Prépa. Agrég écrit d'Analyse, séance n° 1, mercredi 3 Octobre 2001.

1.1. Intégration des fonctions positives

Fonctions et ensembles

Quelques notations : on note A^c le complémentaire d'un sous-ensemble $\mathrm{A} \subset \mathrm{X}$; si f est une fonction réelle sur X et $t \in \mathbb{R}$, on note $\{f > t\}$ l'ensemble $\{x \in \mathrm{X} : f(x) > t\}$; si U est un sous-ensemble de $\mathbb{R}$, on note similairement $\{f \in \mathrm{U}\}$ pour $f^{-1}(\mathrm{U})$. On note $\mathbf{1}_{\mathrm{A}}$ la fonction indicatrice d'un sous-ensemble $\mathrm{A} \subset \mathrm{X}$, fonction égale à 1 sur A et à 0 en dehors de A.

Arithmétique de $[0, +\infty]$

On définit les opérations d'addition et multiplication de la façon suivante : les résultats ont les valeurs habituelles quand les deux nombres sont finis ; on pose en outre $a + (+\infty) = +\infty$ pour tout $a \geq 0$ (ce qui va de soi), $a \times (+\infty) = +\infty$ si $a > 0$, et enfin $0 \times (+\infty) = 0$. Pour comprendre cette dernière convention il faut voir qu'on fait le choix qui rend les opérations continues pour les limites de suites **croissantes** ; ainsi $0 = 0 \times n$ tend vers $0 \times +\infty = 0$ quand n croît vers $+\infty$.

Toute suite croissante admet évidemment une limite dans ce cadre ; on peut donc définir sans problème la somme de n'importe quelle série à termes dans $[0, +\infty]$.

Algèbres, σ-algèbres (ou tribus)

Définition. Une *algèbre* $\mathcal{A}$ de parties d'un ensemble X est un sous-ensemble $\mathcal{A} \subset \mathcal{P}(\mathrm{X})$, qui contient $\emptyset$ et est stable par réunion finie et par passage au complémentaire :

(i) $\emptyset \in \mathrm{A}$;
(ii) $(\mathrm{A}, \mathrm{B} \in \mathcal{A}) \Rightarrow (\mathrm{A} \cup \mathrm{B}) \in \mathcal{A}$;
(iii) $\mathrm{A} \in \mathcal{A} \Rightarrow \mathrm{A}^c \in \mathcal{A}$.

Il en résulte la stabilité par intersection finie. On peut définir la notion d'algèbre, si on préfère, avec la stabilité par intersection finie et par passage au complémentaire.

Définition. Une *tribu* ou *σ-algèbre* de parties de X est une classe $\mathcal{A} \subset \mathcal{P}(\mathrm{X})$ qui contient la partie vide $\emptyset$, et qui est stable par complémentaire et par *union dénombrable*.

Autrement dit, il suffit de remplacer dans la définition d'une algèbre l'axiome (ii) par

$(ii)_\sigma$ pour toute suite (A_n) d'éléments de $\mathcal{A}$, on a $\bigcup_n \mathrm{A}_n \in \mathcal{A}$.

La tribu est le cadre naturel où toutes les opérations dénombrables sont permises. Pour qu'une algèbre soit en fait une tribu, il suffit d'ajouter la stabilité par réunion des suites croissantes (parce que $\bigcup_n \mathrm{A}_n$ est aussi égal à la réunion de la suite croissante $\mathrm{B}_n = \mathrm{A}_0 \cup \ldots \cup \mathrm{A}_n$), ou bien la stabilité par intersection des suites décroissantes.

Pour tout ensemble X, l'ensemble $\mathcal{P}(\mathrm{X})$ est évidemment une tribu de parties de X, de même que $\{\emptyset, \mathrm{X}\}$. Sur $\mathbb{R}$, $\mathbb{C}$ ou $\mathbb{R}^d$, la tribu naturelle à considérer est la tribu engendrée par les ouverts ; on l'appelle la *tribu borélienne*.

Mesures positives sur une tribu

Un *espace mesurable* $(\Omega, \mathcal{A})$ est simplement le couple d'un ensemble Ω et d'une tribu $\mathcal{A}$ de parties de Ω.

Définition. Une *mesure positive* μ sur un espace mesurable $(\Omega, \mathcal{A})$ est une application $\mu : \mathcal{A} \to [0, +\infty]$ telle que :

$\mu(\emptyset) = 0$ et pour toute suite (A_n) d'éléments deux à deux disjoints de $\mathcal{A}$, on a $\mu(\bigcup_{n=0}^{+\infty} \mathrm{A}_n) = \sum_{n=0}^{+\infty} \mu(\mathrm{A}_n)$ (tous les calculs sont faits avec la valeur $+\infty$ admise).

En prenant $\mathrm{A}_0 = \mathrm{A}$, $\mathrm{A}_1 = \mathrm{B}$ et $\mathrm{A}_n = \emptyset$ pour $n \geq 2$ on retrouve l'additivité finie. Si la mesure prend au moins une valeur finie, $\mu(\emptyset) = 0$ résulte de l'additivité. Dans le cas le plus général, on a la mesure pathologique qui vaut $+\infty$ pour tout $\mathrm{A} \neq \emptyset$, qui nous oblige à inclure $\mu(\emptyset) = 0$ dans la définition.

Si on a déjà l'additivité, la σ-additivité s'obtient par la régularité de la mesure pour les limites croissantes : pour toute suite (B_n) croissante d'éléments de $\mathcal{A}$, on a

$$\mu\Big(\bigcup_n \mathrm{B}_n\Big) = \lim_n \mu(\mathrm{B}_n).$$

Pour une mesure finie, on peut aussi passer à la limite pour les suites décroissantes, mais la régularité décroissante n'est pas vraie pour les mesures infinies.

Intégrale des fonctions étagées positives

Le point de vue le plus agréable pour ce paragraphe (et le suivant) est de travailler avec l'ensemble $[0, +\infty]$, muni de l'arithmétique étendue introduite précédemment.

On considère un espace mesuré $(\Omega, \mathcal{A}, \mu)$ et on intègre d'abord les fonctions $\mathcal{A}$-étagées à valeurs dans $[0, +\infty]$. Nous appelons fonction $\mathcal{A}$-étagée une fonction f sur Ω qui ne prend qu'un nombre fini de valeurs, disons $v_1, \ldots, v_n \in [0, +\infty]$, de façon que $\{f = v_j\}$ soit un ensemble de $\mathcal{A}$ pour tout $j = 1, \ldots, n$ (c'est une fonction mesurable de $(\Omega, \mathcal{A})$ dans $[0, +\infty]$ muni de sa tribu borélienne –qui ne sert à rien pour l'instant– et qui ne prend qu'un nombre fini de valeurs).

Si f est une telle fonction étagée, on peut l'exprimer sous la forme $f = \sum_{i=1}^m a_i \, \mathbf{1}_{\mathrm{A}_i}$, où $\mathrm{A}_1, \ldots, \mathrm{A}_m$ est une partition de Ω en ensembles $\mathrm{A}_i \in \mathcal{A}$, et où $a_i \geq 0$ est la valeur constante de f sur A_i (si A_i n'est pas vide ; si A_i est vide, a_i est n'importe quel nombre ≥ 0). Une façon d'obtenir une telle expression pour f est de considérer l'ensemble fini $v_1, \ldots, v_n$ des valeurs de f, de poser $\mathrm{V}_i = \{f = v_i\}$ pour tout i et de voir qu'on a $f = \sum_{i=1}^n v_i \, \mathbf{1}_{\mathrm{V}_i}$; cependant il est essentiel de permettre un peu plus de souplesse dans la représentation de f, en n'exigeant pas que A_i soit précisément de la forme $\{f = v\}$ pour un certain v.

La quantité $\sum_{i=1}^m a_i \, \mu(\mathrm{A}_i)$ peut être $+\infty$ mais on va vérifier qu'elle ne dépend que de la fonction f, et pas de la représentation de f du type précédent. On posera alors

$$\int f \, d\mu = \sum_{i=1}^m a_i \, \mu(\mathrm{A}_i)$$

(élément de $[0, +\infty]$) et on dira que $\int f \, d\mu$ est l'intégrale (par rapport à μ) de la fonction étagée $f \geq 0$.

Détaillons (à l'extrême) la preuve de l'indépendance par rapport à la représentation. Elle est fondée sur un principe que les intégristes tiennent à démontrer par récurrence : si $c_{i,j}$ est une famille de nombres, indexés par $i = 1, \dots, m$ et $j = 1, \dots, n$, on a

$$\sum_{i=1}^{m}\Big(\sum_{j=1}^{n} c_{i,j}\Big) = \sum_{j=1}^{n}\Big(\sum_{i=1}^{m} c_{i,j}\Big).$$

Supposons que $f = \sum_{j=1}^{n} b_j \, \mathbf{1}_{\mathrm{B}_j}$ soit une autre représentation de la même fonction f, avec toujours $b_j \geq 0$ et $\mathrm{B}_j \in \mathcal{A}$ pour tout $j = 1, \dots, n$, qui réalisent une deuxième partition de Ω. Les ensembles $\mathrm{C}_{i,j} = \mathrm{A}_i \cap \mathrm{B}_j$ forment une partition de Ω ; lorsque $\mathrm{C}_{i,j}$ est non vide, $a_i = b_j$ est la valeur de f sur $\mathrm{C}_{i,j}$, qu'on appellera $c_{i,j} = a_i = b_j$; si $\mathrm{C}_{i,j} = \emptyset$, alors $\mu(\mathrm{C}_{i,j}) = 0$, et $c_{i,j}\, \mu(\mathrm{C}_{i,j}) = 0 = a_i\, \mu(\mathrm{C}_{i,j}) = b_j\, \mu(\mathrm{C}_{i,j})$ dans ce cas, pour n'importe quel nombre $c_{i,j} \geq 0$; on aura donc dans tous les cas

$$c_{i,j}\, \mu(\mathrm{C}_{i,j}) = a_i\, \mu(\mathrm{C}_{i,j}) = b_j\, \mu(\mathrm{C}_{i,j})$$

pour tous i, j. On obtient ainsi

$$\sum_{i,j} c_{i,j}\, \mu(\mathrm{C}_{i,j}) = \sum_{i=1}^{m}\Big(\sum_{j=1}^{n} a_i\, \mu(\mathrm{C}_{i,j})\Big) = \sum_{i=1}^{m} a_i \Big(\sum_{j=1}^{n} \mu(\mathrm{A}_i \cap \mathrm{B}_j)\Big) = \sum_{i=1}^{m} a_i\, \mu(\mathrm{A}_i)$$

(on note que $\mathrm{A}_i = \bigcup_{j=1}^{n} \mathrm{A}_i \cap \mathrm{B}_j$) et de même dans l'autre sens,

$$\sum_{i,j} c_{i,j}\, \mu(\mathrm{C}_{i,j}) = \sum_{j=1}^{n}\Big(\sum_{i=1}^{m} b_j\, \mu(\mathrm{C}_{i,j})\Big) = \sum_{j=1}^{n} b_j \Big(\sum_{i=1}^{m} \mu(\mathrm{A}_i \cap \mathrm{B}_j)\Big) = \sum_{j=1}^{n} b_j\, \mu(\mathrm{B}_j).$$

On démontre ensuite que $(f \leq g) \Rightarrow (\int f\, d\mu \leq \int g\, d\mu)$ et l'additivité dans le cas ≥ 0 (ainsi que la propriété triviale $\int (\lambda f)\, d\mu = \lambda \int f\, d\mu$, quand $\lambda \geq 0$) : il suffit de remarquer qu'étant données deux fonctions étagées, on peut raffiner les partitions de façon que les deux fonctions soient constantes sur les atomes $\mathrm{C}_{i,j}$ de la partition raffinée, comme on l'a fait ci-dessus. Si $f = \sum_i a_i \mathbf{1}_{\mathrm{A}_i} \leq \sum_j b_j \mathbf{1}_{\mathrm{B}_j} = g$, on aura $a_i\, \mu(\mathrm{C}_{i,j}) \leq b_j\, \mu(\mathrm{C}_{i,j})$ pour tous i, j et on raisonne comme ci-dessus ; idem pour l'additivité.

Intégrale des fonctions mesurables positives

Dans ce paragraphe, fonction mesurable ≥ 0 signifie fonction mesurable à valeurs dans $[0, +\infty]$. On dit que la fonction $f \geq 0$ sur $(\Omega, \mathcal{A})$ est $\mathcal{A}$-mesurable si l'ensemble $\{f > c\}$ est dans $\mathcal{A}$ pour tout c réel. Si on définit f_n qui prend les valeurs $i/2^n$ pour $i = 0, \dots, 4^n$, avec $\{f_n = i/2^n\} = \{i/2^n \leq f < (i+1)/2^n\}$ quand $0 \leq i < 4^n$ et $\{f_n = 2^n\} = \{f \geq 2^n\}$, on voit qu'on définit ainsi une fonction f_n qui est $\mathcal{A}$-étagée, et que la suite (f_n) tend simplement vers f, en croissant. On peut aussi définir f_n par la formule

$$\forall \omega \in \Omega, \quad f_n(\omega) = \min\big\{2^n,\, 2^{-n}\, \big[2^n f(\omega)\big]\big\}$$

où $[x]$ désigne la partie entière d'un réel x.

Réciproquement, toute limite f d'une suite croissante (f_n) de fonctions mesurables est mesurable. En effet, on a dans ce cas pour tout c réel

$$\{f > c\} = \bigcup_n \{f_n > c\}.$$

Définition. L'intégrale d'une fonction mesurable $f \geq 0$ est le sup des intégrales des fonctions étagées $g \leq f$.

Avec cette définition, il est évident que $0 \leq f_1 \leq f_2$ implique $\int f_1 \, d\mu \leq \int f_2 \, d\mu$; il est clair aussi que $\int (f_1 + f_2) \, d\mu \geq \int f_1 \, d\mu + \int f_2 \, d\mu$. De plus on voit facilement que

$$\int f \, d\mu = \sum_{i=1}^{m} \int \mathbf{1}_{\mathrm{A}_i} f \, d\mu$$

si $\mathrm{A}_1, \ldots, \mathrm{A}_m$ est une partition de Ω en ensembles de la tribu $\mathcal{A}$: en effet dans ce cas toute fonction étagée $g \leq f$ s'écrit comme la somme $\sum_i \mathbf{1}_{\mathrm{A}_i} g$ de fonctions étagées plus petites que chaque $\mathbf{1}_{\mathrm{A}_i} f$.

Lemme. *Soient* $\mathrm{A} \in \mathcal{A}$ *et* $a \geq 0$ *; si une suite croissante* (f_n) *de fonctions mesurables* ≥ 0 *vérifie* $a\mathbf{1}_{\mathrm{A}} \leq \lim_n f_n$*, alors*

$$a \, \mu(\mathrm{A}) \leq \lim_n \int f_n \, d\mu.$$

Démonstration. Si $a = 0$, alors $a \, \mu(\mathrm{A}) = 0$ et l'inégalité est évidente. Supposons donc $a > 0$ et soit $0 < b < a$; l'ensemble $\mathrm{B}_n = \{f_n > b\}$ est croissant avec n, de limite $\mathrm{B} = \{\lim_n f_n > b\}$; l'hypothèse implique que $\mathrm{B} \supset \mathrm{A}$, et on a $b \, \mathbf{1}_{\mathrm{B}_n} \leq f_n$ donc

$$b \, \mu(\mathrm{A}) \leq b \, \mu(\mathrm{B}) = \lim_n b \, \mu(\mathrm{B}_n) \leq \lim_n \int f_n \, d\mu.$$

Proposition. Lemme des suites croissantes. *Si* (f_n) *mesurable* ≥ 0 *tend en croissant vers* f*,*

$$\int f \, d\mu = \lim_n \int f_n \, d\mu.$$

Démonstration. Soit $g = \sum_{j=1}^{m} a_j \mathbf{1}_{\mathrm{A}_j}$ une fonction étagée positive $\leq f$, avec des A_j non vides qui réalisent une partition de Ω, et $a_j \geq 0$ pour $j = 1, \ldots, m$. Pour chaque indice j, la suite $\mathbf{1}_{\mathrm{A}_j} f_n$ est croissante et tend vers $\mathbf{1}_{\mathrm{A}_j} f \geq a_j \mathbf{1}_{\mathrm{A}_j}$, donc

$$a_j \, \mu(\mathrm{A}_j) \leq \lim_n \int \mathbf{1}_{\mathrm{A}_j} f_n \, d\mu$$

d'après le lemme précédent et il n'y a plus qu'à additionner en j pour obtenir

$$\int g \, d\mu \leq \sum_{j=1}^{m} \big(\lim_n \int \mathbf{1}_{\mathrm{A}_j} f_n \, d\mu\big) = \lim_n \sum_{j=1}^{m} \int 1_{\mathrm{A}_j} f_n \, d\mu = \lim_n \int f_n \, d\mu,$$

et ceci pour toute fonction étagée $g \leq f$. Il en résulte que $\int f \, d\mu \leq \lim_n \int f_n \, d\mu$. L'autre inégalité est triviale.

Conséquence : additivité de l'intégrale. Si f et g sont deux fonctions mesurables ≥ 0, on peut trouver deux suites croissantes (f_n) et (g_n) de fonctions étagées qui tendent vers f et g respectivement ; on passe à la limite croissante dans l'additivité pour les fonctions étagées.

Proposition : lemme de Fatou. *Si (f_n) est une suite de fonctions mesurables ≥ 0,*

$$\int \liminf_n f_n \, d\mu \leq \liminf_n \int f_n \, d\mu.$$

Démonstration. Posons $g_n = \inf_{m \geq n} f_m \geq 0$. Pour chaque entier $m \geq n$ on a évidemment $g_n \leq f_m$, donc $\int g_n \leq \int f_m$ et par conséquent $\int g_n \leq \inf_{m \geq n} \int f_m$. D'après le lemme des suites croissantes le premier terme croît vers $\int \liminf_n f_n$ (parce que (g_n) tend en croissant vers $\liminf_n f_n$) alors que le deuxième croît vers $\liminf_n \int f_n$.

Intégrale des fonctions réelles ou complexes

Contrairement aux paragraphes précédents, on considère maintenant des fonctions mesurables à valeurs dans $\mathbb{R}$. On désigne par $\mathcal{L}_1(\Omega, \mathcal{A}, \mu)$ l'espace vectoriel des fonctions mesurables f à valeurs dans $\mathbb{R}$ telles que $\int |f| \, d\mu < +\infty$.

Si $f \in \mathcal{L}_1$, on peut écrire $f = f^+ - f^-$, où $f^+ = \max(f, 0)$ et $f^- = \max(-f, 0)$. Ces deux fonctions f^+ et f^- sont mesurables ≥ 0 (exercice facile), et $f^+, f^- \leq f^+ + f^- = |f|$, donc $\int f^+ \, d\mu$ et $\int f^- \, d\mu$ sont finies.

Si $f \in \mathcal{L}_1$, il y a donc au moins une façon d'écrire $f = f_1 - f_2$ avec $f_1, f_2 \geq 0$ d'intégrale finie. Si on a une autre décomposition $f = g_1 - g_2$ avec $g_1, g_2 \geq 0$ d'intégrale finie, on remarque que $f_1 + g_2 = f_2 + g_1$ et on utilise l'additivité de l'intégrale dans le cas positif pour vérifier que $\int f_1 - \int f_2 = \int g_1 - \int g_2$. On peut donc poser

$$\int f \, d\mu = \int f_1 \, d\mu - \int f_2 \, d\mu.$$

En particulier on a la cohérence avec la définition antérieure du cas positif. Il est facile de vérifier la linéarité de l'intégrale. Notons aussi que

$$\Big| \int f \, d\mu \Big| = \Big| \int f^+ \, d\mu - \int f^- \, d\mu \Big| \leq \int f^+ \, d\mu + \int f^- \, d\mu = \int |f| \, d\mu.$$

Si f est à valeurs complexes, on la décomposera en partie réelle et partie imaginaire, et on posera la définition qui va de soi.

Théorème : théorème de convergence dominée de Lebesgue. *Si f_n tend simplement vers f et $|f_n| \leq g$ pour tout n, avec $\int g < +\infty$, alors $\int f = \lim_n \int f_n$.*

Démonstration à partir du lemme de Fatou. On va montrer un résultat apparemment plus fort, $\int |f - f_n| \, d\mu \to 0$. On a $|f_n - f| \leq 2g$ avec g intégrable. On considère $g_n = 2g - |f - f_n| \geq 0$, qui tend simplement vers $2g$. D'après Fatou

$$\int 2g \, d\mu = \int \liminf \big(2g - |f - f_n|\big) \, d\mu \leq \liminf \int \big(2g - |f - f_n|\big) \, d\mu$$

c'est à dire

$$\int 2g \, d\mu \leq \int 2g \, d\mu - \limsup_n \int |f - f_n| \, d\mu$$

donc, puisque $\int 2g \, d\mu$ est un nombre fini, on obtient $\limsup_n \int |f - f_n| \, d\mu \leq 0$.

1.2. Opérations sur les tribus ; tribu engendrée ; mesurabilité

Toute intersection de tribus est une tribu. Il en résulte que pour toute classe $\mathcal{C}$ de parties de X, il existe une plus petite tribu contenant $\mathcal{C}$. On l'appelle la *tribu engendrée* par la classe $\mathcal{C}$, et on la note $\sigma(\mathcal{C})$.

Evidemment, si $\mathcal{C}$ est déjà une tribu, alors $\sigma(\mathcal{C}) = \mathcal{C}$.

Principe trivial et répété sans cesse :

si $\mathcal{C} \subset \mathcal{A}$, avec $\mathcal{A}$ une tribu, alors $\sigma(\mathcal{C}) \subset \mathcal{A}$;
on peut dire aussi : si $\mathcal{C}_1 \subset \mathcal{C}_2$, alors $\sigma(\mathcal{C}_1) \subset \sigma(\mathcal{C}_2)$.

Mettons ce petit principe en application. La tribu borélienne de $\mathbb{R}$ est la tribu $\mathcal{B}_{\mathbb{R}}$ qui est engendrée par les ouverts de $\mathbb{R}$; c'est aussi la tribu engendrée par la classe $\mathcal{C}$ formée des intervalles $]c, +\infty[$, où c varie dans $\mathbb{R}$; posons $\mathcal{A} = \sigma(\mathcal{C})$, et montrons que $\mathcal{A} = \mathcal{B}_{\mathbb{R}}$. Evidemment $\mathcal{A} \subset \mathcal{B}_{\mathbb{R}}$ puisque $\mathcal{B}_{\mathbb{R}}$ contient tous les générateurs de $\mathcal{A}$. Pour tout c, l'intervalle fermé $[c, +\infty[$ est l'intersection de la suite des $]c-2^{-n}, +\infty[$, donc $[c, +\infty[\in \mathcal{A}$ pour tout c réel ; en passant au complémentaire on voit que $]-\infty, b[$ est dans $\mathcal{A}$, et par intersection avec $]a, +\infty[$ on déduit que tout intervalle $]a, b[$ est dans $\mathcal{A}$. Comme tout ouvert de $\mathbb{R}$ est réunion dénombrable d'intervalles ouverts, on voit que les générateurs de $\mathcal{B}_{\mathbb{R}}$ sont dans $\mathcal{A}$, donc $\mathcal{B}_{\mathbb{R}} \subset \mathcal{A}$, d'où l'égalité des deux tribus.

On montre de la même façon que la tribu borélienne de $\overline{\mathbb{R}}$ est engendrée par les intervalles $]c, +\infty]$.

Produit de tribus

On suppose donnés X_1 et X_2 avec chacun une tribu $\mathcal{A}_1$ et $\mathcal{A}_2$; on définit la *tribu produit* sur l'ensemble produit $\mathrm{X}_1 \times \mathrm{X}_2$, notée $\mathcal{A}_1 \otimes \mathcal{A}_2$: c'est la tribu engendrée par les ensembles de la forme $\mathrm{A}_1 \times \mathrm{A}_2$, avec $\mathrm{A}_j \in \mathcal{A}_j$:

$$\mathcal{A}_1 \otimes \mathcal{A}_2 = \sigma\big(\{\mathrm{A}_1 \times \mathrm{A}_2 : \mathrm{A}_1 \in \mathcal{A}_1, \mathrm{A}_2 \in \mathcal{A}_2\}\big).$$

La mesurabilité abstraite

Définition. Etant donnés deux espaces mesurables $(\Omega_1, \mathcal{A}_1)$ et $(\Omega_2, \mathcal{A}_2)$, on dit qu'une application $f : \Omega_1 \to \Omega_2$ est *mesurable* si $f^{-1}(\mathcal{A}_2) \subset \mathcal{A}_1$.

Propriété évidente : la composition d'applications mesurables fournit une application mesurable.

Le critère suffisant en $f^{-1}(\mathcal{C})$: pour que f soit mesurable de $(\Omega_1, \mathcal{A}_1)$ dans $(\Omega_2, \mathcal{A}_2)$, il suffit que $\{f \in \mathrm{C}\} \in \mathcal{A}_1$ pour tout C d'une classe $\mathcal{C}$ qui engendre $\mathcal{A}_2$.

Par exemple :

pour que f soit mesurable de $(\Omega, \mathcal{A})$ dans $\overline{\mathbb{R}}$ muni de la tribu borélienne, il faut et il suffit que

$$\forall t \in \mathbb{R}, \quad \{f > t\} \in \mathcal{A}.$$

Si f_1 et f_2 sont deux applications mesurables de $(\Omega, \mathcal{A})$ dans $(\Omega_j, \mathcal{A}_j)$, $j = 1, 2$, l'application couple $g : \omega \to (f_1(\omega), f_2(\omega))$ est mesurable de $(\Omega, \mathcal{A})$ dans $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$. En effet, il suffit de tester l'image inverse des générateurs de la forme $\mathrm{A}_1 \times \mathrm{A}_2$, ce qui est facile :

$$g^{-1}(\mathrm{A}_1 \times \mathrm{A}_2) = f_1^{-1}(\mathrm{A}_1) \cap f_2^{-1}(\mathrm{A}_2) \in \mathcal{A}.$$

Topologie et tribus ; tribus boréliennes

Définition. Soit X un espace topologique ; la *tribu borélienne* de X, qu'on notera $\mathcal{B}_{\mathrm{X}}$, est la tribu de parties de X engendrée par la classe $\mathcal{O}_{\mathrm{X}}$ des ouverts de X, c'est à dire $\mathcal{B}_{\mathrm{X}} = \sigma(\mathcal{O}_{\mathrm{X}})$.

Si $\mathcal{C}$ est une classe d'ouverts telle que tout ouvert de X soit réunion dénombrable d'éléments de $\mathcal{C}$, on aura $\mathcal{B}_{\mathrm{X}} = \sigma(\mathcal{C})$. En fait si $\mathcal{C}$ est une classe de *boréliens* telle que tout ouvert de X soit réunion dénombrable d'éléments de $\mathcal{C}$, on aura encore $\mathcal{B}_{\mathrm{X}} = \sigma(\mathcal{C})$.

Topologie et dénombrabilité

Proposition. *Si* (X, d) *est un espace métrique séparable, il existe une suite* (U_n) *d'ensembles ouverts telle que tout ouvert de* X *soit réunion d'une sous-famille de cette famille. On dit que la topologie de* X *est* à base dénombrable. *On peut choisir pour* (U_n) *les boules de rayon rationnel centrées en une suite dense dans* X.

Démonstration. Soit (x_k) une suite dense dans X et considérons la famille dénombrable de toutes les boules ouvertes $\mathrm{U}_{k,j} = \mathrm{B}(x_k, 2^{-j})$, où j varie dans $\mathbb{N}$; on va montrer que pour tout ouvert V de X et pour tout $x \in \mathrm{V}$, il existe un ouvert $\mathrm{U}_{k,j}$ tel que

$$x \in \mathrm{U}_{k,j} \subset \mathrm{V}.$$

Il en résulte évidemment que V est la réunion de ceux des ensembles $\mathrm{U}_{k,j}$ qu'il contient. Vérifions l'affirmation : puisque $x \in \mathrm{V}$ et V ouvert dans l'espace métrique X, il existe $r > 0$ tel que la boule ouverte $\mathrm{B}(x, r)$ soit contenue dans V. Choisissons j tel que $2^{-j} < r/2$, et un entier k tel que $d(x, x_k) < 2^{-j}$. On a donc $x \in \mathrm{B}(x_k, 2^{-j}) = \mathrm{U}_{k,j}$, et on obtient immédiatement avec l'inégalité triangulaire que $\mathrm{B}(x_k, 2^{-j}) \subset \mathrm{B}(x, r)$: si $y \in \mathrm{B}(x_k, 2^{-j})$, on a $d(y, x) \leq d(y, x_k) + d(x_k, x) \leq 2 \,.\, 2^{-j} < r$.

Exercices proposés dans la séance 1.

Exercice 1.1. Si f est mesurable ≥ 0, montrer que $\int f \, d\mu = 0$ si et seulement si $\mu(\{f > 0\}) = 0$.

Montrer que si $\int f \, d\mu < \infty$, alors $\mu(\{f = +\infty\}) = 0$.

Exercice 1.2. Si f est mesurable ≥ 0 montrer que

$$\int f \, d\mu = \int_0^{+\infty} \mu(\{f > t\} \, dt.$$

Exercice 1.3. Si f est intégrable sur $\mathbb{R}$, on pose

$$\forall t \in \mathbb{R}, \quad \widehat{f}(t) = \int_{\mathbb{R}} f(x) \, \mathrm{e}^{-ixt} \, dx.$$

a. Montrer que $\widehat{f}$ est bornée, continue, et tend vers 0 à l'infini.

b. Montrer que si $\int_{\mathbb{R}} |x|^k |f(x)| \, dx < +\infty$, alors $\widehat{f}$ est de classe C^k sur $\mathbb{R}$ (k est un entier ≥ 1).

c. Si μ est une mesure ≥ 0 finie et symétrique sur $\mathbb{R}$, montrer que $\widehat{\mu}$ est de classe C^2 si et seulement si $\int_{\mathbb{R}} x^2 \, d\mu(x) < +\infty$.

Exercice 1.4. On suppose que μ est une probabilité sur $(\Omega, \mathcal{A})$ telle que pour tout $\mathrm{A} \in \mathcal{A}$ tel que $\mu(\mathrm{A}) > 0$, il existe $\mathrm{B} \in \mathcal{A}$, $\mathrm{B} \subset \mathrm{A}$ et $0 < \mu(\mathrm{B}) < \mu(\mathrm{A})$. Montrer que pour tout $c \in [0, 1]$, il existe un ensemble $\mathrm{A}_c \in \mathcal{A}$ tel que $\mu(\mathrm{A}_c) = c$.

Prépa. Agrég écrit d'Analyse, séance n° 2, mercredi 10 octobre 2001.

Tribu borélienne, suite

On pensera principalement aux espaces $\mathbb{R}$, $\mathbb{C}$ ou bien $\mathbb{R}^n$ dans lesquels les fonctions qui nous intéressent prendront leurs valeurs. Tous ces exemples sont des espaces topologiques métrisables séparables.

On a vu que la topologie d'un espace métrisable séparable X est à *base dénombrable*, c'est à dire qu'il existe une suite $(\mathrm{U}_n)_{n\in\mathbb{N}}$ d'ouverts de X telle que tout autre ouvert V de X soit réunion d'une sous-famille de cette suite ; cela revient à dire que pour tout ouvert V de l'espace topologique X, on a

$$\forall x \in \mathrm{V},\ \exists n \in \mathbb{N},\quad x \in \mathrm{U}_n \subset \mathrm{V}.$$

Si (U_m) est une base de la topologie de X et (V_n) une base de la topologie de Y, la famille dénombrable des $\mathrm{U}_m \times \mathrm{V}_n$ est une base de la topologie de $\mathrm{X} \times \mathrm{Y}$ (très facile).

Théorème. *Si* X_1 *et* X_2 *sont deux espaces métrisables séparables, on a*

$$\mathcal{B}_{\mathrm{X}_1\times\mathrm{X}_2} = \mathcal{B}_{\mathrm{X}_1} \otimes \mathcal{B}_{\mathrm{X}_2}.$$

Démonstration. Montrons d'abord que $\mathcal{B}_{\mathrm{X}_1\times\mathrm{X}_2} \subset \mathcal{B}_{\mathrm{X}_1}\otimes\mathcal{B}_{\mathrm{X}_2}$. Pour cela il suffit de montrer que tout ouvert V du produit $\mathrm{X}_1 \times \mathrm{X}_2$ appartient à $\mathcal{B}_{\mathrm{X}_1} \otimes \mathcal{B}_{\mathrm{X}_2}$. Si (U_m) est une base de la topologie de X_1 et (V_n) une base de la topologie de X_2, on a dit que la famille dénombrable des $\mathrm{U}_m\times\mathrm{V}_n$ est une base de la topologie de $\mathrm{X}_1\times\mathrm{X}_2$. Chaque $\mathrm{U}_m\times\mathrm{V}_n$ est le produit d'un borélien de X_1 par un borélien de X_2, donc $\mathrm{U}_m \times \mathrm{V}_n \in \mathcal{B}_{\mathrm{X}_1} \otimes \mathcal{B}_{\mathrm{X}_2}$. Puisque V est réunion dénombrable de certains des $\mathrm{U}_m \times \mathrm{V}_n$, on déduit que $\mathrm{V} \in \mathcal{B}_{\mathrm{X}_1} \otimes \mathcal{B}_{\mathrm{X}_2}$.

L'inclusion inverse $\mathcal{B}_{\mathrm{X}_1} \otimes \mathcal{B}_{\mathrm{X}_2} \subset \mathcal{B}_{\mathrm{X}_1\times\mathrm{X}_2}$ est vraie pour tout couple d'espaces topologiques X_1 et X_2 ; il faut maintenant montrer que pour tous boréliens A_1 de X_1 et A_2 de X_2, le produit $\mathrm{A}_1 \times \mathrm{A}_2$ est un borélien de $\mathrm{X}_1 \times \mathrm{X}_2$. On introduit d'abord, lorsque V_2 est un ouvert fixé de X_2, la classe

$$\mathcal{C}_1 = \{\mathrm{A}_1 \in \mathcal{B}_{\mathrm{X}_1} : \mathrm{A}_1 \times \mathrm{V}_2 \in \mathcal{B}_{\mathrm{X}_1\times\mathrm{X}_2}\}\,;$$

on vérifie que $\mathcal{C}_1$ est une tribu, et $\mathcal{C}_1$ contient les ouverts de X_1, donc $\mathcal{C}_1 = \mathcal{B}_{\mathrm{X}_1}$ pour tout V_2, ce qui veut dire que $\mathrm{A}_1 \times \mathrm{V}_2 \in \mathcal{B}_{\mathrm{X}_1\times\mathrm{X}_2}$ pour tous $\mathrm{A}_1 \in \mathcal{B}_{\mathrm{X}_1}$ et V_2 ouvert de X_2. Dans un deuxième temps on considère pour A_1 fixé

$$\mathcal{C}_2 = \{\mathrm{A}_2 \in \mathcal{B}_{\mathrm{X}_2} : \mathrm{A}_1 \times \mathrm{A}_2 \in \mathcal{B}_{\mathrm{X}_1\times\mathrm{X}_2}\}\,;$$

de même, $\mathcal{C}_2$ est une tribu, qui contient les ouverts V_2 d'après le premier pas, donc $\mathcal{C}_2 = \mathcal{B}_{\mathrm{X}_2}$, ce qui donne ce que nous voulons.

Définition. On dit qu'une application f entre deux espaces topologiques X et Y est *borélienne* si elle est mesurable de $(\mathrm{X}, \mathcal{B}_\mathrm{X})$ dans $(\mathrm{Y}, \mathcal{B}_\mathrm{Y})$. En particulier, toute application continue est borélienne.

Proposition. *Soient* φ *une fonction borélienne de* $\mathbb{R}^n$ *dans* $\mathbb{R}$ *et* $f_1, f_2, \ldots, f_n$ *des fonctions mesurables de* $(\Omega, \mathcal{A})$ *dans* $(\mathbb{R}, \mathcal{B}_\mathbb{R})$. *L'application* $\omega \to \varphi(f_1(\omega), \ldots, f_n(\omega))$ *est mesurable de* $(\Omega, \mathcal{A})$ *dans* $(\mathbb{R}, \mathcal{B}_\mathbb{R})$.

Démonstration. Faisons-la pour $n = 2$. L'application couple $(\omega \to (f_1(\omega), f_2(\omega))$ est mesurable de $(\Omega, \mathcal{A})$ dans $(\mathbb{R}^2, \mathcal{B}_\mathbb{R} \otimes \mathcal{B}_\mathbb{R})$, et φ est par définition mesurable de $(\mathbb{R}^2, \mathcal{B}_{\mathbb{R}^2})$ dans $(\mathbb{R}, \mathcal{B}_\mathbb{R})$. Puisque $\mathcal{B}_\mathbb{R} \otimes \mathcal{B}_\mathbb{R} = \mathcal{B}_{\mathbb{R}^2}$, tout roule.

Proposition. *Si une application f de Ω dans un espace X métrisable est limite simple d'une suite d'applications mesurables de $(\Omega, \mathcal{A})$ dans $(\mathrm{X}, \mathcal{B}_\mathrm{X})$, elle est mesurable.*

Si une application à valeurs dans un X métrisable séparable est mesurable de $(\Omega, \mathcal{A})$ dans $(\mathrm{X}, \mathcal{B}_\mathrm{X})$, elle est limite simple d'une suite d'applications $\mathcal{A}$-étagées.

Démonstration. Supposons que f soit limite simple d'une suite (f_n) d'applications mesurables de $(\Omega, \mathcal{A})$ dans $(\mathrm{X}, \mathcal{B})$. Pour montrer que l'application f est mesurable, il suffit de montrer que $\{f \in \mathrm{U}\} = f^{-1}(\mathrm{U}) \in \mathcal{A}$ pour tout U ouvert de X.

Soit d une distance sur X qui définit la topologie de X ; si $\mathrm{A} \subset \mathrm{X}$ et $x \in \mathrm{X}$, posons $\mathrm{dist}(x, \mathrm{A}) = \inf\{d(x, a) : a \in \mathrm{A}\}$. Considérons pour tout entier $k \geq 0$ l'ensemble ouvert

$$\mathrm{U}_k = \{x \in \mathrm{U} : \mathrm{dist}(x, \mathrm{U}^c) > 2^{-k}\}\,;$$

on obtient une suite croissante dont la réunion est U. Si $f(\omega) \in \mathrm{U}$, il existe un entier k tel que $\mathrm{dist}(f(\omega), \mathrm{U}^c) > 2^{-k}$, donc par la continuité de la distance on a pour n assez grand, disons pour n plus grand qu'un certain m, l'inégalité $\mathrm{dist}(f_n(\omega), \mathrm{U}^c) > 2^{-k}$ c'est à dire que $\omega \in \bigcap_{n \geq m} \{f_n \in \mathrm{U}_k\}$. On a donc

$$\{f \in \mathrm{U}\} \subset \mathrm{B} = \bigcup_k \bigcup_m \bigcap_{n \geq m} \{f_n \in \mathrm{U}_k\}.$$

Inversement si $\omega \in \mathrm{B}$, il existe k et m tels que $\mathrm{dist}(f_n(\omega), \mathrm{U}^c) > 2^{-k}$ pour tout $n \geq m$, ce qui implique que $\mathrm{dist}(f(\omega), \mathrm{U}^c) \geq 2^{-k} > 0$ à la limite, donc $\omega \in \{f \in \mathrm{U}\}$. On a donc bien égalité des deux ensembles $\{f \in \mathrm{U}\}$ et B, et $\mathrm{B} \in \mathcal{A}$ puisqu'il se déduit par des opérations dénombrables des $\{f_n \in \mathrm{U}_k\}$, eux-mêmes dans $\mathcal{A}$ puisque chaque f_n est mesurable et U_k ouvert.

Montrons la partie inverse. Soient $(x_k)_{k \geq 0}$ une suite dense dans X et d une distance sur X qui définisse la topologie de X ; pour tout entier $k \geq 0$ posons $\delta_k(\omega) = \min\{d(f(\omega), x_j) : 0 \leq j \leq k\}$. D'après la densité de la suite, on a $\delta_k(\omega) \to 0$; posons aussi $j_k(\omega) = \min\{j \leq k : d(f(\omega), x_j) = \delta_k(\omega)\}$, et enfin $f_k(\omega) = x_{j_k(\omega)}$. On a bien $d(f_k(\omega), f(\omega)) \to 0$, et on vérifie que f_k est $\mathcal{A}$-étagée puisque

$$\{f_k = x_j\} = \{\omega : d(f(\omega), x_j) = \delta_k(\omega)\} \;\cap \bigcap_{0 \leq i < j} \{\omega : d(f(\omega), x_i) > \delta_k(\omega)\} \in \mathcal{A}$$

pour tout $j = 0, \ldots, k$.

Corrigé de l'exercice 1.1.

a. Si f est mesurable ≥ 0, montrer que $\int f\, d\mu = 0$ si et seulement si $\mu(\{f > 0\}) = 0$.

Supposons que $\{f > 0\}$ soit μ-négligeable. Par définition, $\int f\, d\mu$ est le sup des $\int g\, d\mu$ pour les g étagées telles que $0 \leq g \leq f$. Si on écrit $g = \sum_{i=1}^n a_i \mathbf{1}_{\mathrm{A}_i}$, avec (A_i) partition de Ω en ensembles de $\mathcal{A}$, on aura que $\mathrm{A}_i \subset \{f > 0\}$ si $a_i > 0$, donc $\mu(\mathrm{A}_i) = 0$ et $a_i\, \mu(\mathrm{A}_i) = 0$ dans ce cas ; si $a_i = 0$ on a aussi $a_i\, \mu(\mathrm{A}_i) = 0$, donc $a_i\, \mu(\mathrm{A}_i) = 0$ pour tout i, donc $\int g\, d\mu = \sum_{i=1}^n a_i\, \mu(\mathrm{A}_i) = 0$.

Inversement si $\int f\, d\mu = 0$, on pose $\mathrm{B}_n = \{f \geq 2^{-n}\}$ pour tout entier $n \geq 0$ et on constate que $0 \leq 2^{-n} \mathbf{1}_{\mathrm{B}_n} \leq f$ donc $0 \leq 2^{-n} \mu(\mathrm{B}_n) \leq \int f\, d\mu = 0$, donc $\mu(\mathrm{B}_n) = 0$; ceci étant vrai pour tout n, on aura $\mu(\bigcup_n \mathrm{B}_n) = 0$, et $\bigcup_n \mathrm{B}_n = \{f > 0\}$.

b. Montrer que si $\int f\, d\mu < \infty$, alors $\mu(\{f = +\infty\}) = 0$.

Pour tout entier n on a $2^n \mathbf{1}_{\{f = +\infty\}} \leq f$ donc $2^n \mu(\{f = +\infty\}) \leq \int f\, d\mu$ ce qui donne $\mu(\{f = +\infty\}) \leq 2^{-n} \int f\, d\mu$ pour tout n et $\mu(\{f = +\infty\}) = 0$ à la limite.

Réparons un oubli avant de continuer.

Proposition. *Si (u_k) est une suite de fonctions mesurables ≥ 0, on a*

$$\int_\Omega \Big(\sum_{k=0}^{+\infty} u_k(\omega)\Big)\, d\mu(\omega) = \sum_{k=0}^{+\infty} \int_\Omega u_k(\omega)\, d\mu(\omega)$$

avec les conventions d'usage.

Exercice 2.1.

Si $\int |g_k|\, d\mu \leq 2^{-k}$ pour tout entier $k \geq 0$, montrer que la suite $(g_k(\omega))$ tend vers 0 pour μ-presque tout $\omega \in \Omega$.

Si $(f_n) \subset \mathrm{L}_p(\Omega, \mathcal{A}, \mu)$ tend vers f en norme L_p, trouver une sous-suite (f_{n_j}) qui tend vers f μ-presque partout.

Théorème de Fisher-Riesz

Exercice 2.2. Pour que X normé soit complet, il (faut et il) suffit que : pour toute famille (u_k) de vecteurs de X, la condition $\sum \|u_k\| < +\infty$ implique que la série de vecteurs $\sum u_k$ converge dans X.

Théorème. *Pour tout $p \in [1 + \infty]$ l'espace $\mathrm{L}_p(\Omega, \mathcal{A}, \mu)$ est un espace de Banach.*

Démonstration. Soit $\sum u_k$ une série d'éléments de l'espace $\mathrm{L}_p(\Omega, \mathcal{A}, \mu)$ telle que $\mathrm{M} = \sum_{k=0}^{+\infty} \|u_k\|_p < +\infty$; nous devons montrer que la série $\sum u_k$ converge dans L_p. Posons $v_k = |u_k|$, $g_n = \big(\sum_{k=0}^n v_k\big)^p$, remarquons que $\|v_k\|_p = \|u_k\|_p$ pour obtenir $\int g_n(s)\, d\mu(s) = \|\sum_{k=0}^n v_k\|_p^p \leq \big(\sum_{k=0}^n \|v_k\|_p\big)^p \leq \mathrm{M}^p$. La suite (g_n) est une suite croissante de fonctions mesurables ≥ 0, elle converge vers une fonction mesurable g (valeur $+\infty$ admise) dont la valeur en chaque point est $g(s) = \big(\sum_{k=0}^{+\infty} |u_k(s)|\big)^p$ (valeur $+\infty$ admise à nouveau) ; on sait que $\int g(s)\, d\mu(s) = \lim_n \int g_n(s)\, d\mu(s) \leq \mathrm{M}^p$. La fonction g est donc finie presque partout, donc la série $\sum u_k(s)$ converge absolument pour presque tout s. On pose alors pour presque tout s

$$\mathrm{U}(s) = \sum_{k=0}^{+\infty} u_k(s),$$

on remarque que $|\mathrm{U}(s) - \mathrm{U}_n(s)|^p \leq g(s)$ pour tout s, et que $|\mathrm{U}(s) - \mathrm{U}_n(s)|^p$ tend vers 0 lorsque $n \to +\infty$ pour presque tout s. Comme la fonction g est intégrable, le théorème de convergence dominée de Lebesgue permet de conclure que

$$\|\mathrm{U} - \mathrm{U}_n\|_p^p = \int |\mathrm{U} - \mathrm{U}_n|^p\, d\mu \to 0.$$

L'espace $\mathrm{L}_\infty(\Omega, \mathcal{A}, \mu)$ est l'espace des μ-classes bornées, c'est à dire les classes f contenant une fonction bornée. Si f_1 est un représentant de f tel que $|f_1| \leq \mathrm{M}$ partout, on aura alors $\mu(\{\, |f_2| > \mathrm{M}\}) = 0$ pour tout autre représentant f_2 de f. On peut définir la norme L_∞ de f comme étant la plus petite constante M avec cette propriété,

$$\|f\|_\infty = \min\big\{\mathrm{M} : \mu(\{|f| > \mathrm{M}\}) = 0\big\}.$$

L'espace L_∞ est complet pour cette norme.

Un théorème de classe monotone

Disons (entre nous) que E, espace vectoriel de fonctions réelles bornées sur un ensemble X est un espace vectoriel *monotone* (en abrégé : EVM) s'il est stable par limite simple des suites croissantes bornées, c'est à dire que si $(f_n) \subset \mathrm{E}$ est telle que $|f_n(t)| \leq 1$ pour tout entier n et pour tout $t \in \mathrm{X}$, et si $f_n(t)$ converge en croissant vers $f(t)$ pour tout t, alors $f \in \mathrm{E}$.

Il est clair que toute intersection d'EVM sur X est encore un EVM, d'où la possibilité de définir l'EVM engendré par une classe quelconque $\mathcal{C}$ de fonctions réelles bornées sur X, comme étant l'intersection de tous les EVM qui contiennent $\mathcal{C}$. C'est évidemment le plus petit EVM qui contienne la classe $\mathcal{C}$.

Théorème. *Soit $\mathcal{C}$ une classe de fonctions réelles bornées sur* X*, stable par produit des fonctions ; si* F *désigne le plus petit EVM contenant $\mathcal{C}$, alors* F *est stable par produit.*

Démonstration. Soit φ une fonction réelle bornée sur X ; on pose

$$\mathrm{F}_\varphi = \{f \in \mathrm{F} : \varphi f \in \mathrm{F}\}.$$

Il est clair que F_φ est un espace vectoriel. Supposons φ bornée par le nombre réel $\mathrm{M} \geq 0$; on voit que $(\mathrm{M}-\varphi)$ est ≥ 0 et bornée. Si une suite bornée $(f_n) \subset \mathrm{F}_\varphi$ croît vers f, la suite $((\mathrm{M}-\varphi)f_n)$ est bornée et croît vers $(\mathrm{M}-\varphi)f$, et $((\mathrm{M}-\varphi)f_n) \subset \mathrm{F}$ (parce que $f_n \in \mathrm{F}$, $\varphi f_n \in \mathrm{F}$ et que F est un espace vectoriel), donc $(\mathrm{M}-\varphi)f \in \mathrm{F}$ par la monotonie de F, et $\varphi f \in \mathrm{F}$ puisque F est un espace vectoriel et $\mathrm{M}f \in \mathrm{F}$; finalement, on a $f \in \mathrm{F}_\varphi$. On en déduit que F_φ est un EVM.

Si $\varphi \in \mathcal{C}$, on a d'après l'hypothèse $\varphi f \in \mathcal{C} \subset \mathrm{F}$ pour toute $f \in \mathcal{C}$, donc $\mathcal{C} \subset \mathrm{F}_\varphi$. Puisque F_φ est un EVM et $\mathrm{F}_\varphi \subset \mathrm{F}$, il résulte que $\mathrm{F}_\varphi = \mathrm{F}$, du fait que F est *le plus petit* EVM contenant $\mathcal{C}$; cette égalité $\mathrm{F}_\varphi = \mathrm{F}$ signifie que $\varphi f \in \mathrm{F}$ pour toutes $\varphi \in \mathcal{C}$, $f \in \mathrm{F}$.

Recommençons avec $\varphi \in \mathrm{F}$. D'après le premier point, on a $\varphi f \in \mathrm{F}$ pour toute $f \in \mathcal{C}$, donc à nouveau $\mathcal{C} \subset \mathrm{F}_\varphi$ et $\mathrm{F}_\varphi = \mathrm{F}$. Le décodage de cette formule donne le résultat.

Remarque 1. Si F est un EVM contenant les fonctions constantes et stable par produit, la classe

$$\mathcal{A} = \{\mathrm{A} \subset \mathrm{X} : \mathbf{1}_\mathrm{A} \in \mathrm{F}\}$$

est une tribu de parties de X.

On a d'abord $\emptyset \in \mathcal{A}$ puisque $\mathbf{1}_\emptyset = 0 \in \mathrm{F}$; ensuite si $\mathrm{A} \in \mathcal{A}$ on écrit $\mathbf{1}_{\mathrm{A}^c} = 1 - \mathbf{1}_\mathrm{A} \in \mathrm{F}$ puisque F contient la fonction constante $\mathbf{1}$ égale à 1 en tout point de X. Si $\mathrm{A}, \mathrm{B} \in \mathcal{A}$, alors $\mathbf{1}_{\mathrm{A}\cap\mathrm{B}} = \mathbf{1}_\mathrm{A}\,\mathbf{1}_\mathrm{B}$ montre que $\mathrm{A} \cap \mathrm{B} \in \mathcal{A}$ puisque F est stable par produit ; enfin, si $(\mathrm{A}_n) \subset \mathcal{A}$ admet A pour réunion croissante, la fonction $\mathbf{1}_\mathrm{A}$ est limite croissante bornée de la suite $(\mathbf{1}_{\mathrm{A}_n}) \subset \mathrm{F}$, donc $\mathbf{1}_\mathrm{A} \in \mathrm{F}$ puisque F est monotone, donc $\mathrm{A} \in \mathcal{A}$. On a montré que $\mathcal{A}$ est une tribu.

Remarque 2. Si F est un EVM de fonctions sur Ω, contenant les fonctions constantes et stable par produit, et si $\mathcal{A}$ désigne la tribu $\mathcal{A} = \{\mathrm{A} \subset \Omega : \mathbf{1}_\mathrm{A} \in \mathrm{F}\}$, on voit que $\mathcal{L}_\infty(\Omega, \mathcal{A}) \subset \mathrm{F}$.

Exemple. La mesure de Lebesgue est invariante par translation et par symétrie.

Plus généralement, on va montrer que si deux mesures λ et μ sur la tribu borélienne de $\mathbb{R}^n$ donnent la même mesure finie à tous les pavés (par là nous entendons les ensembles compacts de la forme $\prod_{i=1}^n [a_i, b_i]$), alors $\lambda = \mu$.

Prenons $n = 2$ pour simplifier. Soit E l'espace vectoriel des fonctions boréliennes bornées f telles que

$$\forall \mathrm{K} \text{ pavé}, \quad \int_{\mathrm{K}} f\, d\lambda = \int_{\mathrm{K}} f\, d\mu.$$

L'espace E est un EVM qui contient la classe $\mathcal{C}$ de toutes les fonctions indicatrices $\mathbf{1}_{\mathrm{R}}$ de pavés $\mathrm{R} \subset \mathbb{R}^2$. Cette classe $\mathcal{C}$ est stable par produit (l'intersection de deux pavés est un pavé). D'après le théorème précédent, l'EVM F engendré par $\mathcal{C}$ est stable par produit, et il est contenu dans E.

D'après la remarque **1** précédente, la classe $\mathcal{A}$ des ensembles tels que $\mathbf{1}_{\mathrm{A}} \in \mathrm{F}$ est une tribu ; mais cette tribu contient tous les pavés, donc elle contient la tribu borélienne de $\mathbb{R}^2$ (tout ouvert de $\mathbb{R}^2$ est union dénombrable de pavés). Pour tout $\mathrm{A} \in \mathcal{B}_{\mathbb{R}^2}$, on a $\mathbf{1}_{\mathrm{A}} \in \mathrm{F} \subset \mathrm{E}$, donc $\lambda(\mathrm{K} \cap \mathrm{A}) = \mu(\mathrm{K} \cap \mathrm{A})$ pour tout pavé K ; en utilisant la suite des pavés $\mathrm{K}_n = [-n, n]^2$ on obtient par limite croissante $\lambda(\mathrm{A}) = \mu(\mathrm{A})$ pour tout $\mathrm{A} \in \mathcal{B}_{\mathbb{R}^2}$, ce qu'il fallait démontrer.

Exercice 2.3. Théorème d'Egorov.

On suppose que μ est une mesure ≥ 0 finie sur $(\Omega, \mathcal{A})$.

a. Si la suite (f_n) de fonctions $\mathcal{A}$-mesurables tend simplement vers 0 sur Ω, trouver pour tout $k \geq 0$ un entier n_k tel que

$$\mu\{|f_{n_k}| > 2^{-k}\} \leq 2^{-k}.$$

b. Montrer que pour tout entier k_0, la sous-suite $(f_{n_j})_j$ trouvée en *a* tend uniformément vers 0 sur l'ensemble

$$\mathrm{A}(k_0) = \bigcap_{k \geq k_0} \{|f_{n_k}| \leq 2^{-k}\}$$

et que cet ensemble $\mathrm{A}(k_0)$ a une mesure qui tend vers $\mu(\Omega)$ lorsque $k_0 \to +\infty$.

c. Théorème de Lusin. Montrer que pour toute fonction f borélienne sur $[0, 1]$, il existe pour tout $\varepsilon > 0$ un compact $\mathrm{K}_\varepsilon \subset [0, 1]$ tel que $\lambda([0, 1] \backslash \mathrm{K}_\varepsilon) < \varepsilon$ et que la restriction de f à K_ε soit continue.

Prépa. Agrég écrit d'Analyse, séance n° 3, mercredi 17 octobre 2001.

Exercice 3.1.

a. On désigne par λ la mesure de Lebesgue sur $[0,1]$. Montrer que pour tout borélien B de $[0,1]$ et tout $\varepsilon > 0$ il existe un fermé $\mathrm{F} \subset \mathrm{B}$ tel que $\lambda(\mathrm{B} \setminus \mathrm{F}) < \varepsilon$.

b. Montrer que pour tout $\varepsilon > 0$, il existe une fonction φ continue sur $[0,1]$ telle que $\int_0^1 |\varphi(x) - \mathbf{1}_{\mathrm{B}}(x)|\, dx < \varepsilon$.

c. Montrer que $\mathrm{C}([0,1])$ est dense dans $\mathrm{L}_1([0,1])$.

d. Les résultats obtenus dépendent-ils beaucoup du fait que la mesure soit la mesure de Lebesgue ?

Solution de l'exercice **2.1.**

On suppose $\int |g_k|\, d\mu < 2^{-k}$ pour tout $k \geq 0$. On en déduit

$$\int \Big(\sum_{k=0}^{+\infty} |g_k|\Big)\, d\mu = \sum_{k=0}^{+\infty} \int |g_k|\, d\mu < +\infty$$

donc $\sum_{k=0}^{+\infty} |g_k|$ est fini μ-presque partout, et en particulier $(g_k(\omega))$ tend vers 0 pour μ-presque tout $\omega \in \Omega$.

Si (f_n) tend vers f en norme L_p, on sait que $\int |f_n - f|^p\, d\mu$ tend vers 0, donc on peut trouver pour tout $k \geq 0$ un entier n_k tel que $g_k = |f_{n_k} - f|^p$ vérifie $\int |g_k|\, d\mu < 2^{-k}$; on peut aussi imposer que $n_{k+1} > n_k$; d'après la première partie, on déduit que $(f_{n_k} - f)$ tend vers 0 μ-presque partout.

3.1. Théorème de Fubini

Sur chacun des deux espaces mesurables $(\Omega_i, \mathcal{A}_i)$, $i = 1, 2$, est donnée une mesure (positive) finie μ_i. On considère la classe E des fonctions réelles bornées f sur $\Omega_1 \times \Omega_2$ qui vérifient les deux propriétés suivantes

1– l'application $\omega_1 \in \Omega_1 \to f(\omega_1, \omega_2)$ est $\mathcal{A}_1$-mesurable pour tout $\omega_2 \in \Omega_2$;
2– l'application $\omega_2 \in \Omega_2 \to \int_{\Omega_1} f(\omega_1, \omega_2)\, d\mu_1(\omega_1)$ est $\mathcal{A}_2$-mesurable.

Il est clair que E est un espace vectoriel de fonctions (par les propriétés des fonctions mesurables pour le point 1, et la linéarité de l'intégrale pour le point 2). Si une suite $(f_n) \subset \mathrm{E}$, croissante et bornée, tend vers f, alors f vérifie la première propriété comme limite de fonctions mesurables, et la deuxième par le lemme des suites croissantes appliqué à μ_1. L'espace E est donc un EVM.

Si $f \in \mathrm{E}$ on peut considérer la quantité

$$\mathrm{I}(f) = \int_{\Omega_2} \Big(\int_{\Omega_1} f(\omega_1, \omega_2)\, d\mu_1(\omega_1)\Big)\, d\mu_2(\omega_2).$$

Une deuxième application du lemme des suites croissantes montre que I est continue pour les suites croissantes (bornées) : si $(f_n) \subset \mathrm{E}$ est uniformément bornée et tend en croissant vers f, alors $\mathrm{I}(f) = \lim_n \mathrm{I}(f_n)$. Clairement I est linéaire de E dans $\mathbb{R}$.

Désignons par $\mathcal{C}$ la classe des fonctions $\mathbf{1}_{\mathrm{A}_1 \times \mathrm{A}_2}$, $\mathrm{A}_i \in \mathcal{A}_i$. On vérifie facilement que le produit de deux fonctions de $\mathcal{C}$ est encore dans $\mathcal{C}$ (parce que le produit de $\mathbf{1}_{\mathrm{A}_1 \times \mathrm{A}_2}$ et $\mathbf{1}_{\mathrm{B}_1 \times \mathrm{B}_2}$ est égal à $\mathbf{1}_{\mathrm{A}_1 \cap \mathrm{B}_1 \times \mathrm{A}_2 \cap \mathrm{B}_2}$). On vérifie assez facilement que $\mathcal{C} \subset \mathrm{E}$ et que de

plus $\mathrm{I}(\mathbf{1}_{\mathrm{A}_1\times\mathrm{A}_2}) = \mu_1(\mathrm{A}_1)\mu_2(\mathrm{A}_2)$: en effet, si $f = \mathbf{1}_{\mathrm{A}_1\times\mathrm{A}_2}$, la fonction $f(.,\omega_2)$ est égale à $\mathbf{1}_{\mathrm{A}_2}(\omega_2)\,\mathbf{1}_{\mathrm{A}_1}$, qui est $\mathcal{A}_1$-mesurable pour tout ω_2, et la fonction de la condition 2 est $\mu_1(\mathrm{A}_1)\mathbf{1}_{\mathrm{A}_2}$, qui est $\mathcal{A}_2$-mesurable.

Soit F le plus petit EVM contenant $\mathcal{C}$; on sait qu'il est en plus stable par produit, d'après notre version du lemme des classes monotones, et il contient les fonctions constantes, parce que $\mathbf{1} = \mathbf{1}_{\Omega_1\times\Omega_2} \in \mathcal{C} \subset \mathrm{F}$. D'après la remarque $\mathbf{1}$ de la séance précédente, la classe $\mathcal{A}$ des $\mathrm{A} \subset \Omega_1 \times \Omega_2$ tels que $\mathbf{1}_{\mathrm{A}} \in \mathrm{F}$ est une tribu. Elle contient les pavés mesurables $\mathrm{A} = \mathrm{A}_1 \times \mathrm{A}_2$, donc $\mathcal{A}$ contient toute la tribu $\mathcal{A}_1 \otimes \mathcal{A}_2$.

Evidemment on a $\mathrm{F} \subset \mathrm{E}$, ce qui nous permet de conclure que toutes les fonctions $\mathbf{1}_{\mathrm{A}}$, pour $\mathrm{A} \in \mathcal{A}_1 \otimes \mathcal{A}_2$, sont dans E, c'est à dire qu'elles vérifient les propriétés 1 et 2 de la définition de E, et qu'on peut calculer $\mathrm{I}(\mathbf{1}_{\mathrm{A}})$ pour tout $\mathrm{A} \in \mathcal{A}_1 \otimes \mathcal{A}_2$.

On vérifie facilement que la fonction ν sur $\mathcal{A}_1 \otimes \mathcal{A}_2$ définie par $\nu(\mathrm{A}) = \mathrm{I}(\mathbf{1}_{\mathrm{A}})$ pour tout $\mathrm{A} \in \mathcal{A}_1 \otimes \mathcal{A}_2$ est une mesure sur $\mathcal{A}_1 \otimes \mathcal{A}_2$. Par linéarité on obtient pour toute fonction $\mathcal{A}_1 \otimes \mathcal{A}_2$-étagée $f = \sum_{i=1}^n b_i \mathbf{1}_{\mathrm{B}_i}$ (où $\mathrm{B}_i \in \mathcal{A}_1 \otimes \mathcal{A}_2$, $i = 1, \ldots, n$)

$$\int f\, d\nu = \sum_{i=1}^n b_i\, \nu(\mathrm{B}_i) = \sum_{i=1}^n b_i\, \mathrm{I}(\mathbf{1}_{\mathrm{B}_i}) = \mathrm{I}(f) = \int_{\Omega_2} \Big(\int_{\Omega_1} f(\omega_1, \omega_2)\, d\mu_1(\omega_1) \Big)\, d\mu_2(\omega_2).$$

Puisque l'intégrale par rapport à ν et la forme linéaire I sont continues pour les suites croissantes bornées, on en déduit le même résultat pour toute fonction f qui est $\mathcal{A}_1 \otimes \mathcal{A}_2$-mesurable bornée (en effet, une telle f est limite croissante d'une suite de fonctions $\mathcal{A}_1 \otimes \mathcal{A}_2$-étagées). La formule de définition de I nous donne donc le théorème de Fubini pour les fonctions mesurables bornées.

La mesure ν de la construction précédente est la mesure produit, notée $\mu_1 \otimes \mu_2 = \nu$. Pour les applications en Analyse classique, il est important de ne pas se limiter au cas des mesures finies ; cependant, il n'y a pas de bonne théorie du produit de deux mesures positives quelconques. On va admettre une restriction raisonnable.

Définition. On dit qu'une mesure μ positive sur $(\Omega, \mathcal{A})$ est *σ-finie* s'il existe une suite (A_n) d'éléments de $\mathcal{A}$ telle que $\Omega = \bigcup_{n=0}^{+\infty} \mathrm{A}_n$ et $\mu(\mathrm{A}_n) < +\infty$ pour tout $n \geq 0$.

La mesure de Lebesgue sur $\mathbb{R}^d$, la mesure de comptage sur $\mathbb{N}$ sont des exemples de mesures σ-finies.

Si μ_1 et μ_2 sont σ-finies, on peut trouver deux suites $(\mathrm{A}_{1,n})$ et $(\mathrm{A}_{2,n})$ croissantes d'ensembles tels que $\mu_i(\mathrm{A}_{i,n}) < +\infty$ et $\Omega_i = \bigcup_n \mathrm{A}_{i,n}$. Posons $\mathrm{B}_n = \mathrm{A}_{1,n} \times \mathrm{A}_{2,n}$. On peut appliquer ce qui précède pour former le produit $\nu_n = \mathbf{1}_{\mathrm{A}_{1,n}}\, \mu_1 \otimes \mathbf{1}_{\mathrm{A}_{2,n}}\, \mu_2$; on obtient ainsi une suite croissante (ν_n) de mesures finies, et on pose

$$\forall \mathrm{A} \in \mathcal{A}_1 \otimes \mathcal{A}_2, \quad (\mu_1 \otimes \mu_2)(\mathrm{A}) = \lim_n \nu_n(\mathrm{A}).$$

On obtient ainsi une mesure σ-finie sur le produit.

Passons au théorème de Fubini positif.

Théorème. *Supposons μ_i σ-finie. Pour toute fonction mesurable f de $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$ à valeurs dans $[0, +\infty]$,*

1– l'application $\omega_1 \to f(\omega_1, \omega_2)$ est $\mathcal{A}_1$-mesurable pour tout $\omega_2 \in \Omega_2$;
2– l'application $\omega_2 \to \int_{\Omega_1} f(\omega_1, \omega_2)\, d\mu_1(\omega_1)$ est $\mathcal{A}_2$-mesurable.

Enfin,

$$\int_{\Omega_1\times\Omega_2} f\, d(\mu_1\otimes\mu_2) = \int_{\Omega_2}\Big(\int_{\Omega_1} f(\omega_1,\omega_2)\, d\mu_1(\omega_1)\Big)\, d\mu_2(\omega_2).$$

Démonstration. Il suffit de réaliser f comme limite croissante d'une suite (f_n) de fonctions mesurables bornées, et de réutiliser les mêmes bons vieux arguments. Comme μ_1 et μ_2 sont σ-finies on peut trouver deux suites $(\mathrm{A}_{1,n})$ et $(\mathrm{A}_{2,n})$ croissantes d'ensembles tels que $\mu_i(\mathrm{A}_{i,n}) < +\infty$ et $\Omega_i = \bigcup_n \mathrm{A}_{i,n}$. Posons $\mathrm{B}_n = \mathrm{A}_{1,n}\times\mathrm{A}_{2,n}$. A chaque fonction $\mathbf{1}_{\mathrm{B}_n} f_n$ on peut appliquer le théorème de Fubini déjà vu ; on finit par limite croissante de fonctions mesurables et le lemme des suites croissantes.

Solution de l'exercice **1.2.**

Si on établit que $g(\omega,t) = \mathbf{1}_{\{f(\omega)>t\}}$ est $\mathcal{A}\times\mathcal{B}_{\mathbb{R}}$-mesurable on pourra écrire que

$$\int_0^{+\infty} \mu(\{f>t\})\, dt = \int_0^{+\infty}\int_\Omega \mathbf{1}_{\{f(\omega)>t\}}\, d\mu(\omega)\, dt =$$

$$= \int_\Omega\Big(\int_0^{f(\omega)} dt\Big)\, d\mu = \int_\Omega f(\omega)\, d\mu.$$

Pour la mesurabilité, on voit que $\varphi(x,t) = \mathbf{1}_{\{x>t\}}$ est borélienne sur $\mathbb{R}^2$, comme indicatrice d'un ouvert ; la fonction $\psi(\omega,t) = (f(\omega),t)$ est un couple de fonctions mesurables (le couple de f et de l'identité de $\mathbb{R}$), donc ψ est mesurable de $\mathcal{A}\otimes\mathcal{B}_{\mathbb{R}}$ dans $\mathcal{B}_{\mathbb{R}}\otimes\mathcal{B}_{\mathbb{R}} = \mathcal{B}_{\mathbb{R}^2}$; il n'y a plus qu'à composer ψ et φ.

3.2. Convolution

On se place sur $\mathbb{R}^d$ muni de la mesure de Lebesgue. On commence par le cas de deux fonctions mesurables ≥ 0 et intégrables. On pose

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

Le résultat est à valeurs dans $\overline{\mathbb{R}}$, et les propriétés d'invariance de la mesure de Lebesgue entraînent que

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

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 égale aux deux intégrales

$$\int (f*g)(x)\, dx = \int dy\Big(\int f(x-y)g(y)\, dx\Big).$$

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

$$\int dy\Big(\int f(x-y)g(y)\, dx\Big) = \Big(\int f(x)\, dx\Big)\Big(\int g(y)\, dy\Big) < +\infty.$$

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

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 en majorant par les modules montrent que $f * g$ est intégrable et plus précisément $\|f * g\|_1 \leq \|f\|_1 \|g\|_1$.

Autres cadres pour la convolution

Dans le cas périodique, on s'occupe par exemple 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 $\int_a^{a+2\pi} f(t)\,dt$ ne dépend pas de a. Si f et g sont deux fonctions 2π-périodiques intégrables, on définira leur convolution périodique en posant

$$(f * g)(x) = \int_0^{2\pi} f(x-y)g(y)\,dy.$$

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

Exercice 3.2. Soient $f, g \in \mathrm{L}_1(\mathbb{R}^d)$; calculer la transformée de Fourier de $f * g$.

3.3. Deux ou trois principes de convexité

Soit X un espace vectoriel réel ; un sous-ensemble $\mathrm{C} \subset \mathrm{X}$ est dit *convexe* si $[x, y] \subset \mathrm{C}$ pour tous $x, y \in \mathrm{C}$, où $[x, y]$ désigne le segment fermé entre x et y,

$$[x, y] = \{(1-t)x + ty : 0 \leq t \leq 1\}.$$

Une fonction réelle f définie sur C est convexe si

$$f\big((1-t)x + ty\big) \leq (1-t)f(x) + tf(y)$$

pour tous $x, y \in \mathrm{C}$ et tout $t \in [0, 1]$.

Critères avec dérivée seconde ou moins

Proposition. *Pour que f continue sur un intervalle $[a, b]$ soit convexe sur $[a, b]$ il (faut et il) suffit que*

$$\forall x \in]a, b[, \quad \limsup_{t \to 0} \frac{f(x-t) + f(x+t) - 2f(x)}{t^2} \geq 0.$$

Démonstration. Posons quand $a < x \pm t < b$

$$\mathrm{D}_2 f(x, t) = \frac{f(x-t) + f(x+t) - 2f(x)}{t^2}.$$

En ajoutant εx^2 à f, pour $\varepsilon > 0$, on obtient $\limsup_t \mathrm{D}_2 g(x, t) > 0$ pour la fonction modifiée g définie par $g(x) = f(x) + \varepsilon x^2$, puisque $\mathrm{D}_2 g(x, t) = \mathrm{D}_2 f(x, t) + 2\varepsilon$. On va

montrer que g est convexe, et on concluera ensuite, en faisant tendre ε vers 0, que f elle-même est convexe.

Si $x_0 < x_1$ sont deux points de l'intervalle $[a, b]$, on doit voir que $k = g - h \leq 0$ sur $[x_0, x_1]$, où h est la fonction affine égale à g aux points x_0 et x_1. Comme $\mathrm{D}_2 h(x,t) = 0$ en tout point, on a encore $\limsup_t \mathrm{D}_2 k(x,t) > 0$, et de plus $k(x_0) = k(x_1) = 0$. Nous voulons montrer que $k \leq 0$ sur l'intervalle $[x_0, x_1]$. En un point $\xi \in]x_0, x_1[$ où le max de k serait atteint à l'intérieur, on aurait, pour tout t assez petit pour que $\xi \pm t \in [x_0, x_1]$, les inégalités $k(\xi - t) \leq k(\xi)$, $k(\xi + t) \leq k(\xi)$, donc $\mathrm{D}_2 k(\xi, t) \leq 0$, d'où $\limsup_t \mathrm{D}_2(\xi, t) \leq 0$ ce qui est contradictoire ; donc le max de k est au bord, et par conséquent il est nul.

Corollaire. *Si la dérivée seconde est ≥ 0 sur un intervalle ouvert, la fonction est convexe.*

Exercice 3.3. Soient g et h deux fonctions continues strictement croissantes et inverses l'une de l'autre, telles que $g(0) = 0$; montrer que

$$\int_0^x g(s)\, ds = \int_0^{g(x)} h(t)\, dt = xg(x).$$

Solution de l'exercice **2.3.** *sur Egorov et Lusin*

Commençons par Egorov ; on suppose que μ est une mesure finie sur $(\Omega, \mathcal{A})$ et que la suite de fonctions $\mathcal{A}$-mesurables (f_n) tend vers 0 μ-presque partout.

Pour $\varepsilon > 0$ fixé, la suite (g_n) définie par $g_n(\omega) = \mathbf{1}_{\{|f_n(\omega)| > \varepsilon\}}$ tend vers zéro μ-presque partout : en effet, pour tout ω tel que $\lim_n f_n(\omega) = 0$, on aura $|f_n(\omega)| < \varepsilon$ pour $n \geq n_0(\omega)$, donc $g_n(\omega) = 0$ pour $n \geq n_0(\omega)$. De plus, cette convergence vers 0 presque partout de la suite (g_n) est dominée par la fonction intégrable fixe $\mathbf{1}$ (la mesure est finie), donc $\int g_n\, d\mu = \mu(\{|f_n| > \varepsilon\})$ tend vers 0 par le théorème de convergence dominée. On peut en particulier pour tout m_0 donné, trouver un indice $m = m(\varepsilon) > m_0$ tel que $\mu(\{|f_m| > \varepsilon\}) < \varepsilon$.

En appliquant ce qui précède avec $\varepsilon = 2^{-k}$, on peut construire une suite (n_k) strictement croissante telle que $\mu(\{|f_{n_k}| > 2^{-k}\}) < 2^{-k}$ pour tout $k \geq 0$.

Posons

$$\mathrm{A}(k_0) = \bigcap_{j > k_0} \{|f_{n_j}| \leq 2^{-j}\}.$$

La mesure du complémentaire $\mathrm{A}_{k_0}^c$ vérifie

$$\mu(\mathrm{A}_{k_0}^c) = \mu\Big(\bigcup_{j > k_0} \{|f_{n_j}| > 2^{-j}\}\Big) \leq \sum_{j > k_0} \mu(\{|f_{n_j}| > 2^{-j}\}) \leq \sum_{j > k_0} 2^{-j} = 2^{-k_0}$$

qui tend bien vers 0 lorsque $k_0 \to +\infty$. La convergence uniforme sur $\mathrm{X} = \mathrm{A}_{k_0}$ de la sous-suite (f_{n_j}) est évidente puisque $\|f_{n_j}\|_{\mathrm{L}_\infty(\mathrm{X})} \leq 2^{-j}$.

Passons au théorème de Lusin. Si f est une fonction borélienne bornée sur $[0, 1]$, il existe une suite (φ_n) de fonctions continues sur $[0, 1]$ qui tend vers f dans L_1 ; en passant à une sous-suite on peut supposer que la convergence a lieu presque partout ; ensuite par Egorov appliqué aux différences $(\varphi_n - f)$ on trouve une sous-suite (φ_{n_j}) et un borélien B tels que $\lambda(\mathrm{B}^c) < \varepsilon/2$ et que la convergence vers f soit uniforme sur l'ensemble B.

On sait (exercice **3.1.**) qu'il existe un fermé $\mathrm{F} \subset \mathrm{B}$ tel que $\lambda(\mathrm{B} \setminus \mathrm{F}) < \varepsilon/2$; on voit que $\lambda(\mathrm{F}^c) < \varepsilon$, et la fonction f restreinte à F est limite uniforme d'une suite de fonctions continues, donc elle est continue sur F.

Prépa. Agrég écrit d'Analyse, séance n° 4, mercredi 24 octobre 2001.

4. Autour de la convexité

On va exploiter le résultat suivant, vu dans la séance n° 3.

Proposition. *Pour que* F *continue sur un intervalle* $[a,b]$ *soit convexe sur* $[a,b]$ *il (faut et il) suffit que*

$$\forall x \in]a,b[, \quad \limsup_{t\to 0} \frac{\mathrm{F}(x-t)+\mathrm{F}(x+t)-2\mathrm{F}(x)}{t^2} \geq 0.$$

Remarque. Si F est continue sur $[a,b]$, il suffit de regarder les milieux pour montrer la convexité : si $\mathrm{F}\big((x+y)/2\big) \leq \frac{1}{2}\,\mathrm{F}(x) + \frac{1}{2}\,\mathrm{F}(y)$ pour tous $x,y \in [a,b]$, alors F est convexe sur $[a,b]$. En effet, l'expression à étudier dans la proposition précédente est alors ≥ 0 pour tous x et t tels que $x \pm t \in [a,b]$.

Corollaire. *Si la dérivée seconde est* ≥ 0 *sur un intervalle ouvert, la fonction est convexe.*

Corollaire. *Si* f *est une fonction croissante sur un intervalle* I, *si* $a \in \mathrm{I}$ *et si on pose* $\mathrm{F}(x) = \int_a^x f(t)\,dt$ *pour tout* x *de* I, *alors la fonction* F *est convexe sur* I.

Démonstration. On écrit pour $t > 0$ assez petit pour que $x \pm t \in \mathrm{I}$

$$\mathrm{F}(x-t)+\mathrm{F}(x+t)-2\mathrm{F}(x) = \int_x^{x+t} f(s)\,ds - \int_{x-t}^{x} f(s)\,ds =$$
$$\int_x^{x+t} \big(f(s)-f(s-t)\big)\,ds \geq 0.$$

Les fonctions puissance

La fonction $x \to |x|^p$ est convexe sur $\mathbb{R}$ quand $1 \leq p < +\infty$. On peut par exemple voir que $f(t) = \mathrm{sign}(t)\,|t|^{p-1}$ est croissante sur $\mathbb{R}$ et que

$$|x|^p = \int_0^x f(t)\,dt$$

pour tout $x \in \mathbb{R}$.

Normes sur un espace vectoriel

Lemme. *Soient* X *un espace vectoriel sur* $\mathbb{K} = \mathbb{R}$ *ou* $\mathbb{C}$ *et* q *une fonction* **à valeurs** ≥ 0 *définie sur* X *; pour que* q *soit une semi-norme sur* X, *(il faut et) il suffit que :*

pour tous $x \in \mathrm{X}$ *et* $\lambda \in \mathbb{K}$ *on a* $q(\lambda x) = |\lambda|\,q(x)$, *et l'ensemble* $\{x \in \mathrm{X} : q(x) \leq 1\}$ *est convexe.*

Démonstration. Si q est une semi-norme, il est clair que $\mathrm{C}_q = \{x \in \mathrm{X} : q(x) \leq 1\}$ est convexe. Inversement, supposons que q soit positivement homogène et que C_q soit convexe, et déduisons la sous-additivité de q ; soient x et y deux vecteurs de X, et choisissons $a > q(x) \geq 0$ et $b > q(y) \geq 0$; considérons les deux vecteurs de C_q définis par $x_1 = a^{-1}x$ et $y_1 = b^{-1}y$, puis formons la combinaison convexe

$$z = \frac{a}{a+b}\,x_1 + \frac{b}{a+b}\,y_1,$$

qui est dans C_q d'après l'hypothèse de convexité, c'est à dire que $q(z) \leq 1$. On vérifie immédiatement que $z = (a+b)^{-1}(x+y)$, et l'homogénéité de q transforme alors l'inégalité $q(z) \leq 1$ en $q(x+y) \leq a+b$. En faisant tendre a vers $q(x)$ et b vers $q(y)$ on obtient $q(x+y) \leq q(x)+q(y)$.

Exemple. Pour $1 \leq p < +\infty$, soit $\mathcal{L}_p = \mathcal{L}_p(\Omega, \mathcal{A}, \mu)$ l'espace vectoriel des fonctions f complexes $\mathcal{A}$-mesurables telles que $\int_\Omega |f(s)|^p \, d\mu(s) < +\infty$; la quantité

$$q(f) = \Big(\int_\Omega |f(s)|^p \, d\mu(s) \Big)^{1/p}$$

est une semi-norme sur $\mathcal{L}_p$.

Pour le vérifier, on voit d'abord que $q(\lambda f) = |\lambda| \, q(f)$ (facile), puis on montre que l'ensemble $\{f \in \mathcal{L}_p : q(f) \leq 1\}$ est convexe. Cela provient de la convexité sur $[0, +\infty[$ de la fonction $u \to u^p$; on a alors si f, g sont deux éléments de $\mathcal{L}_p$ tels que $q(f) \leq 1$, $q(g) \leq 1$ et si $0 \leq t \leq 1$,

$$|(1-t)f(s) + tg(s)|^p \leq \big((1-t)|f(s)| + t|g(s)|\big)^p \leq (1-t)|f(s)|^p + t|g(s)|^p$$

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

$$\int_\Omega \big|(1-t)f(s)+tg(s)\big|^p \, d\mu(s) \leq (1-t) \int_\Omega |f(s)|^p \, d\mu(s) + t \int_\Omega |g(s)|^p \, d\mu(s) \leq (1-t)+t = 1.$$

Désignons par $\mathcal{N}$ l'ensemble des fonctions $\mathcal{A}$-mesurables scalaires f telles que $f = 0$ μ-presque partout. On voit facilement que $\mathcal{N}$ est un espace vectoriel, contenu dans $\mathcal{L}_p$ pour tout p, ce qui permet de considérer le quotient $\mathrm{L}_p(\Omega, \mathcal{A}, \mu) = \mathcal{L}_p(\Omega, \mathcal{A}, \mu)/\mathcal{N}$. On vérifie sans peine que la fonction q "passe au quotient", parce que sa valeur est constante sur les classes d'équivalence. Quand on passe aux classes, on obtient une norme sur l'espace $\mathrm{L}_p(\Omega, \mathcal{A}, \mu)$, notée $\|f\|_p$,

$$\forall f \in \mathrm{L}_p(\Omega, \mathcal{A}, \mu), \quad \|f\|_p = \Big(\int_\Omega |f(s)|^p \, d\mu(s) \Big)^{1/p}.$$

Inégalité de Young

Théorème. *On suppose que f est continue, strictement croissante sur $\mathbb{R}$, $f(\mathbb{R}) = \mathbb{R}$ et $f(0) = 0$. On désigne par g la fonction réciproque de f, et on pose $\mathrm{F}(x) = \int_0^x f(s) \, ds$, $\mathrm{G}(y) = \int_0^y g(t) \, dt$. On a alors*

$$\forall x, y \in \mathbb{R}, \quad xy \leq \mathrm{F}(x) + \mathrm{G}(y)$$

avec égalité quand $y = f(x)$.

Démonstration (inspirée de celle qui est déjà sur le web). Commençons par démontrer le cas d'égalité. Supposons $x > 0$ pour fixer les idées, et considérons $y = f(x) > f(0) = 0$. Puisque f et g sont continues, on sait que les intégrales $\int_0^x f(s) \, ds$ et $\int_0^y g(t) \, dt$ sont limites de sommes de Riemann

$$\sum_{i=0}^{n-1} (x_{i+1} - x_i) \, f(\xi_i) \quad \text{et} \quad \sum_{i=0}^{n-1} (y_{i+1} - y_i) \, g(\eta_i)$$

à la seule condition que les pas des deux subdivisions $x_0 = 0 < x_1 < \ldots < x_n = x$ de $[0, x]$ et $y_0 = 0 < y_1 < \ldots < y_n = y = f(x)$ de $[0, y]$, égaux à $\sup_{0 \leq i < n}(x_{i+1} - x_i)$ et $\sup_{0 \leq i < n}(y_{i+1} - y_i)$, tendent vers 0. Soit $n \geq 1$; on découpe $[0, x]$ en n segments de longueur x/n, c'est à dire que $x_i = ix/n$, où $i = 0, \ldots, n$, et on découpe $[0, y]$ au moyen

des points $y_i = f(x_i)$, $i = 0, \ldots, n$. Puisque f est uniformément continue sur le compact $[0, x]$, on sait que pour n assez grand, on aura $\sup_{0 \le i < n}(y_{i+1} - y_i) < \varepsilon$.

Pour f, on choisit la somme de Riemann

$$\sum_{i=0}^{n-1} (x_{i+1} - x_i)\, f(x_i)$$

et pour g

$$\sum_{i=0}^{n-1} (y_{i+1} - y_i)\, g(y_{i+1}) = \sum_{i=0}^{n-1} x_{i+1} \big(f(x_{i+1}) - f(x_i)\big).$$

La somme des deux expressions est égale à

$$\sum_{i=0}^{n-1} \big(x_{i+1} f(x_{i+1}) - x_i f(x_i)\big) = x_n\, f(x_n) - x_0\, f(x_0) = x\, f(x).$$

Quand n tend vers l'infini, il en résulte que

$$\int_0^x f(s)\, ds + \int_0^{f(x)} g(t)\, dt = x\, f(x).$$

Montrons maintenant l'inégalité de Young annoncée. On fixe $\overline{y} \in \mathbb{R}$ et on étudie la fonction $\varphi : x \to x\,\overline{y} - \mathrm{F}(x)$; sa dérivée $\overline{y} - f(x)$ est positive pour $x < \overline{x} = g(\overline{y})$, puis négative, donc la fonction φ admet un maximum pour $\overline{x}$; on a donc pour tout $x \in \mathbb{R}$

$$x\,\overline{y} - \mathrm{F}(x) \le \overline{x}\,\overline{y} - \mathrm{F}(\overline{x}) = \overline{x}\, f(\overline{x}) - \mathrm{F}(\overline{x}) = \mathrm{G}(f(\overline{x})) = \mathrm{G}(\overline{y})$$

ce qui est le résultat voulu.

Définition. On dit que deux fonctions convexes telles que F et G sont deux fonctions convexes conjuguées.

Exemple de fonctions conjuguées : $p^{-1}t^p$ et $q^{-1}t^q$. Pour $p \in [1, +\infty]$, on appelle *exposant conjugué* de p le nombre $q \in [1, +\infty]$ tel que $1/p + 1/q = 1$. Cette relation est symétrique ; on dit que (p, q) est un couple *d'exposants conjugués*. On notera que si $1 < p < +\infty$, cela implique que $q(p-1) = p$ et de façon symétrique, $p(q-1) = q$; on notera aussi que $(p-1)(q-1) = 1$.

Dans le cas $1 < p < +\infty$, la dérivée de $\mathrm{F}(x) = p^{-1}|x|^p$ est égale à $f(x) = \mathrm{sign}(x)\,|x|^{p-1}$ et celle de $\mathrm{G}(y) = q^{-1}|y|^q$ est $g(y) = \mathrm{sign}(y)\,|y|^{q-1}$; puisque $(p-1)(q-1) = 1$, ces deux fonctions (strictement croissantes) f et g sont bien inverses l'une de l'autre. On a donc l'inégalité *de Young*

$$\forall x, y \in \mathbb{R}, \quad x\,y \le \frac{1}{p}\,|x|^p + \frac{1}{q}\,|y|^q,$$

inégalité qu'on peut évidemment démontrer directement et facilement (surtout lorsque $p = q = 2$).

Inégalité de Hölder

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*

$$\Big|\int_\Omega fg\, d\mu\Big| \leq \|f\|_p\, \|g\|_q.$$

Démonstration. Pour alléger un peu, on écrira simplement $\int f$ au lieu de $\int_\Omega f(s)\, d\mu(s)$ chaque fois que possible. 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 $|fg| \leq \mathrm{M}\, |g|$, donc fg est intégrable et

$$\Big|\int fg\Big| \leq \int |fg| \leq \mathrm{M} \int |g| = \|f\|_\infty\, \|g\|_1.$$

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

$$tu \leq \frac{1}{p}\, t^p + \frac{1}{q}\, u^q.$$

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

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

ce qui montre que fg est intégrable, et que

$$\Big|\int fg\Big| \leq \frac{1}{p} \int |f|^p + \frac{1}{q} \int |g|^q.$$

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 $|\int fg| \leq 1/p + 1/q = 1$, ce qui est le résultat voulu.

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

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

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

$$\|f\|_p \geq \sup\{\Big|\int_\Omega fg\, d\mu\Big| : \|g\|_q \leq 1\},$$

le problème est de montrer l'autre direction. On va voir qu'en fait le *maximum* est atteint pour une certaine fonction $g \in \mathrm{L}_q$, $\|g\|_q \leq 1$, lorsque $1 \leq p < +\infty$. Si $f = 0$, le résultat est évident, on supposera donc $f \neq 0$, et par homogénéité on peut se ramener à $\|f\|_p = 1$. Soit $\widetilde{f}$ une "vraie" fonction mesurable de la classe f, et définissons une fonction mesurable g sur l'ensemble Ω en posant $g(s) = |\widetilde{f}(s)|^p / \widetilde{f}(s)$ sur l'ensemble mesurable $\mathrm{A} = \{s \in \Omega : \widetilde{f}(s) \neq 0\}$, 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 $|g|^q = |f|^p$, donc $\int |g|^q = 1$, soit encore $\|g\|_q = 1$; pour $p = 1$, $g(s)$ est de module 1 quand $s \in \mathrm{A}$ donc $\|g\|_\infty = 1$. D'autre part

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

Dans le cas $p = +\infty$, le maximum n'est pas nécessairement atteint ; cependant, si on suppose $\|f\|_\infty = 1$, l'ensemble mesurable $\mathrm{B}_\varepsilon = \{s \in \Omega : |f(s)| > 1 - \varepsilon\}$ est de mesure > 0 pour tout $\varepsilon > 0$. On choisira $\varepsilon < 1$ et 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 $\int_\Omega fg\, d\mu > 1 - \varepsilon$.

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 $g \in \mathrm{L}_q \to \int fg$ 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 $\mu \geq 0$. 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 $\mathrm{L}_\infty \neq \{0\}$ (il contient les fonctions bornées) ne peut pas être le dual de cet L_1.

Exercice 4.1. Transformée de Legendre. On suppose que F est une fonction réelle strictement convexe sur $\mathbb{R}^d$, de classe C^1 et telle que $\lim_{\|x\| \to +\infty} \mathrm{F}(x)/\|x\| = +\infty$.

a. Montrer que pour tout $y \in \mathbb{R}^d$, la fonction $x \to x \,.\, y - \mathrm{F}(x)$ atteint son maximum sur $\mathbb{R}^d$.

b. Montrer que $\nabla\mathrm{F}$ est surjective, injective, propre, et un homéomorphisme de $\mathbb{R}^d$ sur $\mathbb{R}^d$.

c. Montrer que $\mathrm{G}(y) = \max_{x \in \mathbb{R}^d} \big[x \,.\, y - \mathrm{F}(x)\big]$ a les mêmes propriétés que F, et que $\nabla\mathrm{G}$ est la bijection réciproque de $\nabla\mathrm{F}$.

Fonctions affines tangentes, 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*

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

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 $h \leq \varphi$ partout. On a ensuite

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

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 ; en effet, on a $-\varphi(f) \leq -af - b \leq |a|\,|f| + |b|$ qui est intégrable.

Inégalité de Minkowski

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 $p \geq 1$, on a*

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

Démonstration. Posons $\mathrm{M} = \int_\mathrm{Y} \big(\int_\mathrm{X} |f(x,y)|^p\, d\mu(x)\big)^{1/p} d\nu(y)$. Si $\mathrm{M} = +\infty$, la relation à démontrer est évidente. Si $\mathrm{M} = 0$, on a $\int_\mathrm{X} |f(x,y)|^p\, d\mu(x) = 0$ pour presque tout y, donc f est négligeable pour la mesure produit et l'autre expression est nulle aussi. Lorsque $0 < \mathrm{M} < 1$, on peut se ramener à $\mathrm{M} = 1$ par homogénéité, et considérer la densité de probabilité

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

Supposons d'abord que $g(y) > 0$ pour tout $y \in \mathrm{Y}$, pour ne pas obscurcir l'argument par des complications techniques sans grand intérêt. Ecrivons

$$\Big(\int_\mathrm{Y} |f(x,y)|\, d\nu(y)\Big)^p = \Big(\int_\mathrm{Y} |g(y)^{-1} f(x,y)|\, g(y)\, d\nu(y)\Big)^p \leq$$

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

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

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

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

Réglons le petit problème technique qui se pose lorsque g peut s'annuler sur Y. Posons $\mathrm{Y}_0 = \{g > 0\}$ et $\mathrm{B} = \{(x,y) : |f(x,y)| > 0,\ g(y) = 0\}$. On aura par Fubini

$$\iint \mathbf{1}_\mathrm{B}(x,y) |f(x,y)|^p\, d\mu(x) d\nu(y) = \int_{\mathrm{Y}_0^c} \big(\int 1_\mathrm{B}(x,y) |f(x,y)|^p\, d\mu(x)\big)\, d\nu(y) \leq$$

$$\leq \int_{\mathrm{Y}_0^c} g(y)^p\, d\nu(y) = 0$$

ce qui implique que B est $\mu \otimes \nu$-négligeable. On a donc $\int_{\mathrm{Y}} \mathbf{1}_{\mathrm{B}}(x,y)\,d\nu(y) = 0$ pour μ-presque tout x, et $\int_{\mathrm{Y}} |f(x,y)|\,d\nu(y) = \int_{\mathrm{Y}_0} |f(x,y)|\,d\nu(y)$ pour μ-presque tout x, donc

$$\int_{\mathrm{X}} \Big(\int_{\mathrm{Y}} |f(x,y)|\,d\nu(y) \Big)^{1/p} d\mu(x) = \int_{\mathrm{X}} \Big(\int_{\mathrm{Y}_0} |f(x,y)|\,d\nu(y) \Big)^{1/p} d\mu(x).$$

Il suffit donc de remplacer toutes les intégrales sur Y par des intégrales sur Y_0 pour justifier l'écriture avec $g(y)^{-1}$.

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

$$\Big(\int_{\mathrm{X}} \Big| \sum_{j=1}^n g_{y_j}(x) \Big|^p d\mu(x) \Big)^{1/p} = \|h\|_p \leq \sum_{j=1}^n \|g_{y_j}\|_p = \sum_{j=1}^n \Big(\int_{\mathrm{X}} |g_{y_j}(x)|^p\,d\mu(x) \Big)^{1/p}.$$

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

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

L'inégalité de Minkowski dit que la norme de h dans $\mathrm{L}_p(\mathrm{X})$ est majorée par l'intégrale par rapport au paramètre des normes dans $\mathrm{L}_p(\mathrm{X})$ des morceaux g_y.

Exercice 4.2. On suppose $0 < r < p < +\infty$. Montrer que

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

Prépa. Agrég écrit d'Analyse, séance n° 5, mercredi 7 novembre 2001.

Solution de l'exercice 3.2 : Fourier d'une convolution

On a vu que si f_1 et g_1 sont deux fonctions intégrables sur $\mathbb{R}^d$, alors $f_1 * g_1$ est intégrable et $\int f_1 * g_1 = (\int f_1)(\int g_1)$ (application directe de Fubini). On applique ceci, pour t fixé, avec $f_1(x) = f(x)\,\mathrm{e}^{-ixt}$ et $g_1(x) = g(x)\,\mathrm{e}^{-ixt}$. On obtient

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

et

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

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

Exercice 5.1.

a. On considère la fonction croissante continue φ définie sur $[0,+\infty[$ par $\varphi(t) = t$ si $0 \le t \le 1$ et $\varphi(t) = 1 + \ln(t)$ si $t \ge 1$. Calculer $\Phi(x) = \int_0^x \varphi(t)\,dt$ pour $x \ge 0$.

b. On considère l'ensemble X des (classes de) fonctions mesurables f sur $[0,1]$ telles que

$$\int_0^1 |f(t)|\,\ln_+(|f(t)|)\,dt < +\infty$$

où on a posé $\ln_+(x) = \ln(x)$ pour $x \ge 1$ et $\ln_+(x) = 0$ pour $x \le 1$.

Montrer que X est un espace vectoriel, contenu dans $\mathrm{L}_1(0,1)$. Montrer que $f \in \mathrm{X}$ si et seulement s'il existe $\lambda > 0$ tel que

$$\int_0^1 \Phi\Big(\frac{|f(t)|}{\lambda}\Big)\,dt \le \Phi(1).$$

Montrer que

$$\mathrm{N}(f) = \inf\{\lambda > 0 : \int_0^1 \Phi\Big(\frac{|f(t)|}{\lambda}\Big)\,dt \le \Phi(1)\}$$

définit une norme sur X qui le rend complet.

c. Montrer que si $f \in \mathrm{X}$ et si g est une fonction mesurable sur $[0,1]$ telle que

$$\int_0^1 \mathrm{e}^{\varepsilon|g(t)|}\;dt < +\infty$$

pour un $\varepsilon > 0$, la fonction produit fg est intégrable sur $[0,1]$.

d. Montrer que $\mathrm{L}_2(0,1) \subset \mathrm{X}$, et que l'image de $\mathrm{L}_2(0,1)$ est dense dans X. En déduire que toute forme linéaire ℓ continue sur X peut être représentée par une fonction g sous la forme

$$\forall f \in \mathrm{X}, \quad \ell(f) = \int_0^1 f(t)g(t)\,dt.$$

Fonctions affines tangentes, Hahn-Banach

Proposition. *On suppose que* X *est un espace affine réel de dimension finie,* f *une fonction convexe sur* X*, à valeurs dans* $\mathbb{R} \cup \{+\infty\}$*. Soient* Y *un sous-espace affine de* X *et* g *une fonction affine définie sur* Y*, telle que* $g(y) \leq f(y)$ *pour tout* $y \in \mathrm{Y}$ *; on suppose que* f *est finie dans un ouvert de* X *contenant un point* $y_0 \in \mathrm{Y}$ *; il existe alors une fonction affine* h *sur* X *qui prolonge* g *et telle que* $h \leq f$ *sur* X*.*

Démonstration. En procédant par récurrence sur la dimension, il suffit de montrer qu'on peut gagner une dimension : si $\mathrm{Y} \neq \mathrm{X}$, on considère un point $x \in \mathrm{X} \setminus \mathrm{Y}$, et le sous-espace affine Z engendré par Y et x. On cherche un prolongement affine h de g à Z, tel que $h \leq f$ sur Z. Considérons le vecteur non nul $v = x - y_0$; tout point de Z s'écrit $y + tv$ avec $y \in \mathrm{Y}$ et $t \in \mathbb{R}$, et si h est affine et prolonge g on a

$$h(y + tv) = h(y) + th(x) - th(y_0) = g(y) + t\,\big(h(x) - g(y_0)\big).$$

Bien entendu, prolonger g à Z demande seulement de définir $\gamma = h(x) - g(y_0)$. Pour que le prolongement soit convenable, il faut (et il suffit) que $g(y) + t\,\gamma = h(y+tv) \leq f(y+tv)$ pour tout nombre réel t et tout $y \in \mathrm{Y}$. C'est automatique si $t = 0$ (puisque $g \leq f$ sur Y), et nous allons découper la propriété voulue en deux, selon le signe de $t \neq 0$:

$$g(y) + \lambda\gamma \leq f(y + \lambda v), \quad g(y') - \lambda'\gamma \leq f(y' - \lambda' v)$$

pour tous $y, y' \in \mathrm{Y}$ et $\lambda, \lambda' > 0$. Le nombre γ doit donc vérifier les deux inégalités

$$\sup\Big\{\frac{g(y') - f(y' - \lambda' v)}{\lambda'} \;:\; y' \in \mathrm{Y}, \lambda' > 0\Big\} =: \mathrm{S} \leq \gamma,$$

$$\gamma \leq \mathrm{I} := \inf\Big\{\frac{f(y + \lambda v) - g(y)}{\lambda} \;:\; y \in \mathrm{Y}, \lambda > 0\Big\}.$$

Notons que I n'est pas $+\infty$, parce que l'inf porte sur un ensemble non vide de valeurs finies : en effet si $y = y_0$ et si $\lambda > 0$ est petit, on sait que $y_0 + \lambda v$ reste dans l'ouvert de X où f est finie, et de même S n'est pas $-\infty$. Pour que le choix de γ soit possible, il faut et il suffit que $\mathrm{S} \leq \mathrm{I}$, ce qui garantira que I et S sont finis, et il suffira de prendre pour γ n'importe quel nombre réel compris entre le sup et l'inf (bien sûr, si $\mathrm{S} = \mathrm{I}$ on n'a pas le choix : il faut prendre pour γ la valeur commune). Il reste donc à vérifier que

$$\frac{g(y') - f(y' - \lambda' v)}{\lambda'} \leq \frac{f(y + \lambda v) - g(y)}{\lambda}$$

pour tous $y, y' \in \mathrm{Y}$, $\lambda, \lambda' > 0$. On réécrit la propriété voulue sous la forme

$$\lambda'\, g(y) + \lambda\, g(y') \leq \lambda'\, f(y + \lambda v) + \lambda\, f(y' - \lambda' v)$$

et il est alors clair que cette propriété est vraie :

$$\frac{1}{\lambda + \lambda'}\,\big(\lambda'\, g(y) + \lambda\, g(y')\big) = g\Big(\frac{\lambda'}{\lambda + \lambda'}\, y + \frac{\lambda}{\lambda + \lambda'}\, y'\Big) \leq$$

$$f\Big(\frac{\lambda'}{\lambda + \lambda'}\, y + \frac{\lambda}{\lambda + \lambda'}\, y'\Big) = f\Big(\frac{\lambda'}{\lambda + \lambda'}\,(y + \lambda v) + \frac{\lambda}{\lambda + \lambda'}\,(y' - \lambda' v)\Big) \leq$$

$$\leq \frac{\lambda'}{\lambda + \lambda'}\, f(y + \lambda v) + \frac{\lambda}{\lambda + \lambda'}\, f(y' - \lambda' v).$$

La proposition est donc établie.

Remarques.

— Un premier cas important est celui où $\mathrm{Y} = \{y_0\}$ et $g(y_0) = f(y_0)$; dans ce cas le prolongement h est une fonction affine d'appui de la fonction convexe f ; le graphe de h est un hyperplan affine de $\mathrm{X} \times \mathbb{R}$ qui touche le graphe de f au point $(y_0, f(x_0))$, et cet hyperplan est partout "au dessous" du graphe de f.

— On peut passer à la dimension infinie avec le lemme de Zorn ; on peut affaiblir l'hypothèse topologique en une hypothèse quasi-algébrique : pour toute droite D de X passant par y_0, l'ensemble des points de la droite où f est finie est un voisinage (relatif) de y_0. On dit alors que y_0 est un *point interne* de l'ensemble convexe $\mathrm{C} = \{f < +\infty\}$.

— On peut donner une démonstration euclidienne, dans le cas où $\mathrm{Y} = \{y_0\}$ et où f est finie sur $\mathrm{X} = \mathbb{R}^d$ euclidien (donc continue sur X). On considère le convexe fermé C de $\mathbb{R}^d \times \mathbb{R}$ formé des points (x, t) tels que $f(x) \leq t$, on considère le point $\mathrm{Z}_\varepsilon = (y_0, f(y_0) - \varepsilon)$ pour $\varepsilon > 0$. En utilisant la projection de plus courte distance de Z_ε sur C, on construira une fonction affine h_ε telle que $h_\varepsilon \leq f$ sur X et $h_\varepsilon(y_0) = f(y_0) - \varepsilon$. Ensuite, on procédera par compacité pour extraire une limite quand $\varepsilon \to 0$.

Rappel : Minkowski

Si $f(x, y)$ est mesurable sur un espace produit $\mathrm{X} \times \mathrm{Y}$, avec deux mesures σ-finies μ et ν, et si $1 \leq p < +\infty$, on a

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

Convolutions $\mathrm{L}_1 * \mathrm{L}_p$

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, et $\mathrm{L}_p(\mathbb{T}) \subset \mathrm{L}_1(\mathbb{T})$ (en désignant par $\mathbb{T}$ le quotient $\mathbb{R}/\mathbb{Z}$, dont un modèle est le cercle unité du plan complexe).

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 $h(x, y) = f(y)g(x - y)$,

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

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

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

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

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

$$|(f * g)(x)| = \Big|\int f(x-y)g(y)\, dy\Big| \leq \|f\|_p\, \|g\|_q.$$

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* $p, q, r \geq 1$, *alors* $\mathrm{L}_p * \mathrm{L}_q \subset \mathrm{L}_r$.

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

Exercice 5.2.

a. On suppose que $\alpha, \beta, \gamma > 0$ vérifient $\alpha + \beta + \gamma = 1$. Montrer que pour toutes fonctions positives intégrables u, v, w sur un espace mesuré $(\Omega, \mathcal{A}, \mu)$ on a

$$\int u^\alpha v^\beta w^\gamma \, d\mu \leq \Big(\int u \, d\mu\Big)^\alpha \Big(\int v \, d\mu\Big)^\beta \Big(\int w \, d\mu\Big)^\gamma.$$

b. On suppose que $\alpha, \beta, \gamma > 0$ vérifient $\alpha + \beta + \gamma = 2$. Montrer que pour toutes fonctions positives intégrables $\mathrm{U}, \mathrm{V}, \mathrm{W}$ sur $\mathbb{R}^d$ on a

$$\int \mathrm{U}(x-y)^\alpha \mathrm{V}(y)^\beta \mathrm{W}(x)^\gamma \, dx\, dy \leq \Big(\int \mathrm{U}(x) \, dx\Big)^\alpha \Big(\int \mathrm{V}(x) \, dx\Big)^\beta \Big(\int \mathrm{W}(x) \, dx\Big)^\gamma.$$

En déduire que si $p, q, r \geq 1$ et $1/p + 1/q = 1 + 1/r$, on a $\mathrm{L}_p * \mathrm{L}_q \subset \mathrm{L}_r$, et plus précisément $\|f * g\|_r \leq \|f\|_p \, \|g\|_q$.

Indication pour un cas particulier, $\alpha = \beta = \gamma = 2/3$. On écrit le produit

$$\mathrm{U}(x-y)^{2/3} \mathrm{V}(y)^{2/3} \mathrm{W}(x)^{2/3}$$

comme produit des trois termes $(\mathrm{U}(x-y)\mathrm{V}(y))^{1/3}$, $(\mathrm{V}(y)\mathrm{W}(x))^{1/3}$ et $(\mathrm{U}(x-y)\mathrm{W}(x))^{1/3}$. On pose ensuite $u(x,y) = \mathrm{V}(y)\mathrm{W}(x)$, $v(x,y) = \mathrm{U}(x-y)\mathrm{W}(x)$ et pour finir $w(x,y) = \mathrm{U}(x-y)\mathrm{V}(y)$; on applique l'inégalité de Hölder pour trois fonctions à u, v, w avec les exposants $1/3, 1/3, 1/3$ de somme 1, sur l'espace mesuré $\mathbb{R}^d \times \mathbb{R}^d$ muni de la mesure produit $dx\,dy$.

La convolution : régularisation. Continuité

Commençons par les propriétés de continuité. 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.*

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 $|f(y)g(x_n - y)| \leq \mathrm{M}\,|f(y)|$, 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 $(\tau_v f)(x) = f(x - v)$ pour tout $x \in \mathbb{R}^d$. On notera aussi $f_v = \tau_v f$ 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 $\omega_f(\delta) = \sup\{|f(x) - f(x')| : |x - x'| \leq \delta\}$ est égal à $\sup\{\|f_v - f\|_\infty : |v| \leq \delta\}$.

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.*

On a $|g_t - g| \leq \omega_g(|t|)$ et $(f * g)_t - (f * g) = f * (g_t - g)$, il suffit d'appliquer la plus simple des majorations : $\|f * (g_t - g)\|_\infty \leq \|f\|_1 \, \|g_t - g\|_\infty$.

Solution de l'exercice 4.1

Soit $y_0 \in \mathbb{R}^d$ fixé, quelconque ; on pose $\mathrm{H}_{y_0}(x) = x \,.\, y_0 - \mathrm{F}(x)$; il s'agit d'une fonction continue et strictement concave sur $\mathbb{R}^d$, qui tend vers $-\infty$ lorsque $\|x\| \to +\infty$ (parce que $\mathrm{F}(x)/\|x\|$ tend vers $+\infty$ à l'infini ; mettre en facteur $\|x\|$ dans l'expression de H_{y_0} pour le voir). Il en résulte que H_{y_0} atteint son maximum sur $\mathbb{R}^d$, et en un point unique x_0 à cause de la concavité *stricte*. Il en résulte encore que le gradient de H s'annule en ce point x_0, ce qui donne $y_0 - \nabla\mathrm{F}(x_0) = 0$. Par ailleurs, une fonction concave de classe C^1 atteint son maximum en un point si et seulement si son gradient s'annule en ce point, ce qui nous dit que le gradient de F en tout point $x \neq x_0$ est différent de y_0.

On a ainsi démontré en un coup les affirmations suivantes : l'application $\nabla\mathrm{F}$ de $\mathbb{R}^d$ dans $\mathbb{R}^d$ définie par $x \to \nabla\mathrm{F}(x)$ est surjective et injective.

Montrons que $\nabla\mathrm{F}$ est propre, c'est à dire que l'image inverse de tout compact K est compacte ; cette image inverse étant fermée par continuité, il reste à vérifier qu'elle est bornée, et puisque K est contenu dans une boule d'un certain rayon R, il suffit de montrer qu'il existe A tel que $\|x\| > \mathrm{A}$ implique $\|\nabla\mathrm{F}(x)\| > \mathrm{R}$. Autrement dit on doit montrer que $x \to \|\nabla\mathrm{F}(x)\|$ tend vers l'infini à l'infini.

Posons $\varphi(t) = \mathrm{F}(tx)$; pour cette fonction convexe φ d'une variable on a

$$\mathrm{F}(x) - \mathrm{F}(0) = \varphi(1) - \varphi(0) \leq \varphi'(1) = \nabla\mathrm{F}(x) \,.\, x \leq \|\nabla\mathrm{F}(x)\| \, \|x\|$$

donc $\|\nabla\mathrm{F}(x)\| \geq (\mathrm{F}(x) - \mathrm{F}(0))/\|x\|$, et l'hypothèse que $\mathrm{F}(x)/\|x\|$ tend vers l'infini à l'infini fournit la conclusion.

Il en résulte que $\nabla\mathrm{F}$ est un homéomorphisme : pour montrer la continuité de l'inverse en un point y_0, on prend un voisinage borné et fermé V de ce point ; on sait que l'image par $(\nabla\mathrm{F})^{-1}$ est un compact W, et $\nabla\mathrm{F}$ est continue bijective de W sur V, donc d'inverse continue à cause de la compacité, ce qui donne le résultat.

La dernière partie demande de prouver que la fonction G définie par

$$\mathrm{G}(y) = \sup_{x \in \mathbb{R}^d} \big[x \,.\, y - \mathrm{F}(x)\big]$$

est convexe (clair), de classe C^1 (plus délicat), avec $\mathrm{G}(y)/\|y\|$ tendant vers l'infini. Considérons le point précédent y_0. On a vu que $\mathrm{G}(y_0) = x_0 \,.\, y_0 - \mathrm{F}(x_0)$ et par définition de G on a partout $\mathrm{G}(y) \geq x_0 \,.\, y - \mathrm{F}(x_0) = \mathrm{G}(y_0) + x_0 \,.\, (y - y_0)$. Considérons un point y_1 voisin de y_0, correspondant à x_1 voisin de x_0. On aura à la fois $\mathrm{G}(y_1) \geq \mathrm{G}(y_0) + x_0 \,.\, (y_1 - y_0)$ et en inversant les rôles $\mathrm{G}(y_0) \geq \mathrm{G}(y_1) + x_1 \,.\, (y_0 - y_1)$. On a donc l'encadrement

$$0 \leq \mathrm{G}(y_1) - \mathrm{G}(y_0) - x_0 \,.\, (y_1 - y_0) \leq (x_1 - x_0) \,.\, (y_1 - y_0)$$

qui entraîne la différentiabilité de G au point y_0 par la continuité de $y \to x = (\nabla\mathrm{F})^{-1}(y)$. En effet, étant donné $\varepsilon > 0$, il existe $\delta > 0$ tel que la relation $\|y - y_0\| < \delta$ implique $\|x - x_0\| < \varepsilon$. L'encadrement précédent, écrit avec y au lieu de y_1 donne, pour tout y tel que $\|y - y_0\| < \delta$,

$$\Big|\mathrm{G}(y) - \mathrm{G}(y_0) - x_0 \,.\, (y - y_0)\Big| \leq \|x - x_0\| \, \|y - y_0\| < \varepsilon \, \|y - y_0\|$$

ce qui montre que $x_0 = (\nabla\mathrm{F})^{-1}(y_0)$ est le gradient de G au point y_0. La continuité de $(\nabla\mathrm{F})^{-1}$ donne ensuite le caractère C^1 de G.

On a en prenant $x = \mathrm{A}\|y\|^{-1}y$ l'inégalité $\mathrm{G}(y) \geq \mathrm{A}\,\|y\| - \max\{\mathrm{F}(v) : \|v\| \leq \mathrm{A}\}$ qui montre que $\liminf_{\|y\| \to +\infty} \mathrm{G}(y)/\|y\| \geq \mathrm{A}$ pour tout A.

Prépa. Agrég écrit d'Analyse, séance n° 6, mercredi 14 novembre 2001.

Le cadre périodique

Un modèle pour le groupe topologique quotient $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ est le groupe multiplicatif U des nombres complexes de module 1. Munissons 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 plus habituelle

$$(f *_{\mathrm{per}} g)(\theta) = \int_0^{2\pi} f(\theta - s)g(s)\, \frac{ds}{2\pi}\cdot$$

On peut aussi voir f et g comme deux fonctions 2π-périodiques définies sur $\mathbb{R}$, étant entendu que toutes les opérations d'intégration sur ces fonctions seront effectuées sur un intervalle d'une période.

Quand on parlera de $\mathbb{T}$, on choisira selon les besoins l'une des trois présentations précédentes ; on peut généraliser à $\mathbb{T}^d$. 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 $\mathbb{R}^d/\mathbb{Z}^d$, $d(x, y)$ est la plus courte distance (euclidienne) entre deux représentants $x_1, y_1 \in \mathbb{R}^d$ de $x, y \in \mathbb{R}^d/\mathbb{T}^d$.

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_t sur $\mathrm{L}_p(\mathbb{R}^d)$, définies par

$$\forall f \in \mathrm{L}_p,\ \forall x \in \mathbb{R}^d, \quad (\tau_t f)(x) = f(x - t);$$

on les définit aussi sur $\mathbb{T}^d$. On écrira aussi $f_t = \tau_t f$; on notera $|t|$ la norme euclidienne du vecteur $t \in \mathbb{R}^d$.

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

$$\lim_{t \to 0} \|\tau_t f - f\|_p = 0.$$

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

Démonstration. Les translations τ_t sont des isométries de L_p, donc $\|\tau_t\| = 1$ pour tout $t \in \mathbb{R}^d$; 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 $\|g - \tau_t(g)\|_p \to 0$ lorsque $t \to 0$ (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 $|t| < t_0$ on aura $\|\tau_t g - g\|_p < \varepsilon/3$ d'après ce qui précède, et d'autre part $\|\tau_t f - \tau_t g\|_p = \|f - g\|_p < \varepsilon/3$, d'où le résultat par l'inégalité triangulaire.

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* T *sur un sous-ensemble dense* X_0 *de* X, *alors on a convergence de la suite* $(\mathrm{T}_n(x))$ *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* $1 \leq p \leq +\infty$, $1/p+1/q = 1$; *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

$$\|(f * g)_t - (f * g)\|_\infty = \|(f_t - f) * g\|_\infty \leq \|f_t - f\|_p \, \|g\|_q$$

par Hölder, et $\|f_t - f\|_p \to 0$ d'après ce qui précède. Pour tout $\varepsilon > 0$, il existe donc $t_0 > 0$ tel que la condition $|t| < t_0$ entraîne $|(f * g)(x - t) - (f * g)(x)| < \varepsilon$ pour tout $x \in \mathbb{R}^d$: c'est exactement la propriété de continuité uniforme pour la fonction $f * g$: si la distance entre $y = x - t$ et x est $< t_0$, alors $|(f * g)(y) - (f * g)(x)| < \varepsilon$.

Exercice 6.1. 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 $\mathbb{R}^d$.

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 $(f * g)' = f * g'$. On a des résultats analogues pour les dérivées partielles ; il existe un grand nombre de variantes de ces résultats. On va considérer le cas le plus simple. Soit u une direction dans $\mathbb{R}^d$; notons D_u la dérivée dans la direction u ; supposons que $g \in \mathrm{L}_\infty$ et que $\mathrm{D}_u g$ soit une fonction bornée par M ; en appliquant le théorème des accroissements finis aux fonctions d'une variable $t \to g(x + tu)$ on voit que ces fonctions sont M-lipschitziennes ; alors

$$t^{-1}\big((f * g)(x + tu) - (f * g)(x)\big) = \int f(y)\, t^{-1}\big(g(x + tu - y) - g(x - y)\big)\, dy,$$

En posant (pour x fixé) $\mathrm{G}(y, t) = t^{-1}\big(g(x + tu - y) - g(x - y)\big)$ on aura $|f(y)\mathrm{G}(y, t)| \leq \mathrm{M}\,|f(y)|$ qui donne la majoration par une fonction intégrable fixe indépendante du paramètre t, pendant que $\mathrm{G}(y, t)$ tend simplement vers $\mathrm{D}_u g(x - y)$ pour tout y ; le théorème de convergence dominée montre que $\mathrm{D}_u(f * g)(x)$ existe, et

$$\mathrm{D}_u(f * g) = f * \mathrm{D}_u g.$$

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 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 6.2. Montrer que la fonction f définie sur $\mathbb{R}$ par $f(x) = 0$ si $x \leq 0$ et $f(x) = \exp(-1/x)$ si $x > 0$ est de classe C^∞ sur $\mathbb{R}$. Utiliser f pour construire des fonctions C^∞ à support compact (non identiquement nulles) sur $\mathbb{R}$, puis sur $\mathbb{R}^d$.

Il existe des fonctions C^∞ à support compact, et non identiquement nulles sur $\mathbb{R}^d$. On note $\mathcal{D}(\mathbb{R}^d)$ l'espace de ces fonctions. Si φ est une telle fonction et si $f \in \mathrm{L}_1$, la convolée $f * \varphi$ est C^∞ et bornée.

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

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

$$(b) \qquad \lim_k \int \varphi_k = 1,$$

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

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

Démonstration. Soit g une fonction uniformément continue et bornée sur $\mathbb{R}^d$; posons $a_k = \int \varphi_k$ pour tout k ; on a

$$d_k(x) = (g * \varphi_k)(x) - a_k\, g(x) = \int \big(g(x-y) - g(x)\big)\, \varphi_k(y)\, dy$$

et

$$\sup_x |a_k\, g(x) - g(x)| = |1 - a_k|\, \|g\|_\infty$$

tend vers 0 quand $k \to +\infty$; il suffit donc de montrer que $d_k(x)$ tend vers 0, uniformément en $x \in \mathbb{R}^d$. La fonction g est bornée par un certain nombre $\mathrm{K} = \|g\|_\infty$. Puisque g est uniformément continue, on peut trouver $\alpha > 0$ tel que $|g(x-y) - g(x)| < \varepsilon$ 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 en posant $\mathrm{M} = \sup_k \|\varphi_k\|_1$

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

Pour k assez grand on obtiendra avec (c) que $\|d_k\|_\infty \leq \mathrm{M}\varepsilon + \varepsilon$; on obtient ainsi la convergence uniforme de $\varphi_k * g$ vers g.

Démontrons la deuxième partie. Les opérateurs définis sur $\mathrm{L}_p(\mathbb{R}^d)$ par $\mathrm{T}_k g = g * \varphi_k$ vérifient $\|\mathrm{T}_k\| \leq \|\varphi_k\|_1 \leq \mathrm{M}$ pour tout k, donc il suffit de démontrer la convergence vers l'identité pour un sous-espace vectoriel dense. Pour $g \in \mathcal{K}(\mathbb{R}^d)$, on sait que $g * \varphi_k$

converge uniformément vers g d'après ce qui précède. Fixons $\alpha > 0$, $\mathrm{B} = \mathrm{B}(0, \alpha)$; posons $\varphi_{k,1} = \mathbf{1}_{\mathrm{B}}\, \varphi_k$ et $\varphi_{k,2} = \varphi_k - \varphi_{k,1}$. D'après (c), on a $\|\varphi_{k,2}\|_1 \to 0$, ce qui implique que $\varphi_{k,2} * g$ tend vers 0 dans L_p ; on voit aussi que la suite $(\varphi_{k,1})$ vérifie encore les propriétés *a, b, c*. Soit $\mathrm{R} > 0$ tel que le support de g soit contenu dans $\mathrm{B}(0, \mathrm{R})$; on vérifie que $g * \varphi_{k,1}$ est nulle hors de $\mathrm{B}' = \mathrm{B}(0, \mathrm{R} + \alpha)$; on sait que $g * \varphi_{k,1} - g$ tend uniformément vers 0 ; on a donc

$$\begin{aligned} \|g * \varphi_{k,1} - g\|_p^p &= \int_{\mathrm{B}'} \big|(g * \varphi_{k,1})(x) - g(x)\big|^p \, dx \\ &\leq \mathrm{vol}(\mathrm{B}')\, \|g * \varphi_{k,1} - g\|_\infty^p \to 0. \end{aligned}$$

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

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

Par le changement de variable $y = kx$ on obtient les propriétés *a, b, c* voulues.

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. A partir d'une $\varphi \in \mathcal{D}$, positive et d'intégrale 1, on peut construire par la méthode précédente une approximation de l'unité (φ_k) formée de fonctions de $\mathcal{D}$.

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 $\mathcal{D}$ qui ont un support à peine plus grand que celui de g.

Pour toute $g \in \mathrm{L}_p$, 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 $\mathrm{L}_p(\mathbb{R}^d)$, pour $1 \leq p < +\infty$.

Une approximation de l'unité utile

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. On va utiliser une idée simple : les puissances accentuent le caractère maximal de la valeur en 0, et l'intégrale de $f(t)^n$ va se concentrer autour de 0.

On pose

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

et $\varphi_n(t) = \mathrm{I}_n^{-1}(1 + \cos(t))^n$, qui est bien ≥ 0 d'intégrale un pour la probabilité $dt/(2\pi)$. Si on se donne $0 < \alpha < \pi$, on aura en limitant l'intégrale à l'intervalle $[-\alpha/2, \alpha/2]$

$$2\pi\, \mathrm{I}_n \geq \alpha\, (1 + \cos(\alpha/2))^n$$

et

$$\int_{d(y,0)>\alpha} \varphi_n(t)\, \frac{dt}{2\pi} \leq \alpha^{-1} \Big(\frac{1 + \cos(\alpha)}{1 + \cos(\alpha/2)}\Big)^n$$

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

Séries de Fourier

On va travailler avec des séries de Fourier complexes sur l'intervalle $[0, 2\pi]$, muni de la mesure normalisée $dx/(2\pi)$. Pour tout $n \in \mathbb{Z}$, posons $e_n(x) = \mathrm{e}^{inx}$; les fonctions e_n, $n \in \mathbb{Z}$, sont deux à deux orthogonales et de norme 1. Si $f \in \mathrm{L}_2$, les coefficients de Fourier sont obtenus par $c_n(f) = \langle f, e_n \rangle$; la formule garde un sens si $f \in \mathrm{L}_1$,

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

Il est clair que $|c_n(f)| \leq \|f\|_1$ pour tout n. 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 $\mathrm{M} \leq \mathrm{N}$ et $\mathrm{M}, \mathrm{N} \in \mathbb{Z}$), on a

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

L'existence de la suite (φ_n) précédente montre que les combinaisons linéaires du système trigonométrique sont denses dans $\mathrm{C}_{\mathrm{per}}(0, 2\pi)$, donc aussi dans L_2 ; on a une base hilbertienne. On vérifie en effet que

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

est un polynôme trigonométrique pour tout $n \geq 0$. 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* $1 \leq p < +\infty$*, les polynômes trigonométriques sont denses dans* $\mathrm{L}_p(0, 2\pi)$.

Le lemme de Riemann-Lebesgue

Proposition. *Pour toute fonction* $f \in \mathrm{L}_1(0, 2\pi)$ *la suite* $(c_n(f))$ *tend vers* 0 *quand* $|n| \to +\infty$.

On a vu que les fonctions de classe C^1 sont denses dans L_1. On a quand f est C^1

$$c_n(f') e_n = f' * e_n = (f * e_n)' = f * e_n' = f * (in\, e_n) = (in) c_n(f) e_n$$

pour tout $n \in \mathbb{Z}$, donc $c_n(f) = c_n(f')/(in)$ pour $n \neq 0$, et $|c_n(f)| \leq \|f'\|_1/|n|$; ceci montre que les coefficients de Fourier de f sont $\mathrm{O}(|n|^{-1})$ dans le cas où f est C^1, donc ils tendent vers 0. La suite $f \to c_n(f)$ est une suite équicontinue de formes linéaires sur $\mathrm{L}_1(0, 2\pi)$, qui tend vers 0 sur le sous-espace dense formé des fonctions de classe C^1. Il y a donc convergence vers 0 pour toute fonction $f \in \mathrm{L}_1$.

On peut aussi démontrer le lemme en utilisant la densité dans L_1 des fonctions en escalier.

Dans une certaine mesure, on peut reconnaître le caractère C^k d'une fonction f 2π-périodique à la décroissance de ses coefficients de Fourier. En effet, si on sait que $c_n(f) = \mathrm{O}(|n|^{-k-2})$, on pourra dériver k fois la série de Fourier de f, et toutes ces séries dérivées seront normalement convergentes (la k-ième dérivée fait encore intervenir des coefficients en $|n|^{-2}$ qui donnent une série numérique majorante absolument convergente). Inversement, si f est de classe C^k, l'argument d'intégration par parties précédent

montre que $c_n(f) = \mathrm{O}(|n|^{-k})$; il y a bien sûr une différence importante entre les deux résultats, mais qui disparaît pour les fonctions f périodiques de classe C^∞ : elles sont caractérisées par le fait que pour tout entier $k \geq 0$, les coefficients de Fourier de f sont $\mathrm{O}(|n|^{-k})$.

Exercice 6.3. Formule de Poisson.

a. Soit f une fonction continue 2π-périodique sur $\mathbb{R}$; on pose

$$c_n(f) = \frac{1}{2\pi} \int_0^{2\pi} f(x)\, \mathrm{e}^{-inx}\, dx$$

pour tout $n \in \mathbb{Z}$. Montrer que si $\sum_{n\in\mathbb{Z}} |c_n(f)| < +\infty$, on a $f(x) = \sum_{n\in\mathbb{Z}} c_n(f)\, \mathrm{e}^{inx}$ pour tout $x \in \mathbb{R}$.

b. Soient F une fonction continue sur $\mathbb{R}$, telle que $|\mathrm{F}(x)| \leq \mathrm{C}\,(1+|x|)^{-a}$ pour un $a > 1$ et tout $x \in \mathbb{R}$, et $\widehat{\mathrm{F}}$ sa transformée de Fourier, définie par

$$\forall t \in \mathbb{R}, \quad \widehat{\mathrm{F}}(t) = \int_{\mathbb{R}} \mathrm{F}(x)\, \mathrm{e}^{-ixt}\, dx.$$

Montrer que la fonction $f(x) = \sum_{n\in\mathbb{Z}} \mathrm{F}(x + 2\pi n)$ est définie, continue et 2π-périodique. Trouver une relation entre les valeurs $\widehat{\mathrm{F}}(n)$, $n \in \mathbb{Z}$ et les coefficients de Fourier de f.

c. On suppose de plus que $\sum |\widehat{\mathrm{F}}(n)| < +\infty$. Démontrer la *formule de Poisson*,

$$\sum_{n\in\mathbb{Z}} \mathrm{F}(2\pi n) = \frac{1}{2\pi} \sum_{n\in\mathbb{Z}} \widehat{\mathrm{F}}(n).$$

Appliquer avec $\mathrm{F}(x) = \mathrm{e}^{-a|x|}$, $a > 0$.