The Fourier Transform

Weijun Xu

1 Approximations

We first give some useful approximation results.

Proposition 1.1 (**Continuity of translation in** L^p). *Let* $f \in L^p(\mathbb{R}^n)$*, and define* $f_h(x) = f(x+h)$*. Then,*

$$\|f_h - f\|_p \to 0 \text{ as } h \to 0.$$

Proof. $\forall \epsilon > 0$, we can choose $g \in C_c(\mathbb{R}^n)$ with $\|g - f\| < \epsilon$. Now, we have

$$\|f_h - f\|_p \leq \|f_h - g_h\|_p + \|g_h - g\|_p + \|f - g\|_p .$$

The first and third terms are both less than ϵ, while the second term goes to 0 as $h \to 0$ since g is continuous with compact support. $\square$

Theorem 1.2 (**Approximations by good kernels**). *Let* $K \in L^1(\mathbb{R}^n)$ *with* $\int_{\mathbb{R}^n} K(x)dx = 1$*. Define* $K_\epsilon(x) := \epsilon^{-n} K(x/\epsilon)$*. If* $f \in L^p(\mathbb{R}_n)$*, then* $f * K_\epsilon \in L^p(\mathbb{R}^n)$ *with*

$$\|f * K_\epsilon\|_p \leq \|f\|_p \|K\|_1 .$$

Moreover, we have $f * K_\epsilon \to f$ *in* L^p *as* $\epsilon \to 0$*.*

Proof. Note that $\|K_\epsilon\|_1 = \|K\|_1$. The fist assertion $\|f * K_\epsilon\|_p \leq \|f\|_p \|K\|_1$ is an immediate consequence of Minkowski's inequality. For the second assertion, since $\int_{\mathbb{R}^n} K_\epsilon(x)dx = 1$, we have

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

Apply Minkowski's inequality, we get

$$\|f * K_\epsilon - f\|_p \leq \int_{\mathbb{R}^n} \|f_y - f\|_p |K_\epsilon(y)|dy,$$

where $f_y(x) = f(x-y)$. Note that $\|f_y - f\|_p \to 0$ as $y \to 0$, so $\forall \delta > 0$, $\exists \eta$ such that $\|f_y - f\|_p < \delta$ whenever $|y| < \eta$. Thus, separating $\mathbb{R}^n$ into $\{|y| < \eta\}$ and $\{|y| \geq \eta\}$, we have

$$\|f * K_\epsilon - f\|_p \leq \delta \|K\|_1 + 2\|f\|_p \int_{|y|\geq\eta} |K_\epsilon(y)|dy.$$

Note that the second integral goes to 0 as $\epsilon \to 0$ as long as η is positive, so it follows that $\|f * K_\epsilon - f\|_p \to 0$. $\square$

If $f \in L^1_{loc}$ and $\varphi \in C_c^\infty(\mathbb{R}^n)$, then by dominated convergence, $f * \varphi \in C_c^\infty(\mathbb{R}^n)$ and

$$D^\alpha(f * \varphi) = f * (D^\alpha \varphi).$$

Note that such a φ exists, for example

$$\varphi(x) = \begin{cases} k\exp(-\frac{1}{1-|x|^2}), & |x| < 1 \\ 0, & |x| \geq 1 \end{cases},$$

where the constant k is added to ensure $\int_{\mathbb{R}^n} \varphi = 1$. Since all L^p functions are in L^1_{loc}, the previous approximation theorem immediately implies the following.

Corollary 1.3. *Smooth functions with compact support are dense in L^p.*

2 Basics in L^1 and inversion

Let $f \in L^1(\mathbb{R}^n)$. The Fourier transform of f is defined by

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{-2\pi i \xi \cdot x} dx.$$

The following properties of L^1 Fourier transform are easy to prove:

1. $||\hat{f}||_\infty \leq \|f\|_1$, and $\hat{f}$ is uniformly continuous on $\mathbb{R}^n$;
2. (Scaling). If $\lambda > 0$ and $\delta_\lambda f(x) = f(\frac{x}{\lambda})$, then
$$\widehat{\delta_\lambda f}(\xi) = \lambda^n \hat{f}(\lambda \xi).$$
3. (Convolution). If $f, g \in L^1$, then $\widehat{f * g} = \hat{f}\hat{g}$.

Theorem 2.1 (**Riemann-Lebesgue**)**.** *If $f \in L^1$, then $\hat{f}(\xi) \to 0$ as $|\xi| \to \infty$.*

Proof. The theorem is true when f is an indicator function of some rectangle in $\mathbb{R}^n$; so it is also true when f is a linear combination of such functions. Since step functions are dense in L^1, we can find a sequence of step functions g_k such that $\|g_k - f\|_1 \to 0$. Now, we have

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} (f(x) - g_k(x)) e^{-2\pi i \xi \cdot x} dx + \widehat{g_k}(\xi).$$

Choose k large (independent of ξ) so that the first term is smaller than ϵ, and the second term goes to 0 when $|\xi|$ grows to infinity, so the left hand side can be made arbitrarily small $|\xi| \to \infty$. $\square$

We will frequently use the Fourier transform of a Gaussian; in particular, the Fourier transform of $\exp(-\pi|x|^2)$ is itself. To see this, it suffices to consider the case $x \in \mathbb{R}$. We first assume $\eta \in \mathbb{R}$, and we can complete the squares to get

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

Now consider the left hand side as a function of $\eta \in \mathbb{C}$. Since the integral converges uniformly on all compact sets in $\mathbb{C}$, it is an analytic function and we can get its value by analytic continuation. Thus, setting $\eta = i\xi$ gives us the desired answer. Now, let

$$\varphi_\epsilon(x) = \exp(-\epsilon|x|^2),$$

then a simple scaling argument yields

$$\widehat{\varphi_\epsilon}(\xi) = \left(\frac{\pi}{\epsilon}\right)^{n/2} \exp(-\pi^2|\xi|^2/\epsilon). \tag{1}$$

We now turn to the question of inversion. Clearly, we have

$$\int_{\mathbb{R}^n} \hat{f}(\xi) e^{2\pi i\xi\cdot x} d\xi = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(y) e^{-2\pi i\xi\cdot(y-x)} dy d\xi. \tag{2}$$

If we change the order of integration on the right hand side, then we will get $\int_{\mathbb{R}^n} f(y)\delta(x-y)dy$, which is exactly $f(x)$ whenever f is continuous at x. But this change of order of integration is just at an informal level, and we need to carefully justify this heuristic.

Theorem 2.2 (L^1 **Inversion**). *If both f and $\hat{f}$ are in L^1, then*

$$f(x) = \int_{\mathbb{R}^n} \hat{f}(\xi) e^{2\pi i\xi\cdot x} d\xi \quad a.e.$$

Proof. In order to apply Fubini's theorem, we multiply the Gaussian function $\exp[-\epsilon|\xi|^2]$ on both sides of (2). For the left hand side, by dominated convergence theorem, we have

$$\int_{\mathbb{R}^n} \hat{f}(\xi) e^{2\pi i\xi\cdot x} e^{-\epsilon|\xi|^2} d\xi \to \int_{\mathbb{R}^n} \hat{f}(\xi) e^{2\pi i\xi\cdot x} d\xi, \ \ \forall x$$

as $\epsilon \to 0$. Note that this is where we have used $\hat{f} \in L^1$. For the right hand side, we apply Fubini's theorem and implement (1), we get

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(y) e^{-2\pi i\xi\cdot(y-x)} e^{-\epsilon|\xi|^2} dy d\xi = (f * G_\epsilon)(x),$$

where $G_\epsilon(y) = \left(\frac{\pi}{\epsilon}\right)^{n/2} \exp(-\pi^2|y|^2/\epsilon)$, and we have $f * G_\epsilon \to f$ in L^1. Thus, we can substract a subsequence which converges a.e. to f. The proof of the theorem is complete by taking $\epsilon \to 0$ along that subsequence. □

As an immediate consequence, we obtain the following corollary. Note that it is definitely not true for general L^1 functions.

Corollary 2.3. *If both f and $\hat{f}$ are in L^1, then f can be modified on a set of measure* 0 *to be a continuous and bounded function.*

3 L^2 theory

If $f \in L^2$, it is not immediately obvious how to define its Fourier transform. We shall however start with functions in $L^1 \cap L^2$.

Theorem 3.1 (**Plancherel**). *If $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, then $\hat{f} \in L^2$ and we have*

$$||\hat{f}||_2 = \|f\|_2 .$$

Proof. The proof also involves approximating the delta function by Gaussian kernels. By monotone convergence theorem, we have

$$\int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 \exp(-\epsilon|\xi|^2) d\xi \to ||\hat{f}||_2^2, \tag{3}$$

with the right hand side possibly being ∞. We now need to show that it also converges to $\|f\|_2^2$. Note that we can write $|\hat{f}(\xi)|^2 = \hat{f}(\xi)\overline{\hat{f}(\xi)}$ in terms of Fourier integrals, and apply Fubini's theorem to integrate ξ out first to we get

$$\int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 \exp(-\epsilon|\xi|^2) d\xi = \int_{\mathbb{R}^n} f(x) \left(\int_{\mathbb{R}^n} \bar{f}(x-y) G_\epsilon(y) dy \right) dx.$$

By a similar argument as before, we can show that the right hand side converges to $\|f\|_2$. This implies $\hat{f} \in L^2$ and $||\hat{f}||_2 = \|f\|_2$. $\square$

Now let $f \in L^2$, and choose a sequence $\{f_j\} \in L^1 \cap L^2$ such that $\|f_j - f\|_2 \to 0$. Since $||\widehat{f_j} - \widehat{f_k}||_2 = \|f_j - f_k\|_2$, $\{\widehat{f_j}\}$ is also a Cauchy sequence in L^2, and hence we can define

$$\hat{f} := \lim_{j \to +\infty} \widehat{f_j}$$

It is clear that the limit is independent of the choice of the sequence f_j. We thus **define** this $\hat{f}$ to be the Fourier transform of the L^2 function f. Moreover, we have

$$||\hat{f}||_2 = \lim_{j \to +\infty} ||\hat{f}_j||_2 = \lim_{j \to +\infty} \|f_j\|_2 = \|f\|_2 .$$

We have now defined the Fourier transform $\mathcal{F}$ on L^2, which is a linear **isometry** from $L^2(\mathbb{R}^n)$ into Ł$^2(\mathbb{R}^n)$. The next theorem asserts that this map is **unitary**: it is onto and invertible! The proof of this fact is by approximations in a similar way as before.

Theorem 3.2. *The Fourier transform $\mathcal{F}$ on $L^2(\mathbb{R}^n)$ is unitary, with inverse*

$$(\mathcal{F}^{-1} f)(x) = (\mathcal{F} f)(-x), \quad f \in L^2(\mathbb{R}^n).$$

Example 3.3. *Let $f \in L^2(\mathbb{R})$, and suppose there exists $g \in L^2(\mathbb{R})$ such that $f' = g$ weakly. We claim that f can be modified on a set of measure 0 to be a continuous function. In fact, we have*

$$2\pi i \xi \hat{f}(\xi) = \hat{g}(\xi) \in L^2(\mathbb{R}).$$

Write

$$\int_{\mathbb{R}} |\hat{f}| d\xi = \int_{|\xi \leq 1|} |\hat{f}| d\xi + \int_{|\xi|>1} |\hat{f}| d\xi.$$

The first term on the right hand side is finite since $\hat{f} \in L^2$ and the set $\{|\xi| \leq 1\}$ has finite measure. For the second term, we have

$$\int_{|\xi|>1} |\hat{f}| d\xi \leq \left(\int_{|\xi|>1} \frac{1}{\xi^2} d\xi \right)^{\frac{1}{2}} \left(\int_{|\xi|>1} \xi^2 |\hat{f}(\xi)|^2 d\xi \right)^{\frac{1}{2}} < +\infty.$$

Thus, $\hat{f} \in L^1(\mathbb{R})$ and the claim follows. Note that f is more than merely being continuous. In fact, f can be modified on a set of measure 0 such that it is Hölder-$\frac{1}{2}$, a consequence of Morrey's inequality.

Remark 3.4. *The above example is not true if $n = 2$. Let $\alpha \in (0, \frac{1}{2})$, and consider for example the function $f(x) = \big(\log(1/|x|)\big)^{\alpha}$ for $|x| < \frac{1}{2}$, $f(x) = 0$ for $|x| \geq 1$ and 'smoothly interpolate' between. Then, $f \in L^2$, and both $\frac{\partial}{\partial x_1} f$ and $\frac{\partial}{\partial x_2} f$ are in L^2 weakly, but f cannot be modified on a set of measure 0 to a continuous function.*

4 Extension to L^p?

To proceed to L^p, one might think of mimicing the L^2 approach: first defining it on the dense subset $L^1 \cap L^p$, then extending it to all of L^p. The essence is whether there exists a q and $C = C(n, p, q)$ such that the following inequality holds

$$||\hat{f}||_q < C \, \|f\|_p \,, \qquad \forall f \in L^1 \cap L^p. \tag{4}$$

One first observes that if such an inequality holds, then q cannot be arbitrary. In fact, $\frac{1}{p} + \frac{1}{q} = 1$ is the only possible choice of q. To see this, suppose $f \in L^1 \cap L^p$, and let $f_\lambda(x) = f(\lambda x)$, then f_λ must satisfy (4) with the same constant C for all λ. Also, we have

$$||\widehat{f_\lambda}||_q = \lambda^{-n(1-\frac{1}{q})} \left\| \hat{f} \right\|_q, \qquad \|f_\lambda\|_p = \lambda^{-\frac{n}{p}} \|f\|_p \,.$$

By sending λ to 0 and to $+\infty$ respectively, one can easily see that the only choice for q is the conjugate of p. Our second observation is that no such inequality exists for $p > 2$. To see this, consider

$$f(x) = \exp\big[-(a + bi)|x|^2 \big], \;\; a > 0.$$

By sending $b \to \infty$, we can see that (4) cannot hold if $p > 2$. Nevertheless, we have a positive result for $p \in (1, 2)$.

Theorem 4.1 (**Hausdorff - Young**). *Let $p \in (1, 2)$ and $\frac{1}{p} + \frac{1}{q} = 1$. Then, there exists $C = C(n, p, q)$ such that*

$$||\hat{f}||_q < C \, \|f\|_p \,, \qquad \forall f \in L^1 \cap L^p.$$

Remark 4.2. *The Fourier transform for general L^p functions can be defined via tempered distributions. These turn out to be equivalent with the above approach when $p \in [1, 2]$.*

5 The heat equation

We now apply some properties of the Fourier transform to solve the heat equation

$$\partial_t u = \Delta u, \qquad u(x,0) = g(x) \tag{5}$$

for the unknown $u : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}$, where $g \in \mathcal{S}(\mathbb{R}^n)$. Since the Fourier transform turns differentiation into multiplication, taking the transform on both sides with the spatial variable yields

$$\hat{u}'(\xi,t) = -4\pi|\xi|^2\hat{u}(\xi,t), \qquad \hat{u}(\xi,0) = \hat{g}(\xi),$$

where the differentiation on the left hand side is with respect to t. For each $\xi \in \mathbb{R}^n$, this is an ODE in t and the solution is

$$\hat{u}(\xi,t) = \hat{g}(\xi)\exp\big(-4\pi|\xi|^2 t\big). \tag{6}$$

Note that the second term is the Fourier transform of the heat kernel

$$H_t(x) = \big(4\pi t\big)^{-n/2} e^{-\frac{|x|^2}{4t}},$$

so we actually have

$$u(x,t) = (g * H_t)(x).$$

One can check by differentiation under the integral sign that this indeed solves the heat equation (5).

Remark 5.1. *The expression on the right hand side of* (6) *suggests that t cannot be negative. This amounts to the fact that given the initial value at $t = 0$, one cannot solve the heat equation backwards. In other words, the heat flow cannot be reversed.*

Remark 5.2. *If the initial condition $g \geq 0$ and is strictly positive somewhere, then as long as $t > 0$, $u(x,t) = (g * H_t)(x)$ is strictly positive for every x. This says that the heat flow has infinite propogation.*

Let us now consider the inhomogeneous heat equation

$$\partial_t u = \Delta u + f, \qquad u(x,0) = g(x). \tag{7}$$

We again assume for simplicity that both f and g are smooth and all derivatives have rapid decay. Same as before, taking Fourier transform on both sides with respect to spatial variable gives rise to an ODE with unknown $\hat{u}(\xi,t)$ in t. Solving this ODE and inverting the Fourier transform gives

$$u(x,t) = (g * H_t)(x) + \int_0^t \int_{\mathbb{R}_n} f(y,s) H_{t-s}(x-y)\,dyds.$$

Remark 5.3. *Note that for every s, the function*

$$v_s(x,t) = \int_{\mathbb{R}^n} f(y,s)H_{t-s}(x-y)dy = (f(\cdot,s) * H_{t-s})(x)$$

solves the homogeneous problem

$$\partial_t v_s = \Delta v_s, \qquad v_s(x,s) = f(x,s)$$

for $t \geq s$. Integrating v_s from 0 to t gives the solution to the original equation (7) *with initial condition $g = 0$. This is an example of Duhamel's principle.*