Fourier Transforms and Tempered Distributions

Randy S

Abstract Some functions don't have a well-defined Fourier transform when treated as a function, but the Fourier transform may often still be defined by treating things as tempered distributions instead of as functions. This article reviews the basic definitions.

Contents

1 Motivation

Let D be the number of dimensions of space. Use a boldface symbol like $\mathbf{x} = (x_1, ..., x_D)$ to denote a list of D real variables, and use the standard abbreviations $\mathbf{p}\cdot\mathbf{x} \equiv \sum_n p_n x_n$ and $\mathbf{x}^2 \equiv \mathbf{x}\cdot\mathbf{x}$ and $|\mathbf{x}| \equiv \sqrt{\mathbf{x}\cdot\mathbf{x}}$.

Let $\text{poly}(\mathbf{p})$ be a polynomial in the components of $\mathbf{p}$. Formally, the inverse Fourier transform of this polynomial would be

$$\int \frac{d^D p}{(2\pi)^D}\; e^{i\mathbf{p}\cdot\mathbf{x}}\,\text{poly}(\mathbf{p}), \tag{1}$$

but this is undefined as it stands. Article 22050 explains how such an expression can be given a legitimate definition as a function of $\mathbf{x}$ for $\mathbf{x} \neq \mathbf{0}$, by starting with a finite integration domain and then taking a limit. This article reviews a different perspective: the "Fourier transform of a polynomial" can be given a legitimate definition by treating things as tempered distributions instead of trying to treat them as functions.

2 Tempered distributions: definition

Suppose that $s(\mathbf{x})$ rapidly approaches zero as $\mathbf{x}^2 \to \infty$. More precisely, suppose that $s(\mathbf{x})$ is a smooth function for which the magnitude of $p(\mathbf{x})q(\partial_{\mathbf{x}})s(\mathbf{x})$ is finite for all $\mathbf{x}$, whenever p is a polynomial in the components of $\mathbf{x}$ and q is a polynomial in the components of $\partial_{\mathbf{x}}$, the partial derivatives with respect to the components of $\mathbf{x}$. A function $s(\mathbf{x})$ satisfying this condition is called a **Schwartz function** or a **test function**.[1] Example: the function $s(\mathbf{x}) = \exp(-\mathbf{x}^2)$ is a Schwartz function.

Every Schwartz function is absolutely integrable, which means that the integral

$$\int d^D x \; \big|s(\mathbf{x})\big|$$

is finite. Proof: let $p(\mathbf{x})$ be any nonnegative polynomial for which the integral

$$\int d^D x \; \frac{1}{p(\mathbf{x})}$$

is finite,[2] and use the inequality

$$\big|s(\mathbf{x})\big| = \frac{p(\mathbf{x})\big|s(\mathbf{x})\big|}{p(\mathbf{x})} \leq \frac{a}{p(\mathbf{x})},$$

where a is the maximum value of the function $p(\mathbf{x})\big|s(\mathbf{x})\big|$. This implies

$$\int d^D x \; \big|s(\mathbf{x})\big| \leq \int d^D x \; \frac{a}{p(\mathbf{x})},$$

which is finite.

Let $\mathcal{S}$ be the set of Schwartz functions. A continuous linear map from $\mathcal{S}$ to the field $\mathbb{C}$ of complex numbers is called a **tempered distribution**.[3]

---

[1] Hunter (2005), definition 11.1

[2] Example: $p(\mathbf{x}) = (1 + \mathbf{x}^2)^D$

[3] Hunter (2005), beginning of section 11.2. If we only require the map to be defined on test functions that have compact support, then it is called a **distribution** (Hunter (2005), end of section 11.2).

3 Example

The map $\delta : \mathcal{S} \to \mathbb{C}$ defined by

$$s \mapsto s(\mathbf{0}) \qquad\qquad \text{for } s \in \mathcal{S} \tag{2}$$

is one example of a tempered distribution. This is the famous Dirac delta "function" $\delta(\mathbf{x})$, and the map (2) is often expressed formally like this:

$$\int d^D x\ \delta(\mathbf{x})s(\mathbf{x}) = s(\mathbf{0}). \tag{3}$$

The integral-like notation (3) is a convenient way of writing the map (2), even though the thing $\delta(\mathbf{x})$ in the integrand can't be a function. No function can satisfy this condition for all Schwartz functions $s(\mathbf{x})$, but the map (2) is perfectly well-defined. Sections 4-5 explain why the integral-like notation is convenient.

4 Derivatives of a tempered distribution

Given a tempered distribution $T : \mathcal{S} \to \mathbb{C}$, let $\langle T, s\rangle$ denote the result of applying it to the function s. The gradient ∇T of a tempered distribution T is the tempered distribution defined by

$$\langle \nabla T, s\rangle \equiv -\langle T, \nabla s\rangle, \tag{4}$$

where ∇s is the ordinary gradient of the test function $s(\mathbf{x})$ with respect to its argument $\mathbf{x}$.

Here's an example to motivate the negative sign in that definition. For any polynomial $p(\mathbf{x})$, then we can define a corresponding tempered distribution T_p by

$$\langle T_p, s\rangle \equiv \int d^D x\; p(\mathbf{x})s(\mathbf{x}).$$

The integral is finite because $p(\mathbf{x})s(\mathbf{x})$ is still a Schwartz function. Then the gradient ∇T is

$$\langle \nabla T_p, s\rangle \equiv -\langle T_p, \nabla s\rangle = -\int d^D x\; p(\mathbf{x})\nabla s(\mathbf{x}) = \int d^D x\; s(\mathbf{x})\nabla p(\mathbf{x}) = \langle T_{\nabla p}, s\rangle. \tag{5}$$

Altogether,

$$\nabla T_p = T_{\nabla p}.$$

This motivates the sign is the definition (4): that negative sign cancels the one that comes from integrating-by-parts in (5).

A distribution that can be represented in the form $\langle T, s\rangle = \int d^D x\; f(\mathbf{x})s(\mathbf{x})$ for an ordinary function f is called a **regular distribution**. A distribution that is not regular is called **singular**. The δ distribution described in section 3 is a singular distribution. The sign in (4) is motivated by considering regular distributions, but the definition (4) can also be applied to singular distributions. That's one reason why the integral notation used in section 3 is convenient: the definition (4) lets us use integration-by-parts for singular distributions just like we can for regular distributions.

5 Example

Take $D = 1$ and consider the distribution T defined by

$$\langle T, s\rangle = \int dx\ \theta(x)s(x),$$

where $\theta(x)$ is the function

$$\theta(x) \equiv \begin{cases} 1 & \text{if } x \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$\langle \nabla T, s\rangle \equiv -\langle T, \nabla s\rangle = -\int dx\ \theta(x)\frac{d}{dx}s(x) = s(0) - s(\infty) = s(0).$$

This shows that ∇T is the δ distribution that was described in section 3:

$$\langle \nabla T, s\rangle = \langle \delta, s\rangle.$$

6 Convolution with a distribution

Let $U(\mathbf{x})$ denote the translation operator, whose effect on a function $s(\mathbf{x})$ is defined by

$$U(\mathbf{x}')s(\mathbf{x}) = s(\mathbf{x}+\mathbf{x}').$$

Let R denote the reflection operator, whose effect on a function $s(\mathbf{x})$ is defined by

$$Rs(\mathbf{x}) = s(-\mathbf{x}).$$

The **convolution** of a distribution T with a test function s is a function whose value at $\mathbf{x}$ is defined to be $\langle T, RU(\mathbf{x})s\rangle$.

To motivate this definition, suppose that T is a regular distribution:

$$\langle T, s\rangle = \int d^D x\; f(\mathbf{x})s(\mathbf{x}).$$

Then

$$\begin{aligned}\langle T, RU(\mathbf{x})s\rangle &= \int d^D y\; f(\mathbf{y})Rs(\mathbf{x}+\mathbf{y}) \\ &= \int d^D y\; f(\mathbf{y})s(\mathbf{x}-\mathbf{y}) \\ &= \int d^D y\; f(\mathbf{x}-\mathbf{y})s(\mathbf{y}),\end{aligned}$$

which is the usual convolution of two ordinary functions. Example: the convolution of the δ distribution with s is

$$\int d^D y\; \delta(\mathbf{x}-\mathbf{y})s(\mathbf{y}) \equiv \langle \delta, RU(\mathbf{x})s\rangle = s(\mathbf{x}).$$

7 Fourier transform of a Schwartz function

The **Fourier transform** of a Schwartz function $s(\mathbf{x})$ is the function $\tilde{s}(\mathbf{p})$ defined by

$$\tilde{s}(\mathbf{p}) \equiv \int d^D x \; e^{-i\mathbf{p}\cdot\mathbf{x}}\, s(\mathbf{x}).$$

The integral is finite, because

$$\int d^D x \; e^{-i\mathbf{p}\cdot\mathbf{x}}\, s(\mathbf{x}) \leq \int d^D x \; \left|e^{-i\mathbf{p}\cdot\mathbf{x}}\, s(\mathbf{x})\right| = \int d^D x \; \left|s(\mathbf{x})\right|.$$

The right-hand side of this inequality is finite because $s(\mathbf{x})$ is a Schwartz function.

Section 8 shows that the Fourier transform of a Schwartz function is another Schwartz function, so the Fourier transform of $\tilde{s}(\mathbf{p})$ is also finite. The Fourier transform of $\tilde{s}(\mathbf{p})$ turns out to be equal to[4] $(2\pi)^D s(-\mathbf{x})$, which immediately implies

$$\int \frac{d^D p}{(2\pi)^D} \; e^{i\mathbf{p}\cdot\mathbf{x}}\, \tilde{s}(\mathbf{p}) = s(\mathbf{x}). \tag{6}$$

This inverts the Fourier transform.[5]

[4] I won't review the proof here. Proposition 2.5 in Liu (2021) proves it for $D = 1$. (This article is mainly about definitions, so proofs are omitted unless they're relatively short, like the one in section 8.)

[5] The normalization conventions used here are standard in the physics literature. Sometimes the Fourier transform is defined with a prefactor of $1/(2\pi)^D/2$ instead, because this splits the factor of $1/(2\pi)^D$ equally between the Fourier transform and its inverse.

8 Proof

To prove that the Fourier transform of a Schwartz function is another Schwartz function, let p and q be arbitrary polynomials. Then

$$\begin{aligned}
\max_{\mathbf{p}} \big|p(\mathbf{p})q(\partial_{\mathbf{p}})\tilde{s}(\mathbf{p})\big| &= \max_{\mathbf{p}} \left|p(\mathbf{p})q(\partial_{\mathbf{p}}) \int d^D x\; e^{-i\mathbf{p}\cdot\mathbf{x}} s(\mathbf{x})\right| \\
&= \max_{\mathbf{p}} \left|p(\mathbf{p}) \int d^D x\; q(-i\mathbf{x}) e^{-i\mathbf{p}\cdot\mathbf{x}} s(\mathbf{x})\right| \\
&= \max_{\mathbf{p}} \left|\int d^D x\; q(-i\mathbf{x}) p(i\partial_{\mathbf{x}}) e^{-i\mathbf{p}\cdot\mathbf{x}} s(\mathbf{x})\right|. \\
&= \max_{\mathbf{p}} \left|\int d^D x\; e^{-i\mathbf{p}\cdot\mathbf{x}} p(-i\partial_{\mathbf{x}}) q(-i\mathbf{x}) s(\mathbf{x})\right|. \\
&\leq \left|\int d^D x\; p(-i\partial_{\mathbf{x}}) q(-i\mathbf{x}) s(\mathbf{x})\right| \\
&= \left|\int d^D x\; \frac{(1+\mathbf{x}^2)^D p(-i\partial_{\mathbf{x}}) q(-i\mathbf{x}) s(\mathbf{x})}{(1+\mathbf{x}^2)^D}\right| \\
&\leq \max_{\mathbf{x}} \big|(1+\mathbf{x}^2)^D p(-i\partial_{\mathbf{x}}) q(-i\mathbf{x}) s(\mathbf{x})\big| \int d^D x\; \frac{1}{(1+\mathbf{x}^2)^D}.
\end{aligned}$$

The last expression is finite because $s(\mathbf{x})$ is a Schwartz function. This proves that the first expression is also finite, so $\tilde{s}(\mathbf{p})$ is a Schwartz function, too.

9 Fourier transform of a tempered distribution

If T is a tempered distribution, then its Fourier transform $\tilde{T}$ is defined by

$$\langle \tilde{T}, s\rangle \equiv \langle T, \tilde{s}\rangle. \tag{7}$$

As usual, this definition is motivated by the case where T is a regular distribution. In that case,[6]

$$\begin{aligned}
\langle \tilde{T}, s\rangle &\equiv \langle T, \tilde{s}\rangle \\
&= \int d^D x\; f(\mathbf{x})\tilde{s}(\mathbf{x}) \\
&= \int d^D x\; \int d^D y\; f(\mathbf{x})s(\mathbf{y})e^{-i\mathbf{x}\cdot\mathbf{y}} \\
&= \int d^D y\; \int d^D x\; f(\mathbf{x})s(\mathbf{y})e^{-i\mathbf{x}\cdot\mathbf{y}} \\
&= \int d^D y\; \tilde{f}(\mathbf{y})s(\mathbf{y}).
\end{aligned}$$

Altogether, this says that if T a regular distribution with kernel f, then $\tilde{T}$ is the regular distribution whose kernel is the Fourier transform of f.[7]

Example: the Fourier transform of the δ distribution is the distribution $\tilde{\delta}$ given by

$$\langle \tilde{\delta}, s\rangle \equiv \langle \delta, \tilde{s}\rangle = \tilde{s}(\mathbf{0}) = \int d^D x\; s(\mathbf{x}).$$

Intuitively, this says that the Fourier transform of δ is 1.

[6] One of the steps in this derivation assumes that $\int d^D x\; f(\mathbf{x})e^{-i\mathbf{x}\cdot\mathbf{y}}$ is finite. That assumption is harmless in this context, because we know that plenty of functions f with that property do exist, and that's enough functions to provide plenty of motivation the definition (7).

[7] Without invoking the concept of tempered distributions, the definition of the Fourier transform can be extended by continuity to all square-integrable functions (Candes (2021) and section 5.11 in Debnath and Mikusiński (2005)). With that definition, the Fourier transform is a unitary operator on the Hilbert space of square-integrable functions.

10 The motivating example again

Let poly($\mathbf{p}$) be a polynomial in the components of $\mathbf{p}$. If s is a Schwartz function, then the integral

$$\int \frac{d^Dp}{(2\pi)^D}\; s(\mathbf{p})\,\text{poly}(\mathbf{p})$$

is finite, so we can define a distribution Ω by

$$\langle \Omega, s\rangle \equiv \int \frac{d^Dp}{(2\pi)^D}\; s(\mathbf{p})\,\text{poly}(\mathbf{p}).$$

The Fourier transform of this distribution is the distribution $\tilde{\Omega}$ defined by

$$\langle \tilde{\Omega}, s\rangle \equiv \langle \Omega, \tilde{s}\rangle.$$

Explicitly,

$$\begin{aligned}\langle \tilde{\Omega}, s\rangle \equiv \langle \Omega, \tilde{s}\rangle &= \int \frac{d^Dp}{(2\pi)^D}\; \tilde{s}(\mathbf{p})\,\text{poly}(\mathbf{p}) \\ &= \int \frac{d^Dp}{(2\pi)^D}\; \text{poly}(\mathbf{p}) \int d^Dx\; e^{-i\mathbf{p}\cdot\mathbf{x}} s(\mathbf{x}) \\ &= \int \frac{d^Dp}{(2\pi)^D} \int d^Dx\; s(\mathbf{x})\,\text{poly}(i\nabla)\,e^{-i\mathbf{p}\cdot\mathbf{x}} \\ &= \int \frac{d^Dp}{(2\pi)^D} \int d^Dx\; e^{-i\mathbf{p}\cdot\mathbf{x}}\,\text{poly}(-i\nabla) s(\mathbf{x}).\end{aligned}$$

This is the integral over all $\mathbf{p}$ of the Fourier transform of the function poly$(-i\nabla)s(\mathbf{x})$. According to equation (6), the result is the function poly$(-i\nabla)s(\mathbf{x})$ evaluated at $\mathbf{x}=0$, so

$$\langle \tilde{\Omega}, s\rangle = \langle \delta,\, \text{poly}(-i\nabla)s\rangle = \langle \text{poly}(-i\nabla)\delta,\, s\rangle.$$

In this sense, taking the Fourier transform of a polynomial gives a combination of derivatives of the Dirac delta distribution.

11 References

Candes, 2021. "Applied Fourier Analysis and Elements of Modern Signal Processing, Lecture 3" https://statweb.stanford.edu/~candes/teaching/math262/Lectures/Lecture03.pdf

Debnath and Mikusiński, 2005. *Introduction to Hilbert Spaces with Applications (Third Edition).* Academic Press

Hunter, 2005. "Chapter 11: Distributions and the Fourier transform" https://www.math.ucdavis.edu/~hunter/book/pdfbook.html

Liu, 2021. "Supplementary Note on Greens Function Method" http://www.math.nagoya-u.ac.jp/~richard/teaching/f2021/SupplementaryNote_Ziyu_Liu.pdf

12 References in this series

Article **00980** (https://cphysics.org/article/00980):
"The Free Scalar Quantum Field: Vacuum State" (version 2022-08-21)

Article **22050** (https://cphysics.org/article/22050):
"Contour Integrals: Applications to the Free Scalar Model" (version 2022-08-23)