The real Fourier transform

Bill Casselman
University of British Columbia
cass@math.ubc.ca

This essay is a brief introduction to the Fourier transform on $\mathbb{R}^n$. Special features of the transform in dimension one are covered elsewhere, as is the detailed relationship with the multiplicative group.

Contents

Notation. Let $\|x\| = \sqrt{x \bullet x}$ be the Euclidean norm on $\mathbb{R}^m$, $|x| = \sup |x_i|$.

1. Introduction

The theory of Fourier series tells us that if $f(x)$ is a smooth function on $\mathbb{R}^n$ invariant under translations in $\mathbb{Z}^n$ then

$$f(x) = \sum\nolimits_m f_m \, e^{2\pi i m \bullet x} \, ,$$

where

$$f_m = \int_{\mathbb{R}^n/\mathbb{Z}^n} f(x) e^{-2\pi i m \bullet x} \, dx = \int_{|x| \le 1/2} f(x) e^{-2\pi i m \bullet x} .$$

The Fourier coefficients f_m are of rapid decrease with respect to the variable $m \in \mathbb{Z}^n$, so the series certainly converges absolutely and uniformly.

We can generalize this to deal with functions of arbitrary period. If $f(x)$ is invariant under translations in $T\mathbb{Z}^n$ then $\varphi(x) = f(xT)$ is invariant under $\mathbb{Z}^n$, and

$$\varphi(x) = \sum\nolimits_m \varphi_n \, e^{2\pi i m \bullet x}$$

where

$$\begin{aligned} \varphi_n &= \int_{|x| \le 1/2} \varphi(x) e^{-2\pi i m \bullet x} \, dx \\ &= \frac{1}{T^n} \int_{|y| \le T/2} f(y) e^{-2\pi i m \bullet y/T} \, dy \quad (y = xT) \, . \end{aligned}$$

This implies that

$$f(x) = \sum\nolimits_{\lambda \in \mathbb{Z}^n/T} f_\lambda e^{2\pi i \lambda \bullet x} \quad \text{where} \quad f_\lambda = \frac{1}{T^n} \int_{|y| \le T/2} f(y) e^{-2\pi i \lambda \bullet y} \, dy \, .$$

What happens as $T \to \infty$?

Define the **Schwartz space** $\mathcal{S}(\mathbb{R}^n)$ to be the space of all smooth functions f on $\mathbb{R}^n$ such that all the semi-norms

$$\|f\|_{k,m} = \sup_x \big(1 + \|x\|\big)^k \left| \frac{\partial^m f}{\partial x^m} \right|$$

are bounded. In other words, f and all its derivatives decrease more rapidly at infinity than any negative power of $\|x\|$. This definition is independent of the coordinate system. An important class of examples is made up of the functions $P(x)e^{-Q(x)}$, where $P(x)$ is a polynomial and $Q(x)$ is a positive definite quadratic form.

If f is in $\mathcal{S}(\mathbb{R}^n)$ the series

(1.1)
$$[\Theta_T f](x) = \sum\nolimits_{\mathbb{Z}^n} f(x + nT)$$

will be a smooth function invariant with respect to $T\mathbb{Z}^n$ to which we can apply the results just recalled:

$$\begin{aligned} [\Theta_T f](x) &= \sum_{\lambda \in \mathbb{Z}^m/T} f_{T,\lambda} e^{2\pi i \lambda \bullet x} \\ \text{with} \quad f_{T,\lambda} &= \frac{1}{T^n} \int_{|y| \le T/2} [\Theta_T f](y) e^{-2\pi i \lambda \bullet y}\, dy \\ &= \int_{\mathbb{R}^n} f(y) e^{-2\pi i \lambda y}\, dy\,. \end{aligned}$$

As $T \to \infty$ the value of $\Theta_T f$ at any x converges to $f(x)$, while the series expansion converges as a Riemann sum with interval $1/T$ to the integral

$$f(x) = \int_{\mathbb{R}^n} \widehat{f}(y) e^{2\pi i x \bullet y}\, dy$$

with

$$\widehat{f}(y) = \int_{\mathbb{R}^m} f(y) e^{-2\pi i x \bullet y}\, dy\,.$$

For any f in $\mathcal{S}(\mathbb{R}^m)$ the second formula defines its **Fourier transform**, and the first is the Fourier inversion formula. In the next sections we'll derive this formula, and in fact a much more general one, in a different way.

2. The Fourier transform

If f is in $\mathrm{L}^1(\mathbb{R}^n)$, its Fourier transform is

$$\widehat{f}(y) = \int_{\mathbb{R}^n} f(x) e^{-2\pi i x \bullet y}\, dx\,.$$

As an example:

2.1. Proposition. *If $f(x) = e^{-\pi\|x\|^2}$ then it is its own Fourier transform.*

Proof. We have

$$\begin{aligned} \widehat{f}(y) &= \int_{\mathbb{R}^n} e^{-\pi\|x\|^2} e^{-2\pi i x \bullet y}\, dx \\ &= \int_{\mathbb{R}^n} e^{-\pi\|x\|^2 - 2\pi i x \bullet y}\, dx \\ &= e^{-\pi\|y\|^2} \int_{\mathbb{R}^n} e^{-\pi(\|x\|^2 - 2ix \bullet y - \|y\|^2)}\, dx \\ &= e^{-\pi\|y\|^2} \int_{\mathbb{R}^n} e^{-\pi\|x - iy\|^2}\, dx \\ &= e^{-\pi\|y\|^2} \int_{\mathbb{R}^n} e^{-\pi\|x\|^2}\, dx \\ &= e^{-\pi\|y\|^2}\,. \end{aligned}$$

One of these steps involves moving an integration contour in $\mathbb{C}^n$.

For $c \neq 0$ let

$$\mu_c f(x) = f(x/c)\,.$$

Also let

$$\lambda_a f(x) = f(x-a)\,.$$

The following is an elementary exercise:

2.2. Lemma. *Suppose f, g to be in $\mathrm{L}^1(\mathbb{R}^n)$. Then*

(a) *the Fourier transform $\widehat{f}$ is bounded and continuous;*
(b) *the Fourier transform of the convolution $f * g$ is the product $\widehat{f}\,\widehat{g}$;*
(c) *the Fourier transform of $e^{2\pi i c x} f(x)$ is $\widehat{f}(y-c)$;*
(d)

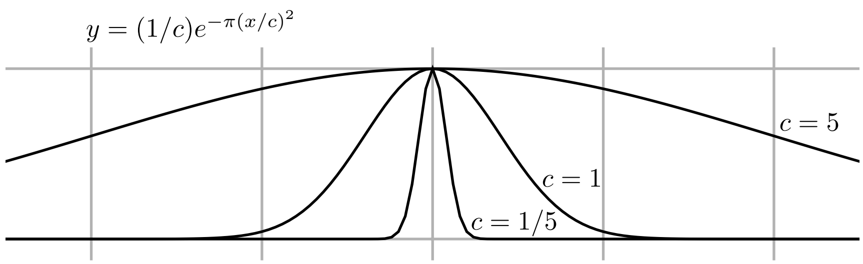
$$\int_{\mathbb{R}^n} f(x)\widehat{g}(x)\,dx = \int_{\mathbb{R}^n} \widehat{f}(x) g(x)\,dx\,.$$

(e) *for $c \neq 0$*

$$\widehat{\mu_c f} = c\,\mu_{1/c}\widehat{f}\,;$$

(f) $\widehat{\lambda_a f} = e^{-2\pi i a y}\widehat{f}(y)$.

For example:



2.3. Theorem. *Suppose that f and $\widehat{f}$ are both in $\mathrm{L}^1(\mathbb{R}^n)$, Then*

$$f(x) = \int_{\mathbb{R}^n} \widehat{f}(y) e^{2\pi i x y}\,dy\,.$$

Proof. Let

$$\varphi(x) = e^{-\pi\|x\|^2}\,,$$

so that $\widehat{\varphi} = \varphi$. Then (d) of Lemma 2.2 tells us that for all c

$$\int_{\mathbb{R}^n} \varphi(x/c)\cdot\widehat{f}(x)\,dx = \int_{\mathbb{R}^n} c\,\varphi(cx)\cdot f(x)\,dx\,.$$

As $c \to \infty$ the function $\varphi(x/c)$ approaches the constant function 1, and the first integral approaches the integral

$$\int_{\mathbb{R}^n} \widehat{f}(x)\,dx\,.$$

The function $c\varphi(cx)$ gets narrower as $c \to \infty$, and the area under its graph remains equal to 1. The second integral therefore approaches $f(0)$.

3. The Schwartz space

Any function f in $\mathcal{S}(\mathbb{R}^n)$ is in $\mathrm{L}^1(\mathbb{R}^n)$, so its Fourier transform

$$\widehat{f}(y) = \int_{\mathbb{R}^n} f(x) e^{-2\pi x \bullet y}\, dx$$

is well defined. More precisely, for $k \gg 0$

$$\begin{aligned} \big|\widehat{f}(y)\big| &\le \int_{\mathbb{R}^n} \big|f(x)\big|\, dx \\ &= \int_{\mathbb{R}^n} \big(1 + \|x\|\big)^k \big|f(x)\big| \, \frac{1}{\big(1+\|x\|\big)^k}\, dx \\ &\le \sup_{\mathbb{R}^n} \big|\big(1+\|x\|\big)^k f(x)\big| \int_{\mathbb{R}^n} \frac{1}{\big(1+\|x\|\big)^k}\, dx\,. \end{aligned}$$

3.1. Lemma. *Suppose f to be in $\mathcal{S}(\mathbb{R}^n)$. The Fourier transform of $\partial f/\partial x_m$ is $2\pi i y_m \widehat{f}(y)$, and $\partial \widehat{f}/\partial y_m$ is the Fourier transform of $-2\pi i x_m f(x)$.*

This is the whole point of the Fourier transform—*it transforms problems in analysis (differentiation) into problems of algebra (multiplication).*

Proof. The first claim follows from integration by parts, the first is straightforward.

3.2. Corollary. *The Fourier transform takes $\mathcal{S}(\mathbb{R}^n)$ to itself.*

In other words, because f is smooth, its Fourier transform is rapidly decreasing, and because f is rapidly decreasing, its Fourier transform is smooth.

I now offer a third proof of Fourier inversion.

3.3. Theorem. *The map $f \mapsto \widehat{f}$ is an isomorphism of $\mathcal{S}(\mathbb{R}^n)$ with $\mathcal{S}(\mathbb{R}^n)$.*

Explicitly, I claim that the inverse map takes $F(y)$ to

$$F^\vee(x) = \int_{\mathbb{R}^n} F(y) e^{2\pi i x \bullet y}\, dy$$

In other words, we want to prove that

$$f(x) = \int_{\mathbb{R}^n} \widehat{f}(y)\, e^{2\pi i x \bullet y}\, dy\,.$$

Proof. I do this in several steps.

Step 1. We know from Proposition 2.1 that the inversion formula is valid for $e^{-\pi\|x\|^2}$, which is clearly in the Schwartz space.

Step 2. In this step I reduce the equation to the special case $x = 0$. If $\varphi(x) = f(x+a)$ then $\widehat{\varphi}(y) = e^{2\pi i a \bullet y} \widehat{f}(y)$. Therefore Fourier inversion holds for f at a if and only if it holds for φ at 0.

Step 3. It remains to show that

$$f(0) = \int_{\mathbb{R}} \widehat{f}(y)\, dy$$

for all f in $\mathcal{S}(\mathbb{R}^n)$. But $e^{-\pi\|x\|^2}$ lies in the Schwartz space, Fourier inversion is valid for it, and $f(x) - f(0)e^{-\pi\|x\|^2}$ vanishes at 0. So it suffices to show that the inversion formula is valid for $x = 0$ when $f(0) = 0$.

Step 4. It must be shown that if $f(0) = 0$ then

$$\int_{\mathbb{R}} \widehat{f}(y)\,dy = 0\,.$$

This will follow in a moment from the observation that if f lies in $\mathcal{S}(\mathbb{R}^n)$ then $f(0) = 0$ if and only if there exist functions f_k in $\mathcal{S}(\mathbb{R}^n)$ with

$$f = \sum_k x_k f_k\,.$$

But for later use, I'll prove something slightly stronger. For $0 < R \le \infty$ let

$$C_R^p = \{x \in \mathbb{R}^p \mid |x| \le R\}\,.$$

Note that the cube C_R^p is a product of p intervals (possibly infinite) of dimension 1.

To go along with this, I want to introduce some new notation. Let

$$\mathcal{S}(C_R^p) = \begin{cases} \mathcal{S}(\mathbb{R}^p) & \text{if } R = \infty \\ C_c^\infty(C_R^p) & \text{otherwise.} \end{cases}$$

In addition, consider $\mathbb{R}^p$ embedded in the first p dimensions of $\mathbb{R}^n$, and let $\overline{\mathbb{R}}^{n-p}$ be its orthogonal complement in $\mathbb{R}^n$, so that $\mathbb{R}^n = \mathbb{R}^p \times \overline{\mathbb{R}}^{n-p}$. Similiarly, let $C_R^n = C_R^p \times \overline{C}_R^{n-p}$.

The basic point of the argument to come can be seen already in dimension one—it a standard result in calculus that if f is in $\mathcal{S}(C_R^1)$ then

$$f(x) - f(0) = x \int_0^1 f'(tx)\,dt\,,$$

so that if $f(0) = 0$ then $f(x)/x$ is in $\mathcal{S}(C_R^1)$. Something like this holds also in any dimension: if $f(x)$ lies in $\mathcal{S}(\overline{C}_R^p)$ and $f|R^{p-1}) = 0$, then f/x_p also belongs to $\mathcal{S}(C_R^p)$. This is because

$$f(x_*, x) = x_p \cdot \int_0^1 \left(\frac{\partial f}{\partial x_p}\right)(x_*, tx)\,dt\,.$$

We can now prove:

3.4. Lemma. *If f is in $\mathcal{S}(C_R^n)$ and $f(0) = 0$ then*

$$f = \sum\nolimits_1^n x_i f_i$$

for functions f_i in $\mathcal{S}(C_R^n)$.

Proof (sketch). We do it by induction on dimension. We can start out in dimension one, to find a function φ_1 in $\mathcal{S}(C_R^1)$ such that

$$f(x_1, 0) = x_1 \varphi(x_1, 0)\,.$$

We may pull this back to all of $\mathbb{R}_n$ and multiply it by a function ψ on $\overline{C}_R^{n-1}$ to get a function $x_1 f_1$ in $\mathcal{S}(C_R^n)$ whose restriction to $\mathbb{R}^1$ agrees with f. So now $f - x_1 f_1$ vanishes on $\mathbb{R}^1$.

Continuing, we can find a function φ_2 in $\mathcal{S}(C_R^2)$ such that $f - x_1 f_1 = x_2 \varphi_2$. We pull this back to $\mathbb{R}^n$ and multiply by a function in $\mathcal{S}(\overline{C}_R^{n-2})$ to get a function f_2 in $\mathcal{S}(C_R^n)$ such that

$$f - x_1 f_1 - x_2 f_2$$

vanishes on $\mathbb{R}^2$. Etc.

Step 5. Why does the Lemma imply the integral formula we want? It suffices now to see that the Fourier transform of each function $x_m f_m(x)$ is a partial derivative whose integral vanishes.

The Fourier inversion theorem may be restated:

3.5. Corollary. *If $\widehat{f}(y) = \varphi(y)$ then $\widehat{\varphi}(x) = f(-x)$.*

4. The Plancherel theorem

We shall see here that the Fourier transform extends to a map from $\mathrm{L}^2(\mathbb{R}^n)$ to itself, and induces an isometry of Hilbert spaces.

As a corollary of (d) in Lemma 2.2 and the inversion theorem:

4.1. Theorem. *For f, g in $\mathcal{S}(\mathbb{R}^n)$*

$$\int_{\mathbb{R}^n} f(x)\overline{g(x)}\,dx = \int_{\mathbb{R}^n} \widehat{f}(x)\overline{\widehat{g}(y)}\,dy\,.$$

4.2. Corollary. (The Plancherel Theorem) *The Fourier transform taking $\mathcal{S}(\mathbb{R}^n)$ to itself extends to an isometry of $\mathrm{L}^2(\mathbb{R}^n)$ with itself.*

This is a bit subtle, because the integral defining the Fourier transform does not make obvious sense for f in $\mathrm{L}^2(\mathbb{R}^n)$. What follows from Theorem 4.1 is that the definition of the Fourier transform on a function f in $\mathrm{L}^2(\mathbb{R}^n)$ is as a limit in a Hilbert space:

$$\widehat{f} = \lim_{R\to\infty} \widehat{f}_R$$

where

$$f_R = \int_{\|x\|\le R} f(x)e^{-2\pi i x\bullet y}\,dx\,.$$

5. Schwartz' Paley-Wiener theorem

There are many theorems characterizing the support of a function on $\mathbb{R}^n$ by analycity properties of its Fourier transform. Questions of this kind were first raised and answered in joint work of R. E. A. C. Paley and Norbert Wiener, particularly in the book [Paley-Wiener:1934], and a certain category of such results (those involving Fourier transforms that are analytic functions) are often called **Paley-Wiener theorems**.

Laurent Schwartz took up such questions in his work on what he calls **Laplace transforms** (see Chaptar VIII in [Schwartz:1952].) One of the most useful of these characterizations is for $C_c^\infty(\mathbb{R}^n)$. If f is integrable and has support in the compact subset Ω of $\mathbb{R}^n$, then the integral

$$\widehat{f}(s) = \int_{\mathbb{R}^n} f(x)e^{-2\pi i s\bullet x}\,dx = \int_\Omega f(x)e^{-2\pi i s\bullet x}\,dx$$

is absolutely and uniformly convergent for all s in $\mathbb{C}^n$.

For $R>0$ let B_R be the Euclidean ball $\|x\|\le R$.

5.1. Proposition. *If f is a smooth function on $\mathbb{R}^n$ with support in B_R, then its Fourier transform $F(s)$ satisfies these conditions:*

PW(a) *$F(s)$ is holomorphic on $\mathbb{C}^n$;*
PW(b) *for all $N>0$*

$$\big|F(s)\big| \ll_{F,N} \frac{e^{2\pi R|\mathrm{IM}(s)|}}{(1+\|s\|)^N}$$

Proof. Since all derivatives of f still satisfy the conditions of the Proposition, this is straightforward. ▯

5.2. Proposition. *Conversely, if $F(s)$ satisfies the conditions PW of the previous proposition, then it is the Fourier transform of a function in $C_c^\infty(\mathbb{R}^n)$ with support in B_R.*

Proof. Just to simplify things slightly, I'll prove this only for $n=1$. It must be shown that the inverse transform of F is well defined, that it is smooth, and that it has support in $[-R,R]$. So suppose F to satisfy conditions PW. I first claim that $t\mapsto F(\sigma+it)$ lies in $\mathcal{S}(\mathbb{R})$ for each σ. It has to be verified that its derivatives are rapidly decreasing as $|t| \to \infty$. But this follows from the Cauchy integral formula expressing $F^{(n)}(s)$ as a path integral around small circles surrounding s.

The inverse Fourier transform of F is the integral

$$f(x) = \int_{-\infty}^{\infty} F(s)e^{2\pi isx}\, ds\,.$$

The Fourier inversion theorem for Schwartz functions implies that this lies in $\mathcal{S}(\mathbb{R})$, and it must be shown that $f(x) = 0$ if $|x| > R$. By assumption, there exists $C_2 > 0$ such that

$$\bigl|F(\sigma + it)\bigr| \le \frac{C_2\, e^{2\pi R|t|}}{1 + |t|^2}$$

for all real σ, t. We may shift the integration path for $f(x)$ to get

$$f(x) = \int_{-\infty+it}^{\infty+it} F(s)e^{2\pi ixs}\, ds\,.$$

But then

$$\begin{aligned}\bigl|f(x)\bigr| &\le C_2 \int_{-\infty+it}^{\infty+it} \frac{e^{2\pi(R-x)t}}{1+|s|^2}\, ds\\ \bigl|f(x)\bigr| &\le C_2 e^{2\pi(R-x)t} \int_{-\infty+it}^{\infty+it} \frac{1}{1+|s|^2}\, ds\end{aligned}$$

But if $x > R$ then letting $t \to \infty$ we see that $f(x)$ must be 0. If we let $t \to -\infty$ we deduce that $f(x) = 0$ for $x < -R$.

6. Tempered distributions

On the vector space $\mathcal{S}(\mathbb{R}^n)$ define the semi-norms

$$\|f\|_{N,n} = \sup_{x\in\mathbb{R}^n} \bigl(1 + \|x\|\bigr)^N \bigl|f^{(n)}(x)\bigr|\,.$$

These make $\mathcal{S}(\mathbb{R}^n)$ into a Fréchet space. A **tempered distribution** is a linear function

$$\Phi\colon\ \mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$$

that's continuous in this topology. In other words, for Φ to be a tempered distribution we require that

$$\bigl|\langle \Phi, f\rangle\bigr| \ll_{N,n} \|f\|_{N,n}$$

for some finite set of N, n. For example, if $F(x)$ is a locally integrable function on $\mathbb{R}^n$ of moderate growth in the sense that

$$\frac{\bigl|F(x)\bigr|}{(1 + \|x\|)^N} \in \mathrm{L}^1(\mathbb{R}^n)$$

for some N, then

$$f \longmapsto \int_{\mathbb{R}^n} \varphi(x)f(x)\, dx$$

defines a tempered distribution.

In particular, if F is an integrable function on $\mathbb{R}^n$, then integration against F defines a tempered distribution. So we can construct lots of tempered distributions from functions with various growth conditions, but as we shall see in the next section there are others of a more exotic kind.

DERIVATIVES. One of the wonderful facts about distributions is that one can define the derivatives of an arbitrary tempered distribution, and it will be is another tempered distribution. Thus every L^1 function has a derivative in this sense. *O brave new world!*

If f is in the Schwartz space and φ is a smooth function of moderate growth whose derivatives are of moderate growth, then integration by parts gives us

$$\langle \varphi, \partial f/\partial x_m \rangle = \int_{\mathbb{R}^n} \varphi(x) \frac{\partial f}{\partial x_m}\, dx = -\int_{\mathbb{R}^n} \frac{\partial \varphi}{\partial x_m} f(x)\, dx = -\langle \partial\varphi/\partial x_m, f \rangle\,.$$

This suggests defining the partial derivative of any tempered distribution Φ by the formula

$$\langle \partial\Phi/\partial x_m, f \rangle = -\langle \Phi, \partial f/\partial x_m \rangle\,.$$

This definition can be extended to any differential operator with constant coefficients.

FOURIER TRANSFORM. According to (d) of Lemma 2.2, if f and g are both in $\mathcal{S}(\mathbb{R}^n)$ then $\langle f, \widehat{g} \rangle = \langle \widehat{f}, g \rangle$, so it is consistent to define the Fourier transform of an arbitrary tempered distribution by duality:

$$\langle \widehat{\Phi}, f \rangle = \langle \Phi, \widehat{f} \rangle\,.$$

Since the Fourier transform is an isomorphism of $\mathcal{S}(\mathbb{R}^n)$ with itself, it induces also an isomorphism of the space of tempered distributions with itself.

We shall see examples in the next section, but we can verify now that in other cases this definition of Fourier transform is consistent with the obvious one:

6.1. Proposition. *If F lies in $\mathrm{L}^1(\mathbb{R}^n)$ then $\widehat{F}$ is the bounded continuous function*

$$\widehat{F}(y) = \int_{\mathbb{R}^n} F(x) e^{-2\pi i x \bullet y}\, dx$$

on $\mathbb{R}^n$.

The Riemann-Lebesgue Lemma asserts that $\widehat{F}(y) \to 0$ as $\|y\| \to \infty$, but I won't prove or need that here.

Proof. The integral

$$G(y) = \int_{\mathbb{R}^n} F(x) e^{-2\pi i x \bullet y}\, dx$$

is a bounded function of y. It remains to show that integration against G defines the distribution $\widehat{F}$. Suppose f to lie in $\mathcal{S}(\mathbb{R})$. By definition

$$\langle \widehat{F}, f \rangle = \langle F, \widehat{f} \rangle$$

and there are no obstacles to expressing this a double integral

$$\langle F, \widehat{f} \rangle = \int_{\mathbb{R}^n} F(y) \widehat{f}(y)\, dy = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} F(y) f(x) e^{-2\pi i x \bullet y}\, dx\, dy = \langle G, f \rangle\,.$$

BASIC OPERATIONS. Certain operations and formulas involving the Fourier transform of functions in $\mathcal{S}(\mathbb{R}^n)$ are valid also for tempered distributions. (a) If $F(x)$ is a smooth function on $\mathbb{R}^n$ all of whose derivatives are of moderate growth, and f lies in $\mathcal{S}(\mathbb{R}^n)$, then $F(x)f(x)$ is again in $\mathcal{S}(\mathbb{R}^n)$. If Φ is a tempered distribution on $\mathbb{R}^n$ then the product $F \cdot \Phi$ is defined by the formula

$$\langle F \cdot \Phi, f \rangle = \langle \Phi, F \cdot f \rangle\,.$$

(b) If g is in $\mathrm{GL}_n(\mathbb{R})$ then the transformation $\mu_g f$ of f by g is defined by

$$[\mu_g f](x) = f(g^{-1}x)\,.$$

If Φ is a tempered distribution then the scale of Φ by g is defined by

$$\langle \mu_g \Phi, f \rangle = \langle \Phi, \mu_{g^{-1}} f \rangle .$$

(c) The translation $\lambda_v f$ of f by v in $\mathbb{R}^n$ is defined by

$$[\lambda_v f](x) = f(x - v) .$$

The translation of Φ by v is defined by

$$\langle \lambda_v \Phi, f \rangle = \langle \Phi, \lambda_{-v} f \rangle .$$

In comparing these to the analogous properties of Schwartz functions, you must keep in mind that it is not really a function $f(x)$ that defines a distribution, but the measure $f(x)dx$.

How do these interact with the Fourier transformation?

6.2. Proposition. *Suppose Φ to be a tempered distribution.*

(a) The Fourier transformation of $x_k \cdot \Phi$ is $(-1/2\pi i)\, \partial\widehat{\Phi}/\partial y_k$;
(b) the Fourier transform of $\mu_g \Phi$ is $|\det(g)|^{-1} \mu_{{}^t g^{-1}} \widehat{\Phi}$;
(c) the Fourier transform of $\lambda_v \Phi$ is $e^{2\pi i v \bullet y} \widehat{\Phi}$;
(d) the Fourier transform of $\partial\Phi/\partial x_k$ is $(2\pi i y_k) \cdot \widehat{\Phi}$.

I leave these as exercises. One interesting consequence of (b) is that the Fourier transform commutes with the orthogonal group.

7. Examples

• **The Dirac delta.** The Dirac delta δ_0 is defined by

$$\langle \delta_0, f \rangle = f(0) .$$

5 If $\Phi = \delta_0$ then

$$\langle \Phi', f \rangle = \langle \Phi, -f' \rangle = -f'(0) ,$$

and continuing on one obtains

$$\langle \Phi^{(n)}, f \rangle = (-1)^n f^{(n)}(0) .$$

What is its Fourier transform?

$$\langle \widehat{\Phi}, f \rangle = \langle \delta_0, \widehat{f} \rangle = \widehat{f}(0) = \int_{-\infty}^{\infty} f(x)\, dx = \langle 1, f \rangle$$

so $\widehat{\Phi}$ is the constant function 1. And conversely, the Fourier transform of 1 is δ_0.

There is a nice characterization of δ_0:

7.1. Lemma. *A distribution Φ satisfies $x_k \Phi = 0$ for all k if and only if it is a scalar multiple of δ_0.*

Proof. This follows directly from Lemma 3.4.

• **The principal value of $1/x$.** The function $\log|x|$ is integrable near 0 and hence defines a tempered distribution. What is its derivative? Its derivative in the naive sense is the function $1/x$. What relationship does it have with the distributional derivative of $\log|x|$?

Let Φ be the distribution defined by integration against $\log|x|$. Then

$$\begin{aligned}
\langle \Phi', f\rangle &= -\langle \Phi, f'\rangle \\
&= -\int_{-\infty}^{\infty} f'(x)\log|x|\,dx \\
&= -\lim_{\varepsilon\to 0}\left(\int_{-\infty}^{-\varepsilon} f'(x)\log|x|\,dx + \int_{\varepsilon}^{\infty} f'(x)\log x\,dx\right) \\
&= \lim_{\varepsilon\to 0}\left(-\big[f(x)\log|x|\big]_{-\infty}^{-\varepsilon} + \int_{-\infty}^{-\varepsilon}\frac{f(x)}{x}\,dx - \big[f(x)\log|x|\big]_{\varepsilon}^{\infty} + \int_{\varepsilon}^{\infty}\frac{f(x)}{x}\,dx\right) \\
&= \lim_{\varepsilon\to 0}\big(f(\varepsilon)-f(-\varepsilon)\big)\log\varepsilon + \lim_{\varepsilon\to 0}\left(\int_{-\infty}^{-\varepsilon}\frac{f(x)}{x}\,dx + \int_{\varepsilon}^{\infty}\frac{f(x)}{x}\,dx\right) \\
&= \lim_{\varepsilon\to 0}\left(\int_{-\infty}^{-\varepsilon}\frac{f(x)}{x}\,dx + \int_{\varepsilon}^{\infty}\frac{f(x)}{x}\,dx\right),
\end{aligned}$$

since $f(\varepsilon) - f(-\varepsilon) \sim 2\varepsilon f'(0)$ as $\varepsilon \to 0$.

The **principal value** distribution $\mathcal{P}(1/x)$ associated to the function $1/x$ is the limit that appears here:

$$\langle \mathcal{P}, f\rangle = \lim_{\varepsilon\to 0}\left(\int_{-\infty}^{-\varepsilon}\frac{f(x)}{x}\,dx + \int_{\varepsilon}^{\infty}\frac{f(x)}{x}\,dx\right).$$

If $f(x) = x\varphi(x)$ then

$$\langle \mathcal{P}(1/x), f\rangle = \lim_{\varepsilon\to 0}\int_{|x|\geq\varepsilon}\frac{f(x)}{x}\,dx = \int_{\mathbb{R}}\varphi(x)\,dx\,.$$

In other words:

(7.2) $$x\mathcal{P}(1/x) = 1\,.$$

7.3. Proposition. *For $a \neq 0$ $\mu_a\mathcal{P}(1/x) = \operatorname{sgn}(a)\mathcal{P}(1/x)$.*

Here

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

Proof. Let $\Phi = \mathcal{P}(1/x)$. Let $a = \sigma|a|$ with $\sigma = \pm 1$. We have

$$\begin{aligned}
\langle \mu_a\Phi, f\rangle &= \langle \Phi, \mu_{1/a} f\rangle \\
&= \lim_{\varepsilon\to 0}\left(\int_{-\infty}^{-\varepsilon}\frac{f(ax)}{x}\,dx + \int_{\varepsilon}^{\infty}\frac{f(ax)}{x}\,dx\right) \\
&= \lim_{\varepsilon\to 0}\left(\int_{-\infty}^{-a\varepsilon}\frac{f(\sigma x)}{x}\,dx + \int_{a\varepsilon}^{\infty}\frac{f(\sigma x)}{x}\,dx\right) \\
&= \lim_{\varepsilon\to 0}\left(\int_{-\infty}^{-\varepsilon}\frac{f(\sigma x)}{x}\,dx + \int_{\varepsilon}^{\infty}\frac{f(\sigma x)}{x}\,dx\right) \\
&= \sigma\cdot\lim_{\varepsilon\to 0}\left(\int_{-\infty}^{-\varepsilon}\frac{f(x)}{x}\,dx + \int_{\varepsilon}^{\infty}\frac{f(x)}{x}\,dx\right).
\end{aligned}$$

The distribution $\mathcal{P}$ has odd parity. If f is decomposed into even and odd components $f = f_+ + f_-$, then $\langle \mathcal{P}, f\rangle = \langle \mathcal{P}, f_-\rangle$. But f_- is divisible by x, so there is a simple formula for this.

The principal value distribution can be generalized to construct similar distributions behaving well with respect to the multiplicative group—the *parties finies* first defined by Cauchy and elaborated by Hadamard. The derivative of $\mathcal{P}$ isone of these, for example. They will be dealt with elsewhere.

The Fourier transform of $\mathcal{P}$ will be seen in a moment.

• **The step function.** Let

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

Thus

$$\langle H, f\rangle = \int_0^\infty f(x)\,dx$$

and

$$\begin{aligned} \langle H', f\rangle &= -\langle H, f'\rangle \\ &= -\int_0^\infty f'(x)\,dx \\ &= f(0)\,. \end{aligned}$$

Thus $H' = \delta_0$. I'll say something in the next section about an intuitive approach to this.

What is the Fourier transform of H? If Φ is any tempered distribution, then according to the Fourier transform of Φ' is $(2\pi i y)\widehat{\Phi}$. Since $H' = \delta_0$ we have

$$2\pi i y \cdot \widehat{H} = 1\,.$$

But $y\mathcal{P}$ is equal to 1, so $y(\mathcal{P} - (2\pi i)\widehat{H}) = 0$. According to Lemma 7.1, this tells us that

$$2\pi i\widehat{H} = \mathcal{P} + c\,\delta_0$$

for some constant c. What is c?

$$\begin{aligned} \langle 2\pi i\widehat{H}, e^{-\pi x^2}\rangle &= \langle \mathcal{P}, e^{-\pi x^2}\rangle + c\,\langle \delta_0, e^{\pi x^2}\rangle \\ &= c \\ &= (2\pi i)\int_0^\infty e^{-\pi x^2}\,dx \\ &= \pi i \end{aligned}$$

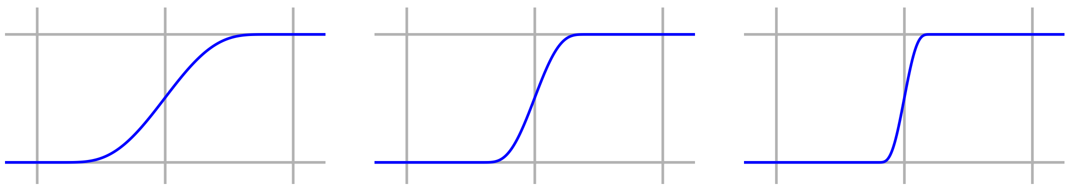
which gives

$$\widehat{H} = \frac{1}{2\pi i}\cdot\mathcal{P} + \frac{1}{2}\cdot\delta_0\,.$$

8. More about the Dirac delta

I attempt in this section to justify the physicists' notion that δ_0 is a function that vanishes everywhere except $x = 0$, but has total area 1 underneath its graph—it is the limit of functions that come closer and closer to that ideal.

The step function $H(x)$ can be well approximated by smooth functions.



What happens to it derivative as this convergence takes place?

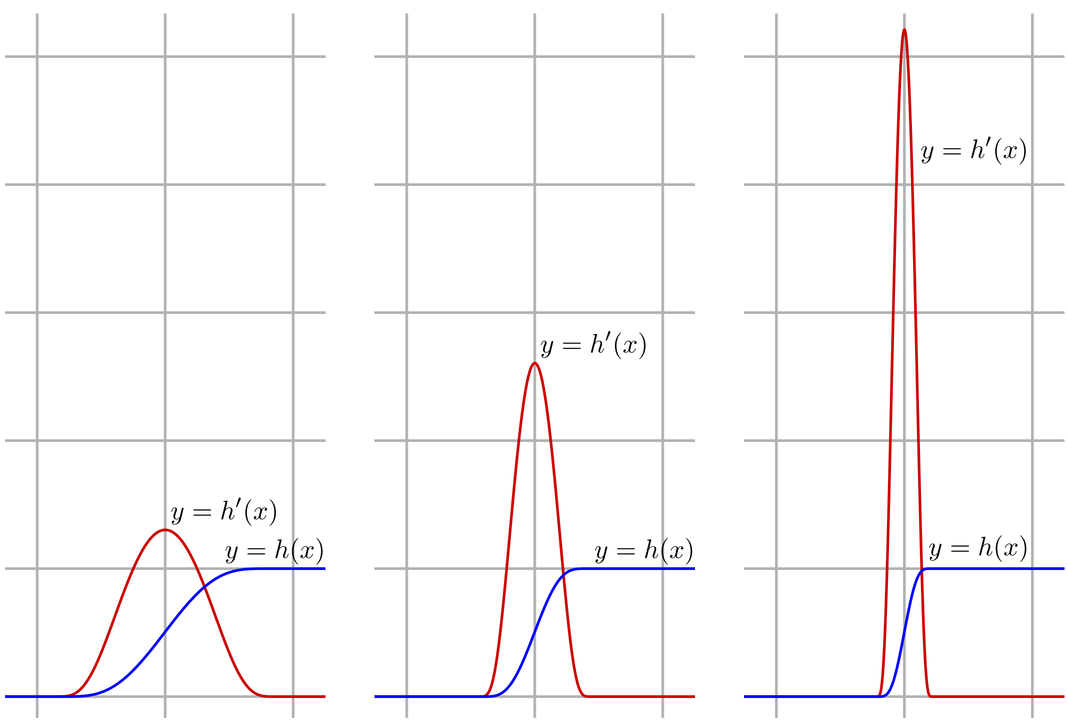
First of all, I recall an explicit construction of functions in $C_c^\infty(\mathbb{R})$. All derivatives of $e^{-1/x}$ vanish at $x = 0$,as do all of $e^{1/(1-x)}$ at $x = 1$. Therefore the function

$$f(x) = \begin{cases} 0 & \text{for} \quad x < 0 \\ e^{-1/x}e^{-1/(1-x)} & \quad 0 \le x \le 1 \\ 0 & \quad 1 < x \end{cases}$$

is smooth on all of $\mathbb{R}$. If we scale it and translate it, we can find a non-negative function $\varphi_1(x)$ which vanishes outside $[-1/2, 1/2]$ and has total area 1 underneath its graph. Now if we define

$$\varphi_c(x) = (1/c)\,\varphi(x/c)$$

then the support of φ_c is contained in $[-c/2, c/2]$ but retains area 1.



As $c \to 0$, the function φ_c has limit δ_0 in the sense that

$$\langle \varphi_c, f \rangle = \int_{-c/2}^{c/2} \varphi_c(x) f(x)\, dx \to f(0)\,.$$

and the function

$$\int_{-\infty}^{x} \varphi_c(t)\,dt\,,$$

whose derivative is φ_c, has as limit the function $H(x)$.

9. Poisson summation

Suppose $f(x)$ to be in $\mathcal{S}(\mathbb{R}^n)$. Then

$$\Theta_f(x) = \sum_n f(x+n)$$

converges absolutely to a smooth function of period 1 on $\mathbb{R}^n$. It may therefore be expanded in a Fourier series

$$\Theta_f(x) = \sum_m \Theta_m e^{2\pi imx}$$

where

$$\begin{aligned}
\Theta_m &= \int_0^1 \Theta_f(x)e^{-2\pi imx}\,dx \\
&= \int_0^1 \sum_n f(x+n)e^{-2\pi imx}\,dx \\
&= \sum_n \int_0^1 f(x+n)e^{-2\pi im(x+n)}\,dx \\
&= \int_{-\infty}^{\infty} f(x)e^{-2\pi imx}\,dx \\
&= \widehat{f}(m)\,.
\end{aligned}$$

9.1. Proposition. (Poisson summation formula) *For f in $\mathcal{S}(\mathbb{R}^n)$*

$$\sum_n f(x+n) = \sum_m \widehat{f}(m)e^{2\pi imx}\,.$$

Suppose we apply this to $f(x) = e^{-\pi tx^2}$. It follows from Proposition 2.1 and (e) of Lemma 2.2 that its Fourier transform is $(1/\sqrt{t})e^{-\pi x^2/t}$. Poisson summation tells us that if

$$\vartheta(t) = \sum_n e^{-\pi n^2 t}$$

then

$$\vartheta_t = \frac{1}{\sqrt{t}}\,\vartheta_{1/t}$$

This functional equation for ϑ will be the justification for the functional equation of Riemann's zeta-function.

10. An application: Riemann's functional equation

Riemann's zeta function is defined to be

$$\zeta(s) = \sum_{n>0} \frac{1}{n^s} \,.$$

The series converges for $\mathrm{RE}(s) > 1$, while for $s = 1$ we have the series

$$\sum_{n>0} \frac{1}{n}$$

which is well known not to converge. For reasons that will become clear in a moment, the more convenient function is

$$\xi(s) = \pi^{-\frac{s}{2}} \Gamma\Big(\frac{s}{2}\Big) \zeta(s)$$

The importance of these functions is that unique factorization of the positive integers allows us to deduce that $\zeta(s)$ may be expanded in the Euler product

$$\zeta(s) = \prod_p \Big(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \cdots \Big) = \prod_p \Big(1 - \frac{1}{p^s}\Big)^{-1} .$$

This suggests that analytic properties of $\zeta(s)$ might imply properties of the prime numbers. As a simple example, that $\zeta(1)$ sums to ∞ implies that there are an infinite number of primes, and even says something, albeit rather weak, about their distribution.

10.1. Theorem. *The function $\xi(s)$ extends meromorphically to all of $\mathbb{C}$, is holomorphic everywhere except for two simple poles at $s = 0$ and $s = 1$, and satisfies the functional equation*

$$\xi(s) = \xi(1 - s) \,.$$

Proof. It all comes down to the consequences of Poisson summation for the function

$$\vartheta(t) = \sum_{n>0} e^{-\pi n^2 t} \,.$$

Poisson summation tells us that

$$1 + 2\vartheta(t) = \frac{1}{\sqrt{t}} \big(1 + 2\vartheta(1/t)\big) \,.$$

The function $\vartheta(t)$ decreases exponentially at ∞. As $t \to 0$ it goes off to ∞, and this equation tells us exactly how:

$$\vartheta(t) = \frac{1}{\sqrt{t}} \, \frac{1 + 2\vartheta(1/t) - \sqrt{t}}{2} \sim \frac{1}{2\sqrt{t}} \,.$$

I recall that the Gamma function is defined to be

$$\Gamma(s) = \int_0^\infty e^{-t} t^s \, \frac{dt}{t} \,.$$

for $\mathrm{RE}(s) > 0$. By a change of variable $t = \pi n^2 x$ this becomes

$$\pi^{-\frac{s}{2}} \Gamma\Big(\frac{s}{2}\Big) \frac{1}{n^s} = \int_0^\infty e^{-\pi n^2 x} x^{s/2} \, \frac{dx}{x} \,.$$

This and the asymptotic behaviour of ϑ at 0 and ∞ tell us that for $\mathrm{RE}(s) > 1$

$$\xi(s) = \int_0^\infty \Big(\sum_{n>0} e^{-\pi n^2 x}\Big) x^{s/2}\,\frac{dx}{x} = \int_0^\infty \vartheta(x) x^{s/2}\,\frac{dx}{x}\,.$$

We can now write (still for $\mathrm{RE}(s) > 1$)

$$\begin{aligned}
\xi(s) &= \int_0^\infty \vartheta(x) x^{s/2}\,\frac{dx}{x} \\
&= \int_0^1 \vartheta(x) x^{s/2}\,\frac{dx}{x} + \int_1^\infty \vartheta(x) x^{s/2}\,\frac{dx}{x} \\
&= \int_1^\infty \vartheta(1/x) x^{-s/2}\,\frac{dx}{x} + \int_1^\infty \vartheta(x) x^{s/2}\,\frac{dx}{x} \\
&= \int_1^\infty \Big(\sqrt{x}\vartheta(x) + \frac{\sqrt{x}}{2} - \frac{1}{2}\Big) x^{-s/2}\,\frac{dx}{x} + \int_1^\infty \vartheta(x) x^{s/2}\,\frac{dx}{x} \\
&= \int_1^\infty \vartheta(x) x^{(1-s)/2}\,\frac{dx}{x} + \int_1^\infty \vartheta(x) x^{s/2}\,\frac{dx}{x} + \frac{1}{2}\int_1^\infty \Big(x^{-(s-1)/2} - x^{-s/2}\Big)\,\frac{dx}{x} \\
&= \int_1^\infty \vartheta(x) x^{(1-s)/2}\,\frac{dx}{x} + \int_1^\infty \vartheta(x) x^{s/2}\,\frac{dx}{x} - \frac{1}{s(1-s)}\,.
\end{aligned}$$

Since $\vartheta(x)$ decreases exponentially fast, the integrals converge for all s. Since the expression is invariant if s and $1-s$ are swapped, we are through.

Furthermore, the pole of $\xi(s)$ at $s = 1$ is simple, with residue 1.

11. References

1. Raymond Edward Alan Christopher Paley and Norbert Wiener, **Fourier transforms in the complex domain**, American Mathematical Society, 1934.

2. Michael Reed and Barry Simon, **Functional analysis**, volume I of the series *Methods of Modern Mathematical Physics*, Academic Press, 1972.

3. ——, **Fourier analysis, self-adjointness**, volume II of the series *Methods of Modern Mathematical Physics*, Academic Press, 1975.

4. Laurent Schwartz, **Théorie des distributions**, Hermann, 1966. This is the first edition of the original two volumes in one.

5. ——, **Méthodes mathématiques pour les sciences physiques**, Hermann, 1966. Translated into English as **Mathematics for the physical sciences**.