Chapter III

Beyond L^2: Fourier transform of distributions

1 Basic definitions and first examples

In this section we generalize the theory of the Fourier transform developed in Section 1 to distributions. The latter are a generalization of ordinary functions which include objects such as the Dirac delta "function".

The definitions at the beginning of this section may at first sight seem difficult and unmotivated, but will ultimately prove extremely useful. They provide the right setting for many applications[34], as well as for natural theoretical questions left open by the classical theory of Section 1[35], such as seeking to Fourier transform the basic building blocks of Fourier theory, plane waves $e^{ik\cdot x}$. The latter do not belong to L^1 or L^2, and the classical definition (46) leads to a divergent integral, but quite remarkably the framework of distributions allows to give rigorous meaning to, and determine, their Fourier transform (see Example 3).

Definition (Schwartz space of smooth, rapidly decaying functions) $\mathcal{S}(\mathbb{R}^n) := \{\varphi : \mathbb{R}^n \to \mathbb{C} \,|\, \varphi$ infinitely differentiable, $||\varphi||_{\alpha,\beta} := \sup_{x\in\mathbb{R}^n} |x^\beta D^\alpha \varphi(x)| < \infty$ for all α, $\beta \in (\mathbb{N} \cup \{0\})^n\}$.

Here and below the following multi-index notation is used: For $\alpha = (\alpha_1, .., \alpha_n)$, $\beta = (\beta_1, .., \beta_n) \in (\mathbb{N} \cup \{0\})^n$,

$$x^\beta := x_1^{\beta_1} ... x_n^{\beta_n} \quad (x \in \mathbb{R}^n),$$
$$D^\alpha \varphi(x) := \frac{\partial^{\alpha_1}}{\partial_{x_1^{\alpha_1}}} \cdots \frac{\partial^{\alpha_n}}{\partial_{x_n^{\alpha_n}}} \varphi(x),$$

[34] such as those presented in Sections ??, ??, ??, ??

[35] see the discussion at the beginning of Section 3

with the convention that $x_j^0 = 1$, $\frac{\partial^0}{\partial x_j^0}\varphi = \varphi$.

The above finiteness conditions mean that all derivatives of φ must decay faster than polynomial as $|x| \to \infty$. For example, C^∞ functions on $\mathbb{R}^n$ with compact support belong to $\mathcal{S}(\mathbb{R}^n)$, as do functions of form $\varphi(x) = p(x)e^{-x^2}$ for some polynomial p, because their nth derivative is of the form $q_n(x)e^{-x^2}$ for some polynomial q_n.

The fact that in the above definition one requires finiteness of *all* norms $||\cdot||_{\alpha,\beta}$ gives rise to two nice structural properties: the space $\mathcal{S}(\mathbb{R}^n)$ is invariant under differentiation $\varphi \mapsto D^\alpha\varphi$, as well as under multiplication by monomials $\varphi(x) \mapsto x^\beta\varphi(x)$.

Definition (Space of tempered distributions)
$\mathcal{S}'(\mathbb{R}^n) := \{L : \mathcal{S}(\mathbb{R}^n) \to \mathbb{C} \,|\, L \text{ linear, } L \text{ continuous}\}$.[36] Here a map $L : \mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$ is called continuous if for any sequence ϕ_j such that $||\varphi_j - \varphi||_{\alpha,\beta} \to 0$ $(j \to \infty)$ for all α, $\beta \in (\mathbb{N} \cup \{0\})^n$, $L\varphi_j \to L\varphi\}$. The elements of $\mathcal{S}'$ are called **(tempered) distributions**.

Thus a distribution is a map from functions to numbers.

The concept of distributions generalizes the concept of ordinary functions. Each ordinary (say L^p) function can be interpreted as a distribution, but in addition there exist distributions not related to any ordinary function.

Example 1 (Ordinary functions as distributions) Each function $f \in L^p(\mathbb{R}^n)$, $1 \le p \le \infty$, gives rise to a tempered distribution $L_f \in \mathcal{S}'(\mathbb{R}^n)$, as follows:

$$L_f\varphi := \int_{\mathbb{R}^n} f(x)\varphi(x)\,dx \quad \text{for all } \varphi \in \mathcal{S}(\mathbb{R}^n).$$

The verification that $L_f \in \mathcal{S}'(\mathbb{R}^n)$ is left to the reader. Often,

[36] In the language of functional analysis, this means that $\mathcal{S}'$ is defined to be the dual space of $\mathcal{S}$.

the distribution L_f is also denoted f.

Example 2 (The delta "function") For $a \in \mathbb{R}^n$, consider the map $\delta_a : \mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$ defined by

$$\delta_a \varphi := \varphi(a) \quad \text{for all } \varphi \in \mathcal{S}(\mathbb{R}^n).$$

We claim that δ_a belongs to the space $\mathcal{S}'(\mathbb{R}^n)$ of tempered distributions. Indeed, linearity is obvious and the required continuity property can be verified as follows. Suppose φ_j, φ are elements of $\mathcal{S}(\mathbb{R}^n)$ such that $||\varphi_j - \varphi||_{\alpha,\beta} \to 0$ as $j \to \infty$ for all α, $\beta \in (\mathbb{N} \cup \{0\})^n$. Then in particular $||\varphi_j - \varphi||_{0,0} = \sup_{x \in \mathbb{R}^n} |\varphi_j(x) - \varphi(x)| \to 0$ (where 0 denotes the zero vector in $(\mathbb{N} \cup \{0\})^n$). Therefore

$$\underbrace{\varphi_j(a)}_{=\delta_a \varphi_j} - \underbrace{\varphi(a)}_{=\delta_a \varphi} \to 0,$$

as required.

Definition (Fourier transform of a distribution) Let $L \in \mathcal{S}'(\mathbb{R}^n)$. The Fourier transform of L is the map $\widehat{L} : \mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$ defined by

$$\widehat{L}\varphi := L\widehat{\varphi} \quad \text{for all } \varphi \in \mathcal{S}(\mathbb{R}^n).$$

The expression $L\widehat{\varphi}$ is well-defined because of

Lemma 3.1 $\varphi \in \mathcal{S}(\mathbb{R}^n) \Rightarrow \widehat{\varphi} \in \mathcal{S}(\mathbb{R}^n)$.

We postpone the (somewhat technical) proof of this lemma, as well as of the fact that the map $\hat{L}$ defined above again belongs to the space $\mathcal{S}'(\mathbb{R}^n)$ of tempered distributions (see Corollary **??**).

To check that the above definition of the Fourier transform makes sense, let us verify that in the case of an ordinary function viewed as a distribution, it coincides with the ordinary definition (given in Section 4).

Example 1, ctd. (Ordinary functions as distributions) Let $f \in L^1(\mathbb{R}^n)$, and let L_f be the distribution defined by $L_f\varphi = \int_{\mathbb{R}^n} f(x)\varphi(x)\,dx$ for all $\varphi \in \mathcal{S}(\mathbb{R}^n)$. The new definition gives

$$\begin{aligned}
\widehat{L_f}\varphi &= \int_{\mathbb{R}^n} f(x) \underbrace{\widehat{\varphi}(x)}_{=\int_{\mathbb{R}^n} e^{-ik\cdot x}\varphi(k)\,dk} dx \\
&= \int_{\mathbb{R}^n} \underbrace{\left(\int_{\mathbb{R}^n} e^{-ik\cdot x} f(x)\,dx\right)}_{=\widehat{f}(k)} \varphi(k)\,dk \\
&= L_{\widehat{f}}\varphi.
\end{aligned}$$

In other words, when you Fourier transform the distribution associated to f, you get the distribution associated to $\widehat{f}$, where $\widehat{f}$ is the ordinary Fourier transform of f.

Next, let us see what the new definition gives in cases where the ordinary Fourier transform is not defined.

Example 2, ctd. (Delta "function") Let $L = \delta_a$, $a \in \mathbb{R}^n$. Compute

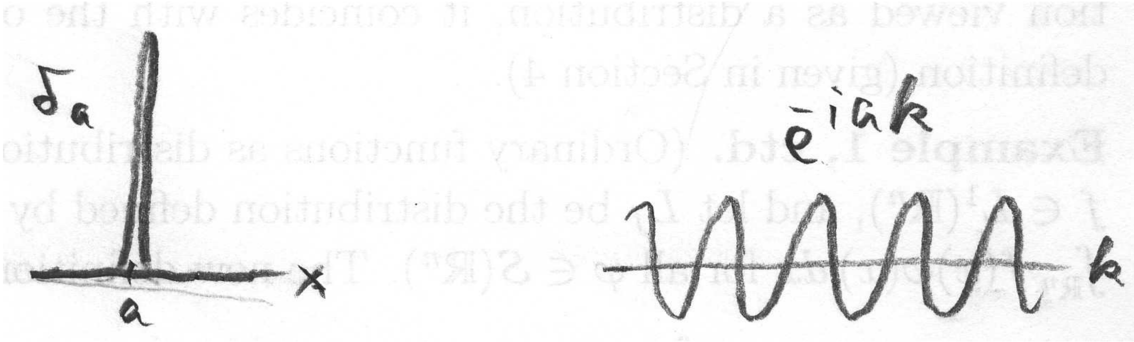
$$\begin{aligned}
\widehat{\delta_a}\varphi &= \delta_a\hat{\varphi}(a) \text{ (by the definition of the F.T. of a distribution)} \\
&= \widehat{\varphi}(a) \text{ (by the definition of the delta "function")} \\
&= \int_{\mathbb{R}^n} e^{-ia\cdot x}\varphi(x)\,dx
\end{aligned}$$

and so

$$\widehat{\delta_a} = L_f, \quad f(x) = e^{-ia\cdot x}. \tag{73}$$

Thus the Fourier transform of a delta function is (the distribution associated to) a plane wave. In common notation (using f instead of L_f), $\widehat{\delta_a}(x) = e^{-ia\cdot x}$.

We remark that this result is a nice manifestation of the Heisenberg uncertainty principle: the delta function has variance



zero, i.e. it is perfectly localized, whereas its Fourier transform has variance infinity, i.e. it is completely delocalized.

Example 3 (Plane wave) Consider a plane wave

$$f(x) = e^{ia\cdot x}.$$

What is its Fourier transform? The classical definition leads to a divergent integral,

$$\text{“}\widehat{f}(k)\text{”} = \text{“}\int_{\mathbb{R}^n} e^{-ik\cdot x} e^{ia\cdot x} dx\text{”}.$$

The occurrence of such a divergent expression stems from the fact that even though f is an ordinary function, it does not satisfy the assumption $\int_{\mathbb{R}^n} |f| < \infty$ required in the classical theory.

Nevertheless, by interpreting f as a distribution we can give rigorous meaning to its Fourier transform. Let L_f be the associated distribution, i.e. $L_f\phi = \int_{\mathbb{R}^n} f(x)\phi(x)\, dx$ for all $\phi \in \mathcal{S}(\mathbb{R}^n)$. The new definition gives

$$\begin{aligned}
\widehat{L_f}\varphi &= L_f\,\widehat{\varphi} \quad \text{(by the definition of the F.T. of a distribution)} \\
&= \int_{\mathbb{R}^n} e^{ia\cdot x}\varphi(x)\, dx \\
&= (2\pi)^n\varphi(a) \quad \text{(by the Fourier representation formula, Thm 2.1)}
\end{aligned}$$

for any $\varphi \in \mathcal{S}(\mathbb{R}^n)$, and therefore

$$\widehat{L_f} = (2\pi)^n\delta_a. \tag{74}$$

Thus the Fourier transform of (the distribution associated to) a plane wave is a "delta" function. In common notation (using f instead of L_f), $\widehat{f} = (2\pi)^n \delta_a$. In particular, by setting $a = 0$,

$$\widehat{1} = (2\pi)^n \delta_0.$$

It is instructive to imagine for a moment that f and $\widehat{f}$ were in L^1 (which they are not...) and think about the result (74) in terms of Fourier decomposition $f(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} \widehat{f}(k) e^{ik\cdot x} dx$. Thus $\widehat{f}(k)$ is the amplitude of the contribution to f by plane waves with wavevector k. Since the plane wave $e^{ia\cdot x}$ consists purely of a contribution with wavevector $k = a$, its Fourier transform should vanish for $k \neq a$, as indeed we have found.

A lengthier but more intuitive approach to deriving the Fourier transforms (73) and (74), based on approximation of the delta "function" by ordinary functions, is described in Appendix A.