WEEK 6: THE FOURIER TRANSFORM

The references for this section is [2] and [1].

The theoretical tool used most often in medical imaging—also important in many other fields, including probability and statistics, signal processing, partial differential equations, certain areas in number theory, and numerical analysis, to name a few—is the Fourier transform. In the medical imaging modalities we discuss in this class, the FFT is important as a practical algorithm, and Fourier series serve as a bridge between the Fourier transform on $\mathbb{R}^n$ and the DFT, but the Fourier transform on $\mathbb{R}^n$ is what motivates the reconstruction methods.

Like the DFT, the Fourier transform takes a function and breaks it into harmonic components, in this case via an integral. Often we refer to the argument of the Fourier transform as a location in 'frequency space', and in many physical applications, the Fourier transform of a function is called its spectrum.

Although the most straightforward definition of the Fourier transform is on L^1 (we only need to have an absolute convergence of the integral), there are three natural domains for the Fourier transform, three spaces that the Fourier transform puts into one-to-one correspondence with themselves. The most restrictive is the Schwarz class, $\mathcal{S}(\mathbb{R}^n)$, which we introduced earlier, the space of rapidly decreasing functions with rapidly decreasing derivatives of all orders. Probably the most often studied is $L^2(\mathbb{R}^n)$, the space of square-integrable functions. The third and most general space on which the Fourier transform is a bijection is the space of tempered distributions, bounded linear functionals on $\mathcal{S}(\mathbb{R}^n)$, which is denoted $\mathcal{S}'(\mathbb{R}^n)$. The elements in this class are not all ordinary functions that can be defined by their coordinates; some are generalize functions which can only be defined by their action (for example, by their action in integrals). Defining the transform on tempered distributions must will be the last step when we have the basic properties of the transform on $\mathcal{S}(\mathbb{R}^n)$.

1. The Fourier transform on functions

Definition 1.1. The *Fourier transform* of $f \in L^1(\mathbb{R}^n)$ is defined to be

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

and is often denoted $\hat{f}(\xi)$. The vector ξ (k and s are also commonly used, depending on the author) is often referred to as a location in 'frequency space', the 'frequency domain', or 'k-space' (the last term is specific to MRI).

It is clear from linearity of the integral that the Fourier transform is linear.

We summarize the basic properties of the Fourier transform $\mathcal{F}$ in the following proposition:

Proposition 1.2. *Let f, $g \in L^1(\mathbb{R}^n)$. Then*

(i) $\widehat{R_k f}(\xi) = e^{-2\pi i k\cdot\xi}\hat{f}(\xi)$, *and* $(e^{2\pi i k\cdot\xi} f)\hat{\ }(\xi) = R_k\hat{f}(\xi)$.

(ii) *Suppose T is an invertible linear transformation on $\mathbb{R}^n$. Let $S = (T^*)^{-1}$ be its transpose inverse. Then* $\widehat{f\circ T} = |\det T|^{-1}(\hat{f})\circ S$.

(iii) $\widehat{f * g} = \hat{f}\hat{g}$, *that is, the Fourier transform converts convolution into multiplication.*

(iv) *If $x^\alpha f \in L^1(\mathbb{R}^n)$ for $|\alpha| \le k$, then $\hat{f}$ is C^k with* $\partial^\alpha \hat{f} = [(-2\pi i x)^\alpha f]\hat{\ }$.

(v) *If $f \in C^k$ with $D^\alpha f \in L^1$ for each $|\alpha| \le k$, then* $\widehat{D^\alpha f}(\xi) = (2\pi i\xi)^\alpha \hat{f}(\xi)$.

(vi) (Riemann-Lebesgue Lemma) $\hat{f} \in C_0(\mathbb{R}^n)$, *i.e., the Fourier transform vanishes at infinity.*

Proof. (i) We prove the first equality,

$$\begin{aligned}\widehat{R_k f}(\xi) &= \int_{-\infty}^{\infty} f(x-k)e^{-2\pi i x\cdot\xi} \\ &= e^{-2\pi i k\cdot\xi}\int_{-\infty}^{\infty} f(x-k)e^{-2\pi i(x-k)\cdot\xi}dx = e^{-2\pi i k\cdot\xi}\hat{f}(\xi),\end{aligned}$$

where we use basic algebraic properties of the exponential and linearity of the integral to get the second equality.

(ii) Using the linear change of variables formula with $y = Tx$,

$$\begin{aligned}\widehat{f\circ T}(\xi) &= \int_{-\infty}^{\infty} f(Tx)e^{-2\pi i x\cdot\xi}dx = |\det T|^{-1}\int_{-\infty}^{\infty} f(y)e^{-2\pi i\xi\cdot(T^{-1}y)}dx \\ &= |\det T|^{-1}\int_{-\infty}^{\infty} f(y)e^{-2\pi i(S\xi)\cdot y}dx = |\det T|^{-1}\hat{f}(S\xi),\end{aligned}$$

(iii) Using Fubini's Theorem, we obtain

$$\begin{aligned}\widehat{f*g}(\xi) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x-y)g(y)dye^{-2\pi i x\cdot\xi}dx \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x-y)g(y)e^{-2\pi i(x-y)\cdot\xi}dxe^{-2\pi i y\cdot\xi}dy \\ &= \hat{f}(\xi)\int_{-\infty}^{\infty} g(y)dye^{-2\pi i y\cdot\xi}dy \\ &= \hat{f}(\xi)\hat{g}(\xi).\end{aligned}$$

(iv) Using Liebniz's rule inductively with partial derivatives,

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

(v) By induction on $|\alpha|$, it suffices to consider $|\alpha| = k = 1$. By Fubini's theorem, we can then reduce to $n = 1$, we can evaluate integrals iteratively. In this case, we integrate by parts and use the fact that for a differentiable L^1 function the boundary terms must vanish to compute

$$\int f'(x)e^{-2\pi i x\xi}dx = -\int f(x)(-2\pi i\xi)e^{2\pi i x\xi}dx = (2\pi i\xi)\hat{f}(\xi).$$

(vi) By (iv), we know that if f is continuously differentiable with compact support, then $\hat{f}$ is infinitely differentiable and $|\xi|\,\hat{f}(\xi)$ is bounded, forcing $\hat{f}$ to vanish at ∞, which shows that $\hat{f}$ is C_0. We now use two facts: that $C_0(\mathbb{R}^n)$ is closed under the uniform norm, and that continuously differentiable functions with compact support are dense in $L^1(\mathbb{R})$. Using the second, for a general $f \in L^1$ we can choose a sequence of continuous functions with compact support f_n converging in the L^1 norm to f, and thus,

$$\left|\hat{f}_n(\xi) - \hat{f}(\xi)\right| = \left|\widehat{f - f_n}(\xi)\right| \le \|f - f_n\|_1 \to 0.$$

Hence, using the fact that C_0 is closed, we get $\hat{f} \in C_0(\mathbb{R}^n)$.

□

Note that if T is a rotation, then (ii) reduces to $\widehat{f\circ T} = (\hat{f})\circ T$, and if T is a dilation, i.e., $T(x) = \lambda x$, then $\widehat{f\circ T}(\xi) = \lambda^{-n}\hat{f}(\lambda^{-1}\xi)$. Also, since radial functions are precisely those functions that are invariant under rotations, a function on $\mathbb{R}^n$ is radial if and only if its Fourier transform is.

Parts (iv) and (v) of Proposition 1.2 indicate that the smoothness properties of a function are reflected in the decay properties of its Fourier transform and vice-versa. Since the Schwarz functions display both smoothness and decay properties, we get the following corollary, which proves critical in that it allows us to extend the Fourier transform to distributions.

Corollary 1.3. *The Fourier transform maps the Schwarz class continuously into itself.*

In order to prove more, we need to have a way to invert the Fourier transform. A fundamental tool in doing so is the Gaussian, $e^{-\pi|x|^2}$. The following technical lemma, whose proof we leave to the reader largely explains why:

Proposition 1.4. *In $\mathbb{R}^n$ we have $(e^{-\pi a|x|^2})^\wedge = a^{-n/2}e^{-\pi|\xi|^2/a}$.*

We now define the inverse Fourier transform:

Definition 1.5. For $f \in L^1(\mathbb{R}^n)$, define the *inverse Fourier transform*

$$\mathcal{F}^{-1}(f)(x) = \int_{\mathbb{R}^n} f(\xi)e^{2\pi i x\cdot\xi}d\xi.$$

The inverse Fourier transform is often denoted $\check{f}$.

It is worth noting that $\check{f}(x) = \hat{f}(-x)$, so that the inverse Fourier transform satisfies most of the same properties as the Fourier transform, suitably adjusted for the opposite symmetry condition.

We would like to be able to say that $(\hat{f})\check{} = f$ whenever $\hat{f} \in L^1$. At first glance, it looks as though we might be able to use Fubini's theorem in the double integral defining

$$(\hat{f})\check{}(x) = \int\int f(y)e^{-2\pi iy\cdot\xi}e^{-2\pi ix\cdot\xi}dyd\xi,$$

but this does not work because the integrand is not in $L^1(\mathbb{R}^n \times \mathbb{R}^n)$. The proof involves limiting using the Gaussian and the following lemma:

Lemma 1.6. *If f, $g \in L^1(\mathbb{R}^n)$, then $\int \hat{f}g = \int f\hat{g}$, or equivalently, $\langle f, g\rangle = \left\langle \hat{f}, \hat{g}\right\rangle$.*

Proof. Note that both integrals are equal to $\int\int f(x)g(\xi)e^{-2\pi ix\cdot\xi}\,dx\,d\xi$. ☐

Proposition 1.7. (Fourier Inversion Formula) *Suppose f, $\hat{f} \in L^1(\mathbb{R}^n)$. Then f agrees almost everywhere with the continuous function f_0 defined by $f_0 = (\hat{f})\check{} = (\check{f})\hat{}$.*

Proof. Let $x \in \mathbb{R}^n$, $t > 0$, and for $\xi \in \mathbb{R}^n$ set $\phi_{x,t}(\xi) = e^{2\pi ix\cdot\xi}e^{-\pi|t\xi|^2}$. It follows from Propositions 1.2(i) and 1.4 that

$$\widehat{\phi_{x,t}}(y) = \frac{1}{t^n}\exp\left(-\pi\left|\frac{x-y}{t}\right|^2\right) = g_t(x-y),$$

where $g(x) = e^{-\pi|x|^2}$ and g_t has the same meaning as in the approximation of the convolution identity. Then from Lemma 1.6 and Proposition 1.2 (iii) it follows that

$$\int e^{-\pi|t(x-y)|^2}e^{2\pi ix\cdot\xi}\hat{f}(\xi)d\xi = \int \hat{f}\phi_{x,t} = \int f\hat{\phi}_{x,t} = f * g_t(x).$$

The fact that the $\{g_t\}_{t>0}$ approximate the convolution identity (as $t \to 0$) yields $\|f * g_t - f\|_1 \to 0$ as $t \to 0$. On the other hand, since $\hat{f} \in L^1$ and the pointwise limit of $e^{-\pi|t(x-y)|^2}$ as $t \to 0$ is 1, it follows from the Lebesgue dominated convergence theorem that

$$\lim_{t\to 0}\int e^{-\pi|t(x-y)|^2}e^{2\pi ix\cdot\xi}\hat{f}(\xi)d\xi = \int e^{2\pi ix\cdot\xi}\hat{f}(\xi)d\xi = (\hat{f})\check{}(x).$$

Thus, $f = (\hat{f})\check{}$ almost everywhere [1]. A similar argument shows that $f = (\check{f})\hat{}$ almost everywhere. Since $f = (\hat{f})\check{}$ and $f = (\check{f})\hat{}$ are continuous by the Riemann-Lebesgue lemma, being Fourier transforms of L^1 functions, the proof is complete. ☐

Corollary 1.8. *If $f \in L^1(\mathbb{R}^n)$ and $\hat{f} = 0$, then $f = 0$ almost everywhere (on $\mathbb{R}^n$) .*

Since the Fourier transform of a Schwarz function is Schwarz and therefore L^1, the inversion formula applies to all Schwarz functions, which shows that the Fourier transform is injective, since a continuous function is determined by its behavior almost everywhere. In addition, since $\check{f}(x) = \hat{f}(-x)$, the inverse Fourier transform also maps the Schwarz space continuously into itself, which shows that the Fourier transform and its inverse are continuous bijections on the Schwarz class. We state this as a corollary:

Corollary 1.9. *The Fourier transform is an isomorphism on the Schwarz class.*

At this point, we could define Fourier transforms of tempered distributions, although we will first list a few results related to the Fourier transform on its intermediate natural domain, $L^2(\mathbb{R}^n)$.

Proposition 1.10. *Let $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. Then $\hat{f} \in L^2(\mathbb{R}^n)$, and for any $g \in L^1$ such that $\hat{g} \in L^1$, we have*

$$\langle f, g\rangle_{L^2} = \left\langle \hat{f}, \hat{g}\right\rangle_{L^2}.$$

In other words, the Fourier transform is unitary on $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$.

[1] We haven't actually proved it, but although L^1 convergence does not imply pointwise convergence anywhere, it is true—and would not be difficult to prove if we had measure theory on hand—that if two functions are both L^1 limits of the same sequence then they must be equal almost everywhere.

Proof. First note that the inner product is well defined: if g, $\hat{g} \in L^1$, then g, $\hat{g}$ are also bounded. Also note that such functions g are dense in both L^1 and L^2, since the Schwarz class is dense in both spaces.

Now, for such a g, set $h = \bar{\hat{g}}$, so that

$$\hat{h}(\xi) = \int_{-\infty}^{\infty} \bar{\hat{g}}(x) e^{-2\pi i x\cdot\xi} dx = \int_{-\infty}^{\infty} \overline{\hat{g}(x) e^{2\pi i x\cdot\xi}} dx = \bar{g}(\xi).$$

Then by Lemma 1.6

$$\int f\bar{g} = \int f\hat{h} = \int \hat{f} h = \int \hat{f}\bar{\hat{g}}.$$

□

Proposition 1.11. *The Fourier transform extends uniquely to a unitary isomorphism on $L^2(\mathbb{R}^n)$.*

Proof. The proof uses the fact that the Schwarz class is dense (twice continuously differentiable functions decaying faster than $\frac{1}{x}$ suffice) in both L^1 and L^2, and from completeness of L^2. The details are left to the reader. □

From the above propositions flow the two familiar results from the discrete cases, where we proved them using linear algebra: for any f, $g \in L^2(\mathbb{R}^n)$.

- (Parseval's relation) $\langle f, g\rangle_{L^2} = \left\langle \hat{f}, \hat{g} \right\rangle_{L^2}$.
- (Plancherel's formula) $\|f\|_2^2 = \left\|\hat{f}\right\|_2^2$.

2. The Fourier Transform on Tempered Distributions

Recall that for f, $g \in L^1(\mathbb{R}^n)$, and hence, for f, $g \in \mathcal{S}(\mathbb{R}^n)$, we have

$$\int \hat{f} g = \int \hat{f}.$$

Also recall that the Fourier transform is a continuous bijection on $\mathcal{S}(\mathbb{R}^n)$. Using these two facts, we can extend the Fourier transform from $\mathcal{S}(\mathbb{R}^n)$ to $\mathcal{S}'(\mathbb{R}^n)$ by defining, for an arbitrary tempered distribution f,

$$\left\langle \hat{f}, \phi \right\rangle = \left\langle f, \hat{\phi} \right\rangle.$$

Clearly, if $f \in L^1$, by Fubini's theorem this new definition is consistent with the old.

It is easy to verify that the basic properties of the Fourier transform continue to hold for tempered distributions. We leave it to the reader to verify that for $f \in \mathcal{S}'(\mathbb{R}^n)$ and $\psi \in \mathcal{S}(\mathbb{R}^n)$,

- $\widehat{R_k f}(\xi) = e^{-2\pi i k\cdot\xi} \hat{f}$,
- $\widehat{e^{2\pi i k\cdot\xi} f} = R_k \hat{f}$,
- $\widehat{f * g} = \hat{f}\hat{g}$,
- $\widehat{f \circ T} = |\det T|^{-1} \hat{f} \circ (T^*)^{-1} \quad (T \in GL(n, \mathbb{R}))$, [2]
- $\widehat{(2\pi i x)^\alpha f} = \partial^\alpha \hat{f}$,
- $\widehat{\partial^\alpha f} = (-2\pi i x)^\alpha \hat{f}$.

The inverse Fourier transform on $\mathcal{S}'(\mathbb{R}^n)$ is defined similarly,

$$\langle f\check{}, \phi\rangle = \langle f, \phi\check{}\rangle,$$

and it follows immediately from the inversion theorem on L^1 that $\mathcal{F}$ is bijective on $\mathcal{S}'(\mathbb{R}^n)$ with inverse $\mathcal{F}^{-1}$.

Perhaps the most useful Fourier transform of a distribution that is not also a function is that of the Dirac delta function. It follows from the formula

$$\left\langle \hat{\delta}, \phi \right\rangle = \left\langle \delta, \hat{\phi} \right\rangle = \hat{\phi}(0) = \int \phi$$

that the Fourier transform of δ is 1, and similarly that the Fourier transform of 1 is δ. This offers another way to look at why δ is the convolution identity: its Fourier transform is the multiplicative identity in the Fourier domain, so it must be the convolution identity in the spatial domain, and vice versa. This also gives us another way to motivate (although not replace, since defining the Fourier transform of distributions requires

[2] $GL(n, \mathbb{R})$ is the set of invertible linear transformations on $\mathbb{R}^n$.

already having the Fourier inversion formula in $\mathcal{S}(\mathbb{R}^n)$) the method used to prove Fourier inversion formula: by convolving the Fourier transform with better and better approximations to the identity, but approximations that were in the Schwarz class and therefore analytically 'nice', we were able to approximate the original function in the spatial domain.

By considering the interactions between polynomial multiplication, derivation, and the Fourier transform, we also get the following interesting result:

Proposition 2.1. *The Fourier transforms of linear combinations of δ and its derivatives are the complex polynomials.*

3. The unitary and angular Fourier Transforms

In these notes, we have presented the *unitary* Fourier transform, which gives the nicest theoretical results. Recall, though, that the term in the exponent changes by a constant when we rescale either the DFT or the interval on which we use Fourier series. In the case of the Fourier transform, since $\mathbb{R}^n$ is continuous and unbounded, we can rescale without changing the domain of the functions. In fact, this was one of the special cases of Proposition **??** (ii).

Another popular version of the Fourier transform, a rescaling of the unitary Fourier transform, is the *angular* Fourier transform, given by

$$\mathcal{F}f(\xi) = \int_{-\infty}^{\infty} f(x)e^{-ix\xi}.$$

This is simply a rescaling by $\frac{1}{2\pi}$. Although it as not as convenient for theory as the unitary Fourier transform—mainly because it is no longer a unitary isomorphism on L^2, but rather a rescaling of one—it is also easier to see the above formula in many applications, such as MRI, where the 2π in the exponent only appears if we force it to by rescaling either the location or the magnetic field. Epstein uses the angular Fourier transform exclusively; we will use whichever is more convenient in a particular setting.

It is possible to work out all the properties of the angular Fourier transform from Proposition **??**, using the rescaling rule, but we collect a list of the basic results below:

- (Inversion Formula) $f = \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} \hat{f}(\xi)e^{ix\cdot\xi}d\xi$,
- $\widehat{R_k f}(\xi) = e^{-ik\cdot\xi}\hat{f}$,
- $\widehat{e^{ik\cdot\xi}f} = R_k\hat{f}$,
- $\widehat{f * g} = \hat{f}\hat{g}$,
- $\widehat{f \circ T} = |\det T|^{-1}\hat{f} \circ (T^*)^{-1} \quad (T \in GL(n, \mathbb{R}))$,
- $\widehat{(ix)^\alpha f} = \partial^\alpha \hat{f}$,
- $\widehat{\partial^\alpha f} = (ix)^\alpha \hat{f}$,
- (Parseval's relation) $\langle f, g\rangle_{L^2} = \frac{1}{(2\pi)^n}\left\langle \hat{f}, \hat{g}\right\rangle_{L^2}$,
- (Plancherel formula) $\|f\|_2^2 = \frac{1}{(2\pi)^n}\left\|\hat{f}\right\|_2^2$.

Also, it can be confusing to keep track of the known transform pairs when dealing with the unitary and angular Fourier transforms. In this case, it is handy to keep in mind that

$$\mathcal{F}_{\text{unitary}}f(\xi) = \mathcal{F}_{\text{angular}}f(2\pi\xi).$$

Example 3.1. For example, we leave it to the reader to show that

$$\mathcal{F}_{\text{angular}}\chi_{[-L,L]}(\xi) = \frac{\sin(L\xi)}{L\xi}.$$

This is equivalent to

$$\mathcal{F}_{\text{angular}}\chi_{[-L,L]}(\xi) = \frac{\sin(2\pi L\xi)}{2\pi L\xi}.$$

The two functions above are together referred to as the sinc function. The particulars of the sinc function depend on the author; generally authors either use the non-normalized sinc function

$$\text{sinc}(x) = \frac{\sin(x)}{x},$$

the angular Fourier transform of $\chi_{[-1,1]}$, or the normalized sinc function

$$\operatorname{sinc}(x) = \frac{\sin(\pi x)}{\pi x},$$

the unitary Fourier transform of $\chi_{[-1/2,1/2]}$, the box function which has norm one in every L^p space. It follows that the 2-norm of the normalized sinc function is 1, and although sinc is not in L^1, its improper Riemann integral converges and is also equal to 1.

In these notes, barring an error on the part of the author, sinc will refer to the normalized sinc function, $\frac{\sin(\pi x)}{\pi x}$.

4. A remark on one- and multi-dimensional Fourier transform

In the notes on the DFT and on Fourier series, we presented the one-dimensional case first, and then showed that multidimensional functions can be dealt with by performing harmonic analysis on each coordinate separately. We did this primarily because vectors in $\mathbb{C}^n$ and sums over $\mathbb{Z}$ are more familiar and hence easier to analyze than in higher dimensions. Since functions on $\mathbb{R}^n$ are the typical setting for integration theory, we dispensed with this strategy and presented the Fourier transform on $\mathbb{R}^n$ directly. However, in applied settings it often proves useful to keep in mind that, as with the DFT and Fourier series, we can perform Fourier transform coordinate-wise.

For simplicity, assume that f is $L^1(\mathbb{R}^n)$. Then Fubini's theorem implies that we can evaluate the integral defining

$$f(\xi) = \int_{\mathbb{R}^n} f(x) e^{-2\pi i (x_1\xi_1 + x_2\xi_2 + \ldots + x_n\xi_n)}\, dx$$

iteratively, integrating a function of only x_j and ξ_j with respect to x_j. The same result holds for square-integrable functions, using the same approximation argument used to extend the Fourier transform from L^1 to L^2, because Fubini's theorem shows that if f is square-integrable then almost every section of f is also square-integrable.

References

[1] Walter Rudin. *Real Analysis: Modern Techniques and Their Applications.* John Wiley and Sons, New York, 2nd edition, 1999.

[2] Elias Stein and Rami Shakarchi. *Fourier Analysis: an Introduction.* Princeton University Press, Princeton, 2003.