

FOURIER INVERSION

1. THE FOURIER TRANSFORM AND THE INVERSE FOURIER TRANSFORM

Consider functions

$$f, g : \mathbf{R}^n \longrightarrow \mathbf{C},$$

and consider the bilinear, symmetric function

$$\psi : \mathbf{R}^n \times \mathbf{R}^n \longrightarrow \mathbf{C}^\times, \quad \psi(x, \xi) = \exp(2\pi i x \cdot \xi),$$

an additive character in each of its arguments. The *Fourier transform* of f is

$$\mathcal{F}f : \mathbf{R}^n \longrightarrow \mathbf{C}, \quad \mathcal{F}f(\xi) = \int_{x \in \mathbf{R}^n} f(x) \overline{\psi(x, \xi)} dx$$

and the *inverse Fourier transform* of g is

$$\mathcal{F}^{-1}g : \mathbf{R}^n \longrightarrow \mathbf{C}, \quad \mathcal{F}^{-1}g(x) = \int_{\xi \in \mathbf{R}^n} g(\xi) \psi(x, \xi) d\xi,$$

provided that the integrals exist. Because $\mathcal{F}^{-1}g(x) = \mathcal{F}g(-x)$, properties of the Fourier transform extend instantly to the inverse Fourier transform, and so the details of the ensuing discussion are limited to the Fourier transform.

2. A HEURISTIC ARGUMENT FOR FOURIER INVERSION

By analogy to familiar symbol-patterns from the context of finite-dimensional vector spaces or Hilbert space, it is tempting to think of the Fourier transform of f at ξ as the inner product of f and the frequency- ξ oscillation $\psi_\xi(x) = \psi(x, \xi)$,

$$\mathcal{F}f(\xi) = \int_{x \in \mathbf{R}^n} f(x) \overline{\psi_\xi(x)} dx = \langle f, \psi_\xi \rangle.$$

And then it is tempting to think of the inverse Fourier transform of the Fourier transform as a synthesis of these coefficient inner products against the relevant oscillations,

$$\mathcal{F}^{-1}\mathcal{F}f(x) = \int_{\xi \in \mathbf{R}^n} \langle f, \psi_\xi \rangle \psi_\xi(x) d\xi.$$

And so our experience suggests that the process should recover the original function,

$$\mathcal{F}^{-1}\mathcal{F}f = f.$$

But this reasoning is only suggestive until we find some mathematical framework where it is meaningful and correct.

3. SCHWARTZ FUNCTIONS, FIRST STATEMENT OF FOURIER INVERSION

Fourier analysis shows that

- *The smoother f is, the faster $\mathcal{F}f$ decays.* Specifically, if all the partial derivatives of f up to some order k exist and are integrable, then $\mathcal{F}f(x)$ decreases at least as quickly as $|x|^{-k}$ as $|x| \rightarrow \infty$.
- *The faster f decays, the smoother $\mathcal{F}f$ is.* Specifically, if $f(x)$ decreases at least as quickly as $|x|^{-k}$ as $|x| \rightarrow \infty$ then $\mathcal{F}f$ has continuous, bounded derivatives up to order $k - n - 1$.

These relations between smoothness and decay are asymmetric in several ways.

The class of *Schwartz functions* is designed to render the asymmetries irrelevant, so that the functions in question should behave well under Fourier analysis. A Schwartz function on \mathbf{R}^n is an infinitely smooth function all of whose derivatives (including itself) decay rapidly. That is, a Schwartz function is a function

$$\varphi : \mathbf{R}^n \longrightarrow \mathbf{C}$$

such that

- φ is a \mathcal{C}^∞ -function.
- For $k \in \mathbf{Z}_{\geq 0}$ and each $d \in \mathbf{Z}_{\geq 1}$, $|\varphi^{(k)}(x)| \leq |x|^{-d}$ for all large enough $|x|$.

The Fourier transform and the inverse Fourier transform of a Schwartz function are again Schwartz functions. The Fourier inversion formula is

$$\boxed{\mathcal{F}^{-1}\mathcal{F}\varphi = \varphi \quad \text{for Schwartz functions } \varphi.}$$

Granting this formula, it follows that also

$$\mathcal{F}\mathcal{F}^{-1}\varphi = \varphi \quad \text{for Schwartz functions } \varphi.$$

Indeed, define the operator $(M\varphi)(x) = \varphi(-x)$. Then $M^2 = 1$ and short calculations show that

$$\mathcal{F}M = \mathcal{F}^{-1}, \quad \mathcal{F}^{-1}M = \mathcal{F}, \quad M\mathcal{F} = \mathcal{F}^{-1}, \quad M\mathcal{F}^{-1} = \mathcal{F}.$$

Thus $\mathcal{F}\mathcal{F}^{-1} = \mathcal{F}^{-1}M\mathcal{F} = \mathcal{F}^{-1}\mathcal{F} = \text{id}$.

Incidentally, Fourier inversion and the formulas in the previous display show that $\mathcal{F}^2 = M$ and so $\mathcal{F}^4 = \text{id}$. Thus the operator $1 + \mathcal{F} + \mathcal{F}^2 + \mathcal{F}^3$ is \mathcal{F} -invariant, i.e., for any Schwartz function φ , the resulting function $(1 + \mathcal{F} + \mathcal{F}^2 + \mathcal{F}^3)\varphi$ is its own Fourier transform. The Gaussian is emphatically not the only such function.

4. REDUCTION TO THE CASE $x = 0$

To establish Fourier inversion we need to show that for any Schwartz function φ and for any point $x \in \mathbf{R}^n$,

$$\mathcal{F}^{-1}\mathcal{F}\varphi(x) = \varphi(x).$$

However, Fourier inversion reduces to the normalized case $x = 0$ as follows. For any $x \in \mathbf{R}^n$, introduce the x -translation operator,

$$T_x : \{\text{Schwartz functions}\} \longrightarrow \{\text{Schwartz functions}\}, \quad T_x\phi(y) = \phi(x + y),$$

and recall the frequency- x oscillation function,

$$\psi_x : \mathbf{R}^n \longrightarrow \mathbf{C}, \quad \psi_x(\xi) = \psi(x, \xi) = \exp(2\pi i x \cdot \xi).$$

Then we have the two identities (exercise)

$$\left\{ \begin{array}{l} T_x \mathcal{F}^{-1} \phi = \mathcal{F}^{-1}(\psi_x \phi) \\ \psi_x \mathcal{F} \phi = \mathcal{F} T_x \phi \end{array} \right\} \text{ for Schwartz functions } \phi.$$

Compute, using the two identities,

$$\begin{aligned} (\mathcal{F}^{-1} \mathcal{F} \varphi)(x) &= (T_x \mathcal{F}^{-1} \mathcal{F} \varphi)(0) \\ &= (\mathcal{F}^{-1}(\psi_x \mathcal{F} \varphi))(0) && \text{by the first identity} \\ &= (\mathcal{F}^{-1} \mathcal{F} T_x \varphi)(0) && \text{by the second identity} \\ &= (T_x \varphi)(0) && \text{granting Fourier inversion at 0} \\ &= \varphi(x). \end{aligned}$$

So indeed we may take $x = 0$.

5. A SECOND HEURISTIC ARGUMENT

Formally, we would like to establish Fourier inversion at 0 as follows:

$$\begin{aligned} \mathcal{F}^{-1} \mathcal{F} \varphi(0) &= \int_{\xi} \mathcal{F} \varphi(\xi) d\xi \\ &= \int_{\xi} \int_x \varphi(x) \bar{\psi}(x, \xi) dx d\xi \\ &= \int_x \varphi(x) \int_{\xi} \bar{\psi}(x, \xi) d\xi dx \\ &= \int_x \varphi(x) \delta(x) dx \quad (\text{Dirac delta}) \\ &= \varphi(0). \end{aligned}$$

But this calculation has many problems. The double integral

$$\iint_{(x, \xi) \in \mathbf{R}^2} \varphi(x) \bar{\psi}(x, \xi) d\mu(x, \xi)$$

is not absolutely convergent (the integrand has absolute value $|\varphi(x)|$), and so interchanging the iterated integrals is unjustified. Furthermore, the integral

$$\int_{\xi} \bar{\psi}(x, \xi) d\xi$$

diverges classically. The ideas that the integral converges to the Dirac delta function $\delta(x)$ and that integrating $\varphi(x)\delta(x)$ gives $\varphi(0)$ are sensible only in the context of distribution theory.

6. THE DAMPENED GAUSSIAN AND ITS FOURIER TRANSFORM

The just-mentioned problems are circumvented by the *Gaussian trick*. It requires the Fourier transform of the n -dimensional dilated Gaussian function. To begin, recall that the one-dimensional Gaussian function,

$$\gamma : \mathbf{R} \longrightarrow \mathbf{R}, \quad \gamma(x) = e^{-\pi x^2},$$

is its own Fourier transform. It follows easily that so is the n -dimensional Gaussian,

$$g : \mathbf{R}^n \longrightarrow \mathbf{R}, \quad g(x) = e^{-\pi |x|^2}.$$

A small calculation gives the Fourier transform of a general dilation. Given a function $f : \mathbf{R}^n \rightarrow \mathbf{C}$ and given $t > 0$, the t -dilation of f is the function

$$f_t : \mathbf{R}^n \rightarrow \mathbf{C}, \quad f_t(x) = f(tx).$$

Compute that

$$\begin{aligned} \mathcal{F}f_t(\xi) &= \int_{x \in \mathbf{R}^n} f_t(x) \overline{\psi}(x, \xi) dx \\ &= \int_{x \in \mathbf{R}^n} f(tx) \overline{\psi}(tx, \xi/t) d(tx)/t^n = t^{-n} \mathcal{F}f(\xi/t). \end{aligned}$$

Now consider a dilated Gaussian function for any $t > 0$,

$$g_{\sqrt{t}} : \mathbf{R}^n \rightarrow \mathbf{C}, \quad g_{\sqrt{t}}(x) = e^{-\pi t|x|^2}.$$

Its Fourier transform is

$$\mathcal{F}g_{\sqrt{t}}(\xi) = t^{-n/2} g_{1/\sqrt{t}}(\xi),$$

or, introducing a new function-name,

$$\mathcal{F}g_{\sqrt{t}}(\xi) = \phi_t(\xi) \quad \text{where} \quad \phi_t(x) = t^{-n/2} e^{-\pi|x|^2/t}.$$

The function ϕ_t is an *approximate identity*, meaning that

- (1) $\phi_t(z) \geq 0$ for all z ,
- (2) $\int_x \phi_t(x) dx = 1$,
- (3) Given any $\varepsilon > 0$ and any $\delta > 0$, if t is close enough to 0 then $\phi_t(x) < \varepsilon$ for all $|x| > \delta$. That is, ϕ_t becomes a tall narrow pulse as $t \rightarrow 0$.

To summarize, the utility of the dilated Gaussian is that

- Both $g_{\sqrt{t}}$ and $\mathcal{F}g_{\sqrt{t}}$ decay rapidly for each $t > 0$, so that multiplying some other function by either of them will dampen the other function.
- Because $g_{\sqrt{t}}$ is a dilated Gaussian, if $t \rightarrow 0$ then $g_{\sqrt{t}}$ approaches the constant function 1 pointwise (despite always decaying rapidly far from the origin) and so the product of some other function and $g_{\sqrt{t}}$ approaches the other function.
- Because the Fourier transform $\mathcal{F}g_{\sqrt{t}} = \phi_t$ is an approximate identity, if $t \rightarrow 0$ then ϕ_t becomes a tall narrow pulse and so some other function times ϕ_t is a weighted average of the other function with the weighting concentrated closely about the origin.

7. PROOF OF FOURIER INVERSION

Recall that the second heuristic argument for Fourier inversion ran up against the absolutely divergent double integral

$$\iint_{(x, \xi) \in \mathbf{R}^2} \varphi(x) \overline{\psi}(x, \xi) d\mu(x, \xi).$$

Consider instead a similar double integral but with its integrand dampened by the t -dilated Gaussian function $g_{\sqrt{t}}$,

$$\iint_{(x, \xi) \in \mathbf{R}^2} \varphi(x) g_{\sqrt{t}}(\xi) \overline{\psi}(x, \xi) d\mu(x, \xi).$$

Now the integral is absolutely convergent thanks to the rapid decay of the Gaussian, and so the two corresponding iterated integrals exist and are equal. One such iterated integral is

$$A_t = \int_{\xi} \int_x \varphi(x) \overline{\psi}(x, \xi) dx g_{\sqrt{t}}(\xi) d\xi = \int_{\xi} \mathcal{F}\varphi(\xi) g_{\sqrt{t}}(\xi) d\xi,$$

and the other is

$$B_t = \int_x \varphi(x) \int_{\xi} g_{\sqrt{t}}(\xi) \overline{\psi}(x, \xi) d\xi dx = \int_x \varphi(x) \mathcal{F}g_{\sqrt{t}}(x) dx = \int_x \varphi(x) \phi_t(x) dx.$$

For the first integral, we have by the Dominated Convergence Theorem that since $g_{\sqrt{t}} \rightarrow 1$ pointwise as $t \rightarrow 0$,

$$A_t = \int_{\xi} \mathcal{F}\varphi(\xi) g_{\sqrt{t}}(\xi) d\xi \xrightarrow{t \rightarrow 0} \int_{\xi} \mathcal{F}\varphi(\xi) d\xi = \mathcal{F}^{-1}\mathcal{F}\varphi(0).$$

For the second integral, a standard argument (that manages the relevant issues by breaking the calculation into three pieces) says that as a special case of the *Approximate Identity Theorem*,

$$B_t = \int_x \varphi(x) \phi_t(x) dx \xrightarrow{t \rightarrow 0} \varphi(0).$$

But $A_t = B_t$ for all $t > 0$, and so the limits of A_t and B_t as $t \rightarrow 0$ are equal. Thus we have the Fourier inversion formula at 0,

$$\mathcal{F}^{-1}\mathcal{F}\varphi(0) = \varphi(0) \quad \text{for Schwartz functions } \varphi.$$

As explained above, general Fourier inversion for Schwartz functions follows.

8. FOURIER INVERSION FOR TEMPERED DISTRIBUTIONS

Now we discuss Fourier inversion in the broader environment \mathcal{S}^* of *tempered distributions*, the space of continuous linear functionals on the space \mathcal{S} of Schwartz functions. The Fourier transform on \mathcal{S}^* is defined by an adjoint-like characterizing property,

$$\langle \mathcal{F}f, \varphi \rangle = \langle f, \mathcal{F}\varphi \rangle \quad \text{for all } f \in \mathcal{S}^* \text{ and } \varphi \in \mathcal{S}.$$

This is the only possible definition compatible with viewing \mathcal{S} as a subspace of \mathcal{S}^* , because if both f and φ are Schwartz functions then

$$\begin{aligned} \langle \mathcal{F}f, \varphi \rangle &= \int_{\xi} \mathcal{F}f(\xi) \varphi(\xi) d\xi = \int_{\xi} \int_x f(x) \overline{\psi}(x, \xi) dx \varphi(\xi) d\xi \\ &= \int_x f(x) \int_{\xi} \varphi(\xi) \overline{\psi}(\xi, x) d\xi dx = \int_x f(x) \mathcal{F}\varphi(x) dx = \langle f, \mathcal{F}\varphi \rangle. \end{aligned}$$

Also, the definition of the Fourier transform on \mathcal{S}^* immediately extends Fourier inversion to \mathcal{S}^* , because for all $f \in \mathcal{S}^*$ and $\varphi \in \mathcal{S}$,

$$\langle \mathcal{F}^{-1}\mathcal{F}f, \varphi \rangle = \langle \mathcal{F}f, \mathcal{F}^{-1}\varphi \rangle = \langle f, \mathcal{F}\mathcal{F}^{-1}\varphi \rangle = \langle f, \varphi \rangle,$$

so that

$$\mathcal{F}^{-1}\mathcal{F}f = f \quad \text{for tempered distributions } f.$$

9. THE ENCODING IDENTITY

Fourier inversion is encoded concisely as one particular equality of tempered distributions,

$$\boxed{\mathcal{F}1 = \delta.}$$

Here 1 is the constant function 1 viewed as a tempered distribution,

$$\langle 1, \varphi \rangle = \int_{x \in \mathbf{R}^n} \varphi(x) dx \quad \text{for Schwartz functions } \varphi,$$

and δ is the Dirac delta distribution,

$$\langle \delta, \varphi \rangle = \varphi(0) \quad \text{for Schwartz functions } \varphi.$$

Thus a small point here is that although 1 is a classical function, its Fourier transform is not.

We know that general Fourier inversion for Schwartz functions follows from Fourier inversion for Schwartz functions at 0. To see that Fourier inversion for Schwartz functions at 0 is precisely the content of the boxed identity, compute for any Schwartz function φ , that on the one hand

$$\langle \mathcal{F}1, \varphi \rangle = \langle 1, \mathcal{F}\varphi \rangle = \mathcal{F}^{-1}\mathcal{F}\varphi(0),$$

while on the other hand

$$\langle \delta, \varphi \rangle = \varphi(0).$$

The left sides are equal if and only if the right sides are equal. That is, the identity $\mathcal{F}1 = \delta$ is equivalent to Fourier inversion for Schwartz functions at 0.

10. ENDING COMMENTS

We end with some comments of what to make of the encoding identity $\mathcal{F}1 = \delta$.

One comment amplifies a remark from the beginning of the writeup. In physical terms, the equality says that the constant function 1 is built entirely from the oscillation of frequency zero, i.e., itself. Thinking in these terms tacitly takes Fourier inversion for granted in some unspecified mathematical environment. The correct environmental formulation is that Fourier inversion encodes the synthesis of *tempered distributions* from oscillations, the oscillations and coefficient “functions” themselves being tempered distributions.

Another point of view is that the equality assigns a distributional value to a divergent integral,

$$\int_{\xi \in \mathbf{R}^n} e^{-2\pi i x \cdot \xi} d\xi = \delta(x) \quad \text{for all } x \in \mathbf{R}^n.$$

This idea is consonant with the observation that although the distribution 1 is a classical function, its Fourier transform δ is not. (The Fourier transform of any *Schwartz function* is again a Schwartz function, but the constant function 1 is not Schwartz.)

A similar-looking identity of distributions,

$$\mathcal{F}\delta = 1,$$

is immediate. Indeed, for any Schwartz function φ ,

$$\langle \mathcal{F}\delta, \varphi \rangle = \langle \delta, \mathcal{F}\varphi \rangle = \mathcal{F}\varphi(0) = \langle 1, \varphi \rangle.$$

However, getting from the trivial identity $\mathcal{F}\delta = 1$ to the encoding identity $\mathcal{F}1 = \delta$ is yet again a matter of Fourier inversion.

Returning to the dampened Gaussian function, assuming that its rapid decay allows us to pass a limit through a Fourier transform, its other two properties then quickly establish the encoding identity,

$$\begin{aligned}\mathcal{F}1 &= \mathcal{F} \lim_{t \rightarrow 0} g_{\sqrt{t}} \quad \text{because } g_{\sqrt{t}} \rightarrow 1 \text{ as } t \rightarrow 0 \\ &= \lim_{t \rightarrow 0} \mathcal{F}g_{\sqrt{t}} \\ &= \delta \quad \text{because } \mathcal{F}g_{\sqrt{t}} = \phi_t \text{ is an approximate identity.}\end{aligned}$$

But justifying the second step of the calculation requires the same work that we did earlier to establish Fourier inversion.