(*November 28, 2016*)

Fourier transforms, I

Paul Garrett garrett@math.umn.edu http://www.math.umn.edu/˜garrett/

[This document is http://www.math.umn.edu/˜garrett/m/fun/notes_2016-17/Fourier_transforms_I.pdf]

The Fourier transform of $f \in L^1(\mathbb{R})$ is [1]

$$\widehat{f}(\xi) \;=\; \int_{\mathbb{R}} e^{-2\pi i \xi x} \cdot f(x)\, dx$$

Since $f \in L^1(\mathbb{R})$, the integral converges absolutely, and uniformly in $\xi \in \mathbb{R}$. Similarly, on $\mathbb{R}^n$, with the usual inner product $\xi \cdot x = \sum_{j=1}^n \xi_j x_j$,

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

An immediately interesting feature of Fourier transform is that *differentiation* is converted to *multiplication*: at first heuristically, but rigorously proven below, imagining that we can integrate by parts,

$$\frac{\partial f}{\partial x_j}\widehat{\ }(\xi) \;=\; \int_{\mathbb{R}^n} e^{-2\pi i \xi\cdot x} \cdot \frac{\partial}{\partial x_j} f(x)\, dx \;=\; \int_{\mathbb{R}^n} \frac{\partial}{\partial x_j} e^{-2\pi i \xi\cdot x} \cdot f(x)\, dx \;=\; \int_{\mathbb{R}^n} (-2\pi i \xi_j) e^{-2\pi i \xi\cdot x} \cdot f(x)\, dx$$

$$=\; (-2\pi i \xi_j) \int_{\mathbb{R}^n} (-2\pi i \xi_j) e^{-2\pi i \xi\cdot x} \cdot f(x)\, dx \;=\; (-2\pi i \xi_j)\widehat{f}(\xi)$$

Thus, the Laplacian $\Delta = \sum_j \frac{\partial^2}{\partial x_j^2}$ is converted to multiplication by $(-2\pi i)^2 \cdot r^2$ where $r^2 = \xi_1^2 + \ldots + \xi_n^2$. Thus, to solve a differential equation such as $(\Delta - \lambda)u = f$, apply Fourier transform to obtain $(-4\pi^2 r^2 - \lambda)\widehat{u} = \widehat{f}$. Divide through by $(-4\pi^2 r^2 - \lambda)$ to obtain

$$\widehat{u} \;=\; \frac{\widehat{f}}{-4\pi^2 r^2 - \lambda}$$

To recover u from $\widehat{u}$, there is *Fourier inversion* (proven below):

$$u(x) \;=\; \int_{\mathbb{R}^n} e^{2\pi i \xi\cdot x}\, \widehat{u}(\xi)\, d\xi$$

There are obvious issues about the integration by parts, the convergence of the relevant integrals, and the inversion formula. In fact, to extend the Fourier transform to $L^2(\mathbb{R}^n)$, the integral definition of the Fourier transform must also be extended to a situation where the literal integral does not converge. Similarly, a bit later, the Fourier transform on the *dual* of the Schwartz space $\mathscr{S}(\mathbb{R}^n)$ (below) is only defined by either an extension by continuity or by a duality relationship.

---

[1] There are other choices of normalizations, that put the 2π in other locations than the exponent, but the differences are inconsequential, so we pick one normalization and use it consistently throughout.

1. *Example computations*

It is useful and necessary to have a stock of explicitly evaluated Fourier transforms, especially on $\mathbb{R}$. In many cases, it is much less obvious how to go in the opposite direction, so *Fourier inversion* (below) has non-trivial content.

[1.1] Characteristic functions of finite intervals It is easy to compute the Fourier transform of the characteristic function $\mathrm{ch}_{[a,b]}$ of an interval $[a,b]$: at least for $\xi \neq 0$, but then extending by continuity (see the *Riemann-Lebesgue Lemma* below),

$$\int_{\mathbb{R}} \mathrm{ch}_{[a,b]}\, e^{-2\pi i\xi x}\, dx \;=\; \int_a^b e^{-2\pi i\xi x}\, dx \;=\; \frac{e^{-2\pi i\xi b} - e^{-2\pi i\xi a}}{-2\pi i\xi}$$

In particular, for a symmetrical interval $[-w,w]$,

$$\int_{\mathbb{R}} \mathrm{ch}_{[-w,w]}\, e^{-2\pi i\xi x}\, dx \;=\; \frac{e^{2\pi i\xi w} - e^{-2\pi i\xi w}}{2\pi i\xi} \;=\; \frac{\sin 2\pi w\xi}{\pi\xi} \;=\; 2w\cdot\frac{\sin 2\pi w\xi}{2\pi w\xi} \;=\; 2w\cdot \mathrm{sinc}(2\pi w\xi)$$

where the (naively-normalized) *sinc* function [2] is $\mathrm{sinc}(x) = \frac{\sin x}{x}$. Anticipating *Fourier inversion* (below), although $\mathrm{sinc}(x)$ is *not* in $L^1(\mathbb{R})$, it is in $L^2(\mathbb{R})$, and its Fourier transform is evidently a characteristic function of an interval. This is not obvious.

[1.2] Tent functions Let $f(x)$ be a piecewise-linear, continuous *tent function* of width $2w$ and height h, symmetrically placed about the origin:

$$f(x) \;=\; \begin{cases} 0 & (\text{for } x \le -w) \\ h - \dfrac{h|x|}{w} & (\text{for } |x| \le w) \\ 0 & (\text{for } x \ge w) \end{cases}$$

Breaking the integral into two pieces and integrating by parts twice, for $\xi \neq 0$ but extending by continuity (see below), we find that

$$\widehat{f}(\xi) \;=\; \frac{h}{\pi^2 w}\Big(\frac{\sin \pi w\xi}{\xi}\Big)^2$$

[1.3] Gaussians With our normalization of the Fourier transform, the best *Gaussian* is $f(x) = e^{-\pi x^2}$, because

$$\int_{\mathbb{R}} e^{-2\pi i\xi x}\, e^{-\pi x^2}\, dx \;=\; e^{-\pi i\xi^2}$$

The sanest proof of this uses *contour shifting* from complex analysis:

$$\int_{\mathbb{R}} e^{-2\pi i\xi x}\, e^{-\pi x^2}\, dx \;=\; \int_{\mathbb{R}} e^{-\pi(x-i\xi)^2 - \pi\xi^2}\, dx \;=\; e^{-\pi\xi^2}\int_{\mathbb{R}} e^{-\pi(x-i\xi)^2}\, dx \;=\; e^{-\pi\xi^2}\int_{-i\xi-\infty}^{-i\xi+\infty} e^{-\pi x}\, dx$$

$$=\; e^{-\pi\xi^2}\int_{-\infty}^{+\infty} e^{-\pi x}\, dx \;=\; e^{-\pi\xi^2}\cdot 1 \;=\; e^{-\pi\xi^2}$$

[2] According to http://en.wikipedia.org/wiki/Sinc_function, the name is a contraction of the Latin name *sinus cardinalis*, bestowed on this function by P. Woodard and I. Davies, *Information theory and inverse probability in telecommunication*, Proc. IEEE-part III: radio and communication engineering **99** (1952), 37-44.

because $\int_{-\infty}^{+\infty} e^{-\pi x}\;dx = 1$. Similarly, in $\mathbb{R}^n$, because the Gaussian and the exponentials both factor over coordinates, the same identity holds:

$$\int_{\mathbb{R}^n} e^{-2\pi i\xi\cdot x}\, e^{-\pi|x|^2}\;dx \;=\; e^{-\pi|\xi|^2}$$

[1.4] Fourier transforms of rational expressions Often, one-dimensional Fourier transforms of relatively elementary expressions can be evaluated *by residues*, meaning via Cauchy's Residue Theorem from complex analysis. Thus, for example,

$$\int_{\mathbb{R}} e^{-2\pi i\xi x}\,\frac{1}{1+x^2}\;dx \;=\; 2\pi i\frac{e^{-2\pi\xi}}{i+i} \;=\; \pi\, e^{-2\pi\xi}$$

by looking at residues in the upper or lower complex half-plane, depending on the sign of ξ. Thinking of Fourier inversion, it is somewhat less obvious how to go in the other direction, to see that the Fourier transform of $e^{-|\xi|}$ is essentially $1/(1+x^2)$. Similarly, for $2 \le k \in \mathbb{Z}$,

$$\int_{\mathbb{R}} e^{-2\pi i\xi x}\,\frac{1}{(x-i)^k}\;dx \;=\; \begin{cases} (2\pi i)(-2\pi i\xi)^{k-1}\,e^{-2\pi|\xi|} & (\text{for } \xi < 0) \\ 0 & (\text{for } \xi > 0) \end{cases}$$

[1.5] Behavior under translations For $f \in L^1(\mathbb{R}^n)$, for $x_o \in \mathbb{R}^n$, certainly $x \to f(x+x_o)$ is still in $L^1(\mathbb{R}^n)$, because Lebesgue measure is translation invariant. Changing variables, replacing x by $x - x_o$,

$$f(*+x_o)\widehat{\;}(\xi) \;=\; \int_{\mathbb{R}^n} e^{-2\pi i\xi\cdot x}\, f(x+x_o)\;dx \;=\; \int_{\mathbb{R}^n} e^{-2\pi i\xi\cdot(x-x_o)}\, f(x)\;dx$$

$$=\; e^{2\pi i\xi\cdot x_o}\int_{\mathbb{R}^n} e^{-2\pi i\xi\cdot x}\, f(x)\;dx \;=\; e^{2\pi i\xi\cdot x_o}\cdot\widehat{f}(\xi)$$

[1.6] Behavior under dilations A similar change of variables applies to dilations $x \to t\cdot x$ with $t > 0$: replacing x by x/t,

$$f(t\cdot *)\widehat{\;}(\xi) \;=\; \int_{\mathbb{R}^n} e^{-2\pi i\xi\cdot x}\, f(t\cdot x)\;dx \;=\; \int_{\mathbb{R}^n} e^{-2\pi i\xi\cdot x/t}\, f(x)\;t^{-n}\,dx$$

$$=\; t^{-n}\int_{\mathbb{R}^n} e^{-2\pi i\frac{\xi}{t}\cdot x}\, f(x)\;dx \;=\; t^{-n}\widehat{f}(t^{-1}\cdot\xi)$$

[1.7] Behavior under linear transformations More generally, with an invertible real matrix A, replacing x by $A^{-1}x$,

$$f(A\cdot *)\widehat{\;}(\xi) \;=\; \int_{\mathbb{R}^n} e^{-2\pi i\xi\cdot x}\, f(Ax)\;dx \;=\; \int_{\mathbb{R}^n} e^{-2\pi i\xi\cdot A^{-1}x}\, f(x)\;(\det A)^{-1}\,dx$$

Since $\xi\cdot A^{-1}x = (A^{-1})^\top\xi\cdot x$, this is

$$(\det A)^{-1}\int_{\mathbb{R}^n} e^{-2\pi i(A^{-1})^\top\xi\cdot x}\, f(x)\;dx \;=\; (\det A)^{-1}\widehat{f}((A^{-1})^\top\xi)$$

2. *Riemann-Lebesgue lemma for* $L^1(\mathbb{R})$

[2.1] Theorem: *(Riemann-Lebesgue)* For $f \in L^1(\mathbb{R})$, the Fourier transform $\widehat{f}$ is in the space $C_o^o(\mathbb{R})$ of continuous functions going to 0 at infinity. In fact, the map $f \to \widehat{f}$ is a continuous linear map from the Banach space $L^1(\mathbb{R})$ to the Banach space $C_o^o(\mathbb{R})$, the latter being the sup-norm completion of $C_c^o(\mathbb{R})$.

Proof: First, for $f \in L^1(\mathbb{R})$,

$$|\widehat{f}(\xi)| \;=\; \Big|\int_{\mathbb{R}} e^{-2\pi i\xi x}\, f(x)\, dx\Big| \;\le\; \int_{\mathbb{R}} |e^{-2\pi i\xi x}|\cdot|f(x)|\, dx \;=\; \int_{\mathbb{R}} |f(x)|\, dx \;=\; |f|_{L^1}$$

Thus, for $|f-g|_{L^1} < \varepsilon$, for all $\xi \in \mathbb{R}$, $|\widehat{f}(\xi) - \widehat{g}(\xi)| < \varepsilon$. Thus, Fourier transform is a continuous map of $L^1(\mathbb{R})$ to the Banach space $C^o_{\text{bdd}}(\mathbb{R})$ of bounded continuous functions with sup norm.

For f the characteristic function of a finite interval, the explicit computation above gives $|\widehat{f}(\xi)| \le 1/\pi|\xi|$ for large $|\xi|$, which certainly goes to 0 at infinity.

The theory of the Riemann integral shows that the space of finite linear combinations of characteristic functions of intervals is L^1-dense in the space $C_c^o(\mathbb{R})$ of compactly-supported continuous functions, which is L^1-dense in $L^1(\mathbb{R})$ itself, by Urysohn's lemma and the definition of integral. That is, every $f \in L^1(\mathbb{R})$ is an L^1-limit of finite linear combinations of characteristic functions of finite intervals. The continuity of the Fourier transform as a map $L^1(\mathbb{R}) \to C^o_{\text{bdd}}(\mathbb{R})$ shows that $\widehat{f}$ is the sup-norm limit of Fourier transforms of finite linear combinations of characteristic functions of finite intervals, which are in $C_c^o(\mathbb{R})$. The sup-norm completion of the latter is $C_o^o(\mathbb{R})$, so $\widehat{f} \in C_c^o(\mathbb{R})$. ///

3. *The Schwartz space* $\mathscr{S} = \mathscr{S}(\mathbb{R}^n)$

The Schwartz space on $\mathbb{R}^n$ consists of all $f \in C^\infty(\mathbb{R}^n)$ such that

$$\sup_{x\in\mathbb{R}^n} (|x|^2)^N \cdot |f^{(\alpha)}(x)| \;<\; \infty \qquad\qquad \text{(for all } N\text{, and for all multi-indices } \alpha)$$

where as usual, for a multi-index $\alpha = (\alpha_1, \ldots, \alpha_n)$ with non-negative integer components,

$$f^{(\alpha)} \;=\; \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}}\frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}}\cdots\frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} f$$

Those supremums

$$\nu_{N,\alpha}(f) \;=\; \sup_{x\in\mathbb{R}^n} (|x|^2)^N \cdot |f^{(\alpha)}(x)|$$

required to be finite for Schwartz functions, are *semi-norms*, in the sense that they are non-negative real-valued functions with properties

$$\begin{cases} \nu_{N,\alpha}(f+g) \;\le\; \nu_{N,\alpha}(f) + \nu_{N,\alpha}(g) & \text{(triangle inequality)} \\ \nu_{N,\alpha}(c\cdot f) \;=\; |c|\cdot\nu_{N,\alpha}(f) & \text{(homogeneity)} \end{cases}$$

In the present context, in fact, these seminorms are genuine *norms*, insofar as no one of them is 0 except for the identically-0 function. That is, this family of seminorms is *separating* in the reasonable sense that, if $\nu_{N,\alpha}(f-g) = 0$ for all N, α, then $f = g$.

The natural *topology* on $\mathscr{S}$ associated to this (separating) family of seminorms can be specified by giving a *sub-basis*[3] *at* $0 \in \mathscr{S}$: in a vector space V, we want a topology to be *translation-invariant* in the sense that vector addition $v \to v + v_o$ is a homeomorphism of V to itself. In particular, for every open neighborhood N of 0, $N + v_o$ is an open neighborhood of v_o, and vice-versa.

Here, take a sub-basis at 0 indexed by N, α, and $\varepsilon > 0$:

$$U_{N,\alpha,\varepsilon} \;=\; \{f \in \mathscr{S} : \nu_{N,\alpha}(f) < \varepsilon\}$$

[3.1] Theorem: With the latter topology, $\mathscr{S}$ is a complete metrizable space. *[... iou ...]*

[3.2] Remark: Since the topology of $\mathscr{S}$ is given by seminorms, the topology is also *locally convex*, meaning that every point has a basis of neighborhoods consisting of *convex* sets. This follows from the convexity of the sub-basis sets, and the fact that an intersection of convex sets is convex. Complete metrizable, locally convex topological vector spaces (with translation-invariant topology, as expected) are *Fréchet spaces*. This is a more general class including Banach spaces. In summary, $\mathscr{S}$ is a Fréchet space.

[3.3] Claim: For $f \in \mathscr{S}$,

$$\Big(\frac{\partial}{\partial x_j} f\Big)\widehat{\;}(\xi) \;=\; (-2\pi i)\cdot \xi_j \cdot \widehat{f}(\xi)$$

Proof: Integration by parts is easily justified for Schwartz functions f, so

$$\Big(\frac{\partial}{\partial x_j} f\Big)\widehat{\;}(\xi) \;=\; \int_{\mathbb{R}^n} e^{-2\pi i\xi\cdot x}\,\frac{\partial f}{\partial x_j}(x)\,dx \;=\; -\int_{\mathbb{R}^n} \frac{\partial f}{\partial x_j} e^{-2\pi i\xi\cdot x}\cdot f(x)\,dx \;=\; -\int_{\mathbb{R}^n} (-2\pi i\xi_j) e^{-2\pi i\xi\cdot x}\cdot f(x)\,dx$$

$$=\; (2\pi i\xi_j)\int_{\mathbb{R}^n} e^{-2\pi i\xi\cdot x}\cdot f(x)\,dx \;=\; (2\pi i)\cdot\xi_j\cdot\widehat{f}(\xi)$$

as claimed. ///

The following claim, essentially the dual or opposite to the previous, has a more difficult proof, a part of which we postpone.

[3.4] Claim: For $f \in \mathscr{S}$,

$$\frac{\partial}{\partial \xi_j}\widehat{f}(\xi) \;=\; (-2\pi i)\cdot (x_j\cdot f)\widehat{\;}(\xi)$$

Proof: The point is that for Schwartz functions, the differentiation in the parameter ξ_j can pass inside the integral:

$$\frac{\partial}{\partial \xi_j}\widehat{f}(\xi) \;=\; \frac{\partial}{\partial \xi_j}\int_{\mathbb{R}^n} e^{-2\pi i\xi\cdot x}\, f(x)\,dx \;=\; \int_{\mathbb{R}^n} \frac{\partial}{\partial \xi_j} e^{-2\pi i\xi\cdot x}\, f(x)\,dx$$

$$=\; \int_{\mathbb{R}^n} (-2\pi i x_j) e^{-2\pi i\xi\cdot x}\, f(x)\,dx \;=\; (-2\pi i)\int_{\mathbb{R}^n} e^{-2\pi i\xi\cdot x}\, x_j f(x)\,dx \;=\; (-2\pi i)\cdot(x_j\cdot f)\widehat{\;}(\xi)$$

Justification for passing the differential operator inside the integral is best given in a slightly more sophisticated context, using Gelfand-Pettis vector-valued integrals, so we will not give any elementary-but-unenlightening argument here. ///

[3] Recall that a set S of sets $U \ni x_o$ is a *sub-basis* at x_o when every neighborhood of x contains a finite intersection of sets from S.

4. *Fourier inversion on $\mathscr{S}$*

In our normalization, the inverse Fourier transform is

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

Of course, this is only slightly different from the *forward* Fourier transform, and sources sometimes do not invent a separate symbol for the inverse transform

[4.1] Theorem: (*Fourier inversion*) $(\widehat{f})^\vee = f$ for $f \in \mathscr{S}$,

Proof: [... iou ...] ///

[4.2] Corollary: Fourier transform is a topological vector space isomorphism $\mathscr{S} \to \mathscr{S}$. [... iou ...]

5. *L^2-isometry of Fourier transform on $\mathscr{S}$*

[5.1] Theorem: (*recast by Schwartz, c. 1950*) For $f, g \in \mathscr{S}$, $\langle f, g\rangle = \langle \widehat{f}, \widehat{g}\rangle$.

Proof: [... iou ...] ///

6. *Isometric extension and Plancherel on $L^2(\mathbb{R}^n)$*

[6.1] Theorem: (*Plancherel, 1910*) There is a unique continuous extension of Fourier transform to an isometry $L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$. In anachronistic terms, the Fourier transform $\mathscr{S} \to \mathscr{S}$ extends by continuity to a map $\mathscr{F} : L^2 \to L^2$, with isometry property

$$\langle \mathscr{F} f, \mathscr{F} g\rangle \;=\; \langle f, g\rangle \qquad\qquad (\text{for all } f, g \in L^2(\mathbb{R}^n))$$

Proof: The L^2 Plancherel theorem on $\mathscr{S}$, and the density of $\mathscr{S}$ in L^2, give the result. ///

7. *Heisenberg uncertainty principle*

This is a theorem about Fourier transforms, once we grant a certain model of quantum mechanics. That is, there is a mathematical mechanism that yields an inequality, which has an interpretation in physics. [4]

For suitable f on $\mathbb{R}$,

$$|f|^2_{L^2} \;=\; \int_{\mathbb{R}} |f|^2 \;=\; -\int_{\mathbb{R}} x(f\cdot \overline{f})' \;=\; -2\,\mathrm{Re} \int_{\mathbb{R}} x f \overline{f}' \qquad\qquad (\text{integrating by parts})$$

That is,

$$|f|^2_{L^2} \;=\; \big||f|^2_{L^2}\big| \;=\; \left|\int_{\mathbb{R}} |f|^2\right| \;=\; \left|-2\,\mathrm{Re}\int_{\mathbb{R}} x f\overline{f}'\right| \;\le\; 2\int_{\mathbb{R}} |x f \overline{f}'|$$

[4] I think I first saw Heisenberg's Uncertainty Principle presented as a theorem about Fourier transforms in Folland's 1983 Tata Lectures on PDE.

Next,

$$2\int_{\mathbb{R}} |xf \cdot \overline{f}'| \;\le\; 2 \cdot |xf|_{L^2} \cdot |f'|_{L^2} \qquad\qquad \text{(Cauchy-Schwarz-Bunyakowsky)}$$

Since Fourier transform is an isometry, and since Fourier transform converts derivatives to multiplications,

$$|f'|_{L^2} \;=\; |\widehat{f'}|_{L^2} \;=\; 2\pi|\xi\widehat{f}|_{L^2}$$

Thus, we obtain the *Heisenberg inequality*

$$|f|^2_{L^2} \;\le\; 4\pi \cdot |xf|_{L^2} \cdot |\xi\widehat{f}|_{L^2}$$

More generally, a similar argument gives, for any $x_o \in \mathbb{R}$ and any $\xi_o \in \mathbb{R}$,

$$|f|^2_{L^2} \;\le\; 4\pi \cdot |(x - x_o)f|_{L^2} \cdot |(\xi - \xi_o)\widehat{f}|_{L^2}$$

Imagining that $f(x)$ is the probability that a particle's *position* is x, and $\widehat{f}(\xi)$ is the probability that its momentum is ξ, Heisenberg's inequality gives a lower bound on how *spread out* these two probability distributions must be. The physical assumption is that position and momentum *are* related by Fourier transform.
