3. Convolution, inégalités, approximation et régularisation

3.1. Convolution dans $\mathrm{L}_1(\mathbb{R}^d)$

On se place sur $\mathbb{R}^d$ muni de la mesure de Lebesgue. On commence par définir le *produit de convolution* $f * g$ de deux fonctions f et g mesurables ≥ 0, en posant

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

Le résultat est à valeurs dans $[0, +\infty]$, et les propriétés d'invariance de la mesure de Lebesgue entraînent que

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

pour tout x. On veut ensuite montrer que l'intégrale ci-dessus est finie pour presque tout $x \in \mathbb{R}^d$, lorsque f et g ont des intégrales finies. On va obtenir ce résultat en montrant que $\int (f * g)(x)\,dx < +\infty$. La fonction $h : (x, y) \to f(x-y)g(y)$ est mesurable ≥ 0 sur le produit $\mathbb{R}^d \times \mathbb{R}^d$, ce qui permet d'appliquer le théorème de Fubini positif. L'intégrale double sur $\mathbb{R}^d \times \mathbb{R}^d$ de cette fonction $h(x, y)$ est donc

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

Par l'invariance de la mesure de Lebesgue par translation on a $\int f(x-y)\,dx = \int f(x)\,dx$ pour tout y, de sorte que

$$(*) \qquad \int_{\mathbb{R}^d} \Bigl(\int_{\mathbb{R}^d} f(x-y)g(y)\,dx \Bigr)\,dy = \Bigl(\int_{\mathbb{R}^d} f(x)\,dx \Bigr) \Bigl(\int_{\mathbb{R}^d} g(y)\,dy \Bigr) < +\infty.$$

Supposons maintenant que f et g soient intégrables, réelles ou complexes. On sait déjà que $\int |f(x-y)|\,|g(y)|\,dy$ est fini pour presque tout x, ce qui montre que la fonction $y \to f(x-y)g(y)$ est intégrable pour presque tout x. On peut donc poser pour presque tout x

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

et les calculs obtenus en majorant par les modules montrent que $f * g$ est intégrable ; plus précisément on obtient l'inégalité de normes

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

Si on reprend avec Fubini le calcul qui menait à la formule $(*)$, on obtient pour deux fonctions intégrables réelles ou complexes f et g que

$$\int_{\mathbb{R}^d} (f * g)(x)\,dx = \Bigl(\int_{\mathbb{R}^d} f(x)\,dx \Bigr) \Bigl(\int_{\mathbb{R}^d} g(x)\,dx \Bigr).$$

On note en passant que pour des fonctions intégrables positives, on a $\|f * g\|_1 = \|f\|_1 \|g\|_1$.

Exercice. Soient $f, g \in \mathrm{L}_1(\mathbb{R}^d)$; calculer la transformée de Fourier de $f * g$.

Solution. On vient de voir que si f_1 et g_1 sont deux fonctions intégrables sur $\mathbb{R}^d$, alors $f_1 * g_1$ est intégrable et $\int f_1 * g_1 = (\int f_1)(\int g_1)$. On applique ceci, pour $t \in \mathbb{R}^d$ fixé, avec $f_1(x) = f(x)\,\mathrm{e}^{-ix\cdot t}$ et $g_1(x) = g(x)\,\mathrm{e}^{-ix\cdot t}$. On obtient

$$(f_1 * g_1)(x) = \int f(x-y)\,\mathrm{e}^{-i(x-y)\cdot t}\, g(y)\,\mathrm{e}^{-iy\cdot t}\; dy = \mathrm{e}^{-ix\cdot t}(f * g)(x)$$

et

$$(\widehat{f * g})(t) = \int (f_1 * g_1)(x)\,dx = \Bigl(\int f_1(x)\,dx\Bigr)\Bigl(\int g_1(x)\,dx\Bigr) = \widehat{f}(t)\,\widehat{g}(t).$$

Quand $f = \frac{1}{2a}\,\mathbf{1}_{[-a,a]}$, $a > 0$, on calcule facilement $\widehat{f}(t) = \sin(at)/at$ (pour $t \neq 0$, et $\widehat{f}(0) = 1$). On obtient donc sans calculs que $\bigl(\sin(at)/at\bigr)^2$ est la transformée de Fourier de $g = f * f$, la fonction en triangle telle que $g(-2a) = g(2a) = 0$, $g(0) = 1/(2a)$, et affine sur $[-2a, 0]$ et $[0, 2a]$, nulle hors de $[-2a, 2a]$.

Exercice. Si $f, g \in \mathrm{L}_1(\mathbb{R}^d)$, si f est nulle en dehors du compact A et g nulle en dehors du compact B, montrer que $f * g$ est nulle hors de la *somme de Minkowski*

$$\mathrm{A} + \mathrm{B} = \{a + b : a \in \mathrm{A},\ b \in \mathrm{B}\}.$$

Reprendre le même résultat si A et B sont deux boréliens tels que

$$\lambda\bigl(\{x \in \mathbb{R}^d : f(x) \neq 0 \text{ et } x \notin \mathrm{A}\}\bigr) = 0,\quad \lambda\bigl(\{x \in \mathbb{R}^d : g(x) \neq 0 \text{ et } x \notin \mathrm{B}\}\bigr) = 0,$$

où λ désigne la mesure de Lebesgue sur $\mathbb{R}^d$.

Autres cadres pour la convolution : le cadre périodique

Dans le cas périodique, on s'occupe en général de fonctions 2π-périodiques sur $\mathbb{R}$; une fonction f est dite intégrable dans ce cadre si elle est intégrable sur une période. Dans ce cas, on vérifie facilement que $\int_a^{a+2\pi} f(t)\,dt = \int_0^{2\pi} f(t)\,dt$ pour tout réel a. Si f et g sont deux fonctions 2π-périodiques intégrables, on définira leur *convolution périodique* en posant

$$(f *_{\mathrm{per}} g)(s) = \int_0^{2\pi} f(s-t)g(t)\,\frac{dt}{2\pi} = \int_a^{a+2\pi} f(s-t)g(t)\,\frac{dt}{2\pi}\cdot$$

On peut généraliser au cas périodique en plusieurs variables ; on peut aussi changer de langage et dire qu'on est en train de travailler avec le groupe compact abélien $\mathbb{R}/2\pi\mathbb{Z}$ et sa probabilité invariante par translation. La convolution se généralise dans le cadre plus général des groupes abéliens localement compacts, munis de leur *mesure de Haar*.

Un modèle pour le groupe topologique quotient $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ est le groupe multiplicatif U des nombres complexes de module 1. Si θ est un nombre réel, il est clair que $\mathrm{e}^{i\theta} \in \mathrm{U}$ ne dépend que de la classe $\widehat{\theta}$ de θ dans $\mathbb{R}/2\pi\mathbb{Z}$, et l'application $\widehat{\theta} \to \mathrm{e}^{i\theta}$ ainsi définie est un isomorphisme du groupe additif $\mathbb{R}/2\pi\mathbb{Z}$ sur le groupe multiplicatif U. Munissons le cercle U de son unique probabilité μ invariante par rotation. La formule de convolution prend alors la forme suivante, pour deux fonctions F et G intégrables sur le cercle unité U,

$$\forall x \in \mathrm{U}, \quad (\mathrm{F} * \mathrm{G})(x) = \int_{\mathrm{U}} \mathrm{F}(xy^{-1})\mathrm{G}(y)\, d\mu(y).$$

C'est la forme qu'il faut utiliser pour définir la convolution sur les groupes non commutatifs, où la loi est toujours notée multiplicativement.

Si on paramétrise le cercle par $x = \mathrm{e}^{i\theta}$, $\theta \in [0, 2\pi]$ et si on pose en plus $f(\theta) = \mathrm{F}(\mathrm{e}^{i\theta})$, $g(\theta) = \mathrm{G}(\mathrm{e}^{i\theta})$ on retrouve l'écriture

$$(f *_{\mathrm{per}} g)(\theta) = \int_0^{2\pi} f(\theta - t) g(t)\, \frac{dt}{2\pi}.$$

Les fonctions f et g sont deux fonctions 2π-périodiques définies sur $\mathbb{R}$, mais on peut aussi les considérer simplement comme deux éléments de $\mathrm{L}_1(0, 2\pi)$.

Quand on parlera de $\mathbb{T}$ ou $\mathbb{T}^d$, on choisira selon les besoins l'une des trois présentations précédentes. On regardera $\mathbb{R}^d$ ou $\mathbb{T}^d$ comme un groupe noté additivement, avec élément neutre noté 0, et muni d'une distance d ; c'est si on veut la distance euclidienne pour $\mathbb{R}^d$, et pour $\mathbb{R}^d/(2\pi\mathbb{Z})^d$, $d(x, y)$ désignera la plus courte distance (euclidienne) entre deux représentants $x_1, y_1 \in \mathbb{R}^d$ des classes $x, y \in \mathbb{R}^d/(2\pi\mathbb{Z})^d$.

Exemple. On a si on pose $e_n(t) = \mathrm{e}^{int}$, $n \in \mathbb{Z}$,

$$(f *_{\mathrm{per}} e_n) = c_n(f)\, e_n,$$

où $c_n(f)$ est le $n$ième coefficient de Fourier complexe de f,

$$c_n(f) = \int_0^{2\pi} f(t)\, \mathrm{e}^{-int}\, \frac{dt}{2\pi}.$$

3.2. Densité des fonctions continues

Lemme. *Pour toute mesure finie μ sur $(\mathbb{R}^d, \mathcal{B})$, pour tout $\varepsilon > 0$ et pour tout borélien $\mathrm{A} \in \mathcal{B}$, il existe une fonction réelle continue φ sur $\mathbb{R}^d$ telle que*

$$\int_{\mathbb{R}^d} \bigl| \mathbf{1}_{\mathrm{A}}(x) - \varphi(x) \bigr|\, d\mu(x) < \varepsilon.$$

Démonstration. Considérons la classe $\mathcal{C}$ des boréliens A de $\mathbb{R}^d$ tels que $\mathbf{1}_{\mathrm{A}}$ soit limite dans $\mathrm{L}_1(\mathbb{R}^d, \mu)$ d'une suite (φ_n) de fonctions continues telles que $0 \le \varphi_n \le 1$. Autrement dit, on considère l'ensemble $\mathcal{E}$ des fonctions réelles continues φ sur $\mathbb{R}^d$ telles que $0 \le \varphi \le 1$, puis les ensembles $\mathrm{A} \subset \mathbb{R}^d$ tels que $\mathbf{1}_{\mathrm{A}}$ soit dans l'adhérence $\overline{\mathcal{E}}$ de $\mathcal{E}$ (pour la norme de $\mathrm{L}_1(\mu)$). On va montrer que $\mathcal{C}$ est une tribu de parties de $\mathbb{R}^d$, qui contient les ouverts. Si U est un ouvert de $\mathbb{R}^d$, considérons pour chaque entier $n \ge 1$ la fonction φ_n définie par

$$\forall x \in \mathbb{R}^d, \quad \varphi_n(x) = \min(1, n\, \mathrm{dist}(x, \mathrm{U}^c)).$$

Cette fonction est continue, la suite (φ_n) tend en croissant vers $\mathbf{1}_{\mathrm{U}}$; par Lebesgue, puisque la mesure μ est finie, on obtient

$$\int \bigl| \mathbf{1}_{\mathrm{U}} - \varphi_n \bigr|\, d\mu \to 0,$$

donc $\mathrm{U} \in \mathcal{C}$. En particulier la classe $\mathcal{C}$ contient $\mathbb{R}^d$ et $\emptyset$. Si A et B sont dans $\mathcal{C}$, considérons $\mathbf{1}_{\mathrm{A} \cap \mathrm{B}} = \mathbf{1}_\mathrm{A} \mathbf{1}_\mathrm{B}$; il existe des suites (φ_n) et (ψ_n) dans $\mathcal{E}$ qui tendent vers $\mathbf{1}_\mathrm{A}$ et $\mathbf{1}_\mathrm{B}$, et

$$\mathbf{1}_\mathrm{A} \mathbf{1}_\mathrm{B} - \varphi_n \psi_n = (\mathbf{1}_\mathrm{A} - \varphi_n) \mathbf{1}_\mathrm{B} + \varphi_n (\mathbf{1}_\mathrm{B} - \psi_n)$$

tend vers 0 dans $\mathrm{L}_1(\mu)$, donc $\mathrm{A} \cap \mathrm{B} \in \mathcal{C}$.

Si $(\varphi_n) \subset \mathcal{E}$ tend vers $\mathbf{1}_\mathrm{A}$, il est clair que $\mathbf{1} - \varphi_n$ tend vers $\mathbf{1}_{\mathrm{A}^c}$, donc $\mathrm{A}^c \in \mathcal{C}$ quand $\mathrm{A} \in \mathcal{C}$. Enfin si une suite (A_n) d'éléments de $\mathcal{C}$ tend en croissant vers A, la suite $(\mathbf{1}_{\mathrm{A}_n})$ tend vers $\mathbf{1}_\mathrm{A}$ dans $\mathrm{L}_1(\mu)$ par convergence dominée ; comme $\mathbf{1}_{\mathrm{A}_n} \in \overline{\mathcal{E}}$ pour tout n, on en déduit $\mathbf{1}_\mathrm{A} \in \overline{\mathcal{E}}$ à la limite, c'est-à-dire $\mathrm{A} \in \mathcal{C}$. On a montré que la classe $\mathcal{C}$ est une tribu qui contient les ouverts. Il en résulte que tout borélien est dans $\mathcal{C}$.

Remarque. La démonstration du lemme précédent fonctionne tout aussi bien si μ est une mesure finie sur la tribu borélienne d'un espace métrique (X, d) quelconque.

Définition. Le *support* d'une fonction continue f définie sur $\mathbb{R}^d$ est l'adhérence de l'ensemble des points où f est non nulle,

$$\operatorname{supp}(f) = \overline{\{x \in \mathbb{R}^d : f(x) \neq 0\}}.$$

Théorème. *Les fonctions continues à support compact sont denses dans* $\mathrm{L}_1(\mathbb{R}^d)$.

Démonstration. Il est clair que les fonctions étagées sont denses dans L_1 : pour toute fonction $f \in \mathrm{L}_1(\mathbb{R}^d)$, il existe une suite de fonctions étagées (f_n) telle que $|f_n| \le |f|$ et $f_n \to f$ simplement ; par convergence dominée, f est limite dans L_1 des f_n. Pour montrer le théorème, il suffit de voir que toute indicatrice de borélien peut être approchée par des fonctions continues à support compact. Par linéarité on obtiendra que toute fonction étagée peut être approchée par une fonction continue à support compact, d'où le résultat puisque les étagées sont denses dans $\mathrm{L}_1(\mathbb{R}^d)$.

Soit A un borélien de $\mathbb{R}^d$; alors $\mathbf{1}_\mathrm{A}$ est limite dans $\mathrm{L}_1(\mathbb{R}^d)$ d'une suite $(\mathbf{1}_{\mathrm{A}_n})$ correspondant à des boréliens bornés (intersections de A avec des boules de rayon n par exemple), et en posant $\mathrm{B} = \mathrm{A}_n$ pour n assez grand on aura

$$\int_{\mathbb{R}^d} |\mathbf{1}_\mathrm{A}(x) - \mathbf{1}_\mathrm{B}(x)| \, dx < \varepsilon/2.$$

Introduisons une fonction continue auxiliaire $\theta \ge 0$, à support compact et égale à 1 en tout point du borné B, par exemple définie par $\theta(x) = \max(1 - \operatorname{dist}(x, \mathrm{B}), 0)$. Posons ensuite $d\mu(x) = \theta(x) \, dx$; c'est une mesure finie sur $\mathbb{R}^d$. D'après le lemme, il existe une fonction continue φ sur $\mathbb{R}^d$ telle que

$$\int_{\mathbb{R}^d} \bigl|\mathbf{1}_\mathrm{B}(x) - \varphi(x)\bigr| \, \theta(x) \, dx < \varepsilon/2.$$

Mais $\theta \, \mathbf{1}_\mathrm{B} = \mathbf{1}_\mathrm{B}$; on a donc trouvé une fonction continue à support compact $\psi = \theta\varphi$ telle que

$$\int_{\mathbb{R}^d} \bigl|\mathbf{1}_\mathrm{B}(x) - \psi(x)\bigr| \, dx = \int_{\mathbb{R}^d} \bigl|\theta \, \mathbf{1}_\mathrm{B} - \theta\varphi\bigr| < \varepsilon/2.$$

Finalement,

$$\int_{\mathbb{R}^d} \bigl|\mathbf{1}_\mathrm{A}(x) - \psi(x)\bigr| \, dx < \varepsilon.$$

Remarque. Les fonctions *en escalier* sont denses aussi (en plusieurs dimensions, on appelle ainsi toute combinaison linéaire de fonctions indicatrices de produits d'intervalles).

Pour montrer cette densité, on peut reprendre ce qui précède en remplaçant fonction continue par fonction en escalier, ou bien constater qu'une fonction continue à support compact peut être approchée dans L_1 par des fonctions en escalier (utiliser la continuité uniforme).

3.3. Inégalités

3.3.1. Inégalité de Jensen

Proposition : inégalité de Jensen. *On suppose que* $(\Omega, \mathcal{A}, \mu)$ *est un espace de probabilité,* f *une fonction réelle* μ*-intégrable sur* Ω *et* φ *une fonction convexe sur* $\mathbb{R}$*. On a*

$$\varphi\Big(\int_\Omega f(\omega)\,d\mu(\omega)\Big) \le \int_\Omega \varphi(f(\omega))\,d\mu(\omega).$$

Il est possible que $\varphi(f)$ *ne soit pas* μ*-intégrable, mais l'intégrale a toujours un sens (en admettant la valeur* $+\infty$*) parce que* $\varphi(f)_- = \max(-\varphi(f), 0)$ *est intégrable.*

Si la fonction f ne prend que deux valeurs u et v sur l'espace Ω, avec probabilités $1-t$ et t, l'inégalité de Jensen se réduit à la définition de la convexité de φ, à savoir $\varphi\big((1-t)u + tv\big) \le (1-t)\varphi(u) + t\varphi(v)$.

Démonstration. On pose $m = \int_\Omega f(\omega)\,d\mu(\omega) \in \mathbb{R}$ (c'est la *valeur moyenne* de f). On considère une fonction affine d'appui h à la fonction convexe φ au point m, $h(x) = ax+b$, avec $h(m) = \varphi(m)$ et $h \le \varphi$ partout. On a ensuite

$$\varphi(m) = h(m) = am + b = \int_\Omega (af(\omega) + b)\,d\mu(\omega) \le \int_\Omega \varphi(f(\omega))\,d\mu(\omega).$$

La fonction $\varphi(f)_- = \max(-\varphi(f), 0)$ est intégrable ; en effet, on a $-\varphi(f) \le -af - b \le |a|\,|f| + |b|$, donc $\max(-\varphi(f), 0) \le |a|\,|f| + |b|$ est intégrable.

Remarque. Si on considère la fonction convexe $\varphi(t) = |t|^p$, $1 \le p < +\infty$ et une mesure finie ν sur Ω, on a

$$\Big|\int_\Omega f(\omega)\,d\nu(\omega)\Big|^p \le \nu(\Omega)^{p-1} \int_\Omega |f(\omega)|^p\,d\nu(\omega).$$

Pour obtenir cette inégalité dans le cas non trivial où $\nu(\Omega) \ne 0$, il suffit de raisonner sur la probabilité $d\mu(\omega) = \nu(\Omega)^{-1} d\nu(\omega)$ et d'utiliser l'homogénéité de la fonction puissance.

3.3.2. Inégalité de Minkowski

Lemme. *Si* ν *est une mesure* σ*-finie sur un espace mesurable* $(\mathrm{Y}, \mathcal{B})$*, on peut trouver une fonction mesurable* k *strictement positive en tout point de* Y*, et telle que* $\int_\mathrm{Y} k\,d\nu = 1$.

Preuve. Il existe une suite $(\mathrm{A}_n) \subset \mathcal{B}$ d'ensembles de mesure finie > 0 qui recouvre Y ; on pose

$$k = \sum_{n=0}^{+\infty} \frac{2^{-n-1}}{\nu(\mathrm{A}_n)}\,\mathbf{1}_{\mathrm{A}_n}.$$

On a $k(y) > 0$ pour tout $y \in \mathrm{Y}$, et $\int_\mathrm{Y} k(y)\,d\nu(y) = 1$.

Théorème. *Si $f(x,y)$ est mesurable sur un espace mesurable produit* $(\mathrm{X}\times\mathrm{Y},\mathcal{A}\otimes\mathcal{B})$, *si μ et ν sont σ-finies sur* X *et* Y *respectivement, et si $p\ge 1$, on a*

$$\Bigl(\int_{\mathrm{X}}\Bigl(\int_{\mathrm{Y}}|f(x,y)|\,d\nu(y)\Bigr)^p\,d\mu(x)\Bigr)^{1/p}\le\int_{\mathrm{Y}}\Bigl(\int_{\mathrm{X}}|f(x,y)|^p\,d\mu(x)\Bigr)^{1/p}\,d\nu(y).$$

Il faut comprendre l'inégalité de Minkowski de la façon suivante : si $f_{y_1},\ldots,f_{y_n}$ est une famille finie de fonctions dans $\mathrm{L}_p(\mathrm{X},\mu)$, l'inégalité triangulaire dans $\mathrm{L}_p(\mathrm{X},\mu)$ nous donne pour la fonction $\mathrm{F}=\sum_{j=1}^n f_{y_j}\in\mathrm{L}_p(\mathrm{X},\mu)$

$$\Bigl(\int_{\mathrm{X}}\Bigl|\sum_{j=1}^n f_{y_j}(x)\Bigr|^p\,d\mu(x)\Bigr)^{1/p}=\|\mathrm{F}\|_p\le\sum_{j=1}^n\|f_{y_j}\|_p=\sum_{j=1}^n\Bigl(\int_{\mathrm{X}}|f_{y_j}(x)|^p\,d\mu(x)\Bigr)^{1/p}.$$

Dans l'inégalité de Minkowski, on a une famille (f_y) de fonctions de $\mathrm{L}_p(\mathrm{X},\mu)$, qui dépend d'un paramètre continu $y\in\mathrm{Y}$, famille définie par $f_y(x)=|f(x,y)|$, et on considère son intégrale $\mathrm{F}\in\mathrm{L}_p(\mathrm{X},\mu)$ par rapport au paramètre y, au lieu de la somme finie considérée précédemment. Cette fonction de la variable x est égale à

$$\mathrm{F}(x)=\int_{\mathrm{Y}}f_y(x)\,d\nu(y)=\int_{\mathrm{Y}}|f(x,y)|\,d\nu(y).$$

L'inégalité de Minkowski dit que la norme de F dans $\mathrm{L}_p(\mathrm{X},\mu)$ est majorée par l'intégrale par rapport au paramètre des normes dans $\mathrm{L}_p(\mathrm{X},\mu)$ des morceaux f_y. Il s'agit donc d'une version continue de l'inégalité triangulaire.

Démonstration. Clairement on peut se limiter au cas $f\ge 0$. Posons

$$\mathrm{M}=\int_{\mathrm{Y}}\bigl(\int_{\mathrm{X}}f(x,y)^p\,d\mu(x)\bigr)^{1/p}\,d\nu(y).$$

Si $\mathrm{M}=+\infty$, la relation à démontrer est évidente. Soit $t<1$ fixé, et supposons $\mathrm{M}\le t<1$; posons

$$g(y)=\Bigl(\int_{\mathrm{X}}f(x,y)^p\,d\mu(x)\Bigr)^{1/p}.$$

On a $\int_{\mathrm{Y}}g(y)\,d\nu(y)=\mathrm{M}<1$. On va suivre une démarche analogue à celle qui nous a permis de voir que la norme de L_p est bien une norme. Dans l'interprétation donnée avant la démonstration, $g(y)$ est la norme de la fonction f_y. Puisque ν est σ-finie, on peut trouver $h>g\ge 0$ telle que $\int h(y)\,d\nu(y)=1$ (prendre $h=g+(1-\mathrm{M})k$, avec k donnée par le lemme). Écrivons

$$\Bigl(\int_{\mathrm{Y}}f(x,y)\,d\nu(y)\Bigr)^p=\Bigl(\int_{\mathrm{Y}}h(y)^{-1}\,f(x,y)\,h(y)\,d\nu(y)\Bigr)^p\le$$

$$\le\int_{\mathrm{Y}}\bigl(h(y)^{-1}\,f(x,y)\bigr)^p\,h(y)\,d\nu(y)=\int_{\mathrm{Y}}f(x,y)^p\,h(y)^{1-p}\,d\nu(y)$$

où on a utilisé Jensen pour la probabilité $h(y)\,d\nu(y)$, puis avec Fubini on obtient

$$\int_{\mathrm{X}}\Bigl(\int_{\mathrm{Y}}f(x,y)\,d\nu(y)\Bigr)^p\,d\mu(x)\le\int_{\mathrm{X}}\int_{\mathrm{Y}}f(x,y)^p\,h(y)^{1-p}\,d\nu(y)d\mu(x)=$$

$$\int_{\mathrm{Y}} \Bigl(\int_{\mathrm{X}} f(x,y)^p \, d\mu(x)\Bigr) \, h(y)^{1-p} \, d\nu(y) = \int_{\mathrm{Y}} g(y)^p h(y)^{1-p} \, d\nu(y) \le \int_{\mathrm{Y}} h(y) \, d\nu(y) = 1.$$

Par homogénéité, on voit qu'on a prouvé que

$$\Bigl(\int_{\mathrm{X}} \Bigl(\int_{\mathrm{Y}} |f(x,y)| \, d\nu(y)\Bigr)^p \, d\mu(x)\Bigr)^{1/p} \le \frac{1}{t} \int_{\mathrm{Y}} \Bigl(\int_{\mathrm{X}} |f(x,y)|^p \, d\mu(x)\Bigr)^{1/p} \, d\nu(y)$$

pour tout $t < 1$, d'où le résultat voulu en faisant tendre t vers 1.

Exercice. On suppose $0 < r < p < +\infty$. Montrer que

$$\Bigl(\int_{\mathrm{X}} \Bigl(\int_{\mathrm{Y}} |f(x,y)|^r \, d\nu(y)\Bigr)^{p/r} \, d\mu(x)\Bigr)^{1/p} \le \Bigl(\int_{\mathrm{Y}} \Bigl(\int_{\mathrm{X}} |f(x,y)|^p \, d\mu(x)\Bigr)^{r/p} \, d\nu(y)\Bigr)^{1/r}.$$

3.3.3. Inégalité de Hölder

On dit que deux nombres p, q de $[1, +\infty]$ forment un couple *d'exposants conjugués* s'ils vérifient la relation

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

Cette relation est à prendre au sens étendu : $(1, +\infty)$ est un cas particulier de couple d'exposants conjugués. Dans le cas $1 < p < +\infty$, on notera que $q(p-1) = p$.

Théorème : inégalité de Hölder. *Soient $p, q \in [1, +\infty]$ tels que $1/p + 1/q = 1$; si $f \in \mathrm{L}_p(\Omega, \mathcal{A}, \mu)$ et $g \in \mathrm{L}_q(\Omega, \mathcal{A}, \mu)$, la fonction produit fg est intégrable et*

$$\Bigl|\int_\Omega fg \, d\mu\Bigr| \le \|f\|_p \, \|g\|_q.$$

Démonstration. Si $p = \infty$, alors $q = 1$; la fonction f est (presque-sûrement) bornée par $\mathrm{M} = \|f\|_\infty$ et g est intégrable ; le produit fg est mesurable et $|fg| \le \mathrm{M}\,|g|$, donc fg est intégrable et

$$\Bigl|\int fg\Bigr| \le \int |fg| \le \mathrm{M} \int |g| = \|f\|_\infty \, \|g\|_1.$$

Supposons maintenant $1 < p < +\infty$. Pour tous nombres réels $t, u \ge 0$, on a la relation

$$tu \le \frac{1}{p}\, t^p + \frac{1}{q}\, u^q \, ;$$

cette relation résulte immédiatement de la convexité de l'exponentielle : on pose $t^p = \mathrm{e}^x$, $u^q = \mathrm{e}^y$, et on obtient que

$$tu = \exp\Bigl(\frac{1}{p}\, x + \frac{1}{q}\, y\Bigr) \le \frac{1}{p}\, \mathrm{e}^x + \frac{1}{q}\, \mathrm{e}^y \, .$$

Il en résulte que pour tout $s \in \Omega$

$$|f(s)g(s)| \le \frac{1}{p}\, |f(s)|^p + \frac{1}{q}\, |g(s)|^q,$$

ce qui montre que fg est intégrable, et que

$$\Bigl|\int fg\Bigr| \le \frac{1}{p} \int |f|^p + \frac{1}{q} \int |g|^q.$$

L'inégalité cherchée est positivement homogène par rapport à f et à g, donc il suffit de la démontrer lorsque $\|f\|_p = \|g\|_q = 1$. Mais dans ce cas, $\int |f|^p = 1$ et $\int |g|^q = 1$, donc l'inégalité précédente donne $|\int fg| \le 1/p + 1/q = 1$, ce qui est le résultat voulu.

Corollaire. *Soient* $p \in [1, +\infty[$ *et* q *tels que* $1/p + 1/q = 1$ *; si* $f \in \mathrm{L}_p(\Omega, \mathcal{A}, \mu)$,

$$\|f\|_p = \sup \{ \Big| \int_\Omega f g \, d\mu \Big| : \|g\|_q \le 1 \}.$$

Dans le cas $p = +\infty$, *le résultat reste vrai si tout ensemble* $\mathrm{B} \in \mathcal{A}$ *tel que* $\mu(\mathrm{B}) = +\infty$ *contient un* $\mathrm{A} \in \mathcal{A}$ *tel que* $0 < \mu(\mathrm{A}) < +\infty$.

Démonstration. L'inégalité de Hölder nous dit déjà que

$$\|f\|_p \ge \sup \{ \Big| \int_\Omega f g \, d\mu \Big| : \|g\|_q \le 1 \},$$

le problème est de montrer l'autre direction. On va voir qu'en fait le *maximum* est atteint pour une certaine fonction $g \in \mathrm{L}_q$, $\|g\|_q \le 1$, lorsque $1 \le p < +\infty$. Si $f = 0$, le résultat est évident, on supposera donc $f \ne 0$, et par homogénéité on peut se ramener à $\|f\|_p = 1$. En choisissant un représentant de la classe f, on se permettra de raisonner sur une “vraie” fonction mesurable f ; définissons une fonction mesurable g sur l'ensemble Ω en posant $g(s) = |f(s)|^p / f(s)$ sur l'ensemble mesurable $\mathrm{A} = \{f \ne 0\}$, et $g(s) = 0$ lorsque $s \notin \mathrm{A}$. Alors $|g(s)| = |f(s)|^{p-1}$ pour tout $s \in \mathrm{A}$; pour $p > 1$, on a $|g|^q = |f|^{q(p-1)} = |f|^p$, donc $\int |g|^q = \int |f|^p = 1$, soit encore $\|g\|_q = 1$; pour $p = 1$, $g(s)$ est de module 1 quand $s \in \mathrm{A}$, nul sinon, donc $\|g\|_\infty = 1$. D'autre part

$$\int_\Omega f g \, d\mu = \int_\mathrm{A} |f(s)|^p \, d\mu(s) = \int_\Omega |f|^p \, d\mu = 1 = \|f\|_p.$$

Dans le cas $p = +\infty$, le maximum n'est pas nécessairement atteint ; cependant, si on suppose $\|f\|_\infty = 1$, l'ensemble mesurable $\mathrm{B}_\varepsilon = \{|f| > 1 - \varepsilon\}$ est de mesure > 0 pour tout $\varepsilon \in]0, 1[$. On prendra $g(s) = \mu(\mathrm{A})^{-1} |f(s)| / f(s)$ si $s \in \mathrm{A}$, $g(s) = 0$ sinon, où A est un sous-ensemble de B_ε tel que $0 < \mu(\mathrm{A}) < +\infty$. Alors $\|g\|_1 = 1$ et $\int_\Omega f g \, d\mu > 1 - \varepsilon$.

3.3.4. Isométrie dans le dual de L_q

Soient $(\Omega, \mathcal{A}, \mu)$ un espace mesuré quelconque, et $p, q \in [1, +\infty]$ un couple d'exposants conjugués. Par Hölder, on peut définir pour toute $f \in \mathrm{L}_p(\Omega, \mathcal{A}, \mu)$ une forme linéaire continue $j_p(f)$ sur $\mathrm{L}_q(\Omega, \mathcal{A}, \mu)$ en posant

$$\forall g \in \mathrm{L}_q, \quad j_p(f)(g) = \int_\Omega f(t) g(t) \, d\mu(t).$$

Le résultat précédent signifie que l'application j_p qui associe à une fonction $f \in \mathrm{L}_p$ la forme linéaire $g \in \mathrm{L}_q \to \int f g \, d\mu$ est une isométrie linéaire de L_p dans le dual de L_q (ici, nous entendons par *dual* d'un espace normé X le *dual topologique*, formé de toutes les formes linéaires continues sur X). On sait en fait que

lorsque $1 < p < +\infty$, *l'application* j_p *précédente est une bijection isométrique de* $\mathrm{L}_p(\Omega, \mathcal{A}, \mu)$ *sur le dual de* $\mathrm{L}_q(\Omega, \mathcal{A}, \mu)$.

Le résultat précédent est valable pour toute mesure $\mu \ge 0$. En revanche,

lorsque $p = +\infty$, *et si la mesure* μ *est* σ*-finie, l'application* j_∞ *est une bijection isométrique de* $\mathrm{L}_\infty(\Omega, \mathcal{A}, \mu)$ *sur le dual de* $\mathrm{L}_1(\Omega, \mathcal{A}, \mu)$.

La condition σ-finie n'est pas une condition nécessaire pour la validité de l'énoncé précédent, mais une condition minimale est nécessaire. En effet, pour la mesure pathologique μ qui vaut toujours $+\infty$ sauf pour $\emptyset$, on a $\mathrm{L}_1 = \{0\}$, alors que $\mathrm{L}_\infty \ne \{0\}$ (il contient les fonctions bornées) ne peut pas être le dual de cet espace L_1.

3.3.5. Convolutions $\mathrm{L}_1 * \mathrm{L}_p$ *lorsque* $1 \le p < +\infty$

Si $f \in \mathrm{L}_1$ et $g \in \mathrm{L}_p$, on va voir que la convolution $f * g$ a un sens ; dans le cas de $\mathbb{R}^d$, ce résultat n'est pas contenu dans ce qui a été fait avec $\mathrm{L}_1 * \mathrm{L}_1$; en revanche, ce résultat n'est pas nouveau dans le cas périodique où la mesure est finie, car $\mathrm{L}_p(\mathbb{T}) \subset \mathrm{L}_1(\mathbb{T})$.

On commence avec $f \in \mathrm{L}_1$, $g \in \mathrm{L}_p$ positives et on écrit, avec valeur infinie admise pour les intégrales, l'inégalité de Minkowski pour la fonction $h(x,y) = f(y)g(x-y)$,

$$\int (f*g)(x)^p \, dx = \Bigl(\int \Bigl(\int f(y)g(x-y)\, dy \Bigr)^p dx \Bigr)^{1/p} \le \int \Bigl(\int f(y)^p g(x-y)^p \, dx \Bigr)^{1/p} dy =$$

$$= \int f(y) \Bigl(\int g(x-y)^p \, dx \Bigr)^{1/p} dy = \|f\|_1 \, \|g\|_p.$$

Si f et g sont à valeurs réelles ou complexes, on utilise d'abord le résultat obtenu pour $|f|$ et $|g|$, qui garantit que $(f*g)(x)$ existe pour presque tout x, et on obtient que $f*g \in \mathrm{L}_p$, avec la majoration de norme

$$\|f*g\|_p \le \|f\|_1 \, \|g\|_p$$

obtenue en majorant les modules d'intégrales par les intégrales de modules.

3.3.6. Convolution $\mathrm{L}_p * \mathrm{L}_q$ *lorsque* $1/p + 1/q = 1$

Ce cas est une conséquence immédiate de l'inégalité de Hölder et des propriétés d'invariance de la mesure de Lebesgue : si f est dans L_p, la fonction $y \to f(x-y)$ est aussi dans L_p et a la même norme que f, donc pour tout x

$$|(f*g)(x)| = \Bigl| \int f(x-y)g(y)\, dy \Bigr| \le \|f\|_p \, \|g\|_q.$$

On a donc $\mathrm{L}_p * \mathrm{L}_q \subset \mathrm{L}_\infty$ quand les exposants p, q sont conjugués, et $\|f*g\|_\infty \le \|f\|_p \, \|g\|_q$.

Les deux cas $\mathrm{L}_1 * \mathrm{L}_p \subset \mathrm{L}_p$ et $\mathrm{L}_p * \mathrm{L}_q \subset \mathrm{L}_\infty$ sont deux cas particuliers dans une échelle continue de résultats du même type :

si $1/p + 1/q = 1 + 1/r$, *avec* $p, q, r \ge 1$, *alors* $\mathrm{L}_p * \mathrm{L}_q \subset \mathrm{L}_r$ *et* $\|f*g\|_r \le \|f\|_p \, \|g\|_q$.

Ce résultat s'appelle *l'inégalité de convolution de Young*.

3.4. Régularisation et approximation

3.4.1. Continuité, opérateurs de translation

Commençons par les propriétés de continuité des convolées. Elles résultent du théorème de convergence dominée de Lebesgue.

Proposition. *Si* $f \in \mathrm{L}_1(\mathbb{R}^d)$ *et si* g *est continue bornée sur* $\mathbb{R}^d$*, la convolée* $f * g$ *est continue bornée sur* $\mathbb{R}^d$*.*

Preuve. Posons $\mathrm{M} = \|g\|_\infty$. Soient x un point de $\mathbb{R}^d$ et (x_n) une suite tendant vers x ; on obtient pour tout $y \in \mathbb{R}^d$ l'inégalité $|f(y)g(x_n - y)| \le \mathrm{M}\,|f(y)|$, qui donne la majoration par une fonction intégrable fixe, et $h_n(y) = f(y)g(x_n - y)$ tend simplement vers $h(y) = f(y)g(x-y)$. La convergence des intégrales, justifiée par le théorème de convergence dominée, donne $(f*g)(x_n) \to (f*g)(x)$.

Notons τ_v l'opérateur de translation par un vecteur $v \in \mathbb{R}^d$: si f est une fonction sur $\mathbb{R}^d$, on posera

$$\forall x \in \mathbb{R}^d, \quad (\tau_v f)(x) = f(x - v)\,;$$

on définit aussi les translations τ_v sur $\mathbb{T}^d$, par la même formule. On emploiera aussi la notation $f_v = \tau_v f$ pour gagner de la place. Il est important de remarquer que les translations commutent avec les convolutions, $(f * g)_v = f * g_v = f_v * g$.

Une fonction définie sur $\mathbb{R}^d$ est uniformément continue si et seulement si $\|f_v - f\|_\infty$ tend vers 0 lorsque la norme $|v|$ du vecteur v de la translation tend vers 0. Le *module de continuité* ω_f de la fonction f, défini par

$$\omega_f(\delta) = \sup\{|f(x) - f(x')| : |x - x'| \le \delta\}$$

est aussi égal à $\sup\{\|f_v - f\|_\infty : |v| \le \delta\}$: en effet, on a $|x - x'| \le \delta$ si et seulement si on peut écrire $x' = x - v$ avec $|v| \le \delta$. Alors

$$|f(x') - f(x)| = |(f_v - f)(x)| \le \|f_v - f\|_\infty.$$

Proposition. *Si $f \in \mathrm{L}_1(\mathbb{R}^d)$ et si g est uniformément continue bornée sur $\mathbb{R}^d$, $f * g$ est uniformément continue bornée.*

Preuve. On a $|g_v - g| \le \omega_g(|v|)$ et $(f * g)_v - (f * g) = f * (g_v - g)$, il suffit d'appliquer la plus simple des majorations : $\|f * (g_v - g)\|_\infty \le \|f\|_1 \, \|g_v - g\|_\infty$.

Remarque. On a obtenu plus précisément l'estimation

$$\omega_{f*g}(\delta) \le \|f\|_1 \, \omega_g(\delta)$$

pour les modules de continuité. De plus cette inégalité subsiste si g est uniformément continue, non nécessairement bornée, mais telle que $f * g$ soit définie en tout point : par exemple, si $f \in \mathrm{L}_1(\mathbb{R}^d)$ est nulle hors d'un borné, et g uniformément continue sur $\mathbb{R}^d$.

Continuité des translations sur $\mathrm{L}_p(\mathbb{R}^d)$, ou bien $\mathrm{L}_p(\mathbb{T}^d)$.

On considère la famille des translations τ_v, $v \in \mathbb{R}^d$, agissant sur $\mathrm{L}_p(\mathbb{R}^d)$; il est clair que les translations définissent des isométries de tous les espaces L_p (à cause de l'invariance de la mesure de Lebesgue par translation).

Proposition. *On suppose $1 \le p < +\infty$. Pour toute fonction $f \in \mathrm{L}_p(\mathbb{R}^d)$ on a*

$$\lim_{v \to 0} \|\tau_v f - f\|_p = 0.$$

Le même résultat est vrai pour $\mathbb{T}^d$.

Démonstration. Les translations τ_v sont des isométries de L_p, donc $\|\tau_v\| = 1$ pour tout $v \in \mathbb{R}^d$; le sous-espace vectoriel $\mathcal{K} = \mathcal{K}(\mathbb{R}^d)$ des fonctions continues à support compact est dense dans $\mathrm{L}_p(\mathbb{R}^d)$ et pour toute fonction $g \in \mathcal{K}$ on a $\|\tau_v g - g\|_p \to 0$ lorsque $v \to 0$ (continuité uniforme plus considérations de support). On en déduit que ce résultat est vrai pour toute fonction de L_p : en effet, on approche $f \in \mathrm{L}_p$ par $g \in \mathcal{K}$, disons $\|f - g\|_p < \varepsilon/3$; pour $|v| < t_0$ on aura $\|\tau_v g - g\|_p < \varepsilon/3$ d'après ce qui précède, et d'autre part $\|\tau_v f - \tau_v g\|_p = \|f - g\|_p < \varepsilon/3$, d'où le résultat $\|\tau_v f - f\|_p < \varepsilon$ par l'inégalité triangulaire.

Justifions plus précisément la convergence L_p dans le cas $g \in \mathcal{K}(\mathbb{R}^d)$. Supposons que le support de g soit contenu dans la boule $\mathrm{B}(0, \mathrm{R})$, et supposons $|v| \le 1$. Les deux fonctions g et $\tau_v g$ sont alors nulles en dehors de l'ensemble $\mathrm{A} = \mathrm{B}(0, \mathrm{R}{+}1)$, par conséquent

$$\|\tau_v g - g\|_p^p = \int_{\mathrm{A}} |g(x-v) - g(x)|^p \, dx \le \lambda(\mathrm{A})\, \omega_g(|v|)^p,$$

qui tend vers 0 quand $|v| \to 0$.

On vient d'utiliser un principe d'équicontinuité bien connu.

Lemme. *Si une suite* (T_n) *d'applications linéaires continues entre deux espaces normés* X *et* Y, *telle que* $\sup \|\mathrm{T}_n\| < +\infty$ *(c'est* l'équicontinuité *de la suite) tend simplement vers une application continue* $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ *sur un sous-ensemble dense* X_0 *de* X, *alors on a convergence de la suite* $(\mathrm{T}_n(x)) \subset \mathrm{Y}$ *vers* $\mathrm{T}(x)$ *pour tout* $x \in \mathrm{X}$.

On a vu que $\mathrm{L}_p * \mathrm{L}_q \subset \mathrm{L}_\infty$ quand p et q sont conjugués. De plus, la fonction $f * g$ est uniformément continue dans ce cas.

Proposition. *On suppose que* $1 \le p \le +\infty$, $1/p + 1/q = 1$; *si* $f \in \mathrm{L}_p(\mathbb{R}^d)$ *et* $g \in \mathrm{L}_q(\mathbb{R}^d)$, *la convolée* $f * g$ *est uniformément continue sur* $\mathbb{R}^d$ *(même résultat sur* $\mathbb{T}^d$*).*

Démonstration. L'un au moins de p ou q est fini, par exemple $p < +\infty$. On a

$$\|(f * g)_v - f * g\|_\infty = \|(f_v - f) * g\|_\infty \le \|f_v - f\|_p \, \|g\|_q$$

par Hölder, et $\|f_v - f\|_p \to 0$ d'après ce qui précède. Pour tout $\varepsilon > 0$, il existe donc $t_0 > 0$ tel que la condition $|v| < t_0$ entraîne $\|(f * g)_v - f * g\|_\infty < \varepsilon$.

Exercice. Soit A un borélien de mesure > 0 dans $\mathbb{R}^d$; montrer que l'ensemble $\mathrm{A} - \mathrm{A} = \{a - b : a, b \in \mathrm{A}\}$ est un voisinage de 0 dans $\mathbb{R}^d$.

3.4.2. Convolution et dérivées

En une variable, si $f \in \mathrm{L}_1$ et si $g \in \mathrm{L}_\infty$ a une dérivée bornée, alors $f * g$ est dérivable et $(f * g)' = f * g'$ (dérivation sous l'intégrale à la Lebesgue). On a des résultats analogues pour les dérivées partielles ; il existe un grand nombre de variantes de ces résultats. On va considérer un cas simple. Soit u une direction dans $\mathbb{R}^d$; notons D_u la dérivée dans la direction u,

$$(\mathrm{D}_u g)(x) = \lim_{t \to 0} \frac{g(x + tu) - g(x)}{t} \ ;$$

supposons que $g \in \mathrm{L}_\infty(\mathbb{R}^d)$, que $\mathrm{D}_u g$ existe sur $\mathbb{R}^d$ et soit une fonction bornée par M ; en appliquant le théorème des accroissements finis aux fonctions d'une variable $t \in \mathbb{R} \to g(x + tu)$, on voit que ces fonctions sont M-lipschitziennes ; alors

$$\frac{(f * g)(x + tu) - (f * g)(x)}{t} = \int f(y) \, \frac{g(x + tu - y) - g(x - y)}{t} \, dy.$$

En posant (pour x fixé) $\mathrm{G}(y, t) = t^{-1}\big(g(x + tu - y) - g(x - y)\big)$ on aura l'inégalité $|f(y)\mathrm{G}(y, t)| \le \mathrm{M}\, |f(y)|$ qui donne la majoration par une fonction intégrable fixe indépendante du paramètre t, pendant que $\mathrm{G}(y, t)$ tend simplement vers $\mathrm{D}_u g(x - y)$ quand

$t \to 0$, pour tout y ; le théorème de convergence dominée montre que $\mathrm{D}_u(f*g)(x)$ existe, et

$$\mathrm{D}_u(f*g)(x) = \int f(y)(\mathrm{D}_u g)(x-y)\,dy = (f*\mathrm{D}_u g)(x).$$

Si $\mathrm{D}_u g$ est de plus continue, la dérivée directionnelle $f*\mathrm{D}_u g$ de $f*g$ est aussi continue. Si toutes les dérivées partielles de g sont des fonctions continues bornées, il en résulte que $f*g$ admet des dérivées partielles continues. Autrement dit : si g est bornée et de classe C^1, à dérivées bornées sur $\mathbb{R}^d$, et si $f \in \mathrm{L}_1$, la convolution $f*g$ est de classe C^1 sur $\mathbb{R}^d$.

Si g est C^∞ à support compact, il résulte de tout ceci que $f*g$ est C^∞ lorsque $f \in \mathrm{L}_1(\mathbb{R}^d)$. En effet, toutes les dérivées partielles de g sont continues à support compact, donc bornées ; d'après ce qui précède toutes les dérivées partielles D_i, $i = 1, \ldots, d$ de la convolée $f*g$ existent, et sont données par les $f*\mathrm{D}_i g$; mais chacune de ces $f*\mathrm{D}_i g$ est à nouveau une convolution d'une fonction f de L_1 par une fonction $\mathrm{D}_i g$ qui est C^∞ à support compact. On voit de proche en proche que $f*g$ admet des dérivées partielles de tous les ordres, donc $f*g$ est C^∞. Pour tout multi-indice α, on a $\mathrm{D}^\alpha(f*g) = f*\mathrm{D}^\alpha g$.

Exercice. Montrer que la fonction f définie sur $\mathbb{R}$ par $f(x) = 0$ si $x \le 0$ et $f(x) = \exp(-1/x)$ si $x > 0$ est de classe C^∞ sur $\mathbb{R}$.

La classe $\mathcal{D}$ de Schwartz

C'est la classe $\mathcal{D}(\mathbb{R}^d)$ formée de toutes les fonctions sur $\mathbb{R}^d$, de classe C^∞ et à support compact ; il n'est pas totalement évident de donner un élément non nul de cet espace. À partir de la fonction f de l'exercice précédent, on peut construire des fonctions C^∞ à support compact, par exemple la fonction φ définie par

$$\forall x \in \mathbb{R}, \quad \varphi(x) = f\bigl(2(1+x)\bigr)\, f\bigl(2(1-x)\bigr).$$

Cette fonction est à support dans $[-1,1]$, et pour $|x| < 1$ on a

$$\varphi(x) = \exp\Bigl(-\frac{1}{2(1+x)} - \frac{1}{2(1-x)}\Bigr) = \exp\Bigl(-\frac{1}{1-x^2}\Bigr).$$

À partir de cette fonction C^∞ à support compact, ≥ 0 et non identiquement nulle, on construit des exemples ψ en toute dimension d, par exemple

$$\psi(x_1, \ldots, x_d) = \prod_{j=1}^d \varphi(x_j) \quad \text{ou bien} \quad \varphi\Bigl(\bigl(\sum_{j=1}^d x_j^2\bigr)^{1/2}\Bigr).$$

En utilisant les changements d'échelle définis par

$$\psi_a(x) = \psi(ax),$$

et en prenant $a > 0$ grand, on pourra trouver des fonctions dans $\mathcal{D}(\mathbb{R}^d)$, non identiquement nulles ≥ 0, et avec un support dans une boule $\mathrm{B}(0,\varepsilon)$ de rayon arbitrairement petit. Retenons le fait essentiel suivant :

si $\varphi \in \mathcal{D}(\mathbb{R}^d)$ et si $f \in \mathrm{L}_1(\mathbb{R}^d)$, la convolée $f*\varphi$ est C^∞ et bornée sur $\mathbb{R}^d$.

Fonctions plateaux

Lemme. *Si* K *est un compact non vide de* $\mathbb{R}^d$ *et* U *un ouvert contenant* K, *il existe une fonction* $\theta \in \mathcal{D}(\mathbb{R}^d)$ *telle que* $0 \le \theta \le 1$, $\theta = 1$ *sur* K *et* $\theta = 0$ *hors de* U.

Preuve. On pose

$$\varepsilon = \min\{\mathrm{dist}(x, \mathrm{U}^c) : x \in \mathrm{K}\} = \min\{\|x - x'\| : x \in \mathrm{K},\ x' \notin \mathrm{U}\} > 0,$$

puis

$$\mathrm{V} = \{x \in \mathbb{R}^d : \mathrm{dist}(x, \mathrm{K}) < \varepsilon/2\},$$

et on choisit une fonction $\varphi \in \mathcal{D}(\mathbb{R}^d)$ à support dans la boule $\mathrm{B}(0, \varepsilon/2)$, $\varphi \ge 0$ et d'intégrale égale à 1. On va montrer que $\theta = \mathbf{1}_{\mathrm{V}} * \varphi$ convient. On a pour x fixé

$$\theta(x) = \int_{\mathbb{R}^d} \mathbf{1}_{\mathrm{V}}(x - y)\varphi(y)\, dy\,;$$

il est clair sur cette formule que $0 \le \theta \le 1$. Supposons que $x' \notin \mathrm{U}$, et montrons que la fonction $h : y \to \mathbf{1}_{\mathrm{V}}(x' - y)\varphi(y)$ est identiquement nulle, ce qui donnera bien sûr

$$\theta(x') = \int_{\mathbb{R}^d} \mathbf{1}_{\mathrm{V}}(x' - y)\varphi(y)\, dy = 0\,;$$

si $|y| \ge \varepsilon/2$, on a $\varphi(y) = 0$, donc $h(y) = 0$; puisque $x' \notin \mathrm{U}$, on a $\mathrm{dist}(x', \mathrm{K}) \ge \varepsilon$. Si $|y| < \varepsilon/2$, on aura $\mathrm{dist}(x' - y, \mathrm{K}) \ge \varepsilon - \varepsilon/2 = \varepsilon/2$ par l'inégalité triangulaire, donc $x' - y \notin \mathrm{V}$ et $\mathbf{1}_{\mathrm{V}}(x' - y) = 0 = h(y)$. Ainsi, $h(y) = 0$ pour tout y.

Si $x \in \mathrm{K}$, on va montrer que $y \to \mathbf{1}_{\mathrm{V}}(x - y)\varphi(y)$ est égale à la fonction $y \to \varphi(y)$, ce qui donnera

$$\theta(x) = \int_{\mathbb{R}^d} \mathbf{1}_{\mathrm{V}}(x - y)\varphi(y)\, dy = \int_{\mathbb{R}^d} \varphi(y)\, dy = 1\,;$$

si $\varphi(y) = 0$, les deux fonctions sont nulles, donc égales au point y ; sinon, on aura $|y| < \varepsilon/2$, $\mathrm{dist}(x - y, \mathrm{K}) \le |y| < \varepsilon/2$, donc $x - y \in \mathrm{V}$ et $\mathbf{1}_{\mathrm{V}}(x - y) = 1$.

3.4.3. Approximations de l'unité

Une *approximation de l'unité* est une suite (φ_k) dans $\mathrm{L}_1(\mathbb{R}^d)$ ou $\mathrm{L}_1(\mathbb{T}^d)$ telle que

$$(a) \qquad \sup_k \int |\varphi_k| < +\infty,$$

$$(b) \qquad \int \varphi_k = 1 \ \text{ pour tout } k \ge 1,$$

$$(c) \qquad \forall \alpha > 0, \quad \lim_k \int_{\{y : d(y,0) > \alpha\}} |\varphi_k| = 0.$$

On rappelle que $d(y, 0)$ est égal à la norme $|y|$ dans le cas $\mathbb{R}^d$, et au minimum des $|y'|$ pour y' appartenant à la classe y, dans le cas du quotient $\mathbb{R}^d/(2\pi\mathbb{Z})^d$.

Théorème. *On se place dans $\mathbb{R}^d$ ou $\mathbb{T}^d$. Pour toute fonction uniformément continue bornée g, la suite $(\varphi_k * g)$ converge uniformément vers g. Pour toute fonction $g \in \mathrm{L}_p$, $1 \le p < +\infty$, la suite $(\varphi_k * g)$ converge vers g dans L_p.*

Démonstration. Soit g une fonction uniformément continue et bornée sur $\mathbb{R}^d$; considérons

$$d_k(x) = (g * \varphi_k)(x) - g(x) = \int \bigl(g(x-y) - g(x)\bigr)\, \varphi_k(y)\, dy\,;$$

il faut montrer que $d_k(x)$ tend vers 0, uniformément en $x \in \mathbb{R}^d$. On a pour tout x

$$|d_k(x)| \le \int \bigl|g(x-y) - g(x)\bigr|\, |\varphi_k(y)|\, dy \le \int \|\tau_y g - g\|_\infty\, |\varphi_k(y)|\, dy.$$

Posons $\mathrm{M} = \sup_k \|\varphi_k\|_1$. Puisque g est uniformément continue, on peut trouver $\alpha > 0$ tel que $\|\tau_y g - g\|_\infty < \varepsilon/(2\mathrm{M})$ pour tout $y \in \mathbb{R}^d$ tel que $|y| < \alpha$. On découpe alors l'intégrale ci-dessus en intégrale sur $\mathrm{B} = \mathrm{B}(0, \alpha)$ et complémentaire et on obtient

$$\begin{aligned}\|d_k\|_\infty &\le \int_{\mathrm{B}} \|\tau_y g - g\|_\infty\, |\varphi_k(y)|\, dy + \int_{\mathrm{B}^c} \|\tau_y g - g\|_\infty\, |\varphi_k(y)|\, dy \\ &\le \mathrm{M}\, \frac{\varepsilon}{2\mathrm{M}} + 2\|g\|_\infty \int_{\mathrm{B}^c} |\varphi_k(y)|\, dy.\end{aligned}$$

Pour k assez grand on obtiendra avec (c) que $\|d_k\|_\infty \le \varepsilon/2 + \varepsilon/2$; on obtient ainsi la convergence uniforme de $\varphi_k * g$ vers g.

Démontrons la deuxième partie. On doit montrer que $\|d_k\|_p \to 0$; on utilise l'inégalité de Minkowski pour la mesure $|\varphi_k(y)|\, dy$ pour obtenir que

$$\|d_k\|_p = \Bigl(\int \Bigl| \int \bigl(g(x-y) - g(x)\bigr)\, \varphi_k(y)\, dy \Bigr|^p dx\Bigr)^{1/p} \le$$

$$\le \int \Bigl(\int \bigl|g(x-y) - g(x)\bigr|^p\, dx\Bigr)^{1/p} |\varphi_k(y)|\, dy = \int \|\tau_y g - g\|_p\, |\varphi_k(y)|\, dy.$$

À partir de ce point la démonstration est identique au cas précédent, si on rappelle qu'on peut choisir $\alpha > 0$ tel que $\|\tau_y g - g\|_p < \varepsilon/(2\mathrm{M})$ pour tout vecteur y tel que $|y| < \alpha$. On obtient alors par un découpage identique à celui de la première partie, où $\mathrm{B} = \mathrm{B}(0, \alpha)$

$$\begin{aligned}\|d_k\|_p &\le \int_{\mathrm{B}} \|\tau_y g - g\|_p\, |\varphi_k(y)|\, dy + \int_{\mathrm{B}^c} \|\tau_y g - g\|_p\, |\varphi_k(y)|\, dy \\ &\le \mathrm{M}\, \frac{\varepsilon}{2\mathrm{M}} + 2\|g\|_p \int_{\mathrm{B}^c} |\varphi_k(y)|\, dy.\end{aligned}$$

Pour k assez grand on obtiendra avec (c) que $\|d_k\|_p \le \varepsilon$; on obtient ainsi la convergence L_p de $\varphi_k * g$ vers g.

Une méthode utile pour construire des approximations de l'unité est la suivante. On se donne une fonction φ intégrable sur $\mathbb{R}^d$ telle que $\int \varphi(x)\,dx = 1$ et on considère la suite (φ_k) définie par

$$\varphi_k(x) = k^d \varphi(kx)$$

pour k entier ≥ 1. Par le changement de variable $y = kx$ on obtient les propriétés *a, b, c* voulues. En effet,

$$\int_{\mathbb{R}^d} |\varphi_k(x)|\,dx = \int_{\mathbb{R}^d} |\varphi(kx)|\,k^d\,dx = \int_{\mathbb{R}^d} |\varphi(y)|\,dy,$$

$$\int_{\mathbb{R}^d} \varphi_k(x)\,dx = \int_{\mathbb{R}^d} \varphi(kx)\,k^d\,dx = \int_{\mathbb{R}^d} \varphi(y)\,dy = 1$$

et

$$\int_{|x|\geq\alpha} |\varphi_k(x)|\,dx = \int_{|y|\geq k\alpha} |\varphi(y)|\,dy$$

tend vers 0 par convergence dominée.

Densité de la classe $\mathcal{D}(\mathbb{R}^d)$

À partir d'une $\varphi \in \mathcal{D}(\mathbb{R}^d)$, positive et d'intégrale 1, on peut construire par la méthode précédente une approximation de l'unité (φ_k) formée de fonctions de $\mathcal{D}(\mathbb{R}^d)$. Les fonctions continues à support compact peuvent être approchées uniformément par des fonctions C^∞ qui ont “presque” le même support : si g est continue à support compact, la suite $g * \varphi_k$ converge uniformément vers g, et elle est formée de fonctions de $\mathcal{D}(\mathbb{R}^d)$ qui ont un support à peine plus grand que celui de g.

Pour toute fonction $g \in L_p$, la suite $\varphi_k * g$ est formée de fonctions C^∞ et converge vers g dans L_p. Mentionnons une dernière méthode d'approximation, la méthode de troncature ; on se donne une fonction $\psi \in \mathcal{D}(\mathbb{R}^d)$, égale à 1 dans un voisinage de 0, et on considère la suite des fonctions $x \to \psi(x/k)$, qui tend vers 1 (uniformément sur tout compact). En associant régularisation et troncature, on voit que $\mathcal{D}(\mathbb{R}^d)$ est dense dans tous les $L_p(\mathbb{R}^d)$, pour $1 \leq p < +\infty$.

Une approximation de l'unité utile

Lemme. *Si φ est une fonction continue ≥ 0 à support compact sur $\mathbb{R}^d$ (ou bien une fonction continue sur $\mathbb{R}/2\pi\mathbb{Z}$), telle que*

$$x \neq 0 \Rightarrow \varphi(x) < \varphi(0),$$

la suite (φ_n) construite à partir des puissances φ^n de la fonction φ par la formule

$$\varphi_n(x) = \Big(\int \varphi(y)^n\,dy \Big)^{-1} \varphi(x)^n$$

est une approximation de l'unité.

Preuve. On a $\int \varphi^n > 0$ puisque φ est continue ≥ 0 et $\varphi(0) > 0$. Il est clair que $\int \varphi_n = 1$, et $\varphi_n \geq 0$, ce qui donne immédiatement que la suite (φ_n) vérifie les conditions (a) et (b) pour une approximation de l'unité. Soit $\alpha > 0$ fixé et soit (t_k) une suite croissante de nombres tendant vers $\varphi(0)$, tels que $0 < t_k < \varphi(0)$; on note que

$$A_k = \{x : d(x,0) \geq \alpha \text{ et } \varphi(x) \geq t_k\}$$

est compact, puisque c'est un fermé contenu dans le support de φ, et la suite décroissante des compacts (A_k) a pour intersection

$$\bigcap_k \mathrm{A}_k = \{x : d(x,0) \geq \alpha \text{ et } \ \varphi(x) \geq \varphi(0)\} = \emptyset$$

d'après l'hypothèse, donc l'un de ces compacts A_k est vide. En posant $\delta = t_k$ pour k assez grand, on aura donc $\mathrm{A}_k = \emptyset$, c'est-à-dire que $\varphi < \delta$ sur le complémentaire de la boule $\mathrm{B}(0,\alpha)$. En désignant par M le volume du support de φ, on aura donc

$$\int_{\{d(x,0)\geq\alpha\}} \varphi(x)^n \, dx \leq \mathrm{M}\, \delta^n \,;$$

d'un autre côté, choisissons δ_1 tel que $\delta < \delta_1 < 1$; l'ouvert $\mathrm{V} = \{\varphi > \delta_1\}$ n'est pas vide, donc $|\mathrm{V}| > 0$, où $|\mathrm{V}|$ désigne le volume de l'ouvert V ; on a de plus

$$c_n^{-1} = \int \varphi(x)^n \, dx \geq \int_{\mathrm{V}} \varphi(x)^n \, dx \geq |\mathrm{V}|\, \delta_1^n.$$

On en déduit pour $\varphi_n = c_n \varphi^n$ l'inégalité

$$\int_{\{d(x,0)\geq\alpha\}} \varphi_n(x) \, dx = c_n \int_{\{d(x,0)\geq\alpha\}} \varphi(x)^n \, dx \leq \mathrm{M}|\mathrm{V}|^{-1} \Bigl(\frac{\delta}{\delta_1}\Bigr)^n$$

qui tend vers 0 quand $n \to +\infty$.

Deux exemples

1. Considérons la fonction 2π-périodique $f(t) = 1 + \cos(t)$; il s'agit d'une fonction ≥ 0, qui atteint sur l'intervalle $[-\pi,\pi]$ un unique maximum au point 0. D'après le lemme précédent, on obtient une approximation de l'unité dans $\mathrm{L}_1(\mathbb{T})$ en posant pour tout $n \geq 0$

$$\mathrm{I}_n = \int_{-\pi}^{\pi} (1+\cos(t))^n \, \frac{dt}{2\pi}$$

et $\varphi_n(t) = \mathrm{I}_n^{-1}(1+\cos(t))^n$, qui est bien ≥ 0 d'intégrale 1 pour la probabilité $dt/(2\pi)$. L'intérêt de cet exemple est que les fonctions de la suite sont des polynômes trigonométriques.

2. Si on considère sur $\mathbb{R}^d$ la fonction

$$\varphi(x) = \max((1-|x|^2/4), 0)$$

on aura le même phénomène. L'intérêt ici est que les fonctions (φ_n) sont polynomiales sur la boule de rayon 2.

3.4.4. Théorème d'approximation de Weierstrass

Lemme. *Si f est une fonction réelle continue sur un compact* K *contenu dans la boule unité ouverte de $\mathbb{R}^d$, on peut trouver une fonction g continue sur $\mathbb{R}^d$, à support dans la boule unité et telle que $|f-g| < \varepsilon$ sur* K.

Preuve. On commence par trouver une fonction g_1 continue sur $\mathbb{R}^d$, et proche de f sur le compact K ; soit $\varepsilon > 0$, soit $\delta > 0$ tel que $|f(y) - f(y')| < \varepsilon$ pour tout couple (y, y') d'éléments de K tels que $|y - y'| < \delta$; choisissons ensuite N assez grand pour que $\mathrm{N}\delta > 2\|f\|_\infty$, et posons

$$g_1(x) = \min_{y \in \mathrm{K}} \{f(y) + \mathrm{N}|x - y|\}.$$

Cette fonction g_1 est lipschitzienne de constante N sur $\mathbb{R}^d$, comme inf de la famille des fonctions $x \to \mathrm{N}|x - y| + f(y)$ dépendant du paramètre $y \in \mathrm{K}$, qui sont toutes N-lipschitziennes sur $\mathbb{R}^d$. Fixons $y_0 \in \mathrm{K}$ quelconque ; on a évidemment $f(y_0) \geq g_1(y_0)$ (en prenant $y = y_0$ comme constituant du min). Si $|y - y_0| \geq \delta$,

$$f(y) + \mathrm{N}|y_0 - y| \geq -\|f\|_\infty + \mathrm{N}\delta > \|f\|_\infty \geq f(y_0),$$

et si $|y - y_0| < \delta$

$$f(y) + \mathrm{N}|y_0 - y| \geq f(y) \geq f(y_0) - \varepsilon\,;$$

il en résulte que $g_1(y_0) \geq f(y_0) - \varepsilon$. La fonction g_1 approche donc f sur K à moins de ε. On choisit ensuite une fonction plateau θ égale à 1 sur K, nulle hors de la boule unité, et on pose $g = \theta g_1$. La fonction g est maintenant à support compact, toujours continue, et elle est égale à g_1 sur K, donc $|g - f| \leq \varepsilon$ sur K.

Remarque. En fait, on peut *prolonger* la fonction continue f définie sur K, en une fonction continue sur $\mathbb{R}^d$, que l'on peut choisir à support compact si on veut (traiter cette remarque en exercice).

Théorème. *Si f est une fonction réelle continue sur un compact* K *de $\mathbb{R}^d$, on peut trouver une fonction* P *polynomiale sur $\mathbb{R}^d$ telle que $|f - \mathrm{P}| < \varepsilon$ sur* K.

Preuve. On peut supposer que le compact K est contenu dans la boule unité ouverte, et remplacer f par g, continue sur $\mathbb{R}^d$, à support dans la boule unité, et qui soit telle que $|f - g| < \varepsilon/2$ sur K. On utilise l'approximation de l'unité (φ_n) fabriquée avec les puissances $((1 - |x|^2/4)_+)^n$. On sait que $g * \varphi_n$ converge uniformément vers g sur $\mathbb{R}^d$, donc aussi sur K. On écrit

$$g_n(x) = \int g(t)\varphi_n(x - t)\,dt.$$

Fixons x tel que $|x| \leq 1$. Puisque $|x| \leq 1$ et que le support de g est contenu dans la boule unité, l'expression sous l'intégrale n'est non nulle que si $|t| \leq 1$, donc $|x - t| \leq 2$, ce qui montre que g_n coïncide sur la boule unité avec la fonction P_n définie par

$$\mathrm{P}_n(x) = \int g(t) c_n (1 - |x - t|^2)^n\,dt,$$

qui est une fonction polynomiale, comme on le vérifie facilement en développant la puissance par la formule du binôme : on trouve ainsi que P_n est combinaison linéaire de fonctions de la forme

$$\mathrm{Q}(x) = \int g(t)\,|x - t|^{2k}\,dt = \int g(t)\,\Bigl(\sum_{j=1}^{d} (x_j - t_j)^2\Bigr)^k\,dt.$$

Continuons l'explication dans le cas $d = 2$ pour simplifier ; par la formule du binôme, chaque puissance $\bigl((x_1-t_1)^2+(x_2-t_2)^2\bigr)^k$ est une combinaison linéaire d'expressions qui sont de la forme $(x_1-t_1)^{2n_1}(x_2-t_2)^{2n_2}$, avec $n_1, n_2 \in \mathbb{N}$, elles-mêmes combinaisons de monômes $t_1^{k_1}t_2^{k_2}x_1^{\ell_1}x_2^{\ell_2}$; finalement, on décompose P_n en combinaison linéaire de fonctions monomiales de deux variables

$$\mathrm{R}(x) = \int g(t)(t_1^{k_1}t_2^{k_2}x_1^{\ell_1}x_2^{\ell_2})\,dt_1\,dt_2 = \Bigl(\int g(t)t_1^{k_1}t_2^{k_2}\,dt_1\,dt_2\Bigr)\,x_1^{\ell_1}x_2^{\ell_2} = a_{\ell_1,\ell_2}\,x_1^{\ell_1}x_2^{\ell_2}.$$

Séries de Fourier

On va travailler avec des séries de Fourier complexes sur l'intervalle $[0, 2\pi]$, muni de la mesure normalisée $dx/(2\pi)$. Si $f \in \mathrm{L}_1(\mathbb{T})$, on définit les coefficients de Fourier (complexes) de la fonction f par

$$\forall n \in \mathbb{Z}, \quad c_n(f) = \int_0^{2\pi} f(x)\,\mathrm{e}^{-inx}\,\frac{dx}{2\pi}.$$

Il est clair que $|c_n(f)| \le \|f\|_1$ pour tout n. Pour tout $n \in \mathbb{Z}$, posons $e_n(x) = \mathrm{e}^{inx}$; on a $f * e_n = c_n(f)e_n$ pour tout n ; ceci entraîne que pour tout polynôme trigonométrique $\mathrm{K} = \sum_{k=\mathrm{M}}^{\mathrm{N}} a_k e_k$ (avec $\mathrm{M} \le \mathrm{N}$ et $\mathrm{M}, \mathrm{N} \in \mathbb{Z}$), on a

$$\mathrm{K} * f = \sum_{k=\mathrm{M}}^{\mathrm{N}} a_k c_k(f) e_k,$$

ce qui montre que la convolution de $f \in \mathrm{L}_1(\mathbb{T})$ par un polynôme trigonométrique donne un polynôme trigonométrique $\mathrm{K} * f$.

L'approximation de l'unité (φ_n) construite avec les puissances de $1 + \cos(t)$ permet de montrer que les polynômes trigonométriques sont denses dans $\mathrm{C}_{\mathrm{per}}(0, 2\pi)$, donc aussi dans tous les L_p. On vérifie en effet que

$$\bigl(1 + \cos(t)\bigr)^n = \Bigl(\frac{1}{2}\,\mathrm{e}^{-it} + 1 + \frac{1}{2}\,\mathrm{e}^{it}\Bigr)^n$$

est un polynôme trigonométrique pour tout $n \ge 0$. Les résultats généraux sur la convolution donnent la proposition suivante.

Proposition. *Les polynômes trigonométriques sont denses dans* $\mathrm{C}(\mathbb{T})$ *(fonctions continues 2π-périodiques) ; pour tout p tel que $1 \le p < +\infty$, les polynômes trigonométriques sont denses dans* $\mathrm{L}_p(0, 2\pi)$.