Intégration

Licence de mathématiques - 2020/2021

6 novembre 2020

Table des matières

Chapitre 1

Tribus

1.1 Ensembles finis, dénombrables et infinis non dénombrables

Soit E un ensemble. On désigne par $\mathcal{P}(E)$ l'ensemble des sous-ensembles de E :

$$\mathcal{P}(E) = \{A \subset E\}.$$

Définition 1.1.1. On dit que deux ensembles E et F sont équipotents s'il existe une bijection de E sur F. On dit alors qu'ils ont même cardinal.

Définition 1.1.2. On dit qu'un ensemble A est fini s'il existe $n \in \mathbb{N}$ tel que A soit en bijection avec $\{1, \dots, n\}$. On dit alors que le cardinal de A est n. On note $\mathrm{Card}(A) = n$ ou $\#(A) = n$.

Définition 1.1.3. Un ensemble E qui n'est pas fini est dit infini.

On peut également définir un ensemble E comme infini en disant qu'il existe $x_0 \in E$ et une injection de E dans $E \backslash \{x_0\}$.

Exemple 1.1.4. L'ensemble $\mathbb{N}$ est infini puisque $n \mapsto n + 1$ est une injection de $\mathbb{N}$ dans $\mathbb{N}^*$.

Définition 1.1.5. Un ensemble E est dit dénombrable s'il existe une bijection de X sur $\mathbb{N}$. Il est dit infini non dénombrable s'il n'est ni fini, ni dénombrable.

Proposition 1.1.6. *L'ensemble des nombres premiers, $\mathbb{Z}$, $\mathbb{N}^2$, $\mathbb{Q}$ sont dénombrables.*

Démonstration. Raisonnement par l'absurde pour le premier. Pour $\mathbb{N}^2$, on peut considérer l'application suivante :

$$f(p, q) = \frac{(p + q)^2 + p + 3q}{2}.$$

On peut construire une injection de $\mathbb{Q}^+$ dans $\mathbb{N}^2$. $\square$

Proposition 1.1.7. *Un sous-ensemble d'un ensemble au plus dénombrable est au plus dénombrable.*

Proposition 1.1.8. *Une réunion au plus dénombrable d'ensembles au plus dénombrables est au plus dénombrable.*

Corollaire 1.1.9. *L'ensemble des nombres algébriques (racines d'un polynôme à coefficients entiers) est dénombrable.*

Proposition 1.1.10. *L'ensemble* $[0,1]$ *est non dénombrable.*

Les ensembles $\mathcal{P}(E)$ et $\{0,1\}^E$ sont équipotents car l'application

$$A \in \mathcal{P}(X) \mapsto \mathbf{1}_A \in \{0,1\}^E \quad \text{où} \quad \mathbf{1}_A(x) = \begin{cases} 1 & \text{si } x \in A, \\ 0 & \text{si } x \notin A, \end{cases}$$

est bijective.

Exemple 1.1.11. L'ensemble $\mathbb{N}$ est équipotent à $2\mathbb{N}$ car $n \in \mathbb{N} \mapsto 2n \in 2\mathbb{N}$ est une bijection.

Proposition 1.1.12. *Soit E un ensemble. Alors E et $\mathcal{P}(E)$ ne sont pas équipotents car il n'existe pas de surjection de E sur $\mathcal{P}(E)$.*

Démonstration. Supposons qu'une telle surjection existe et notons-là f. Alors il existe un antécédent $a \in E$ par f à l'ensemble $A = \{x \in E \,:\, x \notin f(x)\}$. Si $a \in A$ alors $a \notin f(a) = A$ et si $a \notin A$ alors $a \in f(a) = A$. Dans les deux cas, on obtient une contradiction. □

1.2 Opérations sur les ensembles

Définition 1.2.1. Soit E un ensemble.
Si A et B sont inclus dans E on note $A \cup B$ l'union de A et B définie par

$$A \cup B = \{x \in E \,:\, x \in A \text{ ou } x \in B\}.$$

Si A et B sont inclus dans E on note $A \cap B$ l'intersection de A et B définie par

$$A \cap B = \{x \in E \,:\, x \in A \text{ et } x \in B\}.$$

Si A est inclus dans E, on note A^c son complémentaire (dans E) défini par

$$A^c = \{x \in E \,:\, x \notin A\}.$$

Proposition 1.2.2. *Soit A, B et C des sous-ensembles de E. Alors*

$$(A \cup B)^c = A^c \cap B^c, \quad (A \cap B)^c = A^c \cup B^c \qquad \textit{et} \quad (A \cup B) \cap C = (A \cap C) \cup (B \cap C).$$

Définition 1.2.3. Soit f une application de E dans F, A inclus dans E et B inclus dans F. L'image directe de A par f, notée $f(A)$ est le sous-ensemble de F suivant :

$$f(A) = \{y \in F \,:\, \exists x \in A,\ f(x) = y\} = \{f(x) \,:\, x \in A\}.$$

L'image réciproque de B par f, notée $f^{-1}(B)$ est le sous-ensemble de E suivant :

$$f^{-1}(B) = \{x \in E \,:\, f(x) \in B\}.$$

Proposition 1.2.4. *Soit f de E dans F, $A, A' \subset E$ et $B, B' \subset F$. On a*

$$f^{-1}(B \cup B') = f^{-1}(B) \cup f^{-1}(B') \quad et \quad f^{-1}(B \cap B') = f^{-1}(B) \cap f^{-1}(B')$$

$$f(A \cup A') = f(A) \cup f(B') \quad et \quad f(B \cap B') \subset f(B) \cap f(B').$$

L'incluse manquante est fausse en général.

Démonstration. Soit $x \in f^{-1}(B \cup B')$. Alors $f(x) \in B \cup B'$. Si $f(x) \in B$ alors $x \in f^{-1}(B)$ et $x \in f^{-1}(B) \cup f^{-1}(B')$. La conclusion est la même si $f(x) \in B'$. □

1.3 Définitions

Soit E un ensemble. On appelle classe de parties de E un sous-ensemble non vide de $\mathcal{P}(E)$.

Définition 1.3.1. Une tribu $\mathcal{A}$ sur E est un sous-ensemble non vide de $\mathcal{P}(E)$ tel que :

(i) la partie vide appartient à $\mathcal{A}$,

(ii) le complémentaire d'un élément de $\mathcal{A}$ est dans $\mathcal{A}$,

(iii) $\mathcal{A}$ est stable par réunion dénombrable.

Notons immédiatement quelques propriétés satisfaites par les tribus. Si $\mathcal{A}$ est une tribu alors
- $E \in \mathcal{A}$,
- $\mathcal{A}$ est stable par intersection dénombrable,
- $\mathcal{A}$ est stable par différence : $A, B \in \mathcal{A} \Rightarrow A \backslash B \in \mathcal{A}$,
- $\mathcal{A}$ est stable par différence symétrique : $A, B \in \mathcal{A} \Rightarrow A \Delta B \in \mathcal{A}$.

Exemple 1.3.2. La plus petite tribu de E est $\mathcal{A} = \{\emptyset, E\}$, tandis que la plus grande est $\mathcal{P}(E)$.

Définition 1.3.3. On appelle espace mesurable tout couple $(E, \mathcal{A})$ formé par un ensemble E et une tribu $\mathcal{A}$ sur E.

Proposition 1.3.4. *L'intersection de tribus sur E est encore une tribu.*

Démonstration. Soit $(\mathcal{A}_i)_{i \in I}$ une famille de tribus sur E. Notons $\mathcal{A} = \cap_{i \in I} \mathcal{A}_i$ l'intersection de ces tribus. Alors, l'ensemble vide appartient à chaque tribu $\mathcal{A}_i$ et donc à $\mathcal{A}$. Soit $A \in \mathcal{A}$. Pour tout $i \in I$, $A \in \mathcal{A}_i$ donc $A^c \in A_i$: $A^c \in \mathcal{A}$. La réunion dénombrable s'établit de même. □

Proposition 1.3.5 (tribu engendrée)**.** *Soit $\mathcal{E}$ un sous-ensemble de $\mathcal{P}(E)$. Il existe une plus petite tribu (au sens de l'inclusion) contenant tous les éléments de $\mathcal{E}$. Elle est appelée tribu engendrée par $\mathcal{E}$, et est notée $\sigma(\mathcal{E})$.*

Démonstration. Soit $\mathcal{X}$ l'ensemble de toutes les tribus $\mathcal{M}$ sur E contenant $\mathcal{E}$. L'ensemble $\mathcal{X}$ n'est pas vide puisqu'il contient la tribu $\mathcal{P}(E)$. Posons

$$\mathcal{A} = \bigcap_{\mathcal{M} \in \mathcal{X}} \mathcal{M} = \{A \subset E, \ \forall \mathcal{M} \in \mathcal{X}, \ A \in \mathcal{M}\}.$$

Il est clair que $\mathcal{E} \subset \mathcal{A}$. De plus, $\mathcal{A}$ est une tribu comme intersection (quelconque) de tribus et, par définition, elle est contenue dans toute tribu contenant $\mathcal{E}$. □

Remarque 1.3.6. Si A est un sous-ensemble de E, alors

$$\sigma(\{A\}) = \{\emptyset, A, A^c, E\}.$$

Remarque 1.3.7. Si $\mathcal{A}$ est une tribu sur E, alors $\sigma(\mathcal{A}) = \mathcal{A}$.

Exemple 1.3.8. Si E est un ensemble dénombrable alors la tribu engendrée par les singletons $(\{x\})_{x\in E}$ est l'ensemble de toutes les parties de E : $\mathcal{P}(E)$.

Proposition 1.3.9 (image réciproque d'une tribu)**.** *Soit E et F deux ensembles, f une application de E dans F et $\mathcal{A}$ une tribu sur F. Alors*

$$f^{-1}(\mathcal{A}) = \left\{f^{-1}(A),\ A \in \mathcal{A}\right\}$$

est une tribu sur E, appelée tribu image réciproque de $\mathcal{A}$ par f.

Démonstration. La classe $f^{-1}(\mathcal{A})$ contient l'ensemble vide puisque $\emptyset = f^{-1}(\emptyset)$. Soit $B \in f^{-1}(\mathcal{A})$. Alors il existe $A \in \mathcal{A}$ tel que $B = f^{-1}(A)$. Puisque $\mathcal{A}$ est une tribu, $A^c \in \mathcal{A}$. Enfin, remarquons que $f^{-1}(A)^c = f^{-1}(A^c)$. La stabilité par réunion dénombrable s'établit de même. □

Proposition 1.3.10. *Soit f une application de E dans F, $\mathcal{E}$ un sous-ensemble de $\mathcal{P}(F)$. Alors*

$$f^{-1}(\sigma(\mathcal{E})) = \sigma\big(f^{-1}(\mathcal{E})\big).$$

En d'autres termes, l'image réciproque de la tribu engendrée par $\mathcal{E}$ est la tribu engendrée par l'image réciproque de $\mathcal{E}$.

Démonstration. Comme $\mathcal{E} \subset \sigma(\mathcal{E})$, on a $f^{-1}(\mathcal{E}) \subset f^{-1}(\sigma(\mathcal{E}))$ qui est une tribu et ainsi $\sigma(f^{-1}(\mathcal{E}))$ est inclus dans $f^{-1}(\sigma(\mathcal{E}))$.

Montrons l'inclusion inverse. Notons

$$\mathcal{B} = \left\{B \subset F\ :\ f^{-1}(B) \in \sigma(f^{-1}(\mathcal{E}))\right\}.$$

Alors $\mathcal{B}$ est une tribu. De plus, $\mathcal{B}$ contient $\mathcal{E}$ donc contient $\sigma(\mathcal{E})$. Il en résulte que $f^{-1}(\sigma(\mathcal{E}))$ est inclus dans $f^{-1}(\mathcal{B})$. Comme, par définition, $f^{-1}(\mathcal{B}) \subset \sigma(f^{-1}(\mathcal{E}))$, on obtient l'inclusion souhaitée : $f^{-1}(\sigma(\mathcal{E}))$ est inclus dans $\sigma(f^{-1}(\mathcal{E}))$. □

Définition 1.3.11 (tribu induite)**.** Soit B un sous-ensemble de E et $\mathcal{A}$ une tribu sur E. On appelle tribu trace, ou tribu induite, par $\mathcal{A}$ sur B la tribu

$$\mathcal{A}_B = \{A \cap B,\ A \in \mathcal{A}\}.$$

Définition 1.3.12 (tribu produit)**.** Soit $\mathcal{A}$ une tribu sur E et $\mathcal{B}$ une tribu sur F. On appelle tribu produit, et l'on note $\mathcal{A} \otimes \mathcal{B}$, la tribu sur $E \times F$ engendrée par l'ensemble des parties de $E \times F$ qui s'écrivent sous la forme $A \times B$ avec $A \in \mathcal{A}$ et $B \in \mathcal{B}$.

1.4 Tribu borélienne

Rappelons que pour la topologie usuelle de $\mathbb{R}$, un ensemble O de $\mathbb{R}$ est ouvert si

$$\forall x \in O,\ \exists a, b \in O,\ x \in]a, b[\subset O.$$

On note $\mathcal{O}$ l'ensemble des ouverts de $\mathbb{R}$.

Définition 1.4.1. La tribu $\sigma(\mathcal{O})$ engendrée par $\mathcal{O}$ est appelée la tribu borélienne de $\mathbb{R}$. On la note $\mathcal{B}(\mathbb{R})$. Ses éléments sont appelés les boréliens.

Remarque 1.4.2. Même si cela n'est pas évident, on peut montrer que $\mathcal{B}(\mathbb{R})$ est strictement inclus dans $\mathcal{P}(\mathbb{R})$: il existe des parties de $\mathbb{R}$ qui ne sont pas boréliennes.

Proposition 1.4.3. *Sur $\mathbb{R}$, muni de sa topologie usuelle, la tribu borélienne est engendrée par*

1. *la classe des intervalles ouverts bornés,*
2. *la classe des intervalles ouverts bornés à extrémités rationnelles,*
3. *la classe des intervalles de la forme $]-\infty, a[$ avec $a \in \mathbb{R}$,*
4. *la classe des intervalles de la forme $]-\infty, a]$ avec $a \in \mathbb{R}$.*

Démonstration. Prouvons le point 2 qui est plus fort que le point 1. Notons $\mathcal{E}$ la classe des intervalles ouverts bornés à extrémités rationnelles. On a $\mathcal{E} \subset \mathcal{O}$, donc $\sigma(\mathcal{E}) \subset \sigma(\mathcal{O})$. Réciproquement, soit O un ouvert de $\mathbb{R}$. Notons

$$I = \{(\rho, r) \in \mathbb{Q} \times \mathbb{Q}_+^*,\ ]\rho - r, \rho + r[\subset O\}.$$

Alors I est dénombrable et

$$O = \bigcup_{(\rho,r)\in I}]\rho - r, \rho + r[.$$

On voit ainsi que tout ouvert de $\mathbb{R}$ peut s'écrire comme réunion dénombrable d'intervalles ouverts à extrémités rationnelles, d'où $\mathcal{O} \subset \sigma(\mathcal{E})$ et par suite $\sigma(\mathcal{O}) \subset \sigma(\mathcal{E})$.

Prouvons le point 3. Soit $\mathcal{E}'$ la classe des intervalles de la forme $]-\infty, a[$. On a $\sigma(\mathcal{E}') \subset \sigma(\mathcal{O})$. Pour établir l'inclusion inverse, il suffit de montrer que $\mathcal{E} \subset \sigma(\mathcal{E}')$ (puisque la tribu engendrée par $\mathcal{E}$ est la tribu borélienne). Soit $]a, b[\in \mathcal{E}$. On a

$$\begin{aligned}]a, b[&=]-\infty, b[\cap]a, +\infty[=]-\infty, b[\cap]-\infty, a]^c \\ &=]-\infty, b[\cap (\cap_{n\in\mathbb{N}^*}]-\infty, a + 1/n[)^c \in \sigma(\mathcal{E}'). \end{aligned}$$

Tout intervalle $]a, b[$ appartient donc à la tribu engendrée par $\mathcal{E}'$ et donc $\sigma(\mathcal{E}) \subset \sigma(\mathcal{E}')$.

Le point 4 s'établit de manière analogue. □

Remarque 1.4.4. Nous aurons aussi à considérer la droite achevée $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$. Rappelons que sa topologie est définie par la base d'ouverts formés des intervalles ouverts de la forme $]a, b[$, $]a, +\infty]$ et $[-\infty, b[$ avec $a, b \in \overline{\mathbb{R}}$. On démontre de façon analogue que la tribu borélienne de $\overline{\mathbb{R}}$ est engendrée par les classes $\{[-\infty, a[,\ a \in \overline{\mathbb{R}}\}$ ou $\{[-\infty, a],\ a \in \overline{\mathbb{R}}\}$ par exemple.

Proposition 1.4.5. *La tribu borélienne de $\mathbb{R}^d$ est égale à la tribu engendrée par la classe des ouverts de la forme*

$$\prod_{i=1}^{d}]a_i, b_i[\quad \textit{avec } \infty < a_i < b_i < +\infty.$$

Il existe une notion d'espace topologique abstrait. Rappelons que $\mathcal{O} \subset \mathcal{P}(E)$ est une topologie (l'ensemble des ouverts) sur E si

(i) $\emptyset$ et E appartiennent à $\mathcal{O}$,

(ii) $\mathcal{O}$ est stable par intersection finie,

(iii) $\mathcal{O}$ est stable par réunion quelconque.

Il est naturel de munir un espace topologique $(E, \mathcal{O})$ (où $\mathcal{O}$ désigne l'ensemble des ouverts de E) d'une tribu compatible en un certain sens avec la structure topologique préexistante.

Définition 1.4.6. Soit $(E, \mathcal{O})$ un espace topologique. La tribu $\sigma(\mathcal{O})$ engendrée par $\mathcal{O}$ est appelée la tribu borélienne de E. On la note $\mathcal{B}(E)$. Ses éléments sont appelés les boréliens.

Chapitre 2

Mesures positives

2.1 Définitions et propriétés élémentaires

Définition 2.1.1. Soit $(E, \mathcal{A})$ un espace mesurable. On appelle mesure positive sur $(E, \mathcal{A})$ une application μ de $\mathcal{A}$ dans $\overline{\mathbb{R}}_+$ telle que

(i) $\mu(\emptyset) = 0$,

(ii) si $(A_n)_{n\in\mathbb{N}}$ est une suite d'éléments deux à deux disjoints d'éléments de $\mathcal{A}$, alors

$$\mu(\cup_n A_n) = \sum_{n=0}^{\infty} \mu(A_n).$$

On peut parfois préciser le terme de mesure positive :

— s'il existe une suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{A}$ telle que $\cup_n A_n = E$ et, pour tout $n \in \mathbb{N}$, $\mu(A_n)$ est fini, on dit que μ est une mesure σ-finie ;

— si $\mu(E) < +\infty$, on dit que la mesure μ est finie (ou bornée) ;

— si $\mu(E) = 1$, la mesure μ est appelée mesure de probabilité.

Exemple 2.1.2. La mesure de Lebesgue sur $\mathbb{R}$ est σ-finie mais pas finie. La mesure δ_0 est une mesure de probabilité.

Définition 2.1.3. On appelle espace mesuré tout triplet $(E, \mathcal{A}, \mu)$ où $(E, \mathcal{A})$ est un espace mesurable et μ est une mesure positive sur $(E, \mathcal{A})$.

Analysons à présent les propriétés satisfaites par une mesure en commençant par les propriétés faisant intervenir un nombre fini d'ensembles.

Proposition 2.1.4. *Soit $(E, \mathcal{A}, \mu)$ un espace mesuré.*

(i) Si $A_1, \ldots, A_n$ sont des éléments de $\mathcal{A}$ deux à deux disjoints alors

$$\mu(A_1 \cup A_2 \cup \cdots \cup A_n) = \mu(A_1) + \cdots + \mu(A_n).$$

(ii) Si A et B sont deux éléments de $\mathcal{A}$ tels que $A \subset B$, alors $\mu(A) \leq \mu(B)$. De plus, si $\mu(A) < +\infty$, alors $\mu(B \backslash A) = \mu(B) - \mu(A)$.

(iii) Soient A et B deux éléments de $\mathcal{A}$, $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$.

Démonstration. Le point (i) s'établit à partir du point (ii) de la définition d'une mesure en choisissant $A_k = \emptyset$ pour $k \neq 1, \ldots, n$.

Pour (ii), si $A \subset B$, on écrit $B = A \cup (B \backslash A)$. Comme A et $(B \backslash A)$ sont disjoints, $\mu(B)$ est égal à $\mu(A) + \mu(B \backslash A) \geq \mu(A)$.

Pour établir le point (iii), distinguons deux cas. Si $\mu(A \cap B) = +\infty$ alors $\mu(A)$ ou $\mu(B)$ vaut aussi $+\infty$. Sinon, il faut remarquer que $A \cup B$ s'écrit comme la réunion disjointe de $(A \backslash (A \cap B))$, $A \cap B$ et $B \backslash (A \cap B)$. En utilisant le point (i), il vient :

$$\mu(A \cup B) = \mu(A \backslash (A \cap B)) + \mu(A \cap B) + \mu(B \backslash (A \cap B)).$$

Puisque $A \cap B$ est bien entendu inclus dans A et dans B le point (ii) fournit le dernier argument :

$$\begin{aligned} \mu(A \cup B) &= \mu(A) - \mu(A \cap B) + \mu(A \cap B) + \mu(B) - \mu(A \cap B) \\ &= \mu(A) + \mu(B) - \mu(A \cap B), \end{aligned}$$

ce qui est le résultat attendu. □

Donnons une définition équivalente de la notion de mesure (positive).

Proposition 2.1.5. *Une application μ de $\mathcal{A}$ dans $\overline{\mathbb{R}}_+$ est une mesure si et seulement si*

(i) $\mu(\emptyset) = 0$;

(ii) si A et B sont deux éléments disjoints de $\mathcal{A}$, $\mu(A \cup B) = \mu(A) + \mu(B)$;

(iii) pour toute suite croissante $(B_n)_{n \in \mathbb{N}}$ d'éléments de $\mathcal{A}$, $\mu(\cup_n B_n) = \lim_n \mu(B_n)$.

Démonstration. Supposons que les points (i), (ii) et (iii) de la proposition soient vrais. Par récurrence sur le point (ii), on obtient que, si $A_1, \ldots, A_n$ sont des éléments de $\mathcal{A}$ deux à deux disjoints alors

$$\mu(A_1 \cup A_2 \cup \cdots \cup A_n) = \mu(A_1) + \cdots + \mu(A_n).$$

Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'éléments deux à deux disjoints de $\mathcal{A}$. Pour tout $n \in \mathbb{N}$, posons $B_k = \cup_{n \leq k} A_n$. On a $\mu(B_k) = \sum_{n=0}^{k} \mu(A_n)$. De plus, $(B_k)_{k \in \mathbb{N}}$ est une suite croissante et $\cup_{k=0}^{\infty} B_k$ coïncide avec $\cup_{n=0}^{\infty} A_n$. Par hypothèse, on obtient

$$\mu(\cup_{n=0}^{\infty} A_n) = \mu(\cup_{k=0}^{\infty} B_k) = \lim_{k \to \infty} \mu(B_k) = \lim_{n \to \infty} \sum_{k=1}^{n} \mu(A_k) = \sum_{k=0}^{\infty} \mu(A_k).$$

Réciproquement, supposons que μ soit une mesure. Soit $(B_n)_{n \in \mathbb{N}}$ une suite croissante d'éléments de $\mathcal{A}$. Posons $A_0 = B_0$ et, pour tout $n \geq 1$, $A_n = B_n \backslash B_{n-1} \in \mathcal{A}$. Alors $(A_n)_{n \in \mathbb{N}}$ est une suite d'éléments de $\mathcal{A}$ deux à deux disjoints et, pour tout $n \geq 0$, $B_n = \cup_{k=0}^{n} A_k$. Il en résulte que

$$\mu(\cup_{n=0}^{\infty} B_n) = \mu(\cup_{k=0}^{\infty} A_k) = \sum_{k=0}^{\infty} \mu(A_k) = \lim_{n \to \infty} \sum_{k=1}^{n} \mu(A_k) = \lim_{n \to \infty} \mu(B_n),$$

et la proposition est démontrée. □

Proposition 2.1.6. *Soit* $(E, \mathcal{A}, \mu)$ *un espace mesuré.*

(i) Si $(B_n)_{n\in\mathbb{N}}$ *est une suite d'éléments de* $\mathcal{A}$*, alors* $\mu(\cup_{n=0}^{\infty} B_n) \leq \sum_{n=0}^{\infty} \mu(B_n)$.

(ii) Si $(A_n)_{n\in\mathbb{N}}$ *est une suite décroissante d'éléments de* $\mathcal{A}$ *telle qu'il existe* n_0 *avec* $\mu(A_{n_0})$ *fini, alors la suite* $(\mu(A_n))_{n\in\mathbb{N}}$ *converge en décroissant vers* $\mu(\cap_n A_n)$.

Démonstration. Démontrons (i). Posons $A_0 = B_0$ et, pour tout $n \geq 1$, $A_n = B_n \backslash (\cup_{k<n} B_k)$. Les ensembles $(A_n)_{n\in\mathbb{N}}$ sont deux à deux disjoints et $B_n = \cup_{k\leq n} A_k$. Il en résulte que

$$\mu(\cup_{n=0}^{\infty} B_n) = \mu(\cup_{k=0}^{\infty} A_k) = \sum_{k=0}^{\infty} \mu(A_k) \leq \sum_{n=0}^{\infty} \mu(B_n),$$

puisque $A_n \subset B_n$ pour tout $n \in \mathbb{N}$.

Démontrons (ii). Pour $i \geq n_0$, posons $B_k = A_{n_0} \backslash A_k$. La suite $(B_k)_{k\geq n_0}$ est croissante et on a $\cup_{k\geq n_0} B_k = A_{n_0} \backslash (\cap_{k\geq n_0} A_k)$. Puisque $\cap_{k\geq n_0} A_k \subset A_{n_0}$ et $A_k \subset A_{n_0}$, on a

$$\mu(A_{n_0} \backslash (\cap_{k\geq n_0} A_k)) = \mu(A_{n_0}) - \mu(\cap_{k\geq n_0} A_k) \quad \text{et} \quad \mu(B_k) = \mu(A_{n_0}) - \mu(A_k),$$

d'où

$$\begin{aligned} \mu(A_{n_0}) - \mu(\cap_{k\geq n_0} A_k) &= \mu(\cup_{k\geq n_0} B_k) = \lim_{k\to\infty} \mu(B_k) \\ &= \lim_{k\to\infty} (\mu(A_{n_0}) - \mu(A_k)) = \mu(A_{n_0}) - \lim_{k\to\infty} \mu(A_k), \end{aligned}$$

et donc $\mu(\cap_{k\geq 1} A_k) = \lim_{k\to\infty} \mu(A_k)$. □

Remarque 2.1.7. Dans l'énoncé (ii), l'hypothèse de l'existence d'un entier n_0 tel que $\mu(A_{n_0})$ est fini ne peut être supprimée. En effet, si μ est la mesure de comptage sur $\mathbb{N}$ et $A_n = \{n, n+1, \ldots\}$ alors $\mu(A_n) = +\infty$ et $\cap_n A_n = \emptyset$.

2.2 Premières mesures discrètes

Les premiers exemples de mesures que l'on va considérer sont à la fois élémentaires et fondamentaux. Ils correspondent à l'idée intuitive de masses ponctuelles : il va s'agir d'affecter des *poids* à des points de l'espace.

L'exemple le plus naïf consiste à affecter un poids à un seul point.

Définition 2.2.1 (Mesure de Dirac)**.** Soit $(E, \mathcal{A})$ un espace mesuré et $a \in E$. Posons, pour tout $A \in \mathcal{A}$:

$$\delta_a(A) = \begin{cases} 1 & \text{si } a \in A, \\ 0 & \text{si } a \notin A. \end{cases}$$

L'application δ_a est une mesure de probabilité, appelée mesure (ou masse) de Dirac au point a.

Remarque 2.2.2. Si $A \in \mathcal{A}$, $\delta_a(A) = \mathbf{1}_A(a)$.

Pour montrer que δ_a est une mesure, utilisons par exemple la définition alternative d'une mesure fournie par la proposition 2.1.5. Il est clair que $\delta_a(\emptyset)$ est nul. Soit A et B deux ensembles disjoints appartenant à $\mathcal{A}$. Alors

$$\delta_a(A \cup B) = \mathbf{1}_{A\cup B}(a) = \mathbf{1}_A(a) + \mathbf{1}_B(a) = \delta_a(A) + \delta_a(B).$$

Soit à présent une suite croissante $(B_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{A}$. Alors on a

$$a \in \cup_n B_n \iff \exists n_0 \geq 0,\ a \in B_{n_0} \iff \exists n_0 \geq 0,\ \forall n \geq n_0,\ a \in B_n.$$

Ainsi, si $a \in \cup_n B_n$ alors $\delta_a(\cup_n B_n) = 1$ et la suite $(\delta_a(B_n))_{n\in\mathbb{N}}$ (à valeurs dans $\{0,1\}$) vaut 1 à partir d'un certain rang. De même, si $a \notin \cup_n B_n$ alors $\delta_a(\cup_n B_n) = 0$ et la suite $(\delta_a(B_n))_{n\in\mathbb{N}}$ est la suite nulle.

Définition 2.2.3 (Mesure de Bernoulli)**.** Soit $p \in]0,1[$. La mesure de Bernoulli de paramètre p est définie par $\mu = (1-p)\delta_0 + p\delta_1$. C'est une mesure de probabilité.

Définition 2.2.4 (Mesures discrètes finies)**.** Soit $(\mathbb{R},\mathcal{B})$, $N \in \mathbb{N}^*$, $(a_n)_{1\leq n\leq N}$ des réels et $(\alpha_n)_{1\leq n\leq N}$ des réels strictement positifs. Posons pour tout $A \in \mathcal{A}$,

$$\mu(A) = \sum_{n=1}^{N} \alpha_n \delta_{a_n}(A).$$

L'application $\mu : \mathcal{B} \to \overline{\mathbb{R}}_+$ est une mesure positive. Les points $(a_n)_{1\leq n\leq N}$ sont appelés atomes de μ.

Montrons que l'application μ définie sur $\mathcal{B}$ est une mesure. Clairement, $\mu(\emptyset) = 0$. Soit $(A_k)_{k\in\mathbb{N}}$ une suite d'éléments deux à deux disjoints de $\mathcal{B}$. Alors

$$\begin{aligned}
\mu(\cup_k A_k) &= \sum_{n=1}^{N} \alpha_n \delta_{a_n}(\cup_k A_k) = \sum_{n=1}^{N} \alpha_n \mathbf{1}_{\cup_k A_k}(a_n) = \sum_{n=1}^{N} \alpha_n \sum_{k=0}^{\infty} \mathbf{1}_{A_k}(a_n) \\
&= \sum_{k=0}^{\infty} \sum_{n=1}^{N} \alpha_n \mathbf{1}_{A_k}(a_n) = \sum_{k=0}^{\infty} \sum_{n=1}^{N} \alpha_n \delta_{a_n}(A_k) = \sum_{k=0}^{\infty} \mu(A_k).
\end{aligned}$$

Exemple 2.2.5. Voici quelques exemples de mesures discrètes à un nombre fini d'atomes :

— Mesure de comptage sur $\{1, 2, \ldots, N\}$:

$$\mu = \sum_{n=1}^{N} \delta_n \ ;$$

— Mesure binomiale de paramètres $n \in \mathbb{N}^*$ et $p \in [0,1]$:

$$\mu = \sum_{k=0}^{n} \binom{n}{k} p^k (-p)^{n-k} \delta_k.$$

2.3 Mesure de Lebesgue

Théorème 2.3.1. *Il existe une unique mesure* λ *sur* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *telle que*

(i) $\lambda([0,1]) = 1$,

(ii) Pour tout $a \in \mathbb{R}$ *et tout* $B \in \mathcal{B}(\mathbb{R})$, $\lambda(a+B) = \lambda(B)$.

Elle est appelée mesure de Lebesgue sur $\mathbb{R}$.

Remarque 2.3.2. Ce théorème est dificile. Nous verrons lors de la construction de mesures produit les arguments clé de l'unicité.

Cette mesure coïncide avec la notion intuitive de longueur comme le montre le résultat suivant.

Proposition 2.3.3. *Pour tous* $a < b$ *réels,*

$$\lambda([a,b]) = \lambda(]a,b]) = \lambda([a,b[) = \lambda(]a,b[) = b-a.$$

Si I *est un intervalle non borné alors* $\lambda(I) = +\infty$.

Démonstration. Soit $\alpha = \lambda(\{0\})$. Alors, d'après l'invariance par translation de λ, pour tout $x \in \mathbb{R}$,

$$\lambda(\{x\}) = \lambda(x + \{0\}) = \lambda(\{0\}) = \alpha.$$

Ainsi, pour tout $n \geq 1$,

$$n\alpha = \lambda(\{1/k,\ k = 1, \ldots, n\}) \leq \lambda([0,1]) = 1.$$

Ceci assure donc que $\alpha = 0$. On dit que λ ne charge aucun singleton. La mesure des intervalles ne dépend pas du fait qu'ils contiennent ou non leurs extrémités (on utilisera ce fait dans la suite sans le rappeler systématiquement).

Soit $n \geq 1$. Découpons $]0,1]$ en n intervalles disjoints égaux.

$$\begin{aligned} 1 &= \lambda([0,1]) = \lambda(]0,1]) = \lambda(\cup_{k=1}^n](k-1)/n, k/n]) \\ &= \sum_{k=1}^n \lambda((k-1)/n +]0,1/n]) = n\lambda(]0,1/n]). \end{aligned}$$

Ainsi, pour tout $n \geq 1$, $\lambda(]0,1/n] = 1/n$.

Soit à présent $n \geq 1$ et $k_1 \leq k_2 \in \mathbb{Z}$. Alors

$$\lambda(]k_1/n, k_2/n]) = \lambda(\cup_{l=1}^{k_2-k_1}](k_1+l-1)/n, (k_1+l)/n]) = \frac{k_2}{n} - \frac{k_1}{n}.$$

Ainsi, pour tous rationnels $r \leq r'$, $\lambda(]r, r'[) = r' - r$.

Soit $a < b \in \mathbb{R}$. il existe deux suites $(u_n)_n$ et $(v_n)_n$ de rationnels strictement décroissante pour la première et strictement croissante pour la seconde telles que $u_n \leq v_n$ pour tout n et qui convergent respectivement vers a et b. On obtient alors

$$\lambda(]a,b[) = \lambda(\cup_n]u_n, v_n[) = \lim_n \lambda(]u_n, v_n[) = \lim_n (v_n - u_n) = b - a.$$

Soit enfin I un intervalle non borné. Supposons-le de la forme $[a, +\infty[$. Alors, pour tout $n \in \mathbb{N}$, I contient $[a, a+n]$ et ainsi, $\lambda(I) \geq n$. Ceci assure que $\lambda(I) = +\infty$. □

Chapitre 3

Applications mesurables

3.1 Définitions et critères de mesurabilité

Définition 3.1.1. Soit $(E, \mathcal{A})$ et $(F, \mathcal{B})$ deux espaces mesurables et f une application de E dans F. On dit que f est mesurable de $(E, \mathcal{A})$ dans $(F, \mathcal{B})$ si l'image réciproque par f de tout élément de $\mathcal{B}$ est un élément de $\mathcal{A}$. On dira plus simplement que f est mesurable s'il n'y a pas d'ambiguïté sur les tribus considérées.

Autrement dit, f est mesurable si $f^{-1}(\mathcal{B}) \subset \mathcal{A}$. Lorsque E et F sont des espaces topologiques et $\mathcal{A}$ et $\mathcal{B}$ désignent leurs tribus boréliennes respectives, une application mesurable est encore appelée borélienne.

Exemple 3.1.2. Soit $(E, \mathcal{A})$ un espace mesurable. Pour toute partie A de E on note $\mathbf{1}_A$ la fonction indicatrice de l'ensemble A (valant 1 sur A et 0 sur son complémentaire). La fonction $\mathbf{1}_A$ est mesurable de $(E, \mathcal{A})$ dans $\mathbb{R}$ (muni de sa tribu borélienne) si et seulement si $A \in \mathcal{A}$.

Proposition 3.1.3. *Soit $(E, \mathcal{A})$ et $(F, \mathcal{B})$ deux espaces mesurables, f une application de E dans F et $\mathcal{E}$ une classe sur F telle que $\sigma(\mathcal{E}) = \mathcal{B}$. Alors f est mesurable si et seulement si l'image réciproque de tout élément de $\mathcal{E}$ appartient à $\mathcal{A}$.*

Démonstration. La condition est évidemment nécessaire. Réciproquement, si $\mathcal{A}$ contient l'image réciproque de $\mathcal{E}$, elle contient également la tribu engendrée par l'image réciproque de $\mathcal{E}$ qui est encore l'image réciproque de la tribu engendrée par $\mathcal{E}$, c'est-à-dire l'image réciproque de $\mathcal{B}$ par hypothèse. $\square$

Corollaire 3.1.4. *Soit E et F deux espaces topologiques munis de leurs tribus boréliennes respectives. Toute fonction continue f de E dans F est mesurable.*

Démonstration. Soit $\mathcal{O}_E$ (resp. $\mathcal{O}_F$) la classe des ouverts de E (resp. de F). Par définition de la continuité de f, on a $f^{-1}(\mathcal{O}_F) \subset \mathcal{O}_E \subset \mathcal{B}(E)$. La tribu borélienne $\mathcal{B}(E)$ de E contient donc $\sigma(f^{-1}(\mathcal{O}_F)) = f^{-1}(\sigma(\mathcal{O}_F)) = f^{-1}(\mathcal{B}(F))$ et f est mesurable. $\square$

Corollaire 3.1.5. *Soit f une application mesurable de $(E, \mathcal{A})$ à valeurs dans $\mathbb{R}$ muni de sa tribu borélienne. Alors f est mesurable si et seulement si l'une des conditions suivantes est vérifiée :*

(i) $\forall a \in \mathbb{R},\ \{x \in E,\ f(x) < a\} \in \mathcal{A}$,

(ii) $\forall a \in \mathbb{R},\ \{x \in E,\ f(x) \leq a\} \in \mathcal{A}$,

(iii) $\forall a \in \mathbb{R},\ \{x \in E,\ f(x) > a\} \in \mathcal{A}$,

(iv) $\forall a \in \mathbb{R},\ \{x \in E,\ f(x) \geq a\} \in \mathcal{A}$.

Démonstration. En effet, l'une quelconque des classes suivantes de parties de $\mathbb{R}$

$$\{]-\infty, a[\ ;\ a \in \mathbb{R}\}\ ;\quad \{]-\infty, a]\ ;\ a \in \mathbb{R}\}\ ;\quad \{]a, +\infty[\ ;\ a \in \mathbb{R}\}\ ;\quad \{[a, +\infty[\ ;\ a \in \mathbb{R}\}$$

engendre la tribu borélienne de $\mathbb{R}$. □

3.2 Propriétés de stabilité

La mesurabilité est stable par composition.

Proposition 3.2.1. *Soit $(E, \mathcal{A})$, $(F, \mathcal{B})$, $(G, \mathcal{C})$ trois espaces mesurables, f une application mesurable de $(E, \mathcal{A})$ dans $(F, \mathcal{B})$ et g une application mesurable de $(F, \mathcal{B})$ dans $(G, \mathcal{C})$. Alors l'application $g \circ f$ est mesurable de $(E, \mathcal{A})$ dans $(G, \mathcal{C})$.*

Proposition 3.2.2. *Soit $(F_1, \mathcal{B}_1)$ et $(F_2, \mathcal{B}_2)$ deux espaces mesurables et p_1 et p_2 les projections de $F_1 \times F_2$ sur F_1 et F_2 respectivement. On munit $F_1 \times F_2$ de la tribu produit $\mathcal{B}_1 \otimes \mathcal{B}_2$.*

(i) les projections p_1 et p_2 sont mesurables ;

(ii) soit $(E, \mathcal{A})$ un espace mesurable et f une application de E dans $F_1 \times F_2$. Alors f est mesurable si et seulement si les composées $p_1 \circ f : E \to F_1$ et $p_2 \circ f : E \to F_2$ sont mesurables.

Démonstration. Prouvons le point (i). Pour tout $B_1 \in \mathcal{B}_1$, on a $p_1^{-1}(B_1) = B_1 \times F_2 \in \mathcal{B}_1 \otimes \mathcal{B}_2$. Donc p_1 est mesurable. On procède de même pour p_2.

Prouvons le point (ii). Si f est mesurable, il est clair que $p_1 \circ f$ et $p_2 \circ f$ le sont. Réciproquement, supposons que $p_1 \circ f$ et $p_2 \circ f$ soient mesurables. Alors, pour tout $B_1 \in \mathcal{B}_1$, l'ensemble $f^{-1}(B_1 \times F_2)$ n'est autre que $(p_1 \circ f)^{-1}(B_1)$ qui appartient à $\mathcal{A}$. De même, pour tout $B_2 \in \mathcal{B}_2$, on a $f^{-1}(F_1 \times B_2)$ appartient à $\mathcal{A}$. Il en résulte que

$$f^{-1}(B_1 \times B_2) = f^{-1}((B_1 \times F_2) \cap (F_1 \times B_2)) = f^{-1}(B_1 \times F_2) \cap f^{-1}(F_1 \times B_2) \in \mathcal{A}.$$

Comme $\mathcal{B}_1 \otimes \mathcal{B}_2$ est la tribu engendrée par les parties de la forme $B_1 \times B_2$, avec $B_1 \in \mathcal{B}_1$ et $B_2 \in \mathcal{B}_2$, la proposition 3.1.3 permet de conclure que f est mesurable. □

Corollaire 3.2.3. *Pour qu'une fonction à valeurs complexes soit mesurable, il faut et il suffit que ses parties réelle et imaginaire soient mesurables. Si f et g sont des fonctions mesurables de $(E, \mathcal{A})$ dans $\mathbb{C}$, alors $f + g$, fg, $|f|$,... sont mesurables.*

Avant d'étudier la stabilité de la notion de mesurabilité par passage à la limite, rappelons quelques définitions concernant les suites à valeurs dans $\overline{\mathbb{R}}$.

Définition 3.2.4. Soit $(u_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $\overline{\mathbb{R}}$. La plus grande (resp. petite) valeur d'adhérence de la suite $(u_n)_n$ est notée $\limsup u_n$ (resp $\liminf u_n$). Leurs définitions sont données par

$$\limsup u_n = \inf_{n\geq 0} \sup_{k\geq n} u_k \quad \text{et} \quad \liminf u_n = \sup_{n\geq 0} \inf_{k\geq n} u_k.$$

Remarque 3.2.5. Les limites supérieure et inférieure sont a priori des éléments de $\overline{\mathbb{R}}$.

Remarque 3.2.6. Pour tout $x > \limsup u_n$, il existe $\varepsilon > 0$ et $n_0 \in \mathbb{N}$ tels que, pour tout $n \geq n_0$, $u_n \leq x - \varepsilon$.

Remarque 3.2.7. On a toujours $\liminf u_n \leq \limsup_n u_n$ et la suite $(u_n)_{n\in\mathbb{N}}$ converge vers un réel si et seulement si $\liminf u_n = \limsup_n u_n$ et ces quantités sont finies.

Remarque 3.2.8. Si $(f_n)_n$ est une suite d'applications de E dans $\overline{\mathbb{R}}$, on note $\limsup f_n$ la fonction qui à $x \in E$ associe $\limsup f_n(x) \in \overline{\mathbb{R}}$.

Proposition 3.2.9. *La mesurabilité est stable par passage à la limite.*

(i) Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables sur $(E, \mathcal{A})$ à valeurs dans $\overline{\mathbb{R}}$. Les fonctions $\sup f_n$, $\inf f_n$, $\limsup f_n$ et $\liminf f_n$ sont mesurables.

(ii) Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables sur $(E, \mathcal{A})$ à valeurs dans $\mathbb{C}$ telle que, pour tout $x \in E$, la limite $\lim_n f_n(x) = f(x)$ existe. Alors f est mesurable.

Démonstration. Établissons tout d'abord le point (i). Par hypothèse, pour tout $a \in \mathbb{R}$, l'ensemble $\{f_n \leq a\}$ appartient à $\mathcal{A}$. Or,

$$\{\sup f_n \leq a\} = \bigcap_n \{f_n \leq a\}.$$

D'après le corollaire 3.1.5, $\sup f_n$ est mesurable. Comme $\inf f_n = -\sup(-f_n)$, $\inf f_n$ est mesurable. Enfin, $\limsup f_n = \inf_n(\sup_{k\geq n} f_k)$ et $\liminf f_n = \sup_n(\inf_{k\geq n} f_k)$ sont mesurables d'après ce qui précède.

Pour prouver le point (ii), quitte à considérer les parties réelle et imaginaire des fonctions f_n, on peut supposer que f_n est réelle. Mais alors $f = \liminf f_n = \limsup f_n$ est mesurable d'après (i). □

Proposition 3.2.10. *Soit f et g deux applications mesurables de $(E, \mathcal{A})$ dans $\overline{\mathbb{R}}_+$ (muni de sa tribu borélienne). Alors $\{f < g\}$ et $\{f \leq g\}$ sont des éléments de $\mathcal{A}$.*

Démonstration. En décomposant ces ensembles de la façon suivante :

$$\begin{aligned} \{f < g\} &= \cup_{q\in\mathbb{Q}}\{f < q < g\} = \cup_{q\in\mathbb{Q}}(\{f < q\} \cap \{q < g\}), \\ \{f \leq g\} &= \cap_{n\geq 1}\{f < g + 1/n\}, \end{aligned}$$

on obtient leur appartenance à la tribu $\mathcal{A}$. □

Voici un dernier résultat (admis) important.

Définition 3.2.11. Une fonction f est continue par morceaux sur un segment $[a, b]$ s'il existe une subdivision $a < a_0 < \cdots < a_n < b$ telle que, pour tout $0 \leq i \leq n-1$, la restriction de f à $]a_i, a_{i+1}[$ admette un prolongement continu à $[a_i, a_{i+1}]$. Une fonction f est continue par morceaux sur un intervalle I elle est continue par morceaux sur tout segment inclus dans I.

Théorème 3.2.12. *Soit I un intervalle de $\mathbb{R}$ muni de sa tribu borélienne (qui est constituée des intersections des boréliens de $\mathbb{R}$ avec I). Si f est continue par morceaux sur I, alors elle est mesurable de $(I, \mathcal{B}(I))$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.*

3.3 Approximation des fonctions mesurables

L'objet de ce paragraphe est d'établir un résultat d'approximation relativement élémentaire mais fondamental pour la construction de l'intégrale de Lebesgue : toute fonction mesurable à valeurs dans $\overline{\mathbb{R}}_+$ est limite croissante de fonctions élémentaires, appelées fonctions étagées.

Définition 3.3.1. On notera $\mathcal{M}$ (resp. $\mathcal{M}_+$) l'ensemble des fonctions mesurables (resp. mesurables positives) sur $(E, \mathcal{A})$ à valeurs dans $\mathbb{R}$ (resp $\overline{\mathbb{R}}_+$).

Définition 3.3.2. Une fonction mesurable sur $(E, \mathcal{A})$ à valeurs dans $\mathbb{C}$ est dite étagée si elle prend seulement un nombre fini de valeurs distinctes. On notera $\mathcal{E}_+$ (resp. $\mathcal{E}$) l'ensemble des fonctions étagées sur $(E, \mathcal{A})$ à valeurs dans $\mathbb{R}_+$ (resp. $\mathbb{C}$).

Une fonction étagée ne peut prendre que des valeurs finies (dans $\mathbb{C}$) contrairement aux fonctions mesurables à valeurs dans $\overline{\mathbb{R}}$. Soit f une fonction étagée et n le nombre de valeurs distinctes prises par f. Notons $\alpha_1, \ldots, \alpha_n$ ces valeurs et posons, pour $i = 1 \ldots, n$, $A_i = \{f = \alpha_i\}$. Alors les parties $(A_i)_{i=1,\ldots,n}$ sont mesurables et f peut encore s'écrire

$$f = \sum_{i=1}^{n} \alpha_i \mathbf{1}_{A_i}.$$

Réciproquement, toute combinaison linéaire à coefficients réels ou complexes de fonctions caractéristiques d'ensembles mesurables est une fonction étagée. Remarquons de plus que les fonctions étagées sur $(E, \mathcal{A})$ forment un espace vectoriel.

Théorème 3.3.3. *Soit f une fonction mesurable sur $(E, \mathcal{A})$ à valeurs dans $\overline{\mathbb{R}}_+$. Il existe une suite croissante $(f_n)_{n\in\mathbb{N}}$ de fonctions étagées positives qui converge simplement vers f. De plus, la convergence est uniforme sur tout ensemble $B \in \mathcal{A}$ sur lequel f est bornée.*

Démonstration. Pour $n \in \mathbb{N}$ et $k \in \{0, 1, 2, \ldots, n2^n - 1\}$, posons

$$A_n = \{f \geq n\} \quad \text{et} \quad A_{n,k} = \left\{\frac{k}{2^n} \leq f < \frac{k+1}{2^n}\right\}.$$

On définit alors la fonction f_n par :

$$f_n = \sum_{k=0}^{n2^n-1} \frac{k}{2^n} \mathbf{1}_{A_{n,k}} + n \mathbf{1}_{A_n}.$$

Par définition, f_n est une fonction étagée positive telle que $f_n \leq f$. D'autre part, on vérifie que si $x \in A_{n,k}$,

$$f_{n+1}(x) = \begin{cases} f_n(x) & \text{si } \dfrac{2k}{2^{n+1}} \leq f(x) < \dfrac{2k+1}{2^{n+1}} \\ f_n(x) + \dfrac{1}{2^{n+1}} & \text{si } \dfrac{2k+1}{2^{n+1}} \leq f(x) < \dfrac{2(k+1)}{2^{n+1}}. \end{cases}$$

De même, si $x \in A_n$,

$$f_{n+1}(x) = \begin{cases} n+1 & \text{si } f(x) \geq n+1 \\ n + \dfrac{l}{2^{n+1}} & \text{si } n + \dfrac{l}{2^{n+1}} \leq f(x) < n + \dfrac{l+1}{2^{n+1}},\ 0 \leq l \leq 2^{n+1} - 1. \end{cases}$$

Ainsi, pour tout $n \in \mathbb{N}$ et tout $x \in E$, $f_n(x) \leq f_{n+1}(x)$: la suite (f_n) est croissante.

De plus, $(A_n)_n$ est une suite décroissante d'éléments de $\mathcal{A}$ donc si $x \in A_{n_0}^c$, alors pour tout $n \geq n_0$, $x \in A_n^c$ ou encore

$$0 \leq f(x) - f_n(x) \leq \frac{1}{2^n}.$$

Ceci implique que $(f_n(x))_n$ converge vers $f(x)$. Ainsi, la suite (f_n) converge sur l'ensemble $\cup_n A_n^c$ qui n'est autre que $\{f < +\infty\}$.

D'autre part, si $x \in \{f = +\infty\}$ alors, pour tout $n \in \mathbb{N}$, $f_n(x) = n$ qui tend vers $+\infty$ quand n tend vers $+\infty$.

Soit à présent $B \in \mathcal{A}$ tel que f soit bornée sur B. Il existe n_1 tel que, pour tout $x \in B$, $f(x) < n_1$. Alors $B \cap A_{n_1} = \emptyset$ et ainsi,

$$\forall n \geq n_1,\ \forall x \in B, \quad 0 \leq f(x) - f_n(x) \leq \frac{1}{2^n}.$$

La convergence est donc bien uniforme sur B. □

Corollaire 3.3.4. *Toute fonction f mesurable sur $(E, \mathcal{A})$ à valeurs dans $\overline{\mathbb{R}}$ (ou $\mathbb{C}$) est limite simple d'une suite (f_n) de fonctions étagées à valeurs dans $\mathbb{R}$ (ou $\mathbb{C}$).*

Démonstration. Si f est à valeurs dans $\overline{\mathbb{R}}$, on peut l'écrire $f = f^+ - f^-$ avec $f^+ = \sup(f, 0)$ et $f^- = -\inf(0, f)$. Comme f^+ et f^- sont mesurables à valeurs dans $\overline{\mathbb{R}}_+$, il existe des suites (g_n) et (h_n) de fonctions étagées positives tendant simplement vers f^+ et f^- respectivement. La suite (f_n), où $f_n = g_n - h_n$, est formée de fonctions étagées et converge simplement vers f. Si f est à valeurs complexes, on l'écrira comme combinaison de ses parties réelle et imaginaire. □

Chapitre 4

Construction de l'intégrale de Lebesgue

Dans ce chapitre, on se donne un espace mesuré $(E, \mathcal{A}, \mu)$. L'idée est de construire l'intégrale pour des fonctions de plus en plus générales grâce à des passages à la limite.

4.1 Intégration des fonctions étagées positives

Définition 4.1.1. Soit f une fonction étagée positive, prenant les valeurs distinctes $\alpha_1, \ldots, \alpha_n$. On pose $A_i = f^{-1}(\{\alpha_i\})$ pour $1 \leq i \leq n$. On appelle intégrale de f par rapport à μ, et on note $\int f\, d\mu$, le nombre fini ou infini (élément de $\overline{\mathbb{R}}_+$) défini par

$$\int f\, d\mu = \sum_{i=1}^{n} \alpha_i \mu(A_i),$$

avec la convention usuelle en théorie de la mesure : $0 \times \infty = 0$.

Proposition 4.1.2. *L'intégrale de fonctions étagées positives vérifie les propriétés suivantes.*

(i) Si f et g sont deux fonctions étagées positives et $\lambda \in \mathbb{R}_+^*$, alors

$$\int (\lambda f + g)\, d\mu = \lambda \int f\, d\mu + \int g\, d\mu.$$

(ii) Si f et g sont deux fonctions étagées positives telles que $f \leq g$, alors

$$\int f\, d\mu \leq \int g\, d\mu.$$

Démonstration. Montrons la propriété (i) dans le cas où $\lambda = 1$. Le cas général s'en déduit immédiatement. Posons

$$f = \sum_{i=1}^{n} \alpha_i \mathbf{1}_{A_i} \quad \text{et} \quad g = \sum_{i=1}^{m} \beta_j \mathbf{1}_{B_j}$$

où les $(\alpha_i)_i$ (resp. les $(\beta_j)_j$) sont distincts. Notons $\gamma_1, \ldots, \gamma_l$ les valeurs (distinctes) prises par $f+g$ et

$$C_k = (f+g)^{-1}(\{\gamma_k\}) = \bigcup_{(i,j)\in I_k} (A_i \cap B_j),$$

où $I_k = \{(i,j),\ \alpha_i + \beta_j = \gamma_k\}$. Puisque les ensembles $(A_i \cap B_j)_{i,j}$ sont deux à deux disjoints,

$$\mu(C_k) = \sum_{(i,j)\in I_k} \mu(A_i \cap B_j).$$

On a donc par définition de l'intégrale de $f+g$

$$\begin{aligned}
\int (f+g)\,d\mu = \sum_{k=1}^{l} \gamma_k \mu(C_k) &= \sum_{k=1}^{l} \sum_{(i,j)\in I_k} (\alpha_i + \beta_j)\mu(A_i \cap B_j) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} \alpha_i \mu(A_i \cap B_j) + \sum_{j=1}^{m}\sum_{i=1}^{n} \beta_j \mu(A_i \cap B_j) \\
&= \sum_{i=1}^{n} \alpha_i \mu(A_i) + \sum_{j=1}^{m} \beta_j \mu(B_j) \\
&= \int f\,d\mu + \int g\,d\mu.
\end{aligned}$$

Pour établir (ii), il suffit d'appliquer (i), en remarquant que $g-f$ est une fonction étagée positive, pour obtenir

$$\int f\,d\mu \le \int f\,d\mu + \int (g-f)\,d\mu = \int (f + (g-f))\,d\mu = \int g\,d\mu.$$

Ceci achève la preuve. □

Remarque 4.1.3. Soit la fonction $f = \sum_i \alpha_i \mathbf{1}_{A_i}$ où les (α_i) ne sont pas nécessairement distincts et les (A_i) nécessairement disjoints. On a encore $\int f\,d\mu = \sum_i \alpha_i \mu(A_i)$.

4.2 Intégration des fonctions mesurables positives

Définition 4.2.1. Soit f une fonction mesurable à valeurs dans $\overline{\mathbb{R}}_+$. On appelle intégrale de f par rapport à μ, et on note $\int f\,d\mu$ l'élément de $\overline{\mathbb{R}}_+$ défini par

$$\int f\,d\mu = \sup\left\{\int u\,d\mu,\ u \in \mathcal{E}_+ \text{ telle que } u \le f\right\}.$$

Remarque 4.2.2. Si f est une fonction étagée positive alors les deux définitions de son intégrale coïncident car le supremum est atteint pour $u = f$.

Proposition 4.2.3 (Croissance de l'intégrale)**.** *Pour toutes fonctions f et g mesurables positives telles que $f \le g$,*

$$\int f\, d\mu \le \int g\, d\mu.$$

Démonstration. C'est une conséquence immédiate de l'inclusion

$$\{u \in \mathcal{E}_+,\ u \le f\} \subset \{u \in \mathcal{E}_+,\ u \le g\}$$

et de la définition de l'intégrale. □

Voici le premier des grands théorèmes d'intervertion limite-intégrale qui font toute la puissance de la théorie de la mesure.

Théorème 4.2.4 (de convergence monotone ou de Beppo Levi)**.** *Soit $(f_n)_{n\in\mathbb{N}}$ une suite croissante de $\mathcal{M}_+$. Alors $f = \lim_n f_n (= \sup_n f_n)$ est aussi dans $\mathcal{M}_+$ et*

$$\int f\, d\mu = \lim_{n\to+\infty} \int f_n\, d\mu.$$

Démonstration. On sait déjà que le supremum d'éléments de $\mathcal{M}_+$ est encore dans $\mathcal{M}_+$ d'après la proposition 3.2.9. Comme $f_n \le f$, on a $\int f_n\, d\mu \le \int f\, d\mu$. La croissance de l'intégrale assure que la suite $\left(\int f_n\, d\mu\right)_n$ est elle aussi croissante et donc convergente dans $\overline{\mathbb{R}}_+$. On obtient donc

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

Démontrons l'inégalité opposée. Soit u une fonction positive étagée inférieure à f et $\lambda \in]0,1[$. Posons

$$E_n = \{x \in E,\ f_n(x) \ge \lambda u(x)\}.$$

La suite $(E_n)_{n\in\mathbb{N}}$ de sous-ensembles de E est croissante au sens de l'inclusion. Soit $x \in E$. Si $u(x) = 0$ alors $x \in E_n$ pour tout $n \in \mathbb{N}$. Si $u(x) > 0$ alors

$$\lim_n f_n(x) = f(x) \ge u(x) > \lambda u(x),$$

et ainsi $x \in E_n$ pour n assez grand et donc $\cup_n E_n = E$. D'autre part, par définition de E_n, $f_n \ge \lambda u \mathbf{1}_{E_n}$ et donc, pour tout $n \in \mathbb{N}$, par croissance de l'intégrale,

$$\int f_n\, d\mu \ge \int \lambda u \mathbf{1}_{E_n}\, d\mu,$$

La fonction $\lambda u \mathbf{1}_{E_n}$ est étagée positive. On sait donc calculer son intégrale. Si $u = \sum_{i=1}^k \alpha_i \mathbf{1}_{A_i}$ alors

$$\int u\, d\mu = \sum_{i=1}^k \alpha_i \mu(A_i) \quad \text{et} \quad \int u \mathbf{1}_{E_n}\, d\mu = \sum_{i=1}^k \alpha_i \mu(A_i \cap E_n).$$

Or, pour tout $i = 1, \ldots, k$, $\mu(A_i \cap E_n)$ converge en croissant vers $\mu(A_i)$, donc, $\int u \mathbf{1}_{E_n}\, d\mu$ converge vers $\int u\, d\mu$. On a donc établi que, pour tout $u \in \mathcal{E}_+$ tel que $u \leq f$ et tout $\lambda \in]0, 1[$,

$$\lim_n \int f_n\, d\mu \geq \lim_n \lambda \int u \mathbf{1}_{E_n}\, d\mu = \lambda \int u\, d\mu.$$

On obtient donc, en faisant tendre λ vers 1, que, l'intégrale de toute fonction étagée positive u majorée par f est inférieure à la limite des intégrales des fonctions f_n. Il en est donc de même pour l'intégrale de f :

$$\int f\, d\mu = \sup \left\{ \int u\, d\mu,\ u \in \mathcal{E}_+,\ \text{telle que } u \leq f \right\} \leq \lim_n \int f_n\, d\mu,$$

et l'égalité est donc établie. □

Corollaire 4.2.5. *Si f et g sont deux fonctions mesurables positives, alors*

$$\int (f + g)\, d\mu = \int f\, d\mu + \int g\, d\mu.$$

Démonstration. D'après le théorème 3.3.3, il existe des suites $(f_n)_{n\in\mathbb{N}}$ et $(g_n)_{n\in\mathbb{N}}$ croissantes de fonctions étagées positives qui convergent simplement vers f et g respectivement. Alors la suite $(f_n + g_n)_{n\in\mathbb{N}}$ est une suite croissante de fonctions étagées positives qui converge simplement vers $f + g$. La linéarité de l'intégrale de fonctions étagées assure alors, pour tout n,

$$\int (f_n + g_n)\, d\mu = \int f_n\, d\mu + \int g_n\, d\mu.$$

Le théorème de Beppo Levi permet de conclure en passant à la limite. □

Corollaire 4.2.6 (Intervertion du signe somme et du signe intégrale)**.** *Si $(f_n)_{n\in\mathbb{N}}$ est une suite de fonctions mesurables positives, on a*

$$\int \left(\sum_{n=0}^{\infty} f_n \right) d\mu = \sum_{n=0}^{\infty} \int f_n\, d\mu,$$

l'égalité ayant lieu dans $\overline{\mathbb{R}}_+$.

Démonstration. Posons $g_n = \sum_{k=0}^{n} f_k$. La suite $(g_n)_{n\in\mathbb{N}}$ est une suite croissante de fonctions mesurables donc on peut écrire

$$\begin{aligned} \int \left(\sum_{n=0}^{\infty} f_n \right) d\mu &= \int \left(\lim_{n\to\infty} g_n \right) d\mu = \lim_{n\to\infty} \int g_n\, d\mu \\ &= \lim_{n\to\infty} \sum_{k=0}^{n} \left(\int f_k\, d\mu \right) = \sum_{k=0}^{\infty} \left(\int f_k\, d\mu \right), \end{aligned}$$

grâce au théorème de convergence monotone. □

Exemple 4.2.7. Posons, pour $n \geq 3$,

$$g_n(x) = \frac{x}{(1+x)^n} \mathbf{1}_{[1,+\infty[}(x).$$

Les applications (g_n) sont continues par morceaux donc mesurables. Pour tout $x \in \mathbb{R}$, la suite $(g_n(x))_{n\geq 3}$ est croissante : c'est évident si $x < 1$ et si $x \geq 1$, $(1+x)^n \leq (1+x)^{n+1}$. D'après le théorème de convergence monotone,

$$\sum_{n=3}^{\infty} \int_{\mathbb{R}} g_n(x)\lambda(dx) = \int_{[1,+\infty[} \sum_{n=3}^{\infty} \frac{x}{(1+x)^n} \lambda(dx).$$

Par ailleurs, puisque $1 + x > 1$,

$$\sum_{n=3}^{\infty} \frac{x}{(1+x)^n} = \frac{x}{(1+x)^3} \sum_{n=0}^{\infty} \frac{1}{(1+x)^n} = \frac{x}{(1+x)^3} \frac{1}{1 - \frac{1}{1+x}} = \frac{1}{(1+x)^2}.$$

Enfin, pour $A > 1$, on a

$$\int_1^A \frac{1}{(1+x)^2} dx = \left[\frac{-1}{1+x}\right]_1^A = \frac{1}{2} - \frac{1}{1+A}.$$

On a donc

$$\int_1^{+\infty} \frac{1}{(1+x)^2} dx = \lim_{A\to\infty} \int_1^A \frac{1}{(1+x)^2} dx = \frac{1}{2}.$$

En conclusion,

$$\sum_{n=3}^{\infty} \int_1^{\infty} \frac{x}{(x+1)^n} dx = \frac{1}{2}.$$

4.3 Intégration de fonctions mesurables

Définition 4.3.1. Une fonction f définie sur E à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$ est dite intégrable (par rapport à μ) si elle est mesurable et si $\int |f| \, d\mu < +\infty$.

Nous noterons $\mathcal{L}^1_{\mathbb{R}}(\mu)$ (resp $\mathcal{L}^1_{\mathbb{C}}(\mu)$) l'ensemble des fonctions intégrables à valeurs réelles (resp. complexes). Pour être plus précis, nous utiliserons (en cas d'éventuelles confusions) les notations $\mathcal{L}^1_{\mathbb{R}}(E, \mathcal{A}, \mu)$ et $\mathcal{L}^1_{\mathbb{C}}(E, \mathcal{A}, \mu)$.

Remarque 4.3.2. Si f et g sont mesurables telles que $|f| \leq |g|$ et g intégrable alors f est intégrable.

Proposition 4.3.3. *Soit f une fonction mesurable à valeurs dans $\mathbb{R}$. Alors f est intégrable si et seulement si f^+ et f^- le sont.*

Démonstration. Rappelons que $f^+ = \sup(f, 0)$ et $f^- = -\inf(f, 0) = \sup(-f, 0)$. On a alors

$$|f| = f^+ + f^-, \quad f^- \leq |f|, \quad f^+ \leq |f|.$$

La proposition découle de ces relations. □

Définition 4.3.4. Soit $f \in \mathcal{L}^1_{\mathbb{R}}(\mu)$. On appelle intégrale de f, et on note $\int f\, d\mu$, le nombre réel

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

Remarque 4.3.5. On pourra noter encore

$$\int f\, d\mu = \int f(x)\, \mu(dx) = \int f(x)\, d\mu(x).$$

Il s'agit de notations dont aucune n'est ni plus ni moins arbitraire qu'une autre.

Proposition 4.3.6. *L'ensemble $\mathcal{L}^1_{\mathbb{R}}(\mu)$ est un espace vectoriel sur $\mathbb{R}$ et l'application qui à f associe $\int f\, d\mu$ est une forme linéaire sur cet espace. De plus, on a*

(i) l'intégrale conserve la positivité (si $f \in \mathcal{L}^1_{\mathbb{R}}(\mu)$ et $f \geq 0$, alors $\int f\, d\mu \geq 0$),

(ii) l'intégrale conserve les inégalités (si $f, g \in \mathcal{L}^1_{\mathbb{R}}(\mu)$ et $f \leq g$, alors $\int f\, d\mu \leq \int g\, d\mu$),

(iii) si $f \in \mathcal{L}^1_{\mathbb{R}}(\mu)$, $\left|\int f\, d\mu\right| \leq \int |f|\, d\mu$.

Démonstration. On sait déjà que l'ensemble des fonctions réelles mesurables est un espace vectoriel sur $\mathbb{R}$. De plus, si $f, g \in \mathcal{L}^1_{\mathbb{R}}(\mu)$ et $\lambda \in \mathbb{R}$, alors $|\lambda f + g| \leq |\lambda||f| + |g|$. On en déduit que

$$\int |\lambda f + g|\, d\mu \leq |\lambda| \int |f|\, d\mu + \int |g|\, d\mu < +\infty.$$

L'ensemble $\mathcal{L}^1_{\mathbb{R}}(\mu)$ est donc un espace vectoriel sur $\mathbb{R}$.

Soient $f, g \in \mathcal{L}^1_{\mathbb{R}}(\mu)$. On a

$$\begin{cases} f + g = (f+g)^+ - (f+g)^- \\ f + g = f^+ - f^- + g^+ - g^-, \end{cases}$$

d'où $(f+g)^+ + f^- + g^- = (f+g)^- + f^+ + g^+$. On intègre cette égalité par rapport à μ en remarquant que tous les termes sont des fonctions mesurables positives. Il vient donc

$$\int (f+g)^+\, d\mu + \int f^-\, d\mu + \int g^-\, d\mu = \int (f+g)^-\, d\mu + \int f^+\, d\mu + \int g^+\, d\mu.$$

Toutes ces quantités sont finies donc on obtient

$$\int (f+g)^+\, d\mu - \int (f+g)^-\, d\mu = \int f^+\, d\mu - \int f^-\, d\mu + \int g^+\, d\mu - \int g^-\, d\mu,$$

ce qui établit la linéarité de l'intégrale. On montre de même que

$$\int (\lambda f)\, d\mu = \lambda \int f\, d\mu.$$

Pour prouver (i), on remarque que, si $f \in \mathcal{L}^1_{\mathbb{R}}(\mu)$ est positive, alors son intégrale est celle qui a été définie dans la définition 4.2.1. Elle appartient à $\mathbb{R}_+$. Le point (ii) se déduit du point (i) en considérant la fonction positive et intégrable $g - f$. Pour montrer (iii) on écrit tout simplement,

$$\left|\int f\, d\mu\right| = \left|\int f^+\, d\mu - \int f^-\, d\mu\right| \leq \int f^+\, d\mu + \int f^-\, d\mu = \int |f|\, d\mu,$$

ce qui assure la commutation annoncée de la valeur absolue et de l'intégrale. □

Définition 4.3.7. Soit $f \in \mathcal{L}^1_{\mathbb{C}}(\mu)$. On appelle intégrale de f, et on note $\int f\,d\mu$, le nombre complexe

$$\int f\,d\mu = \int \mathrm{Re}(f)\,d\mu + i\int \mathrm{Im}(f)\,d\mu.$$

Proposition 4.3.8. *Soit f une fonction mesurable à valeurs dans $\mathbb{C}$. Alors f est intégrable si et seulement si $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ le sont.*

Proposition 4.3.9. *L'ensemble $\mathcal{L}^1_{\mathbb{C}}(\mu)$ est un espace vectoriel sur $\mathbb{C}$ et l'application qui à f associe $\int f\,d\mu$ est une forme linéaire sur cet espace. De plus,*

$$\left|\int f\,d\mu\right| \leq \int |f|\,d\mu.$$

Démonstration. Soit $\alpha \in \mathbb{C}$ tel que $\left|\int f\,d\mu\right| = \alpha \int f\,d\mu$. On peut toujours choisir α de module 1 et

$$\begin{aligned}\left|\int f\,d\mu\right| &= \int \alpha f\,d\mu = \int \mathrm{Re}\ (\alpha f)\,d\mu + i\int \mathrm{Im}\ (\alpha f)\,d\mu \\ &\leq \int |\mathrm{Re}\ (\alpha f)|\,d\mu + \int |\mathrm{Im}\ (\alpha f)|\,d\mu \leq \int |\alpha f|\,d\mu = \int |f|\,d\mu.\end{aligned}$$

□

4.4 Mesures discrètes

Soit $(E,\mathcal{A})$ un espace mesurable, $(a_k)_{k\in\mathbb{N}^*}$ une suite de points de E telle que, pour tout $k \in \mathbb{N}^*$, $\{\alpha_k\} \in \mathcal{A}$ et $(\alpha_k)_{k\in\mathbb{N}^*}$ une suite de réels positifs. On définit une mesure μ sur $(E,\mathcal{A})$

$$\mu = \sum_{k=1}^{\infty} \alpha_k \delta_{a_k}.$$

Remarque 4.4.1. Cette application est bien une mesure (voir TD).

On souhaite étudier l'ensemble $\mathcal{L}^1(\mu)$ et comprendre l'objet $\int f\,d\mu$ pour $f \in \mathcal{L}^1(\mu)$.

Proposition 4.4.2. *Avec les notations du début du paragraphe.*

(i) Soit f mesurable de $(E,\mathcal{A})$ dans $\overline{\mathbb{R}}_+$. Alors, dans $\overline{\mathbb{R}}_+$, $\displaystyle\int f\,d\mu = \sum_{k=1}^{\infty} \alpha_k f(a_k)$.

(ii) Une fonction f mesurable de $(E,\mathcal{A})$ dans $\mathbb{C}$ est μ-intégrable ssi $\displaystyle\sum_{k=1}^{\infty} \alpha_k |f(a_k)| < +\infty$.

Dans ce cas, $\displaystyle\int f\,d\mu = \sum_{k=1}^{\infty} \alpha_k f(a_k)$.

Démonstration. Démontrons le point (i). On procède en trois étapes. Supposons que $f = \mathbf{1}_A$ avec $A \in \mathcal{A}$. Alors

$$\int f\, d\mu = \mu(A) = \sum_{k=1}^{\infty} \alpha_k \mathbf{1}_A(a_k) = \sum_{k=1}^{\infty} \alpha_k f(a_k).$$

Supposons à présent f étagée positive, alors $f = \sum_{i=1}^n \beta_i \mathbf{1}_{A_i}$. Par linéarité de l'intégrale,

$$\int f\, d\mu = \sum_{i=1}^{n} \beta_i \mu(A_i) = \sum_{i=1}^{n} \beta_i \sum_{k=1}^{\infty} \alpha_k \mathbf{1}_{A_i}(a_k) = \sum_{k=1}^{\infty} \alpha_k \sum_{i=1}^{n} \beta_i \mathbf{1}_{A_i}(a_k) = \sum_{k=1}^{\infty} \alpha_k f(a_k).$$

Enfin, si f est mesurable positive, il existe une suite croissante de fonctions étagées positives $(f_n)_{n\in\mathbb{N}}$ qui converge simplement vers f. Par le théorème de convergence monotone,

$$\int f\, d\mu = \lim_n \int f_n\, d\mu = \lim_n \sum_{k=1}^{\infty} \alpha_k f_n(a_k) = \sum_{k=1}^{\infty} \alpha_k \lim_n f_n(a_k) = \sum_{k=1}^{\infty} \alpha_k f(a_k),$$

ce qui achève la preuve du point (i). L'avant-dernière égalité ci-dessous provient du théorème de convergence monotone pour la mesure sur $(E, \mathcal{A})$ donnée par

$$A \mapsto \sum_{k\geq 1} \alpha_k)\delta_{a_k}$$

Démontrons à présent le point (ii). Soit f mesurable à valeurs dans $\mathbb{C}$. Appliquons le point (i) à $|f|$: f est μ-intégrable ssi $\int |f|\, d\mu$ est fini c'est-à-dire ssi $\sum_k \alpha_k |f(a_k)|$ est fini. Si tel est le cas, on écrit

$$f = (\mathrm{Re}\ f)^+ - (\mathrm{Re}\ f)^- + i(\mathrm{Im}\ f)^+ - i(\mathrm{Im} f)^-.$$

Les quatre fonctions mesurables positives $(\mathrm{Re} f)^+, \ldots, (\mathrm{Im}\ f)^-$ sont intégrables par rapport à μ (puisqu'elles sont toutes majorées par $|f|$). D'après (i) et la linéarité de l'intégrale, on obtient la relation souhaitée. □

Exemple 4.4.3. Soit μ la mesure Bernoulli de paramètre $p \in]0,1[$: $\mu = p\delta_1 + (1-p)\delta_0$. Alors,

$$\int x\, \mu(dx) = 0 \times (1-p) + 1 \times p \quad \text{et} \quad \int \cos(\pi x/4)\, \mu(dx) = 1 - p + p\frac{\sqrt{2}}{2}.$$

Exemple 4.4.4. Soit $\mu = \sum_{k=1}^{\infty} p(1-p)^{k-1}\delta_k$. Alors $f \in \mathcal{L}^1$ si et seulement si

$$\sum_{k=1}^{\infty} |f(k)| p(1-p)^{k-1} < +\infty$$

et dans ce cas,

$$\int f\, d\mu = \sum_{k=1}^{\infty} f(k) p(1-p)^{k-1}.$$

4.5 Mesures à densité

Étant donné un espace mesuré $(E, \mathcal{A}, \mu)$, on peut construire de nombreuses mesures à partir de μ comme le montre la proposition suivante.

Proposition 4.5.1. *Soit $(E, \mathcal{A}, \mu)$ un espace mesuré et g une fonction mesurable positive sur $(E, \mathcal{A})$. Soit ν l'application de $\mathcal{A}$ dans $\overline{\mathbb{R}}_+$ définie par*

$$\nu(A) = \int \mathbf{1}_A g \, d\mu = \int_A g \, d\mu.$$

Alors ν est une mesure sur $(E, \mathcal{A})$.

Démonstration. On a évidemment $\nu(\emptyset) = 0$. Soit (A_n) une suite d'éléments de $\mathcal{A}$ deux à deux disjoints. Posons $A = \cup_n A_n$. On a

$$\nu(A) = \int \mathbf{1}_A g \, d\mu = \int \sum_n \mathbf{1}_{A_n} g \, d\mu = \sum_n \int \mathbf{1}_{A_n} g \, d\mu = \sum_n \nu(A_n),$$

grâce au théorème de convergence monotone. □

Définition 4.5.2. La mesure ν est appelée mesure de densité g par rapport à μ. On la note souvent $g.\mu$. La fonction g est appelée la densité de ν par rapport à μ.

Proposition 4.5.3 (Intégration par rapport à une mesure à densité)**.** *Avec les notations de la proposition 4.5.1.*

(i) Soit f une fonction mesurable positive sur $(E, \mathcal{A})$. Alors, dans $\overline{\mathbb{R}}_+$,

$$\int f \, d\nu = \int (fg) \, d\mu. \tag{4.1}$$

(ii) Soit f une fonction mesurable à valeurs complexes sur $(E, \mathcal{A})$. Alors f est intégrable pour ν si et seulement si fg est intégrable pour μ et on a alors

$$\int f \, d\nu = \int (fg) \, d\mu.$$

Démonstration. Pour démontrer le point (i), on procède en trois étapes. Si $f = \mathbf{1}_A$ avec $A \in \mathcal{A}$, la relation (4.1) découle de la définition de ν. Si f est étagée et positive, l'égalité se déduit de la linéarité de l'intégrale. Supposons enfin que f soit simplement mesurable et positive. Soit $(f_n)_{n\in\mathbb{N}}$ une suite croissante de fonctions étagées et positives qui converge simplement vers f. Par le théorème de convergence monotone,

$$\int f \, d\nu = \lim_n \int f_n \, d\nu = \lim_n \int (f_n g) \, d\mu = \int (fg) \, d\mu.$$

Démontrons à présent le point (ii). Soit f mesurable à valeurs dans $\mathbb{C}$. Appliquons le point (i) à $|f|$: f est ν-intégrable ssi $\int |f| g \, d\nu$ est fini c'est-à-dire ssi fg est μ-intégrable. Si tel est le cas, on écrit

$$f = (\text{Re } f)^+ - (\text{Re } f)^- + i(\text{Im } f)^+ - i(\text{Im } f)^-.$$

Les quatre fonctions mesurables positives $(\text{Re } f)^+, \ldots, (\text{Im } f)^-$ sont intégrables par rapport à ν (puisqu'elles sont toutes majorées par $|f|$). D'après (i) et la linéarité de l'intégrale, on obtient la relation souhaitée. □

4.6 Intégration par rapport à une mesure image

Proposition 4.6.1 (Définition d'une mesure image)**.** *Soit $(E, \mathcal{A})$ et $(F, \mathcal{B})$ deux espaces mesurables et φ une application mesurable de E dans F. Soit μ une mesure sur $(E, \mathcal{A})$. L'application ν qui à $B \in \mathcal{B}$ associe*

$$\nu(B) = \mu(\varphi^{-1}(B))$$

définit une mesure sur $(F, \mathcal{B})$ appelée mesure image *de μ par φ. On la notera μ_φ.*

Démonstration. Comme $\varphi^{-1}(\emptyset) = \emptyset$, on a $\mu(\varphi^{-1}(\emptyset)) = 0$.

Soit $(B_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathcal{B}$ deux à deux disjoints. La suite $(\varphi^{-1}(B_n))$ est une suite d'éléments de $\mathcal{A}$ deux à deux disjoints et $\varphi^{-1}(\cup B_n) = \cup\varphi^{-1}(B_n)$ donc

$$\nu(\cup_n B_n) = \mu(\varphi^{-1}(\cup_n B_n)) = \mu(\cup_n \varphi^{-1}(B_n)) = \sum_n \mu(\varphi^{-1}(B_n)) = \sum_n \nu(B_n),$$

ce qui achève la preuve. □

Proposition 4.6.2. *Avec les notations de la proposition 4.6.1.*

(i) Soit f une fonction mesurable positive définie sur $(F, \mathcal{B})$. Alors (l'égalité a lieu dans $\overline{\mathbb{R}}_+$)

$$\int_F f\, d\mu_\varphi = \int_E f \circ \varphi\, d\mu. \tag{4.2}$$

(ii) Soit f une fonction mesurable à valeurs complexes définie sur $(F, \mathcal{B})$. Alors f est intégrable par rapport à μ_φ si et seulement si $f \circ \varphi$ est intégrable par rapport à μ. Dans ce cas,

$$\int_F f\, d\mu_\varphi = \int_E f \circ \varphi\, d\mu.$$

Démonstration. Démontrons le point (i) en trois étapes. Si f est la fonction indicatrice de $B \in \mathcal{B}$, l'égalité $\mu_\varphi(B) = \mu(\varphi^{-1}(B))$ qui définit la mesure image s'écrit encore

$$\int_Y \mathbf{1}_B\, d\mu_\varphi = \int_X \mathbf{1}_{\varphi^{-1}(B)}\, d\mu = \int_X \mathbf{1}_B \circ \varphi\, d\mu.$$

Si f est étagée positive, la relation (4.2) se déduit du cas précédent par linéarité. Enfin, si f est mesurable positive, d'après le théorème d'approximation 3.3.3, il existe une suite $(f_n)_{n\in\mathbb{N}}$ croissante de fonctions étagées positives qui converge simplement vers f. Alors $(f_n \circ \varphi)_{n\in\mathbb{N}}$ est une suite croissante de fonctions étagées positives qui converge simplement vers $f \circ \varphi$. D'après ce qui précède, on a, pour tout $n \in \mathbb{N}$,

$$\int_Y f_n\, d\mu_\varphi = \int_X f_n \circ \varphi\, d\mu,$$

et l'égalité souhaitée est conséquence du théorème de convergence monotone.

Démontrons à présent le point (ii). Soit f mesurable à valeurs dans $\mathbb{C}$. Le point (i) appliqué à $|f|$ montre que f est intégrable par rapport à μ_φ si et seulement si $f \circ \varphi$ l'est par rapport à μ. Supposons donc f intégrable et écrivons alors

$$f = (\mathrm{Re}\ f)^+ - (\mathrm{Re}\ f)^- + i(\mathrm{Im}\ f)^+ - i(\mathrm{Im}\ f)^-.$$

Les quatre fonctions mesurables positives $(\mathrm{Re}\ f)^+, \ldots, (\mathrm{Im}\ f)^-$ sont intégrables par rapport à μ_φ (puisqu'elles sont toutes majorées par $|f|$). D'après (i) et la linéarité de l'intégrale, on obtient la relation souhaitée. $\square$

4.7 Intégrale de Lebesgue et intégrale de Riemann

Rappelons brièvement les principes fondamentaux de l'intégrale de Riemann.

4.7.1 Intégrale sur un intervalle compact

Soit f une fonction réelle bornée sur $[a, b]$. Soit $\sigma\ :\ a = x_0 < x_1 < \cdots < x_{n+1} = b$ une subdivision de $[a, b]$. On appelle pas de la subdivision le nombre $\delta(\sigma) = \max_{1 \leq k \leq n+1}(x_k - x_{k-1})$. Posons

$$m_k = \inf\{f(t),\ t \in [x_k, x_{k+1}]\} \quad \text{et} \quad M_k = \sup\{f(t),\ t \in [x_k, x_{k+1}]\}.$$

Les sommes de Darboux associées à la subdivision σ sont

$$s(\sigma) = \sum_{k=1}^{n} m_k(x_{k+1} - x_k) \quad \text{et} \quad S(\sigma) = \sum_{k=1}^{n} M_k(x_{k+1} - x_k).$$

Définition 4.7.1. On dit que f est intégrable au sens de Riemann sur $[a, b]$ s'il existe un nombre réel I tel que les sommes $s(\sigma)$ et $S(\sigma)$ tendent vers I quand $\delta(\sigma)$ tend vers 0 :

$$\forall \varepsilon > 0,\ \exists \eta > 0, \quad \forall \sigma \text{ t.q. } \delta(\sigma) \leq \eta, \quad |S(\sigma) - I| \leq \varepsilon \quad \text{et} \quad |s(\sigma) - I| \leq \varepsilon.$$

Le nombre I est alors appelé l'intégrale de Riemann de f sur $[a, b]$ et on le note $\int_a^b f(t)\, dt$.

Considérons à nouveau la subdivision σ et, pour tout k, choisissons $\xi_k \in [x_{k-1}, x_k]$. La somme de Riemann définie par σ et $\xi = (\xi_1, \ldots, \xi_n)$ est par définition

$$S(\sigma, \xi) = \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}).$$

Si f est intégrable au sens de Riemann, les sommes de Riemann convergent vers $\int_a^b f(t)\, dt$ lorsque $\delta(\sigma)$ tend vers 0, uniformément par rapport au choix de ξ. Plus précisément,

$$\forall \varepsilon > 0,\ \exists \eta > 0, \quad \forall \sigma \text{ t.q. } \delta(\sigma) \leq \eta, \quad \forall \xi \text{ associé à } \sigma, \quad |S(\sigma, \xi) - I| \leq \varepsilon.$$

Théorème 4.7.2. *Toute fonction f continue par morceaux sur $[a, b]$ est intégrable au sens de Riemann. De plus, si f est continue, la fonction $x \mapsto F(x) = \int_a^x f(t)\, dt$ est dérivable sur $[a, b]$ de dérivée $F' = f$.*

4.7.2 Intégrale généralisée

Soit $f : [a, b[\to \mathbb{R}$, où b peut être égal à $+\infty$, localement intégrable au sens de Riemann ; c'est-à-dire intégrable au sens de Riemann sur tout intervalle compact $[a, c] \subset [a, b[$.

On dit que f admet une intégrale généralisée sur $[a, b[$ si la fonction $x \mapsto \int_a^x f(t)\,dt$ admet une limite lorsque x tend vers b (avec $x < b$). On pose alors

$$\int_a^b f(t)\,dt = \lim_{x \to b_-} \int_a^x f(t)\,dt.$$

Dans ce cas, on dit encore que l'intégrale $\int_a^b f(t)\,dt$ est convergente.

On dit que l'intégrale généralisée $\int_a^b f(t)\,dt$ est absolument convergente si l'intégrale $\int_a^b |f(t)|\,dt$ est convergente.

Remarque 4.7.3. La convergence absolue entraîne la convergence, mais la réciproque est fausse comme le montre l'exemple classique

$$\int_1^\infty \frac{\sin t}{t}\,dt.$$

4.7.3 Comparaison des intégrales de Riemann et Lebesgue pour une fonction bornée sur un intervalle compact

Proposition 4.7.4. *Soit f une fonction continue sur $[a, b]$. Alors si λ désigne la mesure de Lebesgue sur $\mathbb{R}$, $f\mathbf{1}_{[a,b]} \in \mathcal{L}^1_{\mathbb{R}}(\lambda)$ et*

$$\int_{\mathbb{R}} \mathbf{1}_{[a,b]} f\,d\lambda = \int_a^b f(t)\,dt.$$

Démonstration. Il est clair que $f\mathbf{1}_{[a,b]}$ est borélienne. Soit $M = \sup_{t\in[a,b]} |f(t)|$. La fonction f étant continue sur le compact $[a, b]$, M est un réel positif et

$$\left|f\mathbf{1}_{[a,b]}\right| \leq M\mathbf{1}_{[a,b]} \in \mathcal{L}^1_{\mathbb{R}}(\lambda),$$

ce qui assure que $f\mathbf{1}_{[a,b]}$ est Lebesgue-intégrable. De même, pour tout $x \in [a, b]$, $f\mathbf{1}_{[a,x]}$ est Lebesgue-intégrable. Posons $F(x) = \int f\mathbf{1}_{[a,x]}\,d\lambda$ et montrons que F est dérivable en tout point x_0 de $[a, b]$ de dérivée $f(x_0)$. Soit $h > 0$. Comme

$$\mathbf{1}_{[a,x_0+h]} f = \mathbf{1}_{[a,x_0]} f + \mathbf{1}_{]x_0,x_0+h]} f,$$

on a

$$\frac{F(x_0 + h) - F(x_0)}{h} = \frac{1}{h} \int \mathbf{1}_{]x_0,x_0+h]} f\,d\lambda,$$

d'où

$$\frac{F(x_0 + h) - F(x_0)}{h} - f(x_0) = \frac{1}{h} \int \mathbf{1}_{]x_0,x_0+h]} (f - f(x_0))\,d\lambda.$$

Soit $\varepsilon > 0$. Puisque f est continue en x_0, il existe $\eta > 0$ tel que pour tout x tel que $|x - x_0| \leq \eta$, on ait $|f(x_0) - f(x)| \leq \varepsilon$. Si $0 < h < \eta$ alors

$$\left| \frac{F(x_0 + h) - F(x_0)}{h} - f(x_0) \right| \leq \frac{1}{h} \int \varepsilon \mathbf{1}_{]x_0, x_0+h]} \, d\lambda = \varepsilon.$$

Le cas $h < 0$ se traite de même. Ainsi, F est dérivable sur $[a, b]$ de dérivée f. Comme $F(a) = 0$ (car $\lambda(\{a\}) = 0$), on a $F(x) = \int_a^x f(t)\, dt$ pour tout $x \in [a, b]$ (et notamment pour $x = b$). □

Remarque 4.7.5. La proposition précédente s'étend facilement au cas d'une fonction f continue par morceaux. Elle conduit à noter $\int_{[a,b]} f \, dx$ l'intégrale de Lebesgue $\int f \mathbf{1}_{[a,b]} \, d\lambda$ (et même $\int_a^b f(x)\, dx$). Cette notation est souvent adoptée pour une fonction f intégrable au sens de Lebesgue sur $[a, b]$ sans hypothèse de continuité.

Lorsque l'on sort du cadre des fonctions continues par morceaux, les liens entre intégrales de Riemann et de Lebesgue sont assez subtils. Voici quelques résultats éclairants.

Il existe des fonctions intégrables au sens de Lebesgue qui ne sont pas intégrables au sens de Riemann. Par exemple la fonction $f = \mathbf{1}_{\mathbb{Q} \cap [0,1]}$ est intégrable au sens de Lebesgue et son intégrale est nulle. En revanche, pour toute subdivision σ de $[0, 1]$, on a $S(\sigma) = 1$ et $s(\sigma) = 0$.

Les fonctions intégrables au sens de Riemann sur $[a, b]$ sont connues.

Théorème 4.7.6 (Lebesgue)**.** *Une fonction $f : [a, b] \to \mathbb{R}$ bornée est intégrable au sens de Riemann ssi il existe $N \subset [a, b]$ de mesure de Lebesgue nulle tel que f est continue en tout $x \in [a, b] \backslash N$.*

4.7.4 Intégrale de Riemann généralisée et intégrale de Lebesgue

Proposition 4.7.7. *Soit $f : [a, b[\to \mathbb{R}$ une fonction continue. Alors $f \mathbf{1}_{[a,b[} \in \mathcal{L}^1_{\mathbb{R}}(\lambda)$ si et seulement si $\int_a^b f(t)\, dt$ est absolument convergente et, dans ce cas, on a*

$$\int f \mathbf{1}_{[a,b[} \, d\lambda = \int_a^b f(t)\, dt.$$

Démonstration. Supposons d'abord f positive. Soit $(b_n)_{n \in \mathbb{N}}$ une suite croissante de points de $[a, b[$ qui converge vers b. Pour tout n,

$$\int f \mathbf{1}_{[a,b_n[} \, d\lambda = \int_a^{b_n} f(t)\, dt.$$

En utilisant le théorème de convergence monotone (pour l'intégrale de Lebesgue), on a

$$\int f \mathbf{1}_{[a,b[} \, d\lambda = \lim_{n \to +\infty} \int f \mathbf{1}_{[a,b_n[} \, d\lambda = \lim_{n \to +\infty} \int_a^{b_n} f(t)\, dt \in \overline{\mathbb{R}}_+.$$

Or, par définition, f est intégrable au sens de Lebesgue si et seulement si cette limite est finie, donc si et seulement si f est intégrable au sens de Riemann. De plus les intégrales sont les mêmes.

Dans le cas général, on sait que f est intégrable au sens de Lebesgue si et seulement si $|f|$ l'est, donc si et seulement si $\int_a^b f(t)\,dt$ est absolument convergente. Si c'est le cas, écrivons $f = f^+ - f^-$. On a $f^+ \leq |f|$ et $f^- \leq |f|$ donc f^+ et f^- sont intégrables dans les deux sens et

$$\int f^+ \mathbf{1}_{[a,b[}\,d\lambda = \int_a^b f^+(t)\,dt, \quad \text{et} \quad \int f^- \mathbf{1}_{[a,b[}\,d\lambda = \int_a^b f^-(t)\,dt,$$

d'où le résultat par linéarité. □

Exemple 4.7.8. Soit f définie par $f(x) = x^{100}e^{-x}\mathbf{1}_{\mathbb{R}_+}(x)$. Démontrons qu'elle est intégrable (pour la mesure de Lebesgue). La fonction f est continue sur $\mathbb{R}$ donc mesurable. De plus,

$$f(x) = x^{100}e^{-x/2}e^{-x/2}\mathbf{1}_{\mathbb{R}_+}(x).$$

Par croissance comparée, $x \mapsto x^{100}e^{-x/2}$ converge vers 0 quand $x \to \infty$ donc il existe $M \in \mathbb{R}_+$ tel que pour tout $x \geq M$, $0 \leq x^{100}e^{-x/2} \leq 1$. Ainsi, pour $x \geq M$, $|f(x)| \leq e^{-x/2}$. Puisque f est continue, elle est majorée sur le compact $[0, M]$ par une constante K. On a donc, pour tout $x \in \mathbb{R}$,

$$0 \leq f(x) \leq K\mathbf{1}_{[0,M]}(x) + e^{-x/2}\mathbf{1}_{[M,+\infty[}(x).$$

Le membre de droite est une fonction intégrable donc f l'est aussi.

Chapitre 5

Théorèmes limites et applications

5.1 Lemme de Fatou

Dans le chapitre précédent, nous avons déjà établi un théorème limite fondamental : le théorème de convergence monotone (ou théorème de Beppo Levi).

Théorème 5.1.1. *Soit* $(f_n)_{n\in\mathbb{N}}$ *une suite croissante de* $\mathcal{M}_+$. *Alors* $f = \lim_n f_n \in \mathcal{M}_+$ *et*

$$\int f\,d\mu = \lim_{n\to+\infty} \int f_n\,d\mu.$$

Toutefois, l'hypothèse de croissance, très pratique puisqu'elle assure l'existence de la limite dans $\overline{\mathbb{R}}_+$, est inadaptée dans bien des situations. Nous avons besoin d'un théorème valable pour une suite d'applications générique. Le prix à payer est que l'on ne sera plus assuré de l'existence d'une limite. Par contre l'application $\liminf f_n$ est encore définie et c'est elle qui remplacera avantageusement l'application $\lim f_n$.

Théorème 5.1.2 (Lemme de Fatou)**.** *Si* $(f_n)_{n\in\mathbb{N}}$ *est une suite d'applications mesurables positives, alors*

$$\int \liminf_{n\to\infty} f_n\,d\mu \le \liminf_{n\to\infty} \int f_n\,d\mu.$$

Démonstration. Posons $g = \liminf f_n$. Par définition,

$$g = \lim_{n\to+\infty} \inf_{k\ge n} f_k = \sup_{n\in\mathbb{N}} \inf_{k\ge n} f_k.$$

L'application $g_n = \inf_{k\ge n} f_k$ est une application mesurable positive et la suite $(g_n)_{n\in\mathbb{N}}$ converge en croissant vers g. Le théorème de convergence monotone assure donc :

$$\lim_{n\to\infty} \int g_n\,d\mu = \int \liminf_{n\to\infty} f_n\,d\mu.$$

D'autre part, pour tout $n \in \mathbb{N}$, $g_n \le f_n$, d'où, par croissance de l'intégrale, $\int g_n\,d\mu \le \int f_n\,d\mu$. Le second membre de cette inégalité n'a pas nécessairement de limite mais sa limite inférieure existe toujours. On obtient donc par passage à la limite inférieure :

$$\liminf_{n\to\infty}\int g_n\,d\mu \leq \liminf_{n\to\infty}\int f_n\,d\mu.$$

La limite inférieure du premier membre de l'inégalité ci-dessus est en fait une limite d'après la première partie de la preuve, qui n'est rien d'autre que l'intégrale de la limite inférieure de la suite $(f_n)_{n\in\mathbb{N}}$. Cette remarque achève donc la preuve. □

Remarque 5.1.3. Un des intérêts pratiques du lemme de Fatou est le suivant. Si $(f_n)_n$ est une suite d'applications mesurables positives qui converge simplement vers une application f dont l'intégrale vaut $+\infty$, alors

$$\lim_{n\to\infty}\int_E f_n d\mu = +\infty.$$

C'est le cas par exemple de la suite d'applications données par

$$f_n \;:\; x \mapsto \frac{x^n}{1+x^{n+1}}\mathbf{1}_{\mathbb{R}_+}(x).$$

5.2 Ensembles et applications négligeables

Définition 5.2.1. Soit $(E,\mathcal{A},\mu)$ un espace mesuré.

(i) On dit qu'une partie N de E est négligeable pour μ (ou μ-négligeable) s'il existe $A\in\mathcal{A}$ tel que $N\subset A$ et $\mu(A)=0$.

(ii) On dit que la tribu $\mathcal{A}$ est complète pour μ si toute partie μ-négligeable appartient à $\mathcal{A}$.

Définition 5.2.2. Soit $(E,\mathcal{A},\mu)$ un espace mesuré. On dit qu'une propriété $\mathcal{P}$ sur E est vraie presque partout (en abrégé p.p. ou μ-p.p.) si l'ensemble des points de E où elle est fausse est négligeable.

Une application f définie sur E à valeurs réelles ou complexes est dite μ-négligeable si $\{f\neq 0\}$ est négligeable.

Deux applications f et g définies sur E à valeurs dans un même ensemble F sont dites égales presque partout si $\{f\neq g\}$ est négligeable.

On dit qu'une suite $(f_n)_{n\in\mathbb{N}}$ de applications définies sur E à valeurs dans $\mathbb{C}$ converge vers f μ-presque partout s'il existe un ensemble μ-négligeable N tel que pour tout $x\notin N$, on ait $\lim_n f_n(x)=f(x)$.

Le lemme suivant est très utile en pratique.

Lemme 5.2.3 (Inégalité de Markov). *Soit f une application mesurable positive sur $(E,\mathcal{A})$. Alors pour tout $\lambda>0$, on a*

$$\mu(\{f\geq\lambda\})\leq\frac{1}{\lambda}\int f\,d\mu.$$

Démonstration. Il suffit d'intégrer la relation $\lambda\mathbf{1}_{\{f\geq\lambda\}}\leq f$ qui est vraie puisque f est positive. □

Proposition 5.2.4. *Soit f une application mesurable de $(E, \mathcal{A})$ dans $\overline{\mathbb{R}}$ telle que $\int |f| \, d\mu < +\infty$. Alors f est finie μ-presque partout.*

Démonstration. En effet, pour tout n, on a

$$\frac{1}{n} \int |f| \, d\mu \geq \mu(\{|f| \geq n\}) \geq \mu(\{|f| = +\infty\}).$$

Comme $\int |f| \, d\mu$ est fini, en faisant tendre n vers $+\infty$, on obtient $\mu(\{|f| = +\infty\}) = 0$. □

Remarque 5.2.5. La réciproque de cette proposition est fausse : l'application constante égale à 1 est finie λ-p.p. mais n'est pas intégrable par rapport à la mesure de Lebesgue.

Proposition 5.2.6. *Soit f une application mesurable sur $(E, \mathcal{A})$ à valeurs complexes. Alors f est négligeable si et seulement si $\int |f| \, d\mu = 0$.*

Démonstration. Supposons tout d'abord que f est négligeable. Comme $\min(|f|, n) \leq n\mathbf{1}_{\{f \neq 0\}}$, on a

$$\int \min(|f|, n) \, d\mu \leq n\mu(\{f \neq 0\}) = 0,$$

d'où $\int \min(|f|, n) \, d\mu = 0$ pour tout n. D'après le théorème de convergence monotone, on a alors

$$\int |f| \, d\mu = \int \lim_n \min(|f|, n) \, d\mu = \lim_n \int \min(|f|, n) \, d\mu = 0.$$

Réciproquement, supposons que $\int |f| \, d\mu = 0$. Alors, pour tout $n \geq 1$, on a

$$\mu\left(\left\{|f| \geq \frac{1}{n}\right\}\right) \leq n \int |f| \, d\mu = 0.$$

L'ensemble $\{|f| \neq 0\}$ s'écrit donc comme réunion dénombrable d'ensembles de mesure nulle :

$$\{|f| \neq 0\} = \bigcup_{n \geq 1} \left\{|f| \geq \frac{1}{n}\right\}.$$

Il est donc également de mesure nulle. □

Proposition 5.2.7. *Soit $(E, \mathcal{A}, \mu)$ un espace mesuré.*

(i) Soit f et g deux applications mesurables positives telles que $f \leq g$ presque partout. Alors $\int f \, d\mu \leq \int g \, d\mu$.

(ii) Soit f et g deux applications mesurables positives telles que $f = g$ presque partout. Alors $\int f \, d\mu = \int g \, d\mu$.

(iii) Soit f et g deux applications mesurables complexes telles que $f = g$ presque partout. Alors f est intégrable si et seulement si g l'est et, dans ce cas, $\int f \, d\mu = \int g \, d\mu$.

Démonstration. Pour prouver (i), on écrit

$$f = f\mathbf{1}_{\{f\leq g\}} + f\mathbf{1}_{\{f>g\}},$$

que l'on intègre par rapport à μ :

$$\int f\,d\mu = \int f\mathbf{1}_{\{f\leq g\}}\,d\mu + \int f\mathbf{1}_{\{f>g\}}\,d\mu.$$

Par hypothèse, $f\mathbf{1}_{\{f>g\}}$ est négligeable donc son intégrale est nulle. On a donc

$$\int f\,d\mu = \int f\mathbf{1}_{\{f\leq g\}}\,d\mu.$$

De même, on voit que

$$\int g\,d\mu = \int g\mathbf{1}_{\{f\leq g\}}\,d\mu.$$

Pour conclure, il suffit de remarquer que $f\mathbf{1}_{\{f\leq g\}} \leq g\mathbf{1}_{\{f\leq g\}}$. Le point ($ii$) se déduit de ($i$) par symétrie entre f et g.

Démontrons (iii). Si $f = g$ μ-p.p., alors $|f| = |g|$ μ-p.p., d'où $\int |f|\,d\mu = \int |g|\,d\mu$ par (ii). Par conséquent $f \in \mathcal{L}^1_{\mathbb{C}}(\mu)$ si et seulement si $g \in \mathcal{L}^1_{\mathbb{C}}(\mu)$. On obtient la conclusion par égalité μ-p.p. des parties positives et négatives des parties réelles et imaginaires et en appliquant (ii). □

5.3 Théorème de convergence dominée

Théorème 5.3.1 (de convergence dominée)**.** *Soit $(f_n)_{n\in\mathbb{N}}$ une suite d'applications mesurables sur $(E, \mathcal{A})$ à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$ telle que :*

(i) $(f_n)_{n\in\mathbb{N}}$ converge μ-presque partout vers une application f mesurable,

(ii) il existe une application g positive appartenant à $\mathcal{L}^1_{\mathbb{R}}(\mu)$ telle que

$$\forall n \geq 1, \quad |f_n(x)| \leq g(x) \quad \mu - p.p.$$

Alors les applications $(f_n)_{n\in\mathbb{N}}$ et f sont intégrables et

$$\lim_{n\to\infty} \int f_n\,d\mu = \int f\,d\mu.$$

On a même $\lim_n \int |f_n - f|\,d\mu = 0$.

Démonstration. Supposons tout d'abord que la convergence de $(f_n)_{n\in\mathbb{N}}$ vers f ait lieu partout et que les inégalités (ii) soient vraies pour tout $x \in E$. Posons $g_n = 2g - |f_n - f|$. Alors $(g_n)_{n\in\mathbb{N}}$ est une suite d'applications mesurables positives, et d'après le lemme de Fatou,

$$2\int g\,d\mu = \int \liminf_n g_n\,d\mu \leq \liminf_n \int g_n\,d\mu = 2\int g\,d\mu - \limsup_n \int |f_n - f|\,d\mu.$$

Puisque $\int g\,d\mu < +\infty$, on voit que $\limsup_n \int |f_n - f|\,d\mu \leq 0$. On en déduit donc que

$$\lim_{n\to+\infty} \int |f_n - f|\,d\mu = 0.$$

Il en résulte que $\int f\,d\mu = \lim \int f_n\,d\mu$.

Passons à présent au cas général. Soit $N \in \mathcal{A}$ tel que, si $x \notin N$, $\lim_n f_n(x) = f(x)$ et $\mu(N) = 0$. Choisissons de plus, pour tout $n \in \mathbb{N}$, un ensemble $N_n \in \mathcal{A}$ tel que si $x \notin N_n$ $|f_n(x)| \leq g(x)$ et $\mu(E_n) = 0$. Posons $M = N \cup (\cup_n N_n) \in \mathcal{A}$. On a encore $\mu(M) = 0$. Posons $h_n = f_n \mathbf{1}_{M^c}$ et $h = f\mathbf{1}_{M^c}$. On a, pour tout $x \in E$ et tout n,

$$\lim_{m\to+\infty} h_m(x) = h(x) \quad \text{et} \quad |h_n(x)| \leq g(x).$$

La première partie de la preuve assure donc que $\lim \int |h_n - h|\,d\mu = 0$. Pour conclure, il suffit de remarquer que $h_n = f_n$ μ-p.p. et $h = f$ μ-p.p. □

Exemple 5.3.2. Déterminer la limite éventuelle de la suite de terme général suivant :

$$I_n = \int_0^n \left(1 + \frac{x}{n}\right)^n e^{-2x} dx.$$

Posons, pour $x \in \mathbb{R}$,

$$f_n(x) = \left(1 + \frac{x}{n}\right)^n e^{-2x} \mathbf{1}_{[0,n]}(x).$$

Puisque

$$\left(1 + \frac{x}{n}\right)^n = e^{n\ln(1+x/n)} = e^{n(x/n + o(1/n))},$$

on a

$$f_n(x) \xrightarrow[n\to\infty]{} f(x) = e^{-x}\mathbf{1}_{\mathbb{R}_+}(x).$$

De plus, pour tout $u > -1$, $\ln(1+u) \leq u$ donc $f_n(x) \leq f(x)$ pour tout $x \in \mathbb{R}$ avec f intégrable. Le théorème de convergence dominée assure donc que

$$\lim_{n\to\infty} I_n = \lim_{n\to\infty} \int_{\mathbb{R}} f_n(x)dx = \int_{\mathbb{R}} f(x)dx = 1.$$

Exemple 5.3.3. Déterminer la limite éventuelle de la suite de terme général suivant :

$$J_n = \int_{\mathbb{R}_+} \frac{x^n}{1 + x^{n+2}} dx.$$

Posons, pour $x \in \mathbb{R}$,

$$g_n(x) = \frac{x^n}{1 + x^{n+2}} \mathbf{1}_{\mathbb{R}_+}(x).$$

On a

$$\lim_{n\to\infty} g_n(x) = \frac{1}{2}\mathbf{1}_{\{1\}}(x) + \frac{1}{x^2}\mathbf{1}_{]1,+\infty[}(x).$$

De plus,

$$|g_n(x)| \leq \mathbf{1}_{[0,1]}(x) + \frac{1}{x^2}\mathbf{1}_{]1,+\infty[}(x) \in \mathcal{L}^1(\lambda).$$

Cette majoration a été obtenue ainsi : sur $[0,1]$, $g_n(x)$ est majorée en majorant le numérateur par 1 et en minorant le dénominateur par 1 également ; pour $x > 1$, $g_n(x)$ est majorée en minorant le dénominateur par x^{n+2}. On a donc, d'après le théorème de convergence dominée,

$$\lim_{n\to\infty} J_n = \lim_{n\to\infty} \int_{\mathbb{R}} g_n(x)dx = \int \frac{1}{x^2}\mathbf{1}_{]1,+\infty[}(x)dx = 1.$$

Exemple 5.3.4. Déterminer la limite éventuelle de la suite de terme général suivant :

$$K_n = \int_{\mathbb{R}_+} \frac{\cos(x/n)}{1+x^n}dx.$$

Posons, pour $x \in \mathbb{R}$,

$$h_n(x) = \frac{\cos(x/n)}{1+x^n}\mathbf{1}_{\mathbb{R}_+}(x).$$

On a

$$\lim_{n\to\infty} h_n(x) = \mathbf{1}_{[0,1[}(x) + \frac{1}{2}\mathbf{1}_{\{1\}}(x)$$

et, pour $n \geq 2$,

$$|h_n(x)| \leq \mathbf{1}_{[0,1]}(x) + \frac{1}{1+x^2}\mathbf{1}_{[1,+\infty[}(x) \in \mathcal{L}^1.$$

Pour $x \geq 1$, la majoration de $|h_n(x)|$ ci-dessus est obtenue en majorant $|\cos(x/n)|$ par 1 et en minorant $1 + x^n$ par $1 + x^2$. On a donc, d'après le théorème de convergence dominée,

$$\lim_{n\to\infty} K_n = \lim_{n\to\infty} \int_{\mathbb{R}} h_n(x)dx = \int \mathbf{1}_{[0,1]}(x)dx = 1.$$

Exemple 5.3.5. Pour $n \in \mathbb{N}^*$, notons f_n l'application définie par

$$f_n(x) = \frac{n\sin(x/n)}{x^3}\mathbf{1}_{[1,+\infty[}(x).$$

Pour tout $x \in \mathbb{R}$, on a

$$\lim_{n\to\infty} f_n(x) = \frac{1}{x^2}\mathbf{1}_{[1,+\infty[}(x)$$

car $\sin(u) \sim_0 u$. D'autre part, $|\sin(u)| \leq |u|$ pour tout $u \in \mathbb{R}$ donc

$$|f_n(x)| \leq \frac{1}{x^2}\mathbf{1}[1,+\infty[(x) \in \mathcal{L}^1.$$

D'après le théorème de convergence dominée, on obtient

$$\lim_{n\to\infty} \int_1^\infty \frac{n\sin(x/n)}{x^3}dx = \int_1^\infty \frac{1}{x^2}dx = 1.$$

Corollaire 5.3.6. *Soit $(f_n)_{n\in\mathbb{N}}$ une suite d'applications mesurables sur $(E,\mathcal{A})$ à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$ telle que $\sum_n \int |f_n|\,d\mu < +\infty$. Alors les applications $(f_n)_{n\in\mathbb{N}}$ sont intégrables, la série $\sum_n f_n$ converge μ-p.p. et il existe $f \in \mathcal{L}^1(\mu)$ telle que*

$$f = \sum_{n=1}^{\infty} f_n \quad \mu-p.p., \qquad \lim_{n\to\infty} \int \left| f - \sum_{k=1}^{n} f_k \right| d\mu = 0, \qquad \int f\,d\mu = \sum_{n=1}^{\infty} \int f_n\,d\mu.$$

Démonstration. Posons $g = \sum_{n\geq 1} |f_n|$. D'après le corollaire 4.2.6 (intervertion série-intégrale pour des applications positives),

$$\int g\,d\mu = \sum_{n\geq 1} \int |f_n|\,d\mu < +\infty.$$

L'application g étant intégrable, elle est finie μ-p.p. Posons

$$N = \left\{ x \in E,\ \sum_{n=1}^{+\infty} |f_n(x)| = +\infty \right\}.$$

C'est un ensemble négligeable de $\mathcal{A}$, et si $x \notin N$, la série $\sum f_n(x)$ est absolument convergente, donc convergente. Posons alors

$$f(x) = \begin{cases} \displaystyle\sum_{n=1}^{+\infty} f_n(x) & \text{si } x \notin N, \\ 0 & \text{si } x \in N. \end{cases}$$

Cette application est mesurable comme limite simple de la suite $(\mathbf{1}_{N^c} \sum_{k=1}^{n} f_k)_{n\in\mathbb{N}}$ d'applications mesurables. De plus, comme

$$\forall x \in N^c, \quad |f(x)| \leq \sum_{n=1}^{+\infty} |f_n(x)| = g(x),$$

et comme g est intégrable, f l'est aussi et on a

$$\int |f|\,d\mu \leq \int g\,d\mu = \sum_{n=1}^{+\infty} \int |f_n|\,d\mu.$$

Enfin,

$$\begin{aligned} \int \left| f - \sum_{k=1}^{n} f_k \right| d\mu &= \int \mathbf{1}_{N^c} \left| f - \sum_{k=1}^{n} f_k \right| d\mu = \int \mathbf{1}_{N^c} \left| \sum_{k=n+1}^{+\infty} f_k \right| d\mu \\ &\leq \sum_{k=n+1}^{+\infty} \int \mathbf{1}_{N^c} |f_k|\,d\mu \leq \sum_{k=n+1}^{+\infty} \int |f_k|\,d\mu. \end{aligned}$$

Par hypothèse, le membre de droite tend vers 0, ce qui achève la preuve. □

5.4 Intégrale dépendant d'un paramètre

Théorème 5.4.1 (continuité d'une intégrale à paramètre)**.** *Soit* $(E, \mathcal{A}, \mu)$ *un espace mesuré,* (G, d) *un espace métrique et* f *une application définie sur* $E \times G$*, à valeurs réelles ou complexes. On suppose que*

(i) pour μ*-presque tout* $x \in E$*, l'application* $t \mapsto f(x,t)$ *est continue sur* G*,*

(ii) pour tout $t \in G$*, l'application* $x \mapsto f(x,t)$ *est mesurable sur* $(E, \mathcal{A})$*,*

(iii) il existe une application g *sur* $(E, \mathcal{A})$ *intégrable telle que pour tout* $t \in G$*, on ait*

$$|f(x,t)| \leq g(x) \quad \mu\text{-p.p.}$$

Alors l'application $F : t \mapsto \int f(x,t)\mu(dx)$ *est définie et continue sur* G*.*

Démonstration. Pour tout $t \in G$, l'application $x \mapsto f(x,t)$ est μ-intégrable et F est donc bien définie sur G. Soit $t \in G$. Montrons que F est continue au point t. Comme G est un espace métrique, il suffit de montrer que si $(t_n)_{n\in\mathbb{N}}$ est une suite de G qui converge vers t, alors la suite $(F(t_n))_{n\in\mathbb{N}}$ converge vers t. Notons f_n l'application définie sur E qui à tout $x \in E$ associe $f_n(x) = f(x, t_n)$. La suite $(f_n)_{n\in\mathbb{N}}$ satisfait aux hypothèses du théorème de convergence dominée, d'où la conclusion. $\square$

Théorème 5.4.2 (dérivabilité d'une intégrale à paramètre)**.** *Soit* $(E, \mathcal{A}, \mu)$ *un espace mesuré,* I *un intervalle ouvert de* $\mathbb{R}$ *et* f *une application définie sur* $E \times I$*, à valeurs réelles ou complexes. On suppose que*

(i) pour μ*-presque tout* $x \in E$*, l'application* $t \mapsto f(x,t)$ *est de classe* $\mathcal{C}^1$ *sur* I*,*

(ii) pour tout $t \in I$*, l'application* $x \mapsto f(x,t)$ *est* μ*-intégrable,*

(iii) il existe une application g *sur* $(E, \mathcal{A})$ *intégrable et positive telle que pour* μ*-presque tout* $x \in E$*, on ait*

$$\forall t \in I, \quad \left|\frac{\partial f}{\partial t}(x,t)\right| \leq g(x).$$

Alors, l'application $F : t \mapsto \int f(x,t)\mu(dx)$ *est de classe* $\mathcal{C}^1$ *sur* I *et*

$$\forall t \in I, \quad F'(t) = \int \frac{\partial f}{\partial t}(x,t)\mu(dx).$$

Démonstration. Par hypothèse, il existe un ensemble de mesure nulle $N \in \mathcal{A}$ tel que si $x \notin N$, la dérivée partielle $\frac{\partial f}{\partial t}(x,t)$ existe pour tout point $t \in I$ et

$$\left|\frac{\partial f}{\partial t}(x,t)\right| \leq g(x).$$

Il en résulte que $x \mapsto \frac{\partial f}{\partial t}(x,t)$ est μ-intégrable pour tout $t \in I$.

Étudions la dérivabilité de F en $t \in I$. Soit (t_n) une suite de I qui converge vers t avec $t_n \neq t$ pour tout n. D'après le théorème des accroissements finis, on a, si $x \notin N$,

$$|f(x,t_n) - f(x,t)| \leq |t_n - t| \sup_{t\in I}\left|\frac{\partial f}{\partial \alpha}(x,t)\right| \leq |t_n - t| g(x).$$

On peut donc appliquer le théorème de convergence dominée à la suite $(h_n)_{n\in\mathbb{N}}$ où l'application h_n est définie sur E par

$$h_n(x) = \frac{f(x,t_n) - f(x,t)}{t_n - t}.$$

Cette suite $(h_n)_{n\in\mathbb{N}}$ converge simplement sur $E\backslash N$ vers l'application $x \mapsto \frac{\partial f}{\partial t}(x,t)$. Cette application est donc μ-intégrable. De plus, on a

$$\int \frac{\partial f}{\partial t}(x,t)\,\mu(dx) = \lim_{n\to\infty} \int \frac{f(x,t_n) - f(x,t)}{t_n - t}\mu(dx) = \lim_{n\to\infty} \frac{F(t_n) - F(t)}{t_n - t}.$$

Il en résulte que F est dérivable en t de dérivée

$$F'(t) = \int \frac{\partial f}{\partial t}(x,t)\,\mu(dx).$$

La continuité de F découle du théorème 5.4.1. □

Remarque 5.4.3. Il faut toujours bien faire la différence entre la variable d'intégration et le paramètre : dans l'expression

$$F(x) = \int f(x,t)d\mu(t),$$

la variable d'intégration est t et le paramètre est x.

Exemple 5.4.4. Pour tous $x, t \in [0, +\infty[$, on pose

$$f(t,x) = \frac{\sqrt{x+t}}{1+t^2} \quad \text{et} \quad F(x) = \int_{\mathbb{R}_+} \frac{\sqrt{x+t}}{1+t^2}dt.$$

Pour tout $x \geq 0$, $t \mapsto f(t,x)$ est continue sur $\mathbb{R}_+$ et équivalente en $+\infty$ à $t^{-3/2}$. Elle est donc intégrable. Soit $M > 0$. Pour tout $x \in [0, M]$,

$$0 \leq f(t,x) \leq f(M,x).$$

De plus, pour tout $t \geq 0$, $x \mapsto f(t,x)$ est continue sur $\mathbb{R}_+$. Ainsi, F est continue sur $[0, M]$. Puisque ceci est vrai pour tout $M > 0$, F est continue sur $\mathbb{R}_+$. De plus, $x \mapsto f(t,x)$ est dérivable et

$$\frac{\partial f}{\partial x}(t,x) = \frac{1}{2(1+t^2)\sqrt{x+t}}.$$

Pour tout $x \geq 0$,

$$0 \leq \frac{\partial f}{\partial x}(t,x) \leq g(t) \quad \text{où} \quad g(t) = \frac{1}{2(1+t^2)\sqrt{t}}.$$

L'application g est continue sur $\mathbb{R}_+^*$ et équivalente à $1/(2t^{5/2})$ en $+\infty$ et à $1/(2\sqrt{t})$ en 0. Elle est donc intégrable. Ainsi, F est dérivable et

$$F'(x) = \int_{\mathbb{R}_+} \frac{1}{2(1+t^2)\sqrt{x+t}}dt.$$

Exemple 5.4.5. Posons, pour $t, x \in \mathbb{R}_+$,

$$f(t,x) = \left(\frac{\sin(x)}{x}\right)^2 e^{-tx} \quad \text{et} \quad F(t) = \int_{\mathbb{R}_+} f(t,x)dx.$$

Pour tout $x > 0$, l'application $t \mapsto f(t,x)$ est continue. Pour tout $t > 0$, l'application $x \mapsto f(t,x)$ est continue donc mesurable. Enfin, pour tout $t \geq 0$,

$$|f(t,x)| \leq \left(\frac{\sin(x)}{x}\right)^2 \leq \mathbf{1}_{[0,1]}(x) + \frac{1}{x^2}\mathbf{1}_{\mathbb{R}_+}(x)$$

et ce majorant est intégrable. L'application F est continue sur $\mathbb{R}_+$.

Soit $\varepsilon > 0$. On applique le théorème de dérivation deux fois sur $]\varepsilon, +\infty[$. Outre les hypothèses établies ci-dessus, $t \mapsto f(t,x)$ est de classe $\mathcal{C}^2$ sur $\mathbb{R}_+$. De plus, pour tout $t > \varepsilon$,

$$\left|\frac{\partial f}{\partial t}(t,x)\right| = \left|-\frac{\sin^2(x)}{x}e^{-tx}\right| \leq xe^{-\varepsilon x}$$

et ce majorant est intégrable. De même,

$$\left|\frac{\partial^2 f}{\partial t^2}(t,x)\right| = \left|\sin^2(x)e^{-tx}\right| \leq e^{-\varepsilon x}$$

et ce majorant est intégrable. L'application F est donc de classe $\mathcal{C}^2$ sur $]\varepsilon, +\infty[$ pour tout $\varepsilon > 0$ et donc sur $\mathbb{R}_+^*$ et

$$F'(t) = -\int_{\mathbb{R}_+} \frac{\sin^2(x)}{x}e^{-tx}dx \quad \text{et} \quad F''(t) = \int_{\mathbb{R}_+} \sin^2(x)e^{-tx}dx.$$

On utilise alors la relation classique

$$\sin^2(x) = \frac{1}{2} - \frac{1}{2}\cos(2x).$$

On obtient

$$\begin{aligned} F''(t) &= \frac{1}{2}\int_{\mathbb{R}_+} e^{-tx}dx + \frac{1}{2}\int_{\mathbb{R}_+} \cos(2x)e^{-tx}dx \\ &= \frac{1}{2} \times \frac{1}{t} + \frac{1}{2} \times \frac{t}{t^2+4}. \end{aligned}$$

On peut déterminer F en intégrant deux fois F'' et en utilisant que F' et F tendent vers 0 en $+\infty$...

Exemple 5.4.6. Posons, pour $t, x \in \mathbb{R}_+$,

$$f(t,x) = \frac{e^{-xt^2}}{1+t^2} \quad \text{et} \quad F(x) = \int_{\mathbb{R}_+} \frac{e^{-xt^2}}{1+t^2}dt.$$

Pour tout $t \in \mathbb{R}_+$, $x \mapsto f(t,x)$ est continue et pour tout $x \in \mathbb{R}_+$, on a

$$0 \leq f(t,x) \leq \frac{1}{1+t^2}.$$

Ce majorant est intégrable. La fonction F est continue sur $\mathbb{R}_+$. De plus, pour tout $t \in \mathbb{R}_+$, $x \mapsto f(t,x)$ est dérivable et pour tout $x > \varepsilon$, on a

$$\left|\frac{\partial f}{\partial x}(t,x)\right| = \frac{t^2}{1+t^2}e^{-xt^2} \leq e^{-\varepsilon t^2}$$

et ce majorant est intégrable. L'application F est donc dérivable sur $]\varepsilon, +\infty[$ et

$$F'(x) = -\int_{\mathbb{R}_+} \frac{t^2}{1+t^2}e^{-xt^2}dt.$$

Ceci est encore valable sur $\mathbb{R}_+^*$. On a, pour $x > 0$,

$$F(x) - F'(x) = \int_{\mathbb{R}_+} e^{-xt^2}dt \overset{u=t\sqrt{x}}{=} \frac{1}{\sqrt{x}}\int_{\mathbb{R}_+} e^{-u^2}du = \frac{\sqrt{\pi}}{2\sqrt{x}}.$$

L'application F est continue sur $\mathbb{R}_+$. On a donc

$$F'(x) \underset{0}{\sim} -\frac{\sqrt{\pi}}{2\sqrt{x}}.$$

Par ailleurs,

$$F'(x) \overset{u=t\sqrt{x}}{=} -\frac{1}{x\sqrt{x}}\int_{\mathbb{R}_+} \frac{u^2}{1+u^2/x}e^{-u^2}du.$$

Par convergence monotone, pour toute suite croissante $(x_n)_n$ qui tend vers $+\infty$, on a

$$I(x_n) = \int_{\mathbb{R}_+} \frac{u^2}{1+u^2/x_n}e^{-u^2}du \xrightarrow[x\to+\infty]{} \int_{\mathbb{R}_+} u^2e^{-u^2}du = \frac{\sqrt{\pi}}{4}.$$

En d'autres termes, $I(x)$ tend vers $+\infty$ lorsque x tend vers $+\infty$. On a donc

$$F(x) = \frac{\sqrt{\pi}}{2\sqrt{x}} - \frac{\sqrt{\pi}}{4x\sqrt{x}} + o(x^{-3/2}).$$

En résumé, F est continue et décroissante sur $\mathbb{R}_+$. Elle vaut $\pi/2$ en 0 mais n'est pas dérivable en ce point puisque $F'(x)$ est équivalente à $-\sqrt{\pi/(4x)}$. Enfin, F tend vers 0 en $+\infty$ comme $\sqrt{\pi/(4x)}$.

Exemple 5.4.7. Posons pour $t, x \in \mathbb{R}$

$$f(t,x) = \cos(tx)e^{-x^2} \quad \text{et} \quad F(t) = \int_{\mathbb{R}} f(t,x)dx.$$

L'application $(t,x) \mapsto f(t,x)$ est de classe $\mathcal{C}^\infty$ et, pour tout $t \in \mathbb{R}$,

$$|f(t,x)| \leq e^{-x^2} \in \mathcal{L}^1_\lambda.$$

L'application F est donc continue sur $\mathbb{R}$. De plus,

$$\left|\frac{\partial f}{\partial t}(t,x)\right| = \left|x\sin(xt)e^{-x^2}\right| \leq xe^{-x^2} \in \mathcal{L}^1_\lambda.$$

L'application F est donc dérivable sur $\mathbb{R}$ et

$$F'(t) = -\int_{\mathbb{R}} \sin(xt)\, xe^{-x^2} dx.$$

Grâce à une intégration par parties, on a pour $a < b$

$$-\int_a^b \sin(xt)\, xe^{-x^2} dx = \left[\sin(xt)\frac{1}{2}e^{-x^2}\right]_a^b - \int_a^b t\cos(xt)\,\frac{1}{2}e^{-x^2} dx$$

En passant à la limite, on obtient

$$F'(t) = -\frac{t}{2}F(t).$$

De plus, $F(0) = \sqrt{\pi}$. Cette équation différentielle admet une unique solution :

$$F(t) = \sqrt{\pi}e^{-t^2/4}.$$

Exemple 5.4.8 (Transformée de Laplace)**.** Soit μ une mesure sur $\mathbb{R}$ telle que $x \mapsto e^{\alpha x}$ soit μ intégrable pour tout $\alpha \in \mathbb{R}$. On note

$$L(\alpha) = \int_{\mathbb{R}} e^{\alpha x} d\mu(x).$$

On applique le théorème de dérivation d'une intégrale à paramètre sur l'intervalle $I = [-M, M]$: pour tout $\alpha \in I$,

$$\left|\frac{\partial^n}{\partial \alpha^n} e^{\alpha x}\right| = |x^n e^{\alpha x}| \leq |x|^n e^{Mx} + |x|^n e^{-Mx}.$$

L'application du membre de droite est intégrable. En particulier, la dérivée seconde de L vaut

$$L(\alpha) = \int_{\mathbb{R}} x^2 e^{\alpha x} d\mu(x) \geq 0.$$

L'application L est donc convexe sur $\mathbb{R}$. Voici deux exemples de transformées de Laplace explicites. Si μ est la mesure de densité $x \mapsto \mathbf{1}_{[0,1]}(x)$ (par rapport à la mesure de Lebesgue) alors

$$L(t) = \int_{\mathbb{R}} e^{tx} d\mu(x) = \int_{\mathbb{R}} e^{tx}\mathbf{1}_{[0,1]}(x)dx = \int_0^1 e^{tx}dx = \frac{e^t - 1}{t} \quad \text{pour } t \in \mathbb{R}.$$

Si $\mu = e^{-1}\sum_{n\geq 0}(n!)^{-1}\delta_n$ est la mesure de Poisson (de paramètre 1),

$$L(t) = \int_{\mathbb{R}} e^{tx} d\mu(x) = \sum_{n\geq 0} e^{tn}\frac{1}{n!} = e^{e^t - 1} \quad \text{pour } t \in \mathbb{R}.$$

Exemple 5.4.9 (Fonction Gamma)**.** Pour tout $s > 0$, on pose

$$\Gamma(s) = \int_0^{+\infty} x^{s-1} e^{-x} dx.$$

Soit $0 < a < 1 < b$. On a la majoration suivante :

$$\left|\ln(x)^n x^{s-1} e^{-x}\right| \leq |\ln(x)|^n x^{a-1} e^{-x} \mathbf{1}_{]0,1]}(x) + |\ln(x)|^n x^{b-1} e^{-x} \mathbf{1}_{]1,+\infty[}(x).$$

Le membre de droite est une application intégrable. La fonction Γ est de classe $\mathcal{C}^\infty$ sur $]0, +\infty[$ et

$$\Gamma^{(n)}(s) = \int_0^{+\infty} \ln(x)^n x^{s-1} e^{-x} dx \quad \text{pour } s > 0.$$

5.5 Transformée de Laplace

Cette section propose une étude détaillée de la transformée de Laplace d'une mesure de probabilité dont un cas particulier est donné dans l'exemple 5.4.8. Soit μ une mesure de probabilité sur $\mathbb{R}$. On pose

$$L(t) = \int_{\mathbb{R}} e^{tx} \mu(dx) \in [0, +\infty], \quad \text{pour } t \in \mathbb{R}.$$

Notons

$$\mathcal{D}_\mu = \{t \in \mathbb{R} \,:\, L(t) < +\infty\}.$$

Remarquons tout d'abord que $L(0) = \int_{\mathbb{R}} \mu(dx) = \mu(\mathbb{R}) = 1$. On a toujours $0 \in \mathcal{D}_\mu$. Si μ est la mesure de densité $x \mapsto (\pi(1+x^2))^{-1}$ alors $\mathcal{D}_\mu = \{0\}$. L'exemple 5.4.8 présente des mesures pour lesquelles $\mathcal{D}_\mu = \mathbb{R}$.

Lemme 5.5.1. *L'application L est convexe, $\mathcal{D}_\mu$ est un intervalle de $\mathbb{R}$.*

Démonstration. Soit $t \in \mathbb{R}$, l'application $x \mapsto e^{tx}$ est convexe. Donc, pour $x_1, x_2 \in \mathbb{R}$ et $\alpha \in [0, 1]$, on a

$$e^{t(\alpha x_1 + (1-\alpha)x_2)} \leq \alpha e^{tx_1} + (1-\alpha) e^{tx_2}.$$

En intégrant cette inégalité par rapport à μ, on obtient

$$L(\alpha x_1 + (1-\alpha)x_2) \leq \alpha L(x_1) + (1-\alpha) L(x_2).$$

La transformée de Laplace L est donc convexe. De plus, si x_1 et x_2 sont dans $\mathcal{D}_\mu$ alors c'est aussi le cas de $\alpha x_1 + (1-\alpha)x_2$. Ceci assure que $\mathcal{D}_\mu$ est un intervalle de $\mathbb{R}$. □

Lemme 5.5.2. *L'application L est de classe $\mathcal{C}^\infty$ sur l'intérieur de $\mathcal{D}_\mu$.*

Démonstration. Soit $s_0 < s_1$ dans l'intérieur de $\mathcal{D}_\mu$. Montrons que L est de classe $\mathcal{C}^\infty$ sur $]s_0, s_1[$. Il existe t_0 et t_1 dans l'intérieur de $\mathcal{D}_\mu$ tels que $t_0 < s_0 < s_1 < t_1$. Alors, pour tout $t \in]s_0, s_1[$, et tout $n \in \mathbb{N}$, on a

$$\left|x^n e^{tx}\right| \mathbf{1}_{\mathbb{R}_+}(x) = x^n e^{-(t_1 - t)x} e^{t_1 x} \leq \underbrace{x^n e^{-(t_1 - s_1)x}}_{=f_1(x)} \underbrace{e^{t_1 x}}_{g_1(x)}.$$

L'application f_1 est continue et tend 0 quand x tend vers $+\infty$. Elle est donc bornée sur $\mathbb{R}_+$ par une constante M_1. L'application f_1 est intégrable sur $\mathbb{R}$ donc sur $\mathbb{R}_+$. De même, on a

$$\left|x^n e^{tx}\right|\mathbf{1}_{\mathbb{R}_-}(x) = |x^n|e^{-(t_0-t)x}e^{t_0x} \leq \underbrace{|x^n|e^{-(t_0-s_0)x}}_{=f_0(x)}\underbrace{e^{t_0x}}_{g_0(x)}.$$

L'application f_0 est continue et tend 0 quand x tend vers $-\infty$. Elle est donc bornée sur $\mathbb{R}_-$ par une constante M_0. L'application f_0 est intégrable sur $\mathbb{R}$ donc sur $\mathbb{R}_-$. En conclusion, on a

$$\sup_{t\in]s_0,s_1[}\left|x^n e^{tx}(x)\right| \leq g(x) \quad \text{où} \quad g(x) = M_0e^{t_0x}\mathbf{1}_{\mathbb{R}_-}(x) + M_1e^{t_1x}\mathbf{1}_{\mathbb{R}_+}(x)$$

avec g intégrable. □

Lemme 5.5.3. *L'application L est log-convexe, c'est-à-dire que* $\ln(L)$ *est convexe.*

Démonstration. Soit t dans l'intérieur de $\mathcal{D}_\mu$. Les deux premières dérivées de $\ln(L)$ au point t sont données par

$$\ln(L)'(t) = \frac{L'(t)}{L(t)} \quad \text{et} \quad \ln(L)''(t) = \frac{L''(t)L(t) - L'(t)^2}{L(t)}.$$

Or, on a, d'après l'inégalité de Cauchy-Schwarz (voir la remarque 5.5.4) appliquée à $x \mapsto xe^{tx/2}$ et $x \mapsto e^{tx/2}$,

$$\begin{aligned}
L'(t)^2 &= \left(\int_{\mathbb{R}} xe^{tx}d\mu(x)\right)^2 = \left(\int_{\mathbb{R}} xe^{tx/2}e^{tx/2}d\mu(x)\right)^2 \\
&\leq \left(\int_{\mathbb{R}} x^2e^{tx}d\mu(x)\right)\left(\int_{\mathbb{R}} e^{tx}d\mu(x)\right) \\
&\leq L''(t)L(t).
\end{aligned}$$

Ainsi, la dérivée seconde de $\ln(L)$ est positive. Cette application est donc convexe. □

Remarque 5.5.4. L'inégalité de Cauchy-Schwarz utilisée ci-dessus s'écrit : pour toutes applications f et g telles que f^2 et g^2 sont μ-intégrables,

$$\left(\int_{\mathbb{R}} fgd\mu\right)^2 \leq \left(\int_{\mathbb{R}} f^2d\mu\right)\left(\int_{\mathbb{R}} g^2d\mu\right).$$

Elle se démontre en remarquant que, pour tout $\alpha \in \mathbb{R}$, on a

$$0 \leq \int_{\mathbb{R}} (f+\alpha g)^2d\mu = \int_{\mathbb{R}} f^2d\mu + 2\alpha\int_{\mathbb{R}} fgd\mu + \alpha^2\int_{\mathbb{R}} g^2d\mu.$$

Ainsi, le discriminant de ce polynôme de degré 2 avec au plus une racine réelle est négatif :

$$0 > \Delta := 4\left(\int_{\mathbb{R}} fgd\mu\right)^2 - 4\left(\int_{\mathbb{R}} f^2d\mu\right)\left(\int_{\mathbb{R}} g^2d\mu\right),$$

ce qui démontre l'inégalité.

5.6 Un premier contact avec la transformée de Fourier

La notion de transformée de Fourier d'une application intégrable pour la mesure de Lebesgue ou d'une mesure finie joue un rôle essentiel en analyse et en probabilités. Nous présentons ici quelques unes de ses propriétés et quelques exemples de calculs.

5.6.1 Transformée de Fourier d'une application

Définition 5.6.1. Soit $f \in \mathcal{L}^1_{\mathbb{R}}(\lambda)$. La transformée de Fourier de f est définie sur $\mathbb{R}$ par

$$\hat{f}(t) = \int_{\mathbb{R}} f(x) e^{itx} \, \lambda(dx).$$

Remarque 5.6.2. Comme $\left|e^{itx}\right| = 1$, $x \mapsto e^{itx} f(x)$ est Lebesgue intégrable si et seulement si f l'est.

Remarque 5.6.3. Par définition de l'intégrale d'une application à valeurs complexes, on a aussi :

$$\hat{f}(t) = \int_{\mathbb{R}} \cos(tx) f(x) \, \lambda(dx) + i \int_{\mathbb{R}} \sin(tx) f(x) \, \lambda(dx).$$

Si f est paire alors

$$\hat{f}(t) = 2 \int_{[0,+\infty[} \cos(tx) f(x) \, \lambda(dx),$$

et en particulier, la partie imaginaire de $\hat{f}(t)$ est nulle. De même, si f est impaire alors

$$\hat{f}(t) = 2i \int_{[0,+\infty[} \sin(tx) f(x) \, \lambda(dx),$$

et en particulier, la partie réelle de $\hat{f}(t)$ est nulle.

Les théorèmes sur l'intégrale à paramètre permettent de montrer immédiatement le résultat de dérivabilité suivant.

Proposition 5.6.4. *L'application $\hat{f}$ est continue sur $\mathbb{R}$. De plus, si $x \mapsto x f(x)$ est intégrable sur $\mathbb{R}$ alors $\hat{f}$ est dérivable sur $\mathbb{R}$ et*

$$\hat{f}'(t) = \int_{\mathbb{R}} ix e^{itx} f(x) \, \lambda(dx).$$

Exemple 5.6.5. Soit $f_1(x) = e^{-x} \mathbf{1}_{\{x>0\}}$. Alors

$$\hat{f}_1(t) = \int_{\mathbb{R}} e^{itx} e^{-x} \mathbf{1}_{\{x>0\}} \, \lambda(dx) = \int_{\mathbb{R}^+} e^{-(1-it)x} \, \lambda(dx).$$

L'intégrale (au sens de Riemann) généralisée $\displaystyle\int_0^{+\infty} e^{-(1-it)x} \, dx$ est absolument convergente donc :

$$\int_{\mathbb{R}^+} e^{-(1-it)x} \, \lambda(dx) = \int_0^{+\infty} e^{-(1-it)x} \, dx.$$

La dérivée de $x \mapsto \dfrac{e^{-(1-it)x}}{-1+it}$ est l'application $x \mapsto e^{-(1-it)x}$. De plus, $\left|e^{-(1-it)x}\right| = e^{-x}$ tend vers 0 quand x tend vers l'infini. Donc

$$\hat{f}_1(t) = \lim_{A \to +\infty} \int_0^A e^{-(1-it)x}\, dx = \lim_{A \to +\infty} \left[\frac{e^{-(1-it)x}}{-1+it}\right]_0^A = \frac{1}{1-it}.$$

On en déduit donc que $\hat{f}_t(t) = \dfrac{1}{1-it}$ pour tout $t \in \mathbb{R}$.

Exemple 5.6.6. Soit à présent $f_2(x) = e^{-|x|}$. D'après la remarque 5.6.3 et le calcul précédent,

$$\begin{aligned} \hat{f}_2(t) &= 2\int_{[0,+\infty[} \cos(tx) e^{-|x|}\, \lambda(dx) = 2\int_{[0,+\infty[} \mathrm{Re}\,(e^{itx}) e^{-x}\, \lambda(dx) \\ &= 2\mathrm{Re} \int_{[0,+\infty[} e^{-(1-it)x}\, \lambda(dx) = 2\mathrm{Re}\, \hat{f}_1(t) = \frac{2}{1+t^2}. \end{aligned}$$

Théorème 5.6.7. *Soit f_3 définie sur $\mathbb{R}$ par $f_3(x) = 1/(1+x^2)$. Alors*

$$\hat{f}_3(t) = \pi e^{-|t|} \quad \textit{pour } t \in \mathbb{R}.$$

Remarque 5.6.8. Remarquons que $\hat{f}_3$ n'est pas dérivable en 0, ce qui n'est pas contraire aux conclusions de la proposition 5.6.4 puisque $x \mapsto x/(1+x^2)$ n'est pas λ-intégrable sur $\mathbb{R}$.

Remarque 5.6.9. Les calculs ci-dessus montrent que $\hat{\hat{f}}_2(t) = 2\pi f_2$ et $\hat{\hat{f}}_3(t) = 2\pi f_3$. Nous tirerons ceci au clair plus tard.

Exemple 5.6.10. Soit f l'application définie sur $\mathbb{R}$ par $f(x) = e^{-x^2}$. Puisque f est paire, on a

$$\hat{f}(t) = \int_{\mathbb{R}} \cos(tx) e^{-x^2} dx.$$

D'après l'exemple 5.4.7, on obtient

$$\hat{f}(t) = \sqrt{\pi} e^{-t^2/4}.$$

5.6.2 Transformée de Fourier d'une mesure finie

Définition 5.6.11. Soit μ une mesure finie sur $\mathbb{R}$. La transformée de Fourier de μ (on parle encore de fonction caractéristique de μ) est l'application notée φ_μ définie sur $\mathbb{R}$ par

$$\forall t \in \mathbb{R}, \quad \varphi_\mu(t) = \int_{\mathbb{R}} e^{itx}\, \mu(dx).$$

Remarquons tout de suite que si μ admet f (supposée positive) pour densité par rapport à la mesure de Lebesgue sur $\mathbb{R}$, alors μ est finie si et seulement si f est λ-intégrable et $\hat{f} = \varphi_\mu$. En particulier, la proposition 5.6.4 s'étend à toutes les mesures finies : φ_μ est continue sur $\mathbb{R}$ et si $x \mapsto x$ est μ intégrable, φ_μ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ de dérivée

$$\varphi_\mu'(t) = \int_{\mathbb{R}} ixe^{itx}\,\mu(dx).$$

En particulier, $\varphi_\mu'(0) = i\displaystyle\int_{\mathbb{R}} x\,\mu(dx)$.

Exemple 5.6.12. Donnons quelques exemples de fonctions caractéristiques de mesures discrètes classiques.

1. Si $\mu_1 = (1-p)\delta_0 + p\delta_1$, avec $p \in [0,1]$ alors $\varphi_{\mu_1}(t) = pe^{it} + 1 - p$.
2. Si $\mu_2 = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k}\delta_k$, avec $p \in [0,1]$, alors $\varphi_{\mu_2}(t) = \left(pe^{it} + 1 - p\right)^n$.
3. Si $\nu_\alpha = \displaystyle\sum_{k=0}^{+\infty} e^{-\alpha}\frac{\alpha^k}{k!}\delta_k$, avec $\alpha > 0$ alors $\varphi_{\nu_\alpha}(t) = e^{\alpha(e^{it}-1)}$.

Remarque 5.6.13. On peut utiliser le théorème de dérivation de Lebesgue pour calculer l'intégrale de l'application $x \mapsto x$ en évaluant la dérivée de la transformée de Fourier de μ. On peut aussi obtenir ces résultats par un calcul direct mais un peu plus pénible :

$$\int x\,\mu_1(dx) = p, \quad \int x\,\mu_2(dx) = np, \quad \int x\,\nu_\alpha(dx) = \alpha.$$

Remarque 5.6.14. Les relations $\varphi_{\mu_1}^n = \varphi_{\mu_2}$ et $\varphi_{\nu_\alpha}^n = \varphi_{\nu_{n\alpha}}$ ne sont pas des coïncidences. Il faudra faire un peu de probabilités pour le comprendre...

Exemple 5.6.15 (Transformée de Fourier de la mesure gaussienne standard)**.** Pour $t \in \mathbb{R}$, notons

$$\varphi(t) = \int_{\mathbb{R}} e^{itx}e^{-x^2/2}dx.$$

Comme l'application $x \mapsto e^{-x^2/2}$ est paire, l'application φ est à valeurs dans $\mathbb{R}$ et est paire. Le théorème de dérivation d'une intégrale à paramètre assure que

$$\varphi'(t) = \int_{\mathbb{R}} ie^{itx}xe^{-x^2/2}dx.$$

Une intégration par parties implique que $\varphi'(t) = -t\varphi(t)$. Comme $\varphi(0) = 1$, on en déduit que $\varphi(t) = e^{-t^2/2}$.

Chapitre 6

Intégrales multiples

Est-il possible de construire une mesure sur l'espace mesurable $(E \times F, \mathcal{A} \otimes \mathcal{B})$ qui affecterait aux rectangles $A \times B$ la mesure $\mu(A)\nu(B)$? Si μ et ν étaient toutes deux la mesure de Lebesgue sur $\mathbb{R}$, répondre à cette question donnerait un sens mathématique à la notion intuitive d'aire. Si une telle mesure existe, est-elle unique ? Nous pourrons apporter une réponse positive à ces deux questions dans le cas où les deux mesures sont σ-finies.

Nous étudierons dans son ensemble la preuve de l'existence et l'unicité de la mesure produit. Elle présente l'avantage d'être constructive et nous sera utile pour la suite. L'unicité de la mesure s'établit de manière similaire à celle de la mesure de Lebesgue sur $\mathbb{R}$ que nous avons admise au chapitre 2. La preuve de l'existence se fait quant à elle beaucoup plus facilement.

Dans un deuxième temps, nous étudierons les propriétés de l'intégrale par rapport à la mesure produit $\mu \otimes \nu$ en faisant notamment le lien avec les intégrales par rapport aux mesures μ et ν. Les théorèmes de Tonelli et Fubini joueront ce rôle, le premier pour les fonctions mesurables positives, le second pour les fonctions intégrables à valeurs dans $\mathbb{C}$.

Nous énoncerons le théorème de changement de variables pour l'intégrale de Lebesgue sur $\mathbb{R}^d$ et nous rencontrerons à nouveau à cette occasion la notion de mesure image qui jouera un très grand rôle en probabilités.

6.1 Mesures produit

Rappelons qu'une mesure sur $(E, \mathcal{A})$ est dite σ-finie s'il existe une suite croissante $(E_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{A}$ telle que $\mu(E_n)$ est finie pour tout $n \in \mathbb{N}$ et E soit la réunion (croissante) des ensembles $(E_n)_{n\in\mathbb{N}}$. On dit alors que $(E, \mathcal{A}, \mu)$ est σ-fini. Par exemple, la mesure de Lebesgue sur $\mathbb{R}$ est σ-finie (on peut prendre $E_n = [-n, n]$) tandis que la mesure de comptage sur $[0, 1]$ ne l'est pas.

Soit à présent $(E, \mathcal{A}, \mu)$ et $(F, \mathcal{B}, \nu)$ deux espaces mesurés σ-finis. On dispose déjà d'une tribu naturelle sur $E \times F$ construite à partir de $\mathcal{A}$ et $\mathcal{B}$: la tribu produit $\mathcal{A} \otimes \mathcal{B}$. Rappelons que la tribu produit est la tribu engendrée par la famille $\mathcal{R}$ des rectangles de $\mathcal{A} \times \mathcal{B}$, c'est-à-dire par l'ensemble des parties $E \times F$ de la forme $A \times B$ avec $A \in \mathcal{A}$ et $B \in \mathcal{B}$.

Soit C un élément de $\mathcal{A} \otimes \mathcal{B}$. Nous allons étudier les propriétés des opérations *découpage en tranches horizontales et verticales.*

Lemme 6.1.1. *Si $C \in \mathcal{A} \otimes \mathcal{B}$, notons*

$$C_x = \{y \in F,\ (x,y) \in C\} \quad \text{et} \quad C^y = \{x \in E,\ (x,y) \in C\}$$

les sections verticale et horizontale. Alors, pour tout $x \in E$ et tout $y \in F$, $C_x \in \mathcal{B}$ et $C^y \in \mathcal{A}$.

Remarque 6.1.2. Si C et D sont des éléments de $\mathcal{A} \otimes \mathcal{B}$ alors

$$(C_x)^c = (C^c)_x, \quad C_x \cup D_x = (C \cup D)_x \quad \text{et} \quad C_x \cap D_x = (C \cap D)_x.$$

Il en est de même avec les unions et intersections dénombrables.

Preuve du lemme 6.1.1. Soit $\mathcal{C}$ l'ensemble des parties $C \in \mathcal{A} \otimes \mathcal{B}$ telles que, pour tout $x \in E$ et tout $y \in F$, $C_x \in \mathcal{B}$ et $C^y \in \mathcal{A}$. Il est clair que $\mathcal{C}$ est une tribu. Pour achever la preuve, il reste à remarquer que $\mathcal{C}$ contient les rectangles $C = A \times B$. C'est le cas puisqu'alors, pour tous $x \in E$ et $y \in F$,

$$C_x = \begin{cases} B & \text{si } x \in A \\ \emptyset & \text{si } x \notin A \end{cases} \in \mathcal{B} \quad \text{et} \quad C^y = \begin{cases} A & \text{si } y \in B \\ \emptyset & \text{si } y \notin B \end{cases} \in \mathcal{A}.$$

Ainsi, $\mathcal{C}$ est une tribu qui contient les rectangles : elle contient donc la tribu produit $\mathcal{A} \otimes \mathcal{B}$. L'inclusion inverse découle de la définition de $\mathcal{C}$. □

Théorème 6.1.3. *Soit $(E, \mathcal{A}, \mu)$ et $(F, \mathcal{B}, \nu)$ deux espaces mesurés σ-finis.*

(i) Il existe une unique mesure m sur $(E \times F, \mathcal{A} \otimes \mathcal{B})$ telle que, pour tous $A \in \mathcal{A}$ et $B \in \mathcal{B}$,

$$m(A \times B) = \mu(A)\nu(B),$$

avec la convention $0 \times +\infty = 0$. Cette mesure est σ-finie. On la note généralement $\mu \otimes \nu$ et on l'appelle mesure produit de μ et ν.

(ii) Pour tout $C \in \mathcal{A} \otimes \mathcal{B}$, les applications $x \mapsto \nu(C_x)$ et $y \mapsto \mu(C^y)$ sont respectivement $\mathcal{A}$-mesurable et $\mathcal{B}$-mesurable et

$$\mu \otimes \nu(C) = \int_E \nu(C_x)\, \mu(dx) = \int_F \mu(C^y)\, \nu(dy). \tag{6.1}$$

Avant de nous lancer dans la démonstration, il nous faut introduire une notion très importante en intégration et probabilités il s'agit de la notion de λ-système ou de classe monotone. Notons qu'il existe autant de définitions de ces objets que de mathématiciens. Nous utiliserons ici le terme de λ-système sous l'acception suivante.

Définition 6.1.4. Une famille Λ de parties de E est appelée λ-système si

(i) $E \in \Lambda$,

(ii) si $(A_n)_{n \in \mathbb{N}}$ est une suite croissante d'éléments de Λ, alors $A = \cup_n A_n \in \Lambda$,

(iii) si A et B sont dans Λ avec A inclus dans B, alors $B \backslash A$ est dans Λ.

Remarque 6.1.5. Tout comme pour les tribus, il existe un plus petit λ-système contenant une partie $\mathcal{S}$: il s'agit de l'intersection de tous les λ-systèmes contenant $\mathcal{S}$ (intersection non vide car $\mathcal{P}(E)$ en fait partie). On le note $\Lambda(\mathcal{S})$.

Remarque 6.1.6. Une tribu est un λ-système. En particulier, si $\mathcal{S}$ est une partie de $\mathcal{P}(E)$ alors $\Lambda(\mathcal{S}) \subset \sigma(\mathcal{S})$.

Lemme 6.1.7. *Un λ-système stable par intersection finie est une tribu.*

Démonstration. Comme $E \in \Lambda$, Λ est stable par complémentaire d'après (iii) et, en particulier, $\emptyset \in \Lambda$. Soit A et B dans Λ. Alors $A \cup B = (A^c \cap B^c)^c$ appartient à Λ. Par récurrence, on en déduit que Λ est stable par union finie. Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'éléments de Λ. On peut alors écrire $\cup_n A_n$ comme la réunion croissante des ensembles $(\cup_{k=0}^p A_k)_{p\in\mathbb{N}}$ qui sont dans Λ. Ainsi, Λ est stable par union dénombrable et est donc une tribu. □

Il est bien plus facile en pratique de montrer qu'une famille de parties est un λ-système qu'une tribu. Le résultat suivant montre que cela suffit dans certains cas. On donne à toutes les variantes de ce résultat le nom de théorème des classes monotones bien que la version ci-dessous fasse appel aux λ-systèmes.

Théorème 6.1.8 (Théorème des classes monotones)**.** *Si $\mathcal{S}$ est une famille de parties de E fermée pour l'intersection finie, alors $\Lambda(\mathcal{S}) = \sigma(\mathcal{S})$.*

Démonstration. Il suffit, au vu du lemme 6.1.7, de montrer que $\Lambda(\mathcal{S})$ est stable par intersection finie. Soit $B \in \mathcal{S}$ fixé et

$$\Lambda_B = \{A \in \Lambda(\mathcal{S}) \ ; \ A \cap B \in \Lambda(\mathcal{S})\}.$$

On vérifie sans problème que Λ_B est un λ-système contenant $\mathcal{S}$ et donc $\Lambda(\mathcal{S})$, c'est-à-dire que

$$\forall B \in \mathcal{S}, \quad \forall A \in \Lambda(\mathcal{S}), \quad A \cap B \in \Lambda(\mathcal{S}).$$

Soit maintenant $C \in \Lambda(\mathcal{S})$ et

$$\Lambda_C = \{A \in \Lambda(\mathcal{S}) \ ; \ A \cap C \in \Lambda(\mathcal{S})\}.$$

D'après ce qui précède, Λ_C est un λ-système contenant $\mathcal{S}$. Donc pour tout $C \in \Lambda(\mathcal{S})$, $\Lambda_C = \Lambda(\mathcal{S})$.

Finalement $\Lambda(\mathcal{S})$ est stable par intersection finie, c'est donc une tribu et $\sigma(\mathcal{S}) \subset \Lambda(\mathcal{S})$. Comme une tribu est un λ-système l'inclusion précédente est une égalité. □

La famille fermée pour l'intersection finie est très souvent une algèbre de Boole.

Définition 6.1.9. Une algèbre de Boole $\mathcal{A}$ sur E est un sous-ensemble non vide de $\mathcal{P}(E)$ tel que

(i) la partie vide appartient à $\mathcal{A}$,

(ii) $\mathcal{A}$ est stable par passage au complémentaire,

(iii) $\mathcal{A}$ est stable par réunion finie.

Exemple 6.1.10. L'algèbre de Boole que nous rencontrerons le plus souvent dans ce chapitre est l'algèbre de Boole engendrée par $\mathcal{R}$ c'est-à-dire la plus petite algèbre de Boole contenant $\mathcal{R}$. Nous la noterons $\mathcal{F}$. Il est facile de se convaincre qu'elle est formée des réunions finies de rectangles disjoints.

Théorème 6.1.11 (Unicité du prolongement d'une mesure). *Soit $\mathcal{A}$ une algèbre de Boole sur E et μ et ν deux mesures σ-finies sur $\sigma(\mathcal{A})$ telles que*

(i) pour tout $A \in \mathcal{A}$, $\mu(A) = \nu(A)$,

(ii) il existe $(E_n)_{n\in\mathbb{N}}$ une suite croissante d'éléments de $\mathcal{A}$ telle que $\cup_n E_n = E$ et pour tout $n \in \mathbb{N}$, $\mu(E_n) = \nu(E_n) < +\infty$.

Alors μ et ν coïncident sur $\sigma(\mathcal{A})$.

Démonstration. Supposons dans un premier temps que μ et ν soient finies. On pose

$$\Lambda = \{A \in \sigma(\mathcal{A}) \ ; \ \mu(A) = \nu(A)\}.$$

Montrons que la famille Λ est un λ-système. Puisque $E \in \mathcal{A}$, $E \in \Lambda$. Soit $(A_n)_{n\in\mathbb{N}}$ une suite croissante d'éléments de Λ. Alors

$$\mu(\cup_n A_n) = \lim_{n\to\infty} \mu(A_n) = \lim_{n\to\infty} \nu(A_n) = \nu(\cup_n A_n).$$

Soit A et B dans Λ avec A inclus dans B. Puisque $\mu(E)$ et $\nu(E)$ sont finis, et $B \in \Lambda$, on a $\mu(B) = \nu(B) < +\infty$. On en déduit

$$\mu(B\backslash A) = \mu(B) - \mu(A) = \nu(B) - \nu(A) = \nu(B\backslash A),$$

et donc $B\backslash A$ est dans Λ. Ainsi, Λ est un λ-système qui contient $\mathcal{A}$ donc $\Lambda(\mathcal{A}) \subset \Lambda$. D'autre part, par définition de Λ, $\Lambda \subset \sigma(\mathcal{A})$. Enfin, d'après le théorème des classes monotones, $\Lambda(\mathcal{A}) = \sigma(\mathcal{A})$. On a donc obtenu

$$\Lambda \subset \sigma(\mathcal{A}) = \Lambda(\mathcal{A}) \subset \Lambda,$$

ou encore $\Lambda = \sigma(A)$: les mesures μ et ν coïncident sur la tribu engendrée par $\mathcal{A}$.

Supposons à présent que μ et ν soient uniquement σ-finies et soit $(E_n)_{n\in\mathbb{N}}$ la suite donnée par l'hypothèse (ii). On définit, pour tout $n \in \mathbb{N}$, les mesures μ_n et ν_n sur $\sigma(\mathcal{A})$ par

$$\forall A \in \mathcal{A}, \quad \mu_n(A) = \mu(A \cap E_n) \quad \text{et} \quad \nu_n(A) = \nu(A \cap E_n).$$

Les mesures μ_n et ν_n sont finies et coïncident sur $\mathcal{A}$ donc sur $\sigma(\mathcal{A})$ d'après la première partie de la preuve. Enfin, pour tout $A \in \sigma(\mathcal{A})$, puisque $(A \cap E_n)_{n\in\mathbb{N}}$ une suite croissante d'éléments de $\sigma(\mathcal{A})$,

$$\mu(A) = \lim_{n\to+\infty} \mu_n(A) = \lim_{n\to+\infty} \mu(A \cap E_n) = \lim_{n\to+\infty} \nu(A \cap E_n) = \lim_{n\to+\infty} \mu_n(A) = \nu(A),$$

ce qui achève la preuve. □

Preuve de l'unicité de la mesure produit du le théorème 6.1.3. Supposons qu'il existe une autre mesure m' telle que, pour tous $A \in \mathcal{A}$ et $B \in \mathcal{B}$, on ait $m'(A \times B) = \mu(A)\nu(B)$, c'est-à-dire que m et m' coïncident sur l'ensemble des rectangles. Elles coïncident donc sur l'algèbre de Boole engendrée par les rectangles. Les mesures μ et ν étant σ-finies, il existe deux suites croissantes $(E_n)_{n\in\mathbb{N}}$ et $(F_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{A}$ et $\mathcal{B}$ respectivement telles que, pour tout $n \in \mathbb{N}$, $\mu(E_n)$ et

$\nu(F_n)$ soient finis et $\cup_n E_n = E$ et $\cup_n F_n = F$. La suite de rectangles $(E_n \times F_n)_{n\in\mathbb{N}}$ est telle que $\cup_n(E_n \times F_n) = E \times F$ et

$$m(E_n \times F_n) = \mu(E_n)\nu(F_n) = m'(E_n \times F_n) < +\infty.$$

Les mesures m et m' sont donc σ-finies et coïncident, en vertu du théorème d'unicité 6.1.11, sur la tribu engendrée par l'algèbre de Boole engendrée par les rectangles qui n'est autre que la tribu produit. □

Preuve de l'existence de la mesure produit du théorème 6.1.3. Considérons la fonction suivante :

$$\forall C \in \mathcal{A}\otimes\mathcal{B}, \quad m(C) = \int_E \nu(C_x)\,\mu(dx). \tag{6.2}$$

Dans un premier temps, montrons que cette définition a un sens, c'est-à-dire que $x \mapsto \nu(C_x)$ est une fonction $\mathcal{A}$-mesurable positive (pour que son intégrale par rapport à μ soit définie). Ce résultat nous sera également utile plus tard, nous le mettons donc en exergue ci-dessous.

Lemme 6.1.12. *Si $C \in \mathcal{A}\otimes\mathcal{B}$, la fonction $x \mapsto \nu(C_x)$ est mesurable sur $(E,\mathcal{A})$ et la fonction $y \mapsto \mu(C^y)$ est mesurable sur $(F,\mathcal{B})$.*

Preuve du lemme 6.1.12. Il suffit de démontrer la première assertion. Supposons dans un premier temps que ν soit finie. Soit $\mathcal{C}$ l'ensemble des parties $C \in \mathcal{A}\otimes\mathcal{B}$ telles que $x \mapsto \nu(C_x)$ soit mesurable. Nous allons montrer que $\mathcal{C}$ est un λ-système contenant l'algèbre de Boole $\mathcal{F}$ engendrée par $\mathcal{R}$. Comme le plus petit λ-système contenant $\mathcal{F}$ (qui est stable par intersection finie) est la tribu engendrée par $\mathcal{F}$ (mais aussi par $\mathcal{R}$), c'est donc que $\mathcal{C} = \mathcal{A}\otimes\mathcal{B}$.

Si C est le rectangle $A \times B$, alors $\nu(C_x) = \mathbf{1}_A(x)\nu(B)$ et donc $C \in \mathcal{C}$. Si $C = \cup_{i=1}^n C^i$ où les $(C^i)_{1\le i\le n}$ sont des rectangles deux à deux disjoints, on a $\nu(C_x) = \sum_i \nu(C^i_x)$ et $x \mapsto \nu(C_x)$ est mesurable en tant que somme de fonctions mesurables. D'où $\mathcal{F} \subset \mathcal{C}$.

Montrons à présent que $\mathcal{C}$ est un λ-système. Il est clair que $C = E \times F \in \mathcal{C}$ car alors $C_x = F$ pour tout $x \in E$ et $x \mapsto \nu(F)$ est mesurable. Soit $(C^n)_{n\in\mathbb{N}}$ une suite croissante d'éléments de $\mathcal{C}$ et C sa limite. Pour tout $x \in E$, d'après le théorème de convergence monotone appliqué à la suite croissante $(\mathbf{1}_{C^n_x})_{n\in\mathbb{N}}$, $\nu(C^n_x)$ converge vers $\nu(C_x)$. Donc la fonction $x \mapsto \nu(C_x)$ est mesurable en tant que limite simple d'une suite de fonctions mesurables. Enfin, si C et D sont dans $\mathcal{C}$ avec $C \subset D$ alors $(D\backslash C)_x = D_x\backslash C_x$. Comme ν est supposée finie, $x \mapsto \nu((D\backslash C)_x) = \nu(D_x) - \nu(C_x)$ est mesurable comme différence de fonctions mesurables bornées.

Si ν est seulement σ-finie, soit $(F_n)_{n\in\mathbb{N}}$ une suite croissante d'éléments de $\mathcal{B}$ dont la réunion est F et telle que $\nu(F_n)$ soit fini pour tout $n \in \mathbb{N}$. Soit $C \in \mathcal{A}\otimes\mathcal{B}$ et, pour tout $n \in \mathbb{N}$, posons $C^n = C \cap (E \times F_n)$. D'après la première partie de la démonstration, la fonction $x \mapsto \nu(C^n_x)$ est mesurable. Il en est donc de même pour $x \mapsto \nu(C_x)$ comme limite simple de fonctions mesurables. □

On sait à présent que l'application m est bien définie par (6.2). Montrons qu'il s'agit d'une mesure. Il est clair que $m(\emptyset) = 0$. Soit $(C^n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathcal{A}\otimes\mathcal{B}$ deux à deux disjoints et C leur réunion. On a $C_x = \cup_n C^n_x$ avec $(C^n_x)_{n\in\mathbb{N}}$ deux à deux disjoints dans $\mathcal{B}$, d'où

$\nu(C_n) = \sum_n \nu(C_x^n)$. Mais alors, en utilisant la commutation des signes somme et intégrale pour la suite de fonctions mesurables positives $(\mathbf{1}_{C_x^n})_{n\in\mathbb{N}}$,

$$\int_E \nu(C_x)\,\mu(dx) = \sum_{n\geq 0}\int_E \nu(C_x^n)\,\mu(dx) = \sum_{n\geq 0} m(C^n).$$

Il reste à vérifier que m affecte la mesure souhaitée aux rectangles. Si C est le rectangle $A \times B$, on a

$$m(C) = \int_E \nu(C_x)\,\mu(dx) = \int_E \mathbf{1}_A(x)\nu(B)\,\mu(dx) = \mu(A)\nu(B).$$

De même on montre que $C \mapsto \int_F \mu(C^y)\,\nu(dy)$ définit une mesure sur $\mathcal{A}\otimes\mathcal{B}$ qui coïncide avec m sur les rectangles. Par unicité, cette mesure est égale à m et l'on obtient la relation (6.1). □

Le premier exemple à donner est celui de la mesure de Lebesgue sur $\mathbb{R}^d$ muni de sa tribu borélienne $\mathcal{B}(\mathbb{R}^d) = \mathcal{B}(\mathbb{R})\otimes\cdots\otimes\mathcal{B}(\mathbb{R})$.

Théorème 6.1.13 (Mesure de Lebesgue sur $\mathbb{R}^d$)**.** *Il existe une unique mesure λ_d sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ telle que, pour tout produit d'intervalles $I_1 \times\cdots\times I_d$, $\lambda_d(I_1\times\cdots\times I_d)$ soit égal au produit des longueurs des intervalles (I_j). De plus, λ_d est le produit d fois de la mesure de Lebesgue sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Cette mesure est appelée mesure de Lebesgue sur $\mathbb{R}^d$.*

De plus, la mesure λ_d est l'unique mesure sur $\mathcal{B}(\mathbb{R}^d)$ telle que

1. $\lambda_d([0,1]^d) = 1$,
2. $\forall a \in \mathbb{R}^d, \quad \forall B \in \mathcal{B}(\mathbb{R}^d), \quad \lambda_d(a+B) = \lambda_d(B)$.

Remarque 6.1.14. On pourrait soulever ici le point de rigueur suivant : soit $((E_i, \mathcal{A}_i, \mu_i))_{1\leq i\leq n}$ n espaces mesurés σ-finis. Il existe de nombreuses manières de construire une tribu et une mesure produit sur $E_1\times\cdots\times E_n$ par récurrence. Obtient-on toujours la même ? Rassurons-nous tout de suite : la réponse est oui et la notation $(E_1\times\cdots\times E_n, \mathcal{A}_1\otimes\cdots\otimes\mathcal{A}_n, \mu_1\otimes\cdots\otimes\mu_n)$ a bien un sens (et un seul).

6.2 Théorèmes de Tonelli et Fubini

Remarque 6.2.1. L'égalité (6.1) s'écrit encore

$$\int_{E\times F} \mathbf{1}_C\, d\mu\otimes\nu = \int_E \left[\int_F \mathbf{1}_C(x,y)\nu(dy)\right]\mu(dx) = \int_F \left[\int_E \mathbf{1}_C(x,y)\mu(dx)\right]\nu(dy). \tag{6.3}$$

Ceci signifie que calculer l'intégrale de la fonction indicatrice d'un élément de la tribu produit revient à intégrer l'intégrale des sections, l'ordre des opérations ne jouant pas de rôle.

Comme nous l'avons déjà remarqué lors de la construction de l'intégrale, on est en droit d'espérer que cette propriété s'étende à toutes les fonctions mesurables positives définies sur l'espace produit. Le théorème de Tonelli assure que c'est effectivement le cas.

Théorème 6.2.2 (Tonelli)**.** *Soit f une fonction mesurable de $(E\times F, \mathcal{A}\otimes\mathcal{B})$ dans $\overline{\mathbb{R}}_+$, μ et ν deux mesures σ-finies, respectivement sur $(E, \mathcal{A})$ et $(F, \mathcal{B})$.*

(*i*) *Les fonctions partout définies* $x \mapsto \displaystyle\int_F f(x,y)\,\nu(dy)$ *et* $y \mapsto \displaystyle\int_E f(x,y)\,\mu(dx)$ *sont respectivement* $\mathcal{A}$ *et* $\mathcal{B}$*-mesurables.*

(*ii*) *Les égalités suivantes ont lieu dans* $\overline{\mathbb{R}}_+$ *:*

$$\int_{E\times F} f d\mu\otimes\nu = \int_E\left[\int_F f(x,y)\nu(dy)\right]\mu(dx) = \int_F\left[\int_E f(x,y)\mu(dx)\right]\nu(dy). \tag{6.4}$$

Démonstration. Montrons dans un même temps le point (i) pour la première fonction et la première égalité du point (ii). Le raisonnement se fait comme d'habitude en trois étapes (indicatrice, étagée positive, mesurable positive) et à chaque fois, on doit montrer

(i) $\forall x\in E$, $y\mapsto f(x,y)$ est $\mathcal{B}$-mesurable et positive,

(ii) $x\mapsto \int_F f(x,y)\,\nu(dy)$ est $\mathcal{A}$-mesurable et positive,

(iii) la relation (6.4) est vérifiée par f.

Étape 1. Si f est l'indicatrice d'un élément C de $\mathcal{A}\otimes\mathcal{B}$ alors, (i) est assuré par le lemme 6.1.1 puisque $y\mapsto f(x,y)$ est l'application $y\mapsto \mathbf{1}_{C_x}(y)$, (ii) est assuré par le lemme 6.1.12 et l'égalité (6.4) n'est autre que l'égalité (6.3) qui a été prouvée dans le théorème 6.1.3.

Étape 2. Si f est une fonction étagée positive le résultat découle de la linéarité de l'intégrale et de la stabilité de la mesurabilité par combinaison linéaire.

Étape 3. Si f est mesurable positive, d'après le théorème 3.3.3 d'approximation, il existe une suite $(f_n)_{n\in\mathbb{N}}$ croissante de fonctions étagées positives qui converge simplement vers f. Donc, pour tout $x\in E$, $(y\mapsto f_n(x,y))_{n\in\mathbb{N}}$ est une suite de fonctions mesurables positives qui converge vers $y\mapsto f(x,y)$ qui est donc mesurable positive. Le théorème de convergence monotone assure que

$$\int_F f(x,y)\,\nu(dy) = \int_F \lim_{n\to\infty} f_n(x,y)\,\nu(dy) = \lim_{n\to\infty}\int_F f_n(x,y)\,\nu(dy).$$

La fonction $x\mapsto \int_F f(x,y)\,\nu(dy)$ est donc $\mathcal{A}$-mesurable comme limite de fonctions mesurables. De plus,

$$\begin{aligned}\int_{E\times F} f\,d\mu\otimes\nu &\overset{CM}{=} \lim_n \int_{E\times F} f_n d\mu\otimes\nu \overset{\text{étape2}}{=} \lim_n \int_E\left[\int_F f_n(x,y)\nu(dy)\right]\mu(dx),\\ &\overset{CM}{=} \int_E \lim_n \left[\int_F f_n(x,y)\nu(dy)\right]\mu(dx) \overset{CM}{=} \int_E\left[\int_F f(x,y)\nu(dy)\right]\mu(dx),\end{aligned}$$

ce qui achève la preuve. □

Corollaire 6.2.3. *Soit* f *une fonction mesurable sur* $(E\times F, \mathcal{A}\otimes\mathcal{B})$ *à valeurs complexes. Alors* f *est intégrable pour la mesure* $\mu\otimes\nu$ *si et seulement si l'une des deux conditions suivantes est réalisée :*

$$\int_E\left[\int_F |f(x,y)|\,\nu(dy)\right]\mu(dx) < +\infty \quad \textit{ou} \quad \int_F\left[\int_E |f(x,y)|\,\mu(dx)\right]\nu(dy) < +\infty.$$

Avant d'énoncer un théorème analogue à celui de Tonelli mais adapté aux fonctions non nécessairement positives, donnons quelques précisions. On dit qu'une fonction f est définie presque partout sur un espace mesuré $(E, \mathcal{A}, \mu)$ s'il existe un ensemble négligeable N tel que f soit définie sur $E \backslash N$. On dira alors que f est mesurable pour la tribu induite par $\mathcal{A}$ sur $E \backslash N$. On dira que f est intégrable si $\int_{E \backslash N} |f| \, d\mu < +\infty$, et, dans ce cas, on notera encore $\int_E f \, d\mu$ l'intégrale $\int_{E \backslash N} f \, d\mu$. Cette notation est justifiée par le fait que si g est une fonction mesurable sur $(E, \mathcal{A})$ telle que $g_{|E \backslash N} = f$ alors $f \in \mathcal{L}^1(E \backslash N, \mu_{E \backslash N})$ si et seulement si $g \in \mathcal{L}^1(E, \mu)$ et dans ce cas,

$$\int_{E \backslash N} f \, d\mu = \int_{E \backslash N} g \, d\mu = \int_E g \, d\mu.$$

Théorème 6.2.4 (Fubini). *Soit f une fonction intégrable sur $(E \times F, \mathcal{A} \otimes \mathcal{B}, \mu \otimes \nu)$. Alors,*

(i) pour presque tout $x \in E$, la fonction $y \mapsto f(x, y)$ est dans $\mathcal{L}^1(\nu)$; de plus la fonction $x \mapsto \int_F f(x, y) \, \nu(dy)$, définie μ p.p., est μ-intégrable.

(ii) pour presque tout $y \in F$, la fonction $x \mapsto f(x, y)$ est dans $\mathcal{L}^1(\mu)$; de plus la fonction $y \mapsto \int_E f(x, y) \, \mu(dx)$, définie ν p.p., est ν-intégrable.

(iii) On a enfin

$$\int_{E \times F} f d\mu \otimes \nu = \int_E \left[\int_F f(x, y) \nu(dy) \right] \mu(dx) = \int_F \left[\int_E f(x, y) \mu(dx) \right] \nu(dy).$$

Démonstration. Démontrons par exemple (i) et la première égalité de (iii) pour une fonction à valeurs dans $\mathbb{R}$. D'après le théorème de Tonelli, on a

$$\int_E \left[\int_F |f(x, y)| \nu(dy) \right] \mu(dx) = \int_{E \times F} |f| d\mu \otimes \nu.$$

Il résulte alors de la proposition 5.2.4 que $x \mapsto \int |f(x, y)| \, \nu(dy)$ est finie sauf peut-être sur ensemble μ-négligeable N. Donc si $x \notin N$, la fonction $y \mapsto f(x, y)$ est ν-intégrable. On décompose f en la différence de ses parties positive et négative : $f = f^+ - f^-$. Si $x \notin N$, les fonctions $y \mapsto f^+(x, y)$ et $y \mapsto f^-(x, y)$ sont ν-intégrables et on a

$$\forall x \in N^c, \quad \int f(x, y) \, \nu(dy) = \int f^+(x, y) \, \nu(dy) - \int f^-(x, y) \, \nu(dy).$$

D'après le théorème de Tonelli, les fonctions $x \mapsto \int f^{\pm}(x, y) \, \nu(dy)$ sont mesurables sur $E \backslash N$ muni de la tribu induite et on a

$$\int_{E \backslash N} \left[\int_F f^{\pm}(x, y) \, \nu(dy) \right] \mu(dx) = \int_E \left[\int_F f^{\pm}(x, y) \, \nu(dy) \right] \mu(dx) = \int_{E \times F} f^{\pm} \, d\mu \otimes \nu < +\infty.$$

Par conséquent, $x \mapsto \int_F f(x, y) \nu(dy)$ définie sur $E \backslash N$ est intégrable comme combinaison linéaire de deux fonctions intégrables. On a enfin

$$\begin{aligned}
\int_{E\times F} f\,d\mu\otimes\nu &= \int_{E\times F} f^+\,d\mu\otimes\nu - \int_{E\times F} f^-\,d\mu\otimes\nu \\
&= \int_{E\setminus N}\left[\int_F f^+(x,y)\,\nu(dy)\right]\mu(dx) - \int_{E\setminus N}\left[\int_F f^-(x,y)\,\nu(dy)\right]\mu(dx) \\
&= \int_{E\setminus N}\left[\int_F f^+(x,y)\,\nu(dy) - \int_F f^-(x,y)\,\nu(dy)\right]\mu(dx) \\
&= \int_{E\setminus N}\left[\int_F f(x,y)\,\nu(dy)\right]\mu(dx) = \int_E\left[\int_F f(x,y)\,\nu(dy)\right]\mu(dx),
\end{aligned}$$

ce qui achève la preuve. □

Remarque 6.2.5. Lorsque l'on veut signifier les variables d'intégration dans l'intégrale multiple, on note souvent $\mu(dx)\nu(dy)$ pour $d\mu\otimes\nu$, $\mu\otimes\nu(d(x,y))$ ou encore $\mu\otimes\nu(dx,dy)$, ce qui est justifié par le théorème de Fubini.

6.3 Exemples d'utilisation

6.3.1 Exemples de mesures produit

Soit α et β deux réels strictement positifs et μ et ν les mesures de densités respectives par rapport à la mesure de Lebesgue sur $\mathbb{R}$:

$$\forall x\in\mathbb{R},\quad f_\mu(x) = \alpha e^{-\alpha x}\mathbf{1}_{\{x>0\}},\quad \text{et}\quad f_\nu(x) = \beta e^{-\beta x}\mathbf{1}_{\{x>0\}}.$$

On souhaite déterminer la mesure pour $\mu\otimes\nu$ de l'ensemble $A = \{(x,y)\ ;\ 0<x<y\}$. Il est facile de déterminer les sections horizontale et verticale de A :

$$\forall x\in\mathbb{R},\quad C_x = \begin{cases}]x,+\infty[& \text{si } x>0,\\ \emptyset & \text{sinon,}\end{cases} \quad \text{et}\quad \forall y\in\mathbb{R},\quad C^y = \begin{cases}]0,y[& \text{si } y>0,\\ \emptyset & \text{sinon,}\end{cases}$$

Pour tout $x>0$,

$$\nu(C_x) = \int_{C_x} f_\nu(y)\,\nu(dy) = \int_{]x,+\infty[} \beta e^{-\beta y}\,\lambda(dy) = e^{-\beta x}.$$

Ainsi,

$$\mu\otimes\nu(A) = \int \nu(C_x)\,\mu(dx) = \int_{\mathbb{R}_+} (e^{-\beta x})\alpha e^{-\alpha x}\,\lambda(dx) = \frac{\alpha}{\alpha+\beta}.$$

On a de même $\mu(C^y) = 1-e^{-\alpha y}$ qui redonne bien l'expression de $\mu\otimes\nu(A)$.

6.3.2 Mesure produit de mesures à densité par rapport à la mesure de Lebesgue

Soit μ et ν deux mesures sur $\mathbb{R}$ admettant respectivement pour densité par rapport à la mesure de Lebesgue les fonctions mesurables positives g_μ et g_ν. Soir f mesurable positive de $\mathbb{R}^2$ dans $\mathbb{R}$. Le théorème de Tonelli et la définition d'une mesure à densité assurent que

$$\begin{aligned}\int_{\mathbb{R}^2} f\,d\mu\otimes\nu &= \int\left(\int f(x,y)\,d\mu(x)\right)d\nu(dy) = \int\left(\int f(x,y)g_\mu(x)\,d\lambda(x)\right)g_\nu(y)\,d\lambda(y)\\ &= \int\left(\int f(x,y)g_\mu(x)g_\nu(y)\,d\lambda(x)\right)d\lambda(y) = \int_{\mathbb{R}^2} f(x,y)g_\mu(x)g_\nu(y)\,d(\lambda\otimes\lambda)(x,y).\end{aligned}$$

Ainsi, la mesure produit admet $(x,y)g_\mu(x)g_\nu(y)$ pour densité par rapport à la mesure de Lebesgue sur $\mathbb{R}^2$.

6.3.3 Premiers exemples

Dans les exemples suivants, on cherche l'intégrale de f sur le domaine D inclus dans $\mathbb{R}^2$:

$$I = \int_D f(x,y)\lambda\otimes\lambda(x,y).$$

Exemple 6.3.1. Soit $D = \left\{(x,y)\,/\,x+y\geq 1,\ x^2+y^2\leq 1\right\}$ et $f(x,y) = xy^2$. La fonction f est mesurable et positive sur D. On peut appliquer le théorème de Tonelli pour obtenir que

$$I = \int_0^1\left(\int_{1-y}^{\sqrt{1-y^2}} xy^2dx\right)dy = \int_0^1 (y^3-y^4)dy = \frac{1}{20}.$$

Exemple 6.3.2. Soit $D = \left\{(x,y)\,/\,0\leq x\leq y\leq x+x^2\right\}$ et $f(x,y) = ye^{-x}$. La fonction f est mesurable et positive sur D. On peut appliquer le théorème de Tonelli. Pour tout $x>0$,

$$\int_x^{x+x^2} f(x,y)dy = e^{-x}\left[\frac{y^2}{2}\right]_x^{x+x^2} = \left(\frac{x^4}{2}+x^3\right)e^{-x}.$$

De plus, pour tout $n\in\mathbb{N}$,

$$\int_0^{+\infty} x^ne^{-x}dx = n!$$

donc $I = 18$.

Exemple 6.3.3. Soit $D = [0,1]\times[0,+\infty[$ et $f(x,y) = e^{-y}\sin(2xy)$. On a $|f(x,y)|\leq e^{-y}$ et l'intégrale de la fonction positive $(x,y)\mapsto e^{-y}$ sur D vaut 1. La fonction f est donc intégrable sur D. On peut appliquer le théorème de Fubini. Pour tout $x>0$ et tout $A>0$,

$$\begin{aligned}\int_0^A \sin(2xy)e^{-y}dy &= \left[-\sin(2xy)e^{-y}\right]_0^A + \int_0^A 2x\cos(2xy)e^{-y}dy\\ &= \sin(2xA)e^{-A} + \left[-2x\cos(2xy)e^{-y}\right]_0^A - \int_0^A 4x^2\sin(2xy)e^{-y}dy\\ &= \sin(2xA)e^{-A} + 2x - 2x\cos(2xA)e^{-A} - 4x^2\int_0^A \sin(2xy)e^{-y}dy.\end{aligned}$$

On en déduit que

$$\int_0^A \sin(2xy)e^{-y}dy = \frac{2x + \sin(2xA)e^{-A} - 2x\cos(2xA)e^{-A}}{1+4x^2},$$

puis que

$$\int_0^{+\infty} \sin(2xy)e^{-y}dy = \frac{2x}{1+4x^2}.$$

On a

$$I = \int_0^1 \left(\int_0^{+\infty} f(x,y)dy\right)dx = \int_0^1 \frac{2x}{1+4x^2}dx = \left[\frac{\ln(1+4x^2)}{4}\right]_0^1 = \frac{\ln(5)}{4}.$$

D'autre part,

$$\begin{aligned} I &= \int_0^{+\infty} \left(\int_0^1 f(x,y)dx\right)dy = \int_0^{+\infty} e^{-y}\left[\frac{-\cos(2xy)}{2y}\right]_0^1 dy \\ &= \int_0^{+\infty} e^{-y}\left(\frac{1-\cos(2xy)}{2y}\right)dy = \int_0^{+\infty} e^{-y}\left(\frac{\sin^2(y)}{y}\right)dy. \end{aligned}$$

On peut donc conclure que

$$\int_0^{+\infty} e^{-y}\left(\frac{\sin^2(y)}{y}\right)dy = \frac{\ln(5)}{4}.$$

Exemple 6.3.4. Soit $D = [0,+\infty[\times[0,1]$ et $f(x,y) = \cos(xy)e^{-tx}$ où t est un réel strictement positif fixé. On peut montrer, grâce au théorème de Fubini, que

$$\int_0^{+\infty} \frac{\sin(x)}{x}e^{-tx}dx = \arctan(1/t).$$

6.3.4 Encore des exemples

Exemple 6.3.5. Soit $D = [0,+\infty[\times[a,b]$ avec $0 < a < b$ et f définie sur $\mathbb{R}^2$ par $f(x,y) = e^{-xy}$. La fonction $f\mathbf{1}_D$ est mesurable positive et la mesure de Lebesgue sur $\mathbb{R}^2$ est σ-finie donc

$$\int_0^{+\infty} \frac{e^{-ax} - e^{-bx}}{x}dx = \ln\left(\frac{b}{a}\right).$$

Exemple 6.3.6. Soit f définie par $f(x,y) = 2e^{-2xy} - e^{-xy}$. Est-elle intégrable sur $[0,1]\times\mathbb{R}_+$? Soit $y \in]0,1]$ et $A > 0$.

$$\int_0^A f(x,y)dx = \left[\frac{e^{-2xy} - e^{-xy}}{y}\right]_0^A = \frac{e^{-2Ay} - e^{-Ay}}{y}.$$

Donc

$$\int_0^{+\infty} f(x,y)dx = 0 \quad \text{et} \quad \int_0^1 \left(\int_0^{+\infty} f(x,y)dx\right)dy = 0.$$

D'autre part, pour $x > 0$,

$$\int_0^1 f(x,y)dy = \left[\frac{e^{-2xy} - e^{-xy}}{x}\right]_0^1 = \frac{e^{-2x} - e^{-x}}{x} < 0.$$

L'intégrale de la fonction ci-dessus sur $\mathbb{R}_+$ est donc strictement négative (elle vaut $-\ln(2)$ d'après l'exemple précédent) et le théorème de Fubini ne s'applique pas. Ceci est dû au fait que f n'est pas intégrable sur $[0,1] \times \mathbb{R}_+$.

Exemple 6.3.7. Soit $(E, \mathcal{A}, \mu)$ un espace mesuré σ-fini. Soit $f : E \mapsto \mathbb{R}_+$ mesurable. On a

$$\int_E f d\mu = \int_0^{+\infty} \mu(\{x \in E \text{ t.q. } f(x) > t\})dt.$$

On a, d'après le théorème de Tonelli,

$$\begin{aligned}
\int_0^{+\infty} \mu(\{x \in E \text{ t.q. } f(x) > t\})dt &= \int_0^{+\infty} \left(\int_E \mathbf{1}_{\{x \in E \text{ t.q. } f(x) > t\}} d\mu(x) \right) dt \\
&= \int_E \left(\int_0^{+\infty} \mathbf{1}_{\{t > 0 \text{ t.q. } f(x) > t\}} dt \right) d\mu(x) \\
&= \int_E f(x) d\mu(x).
\end{aligned}$$

6.3.5 Normalisation de la mesure gaussienne

On se propose de montrer que

$$I = \int_0^{+\infty} e^{-x^2/2}\, dx = \sqrt{\frac{\pi}{2}}.$$

Soit f définie sur $\mathbb{R}_+^2$ par $f(x,y) = y \exp(-y^2(1+x^2)/2)$. La fonction f est continue sur $\mathbb{R}_+^2$ donc $\mathcal{B}(\mathbb{R}_+^2)$-mesurable et positive. Le théorème de Tonelli s'applique ici.

$$\int_0^A f(x,y)\, dy = \left[\frac{-\exp(-y^2(1+x^2)/2)}{1+x^2}\right]_0^A = \frac{1}{1+x^2} - \frac{\exp(-A^2(1+x^2)/2)}{1+x^2} \xrightarrow[A \to +\infty]{} \frac{1}{1+x^2}.$$

On a donc

$$\int_{\mathbb{R}_+} \left(\int_{\mathbb{R}_+} f(x,y)\, d\lambda(x)\, d\lambda(y) \right) = \int_{\mathbb{R}_+} \frac{1}{1+x^2}\, d\lambda(x) = \frac{\pi}{2}.$$

D'autre part, avec $y > 0$ fixé,

$$\int_0^A f(x,y)\, dx = e^{-y^2/2} \int_0^A e^{-(yx)^2/2}\, y dx \overset{u=yx}{=} e^{-y^2/2} \int_0^{yA} e^{-u^2/2}\, du \xrightarrow[A \to +\infty]{} I e^{-y^2/2}.$$

On a donc

$$\int_{\mathbb{R}_+} \left(\int_{\mathbb{R}_+} f(x,y)\, d\lambda(x)\, d\lambda(y) \right) = I^2.$$

Le théorème de Tonelli fournit le résultat annoncé.

6.3.6 Transformée de Fourier de la mesure de Cauchy

On souhaite montrer que

$$\int_{\mathbb{R}} e^{itx}\,\frac{1}{\pi}\frac{1}{1+x^2}\,\lambda(dx) = e^{-|t|}.$$

Remarque 6.3.8. La mesure μ de densité $\frac{1}{\pi}\frac{1}{1+x^2}$ par rapport à la mesure de Lebesgue sur $\mathbb{R}$ est appelée mesure de Cauchy. Remarquons que φ_μ n'est pas dérivable en 0 et que la fonction identité n'est pas intégrable par rapport à μ.

Soit $a > 0$. Rappelons que la transformée de Fourier de la fonction $g_a : x \mapsto e^{-a|x|}$ est égale

$$\forall t \in \mathbb{R}, \quad \hat{g}_a(t) = \int_{\mathbb{R}} e^{itx}\, e^{-a|x|}\,\lambda(dx) = \frac{a}{a^2+t^2}.$$

Pour $t \in \mathbb{R}$ fixé, introduisons la fonction continue sur $\mathbb{R}^2$

$$f_t(u,x) = e^{i(u+t)x} e^{-|u|} e^{-a|x|}.$$

La majoration

$$|f_t(u,x)| \le e^{-|u|} e^{-a|x|}$$

assure que f_t est intégrable pour la mesure de Lebesgue λ_2 sur $\mathbb{R}^2$. Le théorème de Fubini assure alors

$$\begin{aligned}
\iint f_t(u,x)\lambda_2(du,dx) &= \int \left[e^{itx} e^{-a|x|} \int e^{ixu} e^{-|u|}\,\lambda(du) \right] \lambda(dx) = \int \frac{e^{itx}}{1+x^2}\, e^{-a|x|}\,\lambda(dx) \\
&= \int \left[e^{-|u|} \int e^{i(t+u)x} e^{-a|x|}\,\lambda(dx) \right] \lambda(du) = \int \frac{a}{a^2+(t+u)^2}\, e^{-|u|}\,\lambda(du).
\end{aligned}$$

On a donc montré que

$$\int_{\mathbb{R}} \frac{e^{itx}}{1+x^2}\, e^{-a|x|}\,\lambda(dx) = \int \frac{a}{a^2+(t+u)^2}\, e^{-|u|}\,\lambda(du).$$

Dans l'intégrale de droite, opérons le changement de variables $z = (t+u)/a$. Il vient

$$\int_{\mathbb{R}} \frac{e^{itx}}{1+x^2}\, e^{-a|x|}\,\lambda(dx) = \int \frac{e^{-|az-t|}}{1+z^2}\,\lambda(dz).$$

Il reste à utiliser le théorème de convergence dominée pour déterminer la limite des deux intégrales lorsque a tend vers 0. Remarquons bien que l'on a les dominations suivantes :

$$\left| \frac{e^{itx}}{1+x^2}\, e^{-a|x|} \right| \le \frac{1}{1+x^2} \quad \text{et} \quad \left| \frac{e^{-|az-t|}}{1+z^2} \right| \le \frac{1}{1+z^2}.$$

On obtient alors

$$\int \frac{e^{itx}}{1+x^2}\,\lambda(dx) = \int \frac{e^{-|t|}}{1+z^2}\,\lambda(dz) = \pi e^{-|t|}.$$

6.4 Mesure image et changement de variables

6.4.1 Mesure image et changement de variables affine

La notion de mesure image joue un rôle central en probabilités. La mesure image, notée μ_φ sur $(F, \mathcal{B})$ par une fonction mesurable h de $(E, \mathcal{A})$ dans $(F, \mathcal{B})$ d'une mesure μ est définie par

$$\forall B \in \mathcal{B}, \quad \mu_\varphi(B) = \mu(\varphi^{-1}(B)).$$

Rappelons également résultat qui relie les intégrales par rapport à ces deux mesures.

Théorème 6.4.1 (Théorème de transfert)**.** *Soit $(E, \mathcal{A})$ et $(F, \mathcal{B})$ deux espaces mesurables et φ une application mesurable de E dans F. Soit μ une mesure sur $(E, \mathcal{A})$. Soit f une fonction mesurable à valeurs complexes définie sur $(F, \mathcal{B})$. Alors f est intégrable par rapport à μ_φ si et seulement si $f \circ \varphi$ est intégrable par rapport à μ. Dans ce cas,*

$$\int_F f \, d\mu_\varphi = \int_E f \circ \varphi \, d\mu.$$

Remarque 6.4.2. Soit $\varphi : (E, \mathcal{A}) \to (E, \mathcal{A})$. Si μ est invariante par h, c'est-à-dire que $\mu_\varphi = \mu$, alors

$$\forall f \in \mathcal{L}^1(\mu), \quad \int_E f \, d\mu = \int_E f \circ \varphi \, d\mu.$$

C'est par exemple le cas pour la mesure de Lebesgue et les fonctions de translation $\tau_a : x \to x+a$.

Cette remarque admet une généralisation importante qui a de nombreuses applications et est à la base de la démonstration du théorème de changement de variables.

Proposition 6.4.3. *Soit $A \in GL_d(\mathbb{R})$ et $b \in \mathbb{R}^d$. Alors*

$$\forall f \in \mathcal{L}^1(\lambda_d), \quad \int_{\mathbb{R}^d} f(Ax + b) \, \lambda(dx) = \frac{1}{|\det A|} \int_{\mathbb{R}^d} f(x) \, \lambda(dx). \tag{6.5}$$

L'égalité est vraie dans $\overline{\mathbb{R}}_+$ pour les fonctions mesurables positives.

Remarque 6.4.4. La formule (6.5) se reformule en termes de mesure image. Soit $A \in GL_d(\mathbb{R})$ et $b \in \mathbb{R}^d$. L'image de la mesure de Lebesgue sur $\mathbb{R}^d$ par l'application $x \mapsto Ax + b$ est la mesure $|\det A|^{-1} \lambda_d$.

Preuve de la proposition 6.4.3. La remarque 6.4.2 montre qu'il suffit de démontrer le résultat lorsque $b = 0$. Il faut donc montrer que la mesure image de λ_d par A est la mesure $|\det A|^{-1} \lambda_d$. Soit ν cette mesure. Remarquons tout d'abord que ν est invariante par translation. Soit $a \in \mathbb{R}^d$ et $B \in \mathcal{B}(\mathbb{R}^d)$,

$$\nu(a + B) = \lambda_d(A^{-1}(a + B)) = \lambda_d(A^{-1}a + A^{-1}(B)) = \lambda_d(A^{-1}(B)) = \nu(B).$$

D'autre part, $\nu([0,1]^d) \neq 0$ car

$$\nu(\mathbb{R}^d) = \lambda_d(\mathbb{R}^d) = +\infty \quad \text{et} \quad \nu(\mathbb{R}^d) \leq \sum_{n \in \mathbb{Z}^d} \nu\Big(n + [0,1]^d\Big) = \sum_{n \in \mathbb{Z}^d} \nu\Big([0,1]^d\Big).$$

Enfin $\nu([0,1]^d) < +\infty$ car $\nu([0,1]^d) = \lambda_d(A^{-1}([0,1]^d))$ et $A^{-1}([0,1]^d)$ est compact comme image du compact $[0,1]^d$ par l'application continue A^{-1}. Ainsi, la mesure $\nu' = \nu/\nu([0,1]^d)$ est invariante par translation et $\nu([0,1]^d) = 1$, il s'agit donc de la mesure de Lebesgue sur $\mathbb{R}^d$. Il existe donc $c \in \mathbb{R}_+^*$ tel que $\nu = c\lambda_d$.

Pour déterminer c, il ne reste donc plus qu'à déterminer $\nu(B)/\lambda_d(B)$ pour un borélien de mesure de Lebesgue non nulle et finie bien choisi. On procède pour cela par étapes selon la forme de A.
- Soit A une matrice orthogonale. Si $A \in \mathcal{O}_d(\mathbb{R}) = \big\{O \in GL_d(\mathbb{R}),\ {}^tOO = I_d\big\}$, il est clair que

$$A^{-1}\big(\big\{x\ ;\ {}^txx \leq 1\big\}\big) = \big\{x\ ;\ {}^txx \leq 1\big\}$$

donc

$$\nu\big(\big\{x\ ;\ {}^txx \leq 1\big\}\big) = \lambda_d\big(\big\{x\ ;\ {}^txx \leq 1\big\}\big) \geq \lambda_d\left(\left[0, 1/\sqrt{d}\right]^d\right) \geq d^{-d/2} > 0.$$

- Soit A une matrice symétrique définie positive. Si $A \in \mathcal{S}_d^+(\mathbb{R}) \cap GL_d(\mathbb{R})$, A est diagonalisable dans le groupe orthogonal : il existe $O \in \mathcal{O}_d(\mathbb{R})$ tel que $A = OD{}^tO$ où $D = \mathrm{Diag}(\alpha_1, \ldots, \alpha_d)$, $\alpha_i \in \mathbb{R}_+^*$. On choisit $B = OD([0,1]^d)$ et on a alors

$$\nu(B) = \lambda_d(A^{-1}(B)) = \lambda_d((OD^{-1t}O)(OD([0,1]^d)) = \lambda_d(O([0,1]^d)) = 1.$$

D'autre part, comme $D([0,1]^d) = \prod_{i=1}^d [0, \alpha_i]$,

$$\lambda_d(B) = \lambda_d\left(O\prod_{i=1}^d [0,\alpha_i]\right) = \lambda_d\left(\prod_{i=1}^d [0,\alpha_i]\right) = \prod_{i=1}^d \alpha_i = \det A.$$

- Soit A une matrice inversible. On peut décomposer A en $A = OS$ avec $O \in \mathcal{O}_d(\mathbb{R})$ et $S \in \mathcal{S}_d(\mathbb{R})$. Il suffit en effet de poser $S = \sqrt{{}^tAA}$ et $O = AS^{-1}$. Grâce aux résultats précédents,

$$\nu(B) = \lambda_d(A^{-1}(B)) = \lambda_d(S^{-1}(O^{-1}(B))) = (\det S)^{-1}\lambda_d((O^{-1}(B)) = (\det S)^{-1}\lambda_d(B).$$

Puisque $\det A = \det O \det S$ et $\det O \in \{-1, 1\}$, on a bien $\nu(B) = |\det A|^{-1}\lambda_d(B)$. □

6.4.2 Théorème général du changement de variables

Soit φ un difféomorphisme de classe $\mathcal{C}^1$ d'un ouvert Δ de $\mathbb{R}^d$ sur un ouvert D de $\mathbb{R}^d$. Si $y \in \Delta$, on note $J_\varphi(y)$ le déterminant jacobien de $\varphi = (\varphi_1, \ldots, \varphi_d)$ en y (c'est-à-dire le déterminant de la matrice jacobienne) :

$$J_\varphi(y) = \det \begin{pmatrix} \dfrac{\partial \varphi_1}{\partial y_1} & \cdots & \dfrac{\partial \varphi_1}{\partial y_d} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial \varphi_d}{\partial y_1} & \cdots & \dfrac{\partial \varphi_d}{\partial y_d} \end{pmatrix}$$

Les ouverts Δ et D sont munis de la mesure de Lebesgue.

Théorème 6.4.5 (Changement de variables)**.** *(i) Soit f une fonction borélienne positive sur D. Alors*

$$\int_D f(x)\,dx = \int_\Delta f(\varphi(y))|J_\varphi(y)|\,dy.$$

(ii) Soit f une fonction borélienne à valeurs complexes sur D. Alors f est intégrable sur D si et seulement si $\int_\Delta |f(\varphi(y))||J_\varphi(y)|\,dy < +\infty$ et dans ce cas,

$$\int_D f(x)\,dx = \int_\Delta f(\varphi(y))|J_\varphi(y)|\,dy.$$

6.4.3 Passage en coordonnées polaires

Passer en coordonnées polaires consiste à considérer le difféomorphisme

$$\begin{array}{rcl} \varphi \ : \ \mathbb{R}_+^* \times]-\pi,\pi[& \to & \mathbb{R}^2 \backslash (\mathbb{R}_- \times \{0\}) \\ (r,\theta) & \mapsto & \varphi(r,\theta) = (r\cos\theta, r\sin\theta). \end{array}$$

La fonction φ est bien un difféomorphisme puisque, d'une part, φ est bijective de réciproque donnée par

$$\varphi^{-1}(x,y) = \left(\sqrt{x^2+y^2}, \mathrm{Arg}(x+iy)\right),$$

où $\mathrm{Arg}(z)$ désigne l'argument principal (dans $[-\pi,\pi[$) de $z \in \mathbb{C}^*$. D'autre part, φ est continûment dérivable :

$$\varphi'(r,\theta) = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix} \quad \text{et} \quad J_\varphi(r,\theta) = r\cos^2\theta + r\sin^2\theta = r.$$

Comme l'ensemble $\mathbb{R}_- \times \{0\}$ est de mesure nulle pour la mesure de Lebesgue sur $\mathbb{R}^2$, pour toute fonction $f \in \mathcal{L}^1_\mathbb{C}(\mathbb{R}^2)$,

$$\int_{\mathbb{R}^2} f(x)\lambda_2(dx) = \int_{\mathbb{R}^2\backslash(\mathbb{R}_-\times\{0\})} f(x)\lambda_2(dx) = \int_0^{2\pi}\int_0^{+\infty} f(r\cos\theta, r\sin\theta) r\,dr\,d\theta,$$

le second membre s'intégrant dans un ordre indifférent d'après le théorème de Fubini.

Exemple 6.4.6 (Encore une fois...)**.** D'après le théorème de Tonelli,

$$\left(\int_0^{+\infty} e^{-x^2}\,dx\right)^2 = \int_0^{+\infty} e^{-x^2}\,dx \int_0^{+\infty} e^{-y^2}\,dy = \int_{\mathbb{R}_+^2} e^{-(x^2+y^2)}\,dx\,dy.$$

On calcule l'intégrale de droite grâce au passage en coordonnées polaires :

$$\int_{\mathbb{R}_+^2} e^{-(x^2+y^2)}\,dx\,dy = \int_0^{\pi/2}\int_0^{+\infty} e^{-r^2} r\,dr\,d\theta = \frac{\pi}{2}\int_0^{+\infty} re^{-r^2}\,dr = \frac{\pi}{2}\left[-\frac{e^{-r^2}}{2}\right]_0^{+\infty} = \frac{\pi}{4}.$$

On retrouve ainsi $\displaystyle\int_0^{+\infty} e^{-x^2}\,dx = \frac{\sqrt{\pi}}{2}$.

Exemple 6.4.7 (Volume de la boule unité)**.** La boule euclidienne unité de $\mathbb{R}^d$ est définie par $B_d = \{(x_1, \ldots, x_d) \in \mathbb{R}^d \ ; \ x_1^2 + \cdots + x_d^2 \leq 1\}$ et son volume v_d est donné par $\lambda_d(v_d)$. On calcule v_d par récurrence. Supposons $d \geq 2$. D'après le théorème de Tonelli,

$$\begin{aligned} v_d &= \int_{\mathbb{R}^d} \mathbf{1}_{\{x_1^2+\cdots+x_d^2 \leq 1\}} \, dx_1 \ldots dx_d \\ &= \int_{\mathbb{R}^2} \left[\int_{\mathbb{R}^{d-2}} \left(\mathbf{1}_{\{x_1^2+\cdots+x_{d-2}^2 \leq 1-(x_{d-1}^2+x_d^2)\}} \right) dx_1 \ldots dx_{d-2} \right] \mathbf{1}_{\{x_{d-1}^2+x_d^2 \leq 1\}} \, dx_{d-1} \, dx_d. \end{aligned}$$

Calculons l'intégrale entre crochet (c'est une fonction de (x_{d-1}, x_d)) en distinguant deux cas.

1. Si $x_{d-1}^2 + x_d^2 = 1$ alors

$$\int_{\mathbb{R}^{d-2}} \mathbf{1}_{\{x_1^2+\cdots+x_{d-2}^2 \leq 1-(x_{d-1}^2+x_d^2)\}} \, dx_1 \ldots dx_{d-2} = \lambda_{d-2}(\{0_d\}) = 0.$$

2. Si $x_{d-1}^2 + x_d^2 < 1$, posons $x_i = u_i \sqrt{1 - (x_{d-1}^2 + x_d^2)}$. Ce changement de variables est une simple homothétie dont le déterminant vaut $\left(\sqrt{1 - (x_{d-1}^2 + x_d^2)} \right)^{d-2}$.

On obtient donc

$$\begin{aligned} v_d &= \int_{\mathbb{R}^2} \left[\int_{B_{d-2}} \left(1 - (x_{d-1}^2 + x_d^2)\right)^{d/2-1} du_1 \ldots du_{d-2} \right] \mathbf{1}_{\{x_{d-1}^2+x_d^2 \leq 1\}} \, dx_{d-1} \, dx_d \\ &= v_{d-2} \int_{\mathbb{R}^2} \left(1 - (x_{d-1}^2 + x_d^2)\right)^{d/2-1} \mathbf{1}_{\{x_{d-1}^2+x_d^2 \leq 1\}} \, dx_{d-1} \, dx_d \\ &= v_{d-2} \int_{-\pi}^{\pi} \int_0^{+\infty} (1 - r^2)^{d/2-1} \mathbf{1}_{\{r^2 \leq 1\}} r \, dr \, d\theta \\ &= 2\pi v_{d-2} \int_0^1 r(1 - r^2)^{d/2-1} \, dr = 2\pi v_{d-2} \left[-\frac{1}{d}(1 - r^2)^{d/2} \right]_0^1 = \frac{2\pi}{d} v_{d-2}. \end{aligned}$$

Comme $v_1 = 2$ et $v_2 = \pi$, on obtient immédiatement le résultat suivant :

$$v_d = \begin{cases} \dfrac{\pi^{d/2}}{(d/2)!} & \text{si } d \text{ est pair,} \\ \\ \dfrac{2^d \pi^{(d-1)2}((d-1)/2)!}{d!} & \text{si } d \text{ est impair.} \end{cases}$$

Chapitre 7

Espaces $\mathcal{L}^p$ et L^p, convolution

Dans toute la suite $\mathbb{K}$ désignera indifféremment le corps des réels ou le corps des complexes.

Définition 7.0.1. Pour tout réel $p > 0$, on définit

$$\mathcal{L}^p_{\mathbb{K}}(E, \mathcal{A}, \mu) = \left\{ f \ : \ (E, \mathcal{A}) \to (\mathbb{K}, \mathcal{B}(\mathbb{K})) \text{ mesurable, } \int_E |f|^p \, d\mu < +\infty \right\}.$$

On utilisera en général la notation $\mathcal{L}^p_{\mathbb{K}}(\mu)$.

Exemple 7.0.2. Si m désigne la mesure de comptage sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$, alors

$$\mathcal{L}^p_{\mathbb{K}}(m) = l^p_{\mathbb{K}}(\mathbb{N}) = \left\{ (a_n)_{n\in\mathbb{N}}, \ \sum_{n=0}^{\infty} |a_n|^p < +\infty \right\}.$$

Proposition 7.0.3. *Pour tout p, $\mathcal{L}^1_{\mathbb{K}}(\mu)$ est un $\mathbb{K}$-e.v.*

Démonstration. On vérifie que $\mathcal{L}^p_{\mathbb{K}}(\mu)$ est un s.e.v. du $\mathbb{K}$- e.v. des fonctions mesurables de E dans $\mathbb{K}$. Il est immédiat que la fonction nulle appartient à $\mathcal{L}^p_{\mathbb{K}}(\mu)$. Soit $\lambda \in \mathbb{K}$ et $f, g \in \mathcal{L}^p_{\mathbb{K}}(\mu)$. Les majorations

$$|\lambda f + g|^p \leq (|\lambda||f| + |g|)^p \leq (2\max(|\lambda||f|, |g|))^p \leq 2^p|\lambda|^p|f|^p + 2^p|g|^p$$

assurent que $\lambda f + g$ est μ-intégrable. □

Proposition 7.0.4. *1. Si $\mu(E) < +\infty$, alors*

$$0 < p \leq q \quad \Longrightarrow \quad \mathcal{L}^q_{\mathbb{K}}(\mu) \subset \mathcal{L}^p_{\mathbb{K}}(\mu).$$

2. Si m est la mesure de comptage sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$, alors

$$0 < p \leq q \quad \Longrightarrow \quad l^p_{\mathbb{K}}(\mathbb{N}) \subset l^q_{\mathbb{K}}(\mathbb{N}).$$

Démonstration. 1) Si $0 < p \leq q$, alors $|f|^p \leq |f|^q \mathbf{1}_{\{|f|\geq 1\}} + \mathbf{1}_{\{|f|\leq 1\}}$. Ainsi, dès que $f \in \mathcal{L}^p_{\mathbb{K}}(\mu)$,

$$\int |f|^p \, d\mu \leq \int |f|^q \, d\mu + \mu(\{|f| \leq 1\}) < +\infty.$$

2) Si $0 < p \leq q$, et $(a_n)_{n\in\mathbb{N}} \in l^p_{\mathbb{K}}(\mathbb{N})$ alors $\lim_n a_n = 0$. Il existe donc $n_0 \in \mathbb{N}$ tel que pour tout $n \geq n_0$, $|a_n| \leq 1$. Ainsi, pour tout $n \geq n_0$, $|a_n|^q \leq |a_n|^p$, d'où $\sum_{n\geq 0} |a_n|^q < +\infty$. □

Remarque 7.0.5. Sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ muni de la mesure de Lebesgue λ, on remarque que

$$x \mapsto \frac{\mathbf{1}_{]0,1]}(x)}{\sqrt{x}} \in \mathcal{L}^1(\lambda) \backslash \mathcal{L}^2(\lambda) \quad \text{et} \quad x \mapsto \frac{1}{\sqrt{x^2+1}} \in \mathcal{L}^2(\lambda) \backslash \mathcal{L}^1(\lambda).$$

Il n'existe donc pas d'inclusion entre $\mathcal{L}^1(\lambda)$ et $\mathcal{L}^2(\lambda)$.

Pour toute fonction $f : (E, \mathcal{A}) \to \mathbb{K}$ et pour tout réel $p > 0$, on définit la quantité

$$\|f\|_p = \left(\int_E |f|^p \, d\mu \right)^{1/p} \in \overline{\mathbb{R}}_+,$$

avec la convention $(+\infty)^{1/p} = +\infty$.

Théorème 7.0.6 (Inégalité de Hölder). *Soient $f, g : E \to \mathbb{K}$ et $p, q > 1$ tels que $1/p + 1/q = 1$ (on dit que p et q sont conjugués).*

1. *Si f et g sont à valeurs dans $\mathbb{R}_+$ alors (dans $\overline{\mathbb{R}}_+$)*

$$0 \leq \int fg \, d\mu \leq \|f\|_p \|g\|_q.$$

   En outre, lorsque $\|f\|_p$ et $\|g\|_q$ sont finis, l'inégalité est une égalité si et seulement si il existe $(\alpha, \beta) \in \mathbb{R}^2_+ \backslash \{(0,0)\}$ tel que $\alpha f^p = \beta g^q$ μ-p.p.

2. *Si $f \in \mathbb{L}^p_{\mathbb{K}}(\mu)$ et $g \in \mathbb{L}^q_{\mathbb{K}}(\mu)$ alors $fg \in \mathbb{L}^1_{\mathbb{K}}(\mu)$ et*

$$\|fg\|_1 \leq \|f\|_p \|g\|_q.$$

   En outre, l'inégalité est une égalité si et seulement si il existe $(\alpha, \beta) \in \mathbb{R}^2_+ \backslash \{(0,0)\}$ tel que $\alpha|f|^p = \beta|g|^q$ μ-p.p.

Démonstration. Commençons par établir une inégalité utile dans la suite. On pose pour tout $\alpha \in]0,1[$ et tout $x \in \mathbb{R}_+$, $\varphi_\alpha(x) = x^\alpha - \alpha x$. La fonction φ_α est dérivable sur $\mathbb{R}^*_+$ et $\varphi'_\alpha(x) = \alpha(x^{\alpha-1} - 1)$. Par suite, $\varphi'_\alpha < 0$ sur $]1, +\infty[$ et $\varphi'_\alpha > 0$ sur $]0, 1[$. Donc, pour tout $x \in \mathbb{R}_+$, $\varphi_\alpha(x) \leq \varphi_\alpha(1)$ avec égalité ssi $x = 1$. En reformulant, $x^\alpha \leq \alpha x + 1 - \alpha$ avec égalité ssi $x = 1$. En posant $x = u/v$ avec $u \geq 0$ et $v > 0$, il vient

$$u^\alpha v^{1-\alpha} \leq \alpha u + (1-\alpha)v \quad \text{avec égalité ssi } u = v. \tag{7.1}$$

Remarquons que cette inégalité est encore vraie pour $u, v \in \mathbb{R}_+$.

Revenons à présent à la preuve de l'inégalité de Hölder. Si $\|f\|_p$ ou $\|g\|_q$ est nulle, alors f ou g est nulle μ-p.p. et il en est de même pour fg. Dans ce cas, l'inégalité est triviale. C'est encore le cas si $\|f\|_p$ ou $\|g\|_q$ vaut $+\infty$. Supposons donc que ces deux quantités sont strictement positives et finies. On pose alors

$$\alpha = \frac{1}{p}, \quad \text{d'où } 1-\alpha = \frac{1}{q}, \quad u = \frac{f^p}{\|f\|_p^p} \quad \text{et} \quad v = \frac{g^q}{\|g\|_q^q}.$$

D'après l'inégalité (7.1),

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

En intégrant cette relation par rapport à μ, il vient

$$0 \leq \int fg\,d\mu \leq \|f\|_p\|g\|_q\left(\frac{1}{p}\int \frac{f^p}{\|f\|_p^p}\,d\mu + \frac{1}{q}\int \frac{g^q}{\|g\|_q^q}\,d\mu\right) = \|f\|_p\|g\|_q.$$

L'égalité a lieu ssi $f/\|f\|_p = g/\|g\|_q$ μ-p.p. □

Corollaire 7.0.7. *Si μ est une mesure de probabilité, l'application $r \mapsto \|f\|_r$ est croissante.*

Corollaire 7.0.8 (Inégalité de Cauchy-Schwarz)**.**

$$\left|\int fg d\mu\right| \leq \|f\|_2\|g\|_2.$$

Théorème 7.0.9 (Inégalité de Minkowski)**.** *Si $p \geq 1$, alors, pour tous $f, g \in \mathcal{L}^p_{\mathbb{K}}(\mu)$,*

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p.$$

L'égalité a lieu ssi

- *$f = 0$ μ-p.p. ou $g = \alpha f$ μ-p.p., pour un $\alpha \geq 0$ si $p \geq 1$.*
- *$f = 0$ μ-p.p. ou $f\overline{g} \geq 0$ μ-p.p. si $p = 1$.*

Démonstration. Si $\|f+g\|_p = 0$, l'inégalité est triviale. Sinon, on intègre par rapport à μ l'inégalité

$$|f+g|^p \leq |f||f+g|^{p-1} + |g||f+g|^{p-1} \quad \text{avec la convention } x^0 = 1 \text{ pour } x \geq 0.$$

On obtient alors

$$\|f+g\|_p^p \leq \int |f||f+g|^{p-1}\,d\mu + \int |g||f+g|^{p-1}\,d\mu.$$

Si $p = 1$, l'inégalité est établie. Sinon, l'inégalité de Hölder assure que (puisque $(p-1)q = p$)

$$\int |f||f+g|^{p-1}\,d\mu \leq \|f\|_p\left(\int |f+g|^{(p-1)q}\,d\mu\right)^{1/q} = \|f\|_p\|f+g\|_p^{p/q}.$$

Ainsi,

$$\|f+g\|_p^p \leq \left(\|f\|_p + \|g\|_p\right)\|f+g\|_p^{p/q}.$$

Il ne reste plus qu'à simplifier par $\|f+g\|_p^{p/q}$ qui est strictement positif et à remarquer que $p - p/q = 1$ pour obtenir l'inégalité souhaitée. □

Remarque 7.0.10. Ainsi, $\|\cdot\|_p$ est une semi-norme sur l'espace $\mathcal{L}^p_{\mathbb{K}}(\mu)$. Pour que ce soit une norme, il faudrait que $\|f\|_p = 0$ implique $f = 0$, ce qui est faut puisque la nullité de $\|f\|_p$ implique seulement $f = 0$ μ-p.p.

Il existe une façon simple de construire un espace vectoriel normé à partir de $\mathcal{L}^p$ et $\|\cdot\|_p$: il suffit de quotienter $\mathcal{L}^p$ par la relation d'équivalence $f \sim g$ ssi $f = g$ μ-p.p.

Théorème 7.0.11. *Soit $p \in [1, \infty]$et $L^p_{\mathbb{K}}(\mu) = \mathcal{L}^p_{\mathbb{K}}(\mu)/\sim$. L'espace $L^p_{\mathbb{K}}(\mu)$ muni de l'application $\|\cdot\|_p$ est un $\mathbb{K}$-espace vectoriel normé.*

On commet l'abus bien pratique d'identifier une fonction à sa classe d'équivalence.

Théorème 7.0.12 (Fisher Riesz)**.** *Soit $p \in [1, \infty]$. L'espace $\mathcal{L}^p(\mu)$ est un espace vectoriel normé complet pour la norme $\|\cdot\|_p$. De plus, si $(f_n)_n$ est une suite de $L^p(\mu)$ qui converge vers f pour la norme $\|\cdot\|_p$, alors il existe une suite extraite $(f_{n_k})_k$ qui converge μ-p.p. vers f.*

Démonstration. Soit $(f_n)_n$ une suite de Cauchy dans $L^p(\mu)$. Il existe une sous-suite $(f_{n_k})_k$ telle que $\left\|f_{n_k} - f_{n_{k+1}}\right\|_p < 2^{-k-1}$. Posons

$$g_k = \sum_{i=0}^{k} \left|f_{n_{i+1}} - f_{n_i}\right| \quad \text{et} \quad g = \sum_{i=0}^{\infty} \left|f_{n_{i+1}} - f_{n_i}\right|.$$

Pour tout k, $\|g_k\|_p < 1$ et $(|g_k|^p)_k$ est une suite croissante de fonctions positives qui tend vers $|g|^p$. Par convergence monotone, $\|g\|_p \leq 1$. En particulier, g est finie μ-p.p. Cela signifie que la série de terme général $f_{n_{k+1}}(x) - f_{n_k}(x)$ est convergente pour presque tout x. On pose alors

$$f(x) = f_{n_0}(x) + \sum_{k=0}^{\infty} \left(f_{n_{k+1}} - f_{n_k}\right)(x).$$

Là où la série ne converge pas absolument, on peut poser f égale à 0. En particulier, f est mesurable en tant que limite p.p. de fonctions mesurables.

Il reste à montrer que $f \in L^p(\mu)$ et que $f = \lim_k f_{n_k} = \lim_n f_n$ au sens de la norme $\|\cdot\|_p$. Puisque $|f| \leq |f_{n_0}| + |g|$, on a

$$\|f\|_p \leq \|f_{n_0}\|_p + \|g\|_p < \infty.$$

Posons enfin

$$g_k^{(l)} = \sum_{i=l}^{l+k} \left|f_{n_{i+1}} - f_{n_i}\right| \quad \text{et} \quad g^{(l)} = \sum_{i=l}^{\infty} \left|f_{n_{i+1}} - f_{n_i}\right|.$$

Alors $\left\|g_k^{(l)}\right\|_p < 2^{-l}$, d'où, par convergence monotone, $\left\|g^{(l)}\right\|_p \leq 2^{-l}$. Or, $|f - f_{n_l}| \leq g^{(l)}$, d'où

$$\|f - f_{n_l}\|_p \leq \left\|g^{(l)}\right\|_p \leq 2^{-l} \xrightarrow[n\to\infty]{} 0.$$

Ainsi, f_{n_k} converge vers f en $\|\cdot\|_p$. Comme $(f_n)_n$ est une suite de Cauchy en $\|\cdot\|_p$, elle converge vers la même limite. $\square$

Corollaire 7.0.13. *Soit $(f_n)_n$ une suite de $L^p(\mu)$ telle que la série de terme général f_n converge normalement en $\|\cdot\|_p$, c'est-à-dire $\sum_{n=0}^{\infty} \|f_n\|_p < \infty$. Alors la série converge dans $L^p(\mu)$, c'est-à-dire qu'il existe $f \in L^p(\mu)$ telle que*

$$\lim_k \left\| f - \sum_{n=0}^{k} f_n \right\|_p = 0.$$

7.1 Convolution

Définition 7.1.1. Soit $f, g : \mathbb{R}^d \mapsto \mathbb{C}$ mesurables. On pose, formellement,

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

là où l'intégrale est définie, et on appelle $f * g$ le produit de convolution de f et g.

Proposition 7.1.2. *Soit p et q conjugués. Si $f \in L^p(\mathbb{R}^d)$ et $g \in L^q(\mathbb{R}^d)$ alors*

$$\left| \int_{\mathbb{R}^d} f(x-y)g(y)dy \right| \leq \|f\|_p \|g\|_q$$

donc $(f * g)(x)$ existe et

$$\|f * g\|_\infty \leq \|f\|_p \|g\|_q.$$

Démonstration. Soit x fixé. On applique l'inégalité de Hölder à $y \mapsto f(x-y)$ et $y \mapsto g(y)$:

$$\left| \int_{\mathbb{R}^d} f(x-y)g(y)dy \right| \leq \left(\int_{\mathbb{R}^d} |f(x-y)|^p dy \right)^{1/p} \left(\int_{\mathbb{R}^d} |g(y)|^q dy \right)^{1/q}$$

puis on effectue le changement de variables $z = x - y$ (à x fixé) pour obtenir

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

ce qui fournit le résultat. □

Proposition 7.1.3. *Si $f, g \in L^1(\mathbb{R}^d)$, alors le théorème de Fubini, $(f * g)(x)$ est défini pour presque tout $x \in \mathbb{R}^d$ et*

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

Démonstration. On a

$$\int_{\mathbb{R}^d} |(f * g)(x)| dx \leq \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |f(x-y)||g(y)| dy \right) dx.$$

D'après le théorème de Tonelli, on a encore

$$\int_{\mathbb{R}^d} |(f * g)(x)| dx \leq \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |f(x-y)| dx \right) |g(y)| dy \leq \|f\|_1 \|g\|_1.$$

Ainsi, la fonction $f * g$ est intégrable. Elle est donc finie presque sûrement et l'inégalité demandée est obtenue. □

Exemple 7.1.4. Si $f = \mathbf{1}_{[0,1]}$ alors

$$f * f(x) = \int_{\mathbb{R}} \mathbf{1}_{[0,1]}(x-y)\mathbf{1}_{[0,1]}(y)dy = \int \mathbf{1}_{[0,1]\cap[x-1,x]}(y)dy = \begin{cases} x & \text{si } x \in [0,1], \\ 2-x & \text{si } x \in [1,2], \\ 0 & \text{sinon.} \end{cases}$$

Exemple 7.1.5. Pour $\alpha > 0$, posons

$$f_\alpha(x) = x^{\alpha-1}e^{-x}\mathbf{1}_{\mathbb{R}^+}(x).$$

On a, pour $\alpha, \beta > 0$ et $x > 0$

$$\begin{aligned} f_\alpha * f_\beta(x) &= \int_{\mathbb{R}} (x-y)^{\alpha-1}e^{-(x-y)}\mathbf{1}_{\mathbb{R}^+}(x-y)y^{\beta-1}e^{-y}\mathbf{1}_{\mathbb{R}^+}(y)dy \\ &= e^{-x}\int_{[0,x]} (x-y)^{\alpha-1}y^{\beta-1}dye^{-x} \\ &= \left(\int_{[0,1]} (1-u)^{\alpha-1}u^{\beta-1}du\right) f_{\alpha+\beta}(x). \end{aligned}$$

Exemple 7.1.6. Posons

$$g(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}.$$

On a alors

$$(g*g)(x) = \int_{\mathbb{R}} e^{-(x-y)^2/2}e^{-y^2/2}\frac{dy}{2\pi} = e^{-x^2/4}\int_{\mathbb{R}} e^{-(y-x/2)^2}\frac{dy}{2\pi} = \frac{1}{\sqrt{4\pi}}e^{-x^2/4}.$$