Chapter 2

Fourier Transform

The Fourier transform is a fundamental tool, in particular for the study of partial derivative equations, because it "diagonalizes" (in a sense to be made precise) the differential operators.

Let us begin by reminding the motivation for diagonalization in finite dimension. Let E be a $\mathbb{C}$-vector space of finite dimension $n \in \mathbb{N}$. Let A be an endomorphism of E and $u_0 \in E$. We consider the following problem on E

$$\begin{cases} u'(t) = Au(t), & \forall t \in \mathbb{R}, \\ u(0) = u_0, \end{cases} \tag{2.1}$$

where the unknown u is a function (say of class C^1) from $\mathbb{R}$ to E. The favorable case is the case where u_0 is an eigenvector of A, associated to an eigenvalue $\lambda \in \mathbb{C}$. In that case, It is easy to see that we get a solution by setting, for any $t \in \mathbb{R}$,

$$u(t) = e^{t\lambda} u_0.$$

Indeed, the function u thus defined is of class C^∞ from $\mathbb{R}$ to E, we have $u(0) = u_0$ and, for any $t \in \mathbb{R}$,

$$u'(t) = e^{t\lambda} \lambda u_0 = e^{t\lambda} A u_0 = Au(t).$$

Since the equation $u' = Au$ is linear with respect to u, we get that if u_0 is a linear combination of eigenvectors, that is if

$$u_0 = \sum_{j=1}^{k} \alpha_j e_j,$$

where $k \leqslant n$, $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ and for any $j \in [\![1, k]\!]$ the vector e_j is an eigenvector A associated to an eigenvalue $\lambda_j \in \mathbb{C}$, then a solution of (2.1) is given by

$$u(t) = \sum_{j=1}^{k} e^{t\lambda_j} \alpha_j e_j.$$

The interest of results about diagonalization is to give criteria ensuring that any vector u_0 of E can be written as a linear combination of eigenvectors of A. Thus, if A is diagonalizable and if we can determine its eigenvalues and eigenvectors, we can *easily* solve (2.1) *for any initial data* u_0.

It turns out that many models from concrete problems can be written in the form (2.1), but in a functionnal space E of infinite dmension (typically $L^2(\mathbb{R}^d)$), and a linear map A that is a differential operator. An important example (for example for the heat equation, but also for Schrödinger equation or the wave equation) is the Laplace operator

$$A = \Delta = \sum_{k=1}^{d} \frac{\partial^2}{\partial x_k^2}.$$

For example, the heat equation on $\mathbb{R}^d$ can be written as

$$\partial_t u(t,x) = \Delta u(t,x), \quad \forall t \geqslant 0, \forall x \in \mathbb{R}^d.$$

The analysis is obviously more complicated in infinite dimension, so it is even more important to use an appropriate setting.

To simplify the rest of the discussion, we assume that $d = 1$. The simplest differential operator on R is

$$Av(x) = v'(x).$$

We look for eigenvectors of A, that is functions v such that $v' = \lambda v$ for some $\lambda \in \mathbb{C}$. The candidates are obviously the exponential functions $x \mapsto e^{\lambda x}$. These functions appears in the Fourier series theory, that have precisely been introduced to solve the heat equation, but on a bounded interval. According to the Fourier series theory, if we set

$$e_n : x \mapsto e^{\frac{2i\pi n x}{T}}, \quad \text{for } n \in \mathbb{Z},$$

then the family $(e_n)_{n\in\mathbb{Z}}$ is a *Hilbert basis* of the Hilbert space $L^2_T(\mathbb{R})$ of locally integrable and T-periodic functions from $\mathbb{R}$ to $\mathbb{C}$, endowed with the norm defined by

$$\|f\|^2_{L^2_T(\mathbb{R})} = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} |f(x)|^2 \, dx.$$

Thus for any $f \in L^2_T(\mathbb{R})$ we have

$$f = \sum_{n\in\mathbb{Z}} c_n(f) e_n \quad \text{where} \quad c_n(f) = \langle f, e_n \rangle = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} e^{-\frac{2i\pi n y}{T}} f(y) \, dy.$$

In addition, each e_n is an eigenvector for the derivative operator. If f is regular we have

$$f' = \sum_{n\in\mathbb{Z}} c_n(f) e_n' = \sum_{n\in\mathbb{Z}} \frac{2i\pi n}{T} c_n(f) e_n.$$

This holds for T-periodic functions, but we can do the same for the functions defined on an interval of length T.

Here, we are interested in functions that are not periodic and are defined on all $\mathbb{R}$. Formally, we want to see what happens if we let T go to $+\infty$ in the previous expressions. Observing that for $\phi \in C_0^\infty(\mathbb{R})$ we have

$$\frac{2\pi}{T} \sum_{n\in\mathbb{Z}} \phi\left(\frac{2\pi n}{T}\right) \xrightarrow[T\to+\infty]{} \int_{\mathbb{R}} \phi(\xi) \, d\xi \tag{2.2}$$

(this is essentially a Riemann sum), we want to write $f(x)$ as

$$\frac{1}{2\pi} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{-iy\xi} f(y) \, dy \right) e^{ix\xi} \, d\xi. \tag{2.3}$$

If this has a meaning, and if it really gives an expression for $f(x)$, then we have written f as an integral (that can be seen as the limit of a series, see (2.2), or as a "continuous sum") of functions of the form $c_\xi(f)e^{ix\xi}$, that are "eigenvectors" (it would be necessary to specify the space in which one works) associated to the derivative operator. In addition the corresponding coefficient $c_\xi(f)$ can be written as a scalar product in $L^2(\mathbb{R})$ of f with the conjugate function $x \mapsto \overline{e^{ix\xi}} = e^{-ix\xi}$. Except that the function $x \mapsto e^{ix\xi}$ is not in $L^2(\mathbb{R})$...

Thus, the definition of the Fourier transform on $\mathbb{R}$ (or $\mathbb{R}^d$ for any $d \in \mathbb{N}^*$) is motivated and guided by good properties of the Fourier series on a bounded interval, but to get similar results in this setting, a specific analysis is necessary. This is the purpose of this chapter.

2.1 The Schwartz space

In this chapter we will rely on Schwartz functions.

Definition 2.1. We note $\mathcal{S}(\mathbb{R}^d)$ and we call Schwartz space the set of functions ϕ of class C^∞ on $\mathbb{R}^d$ such that for any $\alpha, \beta \in \mathbb{N}^d$ there exists $C_{\alpha,\beta} > 0$ for which

$$\forall x \in \mathbb{R}^d, \quad \left| x^\alpha \partial^\beta \phi(x) \right| \leqslant C_{\alpha,\beta}.$$

We say that every derivative of ϕ is a rapidly decreasing function.

Examples 2.2.

- We have $C_0^\infty(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d)$.
- The map $x \mapsto e^{-|x|^2}$ is in $\mathcal{S}(\mathbb{R}^d)$ but not in $C_0^\infty(\mathbb{R}^d)$.

Definition 2.3. We say that $f : \mathbb{R}^d \to \mathbb{C}$ is a slowly increasing function if there exist $N \in \mathbb{N}$ and $C > 0$ such that

$$\forall x \in \mathbb{R}^d, \quad |f(x)| \leqslant C(1 + |x|)^N.$$

Proposition 2.4. *$\mathcal{S}(\mathbb{R}^d)$ is a vector space, stable under the usual product, under convolution, under derivation and multiplication by a function of class C^∞ whose derivatives are slowly increasing.*

Proof. Let us prove that f and g are Schwartz functions then their convolution $(f * g)$ is also in $\mathcal{S}(\mathbb{R}^d)$. Let $k \in \mathbb{N}$ and $x \in \mathbb{R}^d$. Let $y \in \mathbb{R}^d$. If $|x| \geqslant 2\,|y|$ we have

$$|x - y| \geqslant |x| - |y| \geqslant \frac{|x|}{2},$$

so

$$\begin{aligned} |x|^k \, |(f * g)(x)| &\leqslant |x|^k \int_{y \in \mathbb{R}^d} |f(x-y)| \, |g(y)| \, dy \\ &\leqslant \int_{|y| \leqslant \frac{|x|}{2}} 2^k \, |x-y|^k \, |f(x-y)| \, |g(y)| \, dy + \int_{|y| \geqslant \frac{|x|}{2}} |f(x-y)| \, 2^k \, |y|^k \, |g(y)| \, dy. \end{aligned}$$

This proves that f and g are rapidly decreasing functions, then $(f * g)$ is also a rapidly decreasing function.

By derivation under the integral sign, we see that for $f, g \in \mathcal{S}(\mathbb{R}^d)$ the convolution $(f * g)$ is of class C^∞ on $\mathbb{R}^d$ and for $\beta \in \mathbb{N}^d$ we have $\partial^\beta(f * g) = (\partial^\beta f) * g$. The previous result applied to $\partial^\beta f$ and g ensures that the derivatives of $(f * g)$ are rapidly decreasing, and so $(f * g) \in \mathcal{S}(\mathbb{R}^d)$. ∎

Since we will also work in different Lebesgue spaces, we give some useful links between these different spaces.

Proposition 2.5. (i) *For any* $p \in [1, +\infty]$ *we have* $\mathcal{S}(\mathbb{R}^d) \subset L^p(\mathbb{R}^d)$.

(ii) *For* $p \in [1, +\infty[$, $\mathcal{S}(\mathbb{R}^d)$ *is dense in* $L^p(\mathbb{R}^d)$. *More precisely, for* $p, q \in [1, +\infty[$ *and* $f \in L^p(\mathbb{R}^d) \cap L^q(\mathbb{R}^d)$ *there exists a sequence* $(f_n)_{n\in\mathbb{N}}$ *of Schwartz functions that converges to* f *in both* $L^p(\mathbb{R}^d)$ *and* $L^q(\mathbb{R}^d)$.

The first point is easy, for the second one we use the fact that these properties of density are valid for $C_0^\infty(\mathbb{R}^d)$, which is included in $\mathcal{S}(\mathbb{R}^d)$.

We finish with the property of integration by parts on $\mathbb{R}^d$. The fact that Schwartz functions ar rapidly decreasing at infinity ensures that there is no difficulty with the boundary terms to show the following result.

Proposition 2.6. *Let* $u \in C^1(\mathbb{R}^d)$ *be a slowly increasing function whose first order partial derivatives are also slowly increasing. Let* $v \in \mathcal{S}(\mathbb{R}^d)$. *Then for* $j \in [\![1, d]\!]$ *the functions* $(\partial_j u)v$ *and* $u(\partial_j v)$ *are integrable on* $\mathbb{R}^d$ *and we have*

$$\int_{\mathbb{R}^d} (\partial_j u)v \, dx = -\int_{\mathbb{R}^d} u(\partial_j v) \, dx.$$

2.2 Fourier transform in L^1

We begin by defining what will play the role of “Fourier coefficient” for a non periodic function. It is in $L^1(\mathbb{R}^d)$ that the following definition has a natural meaning.

Definition 2.7. Let $f \in L^1(\mathbb{R}^d)$. For $\xi \in \mathbb{R}^d$ we set

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} e^{-ix\cdot\xi} f(x) \, dx. \tag{2.4}$$

This defines a function $\hat{f}$ on $\mathbb{R}^d$ which is called the *Fourier transform* of f. It is also denoted by $\mathcal{F}f$.

This definition is licit since the function $x \mapsto e^{-ix\cdot\xi} f(x)$ is integrable for any $\xi \in \mathbb{R}^d$. Moreover, the definition of $\hat{f}(\xi)$ does not depend of the choice of the representative of f (in other words, the definition does not change if we replace f by a function equal to f almost everywhere). In the sequel we will no longer discuss the distinction between functions of $\mathcal{L}^1(\mathbb{R}^d)$ or equivalence classes modulo almost everywhere equality in $L^1(\mathbb{R}^d)$.

Notice that even if f is only defined for an equivalence class of function, $\hat{f}(\xi)$ is well defined for each fixed. Notice also that $\hat{f}(0)$ is just the integral of f:

$$\hat{f}(0) = \int_{\mathbb{R}^d} f(x) \, dx.$$

Example 2.8. Let $a > 0$. We consider on $\mathbb{R}$ the function $f : x \mapsto \frac{1}{2a}\mathbb{1}_{[-a,a]}$. Then $f \in L^1(\mathbb{R})$, $\hat{f}(0) = 1$ and for $\xi \in \mathbb{R}^*$ we have

$$\hat{f}(\xi) = \frac{1}{2a}\int_{-a}^{a} e^{-ix\xi} \, dx = \frac{1}{2a}\frac{e^{-ia\xi} - e^{ia\xi}}{-i\xi} = \frac{\sin(a\xi)}{a\xi}.$$

Example 2.9. Let $a > 0$. For $x \in \mathbb{R}$ we set $f(x) = e^{-a|x|}$. Then $f \in L^1(\mathbb{R})$ and for $\xi \in \mathbb{R}$ we have (exercise)

$$\hat{f}(\xi) = \frac{2a}{a^2 + \xi^2}.$$

Remark 2.10. Different conventions co-exist for the definition of $\hat{f}$. We can also set

$$\hat{f}(\xi) = \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} e^{-ix\cdot\xi} f(x)\, dx.$$

The advantage of this definition is that the expression of the Fourier transform and its inverse transform (see Proposition 2.14) are symetrical, and there is no 2π factor in the Parseval identity (see Proposition 2.21). However a $(2\pi)^{\frac{d}{2}}$ factor appears for example for the convolution (see Proposition 2.13). Another convention is to set

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} e^{-2i\pi x\cdot\xi} f(x)\, dx.$$

None of these choices is better than the others, if we remove the 2π factor somewhere it will appear somewhere else in the theory. Thus, when we use the Fourier transform, it is better to recall which definition is used and to be consistent afterwards (even if the $(2\pi)^d$ factor is in general harmless for the studied properties).

We now give a number of basic properties for the Fourier transform.

Proposition 2.11. *(i) For $f \in L^1(\mathbb{R}^d)$, $\hat{f}$ is bounded and $\|\hat{f}\|_\infty \leqslant \|f\|_1$.*

(ii) The map $f \mapsto \hat{f}$ is linear from $L^1(\mathbb{R}^d)$ to $L^\infty(\mathbb{R}^d)$.

(iii) For $f \in L^1(\mathbb{R}^d)$, $\hat{f}$ is continuous on $\mathbb{R}^d$ and tends to 0 when $|\xi|$ tends to $+\infty$.

Proof. The first two properties result from the triangular inequality and the linearity of the integral. The continuity of $\hat{f}$ is given by continuity under the integral sign.

Let us prove that $\hat{f}$ tends to 0 when $|\xi|$ tends to $+\infty$. We first assume that f is in $\mathcal{S}(\mathbb{R}^d)$. Let $j \in [\![1, d]\!]$. Let $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^d$ such that $\xi_j \neq 0$. By performing an integration by part, we get

$$\hat{f}(\xi) = \frac{1}{i\xi_j} \int_{\mathbb{R}^d} -\partial_{x_j} e^{-ix\cdot\xi} f(x)\, dx = \frac{1}{i\xi_j} \int_{\mathbb{R}^d} e^{-ix\cdot\xi} \partial_{x_j} f(x)\, dx = \frac{1}{i\xi_j} \widehat{\partial_{x_j} f}(\xi).$$

Since $\partial_{x_j} f \in L^1(\mathbb{R}^d)$ we have $\widehat{\partial_{x_j} f} \in L^\infty(\mathbb{R}^d)$, and so $\hat{f}(\xi)$ tends to 0 when $|\xi_j|$ tends to $+\infty$. This applies to every j, so we get the result for $f \in \mathcal{S}(\mathbb{R}^d)$.

Now let us consider the general case. Let $f \in L^1(\mathbb{R}^d)$. Let $\varepsilon > 0$. Since $\mathcal{S}(\mathbb{R}^d)$ is dense in $L^1(\mathbb{R}^d)$, there exists $g \in \mathcal{S}(\mathbb{R}^d)$ such that $\|f - g\|_{L^1} \leqslant \frac{\varepsilon}{2}$. Then there exists $R > 0$ such that for $|\xi| \geqslant R$ we have $|\hat{g}(\xi)| \leqslant \frac{\varepsilon}{2}$. For $|\xi| \geqslant R$ we have

$$|\hat{f}(\xi)| \leqslant |\hat{f}(\xi) - \hat{g}(\xi)| + |\hat{g}(\xi)| \leqslant \|f - g\|_{L^1} + \frac{\varepsilon}{2} \leqslant \varepsilon.$$

This proves that $\hat{f}(\xi)$ tends to 0 when $|\xi|$ tends to $+\infty$. □

The proof of the following properties are left as exercises for the reader:

Proposition 2.12. *Let $f \in L^1(\mathbb{R}^d)$.*

(i) *If f is an even (respectively odd) function, then $\hat{f}$ is even (respectively odd). More generally, if we note $\mathcal{P}$ the operator that associates $\mathcal{P}f : x \mapsto f(-x)$ to f, then for $f \in L^1(\mathbb{R}^d)$ we have*

$$\mathcal{P}\mathcal{F}f = \mathcal{F}\mathcal{P}f. \tag{2.5}$$

(ii) *If f is a real valued function $\hat{f}(-\xi) = \overline{\hat{f}(\xi)}$ for any $\xi \in \mathbb{R}^d$.*

(iii) *Let* $x_0 \in \mathbb{R}^d$ *et* $g : x \mapsto f(x - x_0)$*. Then* $\hat{g}(\xi) = \hat{f}(\xi)e^{-ix_0\cdot\xi}$ *for any* $\xi \in \mathbb{R}^d$*.*
Let $\xi_0 \in \mathbb{R}^d$ *and* $g : x \mapsto e^{ix\cdot\xi_0}f(x)$*. Then* $\hat{g}(\xi) = \hat{f}(\xi - \xi_0)$ *for any* $\xi \in \mathbb{R}^d$*.*

(iv) *Let* $\alpha > 0$ *and* $g : x \mapsto f(\alpha x)$*. Then* $\hat{g}(\xi) = \frac{1}{\alpha^d}\hat{f}\left(\frac{\xi}{\alpha}\right)$ *for any* $\xi \in \mathbb{R}^d$*.*

A remarkable property (and one that will play an important role in applications) is that the Fourier transform of a convolution product is the usual product of Fourier transforms.

Proposition 2.13. *Let* $f, g \in L^1(\mathbb{R}^d)$*. Then we have*

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

Proof. Let $\xi \in \mathbb{R}^d$. The map $(x, y) \mapsto e^{-ix\cdot\xi}f(x-y)g(y)$ is measurable and using Fubini-Tonelli theorem and convolution properties of $L^1(\mathbb{R}^d)$ functions we have

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} \left|e^{-ix\cdot\xi}f(x-y)g(y)\right| dx\,dy \leqslant \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} |f(x-y)|\,|g(y)|\,dy\right) dx \leqslant \|f\|_1\,\|g\|_1 < +\infty.$$

Using the Fubini-Lebesgue theorem, we then have

$$\begin{aligned}
\widehat{f * g}(\xi) &= \int_{\mathbb{R}^d} e^{-ix\cdot\xi}(f * g)(\xi)\,dx \\
&= \int_{\mathbb{R}^d} e^{-ix\cdot\xi}\left(\int_{\mathbb{R}^d} f(x-y)g(y)\,dy\right) dx \\
&= \int_{\mathbb{R}^d} e^{-iy\cdot\xi}g(y)\left(\int_{\mathbb{R}^d} e^{-i(x-y)\cdot\xi}f(x-y)\,dx\right) dy \\
&= \hat{f}(\xi)\hat{g}(\xi).
\end{aligned}$$

□

2.3 Inversion formula

The property that makes the Fourier transform usable in practice is that we can recover a function if we know its Fourier transform.

Proposition 2.14. *Let* $f \in L^1(\mathbb{R}^d)$ *such that* $\hat{f} \in L^1(\mathbb{R}^d)$*. Then for almost every* $x \in \mathbb{R}^d$ *we have*

$$f(x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{ix\cdot\xi}\hat{f}(\xi)\,d\xi.$$

We observe that the formula giving the inverse Fourier transform is almost the same as the one giving the Fourier transform itself (only the sign in the exponential changes, and with our choice for the definition of $\mathcal{F}$ there is an extra factor $(2\pi)^{-d}$). For $g \in L^1(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$ we set

$$\overline{\mathcal{F}}g(x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{ix\cdot\xi}g(\xi)\,d\xi = \frac{1}{(2\pi)^d}(\mathcal{F}g)(-x).$$

Using the notation $\mathcal{P}$ introduced in proposition 2.12 this can be rewritten as

$$\overline{\mathcal{F}} = \frac{1}{(2\pi)^d}\mathcal{P}\mathcal{F} = \frac{1}{(2\pi)^d}\mathcal{F}\mathcal{P}.$$

Thus we can write

$$f = \overline{\mathcal{F}}\mathcal{F}f = \frac{1}{(2\pi)^d}\mathcal{P}\mathcal{F}\mathcal{F}f = \frac{1}{(2\pi)^d}\mathcal{F}\mathcal{F}\mathcal{P}f = \mathcal{F}\overline{\mathcal{F}}f. \tag{2.6}$$

Proof. • For $\varepsilon > 0$ and $x \in \mathbb{R}^d$ we set

$$F(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{ix\cdot\xi} \hat{f}(\xi)\, d\xi \quad \text{and} \quad F_\varepsilon(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{ix\cdot\xi} e^{-\varepsilon(|\xi_1|+\cdots+|\xi_d|)} \hat{f}(\xi)\, d\xi.$$

By continuity under the integral sign, these two functions are continuous on $\mathbb{R}^d$. Moreover, by the dominated convergence Theorem we have, for all $x \in \mathbb{R}^d$,

$$F_\varepsilon(x) \xrightarrow[\varepsilon\to 0]{} F(x).$$

- Let $\varepsilon > 0$ and $x \in \mathbb{R}^d$. The function

$$(y, \xi) \mapsto e^{i(x-y)\cdot\xi} e^{-\varepsilon(|\xi_1|+\cdots+|\xi_d|)} f(y)$$

is integrable on $\mathbb{R}^d \times \mathbb{R}^d$, so by the Fubini-Lebesgue Theorem we have

$$\begin{aligned} F_\varepsilon(x) &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{ix\cdot\xi} e^{-\varepsilon(|\xi_1|+\cdots+|\xi_d|)} \left(\int_{\mathbb{R}^d} e^{-iy\cdot\xi} f(y)\, dy \right) d\xi \\ &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \left(\prod_{j=1}^d \int_{\mathbb{R}} e^{i(x_j-y_j)\xi_j} e^{-\varepsilon|\xi_j|}\, d\xi_j \right) f(y)\, dy. \end{aligned}$$

Let $j \in [\![1, d]\!]$. The computation of Example 2.9 gives

$$\int_{\mathbb{R}} e^{i(x_j-y_j)\xi_j} e^{-\varepsilon|\xi_j|}\, d\xi_j = \frac{2\varepsilon}{\varepsilon^2 + (x_j - y_j)^2}.$$

For $s = (s_1, \ldots, s_j) \in \mathbb{R}^d$ we have

$$\chi(s) = \prod_{j=1}^d \frac{1}{\pi(1+s_j^2)} \quad \text{and} \quad \chi_\varepsilon(s) = \frac{1}{\varepsilon^d} \chi\left(\frac{s}{\varepsilon}\right).$$

Then

$$F_\varepsilon(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \frac{(2\varepsilon)^d f(y)}{\prod_{j=1}^d (\varepsilon^2 + (x_j - y_j)^2)}\, dy = (\chi_\varepsilon * f)(x).$$

Since $(\chi_\varepsilon)_{\varepsilon\in]0,1]}$ defines an approximation of unity, we have $F_\varepsilon \in L^1(\mathbb{R}^d)$ and

$$\|F_\varepsilon - f\|_{L^1} \xrightarrow[\varepsilon\to 0]{} 0.$$

In particular, there exists a sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ that tends to 0 and such that F_{ε_n} tends to f almost everywhere. This proves that $F = f$ almost everywhere. □

Thus we have proved that under the hypotheses of the proposition 2.14, the expression obtained formally in (2.3) holds for almost any $x \in \mathbb{R}^d$.

We observe that Proposition 2.14 gives injectivity for the Fourier transform in $L^1(\mathbb{R}^d)$.

Corollary 2.15. *If $f \in L^1(\mathbb{R}^d)$ is such that $\hat{f} = 0$ almost everywhere, then $f = 0$.*

Proof. If $\hat{f} = 0$ then in particular $\hat{f} \in L^1(\mathbb{R}^d)$. So we can apply Proposition 2.14, which shows that $f(x) = 0$ for almost every $x \in \mathbb{R}^d$. □

2.4 Derivation and multiplication by a polynomial - Fourier transform in $\mathcal{S}$

We now discuss the property for which the Fourier transform has been introduced, namely the good behavior with respect to differential operators. We work with functions $x \mapsto e^{ix\cdot\xi}$ because computing their derivatives is the same as multiplying them by a scalar. Thus, if

$$f(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{ix\cdot\xi} \hat{f}(\xi)\, d\xi,$$

we expect

$$\partial_{x_j} f(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} i\xi_j e^{ix\cdot\xi} \hat{f}(\xi)\, d\xi.$$

The computations we made for the proof of the proposition 2.11 gives the following property.

Proposition 2.16. *Let $f \in \mathcal{S}(\mathbb{R}^d)$. Let $j \in [\![1, d]\!]$. Then for $\xi \in \mathbb{R}$ we have*

$$\widehat{\partial_{x_j} f}(\xi) = i\xi_j \hat{f}(\xi)$$

In particular, the map $\xi \mapsto \xi_j \hat{f}(\xi)$ is in $L^\infty(\mathbb{R}^d)$.

This property can in fact be extended to any function $f \in L^1(\mathbb{R}^d) \cap C^1(\mathbb{R}^d)$ such that the partial derivative $\partial_{x_j} f$ is also in $L^1(\mathbb{R}^d)$.

Since the inverse Fourier transform has an expression analogous to the Fourier transform itself, it is not surprising to find that we actually have an analogous property in the other direction. That is, the Fourier transform changes the multiplication by the variable x_j into a derivative with respect to ξ_j:

Proposition 2.17. *Let $f \in \mathcal{S}(\mathbb{R}^d)$. Let $j \in [\![1, d]\!]$. Then $\hat{f}$ is derivable with respect to ξ_j and for any $\xi \in \mathbb{R}^d$ we have*

$$\widehat{x_j f}(\xi) = i\partial_{\xi_j} \hat{f}(\xi).$$

Proof. For $(x, \xi) \in \mathbb{R}^d \times \mathbb{R}^d$ we set $\varphi(x, \xi) = e^{-ix\cdot\xi} f(x)$. Then φ is differentiable with respect to ξ_j and for any $(x, \xi) \in (\mathbb{R}^d)^2$ we have

$$\left| \frac{\partial \varphi}{\partial \xi_j}(x, \xi) \right| = \left| -ix_j e^{-ix\xi} f(x) \right| = |x_j f(x)|.$$

Since $x \mapsto |x_j f(x)|$ is integrable on $\mathbb{R}^d$, we get by differentiation under the integral sign that $\hat{f}$ is differentiable with respect to ξ_j in $\mathbb{R}^d$ and, for $\xi \in \mathbb{R}^d$,

$$i\partial_{\xi_j} \hat{f}(\xi) = \int_{\mathbb{R}^d} e^{-ix\cdot\xi} x_j f(x)\, dx = \widehat{x_j f}(\xi).$$ □

Notice that this proof holds for any function $f \in L^1(\mathbb{R}^d)$ such that $x \mapsto x_j f(x)$ is integrable.

The Schwartz space is the space of functions for which one can derive and multiply by the variable as many times as one wants. Moreover, these two operations play symmetric roles with respect to the Fourier transform. With the two previous propositions, everything indicates that it is a space in which the Fourier transform is particularly convenient.

Proposition 2.18. *The Fourier transform $\mathcal{F}$ defines a bijection from $\mathcal{S}(\mathbb{R}^d)$ to itself, and its inverse function is the restiction of $\overline{\mathcal{F}}$ to $\mathcal{S}(\mathbb{R}^d)$. Moreover, for $f \in \mathcal{S}(\mathbb{R}^d)$, $\alpha, \beta \in \mathbb{N}^d$ and $\xi \in \mathbb{R}^d$ we have*

$$\xi^\alpha \partial_\xi^\beta \mathcal{F} f(\xi) = (-i)^{|\alpha|+|\beta|} \mathcal{F}(\partial_x^\alpha x^\beta f)(\xi). \tag{2.7}$$

Proof. Let $f \in \mathcal{S}(\mathbb{R}^d)$ and $\alpha, \beta \in \mathbb{N}^d$. Applying $|\beta|$ times Proposition 2.17, we get that the partial derivative $\partial_\xi^\beta \hat{f}$ exists and

$$\widehat{x^\beta f}(\xi) = i^{|\beta|} \partial_\xi^\beta \hat{f}(\xi).$$

Now applying Proposition 2.16 to the successive derivatives of $x^\beta f$ we get

$$\widehat{\partial_x^\alpha x^\beta f}(\xi) = i^{|\alpha|+|\beta|} \xi^\alpha \partial_\xi^\beta \hat{f}(\xi).$$

In particular, the function $\xi \mapsto \xi^\alpha \partial^\beta \hat{f}(\xi)$ is bounded. this ensures that $\hat{f} \in \mathcal{S}(\mathbb{R}^d)$ and proves (2.7). The restriction $\mathcal{F}_\mathcal{S}$ of $\mathcal{F}$ to $\mathcal{S}(\mathbb{R}^d)$ is then injective by the corollary 2.15. Finally, according to (2.6) we have $f = \frac{1}{(2\pi)^{2d}} \mathcal{F}\mathcal{F}\mathcal{P} f$, so $\mathcal{F}_\mathcal{S}$ is surjective. Thus $\mathcal{F}_\mathcal{S}$ is a bijection in $\mathcal{S}(\mathbb{R}^d)$, and its inverse function is the restriction of $\overline{\mathcal{F}}$ to $\mathcal{S}(\mathbb{R}^d)$. □

Example 2.19. Let $f \in \mathcal{S}(\mathbb{R}^d)$. Then the Fourier transform of $-\Delta f$ is $\xi \mapsto |\xi|^2 \hat{f}(\xi)$.

Example 2.20. We want to compute on $\mathbb{R}$ the Fourier transform of the Gaussian $f : x \mapsto e^{-\frac{x^2}{2}}$. Notice that $f \in \mathcal{S}(\mathbb{R})$. In particular $\hat{f} \in \mathcal{S}(\mathbb{R})$. Since $f'(x) = -x f(x)$ for any $x \in \mathbb{R}$, we have that for $\xi \in \mathbb{R}$

$$\hat{f}'(\xi) = -i\widehat{xf}(\xi) = i\widehat{f'}(\xi) = -\xi \hat{f}(\xi).$$

This proves that

$$\hat{f}(\xi) = e^{-\frac{\xi^2}{2}} \hat{f}(0).$$

But we know that

$$\hat{f}(0) = \int_\mathbb{R} e^{-\frac{x^2}{2}} \, dx = \sqrt{2\pi},$$

So for any $\xi \in \mathbb{R}$ we have

$$\hat{f}(\xi) = \sqrt{2\pi} e^{-\frac{\xi^2}{2}}.$$

More generally, for $\sigma > 0$, if we consider $f_\sigma : x \mapsto e^{-\frac{x^2}{2\sigma^2}}$ then for $\xi \in \mathbb{R}$ we have

$$\widehat{f_\sigma}(\xi) = \sigma\sqrt{2\pi} e^{-\frac{\sigma^2 \xi^2}{2}}.$$

2.5 Fourier transform in L^2

So far we have discussed the properties of the Fourier transform in $L^1(\mathbb{R}^d)$ (because it is the space in which the definition naturally makes sense) or of its restriction to $\mathcal{S}(\mathbb{R}^d)$ (where it is the most comfortable). But our favorite space in the analysis of partial differential equations will be $L^2(\mathbb{R}^d)$, because it is a Hilbert space. The problem is that a function in $L^2(\mathbb{R}^d)$ is not necessarily integrable, and the Fourier transform as defined in (2.4) does not even make sense in this case.

We start by going a little further with the analysis of the Fourier transform for Schwartz functions. Since $\mathcal{S}(\mathbb{R}^d)$ is included in $L^2(\mathbb{R}^d)$, it can be endowed with the corresponding norm. The Fourier transform of $\mathcal{S}(\mathbb{R}^d)$ is then a quasi-isomery for this norm:

Proposition 2.21. *Let $f, g \in \mathcal{S}(\mathbb{R}^d)$. Then we have*

$$\int_{\mathbb{R}^d} \hat{f}(\xi)\overline{\hat{g}(\xi)}\, d\xi = (2\pi)^d \int_{\mathbb{R}^d} f(x)\overline{g(x)}\, dx.$$

In particular,

$$\|\hat{f}\|_{L^2} = (2\pi)^{\frac{d}{2}} \|f\|_{L^2}.$$

Proof. The map $(x,\xi) \mapsto e^{-ix\cdot\xi} f(x)\overline{\hat{g}(\xi)}$ is integrable on $\mathbb{R}^d \times \mathbb{R}^d$, so by the Fubini-Lebesgue Theorem we have

$$\begin{aligned}
\int_{\mathbb{R}^d} \hat{f}(\xi)\overline{\hat{g}(\xi)}\, d\xi &= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} e^{-ix\cdot\xi} f(x)\, dx \right) \overline{\hat{g}(\xi)}\, d\xi \\
&= \int_{\mathbb{R}^d} f(x) \left(\int_{\mathbb{R}^d} e^{-ix\cdot\xi} \overline{\hat{g}(\xi)}\, d\xi \right) dx \\
&= \int_{\mathbb{R}^d} f(x) \overline{\left(\int_{\mathbb{R}^d} e^{ix\cdot\xi} \hat{g}(\xi)\, d\xi \right)} dx \\
&= (2\pi)^d \int_{\mathbb{R}^d} f(x)\overline{g(x)}\, dx.
\end{aligned}$$

The second propery follows by taking $g = f$. □

Considering $g = \overline{\hat{\phi}}$ we can re-write the previous equality as follows.

Corollary 2.22. *For $f, \phi \in \mathcal{S}(\mathbb{R}^d)$ we have*

$$\int_{\mathbb{R}^d} \hat{f}\phi\, dy = \int_{\mathbb{R}^d} f\hat{\phi}\, dy.$$

Now we use the fact that $\mathcal{S}(\mathbb{R}^d)$ is dense in $L^2(\mathbb{R}^d)$ to extend the Fourier transform on $\mathcal{S}(\mathbb{R}^d)$ into an ismoetry of $L^2(\mathbb{R}^d)$.

Theorem 2.23. *There exists a unique isomorphism $\mathcal{F}$ of $L^2(\mathbb{R}^d)$ which coincides with the Fourier transform on $\mathcal{S}(\mathbb{R}^d)$ and such that*

$$\forall f \in L^2(\mathbb{R}^d), \quad \|\mathcal{F}f\|_{L^2(\mathbb{R}^d)} = (2\pi)^{\frac{d}{2}} \|f\|_{L^2(\mathbb{R}^d)}. \tag{2.8}$$

Proof. • Assume that two isomorphisms of $L^2(\mathbb{R}^d)$ satisfy the conditions of the theorem. They are in particular continuous applications on $L^2(\mathbb{R}^d)$ which coincide on the dense subset $\mathcal{S}(\mathbb{R}^d)$. Then they coincide everywhere, which gives uniqueness.
• Let $f \in L^2(\mathbb{R}^d)$. There exists a sequence $(f_n)_{n\in\mathbb{N}}$ of Schwartz functions that converge to f in $L^2(\mathbb{R}^d)$. By Proposition 2.21, the sequence $(\widehat{f_n})_{n\in\mathbb{N}}$ is a Cauchy sequence in $L^2(\mathbb{R}^d)$. Since $L^2(\mathbb{R}^d)$ is complete, $\widehat{f_n}$ has a limit in $L^2(\mathbb{R}^d)$, which we denote by $\mathcal{F}f$. We can check that this definition of $\mathcal{F}f$ does not depend on the choice of the sequence $(f_n)_{n\in\mathbb{N}}$ and that for $f \in \mathcal{S}(\mathbb{R}^d)$ the limit obtained this way coincides with the Fourier transform of f. By linearity of the Fourier transform on $\mathcal{S}(\mathbb{R}^d)$ and by Proposition 2.21, the map $\mathcal{F}$ is linear and satisfies (2.8). By continuity, the equality (2.6) still holds in $L^2(\mathbb{R}^d)$, which proves that $\mathcal{F}$ is surjective. □

Remark 2.24. For $f, g \in L^2(\mathbb{R}^d)$ we have

$$\int_{\mathbb{R}^d} \mathcal{F}f\, \overline{\mathcal{F}g}\, d\xi = (2\pi)^d \int_{\mathbb{R}^d} f\overline{g}\, dx.$$

and

$$\int_{\mathbb{R}^d} (\mathcal{F}f) g\, dy = \int_{\mathbb{R}^d} f(\mathcal{F}g)\, dy.$$

Proposition 2.25. *The isomorphism of Theorem 2.23 coincides with the Fourier transform already defined on $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$.*

Proof. Let $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$. For this proof, we denote by $\hat{f}$ the Fourier transform of f in the sense (2.4), and by $\mathcal{F}f$ the function of $L^2(\mathbb{R}^d)$ given by Theorem 2.23. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of Schwartz functions which converges to f both in $L^1(\mathbb{R}^d)$ and $L^2(\mathbb{R}^d)$. With the same convention as for f, we have $\hat{f_n} = \mathcal{F}f$ for any $n \in \mathbb{N}$. Since $\mathcal{F}$ is continuous on $L^2(\mathbb{R}^d)$ we have

$$\|\mathcal{F}f_n - \mathcal{F}f\|_2 \xrightarrow[n\to+\infty]{} 0.$$

In particular, possibly after extracting a subsequence, we can assume that $\mathcal{F}f_n$ tends to $\mathcal{F}f$ for almost every $x \in \mathbb{R}^d$. On the other hand we have

$$\|\hat{f_n} - \hat{f}\|_\infty \leqslant \|f - f_n\|_1 \xrightarrow[n\to+\infty]{} 0.$$

This implies that $\mathcal{F}f = \hat{f}$ almost everywhere. ☐

Recall that for a non integrable function, we cannot use the expression of $\mathcal{F}f$ given by (2.4). We at least have the following property, where $\tilde{f}_n$ is a kind of Fourier transform for locally integrable functions.

Proposition 2.26. *Let $f \in L^2(\mathbb{R}^d)$. For $n \in \mathbb{N}$ and $\xi \in \mathbb{R}^d$ we set*

$$\tilde{f}_n(\xi) = \int_{B(n)} e^{-ix\cdot\xi} f(x)\,dx,$$

where $B(n)$ denotes the ball of radius n centered at 0. Then we have

$$\left\|\tilde{f}_n - \mathcal{F}f\right\|_2 \xrightarrow[n\to+\infty]{} 0.$$

Proof. Using the doiminated convergence Theorem we have

$$\left\|\mathbb{1}_{B(n)}f - f\right\|_2 \xrightarrow[n\to+\infty]{} 0.$$

Since $\mathbb{1}_{B(n)}f \in L^1 \cap L^2$ for any $n \in \mathbb{N}$, we get using proposition 2.25 that

$$\|\tilde{f}_n - \mathcal{F}f\|_2 = \left\|\mathcal{F}(\mathbb{1}_{B(n)}f) - \mathcal{F}f\right\|_2 \xrightarrow[n\to+\infty]{} 0.$$ ☐

2.6 Example of application

For $f \in \mathcal{S}(\mathbb{R}^d)$, we consider the equation

$$-\Delta u + u = f, \tag{2.9}$$

of unknown $u \in \mathcal{S}(\mathbb{R}^d)$. Taking the Fourier transform, we see that u is a solution if and only if for any $\xi \in \mathbb{R}^d$ we have

$$(|\xi|^2 + 1)\hat{u}(\xi) = \hat{f}(\xi),$$

or equivalently

$$\hat{u}(\xi) = \frac{\hat{f}(\xi)}{1 + |\xi|^2}.$$

We consider the function $\tilde{f}$ which maps $\xi \in \mathbb{R}^d$ to $\frac{\hat{f}(\xi)}{1+|\xi|^2}$. This defines a function $\tilde{f} \in \mathcal{S}(\mathbb{R}^d)$ and we have

$$\|\tilde{f}\|_{L^2} \leqslant \|\hat{f}\|_{L^2} = (2\pi)^{\frac{d}{2}} \|f\|_{L^2}.$$

We then set $Rf = \mathcal{F}^{-1}\tilde{f}$. This also defines a function of $\mathcal{S}(\mathbb{R}^d)$ and we have

$$\|Rf\|_{L^2} = (2\pi)^{-\frac{d}{2}} \|\tilde{f}\|_{L^2} \leqslant \|f\|_{L^2}.$$

It is not very difficult to check that the map $f \mapsto Rf$ is linear. Thus R defines a linear map from $\mathcal{S}(\mathbb{R}^d)$ to itself, continuous for the L^2 norm. It is therefore extended by continuity into a continuous linear map on $L^2(\mathbb{R}^d)$.

For $f \in \mathcal{S}(\mathbb{R}^d)$ we thus have

$$(-\Delta + 1)Rf = f.$$

We can then check that for any $u \in \mathcal{S}(\mathbb{R}^d)$ we have

$$R(-\Delta + 1)u = u.$$

Thus R is on $\mathcal{S}(\mathbb{R}^d)$ the inverse function of $(-\Delta + 1)$ (however, be careful, R is not a bijection on $L^2(\mathbb{R}^d)$).

A final question about this example. Given $f \in \mathcal{S}(\mathbb{R}^d)$, can we give a slightly more explicit expression for the solution Rf of the problem (2.9)? We have to compute the inverse Fourier transform of the product of two functions. This is where the convolution product appears (see Proposition 2.13). Thus, if G is a function which Fourier transform is given by

$$\hat{G}(\xi) = \frac{1}{1 + |\xi|^2},$$

then we will have

$$\widehat{Rf}(\xi) = \hat{G}(\xi)\hat{f}(\xi),$$

and hence

$$Rf(x) = (G * f)(x).$$

It remains to determine G. At least for the dimension $d = 1$, we have already done the computations at example 2.9 and we get

$$G(x) = \frac{e^{-|x|}}{2}.$$

This gives

$$Rf(x) = \int_{\mathbb{R}} \frac{e^{-|x-y|}}{2} f(y)\, dy = \int_{\mathbb{R}} \frac{e^{-|y|}}{2} f(x-y)\, dy.$$

We can now check with an explicit computation (left as an exercise) that if $f \in \mathcal{S}(\mathbb{R})$ then for any $x \in \mathbb{R}$ we have

$$-(Rf)''(x) + Rf(x) = f(x).$$

In dimension 1, the result is not very spectacular, because it is only an ordinary differential equation with constant coefficients, and we could have got the same result using only a double variation of constants, but the method described on this simple example will be useful to understand more subtle problems...