# B4.4 Fourier Analysis HT22

Lecture 1: Introduction and the Fourier transform of integrable functions

1. Why is the Fourier transform so efficient in connection with linear constant-coefficient PDEs?
2. Definition of the Fourier transform on $L^1(\mathbb{R}^n)$
3. The Riemann-Lebesgue Lemma
4. The product rule
5. The quest for an adjoint identity and the need for a modified frame work

The material corresponds to pp. 2-5 in the lecture notes and should be covered in Week 1.

## Intro: Fourier transform and linear constant-coefficient PDEs

The course is a continuation of B4.3 *Distribution Theory*. It is an analysis course with some applications to PDEs and, to a lesser extent, to analytic number theory.

• Why is the Fourier transform so useful in connection with linear constant-coefficient PDEs?

Consider a linear constant-coefficient PDE:

$$p(\partial)u = f \ \text{ in } \ \mathbb{R}^n$$

Here $f$ is given and we want to find $u$. At this stage we shall keep things vague and not specify where $f$ is and where we should be looking for $u$. The *linear constant-coefficient differential operator* $p(\partial)$ is represented conveniently in multi-index notation:

$$p(\partial) = \sum_{|\alpha| \leq d} c_\alpha \partial^\alpha$$

where the coefficients $c_\alpha \in \mathbb{C}$ and $d \in \mathbb{N}_0$.

## Intro: Fourier transform and linear constant-coefficient PDEs

We have seen it before in B4.3 but we emphasize the one-to-one correspondence again:

$$\text{differential operator } p(\partial) \leftrightarrow \text{ polynomial } p(x) \in \mathbb{C}[x]$$

If $c_\alpha \neq 0$ for some multi-index $\alpha$ of length $d$, then we say the differential operator $p(\partial)$ has order $d$ (and the corresponding polynomial $p(x)$ has degree $d$). Furthermore, $p(\partial)$ is homogeneous of order $d$ (and $p(x)$ homogeneous of degree $d$) if $c_\alpha \neq 0$ for some multi-index $\alpha$ of length $d$ and $c_\alpha = 0$ for all multi-indices $\alpha$ with $|\alpha| \neq d$.
Note that in all cases we always have $p(\partial)\colon X \to X$ when

$$X \in \left\{ \mathscr{D}(\mathbb{R}^n),\, \mathrm{C}^\infty(\mathbb{R}^n),\, \mathscr{D}'(\mathbb{R}^n),\, \mathscr{E}'(\mathbb{R}^n) \right\}$$

but that it *fails* when $p(\partial)$ has order $d > 0$ and $X = \mathrm{L}^p(\mathbb{R}^n)$. However, in that case we can replace the Lebesgue space by a corresponding Sobolev space and have then $p(\partial)\colon\ \mathrm{W}^{d,p}(\mathbb{R}^n) \to \mathrm{L}^p(\mathbb{R}^n)$ for any $p \in [1, \infty]$.

## Intro: Fourier transform and linear constant-coefficient PDEs

Given $f$, find $u$ so

$$p(\partial)u = f \text{ in } \mathbb{R}^n$$

Why use the Fourier transform for this?

Recall that in Linear Algebra it was often useful to analyze a linear map by finding its eigenvectors and eigenvalues. Sometimes it was even possible to *diagonalize* the linear map and in such cases it was easier to understand its action. Can we diagonalize $p(\partial)$?

Let $\lambda = \big(\lambda_1, \ldots, \lambda_n\big) \in \mathbb{C}^n$ and define

$$e_\lambda(x) = \mathrm{e}^{\lambda_1 x_1 + \ldots + \lambda_n x_n}, \, x \in \mathbb{R}^n.$$

Note that $e_\lambda \in \mathrm{C}^\infty(\mathbb{R}^n)$ and that for a multi-index $\alpha \in \mathbb{N}_0^n$, $\partial^\alpha e_\lambda = \lambda^\alpha e_\lambda$. Therefore also

$$p(\partial)e_\lambda = p(\lambda)e_\lambda$$

$e_\lambda$ is an eigenvector corresponding to the eigenvalue $p(\lambda)$ for $p(\partial)$

## Intro: Fourier transform and linear constant-coefficient PDEs

Note

- $e_\lambda$ is an eigenvector for all differential operators $p(\partial)$
- $e_\lambda$ is a bounded function if and only if $\lambda \in \mathrm{i}\mathbb{R}^n$.

The boundedness is useful for estimates and analysis!

Furthermore it can be shown that

$$e_{\mathrm{i}\xi} \colon \mathbb{R}^n \to \mathbb{S}^1 \subset \mathbb{C} \qquad (\xi \in \mathbb{R}^n)$$

are precisely the *continuous homomorphisms* from the additive group $\mathbb{R}^n$ to the unimodular multiplicative group $\mathbb{S}^1$ (they are called the *characters* of the additive group $\mathbb{R}^n$).

The Fourier transform allows us to write $f$ as an expansion in the characters $e_{\mathrm{i}\xi}$:

$$f(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} \widehat{f}(\xi) \mathrm{e}^{\mathrm{i}x\cdot\xi} \, \mathrm{d}\xi, \text{ where } \widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x) \mathrm{e}^{-\mathrm{i}x\cdot\xi} \, \mathrm{d}x.$$

Among other things we shall make this precise and give a number of applications.

## Intro: Fourier transform and linear constant-coefficient PDEs

Solution of the linear constant-coefficient PDE $p(\partial)u = f$:

$$u(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} \frac{\widehat{f}(\xi)}{p(\mathrm{i}\xi)} \mathrm{e}^{\mathrm{i}\xi \cdot x} \, \mathrm{d}\xi + h(x),$$

where $h(x)$ is a solution to the homogeneous equation. It is far from clear how this should be interpreted at this stage. (The formula for the particular solution is symbolic and formally follows from the differentiation rule and the Fourier inversion formula–both proved later in this course.)
However, just looking at the '*formula*' it becomes clear that it is desirable to be able to Fourier transform quite general functions and distributions. In this course we will do that in the spirit of distribution theory following Laurent Schwartz.

**Definition of the Fourier transform on $\mathrm{L}^1(\mathbb{R}^n)$:** Let $f \in \mathrm{L}^1(\mathbb{R}^n)$. The *Fourier transform* of $f$ is

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

Note

- $\widehat{f}(\xi)$ is well-defined as a Lebesgue integral for each $\xi \in \mathbb{R}^n$ since

  $$\left|f(x)\mathrm{e}^{-\mathrm{i}\xi\cdot x}\right| = |f(x)| \text{ is integrable over } \mathbb{R}^n$$

  with $|\widehat{f}(\xi)| \leq \|f\|_1$.
- In particular $\widehat{f}(0) = \int_{\mathbb{R}^n} f(x)\,\mathrm{d}x$ so when $f(x) \geq 0$ for all $x$, then

  $$\left|\widehat{f}(\xi)\right| \leq \|f\|_1 = \widehat{f}(0) \quad \forall\, \xi \in \mathbb{R}^n$$

- $\mathcal{F}\colon \mathrm{L}^1(\mathbb{R}^n) \to \mathrm{L}^\infty(\mathbb{R}^n)$ is a bounded linear map with bound

  $$\|\widehat{f}\|_\infty \leq \|f\|_1$$

## Example

Let $a < b$ be real numbers. Then $\mathbf{1}_{(a,b)} \in \mathsf{L}^1(\mathbb{R})$ and

$$\widehat{\mathbf{1}_{(a,b)}}(\xi) = \begin{cases} \frac{\mathrm{e}^{-\mathrm{i}a\xi} - \mathrm{e}^{-\mathrm{i}b\xi}}{\xi} & \text{if } \xi \neq 0 \\ b - a & \text{if } \xi = 0. \end{cases}$$

In particular, if $a = -1$, $b = 1$, then $\widehat{\mathbf{1}_{(a,b)}}(\xi) = 2\mathrm{sinc}(\xi)$ is the *sinus cardinalis* function

$$\mathrm{sinc}(\xi) = \begin{cases} \frac{\sin \xi}{\xi} & \text{if } \xi \neq 0 \\ 1 & \text{if } \xi = 0. \end{cases}$$

Recall that $\mathrm{sinc} \notin \mathsf{L}^1(\mathbb{R})$ and hence similarly $\widehat{\mathbf{1}_{(a,b)}} \notin \mathsf{L}^1(\mathbb{R})$. But it is not difficult to check that

$$\widehat{\mathbf{1}_{(a,b)}} \text{ is continuous and } \widehat{\mathbf{1}_{(a,b)}}(\xi) \to 0 \text{ as } |\xi| \to \infty$$

**Example continued...**

The previous calculation can easily be generalized to $n$ dimensions: Let $R = (a_1, b_1) \times \ \ldots \ \times (a_n, b_n)$ be a rectangle in $\mathbb{R}^n$.
Then $\mathbf{1}_R \in \mathrm{L}^1(\mathbb{R}^n)$ and by Fubini's theorem

$$\widehat{\mathbf{1}_R}(\xi) = \prod_{j=1}^{n} \widehat{\mathbf{1}_{(a_j, b_j)}}(\xi_j)$$

is continuous, tends to 0 as $|\xi| \to \infty$, but isn't integrable over $\mathbb{R}^n$.

## The space of continuous functions on $\mathbb{R}^n$ that vanish at infinity

$$\mathsf{C}_0(\mathbb{R}^n) = \big\{ f \in \mathsf{C}(\mathbb{R}^n) : \ f(x) \to 0 \ \text{ as } \ |x| \to \infty \big\}$$

Because the Fourier transform on $\mathsf{L}^1(\mathbb{R}^n)$ is linear the previous example implies that

$$\mathcal{F}\big(\mathsf{L}^{\mathrm{step}}(\mathbb{R}^n)\big) \subset \mathsf{C}_0(\mathbb{R}^n),$$

where we recall that the set of step functions on $\mathbb{R}^n$ is the linear span of indicator functions $\mathbf{1}_R$ of bounded rectangles $R$ (open, closed or neither and they can be empty or singletons but must be bounded).

**Example** $C_c(\mathbb{R}^n)$ is a subspace of $\mathrm{BC}(\mathbb{R}^n) = \big\{ f \in \mathsf{C}(\mathbb{R}^n) : \ f \ \text{ bounded} \big\}$. If we equip $\mathrm{BC}(\mathbb{R}^n)$ with the sup-norm $\|f\|_\infty = \sup_{x \in \mathbb{R}^n} |f(x)|$, then it is complete and the space $\mathsf{C}_c(\mathbb{R}^n)$ is *not closed* in $\mathrm{BC}(\mathbb{R}^n)$. Its closure is exactly $\mathsf{C}_0(\mathbb{R}^n)$ so the space is also complete in the sup-norm. The details are left as an exercise. Note that in the terminology of Functional Analysis we have that $\mathrm{BC}(\mathbb{R}^n)$ and $\mathsf{C}_0(\mathbb{R}^n)$ are Banach spaces in the sup-norm.

**The Riemann-Lebesgue Lemma:** Let $f \in \mathrm{L}^1(\mathbb{R}^n)$. Then $\widehat{f} \in \mathrm{C}_0(\mathbb{R}^n)$.

*Proof.* Because step functions are dense in $\mathrm{L}^1(\mathbb{R}^n)$ we can for each $j \in \mathbb{N}$ find $s_j \in \mathrm{L}^{\mathrm{step}}(\mathbb{R}^n)$ so $\|f - s_j\|_1 < \frac{1}{j}$. By the previous examples we have that $\widehat{s_j} \in \mathrm{C}_0(\mathbb{R}^n)$ and

$$\|\widehat{f} - \widehat{s_j}\|_\infty = \|\mathcal{F}(f - s_j)\|_\infty \leq \|f - s_j\|_1 < \tfrac{1}{j},$$

hence by the triangle inequality, $\|\widehat{s_j} - \widehat{s_k}\|_\infty \leq \frac{2}{j}$ for all $k > j$. Thus we have a sequence $(\widehat{s_j})$ in $\mathrm{C}_0(\mathbb{R}^n)$ that is Cauchy in the sup-norm, hence, by completeness, is convergent in the sup-norm to an element of $\mathrm{C}_0(\mathbb{R}^n)$. Clearly this element is $\widehat{f}$, concluding the proof. □

**Remark** The set $\mathcal{F}(\mathrm{L}^1(\mathbb{R}^n))$ is called the *Wiener algebra*. By the Riemann-Lebesgue lemma it is contained in $\mathrm{C}_0(\mathbb{R}^n)$. It follows from the convolution rule, proved later in the course, that it *is* an algebra (so a vector space that is closed under multiplication too). We will show on a problem sheet that it is a proper subset of $\mathrm{C}_0(\mathbb{R}^n)$.

**Example** Let $\rho$ be the standard mollifier kernel on $\mathbb{R}$. Then

$$\widehat{\rho}(\xi) = \int_{-1}^{1} \rho(x) \mathrm{e}^{-\mathrm{i}x\xi} \, \mathrm{d}x, \quad \xi \in \mathbb{R}.$$

Using Lebesgue's dominated convergence theorem and an induction argument it follows that $\widehat{\rho} \in \mathrm{C}^{\infty}(\mathbb{R})$. It is not very difficult to show that $\widehat{\rho}$ *does not* have a compact support. This is connected to the so-called *uncertainty principle* that we shall discuss later in the course. You will also be asked to prove it on a problem sheet. We will also show later, and in greater generality, that

$$\xi^{k} \widehat{\rho}^{(l)}(\xi) \to 0 \text{ as } |\xi| \to \infty$$

holds for all $k, l \in \mathbb{N}_0$.

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

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

holds.
The proof is straight forward by use of Fubini's theorem to swap integration orders.

Note that it looks like an adjoint identity! Indeed with $S = T = \mathcal{F}$ we have in particular for $\phi,\, \psi \in \mathscr{D}(\mathbb{R}^n)$ that

$$\int_{\mathbb{R}^n} S(\phi)\psi\,\mathrm{d}x = \int_{\mathbb{R}^n} \phi T(\psi)\,\mathrm{d}x.$$

But is it really an adjoint identity?

## A first failed attempt...

The product rule does not gives us an adjoint identity on $\mathscr{D}(\mathbb{R}^n)$

The trouble is that $\widehat{\psi} \notin \mathscr{D}(\mathbb{R}^n)$ for some $\psi \in \mathscr{D}(\mathbb{R}^n)$. Hence if we try to use the adjoint identity scheme we run into trouble for general distributions $u \in \mathscr{D}'(\mathbb{R}^n)$:

$$\langle \widehat{u}, \psi \rangle := \langle u, \widehat{\psi} \rangle$$

It is not clear how to make sense of the right-hand side at this stage!

However, we do not give up so easily and will instead define a *larger class of test functions* for which we can obtain an adjoint identity from the product rule. This will allow us to extend the Fourier transform to a *smaller class of distributions*. These are respectively the *Schwartz test functions* and the *tempered distributions* that we will develop and explore in this course.

However some extensions are possible without much additional work:

**1. Finite Borel measures:** Let $\mu$ be a finite Borel measure on $\mathbb{R}^n$, so a non-negative, finite $\sigma$-additive set function defined on the Borel $\sigma$-algebra of Borel sets in $\mathbb{R}^n$. Its Fourier transform is defined as

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

Note that the right-hand side is well-defined for each $\xi \in \mathbb{R}^n$ as a Lebesgue integral, and we have $|\widehat{\mu}(\xi)| \leq \mu(\mathbb{R}^n) = \widehat{\mu}(0)$ for all $\xi \in \mathbb{R}^n$.

Using Lebesgue's dominated convergence theorem it is not difficult to show that $\widehat{\mu}$ is uniformly continuous. We therefore have that $\widehat{\mu}$ is a bounded and uniformly continuous function on $\mathbb{R}^n$: $\widehat{\mu} \in \mathrm{BUC}(\mathbb{R}^n)$.

**Example** Note that $\widehat{\delta_0}(\xi) = 1$ for all $\xi \in \mathbb{R}^n$, so we see that

$$\widehat{\delta_0} \in \mathrm{BUC}(\mathbb{R}^n) \setminus \mathsf{C}_0(\mathbb{R}^n).$$

**2. Distributions of compact support:** Let $u \in \mathscr{E}'(\mathbb{R}^n)$. Then its Fourier transform is

$$\widehat{u}(\xi) = \langle u, e_{-\mathrm{i}\xi} \rangle , \quad \xi \in \mathbb{R}^n.$$

By the theorem about differentiation behind the distribution sign it follows that $\widehat{u} \in \mathrm{C}^\infty(\mathbb{R}^n)$. We shall see that this definition is consistent with the definition on $\mathrm{L}^1(\mathbb{R}^n)$ and also that in this case it is possible to extend $\widehat{u}(\xi)$ to all $\xi \in \mathbb{C}^n$ as a holomorphic function (that is, holomorphic as a function in each of the coordinates $\xi_j$ separately). We shall discuss that in connection with the theorem of Paley and Wiener.

**Example** Clearly $\delta_0 \in \mathscr{E}'(\mathbb{R}^n)$ too. In this case it is clear that $\widehat{\delta_0} \equiv 1$ can be extended to all $\xi \in \mathbb{C}^n$ as a holomorphic function.

## The Fourier transform of a nice function might be a distribution of higher order

**Example** In connection with the Fourier inversion formula we will show that if $p(x) \in \mathbb{C}[x]$ is a polynomial on $\mathbb{R}$, then

$$\widehat{p(x)} = 2\pi p \left( \mathrm{i} \frac{\mathrm{d}}{\mathrm{d}x} \right) \delta_0.$$

When we want to define the Fourier transform of, say bounded functions, we need more machinery. The same goes for $\mathrm{L}^p$ functions when $p > 1$. It turns out that the Fourier transforms of $\mathrm{L}^p$ functions are regular distributions when $p \in [1, 2]$, but that they need not be when $p > 2$. We highlight in particular the case $p = 2$ that is particularly nice and that we shall investigate carefully when we prove Plancherel's theorem.