6. Fourier transform - a quick introduction

We give a short introduction to the basic properties of the Fourier transform[6]. For $f \in L^1(\mathbb{R})$ (all our functions are complex-valued), define its Fourier transform $\hat{f} : \mathbb{R} \mapsto \mathbb{C}$ by $\hat{f}(\lambda) = \int\limits_{\mathbb{R}} f(x) e^{-i\lambda x} dx$.

(1) $f \to \hat{f}$ is linear. Further,

$$\hat{\bar{f}}(\lambda) = \overline{\hat{f}(-\lambda)},$$

$$\hat{f}_\tau(\lambda) = \hat{f}(\lambda) e^{-i\lambda\tau} \text{ if } f_\tau(x) = f(x-\tau),$$

$$\hat{f}_\sigma(\lambda) = \hat{f}(\lambda\sigma) \text{ if } f_\sigma(x) = \frac{1}{\sigma} f(x/\sigma).$$

All these are easy exercises.

(2) If $f \in L^1$ then $\hat{f}$ is uniformly continuous on $\mathbb{R}$ and bounded by $\|f\|_{L^1}$.

Boundedness is obvious. To see continuity, note that

$$\begin{aligned} |\hat{f}(\lambda) - \hat{f}(\mu)| &= \left| \int f(x) [e^{-i\lambda x} - e^{-i\mu x}] du \right| \\ &\le \int |f(x)|.|e^{i(\lambda-\mu)x} - 1| du \\ &\le 10M|\lambda - \mu| \int_{-M}^{M} |f(u)| + 2\int_{|x|>M} |f(x)| dx \end{aligned}$$

since $|e^{it} - 1| \le 2 \wedge 10|t|$ for all $t \in \mathbb{R}$ (we use the bound $10t$ for small $|t|$ and the bound 2 for large $|t|$). Thus, given $\varepsilon > 0$, first choose M large so that $\int_{|x|>M} |f(x)| dx < \varepsilon$ (possible since $f \in L^1$) and then choose $\delta = \frac{\varepsilon}{10M\|f\|_{L^1}}$. Then for $|\lambda - \mu| < \delta$, the first term above will be bounded by ε too. Thus, whenever $|\lambda - \mu| < \delta$ we have $|\hat{f}(\lambda) - \hat{f}(\mu)| < 2\varepsilon$.

(3) *Riemann-Lebesgue lemma:* If $f \in L^1$, then $\hat{f}(\lambda) \to 0$ as $|\lambda| \to \infty$.

If f is a step function, $f = \sum_{k=1}^{n} c_k \mathbf{1}_{[a_k, b_k]}$ for some $c_k \in \mathbb{C}$ and $a_1 < b_1 < a_2 < b_2 < \ldots < a_n < b_n$, then by actual integration

$$\hat{f}(\lambda) = \frac{1}{i\lambda} \sum_{k=1}^{n} c_k (e^{-i\lambda a_k} - e^{-i\lambda b_k}).$$

Clearly $\hat{f}(\lambda) \to 0$ as $|\lambda| \to \infty$ (in fact it decays like $1/|\lambda|$). Now if $f \in L^1$ is any function, find a step function g such that $\|f - g\|_{L^1} < \varepsilon$. Then, $|\hat{f}(\lambda) - \hat{g}(\lambda)| \le \varepsilon$. Since g is a step function, there exists a finite M such that $|\hat{g}(\lambda)| < \varepsilon$ if $|\lambda| > M$. Then, for $|\lambda| > M$ we have $|\hat{f}(\lambda)| \le |\hat{g}(\lambda)| + \varepsilon < 2\varepsilon$.

(4) If $f, g \in L^1$, then $f * g$ is also in L^1 and $(f \hat{*} g) = \hat{f}\hat{g}$. Indeed, if we define $h : \mathbb{R}^2 \to \mathbb{C}$ by $h(x,y) = f(x-y)g(y)e^{-i\lambda x}$, then it is easy to see that $\|h\|_{L^1(\mathbb{R}^2)} = \|f\|_{L^1(\mathbb{R})} \|g\|_{L^1(\mathbb{R})}$. Hence, by Fubini's theorem, we may compute $\int\limits_{\mathbb{R}^2} h$ as an iterated integral in two ways. On the one hand, it is equal to

$$\int\int f(x-y)g(y)e^{-i\lambda x} dy dx = \int (f*g)(x) e^{-i\lambda x} dx = (f \hat{*} g)(\lambda).$$

On the other hand it is equal to

$$\int\int f(x-y)g(y)e^{-i\lambda x} dx dy = \int \hat{f}(y) g(y) e^{-i\lambda y} dy = \hat{f}(\lambda)\hat{g}(\lambda).$$

Thus, $(f \hat{*} g) = \hat{f}\hat{g}$.

---

[6]There are many good books. We recommend Katznelson's *An introduction to harmonic analysis*, section 1, chapter VII. Feller's *An introduction to probability theory and its applications, volume-II* also has a nice chapter on this subject, except that Fourier transforms are defined not only for functions but measures.

(5) *Parseval's relation:* If $f,g \in L^1$, then $f\hat{g}$ and $g\hat{f}$ are also in L^1 and $\int_{\mathbb{R}} \hat{f}g = \int_{\mathbb{R}} f\hat{g}$.

Since $f \in L^1$ and g is bounded it follows that $\hat{g}f$ is in L^1. Similarly for $\hat{f}g$. Next, consider $h(x,y) = f(x)g(y)e^{ixy}$. This is in $L^1(\mathbb{R}^2)$ (in fact $\|h\|_{L^1(\mathbb{R}^2)} = \|f\|_{L^1}\|g\|_{L^1}$. By Fubini's theorem, we may get the double integral by integrating w.r.t. x and then w.r.t y or vice versa. We get $\int \hat{f}g$ or $\int f\hat{g}$ in the two cases.

(6) *Plancherel's identity:* Suppose $f \in L^1 \cap L^2$. Then $\hat{f} \in L^2$ and $\|\hat{f}\|_{L^2}^2 = 2\pi\|f\|_{L^2}^2$. If we assume a bit more about f (just assume that $f, \hat{f}$ are both smooth and decay fast), then the Plancherel identity follows from Parseval's relation and Fourier inversion (proved below). The general case follows by approximation.

Exercise 1. If $f,g \in L^1$, use Parseval's relation to show that $(f * \hat{g})(x) = \int_{\mathbb{R}} \hat{f(\lambda)} g(-\lambda)e^{i\lambda x}d\lambda$ (why is $(f*\hat{g})(x)$ well-defined?) .

Example 2. If $f(x) = \frac{1}{2a}\mathbf{1}_{[-a,a]}(x)$, then check that $\hat{f}(\lambda) = \frac{\sin(\lambda a)}{\lambda a}$.

Example 3. If $f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-x^2/2\sigma^2}$, then check that $\hat{f}(\lambda) = e^{-\sigma^2\lambda^2/2}$. [**Hint:** Use contour integration]

Example 4. If $f(x) = e^{-|x|}$, then by direct integration, check that $\hat{f}(\lambda) = \frac{2}{1+\lambda^2}$.

Example 5. If $f(x) = \frac{1}{1+x^2}$, then $\hat{f}(\lambda) = \pi e^{-|\lambda|}$. [**Hint:** Use contour integration]

Fourier inversion: We now show that the Fourier transform is injective. In other words, if $f,g \in l^1$ and $\hat{f} = \hat{g}$, then $f = g$. We don't simply say this abstractly, but give a recipe to recover f from $\hat{f}$. Fix $f \in L^1$.

Step 1: Let $\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ be the standard Gaussian density. Set $\varphi_\sigma(x) = \frac{1}{\sigma}\varphi(x/\sigma)$ (also a probability density). As $\sigma \to 0$, the probability density concentrates more and more near the origin in the sense that $\int_{|x|>\delta}\varphi_\sigma(x)dx \to 0$ as $\sigma \to 0$, for any fixed $\delta > 0$. We claim that $f * \varphi_\sigma \xrightarrow{L^1} f$, as $\sigma \to 0$ and leave it as a guided exercise.

Exercise 6. Show that $f * \varphi_\sigma \xrightarrow{L^1} f$, as $\sigma \to 0$ by following these steps.

(1) Let $f_\tau(x) = f(x-\tau)$. Show that $f_\tau \xrightarrow{L^1} f$, as $\tau \to 0$.
(2) Write $(f*\varphi_\sigma)(x) - f(x) = \int [f(x-y) - f(y)]\varphi_\sigma(y)dy$. Break the integral into $|y| \le \delta$ and $|y| > \delta$. Use the first part to control the first integral and the concentration of φ_σ around the origin to control the second integral. Deduce the claim.

Step 2: Observe that $\varphi_\sigma = \hat{\psi}_\sigma$ where $\psi_\sigma(x) = \frac{1}{2\pi}e^{-\sigma^2x^2/2}$. Use Exercise 1 to write

$$\begin{aligned}(f*\varphi_\sigma)(x) &= (f*\hat{\psi}_\sigma)(x) \\ &= \int_{\mathbb{R}} \hat{f}(\lambda)\psi_\sigma(-\lambda)e^{i\lambda x}d\lambda \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}(\lambda)e^{i\lambda x}e^{-\sigma^2\lambda^2/2}d\lambda.\end{aligned}$$

Conclusion: If we define $g_\sigma(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}(\lambda)e^{i\lambda x}e^{-\sigma^2\lambda^2/2}d\lambda$, then by the above steps, $g_\sigma \xrightarrow{L^1} f$. On the other hand, g_σ is determined entirely by $\hat{f}$. Therefore, we can recover f from $\hat{f}$. The Fourier transform is injective.

Remark 7. Observe that we could also have used a different probability density φ to start with and define φ_σ by scaling. The first step goes through without change. For the second step to go through, we need φ to be the Fourier transform of some function. There are many choices that do this. For example, if $\varphi(x) = \frac{1}{2}e^{-|x|}$ then $\psi_\sigma(t) = \frac{1}{1+\sigma^2 t^2}$.

Under extra assumption, we can write f in terms of $\hat{f}$ more explicitly.

Corollary 8. *Let $f \in L^1$ and suppose $\hat{f} \in L^1$. Then $f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}(\lambda)e^{i\lambda x}d\lambda$ for a.e.x.*

We wrote for a.e. x, but the point is that the integral $\int_{\mathbb{R}} \hat{f}(\lambda)e^{i\lambda x}d\lambda$ makes sense for every x and is a continuous function of x. In other words, the assumption that $\hat{f}$ is integrable ensures that f is almost every equal to a continuous function which is given by the integral.

Proof. Let $F(x) = \int_{\mathbb{R}} \hat{f}(\lambda)e^{i\lambda x}d\lambda$. As we saw, $g_\sigma \to f$ in L^1 as $\sigma \to 0$. But for any x, we claim that $g_\sigma(x) \to F(x)$. To prove the claim, recall that $g_\sigma(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}(\lambda)e^{i\lambda x}e^{-\sigma^2\lambda^2/2}d\lambda$. As $\sigma \to 0$, the integrand converges to $\hat{f}(\lambda)e^{i\lambda x}$ for each λ. Further, the integrand (as a function of λ) is dominated by $|\hat{f}(\lambda)|$ which is assumed to be integrable. By DCT, it follows that $g_\sigma(x) \to F(x)$ for each x.

Thus g_σ converges in L^1 to f and pointwise to F. Argue that $f = F$, a.e. ■

We give exaples to show how the Fourier transform can be helpful in some computations.

Example 9. Suppose we want to compute $\int_{-\infty}^{\infty} \frac{\sin^2(x)}{x^2}dx$ (the integral clearly exists). We observe that if $f(x) = \frac{1}{2}\mathbf{1}_{[-1,1]}(x)$, then $\hat{f}(\lambda) = \frac{\sin\lambda}{\lambda}$. Hence $\frac{\sin^2\lambda}{\lambda^2}$ is the Fourier transform of $(f * f)(x) = \frac{1}{4}(2-|x|)_+$ (do the computation!). By the Fourier inversion formula,

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\sin^2(\lambda)}{\lambda^2}e^{i\lambda x}d\lambda = \frac{1}{4}(2-|x|)_+.$$

Put $x = 0$ to get $\int_{-\infty}^{\infty} \frac{\sin^2(\lambda)}{\lambda^2}d\lambda = \pi$.

Smoothness and decay:

Proposition 10. *Let $f \in L^1$.*

(1) *Assume that f' exists and is in L^1. Then, $\hat{f'}(\lambda) = i\lambda\hat{f}(\lambda)$. In particular, $\hat{f}(\lambda) = o(1/|\lambda|)$ as $\lambda \to \pm\infty$.*

(2) *Assume that $g(x) := xf(x)$ is in L^1. Then, $\hat{f}$ has a continuous derivative and $\hat{f}'(\lambda) = -i\hat{g}(\lambda)$.*

Repeatedly applying the above, we easily get the following corollary. The point is that smoothness of a function implies fast decay of its Fourier transform. And fast decay of the function implies smoothness of the Fourier transform. By the Fourier inversion, two other statements are also valid - smoothness of the Fourier transform implies decay of the function and decay of the Fourier transform implies smoothness of the function.

Corollary 11. *Let $f \in L^1$ and let $k \geq 1$ be an integer.*

(1) *If f is k times differentiable and $f^{(k)} \in L^1$, then $\widehat{f^{(k)}}(\lambda) = (i\lambda)^k\hat{f}(\lambda)$. In particular, $\hat{f}(\lambda) = o(|\lambda|^{-k})$ as $\lambda \to \pm\infty$.*

(2) *If $g(x) := x^k f(x)$ is integrable, then $\hat{f} \in C^k(\mathbb{R})$ and $\hat{f}^{(k)}(\lambda) = i^k\hat{g}(\lambda)$.*

We end our introduction here. One very nice thing that we omitted for lack of time is the *Poisson summation formula*, you may read about it in Katznelson's book.

7. UNCERTAINTY PRINCIPLES

We have see that if $g_\sigma(x) = \frac{1}{\sigma} g\left(\frac{x}{\sigma}\right)$, then $\hat{g_\sigma}(\lambda) = \hat{g}(\lambda\sigma)$ for any $g \in L^1(\mathbb{R})$ and $\sigma > 0$. In other words, if g is more concentrated (by making σ small), then $\hat{g}$ becomes more spread out.

Uncertainty principles are statements to the effect that a function and its Fourier transform cannot both be concentrated (i.e., put most of its mass in a small interval). These statements are about all functions, instead of staying within a class of dilates as in the previous paragraph. We just emntion two.

Heisenberg's uncertainty principle: Suppose $f \in L^1 \cap L^2$. Then we know that $\hat{f} \in L^2$ too and that $\|\hat{f}\|_{L^2}^2 = 2\pi \|f\|_{L^2}^2$.

If $\|f\|_{L^2} = 1$, then $|f|^2$ is a probability density on the line. Its mean and variance are given by

$$M_f = \int_{\mathbb{R}} x|f(x)|^2 dx,$$

$$V_f = \int_{\mathbb{R}} (x - M_f)^2 |f(x)|^2 dx.$$

By the Plancherel identity, $|\frac{1}{\sqrt{2\pi}}\hat{f}|^2$ is also a probability density. Let $M_{\hat{f}}$ and $V_{\hat{f}}$ denote its mean and variance (slight abuse of notation, we should write $M_{\hat{f}/\sqrt{2\pi}}$ and $V_{\hat{f}/\sqrt{2\pi}}$ but that is painful to write and read).

Theorem 12. *Let $f \in L^1 \cap L^2$. Then, $V_f V_{\hat{f}} \geq 1$.*

We presented the proof in class but I don't want to type it here. See the link `http://www.ias.ac.in/resonance/Volumes/04/02/0020-0023.pdf` for a short exposition of the subject by Alladi Sitaram.

Hardy's theorem: Even before Heisenberg's uncertainty principle, Hardy had proved the following theorem.

Theorem 13. *Let $f : \mathbb{R} \to \mathbb{C}$ be such that $|f(x)| \leq Ce^{-ax^2/2}$ for a.e. x for some constants C, a. Then $\hat{f}$ is well-defined and assume that $|\hat{f}(\lambda)| \leq Ce^{-b\lambda^2/2}$ for $\lambda \in \mathbb{R}$.*

(1) If $ab > 1$, then $f = 0$ a.s.

(2) If $ab = 1$, then $f(x) = ce^{ax^2/2}$ for some constant c (the exponent a is the same as above).

We sketched the proof in class. See Sundaram Thangavelu's book, *An introduction to the uncertainty principle*, page 18, for a full proof of the theorem. It depends on the Phragmen-Lindelöf theorem, a version of the maximum modulus principle for holomorphic functions on some unbounded domains. We shall speak about this class of theorems in the next section. To the best of my knowledge, there is no proof of Hardy's theorem that avoids complex analysis. A purely real analysis proof of a somewhat weaker statement was given by Terence Tao in `http://terrytao.wordpress.com/2009/02/18/hardys-uncertainty-principle/`.