(*February 25, 2020*)

06b. Fourier transforms

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

[This document is
http://www.math.umn.edu/˜garrett/m/real/notes_2019-20/06b_intro_Fourier_transforms.pdf]

1. Basic classes of functions $\mathcal{D}, \mathscr{S}, \mathcal{E}$ and their duals
2. Standard example computations
3. Riemann-Lebesgue lemma for $L^1(\mathbb{R})$
4. The Schwartz space $\mathscr{S} = \mathscr{S}(\mathbb{R}^n)$
5. Fourier inversion on $\mathscr{S}$
6. L^2-isometry of Fourier transform on $\mathscr{S}$
7. Isometric extension to Plancherel for $L^2(\mathbb{R}^n)$
8. Heisenberg uncertainty principle
9. Tempered distributions

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 apparently converted to *multiplication*: at first heuristically, but rigorously proven below, imagining that we can integrate by parts,

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

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

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

[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.

later, the Fourier transform on the *dual* of the Schwartz space $\mathscr{S}(\mathbb{R}^n)$ (below), the *tempered distributions* $\mathscr{S}(\mathbb{R}^n)^*$, is only defined by either an *extension by continuity* or by a *duality*.

---

1. *Basic classes of functions $\mathcal{D}, \mathscr{S}, \mathcal{E}$ and their duals*

Even though our immediate discussion will be incomplete, it is worthwhile to introduce some basic, standard function spaces. Recall that a function F on $\mathbb{R}^n$ is *of rapid decay* when $\sup_{x\in\mathbb{R}^n} |x|^N f(x)| < +\infty$ for all positive integers N.

$$\begin{cases} \mathcal{D} \;=\; \mathcal{D}(\mathbb{R}^n) \;=\; \text{test functions} \;=\; C_c^\infty(\mathbb{R}^n) \\ \\ \mathscr{S} \;=\; \mathscr{S}(\mathbb{R}^n) \;=\; \text{Schwartz functions} \;=\; \{f \in C^\infty(\mathbb{R}^n) : f \text{ and all its derivatives are of rapid decay}\} \\ \\ \mathcal{E} \;=\; \mathcal{E}(\mathbb{R}^n) \;=\; \text{smooth functions} \;=\; C^\infty(\mathbb{R}^n) \end{cases}$$

The spaces $\mathscr{S}$ and $\mathcal{E}$ will turn out to be Fréchet spaces, while the appropriate topology on test function $\mathcal{D}$ is somewhat more complicated. Without elaborating on these topologies, the *dual spaces*, that is, the vector spaces of continuous linear functionals $\mathcal{D} \to \mathbb{C}$, $\mathscr{S} \to \mathbb{C}$, and $\mathcal{E} \to \mathbb{C}$, are

$$\begin{cases} \mathcal{D}^* \;=\; \mathcal{D}' \;=\; \mathcal{D}(\mathbb{R}^n)^* \;=\; \text{distributions} \\ \\ \mathscr{S}^* \;=\; \mathscr{S}' \;=\; \mathscr{S}(\mathbb{R}^n)^* \;=\; \text{tempered distributions} \\ \\ \mathcal{E}^* \;=\; \mathcal{E}' \;=\; \mathcal{E}(\mathbb{R}^n)^* \;=\; \text{compactly-supported distributions} \end{cases}$$

For the latter name to make better sense, we'd need to describe the *support* of a distribution, and also prove that this naming convention is correct.

The obvious inclusions $\mathcal{D} \subset \mathscr{S} \subset \mathcal{E}$ do turn out to be *continuous* in the relevant topologies. Thus, we have inclusion-reversing containments of duals: $\mathcal{E}^* \subset \mathscr{S}^* \subset \mathcal{D}^*$.

Thus, *tempered* distributions really are a kind of distribution, and *compactly-supported* distributions are a kind of tempered distribution.

Eventually (below), we refine the chain of containments

$$\mathcal{D} \;\subset\; \mathscr{S} \;\subset\; L^2(\mathbb{R}^n) \;\subset\; \mathscr{S}^* \;\subset\; \mathcal{D}^*$$

in various ways. One such refinement is in terms of *Sobolev spaces*.

---

2. *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.

[2.1] 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.

[2.2] 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

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

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

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

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

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

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

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:

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

[2.4] 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,

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

---

[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.

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 $2 \le k \in \mathbb{Z}$,

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

[2.5] Translations are converted to multiplications 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 $x - x_o$,

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

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

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

[2.7] 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

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

[2.8] Multiplications are converted to differentiation, and vice-versa For *suitable* f, so that integration by parts succeeds,

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

Anticipating Fourier inversion, we would know that, symmetrically, multiplication by x is essentially converted to differentiation. We can also compute this directly, but with a non-trivial issue about moving the differentiation through the integral: [3]

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

[3] For f a *Schwartz function*, that is, smooth and it and all derivatives are of rapid decay (see below), moving the differentiation through the integral is demonstrably legitimate. However, the best proof, which shows that this is a special case of a very general pattern of operators commuting with integrals, is not elementary. It uses Gelfand-Pettis (also called *weak*) vector-valued integrals, which will be discussed later.

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

The issue of moving the differential operator through the integral also arises below in proving that Fourier transform maps the space $\mathscr{S}$ of Schwartz functions to itself.

---

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

Just to be sure that this result is not overlooked, we recall it:

[3.1] 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 $C^o_c(\mathbb{R})$.

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 $C^o_c(\mathbb{R})$ 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 $C^o_c(\mathbb{R})$. The sup-norm completion of the latter is $C^o_o(\mathbb{R})$, so $\widehat{f} \in C^o_c(\mathbb{R})$. ///

---

4. *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. 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*[4] *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 $\varepsilon > 0$:

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

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

[4.2] 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.

[4.3] 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: We've already sketched the integration by parts argument for this, so now we should check in detail that $f \in \mathscr{S}$ is sufficient for that argument to succeed. For notational simplicity, we treat just the one-dimensional case:

$$\widehat{(f')}(\xi) \;=\; \int_{\mathbb{R}} e^{-2\pi i \xi x}\,\frac{\partial}{\partial x} f(x)\,dx \;=\; \lim_{N\to+\infty}\int_{|x|\le N} e^{-2\pi i\xi x}\,\frac{\partial}{\partial x} f(x)\,dx$$

Integrating by parts, the integral is

$$\Big[e^{-2\pi i\xi x}\,f(x)\Big]_{-N}^{N} \;-\; \int_{|x|\le N} \frac{\partial}{\partial x} e^{-2\pi i \xi x}\cdot f(x)\,dx$$

$$=\; e^{-2\pi i\xi N}\,f(N) - e^{2\pi i \xi N}\,f(-N) \;-\; \int_{|x|\le N} (-2\pi i\xi)\,e^{-2\pi i\xi x}\cdot f(x)\,dx$$

The boundary terms go to 0 as $N \to +\infty$, the factor of $-2\pi i\xi$ comes out of the integral, and the limit as $N \to +\infty$ of the integral over $|x| \le N$ becomes the integral over $\mathbb{R}$, as claimed. ///

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

[4.4] Claim: For $f \in \mathscr{S}(\mathbb{R}^n)$,

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

[4] 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.

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

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

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

As remarked earlier, passing the differential operator inside the integral is best justified in a more sophisticated context, so we will not give any elementary-but-unenlightening argument here. ///

Let *translation* by x on $\mathscr{S}$ be written $T_x f$, where

$$(T_x f)(y) \;=\; f(y+x)$$

[4.5] Claim: For each $x\in\mathbb{R}^n$, translation by x is a continuous map $\mathscr{S}\to\mathscr{S}$.

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

---

5. *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

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

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

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

---

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

[6.1] Theorem: (*recast by Schwartz, c. 1950*) For $f,g\in\mathscr{S}$, $\langle f,g\rangle_{L^2} = \langle \widehat{f},\widehat{g}\,\rangle_{L^2}$. In particular, $|\widehat{f}|_{L^2} = |f|_{L^2}$.

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

---

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

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

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

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

[7.2] Remark: Even though the literal integral for the Fourier transform of typical $f \in L^2$ (but not in L^1) need not converge, it is standard to write the Fourier transform as an integral, with the understanding that it is not the *literal* integral, but is an extension-by-continuity of the literal integral, via Plancherel.

8. *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. [5]

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

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

That is,

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

Next,

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

Since Fourier transform is an L^2-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.

9. *Tempered distributions*

Tempered distributions can be first described as the space $\mathscr{S}^*$ of continuous linear function(al)s $\lambda : \mathscr{S} \to \mathbb{C}$.

[9.1] Claim: The Dirac δ, given by $\delta(\varphi) = \varphi(0)$ for $\varphi \in \mathscr{S}$, is a tempered distribution.

Proof: To prove continuity of $\varphi \to \varphi(0)$, it suffices to prove continuity at 0. The easy inequality

$$|f(0)| \;\le\; \sup_{x\in\mathbb{R}^n} |x|^0\cdot|f^{(0)}(x)| \;=\; \nu_{0,0}(f)$$

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

shows that $|f(0)|$ can be made as small as desired by making $\nu_{0,0}(f)$ sufficiently small, proving continuity. ///

The duality approach *does* allow an easy definition of Fourier transform $\widehat{u}$ of $u \in \mathscr{S}^*$, not by an integral, but by

$$\widehat{u}(\varphi) \;=\; u(\widehat{\varphi}) \qquad\qquad (\text{for } \varphi \in \mathscr{S})$$

Similarly for inverse Fourier transform, which we've shown truly is an inverse to the Fourier transform on $\mathscr{S}$. It remains to be shown that it is truly an inverse on $\mathscr{S}^*$. Prior to that, we have a basic example:

[9.2] Claim: $\widehat{\delta} = 1$. That is, the Fourier transform of the Dirac δ is integrate-against 1.

Proof: From the definition of Fourier transform on $\mathscr{S}^*$ via duality, for $\varphi \in \mathscr{S}$,

$$\widehat{\delta}(\varphi) \;=\; \delta(\widehat{\varphi}) \;=\; \widehat{\varphi}(0) \;=\; \int_{\mathbb{R}^n} e^{-2\pi i\, 0\cdot x}\, \varphi(x)\, dx \;=\; \int_{\mathbb{R}^n} 1 \cdot \varphi(x)\, dx \;=\; 1(\varphi)$$

by the literal integral definition of Fourier transform on $\mathscr{S}$. ///

We can give $\mathscr{S}^*$ the *weak dual topology*, also called the *weak *-topology*, by seminorms ν_φ attached to $\varphi \in \mathscr{S}$:

$$\nu_\varphi(u) \;=\; |u(\varphi)| \qquad\qquad (\text{for } u \in \mathscr{S}^* \text{ and } \varphi \in \mathscr{S})$$

This is *not* a topology given by a metric, but is obviously a type of topology that can be given to any dual space. This characterization of tempered distributions by *duality* does not explain their usefulness.

[9.3] Theorem: The definition of Fourier transform on $\mathscr{S}^*$ by duality does map $\mathscr{S}^*$ to itself, and is an isomorphism. Fourier inversion for the extended Fourier transform holds on $\mathscr{S}^*$.

Proof: From above, Fourier transform is a continuous linear map of $\mathscr{S}$ to itself. Thus, $\varphi \to \widehat{\varphi} \to u(\widehat{\varphi})$ is a continuous linear functional on $\mathscr{S}$, for any $u \in \mathscr{S}^*$. To prove continuity of $u \to \widehat{u}$ in the weak dual topology, take $\varphi \in \mathscr{S}$, with associated semi-norm ν_φ as above, and compute

$$\nu_\varphi(\widehat{u}) \;=\; |\widehat{u}(\varphi)| \;=\; |u(\widehat{\varphi})| \;=\; \nu_{\widehat{\varphi}}(u)$$

Thus, making $\nu_{\widehat{\varphi}}(u)$ small makes $\nu_\varphi(\widehat{u})$ small, proving continuity of $u \to \widehat{u}$ in the weak dual topology.

To prove Fourier inversion on $\mathscr{S}^*$, let $\mathscr{F}$ be the extended Fourier transform, and $\mathscr{F}'$ the extension of the inverse transform, *not* denoted $\mathscr{F}^{-1}$, to avoid inadvertently begging the question. Then for $\varphi \in \mathscr{S}$,

$$(\mathscr{F}'(\mathscr{F}u))(\varphi) \;=\; (\mathscr{F}u)(\mathscr{F}'\varphi) \;=\; u(\mathscr{F}(\mathscr{F}'\varphi)) \;=\; u(\varphi)$$

by Fourier inversion on $\mathscr{S}$. Since both the transform and its inverse are continuous, both are isomorphisms. ///

There is also a characterization of $\mathscr{S}^*$ as an *extension* of $\mathscr{S}$. First, there is a inclusion $\mathscr{S} \to \mathscr{S}^*$ by taking $\varphi \in \mathscr{S}$ to the *integrate-against* functional u_φ:

$$u_\varphi(f) \;=\; \int_{\mathbb{R}^n} \varphi \cdot f \;=\; \int_{\mathbb{R}^n} \varphi(x) \cdot f(x)\, dx \qquad\qquad (\text{for } f \in \mathscr{S})$$

For most topological vector spaces V, there is no natural inclusion $V \to V^*$, so such inclusions for *spaces of functions* V distinguishes them from the general abstract scenario.

[9.4] Claim: The inclusion $\mathscr{S} \to \mathscr{S}^*$ is continuous, and has dense image.

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

That is, we can think of $\mathscr{S}^*$ as a sort of *completion* or *extension* of $\mathscr{S}$, in the weak dual topology on $\mathscr{S}$ itself. From this viewpoint, $\mathscr{S}^*$ consists of *generalized functions*. Thus, the definition of Fourier transform on $\mathscr{S}^*$ should be compatible with that defined by the literal integral on $\mathscr{S}$:

[9.5] Claim: The Fourier transform on $\mathscr{S}^*$ defined via duality agrees with the integral definition on $\mathscr{S} \subset \mathscr{S}^*$. That is, with u_φ the integrate-against functional attached to $\varphi \in \mathscr{S}$,

$$\widehat{u_\varphi} \;=\; u_{\widehat{\varphi}}$$

Proof: This compatibility is an easy preliminary form of Plancherel: for $f \in \mathscr{S}$,

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

by Fubini-Tonelli. ///

We can compute Fourier transforms of tempered distributions by taking weak dual topology limits of Schwartz functions and the literal integral form of the Fourier transform:

[9.6] Claim: For a sequence of Schwartz functions φ_i approaching a tempered distribution u in the weak dual topology,

$$(\mathscr{S}^*\text{-}) \lim_i \widehat{\varphi}_i \;=\; \widehat{u}$$

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

We define *derivatives* of tempered distributions in a fashion compatible with the integrate-against inclusion $\mathscr{S} \to \mathscr{S}^*$, specifically, to be compatible with *integration by parts*. That is, for $\varphi, f \in \mathscr{S}$, and integration-by-parts distribution u_φ, in one variable,

$$u_{\varphi'}(f) \;=\; \int_{\mathbb{R}^n} \varphi' \cdot f \;=\; -\int_{\mathbb{R}^n} \varphi \cdot f' \;=\; -u_\varphi(f')$$

Note the sign. Thus, on $\mathbb{R}^n$, for $u \in \mathscr{S}^*$, define u' by

$$\frac{\partial}{\partial x_i} u(f) \;=\; -u(\frac{\partial}{\partial x_i} f) \qquad\qquad (\text{for } f \in \mathscr{S})$$

Similarly, *multiplication* by polynomials can be defined by duality, also in a fashion compatible with $\mathscr{S} \subset \mathscr{S}^*$:

$$(x_i \cdot u)(f) \;=\; u(x_i f) \qquad\qquad (\text{for } f \in \mathscr{S})$$

[9.7] Corollary: Differentiation and multiplication by polynomials are continuous maps $\mathscr{S}^* \to \mathscr{S}^*$, with the weak dual topology.

Proof: Again, continuity of a linear map is equivalent to continuity at 0. Given $\varphi \in \mathscr{S}$ and $u \in \mathscr{S}^*$,

$$\nu_\varphi\Big(\frac{\partial}{\partial x_i} u\Big) \;=\; \Big|\frac{\partial}{\partial x_i} u(\varphi)\Big| \;=\; \Big| - u(\frac{\partial}{\partial x_i}\varphi)\Big| \;=\; \nu_{\frac{\partial}{\partial x_i}\varphi}(u)$$

as desired. Similarly,

$$\nu_\varphi(x_i u) \;=\; |(x_i u)(\varphi)| \;=\; |u(x_i \varphi)| \;=\; \nu_{x_i\varphi}(u)$$

///

[9.8] Corollary: As for Schwartz functions, Fourier transform *intertwines* differentiation and multiplication on $\mathscr{S}^*$.

Proof: For notational simplicity, let's do this just on $\mathbb{R}$. For $u \in \mathscr{S}^*$, the Fourier transform of the derivative is described, for $\varphi \in \mathscr{S}$, as

$$\widehat{u'}(\varphi) \;=\; u'(\widehat{\varphi}) \;=\; -u(\frac{d}{dx}\widehat{\varphi}) \;=\; -u\Big(-2\pi i\widehat{\xi\varphi}\Big) \;=\; 2\pi i\cdot\widehat{u}(\xi\varphi) \;=\; 2\pi i\,\xi\cdot\widehat{u}(\varphi)$$

That is, $\widehat{u'} = 2\pi i\xi\cdot\widehat{u}$. The same sort of computation proves the reverse. ///

[9.9] Remark: Also, this intertwining property can be proven by extending by continuity from $\mathscr{S}\subset\mathscr{S}^*$.

[9.10] Polynomials and derivatives of δ From $\widehat{\delta} = 1$ and the intertwining of differentiation and multiplication by powers of $x_1,\ldots,x_n$, for a multi-index α,

$$\widehat{\delta^{(\alpha)}} \;=\; (2\pi i)^{|\alpha|}\,x^\alpha\cdot\widehat{\delta}(x) \;=\; (2\pi i)^{|\alpha|}\,x^\alpha\cdot 1 \;=\; (2\pi i)^{|\alpha|}\,x^\alpha$$

where $x^\alpha = x_1^{\alpha_1}\ldots x_n^{\alpha_n}$, and $|\alpha| = \alpha_1+\ldots+\alpha_n$. By Fourier inversion,

$$\widehat{x^\alpha} \;=\; \frac{1}{(2\pi i)^{|a|}}\cdot\delta^{(\alpha)}$$

As with differentiation, multiplication by polynomials, and Fourier transform, *translation* of $u\in\mathscr{S}^*$ is defined either by duality or by extension-by-continuity from $\mathscr{S}\subset\mathscr{S}^*$. Just as the possibly unexpected -1 in the derivative, to be compatible with integration by parts, we should see how translation behaves for integrate-against distributions. Let the translate T_xf of $f\in\mathscr{S}$ be defined by $T_xf(y) = f(y+x)$. For $\varphi, f\in\mathscr{S}$,

$$u_{T_x\varphi}(f) \;=\; \int_{\mathbb{R}_n} T_x\varphi\cdot f \;=\; \int_{\mathbb{R}_n} T_x\varphi(y)\cdot f(y)\,dy \;=\; \int_{\mathbb{R}_n}\varphi(y+x)\cdot f(y)\,dy$$
$$=\; \int_{\mathbb{R}_n}\varphi(y)\cdot f(y-x)\,dy \;=\; u_\varphi(T_{-x}f)$$

by replacing y by $y-x$. Thus, again, a sign should enter in the definition of translation of a tempered distribution u:

$$(T_xu)(f) \;=\; u(T_{-x}f) \qquad\qquad (\text{for } f\in\mathscr{S})$$
