Théorie de la Mesure – MATHF-3001
Syllabus / draft

Mitia Duerinckx

Université Libre de Bruxelles, Année 2025–2026

Table des matières

Introduction

L'analyse est consacrée à l'étude des fonctions. En pratique, de nombreux phénomènes sont modélisés par des fonctions, souvent caractérisées par des équations (par exemple des équations différentielles ou aux dérivées partielles), que l'on ne peut généralement pas résoudre explicitement. Dans ce cadre, l'objectif de l'analyse est d'établir des résultats généraux reliant les propriétés des solutions aux propriétés des équations, sans avoir besoin de connaître ces solutions de manière explicite.

Jusqu'à présent, les cours de CDI-1 & 2 ont développé les outils classiques pour l'étude des fonctions, avec la théorie du calcul différentiel et intégral telle qu'elle fut développée jusqu'à la fin du XIXe siècle. Dans ce cours-ci, nous abordons les innovations fondamentales initiés par Henri Lebesgue au début du XXe siècle. Son but initial, dans sa thèse de 1901, était de mieux comprendre et généraliser l'intégrale de Riemann ; mais la théorie qu'il a mise en place constitue aujourd'hui le socle de toute l'analyse réelle moderne.

Pour comprendre cette évolution, rappelons d'abord quelques limitations importantes de la théorie d'intégration de Riemann.

Limites et intégrales de Riemann

On rappelle le résultat habituel de passage à la limite dans l'intégrale de Riemann :

> Si f_n est bornée et Riemann-intégrable sur $[0, 1]$ et si $f_n \to f$ uniformément, alors f est Riemann-intégrable et on a $\int_0^1 f_n \to \int_0^1 f$.

Ce résultat de convergence est satisfaisant mais exige de la convergence uniforme, qui est souvent une hypothèse trop forte. Si l'on remplace par de la convergence simple, le résultat prend la forme moins agréable suivante :

> Si f_n est bornée et Riemann-intégrable sur $[0, 1]$, si $f_n \to f$ simplement, *et si f est Riemann-intégrable,* alors on a $\int_0^1 f_n \to \int_0^1 f$.

Dans ce résultat, on ne peut en général pas enlever l'hypothèse que la limite ponctuelle f soit elle-même Riemann-intégrable : par exemple, pour une énumération $\mathbb{Q} \cap [0, 1] = \{q_n\}_n$, la suite $f_n := \mathbb{1}_{\{q_1,\dots,q_n\}}$ satisfait $f_n \to \mathbb{1}_{\mathbb{Q}\cap[0,1]}$ simplement et chaque f_n est Riemann-intégrable, mais la limite $\mathbb{1}_{\mathbb{Q}\cap[0,1]}$ ne l'est pas. Ceci montre que la théorie de Riemann se comporte mal avec le passage à la limite (hormis convergence uniforme). Nous verrons que la théorie de Lebesgue permet de résoudre entièrement ces problèmes.

Espace L^2

L'espace $L^2(0, 1)$ est le cadre naturel pour étudier les séries de Fourier, et la bonne définition de cet espace fut un enjeu crucial pour la mécanique quantique naissante au début du XXe siècle. Avec la théorie de Riemann, l'espace

$$L^2_{\text{Riemann}}(0, 1) = \left\{ f : [0, 1] \to \mathbb{R} \text{ Riemann-intégrable, t.q. } \int_0^1 |f|^2 < \infty \right\}$$

n'est pas complet : en effet, pour la même suite $(f_n)_n$ construite ci-dessus, on obtient $\int_0^1 |f_n - f_m|^2 = 0$ pour tout n, m, mais la limite $1_{\mathbb{Q}\cap[0,1]}$ n'est pas Riemann-intégrable et n'appartient donc pas à l'espace. Avec l'intégrale de Lebesgue, l'espace $L^2(0,1)$ devient complet — c'est ce qu'on appellera un espace de Hilbert, fondant le point de départ de l'analyse fonctionnelle moderne.

Théorie des probabilités

En probabilités, on souhaite typiquement ne plus intégrer sur un intervalle, mais sur un espace abstrait d'événements. Sans une théorie d'intégration adaptée, la formulation de la théorie est problématique : dans le calcul d'espérances, par exemple, on doit systématiquement séparer dans la définiton le cas de variables aléatoires dites discrètes ou absolument continues :

$$\mathbb{E}[X] = \begin{cases} \sum_t t\mathbb{P}[X=t] & : \text{ si } X \text{ est discrète,} \\ \int_{\mathbb{R}} t f_X(t)dt & : \text{ si } X \text{ est absolument continue avec densité } f_X. \end{cases}$$

Une formulation unifiée est possible via la théorie d'intégration de Riemann-Stieltjes, mais on sent que la théorie manque de généralité. Kolmogorov a reconnu que la théorie de Lebesgue offrait un cadre naturel pour formaliser rigoureusement la notion de probabilité, fondant ainsi le cadre moderne pour la théorie des probabilités.

Riemann-intégrabilité trop restrictive

Afin de calculer l'intégrale d'une fonction $f : [0,1] \to \mathbb{R}$, interprétée comme l'aire sous son graphe, la théorie de Riemann procède en découpant l'axe des abscisses en petits intervalles sur lesquelles f est approchée par une constante (par son infimum ou supremum). Ceci revient à encadrer f par des fonctions en escalier :

$$\sum_{a\in\epsilon\mathbb{Z}\cap[0,1)} \Big(\inf_{[a,a+\epsilon)} f\Big) \mathbb{1}_{[a,a+\epsilon)} \le f \le \sum_{a\in\epsilon\mathbb{Z}\cap[0,1)} \Big(\sup_{[a,a+\epsilon)} f\Big) \mathbb{1}_{[a,a+\epsilon)}.$$

Pour ces fonctions en escalier, l'intégrale est naturellement définie en sommant l'aire des rectangles (sommes de Darboux). Lorsque le maillage ε tend vers 0, on obtient, si tout se passe bien, l'aire sous le graphe de f, et l'on définit

$$\int_0^1 f := \lim_{\varepsilon} \varepsilon \sum_{a\in\epsilon\mathbb{Z}\cap[0,1)} \inf_{[a,a+\epsilon)} f = \lim_{\varepsilon} \varepsilon \sum_{a\in\epsilon\mathbb{Z}\cap[0,1)} \sup_{[a,a+\epsilon)} f.$$

Pour que ces deux limites existent et coïncident, il est essentiel que la fonction f soit “suffisamment continue”. Par exemple, la fonction $\mathbb{1}_{\mathbb{Q}\cap[0,1]}$ échoue spectaculairement : elle est discontinue en tout pointtu. Un résultat fondamental (Lebesgue-Vitali) affirme en fait qu'une fonction est en fait Riemann-intégrable si et seulement si l'ensemble de ses points de discontinuité est “de mesure nulle”. Cela montre à quel point la continuité — ou plutôt l'absence de “trop” de discontinuités — est structurelle dans la théorie de Riemann. Cet aspect disparaît entièrement dans la théorie de Lebesgue, qui autorise l'intégration de fonctions beaucoup plus irrégulières.

L'idée de Lebesgue

Lebesgue propose de changer de point de vue : au lieu de découper l'axe des abscisses, on découpe l'axe des ordonnées. Pour chaque “tranche” de valeurs, on considère sa préimage par la fonction f à intégrer. Au lieu d'encadrer f par des fonctions en escalier, cette approche nous conduit à l'approximation suivante en tranches horizontales,

$$\sum_{b\in\epsilon\mathbb{Z}} b\mathbb{1}_{f^{-1}([b,b+\epsilon))} \le f \le \sum_{b\in\epsilon\mathbb{Z}} (b+\epsilon)\mathbb{1}_{f^{-1}([b,b+\epsilon))} = \sum_{b\in\epsilon\mathbb{Z}} b\mathbb{1}_{f^{-1}([b,b+\epsilon))} + \varepsilon. \tag{0.1}$$

Pour une fonction f continue, les ensembles de niveau $f^{-1}([b, b+\epsilon))$ sont des unions finies d'intervalles. L'approximation obtenue revient donc, dans ce cas, à retrouver des fonctions en escalier ; cependant, on observe que la partition induite en abscisses n'est plus fixée a priori, mais s'adapte automatiquement à la géométrie de f, devenant plus fine là où la pente de f est plus grande. Si f n'est pas continue, les ensembles de niveau peuvent devenir beaucoup plus irréguliers : l'approximation ainsi construite ne correspond alors plus à une simple superposition de rectangles. On parlera d'approximation par "fonctions simples". Cet encadrement de f suggère de définir

$$\int_0^1 f := \lim_{\varepsilon} \sum_{b \in \epsilon\mathbb{Z}} b \int_0^1 \mathbb{1}_{f^{-1}([b,b+\epsilon))}, \tag{0.2}$$

à condition de savoir de donner un sens à l'intégrale des indicatrices $\mathbb{1}_{f^{-1}([b,b+\epsilon))}$, c'est-à-dire à la "longueur" de $f^{-1}([b, b+\epsilon))$. Or, pour une fonction arbitraire, ces ensembles de niveaux peuvent être fracturés, épars, voire très pathologiques, de sorte que leur attribuer une "longueur" n'a rien d'évident. C'est précisément pour répondre à cette question que la *théorie de la mesure* a été développée. Elle fournit une notion robuste de "longueur" valable pour des ensembles bien plus généraux que des intervalles.

Une fois cette notion définie, l'intégrale qui en découle sera l'intégrale de Lebesgue. On remarque que, puisqu'elle est basée sur le découpage de l'axe des ordonnées et qu'elle requiert seulement que l'espace de départ soit muni d'une mesure de longueur, cette intégrale de Lebesgue se généralisera naturellement à des fonctions définies sur des espaces abstraits.

L'encadrement (0.1) garantit par ailleurs automatiquement l'existence de la limite (0.2), à la différence de la construction de Riemann où l'égalité des limites constituait un point délicat très restrictif. Contrairement à l'intégrale de Riemann, l'intégrale de Lebesgue ne requiert ainsi aucune hypothèse de continuité : elle repose uniquement sur notre capacité à mesurer des ensembles de niveau, ce qui permettra d'intégrer des fonctions beaucoup plus générales. On entrevoit ici la puissance de l'approche de Lebesgue, qui donnera accès en particulier à des théorèmes de convergence d'une puissance incomparable, résolvant les limitations structurelles de la théorie de Riemann.

Chapitre 1

Mesures et intégration abstraite

Ce chapitre est dédié à la définition de l'intégrale de Lebesgue et à ses propriétés. Comme expliqué, cette notion d'intégrale nécessite de définir avant tout une notion de mesure, qu'on introduira de façon axiomatique générale. La construction particulière de la mesure dite de Lebesgue, qui généralisera la notion de longueur sur $\mathbb{R}$ (ou d'aire sur $\mathbb{R}^2$, volume sur $\mathbb{R}^3$, etc.), est postposée au chapitre suivant. Dans le cadre général de l'intégration abstraite de Lebesgue par rapport à une mesure, on établira trois théorèmes de convergence majeurs : le théorème de convergence monotone, le théorème de convergence dominée, et le lemme de Fatou. La simplicité de ces trois résultats de convergence démontre la supériorité de la théorie d'intégration de Lebesgue.

1.1 Une approche axiomatique

Pour définir une notion de mesure de longueur sur $\mathbb{R}$, on adopte une approche axiomatique. Il s'agit tout d'abord de préciser ce que l'on veut entendre par "mesure de longueur" $\mu : \mathcal{P}(\mathbb{R}) \to [0, \infty]$. Il semble naturel d'imposer les conditions suivantes :

— *additivité :* $\mu(A \cup B) = \mu(A) + \mu(B)$ pour tous $A, B \subset \mathbb{R}$ disjoints ;

— *compatibilité avec longueur usuelle :* $\mu([a, b]) = b - a$ pour tous $a \leq b$;

— *invariance par translation :* $\mu(a + A) = \mu(A)$ pour tous $A \subset \mathbb{R}$ et $a \in \mathbb{R}$.

Comme l'intégration repose sur des processus limites, il est naturel de renforcer l'additivité en σ-additivité :

— *σ-additivité :* $\mu(\cup_n A_n) = \sum_n \mu(A_n)$ pour toute suite $(A_n)_n \subset \mathcal{P}(\mathbb{R})$ de parties deux à deux disjointes.

Cependant, on verra qu'il est impossible de définir une telle application μ satisfaisant ces propriétés (voir le contrexemple de Vitali, sous l'axiome du choix). Pour contourner cette impossibilité, il faut renoncer à mesurer toutes les parties de $\mathbb{R}$: au lieu de considérer μ comme définie sur tout $\mathcal{P}(\mathbb{R})$, on va la définir seulement sur un sous-ensemble strict de parties. Avant d'introduire la notion de mesure, cela nous conduit à définir d'abord la notion de σ-algèbres, qui seront les domaines de définition des mesures.

1.2 σ-algèbres

Définition 1.1. Soit X un ensemble. Une collection $\mathcal{A} \subset \mathcal{P}(X)$ de sous-ensembles de X est une *σ-algèbre* sur X (ou tribu) si

— $X \in \mathcal{A}$,

— pour tout $A \in \mathcal{A}$, on a $A^c \in \mathcal{A}$,
— pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{A}$, on a $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

Dans ce cas, on dit que $(X, \mathcal{A})$ est un *espace mesurable.*

Lemme 1.2. *Soit* $(X, \mathcal{A})$ *un espace mesurable.*

(i) $\varnothing \in \mathcal{A}$

(ii) $A_1, \dots, A_N \in \mathcal{A}$, $N \geq 1 \implies \cup_{n=1}^N A_n \in \mathcal{A}$

(iii) $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A} \implies \cap_n A_n \in \mathcal{A}$

(iv) $A, B \in \mathcal{A} \implies A \setminus B \in \mathcal{A}$

Démonstration. Par définition, $\varnothing = X^c \in \mathcal{A}$. Pour (ii), si $A_1, \dots, A_N \in \mathcal{A}$, on peut prolonger la suite finie par des ensembles vides : définissons $B_n := A_n$ pour $n \leq N$ et $B_n = \varnothing$ pour $n > N$. Alors $\cup_{n=1}^N A_n = \cup_{n\in\mathbb{N}} B_n$, ce qui permet de conclure. Enfin, les points (iii) et (iv) suivent de $\cap_{n\in\mathbb{N}} A_n = (\cup_{n\in\mathbb{N}} A_n^c)^c$ et $A \setminus B = (A^c \cup B)^c$. □

Exemple 1.3.
— La σ-algèbre triviale sur X est $\{\varnothing, X\}$.
— La σ-algèbre discrète sur X est $\mathcal{P}(X)$.
— Si $A \subset X$, alors $\{\varnothing, A, A^c, X\}$ est une σ-algèbre.
— Si X est infini, la collection de tous les sous-ensembles A de X tels que A est dénombrable ou A^c est dénombrable est une σ-algèbre. Par contre, la collection de tous les sous-ensembles A de X tels que A est fini ou A^c est fini n'est pas une σ-algèbre.
— Si $(X, \mathcal{A})$ est un espace mesurable et si $T \subset X$, alors la restriction $\mathcal{A}_T := \{A \cap T : A \in \mathcal{A}\}$ est une σ-algèbre. Si $T \in \mathcal{A}$, on remarque que $\mathcal{A}_T = \{A \in \mathcal{A} : A \subset T\}$.

Remarque 1.4. Une collection $\mathcal{A}$ de sous-ensembles de X est simplement appelée *algèbre* si
— $X \in \mathcal{A}$,
— pour tout $A \in \mathcal{A}$, on a $A^c \in \mathcal{A}$,
— pour toute suite finie $A_1, \dots, A_n \in \mathcal{A}$, $n \geq 1$, on a $A_1 \cup \dots \cup A_n \in \mathcal{A}$.

Comme on l'a vu, une σ-algèbre est en particulier une algèbre.

1.2.1 σ-algèbre engendrée

Étant donné une collection $\mathcal{F}$ de sous-ensembles de X, il existe bien sûr une σ-algèbre contenant $\mathcal{F}$ (par exemple, la σ-algèbre discrète $\mathcal{P}(X)$). Le résultat suivant permet de construire la plus petite σ-algèbre contenant $\mathcal{F}$.

Proposition 1.5. *Toute intersection non-vide de σ-algèbres sur X est une σ-algèbre sur X.*

Démonstration. Soit $\mathcal{C}$ une collection non-vide de σ-algèbres sur X. Notons $\mathcal{A}$ l'intersection de ces σ-algèbres :

— $X \in \mathcal{A}$ puisque X appartient à toutes les σ-algèbres de $\mathcal{C}$.

— Si $A \in \mathcal{A}$, alors A appartient à toutes les σ-algèbres de $\mathcal{C}$. Toutes ces σ-algèbres contiennent donc A^c, ce qui implique que $A^c \in \mathcal{A}$.

— Enfin, si $(A_n)_{n\in\mathbb{N}}$ est une suite d'éléments de $\mathcal{A}$, alors $\bigcup_{n\in\mathbb{N}} A_n$ appartient à chaque σ-algèbre de $\mathcal{C}$, et donc à $\mathcal{A}$. □

Remarque 1.6. Mentionnons néanmoins que l'union de σ-algèbres n'est pas nécessairement une σ-algèbre. Il suffit par exemple de considérer les σ-algèbres $\{\varnothing, A, A^c, X\}$ et $\{\varnothing, B, B^c, X\}$ pour $A, B \subset X, A \neq B$, dont l'union donne la famille $\{\varnothing, A, A^c, B, B^c, X\}$.

Corollaire 1.7. *Pour $\mathcal{F} \subset \mathcal{P}(X)$, il existe une plus petite (au sens de l'inclusion) σ-algèbre sur X qui contient $\mathcal{F}$. On l'appelle la* σ-algèbre engendrée par $\mathcal{F}$ *et on la note $\sigma(\mathcal{F})$.*

Démonstration. Soit $\mathcal{C}$ la collection non-vide de toutes les σ-algèbres sur X contenant $\mathcal{F}$. Par la proposition précédente, l'intersection $\sigma(\mathcal{F})$ de toutes les σ-algèbres de $\mathcal{C}$ est une σ-algèbre. Bien sûr, $\sigma(\mathcal{F})$ contient $\mathcal{F}$. De plus, si $\mathcal{A}$ est une σ-algèbre sur X qui contient $\mathcal{F}$, alors $\mathcal{A} \in \mathcal{C}$ et $\sigma(\mathcal{F}) \subset \mathcal{A}$. □

1.2.2 σ-algèbres de Borel

Sur $\mathbb{R}$, comme nous souhaitons étendre la notion de longueur définie sur les intervalles, cela nous mène à considérer la σ-algèbre engendrée par les intervalles. Plus généralement, sur un espace topologique, on définit la σ-algèbre de Borel comme la σ-algèbre engendrée par les ouverts. Sur $\mathbb{R}$, on va voir que cela coïncide avec la σ-algèbre induite par les intervalles.

Définition 1.8. Soit $(X, \mathcal{T})$ un espace topologique. La *σ-algèbre de Borel* $\mathcal{B}(X) = \mathcal{B}(X, \mathcal{T})$ de $(X, \mathcal{T})$ est la σ-algèbre engendrée par $\mathcal{T}$. Un élément de $\mathcal{B}(X, \mathcal{T})$ est appelé un *ensemble borélien.*

En particulier, les fermés sont des ensembles boréliens. Pour $d \geq 1$, si $\mathcal{T}$ est la topologie euclidienne sur $\mathbb{R}^d$, on pose $\mathbb{B}^d = \mathcal{B}(\mathbb{R}^d, \mathcal{T})$. Pour $d = 1$, la σ-algèbre de Borel de $\mathbb{R}$ sera simplement notée $\mathbb{B}$. Pour un sous-ensemble quelconque $X \subset \mathbb{R}^d$, on peut considérer la topologie euclidienne induite sur X : la σ-algèbre induite est alors appelée la σ-algèbre de Borel sur X et est notée $\mathbb{B}(X)$. On peut remarquer par ailleurs que $\mathbb{B}(X)$ coïncide avec la restriction $(\mathbb{B}^d)_X = \{A \cap X : A \in \mathbb{B}^d\}$.

Proposition 1.9. *La σ-algèbre de Borel $\mathbb{B}$ sur $\mathbb{R}$ est engendrée par les intervalles. Plus précisément, on a les caractérisations suivantes,*

$$\begin{aligned} \mathbb{B} &= \sigma(\{\textit{fermés de } \mathbb{R}\}) \\ &= \sigma(\{[a, \infty) : a \in \mathbb{R}\}) \\ &= \sigma(\{(a, \infty) : a \in \mathbb{R}\}) \\ &= \sigma(\{(-\infty, b] : b \in \mathbb{R}\}) \\ &= \sigma(\{(-\infty, b) : b \in \mathbb{R}\}) \\ &= \sigma(\{[a, b] : a, b \in \mathbb{R},\, a < b\}) \\ &= \sigma(\{(a, b) : a, b \in \mathbb{R},\, a < b\}). \end{aligned}$$

Démonstration. Notons respectivement $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3, \mathcal{B}_4, \mathcal{B}_5, \mathcal{B}_6, \mathcal{B}_7$ les différentes σ-algèbres engendrées considérées dans l'énoncé. On va montrer que

$$\mathcal{B}_7 \subset \mathcal{B}_6 \subset \mathcal{B}_5 \subset \mathcal{B}_4 \subset \mathcal{B}_3 \subset \mathcal{B}_2 \subset \mathcal{B}_1 \subset \mathbb{B} \subset \mathcal{B}_7,$$

d'où l'égalité entre toutes ces σ-algèbres. Tout d'abord, comme $(a, b) = \cup_n [a + \frac{1}{n}, b - \frac{1}{n}] \in \mathcal{B}_6$, on obtient $\mathcal{B}_7 \subset \mathcal{B}_6$. Ensuite, comme $[a, b] = \cap_n (-\infty, b + \frac{1}{n}) \setminus (-\infty, a) \in \mathcal{B}_5$, on a $\mathcal{B}_6 \subset \mathcal{B}_5$. De même, comme $(-\infty, b) = \cup_n (-\infty, b - \frac{1}{n}] \in \mathcal{B}_4$, on a $\mathcal{B}_5 \subset \mathcal{B}_4$. Comme $(-\infty, b] = (b, \infty)^c \in \mathcal{B}_3$, on a $\mathcal{B}_4 \subset \mathcal{B}_3$. Comme $(a, \infty) = \cup_n [a + \frac{1}{n}, \infty) \in \mathcal{B}_2$, on a $\mathcal{B}_3 \subset \mathcal{B}_2$. Comme $[a, \infty)$ est fermé, on a $\mathcal{B}_2 \subset \mathcal{B}_1$. Comme un fermé est le complémentaire d'un ouvert, on a $\mathcal{B}_1 \subset \mathbb{B}$. Il reste à montrer que tout ensemble ouvert $U \subset \mathbb{R}$ appartient à $\mathcal{B}_7$. Puisque $\mathbb{Q}$ est dense dans $\mathbb{R}$, on peut écrire

$$U = \bigcup_{a, b \in \mathbb{Q},\, (a,b) \subset U} (a, b),$$

et donc $U \in \mathcal{B}_7$. En effet, pour $x \in U$, comme U est ouvert, il existe $\varepsilon > 0$ tel que $(x - \varepsilon, x + \varepsilon) \subset U$, et la densité de $\mathbb{Q}$ permet alors de choisir $a, b \in \mathbb{Q}$ avec $x - \varepsilon < a < x < b < x + \varepsilon$, et donc $x \in (a, b) \subset U$. □

Ce résultat s'étend aisément au cas de dimension supérieure.

Proposition 1.10. *La σ-algèbre de Borel $\mathbb{B}^d$ sur $\mathbb{R}^d$ est engendrée par les rectangles. Plus précisément, on a par exemple les caractérisations suivantes,*

$$\mathbb{B}^d = \sigma\Big(\Big\{ \prod_{k=1}^{d} [a_k, \infty) : a_1, \dots, a_k \in \mathbb{R}\Big\}\Big) = \sigma\Big(\Big\{ \prod_{k=1}^{d} [a_k, b_k] : a_1, b_1, \dots, a_k, b_k \in \mathbb{R},\, a_i < b_i\Big\}\Big).$$

Lorsque nous introduirons la notion d'intégrale, on considérera des fonctions à valeurs dans $\mathbb{R}$ ou plus généralement dans $\overline{\mathbb{R}} := [-\infty, +\infty]$. Rappelons que la topologie de $\overline{\mathbb{R}}$ est la topologie dont une base d'ouverts est donnée par les intervalles de la forme $(a, +\infty]$, $[-\infty, b)$ et (a, b) où $a, b \in \mathbb{R}$. On ajoute donc aux ouverts de $\mathbb{R}$ les voisinages de $\pm\infty$. En particulier, la topologie induite par $\overline{\mathbb{R}}$ sur $\mathbb{R}$ coïncide avec la topologie euclidienne de $\mathbb{R}$. Le résultat suivant élucide la structure des boréliens de $\overline{\mathbb{R}}$.

Lemme 1.11. *On a $\mathcal{B}(\overline{\mathbb{R}}) = \big\{B \cup C : B \in \mathbb{B}, C \subset \{-\infty, +\infty\}\big\}$.*

Démonstration. Notons $\mathcal{A} = \{B \cup C : B \in \mathbb{B}, C \subset \{-\infty, +\infty\}\}$. On vérifie aisément que $\mathcal{A}$ est une σ-algèbre. Soit $B \cup C$ un élément de $\mathcal{A}$. Comme les ouverts de $\mathbb{R}$ sont des ouverts de $\overline{\mathbb{R}}$, B est également un élément de $\mathcal{B}(\overline{\mathbb{R}})$. De plus, puisque $\{-\infty\} = \cap_n [-\infty, -n)$ et $\{+\infty\} = \cap_n (n, +\infty]$, l'ensemble $B \cup C$ appartient bien à la σ-algèbre $\mathcal{B}(\overline{\mathbb{R}})$. On en déduit $\mathcal{A} \subset \mathcal{B}(\overline{\mathbb{R}})$. Réciproquement, si $\omega \subset \overline{\mathbb{R}}$ est ouvert, alors on peut écrire

$$\omega = \big(\omega \cap \mathbb{R}\big) \cup \big(\{-\infty, +\infty\} \cap \omega\big)$$

où $\omega \cap \mathbb{R}$ est un ouvert de $\mathbb{R}$. Donc ω est bien un élément de $\mathcal{A}$. Par conséquent, $\mathcal{B}(\overline{\mathbb{R}}) \subset \mathcal{A}$. □

Proposition 1.12. *La σ-algèbre de Borel $\mathcal{B}(\overline{\mathbb{R}})$ a les caractérisations suivantes,*

$$\mathcal{B}(\overline{\mathbb{R}}) \;=\; \sigma(\{[a, \infty] : a \in \overline{\mathbb{R}}\}) \;=\; \sigma(\{(a, \infty] : a \in \overline{\mathbb{R}}\}).$$

Démonstration. Notons respectivement $\mathcal{B}_1$ et $\mathcal{B}_2$ les σ-algèbres engendrées considérées dans l'énoncé. On va montrer que $\mathcal{B}_1 \subset \mathcal{B}_2 \subset \mathcal{B}(\overline{\mathbb{R}}) \subset \mathcal{B}_1$. Comme $[a, \infty] = \cap_n (a - \frac{1}{n}, \infty]$, on a $\mathcal{B}_1 \subset \mathcal{B}_2$. Puisque les ensembles $(a, \infty]$ sont ouverts dans $\overline{\mathbb{R}}$, on a $\mathcal{B}_2 \subset \mathcal{B}(\overline{\mathbb{R}})$. Il reste à montrer $\mathcal{B}(\overline{\mathbb{R}}) \subset \mathcal{B}_1$. Par le Lemme 1.11, on sait que tout élément de $\mathcal{B}(\overline{\mathbb{R}})$ a la forme $B \cup C$ où B est un borélien de $\mathbb{R}$ et $C \subset \{-\infty, +\infty\}$. Comme $\{+\infty\} = \cap_n [n, +\infty]$ et $\{-\infty\} = \cap_n [-\infty, n) = \cap_n [n, +\infty]^c$, on obtient $C \in \mathcal{B}_1$. De plus, on déduit que $[a, \infty) = [a, \infty] \setminus \{\infty\} \in \mathcal{B}_1$, et donc $\mathbb{B} \subset \mathcal{B}_1$ par la Proposition 1.9. Ainsi, $B \in \mathcal{B}_1$ et donc $B \cup C \in \mathcal{B}_1$. □

1.3 Mesures

Nous introduisons à présent la notion de mesure, définie sur une σ-algèbre. Cela sera l'élément central de ce cours.

Définition 1.13. Soit $(X, \mathcal{A})$ un espace mesurable. On dit qu'une application $\mu : \mathcal{A} \to [0, \infty]$ est une *mesure* sur $(X, \mathcal{A})$ si

- $\mu(\varnothing) = 0$,
- μ est σ-additif : si $(A_n)_{n \in \mathbb{N}}$ est une suite d'ensembles de $\mathcal{A}$ deux à deux disjoints, alors

$$\mu\Big(\bigcup_{n \in \mathbb{N}} A_n\Big) = \sum_{n \in \mathbb{N}} \mu(A_n).$$

On dit alors que le triplet $(X, \mathcal{A}, \mu)$ est un *espace de mesure* (ou *espace mesuré*).

Remarque 1.14. La condition $\mu(\varnothing) = 0$ est parfois supprimée de la définition. En effet, si $\mu : \mathcal{A} \to [0, \infty]$ est σ-additive et n'est pas uniformément égale à ∞, alors on vérifie que nécessairement $\mu(\varnothing) = 0$.

Exemple 1.15. Soit $(X, \mathcal{A})$ un espace mesurable.

— Si μ est une mesure telle que $\mu(X) = 1$, on parle de *mesure de probabilité*. Le triplet $(X, \mathcal{A}, \mu)$ est alors appelé un *espace de probabilités* (ou *espace probabilisé*). On dit que X est l'univers, les éléments de $\mathcal{A}$ sont les événements, et les éléments de X sont les réalisations.

— L'application définie par

$$\mu(A) = \begin{cases} \sharp A & : \text{ si } A \text{ est un ensemble fini} \\ +\infty & : \text{ si } A \text{ est un ensemble infini.} \end{cases}$$

est une mesure, appelée *mesure de comptage*, et notée $\mu = \sharp$.

— Étant donné un élément $x \in X$, l'application définie par

$$\delta_x(A) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{sinon.} \end{cases}$$

est une mesure, appelée *mesure de Dirac* en x.

— Si $X = \mathbb{R}$ et $\mathcal{A} = \mathbb{B}$, nous construirons au chapitre suivant une mesure $\mathcal{L}$ qui assigne à tout intervalle de $\mathbb{R}$ sa longueur. Cette mesure sera appelée la *mesure de Lebesgue*.

On remarque quelques propriétés élémentaires des mesures, qui se déduisent aisément de la définition ci-dessus.

Proposition 1.16. *Soit $(X, \mathcal{A}, \mu)$ un espace de mesure.*

(i) Additivité finie : si $N \geq 1$ et si $A_1, \ldots, A_N \in \mathcal{A}$ sont deux à deux disjoints, alors on a $\mu(\cup_{n=1}^N A_n) = \sum_{n=1}^N \mu(A_n)$.

(ii) Monotonie : si $A \subset B$ avec $A, B \in \mathcal{A}$, alors $\mu(A) \leq \mu(B)$. Si en outre $\mu(A)\infty$, alors on a en plus $\mu(B \setminus A) = \mu(B) - \mu(A)$.

(iii) Sous-additivité : si $(A_n)_n \subset \mathcal{A}$, on a

$$\mu\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) \leq \sum_{n\in\mathbb{N}} \mu(A_n).$$

(iv) Continuité à gauche : si la suite $(A_n)_n \subset \mathcal{A}$ est croissante (i.e. $A_1 \subset A_2 \subset \ldots \subset \cup_n A_n$, ce qu'on notera $A_n \uparrow \cup_n A_n$), alors

$$\mu(A_n) \uparrow \mu\Big(\bigcup_n A_n\Big)$$

(v) Continuité à droite : si la suite $(A_n)_n \subset \mathcal{A}$ est décroissante (i.e. $A_1 \supset A_2 \supset \ldots \supset \cap_n A_n$, ce qu'on notera $A_n \downarrow \cap_n A_n$) et s'il existe n_0 tel que $\mu(A_{n_0}) < \infty$, alors on a

$$\mu(A_n) \downarrow \mu\Big(\bigcap_n A_n\Big).$$

Démonstration. Le point (i) est trivial.

(ii) Puisque les ensembles A et $B \setminus A$ sont disjoints et puisque $B = A \cup (B \setminus A)$, on a $\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A)$. Si $\mu(A) < \infty$, cela entraîne aussi que $\mu(B \setminus A) = \mu(B) - \mu(A)$.

(iii) Posons $B_0 = A_0$ et $B_n = A_n \setminus \bigcup_{j=0}^{n-1} A_j$ pour tout $n \geq 1$. Alors chaque B_n appartient à $\mathcal{A}$, satisfait $B_n \subset A_n$, et ces ensembles B_n sont deux à deux disjoints. On déduit donc

$$\mu\Big(\bigcup_n A_n\Big) = \mu\Big(\bigcup_n B_n\Big) = \sum_n \mu(B_n) \leq \sum_n \mu(A_n).$$

(iv) Soit $A = \bigcup_{n\in\mathbb{N}} A_n$. Posons $B_0 = A_0$ et $B_n = A_n \setminus A_{n-1}$ pour tout $n \geq 1$. Les ensembles construits appartiennent à $\mathcal{A}$ et sont deux à deux disjoints. De plus, pour tout $n \in \mathbb{N}$, on a $A_n = B_1 \cup \cdots \cup B_n$ et donc $A = \bigcup_j B_j$. On en déduit $\mu(A_n) = \sum_{j=0}^{n} \mu(B_j)$ et $\mu(A) = \sum_{j=0}^{\infty} \mu(B_j)$, d'où

$$\lim_n \mu(A_n) = \sum_{j=0}^{\infty} \mu(B_j) = \mu(A).$$

(v) On peut supposer que $n_0 = 0$. Posons $B_n = A_1 \setminus A_n$ pour tout n. Les éléments ainsi construits appartiennent à $\mathcal{A}$ et la suite $(B_n)_{n\in\mathbb{N}}$ est croissante. Par le point (iv), on a

$$\mu\Big(\bigcup_n B_n\Big) = \lim_n \mu(B_n) = \mu(A_1) - \lim_n \mu(A_n).$$

D'autre part, par (ii), on a

$$\mu\Big(\bigcup_n B_n\Big) = \mu\Big(A_1 \setminus \bigcap_n A_n\Big) = \mu(A_1) - \mu\Big(\bigcap_n A_n\Big)$$

et la conclusion s'ensuit. □

Remarque 1.17. Notons que l'hypothèse $\mu(A_{n_0}) < \infty$ au point (v) ne peut pas être enlevée. En effet, sur l'espace de mesure $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \sharp)$, considérons $A_n = \{m \in \mathbb{N} : m \geq n\}$. D'une part, on a $\mu(A_n) = \infty$ pour tout n. D'autre part, puisque $\cap_n A_n = \varnothing$, on a $\mu(\cap_n A_n) = 0$.

Le résultat suivant est un outil de base de la théorie des probabilités. Dans ce contexte, il se lit de la manière suivante : si la série des probabilités d'une suite d'événements converge, alors la probabilité qu'une infinité d'entre eux se réalise simultanément est nulle.

Lemme 1.18 (Lemme de Borel-Cantelli)**.** *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $(A_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathcal{A}$. Si*

$$\sum_n \mu(A_n) < +\infty,$$

alors

$$\mu\Big(\limsup_n A_n\Big) = 0,$$

où $\limsup_n A_n = \cap_{n\in\mathbb{N}} \cup_{k\geq n} A_k$.

Démonstration. Pour tout $n \in \mathbb{N}$, on pose $B_n = \bigcup_{k\geq n} A_k$. Alors $(B_n)_{n\in\mathbb{N}}$ est une suite décroissante d'élements de $\mathcal{A}$. De plus

$$\mu(B_0) \leq \sum_n \mu(A_n) < \infty.$$

La continuité à droite de la mesure μ donne alors

$$\mu\Big(\limsup_n A_n\Big) = \mu\Big(\bigcap_n B_n\Big) = \lim_n \mu(B_n) \leq \lim_n \sum_{k\geq n} \mu(A_k) = 0.$$ □

1.4 Fonctions mesurables

Comme on a vu en introduction, sur un espace de mesure $(X, \mathcal{A}, \mu)$, afin de pouvoir définir l'intégrale de Lebesgue d'une fonction $f : X \to \mathbb{R}$, il est nécessaire de pouvoir donner un sens à la mesure des ensembles de niveau $f^{-1}([a, b)) \subset X$ pour tous a, b. Étant donné que la mesure n'est en général pas définie sur toutes les parties de X, mais seulement sur la σ-algèbre $\mathcal{A}$, cela motive la notion suivante de mesurabilité de fonctions : une fonction $f : X \to \mathbb{R}$ est 'mesurable' si tous ses ensembles de niveau appartiennent à $\mathcal{A}$.

Par souci de généralité, nous commençons par introduire la notion de mesurabilité pour des fonctions entre des espaces mesurables quelconques. Pour des fonctions à valeurs dans $\mathbb{R}$, on verra que cela coïncide avec la notion voulue.

Définition 1.19. Soient $(X, \mathcal{A})$ et $(Y, \mathcal{A}')$ deux espaces mesurables. On dit qu'une fonction $f : X \to Y$ est $(\mathcal{A}, \mathcal{A}')$-*mesurable* (ou simplement *mesurable*) si $f^{-1}(B) \in \mathcal{A}$ pour tout $B \in \mathcal{A}'$.

Le résultat suivant est souvent utilisé pour vérifier la mesurabilité d'une fonction. Sa preuve repose sur un argument que nous emploierons fréquemment par la suite : pour montrer qu'une propriété est satisfaite sur toute une σ-algèbre, on considère l'ensemble des "bons" ensembles vérifiant cette propriété, puis on montre que cet ensemble forme lui-même une σ-algèbre contenant suffisamment d'ensembles pour engendrer la σ-algèbre tout entière.

Lemme 1.20. *Soient $(X, \mathcal{A})$ et $(Y, \mathcal{A}')$ deux espaces mesurables et $f : X \to Y$. Supposons que $\mathcal{F} \subset \mathcal{P}(Y)$ est tel que $\sigma(\mathcal{F}) = \mathcal{A}'$. Alors f est mesurable si et seulement si $f^{-1}(A) \in \mathcal{A}$ pour tout $A \in \mathcal{F}$.*

Démonstration. Si f est mesurable, alors il est clair que $f^{-1}(A) \in \mathcal{A}$ pour tout $A \in \mathcal{F}$. On passe à la preuve de la réciproque : supposons que $f^{-1}(A) \in \mathcal{A}$ pour tout $A \in \mathcal{F}$, et montrons que f est mesurable. Posons $\mathcal{C} = \{A \in \mathcal{A}' : f^{-1}(A) \in \mathcal{A}\}$. On observe que $\mathcal{C}$ est une σ-algèbre. Puisque $\mathcal{F} \subset \mathcal{C}$ par hypothèse, on en tire que $\mathcal{A}' = \sigma(\mathcal{F}) \subset \mathcal{C}$, et la conclusion s'ensuit. $\square$

Avec ce lemme en main, on peut maintenant montrer que la définition ci-dessus de mesurabilité est équivalente à la définition voulue pour des fonctions à valeurs dans $\mathbb{R}$.

Corollaire 1.21. *Soit $(X, \mathcal{A})$ un espace mesurable.*

(i) Soit Y un espace topologique. Une fonction $f : X \to Y$ est $(\mathcal{A}, \mathcal{B}(Y))$-mesurable si et seulement si $f^{-1}(U) \in \mathcal{A}$ pour tout ouvert $U \subset Y$. On dit alors simplement que f est $\mathcal{A}$-mesurable (ou mesurable).

(ii) Une fonction $f : X \to \mathbb{R}$ ou $\overline{\mathbb{R}}$ est $\mathcal{A}$-mesurable
$\iff f^{-1}(U) \in \mathcal{A}$ *pour tout ouvert* $U \subset \mathbb{R}$
$\iff \{x \in X : f(x) \le b\} \in \mathcal{A}$ *pour tout* $b \in \mathbb{R}$, ,
$\iff \{x \in X : f(x) < b\} \in \mathcal{A}$ *pour tout* $b \in \mathbb{R}$,
$\iff \{x \in X : f(x) \ge b\} \in \mathcal{A}$ *pour tout* $b \in \mathbb{R}$,
$\iff \{x \in X : f(x) > b\} \in \mathcal{A}$ *pour tout* $b \in \mathbb{R}$.

(iii) Soient X et Y deux espaces topologiques. Toute fonction continue $f : X \to Y$ est $\mathcal{B}(X)$-mesurable.

Définition 1.22. Si X est un espace topologique, une fonction $X \to \mathbb{R}$ ou $\overline{\mathbb{R}}$ qui est $\mathcal{B}(X)$-mesurable est dite borélienne.

1.4.1 Propriétés élémentaires

Les lemmes suivants listent un nombre de propriétés élémentaires importantes de la mesurabilité de fonctions.

Lemme 1.23.

(i) Soient $(X,\mathcal{A})$, $(Y,\mathcal{A}')$ et $(Z,\mathcal{A}'')$ des espaces mesurables. Si $f : X \to Y$ et $g : Y \to Z$ sont mesurables, alors $g \circ f : X \to Z$ est mesurable.

(ii) Soient $(X,\mathcal{A})$ un espace mesurable, $f : X \to \mathbb{R}^d$ une fonction $\mathcal{A}$-mesurable, et $\phi : \mathbb{R}^d \to \mathbb{R}$ continue. Alors $\phi \circ f : X \to \mathbb{R}$ est $\mathcal{A}$-mesurable.

Démonstration. $(g \circ f)^{-1}(A) = f^{-1}(g^{-1}(A))$ □

Lemme 1.24. *Soit $(X,\mathcal{A})$ un espace mesurable et soient $f, g : X \to \mathbb{R}$. La fonction $(f,g) : X \to \mathbb{R}^2 : x \mapsto (f(x), g(x))$ est $\mathcal{A}$-mesurable si et seulement si f, g sont toutes les deux $\mathcal{A}$-mesurables.*

Démonstration. Supposons (f,g) mesurable. Comme les projections sont continues, on déduit que f, g sont toutes deux mesurables. Supposons à présent f, g mesurables. Alors $(f,g)^{-1}([a,\infty) \times [b,\infty)) = f^{-1}([a,\infty)) \cap g^{-1}([b,\infty)) \in \mathcal{A}$. □

Lemme 1.25. *Soit $(X,\mathcal{A})$ un espace mesurable.*

(i) $f, g : X \to \mathbb{R}$ mesurables $\implies$ $\alpha f, f+g, f-g, fg : X \to \mathbb{R}$ mesurables ; les fonctions mesurables forment donc une algèbre.

(ii) $A \in \mathcal{A} \iff \mathbb{1}_A$ $\mathcal{A}$-mesurable

(iii) si $f_n : X \to \overline{\mathbb{R}}$ sont mesurables pour tous n, alors $\sup_n f_n$ et $\limsup_n f_n$ sont mesurables

(iv) $f, g : X \to \overline{\mathbb{R}}$ mesurables $\implies$ $f \vee g := \max\{f,g\}$ et $f \wedge g := \min\{f,g\}$ mesurables

(v) $f : X \to \overline{\mathbb{R}}$ mesurable $\iff$ $f_+ := f \vee 0$ et $f_- := (-f) \vee 0$ mesurables

(vi) $f : X \to \overline{\mathbb{R}}$ mesurable $\implies$ $|f| = f_+ + f_-$ mesurable

Démonstration. (i) Utiliser les lemmes précédents et noter que $F(x,y) = x + y$ et $G(x,y) = xy$ sont continues $\mathbb{R}^2 \to \mathbb{R}$. (ii) On a $\mathbb{1}_A^{-1}(U) = A$ si $1 \in U \not\ni 0$, $\mathbb{1}_A^{-1}(U) = A^c$ si $0 \in U \not\ni 1$, $\mathbb{1}_A^{-1}(U) = X$ si $0, 1 \in U$, et enfin $\mathbb{1}_A^{-1}(U) = \varnothing$ si $0, 1 \notin U$. (iii) $(\inf_n f_n)^{-1}[a,\infty]) = \{x : f_n(x) \geq a \ \forall n\} = \cap_n f_n^{-1}([a,\infty]) \in \mathcal{A}$ et on a $\sup_n f_n = -\inf_n(-f_n)$, $\limsup_n f_n = \lim_N \sup_{n \geq N} f_n = \inf_N \sup_{n \geq N} f_n$, et $\liminf_n f_n = -\limsup_n(-f_n)$. □

Lemme 1.26. *Soit $(X,\mathcal{A})$ un espace mesurable.*

(i) Pour $f, g : X \to \overline{\mathbb{R}}$ mesurables, l'ensemble $\{x : f(x) < g(x)\}$ (resp. $>$, $=$) est mesurable.

(ii) Pour une suite $(f_n)_n$ de fonctions mesurables $X \to \overline{\mathbb{R}}$, l'ensemble $\{x : \lim_n f_n(x)$ existe$\} = \{x : \limsup_n f_n(x) = \liminf_n f_n(x)\}$ est mesurable.

1.4.2 Approximation par des fonctions simples

On appelle fonctions simples les fonctions pour lesquelles seul un nombre fini de points est atteint en ordonnée. Comme on a vu en introduction, l'idée de Lebesgue est d'approcher les fonctions à intégrer par des fonctions simples plutôt que par des fonctions en escalier.

Définition 1.27. Soit $(X,\mathcal{A})$ un espace mesurable. Une fonction $f : X \to \overline{\mathbb{R}}$ est dite simple si son image $f(X)$ est un ensemble fini. Si $f(X) = \{a_1, \ldots, a_n\}$ avec $a_i \neq a_j$ pour $i \neq j$, on peut alors écrire $f = \sum_{i=1}^n a_i \mathbb{1}_{A_i}$ avec $A_i := f^{-1}(\{a_i\})$.

Lemme 1.28. *Une fonction simple $f : X \to \overline{\mathbb{R}}$ est mesurable si et seulement si elle s'écrit sous la forme $f = \sum_{i=1}^n a_i \mathbb{1}_{A_i}$ avec $a_1, \ldots, a_n \in \overline{\mathbb{R}}$ et $A_1, \ldots, A_n \in \mathcal{A}$ deux à deux disjoints.*

Démonstration. Si f est simple mesurable, on peut écrire $f = \sum_{i=1}^n a_i 1_{A_i}$ avec $A_i = f^{-1}(\{a_i\})$, et donc nécessairement $A_i \in \mathcal{A}$. Réciproquement, une fonction de la forme $f = \sum_{i=1}^n a_i 1_{A_i}$ avec $A_i \in \mathcal{A}$ est clairement simple et mesurable. □

Proposition 1.29. *Soit $(X, \mathcal{A})$ un espace mesurable. Si $f : X \to [0, \infty]$ est mesurable, il existe une suite $(s_n)_n$ de fonctions simples mesurables positives telles que $s_n \uparrow f$ (c'est-à-dire $s_1 \leq s_2 \leq \ldots \leq f$ et $s_n(x) \to f(x)$ pour tout x).*

Démonstration. Comme on a vu en introduction, l'idée est de discrétiser l'espace des ordonnées. Plus précisément, pour tout n, on définit

$$s_n = \sum_{k=0}^{n2^n-1} \frac{k}{n} \mathbb{1}_{f^{-1}([\frac{k}{2^n}, \frac{k+1}{2^n}))} + n \mathbb{1}_{f^{-1}([n,\infty))},$$

et on vérifie que ce choix satisfait les propriétés annoncées. □

1.5 Intégration abstraite

La définition des intégrales se fait en trois étapes : on commence par définir l'intégrale de fonctions simples mesurables positives, puis on passe à des fonctions mesurables positives générales, et enfin à des fonctions mesurables signées.

1.5.1 Intégration de fonctions simples mesurables positives

Définition 1.30. Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $f : X \to [0, \infty]$ une fonction simple mesurable, disons $f = \sum_{i=1}^n a_i \mathbb{1}_{A_i}$. Pour $A \in \mathcal{A}$, l'intégrale (de Lebesgue) de f sur A par rapport à μ est définie par

$$\int_A f d\mu := \sum_{i=1}^n a_i \mu(A_i \cap A).$$

Sur l'espace entier, on note simplement $\int f d\mu = \int_X f d\mu$. Dans cette définition, ainsi que partout dans ce cours, on utilise la convention $0 \cdot (+\infty) = 0$.

1.5.2 Intégration de fonctions positives

Définition 1.31. Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $f : X \to [0, \infty]$ une fonction mesurable. Pour $A \in \mathcal{A}$, l'intégrale (de Lebesgue) de f sur A par rapport à μ est définie par

$$\int_A f d\mu := \sup \left\{ \int_A s d\mu : s \text{ simple mesurable avec } 0 \leq s \leq f \right\}.$$

Sur l'espace entier, on note simplement $\int f d\mu = \int_X f d\mu$.

Lemme 1.32. *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure, $f : X \to [0, \infty]$ une fonction simple mesurable, et $A \in \mathcal{A}$. Alors les deux définitons de $\int_A f d\mu$ données dans les Définitions 1.30 et 1.31 coïncident.*

Démonstration. Notons $I_1(f, A)$ et $I_2(f, A)$ les définitions de $\int_A f d\mu$ données dans les Définitions 1.30 et 1.31, respectivement. Par définition, $I_1(f, A) \leq I_2(f, A)$. Pour l'autre inégalité, il reste à voir que pour s simple mesurable avec $0 \leq s \leq f$ on a $I_1(s, A) \leq I_1(f, A)$. Pour ce faire, on écrit f et s comme

$$f = \sum_{i=1}^n a_i \mathbb{1}_{A_i}, \qquad s = \sum_{i=1}^m b_i \mathbb{1}_{B_i},$$

avec $f(X) = \{a_1, \dots, a_n\}$, $s(X) = \{b_1, \dots, b_m\}$, $A_i = f^{-1}(\{a_i\})$ disjoints deux à deux, et $B_i = s^{-1}(\{b_i\})$ disjoints deux à deux. Comme $s \leq f$, on a $b_j \leq a_i$ sur $A_i \cap B_j$. On peut alors calculer

$$I_1(s, A) = \sum_{j=1}^{m} b_j \mu(B_j \cap A) = \sum_{j=1}^{m} \sum_{i=1}^{n} b_j \mu(A_i \cap B_j \cap A) \leq \sum_{j=1}^{m} \sum_{i=1}^{n} a_i \mu(A_i \cap B_j \cap A) = I_1(f, A),$$

ce qui donne l'inégalité voulue. □

Exercice 1.33. Soient $(X, \mathcal{A}, \mu)$ un espace de mesure, $A, B \in \mathcal{A}$, et $f, g : X \to [0, \infty]$ mesurables.

(i) $f \leq g$ sur $A \implies \int_A f d\mu \leq \int_A g d\mu$

(ii) $A \subset B \implies \int_A f d\mu \leq \int_B f d\mu$

(iii) $c \in [0, \infty) \implies \int_A c f d\mu = c \int_A f d\mu$

(iv) $\mu(A) = 0 \implies \int_A f d\mu = 0$ (même si $f = \infty$ sur A)

(v) $f = 0$ sur $A \implies \int_A f d\mu = 0$ (même si $\mu(A) = \infty$)

(vi) $\int_A f d\mu = \int_X f \mathbb{1}_A d\mu$

Avec le point (iii) ci-dessus, afin de pouvoir conclure la linéarité de l'intégrale de fonctions positives, il reste à voir

$$\int_A (f + g) d\mu = \int_A f d\mu + \int_A g d\mu.$$

Cependant, avec la définition de l'intégrale comme un supremum, ce résultat n'est pas immédiat et demandera un certain travail. On commence par montrer le résultat pour des fonctions simples.

Lemme 1.34. *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure, $A \in \mathcal{A}$, et $f, g : X \to [0, \infty]$ simples mesurables. Alors*

$$\int_A (f + g) d\mu = \int_A f d\mu + \int_A g d\mu.$$

Démonstration. Comme dans la preuve du lemme précédent, on commence par écrire f et g comme

$$f = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i} \quad \text{et} \quad g = \sum_{i=1}^{m} b_i \mathbb{1}_{B_i}$$

avec $f(X) = \{a_1, \dots, a_n\}$, $g(X) = \{b_1, \dots, b_m\}$, $A_i = f^{-1}(\{a_i\})$ disjoints deux à deux, $B_i = g^{-1}(\{b_i\})$ disjoints deux à deux, et $X = \cup_i A_i = \cup_i B_i$. On peut alors écrire $f + g = \sum_{i=1}^{n} \sum_{j=1}^{m} (a_i + b_j) \mathbb{1}_{A_i \cap B_j}$, et donc

$$\begin{aligned}
\int_A (f + g)\, d\mu &= \sum_{i=1}^{n} \sum_{j=1}^{m} (a_i + b_j) \mu(A_i \cap B_j \cap A) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} a_i \mu(A_i \cap B_j \cap A) + \sum_{i=1}^{n} \sum_{j=1}^{m} b_j \mu(A_i \cap B_j \cap A) \\
&= \sum_{i=1}^{n} a_i \mu(A_i \cap A) + \sum_{j=1}^{m} b_j \mu(B_j \cap A) \\
&= \int_A f d\mu + \int_A g\, d\mu,
\end{aligned}$$

comme souhaité. □

Pour montrer que le même résultat est vrai pour des fonctions positives mesurables générales, on va utiliser le résultat de convergence suivant. Ce résultat, dû à Lebesgue et Levi, est fondamental et montre à quel point l'intégrale de Lebesgue se comporte bien avec les limites. On en verra une version légèrement améliorée dans la suite. On rappelle que ce résultat est faux pour l'intégrale de Riemann (voir le contrexemple en introduction) ; ceci prouve en particulier que l'intégrale de Riemann ne peut être vue comme une intégrale de Lebesgue par rapport à une mesure.

Théorème 1.35 (Théorème de convergence monotone, V0)**.** *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure, f une fonction positive mesurable, et $(f_n)_n$ une suite croissante de fonctions simples positives mesurables telles que $f_n \uparrow f$. Alors*

$$\int f_n d\mu \uparrow \int f d\mu.$$

Démonstration. On a $f_1 \leq f_2 \leq \cdots \leq f$ et par monotonie de l'intégrale, il vient

$$\int f_1 d\mu \leq \int f_2 d\mu \leq \cdots \leq \int f d\mu,$$

de sorte que la limite $\alpha := \lim_n \int f_n d\mu = \sup_n \int f_n d\mu$ existe dans $[0, \infty]$ et

$$\alpha := \lim_n \int f_n d\mu \leq \int f d\mu.$$

Il reste à montrer que $\alpha \geq \int f d\mu$. Par définition de l'intégrale, il suffit de montrer que toute fonction simple positive mesurable $s \leq f$ satisfait $\int s d\mu \leq \alpha$.
Soit s une telle fonction simple positive mesurable avec $s \leq f$, et soit $\varepsilon > 0$. Notons $E_n = \{x \in X : f_n(x) \geq (1 - \varepsilon)s(x)\}$. Puisque la suite f_n est croissante, la suite E_n l'est aussi. De plus, on remarque que $\cup_n E_n = X$. En effet, pour tout $x \in X$,

— Si $s(x) = 0$, alors trivialement $x \in E_n$ pour tout n.

— Si $s(x) > 0$, alors on a $0 \leq (1 - \varepsilon)s(x) < s(x) \leq f(x)$, et la convergence $f_n(x) \uparrow f(x)$ assure qu'il existe n assez grand tel que $(1 - \varepsilon)s(x) < f_n(x)$, c'est-à-dire $x \in E_n$.

Or, par définition de E_n et par monotonie de l'intégrale, on peut écrire

$$\underbrace{\int_X f_n d\mu}_{\to\alpha} \geq \int_{E_n} f_n d\mu \geq (1 - \varepsilon) \int_{E_n} s d\mu.$$

Comme s est simple et comme $E_n \uparrow X$, la définition de l'intégrale de fonctions simples et la continuité à gauche des mesures impliquent $\int_{E_n} s d\mu \uparrow \int_X s d\mu$. Ceci prouve $\alpha \geq (1 - \varepsilon) \int s d\mu$. Comme $\varepsilon > 0$ est arbitraire, on conclut $\alpha \geq \int s d\mu$. □

Avec le résultat de convergence ci-dessus, on peut maintenant conclure la linéarité de l'intégrale de fonctions positives.

Corollaire 1.36. *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure, $f, g : X \to [0, \infty]$ mesurables, et $A \in \mathcal{A}$. Alors*

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

Démonstration. Par la Proposition 1.29, il existe des suites croissantes $(f_n)_n$ et $(g_n)_n$ de fonctions simples positives mesurables telles que $f_n \uparrow f$ et $g_n \uparrow g$. La suite $(f_n + g_n)_n$ est alors une suite croissante de fonctions simples positives mesurables avec $f_n + g_n \uparrow f + g$. Le théorème précédent et la linéarité de l'intégrale de fonctions simples donnent alors

$$\int_A (f + g)\, d\mu = \lim_n \int_A (f_n + g_n) d\mu = \lim_n \Big(\int_A f_n d\mu + \lim_n \int_A g_n d\mu \Big) = \int_A f d\mu + \int g d\mu. \qquad \square$$

Un second corollaire montre que séries et intégrales commutent sans hypothèse pour des fonctions positives : en effet, comme d'habitude, dès qu'on a un résultat sur des suites, on a un résultat correspondant sur les séries, via les sommes partielles.

Corollaire 1.37. *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $(f_n)_n$ une suite de fonctions mesurables. Alors*

$$\int_X \Big(\sum_n f_n\Big) d\mu = \sum_n \int_X f_n d\mu.$$

Démonstration. On a $s_N := \sum_{n=1}^N f_n \uparrow \sum_n f_n$ et on peut alors appliquer le théorème de convergence monotone. □

Exemple 1.38. Soit $(X, \mathcal{A}, \mu) = (\mathbb{N}, \mathcal{P}(\mathbb{N}), \sharp)$. Toute fonction $f : \mathbb{N} \to [0, \infty]$ est forcément $\mathcal{P}(\mathbb{N})$-mesurable et on vérifie que

$$\int_{\mathbb{N}} f d\sharp = \sum_{n=0}^{\infty} f(n).$$

En effet, cette égalité est une conséquence du théorème de convergence monotone avec $s_n(x) := f(x)\mathbb{1}_{x\leq n} \uparrow f(x)$ et $\int s_n d\sharp = \sum_{k=0}^n f(k)\sharp(\{k\}) = \sum_{k=0}^n f(k)$. En ces termes, le corollaire ci-dessus nous apprend que pour toute double suite $(a_{n,k})_{n,k} \in [0, \infty]$ on a le théorème de Fubini

$$\sum_{k=0}^{\infty}\sum_{n=0}^{\infty} a_{n,k} = \sum_{n=0}^{\infty}\sum_{k=0}^{\infty} a_{n,k}.$$

En effet, cette égalité suit du Corollaire 1.37 appliqué à la suite de fonctions définies par $f_n(k) = a_{n,k}$. On s'émeuvra devant la vertu unificatrice de la théorie.

1.5.3 Intégration — cas général

Rappelons que les parties positives et négatives d'une fonction f à valeurs dans $\overline{\mathbb{R}}$ sont définies par $f_+ = f \vee 0$ et par $f_- = (-f) \vee 0 = -(f \wedge 0)$ respectivement, et que ces fonctions sont mesurables si et seulement si f l'est. Le seul souci dans la définition est qu'il faut éviter les indéterminations $\infty - \infty$.

Définition 1.39. Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $f : X \to \overline{\mathbb{R}}$. On dit que f est intégrable au sens de Lebesgue sur X par rapport à μ si f est mesurable et si $\int_X |f| d\mu < \infty$. Dans ce cas, pour $A \in \mathcal{A}$, l'intégrale (de Lebesgue) de f sur A par rapport à μ est définie par

$$\int_A f d\mu := \int_A f^+ d\mu - \int_A f^- d\mu \in \mathbb{R}. \tag{1.1}$$

L'ensemble des fonctions intégrables $X \to \mathbb{R}$ (pas $\overline{\mathbb{R}}$) est noté $\mathcal{L}^1(X, \mathcal{A}, \mu)$ (ou simplement $\mathcal{L}^1(X)$ ou $\mathcal{L}^1(\mu)$ s'il n'y a pas de confusion possible). On dit par ailleurs que

— f est intégrable sur A si $f\mathbb{1}_A$ est intégrable ;

— l'intégrale de f sur A existe si $f\mathbb{1}_A$ est mesurable et si soit $\int_A f^+ d\mu$ soit $\int_A f^- d\mu$ est fini ; dans ce cas, on peut également définir l'intégrale de f sur A via (1.1), qui est alors bien défini dans $\overline{\mathbb{R}}$.

Sur l'espace entier, on note simplement $\int f d\mu = \int_X f d\mu$. Par ailleurs, on écrira parfois $\int_A f d\mu = \int_A f(x) d\mu(x) = \int_A f(x)\mu(dx)$.

Lemme 1.40. *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure, $f, g \in \mathcal{L}^1(X)$, et $\alpha \in \mathbb{R}$.*

(i) $\alpha f \in \mathcal{L}^1(X)$ *et* $\int \alpha f d\mu = \alpha \int f d\mu$,

(ii) $f + g \in \mathcal{L}^1(X)$ *et* $\int f + g\, d\mu = \int f d\mu + \int g\, d\mu$,

(iii) $f \leq g$ *sur* $X \implies \int f d\mu \leq \int g\, d\mu$.

En d'autres termes, $\mathcal{L}^1(X)$ *est un espace vectoriel et l'application* $\mathcal{L}^1(X, \mathcal{A}, \mu) \to \mathbb{R} : f \mapsto \int_X f d\mu$ *est une forme linéaire positive.* [1]

Démonstration. (i) Si $\alpha = 0$, alors $\alpha f = 0$ et $\int \alpha f d\mu = \alpha \int f d\mu$. Si $\alpha > 0$, alors $(\alpha f)^+ = \alpha f^+$ et $(\alpha f)^- = \alpha f^-$. Ces deux applications sont donc intégrables, ce qui implique l'intégrabilité de αf. De plus,

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

Si $\alpha < 0$, alors $(\alpha f)^+ = -\alpha f^-$ et $(\alpha f) = -\alpha f^+$, et on procède comme dans le cas précédent.

(ii) Puisque $(f+g)^+ \leq f^+ + g^+$ et $(f+g)^- \leq f^- + g^-$, la monotonie de l'intégrale implique que $(f+g)^+$ et $(f+g)^-$ sont intégrables, donc $f+g$ l'est également. On a

$$(f+g)^+ - (f+g)^- = f + g = f^+ - f^- + g^+ - g^- = (f^+ + g^+) - (f^- + g^-)$$

ou de façon équivalente,

$$(f+g)^+ + (f^- + g^-) = (f^+ + g^+) + (f+g)^-,$$

de sorte que la linéarité de l'intégrale de fonctions positives donne

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

et donc

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

(iii) Si $f \leq g$, alors $g - f \geq 0$ et par monotonie de l'intégrale pour les applications positives on obtient alors $\int g - f d\mu \geq 0$. Vu le point précédent, on en déduit que $\int g\, d\mu \geq \int f d\mu$. □

On montre ensuite l'inégalité triangulaire pour l'intégration. Pour $(X, \mathcal{A}, \mu) = (\mathbb{N}, \mathcal{P}(\mathbb{N}), \sharp)$, ceci correspond à l'inégalité triangulaire habituelle. L'hypothèse de mesurabilité de f est importante : il existe des applications non-mesurables (donc non-intégrables) dont la valeur absolue est intégrable.

Lemme 1.41 (Inégalité triangulaire)**.** *Soient* $(X, \mathcal{A}, \mu)$ *un espace de mesure et* $f : X \to \overline{\mathbb{R}}$ *une fonction mesurable. Alors* f *est intégrable si et seulement si* $|f|$ *est intégrable. Dans ce cas,*

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

Démonstration. Par définition, f est intégrable si et seulement si f^+ et f^- le sont. Puisque $|f| = f^+ + f^-$, la Proposition 1.36 donne que $|f|$ est intégrable si et seulement si f^+ et f^- le sont (un sens est dû à la linéarité, l'autre à la monotonie). Dans ce cas, par l'inégalité triangulaire dans $\mathbb{R}$, on a

$$\left| \int f d\mu \right| = \left| \int f^+ d\mu - \int f^- d\mu \right| \leq \int f^+ d\mu + \int f^- d\mu = \int |f| d\mu. \qquad \square$$

1. Sur un espace topologique Haussdorf localement compact X, on verra en fait que toute forme linéaire positive $C_c(X) \to \mathbb{R}$ peut s'écrire sous cette forme pour une certaine mesure (théorème de représentation de Riesz-Markov).

1.5.4 Intégration de fonctions à valeurs vectorielles

Soit $(X, \mathcal{A}, \mu)$ un espace de mesure. On rappelle qu'une fonction $f = (f_1, \dots, f_n) : X \to \mathbb{R}^n$ est mesurable si et seulement si les composantes $f_1, \dots, f_n : X \to \mathbb{R}$ sont toutes mesurables. Pour des fonctions à valeurs complexes, comme $\mathbb{C} = \mathbb{R}^2$, on déduit qu'une fonction $f : X \to \mathbb{C}$ est mesurable si et seulement si $\Re f$ et $\Im f$ sont mesurables $X \to \mathbb{R}$. L'intégration se généralise alors comme suit.

Définition 1.42. Une fonction $f = (f_1, \dots, f_n) : X \to \mathbb{R}^n$ est dite intégrable si toutes ses composantes le sont, et on définit alors son intégrale

$$\int f d\mu = \Big(\int f_1 d\mu, \dots, \int f_n d\mu \Big) \ \in \ \mathbb{R}^n.$$

Une fonction $f : X \to \mathbb{C}$ est dite intégrable si ses parties réelles et imaginaires le sont, et on définit alors son intégrale

$$\int f d\mu = \int \Re f d\mu + i \int \Im f d\mu \ \in \ \mathbb{C}.$$

De même que sur $\mathbb{R}$, on vérifie facilement la linéarité de l'intégrale de fonctions à valeurs dans $\mathbb{R}^n$ ou dans $\mathbb{C}$. De plus, on peut généraliser l'inégalité triangulaire.

Lemme 1.43 (Inégalité triangulaire)**.** *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $f : X \to \mathbb{R}^n$ ou $\mathbb{C}$ une fonction mesurable. Alors f est intégrable si et seulement si $|f|$ est intégrable. Dans ce cas,*

$$\Big| \int f d\mu \Big| \leq \int |f| d\mu.$$

Démonstration. On se focalise sur le cas d'une fonction mesurable $f = (f_1, \dots, f_n) : X \to \mathbb{R}^n$. Il est facile de voir que $|f|$ est mesurable. Puisque $|f| \leq |f_1| + \dots + |f_n|$, il est clair que $|f|$ est intégrable si les composantes de f le sont. De plus, si $|f|$ est intégrable, alors les composantes de f le sont également puisque $|f_i| \leq |f|$ pour tout i. Enfin, dans ce cas, il reste à montrer l'inégalité triangulaire, et on peut supposer que $\int f d\mu \neq 0$. Notons alors $\alpha = (\alpha_1, \dots, \alpha_n) = \int f d\mu / |\int f d\mu| \in \mathbb{R}^n$. Par définition de la norme sur $\mathbb{R}^n$, on a alors $|\alpha| = 1$ et $\alpha \cdot \int f d\mu = \sum_i \alpha_i \int f_i d\mu = |\int f d\mu|$. Dès lors, en utilisant l'inégalité triangulaire sur $\mathbb{R}$, on obtient

$$\Big| \int f d\mu \Big| = \sum_i \alpha_i \int f_i d\mu = \int \Big(\sum_i \alpha_i f_i \Big) d\mu,$$

et donc, par le Lemme 1.41, la monotonie de l'intégrale, et l'inégalité de Cauchy-Schwarz sur $\mathbb{R}^n$,

$$\Big| \int f d\mu \Big| \leq \int \Big| \sum_i \alpha_i f_i \Big| d\mu \leq \int |f| d\mu. \qquad \square$$

1.6 Ensembles de mesure nulle

On a vu que la valeur des fonctions sur des ensembles de mesure nulle n'a pas d'importance pour l'intégration : pour $f : X \to \overline{\mathbb{R}}$ mesurable et $A \in \mathcal{A}$ avec $\mu(A) = 0$, on a $\int_A f d\mu = 0$. En particulier, changer une fonction sur un ensemble de mesure nulle ne change pas son intégrale. Cela motive la définition suivante.

Définition 1.44. Soit $(X, \mathcal{A}, \mu)$ un espace de mesure. Un sous-ensemble N de X est dit *μ-négligeable* s'il existe un ensemble $A \in \mathcal{A}$ tel que $N \subset A$ et $\mu(A) = 0$.

Par monotonie de la mesure, on remarque qu'un ensemble $N \in \mathcal{A}$ est μ-négligeable si et seulement s'il est de mesure nulle, c'est-à-dire $\mu(N) = 0$. La définition ci-dessus est plus générale comme elle ne suppose par la mesurabilité de N. On définit à présent la notion de 'presque partout'.

Définition 1.45. Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $Y \subset X$. On dit qu'une propriété P est vraie *μ-presque partout dans Y* (on écrit μ-p.p.) si l'ensemble des points de Y pour lesquels cette propriété n'est pas vérifiée est μ-négligeable (c'est-à-dire, s'il existe $N \in \mathcal{A}$ tel que $\mu(N) = 0$ et P est vraie en tout point de $Y \setminus N$).

Lorsque le contexte est clair, on parle d'ensemble négligeable, et on dit qu'une propriété est vérifiée simplement "presque partout", ou "p.p.", sans faire référence à la mesure μ.

Exercice 1.46. Soit $(X, \mathcal{A}, \mu)$ un espace de mesure.

(i) Soient f et g deux fonctions définies sur X.
 — On dit que $f = g$ μ-presque partout si l'ensemble $\{x \in X : f(x) \neq g(x)\}$ est μ-négligeable. On note alors $f = g$ μ-p.p., ou $f(x) = g(x)$ $\mu - \forall' x$.
 — On dit que $f \geq g$ μ-presque partout si l'ensemble $\{x \in X : f(x) < g(x)\}$ est μ-négligeable. On note alors $f \geq g$ μ-p.p., ou $f(x) \geq g(x)$ $\mu - \forall' x$.

 Remarquons que les ensembles $\{x \in X : f(x) \neq g(x)\}$ et $\{x \in X : f(x) < g(x)\}$ appartiennent à $\mathcal{A}$ si f et g sont mesurables.

(ii) Soit $(f_n)_n$ une suite d'applications définies sur X et f une application définie sur X. On dit que la suite $(f_n)_n$ converge μ-p.p. vers f si l'ensemble $\{x \in X : \lim_n f_n(x) \neq f(x)\}$ est μ-négligeable.

On observe que l'égalité p.p. définit une relation d'équivalence sur les fonctions mesurables. Pour l'intégration, comme le montre le lemme suivant, ce sont seulement les classes d'équivalence correspondantes qui importent. En notant $f \sim g$ si $f = g$ p.p., cela motive de définir l'espace quotient $L^1(X, \mathcal{A}, \mu) := \mathcal{L}^1(X, \mathcal{A}, \mu)/\sim$, soit l'ensemble des classes d'équivalence de fonctions intégrables égales presque partout. On reviendra plus en détail sur ce point dans la suite, mais insistons qu'il s'agit ici d'un tournant conceptuel important : on ne parlera plus de fonctions définies en tout point, mais de classes d'équivalences de fonctions égales presque partout.

Lemme 1.47. *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $f, g : X \to \mathbb{R}$ mesurables. Si $f = g$ μ-p.p. et $f \in \mathcal{L}^1$, alors $g \in \mathcal{L}^1$ et $\int_A f = \int_A g$ pour tout $A \in \mathcal{A}$.*

Démonstration. Par hypothèse, il existe $N \in \mathcal{A}$, $\mu(N) = 0$, tel que $f(x) = g(x)$ pour tout $x \in N^c$. On peut alors décomposer

$$\int_X |g| = \int_{N^c} |g| + \underbrace{\int_N |g|}_{=0} = \int_{N^c} |f| < \infty.$$

Par ailleurs,

$$\int_A g = \int_{A \cap N^c} g + \underbrace{\int_{A \cap N} g}_{=0} = \int_{A \cap N^c} f + \underbrace{\int_{A \cap N} f}_{=0} = \int_A f.$$

□

On montre à présent quelques résultats utiles, qui montrent que certaines propriétés des intégrales peuvent impliquer des propriétés presque partout des fonctions.

Lemme 1.48. *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $f : X \to \overline{\mathbb{R}}$ mesurable.*

(i) $\int |f| d\mu = 0 \implies f = 0$ *μ-p.p.*

(ii) $f \in \mathcal{L}^1(X) \implies |f| < \infty$ *μ-p.p.*

Avant de passer à la preuve de ce lemme, on rappelle l'inégalité de Markov, qui est bien connue dans le cadre probabiliste. Il s'agit d'un outil crucial pour des intégrales à la mesure des ensembles de niveau.

Lemme 1.49 (Inégalité de Markov)**.** *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $f : X \to \overline{\mathbb{R}}$ mesurable. Pour tout $t \geq 0$, on a*

$$\mu(\{x : |f(x)| \geq t\}) \leq \tfrac{1}{t} \int_{\{x:|f(x)|\geq t\}} |f| d\mu \leq \tfrac{1}{t} \int |f| d\mu.$$

Démonstration. En notant $A_t := \{x \in X : |f(x)| \geq t\}$ et en remarquant $\mathbb{1}_{A_t} \leq \frac{1}{t}|f|\mathbb{1}_{A_t}$, on obtient par monotonie de l'intégrale,

$$\mu(\{x \in X : |f(x)| \geq t\}) = \mu(A_t) = \int \mathbb{1}_{A_t} d\mu \leq \int_{A_t} \tfrac{1}{t}|f| d\mu \leq \tfrac{1}{t} \int |f| d\mu. \qquad \square$$

Démonstration du Lemme 1.48. (i) Si $\int |f| d\mu = 0$, l'inégalité de Markov donne pour tout $t \geq 0$,

$$\mu(\{x \in X : |f(x)| \geq t\}) \leq \tfrac{1}{t} \int |f| d\mu = 0,$$

et donc, en décomposant

$$\{x \in X : f(x) \neq 0\} = \bigcup_{n=1}^{\infty} \{x \in X : |f(x)| \geq 1/n\},$$

la continuité de la mesure donne

$$\mu(\{x \in X : f(x) \neq 0\}) = \lim_n \mu(\{x \in X : |f(x)| \geq 1/n\}).$$

(ii) Par l'inégalité de Markov, on peut borner

$$\mu(\{x \in X : |f(x)| = \infty\}) \leq \mu(\{x \in X : |f(x)| \geq n\}) \leq \tfrac{1}{n} \int_X |f| d\mu,$$

et on note que le membre de droite tend vers 0 quand $n \uparrow \infty$ pourvu que $\int_X |f| d\mu < \infty$. $\square$

1.6.1 Espace de mesures complets

L'ensemble des points pour lesquels une propriété est vérifiée n'est pas nécessairement un ensemble de $\mathcal{A}$. En particulier, si f, g ne sont pas mesurable, l'ensemble $\{x \in X : f(x) \neq g(x)\}$ n'est en général pas mesurable. De même, dans le Lemme 1.47 ci-dessus, la mesurabilité de f et l'égalité $f = g$ p.p. ne suffit pas à assurer la mesurabilité de g. Cela motive la définition suivante.

Définition 1.50. Soit $(X, \mathcal{A}, \mu)$ un espace de mesure. La mesure μ est dite *complète* si tout ensemble μ-négligeable est un élément de $\mathcal{A}$. En d'autres termes, la mesure μ est complète si pour tout $A \in \mathcal{A}$ avec $\mu(A) = 0$ on a $N \subset A \implies N \in \mathcal{A}$. Alternativement, on dit alors que l'espace de mesure $(X, \mathcal{A}, \mu)$ est complet.

Lemme 1.51. *Soit $(X, \mathcal{A}, \mu)$ un espace de mesure complet.*

(i) Étant donné deux fonctions $f, g : X \to \overline{\mathbb{R}}$, si $f = g$ μ-p.p. et si f est mesurable, alors g est mesurable.

(ii) Si $(f_n)_n$ est une suite de fonctions mesurables $f_n : X \to \overline{\mathbb{R}}$, et si $f_n \to f$ p.p. pour une fonction $f : X \to \overline{\mathbb{R}}$, alors la limite f est mesurable.

Démonstration. (i) Posons $N = \{x : f(x) \neq g(x)\}$. Par hypothèse, N est μ-négligeable et comme la mesure est complète, on en tire que $N \in \mathcal{A}$ et $\mu(N) = 0$. Alors pour tout $b \in \mathbb{R}$, il vient

$$\{x : g(x) < b\} = \big(\{x : f(x) < b\} \cap N^c\big) \cup \big(\{x : g(x) < b\} \cap N\big).$$

Comme μ est complète, on sait que $\{x : g(x) < b\} \cap N \in \mathcal{A}$, d'où la conclusion.

(ii) On sait que $\liminf_n f_n$ est mesurable. Comme f et $\liminf_n f_n$ sont égales μ-presque partout, le Lemme 1.51 implique que f est mesurable. □

Insistons que le résultat précédent est faux si μ n'est pas une mesure complète. En effet, soit N un ensemble μ-négligeable n'appartenant pas à $\mathcal{A}$. Alors la fonction caractéristique $\mathbb{1}_N$ et la fonction constante 0 sont égales μ-p.p. Néanmoins, la fonction $\mathbb{1}_N$ n'est pas mesurable alors que la fonction constante 0 l'est. De même, si on prend la suite de fonctions dont chaque terme est égal à la fonction 0, cette suite converge μ-p.p. vers la fonction $\mathbb{1}_N$.

1.6.2 Complétion de mesures

Au vu de la section précédente et du caractère agréable des espaces de mesure complets, on peut se demander s'il n'est pas possible de toujours compléter un espace de mesure en lui "ajoutant" ses ensembles négligeables.

Définition 1.52. Soit $(X, \mathcal{A}, \mu)$ un espace de mesure. La *complétion* de $\mathcal{A}$ par rapport à μ est la collection $\bar{\mathcal{A}}$ des sous-ensembles A de X pour lesquels il existe $E, F \in \mathcal{A}$ tels que

$$E \subset A \subset F \quad \text{et} \quad \mu(F \setminus E) = 0.$$

Lemme 1.53. *Soit $(X, \mathcal{A}, \mu)$ un espace de mesure. La complétion $\bar{\mathcal{A}}$ de $\mathcal{A}$ est une σ-algèbre qui contient $\mathcal{A}$.*

Démonstration. Il est clair que $\mathcal{A} \subset \bar{\mathcal{A}}$ (en prenant $E = F = A$). En particulier, $X \in \bar{\mathcal{A}}$. De plus, si $E \subset A \subset F$ et $\mu(F \setminus E) = 0$, alors $F^c \subset A^c \subset E^c$ et $\mu(E^c \setminus F^c) = \mu(F \setminus E) = 0$. Donc $\bar{\mathcal{A}}$ est stable par passage au complémentaire. Enfin, si $(A_n)_n$ est une suite d'éléments de $\bar{\mathcal{A}}$, alors pour tout n il existe $E_n, F_n \in \mathcal{A}$ tels que $E_n \subset A_n \subset F_n$ et $\mu(F_n \setminus E_n) = 0$. Posons $E = \cup_n E_n$ et $F = \cup_n F_n$. Clairement, $E, F \in \mathcal{A}$ et $E \subset \cup_n A_n \subset F$. De plus,

$$\mu(F \setminus E) \leq \mu\bigg(\bigcup_n F_n \setminus E_n\bigg) \leq \sum_n \mu(F_n \setminus E_n) = 0$$

et donc $\cup_n A_n \in \bar{\mathcal{A}}$. □

Remarquons que si E et F sont des ensembles comme à la Définition 1.52, on obtient directement que $\mu(E) = \mu(F)$. De plus, pour tout $B \in \mathcal{A}$ avec $B \subset A$, on a $\mu(B) \leq \mu(F) = \mu(E)$, ce qui implique

$$\mu(E) = \sup\big\{\mu(B) : B \in \mathcal{A}, B \subset A\big\},$$

et la valeur commune de $\mu(E)$ et $\mu(F)$ ne dépend donc que de A et μ. Ceci motive la définition suivante pour la complétion de la mesure μ.

Définition 1.54. Soit $(X, \mathcal{A}, \mu)$ un espace de mesure. La *complétion* $\overline{\mu}$ de μ est l'application

$$\overline{\mu} : \bar{\mathcal{A}} \to [0, \infty] : A \mapsto \sup\big\{\mu(B) : B \in \mathcal{A}, B \subset A\big\}.$$

Proposition 1.55. *Soit $(X, \mathcal{A}, \mu)$ un espace de mesure. La complétion $\overline{\mu}$ de μ est une mesure complète sur $(X, \bar{\mathcal{A}})$ telle que $\overline{\mu}(A) = \mu(A)$ pour tout $A \in \mathcal{A}$.*

Démonstration. Il est clair que $\overline{\mu}$ est une application qui prolonge μ. En particulier, $\overline{\mu}(\emptyset) = 0$. Soit $(A_n)_n$ une suite d'éléments deux à deux disjoints de $\bar{\mathcal{A}}$. Pour tout n, il existe $E_n, F_n \in \mathcal{A}$ tels que $E_n \subset A_n \subset F_n$ et $\mu(F_n \setminus E_n) = 0$. Bien sûr, les ensembles $(E_n)_n$ sont deux à deux disjoints et en procédant comme dans la preuve de la Lemme 1.53, on a

$$\overline{\mu}\Big(\bigcup_n A_n\Big) = \mu\Big(\bigcup_n E_n\Big) = \sum_n \mu(E_n) = \sum_n \overline{\mu}(A_n)$$

puisque μ est une mesure. Enfin, $\overline{\mu}$ est complète par construction. □

Exemple 1.56. Un exemple trivial d'espace de mesure non complet est le suivant : $X = \mathbb{R}$, $\mathcal{A} = \{\varnothing, \mathbb{R}\}$, et $\mu \equiv 0$. Sa complétion est $\bar{\mathcal{A}} = \mathcal{P}(\mathbb{R})$ et $\bar{\mu} \equiv 0$.

1.7 Théorèmes limites

Dans cette section, nous démontrons les résultats classiques concernant le passage à la limite dans l'intégrale de Lebesgue. On commence par rappeler le théorème de convergence monotone, cf. Théorème 1.35. On généralise légèrement l'énoncé pour ne demander la validité des hypothèses que p.p.

Théorème 1.57 (Théorème de convergence monotone)**.** *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure, $(f_n)_n$ une suite de fonctions mesurables $X \to [0, \infty]$, et $f : X \to [0, \infty]$ mesurable telles que $f_n \uparrow f$ p.p. (c'est-à-dire $f_n \leq f_{n+1}$ p.p. pour tout n et $f_n \to f$ p.p.). Alors*

$$\int f_n d\mu \uparrow \int f d\mu.$$

Démonstration. Soit $M = \{x \in X : f_n(x) \leq f_{n+1}(x)\ \forall n,\ \lim_n f_n(x) = f(x)\}$. Par mesurabilité de f_n, f, cet ensemble M est mesurable et son complémentaire s'écrit $M^c = \cup_n\{x \in X : f_n(x) > f_{n+1}(x)\} \cup \{x \in X : \liminf_n f_n(x) \neq f(x)\} \cup \{x \in X : \limsup_n f_n(x) \neq f(x)\}$, qui est de mesure nulle par hypothèse. Il suffit à présent d'appliquer Théorème 1.35 à $f_n \mathbb{1}_M$ et $f \mathbb{1}_M$. □

Le deuxième théorème fondamental de convergence est le lemme de Fatou. On remarque que l'égalité dans ce lemme ne peut être espérée en général : par exemple, étant donné $X = A \cup B$ avec $A \cap B = \varnothing$, $A, B \in \mathcal{A}$, et $\mu(A), \mu(B) > 0$, si on considère la suite $f_{2n} = 1_A$, $f_{2n+1} = 1_B$, on a $\liminf_n f_n = 0$ mais $\liminf_n \int_X f_n d\mu = \min\{\mu(A), \mu(B)\} > 0$. Même s'il ne donne qu'une inégalité, ce résultat sera extrêmement utile.

Théorème 1.58 (Lemme de Fatou)**.** *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $(f_n)_n$ une suite de fonctions mesurables $X \to [0, \infty]$. Alors*

$$\int \Big(\liminf_n f_n\Big) d\mu \leq \liminf_n \int f_n d\mu.$$

Démonstration. Pour tout n, posons $g_n = \inf_{k \geq n} f_k$. Alors chaque fonction g_n est mesurable et la suite $(g_n)_n$ est croissante. De plus, par définition, $\liminf_n f_n = \lim_n g_n$. Le théorème de la convergence monotone donne alors

$$\int \liminf_n f_n d\mu = \int \lim_n g_n d\mu = \lim_n \int g_n d\mu \leq \liminf_n \int f_n d\mu$$

où la dernière inégalité suit de la monotonie de l'intégrale avec $g_n \leq f_n$ pour tout n. □

Enfin, on arrive à présent au théorème de convergence dominée de Lebesgue, qui par sa généralité et son applicabilité justifie à lui seul l'intérêt de la théorie de Lebesgue.

Théorème 1.59 (Théorème de convergence dominée de Lebesgue). *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure, $(f_n)_n$ une suite de fonctions mesurables $X \to \overline{\mathbb{R}}$ et $f : X \to \overline{\mathbb{R}}$ une fonction mesurable telles que $f_n \to f$ p.p. Supposons qu'il existe une fonction intégrable $g : X \to [0, \infty]$ telle que $|f_n| \leq g$ p.p. pour tout n. Alors les applications f_n et f sont intégrables pour tout n et l'on a*

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

En particulier,

$$\int f_n d\mu \to \int f d\mu.$$

Démonstration. Comme g est intégrable, on note que $|g| < \infty$ p.p. Par hypothèse, il existe alors un ensemble $N \in \mathcal{A}$ avec $\mu(N) = 0$ tel que $|f_n(x)| \leq g(x) < \infty$ et $f_n(x) \to f(x)$ pour tout $x \in N^c$. On pose pour tout n,

$$h_n(x) = 2g(x) - |f_n(x) - f(x)| \mathbb{1}_{N^c}(x).$$

On remarque que $(h_n)_n$ est une suite de fonctions mesurables à valeurs dans $[0, \infty]$ et que $h_n \to 2g$ simplement. Le lemme de Fatou implique alors

$$\begin{aligned}\int 2g d\mu \;\leq\; \liminf_n \int h_n \, d\mu &= \liminf_n \left(\int 2g d\mu - \int |f_n - f| \, d\mu \right) \\ &= \int 2g \, d\mu - \limsup_n \int |f_n - f| \, d\mu,\end{aligned}$$

d'où

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

et la limite vaut donc également 0. Enfin, l'inégalité triangulaire donne

$$\left| \int f_n d\mu - \int f d\mu \right| = \left| \int (f - f_n) d\mu \right| \leq \int |f - f_n| d\mu,$$

et on conclut $\lim_n \int f_n d\mu = \int f d\mu$. □

1.7.1 Application : dérivation sous le signe intégral

Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $U \subset \mathbb{R}^d$ un ouvert. On considère les propriétés de régularité sur U de fonctions de la forme $F(y) = \int_X f(x, y) d\mu(x)$ avec $f : X \times U \to \mathbb{R}$.

Proposition 1.60. *Étant donné un point $a \in U$, supposons*

(i) $\forall y \in U$, $x \mapsto f(x, y)$ *est dans* $\mathcal{L}^1(X, \mathcal{A}, \mu)$ *;*

(ii) $\forall' x \in X$, $y \mapsto f(x, y)$ *est continue en* a *;*

(iii) $\exists g \in \mathcal{L}^1(X, \mathcal{A}, \mu)$ *tel que* $\forall' x \in X$ *on a* $\sup_{y \in U} |f(x, y)| \leq g(x)$.

Alors la fonction $F(y) := \int_X f(x, y) d\mu(x)$ est aussi continue en a.

Démonstration. Soit $(y_n)_n$ une suite de points de U avec $y_n \to a$. On veut montrer que $F(y_n) \to F(a)$. Par définition, on a $F(y_n) = \int_X \varphi_n d\mu$ où $\varphi_n(x) := f(x, y_k)$ satisfait $|\varphi_n| \leq g$ p.p. et $\varphi_n \to f(\cdot, a)$ p.p. La conclusion est alors une conséquence de la convergence dominée. □

Proposition 1.61. *Supposons*

(i) $\forall y \in U$, $x \mapsto f(x,y)$ *est dans* $\mathcal{L}^1(X, \mathcal{A}, \mu)$ *;*

(ii) $\forall' x \in X$, $y \mapsto f(x,y)$ *est dérivable sur* U *dans la direction* e_1 *;*

(iii) $\exists g \in \mathcal{L}^1(X, \mathcal{A}, \mu)$ *tel que* $\forall' x \in X$ *on a* $\sup_{y \in U} |\partial f / \partial y_1(x,y)| \leq g(x)$.

Alors la fonction $F(y) := \int_X f(x,y) d\mu(x)$ *est aussi dérivable sur* U *dans la direction* e_1 *et on peut dériver sous le signe :* $\partial F / \partial y_1(y) = \int_X \varphi(x,y) d\mu(x)$ *où pour presque tout* x *on a* $\varphi(x, \cdot) = \partial f / \partial y_1(x, \cdot)$.

Démonstration. Soit $a \in U$ et soit $(h_n)_n \subset \mathbb{R} \setminus \{0\}$ avec $h_n \to 0$. On considère le quotient différentiel

$$\frac{F(a + h_n e_1) - F(a)}{h_n} = \int_X \psi_n d\mu, \qquad \psi_n(x) := \frac{f(x, a + h_n e_1) - f(x,a)}{h_n}.$$

Par hypothèse, $\psi_n \to \partial_{y_1} f(\cdot, a)$ p.p. De plus, pour presque tout x, le théorème des accroissements finis donne $|\psi_n(x)| \leq \sup_{\theta \in [0,1]} |\partial_{y_1} f(x, a + \theta h_n e_1)| \leq g(x)$. La conclusion est alors une conséquence de la convergence dominée. $\square$

1.8 Mesure image

On observe qu'une application $f : X \to Y$ permet de “projeter” (ou *push-forward*) toute σ-algèbre sur X en une σ-algèbre sur Y, ainsi que toute mesure sur X en une mesure sur Y.

Lemme 1.62 (Mesure image)**.** *Soit* $(X, \mathcal{A}, \mu)$ *un espace de mesure, soit* Y *un ensemble, et soit* $f : X \to Y$ *une fonction.*

(i) La famille $f_* \mathcal{A} := \{A \subset Y : f^{-1}(A) \in \mathcal{A}\}$ *est une* σ*-algèbre sur* Y*, qu'on appelle la* σ*-algèbre image de* $\mathcal{A}$ *par* f*. Si* Y *est muni d'une* σ*-algèbre* $\mathcal{A}'$ *et si* f *est* $(\mathcal{A}, \mathcal{A}')$*-mesurable, alors* $f_* \mathcal{A} \subset \mathcal{A}'$.

(ii) L'application $f_* \mu : f_* \mathcal{A} \to [0, \infty]$ *définie par* $f_* \mu(A) = \mu(f^{-1}(A))$ *est une mesure sur* $f_* \mathcal{A}$*, qu'on appelle la mesure image de* μ *par* f.

(iii) Une fonction $h : Y \to \overline{\mathbb{R}}$ *est* $f_* \mathcal{A}$*-mesurable si et seulement si* $h \circ f : X \to \overline{\mathbb{R}}$ *est* $\mathcal{A}$*-mesurable. Pour une fonction* $f_* \mathcal{A}$*-mesurable* $h : Y \to [0, \infty]$*, on obtient alors*

$$\int_Y h d(f_* \mu) = \int_X (h \circ f) d\mu. \tag{1.2}$$

Enfin, une fonction $h : Y \to \overline{\mathbb{R}}$ *appartient à* $\mathcal{L}^1(Y, f_* \mathcal{A}, f_* \mu)$ *si et seulement si* $h \circ f \in \mathcal{L}^1(X, \mathcal{A}, \mu)$*, et dans ce cas la même égalité (1.2) a lieu.*

Démonstration. (i) On a $f^{-1}(Y) = X \in \mathcal{A}$, et donc $Y \in f_* A$. Si $A \in f_* \mathcal{A}$, on a $f^{-1}(A^c) = f^{-1}(A)^c \in \mathcal{A}$, et donc $A^c \in f_* \mathcal{A}$. Enfin, si $(A_n)_n \subset f_* \mathcal{A}$, alors $f^{-1}(\cup_n A_n) = \cup_n f^{-1}(A_n) \in \mathcal{A}$, donc $\cup_n A_n \in f_* \mathcal{A}$.

(iii) L'équivalence des notions de mesurabilité suit la définition. On se tourne à présent vers la preuve de (1.2). On commence par montrer ce résultat pour $h = \mathbb{1}_A$ avec $A \in f_* \mathcal{A}$. Dans ce cas, on a $\int_Y \mathbb{1}_A d(f_* \mu) = f_* \mu(A) = \mu(f^{-1}(A))$ et $\int_X (\mathbb{1}_A \circ f) d\mu = \mu(f^{-1}(A))$, ce qui prouve (1.2). Par linéarité, on déduit que (1.2) est valable pour toute fonction h simple mesurable positive. Par convergence monotone, on déduit alors (1.2) pour toute fonction h mesurable positive. L'équivalence des notions d'intégrabilité est une conséquence de (1.2). $\square$

1.8.1 Application : probabilités et espérance

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilité, c'est-à-dire un espace de mesure avec $\mathbb{P}(\Omega) = 1$. On rappelle que l'ensemble Ω est interprété comme l'ensemble des réalisations $\omega \in \Omega$ (l'"univers"), la σ-algèbre comme l'ensemble des événements $A \in \mathcal{A}$, et $\mathbb{P}(A)$ comme la probabilité de l'événement A. Dans ce cadre, une *variable aléatoire* est définie comme une fonction mesurable $X : \Omega \to \overline{\mathbb{R}}$. Celle-ci associe donc à toute réalisation $\omega \in \Omega$ une valeur $X(\omega) \in \overline{\mathbb{R}}$. Si la variable aléatoire X est positive ou intégrable, alors son *espérance* est définie comme son intégrale

$$\mathbb{E}[X] := \int_\Omega X d\mathbb{P}.$$

Par ailleurs, par la construction ci-dessus, on remarque qu'une variable aléatoire X engendre une mesure image $\mathbb{P}_X := X_*\mathbb{P}$ sur $(\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$, que l'on appelle la *loi de* X,

$$\mathbb{P}_X(B) := X_*\mathbb{P}(B) := \mathbb{P}(X^{-1}(B)) = \mathbb{P}(\{\omega : X(\omega) \in B\}), \qquad B \in \mathcal{B}(\overline{\mathbb{R}}).$$

En ces termes, pour X positive ou intégrable, le résultat (1.2) (avec $h(t) = t$) montre que l'espérance peut s'écrire comme une intégrale sur $\overline{\mathbb{R}}$ par rapport à la loi de X,

$$\mathbb{E}[X] = \int_\Omega X d\mathbb{P} = \int_{\overline{\mathbb{R}}} t d\mathbb{P}_X(t).$$

Chapitre 2

Mesure de Lebesgue

Dans ce chapitre, le but est de construire la mesure de Lebesgue $\mathcal{L}$ sur $\mathbb{R}$ (et plus généralement $\mathcal{L}_d$ sur $\mathbb{R}^d$), soit la mesure qui généralise la notion de longueur sur $\mathbb{R}$ (ou d'aire sur $\mathbb{R}^2$, volume sur $\mathbb{R}^3$, etc.) Plus précisément, on va construire une unique mesure $\mathcal{L}$ sur $(\mathbb{R}, \mathbb{B})$ telle que $\mu([a, b]) = b - a$ pour tous $a \leq b$.

2.1 Unicité de mesures : théorème de classe monotone

Sur un espace vectoriel, on rappelle que deux applications linéaires égales sur une base vectorielle coïncident sur l'espace entier. De même, on peut se demander s'il est vrai que deux mesures égales sur un sous-ensemble de parties qui engendrent la σ-algèbre sont forcément égales. En général, la réponse à cette question est cependant négative : on va voir qu'il faut supposer davantage sur la structure du sous-ensemble de parties sur lesquelles les mesures sont supposées égales.

La première question est de comprendre la structure du sous-ensemble de parties sur lesquelles deux mesures peuvent coïncider. Cela nous mène à la définition suivante.

Définition 2.1. Soit X un ensemble. Une collection $\mathcal{M}$ de parties de X est une *classe monotone* (ou *classe de Dynkin*, ou λ-*système*) si

— $X \in \mathcal{M}$,

— $A \in \mathcal{M} \implies A^c \in \mathcal{M}$,

— $(A_n)_n \subset \mathcal{M}$ deux à deux disjoints $\implies \cup_n A_n \in \mathcal{M}$.

Exemple 2.2. Soient $(X, \mathcal{A})$ un espace de mesure et μ, ν des mesures finies sur $(X, \mathcal{A})$ avec $\mu(X) = \nu(X)$. Alors l'ensemble $\mathcal{M} = \{A \in \mathcal{A} : \mu(A) = \nu(A)\}$ est une classe monotone.

Bien sûr, toute σ-algèbre est une classe monotone. La réciproque n'est cependant pas vraie, mais le résultat suivant (dû à Sierpinski et popularisé par Dynkin) caractérise précisément ce qui manque pour qu'une classe monotone soit une σ-algèbre.

Lemme 2.3 (Lemme de classe monotone, ou théorème π-λ)**.** *Toute classe monotone stable par intersection finie est une σ-algèbre. En d'autres termes, un λ-système qui est aussi un π-système (i.e. un sous-ensemble de parties stable par intersection finie) est nécessairement un σ-algèbre.*

Démonstration. Soit $\mathcal{M}$ une classe monotone sur un ensemble X, que l'on suppose stable par intersections finies. Pour $(A_n)_n \subset \mathcal{M}$, on peut alors considérer

$$B_0 = A_0 \quad \text{et} \quad B_{n+1} = A_{n+1} \setminus \bigcup_{j=0}^{n} A_j = A_{n+1} \cap \bigcap_{j=0}^{n} A_j^c,$$

et on remarque que $(B_n)_n$ est une suite d'éléments de $\mathcal{M}$ deux à deux disjoints dont l'union est égale à $\cup_n A_n$. Comme $\mathcal{M}$ est une classe monotone, on déduit $\cup_n A_n = \cup_n B_n \in \mathcal{M}$. $\square$

Comme dans le cas de σ-algèbres, il est aisé de vérifier que l'intersection de classes monotones est encore une classe de Dynkin, ce qui permet de définir la plus petite classe de Dynkin qui contient une collection donnée $\mathcal{F}$ de parties de X.

Lemme 2.4. *Soient X un ensemble et $\mathcal{F} \subset \mathcal{P}(X)$. Il existe une plus petite (au sens de l'inclusion) classe monotone sur X qui contient $\mathcal{F}$. On l'appelle la* classe de Dynkin engendrée par $\mathcal{F}$ *et on la note $\lambda(\mathcal{F})$.*

En ces termes, on peut à présent démontrer la généralisation suivante du lemme de classe monotone : il suffit que $\mathcal{F}$ soit un π-système pour que la classe monotone engendrée $\lambda(\mathcal{F})$ soit également un π-système et donc une σ-algèbre.

Théorème 2.5 (Théorème de classe monotone)**.** *Soit X un ensemble. Si $\mathcal{F} \subset \mathcal{P}(X)$ est stable par intersection finie, alors $\lambda(\mathcal{F}) = \sigma(\mathcal{F})$.*

Démonstration. Comme toute σ-algèbre est une classe monotone, on a $\lambda(\mathcal{F}) \subset \sigma(\mathcal{F})$, et il reste à montrer l'autre inclusion. Par le lemme de classe monotone, il suffit de montrer que $\lambda(\mathcal{F})$ est stable par intersection finie : ce sera alors une σ-algèbre qui contient $\mathcal{F}$, donc aussi $\sigma(\mathcal{F})$. Pour tout $B \in \lambda(\mathcal{F})$, on pose

$$\mathcal{M}_B = \big\{A \in \lambda(\mathcal{F}) : A \cap B \in \lambda(\mathcal{F})\big\}.$$

Il est clair que $\mathcal{M}_B$ contient X et est stable par union disjointe. De plus, si $A \in \mathcal{M}_B$, on a

$$A^c \cap B = (A \cap B)^c \cap B = \big((A \cap B) \cup B^c\big)^c \in \lambda(\mathcal{F}),$$

ce qui montre que $\mathcal{M}_B$ est une classe monotone. Si $B \in \mathcal{F}$, alors $\mathcal{F} \subset \mathcal{M}_B$ puisque $\mathcal{F}$ est stable par intersection finie. Ainsi, $\lambda(\mathcal{F}) \subset \mathcal{M}_B$. Par conséquent, si $A \in \lambda(\mathcal{F})$ et si $B \in \mathcal{F}$, alors $A \cap B \in \lambda(\mathcal{F})$, c'est-à-dire $B \in \mathcal{M}_A$. On en tire que pour tout $A \in \lambda(\mathcal{F})$, $\mathcal{F} \subset \mathcal{M}_A$, d'où $\lambda(\mathcal{F}) \subset \mathcal{M}_A$. Au final, on a donc que pour tous $A, B \in \lambda(\mathcal{F})$, $A \cap B \in \lambda(\mathcal{F})$, d'où la conclusion. $\square$

Grâce à ce théorème, on peut déduire un critère fondamental pour l'unicité de mesures. On commence par le cas de mesures finies.

Corollaire 2.6 (Unicité de mesures, V0)**.** *Soient $(X, \mathcal{A})$ un espace mesurable et $\mathcal{F} \subset \mathcal{A}$ une collection stable par intersection finie, avec $X \in \mathcal{F}$ et $\sigma(\mathcal{F}) = \mathcal{A}$. Si μ_1 et μ_2 sont deux mesures finies sur $(X, \mathcal{A})$ qui sont égales sur $\mathcal{F}$, alors on a $\mu_1 = \mu_2$ sur $\mathcal{A}$.*

Démonstration. Soit $\mathcal{M} = \{A \in \mathcal{A} : \mu_1(A) = \mu_2(A)\}$. On a déjà vu que $\mathcal{M}$ est une classe monotone. Par le théorème de classe monotone, avec les hypothèses sur $\mathcal{F}$, on déduit $\mathcal{A} = \sigma(\mathcal{F}) = \lambda(\mathcal{F}) \subset \mathcal{M}$, d'où la conclusion. $\square$

Or, la mesure de Lebesgue sur $\mathbb{R}$ ne sera bien sûr pas une mesure finie. Il est donc nécessaire d'étendre ce résultat d'unicité à des mesures plus générales. Pour ce faire, on introduit la notion d'espace de mesures σ-finis. C'est en quelque sorte un équivalent en théorie de la mesure de la notion d'espace localement compact en topologie.

Définition 2.7. Soit $(X, \mathcal{A}, \mu)$ un espace de mesure et $\mathcal{F} \subset \mathcal{A}$. On dit que la mesure μ est *σ-finie sur $\mathcal{F}$* s'il existe une suite croissante $(A_n)_n \subset \mathcal{F}$ telle que

$$X = \bigcup_n A_n \qquad \text{et} \qquad \mu(A_n) < \infty \quad \forall n \in \mathbb{N}.$$

Si $\mathcal{F} = \mathcal{A}$, on dit simplement que μ est *σ-finie.* Alternativement, on dira aussi que l'espace de mesure $(X, \mathcal{A}, \mu)$ est *σ-fini.*

Remarque 2.8. De façon équivalente, un espace est σ-fini si et seulement s'il existe une suite $(C_n)_n \subset \mathcal{A}$ d'ensembles deux à deux disjoints tels que $X = \cup_n C_n$ et $\mu(C_n) < \infty$. En effet, il suffit de poser $C_n = A_n \setminus \bigcup_{m<n} A_m$.

Corollaire 2.9 (Unicité de mesures, V1)**.** *Soit $(X, \mathcal{A})$ un espace mesurable et $\mathcal{F} \subset \mathcal{A}$ une collection stable par intersection finie, avec $\sigma(\mathcal{F}) = \mathcal{A}$. Si μ_1 et μ_2 sont deux mesures sur $(X, \mathcal{A})$ qui sont égales sur $\mathcal{F}$, et si μ_1 est σ-finie sur $\mathcal{F}$, alors on a $\mu_1 = \mu_2$ sur $\mathcal{A}$.*

Démonstration. Par hypothèse, nous pouvons considérer une suite croissante $(A_n)_n \subset \mathcal{F}$ telle que $X = \cup_n A_n$ et $\mu_1(A_n) < \infty$ pour tout n. Pour tout n, on pose alors $\mu_{1,n}(A) := \mu_1(A \cap A_n)$ et $\mu_{2,n}(A) := \mu_2(A \cap A_n)$, ce qui définit des mesures finies sur $(X, \mathcal{A})$ qui sont égales sur $\mathcal{F}$. Par le Corollaire 2.6, on obtient $\mu_{1,n} = \mu_{2,n}$ pour tout n. Par continuité à gauche des mesures, on obtient ainsi pour tout $A \in \mathcal{A}$,

$$\mu_1(A) = \lim_n \mu_{1,n}(A) = \lim_n \mu_{2,n}(A) = \mu_2(A).$$ □

2.2 Mesures extérieures

Partant de la notion élémentaire de longueur pour les intervalles sur $\mathbb{R}$, une manière naturelle d'essayer de l'étendre en une mesure sur des ensembles plus généraux consiste à recouvrir l'ensemble considéré par une famille dénombrable d'intervalles, puis à prendre l'infimum des longueurs totales correspondantes. Plus précisément, pour $A \subset \mathbb{R}$, on définit

$$\mathcal{L}^*(A) := \inf \Big\{ \sum_n (b_n - a_n) \,:\, (a_n)_n, (b_n)_n \subset \mathbb{R},\, a_n \le b_n,\, A \subset \cup_n [a_n, b_n] \Big\}. \tag{2.1}$$

Cependant, ceci ne peut pas définir une mesure sur $\mathcal{P}(\mathbb{R})$: si c'était le cas, on obtiendrait une contradiction avec l'exemple de Vitali. Autrement dit, $\mathcal{L}^*$ ne peut pas être additive sur $\mathcal{P}(\mathbb{R})$. La stratégie va consister à partir de $\mathcal{L}^*$ et à montrer qu'elle se restreint naturellement à une vraie mesure sur une sous-σ-algèbre strictement plus petite que $\mathcal{P}(\mathbb{R})$, mais suffisamment riche pour nos besoins.

Pour formaliser cette idée, on adopte à nouveau une approche axiomatique : on identifie les propriétés essentielles vérifiées par $\mathcal{L}^*$, ce qui conduit à la notion abstraite de mesure extérieure. On verra plus loin que $\mathcal{L}^*$ est une mesure extérieure et on l'appellera la mesure extérieure de Lebesgue.

Définition 2.10. Soit X un ensemble. Une application $\mu^* : \mathcal{P}(X) \to [0, \infty]$ est une *mesure extérieure* sur X si

- $\mu^*(\varnothing) = 0$,
- μ^* est monotone : $A \subset B \implies \mu^*(A) \le \mu^*(B)$,
- μ^* est σ-sous-additif : pour tout $(A_n)_n \subset \mathcal{P}(X)$,

$$\mu^*\Big(\bigcup_n A_n\Big) \le \sum_n \mu^*(A_n).$$

Exemple 2.11. Soit X un ensemble.

- L'application μ^* définie par $\mu^*(A) = 1$ si $A \neq \varnothing$ et $\mu^*(\varnothing) = 0$ est une mesure extérieure mais n'est pas une mesure.
- L'application μ^* définie par $\mu^*(A) = 0$ si A est dénombrable et $\mu^*(A) = 1$ si A est non-dénombrable est une mesure extérieure.
- Considérons l'application μ^* définie par $\mu^*(A) = 0$ si A est fini et $\mu^*(A) = 1$ si A est infini. Si X est infini, alors μ^* n'est pas une mesure extérieure.

Ce qui manque à une mesure extérieure pour être une mesure, c'est la σ-additivité. Pour y remédier, on considère une classe d'ensemble importants, sur lesquels une certaine additivité est vérifiée : on dit qu'un ensemble est μ^*-mesurable s'il décompose chaque sous-ensemble de X en deux parties sur lesquelles la mesure extérieure est additive.

Définition 2.12. Soient X un ensemble et μ^* une mesure extérieure sur X. Un ensemble $M \subset X$ est dit μ^**-mesurable* s'il satisfait le critère de Carathéodory, pour tout $A \subset X$,

$$\mu^*(A) = \mu^*(A \cap M) + \mu^*(A \cap M^c).$$

On note $\mathcal{M}_{\mu^*}$ la collection des ensembles μ^*-mesurables de X.

Remarque 2.13. Soient X un ensemble et μ^* une mesure extérieure sur X. On donne deux définitions alternatives de la μ^*-mesurabilité.

(i) $M \in \mathcal{M}_{\mu^*}$ si et seulement si, pour tout $A \subset X$ avec $\mu_*(A) < \infty$,

$$\mu^*(A) \geq \mu^*(A \cap M) + \mu^*(A \cap M^c).$$

Cette inégalité est en effet triviale pour $\mu_*(A) = \infty$, tandis que l'inégalité inverse est toujours vérifiée par sous-additivité.

(ii) $M \in \mathcal{M}_{\mu^*}$ si et seulement si, pour tout $A \subset M$ et $B \subset M^c$,

$$\mu^*(A \cup B) = \mu^*(A) + \mu^*(B). \tag{2.2}$$

En effet, si $M \in \mathcal{M}_{\mu^*}$, on a par définition, pour tout $A \subset M$ et $B \subset M^c$,

$$\mu^*(A \cup B) = \mu^*((A \cup B) \cap M) + \mu^*((A \cup B) \cap M^c) = \mu^*(A) + \mu^*(B).$$

Réciproquement, si M satisfait (2.2), alors, pour tout $C \subset X$, en choisissant $A = C \cap M$ et $B = C \cap M^c$, on obtient $\mu^*(C) = \mu^*(C \cap M) + \mu^*(C \cap M^c)$, et donc $M \in \mathcal{M}_{\mu^*}$.

On va montrer que $\mathcal{M}_{\mu^*}$ constitue une σ-algèbre et que la restriction $\mu^*|_{\mathcal{M}_{\mu^*}}$ est une mesure. Pour ce faire, on commence par un nombre d'éléments préliminaires en ce sens.

Lemme 2.14. *Soient X un ensemble et μ^* une mesure extérieure sur X.*

(i) $M \in \mathcal{M}_{\mu^*} \implies M^c \in \mathcal{M}_{\mu^*}$

(ii) $X, \varnothing \in \mathcal{M}_{\mu^*}$

(iii) $M, N \in \mathcal{M}_{\mu^*} \implies M \setminus N \in \mathcal{M}_{\mu^*}$

(iv) $(M_n)_n \subset \mathcal{M}_{\mu^*}$ *deux à deux disjoints* $\implies \cup_n M_n \in \mathcal{M}_{\mu^*}$ *et de plus, pour tout* $A \subset X$,

$$\mu^*(A) = \sum_n \mu^*(A \cap M_n) + \mu^*\Big(A \cap \bigcap_n M_n^c\Big).$$

Démonstration. Le premier point se déduit directement de la définition.
(ii) On remarque que $\mu^*(\varnothing \cap A) + \mu^*(X \cap A) = \mu^*(\varnothing) + \mu^*(A) = \mu^*(A)$, ce qui montre $X, \varnothing \in \mathcal{M}_{\mu^*}$.
(iii) Soient $M, N \in \mathcal{M}_{\mu^*}$ et considérons $A \subset M \setminus N$ et $B \subset (M \setminus N)^c = M^c \cup N$. Comme $N \in \mathcal{M}_{\mu^*}$, on a

$$\mu^*(A) + \mu^*(B) = \mu^*(A) + \mu^*(B \cap N) + \mu^*(B \cap N^c).$$

Comme $M \in \mathcal{M}_{\mu^*}$, en utilisant Remarque 2.13(ii) avec $A \subset M$ et $B \cap N^c \subset M^c$, on obtient

$$\mu^*(A) + \mu^*(B) \quad = \quad \mu^*(B \cap N) + \mu^*\big(A \cup (B \cap N^c)\big).$$

De même, pour $B \cap N \subset N$ et $A \cup (B \cap N^c) \subset N^c$, on obtient

$$\mu^*(A) + \mu^*(B) = \mu^*\big((B \cap N) \cup A \cup (B \cap N^c)\big) = \mu^*(A \cup B),$$

d'où la conclusion.

(iv) Soit $(M_n)_n \subset \mathcal{M}_{\mu^*}$ deux à deux disjoints. Pour tout n, on pose $A_n = \bigcup_{k \le n} M_k$. Montrons par récurrence $A_n \in \mathcal{M}_{\mu^*}$ et pour tout $A \subset X$,

$$\mu^*(A) = \sum_{k \le n} \mu^*(A \cap M_k) + \mu^*\Big(A \cap \bigcap_{k \le n} M_k^c\Big). \tag{2.3}$$

Le cas de base est clair car $A_0 = M_0 \in \mathcal{M}_{\mu^*}$. Supposons à présent le résultat vérifié pour tout $m \le n$ et montrons-le pour $n+1$. Pour $A \subset X$, on a

$$\begin{aligned}
\mu^*(A) &= \mu^*(A \cap M_{n+1}) + \mu^*(A \cap M_{n+1}^c) \\
&= \mu^*(A \cap M_{n+1}) + \mu^*(A \cap M_{n+1}^c \cap A_n) + \mu^*(A \cap M_{n+1}^c \cap A_n^c) \\
&= \mu^*(A \cap M_{n+1}) + \mu^*(A \cap M_{n+1}^c \cap A_n) + \mu^*(A \cap A_{n+1}^c)
\end{aligned}$$

puisque $M_{n+1}, A_n \in \mathcal{M}_{\mu^*}$ et $A_{n+1} = M_{n+1} \cup A_n$. Comme A_n et M_{n+1} sont disjoints, on a $A_n \subset M_{n+1}^c$ et donc

$$\mu^*(A) = \mu^*(A \cap M_{n+1}) + \mu^*(A \cap A_n) + \mu^*(A \cap A_{n+1}^c).$$

En appliquant à présent l'hypothèse de récurrence avec $A \cap A_n$, on obtient

$$\mu^*(A \cap A_n) = \sum_{k \le n} \mu^*(A \cap A_n \cap M_k) + \mu^*\Big(A \cap A_n \cap \bigcap_{k \le n} M_k^c\Big) = \sum_{k \le n} \mu^*(A \cap M_k)$$

puisque $M_k \subset A_n$ pour $k \le n$ et $\cap_{k \le n} M_k^c = A_n^c$. On obtient donc

$$\begin{aligned}
\mu^*(A) &= \mu^*(A \cap M_{n+1}) + \sum_{k \le n} \mu^*(A \cap M_k) + \mu^*(A \cap A_{n+1}^c) \\
&= \sum_{k \le n+1} \mu^*(A \cap M_k) + \mu^*\Big(A \cap \bigcap_{k \le n+1} M_k^c\Big),
\end{aligned}$$

d'où l'égalité (2.3) souhaitée au niveau $n+1$. Il reste à montrer $A_{n+1} \in \mathcal{M}_{\mu^*}$. En combinant la sous-additivité de μ^* et l'égalité démontrée, on obtient

$$\mu^*(A) = \sum_{k \le n+1} \mu^*(A \cap M_k) + \mu^*(A \cap A_{n+1}^c) \ge \mu^*(A \cap A_{n+1}) + \mu^*(A \cap A_{n+1}^c),$$

ce qui implique $A_{n+1} \in \mathcal{M}_{\mu^*}$ par Remarque 2.13(ii).

Maintenant que (2.3) est démontré pour tout n, il reste à passer à la limite $n \uparrow \infty$ pour conclure (iv). Cependant, ceci n'a rien d'évident comme une mesure extérieure ne satisfait en générale pas de continuité à gauche comme les mesures. Comme $A_n \subset \cup_k M_k$, on a $\cap_k M_k^c \subset A_n^c$, et donc par monotonie

$$\mu^*(A) \ge \sum_{k \le n} \mu^*(A \cap M_k) + \mu^*\Big(A \cap \bigcap_k M_k^c\Big).$$

Puisque cette relation est vraie pour tout n, on déduit

$$\mu^*(A) \ge \sum_k \mu^*(A \cap M_k) + \mu^*\Big(A \cap \bigcap_k M_k^c\Big) \ge \mu^*\Big(A \cap \bigcup_k M_k\Big) + \mu^*\Big(A \cap \Big(\bigcup_k M_k\Big)^c\Big),$$

où la dernière inégalité provient de la σ-sous-additivité de μ^*. Par Remarque 2.13(i), ceci permet de conclure $\cup_k M_k \in \mathcal{M}_{\mu^*}$, et toutes les inégalités ci-dessus sont des égalités. $\square$

À la lumière des observations précédentes, on peut désormais conclure que $\mathcal{M}_{\mu^*}$ est une σ-algèbre sur laquelle μ^* se restreint en une mesure. Ce résultat fournit par conséquent une méthode très efficace et générale pour construire des mesures.

Théorème 2.15 (Théorème fondamental des mesures extérieures)**.** *Soient X un ensemble et μ^* une mesure extérieure sur X.*

(i) $\mathcal{M}_{\mu^*}$ est une σ-algèbre.

(ii) La restriction de $\mu := \mu^*|_{\mathcal{M}_{\mu^*}} : \mathcal{M}_{\mu^*} \to [0, \infty]$ est une mesure sur $(X, \mathcal{M}_{\mu^*})$.

(iii) $(X, \mathcal{M}_{\mu^*}, \mu)$ est complet.

Démonstration. (i) Les points (i), (ii), (iv) du Lemme 2.14 impliquent que $\mathcal{M}_{\mu^*}$ est une classe monotone. Par ailleurs, en combinant les points (i) et (iii) du même lemme, on voit que $\mathcal{M}_{\mu^*}$ est stable par intersections finies. Le lemme de classe monotone permet alors de conclure.
(ii) Pour $(M_n)_n \subset \mathcal{M}_{\mu^*}$ deux à deux disjoints, le Lemme 2.14(iv) donne

$$\mu^*\Big(\bigcup_k M_k\Big) = \sum_n \mu^*\Big(\bigcup_k M_k \cap M_n\Big) + \mu^*\Big(\bigcup_k M_k \cap \bigcap_k M_k^c\Big) = \sum_k \mu^*(M_k).$$

(iii) Soient $M \subset N \in \mathcal{M}_{\mu^*}$ avec $\mu^*(N) = 0$. Par monotonie, pour tout $A \subset X$, on a $\mu^*(A \cap M) \leq \mu^*(M) \leq \mu^*(N) = 0$ et donc

$$\mu^*(A \cap M) + \mu^*(A \cap M^c) \leq \mu^*(A \cap M^c) \leq \mu^*(A).$$

Par Remarque 2.13(i), ceci implique $M \in \mathcal{M}_{\mu^*}$. □

Une question demeure : dans cette construction, on ne sait a priori pas à quel point $\mathcal{M}_{\mu^*}$ est riche. Si $\mathcal{M}_{\mu^*}$ se réduit à la σ-algèbre triviale, la construction ne donne bien sûr rien d'intéressant. Le résultat suivant fournit un critère bien utile pour s'assurer que ce ne soit pas le cas : sur un espace métrique, si μ^* est additive sur les ensembles “lointains”, alors $\mathcal{M}_{\mu^*}$ contient les boréliens.

Proposition 2.16. *Soient (X, d) un espace métrique et μ^* une mesure extérieure sur X. Si pour tous $A, B \subset X$ tels que $d(A, B) > 0$ on a $\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$, alors les boréliens de X sont μ^*-mesurables.*

Démonstration. Nous allons montrer que tout fermé est μ^*-mesurable. Soit $F \subset X$ un fermé. Par Remarque 2.13(i), il suffit de considérer $A \subset X$ avec $\mu^*(A) < \infty$ et montrer que

$$\mu^*(A) \geq \mu^*(A \cap F) + \mu^*(A \cap F^c).$$

Par monotonie, on peut également supposer $A \cap F \neq \varnothing$ et $A \cap F^c \neq \varnothing$. Pour tout k, on considère le fermé épaissi $F_k :=$“$F + \frac{1}{k}B$”, plus précisément

$$F_k = \big\{x \in X : d(x, F) \leq \tfrac{1}{k}\big\}.$$

Comme $d(A \cap F, A \cap F_k^c) \geq \frac{1}{k} > 0$ et $(A \cap F) \cup (A \cap F_k^c) \subset A$, on obtient par hypothèse

$$\mu^*(A) \geq \mu^*\big((A \cap F) \cup (A \cap F_k^c)\big) = \mu^*(A \cap F) + \mu^*(A \cap F_k^c).$$

Pour conclure, il suffit donc de montrer que

$$\lim_k \mu^*(A \cap F_k^c) \geq \mu^*(A \cap F^c).$$

À nouveau, le problème est qu'une mesure extérieure ne satisfait en général pas de continuité à gauche. Pour tout k, considérons les “anneaux”

$$B_k = A \cap (F_k \setminus F_{k+1}).$$

Comme F est fermé, on remarque que $x \in F^c$ si et seulement si $d(x, F) > 0$, ce qui permet d'écrire $A \cap F^c = (A \cap F_k^c) \cup \bigcup_{j \geq k} B_j$. Par σ-sous-additivité, on déduit

$$\mu^*(A \cap F^c) \leq \mu^*(A \cap F_k^c) + \sum_{j \geq k} \mu^*(B_j).$$

Pour conclure, il suffit à présent de montrer $\sum_j \mu^*(B_j) < \infty$. Remarquons que $d(B_{2j}, B_{2k}) > 0$ pour $j \neq k$, et donc par hypothèse, pour tout $K \geq 1$,

$$\sum_{k=1}^{K} \mu^*(B_{2k}) = \mu^*\left(\bigcup_{k=1}^{K} B_{2k} \right) \leq \mu^*(A).$$

En argumentant de même pour les indices impairs, on déduit $\sum_j^\infty \mu^*(B_j) \leq 2\mu^*(A) < \infty$, et la conclusion s'ensuit. □

2.3 Mesure de Lebesgue

Nous allons définir la mesure de Lebesgue sur $\mathbb{R}^d$ pour étendre la notion élémentaire de volume des rectangles fermés,

$$\mathrm{Vol}(I) = \prod_{k=1}^{d} (b_k - a_k) \qquad \text{pour} \quad I = \prod_{k=1}^{d} [a_k, b_k], \quad a_k \leq b_k.$$

Définition 2.17. Si $A \subset \mathbb{R}^d$, on note $\mathcal{C}_A$ la collection de toutes les suites $(I_n)_n$ d'intervalles compacts de $\mathbb{R}^d$ tels que $A \subset \bigcup_n I_n$. La *mesure extérieure de Lebesgue* est définie par

$$\mathcal{L}_d^*(A) = \inf \left\{ \sum_n \mathrm{Vol}(I_n) : (I_n)_n \in \mathcal{C}_A \right\}.$$

Pour $d = 1$, on note $\mathcal{L}^* := \mathcal{L}_1^*$, qui coïncide avec (2.1).

Lemme 2.18. *La mesure extérieure de Lebesgue $\mathcal{L}_d^*$ est une mesure extérieure sur $\mathbb{R}^d$.*

Démonstration. Tout d'abord, $I_0 := \{0\}^d$ est un rectangle fermé (avec $a_k = b_k = 0$) avec $\mathrm{Vol}(I_0) = 0$, et donc $\mathcal{L}_d^*(\varnothing) = 0$. Ensuite, pour $A \subset B \subset \mathbb{R}^d$, si $(I_n)_n \in \mathcal{C}_B$, alors on a aussi $(I_n)_n \in \mathcal{C}_A$, et donc $\mathcal{L}_d^*(A) \leq \mathcal{L}_d^*(B)$. Enfin, pour $(A_n)_n \subset \mathcal{P}(\mathbb{R}^d)$ une suite d'ensembles de $\mathbb{R}^d$, il reste à montrer que

$$\mathcal{L}_d^*\left(\bigcup_n A_n \right) \leq \sum_{n \in \mathbb{N}} \mathcal{L}_d^*(A_n).$$

Pour ce faire, on peut bien sûr supposer $\sum_n \mathcal{L}_d^*(A_n) < \infty$. Soit $\varepsilon > 0$. Pour tout n, on peut choisir $(I_{n,j})_j \in \mathcal{C}_{A_n}$ tel que

$$\sum_j \mathrm{Vol}(I_{n,j}) \leq \mathcal{L}_d^*(A_n) + \frac{\varepsilon}{2^{n+1}}.$$

Dès lors, $\bigcup_n A_n \subset \bigcup_n \bigcup_j I_{n,j}$, et donc

$$\mathcal{L}_d^*\left(\bigcup_n A_n \right) \leq \sum_{n,j} \mathrm{Vol}(I_{n,j}) \leq \sum_n \mathcal{L}_d^*(A_n) + \sum_n \frac{\varepsilon}{2^{n+1}} = \sum_n \mathcal{L}_d^*(A_n) + \varepsilon.$$

Comme ε est arbitraire, on obtient la conclusion. □

On peut à présent utiliser la théorie générale des mesures extérieures développée à la section précédente pour définir la mesure de Lebesgue.

Définition 2.19. On appelle *σ-algèbre de Lebesgue* la σ-algèbre $\mathcal{M}_d := \mathcal{M}_{\mathcal{L}_d^*}$ des ensembles $\mathcal{L}_d^*$-mesurables. La restriction $\mathcal{L}_d := \mathcal{L}_d^*|_{\mathcal{M}_d}$ est alors une mesure sur $(\mathbb{R}^d, \mathcal{M}_d)$, et est appelée la *mesure de Lebesgue sur* $\mathbb{R}^d$. Pour $d = 1$, on note $\mathcal{M} := \mathcal{M}_1$ et $\mathcal{L} = \mathcal{L}_1$.
Les éléments de $\mathcal{M}_d$ sont appelés *Lebesgue-mesurables.* On note par ailleurs $\mathcal{L}^1(\mathbb{R}^d) = \mathcal{L}^1(\mathbb{R}^d, \mathcal{M}_d, \mathcal{L}_d)$ l'espace des fonctions intégrables sur $\mathbb{R}^d$. Pour f positive ou intégrable, on note l'intégrale de Lebesgue sur $\mathbb{R}^d$

$$\int f := \int f dx := \int f(x) dx := \int f d\mathcal{L}_d.$$

La mesure de Lebesgue ainsi définie satisfait toutes les propriétés voulues : les boréliens sont Lebesgue-mesurables et la mesure de Lebesgue étend la notion de volume pour les rectangles.

Théorème 2.20. $(\mathbb{R}^d, \mathcal{M}_d, \mathcal{L}_d)$ *est un espace de mesure complet. De plus :*

(i) $\mathcal{B}(\mathbb{R}^d) \subset \mathcal{M}_d$

(ii) $\mathcal{L}_d(I) = \mathrm{Vol}(I)$ *pour tout rectangle fermé* $I \subset \mathbb{R}^d$.

(iii) $\mathcal{L}_d(x + A) = \mathcal{L}_d(A)$ *pour tout* $x \in \mathbb{R}^d$ *et* $A \in \mathcal{M}_d$.

(iv) $\mathcal{L}_d(K) < \infty$ *pour tout* K *compact.*

(v) Unicité de la construction : si μ *est une mesure sur* $(\mathbb{R}^d, \mathbb{B}_d)$ *avec* $\mu(I) = \mathrm{Vol}(I)$ *pour tout rectangle fermé* $I \subset \mathbb{R}^d$, *alors* $\mu = \mathcal{L}_d|_{\mathbb{B}_d}$.

Démonstration. (i) Par la Proposition 2.16, il suffit de montrer que pour tous $A, B \subset \mathbb{R}^d$ tels que $d(A, B) > 0$, on a

$$\mathcal{L}_d^*(A \cup B) \geq \mathcal{L}_d^*(A) + \mathcal{L}_d^*(B).$$

On peut supposer que $\mathcal{L}_d^*(A \cup B) < \infty$. Fixons $\varepsilon > 0$ et considérons $(I_n)_n \in \mathcal{C}_{A \cup B}$ tel que

$$\sum_n \mathrm{Vol}(I_n) \leq \mathcal{L}_d^*(A \cup B) + \varepsilon.$$

Quitte à subdiviser les intervalles trop grands, on peut supposer que $\mathrm{diam}(I_n) \leq d(A, B)$ pour tout n. La suite $(I_n)_n$ peut alors être décomposée en deux sous-suites $(I'_n)_n$ et $(I''_n)_n$ telles que $A \cap I''_n = \varnothing$ et $B \cap I'_n = \varnothing$ pour tout n. On en déduit

$$\mathcal{L}^*(A) + \mathcal{L}^*(B) \leq \sum_n \mathrm{Vol}(I'_n) + \sum_n \mathrm{Vol}(I''_n) = \sum_n \mathrm{Vol}(I_n) \leq \mathcal{L}_d^*(A \cup B) + \varepsilon.$$

(ii) Soit I un rectangle fermé. Puisque $(I, \varnothing, \ldots) \in \mathcal{C}_I$, on a $\mathcal{L}_d^*(I) \leq \mathrm{Vol}(I)$, et il reste à montrer l'inégalité inverse. Fixons $\varepsilon > 0$ et considérons une suite $(I_n)_n \in \mathcal{C}_I$ telle que

$$\sum_n \mathrm{Vol}(I_n) \leq \mathcal{L}_d^*(I) + \varepsilon.$$

Pour chaque $n \in \mathbb{N}$, on peut choisir un rectangle fermé J_n tel que $I_n \subset J_n^\circ$ et $\mathrm{Vol}(J_n) \leq \mathrm{Vol}(I_n) + \frac{\varepsilon}{2^{n+1}}$. On a alors $I \subset \bigcup_n J_n^\circ$. Comme I est compact, il existe $N \geq 1$ tel que $I \subset \bigcup_{n \leq N} J_n^\circ \subset \bigcup_{n \leq N} J_n$. Une récurrence immédiate montre alors $\mathrm{Vol}(I) \leq \sum_{n \leq N} \mathrm{Vol}(J_n)$, et on obtient ainsi

$$\mathrm{Vol}(I) \leq \sum_{n=0}^{N} \mathrm{Vol}(J_n) \leq \sum_{n=0}^{N} \left(\mathrm{Vol}(I_n) + \frac{\varepsilon}{2^{n+1}}\right) \leq \mathcal{L}_d^*(I) + 2\varepsilon.$$

(iii) L'invariance par translation de Vol s'étend à $\mathcal{L}_d^*$ et donc à $\mathcal{L}_d$.

(iv) Un compact $K \subset \mathbb{R}^d$ est borné, de sorte qu'il existe $n \geq 1$ tel que $K \subset [-n,n]^d$. Dès lors, $\mathcal{L}_d(K) \leq \mathcal{L}_d([-n,n]^d) = \mathrm{Vol}([-n,n]^d) = (2n)^d < \infty$.

(v) On remarque que l'ensemble des rectangles fermés constitue un π-système sur $\mathbb{R}^d$. Par ailleurs, on a vu que ce π-système engendre $\mathbb{B}_d$. De plus, $\mathcal{L}_d|_{\mathbb{B}_d}$ est σ-finie sur l'ensemble des rectangles étant donné que $\mathcal{L}_d([-n,n]^d) = (2n)^d < \infty$ et $[-n,n]^d \uparrow \mathbb{R}^d$. Dans ce cadre, on peut alors appliquer le théorème d'unicité de mesures. $\square$

Remarque 2.21. Pour les rectangles ouverts, on obtient également

$$\mathcal{L}_d\Big(\prod_{k=1}^d (a_k,b_k)\Big) = \mathcal{L}_d\Big(\cup_n \prod_{k=1}^d [a_k+\tfrac{1}{n}, b_k-\tfrac{1}{n}]\Big) = \lim_n \mathcal{L}\Big(\prod_{k=1}^d [a_k+\tfrac{1}{n}, b_k-\tfrac{1}{n}]\Big)$$

$$= \lim_n \prod_{k=1}^d (b_k-a_k-\tfrac{2}{n})_+ = \prod_{k=1}^d (b_k-a_k),$$

qui coïncide avec la notion usuelle de volume.

Remarque 2.22.

— Sous l'axiome du choix, il existe un ensemble $E \in \mathcal{P}(\mathbb{R}) \setminus \mathcal{M}$ (exemple de Vitali, 1905).

— Sous l'axiome du choix, il existe même $E \subset \mathbb{R}$ tel que pour tout $A \in \mathcal{M}$ on a $\mathcal{L}(A) = 0$ pour tout $A \subset E$ et tout $A \subset E^c$.

— Sans l'axiome du choix, on ne peut en fait pas montrer que $\mathcal{P}(\mathbb{R}) \setminus \mathcal{M} \neq \varnothing$ (Solovay, 1964).

— Il existe un ensemble $E \in \mathcal{M} \setminus \mathbb{B}$. Plus généralement, on a $\mathbb{B}_d \subsetneq \mathcal{M}_d$, et la relation précise entre ces deux σ-algèbres sera expliquée plus tard.

Ces différents exemples pathologiques seront construits en exercice.

2.4 Régularité de la mesure de Lebesgue

Le résultat fondamental suivant justifie le 1er principe de Littlewood : "un ensemble Lebesgue-mesurable est 'presque' un ouvert — ou un fermé". Sur $\mathbb{R}$, on obtient de même qu'un ensemble Lebesgue-mesurable est 'presque' une union dénombrable disjointe d'intervalles.

Théorème 2.23. *Soit $A \in \mathcal{M}_d$. Alors, on a :*

(i) régularité extérieure : $\mathcal{L}_d(A) = \inf\big\{\mathcal{L}_d(U) : A \subset U,\ U \text{ ouvert}\big\}$,

(ii) régularité intérieure : $\mathcal{L}_d(A) = \sup\big\{\mathcal{L}_d(K) : K \subset A,\ K \text{ compact}\big\}$.

Si $\mathcal{L}_d(A) < \infty$, pour tout $\varepsilon > 0$, on peut alors trouver $K_\varepsilon \subset A \subset U_\varepsilon$ avec K_ε compact et U_ε ouvert tels que $\mathcal{L}_d(U_\varepsilon \setminus K_\varepsilon) \leq \varepsilon$.

Démonstration. (i) Par monotonie, il suffit de montrer $\mathcal{L}_d(A) \geq \inf\{\mathcal{L}_d(U) : A \subset U,\ U \text{ ouvert}\}$, et on peut supposer que $\mathcal{L}_d(A) < \infty$. Fixons $\varepsilon > 0$ et considérons $(I_n)_n$ une suite de rectangles fermés tels que $A \subset \cup_n I_n$ et

$$\sum_n \mathrm{Vol}(I_n) \leq \mathcal{L}_d(A) + \varepsilon.$$

Pour tout n, on peut choisir un rectangle ouvert J_n avec $I_n \subset J_n$ et $\mathrm{Vol}(J_n) \leq \mathrm{Vol}(I_n) + \frac{\varepsilon}{2^{n+1}}$. En choisissant $U_\varepsilon := \cup_n J_n$, on a dès lors $A \subset U_\varepsilon$ et

$$\mathcal{L}_d(U_\varepsilon) \leq \sum_n \underbrace{\mathcal{L}_d(J_n)}_{=\mathrm{Vol}(J_n)} \leq \sum_n \mathrm{Vol}(I_n) + \varepsilon \leq \mathcal{L}_d(A) + 2\varepsilon.$$

(ii) Par monotonie, il suffit de montrer que $\mathcal{L}_d(A) \leq \sup\{\mathcal{L}_d(K) : K \subset A, K \text{ est compact}\}$ et on peut supposer que $\mathcal{L}_d(A) > 0$. Nous allons distinguer deux cas.
Cas 1 : supposons que A est borné. Dans ce cas, soit K_0 un compact contenant A. Fixons $\varepsilon > 0$. Par régularité extérieure, il existe un ouvert U_ε tel que $K_0 \setminus A \subset U_\varepsilon$ et $\mathcal{L}_d(U_\varepsilon) \leq \mathcal{L}_d(K_0 \setminus A) + \varepsilon$. On déduit alors que $K_\varepsilon := K_0 \setminus U_\varepsilon$ est un compact $\subset A$ avec

$$\mathcal{L}_d(A) = \mathcal{L}_d(K_0) - \mathcal{L}_d(K_0 \setminus A) \leq \mathcal{L}_d(K_0) - \mathcal{L}_d(U_\varepsilon) + \varepsilon \leq \mathcal{L}_d(K_0 \setminus U_\varepsilon) + \varepsilon = \mathcal{L}_d(K_\varepsilon) + \varepsilon.$$

Cas 2 : cas général. Soit $c > 0$ tel que $\mathcal{L}_d(A) > c$. On va construire un compact K inclus dans A tel que $\mathcal{L}_d(K) > c$. Pour tout n, posons $A_n = A \cap \bar{B}(0, n) = \{x \in A : |x| \leq n\} \in \mathcal{M}_d$. On a $A_n \uparrow A$, et donc $\mathcal{L}_d(A_n) \uparrow \mathcal{L}_d(A) > c$. Dès lors, il existe n_0 tel que $\mathcal{L}_d(A_{n_0}) > c$. Par le cas 1, il existe un compact $K \subset A_{n_0} \subset A$ tel que $\mathcal{L}_d(K) > c$. □

De ce résultat pour l'approximation d'ensembles Lebesgue-mesurables, on déduit un résultat pour l'approximation des fonctions intégrables. Il s'agit là du 2e principe de Littlewood : "toute fonction intégrable est 'presque' continue" ; on reviendra sur ce point plus loin.

Corollaire 2.24. *L'espace $C_c(\mathbb{R}^d) = \{f : \mathbb{R}^d \to \mathbb{R}$ continue à support compact$\}$ est "dense" dans $\mathcal{L}^1(\mathbb{R}^d)$. Plus précisément, pour tout $f \in \mathcal{L}^1(\mathbb{R}^d)$ et tout $\varepsilon > 0$, il existe $g_\varepsilon \in C_c(\mathbb{R}^d)$ tel que*

$$\int_{\mathbb{R}^d} |f - g| dx \leq \varepsilon.$$

Démonstration. Quitte à décomposer f en ses parties positive et négative, il suffit de montrer le résultat pour toute fonction f intégrable positive. Or, pour une fonction f intégrable positive, on a $\int f = \sup\{\int s : s \text{ simple mesurable}, 0 \leq s \leq f\} < \infty$. Pour tout ε, il existe donc s_ε simple mesurable avec $0 \leq s_\varepsilon \leq f$ et $\int f \leq \int s_\varepsilon + \varepsilon$. Dès lors,

$$\int |f - s_\varepsilon| = \int f - \int s_\varepsilon \leq \varepsilon.$$

Il suffit donc de montrer le résultat pour toute fonction f simple intégrable positive. Par linéarité, il suffit de montrer le résultat pour $f = \mathbb{1}_A$ avec $A \in \mathcal{M}_d$ et $\mathcal{L}_d(A) < \infty$. Pour tout ε, par régularité de la mesure de Lebesgue, il existe $K_\varepsilon \subset A \subset U_\varepsilon$ avec K_ε compact, U_ε ouvert, et $\mathcal{L}_d(U_\varepsilon \setminus K_\varepsilon) \leq \varepsilon$. Par le lemme d'Urysohn, il existe une fonction continue g_ε telle que

$$g_\varepsilon|_{K_\varepsilon} = 1, \qquad \text{supp}(g_\varepsilon) \subset U_\varepsilon, \qquad 0 \leq g_\varepsilon \leq 1.$$

Sur $\mathbb{R}^d$, comme dans tout espace métrique, on rappelle que le lemme d'Urysohn se montre aisément : il suffit de choisir par exemple

$$g_\varepsilon(x) := \phi\Big(\frac{d(x, U_\varepsilon^c)}{d(K_\varepsilon, U_\varepsilon^c)}\Big), \qquad \phi(t) := \begin{cases} 1 & : \quad t \geq 1, \\ t & : \quad 0 \leq t \leq 1, \\ 0 & : \quad t \leq 0. \end{cases}$$

Cette fonction cut-off satisfait $\mathbb{1}_{K_\varepsilon} \leq g_\varepsilon \leq \mathbb{1}_{U_\varepsilon}$ et donc

$$\int |g_\varepsilon - \mathbb{1}_A| = \int_{U_\varepsilon \setminus K_\varepsilon} |g_\varepsilon - \mathbb{1}_A| \leq \mathcal{L}_d(U_\varepsilon \setminus K_\varepsilon) \leq \varepsilon. \qquad \square$$

2.5 Unicité des mesures invariantes par translation

Théorème 2.25. *Si μ est une mesure sur $(\mathbb{R}^d, \mathbb{B}_d)$ telle que*

— *μ est invariante par translation : $\mu(x + A) = \mu(A)$ pour tout $x \in \mathbb{R}^d$ et $A \in \mathbb{B}_d$;*

— *μ est non-triviale : il existe un ouvert U tel que $\mu(U) < \infty$.*

Alors il existe une constante $c \in [0, \infty)$ telle que $\mu = c\mathcal{L}_d$.

Démonstration. Sans perte de généralité, on peut supposer par exemple $0 < \mu([0, 1)^d) < \infty$. Notons $c := \mu([0, 1)^d)$ et $\nu := \frac{1}{c}\mu$. Considérons cubes "dyadiques", c'est-à-dire de cubes de la forme

$$Q_n(k_1, \dots, k_n) = \prod_{j=1}^{d} [\tfrac{k_j}{2^n}, \tfrac{k_j+1}{2^n}), \qquad n \geq 0, \quad k_1, \dots, k_d \in \mathbb{Z}.$$

L'entier n est appelé l'*échelle* du cube dyadique $Q_n(k_1, \dots, k_n)$. Pour tout n fixé, les cubes dyadiques d'échelle n forment une partition de $\mathbb{R}^d$. De plus, remarquons que tout ouvert de $\mathbb{R}^d$ peut être écrit comme une union dénombrable de cubes dyadiques disjoints. En effet, soit U un ensemble ouvert de $\mathbb{R}^d$. La famille $\mathcal{F}$ de cubes dyadiques qui recouvre U peut être construite récursivement comme suit : on commence par l'ensemble vide et à l'étape n on ajoute les cubes dyadiques d'échelle n qui sont inclus dans U et qui ne sont inclus dans aucun cube déjà sélectionné. Par construction, on a $\cup_{C \in \mathcal{F}} C \subset U$. De plus, si $x \in U$ et si $C_n(x)$ est le cube dyadique d'échelle n qui contient x, alors $C_n(x) \subset U$ pour n suffisamment grand puisque U est ouvert. On conclut $\cup_{C \in \mathcal{F}} C = U$.

Pour tout n, on peut décomposer $[0, 1)^d$ en une union disjointe de 2^{dn} cubes dyadiques d'échelle n. Comme ν est invariante par translation et $\nu([0, 1)^d) = 1$, on déduit $1 = \nu([0, 1)^d) = 2^{dn}\nu([0, \frac{1}{2^n})^d)$, et donc

$$\nu([0, \tfrac{1}{2^n})^d) = 2^{-dn} = \mathcal{L}_d([0, \tfrac{1}{2^n})^d).$$

Pour tout ouvert U, comme on peut l'écrire comme une union dénombrable de cubes dyadiques disjoints, on peut dès lors déduire par invariance par translation et σ-additivité,

$$\nu(U) = \mathcal{L}_d(U).$$

Comme les ouverts forment un π-système et que la mesure de Lebesgue $\mathcal{L}_d$ est σ-finie par rapport aux ouverts, on peut conclure par le théorème d'unicité des mesures. □

2.6 Complétude

On a vu que $\mathbb{B}_d \subsetneq \mathcal{M}_d$. Le résultat suivant affirme que la σ-algèbre de Lebesgue est en fait précisément la complétion de la σ-algèbre de Borel.

Théorème 2.26. *$(\mathbb{R}^d, \mathcal{M}_d, \mathcal{L}_d)$ est la complétion de $(\mathbb{R}^d, \mathbb{B}_d, \mathcal{L}_d|_{\mathbb{B}_d})$*

Démonstration. Notons $\mathcal{L}_d^B := \mathcal{L}_d|_{\mathbb{B}_d}$. On rappelle que la complétion de la σ-algèbre de Borel est définie par

$$\bar{\mathbb{B}}_d := \{A \in \mathbb{R}^d : \exists E, F \in \mathbb{B}_d, E \subset A \subset F, \mathcal{L}_d(F \setminus E) = 0\},$$

et pour $A \in \bar{\mathbb{B}}_d$ la complétion de $\mathcal{L}_d^B$ est donnée par

$$\bar{\mathcal{L}}_d^B(A) = \inf\{\mathcal{L}_d^B(F) : F \in \mathbb{B}_d, F \supset A\}.$$

On commence par montrer que $\bar{\mathbb{B}}_d \subset \mathcal{M}_d$. Pour $A \in \bar{\mathbb{B}}_d$, il existe $E, F \in \mathbb{B}_d$ avec $E \subset A \subset F$ et $\mathcal{L}_d(F \setminus E) = 0$. Comme $A \setminus E \subset F \setminus E$, on déduit que $A \setminus E$ est $\mathcal{L}_d$-négligeable et donc Lebesgue-mesurable comme $\mathcal{L}_d$ est complète. Dès lors, on a $A = (A \setminus E) \cup E \in \mathcal{M}_d$.

Il reste à montrer que $\mathcal{M}_d \subset \bar{\mathbb{B}}_d$. Pour ce faire, il suffit de montrer que pour tout $A \in \mathcal{M}_d$ il existe $E, F \in \mathbb{B}_d$ avec $E \subset A \subset F$ et $\mathcal{L}_d(F \setminus E) = 0$. Pour $A \in \mathcal{M}_d$, on distingue deux cas.

Cas 1 : supposons $\mathcal{L}_d(A) < \infty$. Par régularité, pour tout n, il existe $K_n \subset A \subset U_n$ avec K_n compact, U_n ouvert, et $\mathcal{L}_d(U_n \setminus K_n) \leq \frac{1}{n}$. Choisissons $E = \cup_n K_n$ et $F = \cap_n U_n$. On a bien sûr $E, F \in \mathbb{B}_d$ et $E \subset A \subset F$. De plus, pour tout n, on a $\mathcal{L}_d(F \setminus E) \leq \mathcal{L}_d(F_n \setminus E_n) \leq \frac{1}{n}$, et donc $\mathcal{L}_d(F \setminus E) = 0$.

Cas 2 : cas général. Pour tout n, considérons $A_n = A \cap \bar{B}(0, n) \in \mathcal{M}_d$. Par le cas 1, il existe $E_n, F_n \in \mathbb{B}_d$ avec $E_n \subset A_n \subset F_n$ et $\mathcal{L}_d(F_n \setminus E_n) = 0$. Pour $E = \cup_n E_n$ et $F = \cup_n F_n$, on obtient alors $E, F \in \mathbb{B}_d$, $E \subset A \subset F$, et $\mathcal{L}_d(F \setminus E) \leq \sum_n \mathcal{L}_d(F_n \setminus E_n) = 0$. □

Remarque 2.27. Si $f, g : \mathbb{R} \to \mathbb{R}$ sont Lebesgue-mesurable, on remarque que leur composition $f \circ g$ n'est pas forcement mesurable. Par contre, si f est Lebesgue-mesurable et si g est borélienne, alors la composition $f \circ g$ est bien Lebesgue-mesurable.

2.7 Propriétés de l'intégrale de Lebesgue

Comme annoncé en introduction, l'intégrale de Lebesgue généralise l'intégrale de Riemann : pour une fonction bornée sur un intervalle compact, dès que l'intégrale de Riemann est bien définie, elle coïncide avec celle de Lebesgue. Le théorème suivant précise exactement dans quels cas une fonction est Riemann-intégrable. Sa preuve sera discutée en séance d'exercices. Ce résultat est particulièrement utile : comme on ne dispose pas de techniques de calcul propres à l'intégrale de Lebesgue, il permet souvent de ramener le calcul d'intégrales au cadre classique de Riemann et d'utiliser toutes les règles familières (intégration par parties, etc.).

Théorème 2.28 (Théorème de Lebesgue-Vitali)**.** *Soit $f : [a, b] \to \mathbb{R}$ une fonction bornée. Alors f est Riemann-intégrable si et seulement si f est Lebesgue-mesurable et si l'ensemble des points de discontinuité de f est de mesure de Lebesgue nulle. Dans ce cas, les intégrales de Riemann et de Lebesgue coïncident :*

$$\int_a^b f = (\mathrm{R}) \int_a^b f,$$

où le préfixe '(R)' indique l'intégrale de Riemann.

Exemple 2.29.

- Si $f : [a, b] \to \mathbb{R}$ est continue, alors elle est bornée et l'ensemble de ses points de discontinuité est vide. Le théorème de Lebesgue-Vitali assure alors que $\int_a^b f = (R) \int_a^b f$.
- La fonction de Dirichlet $\mathbb{1}_{\mathbb{Q} \cap [0,1]} : [0, 1] \to \mathbb{R}$ est discontinue en tout point et n'est donc pas Riemann-intégrable, bien qu'elle soit Lebesgue-intégrable (elle est égale à 0 presque partout et son intégrale est donc nulle).

On énonce sans preuve le résultat suivant pour le changement de variable dans l'intégrale de Lebesgue. L'énoncé est identique au cas de l'intégrale de Riemann.

Théorème 2.30 (Changement de variables)**.** *Soient $U, U' \subset \mathbb{R}^d$ ouverts, $T : U \to U'$ un difféomorphisme C^1, et $f : U' \to [0, \infty]$ Lebesgue-mesurable. Alors la fonction $(f \circ T)|\det(\nabla T)| : U \to [0, \infty]$ est aussi Lebesgue-mesurable et on a*

$$\int_{U'} f(x)dx = \int_U (f \circ T)(x)|\det(\nabla T(x))|dx.$$

Chapitre 3

Mesure produit et théorème de Fubini

Dans ce chapitre, pour une fonction Lebesgue-intégrable $f : \mathbb{R}^2 \to \mathbb{R}$, on souhaite déterminer quand on peut calculer son intégrale comme une intégrale itérée,

$$\int_{\mathbb{R}^2} f = \int_{\mathbb{R}} \Big(\int_{\mathbb{R}} f(x,y)dx \Big) dy = \int_{\mathbb{R}} \Big(\int_{\mathbb{R}} f(x,y)dy \Big) dx.$$

Il y a un nombre de questions qui se posent à la lecture de cette ligne : tout d'abord, est-ce que la fonction $x \mapsto f(x,y)$ est forcément Lebesgue-intégrable pour tout y ? Et qu'en est-il de $y \mapsto \int_{\mathbb{R}} f(x,y)dx$?

Avant de répondre à ces questions, on commence par introduire la notion de mesure produit, qui sera le cadre général pour lequel on pourra prouver une première version du théorème de Fubini.

3.1 Mesure produit

Étant donné deux espaces de mesure $(X, \mathcal{A}, \mu)$ et $(Y, \mathcal{A}', \nu)$, on souhaite munir le produit cartésien $X \times Y$ d'une mesure, notée $\mu \otimes \nu$, telle que pour tous $A \in \mathcal{A}$, $A' \in \mathcal{A}'$,

$$\mu \otimes \nu(A \times A') = \mu(A)\nu(A'). \tag{3.1}$$

On va naturellement vouloir définir cette mesure produit sur la σ-algèbre engendrée par les ensembles de la forme $A \times A'$.

Définition 3.1. Soient $(X, \mathcal{A})$ et $(Y, \mathcal{A}')$ deux espaces mesurables. Un ensemble de la forme $A \times A'$ avec $A \in \mathcal{A}$ et $A' \in \mathcal{A}'$ est appelé un *rectangle mesurable*. La *σ-algèbre produit* sur $X \times Y$ est la σ-algèbre engendrée par les rectangles mesurables. On la note $\mathcal{A} \otimes \mathcal{A}'$.

Lemme 3.2.

(i) Soient $(X, \mathcal{A})$ et $(Y, \mathcal{A}')$ deux espaces mesurables. La σ-algèbre produit $\mathcal{A} \otimes \mathcal{A}'$ est la plus petite σ-algèbre sur $X \times Y$ rendant les projections

$$\pi_1 : X \times Y \to X : (x,y) \mapsto x \quad \text{et} \quad \pi_2 : X \times Y \to Y : (x,y) \mapsto y$$

toutes deux mesurables.

(ii) Pour tout $n, m \geq 1$, on a $\mathbb{B}_{n+m} = \mathbb{B}_n \otimes \mathbb{B}_m$ sur $\mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m$.

Démonstration. (i) Si $A \in \mathcal{A}$, alors $\pi_1^{-1}(A) = A \times Y \in \mathcal{A} \otimes \mathcal{A}'$ ce qui montre que $\mathcal{A} \otimes \mathcal{A}'$ rend π_1 mesurable. De même, $\mathcal{A} \otimes \mathcal{A}'$ rend π_2 mesurable. Soit à présent $\mathcal{C}$ une σ-algèbre sur $X \times Y$ rendant π_1 et π_2 mesurables. Alors pour tout $A \in \mathcal{A}$ et tout $A' \in \mathcal{A}'$, on a alors $\pi_1^{-1}(A) \cap \pi_2^{-1}(A') = A \times A' \in \mathcal{C}$. Par conséquent, $\mathcal{C}$ est une σ-algèbre qui contient les rectangles mesurables et on déduit $\mathcal{A} \otimes \mathcal{A}' \subset \mathcal{C}$.

(ii) Disons $n = m = 1$ pour simplifier la notation. On a vu que $\mathbb{B}_2$ est engendrée par les ensembles de la forme $[a,b] \times [c,d]$ avec $a,b,c,d \in \mathbb{R}$. Comme $[a,b] \times [c,d] \in \mathbb{B} \otimes \mathbb{B}$, on déduit $\mathbb{B}_2 \subset \mathbb{B} \otimes \mathbb{B}$. Par ailleurs, comme les projections sur chaque facteur sont continues, elles sont boréliennes, c'est-à-dire $\mathbb{B}_2$-mesurables. Par (i), cela implique $\mathbb{B} \otimes \mathbb{B} \subset \mathbb{B}_2$. $\square$

Étant donné deux espaces de mesure $(X, \mathcal{A}, \mu)$ et $(Y, \mathcal{A}', \nu)$, tournons-nous à présent vers la définition d'une mesure produit $\mu \otimes \nu$ satisfaisant (3.1). Une façon de procéder serait de définir une mesure extérieure à partir de la mesure définies sur les rectangles mesurables, en utilisant les techniques du chapitre précédent. Une autre façon de procéder est de s'inspirer du théorème de Fubini que nous souhaitons également démontrer : pour $E \subset \mathcal{A} \otimes \mathcal{A}'$, si le théorème de Fubini était valable, on aurait

$$\mu \otimes \nu(E) = \int_{X \times Y} \mathbb{1}_E d(\mu \otimes \nu) = \int_Y \Big(\int_X \mathbb{1}_E(x,y) d\mu(x) \Big) d\nu(y) = \int_X \Big(\int_Y \mathbb{1}_E(x,y) d\nu(y) \Big) d\mu(x).$$

En termes des sections verticales et horizontales de E, définies pour tous $x \in X$ et $y \in Y$,

$$E_x = \{y \in Y : (x,y) \in E\} \subset Y, \qquad E^y = \{x \in X : (x,y) \in E\} \subset X,$$

on aurait donc

$$\mu \otimes \nu(E) = \int_Y \mu(E^y) d\nu(y) = \int_X \nu(E_x) d\mu(x). \tag{3.2}$$

Cette façon de calculer par tranches la mesure sur un produit cartésien est le principe de Cavalieri, qui n'est autre que la version de Fubini pour des indicatrices. On va montrer que les intégrales dans (3.2) sont bien définies, sont en effet égales, et peuvent être utilisées pour construite la mesure produit $\mu \otimes \nu$.

La première question est de vérifier que les sections E_x et E^y sont bien toujours mesurables, ce qui permettra alors de définir leurs mesures.

Lemme 3.3. *Soient $(X, \mathcal{A})$ et $(Y, \mathcal{A}')$ deux espaces mesurables. Si $E \in \mathcal{A} \otimes \mathcal{A}'$, alors pour tout $x \in X$ et tout $y \in Y$ on a $E_x \in \mathcal{A}'$ et $E^y \in \mathcal{A}$.*

Démonstration. Fixons $x \in X$ et notons $\mathcal{F}$ la collection des ensembles $E \subset X \times Y$ tels que $E_x \in \mathcal{A}'$. Il suffit de montrer que $\mathcal{F}$ est une σ-algèbre qui contient les rectangles mesurables. Soient donc $A \in \mathcal{A}$ et $A' \in \mathcal{A}'$. On a

$$(A \times A')_x = \{y \in Y : (x,y) \in A \times A'\} = \begin{cases} \varnothing & : \quad \text{si } x \notin A \\ A' & : \quad \text{si } x \in A \end{cases}$$

et donc $(A \times A')_x \in \mathcal{A}'$. Il reste à montrer que $\mathcal{F}$ est une σ-algèbre. Comme $X \times Y$ est un rectangle mesurable, on a $X \times Y \in \mathcal{F}$. Si $E \in \mathcal{F}$, alors $E_x \in \mathcal{A}'$ et en remarquant que $(E_x)^c = (E^c)_x$, on obtient $(E_x)^c \in \mathcal{A}'$. Enfin, puisque $(\cup_n E_n)_x = \cup_n (E_n)_x$, on voit que $\mathcal{F}$ est stable par union dénombrable. $\square$

Afin de donner un sens aux intégrales dans (3.2), il reste à montrer que les applications $x \mapsto \nu(E_x)$ et $y \mapsto \mu(E^y)$ sont mesurables. La σ-finitude des espaces est cruciale ici.

Lemme 3.4. *Soient $(X, \mathcal{A}, \mu)$ et $(Y, \mathcal{A}', \nu)$ deux espaces de mesure σ-finis. Pour tout $E \in \mathcal{A} \otimes \mathcal{A}'$, les applications $X \to [0, \infty] : x \mapsto \nu(E_x)$ et $Y \to [0, \infty] : y \mapsto \mu(E^y)$ sont respectivement $\mathcal{A}$- et $\mathcal{A}'$-mesurables.*

Démonstration. Par le Lemme 3.3, les deux applications considérées sont bien définies. On commence par supposer que ν est une mesure finie, avant de passer au cas σ-fini général.

Cas 1 : supposons $\nu(Y) < \infty$. Notons $\mathcal{D}$ la collection des ensembles $E \in \mathcal{A} \otimes \mathcal{A}'$ pour lesquels l'application $x \mapsto \nu(E_x)$ est mesurable. Alors si $A \in \mathcal{A}$ et $A' \in \mathcal{A}'$, on a

$$\nu((A \times A')_x) = \nu(A') \mathbb{1}_A(x)$$

qui est bien une application mesurable. Donc $A \times A' \in \mathcal{D}$. Pour montrer que $\mathcal{D} = \mathcal{A} \otimes \mathcal{A}'$, il reste à montrer que $\mathcal{D}$ est une σ-algèbre. Comme les rectangles mesurables forment un π-système, le théorème de classe monotone assure qu'il suffit en fait de monter que $\mathcal{D}$ est une classe monotone. Tout d'abord, comme $X \times Y$ est un rectangle mesurable, on a $X \times Y \in \mathcal{D}$. Ensuite, pour $E \in \mathcal{D}$, comme ν est une mesure finie, on peut écrire $\nu((E^c)_x) = \nu(Y) - \nu(E_x)$, et donc $E^c \in \mathcal{D}$. Anfin, pour $(E_n)_n \subset \mathcal{D}$ deux à deux disjoints, on a

$$\nu\left(\left(\bigcup_n E_n\right)_x\right) = \nu\left(\bigcup_n (E_n)_x\right) = \sum_n \nu((E_n)_x),$$

qui est une application mesurable comme limite d'applications mesurables, donc $\cup_n E_n \in \mathcal{D}$.

Cas 2 : supposons ν σ-finie. Soit $(A'_n) \subset \mathcal{A}'$ avec $A'_n \uparrow Y$ et $\nu(A'_n) < \infty$ pour tout n. Définissons alors $\nu_n(A') := \nu(A' \cap A'_n)$ pour $A' \in \mathcal{A}'$. Par définition, ν'_n est une mesure finie sur $(Y, \mathcal{A}')$. Par le Cas 1, on a donc que $x \mapsto \nu_n(E_x)$ est $\mathcal{A}$-mesurable pour tout $E \in \mathcal{A} \otimes \mathcal{A}'$. Comme $\nu_n(E_x) = \nu(E_x \cap A'_n) \uparrow \nu(E_x)$, on déduit que $x \mapsto \nu(E_x)$ est mesurable comme limite de fonctions mesurables. □

Maintenant que les différentes intégrales intervenant dans (3.2) sont bien définies, nous allons montrer qu'elles sont en effet égales et qu'elles peuvent être utilisées pour définir la mesure produit. On insiste que l'hypothèse de σ-finitude est cruciale ici.

Théorème 3.5. *Soient $(X, \mathcal{A}, \mu)$ et $(Y, \mathcal{A}', \nu)$ deux espaces de mesure σ-finis. Il existe une unique mesure $\mu \otimes \nu$ sur l'espace mesurable $(X \times Y, A \otimes \mathcal{A}')$ telle que, pour tous $A \in \mathcal{A}$, $A' \in \mathcal{A}'$,*

$$\mu \otimes \nu(A \times A') = \mu(A)\nu(A').$$

De plus, pour tout $E \in \mathcal{A} \otimes \mathcal{A}'$, on a le principe de Cavalieri

$$\mu \otimes \nu(E) = \int_X \nu(E_x) d\mu(x) = \int_Y \mu(E^y) d\nu(y). \tag{3.3}$$

Démonstration. Posons

$$\xi_1(E) = \int_X \nu(E_x) d\mu(x) \quad \text{et} \quad \xi_2(E) = \int_Y \mu(E^y) d\nu(y).$$

Les lemmes précédents assurent que ces applications sont bien définies. On remarque que pour tous $A \in \mathcal{A}$, $A' \in \mathcal{A}'$,

$$\xi_1(A \times A') = \int_X \nu((A \times A')_x) d\mu(x) = \int_X \nu(A') \mathbb{1}_A(x) d\mu(x) = \mu(A)\nu(A'),$$

et de même $\xi_2(A \times A') = \mu(A)\nu(A')$. On rappelle que les rectangles mesurables forment un π-système et on remarque que la σ-finitude de μ et ν implique la σ-finitude de ξ_1 et ξ_2 sur les rectangles mesurables. Par le théorème d'unicité de mesure, il reste donc à vérifier que ξ_1 et ξ_2 sont des mesures sur $\mathcal{A} \otimes \mathcal{A}'$. On se focalise sur ξ_1. On a $\xi_1(\varnothing) = \int_X \nu(\varnothing) d\mu(x) = 0$. De plus, si $(E_n)_n \subset \mathcal{A} \otimes \mathcal{A}'$ deux à deux disjoints, on obtient

$$\xi_1\left(\bigcup_n E_n\right) = \int_X \nu\left(\bigcup_n (E_n)_x\right) d\mu(x) = \int_X \sum_n \nu((E_n)_x) d\mu(x) = \sum_n \int_X \nu((E_n)_x) d\mu(x) = \sum_n \xi_1(E_n),$$

où la troisième égalité découle du Corollaire 1.37. Ceci montre que ξ_1 est une mesure. Il en est de même pour ξ_2 et la conclusion s'ensuit. □

Remarque 3.6. Pour la mesure de Lebesgue restreinte aux boréliens, on a

$$\mathcal{L}_{n+m}|_{\mathbb{B}_{n+m}} = \mathcal{L}_n|_{\mathbb{B}_n} \otimes \mathcal{L}_m|_{\mathbb{B}_m} \qquad \text{sur } \mathbb{B}_{n+m} = \mathbb{B}_n \otimes \mathbb{B}_m.$$

En effet, ces mesures coïncident sur les rectangles mesurables et la conclusion suit donc du théorème d'unicité de mesures.

3.2 Théorème de Fubini

À partir du principe de Cavalieri (3.3), qui constitue une version du théorème de Fubini pour les fonctions indicatrices, on déduit ensuite le théorème de Fubini dans toute sa généralité. On commence par vérifier la mesurabilité des sections $f(\cdot, y) : x \mapsto f(x, y)$ et $f(x, \cdot) : y \mapsto f(x, y)$ pour tous x, y, lorsque f est une fonction mesurable définie sur un espace produit.

Lemme 3.7. *Soient $(X, \mathcal{A})$ et $(Y, \mathcal{A}')$ deux espaces mesurables et $f : X \times Y \to \overline{\mathbb{R}}$ $\mathcal{A} \otimes \mathcal{A}'$-mesurable. Alors pour tous $x \in X$ et $y \in Y$ les applications $f(x, \cdot)$ et $f(\cdot, y)$ sont mesurables.*

Démonstration. Pour tout $B \in \mathbb{B}$, on a $f_x^{-1}(B) = \{y \in Y : f(x, y) \in B\} = f^{-1}(B)_x$. Puisque f est mesurable, on a $f^{-1}(B) \in \mathcal{A} \otimes \mathcal{A}'$ et la conclusion se déduit en utilisant le Lemme 3.3. □

On peut à présent déduire le théorème de Fubini. On divise l'énoncé en trois parties : la première concerne les fonctions positives, la deuxième donne un critère pour l'intégrabilité sur le produit, et enfin la dernière concerne les fonctions intégrables générales. Notons que la σ-finitude de la mesure est cruciale. Par la Remarque 3.6, on peut appliquer ce résultat à des fonctions boréliennes sur $\mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m$.

Théorème 3.8 (Théorème de Fubini)**.** *Soient $(X, \mathcal{A}, \mu)$ et $(Y, \mathcal{A}', \nu)$ deux espaces de mesure σ-finis.*

(i) Si $f : X \times Y \to [0, \infty]$ est $\mathcal{A} \otimes \mathcal{A}'$-mesurable, alors les applications $x \mapsto \int_Y f(x, y) d\nu(y)$ et $y \mapsto \int_X f(x, y) d\mu(x)$ sont mesurables et l'on a

$$\int_{X\times Y} f d(\mu \otimes \nu) = \int_X \Big(\int_Y f(x, y) d\nu(y) \Big) d\mu(x) = \int_Y \Big(\int_X f(x, y) d\mu(x) \Big) d\nu(y). \tag{3.4}$$

(ii) Une fonction $\mathcal{A} \otimes \mathcal{A}'$-mesurable $f : X \times Y \to \overline{\mathbb{R}}$ est $\mu \otimes \nu$-intégrable si et seulement si

$$\int_X \Big(\int_Y |f(x, y)| d\mu(y) \Big) d\nu(x) < \infty \quad \textit{(ou vice versa).}$$

(iii) Pour $f \in \mathcal{L}^1(X \times Y, \mathcal{A} \otimes \mathcal{A}', \mu \otimes \nu)$, on a $f(\cdot, y) \in \mathcal{L}^1(X)$ pour presque tout y, disons pour tout $y \notin M$ avec $M \in \mathcal{A}'$ et $\nu(M) = 0$. De même $f(x, \cdot) \in \mathcal{L}^1(Y)$ pour presque tout x, disons pour tout $x \notin N$ avec $N \in \mathcal{A}$ et $\mu(N) = 0$. On peut alors définir

$$\phi(x) = \begin{cases} \int_Y f(x, y) d\nu(y) & : \quad x \notin N, \\ 0 & : \quad x \in N, \end{cases} \qquad \psi(y) = \begin{cases} \int_X f(x, y) d\mu(x) & : \quad y \notin M, \\ 0 & : \quad y \in M. \end{cases}$$

Enfin, en ces termes, on a $\phi \in \mathcal{L}^1(X)$, $\psi \in \mathcal{L}^1(Y)$, et

$$\begin{aligned} \int_{X\times Y} f d(\mu \otimes \nu) &= \int_X \phi d\mu = \int_{N^c} \Big(\int_Y f(x, y) d\nu(y) \Big) d\mu(x) \\ &= \int_Y \psi d\nu = \int_{M^c} \Big(\int_X f(x, y) d\mu(x) \Big) d\nu(y). \end{aligned}$$

Démonstration. (i) Soit $f : X \times Y \to [0, \infty]$ est mesurable. Pour tout $x \in X$, comme $f(x, \cdot)$ est mesurable positive, on déduit que $\int_Y f(x, y) d\nu(y)$ est bien définie. Pour montrer le reste de l'énoncé, on considère d'abord le cas où $f = \mathbb{1}_E$ avec $E \in \mathcal{A} \otimes \mathcal{A}'$. Alors $f(x, \cdot) = \mathbb{1}_{E_x}$, et les applications

$$x \mapsto \int_Y f(x, y) d\nu(y) = \nu(E_x) \qquad \text{et} \qquad y \mapsto \int_X f(x, y) d\mu(x) = \mu(E^y)$$

sont mesurables par le Lemme 3.4. De plus, par le principe de Cavalieri (3.3),

$$\int f d(\mu \otimes \nu) = (\mu \otimes \nu)(E) = \int_X \nu(E^x) d\mu(x) = \int_X \Big(\int_Y f(x,y) d\nu(y) \Big) d\mu(x),$$

ce qui montre le résultat voulu pour $f = \mathbb{1}_E$. Par linéarité, le résultat reste vrai pour toute fonction mesurable positive. Pour une fonction positive mesurable quelconque, il suffit de l'approcher par une suite croissante de fonctions simples mesurables et d'utiliser le théorème de la convergence monotone.

(ii) Ce point se déduit de (i) pour la fonction mesurable positive $|f|$.

(iii) Pour $f \in \mathcal{L}^1(X \times Y)$, on a f_+, f_- mesurables positives avec $\int_{X \times Y} f_\pm d(\mu \otimes \nu) < \infty$. Par le point (i), on sait que $x \mapsto \int_Y f_\pm(x,y) d\nu(y)$ est mesurable et que

$$\int_X \Big(\int_Y f_\pm(x,y) d\nu(y) \Big) d\mu(x) = \int_{X \times Y} f_\pm d(\mu \otimes \nu) < \infty. \tag{3.5}$$

Par Lemme 1.48(ii), on déduit $\int_Y f_\pm(x,y) d\nu(y) < \infty$ pour presque tout x, et donc $f(x,\cdot) \in \mathcal{L}^1(Y)$ pour presque tout x, disons pour tout $x \notin N$ avec $N \in \mathcal{A}$ et $\mu(N) = 0$. Définissons alors ϕ comme dans l'énoncé. Comme $\phi(x) = \int_Y f(x,y) d\nu(y) = \int_Y f_+(x,y) d\nu(y) - \int_Y f_-(x,y) d\nu(y)$ pour tout $x \notin N$, on obtient

$$\begin{aligned}
\int_X \phi d\mu &= \int_{N^c} \Big(\int_Y f(x,y) d\nu(y) \Big) d\mu(x) \\
&= \int_{N^c} \Big(\int_Y f_+(x,y) d\nu(y) \Big) d\mu(x) - \int_{N^c} \Big(\int_Y f_-(x,y) d\nu(y) \Big) d\mu(x) \\
&= \int_{X \times Y} f_+ d(\mu \otimes \nu) - \int_{X \times Y} f_- d(\mu \otimes \nu) \\
&= \int_{X \times Y} f d(\mu \otimes \nu),
\end{aligned}$$

ce qui donne la conclusion. □

3.3 Théorème de Fubini — version complétée

Comme le montre le résultat suivant, les mesures produit ne sont typiquement pas complètes. Cela implique en particulier que la σ-algèbre de Lebesgue sur $\mathbb{R}^2$ n'est pas le produit des σ-algèbres de Lebesgue sur $\mathbb{R}$ (contrairement aux σ-algèbres de Borel) ! Dès lors, la version précédente du théorème de Fubini ne peut pas s'appliquer directement à des fonctions Lebesgue-mesurables sur des produits.

Lemme 3.9.

(i) Soient $(X, \mathcal{A}, \mu)$ et $(Y, \mathcal{A}', \nu)$ des espaces de mesures. Supposons qu'il existe $A \in \mathcal{A}$ avec $A \neq \varnothing$ et $\mu(A) = 0$, et supposons qu'il existe $A' \in \mathcal{P}(Y) \setminus \mathcal{A}'$. Alors la mesure $\mu \otimes \nu$ n'est pas complète.

(ii) Sous l'axiome du choix, $\mathcal{M}_n \otimes \mathcal{M}_m \subsetneq \mathcal{M}_{n+m}$ pour $n, m \geq 1$

(iii) $(\mathbb{R}^{n+m}, \mathcal{M}_{n+m}, \mathcal{L}_{n+m})$ est la complétion de $(\mathbb{R}^n \times \mathbb{R}^m, \mathcal{M}_n \otimes \mathcal{M}_m, \mathcal{L}_n \otimes \mathcal{L}_m)$.

Démonstration. (i) On observe d'abord que $A \times A' \notin \mathcal{A} \otimes \mathcal{A}'$. En effet, on a $(A \times A')_{x_1} = A' \notin \mathcal{A}'$ pour $x_1 \in A$, alors que si $A \times A' \in \mathcal{A} \otimes \mathcal{A}'$ on aurait $(A \times A')_{x_1} \in \mathcal{A}'$ pour tout x_1. Par ailleurs, $A \times A' \subset A \times Y$ avec $\mu \otimes \nu(A \times Y) = \mu(A)\nu(Y) = 0$. Ceci montre que $\mu \otimes \nu$ n'est pas complète.

(ii) En appliquant (i) avec l'exemple de Vitali, on voit que la σ-algèbre produit $\mathcal{M}_n \otimes \mathcal{M}_m$ n'est pas complète, alors que $\mathcal{M}_{n+m}$ l'est.

(iii) Ceci se déduit facilement du fait que $(\mathbb{R}^d, \mathcal{M}_d, \mathcal{L}_d)$ est la complétion de $(\mathbb{R}^d, \mathbb{B}_d, \mathcal{L}_d|_{\mathbb{B}_d})$, cf. Théorème 2.26, en rappelant que $\mathbb{B}_{n+m} = \mathbb{B}_n \otimes \mathbb{B}_m$, cf. Lemme 3.2. On laisse le détail en exercice. □

En vertu du point (iii) ci-dessus, afin de généraliser le théorème de Fubini à des fonctions Lebesgue-mesurable, on considère plus généralement des fonctions mesurables sur la *complétion* d'un espace produit. Le lemme suivant va nous permettre de nous ramener au cas de fonctions mesurables sur le produit non complété.

Lemme 3.10.

(i) Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $(X, \bar{\mathcal{A}}, \bar{\mu})$ l'espace de mesure complété. Pour tout $f : X \to \overline{\mathbb{R}}$ $\bar{\mathcal{A}}$-mesurable, il existe $g : X \to \overline{\mathbb{R}}$ $\mathcal{A}$-mesurable telle que $g = f$ μ-p.p.

(ii) Soient $(X, \mathcal{A}, \mu)$ et $(Y, \mathcal{A}', \nu)$ deux espaces de mesure σ-finis. Soit $h : X \times Y \to \overline{\mathbb{R}}$ avec $h = 0$ $\mu \otimes \nu$-p.p. Alors pour presque tous x, y on a $h(x, \cdot) = 0$ p.p. et $h(\cdot, y) = 0$ p.p.

Démonstration. (i) Il suffit de montrer le résultat pour $f = \mathbb{1}_A$ avec $A \in \bar{\mathcal{A}}$. Par définition de la complétion, il existe $E, F \in \mathcal{A}$ avec $E \subset A \subset F$ et $\mu(F \setminus E) = 0$. Dès lors, on a $\mathbb{1}_A = \mathbb{1}_E$ sur $(F \setminus E)^c$, où $\mathbb{1}_E$ est $\mathcal{A}$-mesurable.

(ii) Par définition, il existe $N \in \mathcal{A} \otimes \mathcal{A}'$ avec $h = 0$ sur N^c et $\mu \otimes \nu(N) = 0$. Dès lors, pour tout x, on a $h(x, \cdot) = 0$ sur N_x^c. Or, le principe de Cavalieri donne $0 = \mu \otimes \nu(N) = \int_X \nu(N_x) d\mu(x)$, et donc $\nu(N_x) = 0$ pour presque tout x. □

Pour une fonction mesurable par rapport à la complétion de la σ-algèbre produit, ce lemme va nous permettre de déduire une version correspondante du théorème de Fubini en nous ramenant au cas d'une fonction mesurable par rapport à la σ-algèbre produit non-complétée. Par le Lemme 3.9(iii), on peut appliquer ce résultat à des fonctions Lebesgue-mesurables sur $\mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m$.

Théorème 3.11 (Théorème de Fubini, version complétée)**.** *Soient $(X, \mathcal{A}, \mu)$ et $(Y, \mathcal{A}', \nu)$ deux espaces de mesure σ-finis complets, et notons $(X \times Y, \overline{\mathcal{A} \otimes \mathcal{A}'}, \overline{\mu \otimes \nu})$ l'espace produit complété.*

(i) Si $f : X \times Y \to [0, \infty]$ est $\overline{\mathcal{A} \otimes \mathcal{A}'}$-mesurable, alors $f(\cdot, y)$ est mesurable pour presque tout y, disons pour tout $y \notin M$ avec $M \in \mathcal{A}'$ et $\nu(M) = 0$. De même, $f(x, \cdot)$ est mesurable pour presque tout x, disons pour tout $x \notin N$ avec $N \in \mathcal{A}$ et $\mu(N) = 0$. On peut alors définir les fonctions

$$\phi(x) = \begin{cases} \int_Y f(x,y) d\nu(y) & : \quad x \notin N, \\ 0 & : \quad x \in N, \end{cases} \qquad \psi(y) = \begin{cases} \int_X f(x,y) d\mu(x) & : \quad y \notin M, \\ 0 & : \quad y \in M, \end{cases}$$

qui sont toutes deux mesurables, et l'on a

$$\begin{aligned} \int_{X \times Y} f d(\overline{\mu \otimes \nu}) &= \int_X \phi d\mu = \int_{N^c} \Big(\int_Y f(x,y) d\nu(y) \Big) d\mu(x) \\ &= \int_Y \psi d\nu = \int_{M^c} \Big(\int_X f(x,y) d\mu(x) \Big) d\nu(y). \end{aligned}$$

(iii) Si $f \in \mathcal{L}^1(X \times Y, \overline{\mathcal{A} \otimes \mathcal{A}'}, \overline{\mu \otimes \nu})$, alors on a $f(\cdot, y) \in \mathcal{L}^1(X)$ pour presque tout y, disons pour tout $y \notin M$ avec $M \in \mathcal{A}'$ et $\nu(M) = 0$. De même $f(x, \cdot) \in \mathcal{L}^1(Y)$ pour presque tout x, disons pour tout $x \notin N$ avec $N \in \mathcal{A}$ et $\mu(N) = 0$. On peut alors définir ϕ et ψ comme en (i), on a $\phi \in \mathcal{L}^1(X, \mathcal{A}, \mu)$, $\psi \in \mathcal{L}^1(Y, \mathcal{A}', \nu)$, et les mêmes égalités sont valables qu'en (i).

Démonstration. Focalisons-nous sur le point (i). Par le Lemme 3.10(i), on peut écrire $f = g + h$ sur $X \times Y$ avec g $\mathcal{A} \otimes \mathcal{A}'$-mesurable et $h = 0$ p.p. Par le Lemme 3.7, l'application $g(x, \cdot)$ est mesurable pour tout x. Par le Lemme 3.10(ii), pour presque tout x, on a $h(x, \cdot) = 0$ p.p., et donc par complétude on déduit que $h(x, \cdot)$ est mesurables. Ceci montre que pour presque tout x l'application $f(x, \cdot) = g(x, \cdot) + h(x, \cdot)$ est mesurable. Ceci nous permet alors de définir ϕ comme dans l'énoncé. Pour presque tout x, comme $h(x, \cdot) = 0$ p.p., on remarque $\phi(x) = \int_Y g(x,y) d\nu(y)$. Or, par la première version du théorème de Fubini, cf. Théorème 3.8, la fonction $\phi_0(x) = \int_Y g(x,y)d\nu(y)$ est bien définie pour tout x, est mesurable, et satisfait $\int_X \phi_0 d\mu = \int_{X\times Y} gd(\mu \otimes \nu)$. On déduit donc

$$\int_{X\times Y} gd(\mu \otimes \nu) = \int_X \phi_0 d\mu = \int_X \phi d\mu.$$

Comme g est $\mathcal{A} \otimes \mathcal{A}'$-mesurable, le membre de gauche est égal à $\int_{X\times Y} gd(\overline{\mu \otimes \nu})$, et la conclusion s'ensuit comme $f = g$ p.p. $\square$

Chapitre 4

Décomposition de mesures et théorème de Radon-Nikodym

Ce chapitre est consacré à deux théorèmes de structure fondamentaux en théorie de la mesure : le théorème de décomposition de Lebesgue et le théorème de Radon-Nikodym. Le premier montre que toute mesure peut être scindée de façon canonique en une partie 'absolument continue' et une partie 'singulière' par rapport à une mesure de référence. Le second décrit explicitement la partie absolument continue au moyen d'une densité intégrable. Ensemble, ces résultats rendent la théorie conceptuellement transparente : ils élucident la structure des mesures générales et permettent en partie de réduire leur étude à celle de fonctions mesurables positives.

4.1 Densités, mesures absolument continues et singulières

On commence par l'observation suivante : toute fonction mesurable positive sur un espace de mesure permet de construire une autre mesure dont elle est la fonction de densité.

Lemme 4.1. *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $f : X \to [0, \infty]$ mesurable. Alors l'application*

$$\nu : \mathcal{A} \to [0, \infty] : A \mapsto \int_A f d\mu$$

est une mesure sur $\mathcal{A}$. De plus, on a :

— *pour $g : X \to [0, \infty]$ mesurable, on a $\int g d\nu = \int f g d\mu$;*

— *pour $g : X \to \mathbb{R}$ mesurable, on a $g \in L^1(\nu)$ si et seulement si $fg \in L^1(\mu)$, et dans ce cas on a également $\int g d\nu = \int f g d\mu$.*

On dit que f est une densité *pour ν par rapport à μ, et on note $d\nu = f d\mu$.*

Démonstration. On commence par vérifier que ν est une mesure. On a bien sûr $\nu(\varnothing) = 0$. De plus, pour $(A_n)_n \subset \mathcal{A}$ deux à deux disjoints, on a par le théorème de convergence monotone,

$$\nu\Big(\bigcup_n A_n\Big) = \int_{\cup_n A_n} f d\mu = \int \Big(\sum_n f \mathbb{1}_{A_n}\Big) d\mu = \sum_n \int_{A_n} f d\mu = \sum_n \nu(A_n).$$

Il reste à monter les identités sur les intégrales par rapport à ν. Si $g = \mathbb{1}_A$ avec $A \in \mathcal{A}$, l'identité $\int g d\nu = \int f g d\mu$ découle de la définition de ν. Par linéarité, la même identité vaut pour toute fonction simple mesurable positive. Par convergence monotone, on déduit qu'il en va de même pour toute fonction mesurable positive. Pour g intégrable, le résultat découle de même en divisant en parties positive et négative. $\square$

Lorsqu'une mesure a une densité par rapport à une autre, elle est simplement caractérisée par sa fonction densité, ce qui simplifie grandement sa description. Mais comment savoir si une mesure possède une densité ? Comme on le verra plus loin, le bon critère est donné par la notion d'absolue continuité.

Définition 4.2. Soient $(X, \mathcal{A})$ un espace mesurable et μ, ν deux mesures sur $(X, \mathcal{A})$. On dit que ν est *absolument continue* par rapport à μ si pour tout $A \in \mathcal{A}$ tel que $\mu(A) = 0$, on a $\nu(A) = 0$. Dans ce cas, on écrit $\nu \ll \mu$.
Sur $(\mathbb{R}^d, \mathbb{B}_d)$, on dit simplement qu'une mesure est *absolument continue* si elle est absolument continue par rapport à la mesure de Lebesgue.

Si ν a une densité par rapport à μ, on voit bien sûr que $\nu \ll \mu$ en ce sens. Le théorème de Radon-Nikodym affirmera que la réciproque est également vraie ; l'absolue continuité fournira ainsi un critère simple pour vérifier si une mesure a une densité par rapport à une autre.

Maintenant, qu'en est-il des mesures qui ne sont pas absolument continues l'une par rapport à l'autre ? En prenant le complémentaire de la définition d'absolue continuité, on arrive à la notion de mesures singulières.

Définition 4.3. Soit $(X, \mathcal{A})$ un espace mesurable et soient μ, ν deux mesures sur $(X, \mathcal{A})$. On dit que μ et ν sont *mutuellement singulières* s'il existe un ensemble $D \in \mathcal{A}$ tel que

$$\mu(D) = 0 \quad \text{et} \quad \nu(D^c) = 0.$$

Dans ce cas, on écrit $\mu \perp \nu$. Sur $(\mathbb{R}^d, \mathbb{B}^d)$, une mesure est dite *singulière* si elle est mutuellement singulière avec la mesure de Lebesgue.

Exemple 4.4. Si μ est une mesure discrète sur $(\mathbb{R}, \mathbb{B})$, c'est-à-dire s'il existe D dénombrable tel que $\mu(D^c) = 0$, alors μ est singulière puisque $\mathcal{L}(D) = 0$. Il existe par ailleurs des mesures singulières sur $(\mathbb{R}, \mathbb{B})$ qui ne sont pas discrètes : on peut considérer par exemple la mesure μ donnée par $\mu((-\infty, t]) = F(t)$ avec F l'escalier de Cantor ; on parle dans ce cas de mesure singulière continue. On reviendra sur cette distinction dans la suite.

Comme on le verra plus loin, le théorème de décomposition de Lebesgue affirmera que toute mesure se décompose de façon unique en une partie absolument continue et une partie singulière par rapport à une mesure de référence, élucidant ainsi la structure de mesures générales.

Avant d'énoncer et prouver ce résultat, remarquons quelques propriétés élémentaires, indiquant les relations entre les notions d'absolue continuité et de singularité.

Lemme 4.5. *Soient $(X, \mathcal{A})$ un espace mesurable et μ, ν, ν' des mesures sur $(X, \mathcal{A})$.*

(i) Si $\nu \ll \mu$ et $\nu \perp \mu$, alors $\nu = 0$.

(ii) Si $\nu \perp \mu$ et $\nu' \perp \mu$, alors $\nu + \nu' \perp \mu$.

(iii) Si $\nu \ll \mu$ et $\nu' \perp \mu$, alors $\nu \perp \nu'$.

Démonstration. (i) Soit $A \in \mathcal{A}$ tel que $\mu(A) = 0$ et $\nu(A^c) = 0$. Pour $B \in \mathcal{A}$ avec $B \subset A$, on a $\mu(B) \leq \mu(A) = 0$, et donc $\nu(B) = 0$ comme $\nu \ll \mu$. Si $B \in \mathcal{A}$ avec $B \subset A^c$, on a $\nu(B) \leq \nu(A^c) = 0$ et donc aussi $\nu(B) = 0$. Ainsi, pour tout $B \in \mathcal{A}$, on a $\nu(B) = \nu(A \cap B) + \nu(A^c \cap B) = 0$.
(ii) Soient $A, A' \in \mathcal{A}$ tels que $\mu(A) = 0$, $\nu(A^c) = 0$ et $\mu(A) = 0$, $\nu'((A')^c) = 0$. Alors $A \cup A'$ est tel que $\mu(A \cup A') = 0$ et $(\nu + \nu')((A \cup A')^c) = \nu(A^c \cap (A')^c) + \nu'(A^c \cap (A')^c) \leq \nu(A^c) + \nu_2((A')^c) = 0$.
(iii) Soit $A \in \mathcal{A}$ tel que $\mu(A) = 0$ et $\nu'(A^c) = 0$. Comme $\nu \ll \mu$, on a aussi $\nu(A) = 0$. $\square$

4.2 Décomposition de Lebesgue et théorème de Radon-Nikodym

On commence par une version très affaiblie du théorème de Radon-Nikodym, que l'on déduira ensuite. Notons que l'hypothèse $\nu \leq \mu$ est beaucoup plus forte que $\nu \ll \mu$, et c'est elle qui permet de conclure que la densité de ν est ici à valeurs dans $[0, 1]$.

Lemme 4.6. *Soient $(X, \mathcal{A})$ un espace mesurable et μ, ν deux mesures finies sur $(X, \mathcal{A})$ avec $\nu \leq \mu$. Alors il existe une fonction mesurable $f : X \to [0, 1]$ telle que pour tout $A \in \mathcal{A}$,*

$$\nu(A) = \int_A f d\mu.$$

Démonstration. Considérons

$$\mathcal{H} := \Big\{ h : X \to [0, \infty] \text{ mesurable} : \int_A h \, d\mu \leq \nu(A) \ \forall A \in \mathcal{A} \Big\}.$$

Étape 1 : montrons qu'il existe $f \in \mathcal{H}$ à valeurs dans $[0, 1]$ telle que

$$\int f d\mu = \sup \Big\{ \int h \, d\mu : h \in \mathcal{H} \Big\}.$$

Bien sûr, $\mathcal{H} \neq \varnothing$ puisqu'il contient la fonction nulle. De plus, on vérifie que $\mathcal{H}$ est stable par passage au maximum : en effet, pour $h_1, h_2 \in \mathcal{H}$ et $A \in \mathcal{A}$, en notant $A_1 = \{x \in A : h_1(x) > h_2(x)\}$, on obtient

$$\int_A (h_1 \vee h_2) d\mu = \int_{A_1} h_1 d\mu + \int_{A \setminus A_1} h_2 d\mu \leq \nu(A_1) + \nu(A \setminus A_1) = \nu(A),$$

et donc $h_1 \vee h_2 \in \mathcal{H}$. Par définition du supremum, on peut considérer une suite $(h_n)_n \subset \mathcal{H}$ telle que

$$\lim_n \int h_n d\mu = \sup \Big\{ \int h \, d\mu : h \in \mathcal{H} \Big\}.$$

Quitte à remplacer h_n par $h_1 \vee \cdots \vee h_n \in \mathcal{H}$, on peut supposer que la suite $(h_n)_n$ est croissante. En particulier, on peut considérer $f = \lim_n h_n = \sup_n h_n$. Le théorème de la convergence monotone donne pour tout $A \in \mathcal{A}$,

$$\int_A f d\mu = \lim_n \int_A h_n d\mu \leq \nu(A),$$

et donc $f \in \mathcal{H}$. En prenant $A = X$, on a également

$$\int f d\mu = \sup \Big\{ \int h \, d\mu : h \in \mathcal{H} \Big\}.$$

Il reste à vérifier que f peut être choisie à valeurs dans $[0, 1]$. Pour tout n, considérons les ensemble de niveau $A_n = \{x \in X : f(x) \geq 1 + 1/n\}$. Par l'inégalité de Markov et l'hypothèse $\nu \leq \mu$, on a

$$(1 + 1/n)\mu(A_n) \leq \int_{A_n} f d\mu \leq \nu(A_n) \leq \mu(A_n)$$

et donc $\mu(A_n) = 0$. Dès lors, par continuité à gauche de la mesure,

$$\mu(\{x \in X : f(x) > 1\}) = \lim_n \mu(A_n) = 0.$$

Quitte à remplacer f par une fonction égale presque partout, on peut supposer que $f \leq 1$.

Étape 2 : si $h : X \to [0,\infty]$ est une application mesurable pour laquelle il existe $A \in \mathcal{A}$ tel que $\int_B h\,d\mu \leq \nu(B)$ pour tout $B \in \mathcal{A}$ avec $B \subset A$, alors $h \leq f$ p.p. sur A.
En effet, considérons l'ensemble $B = \{x \in A : h(x) > f(x)\}$. Alors, on a $h\mathbb{1}_A \in \mathcal{H}$ et donc $(h\mathbb{1}_A) \vee f \in \mathcal{H}$. Il s'ensuit que

$$\int_B h\,d\mu + \int_{X\setminus B} fd\mu = \int (h\mathbb{1}_A) \vee fd\mu \leq \int fd\mu = \int_B fd\mu + \int_{X\setminus B} fd\mu,$$

ce qui implique $\int_B fd\mu \geq \int_B h\,d\mu$, et donc $\int_B (f-h)d\mu \geq 0$. Comme on a $f - h < 0$ sur B, cela n'est possible que si $\mu(B) = 0$, d'où $h \leq f$ p.p. sur A.

Étape 3 : pour tout $A \in \mathcal{A}$ on a

$$\nu(A) = \int_A fd\mu.$$

Par définition, on sait déjà que $\nu(A) \geq \int_A fd\mu$, et il suffit de montrer l'autre inégalité. Comme $\nu \leq \mu$, il suffit de le montrer pour tout $A \in \mathcal{A}$ tel que $\mu(A) > 0$. Soit $\varepsilon > 0$. Considérons

$$P_\varepsilon = \left\{ A \in \mathcal{A} \,:\, \int_A (f+\varepsilon)d\mu > \nu(A) \right\}.$$

Remarquons que si $A \in \mathcal{A}$ est tel que $\mu(A) > 0$, alors il existe $B \subset A$ tel que $B \in P_\varepsilon$ (et donc $\mu(B) > 0$) : en effet, sinon, on aurait $\int_B (f+\varepsilon)d\mu \leq \nu(B)$ pour tout $B \subset A$, et l'étape 2 impliquerait alors $f + \varepsilon \leq f$ p.p. sur A. Ceci n'est possible que si $\mu(A) = 0$, une contradiction.

Fixons $A \in \mathcal{A}$ tel que $\mu(A) > 0$ et posons

$$\alpha_0 = \sup\{\mu(B) \,:\, B \in P_\varepsilon, B \subset A\}.$$

On peut trouver $B_0 \in P_\varepsilon$ avec $B_0 \subset A$ tels que $\mu(B_0) \geq \frac{1}{2}\alpha_0$. Considérons à présent $A_1 = A \setminus B_0$. Si $\mu(A_1) = 0$, on s'arrête. Sinon, soit

$$\alpha_1 = \sup\{\mu(B) \,:\, B \in P_\varepsilon,\ B \subset A_1\},$$

et soit $B_1 \in P_\varepsilon$ avec $B_1 \subset A_1$ et $\mu(B_1) \geq \frac{1}{2}\alpha_1$. Considérons ensuite $A_2 = A_1 \setminus B_1$. Par induction, on construit ainsi des suites $(A_i)_{1\leq i<J+1}$, $(B_i)_{0\leq i<J+1}$, et $(\alpha_i)_{0\leq i<J+1}$ avec $J \in \mathbb{N} \cup \{\infty\}$ (le procédé peut ne jamais s'arrêter). Considérons alors

$$A_\infty = A \setminus \bigcup_{0\leq i<J+1} B_i.$$

et montrons que cet ensemble est μ-négligeable. Si la récurrence s'est terminée en un temps fini, $J < \infty$, on a $A_\infty = A_J$ et $\mu(A_J) = 0$. Sinon, puisque les B_i sont des sous-ensembles de A deux à deux disjoints, on a

$$\frac{1}{2}\sum_i \alpha_i \leq \sum_i \mu(B_i) \leq \mu(A) < \infty$$

et donc $\lim_i \alpha_i = 0$. Pour $B \subset A_\infty$ avec $B \in P_\varepsilon$, on a $B \subset A_i$ pour tout i, et donc $\mu(B) \leq \alpha_i$, ce qui implique $\mu(B) = 0$. Par conséquent, les sous-ensembles de A_∞ qui appartiennent à P_ε sont négligeables. Comme on l'a vu, ceci n'est possible que si $\mu(A_\infty) = 0$.

Dans tous les cas, on a donc montrer $\mu(A_\infty) = 0$. Comme $A = A_\infty \cup \cup_i B_i$ et comme les ensembles sont deux à deux disjoints, on en déduit

$$\int_A (f+\varepsilon)d\mu = \sum_i \int_{B_i} (f+\varepsilon)d\mu > \sum_i \nu(B_i) = \nu(A).$$

ce qui signifie $A \in P_\varepsilon$. En conclusion, pour tout $A \in \mathcal{A}$ tel que $\mu(A) > 0$ et tout $\varepsilon > 0$, on a $\int_A (f+\varepsilon)d\mu > \nu(A)$. Comme ε est arbitraire, cela signifie $\int_A fd\mu \geq \nu(A)$, ce qui conclut la preuve. $\square$

À partir du lemme précédent, on peut maintenant déduire le théorème de décomposition de Lebesgue et le théorème de Radon-Nikodym.

Théorème 4.7 (Théorème de décomposition de Lebesgue)**.** *Soient $(X, \mathcal{A})$ un espace mesurable et μ, ν deux mesures σ-finies sur $(X, \mathcal{A})$. Il existe un unique couple (ν_a, ν_s) de mesures sur $(X, \mathcal{A})$ telles que*

$$\nu = \nu_a + \nu_s, \qquad \nu_a \ll \mu, \qquad \nu_s \perp \mu.$$

On appelle le couple (ν_a, ν_s) la décomposition de Lebesgue *de ν par rapport à μ ; on dit que ν_a est la* partie absolument continue *de ν par rapport à μ, et que ν_s est sa* partie singulière *par rapport à μ.*

Démonstration. On décompose la preuve en trois étapes.

Step 1 : Cas de mesures finies.
Soient μ, ν deux mesures finies sur $(X, \mathcal{A})$, et considérons la mesure $\lambda = \mu + \nu$. On a alors $\nu \leq \lambda$ et le Lemme 4.6 nous donne une fonction mesurable $g : X \to [0, 1]$ telle que pour tout $A \in \mathcal{A}$,

$$\nu(A) = \int_A g \, d\lambda = \int_A g \, d\nu + \int_A g \, d\mu,$$

ou de façon équivalente, comme ν est une mesure finie,

$$\int_A (1 - g) \, d\nu = \int_A g \, d\mu.$$

Par un argument usuel, en invoquant la linéarité de l'intégrale et le théorème de convergence monotone, on peut déduire pour toute fonction mesurable positive $h : X \to [0, \infty]$ et pour tout $A \in \mathcal{A}$,

$$\int_A (1 - g)h \, d\nu = \int_A gh \, d\mu.$$

D'autre part, on a aussi

$$\mu(A) = \lambda(A) - \nu(A) = \int_A (1 - g) \, d\lambda.$$

Considérons $D = \{x \in X : g(x) = 1\} \in \mathcal{A}$. L'égalité précédente implique $\mu(D) = 0$. Définissons à présent

$$h = \mathbb{1}_{D^c} \frac{1}{1 - g}.$$

C'est une application mesurable positive et l'on obtient pour tout $A \in \mathcal{A}$,

$$\nu(A \cap D^c) = \int_A (1 - g)h \, d\nu = \int_A gh \, d\mu \, .$$

Pour conclure, il suffit donc de prendre comme densité $f = gh = \frac{g}{1-g} \mathbb{1}_{D^c}$.

Step 2 : Cas de mesures σ-finies.
Soient μ, ν deux mesures σ-finies sur $(X, \mathcal{A})$. Par définition de σ-finitude, on peut trouver deux suites croissantes $(X_n)_n, (Y_n)_n \subset \mathcal{A}$ telles que $X_n, Y_n \uparrow X$ et $\mu(X_n), \nu(Y_n) < \infty$ pour tout n. Considérons alors $Z_n := X_n \cap Y_n \uparrow X$ et $A_n := Z_{n+1} \setminus Z_n$. On obtient ainsi une suite $(A_n)_n \subset \mathcal{A}$ d'ensembles disjoints avec $\cup_n A_n = X$ et $\mu(A_n), \nu(A_n) < \infty$ pour tout n. Définissons alors les mesures

$$\mu_n(A) = \mu(A \cap A_n), \qquad \nu_n(A) = \nu(A \cap A_n).$$

Pour tout n, comme ce sont deux mesures finies, l'étape 1 nous donne un ensemble $D_n \in \mathcal{A}$ et une application mesurable positive f_n tels que pour tout $A \in \mathcal{A}$,

$$\mu_n(D_n) = 0, \qquad \nu_n(A) = \nu_n(A \cap D_n) + \int_A f_n d\mu_n.$$

Comme les A_n sont deux à deux disjoints, on obtient alors

$$\begin{aligned}
\nu(A) = \nu\Big(\bigcup_n A \cap A_n\Big) &= \sum_n \nu(A \cap A_n) \\
&= \sum_n \Big(\nu_n(A \cap D_n) + \int_A f_n d\mu_n\Big) \\
&= \sum_n \nu(A \cap D_n \cap A_n) + \sum_n \int_A f_n \mathbb{1}_{A_n} d\mu \\
&= \nu\Big(\bigcup_n A \cap D_n \cap A_n\Big) + \int_A \sum_n f_n \mathbb{1}_{A_n} d\mu,
\end{aligned}$$

où la dernière égalité découle de la convergence monotone. Pour conclure, il suffit de poser $D = \cup_n D_n \cap A_n$, $f = \sum_n f_n \mathbb{1}_{A_n}$, et de remarquer que

$$\mu(D) = \sum_n \mu(D_n \cap A_n) = \sum_n \mu_n(D_n) = 0.$$

Bref, on a montré qu'il existe $D \in \mathcal{A}$ et $f : X \to [0, \infty]$ mesurable tels que pour tout $A \in \mathcal{A}$,

$$\mu(D) = 0 \quad \text{et} \quad \nu(A) = \nu(A \cap D) + \int_A f d\mu.$$

En posant $\nu_s(A) = \nu(A \cap D)$ et $\nu_a(A) = \int_A f d\mu$, on déduit $\nu = \nu_a + \nu_s$ avec $\nu_a \ll \mu$ et $\nu_s \perp \mu$.

Step 3 : Unicité.
Il reste à montrer l'unicité de la décomposition. Supposons que les couples (ν_a, ν_s) et (ν'_a, ν'_s) satisfont tous deux la décomposition voulue. Alors $\nu_a + \nu_s = \nu'_a + \nu'_s$, d'où

$$\nu_a - \nu'_a = \nu'_s - \nu_s.$$

Posons $\theta = \nu_a - \nu'_a = \nu'_s - \nu_s$. (Notons que θ n'est pas une mesure positive, comme elle est à valeurs dans $\mathbb{R}$; c'est ce qu'on appelle une mesure signée.) En argumentant comme dans le Lemme 4.5(i), on déduit aisément $\theta = 0$. Plus précisément, soient $D, D' \in \mathcal{A}$ tels que $\mu(D) = \mu(D') = 0$ et $\nu_s(D^c) = \nu'_s(D'^c) = 0$. Pour $B := D \cup D'$, on obtient $\mu(B) = 0$ et donc $\nu_a(A \cap B) = \nu'_a(A \cap B) = 0$ pour tout $A \in \mathcal{A}$ comme $\nu_a, \nu'_a \ll \mu$. Par conséquent,

$$\nu_a(A) = \nu_a(A \cap B) + \nu_a(A \cap B^c) = \nu_a(A \cap B^c).$$

et de même $\nu'_a(A) = \nu'_a(A \cap B^c)$. Il vient alors

$$\theta(A) = \nu_a(A) - \nu'_a(A) = \nu_a(A \cap B^c) - \nu'_a(A \cap B^c) = \nu'_s(A \cap B^c) - \nu_s(A \cap B^c) = 0$$

puisque $A \cap B^c = A \cap D^c \cap D'^c$. □

Remarque 4.8. Plus précisément, étant donné deux mesures μ, ν sur un espace mesurable $(X, \mathcal{A})$, on a montré ci-dessus qu'il existe $D \in \mathcal{A}$ et une fonction mesurable $f : X \to [0, \infty]$ tels que pour tout $A \in \mathcal{A}$,

$$\mu(D) = 0 \qquad \text{et} \qquad \nu(A) = \underbrace{\nu(A \cap D)}_{\equiv \nu_s(A)} + \underbrace{\textstyle\int_A f d\mu}_{\equiv \nu_a(A)}.$$

Théorème 4.9 (Théorème de Radon-Nikodym)**.** *Soient* $(X,\mathcal{A})$ *un espace mesurable et* μ,ν *deux mesures* σ*-finies sur* $(X,\mathcal{A})$*. Alors* $\nu \ll \mu$ *si et seulement s'il existe une fonction mesurable* $f : X \to [0,\infty]$ *telle que pour tout* $A \in \mathcal{A}$*,*

$$\nu(A) = \int_A f\, d\mu,$$

c'est-à-dire $d\nu = f d\mu$*. De plus, si* g *est une autre fonction mesurable telle que* $d\nu = g d\mu$*, alors on a* $f = g$ μ*-p.p.*

Démonstration. On sait que si $d\nu = f d\mu$, alors $\nu \ll \mu$. Réciproquement, supposons $\nu \ll \mu$. Par la preuve du Théorème 4.7 ci-dessus, cf. Remarque 4.8, il existe $D \in \mathcal{A}$ et $f : X \to [0,\infty]$ mesurable tels que pour tout $A \in \mathcal{A}$,

$$\mu(D) = 0 \quad \text{et} \quad \nu(A) = \nu(A \cap D) + \int_A f d\mu\,.$$

Comme $\nu \ll \mu$, on a aussi $\nu(D) = 0$, et donc

$$\nu(A) = \int_A f d\mu.$$

Il reste à vérifier l'unicité. Soit g une autre application mesurable telle que $d\nu = g d\mu$. Supposons dans un premier temps que ν est finie. Alors, comme les fonctions f et g sont μ-intégrables, elles sont finies μ-p.p. Soit $N \in \mathcal{A}$ un ensemble μ-négligeable en dehors duquel f et g sont finies. Considérons l'application mesurable $h := (f-g)\mathbb{1}_{N^c}$. On a alors pour tout $A \in \mathcal{A}$,

$$\int_A h\, d\mu = 0$$

En appliquant cette relation avec $A = \{x : h(x) > 0\}$, on obtient $\mu(\{x : h(x) > 0\}) = 0$. De même, $\mu(\{x : h(x) < 0\}) = 0$. Par conséquent, $h = 0$ μ-p.p., c'est-à-dire $f = g$ μ-p.p.

On considère à présent le cas général : supposons que ν est σ-fini. Soit $(A_n)_n \subset \mathcal{A}$ tel que $A_n \uparrow X$ et $\nu(A_n) < \infty$ pour tout n. L'argument précédent montre que $f = g$ μ-p.p. sur chaque ensemble A_n, et la conclusion s'ensuit. □

Définition 4.10. Comme $d\nu = f d\mu$, la fonction f intervenant au théorème précédent est une *densité* pour ν par rapport à μ. Comme elle est unique à un ensemble μ-négligeable près, on l'appelle *la dérivée de Radon-Nikodym* de ν par rapport à μ et on la note $f := \frac{d\nu}{d\mu}$. On peut ainsi écrire $d\nu = \frac{d\nu}{d\mu} d\mu$.

Remarque 4.11. L'hypothèse de σ-finitude est cruciale dans le résultat précédent : par exemple, sur $(\mathbb{R},\mathbb{B})$, la mesure de comptage $\sharp$ n'est bien sûr pas σ-finie et on remarque que $\mathcal{L} \ll \sharp$ mais $\mathcal{L}$ n'a pas de densité par rapport à $\sharp$.

4.3 Décomposition de Lebesgue raffinée sur $\mathbb{R}^d$

Pour terminer ce chapitre par un raffinement de la décomposition de Lebesgue sur $(\mathbb{R}^d,\mathbb{B}_d)$. Dans ce cadre, on rappelle que les mesures discrètes sont un cas particulier de mesures singulières. Comme le montre le résultat suivant, les mesures discrètes sont équivalentes à des sommes de mesures de Dirac. On rappelle par ailleurs que toute mesure singulière n'est pas discrète, voir Exemple 4.4.

Définition 4.12. Sur $(\mathbb{R}^d,\mathbb{B}_d)$, une mesure μ est dite *discrète* (ou *purement atomique*) s'il existe $S \subset \mathbb{R}^d$ dénombrable tel que $\mu(S^c) = 0$. On dit que $x \in \mathbb{R}^d$ est un *atome* de μ si $\mu(\{x\}) > 0$.

Lemme 4.13. *Soit μ une mesure finie sur $(\mathbb{R}^d, \mathbb{B}_d)$. On a l'équivalence suivante :*
μ *est discrète* $\iff \exists (\alpha_n) \subset [0, \infty]$ *et* $(x_n)_n \subset \mathbb{R}^d$ *tels que* $\mu = \sum_n \alpha_n \delta_{x_n}$
$\iff \exists S \in \mathbb{B}_d$ *avec* $\mu(S^c) = 0$ *tel que x est un atome de μ pour tout* $x \in D$

Démonstration. Soit μ une mesure discrète finie sur $(\mathbb{R}^d, \mathbb{B}_d)$. On commence par montrer que forcément l'ensemble des atomes de μ est au plus dénombrable. En effet, pour tout n, l'ensemble $\{x \in \mathbb{R}^d : \mu(\{x\}) \geq 1/n\}$ est forcément fini (car sinon sa mesure serait infinie), et donc l'ensemble

$$S = \{x \in \mathbb{R}^d : \mu(\{x\}) > 0\} = \bigcup_n \{x \in \mathbb{R}^d : \mu(\{x\}) \geq 1/n\}$$

est dénombrable. On peut alors écrire pour tout $A \in \mathbb{B}_d$,

$$\mu(A) = \mu(A \cap S) = \sum_{x \in S} \mu(A \cap \{x\}) = \sum_{x \in S} \mu(\{x\}) \delta_x(A),$$

ce qui implique les deux propriétés annoncées. La réciproque s'obtient de même. □

À l'opposé de la notion de mesure discrète, on définit les mesures continues comme les mesures qui n'ont pas d'atome.

Définition 4.14. Sur $(\mathbb{R}^d, \mathbb{B}_d)$, une mesure μ est dite *continue* (ou *diffuse*) si elle ne possède pas d'atome, c'est-à-dire $\mu(\{x\}) = 0$ pour tout $x \in \mathbb{R}^d$.

Lemme 4.15. *Sur* $(\mathbb{R}^d, \mathbb{B}_d)$,

(i) la mesure de Lebesgue est diffuse, ainsi que toute mesure absolument continue ;

(ii) si une mesure finie μ est discrète et continue, alors $\mu = 0$;

(iii) si μ est une mesure continue et si ν est discrète, alors $\mu \perp \nu$;

(iv) si $d = 1$ et si μ est une mesure continue, alors la fonction de distribution $t \mapsto \mu((-\infty, t])$ est continue.

Démonstration. (ii) Si μ est discrète, par le Lemme 4.13, il existe $S \in \mathbb{B}_d$ avec $\mu(S^c)$ tel que x est un atome de μ pour tout $x \in S$. Comme μ est diffuse, on déduit $S = \varnothing$ et donc $S^c = \mathbb{R}$.

(iii) Si ν est discrète, il existe $S \subset \mathbb{R}^d$ dénombrable tel que $\nu(S^c) = 0$. Si μ est continue, on obtient alors $\mu(S) = \sum_{x \in S} \mu(\{x\}) = 0$. □

Avec ces notions, on peut maintenant montrer le raffinement suivant du théorème de décomposition de Lebesgue. Cette décomposition est très utilisée en probabilités : elle montre que toute variable aléatoire a forcément une loi mixte, donnée par la somme d'une loi discrète, une loi continue singulière, et une loi absolument continue. Une loi discrète est caractérisée par ses atomes et leurs masses, tandis qu'une loi absolument continue est caractérisée par sa densité.

Proposition 4.16. *Soit μ une mesure finie sur $(\mathbb{R}^d, \mathbb{B}^d)$. Il existe un unique triplet (μ_d, μ_s, μ_a) de mesures sur $(\mathbb{R}^d, \mathbb{B}_d)$ telles que μ_d est une mesure discrète, μ_s est une mesure continue singulière, μ_a est une mesure absolument continue, et*

$$\mu = \mu_d + \mu_s + \mu_a.$$

Démonstration. Soit S l'ensemble des atomes de μ. Comme S est dénombrable, la mesure μ_d définie par $\mu_d(A) := \mu(A \cap S)$ est discrète. Considérons à présent la décomposition de Lebesgue (μ_s, μ_a) de la mesure définie par $A \mapsto \mu(A \cap S^c)$, par rapport à la mesure de Lebesgue. On a alors $\mu = \mu_d + \mu_s + \mu_a$. En particulier, par la preuve du Théorème 4.7, cf. Remarque 4.8, il existe $D \in \mathbb{B}_d$ avec $\mu(D) = 0$ tel que $\mu_s(A) = \mu(A \cap S^c \cap D)$. Par définition de S, on voit que μ_s ne peut pas avoir d'atome, c'est-à-dire que μ_s est continue. □

4.4 Application : espérance conditionnelle

L'espérance conditionnelle constitue un outil fondamental en probabilités. Intuitivement, elle représente l'espérance “sachant” une information partielle, mais sa définition rigoureuse est non-triviale et nécessite une formulation via la théorie de la mesure. L'idée centrale est la suivante : lorsqu'on restreint notre “capacité d'observation” à une sous-σ-algèbre $\mathcal{G} \subset \mathcal{A}$, certaines distinctions entre événements deviennent invisibles. On cherche alors à remplacer une variable aléatoire intégrable X par une fonction $\mathcal{G}$-mesurable, notée $\mathbb{E}[X|\mathcal{G}]$, qui reflète au mieux l'information pertinente contenue dans X lorsque seule $\mathcal{G}$ est accessible. Formellement, la construction s'exprime au moyen d'égalités intégrales, ce qui la relie directement au cadre du théorème de Radon-Nikodym.

Définition 4.17. Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilité, soit $\mathcal{G} \subset \mathcal{A}$ une sous-σ-algèbre, et $X : \Omega \to \mathbb{R}$ une fonction $\mathcal{A}$-mesurable $\mathbb{P}$-intégrable. On appelle *espérance conditionnelle de X sachant $\mathcal{G}$* toute fonction $\mathcal{G}$-mesurable Y telle que pour tout $G \in \mathcal{G}$,

$$\mathbb{E}[\mathbb{1}_G Y] = \mathbb{E}[\mathbb{1}_G X].$$

Lemme 4.18. *Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilité, soit $\mathcal{G} \subset \mathcal{A}$ une sous-σ-algèbre, et $X : \Omega \to \mathbb{R}$ une fonction $\mathcal{A}$-mesurable $\mathbb{P}$-intégrable. L'espérance conditionnelle Y de X sachant $\mathcal{G}$ existe et est unique $\mathbb{P}$-p.p. ; on la note $Y = \mathbb{E}[X|\mathcal{G}]$.*

Démonstration. Quitte à décomposer en parties positive et négative, on peut supposer que X est positive. Considérons alors la mesure $d\mu_X = X d\mathbb{P}$ sur $(\Omega, \mathbb{P})$ avec densité X. Cette mesure est bien sûr finie et absolument continue par rapport à $\mathbb{P}$. De même, la restriction $\mu_X|_{\mathcal{G}}$ sur $(\Omega, \mathcal{G})$ est finie et absolument continue par rapport à $\mathbb{P}|_{\mathcal{G}}$. Par définition, l'espérance conditionnelle n'est autre que la dérivée de Radon-Nikodym $\mathbb{E}[X|\mathcal{G}] = \frac{d\mu_X|_{\mathcal{G}}}{d\mathbb{P}|_{\mathcal{G}}}$. $\square$

Remarque 4.19. Supposons que la sous-σ-algèbre $\mathcal{G}$ est engendrée par un événement $A \in \mathcal{A}$, c'est-à-dire que $\mathcal{G} = \sigma(\{A\}) = \{\varnothing, A, A^c, \Omega\}$, et supposons $\mathbb{P}(A) \in (0, 1)$. Dans ce cas, on remarque que toute variable aléatoire est $\mathcal{G}$-mesurable est de la forme $Y = a\mathbb{1}_A + b\mathbb{1}_{A^c}$ avec $a, b \in \mathbb{R}$. En insérant cette forme dans la relation intégrale définissant l'espérance conditionnelle, on déduit explicitement

$$\mathbb{E}[X|\sigma(\{A\})] = \frac{\mathbb{E}[\mathbb{1}_A X]}{\mathbb{P}(A)}\mathbb{1}_A + \frac{\mathbb{E}[\mathbb{1}_{A^c} X]}{\mathbb{P}(A^c)}\mathbb{1}_{A^c}.$$

Cet exemple montre bien que l'espérance conditionnelle est une version “agrégée” de X lorsqu'on ne dispose que de l'information permettant de distinguer A et A^c. Il s'agit de la généralisation naturelle de la formule usuelle pour l'espérance conditionnelle sachant un événement.

Chapitre 5

Espaces L^p

On a déjà introduit l'espace $\mathcal{L}^1$ et son quotient L^1. Dans ce chapitre, on définit plus généralement les espaces L^p, et on étudie leurs propriétés élémentaires. On rappelle l'importance de l'espace L^2, qui sera notamment le cadre naturel pour la théorie de Fourier. Grâce à la théorie de Lebesgue, on montre que ces espaces L^p sont des espaces de Banach, fondant ainsi le point de départ de l'analyse fonctionnelle moderne.

5.1 Espaces $\mathcal{L}^p$ et $\mathcal{L}^\infty$

Définition 5.1. Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $1 \leq p < \infty$. L'espace $\mathcal{L}^p(X, \mathcal{A}, \mu)$ est l'ensemble des fonctions mesurables $f : X \to \mathbb{R}$ telles que $|f|^p$ est intégrable. On note alors

$$\|f\|_{\mathcal{L}^p(X,\mathcal{A},\mu)} := \Big(\int |f|^p d\mu \Big)^{\frac{1}{p}} \in [0, \infty).$$

Lorsqu'il n'y a pas de confusion possible, on note simplement $\mathcal{L}^p = \mathcal{L}^p(X) = \mathcal{L}^p(\mu) = \mathcal{L}^p(X, \mathcal{A}, \mu)$ et $\|f\|_{\mathcal{L}^p} = \|f\|_{\mathcal{L}^p(X)} = \|f\|_{\mathcal{L}^p(\mu)} = \|f\|_{\mathcal{L}^p(X,\mathcal{A},\mu)}$.

Cette définition s'étend naturellement au cas de fonctions à valeurs complexes : $\mathcal{L}^p_{\mathbb{C}}(X, \mathcal{A}, \mu)$ est l'ensemble des fonctions mesurables $f : X \to \mathbb{C}$ telles que $|f|^p$ est intégrable. Dans la suite, on se focalisera sur les espaces de fonctions à valeurs réelles, mais tous les résultats s'adaptent immédiatement au cas complexe.

Lemme 5.2. *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $1 \leq p < \infty$. L'espace $\mathcal{L}^p(X, \mathcal{A}, \mu)$ est un espace vectoriel sur $\mathbb{R}$.*

Démonstration. Soient $f, g \in \mathcal{L}^p(X, \mathcal{A}, \mu)$ et $\alpha \in \mathbb{R}$. Alors, on a $\alpha f \in \mathcal{L}^p(X, \mathcal{A}, \mu)$. De plus, pour tout $x \in X$,

$$|f(x) + g(x)|^p \leq \big(|f(x)| + |g(x)|\big)^p \leq \big(2 \max\{|f(x)|, |g(x)|\}\big)^p \leq 2^p|f(x)|^p + 2^p|g(x)|^p$$

et donc $f + g \in \mathcal{L}^p(X, \mathcal{A}, \mu)$. □

Nous introduisons à présent l'espace correspondant pour $p = \infty$. Intuitivement, on souhaite naturellement définir $\mathcal{L}^\infty$ comme l'espace des fonctions bornées hors d'un ensemble de mesure nulle. La notion de bornitude doit cependant être adaptée pour négliger le comportement des fonctions presque nulle part : pour ce faire, on introduit la notion de *supremum essentiel* d'une fonction.

Définition 5.3. Soit $(X, \mathcal{A}, \mu)$ un espace de mesure. Pour une fonction $g : X \to [0, \infty]$, on considère $S_g := \{\alpha \in [0, \infty) : g \le \alpha\ \mu\text{-p.p.}\}$, et on définit alors

$$\operatorname{supess} g := \begin{cases} \inf S_g & : \quad S_g \ne \varnothing, \\ \infty & : \quad S_g = \varnothing. \end{cases}$$

L'espace $\mathcal{L}^\infty(X, \mathcal{A}, \mu)$ est l'ensemble des fonctions mesurables $f : X \to \mathbb{R}$ telles que $\operatorname{supess} |f| < \infty$. On note alors

$$\|f\|_{\mathcal{L}^\infty(X,\mathcal{A},\mu)} := \operatorname{supess}|f| \ \in \ [0, \infty).$$

Lorsqu'il n'y a pas de confusion possible, on note simplement $\mathcal{L}^\infty = \mathcal{L}^\infty(X) = \mathcal{L}^\infty(\mu) = \mathcal{L}^\infty(X, \mathcal{A}, \mu)$ et $\|f\|_{\mathcal{L}^\infty} = \|f\|_{\mathcal{L}^\infty(X)} = \|f\|_{\mathcal{L}^\infty(\mu)} = \|f\|_{\mathcal{L}^\infty(X,\mathcal{A},\mu)}$.

Lemme 5.4. *Soit $(X, \mathcal{A}, \mu)$ un espace de mesure. Pour des fonctions $f, g : X \to [0, \infty]$, on a*

(i) $g \le \operatorname{supess} g$ *μ-p.p.*

(ii) $\operatorname{supess}(f + g) \le \operatorname{supess} f + \operatorname{supess} g$.

En particulier, on déduit de ce dernier point que $\mathcal{L}^\infty(X, \mathcal{A}, \mu)$ est un espace vectoriel.

Démonstration. (i) On peut écrire

$$\{x : g(x) > \inf S_g\} = \bigcup_n \{x : g(x) \ge \inf S_g + \tfrac{1}{n}\}.$$

Pour tout n, par définition de l'infimum, on a $\inf S_g + \frac{1}{n} \in S_g$. Par définition de S_g, cela signifie que $\{x : g(x) \ge \inf S_g + \frac{1}{n}\}$ est μ-négligeable, et il en va donc de même pour $\{x : g(x) > \inf S_g\}$, ce qui montre que $g \le \inf S_g = \min S_g = \operatorname{supess} g$ μ-p.p.

(ii) Par le point (i), on déduit $f + g \le \operatorname{supess} f + \operatorname{supess} g$ μ-p.p., et la conclusion s'ensuit. $\square$

Remarque 5.5. Sur $(X, \mathcal{A}, \mu) = (\mathbb{N}, \mathcal{P}(\mathbb{N}), \sharp)$, on retrouve les espaces $\ell^p(\mathbb{N})$ usuels.

5.2 Inégalités de Hölder et de Minkovski

On commence par rappeler l'inégalité réelle suivante. Pour $p = 2$, on remarque qu'il s'agit de l'inégalité usuelle $2ab \le a^2 + b^2$.

Lemme 5.6. *Pour tous $a, b \ge 0$ et $1 < p, q < \infty$ tels que $\frac{1}{p} + \frac{1}{q} = 1$, on a*

$$ab \le \frac{a^p}{p} + \frac{b^q}{q}.$$

Démonstration. Bien sûr, on peut supposer que a et b sont non-nuls. En posant $u = a^p$ et $v = b^q$, il suffit de prouver pour tous $u, v > 0$,

$$u^{\frac{1}{p}} v^{\frac{1}{q}} \le \frac{u}{p} + \frac{v}{q}.$$

En posant $t = \frac{u}{v}$, il suffit de montrer pour tout $t > 0$,

$$t^{\frac{1}{p}} \le \frac{t}{p} + \frac{1}{q}.$$

Pour ce faire, il suffit de remarquer que la fonction $\mathbb{R}^+ \to \mathbb{R} : t \mapsto \frac{t}{p} + \frac{1}{q} - t^{\frac{1}{p}}$ atteint son minimum en $t = 1$ et que ce minimum vaut 0. $\square$

Définition 5.7. On dit que $p, q \in [1, \infty]$ sont des *exposants conjugués* si $\frac{1}{p} + \frac{1}{q} = 1$. On note souvent alors $q = p'$.

Remarquons que si p et q sont des exposants conjugués, alors on a notamment $p + q = pq$ et $p(q-1) = q$. Par exemple $(p, q) = (2, 2)$, $(1, \infty)$, $(3, \frac{3}{2})$ sont des paires d'exposants conjugués. En ces termes, on peut à présent montrer deux inégalités fondamentales de la théorie $\mathcal{L}^p$: la première laisse entrevoir la relation de dualité entre $\mathcal{L}^p$ et $\mathcal{L}^{p'}$, qu'on explorera plus loin ; la seconde est l'inégalité triangulaire pour $\| \cdot \|_{\mathcal{L}^p}$.

Théorème 5.8. *Soient $p, p' \in [1, \infty]$ des exposants conjugués et soient $f, g : X \to \mathbb{R}$ mesurables.*

(i) Inégalité de Hölder :

$$\int |fg| d\mu \leq \|f\|_{\mathcal{L}^p} \|g\|_{\mathcal{L}^{p'}}.$$

En particulier, si $f \in \mathcal{L}^p(X)$ et $g \in \mathcal{L}^{p'}(X)$, alors $fg \in \mathcal{L}^1(X)$.

(ii) Inégalité de Minkowski :

$$\|f + g\|_{\mathcal{L}^p} \leq \|f\|_{\mathcal{L}^p} + \|g\|_{\mathcal{L}^p}$$

Démonstration de (i). Si $p = \infty$ (ou $p' = \infty$), alors le résultat est trivial : d'une part, comme $|fg| \leq |g| \operatorname{supess} |f|$ p.p., on obtient

$$\int |fg| d\mu \leq \operatorname{supess} |f| \int |g| d\mu = \|f\|_{\mathcal{L}^\infty} \|g\|_{\mathcal{L}^1}.$$

On peut donc supposer dorénavant $1 < p, p' < \infty$. Notons $A = \|f\|_{\mathcal{L}^p}$ et $B = \|g\|_{\mathcal{L}^{p'}}$. Si $A = 0$ ou $B = 0$, alors on a $f = 0$ p.p. ou $g = 0$ p.p., donc $\int |fg| d\mu = 0$ et le résultat est trivial. Par ailleurs, si $A = \infty$ et $B > 0$, on a $AB = \infty$ et le résultat est trivial également. On peut donc supposer également $0 < A, B < \infty$. Posons alors $F = |f|/A$ et $G = |g|/B$. L'inégalité de Young donne

$$FG \leq \frac{F^p}{p} + \frac{G^{p'}}{p'},$$

et donc

$$\int FG d\mu \leq \frac{1}{p} \int F^p d\mu + \frac{1}{p'} \int G^{p'} d\mu = \frac{1}{pA^p} \int |f|^p d\mu + \frac{1}{p' B^{p'}} \int |g|^{p'} d\mu = \frac{1}{p} + \frac{1}{q} = 1,$$

ce qui donne l'inégalité recherchée. □

Démonstration de (ii). Pour $p = \infty$, le résultat découle du Lemme 5.4,

$$\|f + g\|_{\mathcal{L}^\infty} = \operatorname{supess} |f + g| \leq \operatorname{supess}(|f| + |g|) \leq \operatorname{supess} |f| + \operatorname{supess} |g| = \|f\|_{\mathcal{L}^\infty} + \|g\|_{\mathcal{L}^\infty}.$$

Pour $p = 1$, puisque $|f + g| \leq |f| + |g|$, on a directement

$$\|f + g\|_{\mathcal{L}^1} = \int |f + g| d\mu \leq \int |f| d\mu + \int |g| d\mu = \|f\|_1 + \|g\|_1.$$

Considérons à présent $1 < p, p' < \infty$. On peut supposer $\|f\|_{\mathcal{L}^p}, \|g\|_{\mathcal{L}^p} < \infty$, ce qui implique déjà $\|f + g\|_{\mathcal{L}^p} < \infty$ par le Lemme 5.2. En décomposant $|f + g|^p \leq |f||f + g|^{p-1} + |g||f + g|^{p-1}$ et en appliquant l'inégalité de Hölder, on obtient

$$\begin{aligned}
\int |f + g|^p d\mu &\leq \int |f||f + g|^{p-1} d\mu + \int |g||f + g|^{p-1} d\mu \\
&\leq \|f\|_{\mathcal{L}^p} \Big(\int |f + g|^{p'(p-1)} d\mu \Big)^{\frac{1}{p'}} + \|g\|_{\mathcal{L}^p} \Big(\int |f + g|^{p'(p-1)} d\mu \Big)^{\frac{1}{p'}} \\
&= \Big(\|f\|_{\mathcal{L}^p} + \|g\|_{\mathcal{L}^p} \Big) \|f + g\|_{\mathcal{L}^p}^{1 - \frac{1}{p}}.
\end{aligned}$$

En divisant par $\|f + g\|_{\mathcal{L}^p}^{1 - \frac{1}{p}} < \infty$, on obtient l'inégalité voulue. □

5.3 Espaces L^p et L^∞

Avec l'inégalité de Minkowski, on a montré que $\|\cdot\|_{\mathcal{L}^p}$ est une semi-norme sur l'espace $\mathcal{L}^p(X, \mathcal{A}, \mu)$, c'est-à-dire

— $\|f\|_{\mathcal{L}^p} \geq 0$ pour tout $f \in \mathcal{L}^p$;

— $\|\lambda f\|_{\mathcal{L}^p} = |\lambda| \|f\|_{\mathcal{L}^p}$ pour tous $f \in \mathcal{L}^p$ et $\lambda \in \mathbb{R}$;

— $\|f + g\|_{\mathcal{L}^p} \leq \|f\|_{\mathcal{L}^p} + \|g\|_{\mathcal{L}^p}$ pour tous $f, g \in \mathcal{L}^p$.

Cependant, ce n'est pas une norme en général comme il manque la propriété de séparation : la condition $\|f\|_{\mathcal{L}^p} = 0$ est équivalente à $f = 0$ p.p., ce qui n'implique bien sûr pas que f est la fonction nulle. Pour y remédier, on se tourne naturellement vers le quotient de $\mathcal{L}^p$ par rapport à l'égalité presque partout.

Définition 5.9. Notons $\mathcal{N}(X, \mathcal{A}, \mu)$ l'ensemble des fonctions mesurables $f : X \to \mathbb{R}$ telles que $f = 0$ p.p. Pour $p \in [1, \infty]$, on remarque que $\mathcal{N}(X, \mathcal{A}, \mu)$ est un sous-espace vectoriel de $\mathcal{L}^p(X, \mathcal{A}, \mu)$, et on peut alors définir l'espace quotient

$$L^p(X, \mathcal{A}, \mu) = \mathcal{L}^p(X, \mathcal{A}, \mu)/\mathcal{N}(X, \mathcal{A}, \mu).$$

En d'autres termes, si on note $\sim$ la relation d'équivalence sur $\mathcal{L}^p(X, \mathcal{A}, \mu)$ définie par $f \sim g$ si $f - g \in \mathcal{N}(X, \mathcal{A}, \mu)$, c'est-à-dire si $f = g$ p.p., alors on définit $L^p(X, \mathcal{A}, \mu) = \mathcal{L}^p(X, \mathcal{A}, \mu)/\sim$, qui est donc l'espace des classes d'équivalences d'éléments de $\mathcal{L}^p(X, \mathcal{A}, \mu)$ égaux presque partout.

Remarquons que, si $f, g \in \mathcal{L}^p(X, \mathcal{A}, \mu)$ satisfont $f \sim g$, alors on a $\|f\|_{\mathcal{L}^p} = \|g\|_{\mathcal{L}^p}$. La définition suivante a donc du sens.

Définition 5.10. Pour $p \in [1, \infty]$, on définit

$$\|\cdot\|_{L^p} : L^p(X, \mathcal{A}, \mu) \to [0, \infty) : [f] \mapsto \|f\|_p,$$

où $[f]$ désigne la classe d'équivalence qui contient f.

Grâce au passage au quotient, on peut à présent conclure que $\|\cdot\|_{L^p}$ est une norme sur $L^p(X, \mathcal{A}, \mu)$, qui devient ainsi un espace vectoriel normé.

Corollaire 5.11. *$\|\cdot\|_{L^p}$ est une norme sur l'espace vectoriel $L^p(X, \mathcal{A}, \mu)$.*

Démonstration. Comme $\|\cdot\|_{\mathcal{L}^p}$ est une semi-norme sur $\mathcal{L}^p$, on déduit que $\|\cdot\|_{L^p}$ est également une semi-norme sur L^p. Enfin, si $[f] \in L^p$ satisfait $\|[f]\|_{L^p} = 0$, cela signifie $\|f\|_{\mathcal{L}^p} = 0$, et donc $f = 0$ p.p., c'est-à-dire $[f] = [0]$. □

Pour alléger les notations, on identifiera souvent la classe d'équivalence $[f]$ et son représentant f. Cependant, on insiste sur le tournant conceptuel qui a lieu ici : dans la théorie L^p, la valeur d'une fonction en un point n'aura plus de sens — on parlera désormais seulement de classes d'équivalences de fonctions égales presque partout.

Comme l'espace $L^p(X, \mathcal{A}, \mu)$ est un espace vectoriel normé, la norme $\|\cdot\|_{L^p}$ munit cet espace d'une topologie, associée via la métrique

$$d_p([f], [g]) = \|[f - g]\|_{L^p}.$$

Comme dans tout espace métrique, on a donc sur $L^p(X, \mathcal{A}, \mu)$ les propriétés très utiles suivantes :

— toute suite convergente est une suite de Cauchy ;

— toute suite convergente possède une limite unique.

Dans les sections suivantes, on étudiera d'autres propriétés topologiques des espaces L^p ; on verra notamment qu'ils sont complets. Notons que la notion de convergence sur L^p se traduit en une notion de convergence sur $\mathcal{L}^p$.

Définition 5.12. Soient $1 \leq p \leq \infty$, une suite $(f_n)_n \subset \mathcal{L}^p(X, \mathcal{A}, \mu)$, et $f \in \mathcal{L}^p(X, \mathcal{A}, \mu)$. On dit que f_n converge vers f dans L^p (ce qu'on note $f_n \to f$ en L^p, ou $f_n \xrightarrow{L^p} f$) si la suite de classes d'équivalences $([f_n])_n$ converge vers $[f]$ dans L^p, c'est-à-dire si $\|f_n - f\|_{\mathcal{L}^p} \to 0$.

5.4 Complétude des espaces L^p

Le résultat suivant montre que les espaces $L^p(X, \mathcal{A}, \mu)$ sont complets, c'est-à-dire que toute suite de Cauchy dans $L^p(X, \mathcal{A}, \mu)$ converge dans cet espace. Ces espaces sont donc des espaces vectoriels normés complets, ce qu'on appelle des *espaces de Banach*.

Théorème 5.13 (Théorème de Riesz–Fisher)**.** *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $p \in [1, \infty]$. Muni de la norme $\|\cdot\|_p$, l'espace $L^p(X, \mathcal{A}, \mu)$ est un espace vectoriel normé complet.*

Démonstration. Il suffit de montrer le résultat de complétude au niveau des représentants de classes d'équivalence. Considérons une suite $(f_n)_n \subset \mathcal{L}^p(X, \mathcal{A}, \mu)$ telle que pour tout $\varepsilon > 0$ il existe N tel que $\|f_n - f_m\|_{\mathcal{L}^p} \leq \varepsilon$ pour tous $n, m \geq N$. On va montrer qu'il existe $f \in \mathcal{L}^p(X, \mathcal{A}, \mu)$ tel que $f_n \to f$ en L^p. On considère séparément les cas $p < \infty$ et $p = \infty$.

Cas 1 : $p < \infty$.
Choisissons $(n_j)_j \subset \mathbb{N}$ une suite croissante telle que $\|f_{n_{j+1}} - f_{n_j}\|_{\mathcal{L}^p} \leq 2^{-j}$ pour tout j. On va montrer que la sous-suite $(f_{n_j})_j$ converge en L^p. Pour ce faire, posons

$$g_k = \sum_{j=1}^{k} |f_{n_{j+1}} - f_{n_j}| \quad \text{et} \quad g = \sum_{j=1}^{\infty} |f_{n_{j+1}} - f_{n_j}|.$$

En utilisant l'inégalité de Minkowski, on a

$$\|g_k\|_{\mathcal{L}^p} \leq \sum_{j=1}^{k} \|f_{n_{j+1}} - f_{n_j}\|_{\mathcal{L}^p} \leq \sum_{j=1}^{k} 2^{-j} \leq 1.$$

Par le Lemme de Fatou, on a

$$\int \Big(\liminf_k |g_k|^p\Big) d\mu \leq \liminf_k \int |g_k|^p d\mu \leq 1,$$

et donc $g \in L^p(X, \mathcal{A}, \mu)$. En particulier, g est fini μ-presque partout. Soit N un ensemble de mesure nulle contenant l'ensemble des points où g est infini, et posons

$$f(x) = \begin{cases} f_{n_1}(x) + \sum_{j=1}^{\infty} \big(f_{n_{j+1}}(x) - f_{n_j}(x)\big) & \text{si } x \notin N \\ 0 & \text{si } x \in N. \end{cases}$$

Puisque $f_{n_k} = f_{n_1} + \sum_{j=1}^{k-1} \big(f_{n_{j+1}} - f_{n_j}\big)$, la sous-suite $(f_{n_j})_j$ converge presque partout vers f. Montrons que f est également la limite en L^p de la suite $(f_n)_n$. Soient $\varepsilon > 0$ et $N \geq 1$ tels que $\|f_n - f_m\|_{\mathcal{L}^p} \leq \varepsilon$ pour tout $n, m \geq N$. Par le lemme de Fatou, pour tout $m \geq N$, on a alors

$$\int |f_m - f|^p d\mu = \int \liminf_j |f_n - f_{n_j}|^p d\mu \leq \liminf_j \int |f_n - f_{n_j}|^p d\mu \leq \varepsilon^p.$$

De plus, on en tire également que $f \in \mathcal{L}^p(X, \mathcal{A}, \mu)$ puisque $f = (f - f_m) + f_m$ qui est la somme de deux éléments de $\mathcal{L}^p(X, \mathcal{A}, \mu)$.

Cas 2 : $p = \infty$.
Posons

$$A_n = \{x \in X : |f_n(x)| > \|f_n\|_\infty\}$$

et

$$B_{m,n} = \{x \in X : |f_n(x) - f_m(x)| > \|f_n - f_m\|_\infty\}.$$

Par hypothèse, l'ensemble

$$E = \bigcup_n A_n \cup \bigcup_{m,n} B_{m,n}.$$

est μ-négligeable. En dehors de E, la suite $(f_n)_n$ est uniformément de Cauchy et converge donc vers une fonction bornée f. Posons $f(x) = 0$ pour tout $x \in E$. Alors $f \in \mathcal{L}^\infty(X, \mathcal{A}, \mu)$ est la limite de $(f_n)_n$ dans L^∞. □

La complétude est une propriété particulièrement utile en pratique : quand on cherche à résoudre une équation sur un espace complet, il suffit de trouver une suite dans l'espace qui se rapproche de la solution pour conclure que la limite est également dans le même espace. En particulier, on rappelle le théorème de point fixe suivant.

Théorème 5.14 (Théorème du point fixe de Banach / Picard-Lindelöf)**.** *Soit (X, d) un espace métrique complet, et soit $T : X \to X$ une fonction contractante, c'est-à-dire telle que $d(T(x), T(y)) \leq cd(x, y)$ pour tous $x, y \in X$, pour une constante $c < 1$. Alors il existe un unique $x_* \in X$ tel que $T(x_*) = x_*$. De plus, pour tout $x_0 \in X$, on a $T^n(x_0) \to x_*$ pour $n \uparrow \infty$.*

Exemple 5.15. Comme exemple d'application de la complétude, considérons une équation intégrale de Fredholm du 2e type : sur un espace de mesure $(X, \mathcal{A}, \mu)$, on cherche à trouver une solution $F \in L^p(X)$ telle que

$$F(x) - \int_X K(x, y)F(y)d\mu(y) = G(x), \tag{5.1}$$

où sont données $G \in L^p(X)$, $1 \leq p < \infty$, $K \in L^\infty(X \times X)$, et où l'on suppose

$$\mu(X)\|K\|_{L^\infty} < \infty.$$

Tout d'abord, il s'agit de vérifier que le problème a bien du sens. Étant donné des représentants $f \in F$ et $k \in K$, avec $f \in \mathcal{L}^p(X), k \in \mathcal{L}^\infty(X \times X)$, comme par hypothèse $\mu(X) < \infty$, on vérifie aisément que $(x, y) \mapsto k(x, y)f(y)$ est $\mu \otimes \mu$-intégrable. Par Fubini, on déduit que $\int_X k(x, y)f(y)d\mu(y)$ est bien défini pour presque tout x, disons pour tout $x \notin N$ avec $N \in \mathcal{A}$ et $\mu(N) = 0$. Tous les éléments de l'équation (5.1) ont donc bien du sens presque partout au niveau des représentants. Notons alors

$$sf(x) := \begin{cases} \int_X k(x, y)f(y)d\mu(y) & : \quad x \notin N, \\ 0 & : \quad x \in N, \end{cases}$$

qui définit $sf \in \mathcal{L}^1(X)$ par Fubini. Comme $|k| \leq \|k\|_{\mathcal{L}^\infty}$ p.p., on a $|k(x, \cdot)| \leq \|k\|_{\mathcal{L}^\infty}$ p.p. pour presque tout x, et donc

$$|sf(x)| \leq \|k\|_{\mathcal{L}^\infty}\|f\|_{\mathcal{L}^1} \leq \mu(X)^{\frac{1}{p'}}\|k\|_{\mathcal{L}^\infty}\|f\|_{\mathcal{L}^p},$$

où la dernière inégalité découle de Hölder. En prenant la semi-norme $\mathcal{L}^p$, on obtient ainsi

$$\|sf\|_{\mathcal{L}^p} \leq \mu(X)\|k\|_{\mathcal{L}^\infty}\|f\|_{\mathcal{L}^p}, \tag{5.2}$$

ce qui montre que $s : f \mapsto sf$ est bien définie comme une application $\mathcal{L}^p(X) \to \mathcal{L}^p(X)$. Afin de pouvoir utiliser le bon cadre fonctionnel de L^p, on doit à présent repasser aux classes d'équivalence de fonctions. Si $f' \in F$ et $k' \in K$ sont d'autres représentants, on a $f' = f$ p.p. et $k' = k$ p.p. Si l'on note $s'f'$ comme ci-dessus avec f, k remplacés par f', k', on déduit aisément $sf = s'f'$ p.p. (exercice). Ceci permet donc de définir une application $S : L^p(X) \to L^p(X)$ par $SF := [sf]$, et l'inégalité (5.2) devient

$$\|SF\|_{L^p} \le \mu(X)\|K\|_{L^\infty}\|F\|_{L^p}. \tag{5.3}$$

Comme $\mu(X)\|K\|_{L^\infty} < 1$, ceci montre que l'application $L^p(X) \to L^p(X) : F \mapsto SF + G$ est une contraction. Dès lors, par le théorème de point fixe, comme $L^p(X)$ est un espace de Banach, on déduit qu'il existe un unique $F_* \in L^p(X)$ tel que $F_* = SF_* + G$ — c'est l'unique solution de notre équation (5.1).

Remarque 5.16. Dans le cas $p = 2$, sur un espace de mesure $(X, \mathcal{A}, \mu)$, on remarque que la norme sur $L^2(X)$ s'écrit $\|[f]\|^2_{L^2(X)} = \langle [f], [f]\rangle_{L^2(X)}$, où l'on a défini

$$\langle [f], [g]\rangle_{L^2(X)} = \int fg d\mu.$$

On vérifie que ceci définit un produit scalaire sur $L^2(X)$. Cet espace est donc d'un espace de Banach dont la norme est associée à un produit scalaire — c'est ce qu'on appelle un espace de Hilbert. Ce cadre fonctionnel sera particulièrement pratique comme la présence d'un produit scalaire permet d'utiliser un nombre d'outils géométriques (projections orthogonales, etc.).

5.5 Approximation de fonctions L^p

On a déjà vu deux propriétés de densité de l'espace $\mathcal{L}^1(X)$: pour tout $f \in \mathcal{L}^1(X)$,

— il existe une suite $(f_n)_n$ de fonctions simples mesurables telles que $\|f - f_n\|_{\mathcal{L}^1} \to 0$;

— si $(X, \mathcal{A}, \mu) = (\mathbb{R}^d, \mathcal{M}_d, \mathcal{L}_d)$, il existe une suite $(f_n)_n \subset C_c(\mathbb{R}^d)$ telle que $\|f - f_n\|_{\mathcal{L}^1} \to 0$.

On va montrer que les même propriétés sont valables dans $\mathcal{L}^p$ si $p < \infty$. On commence par la question d'approximation par des fonctions simples. Ce résultat-ci est vrai également pour $p = \infty$, sauf que dans ce cas les fonctions simples ne peuvent en général pas être prises nulles hors d'un ensemble de mesure finie.

Proposition 5.17. *Soit $(X, \mathcal{A}, \mu)$ un espace de mesure.*

(i) Soit $1 \le p < \infty$. Pour tout $f \in \mathcal{L}^p(X)$, il existe une suite $(f_n)_n$ de fonctions simples, mesurables, nulles hors d'un ensemble de mesure finie, telles que $\|f - f_n\|_{\mathcal{L}^p} \to 0$. En particulier, l'ensemble

$$\big\{[f] : f \text{ simple, mesurable } X \to \mathbb{R}, \text{ nulle hors d'un ensemble de mesure finie}\big\}$$

est un sous-espace vectoriel dense dans $L^p(X)$.

(ii) Soit $p = \infty$. Pour tout $f \in \mathcal{L}^\infty(X)$, il existe une suite $(f_n)_n$ de fonctions simples mesurables telles que $\|f - f_n\|_{\mathcal{L}^\infty} \to 0$. En particulier, l'ensemble

$$\big\{[f] : f \text{ simple, mesurable } X \to \mathbb{R}\big\}$$

est un sous-espace vectoriel dense dans $L^\infty(X)$.

Démonstration. (i) Soit $1 \le p < \infty$ et $f \in \mathcal{L}^p(X)$. Quitte à considérer les parties positives et négatives, il suffit de considérer le cas $f \ge 0$. Par la Proposition 1.29, il existe une suite croissante $(f_n)_n$ de fonctions simples positives mesurables telles que $f_n \uparrow f$. Comme la suite est croissante, on a $|f - f_n|^p \le |f|^p$ et le théorème de la convergence dominée donne $\int |f - f_n|^p d\mu \to 0$. Comme f_n est une somme finie d'indicatrices et que $\int |f_n|^p d\mu \le \int |f|^p d\mu < \infty$, il est clair que f_n est nulle hors d'un ensemble de mesure finie.

(ii) Soit $f \in \mathcal{L}^\infty(X)$. Soit $\varepsilon > 0$. Choisissons une partition de l'intervalle $[-\|f\|_{\mathcal{L}^\infty}, \|f\|_{\mathcal{L}^\infty}]$, soit $y_0 := -\|f\|_{\mathcal{L}^\infty} < \ldots < y_N := \|f\|_{\mathcal{L}^\infty}$, telle que $y_n - y_{n-1} \le \varepsilon$ pour tout n. Posons $A_n = f^{-1}((y_{n-1}, y_n])$ et considérons

$$f_\varepsilon = \sum_{n=1}^{N} y_n \mathbb{1}_{A_n}.$$

Alors f_ε est une fonction simple mesurable telle que $\|f - f_\varepsilon\|_\infty \le \varepsilon$. En effet, si $x \in X$ est tel que $f(x) \in [-\|f\|_\infty, \|f\|_\infty]$, alors il existe un n tel que $x \in A_n$ et donc

$$|f(x) - f_\varepsilon(x)| = |f(x) - y_n| \le \varepsilon.$$

Comme l'ensemble $\{x : x \notin [-\|f\|_\infty, \|f\|_\infty]\}$ est μ-négligeable, on conclut $\|f - f_\varepsilon\|_{\mathcal{L}^\infty} \le \varepsilon$. $\square$

On se tourne à présent vers le cas $(X, \mathcal{A}, \mu) = (\mathbb{R}^d, \mathcal{M}_d, \mathcal{L}_d)$, et on montre que, pour $1 \le p < \infty$, tout élément de $\mathcal{L}^p(\mathbb{R}^d)$ peut être approché par des fonctions continues à support compact.

Proposition 5.18. *Soit $1 \le p < \infty$. Pour $f \in \mathcal{L}^p(\mathbb{R}^d)$, il existe une suite $(f_n)_n \subset C_c(\mathbb{R}^d)$ telle que $\|f - f_n\|_{\mathcal{L}^p} \to 0$. En particulier, $C_c(\mathbb{R}^d)/\mathcal{N}(\mathbb{R}^d)$ est dense dans $L^p(\mathbb{R}^d)$.*

Démonstration. Par la proposition précédente, il suffit de montrer le résultat pour une fonction f simple, mesurable, nulle hors d'un ensemble de mesure finie. Par linéarité, on peut alors supposer $f = \mathbb{1}_A$ avec $A \in \mathbb{B}_d$ et $\mathcal{L}_d(A) < \infty$. Soit $\varepsilon > 0$. Par régularité de la mesure de Lebesgue, il existe $K_\varepsilon \subset A \subset U_\varepsilon$ avec K_ε compact et U_ε ouvert tels que $\mathcal{L}_d(U_\varepsilon \setminus K_\varepsilon) \le \varepsilon^p$. De plus, par le Lemme d'Urysohn, on a vu qu'on peut construire une fonction cut-off $g_\varepsilon \in C_c(\mathbb{R}^d)$ avec $g_\varepsilon|_{K_\varepsilon} = 1$, $g_\varepsilon|_{U_\varepsilon^c} = 0$, et $0 \le g_\varepsilon \le 1$. On obtient alors

$$\int |\mathbb{1}_A - g_\varepsilon|^p = \int_{U_\varepsilon \setminus K_\varepsilon} |\mathbb{1}_A - g_\varepsilon|^p \le 2^p \mathcal{L}_d(U_\varepsilon \setminus K_\varepsilon) \le 2^p \varepsilon^p,$$

c'est-à-dire $\|\mathbb{1}_A - g_\varepsilon\|_{\mathcal{L}^p} \le 2\varepsilon$. $\square$

Remarque 5.19. Le résultat précédent est faux pour $p = \infty$!
Pour $f \in \mathcal{L}^\infty(\mathbb{R}^d)$, supposons qu'il existe une suite $(f_n)_n \subset C_c(\mathbb{R}^d)$ telle que $f_n \to f$ en L^∞. En particulier, on déduit que c'est une suite de Cauchy en $\mathcal{L}^\infty$: pour tout ε, il existe N tel que $\|f_n - f_m\|_{\mathcal{L}^\infty} \le \varepsilon$ pour tous $n, m \ge N$. Or, comme $f_n, f_m \in C_c(\mathbb{R}^d)$, on observe que $\|f_n - f_m\|_{\mathcal{L}^\infty} = \sup |f_n - f_m|$. Ceci montre donc que la suite $(f_n)_n$ est en fait de Cauchy uniforme. Dès lors, il existe $g \in C(\mathbb{R}^d)$ tel que $f_n \to g$ uniformément. En particulier, on déduit $f = g$ p.p. avec $g \in C(\mathbb{R}^d)$. Cependant, ceci n'est pas possible pour tout $f \in \mathcal{L}^\infty(\mathbb{R}^d)$: par exemple, on vérifie que la fonction $f = \mathbb{1}_{[0,1]}$ n'est pas égale presque partout à une fonction continue.

5.6 Régularisation par convolution

On va voir que dans le résultat d'approximation précédent les fonctions continues à support compact peuvent être remplacées par des fonctions lisses à support compact. Pour ce faire, on utilise le procédé très utile de régularisation par convolution.

Lemme 5.20.

(i) Pour $f, g \in \mathcal{L}^1(\mathbb{R}^d)$, la fonction $y \mapsto f(x-y)g(y)$ est intégrable pour presque tout x, disons pour tout $x \notin N$ avec $N \in \mathcal{M}_d$ et $\mathcal{L}_d(N) = 0$. On peut alors définir le produit de convolution

$$f * g(x) := \begin{cases} \int_{\mathbb{R}^d} f(x-y)g(y)dy & : \quad x \notin N, \\ 0 & : \quad x \in N, \end{cases}$$

et l'on a $f * g \in \mathcal{L}^1(\mathbb{R}^d)$.

(ii) On a les propriétés suivantes, qui montrent que $(L^1(\mathbb{R}^d), *)$ est ce qu'on appelle une algèbre de Banach : pour tous $f, g, h \in \mathcal{L}^1(\mathbb{R}^d)$,

$$f * g = g * f \text{ p.p.}, \quad f * (g * h) = (f * g) * h \text{ p.p.}, \quad \|f * g\|_{\mathcal{L}^1} \leq \|f\|_{\mathcal{L}^1} \|g\|_{\mathcal{L}^1}.$$

(iii) Si $f \in \mathcal{L}^1(\mathbb{R}^d)$ et $g \in \mathcal{L}^p(\mathbb{R}^d)$ avec $1 \leq p \leq \infty$, alors $f * g$ est également bien défini p.p. par l'intégrale du point (i). On a alors $f * g \in \mathcal{L}^p(\mathbb{R}^d)$ et

$$\|f * g\|_{\mathcal{L}^p} \leq \|f\|_{\mathcal{L}^1} \|g\|_{\mathcal{L}^p}.$$

Démonstration. (i) Pour $f, g \in \mathcal{L}^1(\mathbb{R}^d)$, on vérifie que $(x, y) \mapsto f(x-y)g(y)$ est mesurable, et le théorème de Fubini donne

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} |f(x-y)g(y)| dx dy = \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} |f(x-y)g(y)| dx \Big) dy = \|f\|_{\mathcal{L}^1} \|g\|_{\mathcal{L}^1} < \infty.$$

Le reste de l'énoncé découle du théorème de Fubini.

(iii) Focalisons-nous sur le cas $p < \infty$. Pour $f \in \mathcal{L}^1(\mathbb{R}^d)$ et $g \in \mathcal{L}^p(\mathbb{R}^d)$, comme la fonction $(x, y) \mapsto f(x-y)g(y)$ est mesurable, on déduit par Fubini que l'intégrale partielle $x \mapsto \int_{\mathbb{R}^d} |f(x-y)g(y)| dy$ est bien définie presque partout et définit une fonction mesurable. On peut alors calculer sa semi-norme $\mathcal{L}^p$: par l'inégalité de Hölder et le théorème de Fubini, on obtient

$$\int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} |f(x-y)g(y)| dy \Big)^p dx \leq \|f\|_{\mathcal{L}^1}^{\frac{p}{p'}} \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} |f(x-y)| |g(y)|^p dy \Big) dx = \|f\|_{\mathcal{L}^1}^{1+\frac{p}{p'}} \|g\|_{\mathcal{L}^p}^p.$$

Comme $1 + \frac{p}{p'} = p$, la conclusion s'ensuit. □

Remarque 5.21. En plus du point (iii) ci-dessus, l'inégalité de Young pour la convolution affirme que pour tous $1 \leq p, q, r \leq \infty$ avec $\frac{1}{r} + 1 = \frac{1}{p} + \frac{1}{q}$ on a

$$f \in \mathcal{L}^p(\mathbb{R}^d), \ g \in \mathcal{L}^q(\mathbb{R}^d) \quad \Longrightarrow \quad f * g \in \mathcal{L}^r(\mathbb{R}^d) \ \text{ et } \ \|f * g\|_{\mathcal{L}^r} \leq \|f\|_{\mathcal{L}^p} \|g\|_{\mathcal{L}^q}.$$

On montre à présent que le produit de convolution hérite de la meilleure régularité de ses deux facteurs. C'est cette propriété qui permettra d'utiliser le produit de convolution pour régulariser des fonctions peu régulières.

Lemme 5.22. *Si $f \in \mathcal{L}^1(\mathbb{R}^d)$ et $g \in C_0^m(\mathbb{R}^d)$ avec $m \geq 0$, alors $f * g \in C_0^m(\mathbb{R}^d)$ et $\nabla^\alpha (f * g) = f * \nabla^\alpha g$ pour tout multi-indice $|\alpha| \leq m$.*

Démonstration. Soient $f \in \mathcal{L}^1(\mathbb{R}^d)$ et $g \in C_0^m(\mathbb{R}^d)$. Comme $g \in \mathcal{L}^\infty(\mathbb{R}^d)$, le lemme précédent assure déjà que $f * g$ est bien défini p.p. et appartient à $\mathcal{L}^\infty(\mathbb{R}^d)$. Il reste à voir que $f * g \in C_0^m(\mathbb{R}^d)$. Considérons d'abord le cas $m = 0$. Soient $x_0 \in \mathbb{R}^d$ et $(h_n)_n \subset \mathbb{R}^d$ avec $h_n \to 0$. Par changement de variable dans l'intégrale, on a

$$f * g(x_0 + h_n) = \int f(x_0 + h_n - y) g(y) dy = \int f(y) g(x_0 + h_n - y) dy.$$

La continuité de g assure que $f(y)g(x_0+h_n-y) \to f(y)g(x_0-y)$ pour tout y. De plus, on a $|f(y)g(x_0+h_n-y)| \le |f(y)| \sup|g|$. Comme $f \in \mathcal{L}^1(\mathbb{R}^d)$, on peut appliquer le théorème de convergence dominée pour conclure $f*g(x_0+h_n) \to \int f(y)g(x_0-y)dy = f*g(x_0)$. De plus, pour $(x_n)_n \subset \mathbb{R}^d$ avec $|x_n| \to \infty$, on vérifie de même $f*g(x_n) \to 0$. Ceci montre donc $f*g \in C_0^0(\mathbb{R}^d)$. Pour les dérivées supérieures, on argumente de façon analogue en examinant les quotients différentiels. □

On montre à présent comment choisir un noyau de convolution qui approche l'identité. On commence par le lemme suivant, qui montre la continuité des translations sur $\mathcal{L}^p$. On remarque à nouveau que le résultat est faux pour $p=\infty$ (par exemple pour $f=\mathbb{1}_{[0,1]}$).

Lemme 5.23. *Notons $T_z f(x) = f(x-z)$ l'opérateur de translation sur $\mathbb{R}^d$. Pour $1 \le p < \infty$ et $f \in \mathcal{L}^p(\mathbb{R}^d)$, on a $\|T_z f - f\|_{\mathrm{L}^p} \to 0$ pour $|z| \to 0$.*

Démonstration. On a vu que pour tout $\varepsilon > 0$ il existe $f_\varepsilon \in C_c(\mathbb{R}^d)$ tel que $\|f - f_\varepsilon\|_{\mathcal{L}^p} \le \varepsilon$. Notons $K_\varepsilon := \operatorname{supp}(f_\varepsilon)$ compact. On a alors $\|T_z f_\varepsilon - f_\varepsilon\|_{\mathcal{L}^p}^p = \int |f_\varepsilon(x) - f_\varepsilon(x-z)|^p dx$. Pour $|z| \le 1$, en utilisant la continuité de f et la borne $|f_\varepsilon(x) - f_\varepsilon(x-z)|^p \le 2^p \mathbb{1}_{K_\varepsilon+B}(x) \sup|f_\varepsilon|^p$, on peut appliquer le théorème de convergence dominée pour déduire $\|T_z f_\varepsilon - f_\varepsilon\|_{\mathcal{L}^p} \to 0$ pour $|z| \to 0$. On a donc montré $\limsup_{|z|\to 0} \|T_z f - f\|_{\mathcal{L}^p} \le 2\varepsilon$, et la conclusion s'ensuit. □

Corollaire 5.24. *Soit $\phi \in C_c^\infty(\mathbb{R}^d)$ avec $\phi \ge 0$ et $\int \phi = 1$, et notons $\phi_\varepsilon(x) = \varepsilon^{-d}\phi(x/\varepsilon)$. Alors pour tout $1 \le p < \infty$ et tout $f \in \mathcal{L}^p(\mathbb{R}^d)$ on a $\phi_\varepsilon * f \in C_0^\infty(\mathbb{R}^d) \cap \mathcal{L}^p(\mathbb{R}^d)$ et $\phi_\varepsilon * f \to f$ en L^p. On dit alors que $\{\phi_\varepsilon *\}_\varepsilon$ est une approximation de l'identité.*

Démonstration. On remarque que $\phi_\varepsilon \in C_c^\infty(\mathbb{R}^d)$ satisfait $\int \phi_\varepsilon = \int \phi = 1$ et $\operatorname{supp}\phi_\varepsilon = \varepsilon \operatorname{supp}\phi$. On a $\phi_\varepsilon * f \in \mathcal{L}^p(\mathbb{R}^d)$ par le Lemme 5.20, et $\phi_\varepsilon * f \in C_0^\infty(\mathbb{R}^d)$ par le Lemme 5.22. Par changement de variable, on peut écrire $\phi_\varepsilon * g(x) - g(x) = \int \phi(y)(g(x-\varepsilon y) - g(x))dy$, et donc par l'inégalité triangulaire et l'inégalité de Hölder,

$$\|\phi_\varepsilon * g - g\|_{\mathcal{L}^p}^p \le \int \Big(\int |\phi(y)||g(x-\varepsilon y) - g(x)|dy \Big)^p dx \le \|\phi\|_{\mathcal{L}^1}^{p-1} \int |\phi(y)| \|T_{\varepsilon y} g - g\|_{\mathcal{L}^p} dy.$$

Par le lemme précédent et le théorème de convergence dominée, on déduit que le membre de droite converge vers 0. □

Théorème 5.25. *Soit $1 \le p < \infty$. Pour $f \in \mathcal{L}^p(\mathbb{R}^d)$, il existe une suite $(f_n)_n \subset C_c^\infty(\mathbb{R}^d)$ telle que $\|f - f_n\|_{\mathcal{L}^p} \to 0$. En particulier, $C_c^\infty(\mathbb{R}^d)/\mathcal{N}(\mathbb{R}^d)$ est dense dans $L^p(\mathbb{R}^n)$.*

Démonstration. Par le résultat précédent, il reste à remplacer $\varphi_\varepsilon * g \in C_0^\infty(\mathbb{R}^d) \cap \mathcal{L}^p(\mathbb{R}^d)$ par un élément à support compact. Soit $f \in \mathcal{L}^p(\mathbb{R}^d)$ et soit $\varepsilon > 0$. On commence par remarquer que $\|f - f\mathbb{1}_{B_R}\|_{\mathcal{L}^p} \to 0$ pour $R \uparrow \infty$ en conséquence du théorème de convergence dominée. Il existe donc $R_\varepsilon > 0$ tel que $\|f - f\mathbb{1}_{B_{R_\varepsilon}}\|_{\mathcal{L}^p} \le \varepsilon$. Par le résultat précédent, on a $\phi_\delta * (f\mathbb{1}_{B_{R_\varepsilon}}) \to f\mathbb{1}_{B_{R_\varepsilon}}$ pour $\delta \downarrow 0$. Il existe donc $\delta_\varepsilon > 0$ tel que $\|\phi_{\delta_\varepsilon} * (f\mathbb{1}_{B_{R_\varepsilon}}) - f\mathbb{1}_{B_{R_\varepsilon}}\|_{\mathcal{L}^p} \le \varepsilon$. Pour $f_\varepsilon := \phi_{\delta_\varepsilon} * (f\mathbb{1}_{B_{R_\varepsilon}})$, on obtient donc $\|f - f_\varepsilon\|_{\mathcal{L}^p} \le 2\varepsilon$. □

5.7 Séparabilité des espaces $\mathcal{L}^p$

Revenons au cas d'un espace de mesure général $(X, \mathcal{A}, \mu)$. Dans cette section, on va examiner à quelle condition l'espace $L^p(X)$ est séparable. On commence par montrer le résultat d'approximation suivant, qui montre sur un espace σ-fini on peut approcher tout ensemble mesurable par des ensembles d'une partie génératrice de $\mathcal{A}$ qui est une algèbre. Rappelons qu'une collection de parties de X est une *algèbre* si elle contient X et si elle est stable par passage au complémentaire et par union finie.

Lemme 5.26. *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $\mathcal{C} \subset \mathcal{A}$ une algèbre telle que $\sigma(\mathcal{C}) = \mathcal{A}$. Supposons que μ est σ-finie sur $\mathcal{C}$. Alors pour tout $A \in \mathcal{A}$ tel que $\mu(A) < \infty$ et pour tout $\varepsilon > 0$ il existe $C \in \mathcal{C}$ tel que $\mu(A \triangle C) < \varepsilon$, où l'on définit la différence symétrique $A \triangle C = (A \cup C) \setminus (A \cap C)$.*

Démonstration. On sépare deux cas, selon que μ est finie ou σ-finie.

Cas 1 : μ finie.

Soit $\mathcal{F}$ la collection des ensembles $A \in \mathcal{A}$ tels que pour tout $\varepsilon > 0$ il existe $C \in \mathcal{C}$ avec $\mu(A \triangle C) < \varepsilon$. Bien sûr, on a $\mathcal{C} \subset \mathcal{F}$ et il suffit de montrer que $\mathcal{F}$ est une σ-algèbre pour obtenir la conclusion. Puisque $\mathcal{C} \subset \mathcal{F}$, on a $X \in \mathcal{F}$. De plus, on remarque $A^c \triangle C^c = A \triangle C$, de sorte que $\mathcal{F}$ est stable par passage au complémentaire. Soit $(A_n)_n$ une suite d'éléments de $\mathcal{F}$ et notons $A = \cup_n A_n$. Fixons $\varepsilon > 0$. La continuité de la mesure implique qu'il existe $N \geq 1$ pour lequel

$$\mu\left(A \setminus \bigcup_{n \leq N} A_n\right) < \frac{\varepsilon}{2}.$$

Pour tout $n \leq N$, choisissons $C_n \in \mathcal{C}$ tel que $\mu(A_n \triangle C_n) < \frac{\varepsilon}{2N}$. Pour $C := \bigcup_{n \leq N} C_n \in \mathcal{C}$, on a

$$\mu(A \triangle C) \leq \mu\left(A \setminus \bigcup_{n \leq N} A_n\right) + \mu\left(\bigcup_{n \leq N} A_n \triangle C\right) \leq \mu\left(A \setminus \bigcup_{n \leq N} A_n\right) + \sum_{n \leq N} \mu(A_n \triangle C_n) < \varepsilon,$$

ce qui montre $A \in \mathcal{F}$.

Cas 2 : μ σ-finie.

Soit $(A_n)_n \subset \mathcal{C}$ avec $A_n \uparrow X$ et $\mu(A_n) < \infty$ pour tout n. Soit $A \in \mathcal{A}$ tel que $\mu(A) < \infty$ et soit $\varepsilon > 0$. La continuité de la mesure donne alors l'existence d'un indice $N \geq 1$ pour lequel $\mu(A \cap A_N) > \mu(A) - \varepsilon/2$. L'application $\mathcal{A} \to [0, \infty) : B \mapsto \mu(B \cap A_N)$ défini une mesure finie, de sorte que, par le lemme précédent, il existe $C \in \mathcal{C}$ tel que $\mu((A \triangle C) \cap A_N) < \varepsilon/2$. On a alors $C \cap A_N \in \mathcal{C}$ et

$$\begin{aligned}
\mu(A \triangle (C \cap A_N)) &\leq \mu(A \setminus (A \cap A_N)) + \mu((A \cap A_N) \triangle (C \cap A_N)) \\
&= \mu(A \setminus (A \cap A_N)) + \mu((A \triangle C) \cap A_N) < \varepsilon,
\end{aligned}$$

d'où la conclusion. $\square$

On déduit à présent le critère suivant pour la séparabilité des espaces L^p. Cela s'applique en particulier à $L^p(\mathbb{R}^d)$ pour $1 \leq p < \infty$.

Théorème 5.27. *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $1 \leq p < \infty$. Si $\mathcal{C} \subset \mathcal{A}$ est une famille dénombrable telle que $\sigma(\mathcal{C}) = \mathcal{A}$ et si μ est σ-finie sur $\mathcal{C}$, alors l'espace $L^p(X)$ est séparable.*

Démonstration. Soit $\mathcal{C}^*$ la collection formée des éléments de $\mathcal{C}$, de leurs complémentaires, de leurs unions finies, de X, et de $\varnothing$. Alors $\mathcal{C}^*$ est une algèbre dénombrable qui contient $\mathcal{C}$ et bien sûr, on a toujours $\sigma(\mathcal{C}^*) = \mathcal{A}$. Notons à présent $\mathcal{U}$ la collection des fonctions simples de la forme $\sum_{j=1}^n c_j \mathbb{1}_{C_j}$ où $n \geq 1$, $c_j \in \mathbb{Q}$, et $C_j \in \mathcal{C}^*$ avec $\mu(C_j) < \infty$ pour tout j. Bien sûr, $\mathcal{U}$ est un ensemble dénombrable et $\mathcal{U} \subset \mathcal{L}^p(X)$. Montrons que cela forme un sous-ensemble dense de $\mathcal{L}^p(X)$. Par la Proposition 5.17, il suffit de montrer que cet ensemble est dense dans le sous-espace des fonctions simples de $\mathcal{L}^p(X)$. Soit $g = \sum_{j=1}^n a_j \mathbb{1}_{A_j}$ une telle fonction simple mesurable, avec $a_j \in \mathbb{R}$, $A_j \in \mathcal{A}$, et $\mu(A_j) < \infty$ pour tout j. Fixons $\varepsilon > 0$. Par densité, il existe des rationnels $c_1, \ldots, c_n \in \mathbb{Q}$ tels que

$$\left\| \sum_{j=1}^n a_j \mathbb{1}_{A_j} - \sum_{j=1}^n c_j \mathbb{1}_{A_j} \right\|_{\mathcal{L}^p} \leq \sum_{j=1}^n |a_j - c_j| \mu(A_j)^{\frac{1}{p}} < \frac{\varepsilon}{2}.$$

De plus, par le lemme précédent, il existe des ensembles $C_1, \dots, C_n \in \mathcal{C}$ tels que

$$\left\| \sum_{j=1}^n c_j \mathbb{1}_{A_j} - \sum_{j=1}^n c_j \mathbb{1}_{C_j} \right\|_{\mathcal{L}^p} \le \sum_{j=1}^n |c_j| \|\mathbb{1}_{A_j} - \mathbb{1}_{C_j}\|_{\mathcal{L}^p} = \sum_{j=1}^n |c_j| \mu(A_j \triangle C_j)^{\frac{1}{p}} < \frac{\varepsilon}{2}.$$

Alors, $\sum_{j=1}^n c_j \mathbb{1}_{C_j}$ appartient à $\mathcal{U}$ et satisfait

$$\left\| \sum_{j=1}^n a_j \mathbb{1}_{A_j} - \sum_{j=1}^n c_j \mathbb{1}_{C_j} \right\|_{\mathcal{L}^p} \le \left\| \sum_{j=1}^n a_j \mathbb{1}_{A_j} - \sum_{j=1}^n c_j \mathbb{1}_{A_j} \right\|_{\mathcal{L}^p} + \left\| \sum_{j=1}^n c_j \mathbb{1}_{A_j} - \sum_{j=1}^n c_j \mathbb{1}_{C_j} \right\|_{\mathcal{L}^p} < \varepsilon,$$

ce qui conclut la preuve. □

5.8 Dualité des espaces L^p

L'inégalité de Hölder nous apprend que toute fonction $g \in L^{p'}(X)$ engendre une fonctionnelle linéaire continue $L^p(X) \to \mathbb{R} : f \mapsto \int fgd\mu$. Le théorème suivant montre que toutes les fonctionnelles linéaires continues sur $L^p(X)$ sont de cette forme. Plus précisément, on montre que l'espace de Banach dual de $L^p(X)$ est isométrique à $L^{p'}(X)$. [1]

Théorème 5.28 (Théorème de représentation de Riesz pour L^p)**.** *Soit $(X, \mathcal{A}, \mu)$ un espace de mesure σ-fini, et soit $1 \le p < \infty$. Considérons une fonctionnelle linéaire continue Λ sur $\mathcal{L}^p(X)$, c'est-à-dire une application $\Lambda : \mathcal{L}^p(X) \to \mathbb{R}$ telle que*

— *$\Lambda(\alpha f) = \alpha\Lambda(f)$ et $\Lambda(f + g) = \Lambda(f) + \Lambda(g)$ pour tous $f, g \in \mathcal{L}^p(X, \mathcal{A}, \mu)$ et $\alpha \in \mathbb{R}$;*

— *$|\Lambda(f)| \le C\|f\|_{\mathcal{L}^p}$ pour tout $f \in \mathcal{L}^p(X, \mathcal{A}, \mu)$, pour une certaine constante $C \ge 0$.*

Alors Λ est forcément de la forme $\Lambda(f) = \int_X fgd\mu$ pour une fonction $g \in \mathcal{L}^{p'}(X, \mathcal{A}, \mu)$. De plus, la fonction g est unique à égalité presque partout près, et l'on a

$$\|g\|_{\mathcal{L}^{p'}} = \|\Lambda\|_{(\mathcal{L}^p)'} := \sup_{f \in \mathcal{L}^p(X), \|f\|_{\mathcal{L}^p} = 1} \Lambda(f).$$

Démonstration. Par σ-finitude, on se ramène aisément au cas où μ est une mesure finie ; supposons dorénavant que μ est finie. Pour tout $A \in \mathcal{A}$, on a alors $\mathbb{1}_A \in \mathcal{L}^p(X)$ et l'on peut donc définir

$$\nu(A) := \Lambda(\mathbb{1}_A) \in \mathbb{R}.$$

On remarque que :

— $\nu(\varnothing) = \Lambda(0) = 0$

— ν est σ-additive : pour une suite $(A_n)_n \subset \mathcal{A}$ d'éléments deux à deux disjoints, en notant $A = \cup_n A_n$, on remarque que $\|\mathbb{1}_{\cup_{n \le N} A_n} - \mathbb{1}_A\|_{\mathcal{L}^p} = \mu(\cup_{n>N} A_n)^{1/p} = (\sum_{n>N} \mu(A_n))^{1/p} \to 0$ pour $N \uparrow \infty$ comme $\sum_n \mu(A_n) = \mu(A) \le \mu(X) < \infty$. Dès lors, par continuité et linéarité de Λ, on obtient $\Lambda(\mathbb{1}_A) = \lim_N \Lambda(\mathbb{1}_{\cup_{n \le N} A_n}) = \lim_N \Lambda(\sum_{n \le N} \mathbb{1}_{A_n}) = \sum_n \Lambda(\mathbb{1}_{A_n})$, c'est-à-dire $\nu(A) = \sum_n \nu(A_n)$.

Par ailleurs, par continuité de Λ, notons que

$$|\nu(A)| \le C\|\mathbb{1}_A\|_{\mathrm{L}^p} = C\mu(A)^{\frac{1}{p}}. \tag{5.4}$$

1. On insiste qu'on parle ici de dual topologique : pour un espace vectoriel topologique E, le dual topologique E' est le sous-espace du dual algébrique E^* constitué des formes linéaires continues. Si E est un espace vectoriel normé, on peut munir E' de la norme duale $\|f\|_{E'} := \sup_{v \in E, \|v\|_E = 1} |f(v)|$, et le dual E' muni de cette topologie est forcément complet — c'est un espace de Banach. En dimension finie on a toujours $E' = E^*$, mais ce n'est plus le cas en dimension infinie.

Ceci impliquerait que l'application ν sur $\mathcal{A}$ est une mesure absolument continue par rapport à μ, ce qui permettrait d'appliquer le théorème de Radon-Nikodym. Or, ν n'est pas positive et n'est donc pas une mesure au sens donné dans ce cours (c'est ce qu'on appellerait une mesure signée). On commence donc par la décomposer en une partie positive et une partie négative.

Étape 1 : décomposition de Jordan.
On va montrer que l'application ν peut être décomposée comme $\nu = \nu_+ - \nu_-$, où ν_+ et ν_- sont deux mesures sur $(X, \mathcal{A})$. Pour ce faire, pour tout $A \in \mathcal{A}$, posons

$$\nu_+(A) := \sup\{\nu(B) : B \subset A, B \in \mathcal{A}\}, \qquad \nu_-(A) := \sup\{-\nu(B) : B \subset A, B \in \mathcal{A}\}.$$

Il est clair que $\nu_\pm(A) \geq 0$ et $\nu_\pm(\varnothing) = 0$. Pour montrer que ν_+ et ν_- sont des mesures, il reste à montrer la σ-additivité. Focalisons-nous sur ν_+. Soit $(A_n)_n \subset \mathcal{A}$ deux à deux disjoints et notons $A := \cup_n A_n$. Pour $B \in \mathcal{A}$ avec $B \subset A$, on a $B = \cup_n (B \cap A_n)$ et donc $\nu(B) = \sum_n \nu(B \cap A_n) \leq \sum_n \nu_+(A_n)$. En prenant le supremum par rapport à B, on déduit

$$\nu_+(A) \leq \sum_n \nu_+(A_n).$$

Montrons à présent l'autre inégalité. Étant donné $\varepsilon > 0$, pour tout n, on peut choisir $B_n \in \mathcal{A}$, $B_n \subset A_n$, tel que $\nu(B_n) \geq \nu_+(A_n) - \varepsilon 2^{-n}$. Posons alors $B := \cup_n B_n \subset A$, qui satisfait $\nu(B) = \sum_n \nu(B_n) \geq \sum_n \nu_+(A_n) - \varepsilon$. Ceci conclut la σ-additivité de ν_+, et il en va de même pour ν_-.
Il reste à vérifier que $\nu = \nu_+ - \nu_-$. Pour tout $A, B \in \mathcal{A}$ avec $B \subset A$, en écrivant $\nu(A) = \nu(B) + \nu(A \setminus B)$ et en utilisant respectivement $B \subset A$ et $A \setminus B \subset A$, on peut déduire

$$\nu(B) - \nu_-(A) \leq \nu(A) \leq \nu_+(A) - (-\nu(B)).$$

En prenant respectivement le supremum et l'infimum sur B dans les bornes inférieures et supérieures, on obtient $\nu(A) = \nu_+(A) - \nu_-(A)$.

Étape 2 : conclusion.
Par (5.4) et par la définition des mesures ν_+ et ν_-, on voit que celles-ci sont absolument continues par rapport à μ. Le théorème de Radon-Nikodym donne alors des densités $g_+, g_- : X \to [0, \infty]$ telles que pour tout $A \in \mathcal{A}$,

$$\nu_\pm(A) = \int_A g_\pm d\mu.$$

Comme μ est une mesure finie, on déduit de (5.4) et de la définition des mesures ν_+ et ν_- que celles-ci sont également finies, de sorte que les densités $g_\pm$ sont μ-intégrables, et donc $g_\pm < \infty$ μ-p.p., disons $g_\pm(x) < \infty$ pour tout $x \notin N$ avec $N \in \mathcal{A}$ et $\mu(N) = 0$. On peut alors définir $g := (g_+ - g_-)\mathbb{1}_{N^c}$, et on obtient pour tout $A \in \mathcal{A}$,

$$\Lambda(\mathbb{1}_A) = \nu(A) = \nu_+(A) - \nu_-(A) = \int_A g d\mu.$$

Par linéarité, pour toute fonction f simple mesurable, ceci implique

$$\Lambda(f) = \int f g d\mu. \tag{5.5}$$

Par la continuité de Λ, on peut alors borner pour toute fonction f simple mesurable,

$$\left| \int f g d\mu \right| = |\Lambda(f)| \leq C \|f\|_{\mathcal{L}^p}.$$

Pour $n \geq 1$, considérons la fonction $f_n := \mathbb{1}_{|g|\leq n}\mathbb{1}_{B_n} g|g|^{p'-2} \in \mathcal{L}^p(X)$. Par densité des fonctions simples mesurables, l'estimation ci-dessus vaut également pour $f = f_n$ et donne alors

$$\int \mathbb{1}_{|g|\leq n}\mathbb{1}_{B_n}|g|^{p'} d\mu = |\Lambda(f_n)| \leq C\Big(\int \mathbb{1}_{|g|\leq n}\mathbb{1}_{B_n}|g|^{p'} d\mu\Big)^{\frac{1}{p}},$$

c'est-à-dire $\|\mathbb{1}_{|g|\leq n}\mathbb{1}_{B_n} g\|_{\mathcal{L}^{p'}} \leq C$. Par convergence monotone, on peut conclure $g \in \mathcal{L}^{p'}(X)$. Par l'inégalité de Hölder, on déduit que la fonctionnelle $\mathcal{L}^p(X) \to \mathbb{R} : f \mapsto \int fgd\mu$ est bien définie et continue. En combinant avec la continuité de Λ, cela entraîne que l'égalité (5.5) est valable pour tout $f \in \mathcal{L}^p(X)$. L'unicité et l'identité des normes suivent aisément. □

Remarque 5.29. Contrairement au cas $1 \leq p < \infty$, la structure de l'espace de Banach dual $L^\infty(X)'$ est beaucoup plus délicate. Il est en général faux que toute fonctionnelle linéaire continue sur $L^\infty(X)$ s'obtienne par intégration contre une fonction de $L^1(X)$. On peut en fait identifier $L^\infty(X)' \simeq \mathrm{ba}(\mathcal{A})$, l'espace des mesures signées, bornées, et finiment additives sur la σ-algèbre $\mathcal{A}$. En général, cet espace est strictement plus grand que $L^1(X)$: l'inclusion naturelle $L^1(X) \subset L^\infty(X)'$ est toujours isométrique, mais n'est alors pas surjective. La description complète du dual dépasse le cadre de ce cours.

Chapitre 6

Modes de convergence

Si $(x_n)_n \subset \mathbb{R}$ est une suite de nombres réels, la notion de convergence vers une limite $x \in \mathbb{R}$ est parfaitement univoque : pour tout $\varepsilon > 0$ il existe un indice N tel que $|x_n - x| \leq \varepsilon$ pour tout $n \geq N$. Plus généralement, dans tout espace vectoriel réel de dimension finie, la convergence d'une suite est indépendante du choix de la norme, puisque toutes les normes y sont équivalentes.

La situation change radicalement lorsqu'on considère une suite $(f_n)_n$ de fonctions $f_n : X \to \mathbb{R}$ définies sur un domaine X. Dès que X est infini, l'espace $\mathbb{R}^X$ est de dimension infinie : les fonctions possèdent un nombre illimité de degrés de liberté, offrant de multiples façons d'"approcher" une limite. En analyse, la diversité de ces comportements rend indispensable une étude systématique des différents modes de convergence.

Dans les cours de CDI1&2, on a surtout rencontré les deux notions suivantes :

— **Convergence simple :** On dit que $f_n \to f$ simplement sur X si, pour tout $x \in X$, on a $f_n(x) \to f(x)$. Autrement dit, pour tout $x \in X$ et tout $\varepsilon > 0$ il existe N (dépendant de x) tel que $|f_n(x) - f(x)| \leq \varepsilon$ pour tout $n \geq N$.

— **Convergence uniforme :** On dit que $f_n \to f$ uniformément sur X si $\sup_{x\in X} |f_n(x) - f(x)| \to 0$. Autrement dit, pour tout $\varepsilon > 0$, il existe N tel que $|f_n(x) - f(x)| \leq \varepsilon$ pour tout $n \geq N$ et tout $x \in X$. La différence avec la convergence simple est que l'indice N peut être choisi ici indépendamment de x.

Il est clair que la convergence uniforme implique la convergence simple, et que l'inverse est faux en général. Par exemple, pour la suite $(f_n)_n$ de fonctions $\mathbb{R} \to \mathbb{R}$ définies par $f_n(x) = x/n$, on a $f_n \to 0$ simplement, mais pas uniformément.

Dans cette section, nous introduirons plusieurs modes de convergence propres au cadre de la théorie de la mesure. Ils jouent un rôle central en analyse fonctionnelle et en théorie de l'intégration, et leurs interactions (ou absence d'interactions) constituent l'une des pierres angulaires de la théorie.

6.1 Convergences de fonctions mesurables

On commence par introduire un nombre de notions naturelles de convergence pour des fonctions mesurables. Notons que ces notions de convergence restent inchangées si les fonctions sont modifiées presque nulle part. Remarquons également que pour chacune de ces notions de convergence, la limite est forcément unique presque partout.

Définition 6.1. Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $(f_n)_n$ une suite de fonctions mesurables $X \to \mathbb{R}$, ainsi qu'une fonction mesurable $f : X \to \mathbb{R}$.

(i) On dit que la suite $(f_n)_n$ *converge presque partout* vers f (noté $f_n \to f$ p.p.) s'il existe $N \in \mathcal{A}$ avec $\mu(N) = 0$ tel que $f_n(x) \to f(x)$ pour tout $x \notin N$.

(ii) On dit que la suite $(f_n)_n$ *converge en mesure* vers f (noté $f_n \xrightarrow{\mu} f$) si pour tout $\varepsilon > 0$ on a $\mu\big(\{x \in X : |f_n(x) - f(x)| > \varepsilon\}\big) \to 0$.

(iii) On dit que la suite $(f_n)_n$ *converge en L^p* vers f (noté $f_n \to f$ en L^p) si $\|f_n - f\|_{\mathcal{L}^p} \to 0$.

(iv) On dit que la suite $(f_n)_n$ *converge presque uniformément* vers f (noté $f_n \to f$ p.u.) si pour tout $\varepsilon > 0$ il existe $X_\varepsilon \in \mathcal{A}$ avec $\mu(X \setminus X_\varepsilon) \le \varepsilon$ tel que $f_n \to f$ uniformément sur X_ε, c'est-à-dire $\sup_{x \in X_\varepsilon} |f_n(x) - f(x)| \to 0$.

Par définition, la convergence simple implique la convergence presque partout, et la convergence uniforme implique la convergence presque uniforme : convergence presque partout et presque uniforme peuvent être vus comme les équivalents en théorie de la mesure des notions de convergence simple et uniformes.

Dans la suite de cette section, on va voir un certain nombre d'implications entre ces différentes notions de convergence, que l'on résume dans l'énoncé suivant. Une partie de ces implications sont montrées dans la suite, tandis que d'autres seront vues en séances d'exercices.

Théorème 6.2. *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $(f_n)_n$ une suite de fonctions mesurables $X \to \mathbb{R}$, ainsi qu'une fonction mesurable $f : X \to \mathbb{R}$.*

(i) Convergence en L^p et en mesure : *pour $1 \le p < \infty$,*
- *Si $f_n \to f$ en L^p, alors $f_n \to f$ en mesure.*
- *Si $f_n \to f$ en mesure et si $|f_n| \le g$ p.p. pour tout n avec $g \in \mathcal{L}^p(X)$, alors $f_n \to f$ en L^p.*

(ii) Convergence en L^p et p.p. : *pour $1 \le p < \infty$,*
- *Si $f_n \to f$ en L^p, alors il existe une sous-suite $(f_{n_j})_j$ telle que $f_{n_j} \to f$ p.p.*
- *Si $f_n \to f$ p.p. et si $|f_n| \le g$ p.p. pour tout n avec $g \in \mathcal{L}^p(X)$, alors $f_n \to f$ en L^p.*

(iii) Convergence en mesure et p.p. :
- *Si $f_n \to f$ en mesure, alors il existe une sous-suite $(f_{n_j})_j$ telle que $f_{n_j} \to f$ p.p.*
- *Si $f_n \to f$ p.p. et si la mesure μ est finie, alors $f_n \to f$ en mesure.*

(iv) Convergence p.p. et p.u. :
- *Si $f_n \to f$ p.p. et si la mesure μ est finie, alors $f_n \to f$ p.u.*
- *Si $f_n \to f$ p.u., alors $f_n \to f$ p.p.*

(v) Convergence en mesure et p.u. :
- *Si $f_n \to f$ en mesure, alors il existe une sous-suite $(f_{n_j})_j$ telle que $f_{n_j} \to f$ p.u.*
- *Si $f_n \to f$ p.u., alors $f_n \to f$ en mesure.*

Exemple 6.3. Sur $(X, \mathcal{A}, \mu) = (\mathbb{R}, \mathcal{M}, \mathcal{L})$, on présente quelques exemples illustrant que les hypothèses du résultat précédent sont, en général, indispensables. Chacun met en évidence un "défaut" de convergence possible.

(i) *Évasion horizontale :* La suite $f_n = \mathbb{1}_{[n,n+1]}$ vérifie $f_n \to 0$ p.p. mais ne converge ni en mesure, ni en L^1, ni p.u. Ici, la masse "s'échappe à l'infini" en abscisse.

(ii) *Écrasement :* La suite $f_n = \frac{1}{n}\mathbb{1}_{[0,n]}$ vérifie $f_n \to 0$ uniformément, p.p., p.u., en mesure, et en L^p pour $p > 1$. En revanche, elle ne converge pas en L^1. Ici, les fonctions s'écrasent mais les supports grandissent, de sorte que la masse totale reste constante, empêchant ainsi la convergence en L^1.

(iii) *Concentration :* La suite $f_n = n\mathbb{1}_{[0,1/n]}$ vérifie $f_n \to 0$ simplement, p.p., p.u., et en mesure, mais ne converge pas uniformément ni en L^1. Ici, la masse se concentre sur un support de mesure tendant vers 0 en produisant des pics de plus en plus hauts.

(iv) *Suite du dactylographe :* Définissons les intervalles A_n qui parcourent $[0, 1]$ de façon dyadique : on pose $A_n = [\frac{n-2^k}{2^k}, \frac{n-2^k+1}{2^k}]$ pour $2^k \le n < 2^{k+1}$. La suite $f_n = \mathbb{1}_{A_n}$ vérifie alors $f_n \to 0$ en mesure et en L^1, mais ne converge pas p.p. ni p.u. Ici, malgré la disparition de la masse en mesure, la suite balaye l'espace de façon à empêcher toute convergence presque partout.

6.2 Équivalences sur un espace de mesure finie

Sur un espace de mesure finie (en particulier, sur un espace de probabilité), on évite naturellement le contrexemple ci-dessus de l'évasion horizontale. Dans ce cas, on peut alors montrer que la convergence p.p. implique la convergence en mesure et que les convergences p.p. et p.u. sont équivalentes. On commence par la preuve de ce premier point.

Proposition 6.4. *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $(f_n)_n$ une suite de fonctions mesurables $X \to \mathbb{R}$, ainsi qu'une fonction mesurable $f : X \to \mathbb{R}$. Si μ est une mesure finie et si $f_n \to f$ p.p., alors $f_n \to f$ en mesure.*

Démonstration. Soit $\varepsilon > 0$. Posons $A_n = \{x \in X : |f_n(x) - f(x)| > \varepsilon\}$ et $B_n = \cup_{j \geq n} A_j$. La suite $(B_n)_n$ est une suite décroissante d'éléments de $\mathcal{A}$ et

$$\bigcap_n B_n \subset \{x \in X : (f_n(x))_n \text{ ne converge pas vers } f(x)\}.$$

Comme par hypothèse $f_n \to f$ p.p., on déduit $\mu(\cap_n B_n) = 0$ et la continuité à droite de la mesure implique $\lim_n \mu(B_n) = 0$. Puisque $A_n \subset B_n$, on a aussi $\lim_n \mu(A_n) = 0$, d'où la conclusion. □

On se tourne à présent vers l'équivalence entre convergences p.p. et p.u., qui est l'objet du théorème d'Egorov. Si l'on compare les convergences p.p. et p.u. aux convergences simple et uniforme, respectivement, on comprend l'importance et le caractère non-trivial de ce résultat d'équivalence : quitte à omettre des ensembles de mesure petite, les convergences simple et uniforme deviennent équivalentes.

Théorème 6.5 (Théorème d'Egorov)**.** *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $(f_n)_n$ une suite de fonctions mesurables $X \to \mathbb{R}$, ainsi qu'une fonction mesurable $f : X \to \mathbb{R}$. Si μ est une mesure finie et si $f_n \to f$ p.p., alors $f_n \to f$ p.u.*

Démonstration. Soit $N \in \mathcal{A}$ avec $\mu(N) = 0$ tel que $f_n(x) \to f(x)$ pour tout $x \notin N$. Posons $E_k^n := \{x \in N^c : |f_j(x) - f(x)| < 1/n \ \forall j \geq k\}$. On a $E_{k+1}^n \supset E_k^n$, $\cup_k E_k^n = E$, donc $\mu(E_k^n) \uparrow \mu(X) < \infty$ pour $k \uparrow \infty$. Soit k_n tel que $\mu(X \setminus E_{k_n}^n) < 2^{-n}$ et soit $A_n = \cap_{m \geq n} E_{k_m}^m$. On a alors $\mu(X \setminus A_n) \leq \sum_{m \geq n} 2^{-m} = 2^{1-n} < \epsilon$ pour $n = n_\epsilon$. Enfin, on observe que $f_j \to f$ uniformément sur A_n : pour $x \in A_n$ et $m \geq n$, on a $|f_j(x) - f(x)| < 1/m$ pour tout $j > k_m$, donc $\limsup_j \sup_{A_n} |f_j - f| < 1/m$ pour tout m. □

Une conséquence immédiate de ce théorème d'Egorov est le résultat suivant qui montre que sur $\mathbb{R}^d$ toute fonction borélienne est 'presque' continue, ou plus précisément est continue si on la restreint hors d'un sous-ensemble de mesure arbitrairement petite.

Corollaire 6.6 (Théorème de Lusin)**.** *Sur $\mathbb{R}^d$, soit $E \in \mathcal{M}_d$ avec $\mathcal{L}_d(E) < \infty$, et soit $f : \mathbb{R}^d \to \mathbb{R}$ une fonction mesurable. Alors, pour tout $\epsilon > 0$, il existe un compact $K_\varepsilon \subset E$ avec $\mathcal{L}_d(E \setminus K_\varepsilon) \leq \varepsilon$ tel que la restriction $f|_{K_\varepsilon} : K_\varepsilon \to \mathbb{R}$ est continue.*

Remarque 6.7. Attention, ce résultat ne signifie pas que f est continue en tout point $x \in K_\epsilon$! En utilisant le théorème d'extension de Tietze [1], on peut donner la version équivalente suivante : pour tout $\epsilon > 0$, il existe une fonction continue $f_\epsilon : E \to \mathbb{R}$ telle que $\mathcal{L}_d(\{x : f(x) \neq f_\epsilon(x)\}) < \epsilon$.

Démonstration. Si $f = \mathbb{1}_A$ avec $A \subset E$ et $A \in \mathcal{M}_d$, la régularité intérieure nous permet de choisir, pour tout ε, un compact $K_\varepsilon^0 \subset A$ et un compact $K_\varepsilon^1 \subset E \setminus A$ tels que $\mathcal{L}_d(A \setminus K_\varepsilon^0) \leq \varepsilon$ et $\mathcal{L}_d(E \setminus A) \setminus K_\varepsilon^1) \leq \varepsilon$. Pour $K_\varepsilon := K_\varepsilon^0 \cup K_\varepsilon^1$, on a alors $\mathcal{L}_d(E \setminus K_\varepsilon) \leq 2\varepsilon$ et la fonction $\mathbb{1}_A$ est continue sur K_ε.

1. On rappelle le théorème d'extension de Tietze : sur un espace métrique X (ou plus généralement sur un espace topologique normal), si $F \subset X$ est fermé, si $f : F \to \mathbb{R}$ est une fonction continue, alors il existe une extension $\bar{f}$ de f qui est une fonction continue définie sur X tout entier.

Le résultat est donc démontré pour une fonction f simple mesurable. Soit f une fonction mesurable quelconque et soit $\varepsilon > 0$. Considérons une suite $(f_n)_n$ de fonctions simples mesurables telles que $f_n \to f$ p.p. Pour tout n, par le résultat déjà démontré pour les fonctions simples, on déduit qu'il existe un compact $E_n \subset E$ avec $\mathcal{L}_d(E \setminus E_n) \leq 2^{-n}$ et avec $f_n|_{E_n}$ continue. Le théorème d'Egorov assure par ailleurs qu'il existe $X_\varepsilon \in \mathcal{A}$ avec $\mathcal{L}_d(E \setminus X_\varepsilon) \leq \varepsilon$ tel que $f_n \to f$ uniformément sur X_ε. Pour $F_{\varepsilon,N} := X_\varepsilon \cap \cap_{n \geq N} E_n$, on a $\mathcal{L}_d(E \setminus F_{\varepsilon,N}) \leq \varepsilon + 2^{1-N} \leq 2\varepsilon$ pour un $N = N_\varepsilon$. Comme f_n est continu sur $F_{\epsilon,N}$ et $f_n \to f$ uniformement sur $F_{\epsilon,N}$, on déduit que f est aussi continu sur $F_{\epsilon,N}$. La conclusion suit alors de la régularité intérieure. □

Les résultats ci-dessus constituent les deux derniers des trois principes de Littlewood — principes heuristiques résumant la théorie de la mesure sur $\mathbb{R}$:

— *régularité de Lebesgue :* tout ensemble mesurable est 'presque' une union finie d'intervalles ;

— *théorème de Lusin :* toute fonction mesurable est 'presque' continue ;

— *théorème d'Egorov :* toute suite de fonctions convergeant simplement est 'presque' uniformément convergente.

6.3 Convergence en mesure 'rapide'

On a vu que la convergence en mesure (ou en L^p, $p < \infty$) n'implique en général pas la convergence p.p., mais on montre que l'implication est vraie à une sous-suite près. On remarque que ce résultat

Proposition 6.8. *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $(f_n)_n$ une suite de fonctions mesurables $X \to \mathbb{R}$, ainsi qu'une fonction mesurable $f : X \to \mathbb{R}$. Si $(f_n)_n$ converge en mesure vers f, alors il existe une sous-suite de $(f_n)_n$ qui converge vers f presque partout.*

Démonstration. Puisque la suite converge en mesure, il existe un entier n_1 tel que $\mu(\{x \in X : |f_{n_1}(x) - f(x)| > 1\}) \leq \frac{1}{2}$. Par induction, on construit une suite croissante $(n_j)_j$ telle que pour tout j,

$$\mu(A_j) \leq 2^{-j}, \qquad A_j := \{x \in X : |f_{n_j}(x) - f(x)| > \tfrac{1}{j}\}.$$

Par définition, pour tout $x \notin \cap_k \cup_{j \geq k} A_j$, il existe un indice k tel que pour tout $j \geq k$,

$$|f_{n_j}(x) - f(x)| \leq \tfrac{1}{j}.$$

Par conséquent, la suite $(f_{n_j})_j$ converge vers f hors de $\cap_k \cup_{j \geq k} A_j$. De plus, pour tout k,

$$\mu\Big(\bigcup_{j \geq k} A_j\Big) \leq \sum_{j \geq k} \mu(A_j) \leq \sum_{j \geq k} 2^{-j} = 2^{1-k}.$$

et on en tire que $\mu(\cap_k \cup_{j \geq k} A_j) = 0$. □

En combinant le résultat précédent avec la Proposition 6.4, on déduit aisément la caractérisation suivante de la convergence en mesure. En particulier, cela implique qu'il n'y a pas de topologie pour la convergence presque partout : en effet, dans un espace topologique, une suite converge vers un point si et seulement si toute sous-suite possède une sous-sous-suite qui converge vers ce point.

Corollaire 6.9. *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $(f_n)_n$ une suite de fonctions mesurables $X \to \mathbb{R}$, ainsi qu'une fonction mesurable $f : X \to \mathbb{R}$. Supposons que la mesure μ est finie. La suite $(f_n)_n$ converge vers f en mesure si et seulement si toute sous-suite de $(f_n)_n$ possède une sous-sous-suite qui converge presque partout vers f.*

Il y a une autre façon de voir l'équivalence à sous-suite près de la Proposition 6.8 ci-dessus : si la convergence en mesure (ou en L^p, $p < \infty$) est assez rapide, alors cela doit impliquer la convergence p.p. On a en effet le résultat suivant, qui affirme que cela implique même la convergence p.u., sans avoir besoin de supposer que la mesure est finie.

Théorème 6.10 (Convergence en mesure rapide)**.** *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $(f_n)_n$ une suite de fonctions mesurables $X \to \mathbb{R}$, ainsi qu'une fonction mesurable $f : X \to \mathbb{R}$.*

(i) Si $\sum_n \|f_n - f\|_{\mathcal{L}^1} < \infty$, alors $f_n \to f$ p.u.

(ii) Si pour tout $\varepsilon > 0$ on a $\sum_n \mu(\{x : |f_n(x) - f(x)| \geq \varepsilon\}) < \infty$, alors $f_n \to f$ p.u.

Remarque 6.11. La condition du point (ii) de convergence en mesure 'rapide', c'est-à-dire la condition que pour tout $\varepsilon > 0$ on a $\sum_n \mu(\{x : |f_n(x) - f(x)| \geq \varepsilon\}) < \infty$, est parfois appelée *convergence complète* (voir Hsu-Robbins 1947).

Démonstration. On se focalise sur la preuve de (i). Comme par hypothèse la fonction $\sum_n |f_n - f|$ est intégrable, on déduit $\sum_n |f_n - f| < \infty$ p.p., ce qui montre que $f_n = f_1 + \sum_{k=1}^{n-1}(f_{k+1} - f_k)$ converge p.p. Posons à présent $E_k^n := \{x : |f_j(x) - f(x)| \leq 1/n \ \forall j \geq k\}$. Pour tout n, l'inégalité de Markov permet de borner

$$\mu(X \setminus E_k^n) = \mu(\{x : \exists j \geq k, \, |f_j(x) - f(x)| > 1/n\}) \leq \sum_{j \geq k} n \|f_j - f\|_{\mathcal{L}^1} \xrightarrow{k \uparrow \infty} 0.$$

Soit k_n tel que $\mu(X \backslash E_{k_n}^n) \leq 2^{-n}$ et soit $A_n = \cap_{m \geq n} E_{k_m}^m$. On a alors $\mu(X \backslash A_n) \leq \sum_{m \geq n} 2^{-m} = 2^{1-n} < \varepsilon$ pour $n = n_\varepsilon$. Enfin, on observe que $f_j \to f$ uniformément sur A_n. □

6.4 Domination et uniforme intégrabilité

On a vu que la convergence en mesure ou presque partout n'implique pas la convergence en L^p. En corollaire du théorème de convergence dominée, l'implication est cependant vraie si la suite est dominée.

Lemme 6.12. *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $(f_n)_n$ une suite de fonctions mesurables $X \to \mathbb{R}$, ainsi qu'une fonction mesurable $f : X \to \mathbb{R}$. Si f_n converge vers f presque partout ou en mesure et s'il existe $g \in \mathcal{L}^1(X)$ tel que $|f_n| \leq g$ p.p. pour tout n, alors f_n vers f en L^1.*

Démonstration. Si f_n converge p.p. vers f, la convergence en L^1 est une conséquence directe du théorème de la convergence dominée. Supposons à présent que f_n converge en mesure vers f. Par la Proposition 6.8, on déduit que toute sous-suite possède une sous-sous-suite qui converge presque partout vers f. Par le cas déjà montré, on peut déduire que cette sous-sous-suite converge en L^1. Enfin, si toute sous-suite possède une sous-sous-suite qui converge en L^1 vers f, on déduit aisément que la suite entière converge en L^1. □

Cependant, l'hypothèse de domination est très contraignante, et il est naturel de chercher à l'affaiblir. Comme on l'a déjà observé, deux phénomènes distincts peuvent empêcher la convergence en mesure d'impliquer la convergence en L^1 :

— l'écrasement, lorsque la masse de l'intégrale se disperse sur des régions de plus en plus vastes (par exemple $f_n = \frac{1}{n}\mathbb{1}_{[0,n]}$ sur $\mathbb{R}$) ;

— la concentration, lorsque la masse se concentre sur des pics de plus en plus élevés mais de plus en plus concentrés (par exemple $f_n = n\mathbb{1}_{[0,1/n]}$ sur $\mathbb{R}$).

Pour exclure simultanément ces deux comportements pathologiques, on introduit la notion suivante. La condition (ii) contrôle précisément la concentration, tandis que la condition (iii) empêche l'écrasement.

Définition 6.13. Soit $(X, \mathcal{A}, \mu)$ un espace de mesure. Une suite $(f_n)_n \subset \mathcal{L}^1(X)$ est dite *uniformément intégrable* si les trois conditions suivantes sont satisfaites :

(i) *borne $\mathcal{L}^1$ uniforme :* $\sup_n \int |f_n| d\mu < \infty$;

(ii) *équi-intégrabilité :* $\sup_n \int_{\{x:|f_n(x)|\geq M\}} |f_n| d\mu \to 0$ pour $M \uparrow \infty$;

(iii) *tension :* pour tout $\varepsilon > 0$ il existe $X_\varepsilon \in \mathcal{A}$ avec $\mu(X_\varepsilon) < \infty$ tel que $\sup_n \int_{X \setminus X_\varepsilon} |f_n| d\mu \leq \varepsilon$.

Remarque 6.14. Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $(f_n)_n \subset \mathcal{L}^1(X)$. On vérifie aisément les cas particuliers suivants :

— Si $\mu(X) < \infty$, alors $(f_n)_n$ est uniformément intégrable si et seulement si la condition (ii) ci-dessus est satisfaite.

— S'il existe $g \in \mathcal{L}^1(X)$ tel que $|f_n| \leq g$ p.p. pour tout n, alors $(f_n)_n$ est uniformément intégrable.

En ces termes, on peut à présent montrer que l'hypothèse de domination uniforme du Lemme 6.12 peut être remplacé par l'uniforme intégrabilité. Par le second point de la remarque ci-dessus, on voit que cela constitue bien une généralisation (stricte).

Théorème 6.15 (Théorème de Vitali)**.** *Soient $(X, \mathcal{A}, \mu)$ un espace de mesure et $(f_n)_n$ une suite de fonctions mesurables $X \to \mathbb{R}$, ainsi qu'une fonction mesurable $f : X \to \mathbb{R}$. Supposons que $(f_n)_n$ est uniformément intégrable.*

(i) f_n converge vers f en mesure si et seulement si f_n converge vers f en L^1.

(ii) Si f_n converge vers f p.p., alors f_n converge vers f en L^1.

Démonstration. On se focalise sur la preuve de (i) ; la preuve de (ii) se fait de même. Supposons que $(f_n)_n$ est uniformément intégrable et converge en mesure vers f. Par la Proposition 6.8, il existe une sous-suite $(f_{n_k})_k$ qui converge p.p. vers f. Soit $\varepsilon > 0$. L'uniforme intégrabilité de la suite implique :

— Il existe $A > 0$ tel que $\int |f_n| d\mu \leq A$ pour tout n. Par le lemme de Fatou, on déduit $\int |f| d\mu = \int \liminf_k |f_{n_k}| d\mu \leq \liminf_k \int |f_{n_k}| d\mu \leq A$, et donc $f \in \mathcal{L}^1(X)$.

— Il existe $M_\varepsilon > 0$ tel que $\int |f_n| \mathbb{1}_{|f_n| > M_\varepsilon} d\mu \leq \varepsilon$ pour tout n. Par le lemme de Fatou, on déduit $\int |f| \mathbb{1}_{|f| > M_\varepsilon} d\mu \leq \varepsilon$. En combinant ces informations, on obtient $\int |f_n - f| \mathbb{1}_{|f_n - f| > 2M_\varepsilon} \leq 6\varepsilon$.

— Il existe $X_\varepsilon \in \mathcal{A}$ avec $\mu(X_\varepsilon)$ tel que $\int_{X \setminus X_\varepsilon} |f_n| d\mu \leq \varepsilon$ pour tout n. Par le lemme de Fatou, on déduit $\int_{X \setminus X_\varepsilon} |f| d\mu \leq \varepsilon$.

En combinant ces informations, on peut décomposer pour tout $\kappa > 0$,

$$\begin{aligned}
\int |f_n - f| d\mu &\leq \int_{X_\varepsilon} |f_n - f| d\mu + 2\varepsilon \\
&\leq \int_{X_\varepsilon} |f_n - f| \mathbb{1}_{|f_n - f| \leq M_\varepsilon} d\mu + 8\varepsilon \\
&\leq \int_{X_\varepsilon} |f_n - f| \mathbb{1}_{\kappa \leq |f_n - f| \leq M_\varepsilon} d\mu + \int_{X_\varepsilon} |f_n - f| \mathbb{1}_{|f_n - f| \leq \kappa} d\mu + 8\varepsilon \\
&\leq M_\varepsilon \mu(\{x : |f_n(x) - f(x)| \geq \kappa\}) + \kappa \mu(X_\varepsilon) + 8\varepsilon.
\end{aligned}$$

La convergence en mesure implique alors $\limsup_n \int |f_n - f| d\mu \leq \kappa \mu(X_\varepsilon) + 8\varepsilon$ pour tout $\kappa > 0$. Comme $\mu(X_\varepsilon) < \infty$ et comme κ, ε sont arbitraires, on conclut que f_n converge vers f en L^1. La réciproque est une conséquence directe de l'inégalité de Markov. □

Comme corollaire du théorème de Vitali ci-dessus, on déduit par exemple aisément le résultat très utile suivant, laissé en exercice.

Corollaire 6.16. *Soient* $(X, \mathcal{A}, \mu)$ *un espace de mesure et* $(f_n)_n$ *une suite de fonctions mesurables* $X \to \mathbb{R}$*, ainsi qu'une fonction mesurable* $f : X \to \mathbb{R}$*. Si* $f_n \to f$ *en mesure, si* $\mu(X) < \infty$*, et si* $\sup_n \|f_n\|_{\mathcal{L}^p} < \infty$*, alors* $f_n \to f$ *en* L^r *pour tout* $1 \leq r < p$*.*