B4.4 Fourier Analysis HT21

Lecture 2: Properties of the Fourier transform on L^1 and definition of the Schwartz test functions

1. Invariance and symmetry properties of the Fourier transform
2. The convolution rule
3. The differentiation rules
4. Rapidly decreasing functions and Schwartz test functions
5. Examples

The material corresponds to pp. 5–12 in the lecture notes and should be covered in Week 1.

Invariance and symmetry properties of the Fourier transform

In this connection there are three groups that act naturally on $\mathbb{R}^n$:

- rotations and more generally the orthogonal group: $x \mapsto \theta x$
- dilations: $x \mapsto rx$
- translations: $x \mapsto x + h$

The orthogonal group $\mathrm{O}(n)$: A real $n \times n$ matrix X is orthogonal, $X \in \mathrm{O}(n)$, if its columns form an orthonormal basis for $\mathbb{R}^n$, so precisely when $X^{\mathrm{t}}X = \mathrm{I}$ holds. A rotation is an orthogonal matrix whose determinant is 1. We denote the set of these *special orthogonal* matrices by $\mathrm{SO}(n)$. Of course these notions have intrinsic and invariant meaning too, but we shall use this terminology and concrete representation.

Invariance and symmetry properties of the Fourier transform

Proposition: If $f \in \mathsf{L}^1(\mathbb{R}^n)$ and $\theta \in \mathrm{O}(n)$, then with the notation $(\theta_* f)(x) := f(\theta x)$, $x \in \mathbb{R}^n$, we have

$$\widehat{(\theta_* f)} = \theta_* \widehat{f}.$$

Proof. This is a straight forward calculation using the change-of-variables formula:

$$\begin{aligned}
\widehat{(\theta_* f)}(\xi) &= \int_{\mathbb{R}^n} f(\theta x) \mathrm{e}^{-\mathrm{i}\xi \cdot x} \, \mathrm{d}x \\
&\stackrel{y=\theta x}{=} \int_{\mathbb{R}^n} f(y) \mathrm{e}^{-\mathrm{i}\xi \cdot \theta^{-1} y} \left| \det \theta^{-1} \right| \mathrm{d}y \\
&\stackrel{\theta^{-1}=\theta^{\mathrm{t}}}{=} \int_{\mathbb{R}^n} f(y) \mathrm{e}^{-\mathrm{i}\theta \xi \cdot y} \, \mathrm{d}y \\
&= \left(\theta_* \widehat{f}\right)(\xi),
\end{aligned}$$

as required. □

Special case: reflection through the origin

This is the case with $\theta = -\mathrm{I} \in \mathrm{O}(n)$: for $f \in \mathsf{L}^1(\mathbb{R}^n)$ write

$$\widetilde{f}(x) := f(-x), \quad x \in \mathbb{R}^n.$$

We record that $\mathcal{F}(\widetilde{f}) = \widetilde{\mathcal{F}(f)}$ holds.

Examples

(i) If $f \in \mathsf{L}^1(\mathbb{R}^n)$ is even (odd), then so is $\widehat{f}$.

(ii) If $f \in \mathsf{L}^1(\mathbb{R}^n)$ and $\theta_* f = f$ for all $\theta \in \mathrm{O}(n)$, then also $\theta_* \widehat{f} = \widehat{f}$ holds for all $\theta \in \mathrm{O}(n)$.

In connection with (ii), note that a function is *radial*, meaning that the value $f(x)$ only depends on $|x|$, precisely when $\theta_* f = f$ holds for all $\theta \in \mathrm{O}(n)$. This will be discussed further on a problem sheet.

Dilations

We define two types.

- The dilation d_r of $\mathbb{R}^n$ by factor $r > 0$ is defined by $d_r(x) := rx$ and transferred to functions $f \in \mathrm{L}^1(\mathbb{R}^n)$ by $\big(d_r f\big)(x) := f(rx)$.
- The L^1 dilation with factor $r > 0$ of $f \in \mathrm{L}^1(\mathbb{R}^n)$ is

$$f_r(x) := \frac{1}{r^n} f\left(\frac{x}{r}\right)$$

Note that it is called L^1 dilation because it preserves the L^1 norm: $\|f_r\|_1 = \|f\|_1$.

Proposition: Let $f \in \mathrm{L}^1(\mathbb{R}^n)$ and $r > 0$. Then

$$\widehat{d_r f} = \left(\widehat{f}\right)_r, \; \widehat{f_r} = d_r \widehat{f}.$$

The proof is a straight forward calculation with the change-of-variables formula: see lecture notes for details.

Translation by $h \in \mathbb{R}^n$ is $\tau_h(x) := x + h$, $x \in \mathbb{R}^n$, and transferred to functions $f \in \mathrm{L}^1(\mathbb{R}^n)$ by $\big(\tau_h f\big)(x) := f(x+h)$.

Proposition: Let $f \in \mathrm{L}^1(\mathbb{R}^n)$ and $h \in \mathbb{R}^n$. Then

$$\widehat{(\tau_h f)}(\xi) = \widehat{f}(\xi)\mathrm{e}^{\mathrm{i}h\cdot\xi} \text{ and } \mathcal{F}_{x\to\xi}\Big(\mathrm{e}^{-\mathrm{i}h\cdot x}f(x)\Big) = \big(\tau_h\widehat{f}\big)(\xi).$$

The proof is a straight forward calculation with the change-of-variables formula: see lecture notes for details.

The convolution rule: Let f, $g \in \mathrm{L}^1(\mathbb{R}^n)$. Then $f * g \in \mathrm{L}^1(\mathbb{R}^n)$ and

$$\widehat{f * g} = \widehat{f}\widehat{g}.$$

In fact, there is another related rule, also called the convolution rule, but it will have to wait until we have developed the theory a bit further.

Proof. It is an exercise in using Fubini's theorem to swap integration orders:

$$\begin{aligned}
(\widehat{f * g})(\xi) &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x-y)g(y)\,\mathrm{d}y\,\mathrm{e}^{-\mathrm{i}\xi\cdot x}\,\mathrm{d}x \\
&\overset{\text{Fubini}}{=} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x-y)g(y)\mathrm{e}^{-\mathrm{i}\xi\cdot x}\,\mathrm{d}x\,\mathrm{d}y \\
&\overset{t=x-y}{=} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(t)g(y)\,\mathrm{e}^{-\mathrm{i}\xi\cdot(t+y)}\,\mathrm{d}t\,\mathrm{d}y \\
&= \widehat{f}(\xi)\widehat{g}(\xi),
\end{aligned}$$

as required. □

Example The Wiener algebra $\mathcal{F}\big(\mathsf{L}^1(\mathbb{R}^n)\big)$ is an algebra: if h, k are two functions in the Wiener algebra, also hk is in the Wiener algebra.

Indeed we find f, $g \in \mathsf{L}^1(\mathbb{R}^n)$ so $h = \widehat{f}$, $k = \widehat{g}$, and then by the convolution rule $hk = \widehat{f * g} \in \mathcal{F}\big(\mathsf{L}^1(\mathbb{R}^n)\big)$.

Note that $\mathsf{L}^1(\mathbb{R}^n)$ becomes an algebra if we use convolution as product. The convolution rule is then saying that the Fourier transform is an algebra homomorphism of $\mathsf{L}^1(\mathbb{R}^n)$ onto the Wiener algebra. The Fourier inversion formula, that we will prove in a later lecture, will show that the Fourier transform in fact is an algebra isomorphism of $\mathsf{L}^1(\mathbb{R}^n)$ onto the Wiener algebra.

The differentiation rules

(1) Let $f \in \mathsf{L}^1(\mathbb{R}^n)$ and $\partial_j f \in \mathsf{L}^1(\mathbb{R}^n)$ for some $1 \leq j \leq n$. Then

$$\widehat{\partial_j f}(\xi) = \mathrm{i}\xi_j \widehat{f}(\xi).$$

(2) Let $f \in \mathsf{L}^1(\mathbb{R}^n)$ and $x_j f(x) \in \mathsf{L}^1(\mathbb{R}^n)$ for some $1 \leq j \leq n$. Then

$$\partial_j \widehat{f} = \mathcal{F}_{x \to \xi}\Big(-\mathrm{i} x_j f(x) \Big).$$

Furthermore the partial derivative $\partial_j \widehat{f}$ exists classically and is continuous.

Both rules admit generalizations: Let $p(x) \in \mathbb{C}[x]$.
(G1) Let $f \in \mathsf{L}^1(\mathbb{R}^n)$ and $p(\partial) f \in \mathsf{L}^1(\mathbb{R}^n)$. Then

$$\widehat{p(\partial) f}(\xi) = p(\mathrm{i}\xi)\widehat{f}(\xi).$$

(G2) Let $f \in \mathsf{L}^1(\mathbb{R}^n)$ and $p(-\mathrm{i}x) f(x) \in \mathsf{L}^1(\mathbb{R}^n)$. Then

$$\big(p(\partial)\widehat{f}\big)(\xi) = \mathcal{F}_{x \to \xi}\Big(p(-\mathrm{i}x) f(x) \Big).$$

Proof of (1). First note that

$$\widehat{\partial_j f}(\xi) = \int_{\mathbb{R}^n} \partial_j f(x) \mathrm{e}^{-\mathrm{i}\xi\cdot x} \, \mathrm{d}x.$$

Here $\partial_j f$ is a distributional partial derivative and $x \mapsto \mathrm{e}^{-\mathrm{i}\xi\cdot x}$ is *not* a test function. To overcome this we take $\chi = \rho * \mathbf{1}_{B_2(0)}$, so that $\chi \in \mathscr{D}(\mathbb{R}^n)$ with $\chi = 1$ on $B_1(0)$. Clearly also $\chi_k(x) := \chi(x/k)$ is a test function and $\chi_k = 1$ on $B_k(0)$. Now $x \mapsto \mathrm{e}^{-\mathrm{i}\xi\cdot x}\chi_k(x)$ is a test function. Hence using Lebesgue's dominated convergence theorem and the definition of distributional derivative we find:

$$\begin{aligned} \widehat{\partial_j f}(\xi) &= \lim_{k\to\infty} \int_{\mathbb{R}^n} \partial_j f(x) \mathrm{e}^{-\mathrm{i}\xi\cdot x} \chi_k(x) \, \mathrm{d}x \\ &= -\lim_{k\to\infty} \int_{\mathbb{R}^n} f(x) \partial_j \Big(\mathrm{e}^{-\mathrm{i}\xi\cdot x} \chi_k(x) \Big) \, \mathrm{d}x. \end{aligned}$$

Proof of (1) continued...

Next, note that as $k \to \infty$,

$$\partial_j\left(\mathrm{e}^{-\mathrm{i}\xi\cdot x}\chi_k(x)\right) = -\mathrm{i}\xi_j\mathrm{e}^{-\mathrm{i}\xi\cdot x}\chi\left(\frac{x}{k}\right) + \mathrm{e}^{-\mathrm{i}\xi\cdot x}(\partial_j\chi)\left(\frac{x}{k}\right)\frac{1}{k} \to -\mathrm{i}\xi_j\mathrm{e}^{-\mathrm{i}\xi\cdot x}$$

holds pointwise in $x \in \mathbb{R}^n$. Consequently, using Lebesgue's dominated convergence theorem once more we arrive at

$$\begin{aligned}\widehat{\partial_j f}(\xi) &= -\lim_{k\to\infty}\int_{\mathbb{R}^n} f(x)\partial_j\left(\mathrm{e}^{-\mathrm{i}\xi\cdot x}\chi_k(x)\right)\mathrm{d}x \\ &= \mathrm{i}\xi_j\int_{\mathbb{R}^n} f(x)\mathrm{e}^{-\mathrm{i}\xi\cdot x}\,\mathrm{d}x,\end{aligned}$$

concluding the proof. □

We refer to the lecture notes for the proof of (G1).

Proof of (2). We start by proving the last part of the statement. Fix $\xi \in \mathbb{R}^n$ and put for $h \in \mathbb{R}$, $\Delta_{he_j}\widehat{f}(\xi) := \widehat{f}(\xi + he_j) - \widehat{f}(\xi)$. For $h \neq 0$,

$$\frac{\Delta_{he_j}\widehat{f}(\xi)}{h} = \int_{\mathbb{R}^n} f(x)\frac{\mathrm{e}^{-\mathrm{i}(\xi+he_j)\cdot x} - \mathrm{e}^{-\mathrm{i}\xi\cdot x}}{h}\,\mathrm{d}x \tag{1}$$

and since, using the fundamental theorem of calculus,

$$\left| f(x)\frac{\mathrm{e}^{-\mathrm{i}(\xi+he_j)\cdot x} - \mathrm{e}^{-\mathrm{i}\xi\cdot x}}{h}\right| \leq \left|x_j f(x)\right|$$

and $x_j f(x) \in \mathrm{L}^1(\mathbb{R}^n)$ we can use Lebesgue's dominated convergence theorem, whereby

$$\begin{aligned}\lim_{h\to 0}\frac{\Delta_{he_j}\widehat{f}(\xi)}{h} &= \int_{\mathbb{R}^n} f(x)\left(-\mathrm{i}x_j\mathrm{e}^{-\mathrm{i}\xi\cdot x}\right)\mathrm{d}x \\ &= \mathcal{F}_{x\to\xi}\left(-\mathrm{i}x_j f(x)\right)\end{aligned}$$

Thus the partial derivative $\partial_j\widehat{f}$ exists classically at ξ. It follows from the Riemann-Lebesgue lemma that it is also continuous as a function of ξ.

Proof of (2) continued...

Finally we must show that the formula also holds distributionally. Recall from B4.3 that

$$\lim_{h\to 0} \frac{\Delta_{he_j}\widehat{f}}{h} = \partial_j \widehat{f} \text{ in } \mathscr{D}'(\mathbb{R}^n),$$

where now the right-hand side denotes the distributional partial derivative. Hence the left-hand side of (1) has the correct limit. What about the right-hand side? If we can show that the convergence is *locally uniform* in $\xi \in \mathbb{R}^n$ then we conclude the proof. In order to see that, we let $\xi_h \to \xi$ as $h \to 0$ and consider

$$\int_{\mathbb{R}^n} f(x) \frac{\mathrm{e}^{-\mathrm{i}(\xi_h + he_j)\cdot x} - \mathrm{e}^{-\mathrm{i}\xi\cdot x}}{h}\,\mathrm{d}x.$$

We proceed exactly as before to see that we can use Lebesgue's dominated convergence theorem to conclude that the limit, as $h \to 0$, is $\mathcal{F}_{x\to\xi}\big(-\mathrm{i}x_j f(x)\big)$. □

We refer to the lecture notes for the proof of (G2).

The issue of the adjoint identity revisited

Recall that the product rule implies that

$$\int_{\mathbb{R}^n} \widehat{\phi}\psi \, \mathrm{d}x = \int_{\mathbb{R}^n} \phi\widehat{\psi} \, \mathrm{d}x$$

holds for all ϕ, $\psi \in \mathscr{D}(\mathbb{R}^n)$. The issue here is that the Fourier transform of a test function in general is not a test function: $\widehat{\phi}$ is C^∞ but its support might not be compact. We will prove in the next lecture that it instead has the property: for any multi-indices α, $\beta \in \mathbb{N}_0^n$ we have

$$\xi^\alpha(\partial^\beta\widehat{\phi})(\xi) \to 0 \text{ as } |\xi| \to \infty.$$

We will define a class of test functions requiring this property instead of compact support.

Rapidly decreasing functions: A function $f : \mathbb{R}^n \to \mathbb{C}$ is *rapid decreasing* if for all $m \in \mathbb{N}_0$ there exist $r_m > 0$, $c_m > 0$ so

$$\left|f(x)\right| \leq \frac{c_m}{|x|^m} \text{ for all } |x| > r_m.$$

Remark If $f : \mathbb{R}^n \to \mathbb{C}$ is continuous, then f is rapidly decreasing if and only if

$$\sup_{x \in \mathbb{R}^n} |x|^m \left|f(x)\right| < \infty$$

holds for all $m \in \mathbb{N}_0$. We leave the details as an exercise.

Example The functions

$$\mathrm{e}^{-|x|}, \mathrm{e}^{-x^2} \text{ and } \mathrm{e}^{-x^2} \cos x$$

are rapidly decreasing, while

$$\tfrac{1}{1+x^{127}} \text{ and } \tfrac{1}{1+|x|^{\alpha}} \quad (\alpha > 0)$$

are not. Note also that *rapidly decreasing* does not mean the function need to be decreasing in the usual sense of that word.

Schwartz test functions and the Schwartz space

A function $\phi\colon \mathbb{R}^n \to \mathbb{C}$ is a *Schwartz test function* if

(i) $\phi \in \mathrm{C}^\infty(\mathbb{R}^n)$, and

(ii) all its partial derivatives are rapidly decreasing: for all multi-indices α, $\beta \in \mathbb{N}_0^n$

$$\sup_{x\in\mathbb{R}^n} \left| x^\alpha (\partial^\beta \phi)(x) \right| < \infty.$$

The *Schwartz space* is the set of such functions:

$$\mathscr{S}(\mathbb{R}^n) = \left\{ \phi \in \mathrm{C}^\infty(\mathbb{R}^n) : \partial^\alpha \phi \text{ rapidly decreasing for all } \alpha \in \mathbb{N}_0^n \right\}$$

(Laurent Schwartz 1940s)

Example The functions

$$\mathrm{e}^{-|x|^2} \text{ and } p(x)\mathrm{e}^{-|x|^2}$$

are Schwartz test functions (for any polynomial $p(x) \in \mathbb{C}[x]$). However, the functions

$$\mathrm{e}^{-|x|} \text{ and } \frac{1}{1+x^{127}}$$

are not. We show in the next lecture that $\widehat{\rho} \in \mathscr{S}(\mathbb{R}^n)$.

Proposition With the usual definitions of vector space operations and multiplication the Schwartz space $\mathscr{S}(\mathbb{R}^n)$ is a commutative algebra (without unit).

The only nontrivial issue is to show that $\phi\psi \in \mathscr{S}(\mathbb{R}^n)$ when ϕ, $\psi \in \mathscr{S}(\mathbb{R}^n)$. This is a consequence of the Leibniz rule–see the lecture notes for the details.

Some useful norms

Let $\phi \in \mathrm{C}^\infty(\mathbb{R}^n)$.

Definition For $\alpha,\ \beta \in \mathbb{N}_0^n$ we put

$$S_{\alpha,\beta}(\phi) := \sup_{x\in\mathbb{R}^n} \left| x^\alpha (\partial^\beta \phi)(x) \right|$$

and for $k,\ l \in \mathbb{N}_0$ put

$$\overline{S}_{k,l}(\phi) := \max\left\{ S_{\alpha,\beta}(\phi) :\ |\alpha| \leq k\,,\ |\beta| \leq l \right\}$$

Remark $S_{\alpha,\beta}$ and $\overline{S}_{k,l}$ are norms on $\mathscr{S}(\mathbb{R}^n)$. Note that

$$\begin{aligned} \mathscr{S}(\mathbb{R}^n) &= \left\{ \phi \in \mathscr{S}(\mathbb{R}^n) :\ S_{\alpha,\beta}(\phi) < \infty\, \forall\, \alpha,\ \beta \in \mathbb{N}_0^n \right\} \\ &= \left\{ \phi \in \mathscr{S}(\mathbb{R}^n) :\ \overline{S}_{k,l}(\phi) < \infty\, \forall\, k,\ l \in \mathbb{N}_0 \right\} \end{aligned}$$

As we will see already in the next lecture, many results about Schwartz test functions can be conveniently expressed in terms of these norms.