FOURIER INVERSION

Contents

1. The Fourier transform and the inverse Fourier transform

Let n be a positive integer. Consider functions

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

and consider the bilinear, symmetric function

$$\psi : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{C}^\times, \qquad \psi(x, \xi) = \exp(i\, x \cdot \xi),$$

an additive character in each of its arguments. Introduce a constant and a rescaled measure on $\mathbb{R}^n$,

$$c = (2\pi)^{-n/2}, \qquad \mu_c = c\mu \text{ (where } \mu \text{ is the usual measure)},$$

so that $\mathrm{d}_c x = c\,\mathrm{d}x$, $\mathrm{d}_c\xi = c\,\mathrm{d}\xi$, and so on. The *Fourier transform* of f is

$$\mathcal{F}f : \mathbb{R}^n \longrightarrow \mathbb{C}, \qquad \mathcal{F}f(\xi) = \int_{x \in \mathbb{R}^n} f(x)\overline{\psi}(x, \xi)\,\mathrm{d}_c x$$

and the *inverse Fourier transform* of g is

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

provided that the integrals exist. Here we think of x as a physical space variable and of ξ as a frequency space variable, so that the Fourier transform takes physical space functions to frequency space functions, and the inverse Fourier transform conversely, but we don't always hold to this. 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 discussion to follow are limited to the Fourier transform. The inverse Fourier transform is exactly the Fourier transform for even functions.

Working now with one variable, the Gaussian function $\gamma(x) = e^{-x^2/2}$ is its own Fourier transform. Indeed, it satisfies the differential equation $\gamma'(x) = -x\gamma(x)$ and the initial condition $\gamma(0) = 1$. Its Fourier transform $\mathcal{F}\gamma$ satisfies the same differential equation, as is shown by differentiating under the integral sign and then recognizing $-x\gamma(x)$ as $\gamma'(x)$ and then integrating by parts,

$$\begin{aligned}(\mathcal{F}\gamma)'(\xi) &= \int_{\mathbb{R}} \gamma(x)\frac{\partial}{\partial\xi}\overline{\psi}_\xi(x)\mathrm{d}_c x = \int_{x=-\infty}^{\infty}(-ix)\gamma(x)\overline{\psi}_\xi(x)\mathrm{d}_c x \\ &= i\int_{x=-\infty}^{\infty}\frac{\mathrm{d}}{\mathrm{d}x}\gamma(x)\overline{\psi}_\xi(x)\mathrm{d}_c x = -i\int_{x=-\infty}^{\infty}\gamma(x)\frac{\partial}{\partial x}\overline{\psi}_\xi(x)\mathrm{d}_c x \\ &= -\xi\int_{x=-\infty}^{\infty}\gamma(x)\overline{\psi}_\xi(x)\mathrm{d}_c x = -\xi\,\mathcal{F}\gamma(\xi).\end{aligned}$$

Also $(\mathcal{F}\gamma)(0) = \int_{\mathbb{R}}\gamma(x)\mathrm{d}_c x = 1$, which is to say that $\mathcal{F}\gamma$ satisfies the same initial condition as γ as well. Thus $\mathcal{F}\gamma = \gamma$. Because the Gaussian function is even, also $\mathcal{F}^{-1}\gamma = \gamma$.

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 natural to think of the Fourier transform of f at ξ as the inner product of f and the frequency-ξ oscillation $\psi_\xi(x) = \psi(x,\xi)$,

$$\mathcal{F}f(\xi) = \int_{x\in\mathbb{R}^n} f(x)\overline{\psi}_\xi(x)\,\mathrm{d}_c x = \langle f, \psi_\xi\rangle.$$

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

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

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 absolutely integrable, then $\mathcal{F}f(\xi)$ decreases at least as quickly as $|\xi|^{-k}$ as $|\xi| \to \infty$. The relevant formula here comes from integration by parts,

  $$\mathcal{F}f(\xi) = (i\xi)^{-\alpha}\int_{\mathbb{R}^n} f^{(\alpha)}(x)\overline{\psi}_\xi(x)\,\mathrm{d}_c x, \quad \alpha = (\alpha_1,\dots,\alpha_n),$$

  in which $(i\xi)^{-\alpha}$ abbreviates $\prod_i (i\xi_i)^{-\alpha_i}$.

- *The faster f decays, the smoother $\mathcal{F}f$ is.* Specifically, if $f(x)$ decreases at least as quickly as $|x|^{-k}$ as $|x| \to \infty$ then $\mathcal{F}f$ has continuous, bounded derivatives up to order $k-n-1$. The relevant formula here comes from differentiation under the integral sign,

$$(\mathcal{F}f)^{(\alpha)}(\xi) = \int_{\mathbb{R}^n} (-ix)^\alpha f(x)\overline{\psi}_\xi(x)\,\mathrm{d}_c x, \quad \alpha = (\alpha_1, \dots, \alpha_n).$$

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 $\mathbb{R}^n$ is an infinitely smooth function all of whose derivatives (including itself) decay rapidly. That is, a Schwartz function is a function

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

such that

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

So the Schwartz functions are the collection of functions from $\mathbb{R}^n$ to $\mathbb{C}$ that decay and are preserved under multiplication by the components x_i of their input vectors x and are preserved under differentiation. 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 $M^2 = \mathrm{id}$ 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 $\mathcal{F}\mathcal{F}^{-1} = \mathcal{F}MM\mathcal{F}^{-1} = \mathcal{F}^{-1}\mathcal{F} = \mathrm{id}$.

Incidentally, Fourier inversion and the formulas in the previous display show that $\mathcal{F}^2 = M$ and so $\mathcal{F}^4 = \mathrm{id}$. 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. For a specific example, going back to Hecke, let p be any harmonic homogeneous polynomial of degree d and let φ be the Gaussian; then (not at all trivially) $\mathcal{F}(p\varphi) = i^{-d}p\varphi$, giving $p\varphi$ again if d is a multiple of 4. More generally, because $\mathcal{F}^4 = \mathrm{id}$ the only eigenvalues of $\mathcal{F}$ are $\{\pm 1, \pm i\}$, and the operator decomposition

$$\begin{aligned} 4\,\mathrm{id} = {} & \mathrm{id} + \mathcal{F} + \mathcal{F}^2 + \mathcal{F}^3 \\ & + \mathrm{id} - \mathcal{F} + \mathcal{F}^2 - \mathcal{F}^3 \\ & + \mathrm{id} - i\mathcal{F} - \mathcal{F}^2 + i\mathcal{F}^3 \\ & + \mathrm{id} + i\mathcal{F} - \mathcal{F}^2 - i\mathcal{F}^3 \end{aligned}$$

shows how to decompose any Schwartz function into a sum of functions in the eigenspaces. For more on these topics, see `http://www.math.umn.edu/~garrett/m/number_theory/Notes_2011-12.pdf`

and

`http://www.math.umn.edu/~garrett/m/number_theory/overheads/noth-03-26-2012.pdf`

4. Consequence of Fourier inversion: L^2-isometry

The bilinear inner product on the space of Schwartz functions is

$$\langle f, g\rangle = \int_{\mathbb{R}^n} f \cdot g.$$

The Fourier transform is self-adjoint with respect to this inner product. That is, for Schwartz functions f and g,

$$\begin{aligned}\langle \mathcal{F}f, g\rangle &= \int_{\mathbb{R}^n} (\mathcal{F}f \cdot g)(x)\,\mathrm{d}x = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(y)\overline{\psi}(x,y)\,\mathrm{d}_c y\, g(x)\,\mathrm{d}x \\ &= \int_{\mathbb{R}^n} f(y)\int_{\mathbb{R}^n} \overline{\psi}(x,y)g(x)\,\mathrm{d}_c x\,\mathrm{d}y = \int_{\mathbb{R}^n} (f\cdot \mathcal{F}g)(y)\,\mathrm{d}y = \langle f, \mathcal{F}g\rangle,\end{aligned}$$

with the equality of iterated integrals holding via Fubini's Theorem because the integrand decreases rapidly in both directions. In particular, given a Schwartz function φ, note that $\overline{\mathcal{F}\varphi} = \mathcal{F}^{-1}\overline{\varphi}$ and compute that consequently,

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

or

$$|\mathcal{F}\varphi|^2_{L^2} = |\varphi|^2_{L^2} \qquad \text{(equality of squared } L^2\text{-norms)}.$$

That is, the self-adjointness of the Fourier transform and Fourier inversion quickly show that the Fourier transform is an L^2-isometry of the Schwartz space.

Here we used the bilinear inner product rather than the sesquilinear L^2 inner product $\langle f, g\rangle = \int_{\mathbb{R}^n} f\overline{g}$ on the space of Schwartz functions because the former dovetails with distribution theory, as we will see later in this writeup. The argument has shown that the Fourier transform acts unitarily with respect to the L^2 inner product.

5. Reduction to the case $x = 0$

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

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

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

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

Then we have the two identities (exercise)

$$\begin{Bmatrix} T_x\mathcal{F}^{-1}\phi = \mathcal{F}^{-1}(\psi_x\phi) \\ \psi_x\mathcal{F}\phi = \mathcal{F}T_x\phi \end{Bmatrix} \quad \text{for Schwartz functions } \phi.$$

Compute, using the two identities,

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

So indeed we may take $x = 0$.

6. A second heuristic argument

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

$$\begin{aligned}
\mathcal{F}^{-1}\mathcal{F}\varphi(0) &= \int_\xi \mathcal{F}\varphi(\xi)\,\mathrm{d}_c\xi \\
&= \int_\xi\int_x \varphi(x)\overline{\psi}(x,\xi)\,\mathrm{d}_c x\,\mathrm{d}_c\xi \\
&= \int_x \varphi(x)\,c^2\!\int_\xi \overline{\psi}(x,\xi)\,\mathrm{d}\xi\,\mathrm{d}x \\
&= \int_x \varphi(x)\delta(x)\,\mathrm{d}x \quad \text{(Dirac delta)} \\
&= \varphi(0).
\end{aligned}$$

But this calculation poses several problems. The double integral

$$\iint_{(x,\xi)\in\mathbb{R}^{2n}} \varphi(x)\overline{\psi}(x,\xi)\,\mathrm{d}\mu(x,\xi)$$

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

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

diverges classically. The ideas that after multiplying it by c^2 for some reason, it 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.

7. The dilated 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,

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

is its own Fourier transform under our rescaled measure. (Here is a benefit of the rescaling, because this normalized Gaussian has derivative $\gamma'(x) = -x\gamma(x)$, as compared to some scalar multiple of $-x\gamma(x)$.) It follows easily that so is the n-dimensional Gaussian,

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

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

$$f_t : \mathbb{R}^n \longrightarrow \mathbb{C}, \qquad f_t(x) = f(tx).$$

Compute that

$$\begin{aligned}\mathcal{F}f_t(\xi) &= \int_{x\in\mathbb{R}^n} f_t(x)\overline{\psi}(x,\xi)\,\mathrm{d}_c x \\ &= \int_{x\in\mathbb{R}^n} f(tx)\overline{\psi}(tx,\xi/t)\,\mathrm{d}_c(tx)/t^n = t^{-n}(\mathcal{F}f)_{t^{-1}}(\xi).\end{aligned}$$

That is, $\mathcal{F}f_t = t^{-n}(\mathcal{F}f)_{t^{-1}}$. Now consider a dilated Gaussian function for any $t > 0$,

$$g_{\sqrt{t}} : \mathbb{R}^n \longrightarrow \mathbb{C}, \qquad g_{\sqrt{t}}(x) = e^{-t|x|^2/2}.$$

This is very wide for small positive t, taking values close to 1 near 0. Its Fourier transform is $\mathcal{F}g_{\sqrt{t}} = t^{-n/2}g_{1/\sqrt{t}}$, or, introducing a new function-name,

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

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

(1) $\phi_t(x) \geq 0$ for all x,
(2) $\int_x \phi_t(x)\,\mathrm{d}_c x = 1$,
(3) Given any $\varepsilon > 0$ and any $\delta > 0$, we have $\int_{|x|\geq\delta} \phi_t(x)\,\mathrm{d}_c x < \varepsilon$ if t is close enough to 0. That is, ϕ_t becomes a tall narrow pulse as $t \to 0$.

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 $t \to 0$ 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 $t \to 0$ 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.

8. Analysis proof of Fourier inversion

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

$$\iint_{(x,\xi)\in\mathbb{R}^{2n}} \varphi(x)\overline{\psi}(x,\xi)\,\mathrm{d}_c\mu(x,\xi),$$

where $\mathrm{d}_c\mu(x,\xi) = \mathrm{d}_c x\,\mathrm{d}_c\xi$. Consider instead a similar double integral but with its integrand dampened by the t-dilated Gaussian function $g_{\sqrt{t}}$,

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

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

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

and the other is

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

For the first integral, we have by the Dominated Convergence Theorem that because $g_{\sqrt{t}} \to 1$ pointwise as $t \to 0$ (or, without the Dominated Convergence Theorem, because $g_{\sqrt{t}} \to 1$ uniformly on compacta as $t \to 0$ and $\mathcal{F}\varphi$ is bounded and absolutely integrable),

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

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*,

$$B_t = \int_x \varphi(x)\phi_t(x)\,\mathrm{d}_c x \xrightarrow{t\to 0} \varphi(0).$$

But $A_t = B_t$ for all $t > 0$, and so the limits of A_t and B_t as $t \to 0$ 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.

9. Calculus proof of Fourier inversion

A second argument, which I learned from the book by Richards and Youn, is technically easier and very pretty. The argument relies on one further reduction, and calculus, and the fact that the Fourier transform converts multiplication to differentiation.

As before, to prove that

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

we need only prove that

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

Now we reduce the problem further. Let φ be any Schwartz function, and again let g be the Gaussian function, its own Fourier transform and inverse Fourier transform, with $g(0) = 1$. From these properties of g,

$$(\mathcal{F}^{-1}\mathcal{F}\varphi)(0) - \varphi(0) = \mathcal{F}^{-1}\mathcal{F}(\varphi - \varphi(0)g)(0).$$

We want to show that the left side of the previous display is 0, for which it suffices to show that the right side is 0. Because $\varphi - \varphi(0)g$ vanishes at 0, the problem is thus reduced to proving that

$$(\mathcal{F}^{-1}\mathcal{F}f)(0) = 0 \quad \text{for Schwartz functions } f \text{ that vanish at 0.}$$

Because Schwartz functions are not preserved under vertical translation, this reduction to the case $f(0) = 0$ requires a little more artfulness than merely subtracting off $\varphi(0)$ from φ as can be done for functions on the compact domain $\mathbb{R}^n/\mathbb{Z}^n$, but the idea is essentially the same.

For any Schwartz function f such that $f(0) = 0$, define for $j = 1, \dots, n$,

$$\varphi_j : \mathbb{R}^n \longrightarrow \mathbb{C}, \qquad \varphi_j(x) = \int_{t=0}^{1} \mathrm{D}_j f(xt)\,\mathrm{d}t.$$

Because $f(0) = 0$, the fundamental theorem of calculus gives

$$f(x) = f(x) - f(0) = f(xt)\Big|_{t=0}^{1} = \int_{t=0}^{1} \frac{\mathrm{d}}{\mathrm{d}t} f(xt)\,\mathrm{d}t,$$

and the chain rule gives

$$\frac{\mathrm{d}}{\mathrm{d}t} f(xt) = \sum_{j=1}^{n} \mathrm{D}_j f(xt) x_j,$$

so we have $f(x) = \sum_{j=1}^{n} x_j \int_{t=0}^{1} \mathrm{D}_j f(xt)\,\mathrm{d}t$, and by the definition of the functions φ_j this is

$$f(x) = \sum_{j=1}^{n} x_j \varphi_j(x).$$

Each function φ_j is smooth because f is smooth. This is just differentiation under the integral sign: for any $L = (\ell_1, \dots, \ell_m)$ with each ℓ_i in $\{1, \dots, n\}$,

$$\mathrm{D}_L \varphi_j(x) = \int_{t=0}^{1} \mathrm{D}_{j,L} f(xt)\,\mathrm{d}t.$$

Also we could show that φ_j decays as x grows because $\mathrm{D}_j f$ decays, but we do not need this. More importantly, for $n > 1$ the φ_j need not be Schwartz even though f is Schwartz. As a simple example of a smooth function f all of whose derivatives decay without any of them decaying rapidly, let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be a smooth blend of $1/x$ for $|x|$ large and a constant function for $|x|$ small.

Now take $n = 1$. In this case, φ is Schwartz because f is Schwartz, as follows. Because $f(x) = x\varphi(x)$, the quotient $f(x)/x$ extends smoothly to $f'(0) = \varphi(0)$ at $x = 0$, and this quotient and its derivatives clearly inherit rapid decay from $f(x)$ and its derivatives. With φ known to be Schwartz, the Fourier inversion argument for $n = 1$ begins with the result, left to the reader as a quick calculation, that the Fourier transform converts multiplication to differentiation,

$$(\mathcal{F}f)(\xi) = (\mathcal{F}x\varphi)(\xi) = i(\mathcal{F}\varphi)'(\xi).$$

Thus, as desired,

$$(\mathcal{F}^{-1}\mathcal{F}f)(0) = \int_{\xi \in \mathbb{R}} \mathcal{F}f(\xi)\,\mathrm{d}_c\xi = i \int_{\xi \in \mathbb{R}} (\mathcal{F}\varphi)'(\xi)\,\mathrm{d}_c\xi = i(\mathcal{F}\varphi)\Big|_{-\infty}^{\infty} = 0.$$

For $n > 1$, this Fourier inversion proof requires one more idea. In the decomposition

$$f(x) = \sum_{j=1}^{n} x_j \varphi_j(x),$$

each $\varphi_j(x)$ is smooth but it and its derivatives needn't decay rapidly. In a second decomposition,

$$f(x) = \sum_{j=1}^{n} x_j \cdot x_j f(x)/|x|^2,$$

each $x_j f(x)/|x|^2$ needn't be smooth at 0 but it and its derivatives inherit rapid decay from $f(x)$. Proving Fourier inversion blends these two decompositions in a way that retains the asset of each and suppresses its deficiency. Let $\chi : \mathbb{R}^n \longrightarrow \mathbb{C}$ be a smooth function whose support is compact and contains 0 as an interior point. Define modified versions of the functions φ_j for $j = 1, \ldots, n$ as follows,

$$\tilde{\varphi}_j(x) = \chi(x)\varphi_j(x) + (1 - \chi(x))x_j f(x)/|x|^2.$$

These functions are Schwartz, and we have the same decomposition of f as with the original functions φ_j,

$$\begin{aligned} f(x) &= \chi(x)f(x) + (1 - \chi(x))f(x) \\ &= \chi(x)\sum_{j=1}^{n} x_j\varphi_j(x) + (1 - \chi(x))\sum_{j=1}^{n} x_j^2 f(x)/|x|^2 \\ &= \sum_{j=1}^{n} x_j \left(\chi(x)\varphi_j(x) + (1 - \chi(x))x_j f(x)/|x|^2\right) \\ &= \sum_{j=1}^{n} x_j \tilde{\varphi}_j(x). \end{aligned}$$

From here the Fourier inversion argument for $n > 1$ is the same as for $n = 1$,

$$(\mathcal{F}f)(\xi) = (\mathcal{F}\sum_{j=1}^{n} x_j\tilde{\varphi}_j)(\xi) = i\sum_{j=1}^{n} \mathrm{D}_j(\mathcal{F}\tilde{\varphi}_j)(\xi),$$

and thus

$$(\mathcal{F}^{-1}\mathcal{F}f)(0) = \int_{\xi\in\mathbb{R}^n} \mathcal{F}f(\xi)\,\mathrm{d}_c\xi = i\sum_{j=1}^{n}\int_{\xi\in\mathbb{R}^n} \mathrm{D}_j(\mathcal{F}\tilde{\varphi})(\xi)\,\mathrm{d}_c\xi,$$

and each integral is 0 by integrating first in the jth direction.

10. 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 with viewing $\mathcal{S}$ as a subspace of $\mathcal{S}^*$, because if both f and φ are Schwartz functions then, as already noted above,

$$\langle \mathcal{F}f, \varphi\rangle = \int_{\mathbb{R}^n} \mathcal{F}f \cdot \varphi = \int_{\mathbb{R}^n} f \cdot \mathcal{F}\varphi = \langle f, \mathcal{F}\varphi\rangle.$$

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

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

11. 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,

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

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.

12. An example

We now freely view functions of moderate growth as tempered distributions, so that Fourier inversion applies to them. Consider any $x \in \mathbb{R}_{>0}$, $\xi \in \mathbb{R}$, and $s \in \mathbb{R}_{>1}$. We show that

$$\int_{y=-\infty}^{\infty} \frac{e^{i\xi y}}{(x+iy)^s}\, \mathrm{d}y = \begin{cases} \frac{2\pi}{\Gamma(s)} e^{-x\xi}\xi^{s-1} & \text{if } \xi > 0, \\ 0 & \text{if } \xi \le 0. \end{cases}$$

The idea is to recognize the integral as the inverse Fourier transform $(\mathcal{F}^{-1}f_x)(\xi)$ where $f_x(y) = (2\pi)^{1/2}/(x+iy)^s$. So we are done by Fourier inversion if $f_x(y)$ is in turn the Fourier transform of the function on the right side of the previous display.

To obtain $f_x(y)$ as a Fourier transform, start from the gamma function, replacing ξ by $x\xi$ in the integral to get a variant expression of the gamma integral that incorporates x,

$$\Gamma(s) = \int_{\xi=0}^{\infty} e^{-\xi}\xi^s \frac{\mathrm{d}\xi}{\xi} = x^s \int_{\xi=0}^{\infty} e^{-x\xi}\xi^s \frac{\mathrm{d}\xi}{\xi}.$$

The uniqueness theorem from complex analysis says that this formula extends from the open half-line of positive x-values to the open half-plane of complex numbers $x + iy$ with x positive. That is, for any $y \in \mathbb{R}$,

$$\Gamma(s) = (x+iy)^s \int_{\xi=0}^{\infty} e^{-(x+iy)\xi}\xi^s \frac{\mathrm{d}\xi}{\xi}.$$

This is

$$\Gamma(s) = (x+iy)^s \int_{\xi=0}^{\infty} e^{-iy\xi} \cdot e^{-x\xi}\xi^{s-1}\,\mathrm{d}\xi.$$

That is, if we introduce a function of a variable $\xi \in \mathbb{R}$,

$$\varphi_x(\xi) = \begin{cases} e^{-x\xi}\xi^{s-1} & \text{if } \xi > 0, \\ 0 & \text{if } \xi \leq 0, \end{cases}$$

then, recalling the definition $f_x(y) = (2\pi)^{1/2}/(x+iy)^s$, we have

$$f_x(y) = \frac{2\pi}{\Gamma(s)}(\mathcal{F}\varphi_x)(y), \quad y \in \mathbb{R}.$$

This gives the asserted value of the integral at the beginning of the section.

An n-variable version of this integral goes back at least to Siegel, requiring essentially nothing more than the method here and a standard device called *completing the square* for a positive definite quadratic form. See the online writeup by Paul Garrett,

`http://www.math.umn.edu/~garrett/m/v/siegel_integral.pdf`,

my source for the one-variable treatment here.

13. 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,

$$c^2 \int_{\xi \in \mathbb{R}^n} e^{ix\cdot\xi}\,\mathrm{d}\xi = \delta(x) \quad \text{for all } x \in \mathbb{R}^n.$$

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,

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

But justifying this calculation is not trivial.