(2016) B.I.A.A. no **103** BULLETIN D'INFORMATIQUE APPROFONDIE ET APPLICATIONS
ISSN 0291-5413 (Université d'Aix-Marseille)

Mécanique quantique I : bases mathématiques

Eric OLIVIER [1] [2]

Résumé. – La formulation moderne de « *la mécanique quantique* » est présentée de manière systématique (axiomatique) par Von Neumann au tournant des années 30 : un état quantique est alors vu comme un « *point* » de la sphère unité d'un « *espace de Hilbert complexe* » de dimension infinie. Dans cette première note, nous donnons les repères d'analyse fonctionnelle nécessaires pour comprendre comment cette « *nouvelle mécanique* » s'attache à décrire le réel microscopique. En particulier la transformée de Fourier sur l'espace de Hilbert des signaux d'énergie finie nous permet d'énoncer et de démontrer l'inégalité de Heisenberg-Gabor dans le bon espace de Sobolev. Nous verrons dans les notes suivantes que cette inégalité correspond à l'inégalité de Heisenberg pour la position et l'impulsion des fonctions d'onde de la mécanique ondulatoire.

1. Introduction

Dans le milieu des années 20, *« la mécanique des quanta »* se formalise sous deux formes apparemment différentes (mais en fait équivalentes), avec d'une part *« la mécanique des matrices »* de Heisenberg, Jordan et Born et d'autre part *« la mécanique ondulatoire »* de de Broglie et Schrödinger. La synthèse de ces deux approches donnera *« la mécanique quantique »*. Dans cette série de notes, nous commençons par adopter le point de vue de de Broglie et Schrödinger [3]. Dans une lignée de savants éminents (principalement Descartes, Fermat, Newton, Huyghens, Maupertuis, Hamilton, Jacobi, Maxwell, Fresnel, Einstein,...), l'approche de de Broglie et Schrödinger marque une étape significative de plusieurs siècles de polémiques et de réflexions sur l'ambivalence entre *« l'optique géométrique »* et *« l'optique ondulatoire »* [4] (nous y reviendrons dans [Oli17]). Pour comprendre les idées mises en jeu, nous avons besoin d'un certain nombre de bases d'analyse fonctionnelle et tout spécialement de la théorie des *« espaces de Hilbert »*. Nous considérons ici qu'un espace de Hilbert est un espace de Banach (espace vectoriel normé complet) dans lequel l'identité du parallélogramme est satisfaite (théorème de Fréchet-von Neumann-Jordan) : cela entraîne que la norme est nécessairement déduite d'un produit scalaire (dont l'expression est donnée par l'identité de polarization). Dans cette présentation, *« l'inégalité de Cauchy-Schwarz »* est une conséquence directe de *« l'inégalité triangulaire »* (qui pré-existe comme propriété de la norme) : ce point est particulièrement intéressant dans la mesure où nous verrons (c.f. [Oli16a]) que *« l'inégalité de Heisenberg »* découle directement de l'inégalité de Cauchy-Schwarz et donc de l'inégalité triangulaire !

Dans cette première note, nous donnons une description rapide des principaux espaces fonctionnels classiques : les espaces de Banach avec les espaces L^p (c.f. Section 2), les espaces de Hilbert (c.f. Section 5) et enfin les espaces de Sobolev (c.f. Section 9). La

1. GDAC-I2M UMR 7373 CNRS Université d'Aix-Marseille
2. eric.olivier@univ-amu.fr

3. Nous verrons plus tard comment le point de vue de Heisenberg autorise la modélisation de phénomènes quantiques très généraux (statistique quantique) dans le cadre du formalisme des *« $C^\star$-algèbres »* tel qu'initié par Von-Neumann.
4. On parle aussi *« d'optique physique »*.

Section 8 est consacrée à la théorie des probabilités sur la droite réelle. Nous insistons sur des points élémentaires de la théorie qui jouent un rôle en mécanique quantique : en particulier la signification *intuitive* attachée aux moments d'ordre 1 et 2 d'une distribution de probabilité (i.e. la moyenne et l'écart type) autorise une interprétation du *« théorème d'Ehrenfest »* (sur la correspondance entre les trajectoires classiques et quantiques) ainsi que du *« principe d'incertitude de Heisenberg »* dans le formalisme probabiliste de la mécanique quantique proposée par Born (nous y reviendrons dans [Oli16b]). De manière anachronique et hors du contexte de la mécanique quantique, nous utilisons les propriétés des moments probabilistes pour aborder le principe d'incertitude à partir de la théorie du signal et de la transformée de Fourier des signaux d'énergie finie (c.f. Sections 6 & 7). Plus précisément, nous démontrons *« l'inégalité de Heisenberg-Gabor »* dans le cas où les signaux en question sont dans le bon espace de Sobolev. Cette approche de la mécanique quantique par la théorie du signal est intéressante à deux titres. Premièrement, elle permet de s'abstraite d'un contexte physique où les diverses interprétations compliquent les choses et deuxièmement, elle explique – en un sens – le rôle joué par la transformée de Fourier dans *« le principe de correspondance de Schrödinger »* et la définition des couples d'opérateurs conjugués (c.f. [Oli16b, Oli17]). Typiquement les opérateurs position et impulsion existent déjà – implicitement – dans l'inégalité de Heisenberg-Gabor.

2. Espace de Banach – Algèbre de Banach

2.1. Pour $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, un $\mathbb{K}$-espace vectoriel normé $(\mathscr{E}, \|\cdot\|)$ est appelé un *« espace de Banach »* lorsqu'il est complet, c'est-à-dire lorsque toute suite de points de $\mathscr{E}$ qui est de Cauchy possède une limite dans $\mathscr{E}$. Rappelons qu'une série $\sum_{k=0}^{\infty} a_k$ de $\mathscr{E}$ est dite *« normalement convergente »* (resp. *« absolument convergente »*) ssi les sommes partielles $\sum_{k=0}^{n} a_k$ convergent dans $\mathscr{E}$ quand $n \to +\infty$ (resp. la série numérique $\sum_{k=0}^{\infty} \|a_k\|$ est convergente) : nous utiliserons la caractérisation suivante qui est très pratique.

Lemme 2.1. *L'espace normé $(\mathscr{E}, \|\cdot\|)$ est de Banach ssi toute série absolument convergente de $\mathscr{E}$ est normalement convergente sur $\mathscr{E}$.*

Preuve. Soit $\sum_{k=0}^{\infty} a_k$ une série absolument convergente de l'espace de Banach $(\mathscr{E}, \|\cdot\|)$; alors les $A_n := \sum_{k=0}^{n} a_k$ $(n \geq 0)$ forment une suite de points de $\mathscr{E}$ qui est de Cauchy et nous notons A sa limite : le fait que la série $\sum_{k=0}^{\infty} a_k$ soit normalement convergente sur $(\mathscr{E}, \|\cdot\|)$ découle alors de l'inégalité $\|A_n - A\| \leq \sum_{k=n+1}^{\infty} \|a_k\|$ qui assure que $\|A_n - A\| \to 0$ quand $n \to +\infty$. Réciproquement, supposons que toute série absolument convergente de $(\mathscr{E}, \|\cdot\|)$ est normalement convergente et soit $a_0, a_1, \ldots$ une suite de $\mathscr{E}$ possédant la propriété de Cauchy : il s'agit de montrer que cette suite converge dans $\mathscr{E}$. Pour cela nous partons du fait qu'il existe une sous-suite $a_{n_0}, a_{n_1}, \ldots$ t.q. $\|a_{n_k} - a_{n_{k-1}}\| \leq 1/2^k$. Si nous posons $\alpha_0 := a_{n_0}$ et $\alpha_k := a_{n_k} - a_{n_{k-1}}$, pour tout $k \geq 1$, alors

$$\sum_{i=0}^{k} \alpha_i = a_{n_0} + (a_{n_1} - a_{n_0}) + \cdots + (a_{n_k} - a_{n_{k-1}}) = a_{n_k}$$

Or, par construction $\|\alpha_k\| \leq 1/2^k$, pour tout $k \geq 1$; par suite $\sum_{k=0}^{\infty} \alpha_k$ est une série absolument convergente de $\mathscr{E}$ et donc (hypothèse) normalement convergente : en d'autres termes, $a_* := \sum_{k=0}^{\infty} \alpha_k$ est un élément de $\mathscr{E}$ t.q. $\|a_* - a_{n_k}\| = \|a_* - \sum_{i=0}^{k} \alpha_i\| \to 0$ quand

$k \to +\infty$. Nous venons de démontrer que a_* est une valeur d'adhérence de la suite des a_n : mais cette suite étant de Cauchy, elle converge nécessairement vers a_*.

□

2.2. Soit $\mathscr{A}$ une $\mathbb{K}$-algèbre[5] normée, en ce sens que $\mathscr{A}$ (considéré comme $\mathbb{K}$-espace vectoriel) est muni d'une norme $\|\cdot\|$. Alors $\mathscr{A}$ est appelée *« une algèbre de Banach »* si l'espace vectoriel normé $(\mathscr{A}, \|\cdot\|)$ est un espace de Banach et si de plus la norme est sous multiplicative en ce sens que $\|AB\| \leq \|A\|\,\|B\|$, pour tout $A, B \in \mathscr{A}$; si de plus il existe un élément I in $\mathscr{A}$ t.q. $AI = IA = A$, pour tout $A \in \mathscr{A}$, alors l'algèbre de Banach est dite unitaire. L'espace $(\mathscr{C}[0;1], \|\cdot\|_\infty)$ des fonctions réelles continues sur l'intervalle $[0;1]$ et muni de la norme uniforme est un espace de Banach ; étant données f et g deux fonctions quelconques de $\mathscr{C}[0;1]$, il est immédiat de vérifier $\|fg\|_\infty \leq \|f\|_\infty \|g\|_\infty$: il en découle que $\mathscr{C}[0;1]$ est une algèbre de Banach (unitaire et commutative) pour l'addition et la multiplication des fonctions.

2.3. Soient $\mathscr{E}$ et $\mathscr{F}$ deux $\mathbb{K}$-espaces vectoriels : dans la suite, nous noterons $L(\mathscr{E}, \mathscr{F})$ le $\mathbb{K}$-espace vectoriel des homomorphismes $\mathbb{K}$-linéaires de $\mathscr{E}$ dans $\mathscr{F}$; quand $\mathscr{E} = \mathscr{F}$ nous notons simplement $L(\mathscr{E})$ le $\mathbb{K}$-espace vectoriel des endomorphismes de $\mathscr{E}$. Lorsque $\mathscr{E}$ et $\mathscr{F}$ sont tous deux munis d'une norme, alors $\mathcal{L}(\mathscr{E}, \mathscr{F})$ (resp. $\mathcal{L}(\mathscr{E})$) désignera le sous-espace de $L(\mathscr{E}, \mathscr{F})$ (resp. $L(\mathscr{E})$) des homomorphismes (resp. endomorphismes) de $\mathscr{E}$ dans $\mathscr{F}$ (resp. de $\mathscr{E}$). Si l'axiome du choix fait partie de nos outils mathématiques, alors $\mathcal{L}(\mathscr{E}, \mathscr{F})$ est toujours un sous-espace strict de $L(\mathscr{E}, \mathscr{F})$, dès que $\mathscr{E}$ est de dimension infinie. Il est facile de vérifier que tout endomorphisme linéaire $X \mapsto AX$ de $\mathscr{E}$ est continu (i.e. $A \in \mathcal{L}(\mathscr{E})$) ssi il est borné (sur la sphère unité), c'est-à-dire si le supremum $\|A\|$ des $\|AX\|$, pour $\|X\| = 1$ est fini. L'application $A \mapsto \|A\|$ est une norme sur $\mathcal{L}(\mathscr{E})$ appelée *« norme subordonnée »* (à la norme sur $\mathscr{E}$), ou encore *« la norme opérateur »* : celle-ci possède la propriété importante d'être sous-multiplicative, en ce sens que pour tout[6] $A, B \in \mathcal{L}(\mathscr{E})$

$$\|AB\| \leq \|A\|\,\|B\| \tag{1}$$

2.4. Considérons maintenant que $(\mathscr{E}, \|\cdot\|)$ est un espace de Banach ; alors l'espace normé $(\mathcal{L}(\mathscr{E}), \|\cdot\|)$ est lui aussi un espace de Banach : l'inégalité (1) vérifiée par la norme subordonnée fait alors de $\mathcal{L}(\mathscr{E})$ une algèbre de Banach unitaire, où l'unité est l'identité de $\mathscr{E}$ que nous notons I. (Lorsque $\mathscr{E}$ est dimension au moins égale à 2, l'algèbre $\mathcal{L}(\mathscr{E})$ est non commutative.) Cette structure d'algèbre de Banach unitaire permet en particulier de définir l'image d'un endomorphisme A par une application analytique (comme généralisation des valeurs polynomiales sur A). En effet si $F(z) := \sum_{k=0}^{\infty} a_k z^k$ est une série entière de la variable complexe z (avec $a_k \in \mathbb{K}$) de rayon de convergence $0 < \rho \leq +\infty$ (i.e. $F(z)$ est holomorphe sur l'ensemble des $z \in \mathbb{C}$ t.q. $|z| < r$ ssi $r < \rho$), alors pour tout $A \in \mathcal{L}(\mathscr{E})$ t.q. $\|A\| < \rho$ (et en utilisant la convention que $A^0 = I$)

$$F(A) := \sum_{k=0}^{\infty} a_k A^k$$

5. Où nous notons AB le produite de deux éléments de $\mathscr{A}$.
6. Nous notons AB la composition $A \circ B$.

est un élément de $\mathcal{L}(\mathscr{E})$. Supposons par exemple que $\|A\| < 1$: alors l'endomorphismes $I - A$ est inversible et par *« la formule d'inversion de von Neumann »* :

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k$$

Un autre exemple important est celui de l'exponentielle : pour tout $A \in \mathcal{L}(\mathscr{E})$, nous pouvons définir *« l'exponentielle »* de A, soit

$$\exp(A) := \sum_{k=0}^{\infty} 1/k! A^k$$

avec la propriété importante que pour tout $A, B \in \mathcal{L}(\mathscr{E})$

$$AB = BA \implies \exp(A + B) = \exp(A)\exp(B)$$

(en particulier $\exp(A)$ est toujours inversible avec $\exp(A)^{-1} = \exp(-A)$).

2.5. L'exponentielle permet de résoudre les *« problèmes de Cauchy linéaires »* du premier ordre sur $\mathscr{E}$, du type $dY/dt = AY(t)$ avec la condition initiale $Y(0) = Y_0 \in \mathscr{E}$ et où $A \in \mathcal{L}(\mathscr{E})$. En effet, pour tout $t, h \in \mathbb{R}$, du fait que tA commute avec hA, nous avons

$$\exp\big((t+h)A\big) = \exp(tA)\exp(hA) = \exp(tA)\big(I + hA + \cdots\big) = \exp(tA) + tA\exp(tA) + \cdots$$

de sorte que $d/dt \exp(tA) = A\exp(tA)$. Alors nous avons

$$\frac{dY}{dt} = AY(t) \iff \exp(-tA)\left(\frac{dY}{dt} - AY(t)\right) = 0 \iff \frac{d}{dt}\big(\exp(-tA)Y(t)\big) = 0$$

ce qui (avec la condition $Y(0) = Y_0$) donne finalement la solution

$$Y(t) = \exp(tA)Y_0$$

Remarque 2.2. *Lorsque l'espace de Banach $\mathscr{E}$ est de dimension finie, tout endomorphisme est continu, de sorte que $L(\mathscr{E}) = \mathcal{L}(\mathscr{E})$; quand $\mathscr{E}$ est de dimension infinie, l'axiome du choix permet d'exhiber facilement des endomorphismes non continus, faisant de $\mathcal{L}(\mathscr{E})$ un sous-espace strict de $L(\mathscr{E})$. Un théorème difficile*[7] *montre qu'en un sens, l'axiome du choix est la condition d'existence des endomorphismes non bornés dans les espaces de Banach de dimension infinie. Les « opérateurs non bornés » de la mécanique quantique (c.f.* [Oli16a]*) sont des applications linéaires définies sur un sous-espace dense de l'espace total – en l'occurence un espace de Hilbert $\mathscr{H}$ – à valeurs dans $\mathscr{H}$ mais ne possédant pas de prolongement borné à tout $\mathscr{H}$. Nous verrons (c.f.* [Oli16a]*) comment il est possible de résoudre (dans certains cas) les problèmes de Cauchy du type*

$$\frac{dY}{dt} = AY(t) \quad \textit{avec} \quad Y(0) = Y_0 \in \mathscr{D} \tag{2}$$

lorsque $(A, \mathscr{D})$ est un opérateur non borné défini sur un sous-espace $\mathscr{D}$ dense dans $\mathscr{H}$. « L'équation de Schrödinger » correspond à un problème de Cauchy du type (2) dans le cas où $(A, \mathscr{D})$ est autoadjoint : c'est « le théorème de Stone » (prototype du théorème de « Hille-Yosida ») qui permet la résolution abstraite de l'équation de Schrödinger.

7. Si $\mathscr{E}$ est un espace de Banach de dimension infinie, alors il existe une suite $e_0, e_1, \ldots$ formant une famille dénombrable et libre de vecteurs unitaires : grâce à l'axiome du choix cette famille peut être complétée en une famille $\{e_i\ ;\ I\}$ (avec $I \supset \mathbb{N}$) formant une base (algébrique) de $\mathscr{E}$: si nous posons $f(e_n) = ne_n$, pour tout $n \in \mathbb{N}$ et $f(e_i) = 0$ pour tout $i \in I \setminus \mathbb{N}$, alors f est un endomorphisme non borné. En un sens, c'est le seul moyen d'obtenir un endomorphisme non borné sur un espace de Banach : ce résultat est dû à Solovay.

2.6. Dans cette série de notes, nous supposons connue la *« théorie de la mesure »* et de *« l'intégrale de Lebesgue »* (des présentations simples et complètes sont données dans [ST89, GW90] – pour une présentation plus détaillée c.f. e.g. [KF94]). Nous utiliserons sans plus de précisions, les notions de *« fonction borélienne »* et *« d'égalité presque partout »*, les théorèmes de la *« convergence monotone »* et de la *« convergence dominée »*, le *« théorème de dérivation sous le signe intégral »*, le *« théorème de dérivation de Lebesgue »*, le *« théorème de Radon-Nikodym »* et enfin le *« théorème de Fubini »*. En particulier nous noterons dans la suite $\lambda(dx) = dx$ la mesure de Lebesgue sur $\mathbb{R}$ (qui est caractérisée par le fait que $\lambda([a\,;b]) = b - a$, pour tout $-\infty < a \le b < +\infty$). Afin d'aborder les bases mathématiques de la mécanique quantique, nous aurons justement besoin d'un certain nombres d'espaces de Banach très importants et dont la définition dépend de l'intégrale de Lebesgue sur [8] $\mathbb{R}$. Soit I un intervalle fermé de la droite réelle et $p \ge 1$ un réel donné ; nous allons montrer (c.f. Corollaire 2.8 infra) que l'ensemble des fonctions (classes de fonctions boréliennes) $f : I \to \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$) telle que f^p soit Lebesgue intégrable (sur l'intervalle I) est bien un espace vectoriel : cet espace est classiquement noté $L^p(I)$ (avec le cas spécial de $L^\infty(I)$ qui coïncide par convention avec l'espace des fonctions qui sont bornées presque partout sur I). Chaque espace $L^p(I)$ est associé à une norme qui en fait un espace de Banach (théorème de Riesz-Fisher : c.f. Théorème 2.9 infra). Pour voir tout cela, considérons que $f : I \to \mathbb{K}$ est une fonction borélienne ; si d'une part $f \in L^\infty(I)$, alors nous noterons $\|f\|_\infty$ l'infimum des $M > 0$ t.q. $|f(x)| \le M$ pour presque tout $x \in I$ (par convention $\|f\|_\infty = +\infty$ si f n'est pas dans $L^\infty(I)$) ; si d'autre part $1 \le p < +\infty$ alors nous posons par définition

$$\|f\|_p := \left(\int_I |f(x)|^p dx\right)^{1/p}$$

Définition 2.3. *Soient $0 \le p \le +\infty$ et I un intervalle fermé de $\mathbb{R}$; alors la fonction borélienne $f : I \to \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$) appartient à $L^p(I)$ ssi $\|f\|_p < +\infty$.*

La proposition suivante est *« l'inégalité de Hölder »*.

Proposition 2.4 (Hölder). *Soient $1 \le p, q \le +\infty$ t.q. $1/p + 1/q = 1$ (p et q sont dits conjugués) et I un intervalle fermé de $\mathbb{R}$; si $f, g : I \to \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$) sont deux fonctions boréliennes, alors :*

$$\|fg\|_1 = \int_I |f(x)g(x)|dx \le \left(\int_I |f(x)|^p dx\right)^{1/p} \left(\int_I |g(x)|^q dx\right)^{1/q} = \|f\|_p \cdot \|g\|_q$$

de plus $\|fg\|_1 = \|f\|_p\,\|g\|_q$ si et seulement si $|f|^p = |g|^q$ presque partout.

Lemme 2.5 (Inégalité de Young). *Pour tous réels $a, b \ge 0$ et tous réels $p, q \ge 1$ t.q. $1/p+1/q = 1$:*

$$ab \le a^p/p + b^q/q$$

De plus, l'égalité a lieu si et seulement si $a^p = b^q$.

Preuve. Avec $a^p = \exp(x)$ et $b^q = \exp(y)$ la convexité de la fonction exponentielle donne

$$ab = \exp\left(\frac{x}{p} + \frac{y}{q}\right) \le \frac{1}{p}\exp(x) + \frac{1}{q}\exp(y) = \frac{a^p}{p} + \frac{b^q}{q}$$

Le cas d'égalité est une conséquence de la stricte convexité de l'exponentielle.

□

8. Ou par extension sur $\mathbb{R}^n$.

Preuve de la Proposition 2.4. Dans le cas où $0 < p, q < +\infty$ avec $0 < \|f\|_p, \|g\|_q < +\infty$ (les autres cas étant laissés au soin du lecteur), l'inégalité de Young entraîne que :

$$\begin{aligned}\|fg\|_1 &= \|f\|_p \cdot \|g\|_q \left(\int_I \frac{|f(x)|}{\|f\|_p} \cdot \frac{|g(x)|}{\|g\|_q} dx \right) \\ &\leq \|f\|_p \cdot \|g\|_q \left(\frac{1}{p} \int_I \left(\frac{|f(x)|}{\|f\|_p} \right)^p dx + \frac{1}{q} \int_I \left(\frac{|g(x)|}{\|g\|_q} \right)^q dx \right) = \|f\|_p \cdot \|g\|_q\end{aligned}$$

Le cas d'égalité découle du cas d'égalité de l'inégalité de Young.

□

L'inégalité de Hölder généralisée suivante s'obtient par récurrence.

Corollaire 2.6. *Pour $n \geq 2$ et $1 \leq p_1, \ldots, p_n \leq +\infty$ t.q. $1/p_1 + \cdots + 1/p_n = 1$ et $f_1, \ldots, f_n$ des fonctions (complexes) boréliennes sur $\mathbb{R}$:*

$$\|f_1 \cdots f_n\|_1 \leq \|f_1\|_{p_1} \cdots \|f_n\|_{p_n}$$

Un corollaire important de l'inégalité de Hölder est *« l'inégalité de Minkowski »*.

Corollaire 2.7 (Minkowski). *Si $1 \leq p \leq +\infty$ et si $f, g \in L^p(I)$, alors $f + g \in L^p(I)$ et :*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

Preuve. Nous considérons que $1 < p < +\infty$ (les deux autres cas étant laissés au soin du lecteur). Supposons donc que f et g sont deux fonctions dans $L^p(I)$, i.e. $\|f\|_p, \|g\|_p < +\infty$; alors en utilisant le fait que pour tout $a, b \in \mathbb{R}$

$$|a + b|^p \leq (|a| + |b|)^p \leq 2^p \max\{|a|, |b|\}^p \leq 2^p(|a|^p + |b|^p)$$

nous pouvons intégrer l'inégalité $|f(x) + g(x)|^p \leq 2^p(|f(x)|^p + |g(x)|^p)$ et obtenir que

$$\int_I |f(x) + g(x)|^p dx \leq 2^p \left(\int_I |f(x)|^p dx + \int_I |g(x)|^p dx \right)$$

Cela assure que $f + g$ est bien dans $L^p(I)$, c'est-à-dire que $\|f + g\|_p < +\infty$. Pour l'inégalité de Minkowski, elle est triviale lorsque $\|f + g\|_p = 0$. Lorsque $\|f + g\|_p > 0$, nous notons A l'ensemble des $x \in I$ t.q. $|f(x) + g(x)| > 0$, de sorte que $\|f + g\|_p^p = \int_A |f(x) + g(x)|^p dx$. Grâce à l'inégalité de Hölder (avec $q = p/(p - 1)$) nous pouvons écrire :

$$\begin{aligned}\|f + g\|_p^p &\leq \int_A |f(x)| \cdot |f(x) + g(x)|^{p-1} \mathrm{d}x + \int_A |g(x)| \cdot |f(x) + g(x)|^{p-1} dx \\ &\leq \left(\int_I |f(x)|^p \mathrm{d}x \right)^{1/p} \left(\int_I |f(x) + g(x)|^{(p-1)q} dx \right)^{1/q} \\ &\quad + \left(\int_I |g(x)|^p \mathrm{d}x \right)^{1/p} \left(\int_I |f(x) + g(x)|^{(p-1)q} dx \right)^{1/q}\end{aligned}$$

et donc (avec $(p - 1)q = p$ et $1/q = 1 - 1/p$) il vient : $\|f + g\|_p^p \leq \left(\|f\|_p + \|g\|_p\right) \cdot \|f + g\|_p^{p-1}$; du fait que $0 < \|f + g\|_p < +\infty$, nous pouvons conclure que $\|f + g\|_p \leq \|f\|_p + \|g\|_p$.

□

Corollaire 2.8. *Pour tout $1 \leq p \leq +\infty$ l'ensemble $L^p(I)$ est un espace vectoriel et l'application $f \mapsto \|f\|_p$ est une norme sur $L^p(I)$.*

Le Théorème de Riesz-Fisher ci-dessous est un beau résultat de la théorie de la mesure.

Théorème 2.9. *Soit I un intervalle fermé de la droite réelle et $p \geq 1$ un réel donné ; alors l'espace $L^p(I)$ des fonctions réelles (ou complexes) f définies et boréliennes sur I, telles que $f^p(x)$ soit Lebesgue-intégrable (sur I) est un espace de Banach pour la norme $\|\cdot\|_p$ telle que*

$$\|f\|_p = \left(\int_I |f(x)|^p dx\right)^{1/p}$$

Preuve. D'après le Lemme 2.1, il suffit de montrer que si $\sum_{n=0}^{\infty} f_k$ est une série absolument convergente de $L^p(I)$ (i.e. $f_k \in L^p(I)$ et $\sum_{k=0}^{\infty} \|f_k\|_p < \infty$), alors les sommes partielles $F_n := \sum_{k=0}^{n} f_k$ forment une suite qui converge dans $L^p(I)$ vers une fonction F de $L^p(I)$. Pour voir cela nous pouvons définir pour tout $x \in I$,

$$G(x) := \sum_{k=0}^{\infty} |f_k(x)| \quad (\in [0\,; +\infty])$$

de sorte que $x \mapsto G(x)$ est une fonction borélienne. D'après l'inégalité de Minkowski (inégalité triangulaire pour la norme $\|\cdot\|_p$) nous avons

$$\int G^p(x)dx = \int_I \Big(\sum_{k=0}^{\infty} |f_k(x)|\Big)^p dx = \Big\| \sum_{k=0}^{\infty} |f_k| \Big\|_p^p \leq \Big(\sum_{k=0}^{\infty} \|f_k\|_p\Big)^p < +\infty$$

ce qui signifie que $G \in L^p(I)$. En particulier la valeur de $G(x)$ doit être finie pour presque tout x ; par suite $\sum_{k=0}^{\infty} f_k(x)$ est une série numérique (à valeurs réelles ou complexes) absolument convergente et donc sommable pour presque tout x : il existe donc une fonction F, borélienne sur I et telle que pour presque tout $x \in I$

$$F(x) := \lim_{n\to+\infty} F_n(x) = \sum_{k=0}^{\infty} f_k(x)$$

D'une part, il est facile de vérifier que F est est une fonction de $L^p(I)$ puisque

$$\int_I |F(x)|^p dx \leq \int_I \Big(\sum_{k=0}^{\infty} |f_k(x)|\Big)^p dx = \int_I G^p(x)dx < +\infty$$

Mais d'autre part, les sommes partielles $F_n := \sum_{k=0}^{n} f_k$ sont aussi des fonctions de $L^p(I)$ et nous avons

$$\int_I \Big|F(x) - F_n(x)\Big|^p dx \leq \int_I \Big(\sum_{k=n+1}^{\infty} |f_k(x)|\Big)^p dx$$

Comme $\big(\sum_{k=n+1}^{\infty} |f_k(x)|\big)^p$ est dominée par la fonction $G^p(x)$ qui est Lebesgue intégrable et que (pour presque tout x) $\big(\sum_{k=n+1}^{\infty} |f_k(x)|\big)^p \to 0$ quand $n \to +\infty$, nous pouvons conclure (convergence dominée) que les F_n tendent vers F dans $L^p(I)$.

□

Il y a plusieurs versions du théorème dit de Riesz-Fisher (le Théorème 5.8 ci-dessous en est une version élémentaire [9]). On trouve plusieurs énoncés de ce théorème dans [RSN55, p. 58 & 69-70], dont la formulation primitive due à Riesz : c.f. [RSN55, p. 70]) : l'énoncé de la version primitive utilise implicitement la caractérisation des espaces de Banach donnée par le Lemme 2.1. Nous pointons qu'une version plus complète affirme en plus que si les f_n tendent vers f dans $L^p(I)$, alors il existe une sous-suite $f_{r_0}, f_{r_1}, \ldots$ t.q. $f_{r_0}(x), f_{r_1}(x), \ldots$ qui tend vers $f(x)$ pour presque tout $x \in \mathbb{R}$. Ce résultat est par exemple implicite dans la démonstration [10] donnée dans [RSN55, p. 58] : nous allons démontrer cette dernière proposition à partir de la complétude de l'espace $L^p(I)$.

9. Les démonstrations des Théorèmes 2.9 & 5.8 sont d'ailleurs calquées l'une sur l'autre.
10. Cette démonstration n'utilise pas le Lemme 2.1.

Proposition 2.10. *Soit I un intervalle fermé de $\mathbb{R}$ et $1 \le p \le +\infty$; si $f_0, f_1, \ldots$ est une suite de $L^p(I)$ convergeant vers f dans $L^p(I)$, alors il existe une suite strictement croissante $r_0, r_1, \ldots$ formée d'entiers pour laquelle $f_{r_0}(x), f_{r_1}(x), \ldots$ tend vers $f(x)$ pour presque tout $x \in \mathbb{R}$.*

Preuve. Soit $f_0, f_1, \ldots$ une suite de $L^p(I)$ convergeant vers f dans $L^p(I)$: alors il existe une sous-suite $f_{r_0}, f_{r_1}, \ldots$ t.q. $\|f_{r_k} - f_{r_{k-1}}\|_p \le 1/2^k$. Si nous posons

$$g_n := \sum_{k=1}^{n} |f_{r_k} - f_{r_{k-1}}|$$

alors il est immédiat de vérifier que les g_n forment une suite de Cauchy dans $L^p(I)$; par suite (Théorème de Riesz-Fisher) la suite des g_n converge vers la fonction g_* de $L^p(I)$ t.q $g_*(x) = \sup_n\{g(x)\}$. Mais g_* étant dans $L^p(I)$, il est nécessaire que $g_*(x) < +\infty$ pour presque tout $x \in I$; en particulier, cela signifie que la suite (croissante) $g_n(x)$ est de Cauchy (i.e. bornée), pour presque tout $x \in \mathbb{R}$. Or nous avons

$$\begin{aligned} \left|f_{r_{n+m}}(x) - f_{r_n}(x)\right| &= \left|\left(-f_{n_0}(x) + \sum_{k=1}^{n+m} f_{r_k}(x) - f_{r_{k-1}}(x)\right) - \left(-f_{n_0}(x) + \sum_{k=1}^{n} f_{r_k}(x) - f_{r_{k-1}}(x)\right)\right| \\ &\le \sum_{k=1}^{n+m} |f_{r_k}(x) - f_{r_{k-1}}(x)| - \sum_{k=1}^{n} |f_{r_k}(x) - f_{r_{k-1}}(x)| \\ &= g_{n+m}(x) - g_n(x) \end{aligned}$$

ce qui signifie que $f_{r_n}(x)$ est une suite de Cauchy pour presque tout $x \in I$.

□

3. Convolutions

3.1. La convolution est un outil très puissant d'analyse, en particulier à cause de ses propriétés de régularisation et d'approximation. C'est aussi une opération essentielle à la théorie du signal : la transmission de signaux à travers des systèmes linéaires se traduit par l'action d'opérateurs de convolution [11]. Au lieu d'une présentation systématique, nous donnons quelques exemples d'applications qui couvrirons nos besoins ultérieurs. Pour commencer, nous dirons que deux fonctions réelles f et g définies et boréliennes sur $\mathbb{R}$ sont *« convolables »* ssi $y \mapsto f(x-y)g(y)$ est intégrable pour presque partout $x \in \mathbb{R}$; alors la convolution $f * g$ est la fonction définie pour (presque) tout $x \in \mathbb{R}$ et telle que

$$f * g(x) = \int f(x-y)g(y)dy$$

Par exemple, si f et g sont des fonctions appartenant respectivement aux espaces $L^1(\mathbb{R})$ et $L^\infty(\mathbb{R})$, alors elles sont convolables : de plus, il est immédiat que $f * g(x)$ est bien définie pour tout $x \in \mathbb{R}$ et appartient à $L^\infty(\mathbb{R})$: nous allons voir comment la convolution entraîne ici un effet de régularisation.

Proposition 3.1. *Si f est une fonction de $L^\infty(\mathbb{R})$, alors pour toute fonction $U \in L^1(\mathbb{R})$, la fonction $U * f = f * U$ est un élément uniformément continu de $L^\infty(\mathbb{R})$.*

Preuve. Par les propriétés élémentaires de l'intégrale de Lebesgue (formule de changement de variable linéaire), il est immédiat que $U * f(x) = f * U(x)$ est bien défini pour tout x et que la fonction $U * f = f * U$ est bornée sur $\mathbb{R}$. Pour démontrer l'uniforme continuité

11. Nous n'abordons pas ici ce point très important.

de $U * f$, nous commençons par supposer que U est une fonction de $L^1(\mathbb{R})$ qui est uniformément continue sur $\mathbb{R}$. Pour $\varepsilon > 0$ arbitraire, nous notons $\eta_U(\varepsilon)$ le module de continuité de U pour la jauge ε (i.e. $\eta_U(\varepsilon) := \max\{|U(x) - U(y)|\ ;\ |x - y| \le \varepsilon\}$), de sorte que :

$$\left|U * f(x) - U * f(x+\varepsilon)\right| = \left|\int \big(U(x-t) - U(x+\varepsilon-t)\big) f(t)dt\right| \le \eta_U(\varepsilon)\|f\|_\infty$$

Cela assure l'uniforme continuité de $U * f$. Nous utilisons maintenant la densité [12] dans $L^1(\mathbb{R})$ de l'espace $\mathscr{C}_C(\mathbb{R})$ de fonctions réelles continues sur $\mathbb{R}$ et à support compact : ainsi, pour $U \in L^1(\mathbb{R})$ arbitrairement donnée, nous pouvons considérer une suite $U_1, U_2, \dots$ de fonctions dans $\mathscr{C}_C(\mathbb{R})$ t.q. $\|U - U_k\|_1 \to 0$ quand $k \to +\infty$. Pour tout $x \in \mathbb{R}$, nous avons alors

$$\begin{aligned} \left|U * f(x) - U_k * f(x)\right| &= \left|\int \big(U(x-t) - U_k(x-t)\big) f(t)dt\right| \\ &\le \|f\|_\infty \int |U(x-t) - U_k(x-t)|dt = \|f\|_\infty \|U - U_k\|_1 \end{aligned}$$

Les U_k étant étant uniformément continues sur $\mathbb{R}$, nous savons d'après la première partie de la démonstrations, que chacune des $U_k * f$ est uniformément continue sur $\mathbb{R}$: l'inégalité $\|U * f - U_k * f\|_\infty \le \|f\|_\infty \|U - U_k\|_1$ assure alors que la fonction $U * f$ est elle aussi uniformément continue sur $\mathbb{R}$

□

Les propriétés d'approximation liées à la convolution sont basées sur la notion de *« noyau de sommabilité »* que nous définissons maintenant.

Définition 3.2. *Une suite $(U_0, U_1, \dots)$ de fonctions de $L^1(\mathbb{R})$ est appelée un noyau de sommabilité lorsque les trois propriétés suivantes sont vérifiées, soient :*

- (i) *:* $\int U_n(x)dx = 1$ *;*
- (ii) *:* $\sup_n\{\|U_n\|_1\} < +\infty$ *;*
- (iii) *: pour tout* $\varepsilon > 0$, $\int_{|x|\ge\varepsilon} |U_n(x)|dx = 0 \to 0$ *quand* $\varepsilon \to 0$.

(dans le cas où les fonctions U_k sont positives où nulles, la condition (ii) est redondante).

La proposition suivante affirme que toute fonction uniformément continue $f \in L^\infty(\mathbb{R})$ peut être uniformément approchée par une suite de fonctions uniformément continues de la forme $U_k * f$, où $(U_1, U_2, \dots)$ est un noyau de sommabilité.

Proposition 3.3. *Soit $f : \mathbb{R} \to \mathbb{R}$ uniformément continue et bornée sur $\mathbb{R}$: si $(U_1, U_2, \dots)$ est un noyau de sommabilité, alors $U_n * f$ converge vers f uniformément sur $\mathbb{R}$ quand $n \to +\infty$.*

Preuve. Soit $\varepsilon > 0$ arbitrairement donné : alors par uniforme continuité de f, il existe $\eta > 0$ t.q. $|f(x) - f(y)| \le \varepsilon/2$ dés que $|x - y| \le \eta$. Pour tout $x \in \mathbb{R}$ nous avons

$$\begin{aligned} \left|U_n * f(x) - f(x)\right| &= \left|\int U_n(t)\big(f(x-t) - f(x)\big)dt\right| \\ &\le \int_{|t|\le\eta} U_n(t)\big|f(x-t) - f(x)\big|dt + \int_{|t|>\eta} U_n(t)\big|f(x-t) - f(x)\big|dt \end{aligned}$$

Par définition du module de continuité η et en utilisant le fait que f est bornée, il vient

$$\left|U_n * f(x) - f(x)\right| \le \frac{\varepsilon}{2}\int U_n(t)dt + 2\|f\|_\infty \int_{|t|>\eta} U_n(t)dt$$

12. Résultat de la théorie de la mesure que nous admettons ici.

La conclusion découle de l'existence d'un N t.q. , pour tout $n \geq N$

$$\int_{|t|>\eta} U_n(t)dt \leq \frac{\varepsilon}{4\|f\|_\infty}$$

□

Remarque 3.4. *L'intérêt de la Proposition 3.3 peut sembler (à priori) très discutable du fait (c.f. Proposition 3.1) que nous approximons une fonction bornée et uniformément continue sur $\mathbb{R}$ (soit f) par des fonctions bornées et uniformément continues sur $\mathbb{R}$ (soient $U_k * f$, pour $k = 1, 2, \ldots$) ! Le point important (et que nous utiliserons par exemple pour démontrer le théorème d'approximation de Weierstrass : c.f. § 4.1) est que les fonctions $U_k * f$ ont la même régularité que les fonctions U_k (l'idée générale est que la convolée de deux fonctions possède la régularité de la plus régulière des deux fonctions). Pour illustrer cela, notons par exemple que si U est une une fonction dans l'espace $\mathscr{C}_c^n(\mathbb{R})$ des fonctions n $(= 1, 2, \ldots, +\infty)$ fois continûment dérivables sur $\mathbb{R}$ et à support compact, alors (c.f.* [GW90, § 21.2]*) pour tout $f \in L^1(\mathbb{R})$ la fonction $U * f$ est un élément de $L^1(\mathbb{R}) \cap \mathscr{C}^n(\mathbb{R})$ et que de plus, pour tout $1 \leq k \leq n$:*

$$\frac{d^k}{dx^k} U * f = \left(\frac{d^k}{dx^k} U\right) * f$$

3.2. Considérons maintenant que f et g sont toutes deux des fonctions de $L^1(\mathbb{R})$; alors d'après le théorème de Fubini-Tonelli (et par changement de variable) nous avons

$$\iint |f(x-y)g(y)|dxdy = \int \left(\int |f(x-y)|dx\right)|g(y)|dy = \|f\|_1\,\|g\|_1$$

de sorte que f et g sont convolables (i.e. $y \mapsto |f(x-y)g(y)|$ est dans $L^1(\mathbb{R})$ pour presque tout $x \in \mathbb{R}$) avec

$$\int |f * g(x)|dx = \int \left|\int f(x-y)g(y)dy\right| dx \leq \iint |f(x-y)g(y)|dxdy \leq \|f\|_1\,\|g\|_1 < +\infty$$

Proposition 3.5. *Si f et g sont deux fonctions de $L^1(\mathbb{R})$ alors la convolution $f * g = g * f$ est une fonction de $L^1(\mathbb{R})$ telle que*

$$\|f * g\|_1 \leq \|f\|_1\,\|g\|_1$$

La Proposition 3.5 possède plusieurs extensions possibles. Ainsi notons que pour $1 \leq p, q \leq +\infty$ t.q. $1/p + 1/q = 1$, l'inégalité de Hölder nous donne

$$\begin{aligned}|f| * |g|(x) &\leq \int |f|(x-y)|g|(y)dy \\ &\leq \left(\int |f|^p(x-y)dy\right)^{1/p}\left(\int |g|^q(y)dy\right)^{1/q} \leq \left(\int |f|^p(z)dz\right)^{1/p}\left(\int |g|^q(y)dy\right)^{1/q}\end{aligned}$$

Par suite, lorsque $f \in L^p(\mathbb{R})$ et $g \in L^q(\mathbb{R})$ nous avons $|f| * |g|(x) \leq \|f\|_p \cdot \|g\|_q$: en d'autres termes $f * g$ est une fonction de $L^\infty(\mathbb{R})$. Les deux résultats précédents sur l'existence de la convolution ne sont apparemment pas compatibles (puisque le conjugué de $p = 1$ est $q = +\infty$) : ils sont en fait des cas de *« l'inégalité de Young »* que nous allons démontrer.

Théorème 3.6 (Inégalité de Young). *Soit $0 \leq p, q, r \leq +\infty$ t.q.*

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1$$

Si $f \in L^p(\mathbb{R})$ et $g \in L^q(\mathbb{R})$, alors f et g sont convolables et de plus $f * g \in L^r(\mathbb{R})$ avec

$$\|f * g\|_r \leq \|f\|_p \|g\|_q$$

Preuve. Il suffit de considérer que f et g sont positives ou nulles en presque tout point. Le cas où $r = +\infty$ se traitant directement grâce à l'inégalité de Hölder classique, nous supposons que $r < +\infty$. Alors il est nécessaire que $\max\{p, q\} \leq r < +\infty$. Nous pouvons appliquer l'inégalité de Hölder à trois facteurs (c.f. Corollaire 2.6) avec le triplet d'exposants

$$(\alpha, \beta, \gamma) = \left(r, \frac{rp}{r-p}, \frac{rq}{r-q}\right)$$

En effet on vérifie immédiatement que $1/\alpha + 1/\beta + 1/\gamma = 1$ et par suite :

$$\begin{aligned} f * g(x) &= \int f^{\frac{p}{r}}(x-y) g^{\frac{q}{r}}(y) f^{1-\frac{p}{r}}(x-y) g^{1-\frac{q}{r}}(y) dy \\ &\leq \left(\int \left(f^{\frac{p}{\alpha}}(x-y) g^{\frac{q}{\alpha}}(y)\right)^{\alpha} dy\right)^{\frac{1}{\alpha}} \left(\int \left(f^{\frac{p}{\beta}}(x-y)\right)^{\beta} dy\right)^{\frac{1}{\beta}} \left(\int \left(g^{\frac{q}{\gamma}}(y)\right)^{\gamma} dy\right)^{\frac{1}{\gamma}} \\ &= \left(\int f^{p(x-y)} g^q(y) dy\right)^{\frac{1}{r}} \|f\|_p^{1-\frac{p}{r}} \|g\|_q^{1-\frac{q}{r}} \end{aligned}$$

En élevant à la puissance $r = \alpha$ puis en intégrant, il vient (Fubini-Tonelli)

$$\begin{aligned} \int \left(f * g(x)\right)^r dx &\leq \left(\iint f(x-y)^p g(y)^q dy dx\right) \|f\|_p^{r-p} \|g\|^{r-q} \\ &= \left(\int \left(\int f(x-y)^p dx\right) g(y)^q dy\right) \|f\|_p^{r-p} \|g\|^{r-q} \\ &= \left(\int f(z)^p dz\right) \left(\int g(y)^q dy\right) \|f\|_p^{r-p} \|g\|^{r-q} = \|f\|_p^r \|g\|^r \end{aligned}$$

□

3.3. Comme d'après le théorème de Riesz-Fisher nous savons que l'espace $(L^1(\mathbb{R}), \|\cdot\|)$ est un espace de Banach, nous pouvons déduire de la Proposition 3.5 que :

Corollaire 3.7. *$L^1(\mathbb{R})$ muni de la convolution est une $\mathbb{R}$-algèbre de Banach commutative.*

Notons au passage que $L^1(\mathbb{R})$ est un exemple d'algèbre de Banach (commutative) non unitaire (son unité devrait être l'impulsion de Dirac en 0). Il existe une notion naturelle permettant de palier l'absence d'unité d'une algèbre de Banach : ainsi, une suite $(E_0, E_1, \dots)$ d'éléments d'une algèbre de Banach $\mathscr{A}$ est appelée *« une approximation de l'unité à gauche (resp. à droite) »* lorsque pour tout $A \in \mathscr{A}$ nous avons $\|E_n A - A\| \to 0$ (resp. $\|A E_n - A\| \to 0$) quand $n \to +\infty$. Le théorème suivant est l'analogue de la Proposition 3.3 pour la convergence $L^p(\mathbb{R})$: dans le cas $p = 1$ il signifie que les noyaux de sommabilité tels que donnés dans la Définition 3.2 sont en fait des approximations de l'unité de l'algèbre de Banach $L^1(\mathbb{R})$ pour la convolution.

Théorème 3.8. *Si $(U_0, U_1, \dots)$ est un noyau de sommabilité, alors pour tout $1 \leq p < +\infty$ et $f \in L^p(\mathbb{R})$ nous avons $\|U_n * f - f\|_p \to 0$ quand $n \to +\infty$. (En particulier pour $p = 1$ un noyau de sommabilité est une approximation de l'unité de l'algèbre $L^1(\mathbb{R})$ pour la convolution.)*

Lemme 3.9. *Pour tout $a \in \mathbb{R}$, nous notons $T_a f(x) = f(x-a)$ – opérateur de translation spatiale ; si $f \in L^p(\mathbb{R})$ pour tout $1 \leq p < +\infty$ alors $\|T_a f - f\|_p \to 0$ quad $a \to 0$.*

Preuve. Par uniforme continuité, le résultat est vrai dans l'espace $\mathscr{C}_c(\mathbb{R})$ des fonctions continues sur $\mathbb{R}$ à support compact. Maintenant, considérons que $f \in L^p(\mathbb{R})$ et soit $\varepsilon > 0$ arbitrairement donné. Puisque [13] $\mathscr{C}_c(\mathbb{R})$ est dense dans $L^p(\mathbb{R})$, il existe $g \in \mathscr{C}_c(\mathbb{R})$ telle que $\|f - g\| \leq \varepsilon/4$. Or nous avons vu qu'il existe $\alpha > 0$ tel que $\|T_a g - g\|_p \leq \varepsilon/2$, pour tout $a \in \mathbb{R}$ t.q. $\|a\| \leq \alpha$; pour un tel a nous pouvons alors écrire

$$\|T_a f - f\|_p \leq \|T_a f - T_a g\|_p + \|T_a g - g\|_p + \|g - f\|_p = 2\|f - g\|_p + \|T_a g - g\|_p \leq \varepsilon$$

□

Preuve du Théorème 3.8. Soit $f \in L^p$. Comme $1/1 + 1/p = 1/p + 1$ le fait que U_n soit dans $L^1(\mathbb{R})$ assure (c.f. Théorème 3.6) que $f * U_n$ est une fonction définie presque partout et appartenant à $L^p(\mathbb{R})$. Par définition d'une approximation de l'unité nous avons $f(x) = \int f(x)U_n(y)dy$; par suite, si nous notons q le conjugué de p (i.e. $1/p + 1/q = 1$), alors pour tout rang n et presque tout x nous avons (inégalité de Hölder.)

$$\begin{aligned}
|f * U_n(x) - f(x)| &\leq \int |f(x-y) - f(x)|\,|U_n(y)|dy \\
&= \int \left(|f(x-y) - f(x)|\,|U_n(y)|^{1/p}\right)|U_n(y)|^{1/q}dy \\
&= \left(\int |f(x-y) - f(x)|^p\,|U_n(y)|dy\right)^{1/p}\left(\int |U_n(y)|dy\right)^{1/q}
\end{aligned}$$

En intégrant la puissance p-ème, le théorème de Fubini-Tonelli nous donne :

$$\begin{aligned}
\|f * U_n - f\|_p^p &\leq \left(\iint |f(x-y) - f(x)|^p\,|U_n(y)|dydx\right)\|U_n\|_1^{p/q} \\
&\leq K^{p/q}\int\left(\int |f(x-y) - f(x)|^p dx\right)|U_n(y)|dy
\end{aligned}$$

où nous avons noté $K := \sup_n\{\|U_n\|_1\}$ (constante finie par définition d'une approximation de l'unité). En utilisant l'opérateur de translation nous avons alors :

$$\|f * U_n - f\|_p^p \leq K^{p/q}\int \|T_y f - f\|_p^p\,|U_n(y)|dy \tag{3}$$

Etant donné $\varepsilon > 0$, nous savons (c.f. Lemme 3.9), qu'il existe $\eta > 0$ tel que $\|T_y f - f\|_p \leq \varepsilon$ dès que $\|y\| \leq \eta$. Nous pouvons alors transformer (3) et obtenir pour tout rang n :

$$\begin{aligned}
\|f * U_n - f\|_p^p &\leq K^{p/q}\left(\int_{|y|\leq\eta}\|T_y f - f\|_p^p\,|U_n(y)|dy + \int_{|y|>\eta}\|T_y f - f\|_p^p\,|U_n(y)|dy\right) \\
&\leq K^{p/q}\left(\varepsilon^p\int_{|y|\leq\eta}|U_n(y)|dy + 2^p\|f\|_p^p\int_{|y|>\eta}|U_n(y)|dy\right) \\
&\leq K^{p/q}\left(\varepsilon^p K + 2^p\|f\|_p^p\int_{|y|>\eta}|U_n(y)|dy\right)
\end{aligned}$$

En prenant la limite-sup, quand $n \to +\infty$, la propriété (iii) d'une approximation de l'unité nous permet de conclure que $\limsup_n \|f * U_n - f\|_p^p \leq \varepsilon^p K^{p/q+1}$.

□

3.4. Nous terminons cette courte introduction aux techniques de convolution par une notion de convolution adaptée aux fonctions 2π-périodiques [14]. Deux fonctions réelles f

13. Résultat de la Théorie de la mesure que nous admettons ici.

14. Le choix de la période 2π parmi toutes les périodes possibles est arbitraire ; il est cependant motivé par le fait que la longueur du circle unité vaut 2π ; le terme circulaire est systématiquement utilisé pour préciser (souvent implicitement) que les fonctions considérées sont 2π-périodiques.

et g définies sur $\mathbb{R}$, boréliennes et 2π-périodiques sont dites *« circulairement convolables »* si la fonction $y \mapsto f(x-y)g(y)$ est localement intégrable pour presque tout $x \in \mathbb{R}$: dans ce cas *« la convolution circulaire »* $x \mapsto f \circledast g(x)$ de ces deux fonctions est par définition la fonction 2π-périodique [15]

$$f \circledast g(x) := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-t)g(t)dt = \left(\frac{f}{2\pi}\right) * \left(g \cdot \mathbf{1}_{[-\pi\,;\,\pi]}\right) \tag{4}$$

Dans ce paragraphe, nous identifions l'espace de Banach $L^1[-\pi\,;\pi]$ à l'espace des fonctions réelles définies sur $\mathbb{R}$ boréliennes, 2π-périodiques et localement intégrables. La définition de la convolution circulaire sur les différents espaces $L^p[-\pi\,;\pi]$ (avec $1 \le p \le +\infty$) est simplifiée par le fait que $L^p[-\pi\,;\pi]$ est un sous-espace de $L^1[-\pi\,;\pi]$. La proposition suivante est une conséquence directe de la Proposition 3.5.

Proposition 3.10. *Si f et g sont deux fonctions 2π-périodiques de $L^1[-\pi\,;\pi]$ alors la convolution circulaire $f \circledast g(x)$ de f et g est définie pour presque tout $x \in \mathbb{R}$; de plus $f \circledast g = g \circledast f$ est une fonction 2π-périodique de $L^1[-\pi\,;\pi]$.*

De manière analogue à la notion d'approximation de l'unité sur $\mathbb{R}$ (c.f. Théorème 3.8), nous avons la notion *« un noyau de sommabilité circulaire »*.

Définition 3.11. *Une suite $(U_1(x), U_2(x), \dots)$ de fonctions boréliennes supposées 2π-périodiques et localement intégrables sur $\mathbb{R}$ est appelée un noyau de sommabilité circulaire lorsque* [16]

(i) : $\frac{1}{2\pi}\int_{-\pi}^{\pi} U_n(x)dx = 1$;
(ii) : $\sup_n\{\|U_n\|_1\} < +\infty$;
(iii) : *pour tout* $\varepsilon > 0$, $\int_{\varepsilon}^{2\pi-\varepsilon} |U_n(x)|dx \to 0$ *quand* $n \to +\infty$.

(dans le cas où les fonctions U_n sont positives où nulles, la condition (ii) est redondante).

Le théorème suivant est l'analogue *circulaire* du Théorème 3.8 (dans le cas $p = 1$).

Théorème 3.12. *Si $(U_0, U_1, \dots)$ est un noyau de sommabilité circulaire, alors pour tout $f \in L^1[-\pi\,;\pi]$ nous avons $\|f \circledast U_n - f\|_1 \to 0$ quand $n \to +\infty$: tout noyau de sommabilité circulaire est une approximation de l'unité de l'algèbre $L^1[-\pi\,;\pi]$ pour la convolution circulaire.*

Le proposition suivante est la version circulaire de la Proposition 3.3 pour la convolution circulaire (elle permet de démontrer facilement la version trigonométrique du théorème d'approximation de Weierstrass grâce au théorème de Féjer : c.f. § 4.4).

---

15. Le facteur $\frac{1}{2\pi}$ présent en (4) assure que la transformée de Fourier transforme la convolution circulaire des fonctions en produits des transformées de Fourier. Plus précisément, si nous définissons la transformée de Fourier $\mathfrak{F}[f]$ d'une fonction $f \in L^1[-\pi\,;\pi]$ comme la suite $\mathfrak{F}[f] = (c_n(f))_{n\in\mathbb{Z}}$, où

$$c_n(f) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-2i\pi x}dx$$

alors $\mathfrak{F}[\cdot]$ réalise une application de $L^1[-\pi\,;\pi]$ dans l'espace des suites bilatérales de nombre complexes. D'autre part, si $f, g \in L^1[-\pi\,;\pi]$, alors $f \circledast g$ est dans $L^1[-\pi\,;\pi]$ et (avec les précautions d'usages) :

$$\begin{aligned} c_n(f \circledast g) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f \circledast g(x)e^{-inx}dx \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-y)e^{-in(x-y)}dx\right)g(y)e^{-iny}dy = \frac{1}{2\pi}\int_{-\pi}^{\pi} c_n(f)g(y)e^{-iny}dy \end{aligned}$$

En d'autres termes, $c_n(f \circledast g) = c_n(f)c_n(g)$, soit encore $\mathfrak{F}[f \circledast g] = \mathfrak{F}[f] \cdot \mathfrak{F}[g]$.

16. La présence du facteur $\frac{1}{2\pi}$ dans la condition de normalisation des U_n est due à la définition de la convolution circulaire en (4).

Proposition 3.13. *Soit $f(x)$ 2π-périodique continue sur $\mathbb{R}$ et soit $(U_0(x), U_1(x), \dots)$ une approximation circulaire de l'unité : alors les fonctions $U_n \circledast f$ $(n \geq 0)$ sont des fonctions 2π-périodiques continues sur $\mathbb{R}$ qui convergent uniformément sur $\mathbb{R}$ quand $n \to +\infty$.*

4. Théorème de Stone-Weierstrass

4.1. Le fait que les espaces de Hilbert séparables possèdent tous une base hilbertienne dénombrable et un point particulièrement important : cela prouve en un sens qu'il n'existe qu'un seul espace Hilbert séparable : nous reviendrons sur ce point dans le paragraphe suivant entièrement consacré aux espaces de Hilbert (c.f. Remarque 5.9 infra). Nous commençons par traiter ici les *« théorèmes d'approximation de Weierstrass et de Stone-Weierstrass »*.

Théorème 4.1 (Weierstrass classique)**.** *Si f est une fonction réelle continue sur $[0\,;1]$ alors pour tout $\varepsilon > 0$, il existe un polynôme $P(X) \in \mathbb{R}[X]$ t.q. $|f(x) - P(x)| \leq \varepsilon$, pour tout $x \in [0\,;1]$.*

Nous allons donner plusieurs preuves de ce théorème. La première est basée sur la Proposition 3.3 qui assure l'approximation uniforme des fonctions continues sur la droite réelle par convolution avec *« un noyau de sommabilité »*.

Preuve du Théorème 4.1 (n°1). Etant donnée $f(x)$ une fonction (réelle) définie et continue sur $[0\,;1]$, nous définissons la fonction $g(x)$ telle que $g(x) = 0$ pour $x \in \mathbb{R} \setminus [0\,;1]$ et $g(x) = f(x) - A(x)$ – avec $A(x) = f(0) + \big(f(1) - f(0)\big)x$ – pour $x \in [0\,;1]$. Par définition $g(x)$ est une fonction bornée et uniformément continue sur $\mathbb{R}$. Si pour $n \geq 1$ nous posons

$$U_n(x) := \frac{\max\left\{0\,;(1-x^2)^n\right\}}{\int_{-1}^{+1}(1-y^2)^n dy}$$

alors $(U_1, U_2, \dots)$ est une approximation de l'unité et la Proposition 3.3, assure que

$$U_n * g(x) = \int U_n(y)g(x-y)dy = \int_0^1 U(x-y)g(y)dy$$

converge uniformément vers $g(x)$. Or il existe $2n+1$ fonctions polynomiales en y, soient $a_0(y), \dots, a_{2n}(y)$, telles que

$$\forall x, y \in \mathbb{R},\ x - y \in [-1\,;1] \Longrightarrow U_n(x-y) = \sum_{k=0}^{2n} a_k(y)x^k$$

de sorte que

$$U_n * g(x) = \sum_{k=0}^{2n}\left(\int_0^1 a_k(y)g(y)dy\right)x^k =: B_n(x)$$

Nous avons ainsi obtenu une suite $B_1(x), B_2(x), \dots$ de fonctions polynomiales définies pour tout $x \in [0\,;1]$ qui convergent uniformément vers $g(x)$ sur $[0\,;1]$: cela assure que $A(x) + B_n(x)$ converge uniformément vers $f(x)$ sur $[0\,;1]$ quand $n \to +\infty$.

□

4.2. Une autre démonstration constructive du théorème de Weierstrass et celle proposée par Bernstein : ainsi, toute fonction continue $f : [0\,;1] \to \mathbb{R}$ est-elle associée au polynôme

$$P_n f(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) B_{n,k}(x)$$

où les $B_{n,k}$ sont *« les polynômes de Bernstein* définis en posant

$$B_{n,k}(x) = \mathsf{C}_n^k x^k (1-x)^{n-k}$$

Preuve du Théorème 4.1 (n°2). Comme f est continue sur un intervalle compact, c'est une fonction bornée et uniformément continue. Ainsi $\max\{|f(x)|\ ;\ 0 \leq x \leq 1\} = \|f\|_\infty < +\infty$ et pour tout $\varepsilon > 0$ arbitrairement donné, il existe $\eta > 0$ t.q. $|f(x) - f(y)| < \varepsilon/2$ dès que $|x - y| < \eta$. D'après *« la formule du binôme de Newton »* (et du fait que $0 \leq x \leq 1$) il est clair que $\sum_{k=0}^{n} B_{n,k}(x) = 1$: en fait $B_{n,k}(x)$ s'interprète comme la probability d'avoir k réalisations d'un événement au cours de n expériences indépendantes et où x est la probabilité de l'événement au cours d'une seule expérience. Dans ce modèle probabiliste, nous notons $(\Omega, \mathbb{P})$ l'espace probabilisé associé à la variable aléatoire $X_n : \Omega \to \mathbb{R}$ telle que $\mathbb{P}\{X_n = k\} = B_{n,k}(x)$ pour tout $0 \leq k \leq n$. En utilisant le fait que l'espérance et la variance de X_n valent respectivement nx et $nx(1-x)$ (c.f. Remarque 8.16 infra), nous pouvons déduire de *« l'inégalité de Tchebychev »* (c.f. § 8.14 infra) que

$$\sum_{|k-nx|\geq\lambda} B_{n,k}(x) = \mathbb{P}\Big\{|X_n - nx| \geq \lambda\Big\} \leq \frac{nx(1-x)}{\lambda^2} \leq \frac{n}{\lambda^2} \tag{5}$$

Maintenant nous pouvons écrire que

$$\begin{aligned} |f(x) - P_n f(x)| &= \sum_{|x-k/n|<\eta} \left|f(x) - f\left(\frac{k}{n}\right)\right| B_{n,k}(x) + \sum_{|x-k/n|\geq\eta} \left|f(x) - f\left(\frac{k}{n}\right)\right| B_{n,k}(x) \\ &\leq \frac{\varepsilon}{2} + 2\|f\|_\infty \sum_{|nx-k|\geq n\eta} B_{n,k}(x) \end{aligned}$$

soit encore, d'après (5)

$$|f(x) - P_n f(x)| \leq \frac{\varepsilon}{2} + \frac{2\|f\|_\infty}{n\eta^2}$$

Lorsque $n > 4\|f\|_\infty/(\varepsilon\eta^2)$, nous obtenons bien que $|f(x) - P_n f(x)| \leq \varepsilon$.

□

4.3. Un polynôme trigonométrique est une fonction complexe $T(x)$ de la variable réelle x qui peut s'écrire sous la forme exponentielle

$$T(x) = \sum_{k=-n}^{n} c_k e^{ikx}$$

où les coefficients c_k sont les nombres complexes tels que

$$c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} T(x) e^{-ikx} dx$$

(le polynôme trigonométrique $T(x)$ est dit réel si $T(x) \in \mathbb{R}$, pour tout $x \in \mathbb{R}$).

Théorème 4.2 (Weierstrass trigonométrique). *Si f est une fonction réelle (resp. complexe) définie sur $\mathbb{R}$, continue et 2π-périodique, alors pour tout $\varepsilon > 0$, il existe un polynôme trigonométrique réel (resp. complexe) $T(x)$ telle que $|f(x) - T(x)| \leq \varepsilon$, pour tout $x \in \mathbb{R}$.*

Preuve. Le cas complexe se déduit du cas réel ; il suffit donc d'établir le résultat pour $f(x)$ une fonction 2π-périodique, continue sur $\mathbb{R}$ et à valeur réelle. Pour une telle fonction nous pouvons toujours écrire $f(x) = f_P(x) + f_I(x)$ où $f_P(x)$ (resp. $f_I(x)$) est une fonction continue 2π-périodique paire (resp. impaire) : pour cela il suffit de prendre

$$f_P(x) := \frac{f(x)+f(-x)}{2} \quad \text{et} \quad f_I(x) := \frac{f(x)-f(-x)}{2}$$

Nous pouvons alors écrire $f_P(x) = F_P(\cos(x))$, où $F_P(x)$ est une fonction définie et continue sur $[-1;1]$ et $f_I(x) = F_I(\sin(x))$, où $F_I(x)$ est une fonction définie, continue et impaire sur $[-1;1]$. D'après la version classique du théorème d'approximation de Weierstrass (c.f. Théorème 4.1), pour $\varepsilon > 0$ arbitrairement donné, il existe un polynôme $A(X) = \sum_{k=0}^{a} \alpha_k X^k \in \mathbb{R}[X]$ tel que $|F_P(x) - A(x)| \le \varepsilon/2$, pour tout $-1 \le x \le +1$: par suite $|f_P(x) - A(\cos x)| \le \varepsilon/2$, pour tout $x \in \mathbb{R}$. De même, il existe un polynôme $B(X) = \sum_{k=0}^{b} \beta_k X^k \in \mathbb{R}[X]$ tel que $|f_I(x) - B(\sin x)| \le \varepsilon/2$, pour tout $x \in \mathbb{R}$. Finalement en posant $T(x) := A(\cos x) + B(\sin x)$, il vient pour tout $x \in \mathbb{R}$

$$\left|f(x) - T(x)\right| \le \left|f_P(x) - A(\cos x)\right| + \left|f_I(x) - B(\sin x)\right| \le \varepsilon$$

Or grâce aux *« formules d'Euler »* nous avons

$$T(x) = A(\cos x) + B(\sin x) = \sum_{p=0}^{a} \alpha_p \left(\frac{e^{\boldsymbol{i}x} + e^{-\boldsymbol{i}x}}{2}\right)^p + \sum_{p=0}^{b} \beta_p \left(\frac{e^{\boldsymbol{i}x} - e^{-\boldsymbol{i}x}}{2\boldsymbol{i}}\right)^p$$

où nous reconnaissons un polynôme trigonométrique (nécessairement réel).

□

Nous venons de déduire la version trigonométrique du théorème d'approximation de Weierstrass à partir de sa version classique ; réciproquement il est possible de déduire la version classique de la version trigonométrique. Pour cela, supposons (pour simplifier) que $f : [-\pi;\pi] \to \mathbb{R}$ soit une fonction continue ; quitte à retrancher une fonction affine (i.e. polynomiale de degré 1), nous pouvons supposer que $f(-\pi) = f(\pi) = 0$ et noter $f_P : \mathbb{R} \to \mathbb{R}$ la fonction continue et 2π-périodique t.q. $f_P(x) = f(x)$ pour tout $-\pi \le x \le \pi$. D'après le théorème de Weierstrass trigonométrique, pour tout $\varepsilon > 0$ il existe alors un polynôme trigonométrique à valeurs réelles, soit $T(x) = \sum_{k=0}^{n} c_k e^{\boldsymbol{i}kx}$, t.q. pour tout $x \in \mathbb{R}$.

$$|f_P(x) - T(x)| \le \varepsilon/2 \tag{6}$$

Nous utilisons ici le fait que le polynôme trigonométrique $T(x)$ peut être considéré comme la restriction à l'axe réel d'une fonction $T(z)$ de la variable complexe z, définie et analytique sur tout le plan complexe. En d'autre termes, $T(z)$ coïncide avec la somme $\sum_{k=0}^{\infty} \xi_k z^k$ d'une série entière de rayon de convergence infini : mais par définition [17] le rayon de convergence est le supremum des nombres réels R de $[0;+\infty]$ tels que la suite des sommes partielle $S_r(z) := \sum_{k=0}^{r} \xi_k z^k$ soient uniformément convergentes dans le disque ouvert $\{|z| < R\}$ (lorsque $p \to +\infty$). Mais le rayon de convergence de $T(z) = \sum_{k=0}^{\infty} \xi_k z^k$ étant infini, les fonctions polynomiales $S_r(z)$ convergent uniformément vers $T(z)$ sur le disque fermé $\{|z| \le \pi\}$ quand $r \to +\infty$: par suite, il existe un rang r t.q. $|T(x) - S_r(x)| \le \varepsilon/2$ pour tout $x \in [-\pi;\pi]$, de sorte qu'avec (6) il vient : $|f(x) - S_p(x)| = |f_P(x) - T(x)| + |T(x) - S_p(x)| \le \varepsilon$.

17. Usuellement c'est plutôt une propriété du rayon de convergence (défini par exemple par la formule d'Hadamard) : nous avons choisi de prendre cette propriété comme définition afin de simplifier notre présentation (pour une présentation classique sur les propriétés de convergence uniforme des séries entières, on pourra consulter [Gam01, Chap. V, § 3]).

4.4. Une autre manière d'aborder le théorème d'approximation de Weierstrass est de considérer directement la question de l'approximation uniforme des fonctions 2π-périodiques par des polynômes trigonométriques. Cette approche remonte (au moins) à la polémique entre D'Alembert Euler et Daniel Bernoulli sur la représentations des fonctions permettant une résolution de l'équation de la corde vibrante (voir les mémoires de D'Alembert de 1747 [D'A47b, D'A47a]), puis avec le mémoire [Fou22] publié en 1822 par Fourier sur l'équation de la chaleur (voir aussi le magnifique texte de Jean-Pierre Kahane[18] sur les séries de Fourier dans [KLR98, p. 3 - 279]). Concernant le théorème d'approximation de Weierstrass, nous nous limiterons aux deux piliers de la théorie que sont *« le théorème de Dirichlet »* et *« le théorème de Féjer »*. Pour cela, commençons par noter qu'à toute fonction complexe f localement intégrable et 2π-périodique est associée le polynôme trigonométrique

$$S_n f(x) := \sum_{k=-n}^{n} c_k(f) e^{ikx} \quad \text{où} \quad c_k(f) := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-ikt}\, dt$$

Avec une petite transformation nous pouvons écrire

$$S_n f(x) := \sum_{k=-n}^{n} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-ikt}\, dt \right) e^{ikx} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \left(\sum_{k=-n}^{n} e^{ik(x-t)} \right) dt$$

soit encore, (en utilisant la *« convolution circulaire »* : c.f. § 3.4) $S_n f(x) = D_n \circledast f(x)$ où

$$D_n(x) := \sum_{k=-n}^{n} e^{ikx} = \frac{\sin\left((2n+1)\,\frac{x}{2}\right)}{\sin\left(\frac{x}{2}\right)}$$

est le *« le noyau de Dirichlet »* (d'ordre n). Une version primitive du *« théorème de Dirichlet »* s'énonce comme suit :

Théorème 4.3 (Dirichlet). *Soit f une fonction complexe localement intégrable sur $\mathbb{R}$ et 2π-périodique ; si f est continue par morceau et de classe C^1 par morceau, alors pour tout $x \in \mathbb{R}$*

$$\lim_{n \to +\infty} S_n f(x) = \frac{1}{2}\Big(f(x^+) + f(x^-) \Big)$$

Preuve. Voir [GW90, Théorème 5.2.4 & 5.2.5, p. 40] ou encore [ST90, § 3.3 p. 26]

□

En un sens, *« les moyennes de Cesàro » améliorent la convergence des suites numériques*. En partant du théorème de Dirichlet nous pouvons définir les moyennes de Cesàro successives de la suite $S_0 f(x), S_1 f(x), \ldots$ en posant pour tout $n \geq 1$:

$$\sigma_n f(x) = \frac{1}{n} \sum_{k=0}^{n-1} S_k f(x) = \left(\frac{1}{n} \sum_{k=0}^{n-1} D_k \right) \circledast f(x) = F_n \circledast f(x) \tag{7}$$

où (après calcul) nous obtenons *« le noyau de Féjer »* (fonction continue et 2π-périodique) :

$$F_n(x) = \frac{1}{n} \left(\frac{\sin\left(n \frac{x}{2}\right)}{\sin\left(\frac{x}{2}\right)} \right)^2$$

18. Jean-Pierre Kahane nous a quitté en 2017 : c'est une grande peine pour tous ceux qui ont eu la chance de le croiser. Il a été le premier contact avec l'Académie des Sciences de beaucoup de jeunes mathématiciens, pour la publication de leurs résultats de thèse *aux notes CRAS*. Je me souviens de ce sentiment intimidant et stimulant que j'ai pu ressentir lors de mes échanges avec ce grand mathématicien du 20-ème siècle : ces échanges était toujours empreint d'une sorte de rigueur à visage humain que je ne pourrai pas oublier.

La grande différence entre le noyau de Dirichlet $D_n(x)$ et le noyau de Féjer $F_n(x)$ vient du fait que les $F_1(x), F_2(x), \dots$ forment une approximation circulaire de l'unité (au sens de la Définition 3.11) : Le *« théorème de Féjer »*[19] découle de la Proposition 3.13.

Théorème 4.4 (Fejér – version uniforme)**.** *Si $f(x)$ est une fonction complexe 2π-périodique et continue sur $\mathbb{R}$, alors la série de fonctions $\sigma_n f(x)$ converge uniformément vers $f(x)$.*

Les fonctions $\sigma_n f(x)$ associées en (7) à une fonction f continue et 2π-périodique étant des polynômes trigonométriques, le théorème de Féjer établit de fait la version trigonométrique du théorème d'approximation de Weierstrass.

4.5. *« Le théorème de Stone-Weierstrass »* est une extension abstraite (et significativement plus générale) du théorème d'approximation de Weierstrass. Remarquons d'abord que l'ensemble des fonctions polynômiales (e.g. réelles) sur l'intervalle $[0\,;1]$ forme une sous-algèbre de $\mathscr{C}([0\,,1],\mathbb{R})$ (fonctions réelles) qui est unitaire (i.e. les fonctions constantes sur $[0\,;1]$ sont polynômiales) et qui séparent les points (on dit aussi séparante) en ce sens que pour $x \neq y$, il existe une fonction polynomiale $P(X) \in \mathbb{R}[X]$ t.q. $P(x) \neq P(y)$. Toutes ces remarques sont faciles à noter : il est remarquable qu'elles constituent les bonnes hypothèses d'un théorème très général[20]. Afin d'énoncer la version classique du théorème de Stone-Weierstrass, considérons que (X,d) est un espace métrique compact : alors, muni de la norme uniforme $f \mapsto \|f\|_\infty = \max\{|f(x)|\ ;\ x \in X\}$, l'espace $\mathscr{C}(X,\mathbb{K})$ des fonctions continues sur X et à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ est un espace de Banach. La multiplication des fonctions est une loi de composition interne sur $\mathscr{C}(X,\mathbb{K})$ qui en fait une $\mathbb{K}$-algèbre[21]. On dira qu'une sous-algèbre $\mathscr{A}$ de $\mathscr{C}(X,\mathbb{K})$ sépare les points de X lorsque pour tout couple x, y de points distincts de X il existe $f \in \mathscr{A}$ tel que $f(x) \neq f(y)$.

Théorème 4.5 (Stone-Weierstrass)**.** *Soient (X,d) un espace métrique compact et $\mathscr{A} \subset \mathscr{C}(X,\mathbb{R})$ une sous-algèbre unitaire séparante ; alors $\mathscr{A}$ est uniformément dense dans $\mathscr{C}(X,\mathbb{R})$.*

Preuve. Voir [Rud64, § 7.24 p. 146]

□

L'énoncé du théorème de Stone-Weierstrass sous sa forme réelle donné dans le Théorème 4.5, ne peut s'étendre tel quel au cas complexe. Par exemple si nous considérons $D := \{z \in \mathbb{C}\ ;\ |z| \leq 1\}$ le disque unité fermé du plan complexe et $D^\circ := \{z \in \mathbb{C} : |z| < 1\}$ son intérieur, alors l'ensemble $\mathscr{A}$ formé des polynômes complexes $P(z) = \sum_{k=0}^n \alpha_k z^k$ (avec $n \geq 0$ et $\alpha_k \in \mathbb{C}$) est une algèbre unitaire qui sépare les points de D. Il est cependant assez facile de vérifier que $\mathscr{A}$ n'est pas dense dans $\mathscr{C}(D,\mathbb{C})$. En effet, notons que $\mathscr{A}$ est une sous-algèbre de l'algèbre $\mathscr{H}$ formée des fonctions $f(z)$ qui sont analytiques (holomorphes) sur D° et continues sur D. Or $\mathscr{H}$ est un sous-ensemble propre de $\mathscr{C}(D,\mathbb{C})$ qui est fermé dans l'espace de Banach $(\mathscr{C}(D,\mathbb{C}), \|\cdot\|_\infty)$; en effet si $f_1(z), f_2(z), \dots$ est une suite de fonctions de $\mathscr{H}$ qui converge uniformément vers $f(z)$ dans $\mathscr{C}(D,\mathbb{C})$, alors $f(z)$ est par hypothèse continue sur D, la convergence uniforme sur D° assurant assurant

19. Féjer était âgé de 19 ans lorsqu'il a établi ce résultat.

20. Notons que l'hypothèse de séparation des points est importante : en effet l'algèbre unitaire des fonctions polynômiales $P : [0\,;1] \to \mathbb{R}$ t.q. $P(0) = P(1)$ n'est pas dense dans $\mathscr{C}([0\,,1],\mathbb{R})$ – la clôture uniforme de ces fonctions polynomiales est l'algèbre $\mathscr{C}_P([0\,,1],\mathbb{R})$ des fonctions continues $f : [0\,;1] \to \mathbb{R}$ t.q. $f(0) = f(1)$.

21. De plus, comme $\|fg\|_\infty \leq \|f\|_\infty \|g\|_\infty$, nous pouvons dire que $\mathscr{C}(X,\mathbb{K})$ est une algèbre de Banach.

aussi que $f(z)$ est analytique sur D° (c.f. [Gam01, Chap. V § 3]). La version complexe du théorème de Stone-Weirstrass est un corollaire immédiat de sa version réelle.

Corollaire 4.6 (Stone-Weierstrass complexe)**.** *Soient (X,d) un espace métrique compact et $\mathscr{A} \subset \mathscr{C}(X, \mathbb{C})$ une sous-algèbre unitaire, séparant les points de X et stable par conjugaison (i.e. $\bar{f} \in \mathscr{A}$ dès que $f \in \mathscr{A}$) ; alors $\mathscr{A}$ est uniformément dense dans $\mathscr{C}(X, \mathbb{C})$.*

A partir du théorème de Stone-Weierstrass complexe, il est possible de retrouver le théorème d'approximation de Weierstrass dans sa version trigonométrique pour les fonctions complexes continues 2π-périodiques. En effet si nous notons $\mathbb{T} = \{z \in \mathbb{C}\ ;\ |z| = 1\}$ le 1-tore complexe, alors l'algèbre des polynômes complexes sur $\mathbb{T}$ n'est pas stable par conjugaison complexe (effet $\bar{z}$ n'est pas un polynôme complexe). Par contre en notant $\mathscr{A}$ l'ensemble des fonctions de la forme $g(z) = 1/z^N P(z)$, où $P(X)$ est un polynôme de $\mathbb{C}[X]$ et où N est un entier ≥ 0, alors il est facile de vérifier que $\mathscr{A}$ est une sous-algèbre de $\mathscr{C}(\mathbb{T}, \mathbb{C})$ unitaire séparante et stable par conjugaison. Par une application directe du théorème de Stone-Weierstrass complexe nous retrouvons l'analogue complexe du théorème d'Approximation de Weierstrass affirmant que si $g(z)$ est une fonction complexe définie et continue sur $\mathbb{T}$, alors pour tout $\varepsilon > 0$, il existe une polynôme $P(X) \in \mathbb{C}[X]$ et un entier $N \geq 0$ tels que $\left|g(z) - P(z)/z^N\right| \leq \varepsilon$. Le Théorème 4.2 s'en déduit, puisque toute fonction complexe $f(x)$ continue et 2π-périodique sur $\mathbb{R}$ peut s'écrire $f(x) = g(e^{ix})$, pour $g(z)$ une fonction complexe continue sur $\mathbb{T}$.

5. Espace de Hilbert

5.1. L'utilisation des espaces de Hilbert pour la formulation abstraite de la mécanique quantique s'introduit dans la deuxième moitiés des années 20[22]. Plusieurs mathématiciens et physiciens sont impliqués dans ce travail de formalisation : mentionnons en particulier Wiener, Born, Jordan, Von Neumann et Stone. La présentation systématique/axiomatique de la mécanique quantique est exposée dans le livre de Von Neumann [VN55].

5.2. Dans la suite, nous supposons que les scalaires sont pris dans le corps $\mathbb{C}$ des nombres complexes ; *« le théorème de Fréchet-von Neumann-Jordan »* (c.f. [JvN35]) nous permet de donner un définition très générale des espaces de Hilbert comme cas particuliers des espaces de Banach. Ainsi, nous dirons que le $\mathbb{C}$-espace de Banach $(\mathscr{H}, \|\cdot\|)$ est un *« espace de Hilbert »* (complexe), lorsque *« l'identité du parallélogramme »*

$$\|X\|^2 + \|Y\|^2 = \frac{1}{2}\Big(\|X+Y\|^2 + \|X-Y\|^2\Big) \tag{8}$$

22. On lit dans [Loc94] : *« Ce sont les séries de Fourier qui sont à l'origine de l'espace de Hilbert. En effet, en étudiant les vibrations de membranes et d'enceintes quelconques, les mathématiciens mirent en évidences des vibrations propres de formes variées. Or, au début de notre siècle, surgit d'un autre horizon un nouveau venu en mathématique – l'équation intégrale de Fredholm – qui introduisait sans le dire la notion d'opérateur et exprimait sous une forme nouvelle le problème des vibrations. Grâce à cette équation, Hilbert remarqua que la recherche des vibrations d'une membrane ou d'une cavité possède une signification géométrique remarquable : elle est analogue à la recherche des axes de symétrie d'une conique (ellipse, hyperbole ou parabole).[...]*

La théorie des espaces de Hilbert, déjà très élaborée dans les années 20, était prête à accueillir en son sein la mécanique quantique, de même que les espaces de Riemann semblaient attendre la relativité générale. C'est à John von Neumann que revint le mérite de construire l'algèbre des opérateurs dans cet espace et de créer, dans un ouvrage célèbre [VN55] *la forme mathématique de la mécanique quantique. »*

est satisfaite ; alors (8) assure que l'application $(X,Y) \mapsto \langle X|Y\rangle$ définie par *« l'identité de polarisation (complexe) »*, en posant

$$\langle X|Y\rangle = \frac{1}{4}\big(\|X+Y\|^2 - \|X-Y\|^2 + \boldsymbol{i}\|X+\boldsymbol{i}Y\|^2 - \boldsymbol{i}\|X-\boldsymbol{i}Y\|^2\big) \tag{9}$$

est *« un produit scalaire hermitien (semi-linéaire à droite) sur $\mathscr{H}$ »*, la norme $\|\cdot\|$ étant liée au produit scalaire par la relation $\|X\|^2 = \langle X|X\rangle$. Dans la suite nous travaillons dans un espace de Hilbert (complexe) supposé séparable (i.e. qui possède un sous-ensemble dénombrable dense).

5.3. Partant de l'inégalité triangulaire, nous avons $(\|X\| + \|Y\|)^2 \geq \|X+Y\|^2$ et donc,

$$\|X\|^2 + \|Y\|^2 + 2\|X\|\|Y\| \geq \|X\|^2 + \|Y\|^2 + \langle X|Y\rangle + \langle Y|X\rangle$$

soit encore, après simplification :

$$\|X\|\|Y\| \geq \mathfrak{re}\,\langle X|Y\rangle \tag{10}$$

Par suite, si $\langle X|Y\rangle \neq 0$ et si θ est un argument de $\langle X|Y\rangle$, alors

$$\|X\|\|Y\| = \|X\|\|e^{\boldsymbol{i}\theta}Y\| \geq \mathfrak{re}\big(e^{-\boldsymbol{i}\theta}\langle X|Y\rangle\big) = |\langle X|Y\rangle| \tag{11}$$

Cette inégalité est connue sous plusieurs noms ; en France elle est appelée *« l'inégalité de Cauchy-Schwarz »* (Lusternik et Sobolev l'appelle *« l'inégalité de Bouniakovski-Schwarz »* : c.f. [LS89]). Il peut sembler étrange de démontrer l'inégalité de Cauchy-Schwarz à partir de l'inégalité triangulaire : de fait, nous avons choisi ici de définir un espace de Hilbert comme un espace de Banach spécial dans lequel le produit scalaire se définit à partir de la norme. Dans la pratique (et de manière plus classique : voir par exemple les §§ 5.6 & 5.7) on a souvent affaire à un espace pré-hilbertien, i.e. un espace vectoriel normé dont la norme est déduite d'un produit scalaire $\langle\cdot|\cdot\rangle$: cet espace sera hilbertien si et seulement si il est complet. Dans cette dernière présentation, les propriétés du produit scalaire permettent de déduire successivement l'inégalité de Cauchy-Schwarz puis le fait que l'application $X \mapsto \|X\| := \langle X|X\rangle^{1/2}$ est bien une norme : en particulier, l'inégalité triangulaire s'obtient à partir de l'inégalité de Cauchy-Schwarz en écrivant :

$$\begin{aligned}\|X+Y\|^2 = \|X\|^2 + \|Y\|^2 + 2\mathfrak{re}\,\langle X|Y\rangle &\leq \|X\|^2 + \|Y\|^2 + 2|\langle X|Y\rangle| \\ &\leq \|X\|^2 + \|Y\|^2 + 2\|X\|\,\|Y\| = (\|X\| + \|Y\|)^2\end{aligned}$$

On obtient ainsi un espace pré-hilbertien, dont la complétude doit être vérifiée à postériori (voir par exemple les versions du théorème de Riesz-Fisher données par les Théorèmes 2.9 & 5.8). Dans un espace de Hilbert, l'inégalité triangulaire et l'inégalité de Cauchy Schwarz, sont donc deux formes équivalentes d'une même propriété. Pour conclure ce paragraphe, nous démontrons ci-dessous la forme complète de l'inégalité de Cauchy-Schwarz à partir des propriétés du produit scalaire (et où figure le cas d'égalité).

Théorème 5.1. *Soit $\langle\cdot|\cdot\rangle$ le produit scalaire d'un espace de Hilbert $\mathscr{H}$ et $X \mapsto \|X\| := \langle X|X\rangle^{1/2}$ la norme associée ; alors l'inégalité de Cauchy-Schwarz est satisfaite, i.e. pour tout $X, Y \in \mathscr{H}$:*

$$|\langle X|Y\rangle| \leq \|X\|\,\|Y\| \tag{12}$$

De plus l'égalité en (12) a lieu ssi X et Y sont proportionnels.

Preuve. Le cas $Y = 0$ étant trivial, nous pouvons supposer que $Y \neq 0$; de plus, quitte à multiplier Y par $e^{i\theta}$, pour un réel θ adéquat, nous pouvons supposer sans perte de généralité que $\langle X|Y\rangle$ est un nombre réel. Pour tout λ réel nous posons

$$P(\lambda) = \|X + \lambda Y\|^2 = \|X\|^2 + 2\lambda \langle X|Y\rangle + \lambda^2\|Y\|^2.$$

qui (dans le cas considéré) est un polynôme du second degré en λ (car $\|Y\|^2 > 0$) et dont les coefficients sont réels (car $\langle X|Y\rangle$ est réel). Mais par définition $P(\lambda) \geq 0$, pour tout $\lambda > 0$, ce qui signifie que son discriminant est négatif ou nul, soit encore que

$$4\langle X|Y\rangle^2 - 4\|X\|^2\|Y\|^2 \leq 0,$$

L'inégalité de Cauchy-Schwarz en découle directement. Pour le cas d'égalité, il est immédiat que si X et Y sont proportionnels, alors $|\langle X|Y\rangle| = \|X\|\,\|Y\|$. Réciproquement, si $|\langle X|Y\rangle| = \|X\|\,\|Y\|$ alors le discriminant ci-dessus est nul et donc $P(X)$ admet une racine réelle (double) λ_0, ce qui signifie que $\|X + \lambda_0 Y\| = 0$.

□

Le paragraphe § 10 est consacré au principe d'incertitude de Heisenberg-Gabor, dans le cadre de la théorie des signaux d'énergie finie. Ce résultat est une forme déguisée du principe d'incertitude pour le couple d'observable position-impulsion de la mécanique ondulatoire (c.f. [Oli16b, Théorèmes 5.1 & 5.5]). Cependant, il existe un principe d'incertitude plus général que nous énoncerons dans le cadre abstrait de la mécanique ondulatoire (c.f. [Oli16b, Théorème 6.4]) qui se trouve être une conséquence directe de l'inégalité de Cauchy-Schwarz telle qu'écrite en (10).

5.4. Deux vecteurs X et Y d'un espace de Hilbert $\mathscr{H}$ sont dits orthogonaux lorsque $\langle X|Y\rangle = 0$. L'orthogonal d'un sous-ensemble $K \subset \mathscr{H}$ (non vide) est l'ensemble noté $K^\perp$ formé des $Y \in \mathscr{H}$ qui sont orthogonaux à tous les vecteurs de K, c'est-à-dire t.q. $\langle X|Y\rangle = 0$, pour tout $X \in K$.

Théorème 5.2. *L'orthogonal $K^\perp$ de tout sous-ensemble $K \subset \mathscr{H}$ (non vide) est un sous-espace vectoriel de $\mathscr{H}$ qui est fermé.*

Preuve. Il est immédiat que $K^\perp$ est un sous-espace de $\mathscr{H}$. Pour la fermeture de $K^\perp$, considérons que $X_0, X_1, \ldots$ est une suite de $K^\perp$ qui converge vers une limite X_* (dans $\mathscr{H}$). Si Y est un point arbitraire de K alors, par la continuité de la forme linéaire $\langle Y|\cdot\rangle$, nous savons que $\langle Y|X_n\rangle \to \langle Y|X_*\rangle$ quand $n \to +\infty$; or $\langle Y|X_n\rangle = 0$, pour tout n et donc $\langle Y|X_*\rangle = 0$: la conclusion que $X_* \in K^\perp$ vient du fait que Y est arbitrairement dans K.

□

Le lemme suivant est très important.

Lemme 5.3. *Si K est un sous ensemble dense de $\mathscr{H}$, alors $K^\perp = \{0\}$.*

Preuve. Soit $X \in K^\perp$ et $X_0, X_1, \ldots$ une suite de points de K t.q. $X_n \to X$ quand $n \to +\infty$; alors $\langle X_n|X\rangle = 0$ pour tout n. La conclusion vient de la continuité de l'application $Y \mapsto \langle Y|X_n\rangle$, de sorte que $0 = \langle X_n|X\rangle \to \langle X|X\rangle = \|X\|^2$ quand $n \to +\infty$.

□

5.5. Rappelons qu'une famille $\{X_i\}_{i=M}^{N}$ $(-\infty \leq M \leq N \leq +\infty)$ est *« orthonormée »* ssi $\langle X_i|X_j\rangle = \delta_{ij}$ (symbole de Kchronecker) et que c'est une *« base hilbertienne de $\mathscr{H}$ »*, lorsque le sous-espace vectoriel engendré par les X_n est dense dans $\mathscr{H}$. Dans la suite nous prendrons $(M,N) = (0,+\infty)$ (nous rencontrerons d'autre cas – en particulier le cas $(M,N) = (-\infty,+\infty)$ – qui se traitent de manière analogue).

Proposition 5.4. (i) *: Etant donnée $\{X_n\}_{n=0}^{N}$ (avec $0 \leq N \leq +\infty$) une famille orthonormée de $\mathscr{H}$, nous avons pour tout $Y \in \mathscr{H}$:*

$$\textstyle\sum_{k=0}^{\infty} |\langle Y|X_k\rangle|^2 \leq \|X\|^2 \qquad \text{(« Inégalité de Bessel »)}$$

(ii) *: si $\{X_n\}_{n=0}^{N}$ est une base hilbertienne de $\mathscr{H}$ alors tout $Y \in \mathscr{H}$ s'écrit $Y = \sum_{k=1}^{N} \langle Y|X_k\rangle X_k$ i.e. :*

$$\lim_{n\to N} \left\| Y - \textstyle\sum_{k=0}^{n} \langle Y|X_k\rangle X_k \right\| = 0$$

de plus, $n \mapsto \langle Y|X_n\rangle$ est une suite de nombres complexes de carré sommables t.q. :

$$\textstyle\sum_{n=0}^{N} |\langle Y|X_n\rangle|^2 = \|X\|^2 \qquad \text{(« Identité de Pythagore-Parseval »)}$$

Preuve. (i) : Il suffit de montrer l'inégalité de Bessel lorsque $N < +\infty$; dans ce cas, nous associons à tout $Y \in \mathscr{H}$ donné le vecteur $Y' = \sum_{k=0}^{N} \langle Y|X_k\rangle X_k$ de sorte que $\|Y'\|^2 = \sum_{k=0}^{N} |\langle Y|X_k\rangle|^2$; alors, en posant $Z = Y - Y'$, il vient, pour tout $i = 1,\ldots,N$

$$\langle Z|X_i\rangle = \langle Y|X_i\rangle - \textstyle\sum_j \langle Y|X_j\rangle \langle X_j|X_i\rangle = \langle Y|X_i\rangle - \sum_j \langle Y|X_j\rangle \delta_{ji} = 0$$

Par suite Z étant orthogonal à tous les X_i, il est aussi orthogonal à Y' et donc

$$\|Y\|^2 = \|Y' + Z\|^2 = \|Y'\|^2 + \|Z\|^2 + 2\mathfrak{re}\left\langle Y'|Z\right\rangle = \|Y'\|^2 + \|Z\|^2 \geq \|Y'\|^2$$

(ii) : Nous considérons ici que $N = +\infty$ ($N < +\infty$ étant classique). Pour Y arbitrairement pris dans $\mathscr{H}$, nous savons d'après (i) que $n \mapsto \langle Y|X_n\rangle$ est une suite de nombres complexes de carré sommable : cela entraîne que la suite $n \mapsto \sum_{k=0}^{n} \langle Y|X_k\rangle X_k$ est de Cauchy dans $\mathscr{H}$ (puisque $\|X_{p+q} - X_p\|^2 = \sum_{k=1}^{q} |\langle Y|X_{p+k}\rangle|^2$) : par suite nous pouvons considérer $Y' := \sum_{k=0}^{\infty} \langle Y|X_k\rangle X_k$ (i.e. la limite dans $\mathscr{H}$ de $\sum_{k=0}^{n} \langle Y|X_k\rangle X_k$ quand $n \to +\infty$). En posant $Z := Y - Y'$, nous pouvons montrer par un calcul analogue à celui effectué en (i) que Z est orthogonal à tous les X_n ; par suite Z est orthogonal à l'espace vectoriel engendré par les X_n : or cet espace étant dense dans $\mathscr{H}$ (définition d'une base hilbertienne), nous pouvons en conclure que $Z = 0$, i.e. $Y = Y'$. L'identité de Pythagore-Parseval découle alors de la continuité de la norme et du fait que pour tout $k \geq 0$

$$\left\|\textstyle\sum_{k=0}^{n} \langle Y|X_k\rangle X_k\right\|^2 = \sum_{k=0}^{n} |\langle Y|X_k\rangle|^2$$

□

Théorème 5.5. *Un espace de Hilbert est séparable (de dimension infinie) si et seulement si il possède une base hilbertienne dénombrable.*

Preuve. D'après la partie (ii) de la Proposition 5.4, il est immédiat que si l'espace de Hilbert $\mathscr{H}$ possède une base hilbertienne dénombrable, alors il est nécessairement séparable. Réciproquement si $\mathscr{H}$ est séparable alors il existe une famille dénombrable $\{X_0, X_1, \ldots\} \subset \mathscr{H}$ qui est libre et qui engendre un sous-espace vectoriel dense dans $\mathscr{H}$: en d'autre termes, si $\mathscr{H}_n$ est le sous-espace engendré par $\{X_0,\ldots,X_n\}$, alors $\bigcup_{n=0}^{\infty} \mathscr{H}_n$ est dense dans $\mathscr{H}$. En appliquant l'algorithme d'ortho-normalisation de Gramm-Schmidt, on peut construire une famille dénombrable $\{Y_0, Y_1, \dots\}$ orthonormée de sorte que pour tout $n \geq 0$, le sous-espace $\mathscr{H}_n$ coïncide avec le sous-espace engendré par $\{Y_0, \dots, Y_n\}$: cela entraîne que $\{Y_0, Y_1, \dots\}$ est une base hilbertienne de $\mathscr{H}$.

□

Théorème 5.6. *Soit $\mathscr{H}$ un espace de Hilbert séparable ; si $\mathscr{K}$ est un sous-espace fermé de $\mathscr{H}$ alors $\mathscr{K}^\perp$ est un sous-espace fermé, supplémentaire de $\mathscr{K}$ en ce sens que $\mathscr{H} = \mathscr{K} \oplus \mathscr{K}^\perp$.*

Preuve. Comme $\mathscr{K} \cap \mathscr{K}^\perp = \{0\}$, il suffit de démontrer que tout $X \in \mathscr{H}$ peut s'écrire $X = Y + Z$, avec $Y \in \mathscr{K}$ et $Z \in \mathscr{K}^\perp$. Si $\mathscr{K} = \{0\}$ le résultat est immédiat ; nous supposons donc maintenant que $\mathscr{K} \neq \{0\}$. Comme $\mathscr{K}$ est un sous-espace fermé de $\mathscr{H}$ (espace de Hilbert séparable), c'est lui-même un espace de Hilbert séparable : il possède donc une base hilbertienne (finie ou dénombrable : c.f. Théorème 5.5) soit $(X_i)_{i=0}^N$ (avec $1 \leq N \leq +\infty$). D'après *« l'inégalité de Bessel »* la série $\sum_{i=0}^N |\langle X|X_i\rangle|^2$ est convergente, ce qui entraîne que $Y := \sum_{i=0}^N \langle X|X_i\rangle X_i$ est un élément bien défini de $\mathscr{K}$: or pour tout $1 \leq i_0 \leq N$

$$\langle X - Y|X_{i_0}\rangle = \langle X|X_{i_0}\rangle - \sum_i \langle X|X_i\rangle \langle X_i|X_{i_0}\rangle = \langle X|X_{i_0}\rangle - \sum_i \langle X|X_i\rangle\, \delta^i_{i_0} = 0$$

ce qui entraîne que $Z = X - Y \in \mathscr{K}^\perp$.

□

Lorsque $\mathscr{K}$ est un sous-espace vectoriel fermé de $\mathscr{H}$, on dit aussi que $\mathscr{K}^\perp$ est *« le supplémentaire orthogonal de $\mathscr{K}$ »* ; nous pourrons aussi noter $\mathscr{H} = \mathscr{K} \boxplus \mathscr{G}$ une somme directe orthogonale, ce qui sous-entend que $\mathscr{G} = \mathscr{K}^\perp$.

Corollaire 5.7. *Si $\{0\} \neq \mathscr{K}$ est un sous-espace fermé de $\mathscr{H}$, alors la projection orthogonale sur $\mathscr{K}$ (i.e. la projection sur $\mathscr{K}$ parallèlement à $\mathscr{K}^\perp$) est un endomorphisme de $\mathcal{L}(\mathscr{H})$ de norme 1.*

Preuve. Si $(X_i)_{i=1}^N$ (avec $1 \leq N \leq +\infty$) est une base hilbertienne (finie ou dénombrable) $\mathscr{K}$, alors $PX = \sum_{i=1}^N \langle X|X_i\rangle X_i$ est la projection orthogonale sur $\mathscr{K}$: par suite $\|P\| = 1$, car d'une part (inégalité de Bessel) $\|PX\| \leq \|X\|$ et d'autre part $PX = X$ dès que $X \in \mathscr{K}$.

□

5.6. Les suites de scalaires pris dans le corps $\mathbb{K} = \mathbb{R}$ où $\mathbb{C}$ qui sont de carré sommable, forment un $\mathbb{K}$-espace vectoriel possédant une structure naturelle d'espace de Hilbert : cet espace représente une forme canonique des espaces de Hilbert séparables (c.f. Remarque 5.9). Pour préciser cela, considérons que $a : \mathbb{N} \to \mathbb{K}$ et $b : \mathbb{N} \to \mathbb{K}$ sont deux suites de scalaires ; alors pour tout entier $n \geq 0$, l'inégalité de Minkowski, i.e. l'inégalité triangulaire pour la norme euclidienne (resp. hermitienne) sur $\mathbb{R}^{n+1}$ (resp. $\mathbb{C}^{n+1}$), affirme que

$$\left(\sum_{k=0}^n |a(k) + b(k)|^2\right)^{1/2} \leq \left(\sum_{k=0}^n |a(k)|^2\right)^{1/2} + \left(\sum_{k=0}^n |b(k)|^2\right)^{1/2} \tag{13}$$

Nous en déduisons que l'ensemble $\ell^2 = \ell^2(\mathbb{N}, \mathbb{K})$ formé des suites $a : \mathbb{N} \to \mathbb{K}$ de nombres complexes qui sont de carré sommable (i.e. $\sum_{k=0}^\infty |a(k)|^2 < +\infty$) est un $\mathbb{K}$-espace vectoriel. De même (extension de l'inégalité de Chauchy-Schwarz sur $\mathbb{K}^{n+1}$) l'application $\langle\cdot|\cdot\rangle$ de $\ell^2 \times \ell^2 \to \mathbb{K}$ et t.q. $\langle a|b\rangle := \sum_{k=0}^\infty a(k)\overline{b(k)}$ est bien définie : elle définit un produit scalaire sur ℓ^2 associée à la norme $\|a\|_2 := \langle a|a\rangle^{1/2}$. Nous venons de vérifier que $(\ell^2, \|\cdot\|_2)$ est *« un $\mathbb{K}$-espace préhilbertien »*. Le théorème suivant est le prototype du *« théorème de Riesz-Fisher »*.

Théorème 5.8. *$\ell^2 = \ell^2(\mathbb{N}, \mathbb{K})$ est un $\mathbb{K}$-espace de Hilbert.*

Preuve. Nous devons démontrer que l'espace préhilbertien $(\ell^2, \|\cdot\|_2)$ est complet. D'après le Lemme 2.1, il s'agit de montrer que si $a_0, a_1, \ldots$ est une suite de ℓ^2 telle que la série $\sum_{q=0}^{\infty} a_q$ soit absolument convergente dans ℓ^2 (i.e. $\sum_{q=0}^{\infty} \|a_q\|_2 < \infty$), alors les sommes partielles $\alpha_n := \sum_{q=0}^{n} a_q$ forment une suite qui converge dans ℓ^2. Nous devons donc d'abord vérifier que pour tout p la série scalaire $\sum_{q=0}^{\infty} a_q(p)$ est sommable, puis en notant $\alpha(p)$ sa somme, vérifier que $p \mapsto \alpha(p)$ appartient à ℓ^2. Commençons par définir

$$A(p) := \sum_{q=0}^{\infty} |a_q(p)| \quad (\in [0\,;+\infty])$$

D'après l'inégalité de Minkowski (inégalité triangulaire pour la norme $\|\cdot\|_2$) nous avons

$$\sum_{p=0}^{\infty} A^2(p) = \lim_{n\to+\infty} \sum_{p=0}^{\infty} \left(\sum_{q=0}^{n} |a_q(p)| \right)^2 = \lim_{n\to+\infty} \left\| \sum_{q=0}^{n} |a_q| \right\|_2^2 \leq \lim_{n\to+\infty} \left(\sum_{q=0}^{n} \|a_q\|_2 \right)^2$$

soit encore

$$\sum_{p=0}^{\infty} A^2(p) \leq \left(\sum_{q=0}^{\infty} \|a_q\|_2 \right)^2 < \infty \tag{14}$$

Cela signifie que $A^2(p)$ est sommable, ou encore que $A \in \ell^2$; en particulier $A(p)$ est nécessairement fini pour tout p, ce qui assure à la série $\sum_{q=0}^{\infty} a_q(p)$ d'être absolument sommable pour tout p. Ainsi $p \mapsto \alpha(p) := \sum_{q=0}^{\infty} a_q(p)$ est une suite de scalaires bien définie : or en utilisant (14) nous avons

$$\sum_{p=0}^{\infty} |\alpha(p)|^2 = \sum_{p=0}^{\infty} \left| \sum_{q=0}^{\infty} a_q(p) \right|^2 \leq \sum_{p=0}^{\infty} \left(\sum_{q=0}^{\infty} |a_q(p)| \right)^2 = \sum_{p=0}^{\infty} A^2(p) < +\infty$$

ce qui signifie que $\alpha \in \ell^2$. Comme combinaison linéaire finie d'élément de ℓ^2, la suite $\alpha_n := \sum_{q=0}^{n} a_q$ est aussi dans ℓ^2 et de plus

$$\sum_{p=0}^{\infty} \left| \alpha(p) - \alpha_n(p) \right|^2 \leq \sum_{p=0}^{\infty} \left(\sum_{q=n+1}^{\infty} |a_q(p)| \right)^2$$

Or $p \mapsto \left(\sum_{q=n+1}^{\infty} |a_q(p)| \right)^2$ est dominée par $p \mapsto A^2(p)$ qui d'après (14) est sommable ; d'autre part, pour tout p, nous savons que $0 \leq A(p) < +\infty$ de sorte que $\sum_{q=n+1}^{\infty} |a_q(p)| \to 0$ quand $n \to +\infty$: par suite (convergence dominée) les α_n tendent vers α dans ℓ^2.

□

Remarque 5.9. *En un sens, l'espace ℓ^2 (sur $\mathbb{K} = \mathbb{R}$ où $\mathbb{C}$) est une version canonique des $\mathbb{K}$-espaces de Hilbert séparables. En effet, notons que si ϵ_n est l'élément de ℓ^2 t.q. $\epsilon_n(n) = 1$ et $\epsilon_n(k) = 0$ si $n \neq k$, alors $\mathcal{E} := \{\epsilon_n\ ;\ n = 0, 1, \ldots\}$ est une base hilbertienne de ℓ^2 ; mais $\mathcal{E}$ étant dénombrable, nous en déduisons que ℓ^2 est séparable. Si maintenant $(\mathscr{H}, \|\cdot\|)$ est un $\mathbb{K}$-espace de Hilbert séparable alors (algorithme de Gramm-Scmidt) $\mathscr{H}$ possède une base hilbertienne dénombrable, soit $\mathcal{F} := \{\varepsilon_n\ ;\ n = 0, 1, \ldots\}$. Le $\mathbb{K}$-espace vectoriel ℓ_0^2 engendré par $\mathcal{E}$ est un sous-espace dense de ℓ^2 ; il est immédiat que l'application linéaire $\Phi : \ell_0^2 \to \mathscr{H}$ t.q. $\Phi(\epsilon_n) = \varepsilon_n$ est isométrique (i.e. $\|\Phi(a)\| = \|a\|_2$, pour tout $a \in \ell_0^2$) : par suite Φ s'étend en une unique application linéaire de ℓ^2 sur $\mathscr{H}$ qui est une isométrie de ℓ^2 sur $\mathscr{H}$.*

5.7. Nous considérons maintenant le cas particulier du théorème de Riesz-Fisher sur les espaces $L^p(I)$ (c.f. Théorème 2.9 supra) avec $I = [-\pi\,;\pi]$ et $p = 2$ qui nous permet d'affirmer que $L^2[-\pi,\pi]$ est un espace de Banach. Il est alors immédiat de vérifier que $L^2[-\pi,\pi]$ est un espace de Hilbert pour le norme $\|f\|_2 := \langle f|f\rangle^{1/2}$ avec le produit scalaire

$$(f,g) \mapsto \langle f|g\rangle := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}dx$$

Pour tout entier $n \in \mathbb{Z}$ nous notons

$$\mathbf{e}_n(x) := \cos(nx) + \boldsymbol{i}\sin(nx) = e^{\boldsymbol{i}nx}$$

de sorte que $\mathcal{E} := \{\mathbf{e}_n\ ;\ n \in \mathbb{Z}\}$ est une famille orthonormée de $L^2[-\pi,\pi]$.

Théorème 5.10. *La famille orthonormée $\mathcal{E}$ est totale en ce sens que pour $f \in L^2[-\pi,\pi]$, si $\langle f\,|\,\mathbf{e}_n\rangle = 0$ pour tout $n \in \mathbb{Z}$, alors $f = 0$ presque partout : c'est une base hilbertienne de $L^2[-\pi,\pi]$.*

Une première démonstration du Théorème 5.10 est basée sur le fait [23] que $\mathscr{C}([-\pi\,;\pi],\mathbb{C})$ est dense dans $L^2[-\pi,\pi]$ (pour la topologie de la norme $\|\cdot\|_2$) combiné à la version trigonométrique du théorème d'approximation de Weierstrass affirmant que l'espace des polynômes trigonométriques est uniformément dense dans l'espace $\mathscr{C}_P([-\pi\,;\pi],\mathbb{C})$ des fonctions $f \in \mathscr{C}([-\pi\,;\pi],\mathbb{C})$ t.q. $f(-\pi) = f(\pi)$.

Lemme 5.11. *Soit S_nf la n-ième somme de Dirichlet de la fonction 2π-périodique f t.q. $f(x) = x$, pour tout $-\pi < x < \pi$ et $f(\pi) = 0$: alors $\|S_nf - f\|_2 \to 0$ quand $n \to +\infty$.*

Preuve. La fonction $f(x)$ étant impaire, nous avons $S_nf(x) = \sum_{k=1}^{n} b_k\sin(kx)$ où le coefficient en sinus d'ordre k (pour $k \geq 1$) est

$$b_k := \frac{1}{\pi}\int_{-\pi}^{\pi} x\sin(kx)dx = \frac{1}{\pi}\left[-\frac{x\cos(kx)}{k}\right]_{-\pi}^{\pi} + \frac{1}{k\pi}\int_{-\pi}^{\pi}\cos(kx)dx = 2\frac{(-1)^{k+1}}{k}$$

De plus d'après le théorème de Dirichlet (c.f. Théorème 4.3) nous avons pour tout $x \in \mathbb{R}$

$$f(x) = 2\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}\sin(kx) = 2\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}\left(\frac{e^{\boldsymbol{i}kx} - e^{-\boldsymbol{i}kX}}{2\boldsymbol{i}}\right) \tag{15}$$

La famille $\mathcal{E} := \{e^{\boldsymbol{i}kx}\ ;\ k \in \mathbb{Z}\}$ étant orthonormée, c'est une base hilbertienne de la clôture $\mathscr{H}_0$ du $\mathbb{C}$-espace vectoriel engendré par $\mathcal{E}$. Or d'après (15), nous savons que f et S_nf ($n \geq 1$) sont des éléments de $\mathscr{H}_0$ tels que $\|f - S_n(f)\|_2^2 = 2\sum_{k=n+1}^{\infty} 1/k^2$: cela entraîne [24] que f est la limite dans $\mathscr{H}_0$ (et donc dans $L^2[-\pi\,;\pi]$) de S_nf quand $n \to +\infty$.

□

Preuve du Théorème 5.10 n°1. Nous devons démontrer que l'espace des polynômes trigonométriques est dense dans $L^2[-\pi\,;\pi]$ (muni de la norme $\|\cdot\|_2$). Considérons pour

23. Résultat de la théorie de la mesure que nous admettons (voir par exemple [GW90, Théorème 15.3.3]).
24. Notons au passage que l'identité de *« Pythagore-Parseval »* dans $\mathscr{H}_0$ appliquée à

$$f(x) = \sum_{k=1}^{+\infty}\frac{(-1)^{k+1}}{\boldsymbol{i}k}(e^{\boldsymbol{i}kx} - e^{-\boldsymbol{i}kx})$$

nous donne

$$2\sum_{k=1}^{\infty}\frac{1}{k^2} = \|f\|_2^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} x^2dx = \frac{1}{2\pi}\left[\frac{x^3}{3}\right]_{-\pi}^{\pi} = \frac{\pi^2}{3}$$

d'où *« la formule d'Euler »* $\sum_{k=1}^{\infty} 1/k^2 = \pi^2/6$.

cela $f \in L^2[-\pi\,;\pi]$ et soit $\varepsilon > 0$ arbitrairement donné. Nous partons du fait qu'il existe $g(x) \in \mathscr{C}([-\pi\,;\pi],\mathbb{C})$ t.q. $\|f-g\|_2 \leq \varepsilon/3$. Si nous posons pour tout $-\pi < x < \pi$

$$h(x) = \frac{g(\pi)-g(-\pi)}{2\pi}x$$

alors d'une part (c.f. Lemme 5.11) il existe un polynôme trigonométrique $S(x)$ t.q. $\|h - S\|_2 \leq \varepsilon/3$. Mais d'autre part $g - h$ étant une fonction de $\mathscr{C}_P([-\pi\,;\pi],\mathbb{C})$, nous pouvons appliquer la version trigonométrique du théorème d'approximation de Weierstrass qui nous donne un polynôme trigonométrique $T(x)$ t.q. $\|(g-h)-T\|_\infty \leq 2\pi\varepsilon/3$. En particulier, nous avons $\|(g-h)-T\|_2 \leq \varepsilon/3$ et finalement nous avons

$$\|f-(S+T)\|_2 \leq \|f-g\|_2 + \|(g-h)-T\|_2 + \|h-S\|_2 \leq \varepsilon$$

□

Remarque 5.12. *Le fait que la famille $\mathcal{E}$ formée des exponentielles complexes $\mathbf{e}_k(x) = e^{ikx}$ pour $k \in \mathbb{Z}$ soit une base hilbertienne de $L^2[-\pi\,;\pi]$ entraîne que pour toute fonction $f \in L^2[-\pi\,;\pi]$ la série de Fourier $\sum_{k=-\infty}^{+\infty}\langle f|\mathbf{e}_k\rangle\,\mathbf{e}_k$ converge vers f dans $L^2[-\pi\,;\pi]$: en d'autres termes, si nous notons $c_k := \langle f|\mathbf{e}_k\rangle$, alors le polynôme trigonométrique $S_nf(x) := \sum_{k=-n}^{n} c_k e^{ikx}$ converge vers f dans $L^2[-\pi\,;\pi]$. Nous retrouvons ainsi le résultat du Lemme 5.11, puisque la fonction 2π-périodique $f(x)$ t.q. $f(x) = x$, pour tout $-\pi < x < \pi$ avec $f(\pi) = 0$ appartient à $L^2[-\pi\,;+\infty]$.*

5.8. Pour démontrer le Théorème 5.10, nous avons utilisé que $\mathscr{C}([-\pi\,;\pi],\mathbb{C})$ était un sous-espace dense de $L^2[-\pi\,;\pi]$: c'est un résultat classique de la théorie de la mesure que nous n'avons pas justifié. Nous proposons maintenant une autre approche du Théorème 5.10 en suivant l'argument donné par J. Bass dans [Bas71, § II-3-4, p. 146 - 149]. Il s'agit de vérifier d'abord que la famille $\mathcal{E}$ *sépare* les points du sous-espace $\mathscr{C}_P([-\pi\,;\pi],\mathbb{C})$ de $L^2[-\pi\,;\pi]$ formé des fonctions φ qui sont continues et prolongeables en une fonction continue et 2π-périodique sur $\mathbb{R}$ (i.e. $\varphi \in \mathscr{C}([-\pi\,;\pi],\mathbb{C})$ avec $\varphi(-\pi) = \varphi(\pi)$).

Proposition 5.13. *Si $\varphi \in \mathscr{C}_P([-\pi\,;\pi],\mathbb{C})$, alors $\varphi \equiv 0$ dès que $\langle\varphi\,|\,\mathbf{e}_n\rangle = 0$, pour tout $n \in \mathbb{Z}$.*

Preuve. D'après le théorème de Dirichlet (c.f. Théorème 4.3).

□

Pour $f \in L^2[-\pi\,;\pi]$ nous commençons par remarquer que (inégalité de Cauchy-Schwarz)

$$\int_{-\pi}^{\pi}|f(x)|dx \leq \sqrt{2\pi}\left(\int_{-\pi}^{\pi}|f(x)|^2dx\right)^{1/2} = 2\pi\|f\|_2$$

ce qui assure que $f \in L^1[-\pi\,;\pi]$; par suite, pour tout $x \in [-\pi,\pi]$

$$F(x) := \int_{-\pi}^{x} f(y)dy$$

est un nombre complexe bien définie. De plus, pour tout $0 \leq a \leq b \leq 1$, nous avons (deuxième application de l'inégalité de Cauchy-Schwarz) :

$$|F(b)-F(a)| = \left|\int_a^b f(y)dy\right| \leq \sqrt{(b-a)}\,\left(\int_a^b |f(y)|^2dy\right)^{1/2} \leq \sqrt{2\pi(b-a)}\,\|f\|_2$$

de sorte que $x \mapsto F(x)$ est une fonction de $\mathscr{C}([-\pi\,;\pi],\mathbb{C})$.

Proposition 5.14. *Soit f une fonction de $L^2[-\pi\,;\pi]$ soit $F : [-\pi\,;\pi] \to \mathbb{C}$ la fonction telle que*

$$F(x) = \int_{-\pi}^{x} f(y)dy$$

Si $F \equiv 0$ alors f est nécessairement nulle presque partout.

Pour démontrer la Proposition 5.14, nous avons besoin du fait que si f est une fonction intégrable positive ou nulle sur $[-\pi\,;\pi]$ et telle que $\int_{-\pi}^{\pi} f(x)dx = 0$ alors $f(x) = 0$ pour presque tout $x \in [-\pi\,;\pi]$. Nous utiliserons aussi que la mesure de Lebesgue λ (sur la tribu $\mathcal{B}$ des boréliens inclus dans $[-\pi\,;\pi]$) est *« régulière »*, en ce sens que pour tout $A \in \mathcal{B}$

(16)
$$\lambda(A) = \inf\left\{\lambda(U)\ ;\ A \subset U,\ U \text{ ouvert}\right\}$$

Preuve de la Proposition 5.14. Nous pouvons considérer séparément la partie réelles et la partie imaginaire de f : pour la démonstration, cela revient à supposer que f et F sont à valeurs réelles. Si $F(x) = 0$ pour tout $x \in [-\pi\,;\pi]$, alors nous avons

(17)
$$\int_A f(y)dy = 0$$

pour tout intervalle A ouvert de $[-\pi\,;\pi]$: les ouverts de $[-\pi\,;\pi]$ étant réunions dénombrables d'intervalles ouverts disjoints deux à deux, nous en déduisons (convergence monotone) que (17) est vraie pour tout ouvert de $[-\pi\,;\pi]$. Enfin, d'après (16), l'identité (17) demeure vraie pour tout borelien $A \in \mathcal{B}$: en prenant pour A les deux ensembles boréliens $\{f > 0\}$ et $\{f < 0\}$, nous pouvons conclure que f est nulle presque partout.

$\square$

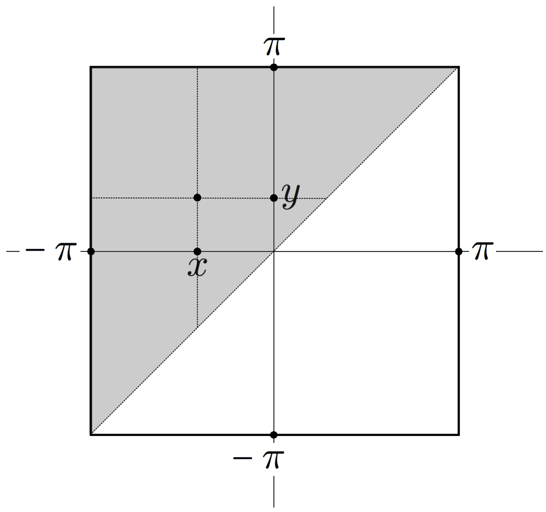

FIGURE 1. *Changement de sommation sur un domaine triangulaire du plan.*

Preuve du Théorème 5.10 n°2. Soit $f \in L^2[-\pi\,;\pi]$ telle que $\langle f\,|\,\mathrm{e}_n\rangle = 0$, pour tout $n \in \mathbb{Z}$ et soit $F : [-\pi\,;\pi] \to \mathbb{C}$ la fonction telle que $F(x) := \int_{-\pi}^{x} f(t)dt$. Nous savons d'une part que $F \in \mathscr{C}([-\pi\,;\pi], \mathbb{C})$; mais d'autre part $F(\pi) = 2\pi\,\langle f|\mathrm{e}_0\rangle = 0$ et donc $F(\pi) = F(-\pi) = 0$:

par suite F appartient à l'espace $\mathscr{C}_P([-\pi\,;\pi],\mathbb{C})$ des fonctions continues prolongeables en une fonction continue 2π-périodique sur $\mathbb{R}$. Nous affirmons que

$$(18) \qquad (\forall n \in \mathbb{Z} \setminus \{0\}) \qquad \int_{-\pi}^{\pi} F(x)e^{inx}dx = 0$$

Si nous admettons (18) et si nous notons $C := \int_{-\pi}^{\pi} F(x)dx$, alors nous avons

$$(\forall n \in \mathbb{Z}) \qquad \int_{-\pi}^{\pi} (F(x) - C)e^{inx}dx = 0$$

Comme F est continue, cela assure que $F(x) = C$ pour tout $x \in [-\pi\,;\pi]$; or par définition $F(-\pi) = 0$ et donc $F(x) = C = 0$ pour tout $x \in [-\pi\,;\pi]$. Or d'après les Propositions 5.13 et 5.14, nous pouvons conclure successivement que F est nulle et que f s'annule presque partout. Il nous reste à établir (18) : en utilisant le fait que $1_{[-\pi\,;x]}(y) = \mathbf{1}_{[y\,;\pi]}(x)$ (voir Fig. 1), nous pouvons écrire que pour tout $n \in \mathbb{Z} \setminus \{0\}$

$$\begin{aligned} \int_{-\pi}^{\pi} F(x)e^{inx}dx &= \int_{-\pi}^{\pi} \left(\int_{-\pi}^{x} f(y)dy \right) e^{inx}dx \\ &= \int_{-\pi}^{\pi} f(y) \left(\int_{y}^{\pi} e^{inx}dx \right) dy \\ &= \int_{-\pi}^{\pi} f(y) \left[\frac{1}{in} e^{inx} \right]_y^{\pi} dy = \frac{1}{in} \left((-1)^n \int_{-\pi}^{\pi} f(y)dy - \int_{-\pi}^{\pi} f(y)e^{iny}dy \right) \end{aligned}$$

soit encore

$$\int_{-\pi}^{\pi} F(x)e^{inx}dx = \frac{2\pi}{in}\Big((-1)^n \langle f|\mathbf{e}_0\rangle - \langle f|\mathbf{e}_{-n}\rangle \Big) = 0$$

$\square$

6. Transformée de Fourier L^1

6.1. La transformée de Fourier permet d'analyser les phénomènes de nature vibratoire. En pratique on se ramène à un signal $f(t)$, i.e. une fonction de la variable temporelle t à valeurs complexes. La transformée de Fourier $f \mapsto \hat{f}$ permettant alors d'obtenir une représentation $\hat{f}(\nu)$ du signal dans la variable de fréquence ν. La *« dualité »* portant sur le couple *« temps-fréquence »* n'est pas la seule possible ; nous verrons qu'en mécanique quantique *« le principe de correspondance de Schrödinger »* est basé sur la transformée de Fourier de sorte que *« les variables conjuguées »* (comme par exemple la position et l'impulsion) sont liées par la transformée de Fourier : ici nous utiliserons simplement le couple (x, y) pour désigner les variables conjuguées. Les transformées de Fourier $\mathfrak{F}_-[f]$ (directe i.e. en analyse) et $\mathfrak{F}_+[f]$ (inverse i.e. en synthèse) d'une fonction complexe $f \in L^1(\mathbb{R})$ (i.e. Lebesgue-intégrable sur $\mathbb{R}$) sont définies en posant, pour tout $p \in \mathbb{R}$ (resp. $q \in \mathbb{R}$) :

$$(19) \qquad \mathfrak{F}_-[f](y) = \frac{1}{\sqrt{2\pi}} \int f(x)e^{-ixy}dx \quad \text{et} \quad \mathfrak{F}_+[f](x) = \frac{1}{\sqrt{2\pi}} \int f(y)e^{ixy}dy$$

Il est immédiat de vérifier que la transformée de Fourier $\mathfrak{F}_\pm[\cdot]$ est $\mathbb{C}$-linéaire et que pour toute fonction $f \in L^1(\mathbb{R})$, la fonction $\mathfrak{F}_\pm[f]$ est borélienne avec

$$\|\mathfrak{F}_\pm[f]\|_\infty \le \|f\|_1$$

Ainsi $\mathfrak{F}_\pm[\cdot]$ est une application linéaire bornée de $L^1(\mathbb{R})$ dans $L^\infty(\mathbb{R})$. Cependant $\mathfrak{F}_\pm[\cdot]$ n'est pas un endomorphisme de $L^1(\mathbb{R})$. En effet, la transformée de Fourier de la fonction

« porte » $\Pi(x) := \mathbf{1}_{[-1/2\,;1/2]}(x)$ s'écrit

$$\mathfrak{F}[\Pi](y) = \frac{1}{\sqrt{2\pi}}\int_{-1/2}^{+1/2} e^{-\boldsymbol{i}xy}dx = \frac{1}{\sqrt{2\pi}}\left[-\frac{1}{y/2}\,\frac{e^{-\boldsymbol{i}xy}}{2\boldsymbol{i}}\right]_{x=-1/2}^{x=+1/2} = \frac{1}{\sqrt{2\pi}}\,\frac{\sin(y/2)}{y/2} = \frac{1}{\sqrt{2\pi}}\mathrm{sinc}\left(\frac{y}{2\pi}\right)$$

où nous avons introduit *« la fonction sinus cardinal »* [25]

$$\mathrm{sinc}(x) = \frac{\sin(\pi x)}{\pi x} = \prod_{n=1}^{+\infty}\left(1-\frac{x^2}{n^2}\right)$$

Il est facile de vérifier que le sinus cardinal est une fonction qui n'est pas Lebesgue intégrable [26]. Le *« lemme de Riemann-Lebesgue »* donne une information importante sur les images des fonctions Lebesgue intégrables par la transformée de Fourier [27].

Théorème 6.1 (Riemann-Lebesgue)**.** *Si $f(x)$ est une fonction de $L^1(\mathbb{R})$ alors sa transformée de Fourier $\hat{f}(y) = \mathfrak{F}[f](y)$ est continue sur $\mathbb{R}$ et tend vers 0 quand $p \to \pm\infty$.*

Preuve. Voir par exemple [GW90, Théorèmes 5.1.1 & 17.1.3] ou [ST90, § 2.3, p. 96].

□

La proposition suivante résume les propriétés basiques de la transformée de Fourier.

Proposition 6.2. *Si f est une fonction dans $L^1(\mathbb{R})$, alors pour tout $\alpha \in \mathbb{R} \setminus \{0\}$:*

$$\begin{array}{rcll} (i) & : & \mathfrak{F}[f(\alpha x)](y) = 1/|\alpha|\,\mathfrak{F}[f]\,(y/\alpha) & (\textit{changement d'échelle}) \\ (ii) & : & \mathfrak{F}[f(x-\alpha)](y) = e^{-\boldsymbol{i}\alpha y}\,\mathfrak{F}[f](y) & (\textit{effet de translation}) \\ (iii) & : & \mathfrak{F}[e^{\boldsymbol{i}\alpha x}f(x)](y) = \mathfrak{F}[f](y-\alpha) & (\textit{effet de modulation}) \end{array}$$

6.2. Les relations entre la transformée de Fourier et la dérivation des fonctions peuvent être exprimées à l'aide de *« l'opérateur position »* $f \mapsto Qf$ t.q. $Qf(x) := xf(x)$. Notons au passage que Q ne peut définir un endomorphisme de $L^1(\mathbb{R})$ (ni de $L^2(\mathbb{R})$: cette remarque deviendra déterminante pour la mécanique ondulatoire de de Broglie-Schrödinger : c.f. [Oli16a]). Par application du théorème de dérivation sous le signe intégral nous pouvons déduire la proposition suivante.

Proposition 6.3. *Soit $f \in L^1(\mathbb{R})$ et $k \geq 1$; (i) : si $x^k f(x) \in L^1(\mathbb{R})$ alors $\mathfrak{F}[f](y) \in \mathscr{C}^k(\mathbb{R})$ et*

$$\partial^k\mathfrak{F}[f](y) = \mathfrak{F}\left[(-\boldsymbol{i}Q)^k f\right](y)$$

(ii) : si $f \in \mathscr{C}^k(\mathbb{R})$ et si $\partial^k f \in L^1(\mathbb{R})$ alors $Q^k\mathfrak{F}[f] \in L^1(\mathbb{R})$ et

$$\mathfrak{F}\left[\partial^k f\right] = (\boldsymbol{i}Q)^k\mathfrak{F}[f]$$

---

25. L'expression de la fonction sinus cardinal en produit eulérien a été introduite et utilisée par Euler dans sa preuve heuristique géniale de la formule $\zeta(2) = \pi^2/6$; rappelons aussi que la fonction sinus cardinal est reliée à la fonction gamma par la formule des compléments d'Euler avec (c.f. [ST89, p. 155-156]) :

$$\Gamma(x)\Gamma(x-1) = \frac{\pi}{\sin(\pi x)}$$

26. Il est aussi facile de vérifier que $\mathrm{sinc}(x)$ est de carré intégrable :. Il est alors possible de démontrer que la famille $(\mathrm{sinc}(x-n)\ ;\ n \in \mathbb{Z})$ est une base hilbertienne de $L^2(\mathbb{R})$: ce résultat important est lié au *« théorème de l'échantillonnage de Shannon »* (c.f. [GW90, Leçon n° 38]).

27. Bien que très fondamental, nous ne développons pas ici la question du lemme de Riemann-Lebesgue. Nous y reviendrons plus en détail à propos de l'intégrale de chemin de Feynamnn (c.f. [Oli17]) et de l'interprétation du phénomène *« des interférences destructrices »*.

Preuve. Voir par exemple [GW90, Théorème 17.2.1, p. 130] ou [ST90, § 2.4, p. 97].

□

6.3. Les fonctions gaussiennes jouent un rôle important pour la transformée de Fourier (nous les retrouverons en particulier dans le cas d'égalité de l'inégalité de Heisenberg-Gabor : c.f. § 10). Le résultat suivant sur les *« fonctions gaussiennes à écart type complexe »* est classique (dans le cas réel).

Proposition 6.4. *Soit $\xi = a + \boldsymbol{i}b$ avec $a, b \in \mathbb{R}$ et $a > 0$; alors*

$$(i) \ : \ \int e^{-\xi x^2} dx = \sqrt{\frac{\pi}{\xi}} \quad et \quad (ii) \ : \ \int e^{-\xi x^2} e^{\boldsymbol{i}yx} dx = \sqrt{\frac{\pi}{\xi}}\, e^{-\frac{1}{4}\frac{y^2}{\xi}}$$

(ici $\sqrt{\xi}$ est le nombre complexe η t.q. $\mathfrak{re}(\eta) > 0$ et $\eta^2 = \xi$).

Preuve. (i) : Pour calculer $\gamma := \int e^{-\xi x^2} dx$, nous remarquons que

$$\begin{aligned}\gamma^2 &= \iint e^{-\xi(x^2+y^2)} dx dy \\ &= \frac{1}{2\xi} \int_{-\pi}^{\pi} \int_0^{\infty} 2\xi \rho e^{-\xi\rho^2} d\rho d\theta = \frac{1}{2\xi} \int_{-\pi}^{\pi} \left[-e^{-\xi\rho^2} \right]_{\rho=0}^{\infty} d\theta = \frac{\pi}{\xi} \lim_{\rho \to +\infty} \left[1 - e^{-a\rho^2} e^{-\boldsymbol{i}b\rho^2} \right]\end{aligned}$$

Or nous avons supposé $a > 0$ et par suite $\gamma^2 = \pi/\xi$.

(ii) : Pour tout $y \in \mathbb{R}$ nous notons

$$\varphi(y) := \int e^{-\xi x^2} e^{\boldsymbol{i}yx} dx$$

alors (dérivation sous le signe intégral puis, intégration par parties) il vient :

$$\begin{aligned}\frac{d\varphi}{dy} &= -\frac{1}{2\xi} \int \left(-2\xi x e^{-\xi x^2} \right) \left(\boldsymbol{i} e^{\boldsymbol{i}xy} \right) dx \\ &= -\frac{1}{2\xi} \left[e^{-\xi x^2} \left(\boldsymbol{i} e^{\boldsymbol{i}xy} \right) \right]_{-\infty}^{+\infty} + \frac{1}{2\xi} \int e^{-\xi x^2} \left(-y e^{\boldsymbol{i}xy} \right) dx\end{aligned}$$

L'hypothèse $a > 0$ assurant que le terme intégré est nul, nous obtenons

$$\frac{d\varphi}{dy} + \frac{y\varphi(y)}{2\xi} = 0 \tag{20}$$

ce qui équivaut à la quadrature, $d/dy\big(e^{\frac{1}{4}y^2/\xi}\varphi(y)\big) = 0$. Par suite $\varphi(y) = \varphi(0)e^{-\frac{1}{4}y^2/\xi}$ et la conclusion vient de (i) du fait que $\varphi(0) = \sqrt{\pi/\xi}$.

□

Le cas particulier de la partie (ii) de la Proposition 6.4 énoncé dans le corollaire suivante est particulièrement important en pratique.

Corollaire 6.5. *Soit $G_\xi(x) := e^{-\xi x^2/2}$, où $\xi = a + \boldsymbol{i}b$ avec $a, b \in \mathbb{R}$ et $a > 0$; alors*

$$\mathfrak{F}[G_\xi](y) = \frac{1}{\sqrt{2\pi}} \int e^{-\xi x^2/2} e^{\boldsymbol{i}yx} dx = \frac{1}{\sqrt{\xi}}\, e^{-\frac{1}{2}\frac{y^2}{\xi}}$$

(ici $\sqrt{\xi}$ est le nombre complexe η t.q. $\mathfrak{re}(\eta) > 0$ et $\eta^2 = \xi$). En particulier

$$\mathfrak{F}[G_1] = G_1$$

6.4. Dans ce paragraphe nous allons démontrer *« le théorème d'inversion de la transformée de Fourier L^1 »* (c.f. Théorème 6.6) affirmant que si f et $\mathfrak{F}_-[f]$ sont toutes deux des fonctions de $L^1(\mathbb{R})$ alors $\mathfrak{F}_+[\mathfrak{F}_-[f]] = f$ (dans $L^1(\mathbb{R})$). Noter au passage (lemme de Riemann-Lebesgue) que dans cette situation les deux fonctions f et $\mathfrak{F}_-[f]$ sont dans $\mathscr{C}_0(\mathbb{R})$ (espace des fonctions continues sur $\mathbb{R}$ et nulle en $\pm\infty$).

Théorème 6.6 (Formule d'inversion). *Si f et $\mathfrak{F}_-[f]$ sont dans $L^1(\mathbb{R})$, alors $\mathfrak{F}_+[\mathfrak{F}_-[f]](x) = f(x)$ pour presque tout $x \in \mathbb{R}$. (De même, si f et $\mathfrak{F}_+[f]$ sont dans $L^1(\mathbb{R})$, alors $\mathfrak{F}_-[\mathfrak{F}_+[f]] = f$ p.p..)*

Preuve. Soit $f \in L^1(\mathbb{R})$ t.q. $\mathfrak{F}_-[f] \in L^1(\mathbb{R})$ et notons $U(x) := 1/\sqrt{2\pi}e^{-x^2/2}$: le facteur $1/\sqrt{2\pi}$ est introduit pour assurer que $\int U(x)dx = 1$. La première étape de l'argument consiste à remarquer que $U(0) = 1/\sqrt{2\pi}$, de sorte que pour tout $x, y \in \mathbb{R}$:

$$\lim_{n\to+\infty} e^{ixy} U(y/n)\,\mathfrak{F}_-[f](y) = \frac{1}{\sqrt{2\pi}} e^{ixy}\mathfrak{F}_-[f](y)$$

Mais d'autre part $|e^{ixy}U(y/n)\mathfrak{F}_-[f](y)| \le 1/\sqrt{2\pi}\,|\mathfrak{F}_-[f](y)|$: or (hypothèse) $\mathfrak{F}_-[f] \in L^1(\mathbb{R})$, de sorte que par le *« théorème de la convergence dominée »*, nous avons pour tout $x \in \mathbb{R}$:

$$\lim_{n\to+\infty} \mathfrak{F}_+\Big[U(\cdot/n)\mathfrak{F}_-[f]\Big](x) = \frac{1}{\sqrt{2\pi}}\mathfrak{F}_+[\mathfrak{F}_-[f]](x) \tag{21}$$

La deuxième étape utilise le fait que les $U_n(x) := nU(nx)$ ($n \ge 1$) définissent une approximation de l'unité (dite gaussienne). Comme U est une fonction bornée et que $\mathfrak{F}_-[f]$ est (par hypothèse) dans $L^1(\mathbb{R})$, il est évident que $U(\cdot/n)\mathfrak{F}_-[f]$ est aussi dans $L^1(\mathbb{R})$, pour tout $n \ge 1$ entier : en utilisant *« le théorème de Fubini »* nous pouvons écrire

$$\begin{aligned}\mathfrak{F}_+\Big[U(\cdot/n)\mathfrak{F}_-[f]\Big](x) &= \frac{1}{\sqrt{2\pi}}\int e^{ixy}U(y/n)\left(\frac{1}{\sqrt{2\pi}}\int e^{-izy}f(z)dz\right)dy\\ &= \frac{1}{\sqrt{2\pi}}\int\left(\frac{1}{\sqrt{2\pi}}\int e^{-iy(z-x)}U(y/n)dy\right)f(z)dz\end{aligned}$$

soit encore

$$\mathfrak{F}_+\Big[U(\cdot/n)\mathfrak{F}_-[f]\Big](x) = \frac{1}{\sqrt{2\pi}}\int \mathfrak{F}_-[U(\cdot/n)](z-x)f(z)dz \tag{22}$$

Du fait que $\mathfrak{F}_-[U] = U$ (c.f. Corollaire 6.5) et par définition de U_n, nous avons

$$\mathfrak{F}_-[U(\cdot/n)](z-x) = n\mathfrak{F}_-[U](n(z-x)) = nU\big(n(z-x)\big) = U_n(z-x)$$

et comme $U_n(-t) = U_n(t)$, nous pouvons récrire (22) sous la forme

$$\mathfrak{F}_+\Big[U(\cdot/n)\mathfrak{F}_-[f]\Big](x) = \frac{1}{\sqrt{2\pi}}\int U_n(x-z)f(z)dz = \frac{1}{\sqrt{2\pi}}U_n * f(x) \tag{23}$$

Le Théorème 3.8 nous assurant que $U_n * f \to f$ dans $L^1(\mathbb{R})$, nous déduisons du corollaire du *« théorème de Riesz-Fisher »* (c.f. Proposition 2.10) qu'il existe une suite strictement croissante d'entiers $n_0, n_1, \ldots$ telle que pour presque tout $x \in \mathbb{R}$

$$\lim_{k\to+\infty} U_{n_k} * f(x) = f(x) \tag{24}$$

La conclusion vient grâce à l'identité (23) en comparant (24) et (21).

$\square$

Nous pouvons reformuler la formule d'inversion de la transformée de Fourier tel qu'énoncée dans le Théorème 6.6 en introduisant *« l'espace de Wiener »* $\mathscr{A}(\mathbb{R})$ c'est-à-dire le sous-espace de $L^1(\mathbb{R})$ constitué des fonctions dont la transformée de Fourier est encore dans $L^1(\mathbb{R})$ (i.e. $f \in \mathscr{A}(\mathbb{R})$ ssi f et $\hat{f} = \mathfrak{F}_-[f]$ sont dans $L^1(\mathbb{R})$). En effet, le Théorème 6.6 assure que $\mathscr{A}(\mathbb{R})$ est stable par transformée de Fourier (directe et inverse) et que la restriction $\mathfrak{F}_\pm[\cdot]$ à $\mathscr{A}(\mathbb{R})$ est un endomorphisme bijectif continu de norme 1 avec $\mathfrak{F}_-[\cdot]^{-1} = \mathfrak{F}_+[\cdot]$. Notons aussi que par le lemme de Riemann-Lebesgue, $\mathscr{A}(\mathbb{R})$ s'identifie à un sous-espace[28] de $\mathscr{C}_0(\mathbb{R})$ (espace des fonctions continues sur $\mathbb{R}$ et nulle en $\pm\infty$)[29] : en d'autres termes, cela signifie que chaque classe de fonctions de $\mathscr{A}(\mathbb{R})$ est identifiée à son unique représentant continu. Du fait que deux fonctions continues qui sont égales presque partout sont nécessairement égales *partout*, nous pouvons récrire le Théorème 6.6 comme suit.

Théorème 6.7 (Formule d'inversion bis)**.** *Si f et appartient à l'espace de Wiener $\mathscr{A}(\mathbb{R})$ alors $\mathfrak{F}_+[\mathfrak{F}_-[f]](x) = f(x)$ pour tout $x \in \mathbb{R}$: ainsi la transformée de Fourier $\mathfrak{F}_-[\cdot]$ réalise un isomorphisme de $\mathscr{A}(\mathbb{R})$ dont l'inverse est la transformée de Fourier inverse $\mathfrak{F}_+[\cdot]$.*

Remarque 6.8. *(1) : Remarquons que le Théorème 3.6 combiné à la Proposition 6.9 entraînent que si f et g sont dans l'espace de Wiener $\mathscr{A}(\mathbb{R})$ alors la convolée $f * g$ est aussi dans $\mathscr{A}(\mathbb{R})$ avec de plus $\|f * g\|_1 \le \|f\|_1 \cdot \|g\|_1$. Ainsi $(\mathscr{A}(\mathbb{R}), *)$ est une algèbre de convolution (ce n'est pas une algèbre de Banach car $(\mathscr{A}(\mathbb{R}), \|\cdot\|_1)$ n'est pas complet).*

(2) : L'espace de Wiener $\mathscr{A}(\mathbb{R})$ est un sous-espace strict de $\mathscr{C}_0(\mathbb{R})$; il est difficile de cerner les contours de $\mathscr{A}(\mathbb{R})$: notons par exemple (c.f. [GW90, Proposition 18.1.2, p. 137]*) que si f appartient à $\mathscr{C}^2(\mathbb{R}) \cap L^1(\mathbb{R})$ et que les dérivées f' et f'' sont toutes les deux dans $L^1(\mathbb{R})$ alors $\mathfrak{F}_-[f]$ est dans $L^1(\mathbb{R})$ et donc $f \in \mathscr{A}(\mathbb{R})$.*

(3) : La transformée de Fourier de la fonction porte $\Pi(x) = \mathbf{1}_{[-1/2\,;1/2]}(x)$ n'est pas une fonction de $L^1(\mathbb{R})$ (c.f. § 6.1) : cela pose la question de la formule d'inversion pour des fonctions de $L^1(\mathbb{R})$ continues par morceau. Mentionnons une version possible de la formule d'inversion qui est – en un sens – l'analogue du théorème de Dirichlet pour les séries de Fourier (c.f. Théorème 4.3) : elle est connue sous le nom de « formule d'inversion en valeurs principales ». On trouve un énoncé de ce résultat dans [GW90, Théorème 18.3.1, p. 139] *affirmant que si f est une fonction de $L^1(\mathbb{R})$ continûment dérivable sur chaque intervalle $]-\infty\,; a_1[,]a_2\,; a_3[, \dots,]a_n\,; +\infty[$ (avec $-\infty < a_1 < a_2 < \cdots < a_N < +\infty$) et que la fonction dérivée f' (définie partout sauf possiblement en un nombre fini de points) est elle-même dans $L^1(\mathbb{R})$, alors*

$$\lim_{a\to+\infty} \frac{1}{\sqrt{2\pi}} \int_{-a}^{+a} e^{ixy}\mathfrak{F}_-[f](y)dy = \frac{1}{2}\big(f(x^+) + f(x^-)\big)$$

Proposition 6.9. *Soient $f, g \in L^1(\mathbb{R})$, de transformée de Fourier $\hat{f} = \mathfrak{F}_-[f]$ et $\hat{g} = \mathfrak{F}_-[g]$; alors*

(i) : $f * g \in L^1(\mathbb{R})$ et $\mathfrak{F}_-[f * g] = \hat{f}\hat{g}$;

(ii) : si de plus $\hat{f}$ et $\hat{g}$ sont aussi dans $L^1(\mathbb{R})$, alors $fg \in L^1(\mathbb{R})$ et $\mathfrak{F}_-[fg] = \hat{f} * \hat{g}$.

28. C'est un sous-espace strict de $\mathscr{C}_0(\mathbb{R})$.
29. Notons aussi que $\mathscr{A}(\mathbb{R})$ est aussi sous-espace de $L^p(\mathbb{R})$ pour tout $1 \le p \le +\infty$.

Preuve. (i) : D'une part, le fait que $f * g$ soit dans $L^1(\mathbb{R})$ est inclus dans l'inégalité de Young (c.f. Théorème 3.6) ; d'autre part, d'après le théorème de Fubini, nous avons

$$\begin{aligned}
\mathfrak{F}_-[f*g](y) &= \int\left(\int f(x-z)g(z)dz\right)e^{-ixy}dx \\
&= \int\left(\int f(x-z)e^{-ixy}dx\right)g(z)dz \\
&= \int\left(\int f(u)e^{-i(u+z)y}du\right)g(z)dz = \left(\int f(u)e^{-iuy}du\right)\left(\int g(z)e^{-izy}dz\right)
\end{aligned}$$

(ii) : Comme $\hat{f}$ et $\hat{g}$ sont dans $L^1(\mathbb{R})$ nous pouvons appliquer la partie (i) avec $\mathfrak{F}_+[\cdot]$ (au lieu de $\mathfrak{F}_-[\cdot]$), de sorte que $\hat{f} * \hat{g}$ est une fonction de $L^1(\mathbb{R})$ t.q.

(25) $$\mathfrak{F}_+[\hat{f}*\hat{g}] = \mathfrak{F}_+[\hat{f}]\cdot\mathfrak{F}_+[\hat{g}] = f\,g$$

où la dernière égalité découle du théorème de l'inversion (c.f. Théorème 6.6). Mais le fait que $f = \mathfrak{F}_+[\hat{f}]$ et $\hat{f} = \mathfrak{F}_-[f]$ (resp. $g = \mathfrak{F}_+[\hat{g}]$ et $\hat{g} = \mathfrak{F}_-[g]$) soient toutes deux des fonctions de $L^1(\mathbb{R})$ assure qu'elles sont aussi des fonctions continues nulles à l'infini (c.f. Lemme de Riemann-Lebesgue). Comme f et g sont des fonctions dans $L^1(\mathbb{R})$, le fait qu'elles soient aussi continues sur $\mathbb{R}$ et nulles en $\pm\infty$ entraîne que $f\,g \in L^1(\mathbb{R})$ (en effet il existe $\alpha > 0$ t.q. $f\,g$ soit bornée sur $[-\alpha\,;\alpha]$ et $|f\,g| \le |f(x)|$ pour $|x| \ge \alpha$). Nous savons donc que d'une part que $\hat{f} * \hat{g} \in L^1(\mathbb{R})$ et comme d'autre part $\mathfrak{F}_+[\hat{f}*\hat{g}] = f\,g$ est aussi une fonction de $L^1(\mathbb{R})$, en partant de l'identité (25) et en appliquant $\mathfrak{F}_-[\cdot]$, nous déduisons (nouvelle application du théorème de l'inversion) que $\hat{f} * \hat{g} = \mathfrak{F}_-[f\,g]$.

□

7. Transformée de Fourier L^2

7.1. Nous nous intéressons maintenant à la transformée de Fourier des *« signaux d'énergie finie »*, c'est-à-dire à la transformée de Fourier (L^2) comme endomorphisme de l'espace de Hilbert (complexe) $L^2(\mathbb{R})$ des fonctions de carré sommable. Une première approche est basée sur l'utilisation de *« l'espace de Schwartz »* (sur $\mathbb{R}$) traditionnellement noté[30] $\mathscr{S}(\mathbb{R})$, c'est-à-dire l'espace constitué des fonctions $\varphi : \mathbb{R} \to \mathbb{R}$ qui sont C^∞ sur $\mathbb{R}$ pour lesquelles chaque dérivée $\partial^p f$ (pour $p \ge 0$) est négligeable en $\pm\infty$ devant toute fonction polynômiale ; cette dernière condition est équivalente à dire que pour tout entier $p, q \ge 0$, il existe une constante $K(p,q)$ t.q.

(26) $$|\partial^p f(x)| \le \frac{K(p,q)}{(1+x^2)^{q/2}}$$

Nous introduisons maintenant *« l'opérateur position »* noté Q (important dans la suite) et t.q. pour toute fonction $x \mapsto f(x)$ définie sur $\mathbb{R}$ et à valeurs dans $\mathbb{C}$,

(27) $$Qf(x) = xf(x)$$

Alors, pour tout entier $n \ge 0$, nous notons

$$\mathcal{N}_n(f) := \sum_{i+j=n} \|Q^i\partial^j f\|_\infty$$

30. Laurent Schwartz utilisait cette notation, appelant ces fonctions tests des fonctions *sphériques*.

La famille $\{\mathcal{N}_n\ ;\ n \geq 0\}$ de semi-normes séparantes ($\mathcal{N}_0(f)$ coïncide à la norme infinie de f) : il est alors facile de vérifier qu'en posant :

$$d(f,g) = \sum_{n=0}^{\infty} \frac{1}{2^n} \frac{\mathcal{N}_n(f-g)}{1+\mathcal{N}_n(f-g)} \tag{28}$$

nous définissons une métrique $d(\cdot,\cdot)$ sur $\mathscr{S}(\mathbb{R})$ (pour l'inégalité triangulaire il suffit de constater que la fonction $f(x) = x/(1+x)$ vérifie $f(a+b) \leq f(a) + f(b)$).

Proposition 7.1. *L'espace $\mathscr{S}(\mathbb{R})$ est complet*[31] *pour la métrique $d(\cdot,\cdot)$ définie en (28).*

Preuve. Il s'agit de montrer que si $\varphi_1, \varphi_2, \ldots$ est une suite de Cauchy de $\mathscr{S}(\mathbb{R})$, alors elle converge dans $\mathscr{S}(\mathbb{R})$ vers une fonction φ. Par définition de la métrique $d(\cdot,\cdot)$ nous savons que (pour tout $p \geq 0$) les fonctions $\partial^p\varphi_1, \partial^p\varphi_2, \ldots$ forment une suite de Cauchy de l'espace de Banach $(\mathscr{C}(\mathbb{R}), \|\cdot\|_\infty)$ – espace des fonctions continues pour la norme sup – et que donc il existe une fonction ϕ_p continue sur $\mathbb{R}$ t.q. $\|\partial^p\varphi_n - \phi_p\|_\infty \to 0$ quand $n \to +\infty$. En particulier, l'ensemble de ces convergences entraînent que pour tout $k \geq 1$ la fonction ϕ_p est continûment dérivable sur $\mathbb{R}$, avec $\partial\phi_p = \phi_{p+1}$: si nous notons $\varphi := \phi_0$, alors il vient par induction que φ est C^∞ sur $\mathbb{R}$ avec $\partial^p\varphi = \phi_p$. Pour conclure, considérons $p, q \geq 0$ (entiers) et $\varepsilon > 0$ (réel) arbitrairement donnés : comme $x^q\partial^p\varphi_n(x) \to x^q\partial^p\varphi(x)$ uniformément sur $\mathbb{R}$, il existe un rang N tel que $|x^q\partial^p\varphi(x)| \leq \varepsilon + |x^q\partial^p\varphi_n(x)|$, pour tout $x \in \mathbb{R}$ et tout $n \geq N$; finalement φ_n étant dans $\mathscr{S}(\mathbb{R})$, nous savons que $x^q\partial^p\varphi_n(x) \to 0$ quand $x \to \pm\infty$, ce qui assure que $|x^q\partial^p\varphi(x)| \leq \varepsilon$ quand $x \to \pm\infty$.

□

Proposition 7.2. *$\mathscr{S}(\mathbb{R})$ est stable par transformée de Fourier (L^1).*

Preuve. En utilisant l'opérateur position $f \mapsto Qf$ t.q. $Qf(x) = xf(x)$ (c.f. § 6.2) nous savons que si $\varphi \in \mathscr{S}(\mathbb{R})$, alors $Q^k\varphi \in L^1(\mathbb{R})$, pour tout $k \geq 0$: la partie (i) de la Proposition 6.3 assure alors que $\hat{\varphi} := \mathfrak{F}_-[\varphi] \in \mathscr{C}^k(\mathbb{R})$. Il nous reste à montrer que $Q^k\hat{\varphi}(x) \to 0$ quand $x \to \pm\infty$: mais comme $\partial^k\varphi \in \mathscr{S}(\mathbb{R}) \subset L^1(\mathbb{R})$, la partie (ii) de la Proposition 6.3 assure que $|Q^k\hat{\varphi}(x)| = |\mathfrak{F}_-[\partial^k\varphi](x)|$: la conclusion vient du lemme de Riemann-Lebesgue.

□

Il découle du théorème d'inversion (c.f. Théorème 6.6) que les restrictions bilatérales $\mathfrak{F}_\pm[\cdot] : \mathscr{S}(\mathbb{R}) \to \mathscr{S}(\mathbb{R})$ sont inverses l'une de l'autre, en ce sens que pour tout f dans $\mathscr{S}(\mathbb{R})$,

$$\mathfrak{F}_+[\mathfrak{F}_-[f]] = \mathfrak{F}_-[\mathfrak{F}_+[f]] = f \tag{29}$$

D'autre part, si g est une autre fonction de $\mathscr{S}(\mathbb{R})$, alors (en utilisant le théorème de Fubini)

$$\int \mathfrak{F}_-[f](y)\overline{g(y)}dy = \int \left(\frac{1}{\sqrt{2\pi}} \int f(x)e^{-ixy}dx\right)\overline{g(y)}dy = \int f(x)\left(\frac{1}{\sqrt{2\pi}} \int \overline{g(y)e^{-ixy}}dy\right)dx$$

soit encore

$$\int \mathfrak{F}_-[f](y)\overline{g(y)}dy = \int f(x)\overline{\mathfrak{F}_+[g](x)}dx \tag{30}$$

31. On peut montrer que c'est *« un espace de Fréchet »*

7.2. L'intégrale de $|f(x)|^2$ est appelée *« l'énergie du signal »* f : les signaux d'énergie finie forment donc l'espace $L^2(\mathbb{R})$ des fonctions de carré intégrable. D'après le *« théorème de Riesz-Fisher »* c'est un *« espace de Hilbert »* pour le produit scalaire $\langle\cdot|\cdot\rangle$ t.q.

$$\langle f|g\rangle = \int f(x)\overline{g(x)}dx$$

Nous notons $\|f\| := \langle f|f\rangle^{1/2}$ la norme hilbertienne sur $L^2(\mathbb{R})$. Une fonction de carré intégrable sur $\mathbb{R}$ n'étant pas nécessairement intégrable, la transformée de Fourier ne peut se définir directement sur $L^2(\mathbb{R})$. Nous allons utiliser le fait que l'espace de Schwarz $\mathscr{S}(\mathbb{R})$ est un sous-espace dense de $L^2(\mathbb{R})$ t.q. $\mathfrak{F}_\pm[\mathscr{S}(\mathbb{R})] \subset \mathscr{S}(\mathbb{R})$. Soient $f, g \in \mathscr{S}(\mathbb{R})$; alors d'après (30), nous pouvons écrire que :

$$\langle \mathfrak{F}_-[f]|g\rangle = \langle f|\mathfrak{F}_+[g]\rangle \tag{31}$$

En combinant (29) et (31) il vient aussi

$$\langle \mathfrak{F}_-[f]|\mathfrak{F}_-[g]\rangle = \langle f|\mathfrak{F}_+[\mathfrak{F}_-[g]]\rangle = \langle f|g\rangle \tag{32}$$

En particulier $\|\mathfrak{F}_-[f](x)\| = \|f\|$, pour tout $f \in \mathscr{S}(\mathbb{R})$: l'espace de Schwarz étant dense dans $L^2(\mathbb{R})$, l'endomorphisme $\mathfrak{F}_\pm[\cdot] : \mathscr{S}(\mathbb{R}) \to \mathscr{S}(\mathbb{R})$ se prolonge de manière unique (c.f. *« théorème de prolongement des opérateurs »*) en un endomorphisme unitaire (isométrie inversible [32]) de $L^2(\mathbb{R})$, que nous continuons à noter $\mathfrak{F}_\pm[\cdot]$. Ce résultat peut s'énoncer (et se démontrer) sous plusieurs formes : il est généralement connu comme de *« théorème de Plancherel »* (ou encore comme le *« théorème de Parseval-Plancherel »* [33]) pour la transformée de Fourier sur $L^2(\mathbb{R})$: on pourra consulter à ce sujet [Wie04, Ch. I, p. 46-71]. Dans la suite nous nous référerons à l'énoncé suivant :

Théorème 7.3. *La transformée de Fourier directe $\mathfrak{F}_-[\cdot]$ qui prolonge par continuité à $L^2(\mathbb{R})$ la transformée de Fourier directe sur $\mathscr{S}(\mathbb{R})$, est un endomorphisme unitaire de $L^2(\mathbb{R})$ dont l'inverse est le prolongement par continuité $\mathfrak{F}_+[\cdot]$ de la transformée de Fourier inverse sur $\mathscr{S}(\mathbb{R})$.*

7.3. Nous allons voir la définition directe de la transformée de Fourier L^2 telle qu'introduite par Wiener en 1933 (c.f. [Wie04]) : outre l'élégance de cette présentation, elle a d'autant plus d'intérêt (relativement aux questions de la mécanique quantique), qu'elle est directement liée à la résolution de l'équation de Schrödinger de l'oscillateur harmonique (c.f. [Oli16b]). En dérivant successivement la fonction e^{-x^2}, nous obtenons

$$\frac{d^n}{dx^n}e^{-x^2} = (-1)^n C_n(x)e^{-x^2}$$

où pour tout entier $n \geq 0$

$$C_n(x) = (-1)^n e^{x^2}\frac{d^n}{dx^n}e^{-x^2}$$

32. Il découle directement de (31) que *l'adjoint* $\mathfrak{F}_-[\cdot]$ coïncide avec $\mathfrak{F}_+[\cdot]$: or (29) entraîne que $\mathfrak{F}_-[\cdot]^{-1} = \mathfrak{F}_+[\cdot]$, ce qui signifie que $\mathfrak{F}_-[\cdot]$ est un endomorphisme unitaire de $L^2(\mathbb{R})$.

33. En fait le résultat de Plancherel (1910) est l'analogue pour la transformée de Fourier dans $L^2(\mathbb{R})$ du résultat de Parseval (1799) pour les séries de Fourier dont les coefficients sont de carré sommable.

le n-ème polynôme d'Hermite (physique [34]). Par des intégrations par parties successives, il est facile de vérifier que $(C_0, C_1, \dots)$ est une famille orthogonale de $L^2(\mu(dx))$ où $\mu(dx) = e^{-x^2}dx$, avec plus précisément :

$$\int C_n(x)C_m(x)e^{-x^2}dx = 2^n n!\sqrt{\pi}\delta^n{}_m$$

Par suite, en définissant les fonctions d'Hermite [35]

$$\psi_n(x) := \frac{1}{\sqrt{2^n n!\sqrt{\pi}}} C_n(x)e^{-x^2/2}$$

nous obtenons une famille $(\psi_0, \psi_1, \dots)$ qui est orthonormée dans $L^2(\mathbb{R}) = L^2(dx)$. Le point important ici est que $(\psi_0, \psi_1, \dots)$ constinue en fait une base hilbertienne de $L^2(\mathbb{R})$ (c.f. [Wie04, § 8, p. 64]). La définition de la transformée de Fourier L^2 donnée par Wiener se justifie alors par la proposition suivante :

Proposition 7.4. $\mathfrak{F}[\psi_n] = (\boldsymbol{i})^n\psi_n$*, pour tout entier* $n \geq 0$.

Preuve. Nous utilisons le fait que la fonction génératrice des polynômes d'Hermite est [36]

$$G(x,t) := \sum_{n=0}^{\infty} C_n(x)\frac{t^n}{n!} = e^{2tx-t^2}$$

Ainsi, pour calculer la transformée de Fourier des fonctions $\psi_n(x) := c_n e^{-x^2/2}C_n(x)$, il est assez naturel de considérer la fonction

$$e^{-x^2/2}G(x,t) = \sum_{n=0}^{\infty}\left(e^{-x^2/2}C_n(x)\right)\frac{t^n}{n!} = e^{-x^2/2+2tx-t^2} = e^{-1/2(x^2-4tx+4t^2)+t^2} = e^{-1/2(x-2t)^2}e^{t^2}$$

de sorte qu'alors [37] (en calculant de la transformée de Fourier relativement à la variable x et t considéré comme paramètre) il vient :

$$\mathfrak{F}\left[e^{-x^2/2}G(x,t)\right](y) = \mathfrak{F}\left[e^{-1/2(x-2t)^2}e^{t^2}\right](y) = e^{2\boldsymbol{i}yt}e^{-y^2/2}e^{t^2} = e^{-y^2/2-2\boldsymbol{i}yt+t^2}$$

34. Mis à part la notation, c'est cette définition qu'utilise Wiener : la lettre « H » pouvant désigner un espace de Hilbert ou encore le hamiltonien d'un système mécanique, nous avons utilisé la lettre « C » (au lieu de la notation traditionnelle avec la lettre « H ») en référence au prénom de l'illustre mathématicien. *« Les polynômes d'Hermite probabilistes »* notés $\mathcal{C}_n(x)$ sont définis pour tout entier $n \geq 0$ par l'identité

$$\mathcal{C}_n(x) := (-1)^n e^{x^2/2}\frac{d^n}{dx^n}(e^{-x^2/2})$$

ce qui donne $C_n(x) = \sqrt{2^n}\,\mathcal{C}_n(x/\sqrt{2})$. Si nous $\mathbf{1}$ est la fonction t.q. $\mathbf{1}(x) = 1$, pour tout $x \in \mathbb{R}$, alors

$$\begin{aligned} C_{n+1}(x) &= (-1)^{n+1}e^{x^2/2}\frac{d}{dx}\left(\frac{d^n}{dx^n}(e^{-x^2/2})\right) \\ &= (-1)^{n+1}e^{x^2/2}\frac{d}{dx}\left((-1)^n e^{-x^2/2}C_n(x)\right) = \left(x - \frac{d}{dx}\right)C_n(x) \end{aligned}$$

de sorte que par récurrence il vient pour tout $n \geq 0$:

$$C_n(x) = \left(x - \frac{d}{dx}\right)^n \mathbf{1}$$

35. Wiener les appelle les fonctions d'Hermite : elles sont aussi appelées les fonctions de Gauss-Hermite.

36. Par développement de Taylor de e^{-x^2} en x (pour l'accroissement $-t$), nous avons

$$e^{-(x-t)^2} = \sum_{n=0}^{\infty}(-1)^n e^{-x^2}C_n(x)\frac{(-t)^n}{n!} = \left(\sum_{n=0}^{\infty}C_n(x)\frac{t^n}{n!}\right)e^{-x^2}$$

37. Nous utilisons le fait que $\mathfrak{F}[e^{-x^2/2}](y) = e^{-y^2/2}$.

soit encore :

$$\mathfrak{F}_{-}\left[e^{-x^2/2}G(x,t)\right](y) = e^{-y^2/2}e^{2y(\boldsymbol{i}t)-(it)^2} = G(y,\boldsymbol{i}t)$$

Mais nous pouvons identifier ce résultat à l'expression de la transformée de Fourier de $e^{-x^2/2}G(x,t)$ obtenue à partir de la série, ce qui nous donne (et en justifiant les opérations)

$$\sum_{n=0}^{\infty}\left(e^{-y^2/2}C_n(y)\right)\frac{(\boldsymbol{i}t^n)}{n!} = \sum_{n=0}^{\infty}\left((\boldsymbol{i})^n e^{-y^2/2}C_n(y)\right)\frac{t^n}{n!} = \sum_{n=0}^{\infty}\left(\mathfrak{F}_{-}\left[e^{-x^2/2}C_n(x)\right](y)\right)\frac{t^n}{n!}$$

□

7.4. Pour terminer cette section, nous allons définir un produit sur les fonctions très analogue à la convolution. Si f et g sont deux fonctions de la variable réelle à valeurs complexes, telles que la fonction $y \mapsto f(x+y)\overline{g(y)}$ soit intégrable pour presque partout $x \in \mathbb{R}$ (par exemple si f et g sont dans $L^2(\mathbb{R})$), alors *« la corrélation croisée »* $f \circledast g$ est la fonction définie pour tout $x \in \mathbb{R}$ et telle que

$$f \circledast g(x) := \int f(x+y)\overline{g(y)}dy = \int f(y)\overline{g(y-x)}dy$$

Si $\varphi \in L^2(\mathbb{R})$, alors $|\hat{\varphi}(y)|^2 = \hat{\varphi}\,\overline{\hat{\varphi}}$ est une fonction de $L^1(\mathbb{R})$ qui s'appelle *« la densité spectrale d'énergie »* ; comme $\mathfrak{F}_{+}[\hat{\varphi}](x) = \varphi(x)$ et que $\mathfrak{F}_{+}\left[\overline{\hat{\varphi}}\right](x) = \overline{\varphi(-x)}$ nous pouvons écrire :

$$\mathfrak{F}_{+}\left[|\hat{\varphi}|^2\right](x) = \mathfrak{F}_{+}\left[\hat{\varphi}\,\overline{\hat{\varphi}}\right](x) = \mathfrak{F}_{+}[\hat{\varphi}] * \mathfrak{F}_{+}\left[\overline{\hat{\varphi}}\right](x) = \int \varphi(y)\overline{\varphi(y-x)}dy = \varphi \circledast \varphi(x) \tag{33}$$

La fonction $\mathsf{C}(x) := \varphi \circledast \varphi(x)$ est généralement appelée la *« fonction d'autocorrélation »* du signal : $\mathsf{C}(x)$ étant dans l'image de $L^1(\mathbb{R})$ par $\mathfrak{F}_{+}[\cdot]$, c'est une fonction continue qui s'annule en $\pm\infty$; remarquons aussi que

$$|\mathsf{C}(x)| = \frac{1}{\sqrt{2\pi}}\left|\int |\hat{\varphi}(y)|^2 e^{\boldsymbol{i}xy}dy\right| \le \frac{1}{\sqrt{2\pi}}\int |\hat{\varphi}(y)|^2 dy = \mathsf{C}(0)$$

La *« fonction d'autocohérence »* du signal φ est par définition $\mathsf{C}(x)/\mathsf{C}(0)$. Lorsque le signal est normalisé – i.e. appartient à la sphère unité de $L^2(\mathbb{R})$ – alors (théorème de Plancherel) $\mathsf{C}(0) = 1$ et donc les fonctions d'autocorrélation et d'autocohérence sont égales.

Remarque 7.5. *L'identité en (33) est une version du « théorème de Wiener-Khintchine ». Nous avons vu que l'autocorrellation $\varphi\circledast\varphi(y)$ est une fonction continue qui est maximale à l'origine. Par un argument analogue à l'argument classique donnant le cas d'égalité de l'inégalité de Cauchy-Schwarz, nous pouvons montrer que $x = 0$ est le seul point où le maximum de $\mathsf{C}(x)$ est atteint. Pour voir cela, considérons que ε est un réel arbitrairement donné. Alors*

$$\begin{aligned}
0 &\le \int \left(\varphi(\tau) + \varepsilon\varphi(\tau+x)\right)^2 d\tau > 0\\
&= \int \varphi^2(\tau)d\tau + 2\varepsilon\int \varphi(\tau)\varphi(\tau+x)d\tau + \varepsilon^2\int \varphi^2(\tau+x)d\tau\\
&= \int \varphi^2(\tau)d\tau + 2\varepsilon\int \varphi(\tau)\varphi(\tau+x)d\tau + \varepsilon^2\int \varphi^2(\tau)d\tau
\end{aligned}$$

Nous obtenons alors $a\varepsilon^2 + b\varepsilon + a > 0$, où nous avons posé $a = \|\varphi\|^2$ et $b = 2\varphi\circledast\varphi(\tau)$: en particulier, nous avons $b^2 - 4ac \le 0$, i.e. $b/2 \le a$, ce qui donne le résultat annoncé.

8. Mesures de Radon, probabilités et variables aléatoires

8.1. Le *« théorème de représentation de Riesz-Markov »*, permet une simplification de la notion de mesure. Notons $\mathscr{M}(\mathbb{R})$ l'ensemble des *« mesures de Radon sur $\mathbb{R}$ »*, c'est-à-dire le demi-cône des formes linéaires positives sur le $\mathbb{R}$-espace vectoriel $\mathscr{C}_c(\mathbb{R})$ des fonctions réelles continues sur $\mathbb{R}$ et de support compact : en d'autres termes, $\mu \in \mathscr{M}(\mathbb{R})$ ssi c'est une forme linéaire sur $\mathscr{C}_c(\mathbb{R})$ telle que

$$\forall f \in \mathscr{C}_c(\mathbb{R}),\ f \geq 0 \implies \mu(f) \geq 0$$

Le théorème de représentation de Riesz-Markov affirme que μ est une mesure de Radon sur $\mathbb{R}$ ssi il existe une mesure borélienne – notée abusivement μ – qui est *« positive »*, *« sigma-finie »*, *« régulière »*[38] et t.q.

$$\forall f \in \mathscr{C}_c(\mathbb{R}), \quad \mu(f) = \int f(x)\mu(dx)$$

(l'intégrale de gauche correspondant à la notation utilisée en théorie de la mesure).

8.2. Dans la suite nous identifions la mesure de Radon μ avec la mesure borélienne $\mu(dx)$ associée ; ainsi μ sera dite discrète s'il existe une suite injective de points $x_1, x_2, \dots$ dans $\mathbb{R}$, t.q. pour tout intervalle compact I la mesure $\mu(I)$ coïncide avec $\sum_{x_k \in I} \mu\{x_k\}$: dans la suite $\mathscr{M}_{\mathrm{d}}(\mathbb{R})$ désignera l'ensemble des mesures de Radon discrètes. Lorsque $\mu\{x\} = 0$, pour tout $x \in \mathbb{R}$, la mesure μ est dite *« continue (ou sans atomes) »* : nous notons $\mathscr{M}_{\mathrm{c}}(\mathbb{R})$ l'ensemble des mesures de Radon continues. La mesure de Lebesgue $\lambda = \lambda(dx)$ – caractérisée par le fait que $\lambda([a\,;b]) = b - a$ pour tout $-\infty < a \leq b < +\infty$ – est une mesure de Radon continue ; $\mu \in \mathscr{M}_{\mathrm{c}}(\mathbb{R})$ sera dite *« absolument continue »* si pour tout borélien A, le fait que $\lambda(A) = 0$ entraine que $\mu(A) = 0$. Le *« théorème de Radon-Nikodym »* affirme alors que μ est absolument continue – et nous notons $\mu \in \mathscr{M}_{\mathrm{ac}}(\mathbb{R})$ – ssi il existe une fonction $f_\mu(x)$ dans $L^1(\mathbb{R})$ – la densité de μ – telle que pour tout $-\infty < a \leq b < +\infty$

$$\mu([a\,;b]) = \int_a^b f_\mu(x)dx$$

Enfin, $\mu \in \mathscr{M}_{\mathrm{c}}(\mathbb{R})$ sera dite *« purement singulière »* – et nous notons $\mu \in \mathscr{M}_{\mathrm{ps}}(\mathbb{R})$ – si elle est portée par un borélien qui est de mesure de Lebesgue nulle. Le théorème suivant est une conséquence du *« théorème de décomposition de Lebesgue »*.

Théorème 8.1. *Pour toute mesure $\mu \in \mathscr{M}(\mathbb{R})$ il existe un unique triplet de mesures de Radon $(\mu_{\mathrm{d}}, \mu_{\mathrm{ac}}, \mu_{\mathrm{ps}}) \in \mathscr{M}_{\mathrm{d}}(\mathbb{R}) \times \mathscr{M}_{\mathrm{ac}}(\mathbb{R}) \times \mathscr{M}_{\mathrm{ps}}(\mathbb{R})$ tel que $\mu = \mu_{\mathrm{d}} + \mu_{\mathrm{ac}} + \mu_{\mathrm{ps}}$.*

« La mesure de Hausdorff de l'ensemble triadique de Cantor » est un exemple de mesure purement singulière ; d'autres exemples classiques importants ont été introduits par Riesz (*« les produits de Riesz »*) afin de donner des exemples de probabilités continues, portées par l'intervalle unité et qui sont purement singulières[39].

---

38. Une mesure borélienne μ sur $\mathbb{R}$ est positive si $\mu(I) \geq 0$, pour tout intervalle I, elle est sigma-finie lorsque $\mu(I)$ est fini quand l'intervalle I est compact ; enfin μ est régulière si pour tout borélien A, la valeur de $\mu(A)$ coïncide à la fois avec l'infimum des $\mu(U)$, pour U ouvert contenant A, mais aussi au supremum des $\mu(K)$, pour K compact contenu dans A.

39. La pure singularité des mesures construites par Riesz est une conséquence du lemme de Riemann-Lebesgue affirmant que la transformée de Fourier d'une fonction Lebesgue intégrable est nulle en $\pm\infty$ (Erdös à utilisé cette propriété pour montrer que la *« la convolution de Bernoulli »* associée au nombre d'or est une mesure purement singulière).

8.3. Il existe aussi la notion de *« fonction absolument continue »* [40]. Pour voir cela, notons qu'il découle du *« le théorème de dérivation de Lebesgue »* que pour toute fonction réelle $f \in L^1(\mathbb{R})$ et tout $a \in \mathbb{R}$, en posant

$$F(x) := F(a) + \int_a^x f(u)du$$

nous définissons une fonction réelle $F \in \mathscr{C}(\mathbb{R})$ dérivable presque partout avec $F'(x) = f(x)$ pour presque tout $x \in \mathbb{R}$: la fonction F est appelée *« une primitive (au sens de Lebesgue) »* de la fonction f. La réciproque n'est pas vraie en général, en ce sens qu'il existe des fonctions réelles F continues sur $\mathbb{R}$ dérivables presque partout, de dérivée F' localement intégrable, mais ne coïncidant pas avec la primitive de leur dérivée (i.e. $F(a) + \int_a^x F'(u)du$ diffère de $F(x)$ sur un ensemble de mesure de Lebesgue strictement positif). L'exemple classique s'obtient à partir de *« la mesure de Hausdorff de l'ensemble triadique de Cantor »* : si nous notons $\mu(dx)$ cette mesure et que pour tout $x \in \mathbb{R}$, nous posons $F(x) := \int \mathbf{1}_{[0\,;x]}(u)\mu(du)$, alors F (i.e. *« l'escalier du diable »*) est une fonction de $\mathscr{C}(\mathbb{R})$ qui est dérivable presque partout avec $F'(x) = 0$ pour presque tout $x \in \mathbb{R}$. Ce préambule justifie la définition suivante.

Définition 8.2. (i) : *Une fonction réelle $f \in \mathscr{C}(\mathbb{R})$ dont la dérivée $f'(x)$ existe presque partout est dite « absolument continue » lorsque sa dérivée f' est identifiable à une fonction de $L^1_{\text{loc}}(\mathbb{R})$ et si de plus f est une primitive de sa dérivée en ce sens pour tout $-\infty < a \leq b < +\infty$*

$$\int_a^b f'(u)du = f(b) - f(a)$$

(ii) : *La fonction complexe $f \in \mathscr{C}(\mathbb{R})$ sera dite « absolument continue » ssi sa partie réelle et sa partie imaginaire sont absolument continues.*

Notons $\Theta(x) := \mathbf{1}_{[0\,;+\infty]}(x)$ la fonction de Heaviside. Alors $x\Theta(x)$ est une fonction absolument continue dont la dérivée s'identifie à la fonction de Heaviside $\Theta(x)$ – ou plus rigoureusement à la classe de $\Theta(x)$ dans $L^1_{\text{loc}}(\mathbb{R})$. Noter que la fonction $\Theta(x)$ est dérivable presque partout et que sa dérivée s'identifie à la fonction nulle. Ainsi la fonction de Heaviside (qui est continue par morceaux) n'est clairement pas une primitive de sa dérivée.

8.4. Nous notons $\mathscr{P}(\mathbb{R})$ l'ensemble convexe des *« probabilités »* sur $\mathbb{R}$, c'est-à-dire des mesures $\mu \in \mathscr{M}(\mathbb{R})$ t.q. $\mu(\mathbb{R}) = 1$. Pour $n = 0, 1, \ldots$ *« le moment d'ordre n »* d'une probabilité $\mu \in \mathscr{P}(\mathbb{R})$ est défini dès que la fonction $x \mapsto x^n$ est μ-intégrable comme la quantité

$$\boldsymbol{m}_n(\mu) := \int x^n \mu(dx)$$

Le moment d'ordre 0 de μ vaut toujours $\boldsymbol{m}_0(\mu) = 1$ et lorsqu'il existe, le moment $\boldsymbol{m}_1(\mu)$ représentant *« la moyenne de μ »*. Les moments d'ordre pair sont toujours définis dans $[0\,;+\infty]$: notons alors que si la moyenne $\boldsymbol{m}_1(\mu)$ existe, alors la variance

$$\mathrm{Var}(\mu) := \int \big(x - \boldsymbol{m}_1(x)\big)^2 \mu(dx) = \int \big(x^2 - 2x\boldsymbol{m}_1(x) + \boldsymbol{m}_1(\mu)^2\big)\mu(dx) = \boldsymbol{m}_2(\mu) - \boldsymbol{m}_1(\mu)^2$$

40. Sans entrer dans les détails, mentionnons qu'une fonction (réelle) de $L^1_{\text{loc}}(\mathbb{R})$ est absolument continue ssi sa dérivée au sens des distributions est identifiable à une mesure de Radon absolument continue (voir aussi la Proposition 8.5).

est une quantité toujours définie [41] dans $[0\,;+\infty]$; par définition *« l'écart type »* de μ est

$$\sigma(\mu) = \sqrt{\mathrm{Var}(\mu)}$$

8.5. Nous définissons maintenant *« la fonction génératrice des moments »* de $\mu \in \mathscr{P}(\mathbb{R})$ comme la fonction $\Psi_\mu(z)$ dépendant de la variable complexe z et telle que

$$\Psi_\mu(z) = \int e^{xz}\mu(dx)$$

Alors, $\Psi_\mu(z)$ est définie sur un sous-ensemble du plan complexe contenant $\boldsymbol{i}\mathbb{R}$. En écrivant le développement de Taylor de e^{zx} en $z = 0$, il vient (formellement)

$$\Psi_\mu(z) = \int \left(\sum_{n=0}^{\infty}\frac{(zx)^n}{n!}\right)\mu(dx) = \sum_{n=0}^{\infty}\left(\int x^n\mu(dx)\right)\frac{z^n}{n!} = \sum_{n=0}^{\infty}\frac{\boldsymbol{m}_n(\mu)}{n!}z^n$$

Par suite, si le rayon de convergence ρ de la série entière $\sum_{n=0}^{\infty}\boldsymbol{m}_n(\mu)z^n/n!$ est positif, alors la fonction $\Psi_\mu(z)$ est holomorphe en tout point z t.q. $|z| < \rho$ avec

$$\Psi_\mu(z) = \sum_{n=0}^{\infty}\frac{\boldsymbol{m}_n(\mu)}{n!}z^n$$

L'existence [42] des moments $\boldsymbol{m}_1(\mu), \boldsymbol{m}_2(\mu), \ldots$ (dans $\mathbb{R}$) est ainsi une condition nécessaire (mais non suffisante) assurant que $\Psi_\mu(z)$ soit holomorphe au voisinage de $z = 0$.

8.6. Nous avons vu que la fonction génératrice des moments $\Psi_\mu(z)$ associée à une probabilité μ est définie sur un sous-ensemble du plan complexe contenant $\boldsymbol{i}\mathbb{R}$: *« la fonction caractéristique »* $\Phi_\mu(t)$ est alors définie pour tout $t \in \mathbb{R}$ de sorte que

$$\Phi_\mu(t) = \Psi_\mu(\boldsymbol{i}t) = \int e^{\boldsymbol{i}tx}\mu(dx)$$

Pour $\mu \in \mathscr{P}(\mathbb{R})$ et pour tout $t \in \mathbb{R}$, la fonction $x \mapsto e^{\pm\boldsymbol{i}tx}$ est μ-intégrable : *« la transformée de Fourier de μ »* (en analyse ou en synthèse) est notée $\mathfrak{F}_\pm[\mu]$ et définie en posant

$$\mathfrak{F}_\pm[\mu](t) = \frac{1}{\sqrt{2\pi}}\int e^{\pm\boldsymbol{i}tx}\mu(dx) \tag{34}$$

Nous utilisons ici les mêmes notations que pour la transformée de Fourier des fonctions de $L^1(\mathbb{R})$ (c.f. § 6 supra) du fait de la compatibilité qui existe entre le deux notions : en effet, si la probabilité μ est absolument continue, c'est-à-dire s'il existe $f \in L^1(\mathbb{R})$ t.q. $\mu(dx) = f(x)dx$, alors $\mathfrak{F}_\pm[\mu] = \mathfrak{F}_\pm[f]$. Notons alors que la fonction caractéristique de μ peut être exprimée en termes de transformée de Fourier, c'est-à-dire en écrivant

$$\Phi_\mu(t) = \sqrt{2\pi}\,\mathfrak{F}_-[\mu](-t) = \sqrt{2\pi}\,\mathfrak{F}_+[\mu](t)$$

Proposition 8.3. *La fonction caractéristique Φ_μ de $\mu \in \mathscr{P}(\mathbb{R})$ vérifie :*

(i) : $|\Phi_\mu(t)| \le 1$ pour tout $t \in \mathbb{R}$ et [43] *$\Phi_\mu(0) = 1$;*

(ii) : Φ_μ est une fonction uniformément continue sur $\mathbb{R}$;

---

41. L'inégalité $\boldsymbol{m}_2(\mu) \ge \boldsymbol{m}_1(\mu)^2$ est aussi un cas particulier de l'inégalité de Jensen, affirmant que si $\mu \in \mathscr{P}(\mathbb{R})$ possède un moment d'ordre 1 et si $f(x)$ est une fonction convexe μ-intégrable, alors $f\left(\int x\mu(dx)\right) \le \int f(x)\mu(dx)$.

42. Si $\mu(dx) = \frac{1}{\pi}dx/(1+x^2)$ est *« la probabilité de Cauchy »* alors $\boldsymbol{m}_1(\mu)$ n'est pas définie : par suite $\Psi_\mu(z)$ n'est pas holomorphe au voisinage de $z = 0$; cependant sa fonction caractéristique Φ_μ est bien définie sur l'ensemble de la droite réelle avec $\Phi_\mu(t) = e^{-|t|}$.

43. La fonction caractéristique de la mesure de Dirac δ_0 concentrée en 0 est $\Phi_{\delta_0}(t) \equiv 1$.

(iii) : si le moment $\boldsymbol{m}_n(\mu)$ existent pour $n \geq 1$, alors les moments $\boldsymbol{m}_k(\mu)$ existent pour $k = 0, \ldots, n$; de plus, $\Phi_\mu(t)$ possède un développement de Taylor-Young d'ordre n en $t = 0$ soit :

$$\Phi_\mu(t) = 1 + \sum_{k=1}^{n} \frac{(\boldsymbol{i}t)^k}{k!}\boldsymbol{m}_k(\mu) + \circ(t^n) \tag{35}$$

En particulier, $\Phi_\mu(t)$ est n fois dérivable en $t = 0$ et pour tout $k = 0, \cdots, n$:

$$\frac{d^k\Phi_\mu}{dt^k}(0) = \boldsymbol{m}_k(\mu)$$

Preuve. (i)-(ii) : La partie (i) est immédiate. Pour (ii), remarquons qu'avec $t, h \in \mathbb{R}$

$$|\Phi_\mu(t+h) - \Phi_\mu(t)| = \left|\int e^{\boldsymbol{i}tx}\big(e^{\boldsymbol{i}hx} - 1\big)\mu(dx)\right| \leq \int \left|e^{\boldsymbol{i}hx} - 1\right|\mu(dx) =: \eta(h)$$

Le *« théorème de la convergence dominée »* assure que $\eta(h) \to 0$ quand $h \to 0$.

(iii) : Remarquons d'abord que si k est un entier positif alors, le fait que x^k soit μ-intégrable entraîne que x^k est aussi μ-intégrable (si $|x| \geq 1$ alors $|x|^{k-1} \leq |x|^k$). L'existence du moment $\boldsymbol{m}_n(\mu)$ entraîne donc l'existence des moments $\boldsymbol{m}_k(\mu)$ pour $0 \leq k \leq n$. Maintenant, fixons $t \in \mathbb{R}$; par la *« formule de Taylor-Lagrange »* (d'ordre n) appliquée séparément à la partie réelle de la fonction $x \mapsto e^{\boldsymbol{i}tx}$ ainsi qu'à sa partie imaginaire, nous obtenons l'existence de deux réels $\alpha = \alpha(x)$ et $\beta = \beta(x)$ dans $[0\,;1]$ tels que

$$e^{\boldsymbol{i}tx} = \left(\sum_{k=0}^{n-1} \frac{(\boldsymbol{i}tx)^k}{k!}\right) + \frac{(\boldsymbol{i}tx)^n}{n!}\Big(\cos(\alpha tx) + \boldsymbol{i}\sin(\beta tx)\Big)$$

soit encore

$$e^{\boldsymbol{i}tx} = \left(\sum_{k=0}^{n} \frac{(\boldsymbol{i}tx)^k}{k!}\right) + \frac{(\boldsymbol{i}tx)^n}{n!}\Big(\cos(\alpha tx) + \boldsymbol{i}\sin(\beta tx) - 1\Big) \tag{36}$$

En intégrant (36) par rapport à $\mu(dx)$, nous obtenons que

$$\Phi_\mu(t) = \sum_{k=0}^{n} \frac{(\boldsymbol{i}t)^k}{k!}\boldsymbol{m}_k(\mu) + t^n\varepsilon(t) \tag{37}$$

où nous avons posé

$$\varepsilon(t) := \int \frac{(\boldsymbol{i}x)^n}{n!}\Big(\cos(\alpha tx) + \boldsymbol{i}\sin(\beta tx) - 1\Big)\mu(dx)$$

Or d'une part pour tout x

$$\lim_{t\to 0} \frac{(\boldsymbol{i}x)^n}{n!}\Big(\cos(\alpha tx) + \boldsymbol{i}\sin(\beta tx) - 1\Big) = 0$$

et d'autre part pour tout $t \in \mathbb{R}$ (et pour $n \geq 3$),

$$\left|\frac{(\boldsymbol{i}x)^n}{n!}\Big(\cos(\alpha tx) + \boldsymbol{i}\sin(\beta tx) - 1\Big)\right| \leq |x^n|$$

Par suite, le fait que $x \mapsto x^n$ soit par hypothèse intégrable (existence du moment $\boldsymbol{m}_n(\mu)$), permet de déduire – par une nouvelle application du *« théorème de la convergence dominée »* – que $\varepsilon(t) \to 0$ quand $t \to 0$: l'identité (37) démontre ainsi la validité de (iii).

□

8.7. Dans la suite nous noterons

$$\mathscr{P}_{\mathrm{d}}(\mathbb{R}) := \mathscr{M}_{\mathrm{d}}(\mathbb{R}) \cap \mathscr{P}(\mathbb{R}) \qquad \mathscr{P}_{\mathrm{c}}(\mathbb{R}) := \mathscr{M}_{\mathrm{c}}(\mathbb{R}) \cap \mathscr{P}(\mathbb{R})$$

$$\mathscr{P}_{\mathrm{ac}}(\mathbb{R}) := \mathscr{M}_{\mathrm{ac}}(\mathbb{R}) \cap \mathscr{P}(\mathbb{R}) \qquad \mathscr{P}_{\mathrm{ps}}(\mathbb{R}) := \mathscr{M}_{\mathrm{ps}}(\mathbb{R}) \cap \mathscr{P}(\mathbb{R})$$

Il est immédiat que ces ensembles sont tous des sous-ensembles convexes de $\mathscr{M}(\mathbb{R})$; le résultat suivant découle directement du Théorème 8.1.

Théorème 8.4. *Pour toute mesure $\mu \in \mathscr{P}(\mathbb{R})$ il existe trois réels $\alpha, \beta, \gamma \geq 0$ avec $\alpha + \beta + \gamma = 1$ ainsi qu'un triplet de probabilités $(\mu_{\mathrm{d}}, \mu_{\mathrm{ac}}, \mu_{\mathrm{ps}}) \in \mathscr{P}_{\mathrm{d}}(\mathbb{R}) \times \mathscr{P}_{\mathrm{ac}}(\mathbb{R}) \times \mathscr{P}_{\mathrm{ps}}(\mathbb{R})$ de sorte que*

$$\mu = \alpha\mu_{\mathrm{d}} + \beta\mu_{\mathrm{ac}} + \gamma\mu_{\mathrm{ps}}$$

et où $\beta\mu_{\mathrm{ac}} + \gamma\mu_{\mathrm{ps}}$ est la partie continue de μ ; de plus cette décomposition est unique.

8.8. Une fonction $F : \mathbb{R} \to [0\,;1]$ croissante, continue à droite et t.q. $(F(-\infty), F(+\infty)) = (0,1)$ est appelée *« une fonction de répartition »*. Toute probabilité μ sur $\mathbb{R}$ est associée à la fonction de répartition $F_\mu : \mathbb{R} \to [0\,;1]$ t.q. $F_\mu(x) = \mu(]-\infty\,;x])$: l'application $\mu \mapsto F_\mu$ ainsi définie réalise une bijection de $\mathscr{P}(R)$ sur l'ensemble convexe des fonctions de répartition. Supposons maintenant que $\mu \in \mathscr{P}_{\mathrm{ac}}(\mathbb{R})$ et notons $f_\mu \in L^1(\mathbb{R})$ la densité de μ ; alors par définition de la fonction de répartition de μ, nous savons que pour tout $-\infty < a \leq b < +\infty$

$$F_\mu(b) - F_\mu(a) = \mu([a\,;b]) = \int_a^b f_\mu(x)dx$$

Par la définition d'une fonction absolument continue (c.f. Définition 8.2) il vient :

Proposition 8.5. *$\mu \in \mathscr{P}(\mathbb{R})$ est absolument continue ssi F_μ est absolument continue.*

8.9. Nous avons déjà noté que la fonction caractéristique d'une probabilité μ absolument continue s'exprime grâce à la transformée de Fourrier de sa densité f_μ, en ce sens que $\Phi_\mu = \sqrt{2\pi}\,\mathfrak{F}_+[f_\mu]$, où $\mathfrak{F}_+[\cdot]$ est la transformée de Fourier L^1. Par suite, si nous supposons que Φ_μ est aussi dans $L^1(\mathbb{R})$, alors nous pouvons appliquer la formule d'inversion, ce qui donne pour presque tout $x \in \mathbb{R}$:

$$f_\mu(x) = \frac{1}{\sqrt{2\pi}}\mathfrak{F}_-[\Phi_\mu](x) = \frac{1}{2\pi}\int \Phi_\mu(t)e^{-itx}dt \tag{38}$$

Proposition 8.6. *Si la fonction caractéristique Φ_μ associée à $\mu \in \mathscr{P}(\mathbb{R})$ est Lebesgue intégrable alors μ est absolument continue et de densité (bornée) $f_\mu(x) = 1/\sqrt{2\pi}\mathfrak{F}_-[\Phi_\mu](x)$.*

8.10. Le cas purement discret est – d'une certaine manière – analogue au cas absolument continue. Pour voir cela considérons $\mu \in \mathscr{P}_{\mathrm{d}}(\mathbb{R})$ et commençons par supposer que le support μ est inclus dans l'ensemble $\mathbb{Z}$ des nombres entiers relatifs. Si pour tout $n \in \mathbb{Z}$ nous posons $p_n := \mu\{n\}$ alors la fonction caractéristique de μ est $\Phi_\mu(t) = \sum_n p_n e^{int}$. Il est remarquable que $\Phi_\mu : \mathbb{R} \to \mathbb{C}$ soit une fonction 2π-périodique dont les coefficients de Fourier $p = (\ldots, p_{-1}, p_0, p_1, \ldots)$ forment un vecteur probabilité sur $\mathbb{Z}$ en ce sens que $\sum_n p_n = 1$: puisque $p_n^2 \leq p_n$, il est alors immédiat que $\sum_n p_n^2 < +\infty$, ce qui signifie que $p \in \ell^2(\mathbb{Z})$. En particulier, nous pouvons dire que $\Phi_\mu \in L^2[0\,;2\pi]$ (espace de Hilbert des fonctions boréliennes 2π-périodiques et de carré Lebesgue-intégrable sur $[0\,;2\pi]$). Par la définition des coefficients de Fourier, nous obtenons ainsi que

$$p_n = \frac{1}{2\pi}\int_0^{2\pi} \Phi_\mu(t)e^{-int}dt \tag{39}$$

Dans le cas général, la probabilité discrète μ n'a aucunes raisons d'avoir son support inclus dans un réseau comme $\mathbb{Z}$: si nous supposons que le support de μ est $\{x_n\}_{n=1}^{\infty}$, alors, en notant $p_n := \mu\{x_n\}$, la fonction caractéristique Φ_μ est en général *« quasi-périodique »* et la formule des coefficients de Fourier (39) est remplacée par

$$p_n = \lim_{n} \frac{1}{2R}\int_{-R}^{R} \Phi_\mu(t) e^{-\boldsymbol{i}x_n t} dt \tag{40}$$

Le résultat peut s'énoncer de manière générale (i.e. pour μ continue, discrète où mixte) de la façon suivante :

Théorème 8.7. *Soit Φ_μ la fonction caractéristique associée à une probabilité $\mu \in \mathscr{P}(\mathbb{R})$: alors*

$$(\forall x \in \mathbb{R}) \qquad \mu\{x\} = \lim_{R\to+\infty} \frac{1}{2R}\int_{-R}^{R} \Phi_\mu(t) e^{-\boldsymbol{i}tx} dt$$

Preuve. En utilisant le *« théorème de Fubini »*, nous obtenons que

$$\begin{aligned}
\frac{1}{2R}\int_{-R}^{R} \Phi_\mu(t) e^{-\boldsymbol{i}tx} dt &= \frac{1}{2R}\int_{-R}^{R} \Big(\int e^{\boldsymbol{i}ty}\mu(dy)\Big) e^{-\boldsymbol{i}xt} dt \\
&= \frac{1}{2R}\int \Big(\int_{-R}^{R} e^{\boldsymbol{i}t(y-x)} dt\Big)\mu(dy) \\
&= \frac{1}{2R}\int \Big[e^{\boldsymbol{i}t(y-x)}\Big]_{t=-R}^{t=R}\mu(dy) \\
&= \int \left(\frac{e^{\boldsymbol{i}R(y-x)} - e^{\boldsymbol{i}R(y-x)}}{2\boldsymbol{i}R(y-x)}\right)\mu(dy) = \int S\big(R(y-x)\big)\mu(dy)
\end{aligned}$$

où nous avons posé $S(u) = \sin(u)/u$. Or dans la limite où $R \to +\infty$, nous avons d'une part $S\big(R(y-x)\big) \to \mathbf{1}_{\{x\}}(y)$; d'autre part, la fonction $y \mapsto S\big(R(y-x)\big)$ est dominée par $\mathbf{1}_{\mathbb{R}}(y)$ $(\equiv 1)$ qui est une fonction μ-intégrable (car μ est une probabilité) : la conclusion vient avec le *« théorème de la convergence dominée »*.

□

Le Théorème 8.7, démontre que la fonction caractéristique d'une probabilité $\mu \in \mathscr{P}_{\mathrm{d}}(\mathbb{R})$ est *caractéristique* de μ ; le théorème suivante généralise cela à tous les types probabilités.

Théorème 8.8 (P. Levy)**.** *Soit μ une probabilité dans $\mathscr{P}(\mathbb{R})$: alors, pour tout $-\infty < a \le b < +\infty$*

$$\lim_{R\to+\infty} \frac{1}{2\pi}\int_{-R}^{R} \frac{e^{-\boldsymbol{i}at} - e^{-\boldsymbol{i}bt}}{\boldsymbol{i}t}\Phi_\mu(t) dt = \mu(]a\,;b[) + \frac{1}{2}\Big(\mu\{a\} + \mu\{b\}\Big) \tag{41}$$

Preuve. Nous commençons en remarquant que

$$\left|\frac{e^{-\boldsymbol{i}ta} - e^{-\boldsymbol{i}tb}}{\boldsymbol{i}t}\right| = \left|\int_a^b e^{\boldsymbol{i}tx} dx\right| \le |b-a|$$

Par suite en utilisant le *« théorème de Fubini »*, nous pouvons écrire que

$$\begin{aligned}
\frac{1}{2\pi}\int_{-R}^{R} \frac{e^{-\boldsymbol{i}ta} - e^{-\boldsymbol{i}tb}}{\boldsymbol{i}t}\Phi_\mu(t) dt &= \frac{1}{2\pi}\int_{-R}^{R} \frac{e^{-\boldsymbol{i}ta} - e^{-\boldsymbol{i}tb}}{\boldsymbol{i}t}\Big(\int e^{\boldsymbol{i}tx}\mu(dx)\Big) dt \\
&= \int \left(\frac{1}{2\pi}\int_{-R}^{R} \frac{e^{\boldsymbol{i}t(x-a)} - e^{\boldsymbol{i}t(x-b)}}{\boldsymbol{i}t} dt\right)\mu(dx) \\
&= \int \left(\frac{1}{2\pi}\int_{-R}^{R} \frac{\sin\big(t(x-a)\big) - \sin\big(t(x-b)\big)}{t} dt\right)\mu(dx)
\end{aligned}$$

(la fonction $t \mapsto \big(\cos t(x-a) - \cos t(x-b)\big)/t$ est impaire et donc son intégrale est nulle sur l'intervalle $[-R\,;R]$). Si nous notons $U_R(x) := \frac{1}{2\pi}\int_{-R}^{R} \frac{\sin(tx)}{t} dt$, alors

$$\int_{-R}^{R} \frac{e^{-ita} - e^{-itb}}{it} \Phi_\mu(t)dt = \int \Big(U_R(x-a) - U_R(x-b)\Big)\mu(dx)$$

Maintenant, nous pouvons écrire en posant $y = tx$

$$U_R(x) := \frac{1}{2\pi}\int_{-R}^{R} \frac{\sin(tx)}{t} dt = \frac{1}{\pi}\int_{0}^{R} \frac{\sin(tx)}{t} d = \frac{1}{\pi}\int_{0}^{xR} \frac{\sin(y)}{y} dy = \frac{\mathrm{sg}(x)}{\pi}\int_{0}^{|x|R} \frac{\sin(y)}{y} dy$$

où $\mathrm{sg}(x)$ est le signe de x (avec $\mathrm{sg}(0) = 0$). Or, avec la valeur de *« l'intégrale de Dirichlet »*

$$\lim_{R\to+\infty} U_R(x) = \frac{\mathrm{sg}(x)}{\pi}\int_{0}^{\infty} \frac{\sin(y)}{y} dy = \frac{\mathrm{sg}(x)}{2}$$

Par suite, il existe $M \geq 0$ t.q. $\|U_R\|_\infty \leq M$, pour tout $R > 0$ et donc, par une application licite du théorème de la convergence dominée, il vient finalement que

$$\lim_{R\to+\infty} \int \Big(U_R(x-a) - U_R(x-b)\Big)\mu(dx) = \mu(]a\,;b[) + \frac{1}{2}\Big(\mu\{a\} + \mu\{b\}\Big)$$

□

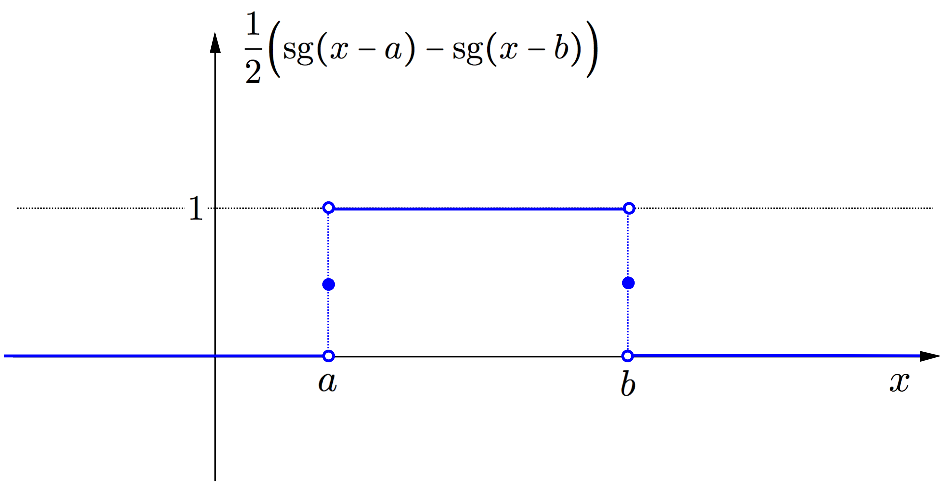

FIGURE 2. *Par définition $U(x) = \mathrm{sg}(x)/2$, où $\mathrm{sg}(x)$ est le signe de x, avec $\mathrm{sg}(0) = 0$. Nous illustrons ici la fonction $U(x-a) - U(x-b)$ afin d'éclairer la présence du terme $\big(\mu\{a\} + \mu\{b\}\big)/2$ dans (41).*

8.11. Il y a plusieurs manières de munir $\mathscr{P}(\mathbb{R})$ d'une topologie, une d'entre-elle étant de considérer les probabilités comme des distributions (voir aussi la Remarque 8.11). Cependant, dans un grand nombre d'applications, il suffit d'avoir une notion adéquate pour la convergence des suites de probabilités : la définition suivante remplit ce rôle.

Définition 8.9. *Une suite $\mu_1, \mu_2, \ldots$ dans $\mathscr{P}(\mathbb{R})$ « convergent faiblement vers μ » dans $\mathscr{P}(\mathbb{R})$ ssi pour toute fonction f dans l'espace $\mathscr{C}_b(\mathbb{R})$ des fonction s réelles continues et bornées sur $\mathbb{R}$*

$$\lim_{n\to+\infty} \int f(x)\mu_n(dx) = \int f(x)\mu(dx)$$

Si la suite de probabilités $\mu_1, \mu_2, \ldots$ tend faiblement vers μ alors les fonctions caractéristiques $\Phi_{\mu_1}, \Phi_{\mu_2}, \ldots$ tendent simplement vers Φ_μ. Le théorème suivant (*« théorème de continuité de Paul Lévy »*) précise les conditions de la réciproque de cette propriété.

Théorème 8.10 (P. Lévy). *Soit $\mu_1, \mu_2, \ldots$ une suite de probabilités sur $\mathbb{R}$; si les fonctions caractéristiques $\Phi_{\mu_1}, \Phi_{\mu_2}, \ldots$ convergent ponctuellement vers $\phi : \mathbb{R} \to \mathbb{C}$ continue en 0, alors les μ_n tendent faiblement vers une probabilité μ telle que $\Phi_\mu(x) = \phi(x)$, pour tout $x \in \mathbb{R}$.*

Preuve. Voir par exemple [Bas71, § V-1-4, p. 332], [Ber92, p. 51] ou [KF94, Chap. VI].

□

Remarque 8.11. *Nous avons défini directement la notion de convergence faible pour les suites de $\mathscr{P}(\mathbb{R})$, sans référence à une quelconque topologie sur $\mathscr{P}(\mathbb{R})$. Pour définir une telle topologie, rappelons que si $(E, \|\cdot\|)$ est un espace de Banach, alors son dual topologique E' peut-être muni de « la topologie $\ast$-faible » (c.f.* [Bre87]*) : une suite $\varphi_1, \varphi_2, \ldots$ de formes linéaires dans E' convergent alors $\ast$-faiblement vers φ ssi pour tout $x \in E$, la suite des $\varphi_n(x)$ tends vers $\varphi(x)$, quand $n \to +\infty$. Si nous considérons comme espace de Banach, l'espace $\mathscr{C}_b(\mathbb{R})$ des fonctions réelles continues et bornées sur $\mathbb{R}$, muni de la norme infinie $\|\cdot\|_\infty$, alors $\mathscr{P}(\mathbb{R})$ est une partie convexe du dual topologique de $\mathscr{C}_b(\mathbb{R})$: il est alors immédiat que la convergence faible d'une suite de probabilités $\mu_1, \mu_2, \ldots$ vers une probabilité μ coïncide avec la convergence $\ast$-faible dans le dual topologique de $\mathscr{C}_b(\mathbb{R})$ – notons au passage que $\mathscr{P}(\mathbb{R})$ est une partie convexe et $\ast$-faiblement fermée de $\mathscr{C}_b(\mathbb{R})$.*

8.12. Un aspect central de la théorie des probabilités (telle que formalisée en particulier par Kolmogorov [44]) porte sur la notion de *« variable aléatoire »*. Nous nous limitons ici aux propriétés élémentaires des variables aléatoires réelles. Etant donné $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé, une variable aléatoire (réelle) est spécifiée par une application borélienne $X : \Omega \to \mathbb{R}$; la probabilité $\mu := \mathbb{P} \circ X^{-1}$ (définie sur les boréliens de $\mathbb{R}$) est par définition *« la loi de X »*. Alors l'espérance de X (si elle est définie) est par définition

$$\mathbb{E}(X) := \int X(\omega)\mathbb{P}(d\omega) = \int x\mu(dx)$$

Pour toute fonction $f : \mathbb{R} \to \mathbb{R}$ borélienne, l'application $Y = f(X) = \Omega \to \mathbb{R}$ est encore une variable aléatoire dont l'espérance (si toutefois elle existe) est

$$\mathbb{E}(Y) = \int Y(\omega)\mathbb{P}(d\omega) = \int f(X(\omega))\mathbb{P}(d\omega) = \int f(x)\mu(dx)$$

Notons alors, que pour tout $k = 0, 1, \ldots$ l'espérance $\mathbb{E}(X^k)$ (si elle existe) coïncide avec le k-ème moment $\boldsymbol{m}_1(\mu)$ de la loi de X. Enfin, si l'espérance de X est bien définie, alors la variance de X est l'espérance de $(X - \mathbb{E}(X))^2$ c'est-à-dire le nombre (de $[0\,;+\infty]$)

$$\mathrm{Var}(X) = \mathbb{E}\big((X - \mathbb{E}(X))^2\big) = \int \big(X(\omega) - \mathbb{E}(X)\big)^2\big)\mathbb{P}(d\omega) = \int (x - \mathbb{E}(X))^2\mu(dx)$$

En d'autres termes, la variance $\mathrm{Var}(X)$ coïncide avec la variance $\mathrm{Var}(\mu)$ de la loi de X. De manière analogue au cas de la variance d'une probabilité nous pouvons utiliser les

44. En 1929 Kolmogorov étend de manière significative *« la loi faible des grands nombres »* sous la forme de *« la lois forte des grands nombres »* : l'énoncé même de cette lois nécessité la notion de convergence presque sûre, qui vient directement de la théorie de la mesure telle que développée à partir de la construction de l'intégrale de Lebesgue (1902).

propriétés de linéarité de l'espérance pour écrire que

$$\begin{aligned}\mathrm{Var}(X) &= \mathbb{E}\big(X^2 - 2X\mathbb{E}(X) + \mathbb{E}(X)^2\big)\\ &= \mathbb{E}\big(X^2\big) - 2\mathbb{E}(X\mathbb{E}(X)) + \mathbb{E}(X)^2\\ &= \mathbb{E}\big(X^2\big) - 2\mathbb{E}(X)\mathbb{E}(X) + \mathbb{E}(X)^2 = \mathbb{E}\big(X^2\big) - \mathbb{E}(X)^2\end{aligned}$$

8.13. Nous nous sommes limités jusqu'ici à des variables aléatoires réelles ; par définition, une variable aléatoire complexe (i.e. à valeurs complexes) s'écrit $Z = X + \boldsymbol{i}Y$, où X et Y sont des variables aléatoire réelles ; l'espérance de Z est alors définie en posant

$$\mathbb{E}(Z) = \mathbb{E}(X) + \boldsymbol{i}\mathbb{E}(Y)$$

En particulier, si X est une variable aléatoire réelle, alors pour tout $z = a + \boldsymbol{i}b \in \mathbb{C}$:

$$e^{zX} = e^{(a+\boldsymbol{i}b)X} = e^{aX}\big(\cos(bX) + \boldsymbol{i}\sin(bX)\big)$$

est une variable aléatoire complexe. Si μ est la loi de X alors $\Psi_X(z) := \Psi_\mu(z)$ (pour z dans le domaine du plan complexe adéquat) est *« la fonction génératrice des moments de X »* ; en utilisant l'espérance des variables aléatoires complexes il vient alors

$$\Psi_X(z) = \mathbb{E}(e^{zX})$$

De même $\Phi_X(t) = \Psi_X(\boldsymbol{i}t) = \Phi_\mu(t)$ est la *« fonction caractéristique de X »* de sorte que

$$\Phi_X(t) = \mathbb{E}(e^{\boldsymbol{i}tX})$$

8.14. Il existe un certain nombre d'inégalités probabilistes qui donnent des indications importantes sur la *répartition* des probabilités. La proposition suivante nous donne deux exemples classiques de telles inégalités.

Proposition 8.12. *Soit X une variable aléatoire réelle possédant une espérance ; (i) : si X est à valeurs positives ou nulle, alors « l'inégalité de Markov » affirme que*

$$(\forall a \geq 0) \qquad \mathbb{P}\big\{Z \geq a\big\} \leq \frac{\mathbb{E}(Z)}{a} \tag{42}$$

(ii) : si X est simplement à valeurs réelles, alors « l'inégalité de Tchebychev » [45] *affirme que*

$$(\forall a \geq 0) \qquad \mathbb{P}\big\{|X - \mathbb{E}(X)| \geq a\big\} \leq \frac{\mathrm{Var}(X)}{a^2} \tag{43}$$

Preuve. (i) : La variable aléatoire X étant positive, nous avons $X(\omega) \geq a\mathbf{1}_{\{X\geq a\}}(\omega)$, pour tout $\omega \in \Omega$. L'existence de l'espérance de X nous permet alors d'écrire que

$$\mathbb{E}(X) \geq \mathbb{E}\big(a\mathbf{1}_{\{X\geq a\}}\big) = a\mathbb{P}\{X \geq a\}$$

Nous en déduisons directement *« l'inégalité de Markov »* telle que donnée en (42).

(ii) : Supposons que X est une variable aléatoire possédant une espérance $\mathbb{E}(X)$. Si la variance de X est infinie alors l'inégalité (43) est évidente. Dans le cas où la variance $\mathrm{Var}(X)$ est finie, $Z := (X - \mathbb{E}(X))^2$ est une variable aléatoire positive dont l'espérance est $\mathbb{E}(Z) = \mathrm{Var}(X)$. En appliquant l'inégalité de Markov à Z avec $a > 0$ il vient $\mathbb{P}\{Z \geq a^2\} \leq \mathbb{E}(Z)/a^2$: du fait que $Z(\omega) \geq a^2$ ssi $|X(\omega) - \mathbb{E}(X)| \geq a$, nous obtenons ainsi *« l'inégalité de Tchebychev »* telle que donnée en (42).

45. L'inégalité de Tchebychev est un résultat simple et très puissant de la théorie des probabilités. Son énoncé est dû à Bienaymée (un ami de Tchebychev) ; elle a été démontrée par Tchebychev. On la présente habituellement comme une conséquence de l'inégalité de Markov (élève de Tchebychev).

$\square$

8.15. Deux variables (réelles)[46] X et Y sur $(\Omega, \mathcal{F}, \mathbb{P})$ sont dites indépendantes – et on note $X \perp Y$ – ssi pour tout couple (A, B) de boréliens de $\mathbb{R}$

$$\mathbb{P}\Big(\{X \in A\} \cap \{Y \in B\}\Big) = \mathbb{P}\{X \in A\}\,\mathbb{P}\{Y \in B\}$$

Proposition 8.13. *Soient X et Y deux variables aléatoires ayant chacune un moment d'ordre 2 ; alors (i) : la variable aléatoire XY possède aussi un moment d'ordre 1 et*

$$\mathbb{E}^2(XY) \leq \mathbb{E}(X^2)\mathbb{E}(Y^2)$$

(ii) : Si de plus X et Y sont indépendantes, alors

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$$

Preuve (esquisse). (i) : Soit $(\Omega, \mathcal{F}, \mathbb{P})$, l'espace probabilisé commun à X et Y. Le fait que X et Y possèdent chacune un moment d'ordre 2 signifie que ce sont toutes deux des fonctions de carré $\mathbb{P}$-intégrable : par suite *« d'après l'inégalité de Cauchy-Schwarz »*, XY est une fonction $\mathbb{P}$-intégrable telle que $\mathbb{E}^2(XY) \leq \mathbb{E}(X^2)\mathbb{E}(Y^2)$.

(ii) : Pour $N \geq 1$ un entier donné, soient $I_0 :=]-1/2^N\,;1/2^N[$ et $I_k := [k/2^N\,;(k+1)/2^N[$, pour tout $k = 1, 2, \ldots$; si $I_{-k} := -I_k$, alors nous avons les *approximations* (c.f. Fig. 3)

$$X(\omega) \approx \sum_{k=-\infty}^{+\infty} \frac{k}{2^N} \mathbf{1}_{X^{-1}(I_k)}(\omega) =: \tilde{X}(\omega) \quad \text{et} \quad Y(\omega) \approx \sum_{k=-\infty}^{+\infty} \frac{k}{2^N} \mathbf{1}_{Y^{-1}(I_k)}(\omega) =: \tilde{Y}(\omega)$$

Alors, par un calcul direct (qu'il faut justifier)

$$\begin{aligned}
\mathbb{E}(\tilde{X}\tilde{Y}) &= \int \tilde{X}(\omega)\tilde{Y}(\omega)\mathbb{P}(d\omega) \\
&= \int \left(\sum_{p,q} \frac{p}{2^N}\frac{q}{2^N} \mathbf{1}_{X^{-1}(I_p) \cap Y^{-1}(I_q)}(\omega) \right) \mathbb{P}(d\omega) \\
&= \sum_{p,q} \frac{p}{2^N}\frac{q}{2^N} \mathbb{P}\Big(\{X^{-1}(I_p)\} \cap \{Y^{-1}(I_q)\} \Big) \\
&= \sum_{p,q} \frac{p}{2^N}\frac{q}{2^N} \mathbb{P}\{X^{-1}(I_p)\}\mathbb{P}\{Y^{-1}(I_q)\} = \left(\sum_p \mathbb{P}\{X^{-1}(I_p)\} \right) \left(\sum_q \frac{q}{2^N} \mathbb{P}\{Y^{-1}(I_q)\} \right)
\end{aligned}$$

soit encore

$$\mathbb{E}(\tilde{X}\tilde{Y}) = \Big(\int \tilde{X}(\omega)\mathbb{P}(d\omega) \Big) \Big(\int \tilde{Y}(\omega)\mathbb{P}(d\omega) \Big) = \mathbb{E}(\tilde{X})\mathbb{E}(\tilde{Y})$$

Nous en déduisons $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ en faisant tendre $N \to +\infty$.

$\square$

Corollaire 8.14. *Si X et Y sont deux variables aléatoires (réelles) indépendantes, alors*

$$\Phi_{X+Y} = \Phi_X \Phi_Y$$

46. Lorsque nous considérons une famille finie, ou une suite de variables aléatoires, nous supposerons toujours qu'elles sont toutes définies sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$.

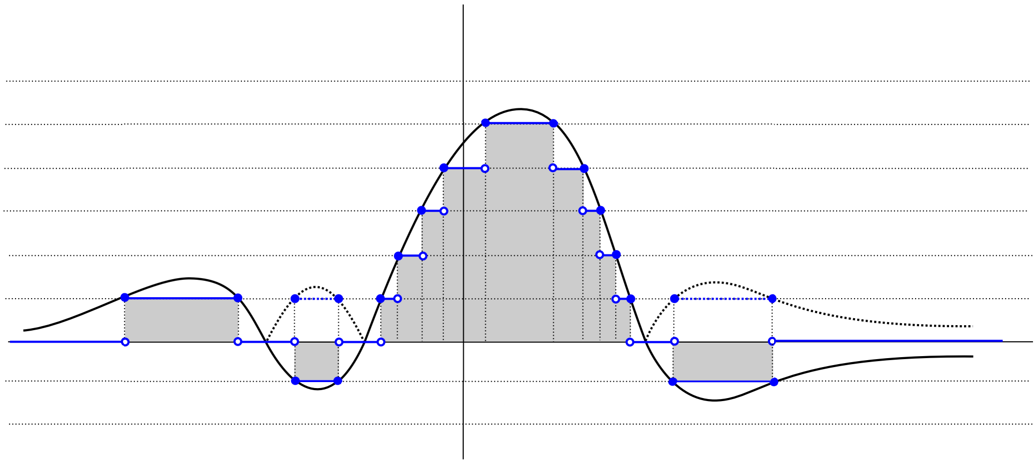

FIGURE 3. *Etant donnée $f \in L^1(\mathbb{R})$ et $N \geq 1$ un entier, nous définissons*

$$f_N := \sum_{k=-\infty}^{\infty} k/2^N \mathbf{1}_{I_k}$$

où $I_0 :=]-1/2^N\,;1/2^N[$, $I_k := [k/2^N\,;(k+1)/2^N[$, pour tout $k = 1, 2, \ldots$ et $I_{-k} := -I_k$; alors pour tout $x \in \mathbb{R}$

$$|f_N(x)| = \sum_{k=-\infty}^{\infty} |k|/2^N \mathbf{1}_{I_k}(x) \leq |f(x)|$$

et d'autre part $f_N(x) \to f(x)$ p.p., quand $N \to +\infty$.

Preuve (esquisse). Soient f et g deux fonctions réelles, boréliennes et définies sur $\mathbb{R}$; si $X \perp Y$ alors $f(X) \perp f(Y)$; le résultat s'obtient alors par une application de la Proposition 8.13 en écrivant

$$e^{\boldsymbol{i}t(X+Y)} = \cos(tX)\cos(tY) - \sin(tX)\sin(tY) + \boldsymbol{i}\Big(\cos(tX)\sin(tY) + \sin(tX)\cos(tY)\Big)$$

de sorte que

$$\begin{aligned}
\Phi_{X+Y}(t) &= \mathbb{E}\big(\cos(tX)\cos(tY)\big) - \mathbb{E}\big(\sin(tX)\sin(tY)\big) \\
&\qquad + \boldsymbol{i}\,\mathbb{E}\big(\cos(tX)\sin(tY)\big) + \boldsymbol{i}\,\mathbb{E}\big(\sin(tX)\cos(tY)\big) \\
&= \mathbb{E}\big(\cos(tX)\big)\mathbb{E}\big(\cos(tY)\big) - \mathbb{E}\big(\sin(tX)\big)\mathbb{E}\big(\sin(tY)\big) \\
&\qquad + \boldsymbol{i}\,\mathbb{E}\big(\cos(tX)\big)\mathbb{E}\big(\sin(tY)\big) + \boldsymbol{i}\,\mathbb{E}\big(\sin(tX)\big)\mathbb{E}\big(\cos(tY)\big) \\
&= \mathbb{E}\big(\cos(tX) + \boldsymbol{i}(\sin(tX)\big)\mathbb{E}\big(\cos(tY) + \boldsymbol{i}\sin(tY)\big)
\end{aligned}$$

soit encore $\Phi_{X+Y}(t) = \Phi_X(t)\Phi_Y(t)$.

□

8.16. Soit $X_1, X_2, \ldots$ une suite de variables aléatoires *« indépendante et identiquement distribuée (i.i.d.) »* : cela signifie que les X_i ont même loi et que $X_i \perp X_j$ dès que $i \neq j$; en particulier (sous réserve) $\mathbb{E}(X_1)$ et $\mathrm{Var}(X_1)$ représente respectivement l'espérance et la variance communes des X_k. Si nous notons $S_n = X_1 + \cdots + X_n$, alors il est immédiat que $\mathbb{E}(S_n) = n\mathbb{E}(X_1)$. Pour calculer le moment d'ordre 2 de S_n, nous utilisons le fait que les variables aléatoires X_k sont deux à deux indépendantes : nous obtenons

$$\begin{aligned}\mathbb{E}(S_n^2) &= \mathbb{E}\left(\sum_{p=1}^{n} X_p^2 + \sum_{1\le p\neq q\le n} X_p X_q\right) \\ &= \sum_{p=1}^{n}\mathbb{E}(X_p^2) + \sum_{1\le p\neq q\le n}\mathbb{E}(X_p)\mathbb{E}(X_q) = n\mathbb{E}(X_1^2) + (n^2-n)\mathbb{E}^2(X_1)\end{aligned}$$

Nous pouvons alors calculer la variance de S_n, en écrivant

$$\mathrm{Var}(S_n) = n\mathbb{E}(X_1^2) + (n^2-n)\mathbb{E}^2(X_1) - \left(n\mathbb{E}(X_1)\right)^2 = n\left(\mathbb{E}(X_1^2) - \mathbb{E}^2(X_1)\right) = n\mathrm{Var}(X_1)$$

Proposition 8.15. *Soit $S_n = X_1+\cdots+X_n$, où $X_1, X_2, \ldots$ forment une suite de variables i.i.d. dont l'espérance commune $\mathbb{E}(X_1)$ est bien définie ; alors $\mathbb{E}(S_n) = n\mathbb{E}(X_1)$ et $\mathrm{Var}(S_n) = n\mathrm{Var}(X_1)$.*

Remarque 8.16. *En vue de « l'argument probabiliste de Bernstein » permettant d'établir « le théorème d'approximation de Weierstrass » (c.f. § 4.2), rappelons que la variable aléatoire X est une variable de Bernoulli de paramètre α (avec $0 \le \alpha \le 1$) si sa loi est $\mu = (1-\alpha)\delta_0 + \alpha\delta_1$ (où nous notons δ_a la mesure de Dirac au point a). Alors X possède des moments de tous ordres, i.e. $\mathbb{E}(X^0) = 1$ et pour $k = 1, 2, \ldots$*

$$\mathbb{E}(X^k) = \int x^k\mu(dx) = (1-\alpha)\int x^k\delta_0(dx) + \alpha\int x^k\delta_1(dx) = \alpha$$

En particulier la moyenne et la variance de X valent respectivement $\mathbb{E}(X) = \alpha$ et $\mathrm{Var}(X) = \alpha - \alpha^2 = \alpha(1-\alpha)$. Notons aussi au passage que la fonction génératrice des moments Ψ_X est holomorphe dans tout le plan complexe, avec

$$\Psi_X(z) = \int e^{zx}\mu(dx) = (1-\alpha) + \alpha e^z = 1 + \sum_{k=1}^{\infty}\frac{\alpha}{k!}z^k$$

(où nous retrouvons bien que $\Psi_X(z) = \sum_{k=0}^{\infty}\mathbb{E}(X^k)z^k/k!$). Maintenant considérons que $S_n = X_1 + \cdots + X_n$, où $X_1, X_2, \ldots$ est une suite de variables i.i.d., la loi commune des X_k étant une loi de Bernoulli de paramètre α. Alors S_n est une variable aléatoire discrète de support inclus dans $\{0, \ldots, n\}$ d'espérance $\mathbb{E}(S_n) = n\alpha$ et de variance $\mathrm{Var}(S_n) = n\alpha(1-\alpha)$. Par définition S_n est une « loi binomiale de paramètre (n, α) » : on vérifie que

$$0 \le \forall k \le n, \qquad \mathbb{P}\{S_n = k\} = \mathrm{C}_n^k\alpha^k(1-\alpha)^{n-k}$$

8.17. Soit $S_n = X_1 + \cdots + X_n$, où $X_1, X_2, \ldots$ est une suite de variables i.i.d. centrées (i.e. d'espérances nulles). Dans le cas (particulier) où les X_k ont un variance $\mathrm{Var}(X_1)$ finie, une application de l'inégalité de Tchebychev à la variable aléatoire S_n donne

$$\mathbb{P}\left\{\left|\frac{S_n(\omega) - n\mathbb{E}(X_1)}{n}\right| < \varepsilon\right\} = 1 - \mathbb{P}\left\{\left|S_n(\omega) - n\mathbb{E}(X_1)\right| \ge n\varepsilon\right\} \le 1 - \frac{\mathrm{Var}(S_n)}{(n\varepsilon)^2} = 1 - \frac{\mathrm{Var}(X_1)}{n\varepsilon^2}$$

En faisant tendre n vers l'infini, il vient :

$$\lim_{n\to+\infty}\mathbb{P}\left\{\left|\frac{S_n(\omega) - n\mathbb{E}(X_1)}{n}\right| < \varepsilon\right\} = 1$$

Nous venons de démontrer *« la loi faible des grands nombres »* dans un cas particulier ; l'hypothèse portant sur la moyenne des X_k (variables centrées) peut être levée très facilement. Par contre la loi faible des grands nombres reste valable lorsque la variance des X_k est infinie (voir [Ber92, p. 45-48] pour la démonstration générale). La *« loi forte des* *grands nombres »* (démontrée en 1929 par Kolmogorov) est d'un autre ordre de difficulté : c'est un théorème très proche du *« théorème ergodique de Birkhoff »* (démontré en 1931) et appartenant pleinement à la théorie de la mesure.

Théorème 8.17. *Soit $S_n := X_1 + \cdots + X_n$, où les X_k forment une suite i.i.d. de variables aléatoires réelles définies sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et supposées $\mathbb{P}$-intégrables ; alors :*

(i) : la loi faible des grands nombres affirme que pour tout $\varepsilon > 0$

$$\lim_{n\to+\infty} \mathbb{P}\left\{\left|\frac{S_n(\omega) - n\mathbb{E}(X_1)}{n}\right| < \varepsilon\right\} = 1$$

(ii) : la loi forte des grands nombres affirme que pour $\mathbb{P}$-presque tout $\omega \in \Omega$

$$\lim_{n\to+\infty} \frac{S_n(\omega) - n\mathbb{E}(X_1)}{n} = 0$$

8.18. A la convergence faible d'une suite de probabilités dans $\mathscr{P}(\mathbb{R})$ correspond une notion de convergence pour les suites de variables aléatoires appelée la *« convergence en loi »*.

Définition 8.18. *Soit $X_1, X_2, \ldots$ une suite de variables aléatoires et soit μ_n la lois de X_n, pour tout $n = 1, 2, \ldots$; on dit alors que la suite des X_n converge en loi vers une variable aléatoire X ssi la suite des probabilités $\mu_1, \mu_2, \ldots$ « converge faiblement » vers la lois μ de X ; on notera alors $X_n \xrightarrow{\mathscr{L}} X$, quand $n \to +\infty$.*

Le résultat suivant est un corollaire du Théorème 8.10 : il est aussi appelé *« le théorème de continuité de Levy »*.

Théorème 8.19. *$X_n \xrightarrow{\mathscr{L}} X$ ssi $\Phi_{X_n}(t) \to \Phi_X(t)$, pour tout $t \in \mathbb{R}$.*

Le *« théorème de la limite centrale »* est essentiellement un corollaire du *« théorème de continuité de Levy »*.

Théorème 8.20. *Soit $S_n := X_1 + \cdots + X_n$, où les X_k forment une suite i.i.d. de variables aléatoires réelles définies sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ supposées $\mathbb{P}$-intégrables et de variance $\sigma^2 = \mathrm{Var}(X_1)$; le théorème de la limite centrale affirme que si $\sigma < +\infty$, alors (pour $n \to +\infty$)*

$$\frac{S_n(\omega) - n\mathbb{E}(X_1)}{\sigma\sqrt{n}} \xrightarrow{\mathscr{L}} \mathrm{N}(0,1)$$

où $\mathrm{N}(0,1)$ est une variable aléatoire « normale centrée réduite » ; autrement dit :

$$\forall f \in \mathscr{C}_b(\mathbb{R}), \qquad \lim_{n\to+\infty} \int f\left(\frac{S_n(\omega) - n\mathbb{E}(X_1)}{\sigma\sqrt{n}}\right)\mathbb{P}(d\omega) = \frac{1}{\sqrt{2\pi}} \int f(t)e^{-t^2/2}dt$$

Preuve (esquisse). Il est possible de se ramener au cas particulier où les variables aléatoires X_k sont centrées et de variance $\sigma^2 = 1$. Dans ce cas, si nous notons $Y_n := S_n/\sqrt{n}$, alors le théorème de la limite centrale affirme que $Y_n \xrightarrow{\mathscr{L}} X$ où X est un variable aléatoire normale centrée et réduite ; en d'autres termes, la loi de X est la mesure gaussienne

$$\mu(dx) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}dx$$

Nous partons du fait (c.f. Corollaire 6.5) que

$$\Phi_X(t) = \sqrt{2\pi}\,\mathfrak{F}_{+}[\mu](t) = e^{-t^2/2}$$

de sorte que par le *« théorème de continuité de Levy »* il s'agit de démontrer que $\Phi_{Y_n}(t)$ converge simplement vers $e^{-t^2/2}$. La fonction caractéristique de Y_n, se calcule grâce à l'indépendance des X_k : ainsi, en utilisant le Corollaire 8.14 nous pouvons écrire que

$$\begin{aligned}\Phi_{Y_n}(t) &= \mathbb{E}\big(e^{it(X_1+\cdots+X_n)/\sqrt{n}}\big)\\ &= \mathbb{E}\big(e^{itX_1/\sqrt{n}}\big)\cdots\mathbb{E}\big(e^{itX_n/\sqrt{n}}\big) = \Big(\mathbb{E}\big(e^{itX_1/\sqrt{n}}\big)\Big)^n = \Big(\Phi_{X_1}(t/\sqrt{n})\Big)^n\end{aligned}$$

Mais les deux premiers moments de X_1 étant $\mathbb{E}(X_1) = 0$ et $\mathbb{E}(X_1^2) = 1$, il vient (c.f. équation (35) in Proposition 8.3) :

$$\Phi_{X_1}(t) = 1 + \mathbb{E}(X_1)\frac{it}{1!} + \mathbb{E}(X_1^2)\frac{(it)^2}{2!} + \circ(t^2) = 1 - \frac{t^2}{2} + \circ(t^2)$$

Finalement nous avons (dans la limite où $n \to +\infty$)

$$\Phi_{Y_n}(t) = \Phi_{X_1}^n(t/\sqrt{n}) \approx \left(1 - \frac{(t/\sqrt{n})^2}{2}\right)^n = \left(1 - \frac{t^2/2}{n}\right)^n \to e^{-t^2/2} = \Phi_X(t)$$

□

8.19. Nous avons défini une variable aléatoire (réelle) comme une application borélienne X définie sur *« un espace probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ ad-hoc »*, à valeurs dans $\mathbb{R}$; nous en avons alors déduit la loi de X comme la probabilité $\mu = \mathbb{P} \circ X^{-1}$, c'est-à-dire un élément de $\mathscr{P}(\mathbb{R})$. Cependant lorsqu'on modélise un phénomène aléatoire physique dont le résultat (obtenu par une mesure) est un nombre réel, le premier objet qu'il est possible de considérer (à partir d'un grand nombre d'expériences successives) est *« une version empirique »* de la fonction de répartition F_μ associée à μ. La question qui se pose alors est de savoir comment il est possible de *« reconstruire une version de l'espace $(\Omega, \mathcal{F}, \mathbb{P})$ »* à partir de la fonction de répartition F_μ. Pour répondre à cette question, le point de départ est de considérer *« la fonction quantile »* $Q_\mu : [0\,;1] \to \mathbb{R}$ t.q.

$$Q_\mu(x) = \inf\left\{x \in \mathbb{R}\ ;\ F_\mu(x) > y\right\}$$

Théorème 8.21 (du quantile)**.** *Notons Q_μ la fonction quantile associée à la fonction de répartition F_μ d'une mesure $\mu \in \mathscr{P}(\mathbb{R})$; si $\lambda(dx) = dx$ est la mesure de Lebesgue sur l'intervalle $[0\,;1]$, alors μ est l'image réciproque de λ par Q_μ, en ce sens que $\mu = \lambda \circ Q_\mu^{-1}$.*

Si $X : \Omega \to \mathbb{R}$ est une variable aléatoire réelle définie sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et de loi $\mu := \mathbb{P} \circ X^{-1}$ alors, d'après le théorème du quantile nous pouvons considérer $(\Omega, \mathcal{F}, \mathbb{P}) = ([0\,;1], \mathcal{B}, \lambda)$ (avec $\mathcal{B}$ la sigma-algèbre des boréliens de $[0\,;1]$) et identifier X avec $Q_\mu : [0\,;1] \to \mathbb{R}$. En d'autres termes, la fonction quantile Q_μ – *vue comme variable aléatoire sur $([0\,;1], \mathcal{B}, \lambda)$* – peut être considérée comme une forme canonique de la variable aléatoire X. En pratique, avec un générateur de nombres aléatoires (uniforme) dans $[0\,;1]$ et une version empirique de F_μ, nous pouvons utiliser la fonction quantile pour obtenir un générateur aléatoire suivant la loi de X.

9. Dérivées faibles – Espaces de Sobolev

9.1. Il est difficile de dater l'émergence de *« la notion de fonction »* ; mentionnons la polémique entre d'Alembert, Euler et Daniel Bernoulli autour de la résolution de *« l'équation* *de la corde vibrante »* et dont un aspect porte sur la définition même de la notion de fonction. Pour fixer les idées, disons que lorsqu'en 1872 Weierstrass analyse les propriétés fines d'une fonction continue sur l'intervalle unité et dérivable en aucun point (*« la fonction de Weierstrass »*), il base son travail sur des définitions rigoureuses que nous utilisons toujours de nos jours. Mais à la fin du 19-ème, il apparaît de nouveaux objets mathématiques qui ressemblent à des fonctions mais ne répondent pas clairement aux définitions courantes. En particulier l'ingénieur-physicien-mathématicien Heaviside utilise des techniques de dérivation de fonctions dans des cas où les notions classiques ne s'appliquent pas. D'autre part, en mathématique, la *« théorie de la mesure »* et de *« l'intégrale de Lebesgue »* élargit encore un peu plus les contours de la notion de fonction : ainsi, une fonction peut-être définie – ou même dérivable – presque partout. L'ensemble de ces développements qui répondent à des questions aussi bien pratiques que théoriques, vont se rejoindre dans une théorie des fonctions généralisées : *« la théorie des distributions »* [47].

9.2. *« L'équation de Schrödinger »* (1926) est une équation aux dérivées partielles (EDP) qui vient se rajouter aux deux EDP classiques : *« l'équation des ondes »* et *« l'équation de la chaleur »* (voir e.g. le paragraphe « *Les trois EDP classiques* » dans [KLR98, p. 231]). A la fin des années 20, le travail de Dirac en mécanique quantique montre que les techniques utilisées par Heaviside ne sont pas de simples artifices de calcul (ce qui avait été reproché à Heaviside). C'est au milieu des années 30 que Sobolev introduit une notion de *« fonctionnelles généralisées »* [48] lui permettant d'exprimer un nouveau type de solutions des EDP [49] : *« les solutions faibles »*. Il y a en germe dans ce travail *« la notion de dérivation au sens faible »*. Nous abordons cette question par une formule d'intégration par parties très générale.

Proposition 9.1. *Soient f et g deux fonctions (complexes) définies sur $\mathbb{R}$; si f et g sont absolument continues (c.f. Définition 8.2) de dérivées respectives f' et g' (fonctions définies presque partout et localement intégrables), alors pour tout $-\infty < a \leq b < +\infty$*

$$\int_a^b f(x)g'(x)dx = \Big[f(x)g(x)\Big]_a^b - \int_a^b f'(x)g(x)dx$$

Preuve. (D'après [GW90, Proposition 14.5.6].) Comme $f(x) = f(a) + \int_a^x f'(u)du$,

$$\begin{aligned}
\int_a^b f(x)g'(x)dx &= \int_a^b \Big(f(a) + \int_a^x f'(y)dy\Big) g'(x)dx \\
&= f(a)\Big[g(b) - g(a)\Big] + \int_a^b \Big(\int_a^x f'(y)dy\Big) g'(x)dx \\
\text{(Théorème de Fubini)} \quad &= f(a)\Big[g(b) - g(a)\Big] + \int_a^b f'(y)\Big(\int_b^y g'(x)dx\Big) dy \\
&= f(a)\Big[g(b) - g(a)\Big] + \int_a^b f'(y)\Big[g(b) - g(y)\Big]dy \\
&= f(a)\Big[g(b) - g(a)\Big] + \Big[f(b) - f(a)\Big]g(b) - \int_a^b f'(y)g(y)dy
\end{aligned}$$

47. Nous n'abordons pas ici la théorie des distributions
48. Comme les appelle Laurent Schwartz c.f. [Kan04].
49. En particulier l'équation des ondes.

soit encore, après simplification :

$$\int_a^b f(x)g'(x)dx = -f(a)g(a) + f(b)g(b) - \int_a^b f'(y)g(y)dy$$

□

Soit f une fonction absolument continue sur $\mathbb{R}$ et φ une fonction continument dérivable sur $\mathbb{R}$ et dont le support est inclus dans un intervalle compact $[a\,;b]$: alors φ est aussi une fonction absolument continue et par une application directe de la Proposition 9.1 (sur l'intervalle $[a\,;b]$), nous obtenons que

$$\int f(x)\varphi'(x)dx = -\int f'(x)\varphi(x)dx$$

Cette remarque nous amène à la définition suivante.

Définition 9.2. *Soit $\mathscr{D}(\mathbb{R})$ l'espace des fonctions réelles de $\mathscr{C}^\infty(\mathbb{R})$ qui sont à support compact (i.e. identiquement nulle en dehors d'un intervalle compact) : une fonction f dans $L^1_{\text{loc}}(\mathbb{R})$ est dite « dérivable au sens faible », s'il existe une fonction $g \in L^1_{\text{loc}}(\mathbb{R})$ t.q.*

$$\forall\varphi \in \mathscr{D}(\mathbb{R}), \qquad \int f(x)\varphi'(x)dx = -\int g(x)\varphi(x)dx \tag{44}$$

La fonction g est alors appelée une dérivée de f au sens faible.

Proposition 9.3. *$f \in L^1_{\text{loc}}(\mathbb{R})$ est dérivable au sens faible ssi elle est absolument continue ; dans ce cas, si g ($\in L^1_{\text{loc}}(\mathbb{R})$) est la dérivée faible de f, alors $f'(x) = g(x)$ pour presque tout $x \in \mathbb{R}$.*

La preuve de la Proposition 5.2 est basée sur la formule de l'intégration par parties généralisée donnée dans la Proposition 9.1 combinée aux deux propriétés de l'espace $\mathscr{D}(\mathbb{R})$ énoncée dans la proposition suivante.

Proposition 9.4. *Si f est une fonction de $L^1_{\text{loc}}(\mathbb{R})$, alors :*

$$\begin{aligned}(i) \quad &: \quad \forall\varphi \in \mathscr{D}(\mathbb{R}), \int f(x)\varphi(x)dx = 0 \implies f(x) = 0 \quad p.p.\\ (ii) \quad &: \quad \forall\varphi \in \mathscr{D}(\mathbb{R}), \int f(x)\varphi'(x)dx = 0 \implies \exists C \in \mathbb{R},\ f(x) = C \quad p.p.\end{aligned}$$

Preuve. (i) : Pour tout $-\infty < a < b < +\infty$ et tout $1 \le p \le +\infty$, nous considérons l'espace $L^p[a\,;b]$ comme un sous-espace de $L^1_{\text{loc}}(\mathbb{R})$. Pour tout $f \in L^1[a\,;b]$, l'application A qui a $\varphi \in L^\infty[a\,;b]$ associe $A\varphi := \int f(x)\varphi(x)dx$, est une forme linéaire continue sur $L^\infty[a\,;b]$: si nous supposons que A s'annule sur $\mathscr{D}(\mathbb{R})$, alors il découle de la densité de $\mathscr{D}(\mathbb{R})$ dans $L^\infty[a\,;b]$ que A est identiquement nulle sur $L^\infty[a\,;b]$: cela entraîne (théorème de représentation de Riesz [50]) que $f(x) = 0$ pour presque tout $x \in [a\,;b]$.

(ii) (d'après [Bre87, p. 122-123]) : Soit $U \in \mathscr{D}(\mathbb{R})$ t.q. $\int U(x)dx = 1$; alors, pour tout $\psi \in \mathscr{D}(\mathbb{R})$, la fonction $\psi - \left(\int \psi(y)dy\right) U$ est dans $\mathscr{D}(\mathbb{R})$ et d'intégrale nulle ; sa primitive

$$\varphi(x) := \int_{-\infty}^x \left(\psi(t) - \left(\int \psi(y)dy\right) U(t)\right) dt \tag{45}$$

50. Nous avons déjà énoncé le théorème de Riesz-Markov (c.f. § 8.1) ; rappelons ici que le *« le Théorème de représentation de Riesz L^p pour $1 \le p < +\infty$ »* affirme que si $(X, \mathfrak{A}, \mu)$ est un espace mesuré sigma-fini et q le conjugué de p (i.e. $1 < q \le +\infty$ t.q. $1/p + 1/q = 1$) : alors, toute forme linéaire A bornée sur $L^p(X, \mathfrak{A}, \mu)$ est associée à une unique fonction f de $L^q(X, \mathfrak{A}, \mu)$ t.q. $\|f\|_q = \|A\|_p$ (norme associée) et pour laquelle $Ag = \int f(x)g(x)\mu(dx)$, pour tout $g \in L^p(X, \mathfrak{A}, \mu)$.

appartient donc elle aussi à $\mathscr{D}(\mathbb{R})$. Si f est une fonction de $L^1_{\text{loc}}(\mathbb{R})$ t.q. $\int f(x)\varphi'(x)dx = 0$, pour tout $\varphi(x) \in \mathscr{D}(\mathbb{R})$, alors dans le cas particulier de la fonction φ en (45) il vient :

$$\begin{aligned} 0 &= \int f(x)\varphi'(x)dx \\ &= \int f(x)\Big(\psi(x) - \Big(\int \psi(y)dy\Big)U(x)\Big)dx \\ &= \int f(x)\psi(x)dx - \iint \psi(y)f(x)U(x)dxdy \\ &= \int f(x)\psi(x)dx - \iint \psi(x)f(y)U(y)dxdy = \int \Big(f(x) - \int f(y)U(y)dy\Big)\psi(x)dx \end{aligned}$$

La fonction ψ étant arbitraire dans $\mathscr{D}(\mathbb{R})$, nous pouvons déduire de la partie (i) que

$$f(x) = \int f(y)U(y)dy \quad \text{p.p.}$$

ce qui prouve (ii).

□

Preuve de la Proposition 5.2. Comme $\mathscr{D}(\mathbb{R})$ est un sous-espace de l'espace des fonctions absolument continues sur $\mathbb{R}$, il découle immédiatement de la Proposition 9.1 que si f est absolument continue de dérivée f' (définie pour presque tout x et dans $L^1_{\text{loc}}(\mathbb{R})$), alors f' est nécessairement une dérivée faible de f. Réciproquement, supposons que $g \in L^1_{\text{loc}}(\mathbb{R})$ soit une dérivée faible de f, en ce sens que pour tout $\varphi \in \mathscr{D}(\mathbb{R})$:

(46) $$\int f(x)\varphi'(x)dx = -\int g(x)\varphi(x)dx$$

La primitive de g, soit $x \mapsto G(x) = f(0) + \int_0^x g(u)du$ est absolument continue et sa dérivée coincide avec g presque partout. Fixons $\varphi \in \mathscr{D}(\mathbb{R})$ arbitrairement donnée : comme G et φ sont des fonctions localement intégrables et absolument continues, nous pouvons appliquer de nouveau la formule de l'intégration par parties généralisée sur un intervalle $[a\,;b]$ contenant le support de φ, de sorte que par (46) il vient

$$\int G(x)\varphi'(x)dx = -\int g(x)\varphi(x)dx = \int f(x)\varphi'(x)dx$$

soit encore :

(47) $$\int (f(x) - G(x))\varphi'(x)dx = 0$$

La fonction $\varphi \in \mathscr{D}(\mathbb{R})$ étant arbitraire, nous déduisons de la partie (ii) de la Proposition 9.4 combinée à (47) que $f(x) - G(x) = f(0) - G(0) = 0$ pour presque tout $x \in \mathbb{R}$.

□

9.3. Notons (par anticipation) $\mathscr{H}^1(\mathbb{R})$ l'espace des fonctions (complexes) absolument continues $f \in L^2(\mathbb{R})$ et dont la dérivée faible f' est aussi dans $L^2(\mathbb{R})$.

Proposition 9.5. *Si f et g sont deux fonctions de $\mathscr{H}^1(\mathbb{R})$, alors fg' et $f'g$ sont dans $L^1(\mathbb{R})$ et*

(48) $$\int f(x)g'(x)dx = -\int f'(x)g(x)dx$$

Preuve. Si f et g sont dans $\mathscr{H}^1(\mathbb{R})$ alors fg est une fonction de $L^1(\mathbb{R})$ qui possède une version dans $\mathscr{C}(\mathbb{R})$. En identifiant fg à sa version continue, il est licite de considérer deux suites de réels $a_1, a_2, \ldots$ et $b_1, b_2, \ldots$ avec $a_n \to -\infty$ et $b_n \to +\infty$, t.q. $f(a_n)g(a_n)$ et

$f(b_n)g(b_n)$ tendent vers 0 quand $n \to +\infty$: d'après la formule d'intégration par parties (c.f. Proposition 9.1) nous avons pour tout $n \geq 1$

$$\int_{a_n}^{b_n} f(x)g'(x)dx = \Big[f(x)g(x)\Big]_{a_n}^{b_n} - \int_{a_n}^{b_n} f'(x)g(x)dx \tag{49}$$

Mais d'autre part, nous savons aussi que fg' et $f'g$ sont dans $L^1(\mathbb{R})$: les suites de fonctions $\mathbf{1}_{[a_n\,;b_n]}fg'$ et $\mathbf{1}_{[a_n\,;b_n]}f'g$ sont donc respectivement dominées par les fonctions intégrables $|fg'|$ et $|f'g|$. Partant de (49), nous pouvons appliquer (deux fois) le théorème de la convergence dominée, de sorte que par définition des suites $a_1, a_2, \ldots$ et $b_1, b_2, \ldots$, nous obtenons (48) en faisant tendre $n \to +\infty$.

□

D'après la Proposition 9.5, si $f, g \in \mathscr{H}^1(\mathbb{R})$, alors $\langle f|g'\rangle + \langle f'|g\rangle = 0$ et donc :

$$\left\langle f + f'|g + g'\right\rangle = \langle f|g\rangle + \left\langle f'|g'\right\rangle := \langle\!\langle f|g\rangle\!\rangle_{\mathscr{H}^1}$$

Il est facile de vérifier que $\langle\!\langle\cdot|\cdot\rangle\!\rangle_{1,2}$ définit un produit scalaire (hermiten) sur $\mathscr{H}^1(\mathbb{R})$: dans la suite, nous noterons $\|\cdot\|_{1,2}$ la norme associée à ce produit produit scalaire.

Proposition 9.6. *Le $\mathbb{C}$-espace vectoriel $\mathscr{H}^1(\mathbb{R})$ formé des fonctions complexes absolument continues sur $\mathbb{R}$ qui sont de carré sommable et dont la dérivée faible f' est aussi de carré sommable, est un espace de Hilbert pour le produit scalaire $\langle\!\langle\cdot|\cdot\rangle\!\rangle_{\mathscr{H}^1}$.*

Preuve. Muni du produit scalaire $\langle\!\langle\cdot|\cdot\rangle\!\rangle_{\mathscr{H}^1}$, l'espace $\mathscr{H}^1(\mathbb{R})$ est un espace pré-hilbertien (complexe) : nous devons montrer que c'est aussi un espace de Banach. Soit $f_1, f_2, \ldots$ une suite de fonctions de $\mathscr{H}^1(\mathbb{R})$ qui est de Cauchy pour la norme $\|\cdot\|_{\mathscr{H}^1}$. Si nous notons f'_n la dérivée faible de f_n, alors $\|f_n\|^2_{\mathscr{H}^1} = \|f_n\|^2_2 + \|f'_n\|^2_2$; il est donc immédiat que les fonctions f_n et f'_n forment des suites de fonctions de $L^2(\mathbb{R})$ qui sont de Cauchy pour la norme $\|\cdot\|_2$: nous notons f et g les fonctions de $L^2(\mathbb{R})$ qui sont les limites (dans $L^2(\mathbb{R})$) de ces deux suites de fonctions. Nous aurons prouvé que $\mathscr{H}^1(\mathbb{R})$ est un espace de Hilbert si nous pouvons vérifier que g est la dérivée faible de f. Comme f_n et f'_n convergent dans $L^2(\mathbb{R})$ vers f et g respectivement, d'après *« le théorème de Riesz-Fisher »* (et plus précisément la Proposition 2.10), il existe une suite croissante d'indices $n_1, n_2, \ldots$ t.q. $f_{n_k}(x)$ et $f'_{n_k}(x)$ convergent respectivement vers $f(x)$ et $g(x)$, pour presque tout $x \in \mathbb{R}$: par suite, pour toute fonction φ dans $\mathscr{D}(\mathbb{R})$ et pour tout $k \geq 1$

$$\int f_{n_k}(x)\varphi'(x)dx = -\int f'_{n_k}(x)\varphi(x)dx$$

de sorte que (*« théorème de la convergence dominée »*) dans la limite où $k \to +\infty$:

$$\int f(x)\varphi'(x)dx = -\int g(x)\varphi(x)dx$$

□

L'espace $\mathscr{H}^1(\mathbb{R})$ appartient à une famille d'espaces de Hilbert introduits par Sobolev.

Définition 9.7. *Pour $p = 1, 2, \ldots,$ l'espace de Sobolev $\mathscr{H}^p(\mathbb{R})$ est le sous-espace de $L^2(\mathbb{R}) =: \mathscr{H}^0(\mathbb{R})$ formé des fonctions f qui sont p fois dérivables au sens faible et t.q. les dérivées $\partial^k f$ (dérivées k-ème) soient dans $L^2(\mathbb{R})$, pout tout $k = 1, \ldots, p$.*

Le théorème suivant s'obtient par induction :

Proposition 9.8. *La suite des espaces de Sobolev $\mathscr{H}^p(\mathbb{R})$ ($p \geq 0$ entier) est décroissante, avec*

$$L^2(\mathbb{R}) = \mathscr{H}^0(\mathbb{R}) \supset \mathscr{H}^1(\mathbb{R}) \supset \mathscr{H}^2(\mathbb{R}) \supset \cdots$$

De plus, pour $p \geq 0$ nous avons $\mathscr{H}^{p+1}(\mathbb{R}) \subset L^2(\mathbb{R}) \cap \mathscr{C}^p(\mathbb{R})$ et plus précisément :

$$f \in \mathscr{H}^{p+1}(\mathbb{R}) \implies \Big(0 \leq \forall k < p,\ \partial^k f \in L^2(\mathbb{R}) \cap \mathscr{C}^{p-k}(\mathbb{R}) \quad et \quad \partial^p f \in \mathscr{H}^1(\mathbb{R})\Big)$$

La généralisation de la Proposition 9.6 s'énonce comme suit :

Proposition 9.9. *Pour tout entier $p \geq 1$, le $\mathbb{C}$-espace vectoriel $\mathscr{H}^p(\mathbb{R})$ est un espace de Hilbert pour le produit scalaire*

$$(f, g) \mapsto \langle\!\langle f | g \rangle\!\rangle_{\mathscr{H}^p} := \sum_{k=0}^{p} \langle \partial^k f | \partial^k g \rangle$$

Si $\|\cdot\|_{1,p}$ dénote la norme associée ce produit scalaire, il vient pour tout $f \in \mathscr{H}^p(\mathbb{R})$

$$\|f\|^2_{\mathscr{H}^p} = \sum_{k=0}^{p} \|\partial^k f\|_2^2$$

Le théorème suivant donne une caractérisation des espaces de Sobolev $\mathscr{H}^p(\mathbb{R})$ grâce à la transformée de Fourier L^2 (ce résultat nous sera très utile dans la démonstration du principe d'incertitude de Heisenberg-Gabor sur $\mathscr{H}^2(\mathbb{R})$: c.f. § 10).

Théorème 9.10. *Soit $p \geq 1$ un entier ; alors les trois propositions suivantes sont équivalentes*
 (i) *: f appartient à $\mathscr{H}^p(\mathbb{R})$;*
 (ii) *: pour tout $0 \leq k \leq p$ la dérivée k-ème $\partial^k f$ est dans $L^2(\mathbb{R})$ et $\mathfrak{F}[\partial^k f] = (\boldsymbol{i}Q)^k \mathfrak{F}[f]$;*
 (iii) *: $Q^k \mathfrak{F}[f]$ est une fonction de $L^2(\mathbb{R})$, pour tout $0 \leq k \leq p$.*

Preuve du Théorème 9.10. $(i) \implies (ii)$: Par définition, si f est dans $\mathscr{H}^p(\mathbb{R})$, alors $f \in \mathscr{C}^k(\mathbb{R})$ pour tout $0 \leq k < p$ avec $\partial^k f \in L^2(\mathbb{R})$; de plus la fonction continue $\partial^{p-1} f$ est dans $\mathscr{H}^1(\mathbb{R})$ et sa dérivée au sens faible $\partial(\partial^{p-1} f) = \partial^p f$ est aussi dans $L^2(\mathbb{R})$, ce qui démontre la première partie de (ii). Pour la deuxième partie de (ii), nous considérons pour tout entier $p \geq 1$ la proposition :

$$(\mathcal{R}_p) \ : \quad f \in \mathscr{H}^p(\mathbb{R}) \implies \Big(0 \leq \forall k \leq p, \quad Q^k \mathfrak{F}[f] \in L^2(\mathbb{R}) \quad \text{et} \quad \mathfrak{F}[\partial^k f] = (\boldsymbol{i}Q)^k \mathfrak{F}[f]\Big)$$

Noue allons démontrer par récurrence sur p que $(\mathcal{R}_p)$ est vraie pour tout $p \geq 1$. Nous commençons par démontrer le lemme suivant que est l'amorce de cette récurrence.

Lemme 9.11. *Si f est une fonction de $\mathscr{H}^1(\mathbb{R})$, alors $\partial f \in L^2(\mathbb{R})$ et $\mathfrak{F}[\partial f] = \boldsymbol{i}Q\mathfrak{F}[f]$.*

Preuve. Pour $f \in \mathscr{H}^1(\mathbb{R})$ (identifiée à sa version continue sur $\mathbb{R}$). Nous utilisons l'existence de deux suites $a_1, a_2, \ldots$ et $b_1, b_2, \ldots$ tendant respectivement vers $-\infty$ et $+\infty$ et telles que $f(a_n)$ et $f(b_n)$ tendent vers 0 quand $n \to +\infty$; en appliquant l'intégration par parties généralisée (c.f. Proposition 9.1), il vient pour tout $n \geq 1$ et $y \in \mathbb{R} \setminus \{0\}$:

$$\frac{1}{\sqrt{2\pi}} \int_{a_n}^{b_n} f(x) e^{-\boldsymbol{i}xy} dx = -\frac{1}{\boldsymbol{i}y\sqrt{2\pi}} \Big[f(x) e^{-\boldsymbol{i}xy} \Big]_{x=a_n}^{b_n} + \frac{1}{\boldsymbol{i}y} \left(\frac{1}{\sqrt{2\pi}} \int_{a_n}^{b_n} \partial f(x) e^{-\boldsymbol{i}xy} dx \right)$$

soit encore, en notant $f_n := \mathbf{1}_{[a_n\,;b_n]} f$ et $\partial f_n := \mathbf{1}_{[a_n\,;b_n]} \partial f$

$$\mathfrak{F}[\partial f_n](y) = \boldsymbol{i}y\mathfrak{F}[f_n](y) + \frac{1}{\sqrt{2\pi}} \Big[f(x) e^{-\boldsymbol{i}xy} \Big]_{x=a_n}^{b_n}$$

(identité qu'il est licite de considérer valable dans $L^2(\mathbb{R})$, pour tout $n = 1, 2, \ldots$). Mais dans la limite où $n \to +\infty$, nous avons [51] $f_n \to f$ et $\partial f_n \to \partial f$ dans $L^2(\mathbb{R})$: la transformée de Fourier L^2 étant un endomorphisme continu de $L^2(\mathbb{R})$, nous pouvons conclure – en utilisant l'opérateur de position – que $\mathfrak{F}_-[\partial f] = \boldsymbol{i}Q\mathfrak{F}_-[f]$: ceci conclut la preuve du lemme. $\square$

Pour compléter la récurrence portant sur les propositions $(\mathcal{R}_p)$, le Lemme 9.11 nous autorise à considérer $p \geq 1$ t.q. $(\mathcal{R}_p)$ soit vraie. Alors d'après l'inclusion $\mathscr{H}^p(\mathbb{R}) \supset \mathscr{H}^{p+1}(\mathbb{R})$, la proposition $(\mathcal{R}_p)$ sera récurrente si nous pouvons vérifier que $\mathfrak{F}_-[\partial^{p+1} f] = (\boldsymbol{i}Q)^{p+1}\mathfrak{F}_-[f]$ dès que $f \in \mathscr{H}^{p+1}(\mathbb{R})$: or, pour une telle fonction f, nous savons (par définition) que $\partial^p f$ appartient à $\mathscr{H}^1(\mathbb{R})$ et que la dérivée faible $\partial(\partial^p f)$ de $\partial^p f$ coïncide avec $\partial^{p+1} f$; grâce au Lemme 9.11 appliqué à $\partial^p f$ et par définition de p (i.e. « *l'hypothèse de récurrence* »), nous pouvons alors écrire que

$$\mathfrak{F}_-[\partial^{p+1} f] = \mathfrak{F}_-[\partial(\partial^p f)] = \boldsymbol{i}Q\mathfrak{F}_-[\partial^p f] = \boldsymbol{i}Q(\boldsymbol{i}Q)^p\mathfrak{F}_-[f] = (\boldsymbol{i}Q)^{p+1}\mathfrak{F}_-[f]$$

$(ii) \implies (iii)$: Comme la transformée de Fourier L^2 est un endomorphisme de $L^2(\mathbb{R})$, le fait que $\partial^k f$ soit une fonction de $L^2(\mathbb{R})$ combiné à l'identité $\mathfrak{F}_-[\partial^k f] = (\boldsymbol{i}Q)^k\mathfrak{F}_-[f]$ assure que $Q^k\mathfrak{F}_-[f]$ est aussi une fonction de $L^2(\mathbb{R})$.

$(i) \impliedby (iii)$: Nous allons effectuer une récurrence sur l'entier $p \geq 1$: pour cela nous considérons la proposition :

$$(\mathcal{R}'_p) \ : \quad \Big(0 \leq \forall k \leq p, \quad Q^k\mathfrak{F}_-[f] \in L^2(\mathbb{R})\Big) \implies f \in \mathscr{H}^p(\mathbb{R})$$

Le lemme suivant que est l'amorce de la récurrence.

Lemme 9.12. *Si f est une fonction de $L^2(\mathbb{R})$ telle que $Q\mathfrak{F}_-[f]$ – i.e. $y\mathfrak{F}_-[f](y)$ – soit aussi dans $L^2(\mathbb{R})$, alors f appartient à $\mathscr{H}^1(\mathbb{R})$ (en d'autres termes la proposition $(\mathcal{R}'_1)$ est vraie).*

Preuve. Soit $f \in L^2(\mathbb{R})$ et notons $\hat{f} := \mathfrak{F}_-[f]$: nous supposons que $Q\hat{f}$ est une fonction de $L^2(\mathbb{R})$. Par définition, la fonction [52]

$$g := \boldsymbol{i}\mathfrak{F}_+\left[Q\hat{f}\right] \tag{50}$$

appartient $L^2(\mathbb{R})$: elle est localement intégrable et en posant

$$G(x) := \int_0^x g(y)dy \tag{51}$$

nous définissons une fonction absolument continue dont la dérivée faible est g. Attention, G est à priori localement intégrable mais pas nécessairement dans $\mathscr{H}^1(\mathbb{R})$; cependant, comme $f \in L^2(\mathbb{R})$, nous aurons prouvé que $f \in \mathscr{H}^1(\mathbb{R})$ si nous parvenons à démontrer

$$\exists C \in \mathbb{R}, \quad f(x) = G(x) + C \quad \text{(pour presque tout } x \in \mathbb{R}) \tag{52}$$

Pour $\varphi \in \mathscr{D}(\mathbb{R})$ arbitrairement donné, le théorème de Plancherel pour la transformée de Fourier L^2, nous donne $\langle f|\varphi'\rangle = \big\langle \mathfrak{F}_-[f]\big|\mathfrak{F}_-[\varphi']\big\rangle = \big\langle \hat{f}\big|\mathfrak{F}_-[\varphi']\big\rangle$: mais comme $\varphi' \in \mathscr{D}(\mathbb{R})$,

51. Pour $g \in L^2(\mathbb{R})$ et $g_n := \mathbf{1}_{[a_n;b_n]}g$ les $|g_n - g|^2$ forment une suite de fonctions intégrables dominée par la fonction intégrable $|g|^2$: comme $|g_n(x) - g(x)|^2 \to 0$ pour presque tout x le théorème de la convergence dominée assure que $\|g_n - g\|_2^2 \to 0$ quand $n \to 0$.

52. La définition de g proposée dans (50) vient par analogie avec la transformée de Fourier L^1 : en effet, nous savons d'après la Proposition 6.3, que si h est une fonction de $L^1(\mathbb{R}) \cap \mathscr{C}^1(\mathbb{R})$ avec $h' \in L^1(\mathbb{R})$ et de transformée de Fourier $\hat{h}$ alors $\partial\mathfrak{F}_-[h'] = \boldsymbol{i}Q\hat{h}$; si de plus $Q\hat{h}$ est aussi dans $L^1(\mathbb{R})$, alors nous pouvons appliquer la formule d'inversion, ce qui nous donne $h' = \boldsymbol{i}\mathfrak{F}_+[Q\hat{h}]$.

nous savons (c.f. Proposition 6.3[53]) que $\mathfrak{F}_-[\varphi'] = iQ\mathfrak{F}_-[\varphi]$ et donc $\langle f|\varphi'\rangle = \langle \hat{f}|iQ\mathfrak{F}_-[\varphi]\rangle$. Or par hypothèse, $Q\hat{f}$ est dans $L^2(\mathbb{R})$, ce qui nous permet d'écrire que $\langle \hat{f}|iQ\mathfrak{F}_-[\varphi]\rangle = -\langle iQ\hat{f}|\mathfrak{F}_-[\varphi]\rangle$; nous pouvons maintenant transposer la transformée de Fourier et utiliser la définition de g donnée en (50) pour obtenir finalement que

$$\langle f|\varphi'\rangle = -\langle i\mathfrak{F}_+[Q\hat{f}]\,|\varphi\rangle = -\langle g|\varphi\rangle \tag{53}$$

Les fonctions G et φ étant toutes deux localement intégrables et absolument continues (de dérivées respectives g et φ'), nous pouvons appliquer la formule d'intégration par parties généralisée (c.f. Proposition 9.1), de sorte que pour $a < b$ tels que l'intervalle $[a\,;b]$ contienne le support de φ :

$$\int G(x)\varphi'(x)dx = \int_a^b G(x)\varphi'(x)dx = \Big[G(x)\varphi(x)\Big]_a^b - \int_a^b g(x)\varphi(x)dx = -\int g(x)\varphi(x)dx$$

Du fait que $\int g(x)\varphi(x)dx = \langle g|\varphi\rangle$, nous pouvons combiner avec (53), ce qui donne :

$$\int G(x)\varphi'(x)dx = -\langle g|\varphi\rangle = \langle f|\varphi'\rangle = \int f(x)\varphi'(x)dx$$

soit encore

$$\int \Big(f(x) - G(x)\Big)\varphi'(x)dx = 0$$

Les fonctions f et G sont dans $L^1_{\text{loc}}(\mathbb{R})$ et φ est une fonction arbitraire de $\mathscr{D}(\mathbb{R})$: la partie (ii) de la Proposition 9.4 entraîne donc la validité de (52) : le lemme est démontré.

□

Pour compléter la récurrence portant sur les propositions $(\mathcal{R}'_p)$, le Lemme 9.12 nous autorise à considérer $p \geq 1$ t.q. $(\mathcal{R}'_p)$ soit vraie. Soit alors f une fonction – nécessairement dans $L^2(\mathbb{R})$ – telle que $Q^k\mathfrak{F}_-[f]$ appartienne à $L^2(\mathbb{R})$, pour tout $0 \leq k \leq p+1$: alors, d'après $(\mathcal{R}'_p)$ il est immédiat que f est dans $\mathscr{H}^p(\mathbb{R})$. Afin de montrer que f appartient à $\mathscr{H}^{p+1}(\mathbb{R})$, il nous suffit donc de montrer que la fonction continue $\partial^p f$ est dans $\mathscr{H}^1(\mathbb{R})$, soit encore d'après le Lemme 9.12 – i.e. d'après $(\mathcal{R}'_1)$ – que $Q\mathfrak{F}_-[\partial^p f]$ est dans $L^2(\mathbb{R})$. Or la fonction f étant dans $\mathscr{H}^p(\mathbb{R})$, l'implication $(i) \Longrightarrow (ii)$ assure que $\mathfrak{F}_-[\partial^p f] = (iQ)^p\mathfrak{F}_-[f]$, de sorte que par l'hypothèse faite sur f, nous avons que

$$L^2(\mathbb{R}) \ni (iQ^{p+1})\mathfrak{F}_-[f] = iQ\big((iQ)^p\mathfrak{F}_-[f]\big) = iQ\mathfrak{F}_-[\partial^p f]$$

Ceci achève la preuve de l'implication $(i) \Longleftarrow (iii)$.

□

Remarque 9.13. *Les coefficients binômiaux C_p^k étant tous compris entre 1 et 2^p, pour tout entier $p \geq 1$ et tout $y \in \mathbb{R}$, nous avons la double inégalité*

$$0 \leq \forall k \leq p, \qquad y^{2k} \leq (1+y^2)^p \leq 2^p(1+y^2+\cdots+y^{2p}) \tag{54}$$

D'après le Théorème 9.10, $f \in \mathscr{H}^p(\mathbb{R})$ ssi pour tout $0 \leq k \leq p$ la fonction $y^k\hat{f}(y)$ est dans $L^2(\mathbb{R})$. D'après la double inégalité dans (54), ceci équivaut au fait que $\sqrt{(1+y^2)^p}\,\hat{f}(y)$ soit dans $L^2(\mathbb{R})$.

53. La transformée de Fourier L^1 et L^2 coïncident sur $\mathscr{S}(\mathbb{R})$; ici, φ étant supposé dans $\mathscr{D}(\mathbb{R})$, c'est aussi un élément de $\mathscr{S}(\mathbb{R})$. En fait on peut montrer que pour $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ les transformées de Fourier L^1 et L^2 sont identifiables à des fonctions égales presque partout (c.f. [GW90, Proposition 22.1.6]).

10. Le principe d'incertitude de Heisenberg-Gabor

10.1. L'espace de Sobolev $\mathscr{H}^2(\mathbb{R})$ joue un rôle important dans la mécanique quantique unidimensionnelle (et par extension dans d'autres cas) ; cela peut se comprendre, en particulier à cause de la présence du laplacien dans l'équation de Schrödinger. Dans ce paragraphe nous nous plaçons dans $\mathscr{H}^2(\mathbb{R})$ pour démontrer une version du *« principe d'incertitude de la théorie du signal sur $L^2(\mathbb{R})$ »* aussi appelé *« principe d'incertitude de Heisenberg-Gabor »*. Considérons pour cela que f soit une fonction de la sphère unité de $L^2(\mathbb{R})$, de sorte que (c.f. théorème de Plancherel) sa transformée de Fourier $\hat{f} = \mathfrak{F}[f]$ est aussi une fonction de la sphère unité de $L^2(\mathbb{R})$; en d'autres termes, les mesures positives finies $\mu_f(dx) = |f(x)|^2dx$ et $\mu_{\hat{f}}(dy) = |\hat{f}(y)|^2dy$ sont toutes les deux des probabilités. Rappelons alors (c.f. § 8.4) que les *« k-ème moments »* de μ_f et $\mu_{\hat{f}}$ (s'ils existent) sont respectivement

$$\boldsymbol{m}_k(\mu_f) = \int x^k|f(x)|^2dx \quad \text{et} \quad \boldsymbol{m}_k(\mu_{\hat{f}}) = \int y^k|\hat{f}(y)|^2dy$$

Par définition les moments d'ordre 0 de μ_f et $\mu_{\hat{f}}$ valent tous deux 1 ; les moments $\boldsymbol{m}_1(\mu_f)$ et $\boldsymbol{m}_1(\mu_{\hat{f}})$ (lorsqu'ils existent) représentent les *« positions moyennes »* respectives de μ_f et $\mu_{\hat{f}}$: cela explique en particulier la dénomination donnée à l'opérateur position Q défini au § 6.2, du fait que (sous réserve que f soit dans le bon sous-espace de $L^2(\mathbb{R})$) :

$$\boldsymbol{m}_1(\mu_f) = \int x|f(x)|^2dx = \int xf(x)\overline{f(x)}dx = \langle Qf|f\rangle = \langle f|Qf\rangle =: \langle f|Q|f\rangle$$

(et de même pour $\boldsymbol{m}_1(\mu_{\hat{f}})$). D'autre part, lorsque les moments d'ordre 1 existent, les *« écartd types »* de μ_f et $\mu_{\hat{f}}$ sont les quantités (finies ou infinies) suivantes :

$$\sigma(\mu_f) := \sqrt{\boldsymbol{m}_2(\mu_f) - \boldsymbol{m}_1(\mu_f)^2} \quad \text{et} \quad \sigma(\mu_{\hat{f}}) := \sqrt{\boldsymbol{m}_2(\mu_{\hat{f}}) - \boldsymbol{m}_1(\mu_{\hat{f}})^2}$$

Théorème 10.1 (Heisenberg-Gabor)**.** *Soit f une fonction de l'espace de Sobolev $\mathscr{H}^2(\mathbb{R})$ t.q. $\int |f(x)|^2dx = 1$ et dont la transformée de Fourier $\hat{f}$ est aussi dans $\mathscr{H}^2(\mathbb{R})$; si $\mu_f(dx) = |f(x)|^2dx$ et $\mu_{\hat{f}}(dx) = |\hat{f}(x)|^2dx$, alors les écarts types $\sigma(\mu)$ et $\sigma(\mu_{\hat{f}})$ sont finis et vérifient :*

$$\sigma(\mu_f)\sigma(\mu_{\hat{f}}) \geq 1/2$$

Nous commençons par démontrer deux lemmes intermédiaires.

Lemme 10.2. *Soit f une fonction de $L^2(\mathbb{R})$ t.q. $\int |f(x)|^2dx = 1$ et dont la transformée de Fourier $\hat{f} = \mathfrak{F}[f]$ appartienne à l'espace de Sobolev $\mathscr{H}^2(\mathbb{R})$; alors les deux premiers moments $\boldsymbol{m}_1(\mu_f)$ et $\boldsymbol{m}_2(\mu_f)$ de la mesure de probabilité $\mu_f(dx) = |f(x)|^2dx$ sont finis.*

Preuve. Le fait que $\hat{f}$ soit dans $\mathscr{H}^2(\mathbb{R})$, assure que $xf(x)$ et $x\overline{f(x)}$ sont des fonctions de $L^2(\mathbb{R})$ (c.f. Théorème 9.10) : cela entraîne que $x|f(x)|^2 = (x\overline{f(x)})f(x)$ et $x^2|f(x)|^2 = (x\overline{f(x)})(xf(x))$ sont toutes deux des fonctions de $L^1(\mathbb{R})$.

□

Lemme 10.3 (Recentrage)**.** *Soient $\mu_f(dx) = |f(x)|^2dx$ et $\mu_{\hat{f}}(dy) = |\hat{f}(y)|^2dy$, où f une fonction de l'espace de Sobolev $\mathscr{H}^2(\mathbb{R})$ t.q. $\int |f(x)|^2dx = 1$ et dont la transformée de Fourier $\hat{f}$ est aussi dans $\mathscr{H}^2(\mathbb{R})$. Si nous notons $\mu_g(dx) = |g(x)|^2dx$ et $\mu_{\hat{g}}(dy) = |\hat{g}(y)|^2dy$ où*

$$g(x) = f\big(x + \boldsymbol{m}_1(\mu_g)\big)e^{-i\boldsymbol{m}_1(\mu_{\hat{g}})x}$$

(avec $\hat{g}(y) = \mathfrak{F}[g](y)$) alors les deux probabilités μ_g et $\mu_{\hat{g}}$ sont centrées[54] *et de plus :*

$$\sigma(\mu_g) = \sqrt{\boldsymbol{m}_2(\mu_g)} = \sigma(\mu_f) \quad et \quad \sigma(\mu_{\hat{g}}) = \sqrt{\boldsymbol{m}_2(\mu_{\hat{g}})} = \sigma(\mu_{\hat{f}})$$

Preuve. Le Lemme 10.2 assure que les deux premiers moments des probabilités μ_f et μ_g existent : par le calcul direct il vient pour $k = 1, 2$ (en posant $u = x + \boldsymbol{m}_1(\mu_f)$) :

$$\boldsymbol{m}_k(\mu_g) = \int x^k \left| f(x + \boldsymbol{m}_1(\mu_f)) e^{-i\boldsymbol{m}_1(\mu_{\hat{f}})x} \right|^2 dx = \int \left(u - \boldsymbol{m}_1(\mu_f)\right)^k |f(u)|^2 du$$

Par suite, $\boldsymbol{m}_1(\mu_g) = 0$ pour $k = 1$ et pour $k = 2$ il vient $\boldsymbol{m}_2(\mu_g) = \sigma(\mu_f)^2$. Par la Proposition 6.2 nous obtenons de même que $\boldsymbol{m}_1(\mu_{\hat{g}}) = 0$ avec $\boldsymbol{m}_2(\mu_{\hat{g}}) = \sigma(\mu_{\hat{f}})^2$ du fait que

$$\hat{g}(y) = \mathfrak{F}\left[f\left(x + \boldsymbol{m}_1(\mu_f)\right)e^{-i\boldsymbol{m}_1(\mu_{\hat{f}})x}\right](y) = \hat{f}\left(y + \boldsymbol{m}_1(\mu_{\hat{f}})\right)e^{-i\boldsymbol{m}_1(\mu_f)p}$$

$\square$

Preuve de Théorème 10.1. Le fait que f et $\hat{f} = \mathfrak{F}[f]$ soient des fonctions de l'espace de Sobolev $\mathscr{H}^2(\mathbb{R})$, assure d'une part (c.f. Lemme 10.2) que les deux premiers moments des probabilités μ_f et $\mu_{\hat{f}}$ existent : en particulier les écarts type $\sigma(\mu_f)$ et $\sigma(\mu_{\hat{f}})$ sont finis. D'autre part, d'après le Théorème 9.10, l'hypothèse $f \in \mathscr{H}^2(\mathbb{R})$ assure aussi que (dans $L^2(\mathbb{R})$) nous avons $\mathfrak{F}[f'](y) = \boldsymbol{i}y\hat{f}(y)$, de sorte que par l'identité de Plancherel

$$\int |f'(x)|^2 dx = \int |\mathfrak{F}[f'](y)|^2 dy = \int y^2 |\hat{f}(y)|^2 dy = \boldsymbol{m}_2(\mu_{\hat{f}}) \tag{55}$$

Enfin (nouvelle application du Théorème 9.10) $xf(x)$ est une fonction de $L^2(\mathbb{R}) \cap \mathscr{C}^1(\mathbb{R})$: mais la fonction $f(x)$ elle même étant aussi dans $L^2(\mathbb{R}) \cap \mathscr{C}^1(\mathbb{R})$, la fonction $x|f(x)|^2$ est dans $L^1(\mathbb{R}) \cap \mathscr{C}^1(\mathbb{R})$: il existe donc deux suites de réels $a_1 a_2, \ldots$ et $b_1, b_2, \ldots$ avec $a_n \to -\infty$ et $b_n \to +\infty$, t.q. $a_n|f(a_n)|^2$ et $b_n|f(b_n)|^2$ tendent vers 0 quand $n \to +\infty$. Comme $f(x)$ est dans $\mathscr{C}^1(\mathbb{R})$, la fonction $x \mapsto |f(x)|^2 = f(x)\overline{f(x)}$ est elle aussi dans $\mathscr{C}^1(\mathbb{R})$ et de plus $(|f(x)|^2)' = 2\mathfrak{re}(f(x)\overline{f(x)'})$. Grâce à une intégration par parties (classique) sur l'intervalle $[a_n ; b_n]$, nous obtenons pour tout $n \geq 1$:

$$\int_{a_n}^{b_n} x\left(|f(x)|^2\right)' dx = \left[x|f(x)|^2\right]_{a_n}^{b_n} - \int_{a_n}^{b_n} |f(x)|^2 dx$$

Dans la limite où $n \to +\infty$, la convergence dominée entraîne que

$$\int x\left(|f(x)|^2\right)' dx = -\int |f(x)|^2 dx = -1$$

Nous pouvons alors écrire :

$$1 = \left| \int x \left(f(x)\overline{f(x)}\right)' dx \right| \leq \int \left| 2x\,\mathfrak{re}\left(f(x)\overline{f'(x)}\right) \right| dx \leq 2 \int \left| xf(x)\overline{f'(x)} \right| dx$$

Or $xf(x)$ et $f'(x)$ étant dans $L^2(\mathbb{R})$, par *« l'inégalité de Cauchy-Schwarz »* il vient :

$$1 \leq 2\left(\int x^2 |f(x)|^2 dx\right)^{1/2} \left(\int |f'(x)|^2 dx\right)^{1/2}$$

c'est-à-dire en utilisant (55) :

$$1 \leq 2\sqrt{\boldsymbol{m}_2(\mu_f)}\sqrt{\boldsymbol{m}_2(\mu_{\hat{f}})} \tag{56}$$

La conclusion découle directement du lemme recentrage (c.f. Lemme 10.3) à partir de l'équation (56) en considérant la fonction $g(x) = f(x + \boldsymbol{m}_1(\mu))e^{-i\boldsymbol{m}_1(\mu_{\hat{f}})x}$.

54. En ce sens que leurs moyennes respectives $\boldsymbol{m}_1(\mu_g)$ et $\boldsymbol{m}_1(\mu_{\hat{g}})$ sont toutes deux nulles.

□

10.2. Le cas d'égalité de l'inégalité de Heisenberg-Gabor – dans $\mathscr{S}(\mathbb{R})$ – peut se déduire simplement de l'argument précédent, du cas d'égalité de l'inégalité de Cauchy-Schwarz.

Corollaire 10.4. *Les fonctions gaussiennes* $\varphi(x) = \varphi(0)\, e^{-ax^2/2}$ *(*$a > 0$*) sont les seuls signaux (normalisés et centrés) de* $\mathscr{S}(\mathbb{R})$ *vérifiant le cas d'égalité de l'inégalité de Heisenberg-Gabor.*

Preuve. Si $\mu_\varphi = |\varphi(x)|^2 dx$ et $\mu_{\hat\varphi} = |\hat\varphi(y)|^2 dy$ sont les probabilités associées à un signal $\varphi(x)$ de $\mathscr{S}(\mathbb{R})$ supposé normalisé et centré, alors par une intégration par partie :

$$\int x\big(2\varphi(x)\varphi'(x)\big)dx = -\int \varphi^2(x)dx = -1$$

de sorte que (c.f. argument du Théorème 10.1)

$$\frac{1}{2} = \Big|\int x\varphi(x)\varphi'(x)dx\Big| \leq \Big(\int x^2\varphi^2(x)dx\Big)^{1/2}\Big(\int \big(\varphi'(x)\big)^2 dx\Big)^{1/2} = \sigma(\mu_\varphi)\,\sigma(\mu_{\hat\varphi}) \tag{57}$$

Nous déduisons alors de la suite des inégalités dans (57) et du cas d'égalité l'inégalité de Cauchy-Schwarz (réel), que le cas d'égalité dans l'inégalité de Heisenberg-Gabor (i.e. si $\sigma(\mu_\varphi)\,\sigma(\mu_{\hat\varphi}) = 1/2$) est équivalent à l'existence d'un nombre α réel t.q. $\varphi'(x) = -\alpha x\varphi(x)$ (l'introduction du signe « – » étant arbitraire) ; mais l'intégration de cette équation différentielle est simples : elle est équivalente à la quadrature $d/dx\big(\varphi(x)e^{\alpha x^2/2}\big) = 0$, ce qui donne finalement $\varphi(x) = \varphi(0)e^{-\alpha x^2/2}$. La condition $\varphi \in \mathscr{S}(\mathbb{R})$ entraîne que $\alpha > 0$.

□

10.3. Nous allons retrouver le cas d'égalité de l'inégalité de Heisenberg-Gabor de manière plus laborieuse, en étudiant directement les *« fonctions gaussiennes à écart type complexe »* (c.f. § 6.3). Pour voir cela, soient a et b sont deux réels quelconques et notons

$$G_{a+ib}(x) := e^{-(a+ib)x^2} \tag{58}$$

Pour vérifier le cas d'égalité de l'inégalité de Heisenberg-Gabor (dans le cas centré), nous considérons le signal gaussien normalisé $\varphi(x) = \varphi(0)\, G_{a+ib}(x)$ avec $a > 0$ et où $\varphi(0) = \varphi(0) > 0$ est la constant de normalisation (i.e. assurant que $\|\varphi\| = 1$), de sorte que

$$\mu_\varphi(dx) = |\varphi(x)|^2 dx = |\varphi(0)|^2\, e^{-2ax^2}\, dx$$

est une probabilité gaussienne centrée (i.e. $\boldsymbol{m}_1(\mu_\varphi) = 0$) et d'écart type

$$\sigma(\mu_\varphi) = \frac{1}{2\sqrt{a}}$$

D'autre part, nous savons (c.f. Lemme 6.4) que

$$\hat\varphi(y) = \mathfrak{F}[\varphi](y) = \frac{\varphi(0)}{\sqrt{2\pi}}\int e^{-(a+ib)x^2} e^{-iyx}dx = \frac{\varphi(0)}{\sqrt{2(a+ib)}}\, e^{-\frac{y^2}{4(a+ib)}}$$

Par suite la probabilité

$$\mu_{\hat\varphi}(dy) = |\hat\varphi(y)|^2 dy = \frac{|\varphi(0)|^2}{2\sqrt{a^2+b^2}}\, e^{-\frac{ay^2}{2(a^2+b^2)}}\, dy$$

est aussi gaussienne centrée et d'écart type

$$\sigma(\mu_{\hat\varphi}) = \sqrt{\frac{a^2+b^2}{a}}$$

Finalement nous obtenons la valeur exact du produit des écarts types de μ_φ et $\mu_{\hat{\varphi}}$, soit :

$$\sigma(\mu_\varphi)\sigma(\mu_{\hat{\varphi}}) = \frac{1}{2}\sqrt{1 + \frac{b^2}{a^2}} \tag{59}$$

Ainsi, pour le signal normalisé $\varphi = \varphi(0)\, e^{-(a+ib)x^2}$ $(a > 0)$, nous avons vérifié directement l'inégalité $\sigma(\mu_\varphi)\sigma(\mu_{\hat{\varphi}}) \geq 1/2$, le cas d'égalité ayant lieu ssi $b = 0$.

RÉFÉRENCES

[Bas71] J. Bass. *Cours de mathématiques : Topologie. Intégration. Distributions. Equations intégrales. Analyse Harmonique.* Cours de mathématiques. Masson et Cie, 1971.

[Ber92] M.A. Berger. *An Introduction to Probability and Stochastic Processes*. Springer Texts in Statistics. Springer New York, 1992.

[Bre87] H. Brezis. *Analyse fonctionnelle : théorie et applications*. Éditions Masson - Paris, 1987.

[D'A47a] Jean Le Rond D'Alembert. *Recherches sur la courbe que forme une corde tendue mise en vibration*. Histoire de l'Académie royale des sciences et des belles lettres de Berlin, T. 1, 1747.

[D'A47b] Jean Le Rond D'Alembert. *Réflexions sur la cause générale des vents*. Paris : David l'aîné, 1747.

[Fou22] J.-B. Fourier. *Théorie analytique de la chaleur*. Firmin Didot, Père et Fils - Paris, 1822.

[Gam01] T.W. Gamelin. *Complex Analysis*. Undergraduate Texts in Mathematics. Springer New York, 2001.

[GW90] C. Gasquet and P. Witomski. *Analyse de Fourier et applications : filtrage, calcul numérique, ondelettes*. Masson, 1990.

[JvN35] P. Jordan and J. von Neumann. On inner products in linear metric spaces. *Ann. of Math.*, 36 no 3 :719–723, 1935.

[Kan04] J.-M. Kantor. Mathématiques d'Est en Ouest – Théorie et pratique : l'exemple des distributions. *Gazette de la SMF*, 100 :33–43, 2004.

[KF94] A.N. Kolmogorov and S. Fomine. *Éléments de la théorie des fonctions et de l'analyse fonctionnelle*. ELLIPSE (Traduit de Éd. Mir en russe), (1994).

[KLR98] J. P. Kahane and P. G. Lemarié-Rieusset. *Séries de Fourier et Ondelettes*. Cassini - Paris, 1998.

[Loc94] G. Lochak. *La Géométrisation de la physique*. Flammarion, Paris, 1994.

[LS89] L. Lusternik and V. Sobolev. *Précis d'analyse fonctionnelle*. Mathématiques (Mir). Mir, 1989.

[Oli16a] E. Olivier. Mécanique quantique II : opérateurs non bornés. *Bull. d'Inf. App. & App.*, 104 :115–153, 2016.

[Oli16b] E. Olivier. Mécanique quantique III : de de Broglie à Schrödinger. *Bull. d'Inf. App. & App.*, 105 :171–209, 2016.

[Oli17] E. Olivier. Mécanique quantique IV : Du principe de Fermat à l'intégrale de chemin de Feynman. *Bull. d'Inf. App. & App.*, 106 :–, 2017.

[RSN55] F. Riesz and B. SZ.-Nagy. *Leçons d'analyse fonctionnelle*. Gauthiers-Villars, 1955.

[Rud64] W. Rudin. *Principles of mathematical analysis*. International series in pure and applied mathematics. McGraw-Hill, 1964.

[ST89] M. Samuelides and L. Touzillier. *Analyse fonctionnelle*. Éditions Cépaduès, 1989.

[ST90] M. Samuelides and L. Touzillier. *Analyse harmonique*. Collection La Chevêche. Cépadues-Éditions, 1990.

[VN55] J. Von Neumann. *Mathematical Foundations of Quantum Mechanics*. Investigations in physics. Princeton University Press, 1955.

[Wie04] N Wiener. *Fourier Transform and Certain of it Applications*. Cambridge Mathematical Library. Dover Publications Inc. New York (First published by The Cambridge Univ. Pr. in 1933), 2004.