The Fourier transform (L^1 and L^2)

1 Approximations

A family of kernels $\{K_\delta\}$ on $\mathbb{R}^d$ is called a family of good kernels if

1. $\int_{\mathbb{R}^d} K_\delta(x)dx = 1$. This is what we mean by 'kernel'.
2. There exists $M > 0$ such that $\int_{\mathbb{R}^d} |K_\delta(x)|dx < M$ for all δ.
3. For any $\eta > 0$, we have $\int_{|x|>\eta} |K_\delta(x)|dx \to 0$.

We recall that both Fejér kernel $\{F_n\}$ and heat kernel on the circle $\{p_t\}$ are families of good kernels. The main use of these kernels is that under suitable conditions of f, $f * K_\delta$ converges to f in various senses.

First of all, if f is bounded, then $(K_\delta * f)(x) \to f(x)$ whenever x is a point of continuity of f. The proof is exactly the same as the case for $\{F_n\}$. Below, we will show that, for $f \in L^p$, we have $f * K_\delta \in L^p$ with $\|f * K_\delta\|_{L^p} \leq \|K_\delta\|_{L_1}\|f\|_{L^p}$, and in fact we have $f * K_\delta \to f$ in L^p. The first bound comes as a direct consequence of Minkowski's inequality, and hence $f * K_\delta \in L^p$. For the convergence, we have

$$\begin{aligned}\|f * K_\delta - f\|_p &= \left(\int \left| \int \big(f(x-y) - f(x)\big)K_\delta(y)dy\right|^p dx\right)^{\frac{1}{p}} \\ &\leq \int |K_\delta(y)| \cdot \|f_y - f\|_p dy,\end{aligned}$$

where we have used the Minkowski's inequality, and set $f_y(x) = f(x-y) - f(x)$ as the translation. Since K_δ concentrates nearly all its mass near the origin as $\delta \to 0$, it is reasonable to separate the above integral into regions near and away from the origin. To bound the integral with domain near the origin, we need the following lemma for the continuity of translations.

Lemma 1 (Continuity of translations in L^p)**.** *Let $f \in L^p$ and set $f_y(x) = f(x-y) - f(x)$, then $\|f_y - f\|_p \to 0$ as $y \to 0$.*

Proof. This lemma clearly holds for any $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$. Since smooth functions with compact support are dense in L^p, we can find such a φ that $\|f - \varphi\|_p < \epsilon$. Now, we have

$$\|f_y - f\|_p \leq \|f_y - \varphi_y\|_p + \|\varphi_y - \varphi\|_p + \|\varphi - f\|_p.$$

The first and third term are both smaller than ϵ because of the choice of φ. The second term goes to 0 as $y \to 0$ as $\varphi \in \mathcal{C}_c^\infty$. □

We can now prove the convergence of $f * K_\delta \to f$ in L^p. For any $\epsilon > 0$, choose η such that $\|f_y - f\|_p < \epsilon$ whenever $|y| \le \eta$. Thus, we have

$$\|f * K_\delta - f\|_p \le \int_{|y|\le\eta} \|f_y - f\|_p |K_\delta(y)| dy + \int_{|y|>\eta} \|f_y - f\|_p |K_\delta(y)| dy.$$

The first term is smaller than $M\epsilon$, uniformly in δ. The second term goes to 0 as $\delta \to 0$ by the third property of good kernels. We have thus established the following.

Proposition 2. *If $\{K_\delta\}$ is a family of good kernels and $f \in L^p$, then $f * K_\delta \to f$ in L^p.*

A common family of good kernels we will be using are generated from rescaling L^1 functions. If $K \in L^1$ with $\int K(x)dx = 1$, then the rescaled version

$$K_\delta(x) = \delta^{-d} K(x/\delta)$$

is actually a family of good kernels.

2 L^1 theory

If $f : \mathbb{R} \to \mathbb{R}$ is periodic with period L, then $f_L(x) = f(Lx)$ has period 1. If f is smooth enough, by a change of variable, we can write

$$f_L(x) = \sum_{n\in\mathbb{Z}} \frac{1}{L} \left(\int_{-\frac{L}{2}}^{\frac{L}{2}} f(y) e^{-2\pi i \frac{n}{L} y} dy \right) e^{2\pi i n x / L}.$$

We see as L becomes very large, the right hand side is an approximation to the Riemann sum on the real line with step size $\frac{1}{L}$. Thus, we formally get

$$f(x) = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(y) e^{-2\pi i \xi y} dy \right) e^{2\pi i \xi x} dx. \tag{2.1}$$

This is the continuous analogue of the case studied in Fourier series. Just as in Fourier series where we explore in what sense do we have

$$f(x) = \sum_n \left(\int_0^1 f(y) e^{2\pi i n y} dy \right) e^{2\pi i n x},$$

the first main issue of its continuous version is whether (2.1) holds, and if so, in what sense.

Below, we will work with $\mathbb{R}^d$ instead of $\mathbb{R}$. For $f \in L^1$, its *Fourier transform* $\hat{f}$ is defined by

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

We will frequently be using the properties below. The proofs are easy and are left as an exercise. All functions concerned are in L^1.

1. $\hat{f}$ is uniformly continuous on $\mathbb{R}^d$, and $\|\hat{f}\|_\infty \le \|f\|_1$.
2. Scaling: if $f_\lambda(x) = \lambda^{-d} f(x/\lambda)$, then $\hat{f}_\lambda(\xi) = \hat{f}(\lambda\xi)$.
3. Convolution: $\widehat{f * g}(\xi) = \hat{f}(\xi)\hat{g}(\xi)$.
4. Riemann-Lebesgue lemma: $\hat{f}(\xi) \to 0$ as $\xi \to +\infty$.
5. Multiplication: $\int \hat{f}(x)g(x)dx = \int f(x)\hat{g}(x)dx$.

It should be noted that although $\hat{f}$ is bounded and continuous, there is no guarantee for its integrability. For example, if $f(x) = 1_{[a,b]}(x)$, then $\hat{f}$ is not in L^1.

Gaussian functions are of interest in many places. In particular, the Fourier transform of $e^{-\pi|x|^2}$ is itself. To see this, it suffices to consider the case $x \in \mathbb{R}$. If $\eta \in \mathbb{R}$, we can complete the squares to get

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

The left hand side (as a function in $\eta \in \mathbb{C}$) converges uniformly on all compact sets in $\mathbb{C}$ (as an integral over the whole real line), thus it is an analytic function in η and its value at $\eta = i\xi$ is just $e^{-\pi\xi^2}$. For the multidimensional case, we simply have

$$\int_{\mathbb{R}^d} e^{-\pi|x|^2} e^{-2\pi i\xi\cdot x} dx = \prod_{j=1}^{d} \int_{\mathbb{R}} e^{-\pi x_j^2} e^{2\pi i\xi_j x_j} dx_j = e^{-\pi|\xi|^2}.$$

We now turn back to the question of Fourier inversion. We have

$$\int \hat{f}(\xi) e^{2\pi i\xi\cdot x} d\xi = \int\int f(y) e^{-2\pi i\xi\cdot(y-x)} dy d\xi.$$

Formally, if we could just change the order of integration, and use the fact that the Fourier transform of a constant function is a delta function, then we get the inversion formula whenever x is a point of continuity of f. This formal reasoning can be justified when both f and $\hat{f}$ are in L^1, in which case we have the following inversion theorem.

Theorem 3. *If both f and $\hat{f}$ are in L^1, then we have*

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

for almost every x. As a consequence, f can be modified on a set of measure 0 to be a uniformly continuous and bounded function.

Proof. We multiply the Gaussian function $e^{-\delta|\xi|^2}$ to the integrand on the right hand side. By dominated convergence theorem, we have

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

as $\delta \to 0$. This is where we need $\hat{f} \in L^1$. On the other hand, we could write down the expression of $\hat{f}(\xi)$, and use Fubini's theorem to change the order of integration to get

$$\int_{\mathbb{R}^d} \hat{f}(\xi) e^{-2\pi i\xi\cdot x} e^{-\delta|\xi|^2} d\xi = (f * \hat{G}_\delta)(x),$$

where $\hat{G}_\delta$ is the Fourier transform of the scaled Gaussian function $e^{-\delta|\xi|^2}$, so it has the form

$$\hat{G}_\delta(y) = (\pi/\delta)^{\frac{d}{2}} \exp(-\pi^2|y|^2/\delta).$$

It is also a rescaled Gaussian function by a scaling of order $\sqrt{\delta}$. Clearly, $\hat{G}_\delta$ is a family of good kernels and we have $f * \hat{G}_\delta \to f$ in L^1. It follows that there is a subsequence that converges almost everywhere. Taking the limit along that subsequence gives the inversion for almost every x. □

3 L^2 theory

The main tool for defining Fourier transform in L^2 is the following.

Theorem 4. *If $f \in L^1 \cap L^2$, then $\hat{f} \in L^2$ with*

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

Proof. Similar as before, we multiply $e^{-\delta|\xi|^2}$ to the integrand on the right hand side, and send δ to 0 to obtain

$$\int |\hat{f}(\xi)|^2 e^{-\delta|\xi|^2} d\xi \to \int |\hat{f}(\xi)|^2 d\xi.$$

Also, one could write down the integrals defining $\hat{f}$, and use Fubini's theorem to change the order of integration to get

$$\int |\hat{f}(\xi)|^2 d\xi = \int f(x)(\bar{f} * \hat{G}_\delta)(x) dx,$$

where $\hat{G}_\delta$ is the same kernel as before. It thus follows that $\bar{f} * \hat{G}_\delta \to \bar{f}$ in L^2. Finally, by Cauchy-Schwarz, we have

$$\left| \int f(x)\big(\bar{f} * \hat{G}_\delta(x) - \overline{f(x)}\big) dx \right| \leq \|f\|_2 \cdot \|\bar{f} * \hat{G}_\delta - \bar{f}\|_2 \to 0.$$

This completes the proof. □

We now use a density argument to define $\hat{f}$ for $f \in L^2$. In fact, there exists a sequence $\{f_n\}$ belonging to $L^1 \cap L^2$ such that $\|f_n - f\|_2 \to 0$. Since Fourier transform is a linear operation, we have

$$\|\hat{f}_n - \hat{f}_m\|_2 = \|\widehat{f_n - f_m}\|_2 = \|f_n - f_m\|_2 \to 0,$$

so $\{\hat{f}_n\}$ is a Cauchy sequence in L^2, and there exists a unique $\hat{f} \in L^2$ such that

$$\hat{f} = \lim_{n\to+\infty} \hat{f}_n \quad \text{in } L^2.$$

It now remains to show that this limit is *independent* of the approximating sequence. If $\{g_n\}$ is another approximating sequence to f in L^2 with $\hat{g}$ being the L^2 limit of $\{\hat{g_n}\}$, then we have

$$\|\hat{f}_n - \hat{g_n}\|_2 = \|f_n - g_n\|_2 \to 0,$$

which gives $\hat{f} = \hat{g}$.

The Plancherel's identity $\|\hat{f}\|_2 = \|f\|_2$ obviously extends to $f \in L^2$. Moreover, if $g \in L^2$, then by writing

$$\langle f, g\rangle = \frac{1}{4}\Big(\|f+g\|_2^2 - \|f-g\|_2^2 + i\big(\|f+ig\|_2^2 - \|f-ig\|_2^2\big)\Big)$$

and applying the special case of Plancherel's identity to each of the four terms on the right hand side, we obtain

$$\langle f, g\rangle = \langle \hat{f}, \hat{g}\rangle$$

for any $f, g \in L^2(\mathbb{R}^d)$. We now summarize what we have proved in the theorem below.

Theorem 5 (**Plancherel's theorem**)**.** *The map $f \mapsto \hat{f}$ defined initially from $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ to $L^2(\mathbb{R}^d)$ has a unique extension of its domain to the whole of $L^2(\mathbb{R}^d)$. In addition, the extended map $\mathcal{F}$ is a continuous linear isometry from L^2 to L^2 with*

$$\langle f, g\rangle = \langle \mathcal{F}f, \mathcal{F}g\rangle.$$

As it turns out, the Fourier transform map $\mathcal{F} : L^2 \to L^2$ defined above is not only an isometry, but also a unitary transformation. Its range is also whole of L^2 and the inverse is given by

$$(\mathcal{F}^{-1}f)(x) = (\mathcal{F}f)(-x).$$

This amounts to changing the $-i$ into i, and is well-defined for any $g \in L^2$, as can be shown using the same approximation arguments above. It now remains to show that it indeed defines an inverse; that is, we want to show

$$\mathcal{F}^{-1}\mathcal{F}f = \mathcal{F}\mathcal{F}^{-1}f = f$$

It is clear that $\mathcal{F}^{-1}$ defines an inverse if $f \in \mathcal{S}(\mathbb{R}^d)$. For general $f \in L^2$, we choose a sequence $\{f_n\} \in \mathcal{S}(\mathbb{R}^d)$ such that $f_n \to f$ in L^2. Then, we have

$$\mathcal{F}^{-1}\mathcal{F}f_n = f_n$$

for all n. The right hand side converges to f in L^2 as $n \to +\infty$. For the right hand side, we have $\mathcal{F}f_n \to \mathcal{F}f \in L^2$, so the continuity of $\mathcal{F}^{-1}$ gives $\mathcal{F}^{-1}\mathcal{F}f = f$. The same is true for $\mathcal{F}\mathcal{F}^{-1}f$. We have thus proved the following.

Theorem 6 (**L^2 inversion**)**.** *The map $\mathcal{F} : L^2 \to L^2$ is a unitary transformation with inverse $(\mathcal{F}^{-1}f)(x) = (\mathcal{F}f)(-x)$.*