Part II

Fourier-Laplace Transforms

The idea of Fourier analysis is to exploit the underlying symmetry of the space to study functions. This is done by decomposing the functions of interest as sums or integrals of functions that transform in simple ways under the action of the underlying symmetry group.

1 | Continuous Characters

The space of interest to us is $\mathbb{R}^n$ which acts on itself by translation. The functions of interest to us are eigenfunctions for translations or transform under translation by multiplication by a factor. These are called characters. A character c on $\mathbb{R}^n$ is such that for every $y \in \mathbb{R}^n$,

$$c(x+y) = k(y)c(x),$$

for all $x \in \mathbb{R}^n$. c is completely determined by c once $c(0)$ is known. A character c is said to be normalized if $c(0) = 1$. In which case, we have $c(x) = c(x+0) = k(x)c(0) = k(x)$. We will denote normalised characters by ξ. For normalised characters, we have

$$\xi(x+y) = \xi(x)\xi(y). \tag{character}$$

The characters should be expected to have good behavior under convolutions as their value at translations is given by product of its value at these points. Suppose ξ is a continuous normalized character, and $f \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ such that

$$\int_{\mathbb{R}^n} \xi(-y)f(y)d\lambda(y) = 1$$

then,

$$\xi(x) = \int_{\mathbb{R}^n} \xi(x)\xi(-y)f(y)d\lambda(y) = \int_{P_f} \xi(x-y)f(y)d\lambda(y) = \xi * f(x).$$

where λ is the Lebesgue measure and P_f is the support of f. So, ξ can be expressed as a convolution with the test function f[1]. Since convolutions with test functions smooth functions the continuous normalised character ξ is differentiable, $\xi \in \mathcal{C}^\infty(\mathbb{R}^n)$. If e_j is an orthonormal basis on $\mathbb{R}^n$,

$$\xi(x) = \xi\big(\textstyle\sum_j x_j e_j\big) = \prod_j \xi_j(x_j),$$

[1] Such test functions exist and are abundant, we can choose any compactly supported smooth function f. If $0 < \left|\int \xi(-y)f(y)d\lambda(y)\right| = K < \infty$, the smooth f/K then will do the job

where $\xi_j(x_j) = \xi(x_j e_j)$. ξ_j are continuous functions, and are themselves continuous characters on $\mathbb{R}$. Since any continuous character can be written as a convolution with a suitable test function, ξ_js are smooth functions. Differentiating, we get,

$$d\xi_j(x_j + y_j)/dy_j = d\big(\xi_j(x_j)\xi_j(y_j)\big)/dy_j = \xi_j(x_j)\big[d\xi_j(y_j)/dy_j\big].$$

At $y_j = 0$, we have

$$\left[\frac{d\xi_j(x_j)}{dy_j}\right] = \xi_j(x_j)\left[\frac{d\xi_j(0)}{dy_j}\right] = -i\chi_j\xi_j(x_j).$$

So, the value of the derivative at each point is also determined by its value at origin. The choice $-i\chi_j$ is used for notational convenience for later.

The unique differentiable function that satisfies this ordinary differential equation is the exponential function, so, we have $\xi_j(x_j) = \xi_i(0)e^{-ix_j\chi_j}$. If ξ is a normalised continuous character, then $\xi_i(0) = 1$. So we have

$$\xi_j(x_j) = e^{-ix_j\chi_j}.$$

Any continuous normalised character on $\mathbb{R}^n$ is of the form,

$$\xi(x) = \prod_i \xi_i(x_i) = e^{-i\sum_j x_j\chi_j} := e^{-i\langle x,\chi\rangle}.$$

The collection of all continuous normalised characters on $\mathbb{R}^n$ is an abelian group, and we denote it by $\widehat{\mathbb{R}}^n$. It is a locally compact abelian group. We have an isomorphism of locally compact abelian groups which assigns to each $(\chi_j)_{n\geq j\geq 1}$ a continuous normalised character where χ_j is a complex number.

Since we are interested in studying functions and distributions that are fairly well-behaved at ∞ we would only need bounded characters. If $i\chi_j$ is not purely complex, then $e^{-ix_j\chi_j}$ is unbounded. If the character is bounded then χ_j will have to be real. Denote the collection of continuous bounded normalised characters by $(\mathbb{R}^n)'$. The above discussion gives us an isomorphism,

$$e^{-i\langle\cdot,\cdot\rangle} : \mathbb{R}^n \to (\mathbb{R}^n)'$$

which sends $\chi \in \mathbb{R}^n$ to the character $e^{-i\langle\cdot,\chi\rangle}$. If the boundedness requirement is removed, then we have a character for each $\chi \in \mathbb{C}^n$. As locally compact abelian groups we have $\mathbb{R}^n \cong (\mathbb{R}^n)' \subset \mathbb{C}^n \cong \widehat{\mathbb{R}}^n$.

The idea of Fourier transform is to decompose a given function into a continuous family of normalised characters. If f is an integrable function on $\mathbb{R}^n$ then its decomposition in terms of characters is,

$$\mathcal{F}f(\chi) \equiv \widehat{f}(\chi) = \int_{\mathbb{R}^n} e^{-i\langle x,\chi\rangle} f(x)d\lambda(x). \qquad \text{(Fourier transform)}$$

The pairing $\langle x,\chi\rangle$ allows for easier manipulations. One starts with a function, works with its Fourier transform, and inverts back the manipulated function via the Fourier inversion,

$$f(x) = (2\pi)^{-n}\int_{\mathbb{R}^n} e^{i\langle x,\chi\rangle}\widehat{f}(\chi)d\lambda(\chi). \qquad \text{(Fourier Inversion)}$$

1.1 | The Fourier Transform of Distributions

To be able to use Fourier transform as a tool for studying distributions, the idea for defining Fourier transform on distributions is to define it for test functions, and define Fourier transform of distributions as the adjoint of its action on test functions via the pairing, $\langle \mathcal{F}\kappa, f\rangle = \langle \kappa, \mathcal{F}f\rangle$. For this to make sense, the term $\langle \kappa, \mathcal{F}f\rangle$ should be well defined. In particular, $\mathcal{F}f$ should be a test function. So, the space of test functions should be invariant under Fourier transform.

Lemma 1.1. *There exist some $f \in \mathcal{C}_c^\infty(\mathbb{R})$ such that $\mathcal{F}(f) \notin \mathcal{C}_c^\infty(\mathbb{R})$.*

Proof

Let f be a smooth function with compact support. Without loss of generality assume the support is contained in the closed interval $[-\epsilon, \epsilon]$. The Fourier transform of f is given by,

$$\begin{aligned}
\mathcal{F}(f)(\chi) &= \int_{-\epsilon}^{\epsilon} e^{-i\langle x,\chi\rangle} f(x) d\lambda(x) \\
&= \int_{-\epsilon}^{\epsilon} \Big[\textstyle\sum_k (-i\langle x,\chi\rangle)^k (k!)^{-1} \Big] f(x) d\lambda(x) \\
&= \sum_k \Big[(-i)^k (k!)^{-1} \int_{-\epsilon}^{\epsilon} x^k f(x) d\lambda(x) \Big] \chi^k = \textstyle\sum_k c_k \chi^k .
\end{aligned}$$

where $c_k = (-i)^k (k!)^{-1} \int_{[-\epsilon,\epsilon]} x^k f(x) d\lambda(x)$. We have,

$$|c_k| = (k!)^{-1} \Big| \int_{-\epsilon}^{\epsilon} x^k f(x) d\lambda(x) \Big| \leq (k!)^{-1} \|f\|_{\sup} \Big| \int_{-\epsilon}^{\epsilon} x^k d\lambda(x) \Big| \leq 2(k!)^{-1} \epsilon^{k+1} \|f\|_{\sup} .$$

By the ratio test, the radius of convergence is given by $\lim_k |c_k|/|c_{k+1}| = 2\epsilon(k+1) \to \infty$. So, the Fourier transform of f can be written as a power series with infinite radius of convergence. So, the Fourier transform of f does not have a compact support, or $\mathcal{F}(f) \notin \mathcal{C}_c^\infty(\mathbb{R})$. □

So, the space of compactly supported smooth functions cannot be used to define Fourier transform on distributions. The reason for the failure is that the radius of convergence of the Fourier transform could not be bounded.

2 | The Space of Tempered Distributions

We need a space of functions larger than $\mathcal{C}_c^\infty(\mathbb{R}^n)$ such that the Fourier transform also belongs to it. We need to find the small enough space of functions which contains $\mathcal{C}_c^\infty(\mathbb{R}^n)$ and remains closed under Fourier transforms. The smallness here is needed to ensures that we can Fourier transform on a dense subset of the space of distributions.

Since the purpose of theory of distributions to study differentiability properties, we first expect the space of test functions $\mathcal{S}$ to have the nice differentiability properties. For any smooth function f, using translation invariance of λ and the fact that the integral is a constant function in x, we have

$$0 = (\partial/\partial x_i) \Big[\int_{\mathbb{R}^n} e^{-i\langle x,\chi\rangle} f(x) d\lambda(x) \Big] = \int_{\mathbb{R}^n} (\partial/\partial x_i) \big(e^{-i\langle x,\chi\rangle} f(x) \big) d\lambda(x),$$

which gives us,

$$\int_{\mathbb{R}^n} e^{-i\langle x,\chi\rangle}\Big[(-i\chi_j)f(x)+\big(\partial f/\partial x_j\big)(x)\Big]d\lambda(x)=0.$$

In the second step we took the differentiation inside the integral sign. Using this, we have

$$\begin{aligned}\mathcal{F}(\partial f/\partial x_i)(\chi) &= \int_{\mathbb{R}^n} e^{-i\langle x,\chi\rangle}\big(\partial f/\partial x_j\big)(x)d\lambda(x)\\ &= i\chi_j\int_{\mathbb{R}^n} e^{-i\langle x,\chi\rangle} f(x)d\lambda(x) = i\chi_j\big(\mathcal{F}f(\chi)\big). \end{aligned}\tag{2.1}$$

Similarly,

$$\begin{aligned}\mathcal{F}(x_j f)(\chi) &= \int_{\mathbb{R}^n} e^{-i\langle x,\chi\rangle}\big(x_i f(x)\big)d\lambda(x) = -i\int_{\mathbb{R}^n}\big(\partial e^{-i\langle x,\chi\rangle}/\partial \chi_j\big)f(x)d\lambda(x)\\ &= -i(\partial/\partial\chi_j)\Big[\int_{\mathbb{R}^n} e^{-i\langle x,\chi\rangle}f(x)d\lambda(x)\Big] = -i\big(\partial\mathcal{F}f/\partial\chi_j\big)(\chi).\end{aligned}\tag{2.2}$$

We want these functions to also be in the space of test functions.

2.1 | The Schwartz Space $\mathcal{S}(\mathbb{R}^n)$

We want the space of test functions to be such that if f is a test function so is $x^\beta\partial^\alpha f$. Such functions are called rapidly decreasing functions or Schwartz functions. By Schwartz space, we denote the set of all smooth functions f such that,

$$\sup_x \big|x^\beta\partial^\alpha\big(f(x)\big)\big| < \infty,$$

for all multi-indices α and β. The collection of all Schwartz functions is denoted by $\mathcal{S}(\mathbb{R}^n)$. If φ is a compactly supported smooth function then $x^\beta\partial^\alpha\varphi(x)$ with the same support as φ and it attains a maxima.

$$\mathcal{C}_c^\infty(\mathbb{R}^n)\subset\mathcal{S}(\mathbb{R}^n).$$

$\mathcal{S}(\mathbb{R}^n)$ consists of all smooth functions on $\mathbb{R}^n$ which vanish faster than than any polynomial and is closed under taking derivatives. The linearity of derivatives ∂^α, and multiplication by x^β, combined with the fact that modulus function satisfying triangle inequality implies that $\mathcal{S}(\mathbb{R}^n)$ is a vector space.

To build a topology on $\mathcal{S}(\mathbb{R}^n)$, consider the map,

$$\|f\|_{\alpha,\beta} \coloneqq \sup_x \big|x^\beta\partial^\alpha\big(f(x)\big)\big|.$$

for every multi-indices α and β. The properties of $|\cdot|$ imply, $\|\lambda f\|_{\alpha,\beta} = |\lambda|\,\|f\|_{\alpha,\beta}$, and $\|(f+g)\|_{\alpha,\beta} \le \|f\|_{\alpha,\beta}+\|g\|_{\alpha,\beta}$. Hence $\|\cdot\|_{\alpha,\beta}$ is a semi-norm for every multi-indices α and β.

Since there are only countably many αs and βs, it follows that $\{\|\cdot\|_{\alpha,\beta}\}$ is a countable collection of semi-norms and hence generates a topology that makes $\mathcal{S}(\mathbb{R}^n)$ a metric space. It is hence separable, that is, it has a countable dense set. In this topology, a sequence $\{f_i\}$ converges to f as $i\to\infty$ if and only if

$$\|f_i - f\|_{\alpha,\beta} = \sup_x \big|x^\beta\partial^\alpha\big(f_i(x)-f(x)\big)\big| \to 0,$$

for all multi-indices α and β.

Hence every compactly supported smooth function is Schwartz.

If f is a Schwartz function then there exists $M_{\alpha,\beta}$ such that, $\left|x^\beta\right|\left|\partial^\alpha f(x)\right| \leq M_{\alpha,\beta}$ for every multi-indices α and β. Let K be the unit ball around the origin $0 \in \mathbb{R}^n$.

$$\int_{\mathbb{R}^n} \left|f(x)\right| d\lambda(x) = \left[\int_K + \int_{K^c}\right]\left|f(x)\right| d\lambda(x),$$

and by choosing β so that $|x^\beta| > |x|^2$ for all $|x| > 1$, we have

$$\int_K \left|\partial^\alpha f(x)\right| d\lambda(x) \leq \lambda(K) M_{\alpha,0},$$

$$\int_{K^c} \left|\partial^\alpha f(x)\right| d\lambda(x) < \mu M_{\alpha,\beta}.$$

where μ is some constant which depends on changing the variable from cartesian to polar coordinates and $\int |\partial^\alpha f(x)| d\lambda(x) < \infty$.[2] Similarly, $\int |x^\beta f(x)| d\lambda(x) < \infty$. Hence we have proved that,

Lemma 2.1. *For every $f \in \mathcal{S}(\mathbb{R}^n)$ the functions $f, \partial^\alpha f, x^\beta f$ are integrable.*

Since Fourier transform can be defined on integrable functions we can define Fourier transform on the Schwartz space. Since $\partial^\alpha f, x^\beta f$ are integrable, it follows that we can define Fourier transform of $\partial f/\partial x_j$ and $x^\beta f$, and are given by (2.1) and (2.2).

$$\mathcal{F}(\partial f/\partial x_j)(\chi) = i\chi_j(\mathcal{F}f)(\chi), \quad \mathcal{F}(x_j f)(\chi) = -i\partial(\mathcal{F}f)/\partial\chi_i(\chi). \tag{exchange}$$

We however need to verify that the Fourier transform keeps $\mathcal{S}(\mathbb{R}^n)$ invariant.

Lemma 2.2. *$\mathcal{F}$ is continuous, and $\mathcal{F}(\mathcal{S}(\mathbb{R}^n)) \subseteq \mathcal{S}(\mathbb{R}^n)$.*

Proof

By (exchange), it follows that, $|\chi^\beta \partial^\alpha(\mathcal{F}f)(\chi)| = |\chi^\beta \mathcal{F}(x^\alpha f)(\chi)| = |\mathcal{F}(\partial^\beta(x^\alpha f))(\chi)|$. If we let $g(x) = \partial^\beta\big(x^\alpha f(x)\big)$, we have

$$|\mathcal{F}(\partial^\beta(x^\alpha f))(\chi)| = \left|(\mathcal{F}g)(\chi)\right|$$

If f is a Schwartz function g is also a Schwartz function. In such a case,

$$\begin{aligned}\left|(\mathcal{F}g)(\chi)\right| &= \left|\int_{\mathbb{R}^n} e^{-i\langle x,\chi\rangle} g(x) d\lambda(x)\right| \\ &\leq \int_{\mathbb{R}^n} \left|e^{-i\langle x,\chi\rangle}\right|\left|g(x)\right| d\lambda(x) \\ &= \int_{\mathbb{R}^n} \left[\tfrac{(1+|x|)^{n+1} g(x)}{(1+|x|)^{n+1}}\right] d\lambda(x)\end{aligned}$$

[2] Note that with a small modification, by changing α from 0 to α and choosing β such that $|x^\beta|^p > |x|^2$ for all $|x| > 1$, we can show that $\int_{\mathbb{R}^n} \left|\partial^\alpha f(x)\right|^p d\lambda(x) < \infty$, or $\partial^\alpha f \in L^p(\mathbb{R}^n)$ whenever $f \in \mathcal{S}(\mathbb{R}^n)$.

Since Schwartz functions decrease faster than any polynomials, $h(x) = (1+|x|)^{n+1}g(x) = (1+\cdots+|x|^{n+1})g(x)$ is bounded. Then by definition,

$$\sup_x \left|(1+|x|)^{n+1}g(x)\right| = \left\|h\right\|_{\sup} < \infty.$$

In polar coordinates we have,

$$\begin{aligned}
\left|\chi^\beta \partial^\alpha (\mathcal{F}f)(\chi)\right| &\leq \left\|h\right\|_{\sup} \int_{\mathbb{R}^n} \tfrac{d\lambda(x)}{(1+|x|)^{n+1}} \\
&= \left\|h\right\|_{\sup} \int_0^\infty \int_{S^{n-1}} \left[\tfrac{r^{n-1}}{(1+r)^{n+1}}\right] dr d\omega \\
&= \left\|h\right\|_{\sup} \int_0^\infty \left[\tfrac{r^{n-1}}{(1+r)^{n+1}}\right] dr \int_{S^{n-1}} d\omega.
\end{aligned}$$

In the last step we used Fubini's rule. $n+1$ is chosen so the integral is well behaved and $r^{n-1}/(1+r)^{n+1}$ is integrable. We have

$$\sup_\chi \left|\chi^\beta \partial^\alpha (\mathcal{F}f)(\chi)\right| < \infty.$$

Which implies that $\mathcal{F}(\mathcal{S}(\mathbb{R}^n)) \subseteq \mathcal{S}(\mathbb{R}^n)$. Continuity follows because if $\{f_i\}$ is a sequence of Schwartz functions such that $f_i \to 0$ then,

$$\left|\mathcal{F}f_i(\chi)\right| = \left|\int_{\mathbb{R}^n} e^{-i\langle x,\chi\rangle} f_i(x) d\lambda(x)\right| \leq \int_{\mathbb{R}^n} \underbrace{\left|e^{-i\langle x,\chi\rangle}\right|}_{1} \left|f_i(x)\right| d\lambda(x) \to 0.$$

Hence $\mathcal{F}f_i \to 0$ or $\mathcal{F}$ is continuous. □

Heuristically speaking, ∂_j tells us how a function changes at different points, and x_j tells us how the value of the function varies in coordinates. If both are known, then the function is fixed upto constant multiple.

Lemma 2.3. *Let $L : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ be linear, suppose*

$$\partial_j L = L\partial_j,\ x_j L = L x_j$$

for all j then $\exists! c$ such that $L(f) = cf$ for all $f \in \mathcal{S}(\mathbb{R}^n)$.

Proof

Theorem 2.4. *$\mathcal{F} : f \mapsto \mathcal{F}f$ is an isomorphism.*

Proof

Note that $\mathcal{F} \circ \mathcal{F}$ maps

2.2 | The Tempered Distributions

The idea to define Fourier transform on distributions is to first define it Schwartz functions, transfer it onto the space of linear transformations on $\mathcal{S}(\mathbb{R}^n)$ via the adjoint map.

A continuous linear functional τ on the Schwartz space $\mathcal{S}(\mathbb{R}^n)$ is called a tempered distribution. The set of all tempered distributions is denoted by $\mathcal{T}(\mathbb{R}^n)$. Since every compactly supported smooth function is also a Schwartz function, the restriction of a tempered distribution τ restricted to $\mathcal{C}_c^\infty(\mathbb{R}^n)$ is a continuous linear functional on $\mathcal{C}_c^\infty(\mathbb{R}^n)$. Hence we have,

$$\mathcal{T}(\mathbb{R}^n) \subset \mathcal{D}(\mathbb{R}^n).$$

Once Fourier transform is defined on $\mathcal{T}(\mathbb{R}^n)$, the Fourier transform is then extended to other distributions via limits using completeness of $\mathcal{D}(\mathbb{R}^n)$. For this idea to work, we need tempered distributions to be dense in $\mathcal{D}(\mathbb{R}^n)$.

2.2.1 | Density of Tempered Distributions

Theorem 2.5. *$\mathcal{C}_c^\infty(\mathbb{R}^n)$ is dense in $\mathcal{S}(\mathbb{R}^n)$.*

Proof

Since the convolution of two functions combines the nice properties of both functions, it can be used to approximate a given function by functions with nicer properties.

Given a positive function $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$, the function $\varphi_t(x) = \varphi\big(x/t\big)/t^n$ defines a positive function in $\mathcal{C}_c^\infty(\mathbb{R}^n)$ for all $t > 0$. For any integrable function f, define the scaled convolution with φ to be the function,

$$f_{\varphi_t}(x) = f * \varphi_t(x) = \int_{\mathbb{R}^n} f(x-y)\varphi_t(x)d\lambda(y).$$

The product inside the integral is nonzero only on the compact support of φ_t. Hence $f_{\varphi_t} \in \mathcal{C}_c^\infty(\mathbb{R}^n)$. Note first that by change of variable,

$$\int_{\mathbb{R}^n} \varphi_t(x)d\lambda(x) = \int_{\mathbb{R}^n} \varphi(y)d\lambda(y)$$

Hence for a positive function $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} \varphi(x)d\lambda(x) = 1$, we have,

$$\begin{aligned}\big|f(x) - f_{\varphi_t}(x)\big| &= \left|f(x)\Big[\int_{\mathbb{R}^n} \varphi_t(x)d\lambda(x)\Big] - \int_{\mathbb{R}^n} f(x-y)\varphi_t(y)d\lambda(y)\right| \\ &= \left|\int_{\mathbb{R}^n} \big(f(x) - f(x-y)\big)\varphi_t(y)d\lambda(y)\right| \leq \sup_{y\in P_{\varphi_t}} \big|f(x) - f(x-y)\big|.\end{aligned}$$

In the second step, we used the fact that outside the support P_{φ_t} of φ_t the product of the two functions will be zero. Now by continuity of f, and as support of φ_t decreases as $t \to 0$, the value of

$$\sup_{y\in P_{\varphi_t}} \big|f(x) - f(x-y)\big| \to 0.$$

Since this holds for every x it follows that $f_{\varphi_t} \to f$. □

2.2.2 | The Fourier Transformation on $\mathcal{T}(\mathbb{R}^n)$

2.2.3 | The Convolution Theorem

3 | The Fourier-Laplace Transform

Theorem 3.1. (Paley-Wiener-Schwartz)

4 | Analyticity-Fourier Transforms Relations

References

[1] L Hörmander, The Analysis of Linear Partial Differential Operators I, Second Edition, Springer-Verlag, 1989