2007-12-08 by BM

3. Convolution, inégalités, approximation et régularisation

Contenu du chapitre

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

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

Le résultat est à valeurs dans $[0, +\infty]$; la fonction $f * g$ est mesurable d'après les résultats de la section 2.2 (théorème de Fubini). Les propriétés d'invariance de la mesure de Lebesgue entraînent que

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

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 $\int (f * g)(x) \, \mathrm{d}x < +\infty$. 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

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

L'invariance par translation de la mesure de Lebesgue donne $\int f(x-y) \,\mathrm{d}x = \int f(x) \,\mathrm{d}x$ pour tout y, de sorte que

$$\int_{\mathbb{R}^d} \Bigl(\int_{\mathbb{R}^d} f(x-y) g(y) \,\mathrm{d}x \Bigr) \,\mathrm{d}y = \Bigl(\int_{\mathbb{R}^d} f(x) \,\mathrm{d}x \Bigr) \Bigl(\int_{\mathbb{R}^d} g(y) \,\mathrm{d}y \Bigr) < +\infty. \tag{1}$$

On vient de montrer le résultat qui suit.

Lemme 3.1.1. *Si f et g sont deux fonctions mesurables positives sur $\mathbb{R}^d$, la fonction $f * g$ est mesurable à valeurs dans $[0, +\infty]$, et*

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

Supposons maintenant que f et g soient intégrables, réelles ou complexes. On vient de voir dans le lemme précédent que

$$\int_{\mathbb{R}^d} \Bigl(\int_{\mathbb{R}^d} |f(x-y)| \, |g(y)| \,\mathrm{d}y \Bigr) \,\mathrm{d}x = \Bigl(\int_{\mathbb{R}^d} |f(x)| \,\mathrm{d}x \Bigr) \Bigl(\int_{\mathbb{R}^d} |g(y)| \,\mathrm{d}y \Bigr) < +\infty,$$

ce qui implique que $\int |f(x-y)| \, |g(y)| \,\mathrm{d}y$ est fini pour presque tout x ; ceci 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

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

et on obtient ainsi une fonction mesurable *définie presque-partout* (c'est une partie de l'énoncé du théorème de Fubini 2.2.2) ; les calculs obtenus en majorant par les modules montrent que $f * g$ est intégrable ; plus précisément on obtient l'inégalité de normes

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

Si on reprend avec Fubini le calcul qui menait à la formule (1), on obtient pour deux fonctions intégrables réelles ou complexes f et g que

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

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 $f_1(x) = f(x)\,\mathrm{e}^{-\mathrm{i}\,x\cdot t}$ et $g_1(x) = g(x)\,\mathrm{e}^{-\mathrm{i}\,x\cdot t}$. On obtient

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

et

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

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 supplémentaires 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 3.1.2. Si f est une fonction borélienne sur $\mathbb{R}^d$, disons que f *est supportée* par un sous-ensemble A de $\mathbb{R}^d$ si f est nulle en dehors de A,

$$f(x) \neq 0 \Rightarrow x \in \mathrm{A}$$

(ne pas confondre avec la notion de support d'une fonction continue, qui est par définition *l'adhérence* de l'ensemble $\{f \neq 0\}$). Si $f, g \in \mathrm{L}^1(\mathbb{R}^d)$, si f est supportée par le borélien A et g par le borélien B, montrer que la convolution $f * g$ est supportée par la *somme de Minkowski* de A et de B, définie par

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

(on n'essaiera pas de prouver que $\mathrm{A} + \mathrm{B}$ est borélien ; on pourra cependant vérifier, bien que la question n'ait pas été posée, qu'il est Lebesgue-mesurable).

Autres cadres pour la convolution : le cadre périodique

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

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

pour tout a. 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 fourni par le groupe multiplicatif U des nombres complexes de module 1. Si θ est un nombre réel, il est clair que $\mathrm{e}^{\mathrm{i}\theta} \in \mathrm{U}$ ne dépend que de la classe $\widehat{\theta}$ de θ dans $\mathbb{R}/2\pi\mathbb{Z}$, et l'application $\widehat{\theta} \to \mathrm{e}^{\mathrm{i}\theta}$ 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,

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

C'est une formule de ce type qu'il faut utiliser pour définir la convolution sur certains groupes non commutatifs (comme $\mathrm{O}(n)$ par exemple), où la loi est toujours notée multiplicativement.

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

$$(f \underset{\mathrm{per}}{*} g)(\theta) = \int_0^{2\pi} f(\theta - t)g(t)\,\frac{\mathrm{d}t}{2\pi}\cdot$$

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 dira qu'une fonction 2π-périodique f sur $\mathbb{R}$ est dans $\mathrm{L}^p(\mathbb{T})$ si sa restriction à une période quelconque $[a, a + 2\pi]$ est dans $\mathrm{L}^p([a, a + 2\pi], \mathrm{d}x/(2\pi))$. 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 $\mathbb{R}^d/(2\pi\mathbb{Z})^d$, la notation $d(x, y)$ 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. Si on pose $e_n(t) = \mathrm{e}^{\mathrm{i}nt}$, pour $n \in \mathbb{Z}$, on a

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

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

$$c_n(f) = \int_0^{2\pi} f(t)\,\mathrm{e}^{-\mathrm{i}nt}\,\frac{\mathrm{d}t}{2\pi}\cdot$$

On peut remarquer que la convolution périodique peut se ramener à la convolution sur $\mathbb{R}$ de la façon suivante : si f et g sont 2π-périodiques sur $\mathbb{R}$ et appartiennent à $\mathrm{L}^1(\mathbb{T})$, posons $f_1 = \mathbf{1}_{(0,2\pi)} f$; on voit que

$$(f_1 * g)(x) = \int_{\mathbb{R}} f_1(y)g(x - y)\,\mathrm{d}y = \int_0^{2\pi} f(y)g(x - y)\,\mathrm{d}y = 2\pi\,(f \underset{\mathrm{per}}{*} g)(x).$$

Ainsi, plusieurs des propriétés de la convolution périodique ne nécessitent pas de preuve séparée de celle qu'on donnera pour $\mathbb{R}$.

Pour finir, il existe un dernier cadre de convolution qui se révèle intéressant, mais où on ne rencontre pas de finesses de théorie de l'intégration : c'est le cas de la mesure de comptage sur $\mathbb{Z}$. Si $\mathbf{a} = (a_n)_{n\in\mathbb{Z}}$ et $\mathbf{b} = (b_n)_{n\in\mathbb{Z}}$ sont deux éléments de $\ell^1(\mathbb{Z})$, leur convolution est définie par

$$(\mathbf{a} * \mathbf{b})_n = \sum_{k\in\mathbb{Z}} a_{n-k} b_k,$$

qui est l'analogue exact pour la mesure de comptage de la formule utilisée sur $\mathbb{R}$.

3.2. Densité des fonctions continues

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 3.2.1. *Les fonctions continues à support compact sont denses dans* $\mathrm{L}^p(\mathbb{R}^d)$, *pour tout p tel que* $1 \leq p < +\infty$.

Il est évident que le résultat ne peut pas être vrai dans L^∞ : une suite de Cauchy dans L^∞, formée de fonctions continues, est une suite de Cauchy uniforme, et la limite sera toujours continue. Le cas de tous les espaces $\mathrm{L}^p(\mathbb{R}^d)$, pour $1 \leq p < +\infty$ est essentiellement le même et on écrira la preuve dans le cas $p = 1$.

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 $|f_n| \leq |f|$ et $f_n \to f$ simplement (proposition 1.1.2) ; 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 fonctions 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 pour n assez grand on aura

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

Finalement, il suffit pour montrer le théorème d'approcher, pour la norme de $\mathrm{L}^1(\mathbb{R}^d)$, toute indicatrice de borélien borné par une fonction continue à support compact. Si A est un borélien borné, on peut le placer dans un ouvert borné de la forme $\mathrm{W} =]-a, a[^d$, qui est de mesure finie pour la mesure de Lebesgue. D'après le théorème d'approximation 1.2.4 appliqué à $\mathrm{X} = \mathbb{R}^d$, il existe une fonction continue φ sur $\mathbb{R}^d$, nulle hors de W et telle que

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

ce qui termine la preuve, car le support de φ est contenu dans $[-a, a]^d$, qui est compact.

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 constater qu'une fonction continue à support compact peut être approchée dans L^1 par des fonctions en escalier (utiliser la continuité uniforme ; procéder ainsi, c'est utiliser la stratégie mathématique qui consiste à toujours commencer par raccrocher la casserole au mur, quand on ne sait faire cuire l'œuf qu'à partir de cet état initial ; sinon, on peut vérifier que les boréliens de $[0,1]$ approchables dans L^1 par des fonctions en escalier forment une tribu qui contient les intervalles).

3.3. Inégalités

3.3.a. *Inégalité de Jensen*

Proposition 3.3.1 : 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*

$$\varphi\Bigl(\int_\Omega f(\omega)\,\mathrm{d}\mu(\omega)\Bigr) \le \int_\Omega \varphi(f(\omega))\,\mathrm{d}\mu(\omega).$$

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 $\varphi\bigl((1-t)u + tv\bigr) \le (1-t)\varphi(u) + t\varphi(v)$.

Démonstration. Posons $m = \int_\Omega f(\omega)\,\mathrm{d}\mu(\omega) \in \mathbb{R}$ (c'est la *valeur moyenne* de la fonction 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, compte tenu du fait que $\int \mathrm{d}\mu = \mu(\Omega) = 1$,

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

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

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

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

3.3.b. *Inégalité de Minkowski*

Lemme. *Si ν est une mesure σ-finie non nulle 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 $\int_{\mathrm{Y}} k \,\mathrm{d}\nu = 1$.*

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 $\int_{\mathrm{Y}} k(y) \,\mathrm{d}\nu(y) = 1$.

Théorème 3.3.2. *Si $f(x,y)$ est mesurable sur un espace 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 $p \geq 1$, on a*

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

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 $\mathrm{F} = \sum_{j=1}^n f_{y_j} \in \mathrm{L}^p(\mathrm{X}, \mu)$

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

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

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

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 y 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 $f \geq 0$. Posons

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

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

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

On a $\int_{\mathrm{Y}} g(y) \,\mathrm{d}\nu(y) = \mathrm{M} < 1$; l'ensemble N des $y \in \mathrm{Y}$ tels que $g(y) = +\infty$ est donc ν-négligeable, et on peut se ramener à $\mathrm{N} = \emptyset$ en modifiant f sur $\mathrm{X} \times \mathrm{N}$: si on pose $\widetilde{f}(x,y) = 0$ lorsque $y \in \mathrm{N}$ et $\widetilde{f}(x,y) = f(x,y)$ quand $y \notin \mathrm{N}$, la fonction $\widetilde{g}$ correspondante est finie partout, et les deux membres de l'inégalité de Minkowski sont inchangés. On supposera donc que $g(y) < +\infty$ pour tout $y \in \mathrm{Y}$.

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 $h > g \geq 0$ telle que $\int h(y)\, \mathrm{d}\nu(y) = 1$ (prendre $h = g + (1 - \mathrm{M})k$, avec k donnée par le lemme). Écrivons

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

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

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

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

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

Par homogénéité, on voit qu'on a prouvé que

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

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

Remarque. Dans la preuve, on a utilisé le théorème de Fubini, qui nous a imposé des mesures σ-finies, mais on peut se demander si des restrictions dans l'énoncé de Minkowski sont vraiment nécessaires. Et en effet, comme le théorème de Fubini, l'inégalité de Minkowski n'est pas vraie sans un minimum d'hypothèses : considérons sur $[0,1]^2$ la fonction borélienne f qui est l'indicatrice de la diagonale, $f(x,y) = 1$ si $x = y$ et $f(x,y) = 0$ sinon. Prenons pour μ la mesure de Lebesgue sur $\mathrm{X} = [0,1]$ et pour ν la mesure de comptage sur $\mathrm{Y} = [0,1]$; on a alors

$$\int_{\mathrm{Y}} |f(x,y)|\, \mathrm{d}\nu(y) = \sum_{y \in \mathrm{Y}} |f(x,y)| = 1$$

pour tout $x \in [0,1]$, donc

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

mais pour la quantité qui devrait majorer ce premier résultat on trouve

$$\int_{\mathrm{X}} |f(x,y)|^p\, \mathrm{d}\mu(x) = 0, \quad \int_{\mathrm{Y}} \Bigl(\int_{\mathrm{X}} |f(x,y)|^p\, \mathrm{d}\mu(x)\Bigr)^{1/p} \mathrm{d}\nu(y) = 0.$$

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

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

3.3.c. *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 3.3.3 : 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*

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

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 $|fg| \le \mathrm{M}\,|g|$ presque partout, donc fg est intégrable et

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

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

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

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

$$\Bigl|\int_\Omega f(s)g(s) \,\mathrm{d}\mu(s)\Bigr| \le \frac{1}{p} \int_\Omega |f(s)|^p \,\mathrm{d}\mu(s) + \frac{1}{q} \int_\Omega |g(s)|^q \,\mathrm{d}\mu(s).$$

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 $\int |g|^q = 1$, et l'inégalité précédente donne $|\int_\Omega fg| \le 1/p + 1/q = 1$, ce qui est le résultat voulu.

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

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

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

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

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

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

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 $\int_\Omega f g \,\mathrm{d}\mu > 1 - \varepsilon$.

3.3.d. *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

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

Le corollaire ci-dessus indique que l'application j_p qui associe à une fonction $f \in \mathrm{L}^p$ la forme linéaire $g \in \mathrm{L}^q \to \int f g \,\mathrm{d}\mu$ 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, le cas $p = +\infty$ demande une hypothèse minimale,

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 il faut imposer une condition minimale. En effet, pour la mesure pathologique μ qui vaut toujours $+\infty$ sauf pour $\emptyset$, on a $\mathrm{L}^1 = \{0\}$, alors que $\mathrm{L}^\infty \neq \{0\}$ (il contient les fonctions bornées) ne peut pas être le dual de cet espace L^1. Pour terminer cette discussion sur l'identification des duaux (topologiques) des espaces L^p, rappelons que L^1 **n'est pas** le dual de L^∞, sauf circonstance extraordinaire (quand L^∞ est de dimension finie, parce que l'ensemble Ω est fini par exemple).

3.3.e. *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 $h(x,y) = f(y)g(x-y)$,

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

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

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. On notera que le résultat est valable pour la mesure de comptage sur $\mathbb{Z}$: on a $\ell^1(\mathbb{Z}) * \ell^p(\mathbb{Z}) \subset \ell^p(\mathbb{Z})$, avec l'inégalité de norme précédente.

3.3.f. *Convolutions* $\mathrm{L}^p * \mathrm{L}^q$ *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 et g dans L^q (sur $\mathbb{R}^d$, $\mathbb{T}$ ou $\mathbb{Z}$), la fonction $y \to f(x-y)$ est aussi dans L^p et a la même norme que f, donc pour tout x

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

On a donc $\mathrm{L}^p * \mathrm{L}^q \subset \mathrm{L}^\infty$ quand les exposants p, q sont conjugués, et $\|f*g\|_\infty \le \|f\|_p \, \|g\|_q$. On remarquera qu'ici, la fonction $f*g$ est définie *en tout point* ; ce fait est à rapprocher de la proposition 3.4.4, qui dira que $f*g$ est en fait une fonction continue.

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*. Il est possible de le démontrer par des applications astucieuses de l'inégalité de Hölder.

3.4. Régularisation et approximation

3.4.a. *Continuité, opérateurs de translation*

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

Proposition 3.4.1. *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 $\mathbb{R}^d$. Le même résultat est valable pour $\mathbb{T}^d$.*

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 3.4.2. *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. Plus précisément, on vient d'obtenir 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 $\mathbb{R}^d$.

Dans le cas de $\mathbb{T}$, on a affaire à des fonctions continues périodiques, qui sont toujours uniformément continues ; la proposition 3.4.2 et sa démonstration sont valables dans le cas périodique, mais elles ne disent rien de plus que le premier résultat de continuité 3.4.1. En revanche, l'information sur le module de continuité reste valable et pertinente dans le cas périodique.

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 3.4.3. *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 $\mathrm{L}^p(\mathbb{R}^d)$ (théorème 3.2.1) 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, voir plus bas). 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

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

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 3.4.4. *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 *différence de Minkowski* $\mathrm{A} - \mathrm{A} = \{a - b : a, b \in \mathrm{A}\}$ est un voisinage de 0 dans $\mathbb{R}^d$.

3.4.b. *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 commencer par un cas simple. Soient u un vecteur $\neq 0$ dans $\mathbb{R}^d$ et g une fonction sur $\mathbb{R}^d$; notons $\mathrm{D}_u g$ la *dérivée de g dans la direction u*, définie au point x par

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

supposons que g soit mesurable, définie et bornée sur $\mathbb{R}^d$, 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 ; on a

$$\frac{(f * g)(x + tu) - (f * g)(x)}{t} = \int_{\mathbb{R}^d} f(y) \, \frac{g(x + tu - y) - g(x - y)}{t} \, \mathrm{d}y.$$

En posant (pour x fixé) $\mathrm{G}(y, t) = t^{-1}\big(g(x + tu - y) - g(x - y)\big)$ on aura l'inégalité $|f(y)\mathrm{G}(y, t)| \leq \mathrm{M}\,|f(y)|$ 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

$$\mathrm{D}_u(f * g)(x) = \int_{\mathbb{R}^d} f(y)(\mathrm{D}_u g)(x - y) \, \mathrm{d}y = (f * \mathrm{D}_u g)(x).$$

Si $\mathrm{D}_u g$ est de plus continue sur $\mathbb{R}^d$, la dérivée directionnelle $f * \mathrm{D}_u g$ de $f * g$ est aussi continue sur $\mathbb{R}^d$ par la proposition 3.4.1.

Rappelons que la $j$ième dérivée partielle $\mathrm{D}_j g$ de la fonction g est la dérivée dans la direction du vecteur $\mathbf{e}_j$ de la base canonique de $\mathbb{R}^d$. 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 partielles bornées sur $\mathbb{R}^d$, et si $f \in \mathrm{L}^1$, la convolution $f * g$ est de classe C^1 sur $\mathbb{R}^d$. On a ainsi obtenu le résultat qui suit.

Proposition 3.4.5. *On suppose que $f \in \mathrm{L}^1(\mathbb{R}^d)$ et que g est mesurable bornée sur $\mathbb{R}^d$; on suppose que u est un vecteur non nul fixé, et que g admet en tout point de $\mathbb{R}^d$ une dérivée directionnelle $\mathrm{D}_u g$, bornée sur $\mathbb{R}^d$. Il en résulte que $f * g$ admet $f * \mathrm{D}_u g$ comme dérivée directionnelle $\mathrm{D}_u(f * g)$. En particulier, si $f \in \mathrm{L}^1(\mathbb{R}^d)$, si g est bornée, de classe C^1 sur $\mathbb{R}^d$ et si toutes les dérivées partielles de g sont bornées sur $\mathbb{R}^d$, la fonction $f * g$ est de classe C^1 et ses dérivées partielles sont données par*

$$\mathrm{D}_j(f * g) = f * \mathrm{D}_j g, \quad j = 1, \ldots, d.$$

On va considérer un énoncé un peu différent, où on demandera un peu plus que la simple dérivabilité directionnelle.

Proposition 3.4.6. *On suppose que $f \in \mathrm{L}^p(\mathbb{R}^d)$ et que g, définie en tout point de $\mathbb{R}^d$, appartient à $\mathrm{L}^q(\mathbb{R}^d)$ avec $1/p + 1/q = 1$; on suppose que $u \neq 0$ est fixé, et que g admet* *une dérivée directionnelle* $\mathrm{D}_u g \in \mathrm{L}^q(\mathbb{R}^d)$*, définie en tout point* $x \in \mathbb{R}^d$ *et telle que pour tout* $t > 0$ *on ait* $s \to (\mathrm{D}_u g)(x + su)$ *intégrable sur* $[0, t]$ *et*

$$g(x + tu) - g(x) = \int_0^t (\mathrm{D}_u g)(x + su)\,\mathrm{d}s. \tag{2}$$

Il en résulte que $f * g$ *admet* $f * \mathrm{D}_u g$ *comme dérivée directionnelle. En particulier, si* $f \in \mathrm{L}^p$*, si* $g \in \mathrm{L}^q$ *est de classe* C^1 *et si toutes les dérivées partielles de* g *sont dans* L^q*, la fonction* $f * g$ *est de classe* C^1 *et ses dérivées partielles sont données par*

$$\mathrm{D}_j(f * g) = f * \mathrm{D}_j g, \quad j = 1, \ldots, d.$$

Démonstration. Fixons $x \in \mathbb{R}^d$, quelconque. D'après la proposition 3.4.4, la convolée $f * \mathrm{D}_u g$ est continue sur $\mathbb{R}^d$, donc

$$s \in \mathbb{R} \to h(s) = (f * \mathrm{D}_u g)(x + su)$$

est continue. D'après le théorème de Fubini et l'hypothèse (2),

$$\mathrm{H}(t) = \int_0^t h(s)\,\mathrm{d}s = \int_{\mathbb{R}^d} f(y)\Bigl(\int_0^t (\mathrm{D}_u g)(x + su - y)\,\mathrm{d}s\Bigr)\mathrm{d}y$$

$$= \int_{\mathbb{R}^d} f(y)\bigl(g(x + tu - y) - g(x - y)\bigr)\,\mathrm{d}y = (f * g)(x + tu) - (f * g)(x).$$

Comme h est continue, la fonction H est dérivable à l'origine, de dérivée $h(0)$, ce qui signifie que $f * g$ admet au point x une dérivée dans la direction u, qui est donnée par $(f * \mathrm{D}_u g)(x)$. Justifions maintenant l'utilisation de Fubini, en montrant l'intégrabilité en $(s, y) \in [0, t] \times \mathbb{R}^d$ de la fonction considérée,

$$\int_0^t \int_{\mathbb{R}^d} |f(y)|\,|(\mathrm{D}_u g)(x + su - y)|\,\mathrm{d}s\mathrm{d}y$$

$$\leq \int_0^t \int_{\mathbb{R}^d} \Bigl(\frac{|f(y)|^p}{p} + \frac{|(\mathrm{D}_u g)(x + su - y)|^q}{q}\Bigr)\,\mathrm{d}s\mathrm{d}y$$

$$= \frac{t}{p} \int_{\mathbb{R}^d} |f(y)|^p\,\mathrm{d}y + \frac{1}{q} \int_0^t \mathrm{d}s\Bigl(\int_{\mathbb{R}^d} |(\mathrm{D}_u g)(x + su - y)|^q\,\mathrm{d}y\Bigr) = \frac{t}{p}\,\|f\|_p^p + \frac{t}{q}\,\|\mathrm{D}_u g\|_q^q$$

qui est fini.

Montrons le dernier point. Si g est de classe C^1, chaque dérivée partielle $\mathrm{D}_j g$ est continue et on a la formule (2) (appliquée avec le vecteur $u = \mathbf{e}_j$, $j$ième vecteur de la base canonique). On voit donc d'après ce qui précède que $f * g$ admet $f * \mathrm{D}_j g$ comme dérivée partielle. Comme toutes ces dérivées partielles sont continues d'après 3.4.4, la fonction $f * g$ est de classe C^1.

Si g est C^∞ à support compact, il résulte de la proposition 3.4.6 que $f * g$ est C^∞ lorsque $f \in \mathrm{L}^p(\mathbb{R}^d)$. En effet, toutes les dérivées partielles de g sont continues à support compact, donc appartiennent à L^q ; 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^p 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 $f(x) = \exp(-1/x)$ si $x > 0$ et $f(x) = 0$ si $x \le 0$ 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 sur $\mathbb{R}$, par exemple la fonction φ définie par

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

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

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

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

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

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 $f \in \mathrm{L}^p(\mathbb{R}^d)$, $1 \le p < +\infty$, la convolée $f * \varphi$ est de classe 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 tout x

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

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

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

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

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

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.c. *Approximations de l'unité*

Définition. 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 3.4.7. *On se place dans $\mathbb{R}^d$ ou $\mathbb{T}^d$, et on suppose que la suite (φ_k) est une approximation de l'unité, vérifiant les propriétés (a), (b) et (c) ci-dessus.*

– Pour toute fonction mesurable bornée g et pour tout point de continuité x_0 de g, la suite $(\varphi_k * g)(x_0)$ converge vers $g(x_0)$.

– 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 des convolées $(\varphi_k * g)$ converge vers g dans L^p.

Démonstration. Soit g une fonction mesurable bornée sur $\mathbb{R}^d$ ou $\mathbb{T}^d$, et qui soit continue au point x_0 ; considérons

$$d_k(x_0) = (\varphi_k * g)(x_0) - g(x_0) = \int \big(g(x_0 - y) - g(x_0)\big)\,\varphi_k(y)\,\mathrm{d}y\,;$$

il faut montrer que $d_k(x_0)$ tend vers 0 quand $k \to \infty$. On a

$$|d_k(x_0)| \le \int \bigl|g(x_0 - y) - g(x_0)\bigr| \, |\varphi_k(y)| \, \mathrm{d}y.$$

Posons $\mathrm{M} = \sup_k \|\varphi_k\|_1$. Puisque g est continue au point x_0, on peut trouver $\alpha > 0$ tel que $|g(x_0 - y) - g(x_0)| < \varepsilon/(2\mathrm{M})$ pour tout y tel que $d(y, 0) < \alpha$. On découpe alors l'intégrale ci-dessus en intégrale sur $\mathrm{B} = \mathrm{B}(0, \alpha)$ et complémentaire et on obtient

$$\begin{aligned} |d_k(x_0)| &\le \int_{\mathrm{B}} |g(x_0 - y) - g(x_0)| \, |\varphi_k(y)| \, \mathrm{d}y + \int_{\mathrm{B}^c} |g(x_0 - y) - g(x_0)| \, |\varphi_k(y)| \, \mathrm{d}y \\ &\le \mathrm{M} \, \frac{\varepsilon}{2\mathrm{M}} + 2\|g\|_\infty \int_{\mathrm{B}^c} |\varphi_k(y)| \, \mathrm{d}y. \end{aligned}$$

La propriété (c) implique que pour k assez grand,

$$\int_{\mathrm{B}^c} |\varphi_k(y)| \, \mathrm{d}y < \frac{\varepsilon}{1 + 4\|g\|_\infty}$$

donc $|d_k(x_0)| \le \varepsilon/2 + \varepsilon/2$; on a ainsi montré la convergence de $(\varphi_k * g)(x_0)$ vers $g(x_0)$.

Pour montrer le deuxième point on va devoir se répéter un peu. Soit g une fonction uniformément continue et bornée sur $\mathbb{R}^d$; il faut montrer que $d_k(x) = (\varphi_k * g)(x) - g(x)$ tend vers 0, uniformément en $x \in \mathbb{R}^d$. On a pour tout x

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

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 $|y| < \alpha$. En découpant comme avant, on obtient

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

qui sera $< \varepsilon$ pour k assez grand d'après la condition (c) ; on obtient ainsi la convergence uniforme de $\varphi_k * g$ vers g.

Démontrons la troisième partie. On doit montrer que $\|d_k\|_p \to 0$; on utilise l'inégalité de Minkowski (théorème 3.3.2) pour la mesure $|\varphi_k(y)| \, \mathrm{d}y$ pour obtenir que

$$\begin{aligned} \|d_k\|_p = \Bigl(\int \Bigl| \int \bigl(g(x - y) - g(x)\bigr) \, \varphi_k(y) \, \mathrm{d}y \Bigr|^p \, \mathrm{d}x \Bigr)^{1/p} \le \\ \le \int \Bigl(\int \bigl|g(x - y) - g(x)\bigr|^p \, \mathrm{d}x \Bigr)^{1/p} |\varphi_k(y)| \, \mathrm{d}y = \int \|\tau_y g - g\|_p \, |\varphi_k(y)| \, \mathrm{d}y. \end{aligned}$$

À partir de ce point la démonstration est identique à celle du 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 $|y| < \alpha$ (proposition 3.4.3). On obtient alors par un découpage identique à celui de la première partie, en posant $\mathrm{B} = \mathrm{B}(0, \alpha)$

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

Pour k assez grand on obtiendra encore avec (c) que $\|d_k\|_p \le \varepsilon$; on a ainsi montré la convergence L^p de $\varphi_k * g$ vers g.

Remarques.

– Dans le cas périodique, la deuxième partie de l'énoncé du théorème 3.4.7 s'applique à toutes les fonctions continues périodiques, qui sont automatiquement uniformément continues.

– Dans le cas de $\mathbb{R}$, si la fonction mesurable bornée g admet des limites à droite et à gauche au point x_0 et si les fonctions (φ_k) sont paires, on voit facilement par la preuve ci-dessus que $(\varphi_k * g)(x_0)$ tend vers la demi-somme $\frac{1}{2}(g(x_0+) + g(x_0-))$ des limites.

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 $\int \varphi(x)\,\mathrm{d}x = 1$ et on considère la suite (φ_k) définie par

$$\varphi_k(x) = k^d \varphi(kx) \tag{3}$$

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

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

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

et

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

tend vers 0 quand $k \to +\infty$, par convergence dominée.

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

À partir d'une fonction $\varphi \in \mathcal{D}(\mathbb{R}^d)$, 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 $\varphi_k * g$ 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 $g \in \mathrm{L}^p(\mathbb{R}^d)$, $1 \leq p < +\infty$, la suite $(\varphi_k * g)$ est formée de fonctions C^∞ d'après la proposition 3.4.6, et converge vers g dans L^p par le théorème 3.4.7. 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 $x \to \psi(x/k)$, $k \geq 1$, qui tend vers 1 (uniformément sur tout compact).

Lemme. *Soit $\psi \in \mathcal{D}(\mathbb{R}^d)$, égale à 1 dans un voisinage de 0 ; si f est dans $\mathrm{L}^p(\mathbb{R}^d)$, $1 \leq p < +\infty$, la suite des fonctions*

$$x \to \psi(x/k)\,f(x)$$

tend vers f dans L^p.

La preuve du lemme est une simple application du théorème de convergence dominée. Puisque ψ est continue à support compact, elle est bornée, disons par M. On voit que

$$g_k(x) = \bigl|\psi(x/k)f(x) - f(x)\bigr|^p$$

est bornée par $(\mathrm{M}+1)^p|f(x)|^p$, fonction intégrable fixe. Par ailleurs pour tout x on a $g_k(x)=0$ pour k assez grand, puisque x/k finit par entrer dans le voisinage de 0 dans lequel $\psi=1$.

En associant régularisation et troncature, on voit que $\mathcal{D}(\mathbb{R}^d)$ est dense dans tous les $\mathrm{L}^p(\mathbb{R}^d)$, pour $1\le p<+\infty$: si f est dans L^p, on trouve d'abord une fonction C^∞ de la forme $f_1=\varphi_{k_0}*f$, telle que $\|f-f_1\|_p<\varepsilon/2$; ensuite, par le lemme, on trouve $f_2(x)=\psi(x/k_1)f_1(x)$ telle que $\|f_1-f_2\|_p<\varepsilon/2$; finalement, f_2 est C^∞ à support compact, et $\|f-f_2\|_p<\varepsilon$.

Une approximation de l'unité utile

Lemme 3.4.8. *Si φ est une fonction continue ≥ 0 à support compact sur $\mathbb{R}^d$ (ou bien une fonction continue ≥ 0 sur $\mathbb{T}=\mathbb{R}/2\pi\mathbb{Z}$), telle que*

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

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

$$\varphi_n(x)=\Bigl(\int\varphi(y)^n\,\mathrm{d}y\Bigr)^{-1}\varphi(x)^n$$

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 $\varphi_n\ge 0$, 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

$$\mathrm{A}_k=\{x:d(x,0)\ge\alpha\ \text{et}\ \ \varphi(x)\ge t_k\}$$

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

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

d'après l'hypothèse, donc l'un de ces compacts A_k est vide. En posant $\delta=t_k$ pour un k choisi 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

$$\int_{\{d(x,0)\ge\alpha\}}\varphi(x)^n\,\mathrm{d}x\le\mathrm{M}\,\delta^n\ ;$$

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

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

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

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

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 3.4.8, on obtient une approximation de l'unité dans $\mathrm{L}^1(\mathbb{T})$ en posant pour tout $n \geq 0$

$$\mathrm{I}_n = \int_{-\pi}^{\pi} (1 + \cos(t))^n \,\frac{\mathrm{d}t}{2\pi} \quad \text{et} \quad \varphi_n(t) = \mathrm{I}_n^{-1} \bigl(1 + \cos(t)\bigr)^n,$$

qui est bien ≥ 0 d'intégrale 1 pour la probabilité $\mathrm{d}t/(2\pi)$. L'intérêt de cet exemple est que les fonctions de la suite sont des polynômes trigonométriques. On voit en effet que

$$1 + \cos(t) = \frac{1}{2}\,\mathrm{e}^{-\mathrm{i}\,t} + 1 + \frac{1}{2}\,\mathrm{e}^{\mathrm{i}\,t} = \frac{1}{2}\,(\mathrm{e}^{-\mathrm{i}\,t/2} + \mathrm{e}^{\mathrm{i}\,t/2})^2$$

ce qui implique que $\bigl(1 + \cos(t)\bigr)^n$ est un polynôme trigonométrique pour tout $n \geq 0$,

$$\bigl(1 + \cos(t)\bigr)^n = 2^{-n}\,(\mathrm{e}^{-\mathrm{i}\,t/2} + \mathrm{e}^{\mathrm{i}\,t/2})^{2n} = 2^{-n} \sum_{k=-n}^{n} \binom{2n}{n+k} \mathrm{e}^{\mathrm{i}\,kt}.$$

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

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

on aura le même phénomène : la fonction est ≥ 0 sur $\mathbb{R}$, continue à support compact, et atteint un maximum unique au point 0. L'intérêt ici est que les fonctions (φ_n) sont polynomiales sur la boule de rayon 2.

On peut aussi traiter les exemples **1** et **2** au moyen du théorème de Lebesgue. Pour l'exemple **2**, on voit pour tout $\delta \in]0, 2]$ que

$$\mathrm{J}_n(\delta) = \sqrt{n} \int_{-\delta}^{\delta} (1 - x^2/4)^n \,\mathrm{d}x = \int_{-\delta\sqrt{n}}^{\delta\sqrt{n}} \Bigl(1 - \frac{y^2}{4n}\Bigr)^n \,\mathrm{d}y$$

converge vers la limite $\ell = \int_{\mathbb{R}} \mathrm{e}^{-y^2/4} \,\mathrm{d}y$, qui est indépendante de δ et non nulle ; cette indépendance de δ donne la condition (c) des approximations de l'unité pour la suite des fonctions $\varphi_n = \mathrm{J}_n(2)^{-1} \sqrt{n}\, \mathbf{1}_{|x| \leq 2}\, (1 - x^2/4)^n$. En effet, on a $\varphi_n \geq 0$, $\int_{\mathbb{R}} \varphi_n = 1$ et de plus, pour tout $\delta \in]0, 2]$,

$$\int_{-\delta}^{\delta} \varphi_n(x) \,\mathrm{d}x = \mathrm{J}_n(\delta)/\mathrm{J}_n(2)$$

tend vers 1, ce qui implique que $\int_{|x|>\delta} \varphi_n(x) \,\mathrm{d}x$ tend vers 0 pour tout $\delta > 0$.

3.4.d. *Théorème d'approximation de Weierstrass*

On va commencer par le cas le plus simple et le plus important, celui de la droite réelle. On suivra essentiellement la méthode présentée par Weierstrass lui même vers 1880. Commençons par un peu de terminologie : une *fonction entière* est une fonction g sur $\mathbb{R}$ qui est la somme d'une série entière de rayon de convergence infini : il existe des coefficients (a_n) tels que

$$\forall x \in \mathbb{R}, \quad g(x) = \sum_{n=0}^{+\infty} a_n x^n.$$

C'est le cas de e^x, $\cos x$, $\sin x$, etc. Il est bien connu que la série converge alors pour tout nombre complexe z et permet de prolonger la fonction g au plan complexe, mais ce n'est pas ce qui nous intéressera ici. On sait aussi que les sommes partielles

$$\mathrm{P}_n(x) = \sum_{k=0}^{n} a_k x^k$$

de la série entière convergent uniformément vers $g(x)$ sur tout compact de $\mathbb{R}$. Or ce sont des fonctions polynomiales ; donc, pour réaliser l'approximation polynomiale du théorème d'approximation de Weierstrass, il suffit d'approcher par des fonctions entières. Ces fonctions entières seront obtenues par convolution avec des noyaux gaussiens.

Théorème. *Toute fonction f uniformément continue et bornée sur la droite $\mathbb{R}$ peut être approchée uniformément sur $\mathbb{R}$ par des fonctions entières. Toute fonction continue sur un intervalle fermé borné $[a, b]$ peut être approchée uniformément par des fonctions polynomiales.*

Démonstration. Posons pour $x \in \mathbb{R}$ et $n > 1$

$$g_1(x) = \frac{\mathrm{e}^{-x^2/2}}{\sqrt{2\pi}}, \quad g_n(x) = n g_1(nx).$$

On sait d'après (3) qu'on obtient ainsi une approximation de l'unité. D'après le théorème 3.4.7, la suite $h_n = g_n * f$ converge uniformément vers f sur $\mathbb{R}$. Il reste à montrer que h_n est une fonction entière. On a pour tout $n \geq 1$

$$h_n(x) = n \int_{\mathbb{R}} \mathrm{e}^{-n^2(x-y)^2/2} \, f(y) \, \frac{\mathrm{d}y}{\sqrt{2\pi}} = n \, \mathrm{e}^{-n^2x^2/2} \int_{\mathbb{R}} \mathrm{e}^{n^2xy} \, \mathrm{e}^{-n^2y^2/2} \, f(y) \, \frac{\mathrm{d}y}{\sqrt{2\pi}}.$$

Comme $x \to \mathrm{e}^{-n^2x^2/2}$ est entière, il suffit de montrer (par le théorème sur les produits de séries entières) que

$$x \to \int_{\mathbb{R}} \mathrm{e}^{n^2xy} \, \mathrm{e}^{-n^2y^2/2} \, f(y) \, \mathrm{d}y$$

est entière, et pour le faire il suffit d'intervertir l'intégrale et le développement en série entière de e^{n^2xy}. On a vu au théorème 1.3.2 qu'il suffit, pour pouvoir intervertir, que

$$\int_{\mathbb{R}} \Bigl(\sum_{k=0}^{+\infty} \frac{(n^2|xy|)^k}{k!} \Bigr) \, \mathrm{e}^{-n^2y^2/2} \, |f(y)| \, \mathrm{d}y = \int_{\mathbb{R}} \mathrm{e}^{n^2|xy|} \, \mathrm{e}^{-n^2y^2/2} \, |f(y)| \, \mathrm{d}y$$

soit fini ; or, puisque f est bornée, disons par M, l'intégrale précédente est finie (on pourra noter que $|xy| \leq x^2 + y^2/4$ pour majorer par

$$\mathrm{M}\, \mathrm{e}^{n^2 x^2} \int_{\mathbb{R}} \mathrm{e}^{-n^2 y^2/4}\, \mathrm{d}y < +\infty\,).$$

L'interversion étant justifiée, nous voyons que

$$h_n(x) = n\, \mathrm{e}^{-n^2 x^2/2} \Big(\sum_{k=0}^{+\infty} x^k \int_{\mathbb{R}} \frac{(n^2 y)^k}{k!}\, \mathrm{e}^{-n^2 y^2/2} f(y) \frac{\mathrm{d}y}{\sqrt{2\pi}} \Big)$$

est une fonction entière, ce qui achève la preuve de la première partie.

Si f_0 est continue sur $[a, b]$, on peut la prolonger en fonction uniformément continue bornée sur $\mathbb{R}$ en posant $f(x) = f(a)$ si $x < a$, $f(x) = f_0(x)$ si $a \leq x \leq b$ et $f(x) = f(b)$ si $x > b$. On aura d'abord une approximation de f_0 sur $[a, b]$ par une fonction entière, dont on prendra les sommes partielles comme expliqué précédemment.

Remarque. On pourrait donner une variante plus terre à terre qui utiliserait les noyaux de la forme $(1 - x^2/4)^n$ sur $[-2, 2]$: pour tout $n \geq 1$, définissons c_n par

$$\frac{1}{c_n} = 4^n \int_{-2}^{2} \big(1 - x^2/4\big)^n\, \mathrm{d}x = \int_{-2}^{2} \big(4 - x^2\big)^n\, \mathrm{d}x.$$

Définissons ensuite la fonction φ_n d'intégrale 1 sur $\mathbb{R}$ par $\varphi_n(x) = c_n\, (4 - x^2)^n$ lorsque $|x| \leq 2$, et $\varphi_n(x) = 0$ sinon. On a vu au lemme 3.4.8 qu'on obtient ainsi une approximation de l'unité. Si f est une fonction continue sur $\mathbb{R}$, à support dans $[-1, 1]$, on sait donc que $\varphi_n * f$ converge uniformément vers f sur $\mathbb{R}$, et en particulier sur $[-1, 1]$; pour tout x, on a

$$(\varphi_n * f)(x) = \int_{\mathbb{R}} \varphi_n(x - t) f(t)\, \mathrm{d}t = \int_{-1}^{1} \varphi_n(x - t) f(t)\, \mathrm{d}t\, ;$$

mais quand on suppose que $|x| \leq 1$, on voit que l'intégrale précédente est égale à

$$\int_{-1}^{1} c_n \big(4 - (x - t)^2\big)^n f(t)\, \mathrm{d}t\, ;$$

en effet, pour les t qui interviennent dans l'intégrale, on a $|x - t| \leq |x| + |t| \leq 2$, et $x - t$ se trouve dans la région où φ_n est polynomiale. On développe maintenant

$$c_n \big(4 - (x - t)^2\big)^n = \sum_{k=0}^{2n} a_{k,n}(t)\, x^k$$

et on obtient que dans l'intervalle $[-1, 1]$, l'approximant uniforme $\varphi_n * f$ coïncide avec la fonction polynomiale

$$\mathrm{P}_n(x) = \sum_{k=0}^{2n} \Big(\int_{-1}^{1} a_{k,n}(t) f(t)\, \mathrm{d}t \Big) x^k,$$

ce qui termine la description de cette variante.

Le théorème précédent ne règle pas le cas d'une fonction continue qui ne serait pas définie sur un intervalle, mais sur un compact $\mathrm{K} \subset \mathbb{R}$ peut-être compliqué, comme l'ensemble de Cantor. Le lemme qui suit s'occupe de cette question, même en plusieurs dimensions.

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. Introduisons le fermé F réunion de K et du complémentaire B^c de la boule unité ouverte B, et la fonction f_1 sur F, égale à f sur K et à 0 sur B^c. On vérifie facilement que f_1 est uniformément continue sur F. Soit $\varepsilon > 0$, soit $\delta > 0$ tel que $|f_1(y) - f_1(y')| < \varepsilon$ pour tout couple (y, y') d'éléments de F tels que $|y - y'| < \delta$; supposons aussi que

$$\delta < \min\{|x - y| : x \in \mathrm{K},\ y \in \mathrm{B}^c\} \,;$$

choisissons ensuite N assez grand pour que $\mathrm{N}\delta > 2\|f\|_\infty = 2\|f_1\|_\infty$, et posons

$$g(x) = \min_{y \in \mathrm{F}} \{f_1(y) + \mathrm{N}\,|x - y|\}.$$

On va montrer que la fonction g convient : cette fonction g est lipschitzienne de constante N sur $\mathbb{R}^d$, comme inf de la famille des fonctions $x \to \mathrm{N}|x - y| + f_1(y)$ dépendant du paramètre $y \in \mathrm{F}$, qui sont toutes N-lipschitziennes sur $\mathbb{R}^d$. Fixons $y_0 \in \mathrm{F}$ quelconque ; on a évidemment $f_1(y_0) \ge g(y_0)$ (en prenant $y = y_0$ comme constituant du min). Si $|y - y_0| \ge \delta$,

$$f_1(y) + \mathrm{N}|y_0 - y| \ge -\|f\|_\infty + \mathrm{N}\delta > \|f\|_\infty \ge f_1(y_0),$$

et si $|y - y_0| < \delta$ l'uniforme continuité de f_1 donne

$$f_1(y) + \mathrm{N}|y_0 - y| \ge f_1(y) \ge f_1(y_0) - \varepsilon \,;$$

il en résulte que $g(y_0) \ge f_1(y_0) - \varepsilon$. La fonction g approche donc f sur K à moins de ε ; de plus, g est nulle hors de B : si x est hors de la boule unité et $y \in \mathrm{K}$, on a $|x - y| \ge \delta$ et

$$f_1(y) + \mathrm{N}\,|x - y| = f(y) + \mathrm{N}\,|x - y| \ge -\|f\|_\infty + \mathrm{N}\,\delta \ge 0 \,;$$

si $y \in \mathrm{F} \setminus \mathrm{K}$, on a $f_1(y) + \mathrm{N}\,|x - y| = \mathrm{N}\,|x - y| \ge 0$, c'est-à-dire que $g \ge 0$ hors de la boule unité, mais on a vu aussi que $g(x) \le f_1(x) = 0$, donc g est bien nulle hors de B.

Une autre approche, qui utilise l'idée de *partition de l'unité*, fonctionne ainsi : soit $\mathrm{B}_i = \mathrm{B}(x_i, \delta)$, pour $i = 1, \ldots, \mathrm{N}$, une famille finie de boules ouvertes, centrées en des points $x_i \in \mathrm{K}$ et recouvrant K, et telles que $|f(x) - f(x_j)| < \varepsilon$ pour tout $x \in \mathrm{B}_j$; pour chaque i, soit φ_i une fonction continue telle que $\varphi_i > 0$ sur B_i et $\varphi_i = 0$ hors de B_i ; la fonction continue $\Phi = \sum_{j=1}^{\mathrm{N}} \varphi_j$ est > 0 sur K, donc il existe $\delta > 0$ tel que $\Phi \ge \delta$ sur le compact K. On pose ensuite

$$\psi_i = \frac{\varphi_i}{\varepsilon\delta + \sum_{j=1}^{\mathrm{N}} \varphi_j} \quad \text{et} \quad g = \sum_{i=1}^{\mathrm{N}} f(x_i)\psi_i.$$

La fonction g est définie sur $\mathbb{R}^d$, continue à support compact. On note que la fonction Ψ définie par $\Psi(x) = \sum_{j=1}^{\mathrm{N}} \psi_j(x)$ vérifie

$$0 \le 1 - \Psi(x) = \frac{\varepsilon\delta}{\varepsilon\delta + \Phi(x)} \le \varepsilon.$$

donc $|f(x) - \Psi(x)f(x)| \le \varepsilon \|f\|_\infty$, et

$$|g(x) - \Psi(x)f(x)| = \Bigl|\sum_{j=1}^{\mathrm{N}} (f(x) - f(x_j))\, \psi_j(x)\Bigr| \le \varepsilon \Psi(x) \le \varepsilon.$$

Remarques.

– On pourrait définir autrement une partition de l'unité formée de fonctions à petit support. Découpons la boule $\mathrm{B} = \mathrm{B}(0, \mathrm{R} + \varepsilon)$ de rayon $\mathrm{R} + \varepsilon$ en ensembles boréliens A_j deux à deux disjoints, de petit diamètre $\le \delta$, de sorte que

$$\mathbf{1}_{\mathrm{B}} = \sum_{j=1}^{n} \mathbf{1}_{\mathrm{A}_j}.$$

Si ψ est une fonction de $\mathcal{D}(\mathbb{R}^d)$ à support dans la boule $\mathrm{B}(0, \varepsilon)$, on vérifie que $\mathbf{1}_{\mathrm{B}} * \psi$ est C^∞ à support compact, égale à 1 sur $\mathrm{B}(0, \mathrm{R})$. Mais $\mathbf{1}_{\mathrm{B}} * \psi$ est la somme des $\varphi_j = \mathbf{1}_{\mathrm{A}_j} * \psi$, et chaque φ_j a un support de diamètre $\le \delta + \varepsilon$ (voir l'exercice 3.1.2).

– En fait, on peut *prolonger* la fonction continue f définie sur K, en une fonction continue sur $\mathbb{R}^d$ (c'est un cas particulier d'un théorème d'Urysohn, et c'est l'objet de l'exercice qui suit), que l'on peut choisir à support compact si on veut.

Exercice. Si (X, d) est un espace métrique, si Y est un sous-ensemble fermé de X et si f une fonction continue de Y dans $[-1, 1]$, il existe un prolongement de f à X, continu sur X.

La première étape est de trouver une fonction continue g_0 de X dans $[-1, 1]$, telle que $|f - g_0| \le 2/3$ sur Y. On recommence ensuite avec f_1 définie sur Y par $(3/2)(f - g_0)$, qu'on approche par g_1 définie sur X, à valeurs dans $[-1, 1]$, telle que $|f_1 - g_1| \le 2/3$ sur Y, ce qui donne sur Y

$$|f - g_0 - (2/3)g_1| \le (2/3)^2,$$

etc... La fonction g égale à $\sum_{n=0}^{+\infty} (2/3)^n g_n$ est continue sur X et prolonge f. Pour construire g_0, on introduira les deux fermés F_0 et F_1 de X, respectivement formés des points de Y où $f \le -1/3$ et où $f \ge 1/3$, on prendra g_0 égale à $-1/3$ sur F_0, à $1/3$ sur F_1, et $-1/3 \le g_0 \le 1/3$ partout, par exemple

$$g_0(x) = \frac{\mathrm{dist}(x, \mathrm{F}_0) - \mathrm{dist}(x, \mathrm{F}_1)}{3(\mathrm{dist}(x, \mathrm{F}_0) + \mathrm{dist}(x, \mathrm{F}_1))}.$$

Théorème 3.4.9. *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.

Esquisse. On peut supposer que le compact K est contenu dans la boule unité ouverte, et remplacer f par f_1, continue sur $\mathbb{R}^d$, à support dans la boule unité, et qui soit telle que $|f - f_1| < \varepsilon/2$ sur K. Dans le cas $d = 1$, on a montré que f_1 peut être approchée uniformément sur $\mathbb{R}$ par une fonction entière, donc approchée uniformément sur K par des fonctions polynomiales.

Pour traiter le cas $d > 1$, on pourrait généraliser la preuve utilisée en dimension 1, en convolant f_1 avec une approximation de l'unité gaussienne. Le principe est exactement le même mais l'écriture pénible. On peut aussi déduire la dimension d de la dimension 1 avec des partitions de l'unité ; esquissons le cas $d = 2$: soit f une fonction continue sur $\mathbb{R}^2$, à support dans un carré $[-\mathrm{A}, \mathrm{A}]^2$, et telle que $|f(x_1, y_1) - f(x, y)| < \varepsilon$ quand $|x_1 - x| < \delta$ et $|y_1 - y| < \delta$. On peut trouver des fonctions continues $\varphi_1, \ldots, \varphi_\mathrm{N} \geq 0$ sur $\mathbb{R}$, chacune à support dans un intervalle de longueur $< \delta$, et telles que

$$\forall x \in [-\mathrm{A}, \mathrm{A}], \quad \sum_{j=1}^{\mathrm{N}} \varphi_j(x) = 1.$$

Les N^2 fonctions $\varphi_i(x)\varphi_j(y)$ ont des supports contenus dans des carrés de côté $< \delta$, et leur somme vaut 1 sur $[-\mathrm{A}, \mathrm{A}]^2$. Pour chaque $i = 1, \ldots, \mathrm{N}$, soit x_i un point du support de φ_i ; on vérifie que la fonction g définie par

$$g(x, y) = \sum_{i,j=1}^{\mathrm{N}} f(x_i, x_j)\varphi_i(x)\varphi_j(y)$$

approche f à ε près sur le carré $[-\mathrm{A}, \mathrm{A}]^2$. D'après le théorème de Weierstrass en dimension 1, on peut ensuite approcher chaque fonction φ_i uniformément sur $[-\mathrm{A}, \mathrm{A}]$ par une fonction polynomiale P_i. On en déduira une approximation de f sur $[-\mathrm{A}, \mathrm{A}]^2$ par la fonction polynomiale de deux variables

$$\mathrm{Q}(x, y) = \sum_{i,j=1}^{\mathrm{N}} f(x_i, x_j)\mathrm{P}_i(x)\mathrm{P}_j(y).$$

Des polynômes aux polynômes trigonométriques

Pour tout entier $n \in \mathbb{Z}$, on considère la fonction exponentielle complexe e_n définie par $e_n(x) = \mathrm{e}^{\mathrm{i}nx}$. Un *polynôme trigonométrique* est une combinaison linéaire de ces fonctions (e_n). On notera $\mathcal{P}_\mathrm{N}$ l'espace engendré par les e_k, $|k| \leq \mathrm{N}$.

Théorème. *Les polynômes trigonométriques sont denses dans l'espace des fonctions continues 2π-périodiques.*

Ce résultat, analogue au théorème de Weierstrass, peut être démontré par convolution périodique avec une approximation de l'unité périodique convenable (par exemple de la forme $c_n(1+\cos t)^n$), mais on peut aussi passer de Weierstrass à Weierstrass périodique, et inversement. Cette approche a le mérite de mettre en évidence les *polynômes de Tchebychev*.

Supposons d'abord que f soit une fonction continue sur $[-1, 1]$. On lui associe une fonction périodique continue g en posant

$$\forall \theta \in \mathbb{R}, \quad g(\theta) = f(\cos(\theta)).$$

Si on connaît le théorème de Weierstrass périodique, on sait qu'on peut trouver un polynôme trigonométrique Q, de la forme $\mathrm{Q} = \sum_{k=-\mathrm{N}}^{\mathrm{N}} c_k e_k$, tel que $|\mathrm{Q} - g| < \varepsilon$; comme on a $g(-\theta) = g(\theta)$ pour tout θ, on peut remplacer Q par le polynôme trigonométrique P défini par $\mathrm{P}(\theta) = \frac{1}{2}(\mathrm{Q}(\theta) + \mathrm{Q}(-\theta))$, qui est un polynôme de cosinus,

$$\mathrm{P}(\theta) = \sum_{k=0}^{\mathrm{N}} a_k \cos(k\theta),$$

et satisfait la même approximation $|\mathrm{P} - g| < \varepsilon$. La relation entre x et θ est $\theta = \arccos x$; on introduit la fonction t_k définie par

$$\forall x \in [-1, 1], \quad t_k(x) = \cos(k \arccos(x)),$$

qui est polynomiale en x (voir plus loin ; son degré est k) ; le polynôme p défini par $p(x) = \sum_{k=0}^{\mathrm{N}} a_k t_k(x)$ approche uniformément la fonction f : tout point x de $[-1, 1]$ peut s'écrire sous la forme $x = \cos(\theta)$, et

$$|f(x) - p(x)| = \Bigl|g(\theta) - \sum_{k=0}^{\mathrm{N}} a_k t_k(\cos(\theta))\Bigr| = \Bigl|g(\theta) - \sum_{k=0}^{\mathrm{N}} a_k \cos(k\theta)\Bigr| = |g(\theta) - \mathrm{P}(\theta)| < \varepsilon.$$

Pour voir que t_k est un polynôme en $x = \cos\theta$, on écrit

$$2\cos(k\theta) = \mathrm{e}^{\mathrm{i}\,k\theta} + \mathrm{e}^{-\mathrm{i}\,k\theta} = (\cos\theta + \mathrm{i}\sin\theta)^k + (\cos\theta - \mathrm{i}\sin\theta)^k,$$

et en utilisant $\sin\theta = \sqrt{1 - x^2}$ si $0 < \theta < \pi$, on voit que

$$t_k(x) = \frac{(x + \mathrm{i}\sqrt{1 - x^2})^k + (x - \mathrm{i}\sqrt{1 - x^2})^k}{2} \ ;$$

on utilise la formule du binôme et la disparition des puissances impaires de $\sin\theta$,

$$t_k(x) = \cos(k\theta) = \sum_{0 \le 2j \le k} \binom{k}{2j} (\cos\theta)^{k-2j} \, (\mathrm{i}\sin\theta)^{2j} = \sum_{0 \le 2j \le k} (-1)^j \binom{k}{2j} x^{k-2j} \, (1 - x^2)^j.$$

Inversement, supposons connu le théorème de Weierstrass, et soit g une fonction 2π-périodique paire continue ; on peut introduire une fonction f continue sur $[-1, 1]$ telle que $f(\cos\theta) = g(\theta)$ pour tout θ ; par Weierstrass, on peut approcher f par un polynôme

$$p(x) = \sum_{k=0}^{\mathrm{N}} c_k x^k$$

et $\mathrm{P}(\theta) = p(\cos(\theta))$ est un polynôme trigonométrique qui approche g. Si g est impaire, on voit que $g(0) = g(\pi) = 0$. On peut approcher g par une fonction impaire g_1 nulle au voisinage de 0 et π ; alors

$$g_2(\theta) = \frac{g_1(\theta)}{\sin\theta}$$

définit une fonction périodique paire continue, qu'on peut approcher par un polynôme trigonométrique P d'après ce qui précède, et $\mathrm{P}_1(\theta) = \mathrm{P}(\theta)\sin\theta$ est un autre polynôme trigonométrique qui approche g_1, donc g. Comme toute fonction périodique continue g est la somme d'une fonction paire et d'une fonction impaire, le résultat est établi : le théorème de Weierstrass pour les polynômes implique le « théorème de Weierstrass trigonométrique ».

Séries de Fourier et théorème de Fejér

On va travailler avec des séries de Fourier complexes sur l'intervalle $[0, 2\pi]$, muni de la mesure normalisée $\mathrm{d}x/(2\pi)$. On identifiera au besoin les fonctions définies sur $[0, 2\pi[$ à des fonctions 2π-périodiques sur $\mathbb{R}$. Le produit scalaire de deux fonctions f et g de $\mathrm{L}^2(\mathbb{T}) \simeq \mathrm{L}^2([0, 2\pi], \mathrm{d}x/(2\pi))$ sera donné par

$$\langle f, g \rangle = \int_0^{2\pi} f(t)\overline{g(t)}\,\frac{\mathrm{d}t}{2\pi}.$$

Pour tout $n \in \mathbb{Z}$, posons $e_n(x) = \mathrm{e}^{\mathrm{i}nx}$; pour ce produit scalaire, la suite (e_n) des exponentielles complexes est orthonormée, puisque

$$\langle e_m, e_n \rangle = \int_0^{2\pi} \mathrm{e}^{\mathrm{i}nx}\,\overline{\mathrm{e}^{\mathrm{i}mx}}\,\frac{\mathrm{d}x}{2\pi} = \int_0^{2\pi} \mathrm{e}^{\mathrm{i}(n-m)x}\,\frac{\mathrm{d}x}{2\pi}$$

est nul quand $m \neq n$, et vaut 1 si $m = n$. Si $f \in \mathrm{L}^1(\mathbb{T})$, on définit les coefficients de Fourier (complexes) de la fonction f par

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

Il est clair que $|c_n(f)| \le \|f\|_1$ pour tout n. Si la fonction f est dans $\mathrm{L}^2(\mathbb{T})$, on peut aussi écrire $c_n(f) = \langle f, e_n \rangle$.

La convolution utilisée ici sera la convolution périodique

$$(f * g)(x) = \int_0^{2\pi} f(x - t)g(t)\,\frac{\mathrm{d}t}{2\pi}.$$

On note que $f * e_n = c_n(f)e_n$ 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 peut aussi utiliser le passage de Weierstrass à Weierstrass périodique expliqué ci-dessus). Les résultats généraux sur la convolution donnent la proposition suivante.

Proposition. *Les polynômes trigonométriques sont denses dans l'espace* $\mathrm{C}(\mathbb{T})$ *des fonctions continues 2π-périodiques ; pour tout p tel que $1 \le p < +\infty$, les polynômes trigonométriques sont denses dans* $\mathrm{L}^p(0, 2\pi)$.

Il est intéressant d'introduire une autre approximation périodique de l'unité, qui a un statut historique plus affirmé et qui a aussi d'autres vertus, qui seront utilisées au chapitre suivant sur les séries de Fourier : il s'agit du *noyau de Fejér*. Partons du polynôme trigonométrique

$$f_n = e_0 + \cdots + e_{n-1},$$

pour lequel on a la formule

$$f_n(x) = \frac{\mathrm{e}^{\mathrm{i}\, nx} - 1}{\mathrm{e}^{\mathrm{i}\, x} - 1} = \frac{\mathrm{e}^{\mathrm{i}\, nx/2} \sin(nx/2)}{\mathrm{e}^{\mathrm{i}\, x/2} \sin(x/2)}\,;$$

le carré du module de f_n est un polynôme trigonométrique, égal à

$$|f_n(x)|^2 = \Bigl(\sum_{k=0}^{n-1} \mathrm{e}^{\mathrm{i}\, kx}\Bigr)\Bigl(\sum_{k=0}^{n-1} \mathrm{e}^{-\mathrm{i}\, kx}\Bigr) = \Bigl(\frac{\sin(nx/2)}{\sin(x/2)}\Bigr)^2.$$

L'expression initiale de f_n comme somme des n fonctions orthogonales $e_0, \ldots, e_{n-1}$ de norme 1 montre que

$$\int_0^{2\pi} |f_n(x)|^2 \,\frac{\mathrm{d}x}{2\pi} = n.$$

Pour $n \ge 0$, on posera

$$\mathrm{K}_n(x) = \frac{|f_{n+1}(x)|^2}{n+1} = \frac{1}{n+1}\Bigl(\frac{\sin\bigl[(n+1)x/2\bigr]}{\sin(x/2)}\Bigr)^2.$$

La fonction K_n est un polynôme trigonométrique, appelé *noyau de Fejér* ; c'est une fonction réelle ≥ 0, paire, d'intégrale 1 sur chaque période (pour la mesure normalisée). On voit facilement que la suite (K_n) est une approximation de l'unité ; en effet, pour tout $\delta > 0$ fixé, plus petit que π, on a

$$\int_{\{\delta<|x|<\pi\}} \mathrm{K}_n(x) \,\frac{\mathrm{d}x}{2\pi} = 2\int_\delta^\pi \mathrm{K}_n(x) \,\frac{\mathrm{d}x}{2\pi} \le \frac{1}{(n+1)\pi} \int_\delta^\pi \frac{\mathrm{d}x}{\sin(x/2)^2} \le \frac{1}{(n+1)\sin(\delta/2)^2}$$

quantité qui tend vers 0 quand $n \to +\infty$. On a par ailleurs

$$(n+1)\mathrm{K}_n(x) = \Bigl(\sum_{j=0}^{n} \mathrm{e}^{\mathrm{i}\, jx}\Bigr)\Bigl(\sum_{k=0}^{n} \mathrm{e}^{-\mathrm{i}\, kx}\Bigr) = \sum_{j,k=0}^{n} \mathrm{e}^{\mathrm{i}\,(j-k)x}\,;$$

la différence $d = j - k$ varie de $-n$ à n, et $j = d + k$ doit rester entre 0 et n ; si $d < 0$, cela impose $|d| \le k \le n$, et $0 \le k \le n - d$ si $d \ge 0$; il en résulte que

$$(n+1)\mathrm{K}_n(x) = \sum_{d=-n}^{n} \sum_{k=0}^{n} \mathbf{1}_{\{0 \le k+d \le n\}}\, \mathrm{e}^{\mathrm{i}\, dx} = \sum_{d=-n}^{n} \bigl(n+1-|d|\bigr)\, \mathrm{e}^{\mathrm{i}\, dx}\,.$$

On a donc

$$\mathrm{K}_n = \sum_{i=-n}^{n} \Bigl(1 - \frac{|i|}{n+1}\Bigr) e_i.$$

Si f est 2π-périodique intégrable sur chaque période, la convolution de f avec K_n est un polynôme trigonométrique appelé $n$ième *somme de Fejér* de f,

$$(\sigma_n f)(x) = (\mathrm{K}_n * f)(x) = \sum_{k=-n}^{n} \Big(1 - \frac{|k|}{n+1}\Big)\, c_k(f)\, \mathrm{e}^{\mathrm{i}kx}\,.$$

Puisque la suite (K_n) est une approximation de l'unité, le théorème 3.4.7 fournit le *théorème de Fejér.*

Théorème de Fejér. *Pour toute fonction 2π-périodique continue f, la suite des sommes de Fejér $(\sigma_n * f)$ converge uniformément vers f. Pour toute fonction 2π-périodique f dont la restriction à chaque période est dans L^p, $1 \le p < +\infty$, la suite $(\sigma_n * f)$ converge vers f dans $\mathrm{L}^p(0, 2\pi)$.*

Base hilbertienne de $\mathrm{L}^2(\mathbb{T})$

Si $f \in \mathrm{L}^2(\mathbb{T})$, on considère pour tout entier $\mathrm{N} \ge 0$

$$f_{\mathrm{N}} = \sum_{k=-\mathrm{N}}^{\mathrm{N}} \langle f, e_k\rangle\, e_k = \sum_{k=-\mathrm{N}}^{\mathrm{N}} c_k(f)\, e_k \in \mathcal{P}_{\mathrm{N}}\,;$$

la différence $f - f_{\mathrm{N}}$ est orthogonale à $\mathcal{P}_{\mathrm{N}}$: en effet, pour tout j tel que $|j| \le \mathrm{N}$ on a

$$\langle f - f_{\mathrm{N}}, e_j\rangle = \langle f, e_j\rangle - \sum_{k=-\mathrm{N}}^{\mathrm{N}} \langle f, e_k\rangle\, \langle e_k, e_j\rangle = \langle f, e_j\rangle - \langle f, e_j\rangle\, \langle e_j, e_j\rangle = 0,$$

donc $f - f_{\mathrm{N}}$ est orthogonale à $\mathrm{Vect}(e_j : |j| \le \mathrm{N}) = \mathcal{P}_{\mathrm{N}}$. Si g est un élément quelconque de $\mathcal{P}_{\mathrm{N}}$, on a $f_{\mathrm{N}} - g \in \mathcal{P}_{\mathrm{N}}$, donc $f - f_{\mathrm{N}}$ lui est orthogonal et

$$\|f - g\|_2^2 = \|(f - f_{\mathrm{N}}) + (f_{\mathrm{N}} - g)\|_2^2 = \|f - f_{\mathrm{N}}\|_2^2 + \|f_{\mathrm{N}} - g\|_2^2 \ge \|f - f_{\mathrm{N}}\|_2^2.$$

D'après le théorème de Weierstrass trigonométrique et la densité des fonctions continues dans $\mathrm{L}^2(\mathbb{T})$ (ou bien d'après Fejér), il existe pour tout $\varepsilon > 0$ un polynôme trigonométrique g tel que $\|f - g\|_2 < \varepsilon$; cette fonction g est dans un certain espace $\mathcal{P}_{\mathrm{N}_0}$, donc aussi dans $\mathcal{P}_{\mathrm{N}}$ pour tout $\mathrm{N} \ge \mathrm{N}_0$; il en résulte que pour tout $\mathrm{N} \ge \mathrm{N}_0$,

$$\|f - f_{\mathrm{N}}\|_2 \le \|f - g\|_2 < \varepsilon,$$

ce qui signifie précisément que la suite (f_{N}) converge vers f dans $\mathrm{L}^2(\mathbb{T})$. On en déduit que

$$\|f\|_2^2 = \lim_{\mathrm{N}} \|f_{\mathrm{N}}\|_2^2 = \sum_{k\in\mathbb{Z}} |c_k(f)|^2.$$

et on a aussi

$$f = \sum_{n\in\mathbb{Z}} c_n(f) e_n,$$

où la série converge au sens de L^2. Ceci ne donne *a priori* aucune convergence ponctuelle de la série de Fourier. Cette question de la convergence ponctuelle sera envisagée au chapitre suivant.