v = v * u.$$

Proof. Note that by Theorem 7.12 we have compactly supported functions $\psi_n \in C_0^\infty(\mathbb{R}^d)$ with support inside a common compact set K containing 0 such that $\psi_n \to \delta_0$ in $\mathcal{E}'(\mathbb{R}^d)$. Now by Theorem 7.8 we have that $\psi_n * v \in C^\infty(\mathbb{R}^d)$ and in fact $\psi_n * v \in C_0^\infty(\mathbb{R})$ has compact support as both ψ_n and v have compact support. This means that

$$(\psi_n * v) * w = w * (\psi_n * v)$$

and $\psi_n * w = w * \psi_n$ for any $w \in D'(\mathbb{R}^d)$ by the previous Lemma. We may now use associativity and these two identities to see that

$$\begin{aligned}
\psi_n * (u * v) &= (\psi_n * u) * v \\
&= (u * \psi_n) * v \\
&= u * (\psi_n * v) \\
&= (\psi_n * v) * u \\
&= \psi_n * (v * u).
\end{aligned}$$

Now finally the proof is complete by using Lemma 7.11 which says that $\psi_n * w \to \delta_0 * w = w$ for any $w \in D'(\mathbb{R}^d)$. $\qquad\square$

Part 2. **Fourier Transform**

8. Fourier Transform of Schwarz Functions

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 $\mathcal{F}u : \mathbb{R}^d \to \mathbb{C}$ by

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

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

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

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 $\mathcal{F}u : \mathbb{R}^d \to \mathbb{C}$ is uniformly continuous bounded function on $\mathbb{R}^d$.

Proof. We see that

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

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 $\eta \to 0$. But this follows from the dominated convergence theorem as the integrand converges pointwise to 0 as $\eta \to 0$ and the integrand is bounded for all η by the integrable function $2|u|$. $\qquad\square$

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

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

We now show that

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

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

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

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

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

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

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

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

Thus $\mathcal{F}u_j \to \mathcal{F}u$ uniformly. But it follows that if $g_j : \mathbb{R}^d \to \mathbb{C}$ are continuous functions that vanish at ∞, then any uniform limit $g : \mathbb{R}^d \to \mathbb{C}$ 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 $|x| > R$. $\qquad\square$

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 $w \to 0$ 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

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

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

$$\left| e^{-2\pi i x z} \left(\frac{e^{-2\pi x w i} - 1}{w} \right) \right| \leq |e^{-2\pi i x z}| \cdot 2\pi |x| e^{2\pi x w}$$

$$\leq 2\pi C_\epsilon(|e^{-2\pi i x z + 2\pi x w - 2\pi \epsilon x}| + |e^{-2\pi i x z + 2\pi x w + 2\pi \epsilon x}|).$$

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. $\qquad\square$

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

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

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,

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

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 $C = \frac{1}{\sqrt{\pi}}$.

By using Fubini's theorem, we can extend to the multidimensional case, giving

$$\mathcal{F}\left(\pi^{-d/2} e^{-\|x\|^2}\right)(\omega) = e^{-\pi^2 \|\omega\|^2}.$$

8.3. **Schwarz space of functions.**

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

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

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 $f : \mathbb{R}^d \to \mathbb{C}$ 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 $\|x\|^{2N}|f(x)| \to 0$ thus there exists large enough $R > 1$ so that

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

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

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

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

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

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

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

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

for some constant $C > 0$ where $M = \max\{N + 1, d + 2\}$. It is thus enough to show that as $x \to 0$ 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

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

and

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

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

$$\int_{A_i} (1 + \|y\|)^{-M}(1 + \|x - y\|)^{-M} \le \int_{A_i} (1 + \|y\|)^{-M}\left(1 + \frac{\|x\|}{2}\right)^{-M}$$

$$= \left(1 + \frac{\|x\|}{2}\right)^{-M} \int_{A_i} (1 + \|y\|)^{-M}$$

$$= O(\|x\|^{-M}),$$

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

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

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

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

As f is integrable, this means that

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

Thus we have shown that

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

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 $f * g \in \mathcal{S}(\mathbb{R}^d)$. $\qquad \square$

We define the operator M_j on the set of functions $\mathbb{R}^d \to \mathbb{C}$ given by

$$M_j f(x) = x_j f(x) \quad \text{for } x = (x_1, \ldots, x_d) \in \mathbb{R}^d.$$

Clearly the M_j preserve Schwarz space.

Theorem 8.11. If $\phi : \mathbb{R}^d \to \mathbb{C}$ is Schwarz, then so is its Fourier transform $\mathcal{F}\phi : \mathbb{R}^d \to \mathbb{C}$. Moreover, for $\phi \in \mathcal{S}(\mathbb{R}^d)$ we have that

$$\mathcal{F}(\partial_j \phi) = 2\pi i M_j (\mathcal{F}\phi)$$

and

$$\partial_j (\mathcal{F}\phi) = -2\pi i \mathcal{F}(M_j \phi).$$

Proof. First note that these equations will show that the Fourier transform is a Schwarz function as follows. If $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{Z}_{\geq 0}^d$ is an index then by induction and the first equation we get that

$$\mathcal{F}(\partial^\alpha \phi)(\omega) = (2\pi i)^{\|\alpha\|_1} \omega^\alpha (\mathcal{F}\phi)(\omega).$$

But by the Riemann-Lebesgue lemma, we have that the right hand side converges to 0 and $\omega \to \infty$. Thus we have shown that

$$\lim_{\omega \to \infty} |P(\omega)|(\mathcal{F}(\omega)) = 0$$

for all polynomials P. Thus the Fourier transform will be shown to be rapidly decreasing by the first identity. The second identity then shows that every element in the space $\mathcal{F}(\mathcal{S}(\mathbb{R}^d))$ has a continuous partial derivative and that this space is invariant under partial differentiation, which shows that the Fourier transform of a Schwarz function is smooth. Thus this will show that the Fourier transform of a Schwarz function is Schwarz.

The first identity is an easy application of inegration by parts as follows:

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

$$= -\int_{\mathbb{R}^d} \phi(x) \partial_j (e^{-2\pi i \langle x, \omega \rangle}) dx$$

$$= 2\pi i \omega_j \int_{\mathbb{R}^d} \mathcal{F}(\phi)(\omega)$$

$$= 2\pi i M_j (\mathcal{F}(\phi))(\omega).$$

For the second identity, we apply differentiation under the integral sign (Proposition A.1) as follows:

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

$$= -2\pi i \mathcal{F}(M_j \phi)(\omega)$$

where the differentation under the integral sign is justified since

$$|-2\pi i x_j \phi(x) e^{-2\pi i \langle x, \omega \rangle}| = 2\pi |x_j \phi(x)|$$

and as ϕ is rapidly decreasing and measurable, we have that $x \mapsto x_j \phi(x)$ is integrable. Thus the integrand is bounded by an integrable function that is independent of ω. $\qquad \square$

8.4. Fourier inversion of Schwarz functions.

Theorem 8.12 (Multiplication formula). *If $\phi, \psi \in S(\mathbb{R}^d)$ then*

$$\int_{\mathbb{R}^d} \phi(x) \cdot \mathcal{F}\psi(x) d(x) = \int_{\mathbb{R}^d} \mathcal{F}\phi(x) \cdot \psi(x) dx.$$

Proof. We may apply Fubini's theorem to show that the left hand side is equal to

$$\int_{\mathbb{R}^d} \phi(x) \left(\int_{\mathbb{R}^d} \psi(y) e^{-2\pi i \langle x, y \rangle} dy \right) dx = \int_{\mathbb{R}^d \times \mathbb{R}^d} \phi(x) \psi(y) e^{-2\pi i \langle x, y \rangle} dx dy$$

where the use of Fubini's theorem is justified as $|\phi(x)\phi(y)|$ is in $L^1(\mathbb{R}^d \times \mathbb{R}^d)$. Now this expression is symmetric in ϕ and ψ. $\qquad \square$

Theorem 8.13 (Fourier inversion). *For $\phi \in S(\mathbb{R}^d)$ we have that*

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

Proof. Let

$$K(x) = Ce^{-\|x\|^2}$$

where $C > 0$ is chosen so that

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

We know from Example 8.7 that

$$\mathcal{F}K(\omega) = e^{-\pi^2 \|\omega\|^2}$$

is also a Gaussian. Now for $\alpha > 0$ let

$$K_\alpha(x) = K(\alpha x).$$

Thus we know that

$$\mathcal{F}(K_\alpha)(\omega) = \alpha^{-d} \mathcal{F}(K)(\alpha^{-1}\omega) = \alpha^{-d} e^{-\pi^2 \|\alpha^{-1}\omega\|^2}.$$

Now we apply the multiplication formula to ϕ and K_α to get

$$\int_{\mathbb{R}^d} \phi(x) \cdot \mathcal{F} K_\alpha(x) d(x) = \int_{\mathbb{R}^d} \mathcal{F} \phi(x) \cdot K_\alpha(x) dx.$$

Now as $\mathcal{F} \phi$ is Schwarz and thus integrable, we can apply the dominated converge theorem to show that as $\alpha \to 0$, the right hand side converges to

$$C \int_{\mathbb{R}^d} \mathcal{F} \phi(\omega) d\omega.$$

On the other hand, the left hand side converges to

$$\phi(0) \int_{\mathbb{R}^d} \mathcal{F} K(\omega) d\omega$$

since for all $\alpha > 0$ we have

$$\int_{\mathbb{R}^d} \mathcal{F} K_\alpha(\omega) d\omega = \int_{\mathbb{R}^d} \mathcal{F} K(\omega) d\omega$$

but for all $\epsilon > 0$ we have that

$$\int_{|x| > \epsilon} K_\alpha(x) dx \to 1 \quad \text{as } \alpha \to 0.$$

Now we compute this integral as

$$\int_{\mathbb{R}^d} e^{-\pi^2 x^2} dx = \pi^{-d} \int_{\mathbb{R}^d} e^{-x^2} dx = \pi^{-d} C^{-1}.$$

This means that

$$\pi^{-d} C^{-1} \phi(0) = C \int_{\mathbb{R}^d} \mathcal{F} \phi(\omega) d\omega.$$

As $C = \pi^{-d/2}$ this means

$$\phi(0) = \int_{\mathbb{R}^d} \mathcal{F} \phi(\omega) d\omega,$$

which is the $x = 0$ case of the Theorem. To derive the general case, we may consider the Schwarz function $\phi_y(x) = \phi(y + x)$ and apply the already proven $x = 0$ case to obtain that

$$\phi(y) = \phi_y(0) = \int_{\mathbb{R}^d} \mathcal{F} \phi_y(\omega) d\omega.$$

But

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

as desired.

$\square$

Corollary 8.14. The Fourier Transform is a bijection $\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$ with inverse given by

$$\mathcal{F}^{-1} = \mathcal{F} \circ S = S \circ \mathcal{F}$$

where $S\phi(x) = \phi(-x)$.

Proof. The Fourier Inversion formula gives that

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

This shows that

$$\mathcal{F} \circ S \circ \mathcal{F} = \mathrm{Id}_{\mathcal{S}(\mathbb{R}^d)}.$$

Thus if we let $\mathcal{G} = S \circ \mathcal{F}$ and $\mathcal{G}' = \mathcal{F} \circ S$ then we see that $\mathcal{G}$ is a right inverse and $\mathcal{G}'$ is a left inverse of $\mathcal{F}$. But we must have $\mathcal{G} = \mathcal{G}'$ by simple algebra (i.e., we have $\mathcal{F}\mathcal{G} = 1$ and $\mathcal{G}'\mathcal{F} = 1$, now multiplying by $\mathcal{G}$ on both sides of of the second equation we get $\mathcal{G}'\mathcal{F}\mathcal{G} = \mathcal{G}$ and then using the first equation we get $\mathcal{G}' = \mathcal{G}$). $\square$

9. Tempered Distributions and their Fourier Transform

9.1. **Topology on Schwarz space.**

Definition 9.1. We say that $\phi_1, \phi_2, \ldots \in \mathcal{S}(\mathbb{R}^d)$ converge in $\mathcal{S}(\mathbb{R}^d)$ to some $\phi \in \mathcal{S}(\mathbb{R}^d)$ if for all all indices $\alpha, \beta \in \mathbb{Z}_{\geq 0}^d$ we have that

$$x^\beta \partial^\alpha \phi_i(x) \to x^\beta \partial^\alpha \phi(x) \quad \text{uniformly on } \mathbb{R}^d.$$

We observe that if $\phi_i \to \phi$ in $D(\mathbb{R}^d)$ then they also converge in $\mathcal{S}(\mathbb{R}^d)$. Thus we have a continuous embedding (preservers limits)

$$D(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d).$$

However if one considers the functions $\phi_1, \phi_2, \ldots \in D(\mathbb{R}^d)$ given by $\phi_n(x) = \phi_1(x-n)$ (a gliding hump) then although in $C^\infty(\mathbb{R}^d)$ this sequence of functions converges to 0 (i.e., all its partial derivatatives converge to 0 uniformly on compact sets) but not in $\mathcal{S}(\mathbb{R}^d)$ as the convergence is not uniform on all of $\mathbb{R}^d$.

Example 9.2 (Multiplcation by x^β is not reduntant)**.** Let $\phi \in D(\mathbb{R})$ be any compactly supported non-zero smooth function. Let $\phi_n = 2^{-n}\phi(x-n)$. Say that support of ϕ is in $[0,1]$, thus the support of ϕ_n is $[n, n+1]$. Then $\phi_n \to 0$ in $\mathcal{S}(\mathbb{R})$, which we justify as follows. Observe that

$$\sup_{x \in \mathbb{R}^d} |x^\beta \partial^\alpha \phi_n(x)| \leq 2^{-n} n^{|\beta|} \sup_{x \in [0,1]} |\partial^\alpha \phi(x)|$$

and this converges to 0 uniformly since exponentials dominate polynomials. On the other hand, if we define $\psi_n(x) = \frac{1}{n}\phi(x-n)$ then ψ_n does not converge to 0 in $\mathcal{S}(\mathbb{R})$, even though all partial derivatives converge uniformly on $\mathbb{R}$ to 0. This is because

$$\sup_{x \in \mathbb{R}^d} |x^2 \psi_n(x)| \geq n^2 n^{-1} \sup_{x \in \mathbb{R}^d} |\phi(x)| = n \sup_{x \in \mathbb{R}^d} |\phi(x)|$$

which does not converge to 0 as the lower bound is linear in n.

Definition 9.3 (Norms on Schwartz space)**.** It will be convenient to define for $\phi \in \mathcal{S}(\mathbb{R}^d)$ the norm

$$\|\phi\|_{\mathcal{S}(k,N)} = \sup_{|\alpha| \leq k, |\beta| \leq N, x \in \mathbb{R}^d} |x^\beta \partial^\alpha \phi|.$$

Thus we have that $\phi_i \to \phi$ in $\mathcal{S}(\mathbb{R}^d)$ if and only if $\|\phi_i - \phi\|_{\mathcal{S}(k,N)} \to 0$ for all non-negative integers k, N.

Proposition 9.4. If $\phi \in \mathcal{S}(\mathbb{R}^d)$ then there are $\phi_i \in \mathcal{S}(\mathbb{R}^d)$ that are compactly supported and $\phi_i \to \phi$ in $\mathcal{S}(\mathbb{R}^d)$.

Proof. Let $\psi : \mathbb{R}^d \to \mathbb{R}$ be a smooth function with compact support such that $\psi(x) = 1$ for all $\|x\| \leq 1$. Now let

$$\phi_n(x) = \psi(x/n)\phi(x).$$

Clearly $\phi_n \in \mathcal{S}(\mathbb{R}^d)$ is compactly supported. To show that $\phi_n \to \phi$ in $\mathcal{S}(\mathbb{R}^d)$, we will need to estimate the norms above as follows. First use the product rule to write

$$x^\beta \partial^\alpha \left((\psi(x/n) - 1)\phi(x) \right) = x^\beta (\psi(x/n) - 1)\partial^\alpha \phi(x) + \sum_{0 \neq \gamma \leq \alpha} O(n^{-|\gamma|}) x^\beta \partial^{\alpha-\gamma} \phi(x)$$

where the implicit O constant depends on γ (and thus can be chosen so it depends on α only). Thus we now wish to show that both terms convergence uniformly as $n \to \infty$. The first term is bounded by

$$\|\psi\|_\infty \sup_{\|x\| > n} x^\beta |\partial^\alpha \phi(x)|$$

and this converges to 0 uniformly as $n \to \infty$ by definition of Schwarz function. The second term clearly converges to 0 as $x^\beta \partial^{\alpha - \gamma} \phi(x)$ is bounded on $\mathbb{R}^d$. $\square$

We say that an operator $\mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$ is continuous if it preserves limits in $\mathcal{S}(\mathbb{R}^d)$.

Lemma 9.5. The operators $\phi \mapsto \partial^\alpha \phi$ and $\phi \mapsto x^\beta \phi$ are continuous functions on $\mathcal{S}(\mathbb{R}^d)$.

Proof. The claim about partial derivatives is obvious. For the second claim, assume that $\phi_n \to 0$ in $\mathcal{S}(\mathbb{R}^d)$. Then

$$x^{\beta'} \partial^\alpha (x^\beta \phi_n(x))$$

is a finite sum of terms of the form $P(x) \partial^\gamma \phi_n(x)$ for some polynomials P and indices γ. Thus by definition of Schwarz functions, we have that such terms converge to 0 in $\mathcal{S}(\mathbb{R}^d)$. $\square$

Proposition 9.6. The Fourier transform $\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$ is continuous. So is its inverse.

Proof. Suppose that $\phi_n \to 0$ in $\mathcal{S}(\mathbb{R}^d)$. We first show that

$$\|\mathcal{F}(\phi_n)\|_\infty \to 0.$$

To see this, note that

$$|\mathcal{F}(\phi_n)(\omega)| \leq \int_{\mathbb{R}^d} |\phi_n(x)| dx$$

so it is enough to show that this upper bound, which is independent of ω, converges to 0. Letting $P(x) = (1 + \|x\|^2)^{d+1}$ we have that

$$\int_{\mathbb{R}^d} |\phi_n(x)| dx = \int_{\mathbb{R}^d} \frac{1}{P(x)} |\phi_n(x) P(x)| dx \leq \sup_{x \in \mathbb{R}^d} |\phi_n(x) P(x)| \int_{\mathbb{R}^d} \frac{1}{P(x)} dx.$$

On the other hand

$$\sup_{x \in \mathbb{R}^d} |\phi_n(x) P(x)| \to 0 \quad \text{uniformly for } x \in \mathbb{R}^d.$$

Thus we have shown that $\|\mathcal{F}(\phi_n)\|_\infty \to 0$ whenever $\phi_n \to 0$ in $\mathcal{S}(\mathbb{R}^d)$. It now remains to show that

$$\lim_{n \to \infty} \|x^\beta \partial^\alpha \mathcal{F}(\phi_n)\|_\infty = 0.$$

But

$$x^\beta \partial^\alpha \mathcal{F}(\phi_n) = (2\pi i)^{|\alpha|} (-2\pi i)^{-|\beta|} \mathcal{F}(\partial^\beta (x^\alpha \phi_n)).$$

By Lemma 9.5 $\partial^\beta x^\beta \phi_n \to 0$ in $\mathcal{S}(\mathbb{R}^d)$ and so by what we have already shown (the $\alpha = \beta = 0$ case) we see that indeed

$$\|\mathcal{F}(\partial^\beta (x^\alpha \phi_n))\|_\infty \to 0.$$

Thus the Fourier transform is continuous. Moreover the inverse of the fourier transform is $\mathcal{F}^{-1} = \mathcal{F} \circ S$ where S is clearly continuous in $\mathcal{S}(\mathbb{R}^d)$. Thus the $\mathcal{F}^{-1}$ is continuous. $\square$

9.2. **Tempered distributions.**

Definition 9.7. A *tempered distribution* is a linear functional $u : \mathcal{S}(\mathbb{R}^d) \to \mathbb{C}$ that is continuous in the sense that if $\phi_n \to \phi$ in $\mathcal{S}(\mathbb{R}^d)$ then $u(\phi_n) \to u(\phi)$. We let $\mathcal{S}'(\mathbb{R}^d)$ denote the space of tempered distributions. We say that a sequence of tempered distributions $u_1, u_2, \ldots$ converges to a tempered distribution u *in* $\mathcal{S}'(\mathbb{R}^d)$ if for all $\phi \in \mathcal{S}(\mathbb{R}^d)$ we have that $u_j(\phi) \to u(\phi)$.

We now give some examples and non-examples.

Proposition 9.8.

- Every tempered distribution $u \in \mathcal{S}'(\mathbb{R}^d)$ restricts to a distribution in $D'(\mathbb{R}^d)$. This restriction mapping is injective and continuous.
- Every compactly supported distribution $u \in \mathcal{E}'(\mathbb{R}^d)$ restricts to a tempered distribution. This restriction mapping is injective and continuous.

Proof. First suppose that $u \in \mathcal{S}'(\mathbb{R}^d)$ and $\phi_n \to \phi$ in $D(\mathbb{R}^d)$. This means there is a compact set K containing the support of all ϕ_n and ϕ and for each index α we have $\partial^\alpha \phi_n \to \partial^\alpha \phi$ uniformly on K and thus uniformly on $\mathbb{R}^d$. So we have $\phi_n \to \phi$ in $\mathcal{S}(\mathbb{R}^d)$ and so indeed $u(\phi_n) \to u(\phi)$. Thus the restriction of u to $D(\mathbb{R}^d)$ gives a well defined continuous distribution $u \in D'(\mathbb{R}^d)$. Clearly this restriction mapping is continuous since the definition of convergence of the distributions is the same. To show injectivity, note that by Proposition 9.4 any tempered distribution u is determined by its values on comapctly supported smooth functions. Thus if $u_1, u_2 \in \mathcal{S}'(\mathbb{R}^d)$ restrict to the same element in $D'(\mathbb{R}^d)$, then they must be equal.

Now suppose $u \in \mathcal{E}'(\mathbb{R}^d)$ and $\phi_n \to \phi$ in $\mathcal{S}(\mathbb{R}^d)$. Then in particular we may view $\phi_n, \phi \in C^\infty(\mathbb{R}^d)$, thus u restricts to a well defined linear map functional on $\mathcal{S}(\mathbb{R}^d)$. Now $\phi_n \to \phi$ uniformly on $\mathbb{R}^d$ and so it clearly converges in $C^\infty(\mathbb{R}^d)$ (which is uniform convergence on compact subsets). Continuity is clear as the definitions of convergence coincide. To show injectivity, suppose $u_1, u_2 \in \mathcal{E}'(\mathbb{R}^d)$ restrict to the same element of $\mathcal{S}'(\mathbb{R}^d)$. Thus u_1, u_2 restrict to the same element of $D'(\mathbb{R}^d)$. By Lemma 4.5 that this means that $u_1 = u_2$ are the same element of $\mathcal{E}'(\mathbb{R}^d)$.

$\square$

In summary, this means that we have natural inclusions of distributions

$$\mathcal{E}'(\mathbb{R}^d) \hookrightarrow \mathcal{S}'(\mathbb{R}^d) \hookrightarrow D'(\mathbb{R}^d).$$

which are continuous in the sense that they preserve limits (i.e., if $u_j \to u$ then it does so when embedded in the larger space). We now show that these inclusions are proper.

Example 9.9 (A non-tempered distribution)**.** Consider the distribution $u \in D'(\mathbb{R})$ given by

$$u(\phi) = \int e^x \phi(x) dx.$$

We claim that there is no tempered distribution that restricts to this distribution. Suppose there was such a distribution u. Now choose a compactly supported non-negative and non-zero smooth function $\phi \in D(\mathbb{R})$. Let us assume it is supported on $[0, 1]$. We already know that

$$\phi_n(x) = e^{-n/2} \phi(x - n)$$

is a sequence of compactly supported Schwarz functions that converge to 0 in $\mathcal{S}(\mathbb{R}^d)$ (see Example 9.2). However

$$u(\phi_n) = \int_n^{n+1} e^{x-n/2}\phi(x-n) \geq e^{\frac{n}{2}} \int_0^1 \phi(x) \to \infty \neq 0 = u(0),$$

thus u is not continuous in $\mathcal{S}(\mathbb{R}^d)$.

We now demonstrate many tempered distributions that are not compactly supported.

Proposition 9.10. Let μ be a finite Borel measure on $\mathbb{R}^d$ (i.e., $\mu = \mu_1 - \mu_2$ where μ_1, μ_2 are finite positive measure), then $\mu \in \mathcal{S}'(\mathbb{R}^d)$ is a tempered distribution.

Proof. As Schwarz functions are bounded,

$$\mu(\phi) := \int_{\mathbb{R}^d} \phi(x)d\mu(x)$$

is well defined linear functional on $\mathcal{S}(\mathbb{R}^d)$. If $\phi_n \to \phi$ in $\mathcal{S}(\mathbb{R}^d)$ then in particular $\phi_n \to \phi$ uniformly, thus $\mu(\phi_n) \to \mu(\phi)$ and so μ is continuous on $\mathcal{S}(\mathbb{R}^d)$. $\square$

In particular, we get that $L^1(\mathbb{R}^d) \subset \mathcal{S}'(\mathbb{R}^d)$ as $L^1(\mathbb{R}^d)$ can be identified with those finite measures that are absolutely continuous (have a density with respect to the Lebesgue measure). In fact, we can do better.

Proposition 9.11. For $1 \leq p \leq \infty$ we have that any $u \in L^p(\mathbb{R}^d)$ gives a tempered distribution, i.e., the distribution

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

Proof. We first consider the $p = \infty$ case. Note that for $\phi \in \mathcal{S}(\mathbb{R}^d)$ we have that $|u(x)\phi(x)| = \|u\|_\infty |\phi(x)|$ is integrable, thus $u(\phi)$ is a well defined linear functional on $\mathcal{S}(\mathbb{R}^d)$. Now suppose that $\phi_n(x) \to 0$ in $\mathcal{S}(\mathbb{R}^d)$. Letting $P(x) = |1 + \|x\|^2|^{d+2}$ we see that

$$|u(\phi_n)| \leq \|u\|_\infty \|P\phi_n\|_\infty \int_{\mathbb{R}^d} (P(x))^{-1}dx.$$

But as $\phi_n \to 0$ in $\mathcal{S}(\mathbb{R}^d)$ we have that the term $\|P\phi_n\|_\infty \to 0$ while the other terms are finite.

Now it remains to demonstrate this result for the case $1 < p < \infty$. For this, we let $P_D(x) = (1 + \|x\|^2)^D$ for integers $D > 0$ Let q by such that $p^{-1} + q^{-1} = 1$. Now using Holder's inequality we get that

$$|u(\phi_n)| = \|u\phi_n\|_1 \leq \|u\|_p \|\phi_n\|_q$$

and so we just need to show that $\|\phi_n\|_q \to 0$. To see this note that

$$\|\phi_n\|_q^q = \int_{\mathbb{R}^d} |\phi_n(x)|^{1/q}dx \leq \|P_D(x)\phi_n^{1/q}\|_\infty \int_{\mathbb{R}^d} P_D(x)dx.$$

Now $P_{qD}(x)\phi_n \to 0$ uniformly (as $\phi_n \to 0$ in Schwartz space) thus $P_D(x)\phi_n^{1/q} \to 0$ uniformly for $x \in \mathbb{R}^d$. Finally, if we choose D large enough so that $\int_{\mathbb{R}^d} P_D(x)dx < \infty$ then we are done. Note that this proof also showed that $|u(\phi)| < \infty$ for any Schwarz function ϕ. $\square$

9.3. **Fourier transform of a tempered distribution.**

Definition 9.12. Let $u \in \mathcal{S}'(\mathbb{R}^d)$ be a tempered distribution. Then define the Fourier transform $\mathcal{F}u \in \mathcal{S}'(\mathbb{R}^d)$ by

$$(\mathcal{F}u)(\phi) = u(\mathcal{F}\phi) \quad \text{for } \phi \in \mathcal{S}(\mathbb{R}^d)$$

Let us check that indeed this defines a tempered distribution, i.e., we need to show that $\mathcal{F}u$ is well defined and continuous. To show that it is well defined, note that by Theorem 8.11 we have that if $\phi \in \mathcal{S}(\mathbb{R}^d)$ then $\mathcal{F}\phi$ is indeed a Schwartz function, thus $u(\mathcal{F}\phi)$ make sense. Continuity follows from Proposition 9.6, which says that if $\phi_i \to \phi$ in $\mathcal{S}(\mathbb{R}^d)$ then $(\mathcal{F}u)(\phi_i) \to (\mathcal{F}u)(\phi)$ and so by continuity of u we indeed that that $\mathcal{F}u$ is continuous.

Let us now show that the Fourier transform of a tempered distribution coincides with previous definitions of the Fourier transform defined on various subspaces of $\mathcal{S}'(\mathbb{R}^d)$. First, if μ is a finite Borel measure on $\mathbb{R}^d$ then, viewed as a distribution, we have that

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

But now since $(x, \omega) \mapsto \phi(x)$ is an integrable function with respect to the Lebesgue measure times μ, Fubini's theorem applies and we can interchange the order of integration. We thus get that

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

Thus $\mathcal{F}\mu$ is a distribution given by the function

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

If we swap the names of the variables ω and x, we get that

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

which makes obvious the fact that this definition coincides with the previous definition given for finite Borel measures. Moreover, it also shows that if $u \in L^1(\mathbb{R}^d)$ then $\mathcal{F}u$ corresponds to the original definition of a Fourier transform of an integrable functions given in Definition 8.1 (i.e., because $\phi \mapsto \int u(x)\phi(x)dx$ is a finite Borel measure).

Example 9.13. Let $u(x) = 1$ by the constant function. It is not integrable and definitely not Schwarz, thus we have not been able to take its Fourier transform. We now can if we view it as a tempered distribution (which we can by Proposition 9.11 since $u \in L^\infty(\mathbb{R}^d)$). For Schwarz functions ϕ we have that

$$(\mathcal{F}u)(\phi) = u(\mathcal{F}\phi) = \int_{\mathbb{R}^d} (\mathcal{F}\phi)(\omega)d\omega = \phi(0)$$

where in the last equality we used the Fourier inversion formula for Schwarz functions given in Theorem **??**.

We can extend Fourier inversion to tempered distributions. We can be a bit more general. Suppose that $\mathcal{G} : S(\mathbb{R}^d) \to S(\mathbb{R}^d)$ is a continuous linear operator (for example, $\mathcal{F}$ of $\mathcal{F}^{-1}$). Then we can define a dual operator $\mathcal{G}^* : S'(\mathbb{R}^d) \to S'(\mathbb{R}^d)$, defined for $u \in \mathcal{S}'(\mathbb{R}^d)$ by

$$(\mathcal{G}^*u)(\phi) = u(\mathcal{G}\phi) \quad \text{for } \phi \in \mathcal{S}(\mathbb{R}^d).$$

So the fourier transform on $\mathcal{S}'(\mathbb{R}^d)$ is just the dual of the fourier transform on $\mathcal{S}(\mathbb{R}^d)$, but we drop the * for convenience.

Note that $(\mathcal{G}_1 \mathcal{G}_2)^* = \mathcal{G}_2^* \mathcal{G}_1^*$ and $\mathrm{Id}_{\mathcal{S}(\mathbb{R}^d)}^* = \mathrm{Id}_{\mathcal{S}'(\mathbb{R}^d)}$. Thus if $\mathcal{G}$ is invertible then $(\mathcal{G}^{-1})^* = (\mathcal{G}^*)^{-1}$. This shows that the Fourier Inversion formula naturally extends to tempered distributions as follows.

Theorem 9.14. The Fourier transform $\mathcal{F} : \mathcal{S}'(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$ is a continuous bijection with inverse given by $\mathcal{F}^{-1} : S'(\mathbb{R}^d) \to S'(\mathbb{R}^d)$ given by

$$(\mathcal{F}^{-1} u)(\phi) = u(\mathcal{F}^{-1}\phi) \quad \text{for } \phi \in \mathcal{S}(\mathbb{R}^d).$$

Example 9.15. We already saw that $\mathcal{F}1 = \delta$ as distributions. Thus we have $\mathcal{F}^{-1}\delta = 1$. But $\mathcal{F}^{-1} = \mathcal{F}S^*$ where $S\phi(x) = \phi(-x)$. Now

$$(S^*\delta)(\phi) = \delta(S\phi) = (S\phi)(0) = \phi(0) = \delta(\phi).$$

Thus δ is symmetric, i.e., invariant number $x \mapsto -x$. Hence $\mathcal{F}\delta = \mathcal{F}^{-1}\delta = 1$.

In summary, Fourier Transform maps 1 to δ and vice versa.

Now let $e_\omega : \mathbb{R}^d \to \mathbb{C}$ be given by

$$e_\omega(x) = e^{2\pi i \langle \omega, x \rangle}$$

and let us compute its Fourier Transform. We have for $\phi \in \mathcal{S}(\mathbb{R}^d)$ that

$$(\mathcal{F}e_{\omega_0})\phi = \int_{\mathbb{R}^d} e_{\omega_0}(\omega)(\mathcal{F}\phi)(\omega)d\omega = \int_{\mathbb{R}^d} e_\omega(\omega_0)(\mathcal{F}\phi)(\omega)d\omega = \phi(\omega_0)$$

where in the last equality we have used the Fourier Inversion formula. Thus it follows that

$$\mathcal{F}e_{\omega_0} = \delta_{\omega_0}$$

where δ_{ω_0} is the point mass at ω. More generally, this means that if we have a trigonometric polynomial

$$f = \sum_{i=1}^{n} c_i e_{\omega_i}$$

then we have

$$\mathcal{F}f = \sum_{i=1}^{n} c_i \delta_{\omega_i}.$$

In other words Fourier transform consists of point masses at frequencies with weights given by the coefficients of this trigonometric polynomial. In poetic terms, the Fourier Transform gives the *frequency distribution*, which is the main philosophy of Fourier Analysis.

9.4. **Fourier transform on $L^2(\mathbb{R}^d)$.** By observing $u \in L^2(\mathbb{R}^d)$ as a tempered distribution we may take its Fourier transform, which we now show is also in $L^2(\mathbb{R}^d)$.

First observe that we have a natural inner product on Schwartz space given by

$$\langle \phi, \psi \rangle = \int_{\mathbb{R}^d} \phi(x) \cdot \overline{\psi(x)} d(x).$$

Theorem 9.16. The fourier transform $\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$ is an isometry for this inner product. That is, for $\phi, \psi \in \mathcal{S}(\mathbb{R}^d)$ we have that

$$\langle \phi, \psi \rangle = \langle \mathcal{F}\phi, \mathcal{F}\psi \rangle.$$

Proof. We only need to show the $\phi = \psi$ case. Note that the multiplication formula can be written as

$$\langle \mathcal{F}\phi, \overline{\psi} \rangle = \langle \phi, \overline{\mathcal{F}\psi} \rangle.$$

Now we let $\psi = \overline{\mathcal{F}\phi}$ to obtain that

$$\langle \mathcal{F}\phi, \mathcal{F}\phi \rangle = \langle \phi, \overline{\mathcal{F}\psi} \rangle.$$

Observe that

$$\overline{\mathcal{F}(\psi)(\omega)} = \int_{\mathbb{R}^d} \overline{\psi(x)} e^{2\pi i \omega \cdot x} dx = \mathcal{F}^{-1}\overline{\psi}(\omega)$$

by the Fourier inversion formula (with x and ω swapped). In other words

$$\overline{\mathcal{F}(\psi)} = \mathcal{F}^{-1}(\overline{\psi}) = \mathcal{F}^{-1}\mathcal{F}\phi = \phi$$

which completes the proof. $\qquad\square$

Theorem 9.17. If $u \in L^2(\mathbb{R}^d)$ is considered as a tempered distribution, then $\mathcal{F}u \in L^2(\mathbb{R}^d)$ and $\mathcal{F}$ thus becomes an isometry on $L^2(\mathbb{R}^d)$.

Proof. We first note that $\mathcal{S}(\mathbb{R}^d)$ is L^2-dense in the complete metric space $L^2(\mathbb{R}^d)$ and thus $\mathcal{F}$ extends uniquely to an isometry $\widetilde{\mathcal{F}} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$. Now Let $u \in L^2(\mathbb{R}^d)$ and choose $u_j \in \mathcal{S}(\mathbb{R}^d)$ so that $u_j \to u$ in $L^2(\mathbb{R}^d)$. Thus $\mathcal{F}u_j \to \widetilde{\mathcal{F}}u$ in $L^2(\mathbb{R}^d)$. By Cauchy-Schwartz we have that $u_j(\phi) \to u(\phi)$ for all Schwartz functions ϕ. For any $\phi \in \mathcal{S}(\mathbb{R}^d)$ we now have that

$$\mathcal{F}u(\phi) = u(\mathcal{F}\phi) = \lim_{j \to \infty} u_j(\mathcal{F}\phi) = \lim_{j \to \infty} \mathcal{F}u_j(\phi) = \widetilde{\mathcal{F}}u(\phi).$$

This shows that $\mathcal{F}u = \widetilde{\mathcal{F}}u$ and thus $\mathcal{F}u \in L^2(\mathbb{R}^d)$. $\qquad\square$

Part 3. **Sobolev Spaces**

10. Sobolev spaces definition

Definition 10.1. For integers $k \geq 0$ and $\Omega \subset \mathbb{R}^d$ open, we define the *Sobolev space* $H^k(\Omega)$ to be the space of those $u \in L^2(\Omega)$ such that whenever $|\alpha| \leq k$ then the distribution $D^\alpha u$ is an element of $L^2(\Omega)$.

We define an inner product on $H^k(\Omega)$ given by

$$\langle u, v \rangle_{H_k(\Omega)} = \sum_{0 \leq |\alpha| \leq k} \langle D^\alpha u, D^\alpha v \rangle.$$

Lemma 10.2. The inner product $H_k(\Omega)$ is a Hilbert space (metric is complete).

Proof. We may isometrically embed $H_k(\Omega)$ into the direct sum of finitely many Hilbert spaces $\oplus_\alpha L^2(\Omega)$ via the mapping

$$\iota : u \mapsto \oplus_\alpha (D^\alpha u).$$

It thus remains to show that the image of this embedding is closed. To see this, suppose that $u_1, u_2, \ldots \in H^k(\Omega)$ is such that $v_\alpha := \lim_{i \to \infty} D^\alpha u_i$ exists in $L^2(\Omega)$ for all $0 \leq |\alpha| \leq k$. Then $u_i \to v_0 \in L^2(\Omega)$ and thus $u_i \to v_0$ as a distribution as well. It follows that the distribution $D^\alpha u_i$ converges to the distribution $D^\alpha v_0$. But we know that $D^\alpha u_i$ converges to v_α in L^2 and thus also the convergence takes place as a distribution. This means that $v_\alpha = D^\alpha v_0$ as distributions and thus also as element of L^2. This means that $\iota(u_1), \iota(u_2), \ldots$ converges to $\iota(v_0)$ as desired. $\qquad\square$

Example 10.3. Let $I = (-1, 1)$ be an open interval. Note that the function $\mathbb{1}_I$ is in $H^k(I)$ for all $k \geq 0$ as all its derivatives are 0. But it is not in $H^k(\mathbb{R})$ for $k \geq 1$ since its derivative is a signed measure $\delta_{-1} - \delta_1$ supported on $\{-1, 1\}$. So the underlying domain matters. The function $u(x) = \mathbb{1}_I |x|$ is in $H^1(-1, 1)$ even though it is not differentiable. It is not in $H^2(-1, 1)$ as the second derivatice is $2\delta_0$.

Example 10.4. Let $B = B(0, 1) \subset \mathbb{R}^d$ be the open ball of radius 1 around $0 \in \mathbb{R}^d$. Fix $\alpha > 0$ and let $u : B \to \mathbb{R} \cup \{\infty\}$ be the map given by $u(x) = \|x\|^{-\alpha}$. We ask now for which $\alpha > 0$ we have that $u \in H^1(B)$. Note if S_r is a sphere of radius r then $\int_{S_r} u(x)^2 dS = Cr^{-2\alpha + d - 1}$ for some constant C. Thus by the coarea formula (integrate of S_r then over r) we have that

$$\int_B u(x)^2 dx = \int_0^1 Cr^{-2\alpha + d - 1} dr$$

and this is finite if $-2\alpha + d > 0$. So for $\alpha < \frac{d}{2}$ we have that $u \in L^2(B)$.

We now want $\partial_j u \in L^2(B)$. For this, we can estimate for non-zero $x = (x_1, \ldots, x_n)$ that

$$|\partial_j u(x)| = |\partial_j (\sum x_i^2)^{-\alpha/2}| \leq C|x_j|(\sum x_i^2)^{-\alpha/2 - 1} = C|x_j| \|x\|^{-\alpha - 2} \leq C\|x\|^{-\alpha - 1}$$

thus it is sufficient to have $\alpha + 1 < \frac{d}{2}$.

It remains to now compute the weak partial derivative $\partial_j u$ and show it coincides with the actual function $\partial_j u$. Let $\phi \in D(B(0, 1))$ be a test function. Then

$$\partial_j u(\phi) = -\int_B u(x)(\partial_j \phi)(x) dx$$
$$= -\int_B \partial_j (u\phi)(x) dx + \int_B (\partial_j u)(x)\phi(x) dx.$$

Note that by dominated convergence theorem we have that

$$\int_B \partial_j (u\phi)(x) dx = \lim_{r \to 0} \int_{B_r} \partial_j (u\phi)(x) dx$$

where $B_r = \{x \in B \mid \|x\| \geq r\}$. We can compute this using the divergence theorem (applied to the vector field $u(x)\phi(x)e_j$ where e_j is the j-th basis vector for $\mathbb{R}^d$) to get that

$$\left| \int_{B_r} \partial_j (u\phi)(x) dx \right| = \left| \int_{|x| = r} u\phi(x)e_j \cdot \hat{n} dS \right| \leq Cr^{-\alpha} r^{d - 1} \|\phi\|_\infty.$$

Now notice that this converges to 0 as $r \to 0$ provided that $\alpha < d - 1$, which holds as $\alpha < d/2 - 1$. Thus we have that the weak derivative of u is the usual derivative $\partial_j u$ (undefined at 0).

In summary, this means that $u \in H^1(B)$ if $\alpha < \frac{d}{2} - 1$. For example, the Newton potential in $\mathbb{R}^3$ given by $u(x) = \|x\|^{-1}$ is in $H^1(B)$.

Example 10.5. Letting $B = B(0, 1)$, we can give an example of $u \in H^1(B)$ such that u is unbounded on any non-empty open subset of B, thus at no point differentiable despite having a weak derivative in $L^2(B)$. First, choose $0 < \alpha < \frac{d}{2} - 1$ as in the previous example. Hence $|x - q_k|^{-\alpha} \in H^1(B)$ for each point $q_k \in B(0, 1)$. Thus if $q_1, q_2, \ldots$ is a dense subset of $B(0, 1)$ then

$$u(x) = \sum_{k=1}^\infty \frac{1}{2^k} |x - q_k|^{-\alpha} \in H^1(B).$$

However, u is unbounded on any open subset as $\lim_{x \to q_k} u(x) = \infty$ for all k.

Note that if $\Omega = \mathbb{R}^d$, we may replace distribution with tempered distribution in the above definition. This is convenient since we may then take the Fourier transform of an element in $H^k(\mathbb{R}^d)$.

11. Sobolev subspace H_0^k

Definition 11.1. For an open set U let $H_0^k(U)$ denote the closure of $D(U) = C_c^\infty(U)$ in $H^k(U)$.

Example 11.2. Let us show that the closed subspace $H_0^k(U)$ can be a proper subspace of $H_k(U)$. Consider $U = (0, 1)$ and $k = 1$. We wish to construct a smooth $u \in H^1(U)$ so that $u \neq 0$ and $\langle u, \phi \rangle = 0$ for all $\phi \in D(U)$, as it will then follow that $\langle u, v \rangle = 0$ for all $v \in H_0^1(U)$ and hence $u \notin H_0^1(U)$. To do this, note that

$$\langle u, \overline{\phi} \rangle = \int_0^1 u\phi \, dx + \int_0^1 u'\phi' \, dx$$

and by the identity $(u'\phi)' = u''\phi + u'\phi'$ together with the fact that $\phi(0) = \phi(1) = 0$ we have

$$\langle u, \overline{\phi} \rangle = \int_0^1 (u - u'')\phi \, dx.$$

So it suffices to select a smooth function so that $u = u''$.

12. Fourier Transform characterization of Sobolev Space

We now seek to classify $H^k(\mathbb{R}^d)$ in terms of the Fourier transform. Recall that if u is a tempered distribution, then so is $M_j u$ where $M_j u(\phi) := u(M_j \phi)$ where $M_j \phi(x) = x_j \phi(x)$ is the multiplication operator. We define $M^\alpha = M_1^{\alpha_1} \cdots M_d^{\alpha_d}$ for $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{Z}_{\geq 0}^d$.

Lemma 12.1. Let u be a tempered distribution. Then

$$\mathcal{F}(D^\alpha u) = (2\pi i)^{|\alpha|} M_j^\alpha \mathcal{F}u.$$

Corollary 12.2. Let $u \in L^2(\mathbb{R}^d)$. Then $D^\alpha u \in L^2(\mathbb{R}^d)$ if and only if

$$\int_{\mathbb{R}^d} |\mathcal{F}(\omega)|^2 |\omega^\alpha|^2 \, d\omega < \infty.$$

Proof. First suppose $D^\alpha u \in L^2(\mathbb{R}^d)$. This means that $\mathcal{F}(D^\alpha u) \in L^2(\mathbb{R}^d)$. Then we have by the lemma above that $M_j^\alpha \mathcal{F}u \in L^2(\mathbb{R}^d)$ and thus, by considering this tempered distribution as a function, we see that it is square-integrable, as desired. The converse follows from also applying this lemma and using the fact that the Fourier transform is a bijection on $L^2(\mathbb{R}^d)$. $\square$

Corollary 12.3. Let $u \in L^2(\mathbb{R}^d)$. Then $u \in H^k(\mathbb{R}^d)$ if and only if

$$\int_{\mathbb{R}^d} |\mathcal{F}u(\omega)|^2 (1 + \|\omega\|)^{2k} \, d\omega < \infty.$$

We let $L^2(\mathbb{R}^d, \mu)$ denote the L^2 space for a measure μ on $\mathbb{R}^d$. We let $L^2(\mathbb{R}^d, f(\omega)d\omega)$ denote the L^2 space of the measure μ given by

$$\int \phi(\omega) d\mu(\omega) = \int \phi(\omega) f(\omega) d\omega.$$

Proposition 12.4. The Fourier Transform provides a linear homeomorphism

$$\mathcal{F} : H^k(\mathbb{R}^d) \to L_k^2(\mathbb{R}^d) := L^2(\mathbb{R}^d, (1 + \|\omega\|)^{2k} d\omega).$$

Proof. It is a linear isomorphism by the above corollary. Let us now show continuity of this map and its inverse. For $u \in H^k(\mathbb{R}^d)$ we have that

$$
\begin{aligned}
\|u\|_{H_k}^2 &= \sum_\alpha \|D^\alpha u\|_{L^2}^2 \\
&= \sum_\alpha \|\mathcal{F} D^\alpha u\|_{L^2}^2 \\
&= \sum_\alpha (2\pi)^{|\alpha|} \|M^\alpha \mathcal{F} u\|_{L^2}^2 \\
&= \int |\mathcal{F} u(\omega)|^2 \sum_\alpha 2\pi^{|\alpha|} |\omega^{2\alpha}| d\omega
\end{aligned}
$$

where we always sum over those α such that $0 \le |\alpha| \le 2k$.

So it is enough to show that $\sum_\alpha |\omega^{2\alpha}|$ and $(1 + \|\omega\|)^{2k}$ are Lipschitz equivalent. This can be seen from the estimate

$$
|\omega^{2\alpha}| \le 1 + |\omega|^{2\alpha} \le (1 + \|\omega\|)^{2\alpha} \le (1 + \|\omega\|)^{2k}
$$

as well as from an expansion of $(1 + \|\omega\|)^{2k}$ into monomials. $\qquad \square$

Part 4. **Appendices**

APPENDIX A. DIFFERENTIATION UNDER THE INTEGRAL SIGN

Proposition A.1. Let X be a measurable space with a finite signed measure (a different of finite positive measures) μ on X. Suppose $U \subset \mathbb{R}$ is open and $K : X \times U \to \mathbb{C}$ is measurable such that

(1) For each $x \in X$, the maps $K_x : U \to \mathbb{C}$ given by $K_x(t) = K(x, t)$, are differentiable in on U.
(2) There exists a μ-integrable function $g : X \to [0, \infty)$ (i.e., $\int_X g(x) d\mu(x) < \infty$) such that $|K_x'(t)| \le g(x)$ for all $t \in U, x \in X$.
(3) For each fixed $t \in U$, the map $x \mapsto K(x, t)$ is μ-integrable.

Then the map $F : U \to \mathbb{C}$ given by

$$
F(t) = \int_X K(x, t) d\mu(x)
$$

defines a differentiable function with derivative equal to

$$
F'(t) = \int_X K_x'(t) d\mu(x) = \int_X \frac{\partial}{\partial s} |_{s=t} K(x, s) d\mu(x)
$$

Proof. By considering the real and imaginary parts of $K(x, t)$, we can reduce to the case where K is a real valued function. Indeed, if $K(x, t) = K_1(x, t) + i K_2(x, t)$ where K_1, K_2 are real valued then K_1, K_2 are clearly differentiable in t and $|K_1(x, t)|, |K_2(x, t)| \le |K(x, t)|$ thus the real and imaginary parts are integrable. Likewise, the derivatives of the real and imaginary parts are at most $|K_x'(t)|$ and so are at most g. So all the assumptions are preserved for the real and imaginary parts.

Now to $t \in U$ fixed and sufficiently small h we have that

$$
\frac{F(t+h) - F(t)}{h} = \int_X \frac{K_x(t+h) - K_x(t)}{h} d\mu(x).
$$

We wish to show that as $h \to 0$, both the left hand side and right hand side converge to the same limit, which will complete the proof. To justifty this we use Lebesgue's Dominated Convergence theorem, which

may be justified since by the mean value theorem, there exists for each $x \in X$ and h sufficiently small some $t_{h,x} \in [t, t+h]$ such that

$$\frac{K_x(t+h) - K_x(t)}{h} = K'_x(t_{h,x}).$$

But by assumption $|K'_x(t_{h,x})| \leq g(x)$ and so the integrand is bounded, for each such h, by the integrable function g. Thus by Lebesgue's dominated convergence theorem, we have the desired limiting behaviour. $\quad\square$

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