Prépa. Agrég écrit d'Analyse, séance n° 1, Mercredi 4 Octobre 2000.

Mesure et Intégration

Fonctions et ensembles

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

Les deux notions $\limsup_n x_n$ et $\liminf_n x_n$ pour une suite (x_n) de nombres réels doivent être sues sans hésitation ; la définition est plus agréable dans $\overline{\mathbb{R}}$. Pour $\limsup_n$, on prend la limite de la suite décroissante $y_n = \sup_{m \ge n} x_m$,

$$\limsup_n x_n = \lim_n \big(\sup_{m \ge n} x_m \big).$$

Une suite numérique (x_n) converge dans $\overline{\mathbb{R}}$ si et seulement si $\limsup_n x_n = \liminf_n x_n$.

Exercice-Propriétés. Si $x_n \le y_n$ pour tout $n \ge 0$, alors $\limsup_n x_n \le \limsup y_n$. Le nombre $\limsup x_n$ est la plus grande valeur d'adhérence de la suite. On a $\limsup(x_n + y_n) \le \limsup_n x_n + \limsup_n y_n$.

Exercices tirés de Zuily-Queffélec.

1. Si $u_n \ge 0$ et $u_{n+p} \le u_n + u_p$ pour tous entiers $p, n \ge 1$, alors u_n/n converge.
2. Montrer que $\limsup_n \varphi(n)/n = 1$, $\liminf_n \varphi(n)/n = 0$, où φ désigne l'indicatrice d'Euler.

Etant donnée une suite (f_n) de fonctions numériques (à valeurs dans $\overline{\mathbb{R}}$) définies sur un ensemble quelconque X, on définit la fonction $f = \limsup_n f_n$ sur X en posant $f(x) = \limsup_n f_n(x)$ pour tout $x \in X$. On note que pour tout t réel

$$(\limsup_n f_n)(x) > t \Rightarrow x \in \bigcap_n \Big(\bigcup_{m \ge n} \{f_m > t\} \Big) \Rightarrow (\limsup_n f_n)(x) \ge t$$

de sorte qu'en utilisant l'ensemble $\limsup_n A_n = \bigcap_n \bigcup_{m \ge n} A_m$ on obtient la chaîne d'inclusions

$$\{\limsup_n f_n > t\} \subset \limsup_n \{f_n > t\} \subset \{\limsup_n f_n \ge t\}.$$

On a l'analogue avec $\liminf$.

Algèbres, σ-algèbres (ou tribus)

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

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

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 passage au complémentaire.

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

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 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.

$\mathcal{P}(\mathrm{X})$ est évidemment une tribu, de même que $\{\emptyset, \mathrm{X}\}$.

Opérations sur les tribus ; tribu engendrée

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 $\sigma(\mathcal{C}) = \mathcal{C}$.

Conception de la génération par l'intérieur : on répète les opérations suivantes sur une classe initiale : ajouter tous les complémentaires des éléments de la classe ; ajouter toutes les réunions dénombrables des suites d'éléments de la classe ; on continue de façon transfinie… Quand on arrive au premier ordinal non dénombrable, on est sûr que l'on a fabriqué la tribu !

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 : $\mathcal{C}_1 \subset \mathcal{C}_2$, alors $\sigma(\mathcal{C}_1) \subset \sigma(\mathcal{C}_2)$.

Produit de tribus

On suppose donnés X_1 et X_2 avec chacun une tribu $\mathcal{A}_1$ ou $\mathcal{A}_2$; 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 $\mathrm{A}_j \in \mathcal{A}_j$:

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

Exercice. Montrer que les sections d'un ensemble de $\mathcal{A}_1 \otimes \mathcal{A}_2$ sont des ensembles de A_1 ou A_2 (remarque utile pour le théorème de Fubini).

Proposition. *On considère une application φ de $\mathcal{P}(\mathrm{X})$ dans $\mathcal{P}(\mathrm{Y})$ telle que*

(i) $\varphi(\bigcup_n \mathrm{A}_n) = \bigcup_n \varphi(\mathrm{A}_n)$ *pour toute suite (A_n) de sous-ensembles de X ;*
(ii) $\varphi(\mathrm{B} \setminus \mathrm{A}) = \varphi(\mathrm{B}) \setminus \varphi(\mathrm{A})$ *lorsque $\mathrm{A} \subset \mathrm{B} \subset \mathrm{X}$.*

1. *Si $\mathcal{B}$ est une tribu de parties de Y et si $\varphi(\mathrm{X}) \in \mathcal{B}$, l'ensemble*

$$\mathcal{A} = \{\mathrm{A} \in \mathcal{P}(\mathrm{X}) : \varphi(\mathrm{A}) \in \mathcal{B}\}$$

est une tribu de parties de X.

2. *Si $\mathcal{A}$ est une tribu de parties de X et si $\varphi(\mathrm{X}) = \mathrm{Y}$, l'ensemble $\mathcal{B} = \{\varphi(\mathrm{A}) : \mathrm{A} \in \mathcal{A}\}$ est une tribu de parties de Y. De plus, si $\mathcal{C}$ engendre $\mathcal{A}$, alors la classe $\mathcal{D} = \{\varphi(\mathrm{C}) : \mathrm{C} \in \mathcal{C}\}$ engendre $\mathcal{B}$.*

On note que (ii) implique $\varphi(\emptyset) = \emptyset$. La démonstration de la proposition est un exercice facile. Un commentaire sur le dernier point : considérons la tribu $\sigma(\mathcal{D})$ de parties de Y engendrée par la classe $\mathcal{D}$. D'après le point (i), l'ensemble $\{\mathrm{U} \in \mathcal{P}(\mathrm{X}) : \varphi(\mathrm{U}) \in \sigma(\mathcal{D})\}$ est une tribu de parties de X, et elle contient $\mathcal{C}$, donc elle contient $\sigma(\mathcal{C}) = \mathcal{A}$. Cela signifie que $\mathcal{B} \subset \sigma(\mathcal{D})$. Inversement $\mathcal{D} \subset \mathcal{B}$ est évident, donc $\sigma(\mathcal{D}) \subset \mathcal{B}$.

Remarques. On comprend mieux les propriétés de φ si on passe aux espaces de fonctions réelles bornées définies sur X et sur Y : φ correspond alors à une application *linéaire* Φ entre espaces vectoriels de fonctions, qui enverrait la fonction $\mathbf{1}_\mathrm{A}$, pour $\mathrm{A} \in \mathcal{P}(\mathrm{X})$, sur $\mathbf{1}_{\varphi(\mathrm{A})}$, et les deux cas de la proposition précédente correspondent à une image directe ou une image inverse de sous-espace vectoriel par une application linéaire (bien sûr, il y a une opération en plus, qui correspond à une prise de limite simple d'une suite de fonctions).

Exemples d'application de la proposition.

1. L'image inverse d'une tribu de parties de Y par une application $f : \mathrm{X} \to \mathrm{Y}$ est une tribu de parties de X.

Quelques trivialités en $f^{-1}(\bigcup_n \mathrm{A}_n) = \ldots$, $f^{-1}(\mathrm{A}^c) = \ldots$ montrent en effet que l'application $\mathrm{B} \to f^{-1}(\mathrm{B})$ est une application φ du type voulu (de $\mathcal{P}(\mathrm{Y})$ dans $\mathcal{P}(\mathrm{X})$).

La tribu engendrée par $f^{-1}(\mathcal{C})$ est l'image inverse de la tribu engendrée par $\mathcal{C}$,

$$(*) \qquad \sigma(f^{-1}(\mathcal{C})) = f^{-1}(\sigma(\mathcal{C}))$$

2. Image “directe” : on utilise le cas **1** de la proposition avec l'application φ précédente. Si $\mathcal{A}$ est une tribu de parties de l'espace de départ X, la tribu image directe est définie par

$$\mathcal{B} = \{\mathrm{B} \in \mathcal{P}(\mathrm{Y}) : f^{-1}(\mathrm{B}) \in \mathcal{A}\}.$$

Pour montrer la partie non-triviale de $(*)$, on considérera l'image directe de la tribu $\sigma(f^{-1}(\mathcal{C}))$.

3. La trace d'une tribu sur un sous-ensemble $\mathrm{Y} \subset \mathrm{X}$ est une tribu de parties de Y. L'application φ est ici $\mathrm{A} \subset \mathrm{X} \to \mathrm{A} \cap \mathrm{Y}$.

4. Encore un exemple d'application φ : si on fixe $\mathrm{A}_2 \subset \mathrm{X}_2$, l'application $\mathrm{A} \to \mathrm{A} \times \mathrm{A}_2$ de $\mathcal{P}(\mathrm{X}_1)$ dans $\mathcal{P}(\mathrm{X}_1 \times \mathrm{X}_2)$ est une application φ qui a les propriétés précédentes.

Topologie et dénombrabilité

Si (X, d) est un espace métrique séparable, il existe une suite (U_n) d'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.

Exercice. Tout ouvert de $\mathbb{R}$ est réunion dénombrable d'intervalles ouverts (disjoints).

Tribus boréliennes

Définition. Soit X un espace topologique ; la *tribu borélienne* de X, qu'on peut raisonnablement noter $\mathcal{B}_\mathrm{X}$, 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 $\mathcal{B}_\mathrm{X} = \sigma(\mathcal{C})$.

Mettons ces petits principes en application. La tribu borélienne de $\mathbb{R}$ ou de $\overline{\mathbb{R}}$ est la tribu engendrée par les ouverts de $\mathbb{R}$; mais tout ouvert de $\mathbb{R}$ est réunion dénombrable d'intervalles ouverts, donc $\mathcal{B}_{\mathbb{R}}$ est la tribu engendrée par les intervalles ouverts. La tribu $\mathcal{B}_{\mathbb{R}}$ est aussi la tribu engendrée par les intervalles fermés bornés, parce que tout intervalle ouvert est réunion dénombrable d'intervalles fermés bornés. La tribu $\mathcal{B}_{\mathbb{R}}$ est aussi la tribu engendrée par les intervalles $(c, +\infty)$.

La tribu borélienne d'un sous-espace topologique X_0 *d'un espace* X_1 *est la trace de la tribu borélienne de* X_1*.*

Pour le voir, on applique le cas **2** de la proposition précédente à l'application φ de $\mathcal{P}(\mathrm{X}_1)$ dans $\mathcal{P}(\mathrm{X}_0)$ définie par $\varphi(\mathrm{A}_1) = \mathrm{X}_\cap \mathrm{X}_0$. La classe image de la classe des ouverts de X_1 engendre la tribu image de la tribu engendrée par $\mathcal{O}_{\mathrm{X}_1}$.

Si X_0 est un borélien de X_1, les boréliens de l'espace topologique X_0 sont exactement les boréliens de X_1 qui sont contenus dans X_0. Par exemple, les boréliens de $\mathbb{R}$ sont exactement les boréliens de $\overline{\mathbb{R}}$ qui sont contenus dans $\mathbb{R}$, et un borélien de $\overline{\mathbb{R}}$ est simplement obtenu en ajoutant à un borélien de $\mathbb{R}$ un, ou deux (ou aucun) des points $+\infty$ et $-\infty$.

Les ensembles de la forme $(c_1, +\infty] \times (c_2, +\infty]$ engendrent la tribu borélienne de $\overline{\mathbb{R}}^2$. Avec une petite manipulation on voit que l'algèbre engendrée par ces ensembles doit contenir les $(c_1, d_1] \times (c_2, d_2]$; il est assez facile de voir que tout ouvert de $\mathbb{R}^2$ est réunion dénombrable de tels pavés.

Proposition. Tribu produit et tribu borélienne produit : *si* X_1 *et* X_2 *sont deux espaces métriques séparables, la tribu borélienne produit* $\mathcal{B}_1 \otimes \mathcal{B}_2$ *est égale à la tribu borélienne* $\mathcal{B}_\times$ *de l'espace topologique* $\mathrm{X}_1 \times \mathrm{X}_2$*.*

Etape 1 : si $\mathcal{T}$ est une tribu de parties de $\mathrm{X}_1 \times \mathrm{X}_2$ et si $\mathrm{X}_1 \times \mathrm{B} \in \mathcal{T}$, l'ensemble des A_1 tels que $\mathrm{A}_1 \times \mathrm{B} \in \mathcal{T}$ est une tribu de parties de X_1. Il en résulte d'abord que pour tout ouvert ω_2 de X_2, on a $\mathrm{A}_1 \times \omega_2 \in \mathcal{B}_2$ pour tous $\mathrm{A}_1 \in \mathcal{B}_1$, donc après un deuxième coup on aura $\mathrm{A}_1 \times \mathrm{A}_2 \in \mathcal{B}_2$ pour tous $\mathrm{A}_j \in \mathcal{B}_j$, donc $\mathcal{B}_1 \otimes \mathcal{B}_2 \subset \mathcal{B}_\times$. Cette inclusion est toujours vraie.

Si les espaces sont métriques séparables, on peut trouver une famille $\mathcal{F}$ d'ouverts $\omega_i \times \omega'_i$ telle que tout ouvert du produit soit réunion dénombrable d'une sous-famille d'éléments de $\mathcal{F}$. Comme $\mathcal{F} \subset \mathcal{B}_1 \otimes \mathcal{B}_2$, il en résulte que $\mathcal{O}_\times \subset \mathcal{B}_1 \otimes \mathcal{B}_2$, donc $\mathcal{B}_\times \subset \mathcal{B}_1 \otimes \mathcal{B}_2$.

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 $f^{-1}(\mathcal{A}_2) \subset \mathcal{A}_1$.

Propriétés évidentes ; la composition d'applications mesurables fournit des applications mesurables.

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 $\mathcal{A}_2$.

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

$$(\mathrm{M}) \qquad \forall t \in \mathbb{R}, \quad \{f > t\} \in \mathcal{A}.$$

Si f_1, f_2 sont deux applications mesurables de $(\Omega, \mathcal{A})$ dans $(\Omega_j, \mathcal{A}_j)$, $j = 1, 2$, l'application $\omega \to (f_1(\omega), f_2(\omega))$ est mesurable de $(\Omega, \mathcal{A})$ dans $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$.

Continuité et mesurabilité ;

Toute application continue de $\mathbb{R}^m$ dans $\mathbb{R}^n$ est borélienne : en effet, les images inverses des ouverts sont ouvertes, donc boréliennes.

Si f_1, f_2 sont deux applications mesurables de $(\Omega, \mathcal{A})$ dans $(\mathbb{R}^{d_j}, \mathcal{B}_j)$, $j = 1, 2$ (tribus boréliennes) alors l'application $\omega \to (f_1(\omega), f_2(\omega))$ est mesurable de $(\Omega, \mathcal{A})$ dans $(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}, \mathcal{B})$ (tribu borélienne du produit ; tout marche bien puisque la tribu borélienne du produit est le produit des tribus boréliennes).

On en déduit la mesurabilité de tous les exemples imaginables.

On dira qu'une fonction mesurable f est *étagée* si f ne prend qu'un nombre fini de valeurs.

Une fonction réelle (étendue) minorée est mesurable si et seulement si elle est limite croissante d'une suite d'étagées.

Ecrivons le pour une $f \geq 0$. Pour chaque entier $n \geq 0$ on pose

$$f_n = \sum_{j=0}^{4^n} j 2^{-n} \mathbf{1}_{\{j2^{-n} \leq f < (j+1)2^{-n}\}}.$$

La suite (f_n) converge simplement vers f.

Inversement si f_n croît vers f on a pour tout t réel

$$\{f > t\} = \bigcup_n \{f_n > t\}$$

donc f est mesurable par le critère usuel.

Exercice. Si (f_n) est une suite de fonctions réelles mesurables sur $(\Omega, \mathcal{A})$, l'ensemble A des points ω où la limite $\lim_n f_n(\omega)$ existe est dans la tribu $\mathcal{A}$.

Une fonction à valeurs dans $\mathbb{R}^d$ est mesurable si et seulement si elle est limite d'une suite de fonctions étagées.

Toute limite simple d'une suite (f_n) de fonctions mesurables de $(\Omega, \mathcal{A})$ dans $(\mathrm{X}, \mathcal{B})$ (tribu borélienne) est mesurable.

Il suffit de tester $f^{-1}(\mathrm{U})$ pour un ouvert U quelconque de X. La convergence simple permet d'exprimer cet ensemble par des opérations ensemblistes dénombrables sur les $\{f_n \in \mathrm{U}\}$.

Exercice. Soit $g : \Omega \to \mathbb{R}$ une application quelconque, $\mathcal{A} = g^{-1}(\mathcal{B}_{\mathbb{R}})$ (qu'on appelera en proba la tribu engendrée par la v.a. g) ; montrer que f est mesurable de $(\Omega, \mathcal{A})$ dans $\mathbb{R}$ si et seulement si $f = \varphi(g)$, avec φ borélienne de $\mathbb{R}$ dans $\mathbb{R}$.

Mesures positives sur une tribu

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 $\mu(\bigcup_n \mathrm{A}_n) = \sum_{n=0}^{+\infty} \mu(\mathrm{A}_n)$ (tous les calculs sont fait avec la valeur $+\infty$ admise).

En prenant $A_0 = A$, $A_1 = B$ et $A_n = \emptyset$ 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 $A \neq \emptyset$, 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.

Une mesure σ-finie sur Ω est une mesure μ telle qu'il existe un recouvrement dénombrable de Ω par une suite d'ensembles de μ-mesure finie.

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

Intégrale

On considère un espace mesuré $(\Omega, \mathcal{A}, \mu)$. Toutes les fonctions mentionnées sont supposées mesurables de $(\Omega, \mathcal{A})$ dans $\overline{\mathbb{R}}$.

Si f est étagée positive, on peut l'exprimer sous la forme $f = \sum_{i=1}^{n} c_i \mathbf{1}_{A_i}$, où $c_i \geq 0$ et $A_i \in \mathcal{A}$. L'expression $\sum_i c_i \mu(A_i)$ peut être $+\infty$ mais ne dépend que de la fonction f (micro-exercice). On l'appelle l'intégrale de la fonction $f \geq 0$.

On démontre que $(f \leq g) \Rightarrow (\int f \leq \int g)$ et l'additivité pour les fonctions étagées positives. 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 de la partition raffinée.

Lemme. *Si une suite croissante (f_n) de fonctions étagées ≥ 0 dépasse $\mathbf{1}_A$, alors*

$$\lim_n \int f_n \geq \mu(A).$$

Prépa. Agrég. écrit d'Analyse, séance n° 2, Mercredi 11 Octobre 2000.

Mesure et Intégration, suite

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 d'une arithmétique étendue de la façon suivante : on pose $a + (+\infty) = +\infty$ pour tout $a \geq 0$ (ce qui va de soi), $a \times (+\infty) = +\infty$ si $a > 0$ et $0 \times (+\infty) = 0$. 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 $+\infty$.

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 que $f = \sum_{i=1}^n v_i \, \mathbf{1}_{\mathrm{V}_i}$; 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 de 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 $i = 1, \ldots, m$ et $j = 1, \ldots, n$, on a

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

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 $j = 1, \ldots, n$, 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 $\mathrm{A}_i = \bigcup_j \mathrm{A}_i \cap \mathrm{B}_j$) et de même dans l'autre sens,

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

On démontre ensuite que $(f \le g) \Rightarrow (\int f\, d\mu \le \int g\, d\mu)$ et l'additivité dans le cas ≥ 0 (ainsi que la propriété triviale $\int (\lambda f)\, d\mu = \lambda \int f\, d\mu$, quand $\lambda \ge 0$) : 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 $f = \sum_i a_i \mathbf{1}_{\mathrm{A}_i} \le \sum_j b_j \mathbf{1}_{\mathrm{B}_j} = g$, on aura $a_i\, \mu(\mathrm{C}_{i,j}) \le b_j\, \mu(\mathrm{C}_{i,j})$ 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 $[0, +\infty]$.

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

Avec cette définition, il est évident que $0 \le f_1 \le f_2$ implique $\int f_1\, d\mu \le \int f_2\, d\mu$; il est clair que $\int (f_1 + f_2)\, d\mu \ge \int f_1\, d\mu + \int f_2\, d\mu$. 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 $\mathcal{A}$: dans ce cas toute fonction étagée $g \le f$ 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 $a \ge 0$; si une suite croissante (f_n) de fonctions mesurables ≥ 0 vérifie $a \mathbf{1}_{\mathrm{A}} \le \lim_n f_n$, alors*

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

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 $b\, \mathbf{1}_{\mathrm{B}_n} \le f_n$ donc

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

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 $g = \sum_{j=1}^n a_j \mathbf{1}_{\mathrm{A}_j}$ une fonction étagée positive $\leq f$, avec des A_j non vides qui réalisent une partition de Ω, et $a_j \geq 0$ pour $j = 1, \ldots, n$. 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

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

et ceci pour toute $g \leq f$. 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.

Corollaire : intégrale et séries dans le cas ≥ 0. *Si les* (u_n) *sont mesurables positives,*

$$\int \bigl(\sum_{n=0}^{+\infty} u_n\bigr)\,d\mu = \sum_{n=0}^{+\infty} \int u_n\,d\mu.$$

Il s'agit bien entendu de calculs de séries avec valeur $+\infty$ admise.

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 $\int g_n \leq \inf_{m \geq n} \int f_m$. 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, 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

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

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 $\limsup_n \int |f - f_n| \, d\mu \leq 0$.

Prépa. Agrég. écrit d'Analyse, séance n° 3, Mercredi 18 Octobre 2000.

Mesure et Intégration, suite II

Classes de fonctions, espace L_1

Deux remarques avant de se lancer. On considère un espace mesuré $(\Omega, \mathcal{A}, \mu)$, avec $\mu \geq 0$.

1. *Si f est mesurable ≥ 0 sur $(\Omega, \mathcal{A})$ et si $\int f\,d\mu < +\infty$, alors f est finie μ-presque partout* : la fonction égale à $+\infty$ sur l'ensemble $\mathrm{A} = \{f = +\infty\} \in \mathcal{A}$ et à 0 ailleurs est $\leq f$, donc son intégrale $+\infty \times \mu(\mathrm{A})$ est finie, donc $\mu(\mathrm{A}) = 0$.

2. *Si $f \geq 0$ et $\int f\,d\mu = 0$, alors f est nulle μ-presque partout* : pour tout entier $n \geq 0$, la fonction $2^{-n}\mathbf{1}_{\{f > 2^{-n}\}}$ est $\leq f$, donc $0 \leq 2^{-n}\mu(\{f > 2^{-n}\}) \leq 0$. On a donc $\mu(\{f > 2^{-n}\}) = 0$, et par réunion dénombrable d'ensembles μ-négligeables on obtient $\mu(\{f > 0\}) = \mu(\{f \neq 0\}) = 0$.

Désignons par $\mathcal{N}$ le sous-ensemble de $\mathcal{L}_1(\Omega, \mathcal{A}, \mu)$ défini par

$$\mathcal{N} = \{f \in \mathcal{L}_1 : \int |f|\,d\mu = 0\}.$$

D'après ce qui précède, une fonction f de $\mathcal{L}_1$ est dans $\mathcal{N}$ si et seulement si elle est nulle μ-presque partout. Il est clair que $\mathcal{N}$ est un sous-espace vectoriel de $\mathcal{L}_1$, et on peut considérer l'espace vectoriel quotient $\mathrm{L}_1(\Omega, \mathcal{A}, \mu) = \mathcal{L}_1(\Omega, \mathcal{A}, \mu)/\mathcal{N}$.

Deux fonctions f_1 et f_2 sont dans la même classe si et seulement si $f_1(\omega) = f_2(\omega)$ pour μ-presque tout $\omega \in \Omega$. Il est alors clair que pour toute fonction φ de $\mathbb{R}$ dans $\mathbb{R}$, tous les $\varphi(f)$ restent dans la même classe quand f varie dans une classe $\widehat{f}$, et on appellera cette classe $\varphi(\widehat{f})$; en particulier on désignera par $|\widehat{f}|$ la classe qui contient toutes les fonctions $|f|$ lorsque f varie dans $\widehat{f}$.

Toutes les fonctions d'une même classe $\widehat{f} \in \mathrm{L}_1$ ont la même intégrale, qu'on notera $\int \widehat{f}\,d\mu$. On posera

$$\|\widehat{f}\|_1 = \int |\widehat{f}|\,d\mu$$

qui vaut $\int |f|\,d\mu$ pour tout représentant f de $\widehat{f}$.

L'application $\widehat{f} \to \|\widehat{f}\|_1$ est une norme sur L_1. Le fait que $f_1 + f_2$ soit un représentant de $\widehat{f_1} + \widehat{f_2}$ entraîne que $\|\widehat{f_1} + \widehat{f_2}\| \leq \|\widehat{f_1}\| + \|\widehat{f_2}\|$. La relation $\|\widehat{f}\|_1 = 0$ signifie que les éléments de $\widehat{f}$ sont négligeables donc implique que $\widehat{f}$ est la classe nulle,

$$\big(\|\widehat{f}\|_1 = 0\big) \Rightarrow \big(\widehat{f} = 0_{\mathrm{L}_1}\big).$$

On en déduit bien que la quantité $\|f\|_1$ définit une norme sur L_1.

On va démontrer en faisant attention aux classes (pour une fois et pour la dernière fois) que L_1 est un espace de Banach.

Théorème. *L'espace* $\mathrm{L}_1(\Omega, \mathcal{A}, \mu)$ *est un espace de Banach.*

Démonstration. Pour montrer qu'un espace normé E est complet, il suffit de montrer que toute série de vecteurs de E, dont la série des normes converge, est une série convergente dans l'espace E. Soit donc $(\widehat{f_n})$ une suite d'éléments de L_1 telle que $\sum_k \|\widehat{f_k}\|_1 < +\infty$; nous devons montrer que la série $\sum_k \widehat{f_k}$ converge dans l'espace L_1. Pour chaque entier

$n \geq 0$, soit f_n une fonction réelle mesurable qui est dans la classe $\widehat{f}_n$. Considérons la fonction $g \geq 0$ définie en chaque point $\omega \in \Omega$ par la somme de la série (valeur $+\infty$ admise)

$$g(\omega) = \sum_{k=0}^{+\infty} |f_k(\omega)|.$$

On définit ainsi une fonction mesurable g à valeurs dans $[0, +\infty]$. On a vu que dans le cas positif on a $\int g \, d\mu = \sum_{k=0}^{+\infty} \int |f_k| \, d\mu = \sum_{k=0}^{+\infty} \|\widehat{f}_k\|_1 < +\infty$. Puisque l'intégrale de la fonction g est finie, il en résulte que $\mathrm{B} = \{g = +\infty\}$ est un ensemble (de la tribu) de mesure nulle. Sur le complémentaire, la série converge absolument. Définissons $\mathrm{F}(\omega) = \sum_{k=0}^{+\infty} f_k(\omega)$ si $\omega \notin \mathrm{B}$ et (par exemple) $\mathrm{F}(\omega) = 0$ sinon. La fonction F est mesurable, $|\mathrm{F}| \leq g$ donc $\int |\mathrm{F}| \, d\mu < +\infty$ et F est intégrable. Pour finir on montre que les sommes partielles $\mathrm{S}_n = \sum_{k=0}^{n} \widehat{f}_k$ convergent dans L_1 vers la classe de F, quand $n \to +\infty$. La fonction égale à $\sum_{k>n} f_k(\omega)$ si $\omega \notin \mathrm{B}$ et à 0 sinon est dans la classe $\widehat{\mathrm{F}} - \sum_{k=0}^{n} \widehat{f}_k$, donc

$$\begin{aligned}\Big\|\widehat{\mathrm{F}} - \sum_{k=0}^{n} \widehat{f}_k\Big\| &= \int_{\Omega\setminus\mathrm{B}} \Big|\sum_{k>n} f_k(\omega)\Big| \, d\mu(\omega) \leq \int_{\Omega\setminus\mathrm{B}} \sum_{k>n} |f_k(\omega)| \, d\mu(\omega) \\ &= \sum_{k>n} \int |f_k| \, d\mu = \sum_{k>n} \|\widehat{f}_k\|_1,\end{aligned}$$

reste d'une série numérique convergente qui tend donc vers 0 quand $n \to +\infty$.

Il serait insupportable de continuer à distinguer aussi pointilleusement fonctions et classes de fonctions, et on va arrêter tout de suite. On a vu en passant le principe suivant, qui est une forme de Fubini facile.

Corollaire. *Soit (f_n) une suite d'éléments de L_1 telle que $\int \big(\sum_n |f_n|\big) \, d\mu < +\infty$; alors la série $\sum f_n$ converge μ-presque partout et la fonction ainsi définie presque partout vérifie*

$$\int \Big(\sum_{n=0}^{+\infty} f_n\Big) \, d\mu = \sum_{n=0}^{+\infty} \Big(\int f_n \, d\mu\Big).$$

Prépa. Agrég. écrit d'Analyse, séance n$^{\text{o}}$ 4, Mercredi 25 Octobre 2000.

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 pourra aussi noter que $(p-1)(q-1) = 1$ dans ce cas.

Si f est une fonction mesurable sur $(\Omega, \mathcal{A}, \mu)$ on pose

$$\|f\|_p = \Big(\int |f(\omega)|^p \, d\mu(\omega)\Big)^{1/p}.$$

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

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

Démonstration. Pour alléger un peu, on écrira simplement $\int f$ au lieu de $\int_\Omega f(\omega)\, d\mu(\omega)$. Si $p = \infty$, alors $q = 1$; la fonction f est (presque-sûrement) bornée par $\mathrm{M} = \|f\|_\infty$ et g est intégrable ; le produit fg est mesurable et $|fg| \le \mathrm{M}\, |g|$, donc fg est intégrable et

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

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

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

(pour le voir, on pourra maximiser la fonction $t \to tu - t^p/p$ sur $[0, +\infty[$). Il en résulte que pour tout $\omega \in \Omega$

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

ce qui montre que fg est intégrable, et que

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

Si $\|f\|_p = 0$, la fonction f est nulle μ-presque partout, donc $\int fg = 0$ et l'inégalité voulue est vraie, et de même si $\|g\|_q = 0$. Supposons donc $\|f\|_p > 0$ et $\|g\|_q > 0$; l'inégalité cherchée est positivement homogène par rapport à f et à g, donc il suffit de la démontrer lorsque $\|f\|_p = \|g\|_q = 1$. Mais dans ce cas, $\int |f|^p = 1$ et $\int |g|^q = 1$, donc l'inégalité précédente donne $|\int fg| \le 1/p + 1/q = 1$, ce qui est le résultat voulu.

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

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

Démonstration. L'inégalité de Hölder nous dit déjà que

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

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

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

Remarque. Si par hasard on n'avait pas déjà démontré que $f \to \|f\|_p$ est une norme sur L_p, on pourrait le déduire de l'inégalité de Hölder ; supposons $1 < p < +\infty$, soient f_1, f_2 telles que $\|f_1\|_p, \|f_2\|_p < +\infty$; pour $j = 1, 2$ soit $\mathrm{A}_{n,j} = \{2^{-n} < |f_j| < 2^n\}$, pour tout $n \ge 0$. La fonction $f_{n,j} = \mathbf{1}_{\mathrm{A}_{n,j}} f_j$ est mesurable bornée et nulle en dehors de l'ensemble de mesure finie $\mathrm{A}_{n,j}$, donc $f_{n,1} + f_{n,2}$ est dans $\mathrm{L}_p(\mu)$; il existe donc g telle que $\|g\|_q = 1$ et

$$\|f_{n,1} + f_{n,2}\|_p = \int (f_{n,1} + f_{n,2}) g = \int f_{n,1} \, g + \int f_{n,2} \, g \le \|f_1\|_p + \|f_2\|_p.$$

On conclut quand $n \to +\infty$ (le pinaillage précédent avec la suite dépendant de n sert seulement à éviter de montrer à part que $f_1 + f_2$ est dans L_p).

Le dual de $\mathrm{L}_p(\Omega, \mu)$

Soit q le nombre tel que $1/q + 1/p = 1$; d'après l'inégalité de Hölder, on a pour toutes fonctions $f \in \mathrm{L}_p$, $g \in \mathrm{L}_q$

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

Ceci signifie que si g est fixée dans L_q, on peut définir une forme linéaire continue ℓ_g sur L_p par la formule

$$\ell_g(f) = \int_\Omega fg \, d\mu.$$

De plus on a vu que

$$\|\ell_g\|_{\mathrm{L}_p^*} = \|g\|_q.$$

On a donc une isométrie $j_q : \mathrm{L}_q \to (\mathrm{L}_p)^*$. Lorsque $1 \le p < +\infty$ on montre (par exemple avec Radon-Nikodym) que cette application est surjective.

Invariance de la mesure de Lebesgue

On commence par vérifier que pour tout borélien A, la mesure de Lebesgue de A est égale à celle de son image par une translation, ou bien par l'application $x \to -x$. Cela provient du fait que la mesure de Lebesgue sur un intervalle $[a, b]$ est l'unique mesure sur la tribu borélienne telle que la mesure de tout intervalle $[c, d] \subset [a, b]$ soit égale à $d - c$ (exercice). On vérifie facilement que la mesure sur l'intervalle translaté $[a + c, b + c]$ qui est la translatée de la mesure de Lebesgue sur $[a, b]$ possède cette propriété, donc c'est la mesure de Lebesgue de l'intervalle $[a + c, b + c]$. Le raisonnement est identique pour l'opération $x \to -x$. On généralise facilement tout ce paragraphe à $\mathbb{R}^n$.

Si f est une fonction étagée positive sur $\mathbb{R}^n$, on vérifie alors facilement que les fonctions $x \to f(x-a)$, ou bien $x \to f(-x)$ ont la même intégrale sur $\mathbb{R}^n$ que f. Par passage à la limite croissante on obtient le même résultat pour f mesurable positive, puis pour f intégrable.

La convolution : inégalités $\mathrm{L}_1 * \mathrm{L}_p$ *et* $\mathrm{L}_p * \mathrm{L}_q$

On commence par le cas de deux fonctions mesurables ≥ 0 et intégrables. On pose

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

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 montrer que l'intégrale ci-dessus est finie pour presque tout $x \in \mathbb{R}^n$. 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 $\mathbb{R}^2$, ce qui permet d'appliquer le théorème de Fubini positif. L'intégrale double sur $\mathbb{R}^2$ de cette fonction $h(x,y)$ est égale aux deux intégrales

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

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

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

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

Soient maintenant $f \in \mathrm{L}_p(\mathbb{R}^n)$ et $g \in \mathrm{L}_1(\mathbb{R}^n)$, avec $1 \leq p < +\infty$, et supposons que f, g soient ≥ 0. On pose toujours

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

à valeurs dans $\overline{\mathbb{R}}$. On veut encore montrer que l'intégrale ci-dessus est finie pour presque tout $x \in \mathbb{R}^n$, et ensuite montrer que la fonction $f * g$ est dans L_p. Si g est nulle presque partout, l'intégrale est nulle. Sinon, on peut supposer en multipliant par une constante que $\int g(y)\,dy = 1$. D'après les propriétés des normes L_p pour la probabilité $g(y)\,dy$, on a pour tout x fixé

$$(\int f(x-y)g(y)\,dy)^p \leq \int f(x-y)^p g(y)\,dy = (f^p * g)(x).$$

D'après ce qui précède, cette quantité est finie pour presque tout x, et $\int (f*g)^p(x)\,dx \leq \|f^p * g\|_1 = \|f^p\|_1 \|g\|_1 = \int f^p(x)\,dx$ (toujours sous l'hypothèse $\int g(y)\,dy = 1$). Par homogénéité on obtient aussi la majoration

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

Le cas des fonctions non nécessairement positives (ou complexes) de L_1 en résulte par majoration du module. On note que pour tout x tel que $(|f| * |g|)(x) < +\infty$, l'intégrale qui définit $(f * g)(x)$ converge ; par conséquent on peut poser

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

et on obtient une fonction définie pour presque tout x, mesurable et qui vérifie la majoration $\|f * g\|_p \le \|f\|_p \|g\|_1$. Il est important de noter que

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

comme on l'a déjà expliqué dans le cas positif ; ici, pour chaque x, ou bien aucune des deux intégrales qui définissent $(f * g)(x)$ et $(g * f)(x)$ n'a de sens (ce qui n'arrive que sur un ensemble négligeable), ou bien les deux intégrales ont un sens (pour Lebesgue) et sont égales. Il existe des variantes importantes de la convolution pour les fonctions périodiques, ou ce qui revient au même, pour les fonctions sur le cercle unité $\mathbb{T}$ du plan complexe.

Il est facile aussi de montrer que $\mathrm{L}_p * \mathrm{L}_q \subset \mathrm{L}_\infty$ quand p et q sont conjugués. En effet

$$\Big| \int f(x-y)g(y)\,dy \Big| \le \|f\|_p \|g\|_q$$

d'après l'inégalité de Hölder et les propriétés d'invariance de la mesure de Lebesgue (la fonction $y \to f(x-y)$ a la même norme dans que f dans L_p). On a donc $\|f * g\|_\infty \le \|f\|_p \|g\|_q$ quand p et q sont conjugués. De plus le résultat $f * g$ est une fonction continue dans ce cas (à voir plus tard).

Les cas $\mathrm{L}_1 * \mathrm{L}_p$ et $\mathrm{L}_p * \mathrm{L}_q$ sont deux cas particuliers d'un résultat plus général, l'inégalité de convolution de Young : si $1/p + 1/q = 1 + 1/r$, on a $\mathrm{L}_p * \mathrm{L}_q \subset \mathrm{L}_r$ (voir exercices).

Prépa. Agrég. écrit d'Analyse, séance n° 5, Mercredi 8 Novembre 2000.

La convolution : régularisation. Continuité et dérivabilité

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 $|f(y)g(x_n - y)| \le \mathrm{M}\,|f(y)|$, qui donne la majoration par une fonction intégrable fixe, et $h_n(y) = f(y)g(x_n - y)$ tend simplement vers $h(y) = f(y)g(x-y)$. La convergence des intégrales 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 $\omega_f(\delta) = \sup\{f(x) - f(x')| : |x - x'| \le \delta\}$ est égal à $\sup\{\|f_v - f\|_\infty : |v| \le \delta\}$.

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

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 $\mathbb{R}^n$; 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 $|f(y)\mathrm{G}(y,t)| \le \mathrm{M}\,|f(y)|$ qui donne la majoration par une fonction intégrable fixe indépendante du paramètre t, pendant que $\mathrm{G}(y,t)$ tend simplement vers $\mathrm{D}_u g(x-y)$ 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 $f \in \mathrm{L}_1(\mathbb{R}^n)$. 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 $\mathrm{D}^\alpha(f * g) = f * \mathrm{D}^\alpha g$.

Il existe des fonctions C^∞ à support compact, et non identiquement nulles sur $\mathbb{R}^d$ (voir exercice 5.3 par exemple). 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.

Approximation et convolution

On utilise souvent le principe de base qui suit.

Théorème : Principe de prolongement par continuité uniforme. *Soient* Z *un sous-ensemble dense d'un espace métrique* (X, d) *et* f *une application* **uniformément** *continue de* (Z, d) *dans un espace métrique* **complet** (Y, δ) *; il existe un unique prolongement continu* $\widetilde{f} : \mathrm{X} \to \mathrm{Y}$ *de* f *; ce prolongement est uniformément continu.*

Démonstration. Si $f_1, f_2 : \mathrm{X} \to \mathrm{Y}$ sont deux prolongements continus de f à X tout entier, l'ensemble $\{x \in \mathrm{X} : f_1(x) = f_2(x)\}$ est un fermé (par continuité de f_1, f_2), qui contient Z (prolongement), donc il est égal à X et $f_1 = f_2$. Cette partie du résultat est une évidence purement topologique (pas d'uniformité là-dedans).

Pour la réciproque on peut raisonner ainsi. Pour chaque $x \in \mathrm{X}$, chaque boule ouverte $\mathrm{B}(x, \alpha)$ (avec $\alpha > 0$) rencontre Z et l'ensemble $\mathrm{C}_\alpha(x) = \{f(z) : z \in \mathrm{Z} \cap \mathrm{B}(x, \alpha)\}$ est un sous-ensemble non vide de Y, de diamètre majoré par la valeur $\omega_f(2\alpha)$ au point 2α du module de continuité uniforme ω_f de f, défini par

$$\forall t > 0, \quad \omega_f(t) = \sup\{\delta(f(z_1), f(z_2)) : z_1, z_2 \in \mathrm{Z},\ d(z_1, z_2) < t\}.$$

Par continuité de la fonction distance δ sur Y il en résulte que le diamètre de l'adhérence $\mathrm{F}_\alpha(x) = \overline{\mathrm{C}_\alpha(x)}$ est aussi $\leq \omega_f(2\alpha)$. La continuité uniforme de f se traduit par le fait que $\lim_{t \to 0} \omega_f(t) = 0$; cela signifie que les fermés non vides monotones $\mathrm{F}_\alpha(x)$ de l'espace complet Y ont un diamètre qui tend vers 0 lorsque $\alpha \to 0$, ce qui entraîne que leur intersection contient un point unique, qu'on appellera bien sûr $\widetilde{f}(x)$. On a $\delta(\widetilde{f}(x), f(z)) \leq \omega_f(2\alpha)$ pour tout $\alpha > 0$ et $z \in \mathrm{Z} \cap \mathrm{B}(x, \alpha)$, puisque $f(z) \in \mathrm{F}_\alpha(x)$ et que le diamètre de $\mathrm{F}_\alpha(x)$ est majoré par $\omega_f(2\alpha)$.

On vérifie pour finir que le module de continuité de $\widetilde{f}$ est le même que celui de f : si $d(x_1, x_2) < t$, on pourra considérer deux points z_1, z_2 de Z tels que pour $j = 1, 2$ on ait $d(z_j, x_j) < \varepsilon < \frac{1}{2}\bigl(t - d(x_1, x_2)\bigr)$. Alors $d(z_1, z_2) < t$, donc $\delta(f(z_1), f(z_2)) \leq \omega_f(t)$, et $\delta(\widetilde{f}(x_1), \widetilde{f}(x_2)) \leq \omega_f(t) + 2\omega_f(2\varepsilon)$ qui tend vers $\omega_f(t)$ quand ε tend vers 0.

Corollaire. *Soient* Z *un sous-espace vectoriel dense d'un espace vectoriel normé* X *et* f *une application linéaire continue de* Z *dans un espace de Banach* Y *; il existe un unique prolongement continu* $\widetilde{f} : \mathrm{X} \to \mathrm{Y}$ *de* f *; ce prolongement est linéaire et* $\|\widetilde{f}\| = \|f\|$.

Linéaire et continue donne uniformément continue. Il est clair que l'extension est linéaire : on prend deux suites $(z_{j,n})$ dans Z qui convergent vers $x_j \in \mathrm{X}$, $j = 1, 2$. Alors $z_{1,n} + z_{2,n}$ tend vers $x_1 + x_2$ donc

$$\widetilde{f}(x_1 + x_2) = \lim_n f(z_{1,n} + z_{2,n}) = \lim_n f(z_{1,n}) + \lim_n f(z_{2,n}) = \cdots$$

Corollaire. *Soient* Z *un sous-espace vectoriel dense d'un espace vectoriel normé* X *et* (f_n) *une suite d'applications linéaires continues de* Z *dans un espace de Banach* Y*, telles que* $\sup \|f_n\| < +\infty$ *(équicontinuité) et que* $f(z) = \lim_n f_n(z)$ *existe pour tout* $z \in \mathrm{Z}$ *; pour tout* n *soit* $\widetilde{f}_n : \mathrm{X} \to \mathrm{Y}$ *l'unique prolongement de* f_n *; si* $\widetilde{f}$ *désigne le prolongement de* f *à* X *on a* $\widetilde{f}(x) = \lim_n \widetilde{f}_n(x)$ *pour tout* $x \in \mathrm{X}$.

Une démonstration un peu pédante consiste à dire que l'espace $c(\mathrm{Y})$ des suites convergentes de vecteurs de Y est un espace de Banach $\mathcal{Y}$ et que l'hypothèse du corollaire donne une application linéaire continue $\mathcal{F}$ de Z dans $\mathcal{Y}$ (attention : c'est ici qu'on utilise $\sup_n \|f_n\| < +\infty$). L'existence de l'extension de $\mathcal{F}$ à X implique que $\lim_n f_n(x)$ existe pour tout $x \in \mathrm{X}$; l'hypothèse $\sup_n \|f_n\| < +\infty$ montre que l'application $x \in \mathrm{X} \to \lim_n f_n(x)$ est continue sur X, c'est donc l'unique prolongement de f.

Exemple. On considère la famille des translations τ_t sur $\mathrm{L}_p(\mathbb{R}^d)$, définies par

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

ce sont toutes des isométries de L_p, donc $\|\tau_t\| = 1$ pour tout $t \in \mathbb{R}^d$; le sous-espace vectoriel $\mathrm{Z} = \mathcal{K}(\mathbb{R}^d)$ des fonctions continues à support compact est dense dans $\mathrm{L}_p(\mathbb{R}^d)$ (voir exercice 3.5) et pour tout $z \in \mathrm{Z}$ on a $\|z - \tau_t(z)\|_p \to 0$ lorsque $t \to 0$ (continuité uniforme plus considérations de support). Il en résulte que ce résultat est vrai pour toute fonction de L_p.

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.

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

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

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

par Hölder, et $\|f_t - f\|_p \to 0$ d'après ce qui précède.

Approximations de l'identité

Une *approximation de l'identité* est une suite (φ_k) dans $\mathrm{L}_1(\mathbb{R}^d)$ telle que

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

$$(b) \qquad \forall k \ge 0, \quad \int \varphi_k = 1,$$

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

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

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

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

parce que $\int \varphi_k = 1$. La fonction g est bornée par un certain nombre K. Puisque g est uniformément continue, on peut trouver $\alpha > 0$ tel que $|g(x-y) - g(x)| < \varepsilon$ pour tout $y \in \mathbb{R}^n$ 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 $\mathrm{M} = \sup_k \|\varphi_k\|_1$

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

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

Démontrons la deuxième partie. Les opérateurs définis sur $\mathrm{L}_p(\mathbb{R}^d)$ par $\mathrm{T}_k g = g * \varphi_k$ vérifient $\|\mathrm{T}_k\| \le \|\varphi_k\|_1 \le \mathrm{M}$ pour tout k, donc il suffit de démontrer la convergence vers l'identité pour un sous-espace vectoriel dense. Pour $g \in \mathcal{K}(\mathbb{R}^n)$, 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 $\varphi'_k = \mathbf{1}_{\mathrm{B}}\, \varphi_k$ et $\varphi''_k = \varphi_k - \varphi'_k$. D'après ($c$), on a $\|\varphi''_k\|_1 \to 0$, ce qui implique que $\varphi''_k * g$ tend vers 0 uniformément et aussi dans L_p (parce que g est à la fois dans L_∞ et dans L_p). Soit $\mathrm{R} > 0$ tel que le support de g soit contenu dans $\mathrm{B}(0, \mathrm{R})$; on vérifie que $g * \varphi'_k$ est nulle hors de $\mathrm{B}' = \mathrm{B}(0, \mathrm{R} + \alpha)$; on sait que $g * \varphi'_k - g$ tend uniformément vers 0 ; on a donc

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

Une méthode utile pour construire des approximations de l'identité 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 $x' = kx$ on obtient les propriétés voulues.

La classe $\mathcal{D}$ de Schwartz

C'est la classe $\mathcal{D}(\mathbb{R}^d)$ formée de toutes les fonctions C^∞ 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'identité (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 $\psi \in \mathcal{D}(\mathbb{R}^n)$, 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 $1 \le p < +\infty$.

La convolution : suite

La convolution périodique. Dans le cadre 2π-périodique (dimension un) on définit la convolution par

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

Le résultat ne dépend pas de a. Les résultats de régularisation et d'approximation sont faciles à adapter à cette situation.

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 $|c_n(f)| \le \|f\|_1$ 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 $\mathrm{M} \le \mathrm{N}$ et $\mathrm{M}, \mathrm{N} \in \mathbb{Z}$), on a

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

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 $n \ne 0$, et $|c_n(f)| \le \|f'\|_1/|n|$; 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 L_1, 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 des fonctions en escalier.

Théorème. *Soit f une fonction 2π-périodique et intégrable ; pour tout $s \in \mathbb{R}$ tel que $\int_{-\pi}^{\pi} |f(s+h) - f(s)|\,dh/|h| < +\infty$, on a*

$$f(s) = \sum_{n \in \mathbb{Z}} c_n(f)\,\mathrm{e}^{ins}.$$

En remplaçant f par $f - f(s)$ on se ramène immédiatement à démontrer le résultat qui suit.

Lemme. *Soit f une fonction 2π-périodique et intégrable ; on suppose que $s \in \mathbb{R}$ est tel que $\int_{-\pi}^{\pi} |f(s+h)|\, dh/|h| < +\infty$; on a alors*

$$\sum_{n \in \mathbb{Z}} c_n(f)\, \mathrm{e}^{ins} = 0.$$

Démonstration. Dans la situation de ce lemme, on va montrer que la série indexée par $\mathbb{Z}$ peut se diviser en deux séries convergentes, la série des $n \geq 0$ et celle des $n < 0$, c'est à dire que $\sum_{-\mathrm{M}}^{\mathrm{N}}$ converge quand $\mathrm{M}, \mathrm{N} \to +\infty$ indépendamment (ce n'est pas toujours le cas, par exemple dans le théorème classique de Dirichlet). Posons

$$f_{\mathrm{M,N}}(s) = \mathrm{S}_{\mathrm{M,N}}(f)(s) = \sum_{n=-\mathrm{M}}^{\mathrm{N}} c_n(f)\, \mathrm{e}^{ins}\,.$$

D'après les remarques précédentes sur les convolutions (équation (N)), on obtient que $f_{\mathrm{M,N}} = \mathrm{K}_{\mathrm{M,N}} * f$, où

$$\mathrm{K}_{\mathrm{M,N}}(t) = \sum_{n=-\mathrm{M}}^{\mathrm{N}} \mathrm{e}^{int} = \mathrm{e}^{-i\mathrm{M}t}\ \frac{1 - \mathrm{e}^{i(\mathrm{M+N}+1)t}}{1 - \mathrm{e}^{it}}.$$

On a

$$f_{\mathrm{M,N}}(s) = \int_0^{2\pi} \mathrm{K}_{\mathrm{M,N}}(t)\, f(s-t)\, \frac{dt}{2\pi} = \int_0^{2\pi} (\mathrm{e}^{-i\mathrm{M}t} - \mathrm{e}^{i(\mathrm{N}+1)t})\, \frac{f(s-t)}{1 - \mathrm{e}^{it}}\, \frac{dt}{2\pi}.$$

Posons $g(t) = f(s-t)/(1 - \mathrm{e}^{it})$. D'après l'hypothèse (et le fait que $1 - \mathrm{e}^{it} \sim it$ pour t voisin de 0) cette fonction g est intégrable sur $(0, 2\pi)$, ce qui entraîne que les coefficients de Fourier de g tendent vers 0 d'après le lemme de Riemann-Lebesgue. Or nous avions

$$f_{\mathrm{M,N}}(s) = c_{\mathrm{M}}(g) - c_{-\mathrm{N}-1}(g)$$

qui tend donc vers 0 quand $\mathrm{M}, \mathrm{N} \to +\infty$.

Prépa. Agrég. écrit d'Analyse, séance n° 6, Mercredi 15 Novembre 2000.

Séries de Fourier : suite

Le résultat de convergence ponctuelle de la séance précédente peut s'énoncer ainsi : si f est une fonction 2π-périodique, intégrable sur $[0, 2\pi]$, et s'il existe une valeur a telle que

$$\int_{-\pi}^{+\pi} \Big| \frac{f(s+t) - a}{t} \Big| \, dt < +\infty,$$

alors la série de Fourier de f converge au point s (aux deux infinis) et on a

$$a = \sum_{n=-\infty}^{+\infty} c_n(f) \, \mathrm{e}^{ins} \, .$$

La démonstration qu'on a vue est en fait extrêmement simple. On a vu qu'en posant $g_s(t) = (f(s-t) - a)/(1 - \mathrm{e}^{it})$, on a

$$(\mathrm{S}_{\mathrm{M,N}} f)(s) - a = c_{\mathrm{M}}(g_s) - c_{-\mathrm{N}-1}(g_s)$$

qui tend vers 0 quand $\mathrm{M}, \mathrm{N} \to +\infty$ parce que g_s est intégrable (lemme de Riemann-Lebesgue).

En général, ce résultat sera appliqué avec $a = f(s)$, et il montrera que pour tout point s au voisinage duquel la fonction f est assez régulière (et ceci s'applique en particulier quand $f'(s)$ existe), la série de Fourier converge (bilatéralement) et sa somme vaut $f(s)$.

Cette convergence aux deux côtés n'est absolument pas la situation générale. Dans la situation du théorème de Dirichlet, où une fonction de classe C^1 par morceaux admet au point s une limite à droite et une limite à gauche, il faut prendre des sommes de Fourier symétriques pour obtenir un résultat.

Par exemple, si $f_0 = \mathbf{1}_{(0,\pi)}$, on voit que $c_0(f_0) = 1/2$, $c_{2k}(f_0) = 0$ si $k \neq 0$ et si $n = 2k+1$, $c_n(f_0) = c_{2k+1}(f) = \frac{1}{\pi i n}$, ce qui montre à l'évidence que la série de Fourier $\sum_{n=0}^{+\infty} c_n(f_0) \, \mathrm{e}^{ins}$ ne peut pas converger au point $s = 0$, alors que le théorème de Dirichlet s'applique (pour les points différents de 0 et de π, la condition intégrale précédente s'applique, et on a la convergence bilatérale de la série). On pose donc classiquement

$$\mathrm{S}_n(f; t) = \sum_{k=-n}^{n} c_k(f) \, \mathrm{e}^{ikt} \, .$$

Dans l'exemple de f_0, on obtient ainsi, en regroupant les indices k et $-k$

$$\mathrm{S}_n(f_0; s) = \frac{1}{2} + \frac{2}{\pi} \sum_{k : 2k+1 \le n} \frac{\sin(2k+1)s}{2k+1}.$$

On constate bien qu'au point $s = 0$, la série converge (en stationnant) vers la valeur $1/2$, qui est la demi-somme des limites à droite et gauche de f_0 au point 0. Par ailleurs, d'après le résultat rappelé au début, on a pour $0 < s < \pi$

$$1 = f_0(s) = \frac{1}{2} + \frac{2}{\pi} \sum_{k=0}^{+\infty} \frac{\sin(2k+1)s}{2k+1}$$

alors que pour $\pi < s < 2\pi$

$$0 = f_0(s) = \frac{1}{2} + \frac{2}{\pi} \sum_{k=0}^{+\infty} \frac{\sin(2k+1)s}{2k+1}.$$

Condition de Dini

Conformément à la discussion qui précède, on va se limiter maintenant aux sommes de Fourier S_n usuelles, symétriques, pour obtenir un théorème à la Dirichlet, sous la condition dite *de Dini*. Dans ce cas le noyau utilisé sera une fonction paire, ce qui permet d'obtenir le critère suivant :

Théorème. *Si f 2π-périodique est intégrable sur $(0, 2\pi)$ et si a est une valeur telle que*

$$\int_0^{+\pi} \Bigl| \frac{f(s+t) + f(s-t) - 2a}{t} \Bigr| \, dt < +\infty,$$

alors $\lim_n \mathrm{S}_n(f; s) = a$.

Cela s'applique en particulier dans le cas suivant : si f admet deux "limites" f_+ et f_- à droite et à gauche du point s, telles que

$$\int_0^{+\pi} \Bigl| \frac{f(s+t) - f_+}{t} \Bigr| \, dt + \int_0^{+\pi} \Bigl| \frac{f(s-t) - f_-}{t} \Bigr| \, dt < +\infty,$$

alors $\lim_n \mathrm{S}_n(f; s) = \frac{1}{2}(f_+ + f_-)$. On retrouve de cette façon l'énoncé le plus habituel du théorème de Dirichlet.

Convergence uniforme de la série de Fourier

On utilise le module de continuité d'une fonction 2π-périodique continue f,

$$\omega_f(t) = \sup\{|f(x) - f(y)| : |x - y| < t\}.$$

En modifiant légèrement la démonstration du critère de Dini, on obtient :

Théorème. *Soit f une fonction 2π-périodique. Si*

$$(*) \qquad \int_0^{\pi} \frac{\omega_f(t)}{t} \, dt < +\infty$$

la convergence de la série de Fourier de f est uniforme.

Le ressort supplémentaire pour la démonstration est d'obtenir un lemme de Riemann-Lebesgue qui s'applique uniformément aux fonctions (g_s) mentionnées plus haut, qui seront ici données par $g_s(t) = (f(s-t) - f(s))/(1 - \mathrm{e}^{it})$. On remarque que si on a un compact dans L_1, on aura Riemann-Lebesgue uniformément pour les éléments de K. Le compact est l'image de $[0, 2\pi]$ par la fonction continue $s \to g_s \in \mathrm{L}_1$. La continuité de l'application $s \to g_s$ vient du critère $(*)$ et de Lebesgue dominé.

En adaptant la démonstration du théorème précédent, on peut obtenir une convergence uniforme "locale", c'est à dire sur un intervalle $[a, b] \subset]c, d[$, en utilisant le module de continuité restreint à l'ouvert $]c, d[$.

Le théorème précédent n'est pas le meilleur critère pour la convergence uniforme. Il suffit que $\omega_f(1/n) = o(1/\ln n)$ (voir Z-Q ; mais ce critère n'est pas plus fin dans tous les cas ; on peut trouver des cas où l'intégrale de $(*)$ converge sans que $\ln(n)\, \omega_f(1/n)$ tende vers 0).

Corollaire. *Soit f une fonction 2π-périodique. Si f est α-hölderienne pour un $\alpha > 0$, la convergence de la série de Fourier est uniforme.*

La théorie L_2

Le premier fait bien connu est que la suite $(e_n)_{n\in\mathbb{Z}}$ est une suite orthonormée. On en déduit facilement que

Lemme. *Pour toute fonction $f \in \mathrm{L}_2(0, 2\pi)$,*

$$\forall n \geq 0, \quad \|\mathrm{S}_n f\|_2 \leq \|f\|_2.$$

Avec l'orthogonalité et Cauchy-Schwarz on trouve

$$\|\mathrm{S}_n f\|_2^2 = \langle \mathrm{S}_n f, \mathrm{S}_n f\rangle = |\langle f, \mathrm{S}_n f\rangle| \leq \|f\|_2 \, \|\mathrm{S}_n f\|_2,$$

et il suffit de diviser par $\|\mathrm{S}_n f\|_2$, s'il est non nul (sinon l'inégalité cherchée est évidente).

La suite des e_n est totale dans L_2 ; une méthode consiste à appliquer le théorème de Stone-Weierstrass (Weierstrass devrait suffire). En effet l'ensemble des combinaisons linéaires est une algèbre stable par conjugaison et qui sépare les points du compact $\mathbb{R}/\mathbb{Z}$.

Mais en fait, les résultats précédents sur la convergence uniforme redémontrent le fait que la suite exponentielle est totale dans L_2. En effet, le critère $(*)$ s'applique aux fonctions continues, périodiques et affines par morceaux ; leur série de Fourier converge uniformément, donc aussi dans L_2 vers f. Comme les opérateurs $f \to \mathrm{S}_n f$ sont équicontinus sur L_2 d'après le lemme précédent, et que les fonctions affines par morceaux sont denses dans L_2, il en résulte que $\mathrm{S}_n f$ tend vers f pour toute fonction $f \in \mathrm{L}_2$.

On obtient en particulier par Parseval, pour toute fonction $f \in \mathrm{L}_2$

$$\int_0^{2\pi} |f(t)|^2 \, \frac{dt}{2\pi} = \sum_{n\in\mathbb{Z}} |c_n(f)|^2.$$

Comme chacun sait, une masse d'exercices calculatoires est fondée sur la relation précédente. On a aussi

Théorème. *Pour tout p tel que $1 < p < +\infty$ et toute fonction $f \in \mathrm{L}_p$, la suite $(\mathrm{S}_n f)$ converge vers f dans L_p.*

Le point qui manque, pour pouvoir procéder comme avec L_2, est de savoir que les projecteurs S_N sont uniformément bornés sur L_p. C'est vrai, mais plus compliqué que ce qui a été vu ici.

Il existe aussi un théorème de convergence presque partout de la série de Fourier pour les fonctions de L_2, le théorème de Carleson, qui date des années 1960 et qui reste extrêmement difficile. Généralisé par Hunt au cas de L_p, $1 < p < +\infty$, il s'énonce ainsi.

Théorème. *Pour tout p tel que $1 < p < +\infty$ et toute fonction $f \in \mathrm{L}_p$, on a*

$$f(t) = \sum_{n\in\mathbb{Z}} c_n(f) \, \mathrm{e}^{int}$$

pour presque tout $t \in [0, 2\pi]$.

En revanche, rien ne va plus dans L_1 (exercice à venir, fondé sur Banach-Steinhaus).

Les noyaux classiques. Théorèmes de Fejér

Posons pour tout entier $k \geq 0$

$$L_k(t) = e^{-ikt/2} + e^{it-ikt/2} + e^{2it-ikt/2} + \cdots + e^{ikt/2}.$$

En faisant apparaître une progression géométrique on voit facilement que

$$L_k(t) = e^{-ikt/2} \frac{e^{i(k+1)t} - 1}{e^{it} - 1} = \frac{\sin(k+1)t/2}{\sin(t/2)}.$$

Quand $k = 2n$ on obtient le noyau de Dirichlet,

$$D_n(t) = \sum_{k=-n}^{n} e^{ikt} = \frac{\sin(n+1/2)t}{\sin(t/2)}.$$

Le noyau de Fejér est donné par

$$K_n = \frac{1}{n+1}(D_0 + \cdots + D_n)$$

(tout le monde n'est pas d'accord sur la numérotation : comparer Chatterji et Zuily-Queffélec). Puisque $D_k = \sum_{j=-k}^{k} e_j$ on obtient facilement (à la Fubini)

$$K_n = \sum_{k=-n}^{n} \Big(1 - \frac{|k|}{n+1}\Big) e_k. \tag{1}$$

Par ailleurs

$$L_n^2(t) = e^{-int}\big(1 + \cdots + e^{int}\big)^2 = \sum_{k=-n}^{n} (n+1-|k|)\, e^{ikt},$$

ce qui donne $(n+1)K_n = L_n^2$,

$$K_n(t) = \frac{1}{n+1} \frac{\sin^2(n+1)t/2}{\sin^2(t/2)}.$$

On a donc $K_n(t) \geq 0$. En utilisant la forme (1) de K_n on voit que $\int_0^{2\pi} K_n(t) \frac{dt}{2\pi} = 1$. De plus, pour tout α tel que $0 < \alpha < \pi$, on constate que

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbf{1}_{\{|t|>\alpha\}} K_n(t)\, dt \leq \frac{1}{(n+1)\pi} \int_{\alpha}^{\pi} \frac{1}{\sin^2(\alpha/2)}\, dt \leq \frac{1}{(n+1)\sin^2(\alpha/2)}$$

qui tend vers 0. On voit donc que la suite (K_n) fournit une approximation de l'unité. Par les résultats généraux, on a donc des résultats de convergence pour les *sommes de Fejér*

$$\sigma_n(f;t) = (K_n * f)(t) = \frac{1}{n+1} \sum_{k=-n}^{n} \Big(1 - \frac{|k|}{n+1}\Big) c_k(f)\, e^{ikt}.$$

Théorème. *Si f est continue périodique, $\sigma_n f$ converge uniformément vers f. Si $f \in L_p$, avec $1 \leq p < +\infty$, on a $\|\sigma_n f - f\|_p \to 0$.*