Language of Tempered Distribution

Hsin-Jui, Chen

June 11, 2023

Contents

1 Rapidly Decreasing Functions

We shall introduce some notations.

Notation 1.

1. Let $C^\infty(\mathbb{R}^n)$ be the set of all smooth functions.

2. Let $C_0^\infty(\mathbb{R}^n)$ be the set of all smooth functions with compact support. (c is the subscript denote "compact".)

We now could give the definition of rapidly decreasing functions.

Definition 1 (Rapidly decreasing functions)**.** Let $f \in C^\infty(\mathbb{R}^n)$. We said f is rapidly decreasing if

$$\sup_{x\in\mathbb{R}^n} \left| \left(\prod_{j=1}^{n} x_j^{\beta_j} \right) D^\alpha f(x) \right| < \infty \tag{1}$$

holds for any nonnegative integers n-tuples $\alpha = (\alpha_1, \dots, \alpha_n)$ and $\beta = (\beta_1, \dots, \beta_n)$. In (1), $D^\alpha f(x)$ means

$$\frac{\partial^{|\alpha|}}{\partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n}} f(x), \text{ where } |\alpha| = \sum_{i=1}^{n} \alpha_i.$$

We often write $f \in \mathfrak{S}(\mathbb{R}^n)$. There are also some books calling $\mathfrak{S}(\mathbb{R}^n)$ the Schwartz space and denote it by $\mathcal{S}$.

I now give the motivation why we need to study the Schwartz space $\mathfrak{S}(\mathbb{R}^n)$. Later, we will find out that Fourier transform on $\mathfrak{S}(\mathbb{R}^n)$ is an automorphism. Hence the problem in Fourier transform may be converted into the automorphism between the dual space of Schwartz space. Here are some propositions of the space $\mathfrak{S}(\mathbb{R}^n)$.

Proposition 1. *$\mathfrak{S}(\mathbb{R}^n)$ is vector space with the the standard addition and scalar multiplication. It is also a topological space, with the topology defined by the semi-norms of the form*

$$p(f) := \sup_{x\in\mathbb{R}^n} \left| P(x) D^\alpha f(x) \right|,$$

where α is a fixed n-tuple and $P(x)$ is a fixed non-zero polynomial. Together with the topological structure and linear structure, it is locally convex.

We sometimes write $\|f\|_{P,\alpha}$ for $p(f)$. In particular, if $P(x) = x^\beta$, then we simply write $\|f\|_{\beta,\alpha} = p(f)$. We shall emphasize that the topology is defined by the system of those semi-norms, not only one of them.

Proposition 2. *$\mathfrak{S}(\mathbb{R}^n)$ is closed under the linear partial operations of the form*

$$P(x) \cdot D^\alpha$$

where P is a polynomial in $\mathbb{R}$.

Proposition 3. *$C_0^\infty(\mathbb{R}^n) \subset \mathfrak{S}(\mathbb{R}^n)$. Moreover, $C_0^\infty(\mathbb{R}^n)$ is dense in $\mathfrak{S}(\mathbb{R}^n)$ with respect to the topology of $\mathfrak{S}(\mathbb{R}^n)$.*

This proposition is not quite trivial, we shall give a proof.

Proof. Let $f \in \mathfrak{S}(\mathbb{R}^n)$ and take $\psi \in C_0^\infty(\mathbb{R}^n)$ such that $\psi(x) = 1$ when $|x| \leq 1$. Then for any $\epsilon > 0$, $f_\epsilon(x) := f(x) \cdot \psi(\epsilon x) \in C_0^\infty(\mathbb{R}^n)$. By applying the Leibniz' rule, we see that

$$D^\alpha(f_\epsilon(x) - f(x)) = D^\alpha\Big(f(x)(\psi(\epsilon x) - 1)\Big)$$

is a finite linear combination of terms of the form

$$D^\beta f(x) \cdot \epsilon^{|\gamma|} \cdot \left(D^\gamma \psi(y)\right)\Big|_{y=\epsilon x},$$

where α, β, and γ are three n-tuples such that $|\beta| + |\gamma| = |\alpha|$ and $|\gamma| > 0$, and the term $D^\alpha f(x) \cdot (\psi(\epsilon x) - 1)$. Thus it is clear that given any n-tuple α, we have

$$\lim_{\epsilon \to 0} D^\beta f(x) \cdot \epsilon^{|\gamma|} \cdot \left(D^\gamma \psi(y)\right)\Big|_{y=\epsilon x} = 0.$$

We conclude that

$$\lim_{\epsilon \to 0} D^\alpha(f_\epsilon(x) - f(x)) = \lim_{\epsilon \to 0}(D^\alpha f(x)) \cdot (\psi(\epsilon x) - 1) = 0.$$

This implies that $\|f_\epsilon - f\|_{1,\alpha} \to 0$ where $\|\cdot\|_{1,\alpha}$ is the semi-norm defined by the polynomial $P(x) = 1$ and the n-tuple α. □

Proposition 4. $\mathfrak{S}(\mathbb{R}^n) \subset L^p(\mathbb{R}^n)$ *for all* $1 \leq p \leq \infty$. *In fact,* $\mathfrak{S}(\mathbb{R}^n)$ *is dense in* $L^p(\mathbb{R}^n)$ *with respect to the* L^p *norm.* (*Note: We will not prove the second assertion here.*)

Proof. Let $f \in \mathfrak{S}(\mathbb{R}^n)$. Then the theorem follows from

$$\begin{aligned}
\|f\|_p^p &= \int_{\|x\|_1 \leq 1} |f(x)|^p \, dx + \int_{\|x\|_1 > 1} |f(x)|^p \, dx \\
&\leq \left(\sup_{x \in \mathbb{R}^n} |f(x)|\right)^p \cdot 2^n + \left(\sup_{x \in \mathbb{R}^n} (|x|^N |f(x)|)^p \int_{|x|_1 > 1} \frac{1}{|x|^{Np}}\right) \\
&< C_1 + C_2 < \infty.
\end{aligned}$$

C_1 and C_2 are two constant, whose existence follows from (1). □

Definition 2 (Fourier transform and inverse Fourier transform)**.** For any $f \in \mathfrak{S}(\mathbb{R}^n)$ define its Fourier transform $\widehat{f}$ by

$$\widehat{f}(\xi) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-i\langle \xi, x \rangle} f(x) dx, \tag{2}$$

and the inverse Fourier transform $\tilde{g}$ of $g \in \mathfrak{S}(\mathbb{R}^n)$ by

$$\tilde{g}(x) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{i\langle x, \xi \rangle} g(\xi) d\xi. \tag{3}$$

In the above notations, $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$ and $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. $\langle -, - \rangle$ denotes the standard inner product on $\mathbb{R}^n$.

Proposition 5. *The Fourier transform* $f \mapsto \widehat{f}$ *maps* $\mathfrak{S}(\mathbb{R}^n)$ *linearly and continuously into* $\mathfrak{S}(\mathbb{R}^n)$. *The inverse Fourier transform* $g \mapsto \tilde{g}$ *also maps* $\mathfrak{S}(\mathbb{R}^n)$ *linearly and continuously into* $\mathfrak{S}(\mathbb{R}^n)$.

Proof. It is clear that the Fourier transform and the inverse Fourier transform are both linear. It suffices to show that Fourier transform is continuous. Recall that for any continuously differentiable function $\phi(x, y) \in C^1(\mathbb{R}^{2n})$, we have

$$\frac{\partial}{\partial y_i} \int_{\mathbb{R}^n} \phi(x, y) dx = \int_{\mathbb{R}^n} \frac{\partial}{\partial y_i} \phi(x, y) dx.$$

The smoothness of f gives us

$$D^\alpha \widehat{f}(\xi) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-i\langle \xi, x \rangle} (-i)^{|\alpha|} x^\alpha f(x) dx.$$

Proposition 4 implies that $D^\alpha \widehat{f}(\xi)$ exists and hence $\widehat{f} \in C^\infty(\mathbb{R}^n)$. By integration by parts, we have

$$(i)^{|\beta|} \xi^\beta \widehat{f}(\xi) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-i\langle \xi, x \rangle} D^\beta f(x) dx.$$

We conclude

$$(i)^{|\alpha|+|\beta|} \xi^\beta D^\alpha \widehat{f}(\xi) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-i\langle \xi, x \rangle} D^\beta (x^\alpha f(x)) dx.$$

Since the Schwartz space is a topological vector space, it suffices to show the Fourier transform is continuous at 0. Note that $D^\beta(x^\alpha f(x))$ is a linear combination of the forms

$$x^\kappa \cdot D^\lambda f(x),$$

where κ and λ are two n-tuples. For convenient, we may write

$$D^\beta(x^\alpha f(x)) = \sum_{j=1}^{m} c_j x^{\kappa_j} \cdot D^{\lambda_j} f(x).$$

Let $\epsilon > 0$ be given, then let

$$\delta := \frac{\epsilon}{m} \cdot \left(\int_{\mathbb{R}^n} \prod_{k=1}^{n} \left(\frac{1}{1 + x_j^2} \right) e^{-i\langle \xi, x \rangle} dx \right)^{-1}.$$

Consider the finite intersection B of open balls in the Schwartz space defined by

$$B := \bigcap_{j=1}^{m} B_{P_j, \lambda_j}(0; \delta), \qquad \text{where} \qquad P_j(x) := c_j \cdot \left(\prod_{k=1}^{n} (1 + x_j^2) \right) \cdot x^{\kappa_j}.$$

Then it is clear that for any $f \in B$, we have

$$\|\widehat{f}\|_{\beta,\alpha} < (2\pi)^{-n/2} \cdot \epsilon.$$

This proves the Fourier transform is continuous at 0. □

Theorem 3 (Fourier's integral theorem)**.** *The inverse Fourier transform is the inverse mapping of the Fourier transform. In other words, we have*

$$\tilde{\widehat{f}} = (2\pi)^{-n/2} \int e^{i\langle x, \xi \rangle} \widehat{f}(\xi) d\xi = f(x)$$

and similarly, $\widehat{\tilde{f}} = f$.

Therefore together with Proposition 5, we see that the Fourier transform is an automorphism on the Schwartz space.

Proof. We first note that

$$\begin{aligned}\int g(\xi)\widehat{f}(\xi)e^{i\langle x,\xi\rangle}d\xi &= \int g(\xi)\left((2\pi)^{-n/2}\int f(y)e^{-i\langle \xi,y\rangle}dy\right)e^{i\langle x,\xi\rangle}d\xi \\ &= (2\pi)^{-n/2}\int\left(\int g(\xi)e^{-i\langle \xi,y-x\rangle}d\xi\right)f(y)dy \\ &= \int \widehat{g}(y-x)f(y)dy = \int \widehat{g}(y)f(x+y)dy.\end{aligned} \tag{4}$$

If we replace $g(\xi)$ with $g(\epsilon\xi)$ $(\epsilon > 0)$, then

$$(2\pi)^{-n/2}\int e^{-i\langle y,\xi\rangle}g(\epsilon\xi)d\xi = (2\pi)^{-n/2}\epsilon^{-n}\int e^{-i\langle y,x/\epsilon\rangle}g(z)dz = \epsilon^{-n}\widehat{g}(y/\epsilon).$$

By (4), we obtain that

$$\int g(\epsilon\xi)\widehat{f}(\xi)e^{i\langle x,\xi\rangle}d\xi = \int \epsilon^{-n}\widehat{g}(y/\epsilon)f(x+y)dy = \int \widehat{g}(y)f(x+\epsilon y)dy.$$

We now take $g(x) = \exp(-|x|^2/2)$ and let $\epsilon \to 0$. By Lebesgue's dominated convergence theorem, we have

$$g(0)\int \widehat{f}(\xi)e^{i\langle x,\xi\rangle}d\xi = f(x)\int \widehat{g}(y)dy.$$

This proves the theorem since $g(0) = 1$ and $\int \widehat{g}(y)dy = (2\pi)^{n/2}$ by the well-known facts:

$$(2\pi)^{-n/2}\int \exp(-|x|^2/2)e^{-i\langle y,x\rangle}dx = \exp(-|y|^2/2),$$
$$(2\pi)^{-n/2}\int \exp(-|x|^2/2)dx = 1.$$

The first one has been proved before. This proves the theorem. □

Corollary. We have

$$\int \widehat{f}(\xi)g(\xi)d\xi = \int f(x)\widehat{g}(x)dx, \tag{5}$$

$$\int f(\xi)\overline{g}(\xi)d\xi = \int \tilde{f}(x)\overline{\tilde{g}}(x)dx, \tag{6}$$

$$\widehat{(f*g)} = (2\pi)^{n/2}\widehat{f}\cdot\widehat{g} \qquad \text{and} \qquad (2\pi)^{n/2}\widehat{(f\cdot g)} = \widehat{f}*\widehat{g}. \tag{7}$$

Recall that $*$ is the convolution, more precisely, $f*g(x)$ is defined as

$$f*g(x) = \int f(x-y)g(y)dy = \int g(x-y)f(y)dy = g*f(x).$$

Proof. (5) can be obtained by plugging in $x = 0$ to (4). For (6), it suffices to show that $\widehat{\overline{g}} = \overline{\tilde{g}}$, which is true by

$$\widehat{\overline{g}}(\xi) = \int e^{-i\langle \xi,x\rangle}\overline{g}(x)dx = \overline{\int e^{i\langle x,\xi\rangle}g(x)dx} = \overline{\tilde{g}}(\xi).$$

We now show (7).

$$\begin{aligned}&(2\pi)^{-n/2}\int (f*g)(x)e^{-i\langle \xi,x\rangle}dx \\ &= (2\pi)^{-n/2}\int g(y)e^{-i\langle \xi,y\rangle}\left(\int f(x-y)e^{-i\langle \xi,x-y\rangle}dx\right)dy \\ &= (2\pi)^{n/2}\widehat{f}(\xi)\widehat{g}(\xi).\end{aligned}$$

Since the product of two rapidly decreasing function is also rapidly decreasing (Leibniz rule). It is also clear that $f * g \in \mathfrak{S}(\mathbb{R}^n)$ if both f and g belong to $\mathfrak{S}(\mathbb{R}^n)$. This shows the first statement of (7). □

Theorem 4 (Poisson's Summation Formula). *Let $\phi \in \mathfrak{S}(\mathbb{R}^1)$ and let $\widehat{\phi} \in \mathfrak{S}(\mathbb{R}^1)$ be its Fourier transform. Then we have*

$$\sum_{n=-\infty}^{\infty} \phi(2\pi n) = \sum_{n=-\infty}^{\infty} \widehat{\phi}(n). \tag{8}$$

Proof. Set $f(x) = \sum_{n=-\infty}^{\infty} \phi(x + 2\pi n)$. The series is absolutely convergent since

$$|\phi(x)| \le C/x^2, \qquad \text{for some } C \in \mathbb{R}^+.$$

Similar argument also holds for the series

$$\sum_{n=-\infty}^{\infty} \widehat{\phi}(n).$$

Hence both sides of (8) are absolutely convergence. Also it is clear that $f(x + 2\pi) = f(x)$. We compute the Fourier coefficient.

$$\begin{aligned} c_k(f) = (2\pi)^{-1/2} \int_0^{2\pi} f(x) e^{-ikx} dx &= \sum_{n=-\infty}^{\infty} (2\pi)^{-1/2} \int_0^{2\pi} \phi(x + 2\pi n) e^{-ikx} dx \\ &= \sum_{n=-\infty}^{\infty} (2\pi)^{-1/2} \int_{2\pi n}^{2\pi(n+1)} \phi(x) e^{-ikx} dx = \widehat{\phi}(k). \end{aligned}$$

Since $f \in L^2(0, 2\pi)$, we know that its Fourier series converge to itself in the L^2 norm. That is

$$\sum_{k=-s}^{s} \widehat{\phi}(k) e^{ikx} \xrightarrow{L^2} f(x),$$

as $s \to \infty$. However it is clear that the convergence on the right hand side is uniform by the Weierstrass M-test, hence the limit function is continuous. This shows that

$$g(x) := \sum_{k=-\infty}^{\infty} \widehat{\phi}(k) e^{ikx}$$

is a continuous function of x and

$$\|g - f\|_{L^2} = 0$$

and hence $f = g$. We obtain that

$$\sum_{n=-\infty}^{\infty} \phi(x + 2\pi n) = \sum_{k=-\infty}^{\infty} \widehat{\phi}(k) e^{ikx}.$$

Plug in $x = 0$ gives the desired result. □

2 The Fourier Transform of Tempered Distributions

In this section, we shall define a concept called tempered distribution. We later will see that we can define a Fourier transform on a tempered distribution.

Definition 5 (tempered distribution on $\mathbb{R}^n$)**.** A bounded (continuous) linear function T on $\mathfrak{S}(\mathbb{R}^n)$ is called a tempered distribution in $\mathbb{R}^n$. The set of all tempered distributions is denoted by $\mathfrak{S}(\mathbb{R}^n)'$.

We can consider the strong topology on $\mathfrak{S}(\mathbb{R}^n)'$.

Definition 6 (strong topology)**.** The strong dual topology is a topology defined by the uniform convergence on some bounded sets. More precisely, let $\mathcal{A}$ be the collection of all bounded sets on $\mathfrak{S}(\mathbb{R}^n)$. Then each set $A \subset \mathcal{A}$ define a semi-norm

$$\left\|\phi(x)\right\|_A = \sup_{x \in A} \left|\phi(x)\right|.$$

The topology generated by all semi-norms of this form is called the strong topology on $\mathfrak{S}(\mathbb{R}^n)$. However we did not define the concept of boundedness of a set $A \subset \mathfrak{S}(\mathbb{R}^n)$, we may understand as A is bounded if and only if $p(A)$ is bounded for each semi-norm defined on $\mathfrak{S}(\mathbb{R}^n)$.

The strong dual topology on $\mathfrak{S}(\mathbb{R}^n)$ makes $\mathfrak{S}(\mathbb{R}^n)'$ an locally convex linear topological space.

Proposition 6. *Recall that C_0^∞ is all the smooth functions that have compact support. We have $C_0^\infty \subset \mathfrak{S}(\mathbb{R}^n) \subset C^\infty$ as abstract set. It is worth noting that the topology of $C_0^\infty(\mathbb{R}^n)$ is stronger than that of $\mathfrak{S}(\mathbb{R}^n)$ thus*

$$\mathfrak{S}(\mathbb{R}^n)' \subset \mathfrak{D}(\mathbb{R}^n)'.$$

Similarly, we have

$$\mathfrak{E}(\mathbb{R}^n)' \subset \mathfrak{S}(\mathbb{R}^n)',$$

where $\mathfrak{E}(\mathbb{R}^n)'$ denotes the (strong) dual space of $\mathfrak{E}$.

Since C_0^∞ is dense in $\mathfrak{S}(\mathbb{R}^n)$, any tempered distribution $T \in \mathfrak{S}(\mathbb{R}^n)'$ restrict to C_0^∞ is also continuous and linear. For more information about the topology of $\mathfrak{D}(\mathbb{R}^n)$ and $\mathfrak{E}(\mathbb{R}^n)$, please refer to the appendix.

Example 7. *For any $f \in L^p(\mathbb{R}^n)$ $(p \geq 1)$, it defines a tempered distribution*

$$T_f(\phi) = \int_{\mathbb{R}^n} \phi(x) f(x) dx.$$

This example might help us understand some concepts later, it help us to define the derivative of a tempered distribution.

We shall now define the concept of slowly increasing function.

Definition 8 (slowly increasing function)**.** A function $f \in C^\infty(\mathbb{R}^n)$ is called slowly increasing at ∞ if for any non-negative integer n-tuple j, there is a non-negative integer N such that

$$\lim_{|x| \to \infty} |x|^{-N} \left|D^j f(x)\right| = 0.$$

The set of all slowly increasing function is denoted by $\mathfrak{O}_M(\mathbb{R}^n)$.

We may do similar things that we have done when we discussed rapidly decreasing functions. We can easily see that $\mathfrak{O}_M(\mathbb{R}^n)$ is a vector space by the function sum and scalar multiplication. It is also a topological space by the topology defined by the system of semi-norms of the form

$$p(f) = p_{h,j}(f) = \sup_{x\in\mathbb{R}^n} \left|h(x)D^j f(x)\right|,$$

where h is a fixed function in $\mathfrak{S}(\mathbb{R}^n)$ and j is a fixed nonnegative integer n-tuple. Together with the linear structure and the topological structure, the function space $\mathfrak{O}_M(\mathbb{R}^n)$ is a locally convex linear topological space.

Proposition 7. *C_0^∞ is dense in $\mathfrak{O}_M(\mathbb{R}^n)$ with respect to the topology in $\mathfrak{O}_M(\mathbb{R}^n)$.*

Proposition 8. *Any function $f \in \mathfrak{O}_M(\mathbb{R}^n)$ defines a tempered distribution*

$$T_f(\phi) = \int_{\mathbb{R}^n} f(x)\phi(x)dx.$$

If $f \in \mathfrak{O}_M(\mathbb{R}^n)$, then we might define

$$D^j T_f(\phi) := T_{D^j f}(\phi) = \int_{\mathbb{R}^n} D^j f(x)\phi(x)dx = (-1)^{|j|}\int_{\mathbb{R}^n} f(x)D^j\phi(x)dx = (-1)^{|j|}T_f(D^j\phi).$$

This suggests us define the derivative of a tempered distribution T by

$$D^j T(\phi) = (-1)^{|j|}T(D^j\phi).$$

This is reasonable since the differential operator D^j is linear on $\mathfrak{S}(\mathbb{R}^n)$. Similarly, for any functions $f, g \in \mathfrak{O}_M(\mathbb{R}^n)$, we can define the multiplication by g to the tempered distribution T_f through

$$g(T_f)(\phi) = T_{fg}(\phi) = \int f(x)g(x)\phi(x)dx.$$

This suggests us we may define the multiplication by a function g to any tempered distribution through

$$(g \cdot T)(\phi) = T(g \cdot \phi).$$

Since the mapping $\phi \mapsto g \cdot \phi$ is linear and continuous, $g(T)$ is also a continuous linear functional.

2.1 The Fourier Transform of Tempered Distributions

Recall that in the linear algebra class, we have learned that for any linear transformation $T : V \to W$, it induces a linear transformation on the dual spaces $T^\vee : W_b^\vee \to V_b^\vee$ by

$$T^\vee(\ell)(v) = \ell(T(v)).$$

Here $V_b^\vee$ is the dual space of V that consists of continuous linear functional on V. This motivates us to define the Fourier transform on $\mathfrak{S}(\mathbb{R}^n)'$ as

Definition 9 (The Fourier Transform of a tempered distribution)**.** Let T be a tempered distribution. We define the Fourier transform $\widehat{T}$ of T by

$$\widehat{T}(\phi) = T(\widehat{\phi}).$$

Example 10. *If $f \in L^1(\mathbb{R}^n)$, then $\widehat{T_f} = T_{\widehat{f}}$. This can be seen from*

$$\widehat{T_f}(\phi) = T_f(\widehat{\phi}) = \int f(x)\widehat{\phi}(x)dx = (2\pi)^{-n/2}\int f(x)\left(\int e^{-i\langle x,\xi\rangle}\phi(\xi)d\xi\right)dx$$
$$= \int \widehat{f}(\xi)\phi(\xi)d\xi.$$

Proposition 9. *Let $\check{f}(x) = f(-x)$. Then we have*

$$\widehat{\widehat{f}} = \check{f}.$$

The proposition is true by

$$\tilde{g}(x) = \int e^{i\langle x,\xi\rangle}g(\xi)d\xi = \int e^{-i\langle -x,\xi\rangle}g(\xi)d\xi = \widehat{g}(-x).$$

This identity helps us to generalize the Fourier's integral theorem.

Theorem 11 (Fourier's integral theorem on Schwartz space)**.** *If we define $\check{T}(\phi) = T(\check{\phi})$, then we have*

$$\widehat{\widehat{T}} = \check{T}.$$

We infer that the map $T \mapsto \check{T}$ is linear.

Proof. By definition, we have

$$\widehat{\widehat{T}}(\phi) = T(\widehat{\widehat{\phi}}) = T(\check{\phi}) = \check{T}(\phi)$$

holds for all $\phi \in \mathfrak{S}(\mathbb{R}^n)$. □

Proposition 10. *The Fourier transform $T \mapsto \widehat{T}$ and its inverse are linear and continuous on $\mathfrak{S}(\mathbb{R}^n)'$ onto $\mathfrak{S}(\mathbb{R}^n)'$ with respect the weak-* topology.*

For clarity, we give the definition of the weak-* topology.

Definition 12 (weak-* topology)**.** Let V be a topological vector space. Then the weak-* topology is the weakest topology defined on the continuous dual space $V_{\mathrm{b}}^{\vee}$ of V, which satisfies the property: for every $x \in V$ the map

$$\phi_x : V_{\mathrm{b}}^{\vee} \to \mathbb{C}$$
$$\ell \mapsto \ell(x) \qquad (\ell \in V_{\mathrm{b}}^{\vee})$$

is continuous.

The Fourier's integral theorem shows that $T_n(\phi) \to T(\phi)$ for all $\phi \in \mathfrak{S}(\mathbb{R}^n)$ implies $\widehat{T_n}(\phi) = T_n\left(\widehat{\phi}\right) \to T\left(\widehat{\phi}\right) = \widehat{T}(\phi)$.

Theorem 13 (Plancherel's Theorem)**.** *If $f \in L^2(\mathbb{R}^n)$ then the Fourier transform $\widehat{T}_f$ of T_f is defined by a function $\widehat{f} \in L^2(\mathbb{R}^n)$, that is, there is a function $\widehat{f} \in L^2(\mathbb{R}^n)$ such that*

$$\widehat{T}_f = T_{\widehat{f}},$$

and

$$\|\widehat{f}\|_{L^2} = \left(\int \left|\widehat{f}\right|^2 dx\right)^{1/2} = \left(\int |f|^2 dx\right)^{1/2} = \|f\|_{L^2}.$$

Proof. By the Hölder's inequality, we have

$$\left|\widehat{T}_f(\phi)\right| = \left|T_f(\widehat{\phi})\right| = \left\|\int f(x)\widehat{\phi(x)}dx\right\| \le \|f\|_{L^2} \cdot \|\widehat{\phi}\|_{L^2} = \|f\|_{L^2} \cdot \|\phi\|_{L^2}.$$

Note that we have used the fact that $\|\widehat{\phi}\|_{L^2} = \|\phi\|_{L^2}$ by (6). Hence, this shows that $\widehat{T}_f$ is a bounded linear functional on $\mathfrak{S}(\mathbb{R}^n)$, and it extends continuously to a bounded linear functional on $L^2(\mathbb{R}^n)$ since it is well-known that the Schwartz space is dense in the $L^2(\mathbb{R}^n)$ space (with respect to the L^2 norm). Thus, by Riesz' representation theorem, there exists a unique function $\widehat{f} \in L^2(\mathbb{R}^n)$ such that

$$\widehat{T}_f(\phi) = \int \phi(x)\widehat{f}(x)dx = T_{\widehat{f}}(\phi) \qquad \text{or} \qquad \int \widehat{f}\phi = \int f\widehat{\phi}, \qquad \forall \phi \in \mathfrak{S}(\mathbb{R}^n). \tag{9}$$

Now we use the fact that $\mathfrak{S}(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$ again, we let $\phi_n \in L^2(\mathbb{R}^n)$ be a sequence of functions converging to $\overline{\widehat{f}}$ (in the L^2 sense). Then we have

$$\left|\int f(x)\phi_n(x)dx\right| \le \|f\|_{L^2} \cdot \|\phi_n\|_{L^2}.$$

Taking limit $n \to \infty$ gives us

$$\|\widehat{f}\|_{L^2}^2 \le \|f\|_{L^2} \cdot \|\widehat{f}\|_{L^2},$$

which implies $\|\widehat{f}\|_{L^2} \le \|f\|_{L^2}$. We thus have $\|\widehat{\widehat{f}}\|_{L^2} \le \|\widehat{f}\|_{L^2} \le \|f\|_{L^2}$. It is worth noting that

$$\int_{\mathbb{R}^n} \widehat{\widehat{f}}(x)\phi(x)dx = \int_{\mathbb{R}^n} f(x)\phi(-x)dx = \int_{\mathbb{R}^n} f(-x)\phi(x)dx$$

for all $\phi \in \mathfrak{S}(\mathbb{R}^n)$. We conclude that

$$\widehat{\widehat{f}}(x) = f(-x)$$

for almost every x. Thus $\|\widehat{\widehat{f}}\|_{L^2} \le \|f\|_{L^2}$ and therefore $\|\widehat{f}\|_{L^2} = \|f\|_{L^2}$. □

From the theorem, we can define the Fourier transform by the following.

Definition 14 (Fourier transform on the L^2 space)**.** For a function $f \in L^2(\mathbb{R}^n)$ we define its Fourier transform be the above obtained $\widehat{f}$.

Corollary. We have, for any $f \in L^2(\mathbb{R}^n)$,

$$\widehat{f}(x) = \lim_{h\to\infty} (2\pi)^{-n/2} \int_{|x|\le h} e^{-i\langle x,y\rangle} f(y)dy. \tag{10}$$

Proof. Let

$$f_h(x) := \begin{cases} f(x) & \text{, if } |x| \le h \\ 0 & \text{, if } |x| > h \end{cases}.$$

Then $\lim_{h\to\infty} \|f_h - f\|_{L^2} = 0$ and thus we have $\lim_{h\to\infty} \|\widehat{f}_h - \widehat{f}\|_{L^2} = 0$ and $\widehat{f}(x) = \lim_{h\to\infty} \widehat{f}_h(x)$ for almost every x. By (9), we have

$$\begin{aligned}\int_{\mathbb{R}^n} \widehat{f}_h(x)\phi(x)dx &= \int_{\mathbb{R}^n} f_h(x)\widehat{\phi}(x)dx \\ &= \int_{|x|\le h} f(x)\left(\int (2\pi)^{-n/2}\int_{\mathbb{R}^n} e^{-i\langle x,y\rangle}\phi(y)dy\right)dx \\ &\overset{\spadesuit}{=} \int_{\mathbb{R}^n}\left((2\pi)^{-n/2}\int_{|x|\le h} e^{-i\langle x,y\rangle}f(x)dx\right)\phi(y)dy.\end{aligned} \tag{11}$$

Changing the order of integration (♠) is valid by Fubini's theorem. Since (11) is true for all $\phi \in \mathfrak{S}(\mathbb{R}^n)$, we conclude that

$$\widehat{f}_h(x) = (2\pi)^{-n/2}\int_{|x|\le h} e^{-i\langle x,y\rangle}f(x)dx.$$

Together with $\widehat{f}(x) = \lim_{h\to\infty} \widehat{f}_h(x)$, we obtain the desired result. □

Remark. However,the equation

$$\widehat{f}(x) = (2\pi)^{-n/2}\int_{\mathbb{R}^n} e^{-i\langle x,y\rangle}f(y)dy. \tag{12}$$

is not true, since the integral on the right hand side might diverge. (Recall that in one of the homework, we are asked to show that one improper integral might not be Lebesgue integrable, but converge in the sense of Riemann integral.)

Corollary. The Fourier transform on the L^2 space is bijective, moreover, we have

$$\langle f,g\rangle = \left\langle \widehat{f},\widehat{g}\right\rangle \tag{13}$$

for all $f,g \in L^2(\mathbb{R}^n)$.

Proof. Similar to Fourier transform on the Schwartz space, we consider the inverse Fourier transform $f \to \tilde{f}$. We can define the inverse Fourier transform $\tilde{f}$ of a function $f \in L^2(\mathbb{R}^n)$ by

$$\tilde{f} \in L^2(\mathbb{R}^n) \text{ is the unique function such that } \widetilde{T}_f = T_{\tilde{f}}.$$

Then, similar argument shows that

$$\tilde{f}(x) = \lim_{h\to\infty}(2\pi)^{-n/2}\int_{|x|\le h} e^{i\langle x,y\rangle}f(y)dy. \tag{14}$$

It is also clear that $\|\tilde{f}\|_{L^2} = \left\|f\right\|_{L^2}$ and $\widehat{\tilde{f}}(x) = \tilde{\widehat{f}}(x) = f(x)$ for almost every x. Therefore $f \to \widehat{f}$ is bijective. The corollary now follows by the identity

$$\langle x,y\rangle = \frac{1}{4}(\left\|x+y\right\|^2 - \left\|x-y\right\|^2) + \frac{i}{4}(\left\|x+iy\right\|^2 - \left\|x-iy\right\|^2),$$

where x,y are in some inner product spaces. □

Theorem 15 (Parseval's Theorem)**.** *Let $f,g \in L^2(\mathbb{R}^n)$ and let $\widehat{f},\widehat{g}$ be their Fourier transform. Then*

$$\int_{\mathbb{R}^n} \widehat{f}(u)\widehat{g}(u)du = \int_{\mathbb{R}^n} f(x)g(-x)dx, \tag{15}$$

and therefore

$$\int_{\mathbb{R}^n} \widehat{f}(u)\widehat{g}(u)e^{i\langle u,x\rangle}du = \int_{\mathbb{R}^n} f(y)g(x-y)dy. \tag{16}$$

Thus, if $\widehat{f}, \widehat{g}$ and $\widehat{f} \cdot \widehat{g}$ both belong to $L^2(\mathbb{R}^n)$, then $\widehat{f} \cdot \widehat{g}$ is the Fourier transform of the function

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

Similarly, if f, g and $f * g$ both belong to $L^2(\mathbb{R}^n)$, then the same conclusion also hold.

Proof. It is easy to see that

$$(2\pi)^{-n/2} \int_{\mathbb{R}^n} \overline{g(-x)} e^{-i\langle u,x\rangle} dx = \overline{\widehat{g}(u)}.$$

Thus we have $\overline{\widehat{g}} = \widehat{\overline{\check{g}}}$. Now it follows that

$$\int \widehat{f}(u)\widehat{g}(u)du = \left\langle \widehat{f}, \overline{\widehat{g}} \right\rangle = \left\langle \widehat{f}, \widehat{\overline{\check{g}}} \right\rangle = \left\langle f, \overline{\check{g}} \right\rangle = \int f(x)g(-x)dx.$$

Next note that we have shown that, in the proof of (6), $\widetilde{\overline{g}} = \overline{\widehat{g}}$. (Although we are focusing on $\mathfrak{S}(\mathbb{R}^n)$ not L^2 in that proof, it still can be proved similarly.) Therefore, we conclude that

$$\int \overline{g(x-y)} e^{-i\langle u,y\rangle} dy = \int \overline{g(t)} e^{-i\langle u,x-t\rangle} dt = e^{-i\langle u,x\rangle} \cdot \widetilde{\overline{g}}(u) = e^{-i\langle u,x\rangle} \cdot \overline{\widehat{g}}(u).$$

Thus we obtain

$$\int_{\mathbb{R}^n} \widehat{f}(u)\widehat{g}(u)e^{i\langle u,x\rangle} du = \left\langle \widehat{f}, e^{-i\langle u,x\rangle} \overline{\widehat{g}}(u) \right\rangle_{[u]} = \left\langle f(y), \overline{g(x-y)} \right\rangle_{[y]} = \int_{\mathbb{R}^n} f(y)g(x-y)dy.$$

We use $\langle \cdot, \cdot \rangle_{[u]}$ to denote the inner product of two functions of the variable u. □

3 Convolutions

We have already define the convolution of two functions $f, g \in C^\infty(\mathbb{R}^n)$, one of which has a compact support, by

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

This formula suggests us that we may define the convolution of $T \in \mathfrak{D}(\mathbb{R}^n)'$ and a function $\phi \in \mathfrak{D}(\mathbb{R}^n)$ (or $T \in \mathfrak{E}(\mathbb{R}^n)'$ and a $\phi \in \mathfrak{E}(\mathbb{R}^n)$) by

$$(T * \phi)(x) := T_{[y]}(\phi(x-y)), \tag{17}$$

where $T_{[y]}$ indicates that we apply the distribution T on test functions of y. It is worth noting that $T * \phi$ is a function on $\mathbb{R}^n$. In fact, we have the following proposition.

Proposition 11. *$(T * \phi) \in C^\infty(\mathbb{R}^n)$ and $\mathrm{supp}(T * \phi) \subset \mathrm{supp}(T) + \mathrm{supp}(\phi)$. In other words,*

$$\mathrm{supp}(T * \phi) \subset \{w \in \mathbb{R}^n : w = x + y, x \in \mathrm{supp}(T), y \in \mathrm{supp}(\phi)\}.$$

Moreover, we have

$$D^\alpha(T * \phi) = T * (D^\alpha \phi) = (D^\alpha T) * \phi. \tag{18}$$

This notation is quite confusing for me at the first glance since my instinct thinks that

$$\mathrm{supp}(T) := \mathrm{cl}\{f \in \mathfrak{D}(\mathbb{R}^n) : T(f) \neq 0\},$$

where the closure is taken on the topology of $\mathfrak{D}(\mathbb{R}^n)$. After I searched some information on the internet, I found that my instinct was totally wrong. For clarity, let me write down the definition of the support of a distribution.

Definition 16 (support of a distribution)**.** Let $T \in \mathfrak{D}(\mathbb{R}^n)'$ be a distribution. Then an open set $\omega \subset \mathbb{R}^n$ is said to be an annihilation set if for every $\phi \in C_0^\infty(\mathbb{R}^n)$ with $\mathrm{supp}(\phi) \subset \omega$, we have $T(\phi) = 0$. Then the support $\mathrm{supp}(T)$ of T is defined to be the complement of the union of all open annihilation sets.

Proposition 12. *Given a distribution $T \in \mathfrak{D}(\mathbb{R}^n)'$. If ω_1 and ω_2 are two open annihilation sets of T, then so is $\omega_1 \cup \omega_2$.*

Proof. This is just a corollary of "partition of unity", that is, for any given bounded open set $U \subset \mathbb{R}^n$, there are countably many smooth function $\rho_i : \mathbb{R}^n \to [0,1]$ such that

$$\sum_{i=1}^{\infty} \rho_i = \chi_U.$$

□

Proof of Proposition 11. Let $x \in \mathbb{R}^n$ such that $T * \phi(x) \neq 0$. Then

$$T_{[y]}(\phi(x-y)) \neq 0.$$

We claim that $(x - \mathrm{supp}(\phi)) \cap \mathrm{supp}(T) \neq \varnothing$. If not, then the support of $\phi(x-y)$ as a function of y is an annihilation set, that is, $(x - \mathrm{supp}(\phi)) \cap \mathrm{supp}(T) = \varnothing$, and hence $T_{[y]}(\phi(x-y)) = 0$, which leads to a contradiction. Thus, we have $x \in \mathrm{supp}(\phi) + \mathrm{supp}(T)$. Since $K + F$ is closed if K is compact and F is closed in $\mathbb{R}^n$, we conclude that $\mathrm{supp}(\phi) + \mathrm{supp}(T)$ is closed. Deduction above shows that

$$\{x : (T * \phi)(x) \neq 0\} \subset \mathrm{supp}(\phi) + \mathrm{supp}(T),$$

and therefore

$$\mathrm{supp}(T * \phi) := \mathrm{cl}\{x : (T * \phi)(x) \neq 0\} \subset \mathrm{supp}(\phi) + \mathrm{supp}(T).$$

Next, for the second part of the proposition, it suffices to show that $|\alpha| = 1$. Let e_j be the unit vector of $\mathbb{R}^n$ along the x_j-axis and consider the equation

$$\frac{(T * \phi)(x + he_j) - (T * \phi)(x)}{h} = T_{[y]}((\phi(x + he_j) - \phi(x))/h).$$

When $h \to 0$, the function enclosed by the outer parenthesis (on the right side) converges, as a function of y, to $\left(\frac{\partial \phi}{\partial x_j}\right)(x-y)$ in $\mathfrak{D}(\mathbb{R}^n)$ (or in $\mathfrak{E}(\mathbb{R}^n)$). Therefore we obtain

$$\frac{\partial}{\partial x_j}(T * \phi)(x) = \left(T * \frac{\partial \phi}{\partial x_j}\right)(x).$$

Moreover, recall that we have defined the derivatives of distributions by

$$(D^\alpha)T(\phi) := (-1)^{|\alpha|} T(D^\alpha(\phi)).$$

Thus we have

$$\left(\frac{\partial T}{\partial x_j} * \phi\right)(x) = \frac{\partial T_{[y]}}{\partial y_j}(\phi(x-y)) := T_{[y]}\left(-\frac{\partial \phi(x-y)}{\partial y_j}\right) = \left(T * \frac{\partial \phi}{\partial x_j}\right)(x).$$

This prove the second statement of the proposition. □

Corollary. If $T \in \mathfrak{E}(\mathbb{R}^n)'$ and $\phi \in \mathfrak{D}(\mathbb{R}^n) \subset \mathfrak{E}(\mathbb{R}^n)$, then

$$\operatorname{supp}(T * \phi) \text{ is compact.}$$

Proof. It is clear that $K_1 + K_2$ is compact for any two compact subsets $K_1, K_2 \subset \mathbb{R}^n$. Note that $\mathfrak{E}(\mathbb{R}^n)'$ is defined as the subset of $\mathfrak{D}(\mathbb{R}^n)'$ consisting of compactly supported distributions. □

Proposition 13. *Suppose ϕ, ψ are in $\mathfrak{D}(\mathbb{R}^n)$ and $T \in \mathfrak{D}(\mathbb{R}^n)'$ (or $\phi \in \mathfrak{E}(\mathbb{R}^n)$, $\psi \in \mathfrak{D}(\mathbb{R}^n)$, and $T \in \mathfrak{E}(\mathbb{R}^n)'$), then we have*

$$(T * \phi) * \psi = T * (\phi * \psi). \tag{19}$$

Proof. We approximate the function $(\phi * \psi)(x)$ by the Riemann sum

$$f_h(x) = h^n \sum_{k \in \mathbb{Z}^n} \phi(x - kh)\psi(kh),$$

where $h > 0$. For every differentiation D^α and for every compact set K, the functions

$$D^\alpha f_h(x) = h^n \sum_{k \in \mathbb{Z}^n} D^\alpha \phi(x - kh)\psi(kh)$$

converges uniformly in x (on K). The limit function is

$$\left(D^\alpha \phi) * \psi\right)(x) = \left(D^\alpha(\phi * \psi)\right)(x).$$

Hence we see that $\lim_{h \to 0} f_h = \phi * \psi$ in $\mathfrak{D}(\mathbb{R}^n)$ (or in $\mathfrak{E}(\mathbb{R}^n)$). Therefore, by the linearity and continuity of T, we have

$$T * (\phi * \psi)(x) = \lim_{h \to 0}(T * f_h)(x) = \lim_{h \to 0} h^n \sum_{k \in \mathbb{Z}^n} (T * \phi)(x - kh)\psi(kh) = ((T * \phi) * \psi)(x).$$

This gives the desired result. □

Next we are going to introduce an useful definition.

Definition 17 (Regularization)**.** Let $\phi \in \mathfrak{D}(\mathbb{R}^n)$ be a non-negative function such that

1. $\displaystyle\int_{\mathbb{R}^n} \phi(x)dx = 1.$
2. $\operatorname{supp}(\phi) \subset \{x \in \mathbb{R}^n : |x| \le 1\}$.

We write $\phi_\epsilon(x) := \epsilon^{-n}\phi(x/\epsilon)$, for $\epsilon > 0$. We call $T * \phi_\epsilon$ the regularization of $T \in \mathfrak{D}(\mathbb{R}^n)'$ (or $\mathfrak{E}(\mathbb{R}^n)'$) through $\phi_\epsilon(x)$.

In fact, this concept have been taught implicitly in the class, which we were constructing the Kolmogorov's function with Fourier series diverging almost everywhere. In that class we have shown that for any $f \in L^p$, we have

$$\left\| f * \phi_\epsilon - f \right\|_p \to 0$$

as $\epsilon \to 0$. Here we have some very similar result, even f is replaced by a distribution T.

Theorem 18. *Let $T \in \mathfrak{D}(\mathbb{R}^n)'$ (or $\mathfrak{E}(\mathbb{R}^n)'$). Then we have*

$$\lim_{\epsilon \to 0}(T * \phi_\epsilon) = T$$

in the weak-* topology. If ϕ_ϵ is chosen for the approximation, then it is called the approximate identity.

This theorem is also very confusing for me when I first heard this in the meeting of our study group. How come the limit of a sequence of functions becomes a distributions? After some discussions, we figured out that we may use the natural inclusion from $\mathfrak{E}(\mathbb{R}^n)$ to $\mathfrak{D}(\mathbb{R}^n)'$ (or from $\mathfrak{D}(\mathbb{R}^n)$ to $\mathfrak{E}(\mathbb{R}^n)'$). For each $f \in \mathfrak{E}(\mathbb{R}^n)$, we can associate it with a natural distribution $T_f \in \mathfrak{D}(\mathbb{R}^n)'$ defined by

$$T_f : \phi \mapsto \int f\phi.$$

Similarly, for $f \in \mathfrak{D}(\mathbb{R}^n)$, we can associate it with a natural distribution on $T_f \in \mathfrak{E}(\mathbb{R}^n)'$ by

$$T_f : \phi \mapsto \int f\phi.$$

By using this identification and corollary 3, we know that $T * \phi_\epsilon \in \mathfrak{D}(\mathbb{R}^n) \subset \mathfrak{E}(\mathbb{R}^n)'$ if $T \in \mathfrak{E}(\mathbb{R}^n)'$. And if $T \in \mathfrak{D}(\mathbb{R}^n)'$, then we only have $T * \phi_\epsilon \in \mathfrak{E}(\mathbb{R}^n) \subset \mathfrak{D}(\mathbb{R}^n)'$. Before proving this theorem, we shall prove some lemmas.

Lemma 1. For any $\psi \in \mathfrak{D}(\mathbb{R}^n)$ (or $\mathfrak{E}(\mathbb{R}^n)$), we have

$$\lim_{\epsilon \to 0}(\psi * \phi_\epsilon) = \psi$$

in $\mathfrak{D}(\mathbb{R}^n)$ (or in $\mathfrak{E}(\mathbb{R}^n)$).

Proof. We first observe that

$$\text{supp}(\psi * \phi_\epsilon) \subset \text{supp}(\psi) + \text{supp}(\phi_\epsilon) = \text{supp}(\psi) + \widehat{B}(0;\epsilon),$$

where $\widehat{B}(0;\epsilon)$ denotes the closed ball with center at 0 and radius ϵ. From the definition of functions convolution, we have

$$D^\alpha(\psi * \phi_\epsilon) = (D^\alpha \psi) * \phi_\epsilon.$$

Hence we have to show that $\lim_{\epsilon \to 0}(\psi * \phi_\epsilon)(x) = \psi(x)$ uniformly on any compact set. Note that $\int \phi_\epsilon(y)dy = 1$, therefore

$$(\psi * \phi_\epsilon)(x) - \psi(x) = \int_{\mathbb{R}^n} \big(\psi(x-y) - \psi(x)\big) \cdot \phi_\epsilon(y)dy.$$

Since ϕ_ϵ is non-negative, $\int_{\mathbb{R}^n} \phi_\epsilon(y)dy = 1$ and the uniform continuity of $\psi(x)$ on any compact set on x, we obtain the desired result. □

We now could give the proof of Theorem 18.

Proof of Theorem 18. First note that

$$(T * \check{\psi})(0) = T_{[y]}(\check{\psi})(-y)) = T_{[y]}(\phi(y)) = T(\phi). \tag{20}$$

Hence it suffices to show that

$$\lim_{\epsilon\to 0}((T * \phi_\epsilon) * \check{\psi})(0) = (T * \check{\psi})(0).$$

By proposition 13 and (20), we have

$$((T * \phi_\epsilon) * \check{\psi})(0) = (T * (\phi_\epsilon * \check{\psi}))(0) = T\left((\phi_\epsilon * \check{\psi})\check{}\right).$$

The continuity of T and Lemma 1 gives

$$\lim_{\epsilon\to 0} T\left((\phi_\epsilon * \check{\psi})\check{}\right) = T((\check{\psi})\check{}) = T(\psi).$$

This completes the proof. □

The next theorem characterizes the operation of convolution.

Theorem 19 (L. Schwartz' Theorem). *Let L be a continuous linear mapping on $\mathfrak{D}(\mathbb{R}^n)$ into $\mathfrak{E}(\mathbb{R}^n)$ such that*

$$L\tau_h\phi = \tau_h L\phi \tag{21}$$

for any $h \in \mathbb{R}^n$ and $\phi \in \mathfrak{D}(\mathbb{R}^n)$. Here τ_h denotes the translation operator, where it is defined by $\tau_h\phi(x) := \phi(x-h)$. Then there exists a unique $T \in \mathfrak{D}(\mathbb{R}^n)'$ such that $L * \phi = T * \phi$. Conversely, for any $T \in \mathfrak{D}(\mathbb{R}^n)'$ defines a continuous linear map L on $\mathfrak{D}(\mathbb{R}^n)$ into $\mathfrak{E}(\mathbb{R}^n)$ by $L\phi = T * \phi$ such that L satisfies (21).

Proof. Since $\phi \mapsto \check{\phi}$ is a continuous linear map of $\mathfrak{D}(\mathbb{R}^n)$ onto itself, the linear map $T : \check{\phi} \mapsto (L\psi)(0)$ defines a distribution $T \in \mathfrak{D}(\mathbb{R}^n)'$. Then by (20), we have

$$(L\phi)(0) = T(\check{\phi}) = (T * \phi)(0).$$

Replacing ϕ be $\tau_h\phi$ and make use of the condition (21), then we obtain

$$(L\phi)(-h) = (\tau_h L\phi)(0) = (L\tau_h\phi)(0) = (T * (\tau_h\phi))(0) = (T * \phi)(-h).$$

Thus $L\phi = T * \phi$. The converse part follows by

$$T * (\tau_h\phi) = T_{[y]}(\phi(x-h-y)) = \tau_h \cdot T_{[y]}(\phi(x-y)) = \tau_h(T * \phi).$$

□

Corollary. Let $T_1 \in \mathfrak{D}(\mathbb{R}^n)'$ and $\mathfrak{E}(\mathbb{R}^n)'$. We consider the linear continuous map L from $\mathfrak{D}(\mathbb{R}^n)$ to $\mathfrak{E}(\mathbb{R}^n)$ defined by

$$L : \phi \mapsto T_1 * (T_2 * \phi).$$

Then Theorem 19 asserts that there is a unique distribution $T \in \mathfrak{D}(\mathbb{R}^n)'$ such that $T * \phi = L(\phi)$. We then may define the convolution $T_1 * T_2$ of T_1 and T_2 as the obtained T. In other words, we may define

$$(T_1 * T_2) * \phi = L(\phi) = T_1 * (T_2 * \phi).$$

Proof. We shall show that the map L defined by

$$L : \phi \mapsto T_1 * (T_2 * \phi)$$

is from $\mathfrak{D}(\mathbb{R}^n)$ to $\mathfrak{E}(\mathbb{R}^n)$, and it is linear and continuous. since we require that $T_2 \in \mathfrak{E}(\mathbb{R}^n)'$, $\text{supp}(T_2)$ is compact. Therefore, $\text{supp}(T_2 * \phi)$ is compact by proposition 11, that is, $T_2 * \phi \in \mathfrak{D}(\mathbb{R}^n)$. It is clear that $T_1 * (T_2 * \phi) \in \mathfrak{E}(\mathbb{R}^n)$. To show the map is linear and continuous, we consider the composition of two maps

$$\phi \mapsto T_2 * \phi \mapsto T_1 * (T_2 * \phi).$$

The first map is linear and continuous from $\mathfrak{D}(\mathbb{R}^n)$ to itself, and the second map is also linear and continuous. □

The next theorem shows that the convolution is actually commutative.

Theorem 20 (Commutativity of the convolution). *Let $T_1 \in \mathfrak{D}(\mathbb{R}^n)'$ and $T_2 \in \mathbb{E}(\mathbb{R}^n)'$. Then we also can consider the continuous linear map*

$$L : \phi \mapsto T_2 * (T_1 * \phi)$$

from $\mathfrak{D}(\mathbb{R}^n)$ to $\mathfrak{E}(\mathbb{R}^n)$. By Theorem 19, there is a distribution $T \in \mathfrak{D}(\mathbb{R}^n)'$ such that $T * \phi = L(\phi)$. Then we define another convolution $T_2 \boxtimes T_1$ to be the just obtained T.

Then, we have $T_1 * T_2 = T_2 \boxtimes T_1$.

Proof. The map $\phi \mapsto T_1 * \phi$ is linear and continuous on $\mathfrak{D}(\mathbb{R}^n)$ into $\mathfrak{E}(\mathbb{R}^n)$, which is again by Proposition 11. Hence the map $\phi \mapsto T_2 * (T_1 * \phi)$ is linear and continuous on $\mathfrak{D}(\mathbb{R}^n)$ into $\mathfrak{E}(\mathbb{R}^n)$. Thus, $T_2 \boxtimes T_1$ is well-defined by Theorem 19. Next we shall show that $T_2 \boxtimes T_1 = T_1 * T_2$.

Let $\phi_1, \phi_2 \in \mathfrak{D}(\mathbb{R}^n)$. Then we have,

$$\begin{aligned}(T_1 * T_2) * (\phi_1 * \phi_2) &\overset{(a)}{=} T_1 * (T_2 * (\phi_1 * \phi_2)) \overset{(b)}{=} T_1 * ((T_2 * \phi_1) * \phi_2) \\ &\overset{(c)}{=} T_1 * (\phi_2 * (T_2 * \phi_1)) \overset{(d)}{=} (T_1 * \phi_2) * (T_2 * \phi_1).\end{aligned}$$

We have used Proposition 13 on the equality (b) and (d), and the commutativity of the convolution of functions on (c). It is worth noting that all computations are valid since $T_2 * \phi_1 \in \mathfrak{D}(\mathbb{R}^n)$ by Corollary 3. Now we compute the expression $(T_2 \boxtimes T_1) * (\phi_1 * \phi_2)$,

$$\begin{aligned}(T_2 \boxtimes T_1) * (\phi_1 * \phi_2) &\overset{(e)}{=} T_2 * (T_1 * (\phi_2 * \phi_1)) \overset{(f)}{=} T_2 * ((T_1 * \phi_2) * \phi_1) \\ &\overset{(g)}{=} T_2 * (\phi_1 * (T_1 * \phi_2)) \overset{(h)}{=} (T_2 * \phi_1) * (T_1 * \phi_2).\end{aligned}$$

Now let $\phi_1 = \psi \in \mathfrak{D}(\mathbb{R}^n)$ and let $\phi_2 = \phi_\epsilon$ be an approximate identity (recall Definition 17). Since both $T_1 * T_2$ and $T_2 \boxtimes T_1$ are belong to $\mathfrak{D}(\mathbb{R}^n)'$, we have

$$(T_1 * T_2) * \psi = \lim_{\epsilon \to 0} (T_1 * T_2) * (\psi * \phi_\epsilon) = \lim_{\epsilon \to 0} (T_2 \boxtimes T_1) * (\psi * \phi_\epsilon) = (T_2 \boxtimes T_1) * \psi,$$

by applying Lemma 1. The theorem now follows from the "uniqueness" mentioned in Theorem 19. □

Remark. We shall use the symbol $T_2 \boxtimes T_1$, since $T_1 * T_2$ is only defined when $T_1 \in \mathfrak{D}(\mathbb{R}^n)'$ and $T_2 \in \mathfrak{E}(\mathbb{R}^n)'$. However, after proving Theorem 21, it is clear that we could write $T_2 * T_1$ to denote $T_2 \boxtimes T_1$, if either T_1 or T_2 is in $\mathfrak{E}(\mathbb{R}^n)'$

Corollary.

1. Suppose both T_1 and T_2 belong to $\mathfrak{D}(\mathbb{R}^n)'$ and at least one of them has compact support, then

$$\text{supp}(T_1 * T_2) \subset \text{supp}(T_1) + \text{supp}(T_2).$$

In particular, if both T_1 and T_2 belong to $\mathfrak{E}(\mathbb{R}^n)'$, then so do $T_1 * T_2$.

2. $T_1 * (T_2 * T_3) = (T_1 * T_2) * T_3$ if all T_j except one have compact support.

3. $D^\alpha(T_1 * T_2) = (D^\alpha T_1) * T_2 = T_1 * (D^\alpha T_2)$.

Proof.

1. Suppose $\phi \in \mathfrak{D}(\mathbb{R}^n)$ are distribution such that

$$(T_1 * T_2)(\phi) \neq 0,$$

then by (20) we have

$$\begin{aligned}(T_1 * T_2)(\phi) \neq 0 &\implies (T_1 * (T_2 * \check{\phi}))(0) \neq 0 \\ &\implies 0 \in \text{supp}(T_1) + \text{supp}(T_2 * \check{\phi}) \\ &\implies 0 \in \text{supp}(T_1) + \text{supp}(T_2) - \text{supp}(\phi) \\ &\implies \text{supp}(\phi) \cap (\text{supp}(T_1) + \text{supp}(T_2)) \neq \varnothing\end{aligned}$$

Thus for any $\text{supp}(\phi) \subset \mathbb{R}^n \setminus (\text{supp}(T_1) + \text{supp}(T_2))$, we have $(T_1 * T_2)(\phi) = 0$. Hence, any open subset ω of $(\text{supp}(T_1) + \text{supp}(T_2))$ is an annihilation set of $T_1 * T_2$, therefore we conclude that $\text{supp}(T_1 * T_2) \subset \text{supp}(T_1) + \text{supp}(T_2)$.

2. By (20), we have

$$\begin{aligned}(T_1 * (T_2 * T_3))(\phi) &= ((T_1 * (T_2 * T_3)) * \check{\phi})(0) \\ &= (T_1 * ((T_2 * T_3) * \check{\phi}))(0) \\ &= (T_1 * (T_2 * (T_3 * \check{\phi})))(0)\end{aligned}$$

and similarly,

$$((T_1 * T_2) * T_3)(\phi) = (T_1 * (T_2 * (T_3 * \check{\phi})))(0).$$

Note that we have used the fact that if both T_1 and T_2 belong to $\mathfrak{E}(\mathbb{R}^n)'$, then so do $T_1 * T_2$.

3. To show the third assertion, we observe that, by proposition 11,

$$(D^\alpha T_\delta) * \phi = T_\delta * (D^\alpha \phi) = D^\alpha (T_\delta * \phi) \overset{\heartsuit}{=} D^\alpha \phi, \tag{22}$$

where $T_\delta(\phi) := \phi(0)$ is the evaluation functional at 0. Note that ($\heartsuit$) holds since

$$T_\delta * \phi(x) = T_{\delta,[y]}(\phi(x-y)) = \phi(x).$$

Now (22) implies that

$$(D^\alpha T) * \phi = T * (D^\alpha \phi) = T * ((D^\alpha T_\delta) * \phi) = (T * D^\alpha T_\delta) * \phi. \tag{23}$$

Therefore, we obtain $D^\alpha T = (D^\alpha T_\delta) * T$. Now by the commutativity (Theorem 20) and the associativity (Corollary 3 (2)), we obtain

$$\begin{aligned} D^{\alpha}(T_1 * T_2) &= (D^{\alpha}T_{\delta}) * (T_1 * T_2) = ((D^{\alpha}T_{\delta}) * T_1) * T_2 = (D^{\alpha}T_1) * T_2 \\ &= (D^{\alpha}T_{\delta}) * (T_2 * T_1) = ((D^{\alpha}T_{\delta}) * T_2) * T_1 = (D^{\alpha}T_2) * T_1. \end{aligned}$$

Discussions above proves the corollaries. □

The Fourier transform and the convolution

Theorem 21. *The Fourier transform of a compactly supported distribution $T \in \mathfrak{E}(\mathbb{R}^n)'$ is given by the equation:*

$$\widehat{T}(\xi) := \lim_{\epsilon\to 0}(\widehat{T * \phi_{\epsilon}})(\xi) = (2\pi)^{-n/2}T_{[x]}(e^{-i\langle x,\xi\rangle}). \tag{24}$$

Recall that $T * \phi_{\epsilon}$ is the regularization defined in Definition 17.

Proof. Recall Theorem 18, stating that any distribution $T \in \mathfrak{D}(\mathbb{R}^n)'$, it can be approximate by some smooth functions, that is

$$\lim_{\epsilon\to 0}(T * \phi_{\epsilon}) = T.$$

However, it is worth noting that this is true under the weak-* topology in $\mathfrak{E}(\mathbb{R}^n)'$. Since the topology of $\mathfrak{S}(\mathbb{R}^n)'$ is stronger than $\mathfrak{E}(\mathbb{R}^n)'$, thus the limit is true in the weak-* topology of $\mathfrak{S}(\mathbb{R}^n)'$.

Therefore, by the continuity of the Fourier transform in the weak-* topology of $\mathfrak{S}(\mathbb{R}^n)'$, we obtain

$$\lim_{\epsilon\to 0}(\widehat{T * \phi_{\epsilon}}) = \widehat{T},$$

where the limit converge in the weak-* topology of $\mathfrak{S}(\mathbb{R}^n)'$. Now by definition,

$$(2\pi)^{n/2}(\widehat{T * \phi_{\epsilon}})(\xi) = \int_{\mathbb{R}^n} e^{-i\langle\xi,x\rangle}(T * \phi_{\epsilon})(x)dx = (T * \phi_{\epsilon})_{[x]}\left(e^{-i\langle\xi,x\rangle}\right).$$

Here we have used the identification mentioned in Theorem 18, since $(T * \phi_{\epsilon})(x) \in \mathfrak{D}(\mathbb{R}^n)$ and it can defined a linear continuous functional in $\mathfrak{E}(\mathbb{R}^n)'$. It follows from (20) that

$$\begin{aligned} (2\pi)^{n/2}(\widehat{T * \phi_{\epsilon}})(\xi) &= (T * \phi_{\epsilon})_{[x]}\left(e^{-i\langle\xi,x\rangle}\right) = \left((T_{[x]} * \phi_{\epsilon}) * e^{i\langle\xi,x\rangle}\right)(0) \\ &= \left(T_{[x]} * (\phi_{\epsilon} * e^{i\langle\xi,x\rangle})\right)(0) = T_{[x]}\left(\widetilde{\phi_{\epsilon}} * e^{-i\langle\xi,x\rangle}\right). \end{aligned}$$

The last expression converge to $T_{[x]}\left(e^{-i\langle\xi,x\rangle}\right)$ uniformly in ξ on any bounded set of ξ of the complex n-space. This proves (24). □

The next theorem is a generalization of the Schwartz' Theorem (Theorem 19).

Theorem 22 (Schwartz' Theorem on the Schwartz' space)**.** *If we define the convolution of a tempered distribution $T \in \mathfrak{S}(\mathbb{R}^n)'$ and a rapidly decreasing function $\phi \in \mathfrak{S}(\mathbb{R}^n)$ by*

$$(T * \phi)(x) = T_{[y]}(\phi(x - y)),$$

then the linear map L on $\mathfrak{S}(\mathbb{R}^n)$ into $\mathfrak{E}(\mathbb{R}^n)$

$$L : \phi \mapsto T * \phi$$

is characterized by the continuity and the translation invariance $\tau_h L\phi = L\tau_h\phi$ for any $\phi \in \mathfrak{S}(\mathbb{R}^n)$.

The proof is omitted since it is very similar to the proof of Theorem 19. Now we shall put more focus back to the Fourier transform of tempered distributions.

Recall that we have shown that for any $f, g \in \mathfrak{S}(\mathbb{R}^n)$, we have

$$\widehat{f * g} = (2\pi)^{n/2}\widehat{f} \cdot \widehat{g}.$$

Similarly, we have the following

Theorem 23 (Convolution and the Fourier transform)**.** *If $T \in \mathfrak{S}(\mathbb{R}^n)'$ and $\phi \in \mathfrak{S}(\mathbb{R}^n)$, then*

$$(\widehat{T * \phi}) = (2\pi)^{n/2}\widehat{\phi} \cdot \widehat{T}. \tag{25}$$

If $T \in \mathfrak{S}(\mathbb{R}^n)'$, and $S \in \mathfrak{E}(\mathbb{R}^n)'$, then

$$(\widehat{T * S}) = (2\pi)^{n/2}\widehat{S} \cdot \widehat{T}. \tag{26}$$

The second equation is reasonable since $\widehat{S}$ can be viewed as a smooth function by (24).

Proof. Let $\psi \in \mathfrak{S}(\mathbb{R}^n)$. Then the Fourier transform of $\widehat{\phi} \cdot \psi$ is equal to

$$(2\pi)^{-n/2}\widehat{\widehat{\phi}} * \widehat{\psi} = (2\pi)^{-n/2}\check{\phi} * \widehat{\psi}.$$

Thus,

$$\begin{aligned}
(\widehat{T * \phi})(\psi) &= (T * \phi)(\widehat{\psi}) = ((T * \phi) * \check{\widehat{\psi}})(0) = (T * (\phi * \check{\widehat{\psi}}))(0) \\
&= T\left((\phi * \check{\widehat{\psi}})^{\check{}}\right) = T\left(\check{\phi} * \widehat{\psi}\right) = T\left((2\pi)^{n/2}(\widehat{\phi} \cdot \psi)^{\hat{}}\right) \\
&= (2\pi)^{n/2}\widehat{T}\left(\widehat{\phi} \cdot \psi\right) = (2\pi)^{n/2}\widehat{\phi}\widehat{T}(\psi).
\end{aligned}$$

Note that we have defined the multiplication of a function and a tempered distribution by $g \cdot T(\psi) := T(g\psi)$ for a $g \in \mathfrak{S}(\mathbb{R}^n)$ and $T \in \mathfrak{S}(\mathbb{R}^n)'$. Also we have applied the identity

$$(\widetilde{f * g})(x) = (\check{f} * \check{g})(x)$$

in our derivation, which could be easily be verified. It now remains to show the second equation. Let S_ϵ be the regularization $S * \phi_\epsilon$. Then, by (25), the Fourier transform of $T * S_\epsilon$ is equal to

$$(2\pi)^{n/2} \cdot \widehat{S_\epsilon} \cdot \widehat{T} = \left((2\pi)^{n/2} \cdot \widehat{\phi_\epsilon} \cdot \widehat{S}\right) \cdot (2\pi)^{n/2} \cdot \widehat{T}.$$

On the other hand, $T * S_\epsilon = (T * S) * \phi_\epsilon$, therefore the Fourier transform of $T * S_\epsilon$ is also equal to

$$(2\pi)^{n/2} \cdot \widehat{\phi_\epsilon} \cdot \widehat{T * S}.$$

We conclude that

$$(2\pi)^{n/2}\left(\widehat{\phi_\epsilon} \cdot \widehat{S}\right) \cdot \widehat{T} = \widehat{\phi_\epsilon} \cdot \widehat{T * S}.$$

Taking $\epsilon \to 0$ and using $\lim_{\epsilon \to 0}\widehat{\phi_\epsilon}(x) = 1$ give us

$$(\widehat{T * S}) = (2\pi)^{n/2} \cdot \widehat{S} \cdot \widehat{T}.$$

This completes the proof. □

4 Appendix

In this part, I am going to deal with those important stuff which I actually do not want to. For instance, I will introduce the inductive limit and the distribution space $\mathfrak{D}(\mathbb{R}^n)$.

Definition 24 (inductive limit)**.** Let X be a vector space, and let $\{X_\alpha\}_{\alpha\in\Lambda}$ be a family of vector subspaces of X such that $\bigcup_{\alpha\in\Lambda} X_\alpha = X$. Suppose that each X_α is a locally convex topological vector space. The family $\{X_\alpha\}$ satisfies the property that, if $X_\alpha \subset X_\beta$, then the an open subset $U \subset X_\alpha$ of X_α is also open in the relative topology on X_β. In other words, the topology of X_β restricted to X_α is identical with X_α. For every convex balanced and absorbing set U, it is defined to be open in X if $U \cap X_\alpha$ is an open subset of X_α. If X is a locally convex linear topological vector space whose topology is defined this way, then we call the inductive limit of X_α.

We now can give the definition of $\mathfrak{D}(\mathbb{R}^n)$.

Definition 25 (distributions)**.** Let K be a compact set in $\mathbb{R}^n$. Let $C_K^\infty(\mathbb{R}^n) \subset C_0^\infty(\mathbb{R}^n)$ be the set of all compactly supported smooth function ϕ whose support $\text{supp}(\phi)$ is a subset of K. We can define the topology on each C_K^∞ through the semi-norms

$$\left\|f\right\|_{K,m} = \sup_{|s|\leq m;\, x\in K} \left|D^s f(x)\right|,$$

then $\mathfrak{D}_K(\mathbb{R}^n)$, the space $C_K^\infty(\mathbb{R}^n)$ with the topology defined by all the semi-norms $\|\cdot\|_{K,m}$, is a locally convex topological vector space.

We define $\mathfrak{D}(\mathbb{R}^n)$ as the inductive limit of $\mathfrak{D}_K(\mathbb{R}^n)$.

Remark. The convergence on $\mathfrak{D}(\mathbb{R}^n)$ is stronger than uniform topology on $C_0^\infty(\mathbb{R}^n)$, that is, if $\lim_{n\to\infty} f_n = f$ in $\mathfrak{D}(\mathbb{R}^n)$ then $f_n \rightrightarrows f$.

Theorem 26. *The convergence in $\mathfrak{D}(\mathbb{R}^n)$ can be characterized by the following conditions:*

1. *There exists a compact set of $\mathbb{R}^n$ such that $\text{supp}(f_n) \subset K$.*
2. *For any n-tuple s, $D^s f_n(x) \rightrightarrows D^s f(x)$ on K.*

The proof is omitted here, however the result is quite useful. Next we are going to define the $\mathfrak{E}(\mathbb{R}^n)$.

Definition 27 (The $\mathfrak{E}^k(\mathbb{R}^n)$ space)**.** Let $C^k(\mathbb{R}^n)$ be the space of all k-th continuously differentiable functions. For any compact set $K \subset \mathbb{R}^n$ and a non-negative integer $m \leq k$, we define the semi-norm by

$$\left\|f\right\|_{K,m} = \sup_{|s|\leq m;\, x\in K} \left|D^s f(x)\right|.$$

Then the set $C^k(\mathbb{R}^n)$ equipped with the topology defined by the semi-norms of the form is denoted by $\mathfrak{E}^k(\mathbb{R}^n)$. We simply write $\mathfrak{E}(\mathbb{R}^n)$ to denote $\mathfrak{E}^\infty(\mathbb{R}^n)$.

Proposition 14. *The convergence $f_n \to f$ in $\mathfrak{E}^k(\mathbb{R}^n)$ is equivalent to the uniform convergence*

$$\lim_{n\to\infty} D^s f_n(x) = f(x)$$

in every compact set $K \subset \mathbb{R}^n$.

Proposition 15. *Note that the space of all compactly supported distributions can be 1-1 correspondence to a continuously linear functional on $\mathfrak{E}(\mathbb{R}^n)$, so we will write $\mathfrak{E}(\mathbb{R}^n)'$ for the set of all compactly supported distributions.*

Reading Reflection

In this report, we explored one of the sections in Yosida's *Functional Analysis* [1]. However, I found this book to be a particularly challenging reference. The content is written in a very concise manner, often leaving me unable to grasp the author's intentions behind certain proofs. As a result, I frequently discussed the parts I didn't understand with my teammates, and through these discussions, I gained a lot of valuable insights.

Beyond these difficulties, understanding unfamiliar sections required me to learn many concepts beyond the scope of the text, such as inductive limits and the support of distributions. This learning process was highly challenging but also expanded my knowledge significantly. Through this report, I came to deeply appreciate that studying mathematics is not always an easy journey. Sometimes, it requires us to confront difficulties and setbacks. Nevertheless, through collaborative discussions with my teammates and thorough study, we were able to overcome these challenges and achieve a deeper understanding.

I believe this experience will have a positive impact on my future learning and research. Despite the obstacles, I am determined to maintain my curiosity and passion for learning while continuously improving my mathematical skills. Lastly, I would like to express my gratitude to our instructor and my teammates for their support and assistance throughout this process. Their encouragement and guidance enabled me to overcome challenges and keep moving forward. This experience has further emphasized the importance of teamwork, and I am convinced that through collective effort, we can achieve even greater accomplishments.

References

[1] K. Yosida. *Functional Analysis*. Classics in Mathematics. Springer Berlin Heidelberg, 2012.