Fourier transform of tempered distributions

1 Test functions and distributions

As we have seen before, many functions are not classical in the sense that they cannot be evaluated at any point. For example, a function in L^p is an equivalence class of functions, and one can change its value at any point without changing the function itself. In fact, an L^p function can be defined as a linear functional on its dual space L^q.

Let $\mathcal{D} = \mathcal{D}(\mathbb{R}^d)$ denote the space of functions in $\mathcal{C}_c^\infty(\mathbb{R}^d)$ with the following notion of convergence: a sequence φ_n converges to φ in $\mathcal{D}$ if there exists a compact set K such that φ and all φ_n's vanish outside K, and all derivatives of $(\varphi_n - \varphi)$ converges to 0 *uniformly.* It is clear that $\mathcal{D}$ is a linear space. A *distribution* is a continuous linear map from $\mathcal{D}$ to $\mathbb{C}$. The space of distributions is the dual of $\mathcal{D}$, denoted by $\mathcal{D}' = \mathcal{D}'(\mathbb{R}^d)$. We say a sequence of distributions $\{T_n\}$ converges to T in $\mathcal{D}'$ if for every $\varphi \in \mathcal{D}$, we have

$$T_n(\varphi) \to T(\varphi).$$

It is clear that a distribution is defined via its effects on $\mathcal{D}$, and this is the reason why elements in $\mathcal{D}$ are called test functions. Any locally integrable functions are distributions. In fact, if $f \in L^1_{\text{loc}}$, then T_f defined by

$$T_f(\varphi) = \int_{\mathbb{R}^d} \varphi(x)\overline{f(x)}dx$$

is a distribution. But on the other hand, not all distributions have representatives in classical function space, and a primary example is the so-called Dirac delta function, defined by

$$\delta(\varphi) = \varphi(0).$$

It assigns every test function φ its value at 0. One can easily verify that this is indeed a distribution. Another example is the one-dimensional function $\frac{1}{x}$, which is not integrable near 0. Nevertheless, one can still define a distribution $pv(\frac{1}{x})$ by setting

$$pv(\frac{1}{x})(\varphi) = \int_0^{+\infty} \frac{1}{x}\big(\varphi(x) - \varphi(-x)\big)dx.$$

Here, the notion pv stands for principal value.

One major advantage of distribution over usual functions is that they can be differentiated as many times as we want. To see how to define the derivative of a distribution T, let us first suppose T is given by a smooth function f, so we have

$$T(\varphi) = \int_{\mathbb{R}^d} \varphi(x)\bar{f}(x)dx.$$

A reasonable definition of T' should meet the requirement that the effect of D^jT should be given by D^jf. Since φ is smooth with compact support, an integration by parts yields

$$\int \varphi(x)(D^j\bar{f})(x)dx = -\int (D^j\varphi)(x)\bar{f}(x)dx,$$

which suggests us to define D^jT by

$$(D^jT)(\varphi) = -T(D^j\varphi).$$

We see that $D^j\varphi$ is again smooth and has compact support, so D^jT is also a well-defined distribution. This is called the *weak derivative* or *distributional derivative* of T.

For any multi-index $\alpha = (\alpha_1, \cdots, \alpha_d)$ with α_j's being non-negative integers, the differentiation operator D^α is defined by

$$D^\alpha\varphi = \frac{\partial^\alpha}{\partial x_1^{\alpha_1}\cdots\partial x_d^{\alpha_d}}\varphi.$$

One could then repeat the above integration by parts to define

$$T^\alpha(\varphi) := (-1)^{|\alpha|}T(D^\alpha\varphi),$$

where $|\alpha| = \alpha_1+\cdots+\alpha_d$ counts the number of times of integrating by parts. Furthermore, differentiation is a continuous operation, as can be seen by

$$(D^\alpha T_n)(\varphi) = (-1)^{|\alpha|}T_n(D^\alpha\varphi) \to (-1)^{|\alpha|}T(D^\alpha\varphi).$$

The Dirac delta distribution on $\mathbb{R}$ is the derivative of the Heaviside function

$$H(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases}, \tag{1.1}$$

which is easy to check by definition. Also, the distributional derivative of $\log|x|$ is $pv(1/x)$.

Finally, one should keep in mind that it is in general not possible to define multiplications between distributions.

2 Schwartz space and tempered distributions

The space $\mathcal{D}(\mathbb{R}^d)$ of test functions have all the nice properties, but unfortunately it is not closed under Fourier transform. As you may remember, the Fourier transform of a compactly supported function cannot have compact support, unless it is identically 0. For this reason, we would introduce a new space of test functions, namely the Schwartz space $\mathcal{S}(\mathbb{R}^d)$. It consists of smooth functions on $\mathbb{R}^d$ whose derivatives of all order decay faster than any polynomials. More precisely, $\varphi \in \mathcal{S}(\mathbb{R}^d)$ if

$$\|\varphi\|_{\alpha,\beta} := \sup_{x\in\mathbb{R}^d} |x^\alpha(D^\beta\varphi)(x)| < +\infty$$

for all multi-indices α and β with components non-negative integer valued. Or equivalently, we have $|x^\alpha(D^\beta\varphi)(x)| \to 0$ as $x \to +\infty$.

The topology on $\mathcal{S}$ is given by the following notion of convergence: $\varphi_n \to \varphi$ in $\mathcal{S}$ if $\|\varphi_n - \varphi\|_{\alpha,\beta} \to 0$ for all α, β. Similar as the definition of a distribution, a *tempered distribution* is then defined as a continuous linear functional on $\mathcal{S}$, with the obvious notion of weak convergence as before. The space of tempered distributions is denoted by $\mathcal{S}'(\mathbb{R}^d)$.

Not all distributions are tempered distributions. For example, the function $e^{|x|^2}$ is locally integrable and is hence a distribution, but it is not a tempered distribution (think of why).

Here is the first theorem about the Schwartz space $\mathcal{S}$.

Theorem 1. *If $\varphi \in \mathcal{S}$, then its Fourier transform $\hat{\varphi}$ also belongs to $\mathcal{S}$.*

In fact, if $\varphi \in \mathcal{S}$, one can integrate by parts as many time as one wants to obtain that $\hat{\varphi}$ decays faster than any polynomials. Also, it is easy to see that $\hat{\varphi}$ is smooth and one has

$$(D^\alpha \hat{\varphi})(\xi) = \int_{\mathbb{R}^d} (-2\pi i x)^\alpha \varphi(x) e^{-2\pi i \xi \cdot x} dx,$$

which is again the Fourier transform of another Schwartz function $(-2\pi i x)^\alpha \varphi(x)$ and thus also decays faster than any polynomials. This shows the Fourier transform $\hat{\varphi}$ is again an element in $\mathcal{S}$. With a little more effort, one can also show that the Fourier transform map $\mathcal{F} : \mathcal{S} \to \mathcal{S}$ is a continuous bijection.

We now turn to defining the Fourier transform for elements in $\mathcal{S}'(\mathbb{R}^d)$. In the case $f \in \mathcal{S}'$ is a smooth function, then the multiplication formula says

$$\int \varphi \hat{\bar{f}} dx = \int \hat{\varphi} \bar{f} dx.$$

This suggests us to define the Fourier transform of any element $T \in \mathcal{S}'$ as

$$\hat{T}(\varphi) := T(\hat{\varphi}). \tag{2.1}$$

It can be easily checked that $\hat{T}$ is also an element in $\mathcal{S}'$. Also, one can define in a similar way for element $T \in \mathcal{S}'$ another tempered distribution S such that $\hat{S} = T$. This shows the Fourier transform map is also a bijection from the space of tempered distributions to itself.

We now look at the Fourier transforms of the distributions 1 and δ. By definition, we have

$$\begin{aligned} \hat{\delta}(\varphi) &= \delta(\hat{\varphi}) = \hat{\varphi}(0) = \int \varphi(x) dx = 1(\varphi); \\ \hat{1}(\varphi) &= 1(\hat{\varphi}) = \int \hat{\varphi}(\xi) d\xi = \varphi(0) = \delta(\varphi), \end{aligned}$$

which verifies our relation $\hat{\delta} = 1$ and $\hat{1} = \delta$.

3 Poisson's equation

Another very important example of tempered distribution is the function

$$g(x) = |x|^{-d+\alpha},$$

where $\alpha \in (0, d)$. This function is smooth everywhere except having an integrable singularity at the origin. It is not in L^p for any p, but it belongs to $\mathcal{S}'$, and thus has a Fourier transform. Formally, by scaling, we see

$$\hat{g}(\lambda\xi) = \int |x|^{-d+\alpha} e^{-2\pi i\lambda\xi\cdot x} dx = \lambda^{-\alpha}\hat{f}(\xi),$$

which suggests $\hat{g}(\xi) \sim |\xi|^{-\alpha}$, again a tempered distribution that does not belong to L^p for any p. We now make a rigorous justification according to the definition above, and also get the explicit value of the proportional constant.

Let T be the tempered distribution given by the function $f(x) = |x|^{-d+\alpha}$, and $\hat{T}$ be its Fourier transform defined as (2.1). If the formal relation

$$\hat{f}(\xi) = C|\xi|^{-\alpha} \tag{3.1}$$

holds, then the distribution $\hat{T}$ should be given by the function $\hat{f}$. Thus, in order to justify (3.1), it suffices to prove

$$\int_{\mathbb{R}^d} |x|^{-d+\alpha}\hat{\varphi}(x)dx = C\int_{\mathbb{R}^d} |x|^{-\alpha}\varphi(x)dx.$$

To start, we again use our familiar Gaussian function and have the multiplication formula

$$\int_{\mathbb{R}^d} e^{-\pi|x|^2 t}\hat{\varphi}(x)dx = t^{-\frac{d}{2}}\int_{\mathbb{R}^d} e^{-\pi|x|^2/t}\varphi(x)dx,$$

which holds for every $t > 0$. Here, we have taken advantage of the fact that the Fourier transform of $e^{-\pi|x|^2}$ is itself. Now, multiply both sides by $t^{\lambda-1}$ with λ to be determined later, and integrate t from 0 to $+\infty$, we get

$$\int_{\mathbb{R}^d}\hat{\varphi}(x)\bigg(\int_0^{+\infty} t^{\lambda-1}e^{-\pi|x|^2 t}dt\bigg)dx = \int_{\mathbb{R}^d}\varphi(x)\bigg(\int_0^{+\infty} t^{\lambda-1-\frac{d}{2}}e^{-\pi|x|^2/t}dt\bigg)dx. \tag{3.2}$$

The main goal to introduce the new integration variable t is to get the correct power of $|x|$ out of the integral. This suggests us to take $\lambda = \frac{d-\alpha}{2}$. Combining this choice of λ with the expression of Gamma function

$$\Gamma(\lambda) = \int_0^{+\infty} t^{\lambda-1}e^{-t}dt,$$

and integrating t out, we see (3.2) is simplified to

$$\pi^{-\frac{d-\alpha}{2}}\Gamma\big((d-\alpha)/2\big)\int_{\mathbb{R}^d}|x|^{-d+\alpha}\hat{\varphi}(x)dx = \pi^{-\frac{\alpha}{2}}\Gamma(\alpha/2)\int_{\mathbb{R}^d}|x|^{-\alpha}\varphi(x)dx.$$

This verifies the relation

$$\widehat{|\cdot|^{-d+\alpha}}(\xi) = C_{\alpha,d}|\xi|^{-\alpha},$$

with the constant given by $C_{\alpha,d} = \pi^{\frac{d}{2}-\alpha}\Gamma(\alpha/2)/\Gamma(\frac{d-\alpha}{2})$.

The case of particular interest is $d \geq 3$ and $\alpha = 2$, as the inverse transform of $-\frac{1}{4\pi^2|\xi^2|}$ is the Green's function of the Laplacian. In this case, the above statement says the Fourier transform of the function (viewed as a tempered distribution)

$$K(x) = -\frac{\Gamma(\frac{d}{2}-1)}{4\pi^{\frac{d}{2}}} \cdot |x|^{-d+2}$$

is exactly $-\frac{1}{4\pi^2|\xi|^2}$ for $d \geq 3$. In fact, if we consider the equation

$$\Delta u = f$$

on $\mathbb{R}^d$ with $d \geq 3$, then we can take the Fourier transform in spacial variables to get

$$-4\pi^2|\xi|^2\hat{u}(\xi) = \hat{f}(\xi),$$

and taking the inversion Fourier transform according to the calculations above suggests that the solution could be written as

$$u = K * f.$$

Again, applying Δ to both sides, the left hand side becomes f by definition, and the right hand side equals

$$\Delta(K * f) = (\Delta K) * f = f,$$

which suggests the relation $\Delta K = \delta$. This leads to the concept of a fundamental solution.

Definition 2. *Let $\mathcal{L}$ be a partial differential operator given by*

$$\mathcal{L} = \sum_{|\alpha| \leq n} a_\alpha D^\alpha,$$

where a_α's are constant coefficients. A fundamental solution of $\mathcal{L}$ is a distribution T such that $\mathcal{L}T = \delta$.

Note that the fundamental solution is not unique, as one can always add a constant and obtain another fundamental solution. The notion of fundamental solution is important in the sense that the map

$$T : \mathcal{S} \to \mathcal{S}, \qquad f \mapsto T * f$$

gives an 'inverse' to the operator $\mathcal{L}$. Indeed, one can easily verify the relation $\mathcal{L}T = T\mathcal{L} = \text{id}$. If T is a fundamental solution of $\mathcal{L}$, then $u = T * f$ is a real solution to the equation $\mathcal{L}u = f$.

We now rigorously justify that the function $K(x) = -\frac{1}{4\pi|x|}$ is a fundamental solution of Δ in $d = 3$ dimensions. We need to show that

$$\langle \Delta K, \varphi \rangle := \langle K, \Delta\varphi \rangle = \varphi(0).$$

Since K is smooth everywhere except the origin, we consider the domain Ω given by

$$\Omega = \{x : \epsilon < |x| < R\}$$

with R large enough such that φ vanishes outside $\frac{R}{2}$. Integration by parts gives

$$\int_{\Omega} K\Delta\varphi dx = \int_{\Omega} \Delta K\varphi + \int_{\partial\Omega} K \cdot \frac{\partial\varphi}{\partial\vec{n}} dS - \int_{\partial\Omega} \frac{\partial K}{\partial\vec{n}} \cdot \varphi dS,$$

where $\vec{n}$ is the outward normal direction, and the largeness of R suggests we only need to consider the inner sphere of radius ϵ as $\partial\Omega$. The first term is identically 0 as $\Delta K = 0$ away from the origin. The second one is $\mathcal{O}(\epsilon)$ as $K(x) = \mathcal{O}(1/\epsilon)$ on $\partial\Omega$, while $|\partial\Omega| = 4\pi\epsilon^2$. So it also vanishes as $\epsilon \to 0$. As for the third one, one can easily compute $\frac{\partial K}{\partial\vec{n}}$, and obtain a limit $\varphi(0)$ as $\epsilon \to 0$. The case for $d > 3$ dimensions is essentially the same. We have thus verified that

$$K(x) = -\frac{\Gamma(\frac{d}{2} - 1)}{4\pi^{\frac{d}{2}}} \cdot |x|^{-d+2}$$

is a fundamental solution of Δ in $d \geq 3$ dimensions.

However, the above argument fails as soon as $d = 2$, as $|\xi|^{-2}$ is not locally integrable in two dimensions, and thus does not define a tempered distribution. Nevertheless, by the same integration by parts arguments, we see that

$$K(x) = \frac{1}{2\pi} \log|x|$$

is a fundamental solution of Δ in $d = 2$, and this locally integrable function gives a well defined tempered distribution. The question then is what is its Fourier transform. It turns out that it is a renormalised version of $|x|^{-2}$, defined as

$$\mathscr{R}(|x|^{-2})(\varphi) := \int_{|x|\leq 1} \frac{\varphi(x) - \varphi(0)}{|x|^2} dx + \int_{|x|>1} \frac{\varphi(x)}{|x|^2} dx.$$

It is easy to verify this distribution coincides with $|x|^{-2}$ when applied to test functions φ that vanishes at 0. Also, we see that the renormalised distribution $\mathscr{R}(|x|^{-2})$ could be written as

$$\mathscr{R}(|x|^{-2}) = T(|x|^{-2}) - \infty \cdot \delta,$$

where this '∞' is formally given by $\int_{|x|\leq 1} |x|^{-2} dx$. One might wonder why the cutoff is chosen to be 1; in fact, one can choose any positive real number as a cutoff, and the effect is just adding a finite quantity C to the '∞'. To see how large this ∞ is, we approximate $|x|^{-2}$ by $|x|^{-2+\lambda}$ and send $\lambda \to 0$. We can do it in the general case d. Fix a test function φ, we write

$$\begin{aligned} I(\lambda) &= \int_{\mathbb{R}^d} |x|^{-d+\lambda} \varphi(x) dx \\ &= \frac{A_d}{\lambda} \cdot \varphi(0) + \int_{|x|\leq 1} |x|^{-d+\lambda} \big(\varphi(x) - \varphi(0)\big) dx + \int_{|x|>1} |x|^{-d+\lambda} \varphi(x) dx, \end{aligned}$$

where A_d is the area of the unit sphere in d dimensions. Multiplying λ on both sides and send $\lambda \to 0$, we see that

$$\mathscr{R}(|x|^{-d}) = \big(\lambda I(\lambda)\big)'|_{\lambda=0}, \tag{3.3}$$

the first-order term in the Taylor expansion of $\lambda I(\lambda)$ near $\lambda = 0$. We are now ready to verify the relation

$$\widehat{\log|\cdot|} \sim \mathscr{R}(|x|^{-2}).$$

First, for $\lambda > 0$, we have

$$\int_{\mathbb{R}^2} |x|^{-\lambda}\hat{\varphi}(x)dx = C(\lambda)\int_{\mathbb{R}^2} |x|^{-2+\lambda}\varphi(x)dx = C(\lambda)I(\lambda),$$

where $C(\lambda) = \frac{\lambda}{2\pi} + C'\lambda^2 + \mathcal{O}(\lambda^3)$. Differentiating both sides with respect to λ, and then evaluate at $\lambda = 0$, we obtain (by using (3.3))

$$\int_{\mathbb{R}^2} \hat{\varphi}(x)\log|x|dx = -\frac{1}{2\pi}\left(\int_{|x|\leq 1}\frac{\varphi(x)-\varphi(0)}{|x|^2}dx + \int_{|x|>1}\frac{\varphi(x)}{|x|^2}dx\right) + 2\pi C'\varphi(0).$$

This shows that the Fourier transform of $\frac{1}{2\pi}\log|x|$ is the renormalised distribution

$$-\frac{1}{4\pi^2}\mathscr{R}(|x|^{-2}) + C'\delta,$$

where C' is the coefficient of the second order Taylor expansion of $C(\lambda)$. The appearance of C' indicates the cutoff in defining $\mathscr{R}(|x|^{-2})$ is some other real number instead of 1.

Finally, we prove that the fundamental solution of the heat operator $\mathcal{L} = \partial_t - \Delta$ is the function $F(t,x)$ which equals the heat kernel $H_t(x)$ for $t \geq 0$, and 0 for negative values of t. To see this, let us first note that the distribution $\mathcal{L}F$ is defined via

$$(\mathcal{L}F)(\varphi) := -\langle F, (\partial_t + \Delta)\varphi\rangle,$$

where the integration is understood to be in both space and time. Thus, we have

$$(\mathcal{L}F)(\varphi) = -\lim_{\epsilon\to 0}\int_{t\geq\epsilon}\int_{\mathbb{R}^d} H_t(x)\varphi(t,x)dxdt.$$

Fix positive ϵ, using integration by parts and the relation $\partial_t H = \Delta H$, we get

$$\begin{aligned} -\int_{t\geq\epsilon}\int_{\mathbb{R}^d} H(\partial_t+\Delta)\varphi dxdt &= -\int_{\mathbb{R}^d}\left(\int_{t\geq\epsilon}\frac{\partial}{\partial_t}(H\varphi)dt\right)dx \\ &= \int_{\mathbb{R}^d} H_\epsilon(x)\varphi(\epsilon,x)dx. \end{aligned}$$

Since $\varphi \in \mathcal{S}(\mathbb{R}^{1+d})$, we have

$$\sup_x |\varphi(\epsilon,x) - \varphi(0,x)| \to 0$$

as $\epsilon \to 0$. Also, H_ϵ approximates the delta function (in space), so the last line above converges to $\varphi(0,0)$ as $\epsilon \to 0$, which shows $\mathcal{L}F = \delta$.