MATH 262/CME 372: Applied Fourier Analysis and Elements of Modern Signal Processing Winter 2016

Lecture 3 — January 12, 2016

Prof. Emmanuel Candes

Scribe: Carlos A. Sing-Long, Edited by E. Bates

1 Outline

Agenda: Fourier Integrals (continued)

1. Parseval-Plancherel theorem
2. Fourier transform of square-integrable functions
3. Fourier transform of generalized functions
4. Fourier transform in higher dimensions

Last Time: We proved the Fourier inversion formula for functions with $f, \hat{f} \in L^1(\mathbb{R})$. We also proved the convolution theorem, which allowed us to describe time-invariant operators (TIOs) as a multiplication by the transfer function of the TIO in the Fourier domain.

2 Square-integrable functions

Recall from Lecture 2 the troublesome example of the boxcar function $f(t) = \mathbb{I}_{\{-1/2 \leq t \leq 1/2\}}$, whose Fourier transform is $\hat{f}(\omega) = \text{sinc}(\omega/2\pi)$. To our disappointment, the inversion formula as proved did not apply, as $\hat{f}$ is not integrable. But $\hat{f}$ happens to be *square*-integrable; that it, it is an element of the vector space

$$L^2(\mathbb{R}) = \left\{ f : \mathbb{R} \mapsto \mathbb{C} : \int |f(t)|^2 \, dt < \infty \right\}.$$

When endowed with the inner product

$$\langle f, \, g \rangle = \int \overline{f(t)} g(t) \, dt,$$

$L^2(\mathbb{R})$ is a **Hilbert space**. In other words, $L^2(\mathbb{R})$ is complete under the norm induced by the inner product

$$\|f\|_{L^2} = \sqrt{\langle f, \, f \rangle} = \sqrt{\int |f(t)|^2 \, dt}.$$

Without fear of confusion, $\| \cdot \|$ will always stand for $\| \cdot \|_{L^2}$. We call $\langle f, \, f \rangle$ the **energy** of f, and so L^2 is exactly the space of functions with finite energy.

Now we would like to define the Fourier transform of L^2 functions. To do so, we need to make sense of the integrals

$$\int f(t)e^{-i\omega t}\,dt \quad \text{and} \quad \frac{1}{2\pi}\int \hat{f}(\omega)e^{i\omega t}\,dt,$$

when $f, \hat{f} \notin L^1(\mathbb{R})$ but $f, \hat{f} \in L^2(\mathbb{R})$. In particular, we want to define an extension for which the inversion formula

$$\underbrace{f(t)}_{\text{e.g., } \mathbb{I}_{\{-1/2\leq t\leq 1/2\}}} = \frac{1}{2\pi}\int \underbrace{\hat{f}(\omega)}_{\text{e.g., } \frac{\sin(\omega/2)}{\omega/2}} e^{i\omega t}\,d\omega,$$

still holds. The key tool will be the Parseval-Plancherel theorem.

3 The Parseval-Plancherel theorem

The result is a simple yet important calculation.

Theorem 1 (Parseval-Plancherel)**.** *Let $f, g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Then*

$$\langle f,\, g\rangle = \frac{1}{2\pi}\langle \hat{f},\, \hat{g}\rangle.$$

Note: The above clearly implies $\|f\|^2 = (2\pi)^{-1}\|\hat{f}\|^2$. This implies the Fourier transform preserves the $L^2(\mathbb{R})$ norm (modulo a 2π factor), that is, it is (a multiple of) an **isometry**. Consequently, its inverse is (a multiple of) its adjoint.

Proof. Define $F(t) = \overline{f(-t)}$. Then a direct computation shows that $\hat{F}(\omega) = \overline{\hat{f}(\omega)}$, giving

$$\langle f,\, g\rangle = \int \overline{f(t)}g(t)\,dt = \int F(0-t)g(t)\,dt = (F * g)(0).$$

Since $F, g \in L^1(\mathbb{R})$, we can apply the convolution theorem (Lecture 2) to obtain

$$(F * g)(0) = \frac{1}{2\pi}\int \overline{\hat{f}(\omega)}\hat{g}(\omega)\,d\omega = \frac{1}{2\pi}\langle \hat{f},\, \hat{g}\rangle,$$

proving the claim. □

4 Fourier transform of square-integrable functions

Now we will use the Parseval-Plancherel theorem to extend the Fourier transform to $L^2(\mathbb{R})$ via a **density** argument. For any fixed $f \in L^2(\mathbb{R})$, we can construct a sequence $\{f_n\} \subset L^2 \cap L^1(\mathbb{R})$ so that $f_n \xrightarrow{L^2} f$. To see this, define $f_n(t) = \mathbb{I}_{\{-n\leq t\leq n\}}f(t)$. Then

- $f_n \in L^2(\mathbb{R})$ as clearly $\|f_n\| \leq \|f\| < \infty$.
- $f_n \in L^1(\mathbb{R})$ as $\|f_n\|_{L^1} \leq \sqrt{2n}\|f\| < \infty$ by the Cauchy-Schwarz inequality.

The Parseval-Plancherel theorem gives $\|\hat{g}\|^2 = \|g\|^2 < \infty$ for $g \in L^2 \cap L^1(\mathbb{R})$, and so

$$\|\hat{f}_n - \hat{f}_m\|^2 = 2\pi \|f_n - f_m\|^2.$$

Now, for $n \geq m$,

$$\|f_n - f_m\|^2 = \|\mathbb{I}_{\{m<t\leq n\}} f(t)\|^2 \leq \|\mathbb{I}_{\{m<t\}} f(t)\|^2 \quad \to \quad 0 \qquad \text{as } m \to \infty.$$

Consequently, $\{\hat{f}_n\}$ is a Cauchy sequence in $L^2(\mathbb{R})$ and thus converges to some $\hat{f} \in L^2(\mathbb{R})$ (since $L^2(\mathbb{R})$ is complete). We *define* the Fourier transform of f to be this limit; that is,

$$\hat{f}(\omega) := \lim_{n\to\infty} \int f_n(t) e^{-i\omega t}\, dt.$$

Remark: This limit does not depend on the sequence $\{f_n\}$!

This extension satisfies all the properties discussed so far, since these properties carry through to limits. In particular,

(i) $\|f\|^2 = (2\pi)^{-1} \|\hat{f}\|^2$ for $f \in L^2(\mathbb{R})$,

(ii) $\langle f,\, g\rangle = (2\pi)^{-1} \langle \hat{f},\, \hat{g}\rangle$ for $f, g \in L^2(\mathbb{R})$,

(iii) $f(t) = (2\pi)^{-1} \int \hat{f}(\omega) e^{i\omega t}\, d\omega,\ f \in L^2(\mathbb{R})$.

Note the inversion formula holds as an equality almost everywhere, which is sufficient for equality in the L^2 sense.

5 Fourier transform of distributions

In bold fashion, we shall extend our definition of the Fourier transform beyond even functions themselves! Consider, for instance, the Dirac δ function. We saw in Lecture 1 that this impulse has the property

$$\int \delta(t)\varphi(t)\; dt = \varphi(0), \tag{1}$$

for any sufficiently nice function φ. However, there is no actual function $\delta = \delta(t)$ for which (1) would be true for every φ. For this reason, we say δ is only a **distribution** or **generalized function**. That is, it is not a function, but rather a (linear) map of functions:

$$\delta: \quad \varphi \quad \mapsto \quad \delta(\varphi) \quad \triangleq \quad \varphi(0).$$

We can attempt to define the Fourier transform of this map. Indeed,

$$\int \delta(t) e^{-i\omega t} = e^0 = 1,$$

and so a reasonable guess is that the Fourier transform of δ is the constant function 1. Next we will define the Fourier transform of a distribution in general and check that our intuition is correct. We proceed by invoking a **duality** argument.

We start by considering really nice functions, say those that are smooth and rapidly decaying at infinity (certainly elements of L^1 and L^2). And from any function $f \in L^2(\mathbb{R})$ comes an associated linear map L_f acting on these nice functions:

$$L_f(\varphi) := \int f(t)\phi(t)\ dt = \langle \overline{f},\, \varphi \rangle.$$

This is how we think of functions as being a subset of generalized functions.

Next we consider an identity we know for nice functions, something that any reasonable definition of the Fourier transform must satisfy. Note that for $f, \varphi \in L^1(\mathbb{R})$, we have

$$\int \hat{f}(t)\varphi(t)\, dt = \iint f(u)\varphi(t)e^{-iut}\, dudt = \int f(u)\hat{\varphi}(u)\, du,$$

or

$$L_{\hat{f}}(\varphi) = L_f(\hat{\varphi}). \tag{2}$$

Accordingly, we define the Fourier transform of a distribution μ to be the distribution $\hat{\mu}$ that satisfies

$$\hat{\mu}(\varphi) = \mu(\hat{\varphi}) \qquad \text{for all smooth and rapidly decaying } \varphi \text{ (a \textbf{test function})}. \tag{3}$$

The identity (2) ensures that (3) is in fact an *extension* of our definition of Fourier transform for actual functions. By this construction, the Fourier transform of the delta distribution is given by

$$\hat{\delta}(\varphi) = \delta(\hat{\varphi}) = \hat{\varphi}(0) = \int \varphi(t)\, dt = \langle \overline{1},\, \phi \rangle = L_1(\phi).$$

The conclusion is that $\hat{\delta}$ is, indeed, the constant function 1.

Conveniently, the Fourier transform defined as in (3) maintains all the properties from Table 2, the reason being that the properties hold for all f and φ as above.

$F(t)$	$\hat{F}(\omega)$	
$(f * g)(t)$	$\hat{f}(\omega)\hat{g}(\omega)$	(convolution)
$f(t)g(t)$	$(2\pi)^{-1}(\hat{f} * \hat{g})(\omega)$	(multiplication)
$f(t-\tau)$	$e^{-i\omega\tau}\hat{f}(\omega)$	(time-shift)
$e^{i\omega_0 t}f(t)$	$\hat{f}(\omega - \omega_0)$	(modulation)
$f(t/a)$	$\lvert a\rvert\hat{f}(a\omega)$	(scaling)
$f^{(p)}(t)$	$(i\omega)^p\hat{f}(\omega)$	(differentiation)
$(-it)^p f(t)$	$\hat{f}^{(p)}(\omega)$	(multiplication by a monomial)

Table 1: Fourier tranform identities

Let us practice computing with some more examples:

1. Let δ_a be the shifted delta function: $\delta_a(\varphi) = \varphi(a)$. Then

$$\widehat{\delta_a}(\varphi) = \delta_a(\hat{\varphi}) = \hat{\varphi}(a) = \int \varphi(t)e^{-iat}\ dt,$$

and so $\hat{\delta}_a$ is the sine wave $e^{-ia\omega}$. We could have also seen this from the time-shift identify and the knowledge that $\hat{\delta} = 1$.

2. In turn, the Fourier transform of $f(t) = e^{iat}$ is the Fourier transform of 1 shifted by $\omega_0 = a$. That is,

$$\widehat{e^{iat}} = 2\pi\delta_a.$$

3. To compute $\hat{1}$, we can apply the inversion formula:

$$\hat{1}(\varphi) = L_1(\hat{\varphi}) = \int \hat{\varphi}(\omega)\, d\omega = 2\pi\varphi(0),$$

and so $\hat{1} = 2\pi\delta$.

4. Consider the Heaviside step function

$$H(t) = \begin{cases} 0 & t < 0 \\ 1 & t \geq 0. \end{cases}$$

The integral $\int H(t)e^{-i\omega t}\, dt$ does not converge, and so a more sophisticated approach is required to compute $\hat{H}$. One can try approximating $H(t)$ by truncations

$$H_n(t) = \begin{cases} 0 & t < 0 \text{ or } t > n \\ 1 & 0 \leq t \leq n, \end{cases}$$

but computing $\hat{H}_n$ poses the same problem. Let us instead approximate $H(t)$ by

$$f_a(t) = \begin{cases} 0 & t < 0 \\ e^{-\alpha t} & t \geq 0, \end{cases}$$

with $\alpha \downarrow 0$. The Fourier transform of f_a is given by

$$\int f_a(t)e^{-i\omega t}\, dt = \frac{-1}{a + i\omega}\bigg|_{t=0}^{\infty} = \frac{1}{a + i\omega} \quad \to \quad \frac{1}{i\omega} \qquad \text{as } a \downarrow 0.$$

The calculation fails, however, for $\omega = 0$. For this reason we return to our definition by duality:

$$\begin{aligned} \int \frac{1}{a + i\omega}\varphi(\omega)\, d\omega &= \int \frac{a - i\omega}{a^2 + \omega^2}\varphi(\omega)\, d\omega \\ &= -\int \arctan\left(\frac{\omega}{a}\right)\varphi'(\omega)\, d\omega - \int \frac{i\omega}{a^2 + \omega^2}\varphi(\omega)\, d\omega \end{aligned}$$

Now, considered in the principal value sense,

$$-\int \frac{i\omega}{a^2 + \omega^2}\varphi(\omega)\, d\omega \quad \to \quad \int \frac{1}{i\omega}\varphi(\omega)\, d\omega \qquad \text{as } a \downarrow 0.$$

And

$$\begin{aligned} -\int \arctan\left(\frac{\omega}{a}\right)\varphi'(\omega)\, d\omega &= -\int_{-\infty}^{0} \arctan\left(\frac{\omega}{a}\right)\varphi'(\omega)\, d\omega - \int_{0}^{\infty} \arctan\left(\frac{\omega}{a}\right)\varphi'(\omega)\, d\omega \\ &\to \int_{-\infty}^{0} \frac{\pi}{2}\varphi'(\omega)\, d\omega - \int_{0}^{\infty} \frac{\pi}{2}\varphi'(\omega)\, d\omega \qquad \text{as } a \downarrow 0 \\ &= \pi\varphi(0). \end{aligned}$$

The two limits together show

$$\hat{H}(\omega) = \frac{1}{i\omega} + \pi\delta(\omega).$$

5. Now we can compute the Fourier transform of the sign function

$$\operatorname{sgn}(t) = \begin{cases} -1 & t < 0 \\ 0 & t = 0 \\ 1 & t > 0. \end{cases}$$

Indeed, $\operatorname{sgn}(t) = 2H(t) - 1$, and so

$$\widehat{\operatorname{sgn}}(\omega) = \frac{2}{i\omega} + 2\pi\delta(\omega) - 2\pi\delta(\omega) = \frac{2}{i\omega}.$$

5.1 Derivative of distributions

Mimicking the duality argument for Fourier transforms above, we can also extend the notion of derivatives to generalized functions. Rather than Parseval-Plancherel, the relative identity now is integration by parts. For smooth test functions ψ and φ that decay at infinity, we know

$$\int \psi'(t)\varphi(t)\,dt = -\int \psi(t)\varphi'(t)\,dt, \qquad \text{or} \qquad L_{\psi'}(\varphi) = -L_{\psi}(\varphi').$$

So the **derivative** of a distribution μ should be the distribution μ' such that

$$\mu'(\varphi) = -\mu(\varphi').$$

In particular, this allows us to compute δ' as

$$\delta'(\varphi) = -\delta(\varphi') = -\varphi'(0).$$

So δ' is the distribution that maps a function φ to the negative of its derivative at the origin. We can now conclude with one more example: What is the Fourier transform of $f(t) = t$? We have

$$L_{\hat{t}}(\varphi) = L_t(\hat{\varphi}) = \int \omega\hat{\varphi}(\omega)\,d\omega = -i\,\frac{d}{dt}\int \hat{\varphi}(\omega)e^{i\omega t}\,d\omega\Big|_{t=0} = -2\pi i\varphi'(0),$$

and so $\hat{t} = 2\pi i\delta'$. Notice that this calculation obeys the identity

$$F(t) = tf(t) \quad \Rightarrow \quad \hat{F}(\omega) = i\hat{f}'(\omega)$$

with $f(t) = 1$.

6 The Fourier transform in higher dimensions

We can define the Fourier transforms of functions in d variables as

$$\hat{f}(\omega_1, \ldots, \omega_d) = \int f(x_1, \ldots, x_d)e^{-i(\omega_1 x_1 + \ldots \omega_d x_d)}\,dx_1 \ldots dx_d.$$

This is equivalent to taking the Fourier transform in each argument independently. Using the notation

$$\begin{aligned} \boldsymbol{x} &= (x_1, \ldots, x_d), \\ \boldsymbol{\omega} &= (\omega_1, \ldots, \omega_d), \\ d\boldsymbol{x} &= dx_1 \ldots dx_d, \\ d\boldsymbol{\omega} &= d\omega_1 \ldots d\omega_d, \end{aligned}$$

and endowing $\mathbb{R}^d$ with the standard inner product $\langle \cdot , \cdot \rangle$, we can write the simpler expression

$$\hat{f}(\boldsymbol{\omega}) = \int f(\boldsymbol{x}) e^{-i\langle \boldsymbol{\omega}, \boldsymbol{x} \rangle} \, d\boldsymbol{x}.$$

This extension satisfies all the properties discussed so far. In particular,

(i) **Inversion formula:** If $f, \hat{f} \in L^1(\mathbb{R}^d)$ then

$$f(\boldsymbol{x}) = \frac{1}{(2\pi)^d} \int \hat{f}(\boldsymbol{\omega}) e^{i\langle \boldsymbol{\omega}, \boldsymbol{x} \rangle} \, d\boldsymbol{x},$$

(ii) **Convolution theorem:** If $f, g \in L^1(\mathbb{R}^d)$ then

$$h(\boldsymbol{x}) = \int f(\boldsymbol{y}) g(\boldsymbol{x} - \boldsymbol{y}) \, d\boldsymbol{y} \quad \Longrightarrow \quad \hat{h}(\boldsymbol{\omega}) = \hat{f}(\boldsymbol{\omega}) \hat{g}(\boldsymbol{\omega}),$$

(iii) **Parseval-Plancherel theorem:** If $f \in L^2(\mathbb{R}^d)$, then

$$\int |f(\boldsymbol{x})|^2 \, d\boldsymbol{x} = \frac{1}{(2\pi)^d} \int |\hat{f}(\boldsymbol{\omega})|^2 \, d\boldsymbol{\omega}.$$

Furthermore, the extension to square-integrable functions is the same. Table 2 shows the multi-dimensional versions of some of the familiar properties.

$F(\boldsymbol{x})$	$\hat{F}(\boldsymbol{\omega})$	
$\partial^\alpha f(\boldsymbol{x})$	$i^{\lvert\alpha\rvert} \boldsymbol{\omega}^\alpha \hat{f}(\boldsymbol{\omega})$	(differentiation)
$f(\boldsymbol{x} - \boldsymbol{x}_0)$	$e^{-i\langle \boldsymbol{\omega}, \boldsymbol{x}_0 \rangle} \hat{f}(\boldsymbol{\omega})$	(translation)
$\boldsymbol{x}^\alpha f(\boldsymbol{x})$	$i^{\lvert\alpha\rvert} \partial^\alpha \hat{f}(\boldsymbol{\omega})$	(multiplication by a monomial)

Table 2: Some properties of the Fourier transform

7 The Fourier transform in optics

Here we consider the experimental setup in Fig. 1a. We have a screen with an aperture where the object has been placed. Monochromatic light shines through this object, and we observe the scattered light on a screen at a distance z_1. We model the object and aperture as a function $f = f(x_1, x_2)$ and the pattern of scattered light observed at the screen as $g = g(x_1, x_2)$. If $z_1 \gg 1$

then we observe the **far-field** of the scattered wave, which corresponds in optic to what is known as **Fraunhoffer diffraction**. Surprisingly, we will see later on the course that

$$g(x_1, x_2) \propto \hat{f}\left(\frac{2\pi x_1}{\lambda_0 z_1}, \frac{2\pi x_2}{\lambda_0 z_1}\right),$$

when z_1 is large (here λ_0 is the wavelength of the incident wave).

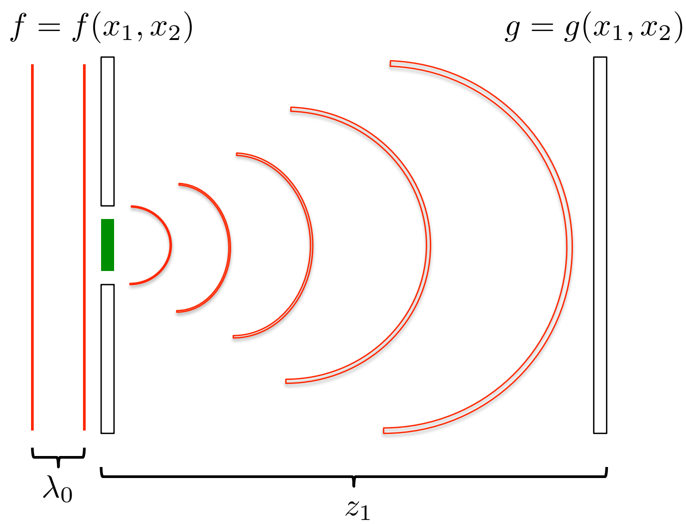
From this we see that if the aperture is square we can write $f(x_1, x_2) = \mathbb{I}_{\{|x_1| \leq R/2\}} \mathbb{I}_{\{|x_2| \leq R/2\}}$ and thus

$$g(x_1, x_2) \propto R^2 \operatorname{sinc}\left(\frac{x_1}{\lambda_0 z_1/R}\right) \operatorname{sinc}\left(\frac{x_2}{\lambda_0 z_1/R}\right).$$

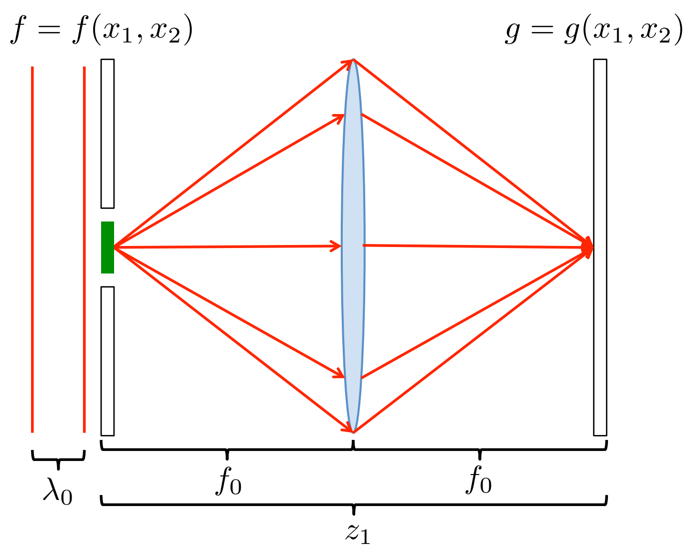
In practical applications z_1 does not need to be very large. The optical setup in Fig. 1b produces similar results. In many cases we do not measure the scattered field g at the screen, but rather its intensity, that is, we measure

$$I(x_1, x_2) \propto |g(x_1, x_2)|^2,$$

which is the case of X-ray crystallography, among others.



(a) Typical experimental setup.



(b) Modified setup with a lens. Here f_0 denotes the focal distance of the lens.

Figure 1: Experimental setup used in diffraction experiments.