Université Paris-Dauphine **2018-2019**
Cours de “Intégration et probabilités" Novembre 2018

Chapitres 6 - Intégration dans $\mathbb{R}^d$: convolution, densité, transformation de Fourier

Table des matières

Dans ce chapitre, on munit $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ de la mesure de Lebesgue, notée $d\lambda$ ou simplement dx. Sont écrites en rouge les parties hors programme, en violet les parties traitées en TD (résulats à connaitre pour sa culture) et en bleu les parties modifiées par rapport au cours en amphi.

1 Convolution dans $M^1_+(\mathbb{R}^d)$ et $L^1(\mathbb{R}^d)$

Pour deux fonctions f et g définies sur $\mathbb{R}^d$, on définit la convolée $f * g$ par

$$(f * g)(x) := \int_{\mathbb{R}^d} f(x-y)g(y)\, dy, \quad x \in \mathbb{R}^d,$$

lorsque cette dernière expression a un sens. De la même manière, pour une fonction f définies sur $\mathbb{R}^d$ et une mesure μ sur $\mathbb{R}^d$, on définit la convolée $f * \mu$ par

$$(f * \mu)(x) := \int_{\mathbb{R}^d} f(x-y) d\mu(y), \quad x \in \mathbb{R}^d,$$

lorsque cette dernière expression a un sens. En observant que $g\lambda$ définit une mesure sur $\mathbb{R}^d$, on voit que cette deuxième définition est une généralisation de la première définition. Nous donnons quelques résultats sur cette opération remarquable et la nouvelle fonction obtenue. Beaucoup d'autres résultats peuvent être obtenus en jouant sur les espaces fonctionnels dans lesquels sont pris f d'une part et g (ou μ) d'autres part.

On définit les espaces suivant :

- $C_b(\mathbb{R}^d)$ l'espace des fonctions continues et bornées, on a donc $C_b(\mathbb{R}^d) \subset L^\infty(\mathbb{R}^d)$, et de la même manière, $C^k_b(\mathbb{R}^d)$, $k \geq 1$, l'espace des fonctions k fois dérivable telles que toutes les dérivées partielles appartiennent à $C_b(\mathbb{R}^d)$;
- $M^1_+(\mathbb{R}^d)$ l'ensemble des mesures μ positives et bornées, au sens où $\mu(\mathbb{R}^d) < \infty$. En particulier, en notant $\mathcal{P}(\mathbb{R}^d)$ l'ensemble des mesures de probabilité sur $\mathbb{R}^d$, on a $\mathcal{P}(\mathbb{R}^d) \subset M^1_+(\mathbb{R}^d)$.

Théorème 1.1 (Convolution $C_b - M^1_+$) *Pour $f \in C_b(\mathbb{R}^d)$ et μ une mesure finie de $\mathbb{R}^d$, on a $f * \mu \in C_b(\mathbb{R}^d)$ et $\|f * \mu\|_\infty \leq \|f\|_\infty \, \mu(\mathbb{R}^d)$.*

Preuve du Théorème 1.1. Pour $x, y \in \mathbb{R}^d$, on note $F_x(y) := f(x-y)$, de sorte que $F_x \in C_b(\mathbb{R}^d)$ pour tout $x \in \mathbb{R}^d$, et on écrit

$$(f * \mu)(x) := \int_{\mathbb{R}^d} F_x(y) d\mu(y), \quad x \in \mathbb{R}^d.$$

Cette expression a bien un sens puisque F_x est borélienne (puisque continue) et donc $F_x \in L^1(d\mu)$ (puisque F_x est bornée). On observe également que
(i) $F_{x'} \to F_x$ ponctuellement si $x' \to x$,
(ii) $|F_x| \leq \|f\|_\infty \in L^1(d\mu)$,
de sorte que $x \mapsto (f * \mu)(x)$ est continue d'après le Théorème de continuité sous le signe somme (Proposition III-2.3 de continuité par rapport au paramètre qui n'est autre que le théorème de convergence dominée). Enfin, on a

$$|(f * \mu)(x)| \leq \int_{\mathbb{R}^d} |F_x(y)| d\mu(y) \leq \|f\|_\infty \mu(\mathbb{R}^d),$$

pour tout $x \in \mathbb{R}^d$. $\square$

Corollaire 1.2 (Convolution $C_b^1 - M_+^1$) *Si de plus $f \in C_b^1(\mathbb{R}^d)$, alors $f * \mu \in C_b^1$ et $\partial_i(f * \mu) = (\partial_i f) * \mu$.*

Preuve du Corollaire 1.2. On se place en dimension $d = 1$ pour simplifier. On observe que sous ces conditions supplémentaires, on a
(iii) $x \mapsto F_x(y)$ est dérivable de dérivée $\partial_x(F_x(y)) = f'(x-y)$ pour tout $y \in \mathbb{R}$,
de sorte que $x \mapsto (f * \mu)(x)$ est de classe C^1 d'après le Théorème de dérivabilité sous le signe somme (Proposition III-2.4 de dérivabilité par rapport au paramètre). $\square$

Corollaire 1.3 (Convolution $C_b - L^1$) *Si maintenant $f \in C_b(\mathbb{R}^d)$ et $g \in L^1(\mathbb{R}^d)$, on a alors $f * g \in C_b(\mathbb{R}^d)$ et $f * g = g * f$.*

Preuve du Corollaire 1.3. On observe que l'on peut définir

$$(f * g)(x) := (f * (g_+ \lambda))(x) - (f * (g_- \lambda))(x),$$

puisque $g_\pm \lambda$ est une mesure (positive) finie. On a donc encore $f * g \in C_b(\mathbb{R}^d)$. En effectuant le changement de variables $z := x - y$, $dz := dy$, on obtient

$$(f * g)(x) = \int_{\mathbb{R}^d} f(x-y) g(y)\, dy = \int_{\mathbb{R}^d} f(z) g(x-z)\, dz = (g * f)(x).$$

Pour être plus précis, on définit $y \mapsto z = \phi_x(y) := x - y$ et on calcule $D\phi_x = -I$ puis la valeur absolue du Jacobien $|\det D\phi_x(y)| = 1$. $\square$

Exemple 1 : Pour $a \in \mathbb{R}^d$, on se rappelle que la mesure de Dirac δ_a est définie par $\delta_a(A) = 1$ si $a \in A$, $\delta_a(A) = 0$ si $a \notin A$. Pour $f \in C_b(\mathbb{R}^d)$ (ou même juste f mesurable), on a

$$(f * \delta_a)(x) = \int_{\mathbb{R}^d} f(x-y)\, \delta_a(dy) = f(x-a), \quad \forall\, x \in \mathbb{R}^d,$$

et en particulier

$$f * \delta_0 = f.$$

On dit que δ_0 est l'identité (ou l'élément neutre) pour le produit de convolution. Justifions la première identité. Pour $f := \mathbf{1}_A$ et pour tout $x \in \mathbb{R}^d$, on a

$$(f * \delta_a)(x) := \int_{\mathbb{R}^d} \mathbf{1}_A(x-y)\, \delta_a(dy) = \int_{\mathbb{R}^d} \mathbf{1}_{x+A}(y)\, \delta_a(dy) = \mathbf{1}_{x+A}(a) = \mathbf{1}_A(x-a),$$

ce qui n'est rien d'autre que l'identité souhaitée dans ce cas particulier. On raisonne par approximations successives (fonctions étagées, fonctions mesurables positives, différences de deux fonctions mesurables positives) afin d'obtenir cette même identité dans le cas général. Il est à retenir de ce calcul que

$$\int_{\mathbb{R}^d} \varphi(y)\, \delta_a(dy) = \varphi(a), \quad \forall\, a \in \mathbb{R}^d,\ \forall\, \varphi \in \mathscr{M}_+(\mathbb{R}^d) \cup \mathcal{L}^\infty(\mathbb{R}^d).$$

Théorème 1.4 (Convolution $L^1 - L^1$) *Soient $f, g \in L^1(\mathbb{R}^d)$. Pour presque tout $x \in \mathbb{R}^d$, la fonction $y \mapsto f(x-y)g(y)$ est intégrable sur $\mathbb{R}^d$, de sorte que $(f*g)(x)$ est presque partout définie. De plus, $f*g = g*f \in L^1(\mathbb{R}^d)$ et*

$$\|f*g\|_{L^1} \le \|f\|_{L^1}\|g\|_{L^1}. \tag{1.1}$$

Preuve du Théorème 1.4. Pour $x, y \in \mathbb{R}^d$, on note $F(x,y) := f(x-y)g(y)$. Pour presque tout $y \in \mathbb{R}^d$, on a

$$\int |F(x,y)|\, dx = \int |f(x-y)| dx |g(y)| = \|f\|_{L^1}\, |g(y)| < \infty$$

et

$$\int \Big[\int |F(x,y)|\, dx\Big] dy = \|f\|_{L^1}\|g\|_{L^1}.$$

Appliquant le théorème V.3.2 de Tonelli on obtient $F \in L^1(\mathbb{R}^d \times \mathbb{R}^d)$. Appliquant ensuite le théorème V.3.3 de Fubini, on a

$$\int |F(x,y)|\, dy < \infty, \quad p.p.\ x \in \mathbb{R}^d$$

et

$$\int \Big| \int F(x,y)\, dy \Big| dx \le \|f\|_{L^1}\|g\|_{L^1},$$

ce qui est exactement la conclusion recherchée. ◻

Il existe beaucoup d'autres résultats sur la convolution, et pour n'en citer que deux, on peut prouver $L^1 * L^p \subset L^p$, $1 \le p \le \infty$, et $L^p * L^{p'} \subset C_0$, $1 < p < \infty$. Ces résultats seront vus en TD ou dans le cours de Topologie et analyse fonctionnelle du 2nd semestre.

Exemple 2 : Dans $\mathbb{R}$, on définit les fonctions gaussiennes (centrées) par

$$G_\lambda(x) := (2\pi\lambda)^{-1/2} \exp(-x^2/(2\lambda)).$$

On appelle gaussienne standard $G := G_1$. On a alors

$$G_s * G_t = G_{s+t}, \quad \forall\, s, t > 0.$$

On commence par observer que

$$e^{-\frac{(x-y)^2}{2s}} e^{\frac{-y^2}{2t}} = \exp\Big\{-\frac{1}{2}\Big[\frac{x^2}{s+t} + \Big(\big(\frac{1}{s}+\frac{1}{t}\big)^{1/2} y - \frac{1}{s}\big(\frac{1}{s}+\frac{1}{t}\big)^{-1/2} x\Big)^2\Big]\Big\},$$

de sorte qu'en effectuant successivement deux changements de variables

$$\begin{aligned} \int_{\mathbb{R}} e^{-\frac{(x-y)^2}{2s}} e^{\frac{-y^2}{2t}}\, dy &= \exp\Big(-\frac{1}{2}\frac{x^2}{s+t}\Big) \int_{\mathbb{R}} \exp\Big(-\frac{1}{2}\big(\frac{1}{s}+\frac{1}{t}\big) y^2\Big) dy \\ &= \exp\Big(-\frac{1}{2}\frac{x^2}{s+t}\Big) \sqrt{2\pi}\big(\frac{1}{s}+\frac{1}{t}\big)^{-1/2}. \end{aligned}$$

On conclut alors aisément.

Nous donnons une dernière définition, pus générale, de la convolution. Lorsque μ et ν sont deux mesures finies sur $\mathbb{R}^d$, on peut définir la convolution $\mu * \nu$ par

$$(\mu * \nu)(A) := \int_{\mathbb{R}^d \times \mathbb{R}^d} \mathbf{1}_{x+y \in A}\, d(\mu \otimes \nu)(x,y),$$

pour tout borélien $A \subset \mathbb{R}^d$.

2 Densité dans $L^1(\mathbb{R}^d)$ et $C(K)$

On accepte le fait suivant de "régularité" de la mesure de Lebesgue : pour tout $A \in \mathscr{B}(\mathbb{R})$ et $\varepsilon > 0$, il existe un compact K et un ouvert $\mathcal{O} \subset [a,b]$ tels que $K \subset A \subset \mathcal{O}$ et $\lambda(\mathcal{O}\backslash K) < \varepsilon$.

Théorème 2.1 (densité des fonctions en escalier dans $\mathcal{L}^1(a,b)$) *Pour toute fonction $f \in \mathcal{L}^1(a,b)$, $-\infty \le a < b \le +\infty$, il existe une suite (f_n) de fonctions en escalier telle que* $\lim \|f - f_n\|_1 = 0$.

Preuve du Théorème 2.1. Etape 1. On suppose $a, b \in \mathbb{R}$. Soit $A \subset [a,b]$. D'après la "régularité" de la mesure de Lebesgue, pour tout $\varepsilon > 0$, il existe un ouvert $\mathcal{O} \subset (a,b)$ tel que

$$\mathcal{O} = \bigcup_{n=1}^{\infty} (a_n, b_n), \quad \lambda(\mathcal{O}\backslash A) < \varepsilon.$$

On pose

$$\mathcal{O}_N := \bigcup_{n=1}^{N} (a_n, b_n), \quad \lambda(\mathcal{O}\backslash \mathcal{O}_N) = \sum_{n=N+1}^{\infty} (b_n - a_n) < \varepsilon.$$

On a alors

$$\begin{aligned} \|\mathbf{1}_A - \mathbf{1}_{\mathcal{O}_N}\|_1 &\le \|\mathbf{1}_A - \mathbf{1}_{\mathcal{O}}\|_1 + \|\mathbf{1}_{\mathcal{O}} - \mathbf{1}_{\mathcal{O}_N}\|_1 \\ &\le \int \mathbf{1}_{\mathcal{O}\backslash A}\, dx + \int \mathbf{1}_{\mathcal{O}\backslash \mathcal{O}_N}\, dx \le 2\varepsilon, \end{aligned}$$

ce qui donne une approximation de $\mathbf{1}_A$ dans L^1 à 2ε près par une fonction en escalier. On traite ensuite le cas d'une fonction étagée par additivité et le cas d'une fonction quelconque par approximation par une suite de fonctions étagées.

Etape 2. On suppose $a, b \in \bar{\mathbb{R}}$. Quitte à définir la fonction $\bar{f}(x) = f(x)$ si $x \in (a,b)$, $\bar{f}(x) = 0$ si $x \notin (a,b)$, il suffit de traiter le cas $f \in \mathcal{L}^1(\mathbb{R})$. D'après le théorème de convergence dominée, il existe $R > 0$ tel que $\|f - f\,\mathbf{1}_{[-R,R]}\|_1 = \|f\,\mathbf{1}_{[-R,R]^c}\|_1 < \varepsilon$ et on peut appliquer l'étape 1 à la fonction $f\,\mathbf{1}_{[-R,R]} \in \mathcal{L}^1(-R,R)$. ◻

Définition 2.2 (Approximation de l'identité) *On appelle approximation de l'identité une suite (ρ_n) bornée de $L^1(\mathbb{R}^d)$ (de borne C) telle que*

$$\int \rho_n\, dx = 1, \quad \forall\, \varepsilon > 0,\ \int_{B_\varepsilon^c} |\rho_n|\, dx \underset{n\to\infty}{\longrightarrow} 0.$$

Lemme 2.3 *Soit (ρ_n) est une approximation de l'identité. Si $f \in C_b(\mathbb{R}^d)$ alors $f * \rho_n \to f$ ponctuellement. Si $f \in \mathcal{L}^1(\mathbb{R}^d)$ alors $f * \rho_n \to f$ dans $\mathcal{L}^1(\mathbb{R}^d)$.*

Preuve du Lemme 2.3. Etape 1. On suppose $f \in C_b(\mathbb{R}^d)$. Pour $x \in \mathbb{R}^d$, on écrit

$$\begin{aligned} f * \rho_n(x) - f(x) &= \int_{\mathbb{R}^d} (f(x-y) - f(x))\rho_n(y)\, dy \\ &= \int_{B_\varepsilon} (f(x-y) - f(x))\rho_n(y)\, dy + \int_{B_\varepsilon^c} (f(x-y) - f(x))\rho_n(y)\, dy = I_1 + I_2. \end{aligned}$$

On a d'une part

$$|I_1| \le \int |\rho_n(y)|\, dy \sup_{|z-x|\le\varepsilon} |f(z) - f(x)| \le C\,\omega_x(\varepsilon) \underset{\varepsilon\to 0}{\longrightarrow} 0,$$

où ω_x désigne le module de continuité de f au point x. On a d'autre part

$$|I_2| \le 2\|f\|_\infty \int_{B_\varepsilon^c} |\rho_n(y)|\, dy \underset{n\to\infty}{\longrightarrow} 0.$$

Etape 2. On suppose f en escalier et donc nulle en dehors d'une boule, et $d = 1$ pour simplifier. L'étape précédente montre que $(f * \rho_n)$ est uniformément bornée dans $\mathcal{L}^\infty$ et $f * \rho_n(x) \to f(x)$ en tout point de continuité x de la fonction f, donc presque partout. En particulier, on a $f * \rho_n \to f$ dans $\mathcal{L}^1([-R, R])$ pour tout $R > 0$. D'autre part, en notant $A > 0$ tel que $f(x) = 0$ pour tout $x \notin [-A, A]$, pour $R > A$, on calcule

$$\begin{aligned}
\int_{|x|\ \geq R} |f * \rho_n - f|\, dx &= \int_{|x|\ \geq R} |f * \rho_n|\, dx \\
&\leq \int_{|x| \geq R} \int_{|y| \leq A} |\rho_n(x-y)|\, |f(y)|\, dydx \\
&\leq \|f\|_\infty \int_{|x| \geq R} \int_{|y| \leq A} |\rho_n(x-y)|\, \mathbf{1}_{|x-y| \geq R-A} dydx \\
&= \|f\|_\infty \int_{\mathbb{R}^d} \int_{|y| \leq A} |\rho_n(z)|\, \mathbf{1}_{|z| \geq R-A} dydz \\
&\leq \|f\|_\infty \lambda(B(0,A)) \int_{|z| \geq R-A} |\rho_n(z)|\, dz \to 0,
\end{aligned}$$

lorsque $n \to \infty$. Cela termine la preuve de $\lim \|f * \rho_n - f\|_{\mathcal{L}^1(\mathbb{R}^d)} = 0$

Etape 3. On suppose $f \in \mathcal{L}^1$. Pour tout $\varepsilon > 0$, il existe une fonction g_ε en escalier telle que $\|f - g_\varepsilon\|_1 \leq \varepsilon$. On écrit alors

$$f - f * \rho_n = (f - g) + (g - g * \rho_n) + (g - f) * \rho_n,$$

de sorte que

$$\|f - f * \rho_n\|_1 \leq \|f - g\|_1 + \|g - g * \rho_n\|_1 + \|f - g\|_1 \|\rho_n\|_1$$

et $\limsup \|f - f * \rho_n\|_1 \leq \varepsilon(1 + C)$ d'après l'étape 2. Il suffit alors de faire tendre $\varepsilon \to 0$ pour conclure. ◻

- Soit f une fonction mesurable. On appelle support de f, on note $\mathrm{supp} f$, le plus petit fermé de $\mathbb{R}^d$ tel que $f(x) = 0$ pour presque tout $x \notin F$. Lorsque f est une fonction continue, on a donc $\mathrm{supp}\, f := \{x \in \mathbb{R}^d;\ f(x) \neq 0\}$. On dit qu'une fonction est à support compact si $\mathrm{supp} f \subset B(0, R)$ pour un certain $R > 0$ (donc $\mathrm{supp} f$ est un compact).
- On note $C_c(\mathbb{R}^d)$ l'espace des fonctions continues et à support compact. On a donc $C_c(\mathbb{R}^d) \subset C_b(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$. De la même manière, on note $C_c^k(\mathbb{R}^d) = C^k(\mathbb{R}^d) \cap C_c(\mathbb{R}^d)$, pour $k \in \{1, \dots, \infty\}$.
- On note $C_0(\mathbb{R}^d)$ l'espace de fonctions $f \in C(\mathbb{R}^d)$ telles que $f(x) \to 0$ lorsque $|x| \to \infty$.

Définition 2.4 (noyau régularisant) *On appelle suite régularisante toute suite (ρ_n) de fonctions telle que*

$$0 \leq \rho_n \in C_c^\infty(\mathbb{R}^d), \quad \mathrm{supp}\rho_n \subset B(0, C/n), \quad \int \rho_n\, dx = 1.$$

En pratique, on prendra $\rho_n(x) := n^d \rho(nx)$ avec $0 \leq \rho \in C_c^\infty(\mathbb{R}^d) \cap \mathcal{P}(\mathbb{R}^d)$.

Il existe bien des suites régularisantes. En effet, on commence par définir

$$\eta(y) := e^{-1/y},\ \forall\, y > 0, \quad \eta(y) := 0,\ \forall\, y \leq 0,$$

et l'on vérifie que $\eta \in C^\infty(\mathbb{R})$, $0 \leq \eta \leq 1$. On pose alors $\rho(x) := \eta(1 - \|x\|^2)$ et $\rho_n(x) := n^d\, C\, \rho(nx)$ avec $C := \Big(\int \rho\Big)^{-1}$.

Théorème 2.5 (Densité $C_c^\infty(\mathbb{R}^d) \subset C_c(\mathbb{R}^d)$) *Soit (ρ_n) une suite régularisante. Pour toute fonction $f \in C_c(\mathbb{R}^d)$, la suite $(f * \rho_n)$ satisfait $f * \rho_n \in C_c^\infty(\mathbb{R}^d)$ et $f * \rho_n \to f$ uniformément.*

Preuve du Théorème 2.5. D'après le Corollaire 1.2 et un argument de récurrence, on a clairement $f * \rho_n \in C^\infty(\mathbb{R}^d)$. Si supp$f \subset B(0,R)$ et supp$\rho_n \subset B(0,C/n)$, on a supp$f*\rho_n \subset B(0,R+C/n)$ et donc $f*\rho_n \in C^\infty_c(\mathbb{R}^d)$. En reprenant la preuve du Lemme 2.3 et en observant qu'il existe un module de continuité uniforme ω tel que $\forall\, x \in \mathbb{R}^d$, $\omega_x(\delta) \le \omega(\delta) \to 0$ lorsque $\delta \to 0$ (en d'autres termes f est uniformément continue, ce qui se démontre classiquement par un argument de compacité) on obtient que $f * \rho_n \to f$ uniformément sur $\mathbb{R}^d$. ◻

Théorème 2.6 (Densité $C^\infty_c(\mathbb{R}^d) \subset \mathcal{L}^1(\mathbb{R}^d)$) *Pour toute fonction $f \in \mathcal{L}^1\mathbb{R}^d)$, il existe une suite (f_n) de $C^\infty_c(\mathbb{R}^d)$ telle que $f_n \to f$ en norme $\mathcal{L}^1$.*

Preuve du Théorème 2.6. On considère (ρ_n) une suite régularisante et on pose $\chi_n := \mathbf{1}_{B(0,n)}$ puis $f_n := (f\,\chi_n) * \rho_n$. En utilisant les mêmes arguments que dans la preuve du Théorème 2.5, on a $f_n \in C^\infty_c(\mathbb{R}^d)$. On écrit enfin

$$f_n - f = (f\,(\chi_n - 1)) * \rho_n + (f * \rho_n - f),$$

et on observe que les deux termes tendent vers 0 dans $\mathcal{L}^1$ (en utilisant notamment (1.1) pour borner le premier terme). ◻

Théorème 2.7 (Stokes en dimension $d = 1$) *Pour toute fonction $f \in C^1(\mathbb{R})$ telle que $f, f' \in \mathcal{L}^1\mathbb{R})$, on a*

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

Preuve du Théorème 2.7. *On fixe $\chi \in C^1(\mathbb{R})$ telle que $\mathbf{1}_{[-1,1]} \le \chi \le \mathbf{1}_{[-2,2]}$ et on note $\chi_R(x) := \chi(x/R)$. Comme $f\chi_R \in C_c(\mathbb{R})$, on a*

$$\int_{\mathbb{R}} (f\chi_R)' dx = [f]^{2R}_{-2R} = 0.$$

En développant, il vient

$$\int_{\mathbb{R}} f'\chi_R dx + \int_{\mathbb{R}} f(\chi_R)' dx = 0.$$

La deuxième intégrale tend vers 0 lorsque $R \to \infty$ puisque $|(\chi_R)'| \le R^{-1}\|\chi'\|_\infty$. On conclut en faisant tendre $R \to \infty$ et en passant à la limite dans la première intégrale grâce au théorème de convergence dominée (on notera que $\chi_R \to 1$ p.p., $0 \le \chi_R \le 1$). ◻

Théorème 2.8 (de Weierstrass de densité des polynômes dans $C(K)$) *Pour tout compact $K \subset \mathbb{R}^d$, l'ensemble des polynômes est dense dans $C(K)$ au sens de la convergence uniforme, où $C(K)$ désigne l'espace des fonctions continues (bornées) de K dans $\mathbb{R}$.*

Preuve du Théorème 2.8. On ne présente la preuve qu'en dimension $d = 1$ et pour le compact $K = [-1/4, 1/4]$, le cas général s'en déduit assez simplement (la preuve complète est donc laissée en exercice). On définit la suite (η_n) par

$$\eta_n(x) := (1 - x^2)^n \text{ si } x \in [-1,1], \quad \eta_n(x) := 0 \text{ si } x \notin [-1,1].$$

On considère alors la suite (ρ_n) définie par

$$\rho_n(x) = a_n^{-1}\eta_n(x), \quad a_n := \int_{\mathbb{R}} \eta_n(x)\,dx.$$

En observant en particulier que $a_n \ge 2/(n+1)$ et donc $\rho_n(x) \le (n+1)(1-\varepsilon^2)^n/2 \to 0$ uniformément sur $B(0,\varepsilon)^c$ pour tout $\varepsilon \in]0,1[$, on vérifie sans difficulté que (ρ_n) est une approximation de l'identité. A une fonction $g \in C([-1/4,1/4])$, on associe la fonction $\tilde g : \mathbb{R} \to \mathbb{R}$ définie par

$$\begin{cases} \tilde g(x) = g(x), \ \forall\, x \in [-1/4, 1/4], \\ \tilde g(x) = 0, \ \forall\, x \notin [-1/2, 1/2]^c, \\ \tilde g \text{ affine sur } [-1/2,-1/4] \cup [1/4,1/2], \end{cases}$$

de sorte que $\tilde g \in C_c(\mathbb{R})$, $\operatorname{supp} g \subset [-1/2, 1/2]$. On définit $g_n := \tilde g * \rho_n$. La suite (g_n) converge uniformément vers la fonction $\tilde g$ d'après le Lemme 2.3 (et l'argument d'uniforme continuité déjà utilisé dans la preuve du Théorème 2.5). On observe enfin

$$\begin{aligned}\tilde g * \rho_n(x) &= \frac{1}{a_n}\int_{-1/2}^{1/2} [1-(x-y)^2]^n \mathbf{1}_{|x-y|\le 1}\, \tilde g(y)\, dy \\ &= \frac{1}{a_n}\int_{-1/2}^{1/2} [1-(x-y)^2]^n \tilde g(y)\, dy =: p_n(x), \quad \forall\, x \in [-1/4, 1/4].\end{aligned}$$

Or p_n est clairement un polynôme, ce qui termine la preuve. $\square$

Corollaire 2.9 (Moments) *Soit $f \in L^1([a,b])$, $a, b \in \mathbb{R}$. On a*

$$M_k(f) := \int_{\mathbb{R}} x^k\, f(x)\, dx = 0,\ \forall\, k \in \mathbb{N}, \quad \textit{implique} \quad f = 0.$$

De la même manière, si $\mu, \nu \in M^1_+([a,b])$, $a, b \in \mathbb{R}$, et $M_k(\mu) = M_k(\nu)$ pour tout $k \ge 0$, alors $\mu = \nu$.

Preuve du Corollaire 2.9. *On fixe $\varphi \in C_c(\mathbb{R})$, de sorte que $\varphi \in C([a,b])$. Pour tout $\varepsilon > 0$, d'après le Théorème 2.8, il existe un polynôme p tel que $\|\varphi - p\|_{L^\infty(a,b)} \le \varepsilon/(b-a)$. On en déduit*

$$\Big| \int_a^b f(x)\varphi(x)\, dx \Big| \le \varepsilon + \Big| \int_a^b f(x) p(x)\, dx \Big| = \varepsilon.$$

En faisant tendre $\varepsilon \to 0$, on en conclut

$$\int_{\mathbb{R}} \bar f(x)\varphi(x)\, dx = \int_a^b f(x)\varphi(x)\, dx = 0, \quad \forall\, \varphi \in C_c(\mathbb{R}),$$

où $\bar f \in L^1(\mathbb{R})$ est l'extension de f par 0 en dehors de $[a,b]$. On a donc

$$\int_{\mathbb{R}} |\bar f(x)|\, dx = \lim_{n\to\infty} \int_{\mathbb{R}} \bar f(x)((\operatorname{sign} \bar f) * \rho_n) = 0,$$

ce qui prouve bien $f = 0$. $\square$

3 Transformation de Fourier dans $M^1_+(\mathbb{R}^d)$ et $L^1(\mathbb{R}^d)$

Pour f une fonction intégrable sur $\mathbb{R}^d$, on définit sa transformée de Fourier

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

Plus généralement, pour μ une mesure bornée sur $\mathbb{R}$, on définit sa transformée de Fourier

$$\hat \mu(\xi) = (\mathcal{F} \mu)(\xi) := \int_{\mathbb{R}^d} e^{-ix\cdot\xi} \mu(dx).$$

Proposition 3.1 *On a $\hat\mu \in C_b(\mathbb{R}^d)$ et $\|\hat\mu\|_\infty \le \mu(\mathbb{R}^d)$ si $\mu \in M^1_+(\mathbb{R}^d)$. On a également $\hat f \in C_0(\mathbb{R}^d)$ et $\|\hat f\|_\infty \le \|f\|_1$ si $f \in \mathcal{L}^1(\mathbb{R}^d)$.*

On n'a pas nécessairement $\hat\mu \in C_0(\mathbb{R}^d)$ si $\mu \in M^1_+(\mathbb{R}^d)$. Par exemple, $\mathcal{F}\delta_0 = 1 \notin C_0(\mathbb{R})$.

Preuve de la Proposition 3.1. *On définit $e_\xi(x) := e^{-ix\,\xi}$. La fonction $e_\xi \in C_b(\mathbb{R}^d)$ et $\xi \mapsto e_\xi(x)$ est continue pour tout $x \in \mathbb{R}^d$, de sorte que $\xi \mapsto (\mathcal{F}\mu)(\xi)$ est continue d'après le Théorème de continuité sous le signe somme. De plus*

$$|\hat\mu(\xi)| \le \int_{\mathbb{R}} |e^{-ix\,\xi}|\mu(dx) \le \mu(\mathbb{R}^d)$$

et de la même manière on trouve

$$\|\hat f\|_\infty \le f\|_{L^1}. \tag{3.1}$$

En dimension $d = 1$, on calcule

$$\begin{aligned}(\mathcal{F}\mathbf{1}_{[a,b]})(\xi) &= \int_a^b e^{-ix\xi}\,dx = \frac{e^{-ib\xi} - e^{-ia\xi}}{-i\xi} \\ &= e^{-i\frac{a+b}{2}\xi}\,\frac{e^{i\frac{b-a}{2}\xi} - e^{-i\frac{b-a}{2}\xi}}{i\xi} = 2\,e^{-i\frac{a+b}{2}\xi}\,\frac{\sin(\xi(b-a)/2)}{\xi} \in C_0(\mathbb{R}).\end{aligned}$$

Ainsi, pour toute fonction en escalier f on a également $\hat f \in C_0(\mathbb{R})$. Par densité des fonctions en escalier et (3.1), on obtient $\hat f \in C_0(\mathbb{R})$. □

La transformation de Fourier jouit d'un certain nombre de propriétés simples et remarquables dont nous donnons une liste maintenant.

Propriétés algébriques de la transformation de Fourier :

- **(F1)** *$\mathcal{F}$ est linéaire au sense où $\mathcal{F}(f + tg) = \mathcal{F}(f) + t\mathcal{F}(g)$.*
- **(F2)** *$\mathcal{F}(f')(\xi) = i\xi\hat f(\xi)$ pour $f \in C^1(\mathbb{R})$ tel que $f, f' \in \mathcal{L}^1(\mathbb{R})$, puisque*

$$\int_{\mathbb{R}} e^{-ix\,\xi} f'(x)\,dx = -\int_{\mathbb{R}} [\frac{d}{dx}e^{-ix\,\xi}]f(x)\,dx = i\xi\int_{\mathbb{R}} e^{-ix\,\xi} f(x)\,dx,$$

où on a effectué une intégration par parties grâce au Théorème 2.7 (de Stokes).

- **(F3)** *$\mathcal{F}(xf)(\xi) = i(\hat f(\xi))'$ pour $f \in \mathcal{L}^1(\mathbb{R})$ tel que $x \mapsto xf(x) \in \mathcal{L}^1(\mathbb{R})$, puisque*

$$\int_{\mathbb{R}} e^{-ix\,\xi} x\,f(x)\,dx = i\int_{\mathbb{R}} [\frac{d}{d\xi}e^{-ix\,\xi}]f(x)\,dx = i\frac{d}{d\xi}\Big[\int_{\mathbb{R}} e^{-ix\,\xi} f(x)\,dx\Big].$$

- **(F4)** *$\widehat{f * g} = \widehat f\,\widehat g$ pour $f, g \in \mathcal{L}^1(\mathbb{R})$, puisque*

$$\begin{aligned}\int_{\mathbb{R}}\Big[\int_{\mathbb{R}} f(x-y)g(y)\,dy\Big]e^{-ix\xi}\,dx &= \int_{\mathbb{R}}\int_{\mathbb{R}} f(x-y)g(y)\,e^{-i(x-y+y)\xi}\,dxdy \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} f(z)g(y)\,e^{-iz\xi}e^{-iy\xi}\,dzdy.\end{aligned}$$

- **(F5)** *Pour $\mathcal{F}(f_\lambda) = \lambda^d(\mathcal{F}f)_{\lambda^{-1}}$, où on définit $g_s(x) := g(x/s)$.*
- **(F6)** *Pour $a \in \mathbb{R}^d$ et une fonction $f : \mathbb{R}^d \to \mathbb{R}$, on note $\tau_a f : \mathbb{R}^d \to \mathbb{R}$ la fonction translatée définie par $(\tau_a f)(x) := f(x-a)$. On observe que si $f \in L^1(\mathbb{R}^d)$ alors*

$$(\mathcal{F}\tau_a f)(\xi) = \int_{\mathbb{R}} e^{-ix\,\xi} f(x-a)\,dx = \int_{\mathbb{R}} e^{-i(y+a)\,\xi} f(y)\,dy = e^{-ia\xi}\hat f(x).$$

- **(F7)** *$\mathcal{F}(e_a f) = \tau_a\hat f$, pour tout $a \in \mathbb{R}^d$, où ici $e_a(x) := e^{iax}$.*
- **(F8)** *En notant $\check g(x) := g(-x)$, on a $(\check{\mathcal{F}} f)(\xi) := (\mathcal{F}f)(-\xi) = (\mathcal{F}\check f)(\xi)$. En particulier, si f est une fonction paire alors $\mathcal{F}f$ est également une fonction paire.*

Théorème 3.2 (Gaussienne) *Pour G la gaussienne standard, on a*

$$\hat{G}(\xi) = e^{-|\xi|^2/2}.$$

Preuve du Théorème 3.2. On considère seulement le cas $d = 1$. Par définition, on a

$$\hat{G}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-ix\xi} e^{-x^2/2}\, dx = \varphi(\xi) e^{-|\xi|^2/2},$$

avec

$$\varphi(\xi) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{1}{2}(x+i\xi)^2} dx.$$

On observe que $\varphi(0) = 1$ (c'est un calcul que nous avons fait au chapitre précédent) et que

$$\begin{aligned} \frac{d}{d\xi}\varphi(\xi) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \frac{d}{d\xi}\left[e^{-\frac{1}{2}(x+i\xi)^2}\right] dx \\ &= \frac{i}{\sqrt{2\pi}} \int_{\mathbb{R}} \frac{d}{dx}\left[e^{-\frac{1}{2}(x+i\xi)^2}\right] dx \\ &= \frac{i}{\sqrt{2\pi}} \left[e^{-\frac{1}{2}(x+i\xi)^2}\right]_{-\infty}^{\infty} = 0, \end{aligned}$$

grâce au Théorème 2.7 (de Stokes). On en déduit $\varphi(\xi) = 1$, et la conclusion. □

Théorème 3.3 (Inversion) *Soit $f \in L^1(\mathbb{R}^d)$ telle que $\hat{f} \in L^1(\mathbb{R}^d)$. Alors $f = \mathcal{F}^{-1}\hat{f}$ avec*

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

En d'autres termes, on a $\mathcal{F}^{-1}\mathcal{F}f = \mathcal{F}\mathcal{F}^{-1}f = f$, pour toute fonction f "convenable".

Preuve du Théorème 3.3. On définit $G_\lambda(x) := G(x/\lambda)$ et on observe que $\hat{G}_\lambda = (2\pi)^{d/2}\lambda^d G_{\lambda^{-1}}$ d'après la propriété **(F5)** et le Théorème 3.2. En utilisant le Théorème de Fubini (comme dans la formule **(F4)**) et la parité de G_λ, on calcule

$$\begin{aligned} (\hat{G}_\lambda * f)(\xi) &= \int_{\mathbb{R}^d} \hat{G}_\lambda(y-\xi) f(y)\, dy \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} G_\lambda(x) e^{-i(y-\xi)\cdot x} f(y)\, dydx \\ &= \int_{\mathbb{R}^d} G_\lambda(x) e^{i\xi\cdot x} \hat{f}(x) dx, \end{aligned}$$

soit donc

$$(\lambda^d G_{\lambda^{-1}} * f)(x) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} G(\xi/\lambda) e^{i\xi\cdot x} \hat{f}(\xi) d\xi.$$

La suite $(\lambda^d G_{\lambda^{-1}})_\lambda$ est une approximation de l'identité et $G_\lambda(x) \nearrow (2\pi)^{-d/2}$ lorsque $\lambda \to \infty$. On obtient l'identité souhaitée en passant à la limite $\lambda \to \infty$ dans cette dernière équation. □

Théorème 3.4 (Unicité) *Soit $f \in \mathcal{L}^1(\mathbb{R}^d)$ telle que $\hat{f} = 0$. Alors $f = 0$.*

Preuve du Théorème 3.3. On applique le théorème Théorème 3.3. □

On dit qu'une suite de mesures de probabilités (μ_n) converge faiblement vers une mesure de probabilité μ, on note $\mu_n \rightharpoonup \mu$, si

$$\int_{\mathbb{R}^d} \varphi(x) d\mu_n(x) \underset{n\to\infty}{\longrightarrow} \int_{\mathbb{R}^d} \varphi(x) d\mu(x), \quad \forall\, \varphi \in C_0(\mathbb{R}^d).$$

Lemme 3.5 *Dans les conditions de la définition précédente, on a $\mu_n \rightharpoonup \mu$ si, et seulement, si*

$$\int_{\mathbb{R}^d} \varphi(x) d\mu_n(x) \underset{n\to\infty}{\longrightarrow} \int_{\mathbb{R}^d} \varphi(x) d\mu(x), \quad \forall \varphi \in \mathcal{X}, \tag{3.2}$$

pour $\mathcal{X} := C_c(\mathbb{R}^d)$ ou $\mathcal{X} := C_b(\mathbb{R}^d)$.

Preuve du Lemme 3.5. Il suffit de montrer que si la convergence a lieu pour tout $\varphi \in C_c(\mathbb{R}^d)$ alors elle a aussi lieu pour tout pour tout $\varphi \in C_b(\mathbb{R}^d)$, puisque les autres implications sont triviales grâce aux inclusions $C_c(\mathbb{R}^d) \subset C_0(\mathbb{R}^d) \subset C_b(\mathbb{R}^d)$. On suppose donc la convergence vraie pour tout $\varphi \in C_c(\mathbb{R}^d)$. Fixons une fonction $\chi \in C_c(\mathbb{R}^d)$, $\mathbf{1}_{B(0,1)} \le \chi \le \mathbf{1}_{B(0,2)}$, et posons $\chi_R(x) := \chi(x/R)$, de sorte que $\chi_R \le 1$ pour tout $R > 0$ et $\chi_R \to 1$ lorsque $R \to \infty$. Par le théorème de convergence dominée, on a

$$\langle \mu, \chi_R \rangle := \int_{\mathbb{R}^d} \chi_R(x) d\mu_n(x) \underset{R\to\infty}{\longrightarrow} 1,$$

et pour tout $\varepsilon > 0$, on peut donc fixer R tel que $\langle \mu, \chi_R \rangle \ge 1 - \varepsilon/2$. Par hypothèse, il existe N tel que pour tout $n \ge N$, on a

$$\int_{\mathbb{R}^d} \chi_R(x) d\mu_n(x) \ge \int_{\mathbb{R}^d} \chi_R d\mu(x) - \varepsilon/2 \ge 1 - \varepsilon.$$

En particulier, on a

$$\int_{\mathbb{R}^d} (1 - \chi_R(x)) d\mu(x) \le \varepsilon \quad \text{et} \quad \int_{\mathbb{R}^d} (1 - \chi_R(x)) d\mu_n(x) \le \varepsilon, \ \forall n \ge N.$$

On considère maintenant $\varphi \in C_b$, $\|\varphi\|_\infty \le 1$, et on écrit

$$\int_{\mathbb{R}^d} \varphi d\mu_n - \int_{\mathbb{R}^d} \varphi d\mu = \int_{\mathbb{R}^d} \varphi(1-\chi_R) d\mu_n + \int_{\mathbb{R}^d} \varphi \chi_R d\mu_n - \int_{\mathbb{R}^d} \varphi \chi_R d\mu + \int_{\mathbb{R}^d} \varphi(1-\chi_R) d\mu.$$

Comme $\varphi \chi_R \in C_c(\mathbb{R}^d)$, on peut fixer $N' \ge N$ de sorte que

$$\Big| \int_{\mathbb{R}^d} \varphi \chi_R d\mu_n - \int_{\mathbb{R}^d} \varphi \chi_R d\mu \Big| \le \varepsilon, \quad \forall n \ge N'.$$

On conclut

$$\Big| \int_{\mathbb{R}^d} \varphi d\mu_n - \int_{\mathbb{R}^d} \varphi d\mu \Big| \le 3\varepsilon, \quad \forall n \ge N',$$

en combinant les différentes informations précédentes. Cela finit la démonstration de (3.2) pour $\mathcal{X} = C_b(\mathbb{R}^d)$. ◻

Théorème 3.6 (de Lévy - version faible) *Soit (μ_n) une suite de mesures de probabilité et μ une mesure de probabilité. On a $\mu_n \rightharpoonup \mu$ faiblement si, et seulement si, $\hat{\mu}_n$ converge (ponctuellement) vers $\hat{\mu}$.*

Preuve du Théorème 3.6. Le sens direct est clair puisque $x \mapsto e^{ix\cdot\xi} \in C_b(\mathbb{R}^d)$ pour tout $\xi \in \mathbb{R}^d$. On suppose inversement que $\hat{\mu}_n \to \hat{\mu}$ ponctuellement. Fixons $\varphi \in C_c^{d+1}(\mathbb{R}^d)$ de sorte que $D^{d+1}\varphi \in C_c(\mathbb{R}^d) \subset L^1(\mathbb{R}^d)$, $(1+|\xi|)^{d+1}\hat{\varphi} \in C_b$ et donc $\hat{\varphi} \in L^1$. D'après le Théorème d'inversion de la transformation de Fourier, on a $\mathcal{F}\tilde{\varphi} = \varphi$, en notant $\tilde{\varphi} := \mathcal{F}^{-1}\varphi$. Par Fubini, on en déduit que

$$\int_{\mathbb{R}^d} \varphi \, d\mu_n = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{-ix\cdot\xi} \tilde{\varphi}(\xi) d\xi d\mu_n(x) = \int_{\mathbb{R}^d} \hat{\mu}_n \tilde{\varphi} \, d\xi.$$

En utilisant le théorème de convergence dominée dans le dernier terme, on obtient

$$\int_{\mathbb{R}^d} \varphi \, d\mu_n \underset{n\to\infty}{\longrightarrow} \int_{\mathbb{R}^d} \hat{\mu} \tilde{\varphi} d\xi = \int_{\mathbb{R}^d} \varphi \, d\mu.$$

On conclut en utilisant la densité $C_c^2(\mathbb{R}^d) \subset C_c(\mathbb{R}^d)$. ◻

4 Autres transformations

On se rappelle le résultat suvant.

Lemme 4.1 *Soient μ, ν deux mesures finies sur $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$. On a $\mu = \nu$ si, et seulement*

$$\forall\, x \in \mathbb{R}, \quad \mu(]-\infty, x]) = \nu(]-\infty, x]).$$

Cela permet de motiver la notion de fonction de répartition qui suit.

Définition 4.2 *Soit μ une mesure de probabilité sur $\mathbb{R}$. On appelle fonction de répartition de μ, la fonction $F : \mathbb{R} \to \mathbb{R}$ déffinie par*

$$\forall\, x \in \mathbb{R}, \quad F(x) = \mu(]-\infty, x]).$$

Plus généralement, on peut associer à une mesure μ sur $\mathbb{R}^d$ une fonction de répartition $F : \mathbb{R}^d \to \mathbb{R}$ en posant

$$\forall\, x = (x_1, \ldots, x_d) \in \mathbb{R}^d, \quad F(x) = \mu(]-\infty, x_1] \times \ldots \times]-\infty, x_d]).$$

De la même manière qu'en dimension $d = 1$, la fonction de répartition F caractérise la mesure μ.

Théorème 4.3 (Admis) *Une fonction $F : \mathbb{R} \to \mathbb{R}$ est la fonction de répartition d'une mesure de probabilité μ sur $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ si, et seulement si, F est continue à droite, croissante, $F(-\infty) = 0$ et $F(+\infty) = 1$. De plus, F est absolument continue, au sens où il existe un module de continuité ω_F tel que pour toute suite $[a_i, b_i]$ d'intervalles d'interieurs disjoints*

$$\sum_i (b_i - a_i) < \delta \quad \textit{implique} \quad \sum_i (F(b_i) - F(a_i)) \leq \omega_F(\delta),$$

si, et seulement si, μ est absolument continue par rapport à la mesure de Lebesgue λ sur $\mathbb{R}$.

Définition 4.4 *Soit μ une mesure de probabilité sur $\mathbb{R}$. On appelle transformation de Laplace de μ, la fonction $M : \mathbb{R} \to [0, \infty]$ déffinie par*

$$\forall\, s \in \mathbb{R}, \quad M(s) := \int_{\mathbb{R}} e^{sx} \mu(dx).$$

On appelle domaine de M l'ensemble $\operatorname{dom} M := \{s \in \mathbb{R},\, M(s) < \infty\}$.

Lemme 4.5 (Admis) *$\operatorname{dom} M$ est un intervalle contenant 0. La transformation de Laplace caractérise la mesure de probabilité.*

Définition 4.6 *Soit μ une mesure de probabilité sur $\mathbb{N}$. On appelle fonction génératrice la série entière*

$$\forall\, s,\, |s| \leq 1, \quad G(s) := \sum_{k=0}^{\infty} s^k \mu(\{k\}).$$

Lemme 4.7 (Admis) *La fonction génératrice caractérise la mesure de probabilité.*