

Cours de Licence 3: Théorie de l'intégration

Cours de: Franck Sueur
Rédigé par Hugo Clouet*

Année universitaire 2023 - 2024
Version: 9 novembre 2025

Table des matières

1 Tribus	3
2 Mesures	5
3 Fonctions mesurables	7
4 Fonctions simples et leur intégrale de Lebesgue	11
5 Intégrale de Lebesgue de fonctions mesurables positives	15
6 Théorème de convergence monotone de Lebesgue	17
7 Théorème de convergence dominée de Lebesgue	21
8 Ensembles négligeables et espace $L^1(\mu)$	23
9 Lien avec l'intégrale de Riemann	27
10 Mesure produit, théorème de Fubini et Tonelli	29
11 Théorème fondamental de l'analyse et changement de variable	31
12 Intégrales à paramètre	33

*email: hugo.clouet@etu.u-bordeaux.fr

Introduction

Le but majeur de ce cours est la présentation de la construction et de l'utilisation de l'intégrale de Lebesgue, qui généralise la théorie de l'intégration de Riemann. Il s'agira de définir un espace mesuré :

$$(X, \mathcal{A}, \mu)$$

où :

- (1) l'ensemble X est un ensemble quelconque,
- (2) $\mathcal{A}$ désigne un ensemble, qui est appelé tribu, de partie de X qui sont considérées comme mesurables,
- (3) μ est une mesure, ce qui signifie en particulier que c'est une application qui associe à chaque ensemble E de la tribu $\mathcal{A}$ une valeur positive ($+\infty$ inclus).

L'intégrale

$$\int_E f \, d\mu$$

de f sur E relativement à μ , pour une classe de fonction $f: X \rightarrow \mathbf{R}$ aussi dites mesurables.

La généralité tient à ce que X peut être n'importe quel ensemble et que la plupart des fonctions que l'on a envie de considérer sont mesurables. Nous verrons le cas particulier de l'espace euclidien $\mathbf{R}^n$, muni

- (1) d'une tribu particulière, associée à la topologie standard de $\mathbf{R}^n$ (*i.e.* la topologie associée à n'importe quelle norme), qui est appelée tribu borélienne et notée $\mathcal{B}(\mathbf{R}^n)$,
- (2) d'une mesure particulière, homogène et invariante par translation, appelée mesure de Lebesgue et qui est notée λ_n .

1 Tribus

Il s'agit de la première notion fondamentale en théorie de la mesure. Les éléments d'une tribu sont précisément les parties d'un ensemble qu'il va nous être possible de mesurer.

Dans la suite du cours, sauf indication contraire, X est un ensemble, on note $\mathcal{P}(X)$ l'ensemble des parties de X . Si $A \subset X$, on note A^c ou $X \setminus A$ son complémentaire dans X .

Définition 1.1. Soit X un ensemble. Une collection $\mathcal{A}$ de parties de X est une *tribu* sur X si elle vérifie les propriétés suivantes :

- (i) on a $X \in \mathcal{A}$;
- (ii) si $A \in \mathcal{A}$, alors $A^c \in \mathcal{A}$;
- (iii) pour toute famille $(X_i)_{i \in I}$ d'éléments de $\mathcal{A}$ indexée par un ensemble dénombrable I , on a $\bigcup_{i \in I} X_i \in \mathcal{A}$.

Si $\mathcal{A}$ est une tribu sur un ensemble X , on dit que le couple $(X, \mathcal{A})$ est un *espace mesurable*.

On rappelle qu'un ensemble E est *dénombrable* s'il existe $A \subset \mathbf{N}$ et une bijection $\varphi: E \rightarrow A$. Toute partie finie ou infinie de $\mathbf{N}$ (en particulier l'ensemble des entiers pairs, l'ensemble des entiers impairs, l'ensemble des nombres premiers) est dénombrable, tout ensemble fini est dénombrable, le produit d'un nombre fini d'ensembles dénombrables (en particulier $\mathbf{N} \times \mathbf{N}$) est dénombrable, une réunion dénombrable d'ensembles dénombrables est dénombrable, l'ensemble des nombres rationnels $\mathbf{Q}$ est dénombrable.

En revanche, l'ensemble des nombres réels $\mathbf{R}$ n'est pas dénombrable. C'est un théorème que l'on doit à Cantor. En conséquence, l'ensemble $\mathbf{R} \setminus \mathbf{Q}$ des nombres irrationnels est non-dénombrable.

Enfin, notons que si E est un ensemble infini dénombrable alors il existe une bijection $\varphi: \mathbf{N} \rightarrow E$ de sorte que l'on peut énumérer les éléments de l'ensemble E , *i.e.* on peut écrire $E = \{e_0, e_1, \dots, e_n, \dots\}$.

Donnons deux exemples universels (*i.e.* qui s'applique pour tout ensemble) de tribus; pour tout ensemble X , l'ensemble $\mathcal{A} = \mathcal{P}(X)$ des parties de X est une tribu sur X , c'est la *tribu discrète*. L'ensemble $\mathcal{A} = \{\emptyset, X\}$ est également une tribu sur X , c'est la *tribu grossière*.

On énonce quelques propriétés découlant facilement de la définition.

Lemme 1.2. Si X est un ensemble et $\mathcal{A}$ une tribu sur X , on a les propriétés suivantes :

- (i) on a $\emptyset \in \mathcal{A}$;
- (ii) toute intersection dénombrable d'éléments de $\mathcal{A}$ est élément de $\mathcal{A}$: $(\forall (A_i)_{i \in \mathbf{N}} \in \mathcal{A}) \bigcap_{i \in \mathbf{N}} A_i \in \mathcal{A}$.
- (iii) Si $A, B \in \mathcal{A}$, alors $A \setminus B \in \mathcal{A}$.

Démonstration. (i) Il suffit de remarquer que $\emptyset = X^c$.

(ii) On a $\bigcap_{i \in \mathbf{N}} A_i = \left(\bigcup_{i \in \mathbf{N}} (A_i)^c \right)^c$ donc si pour tout $i \in \mathbf{N}$, $A_i \in \mathcal{A}$, alors $A_i^c \in \mathcal{A}$, donc $\bigcup_{i \in \mathbf{N}} (A_i)^c \in \mathcal{A}$ et enfin $\left(\bigcup_{i \in \mathbf{N}} (A_i)^c \right)^c \in \mathcal{A}$.

(iii) On a : $A \setminus B := A \cap (B^c)$ et il est facile de voir par le second axiome que $A \cap (B^c) \in \mathcal{A}$, d'où $A \setminus B \in \mathcal{A}$. $\square$

Le résultat suivant énonce qu'une intersection quelconque de tribus est une tribu.

Lemme 1.3. Soit X un ensemble. Si $I \neq \emptyset$ est un ensemble quelconque d'indices et $(\mathcal{A}_i)_{i \in I}$ une famille de tribus sur X , alors l'intersection

$$\bigcap_{i \in I} \mathcal{A}_i := \{A \subset X; (\forall i \in I) A \in \mathcal{A}_i\}$$

est une tribu sur X .

Démonstration. Posons $\mathcal{A} := \bigcap_{i \in I} \mathcal{A}_i$. On a $\emptyset \in \mathcal{A}_i$ pour tout $i \in I$, donc $\emptyset \in \mathcal{A}$. Considérons $(A_n)_{n \in \mathbf{N}}$ de $\mathcal{A}$ et montrons que $\bigcup_{n \in \mathbf{N}} A_n \in \mathcal{A}$. Pour tout $i \in I$ et $n \in \mathbf{N}$, $A_n \in \mathcal{A}_i$, donc $\bigcup_{n \in \mathbf{N}} A_n \in \mathcal{A}_i$, d'où $\bigcup_{n \in \mathbf{N}} A_n \in \mathcal{A}$. Enfin, prenons $A \in \mathcal{A}$ et montrons que $A^c \in \mathcal{A}$, on a $A \in \mathcal{A}_i$ pour tout i , donc $A^c \in \mathcal{A}_i$ pour tout i , d'où $A^c \in \mathcal{A}$. $\square$

Ce résultat nous permet notamment de montrer la conséquence suivante.

Lemme 1.4. Si X est un ensemble et $\mathcal{E} \subset \mathcal{P}(X)$, alors :

$$\mathcal{A}_{\mathcal{E}} := \bigcap_{\mathcal{B} \in \mathcal{S}} \mathcal{B}$$

où $\mathcal{S}$ désigne l'ensemble des tribus sur X contenant $\mathcal{E}$, est la plus petite tribu sur X contenant $\mathcal{E}$.

Avant de commencer la preuve, précisons que dans l'énoncé ci-dessus, dire que $\mathcal{A}_{\mathcal{E}}$ est la plus petite tribu sur X contenant $\mathcal{E}$ signifie que $\mathcal{A}_{\mathcal{E}}$ est une tribu sur X et si $\mathcal{A}$ est une autre tribu sur X contenant $\mathcal{E}$, alors $\mathcal{A}_{\mathcal{E}} \subset \mathcal{A}$.

La tribu $\mathcal{A}_{\mathcal{E}}$ s'appelle la *tribu engendrée* par $\mathcal{E}$ sur X .

Démonstration. On observe que $\mathcal{S} \neq \emptyset$ car $\mathcal{P}(X)$ est une tribu contenant $\mathcal{E}$. D'après le lemme 1.3, $\mathcal{A}_{\mathcal{E}}$ est une tribu sur X . De plus, pour tout $\mathcal{B} \in \mathcal{S}$, on a $\mathcal{E} \subset \mathcal{B}$ et par conséquent, $\mathcal{E} \subset \mathcal{A}_{\mathcal{E}}$. Ainsi $\mathcal{A}_{\mathcal{E}}$ est une tribu sur X et contient $\mathcal{E}$. De plus, si $\mathcal{A}$ est une autre tribu sur X contenant $\mathcal{E}$, i.e. $\mathcal{A} \in \mathcal{S}$, alors comme $\mathcal{A}_{\mathcal{E}}$ est par définition l'intersection de $\mathcal{A}$ et des autres tribus sur X contenant $\mathcal{E}$, on a $\mathcal{A}_{\mathcal{E}} \subset \mathcal{A}$. $\square$

Il est important de ne pas mélanger les axiomes définissant la notion de tribu de ceux définissant la notion de topologie, d'autant plus que la définition suivante associe, canoniquement, à toute topologie une tribu.

Définition 1.5. Soit $(X, \mathcal{U})$ un espace topologique (i.e. $\mathcal{U}$ forme la collection des ouverts de X). La tribu $\mathcal{B}$ engendrée par les ouverts $\mathcal{U}$ sur X est la *tribu borélienne* (ou tribu de Borel) de X relativement à la topologie $\mathcal{U}$. Un élément de $\mathcal{B}$ est un *borélien* de X (relativement à $\mathcal{U}$).

Proposition 1.6. Soit $(X, \mathcal{U})$ un espace topologique et soit $\mathcal{B}$ la tribu borélienne associée. On a les propriétés suivantes.

- (i) Tout fermé de X est élément de $\mathcal{B}$.
- (ii) Toute réunion dénombrable de fermés de X est élément de $\mathcal{B}$.
- (iii) Toute intersection dénombrable d'ouverts de X est élément de $\mathcal{B}$.

Dans le cas particulier de l'espace euclidien $\mathbf{R}^n$ pour $n \geq 1$, on peut montrer que la tribu borélienne $\mathcal{B}(\mathbf{R}^n)$ de $\mathbf{R}^n$ est la tribu engendrée par les pavés ouverts :

$$D((x_i)_i, (y_j)_j) := \{(z_1, z_2, \dots, z_n) \in \mathbf{R}^n; (\forall i \in I) x_i < z_i < y_i\}$$

lorsque $((x_i)_i, (y_j)_j)$ parcourent les couples de n -uplets de $\mathbf{R}^n$ tels que $x_i < y_i$ pour $1 \leq i \leq n$.

En particulier, dans le cas $n = 1$, la tribu borélienne de $\mathbf{R}$ est la tribu engendrée par les intervalles ouverts. On rappelle que la valeur absolue usuelle $|\cdot|$ munit $\mathbf{R}$ d'une topologie métrique qu'on appelle *topologie standard* de $\mathbf{R}$. Tout intervalle ouvert est un ouvert pour cette topologie et réciproquement tout ouvert de $\mathbf{R}$ est réunion dénombrable d'intervalles ouverts. Ce fait se généralise pour tout $n \geq 1$: si l'on fixe une norme quelconque $\|\cdot\|$ sur $\mathbf{R}^n$, on peut munir $\mathbf{R}^n$ de la topologie métrique associée.

Celle-ci est indépendante de la norme choisie puisque toutes les normes sont équivalentes sur un même espace de dimension finie. On appelle *topologie standard* de $\mathbf{R}^n$ la topologie associée.

Tout ouvert pour cette topologie est réunion dénombrable de pavés ouverts $D((x_i)_i, (y_j)_j)$.

Dans la suite, on aura à considérer des fonctions pouvant prendre des « valeurs infinies » (penser par exemple à la fonction $x \mapsto 1/\sqrt{x}$ que l'on sait intégrer sur $]0, 1]$ mais qui prend la valeur $+\infty$ en 0^+). On définit :

$$\overline{\mathbf{R}} = \mathbf{R} \cup \{-\infty, +\infty\}$$

et les intervalles généralisés.

On dit qu'une partie $\mathcal{U}$ de $\overline{\mathbf{R}}$ est un ouvert si $\mathcal{U}$ est réunion dénombrable d'intervalles ouverts de $\mathbf{R}$ et d'intervalles généralisés $]a, +\infty]$ et $[-\infty, b[$ pour $a, b \in \mathbf{R}$. La famille des ouverts $\mathcal{U}$ de $\overline{\mathbf{R}}$ ainsi définie, définit une topologie sur $\overline{\mathbf{R}}$, appelée *topologie standard* de $\overline{\mathbf{R}}$.

On peut aussi montrer que la tribu borélienne $\mathcal{B}(\overline{\mathbf{R}})$ de $\overline{\mathbf{R}}$ est la plus petite tribu contenant à la fois les intervalles ouverts de $\mathbf{R}$ et les intervalles $]a, +\infty]$ et $[-\infty, b[$, pour tout $a, b \in \mathbf{R}$.

2 Mesures

On introduit dans cette section le concept de mesure sur un espace mesurable. Une mesure peut prendre des valeurs infinies, *i.e.* des valeurs dans $[0, +\infty] \subset \overline{\mathbf{R}}$.

Définition 2.1. Soit $(X, \mathcal{A}_X)$ un espace mesurable. Une *mesure* sur $(X, \mathcal{A}_X)$ est une application $\mu: \mathcal{A}_X \rightarrow [0, +\infty]$ vérifiant :

- (i) $\mu(\emptyset) = 0$;
- (ii) pour toute suite $(A_i)_i$ de $\mathcal{A}_X$ telle que $A_i \cap A_j = \emptyset$ dès que $i \neq j$, on a

$$\mu\left(\bigcup_{i \geq 1} A_i\right) = \sum_{i \geq 1} \mu(A_i).$$

Un *espace mesuré* est un triplet $(X, \mathcal{A}_X, \mu)$ où $(X, \mathcal{A}_X)$ est un espace mesurable et μ est une mesure sur $(X, \mathcal{A}_X)$.

Exemple 2.2. Donnons quelques exemples de mesures.

(1) En premier lieu, $(\mathbf{N}, \mathcal{P}(\mathbf{N}))$ est un espace mesurable, cela correspond au cas de la tribu dite discrète. Le cardinal permet alors de définir une mesure μ que l'on appelle *mesure de dénombrement* : si $E \subset \mathbf{N}$, on pose $\mu(E) = +\infty$ si E est infini et $\mu(E) = \text{Card}(E)$ si E est fini.

(2) Soit $(X, \mathcal{A})$ un espace mesurable. Soit $x_0 \in X$ et $A \in \mathcal{A}$. On définit $\delta_{x_0}: \mathcal{A} \rightarrow [0, +\infty]$ par

$$\delta_{x_0}(A) = \begin{cases} 1 & \text{si } x_0 \in A \\ 0 & \text{si } x_0 \notin A \end{cases}$$

L'application δ_{x_0} est une mesure appelée *mesure de Dirac* en x_0 sur X . On a déjà $\delta_{x_0}(\emptyset) = 0$ pour tout $x_0 \in X$ car $x_0 \notin \emptyset$. Si $(A_n)_n \in \mathcal{A}$, avec les $(A_n)_n$ deux à deux disjoints, on a

$$\delta_{x_0}\left(\bigcup_n A_n\right) = \begin{cases} 1 & \text{si } x_0 \in \bigcup_n A_n \\ 0 & \text{sinon} \end{cases}$$

donc

$$\sum_{n \geq 0} \delta_{x_0}(A_n) = \begin{cases} 1 & \text{si } x_0 \in \bigcup_n A_n \\ 0 & \text{sinon} \end{cases}$$

Théorème 2.3. Soit $(X, \mathcal{A}, \mu)$ un espace mesuré.

(i) Si $n \in \mathbf{N}_{>0}$ et $A_1, \dots, A_n \in \mathcal{A}$ sont des parties deux à deux disjointes de X , alors

$$\mu\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n \mu(A_i).$$

En particulier si $B_1, \dots, B_n \in \mathcal{A}$ sont quelconques, alors

$$\mu\left(\bigcup_{i=1}^n B_i\right) \leq \sum_{i=1}^n \mu(B_i).$$

(ii) Si $A, B \in \mathcal{A}$ tels que $A \subset B$, alors $\mu(A) \leq \mu(B)$.

(iii) Si $(A_i)_{i \geq 1}$ est une suite croissante d'éléments de $\mathcal{A}$ (*i.e.* $A_i \subset A_{i+1}$ pour tout i), alors la suite $(\mu(A_i))_i$ converge dans $[0, +\infty]$ et

$$\lim_{N \rightarrow +\infty} \mu(A_N) = \mu\left(\bigcup_{i \geq 1} A_i\right).$$

(iv) Si $(A_i)_{i \geq 1}$ est une suite décroissante d'éléments de $\mathcal{A}$ (*i.e.* $A_i \supset A_{i+1}$ pour tout i), alors la suite $(\mu(A_i))_i$ converge dans $[0, +\infty]$ et

$$\mu(A_1) < +\infty \Rightarrow \lim_{i \rightarrow +\infty} \mu(A_i) = \mu\left(\bigcap_{i \geq 1} A_i\right).$$

Démonstration. (i) Pour la première partie de l'assertion, on complète les parties $A_1, \dots, A_n$ avec $A_i = \emptyset$ pour $i > n$, de sorte que l'on peut appliquer le deuxième axiome de la définition à la suite $(A_i)_{i \in \mathbf{N}}$. Démontrons la seconde partie de l'assertion dans le cas $n = 2$ (le cas général s'en déduit par récurrence) :

$$B_1 \cup B_2 = (B_1 \setminus B_2) \cup (B_2 \setminus B_1) \cup (B_1 \cap B_2)$$

et les 3 parties dont on fait la réunion dans le membre de droite sont dans $\mathcal{A}$ et deux à deux disjointes. La première partie de (i) affirme donc :

$$\begin{aligned}\mu(B_1 \cup B_2) &= \mu(B_1 \setminus B_2) + \mu(B_2 \setminus B_1) + \mu(B_1 \cap B_2) \\ &\leq (\mu(B_1 \setminus B_2) + \mu(B_1 \cap B_2)) + (\mu(B_2 \setminus B_1) + \mu(B_1 \cap B_2)) \\ &= \mu(B_1) + \mu(B_2).\end{aligned}$$

(ii) Comme A et B sont éléments de $\mathcal{A}$, alors $B \setminus A$ est encore élément de $\mathcal{A}$ et on applique (i) avec $n = 2$, $A_1 = A$, $A_2 = B \setminus A$. La positivité de μ permet de conclure.

(iii) La convergence de $(\mu(A_i))_i$ est justifiée par sa croissance dans $[0, +\infty]$. Notons ensuite $B_1 = A_1$, et $B_i = A_i \setminus A_{i-1}$ si $i \geq 2$. Les B_i sont des éléments de $\mathcal{A}$ et par construction, A_n est la réunion disjointe des B_k pour $1 \leq k \leq n$. Par conséquent

$$\mu\left(\bigcup_{i \geq 1} A_i\right) = \mu\left(\bigcup_{k \geq 1} B_k\right) = \sum_{k \geq 1} \mu(B_k) = \lim_{N \rightarrow +\infty} \sum_{k=1}^N \mu(B_k) = \lim_{N \rightarrow +\infty} \mu(A_N)$$

où l'on a utilisé le point (i) du théorème et le point (ii) de la définition.

(iv) L'idée est proche de celle utilisée pour (iii). Posons $A := \bigcap_{i \geq 1} A_i$ et $C_i := A_i \setminus A_{i+1}$, pour $i \geq 1$. On a $A \in \mathcal{A}$, $C_i \in \mathcal{A}$ pour tout i et $A_n = A \cup \left(\bigcup_{i \geq n} C_i\right)$. Comme A est disjoint de tous les C_i et que les C_i sont eux-mêmes deux à deux disjoints, on déduit :

$$\mu(A_n) = \mu(A) + \sum_{i \geq n} \mu(C_i) \quad \text{pour } n \geq 1.$$

On a $A \subset A_1$ et $\mu(A_1) < +\infty$. L'égalité ci-dessus pour $n = 1$ montre que $\sum_{i \geq 1} \mu(C_i) < +\infty$. En reprenant cette même égalité et en faisant cette fois tendre n vers $+\infty$, on voit que

$$\mu(A_n) \rightarrow \mu(A) + \lim_{n \rightarrow +\infty} \sum_{i \geq n} \mu(C_i) = \mu(A).$$

□

Remarque 2.4. On peut se convaincre de la nécessité de l'hypothèse $\mu(A_1) < +\infty$ dans (iv) ci-dessus, en considérant l'exemple de l'espace mesurable $(\mathbf{N}, \mathcal{P}(\mathbf{N}))$ muni de la mesure de dénombrement $\#$ (i.e. $\#A$ est le cardinal de A). Bien sûr pour tout $n \geq 1$, on a $\#\{n, n+1, n+2, \dots\} = +\infty$. Par ailleurs $\bigcap_{n \geq 1} \{n, n+1, n+2, \dots\} = \emptyset$, qui est de mesure nulle.

Un autre exemple fondamental de mesure est donné par la *mesure de Lebesgue* sur $\mathbf{R}^n$, que nous allons maintenant considérer. Soit $n \in \mathbf{N}_{>0}$ et soit $\mathcal{B}(\mathbf{R}^n)$ la tribu borélienne de $\mathbf{R}^n$.

Étant donné $B \in \mathcal{B}(\mathbf{R}^n)$ et $x \in \mathbf{R}^n$, on définit $x + B := \{x + y; y \in B\}$ et on vérifie que $x + B \in \mathcal{B}(\mathbf{R}^n)$.

Définition 2.5. Une mesure μ sur $\mathcal{B}(\mathbf{R}^n)$ est *invariante par translation* si

$$(\forall B \in \mathcal{B}(\mathbf{R}^n)) (\forall x \in \mathbf{R}^n) \quad \mu(B) = \mu(x + B).$$

La mesure de Lebesgue λ_n sur $(\mathbf{R}^n, \mathcal{B}(\mathbf{R}^n))$ est caractérisée par les propriétés suivantes.

Théorème 2.6. La *mesure de Lebesgue* λ_n est l'unique mesure définie sur $\mathcal{B}(\mathbf{R}^n)$ satisfaisant :

- (i) λ_n est invariante par translation.
- (ii) $\lambda_n([0, 1]^n) = 1$, où l'on note $[0, 1]^n = \{(a_1, \dots, a_n); 0 \leq a_i \leq 1 \ (\forall i \in \llbracket 1, n \rrbracket)\}$.

Le théorème est admis, il y a deux types de preuve : l'une constructive par Carathéodory et une autre d'analyse fonctionnelle.

La mesure de Lebesgue λ_1 doit être considérée comme une extension de la notion de longueur, λ_2 comme une extension de la notion d'aire et λ_3 comme une extension de la notion de volume.

Le théorème suivant découle du précédent.

Théorème 2.7. La mesure de Lebesgue λ_n d'un pavé (ouvert, fermé ou semi-ouvert...) est le produit des longueurs de ses côtés i.e.

$$(\forall d \in \mathbf{N}_{>0}) (\forall i \in \llbracket 1, d \rrbracket) (-\infty < a_i < b_i < +\infty) \quad \lambda_d\left(\prod_{i=1}^d [a_i, b_i]\right) = \prod_{i=1}^d \lambda_d([a_i, b_i]).$$

À titre d'exemple, notons que $\lambda_1(\mathbf{Q}) = 0$.

3 Fonctions mesurables

Faisons d'abord de simples rappels et notations : soit $f: X \rightarrow Y$ une application d'un ensemble X vers un ensemble Y . Pour tout $B \subset Y$, on définit l'*image réciproque* de B par f :

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

On rappelle qu'il s'agit là d'une simple notation qui n'implique pas en général que f est inversible. Bien sûr, on a $f^{-1}(Y) = X$, $f^{-1}(\emptyset) = \emptyset$ et on rappelle les formules valables pour tout $A, B \subset Y$:

$$f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$$

$$f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$$

$$f^{-1}(A^c) = (f^{-1}(A))^c.$$

Si I est un ensemble quelconque d'indices, les deux premières formules se généralisent à $\bigcap_{i \in I} A_i$ et $\bigcup_{i \in I} A_i$, où $A_i \subset Y$ pour tout i .

Définition 3.1. Soit $(X, \mathcal{A}_X)$ et $(Y, \mathcal{A}_Y)$ des espaces mesurables et soit $f: X \rightarrow Y$. On dit que f est $(\mathcal{A}_X, \mathcal{A}_Y)$ -*mesurable* (ou simplement mesurable s'il n'y a pas d'ambiguïté sur les tribus concernées) si on a

$$(B \in \mathcal{A}_Y) \implies f^{-1}(B) \in \mathcal{A}_X.$$

Remarque 3.2. (1) Cette notion ne fait pas intervenir de mesures.

(2) Pour les images directes, on n'a pas égalité en général.

Exemple 3.3. (1) Si $(X, \mathcal{A}_X)$ est un espace mesurable, alors $\text{Id}: X \rightarrow X$, définie par $\text{Id}(x) = x$, est mesurable.

(2) Soit $(X, \mathcal{A})$ un espace mesurable et soit $A \subset X$. On note $\mathbf{1}_A$ la *fonction indicatrice* de A définie par

$$\mathbf{1}_A: X \rightarrow \mathbf{R}$$

$$x \mapsto \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A \end{cases}$$

Remarque 3.4. (1) En topologie, on a une caractérisation des fonctions continues, qui est aussi basée sur les préimages. Si $(X, \mathcal{T})$ et $(Y, \mathcal{T}')$ sont des espaces topologiques et $f: (X, \mathcal{T}) \rightarrow (Y, \mathcal{T}')$, alors :

$$f \text{ continue} \iff (\forall B \in \mathcal{T}') f^{-1}(B) \in \mathcal{T}.$$

(2) Il y a une notion d'ensemble mesurable (les éléments d'une tribu) à ne pas confondre avec la notion de fonction mesurable.

Proposition 3.5. L'application $\mathbf{1}_A$ est mesurable si et seulement si A est mesurable.

Démonstration. Les préimages possibles par $\mathbf{1}_A$ d'un ensemble $B \in \mathcal{B}(\mathbf{R})$ sont :

- $\emptyset$ si $B \cap \{0, 1\} = \emptyset$;
- A si $1 \in B$ et $0 \notin B$;
- A^c si $1 \notin B$ et $0 \in B$;
- X si $\{0, 1\} \subset B$.

Si A est mesurable, i.e. $A \in \mathcal{A}$, alors comme $\mathcal{A}$ est stable par complémentaire, on a aussi que $A^c \in \mathcal{A}$. Par ailleurs, $\mathcal{A}$ contient aussi $\emptyset$ et X puisque c'est une tribu. Ainsi, les 4 préimages possibles sont dans la tribu et f est bien mesurable.

Réciproquement, si $\mathbf{1}_A$ est mesurable, alors on a par définition pour tout $B \in \mathcal{B}(\mathbf{R})$, $\mathbf{1}_A^{-1}(B) \in \mathcal{A}$, donc en particulier, pour $B = \{1\}$, on a $\mathbf{1}_A^{-1}(\{1\}) = A \in \mathcal{A}$. On a donc bien que A est mesurable. $\square$

Énonçons maintenant quelques propriétés de stabilité des fonctions mesurables.

Proposition 3.6. Soit $(X, \mathcal{A}_X)$, $(Y, \mathcal{A}_Y)$ et $(Z, \mathcal{A}_Z)$ des espaces mesurables. Si $f: X \rightarrow Y$ et $g: Y \rightarrow Z$ sont mesurables, alors la composition $g \circ f: X \rightarrow Z$ est mesurable.

Démonstration. Soit $C \in \mathcal{A}_Z$. On a directement $(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C)) \in \mathcal{A}_X$. $\square$

Proposition 3.7. On considère $(X, \mathcal{A}_X)$ et $(Y, \mathcal{A}_Y)$ deux espaces mesurables. Soit $f: X \rightarrow Y$ une application (que l'on ne suppose pas nécessairement mesurable). Si on définit

$$f_* \mathcal{A}_X := \{B \subset Y; f^{-1}(B) \in \mathcal{A}_X\}$$

qui est une partie de $\mathcal{P}(Y)$, alors

- (i) $f_* \mathcal{A}_X$ est une tribu sur Y ;
- (ii) une application $f: X \rightarrow Y$ est $(\mathcal{A}_X, \mathcal{A}_Y)$ -mesurable si et seulement si $\mathcal{A}_Y \subset f_* \mathcal{A}_X$;
- (iii) $f_* \mathcal{A}_X$ est la plus grande tribu $\mathcal{S}$ sur Y telle que f soit $(\mathcal{A}_X, \mathcal{S})$ -mesurable de $(X, \mathcal{A}_X)$ vers $(Y, \mathcal{S})$.

Remarque 3.8. La notion de fonction mesurable est sensible aux choix des deux tribus.

La proposition, dont la preuve est à chercher, montre que l'on peut construire une tribu sur l'espace d'arrivée, qui est faite « sur mesure » pour f , *i.e.* pour la rendre mesurable.

Remarque 3.9. En calcul différentiel, il existe aussi une notion de poussée en avant de forme différentielle notée $f_* d\omega$

Proposition 3.10. Soit $(X, \mathcal{A}_X)$ un espace mesurable, Y un ensemble, $\mathcal{S}$ un ensemble de partie de Y et $\mathcal{A}_Y$ la tribu sur Y engendrée par $\mathcal{S}$. Une fonction $f: X \rightarrow Y$ est $(\mathcal{A}_X, \mathcal{A}_Y)$ -mesurable si et seulement si

$$(\forall B \in \mathcal{S}) \quad f^{-1}(B) \in \mathcal{A}_X.$$

Démonstration. Seule l'implication réciproque nécessite une explication. On a $\mathcal{S} \subset f_* \mathcal{A}_X$ par hypothèse, donc comme $f_* \mathcal{A}_X$ est une tribu sur Y , on déduit que $\mathcal{A}_Y \subset f_* \mathcal{A}_X$. Ainsi f est $(\mathcal{A}_X, \mathcal{A}_Y)$ -mesurable. $\square$

Définition 3.11. Soit $(X, \mathcal{U}_X)$ et $(Y, \mathcal{U}_Y)$ des espaces topologiques et soit $\mathcal{B}_X, \mathcal{B}_Y$ les tribus boréliennes respectivement associées à $\mathcal{U}_X$ et $\mathcal{U}_Y$. On dit que $f: X \rightarrow Y$ est *borélienne* si f est $(\mathcal{B}_X, \mathcal{B}_Y)$ -mesurable.

En particulier, si l'image de f est incluse dans $\mathbf{R}$, alors f est borélienne si et seulement si f est $(\mathcal{A}, \mathcal{B}(\mathbf{R}))$ -mesurable.

Théorème 3.12. On a les deux propriétés suivantes, la première étant un cas particulier de la seconde :

(i) Soit $n \in \mathbf{N}_{>0}$. Si $f: \mathbf{R}^n \rightarrow \overline{\mathbf{R}}$ est une fonction continue, alors f est borélienne *i.e.* f est $(\mathcal{B}(\mathbf{R}^n), \mathcal{B}(\overline{\mathbf{R}}))$ -mesurable.

(ii) Soit $(X, \mathcal{U}_X)$ et $(Y, \mathcal{U}_Y)$ des espaces topologiques et soit $\mathcal{B}_X, \mathcal{B}_Y$ les tribus boréliennes respectivement associées à $\mathcal{U}_X, \mathcal{U}_Y$. Si $f: X \rightarrow Y$ est continue, alors f est $(\mathcal{B}_X, \mathcal{B}_Y)$ -mesurable.

Remarque 3.13. La réciproque à cette implication est fautive en général ; par exemple l'indicatrice $1_{]0,1[}: \mathbf{R} \rightarrow \mathbf{R}$ de l'intervalle $]0,1[$, est borélienne mais discontinue en 0 et 1.

Démonstration. (i) Soit I un intervalle ouvert de $\mathbf{R}$. Si $f^{-1}(I) \neq \emptyset$ (qui est le seul cas intéressant), alors fixons $x_0 \in f^{-1}(I)$. On a $f(x_0) \in I$, donc pour un $\varepsilon > 0$, l'un des intervalles

$$I_\varepsilon :=]f(x_0) - \varepsilon, f(x_0) + \varepsilon[\quad I_\varepsilon := [-\infty, f(x_0) + \varepsilon[\quad I_\varepsilon :=]f(x_0) - \varepsilon, +\infty[$$

est inclus dans I . Comme f est continue, il existe $\eta > 0$ tel que si $x \in \mathbf{R}^n$ vérifie $\|x - x_0\| < \eta$, alors $f(x) \in I_\varepsilon$ (où l'on a fixé une norme quelconque $\|\cdot\|$ sur $\mathbf{R}^n$). Ainsi la boule ouverte de centre x_0 et de rayon η est incluse dans $f^{-1}(I)$, donc $f^{-1}(I)$ est un ouvert de $\mathbf{R}^n$ *i.e.* un élément de $\mathcal{B}(\mathbf{R}^n)$.

(ii) On utilise également la proposition : si $\mathcal{U}$ est un ouvert de Y , alors comme f est continue, on a $f^{-1}(U) \in \mathcal{U}_X \subset \mathcal{B}_X$. Ainsi, f est $(\mathcal{B}_X, \mathcal{B}_Y)$ -mesurable. $\square$

Proposition 3.14. Soit $(X, \mathcal{A})$ un espace mesurable. Pour tout $k \geq 1$, on équipe $\mathbf{R}^k$ de sa tribu borélienne $\mathcal{B}(\mathbf{R}^k)$. Soit $n \in \mathbf{N}_{>0}$ et soit

$$f = (f_1, \dots, f_n): X \rightarrow \mathbf{R}^n \\ x \mapsto (f_1(x), \dots, f_n(x))$$

une application. La fonction f est mesurable si et seulement si chaque f_i est mesurable.

Démonstration. • Pour tout $1 \leq i \leq n$, notons $\pi_i: \mathbf{R}^n \rightarrow \mathbf{R}$ la projection sur la i -ème coordonnée. L'application π_i est continue donc borélienne. Ainsi, $f_i = \pi_i \circ f$ est $(\mathcal{A}, \mathcal{B}(\mathbf{R}))$ -mesurable.

• Réciproquement, si chaque f_i est mesurable, on voit qu'en fixant deux n -uplets de réels $(x_1, \dots, x_n)$ et $(y_1, \dots, y_n)$ avec $x_j < y_j$ pour tout j , on a :

$$f^{-1}(D((x_i)_i, (y_i)_i)) = \bigcap_{k=1}^n f_k^{-1}(]x_k, y_k[).$$

La mesurabilité de f se déduit alors de celle des f_i . $\square$

Proposition 3.15. Soit $(X, \mathcal{A})$ un espace mesurable et soit $u, v: X \rightarrow \mathbf{R}$ deux applications mesurables ($\mathbf{R}$ est muni de sa tribu de Borel). Pour toute application $\phi: \mathbf{R}^2 \rightarrow \mathbf{R}$ borélienne (en particulier, pour tout ϕ continue), l'application

$$h: X \rightarrow \mathbf{R} \\ x \mapsto \phi(u(x), v(x))$$

est mesurable.

On peut maintenant énoncer le théorème suivant qui est constamment utilisé dans les exercices pratiques.

Théorème 3.16. Soit $(X, \mathcal{A})$ un espace mesurable. On munit $\mathbf{R}$ et $\overline{\mathbf{R}}$ de leurs tribus boréliennes respectives.

(i) Si $f, g: X \rightarrow \mathbf{R}$ sont mesurables, alors les applications :

$$f + g \quad f \times g \quad \max(f, g) \quad \min(f, g) \quad |f|$$

sont mesurables.

(ii) Si $f_k: X \rightarrow \overline{\mathbf{R}}$ est le terme général d'une suite $(f_k)_{k \geq 1}$ de fonctions mesurables, alors les applications suivantes de X dans $\overline{\mathbf{R}}$:

$$\inf_k f_k \quad \sup_k f_k \quad \limsup_k f_k \quad \liminf_k f_k$$

sont mesurables. En particulier, si $\lim_k f_k(x)$ existe pour tout $x \in X$, alors la fonction $f: X \rightarrow \overline{\mathbf{R}}$ définie pour tout $x \in X$ par $f(x) = \lim_k f_k(x)$ est également mesurable.

Démonstration. Pour (i) il suffit d'appliquer la proposition 3.15 avec une fonction ϕ continue bien choisie. La preuve de (ii) est laissée en exercice. $\square$

4 Fonctions simples et leur intégrale de Lebesgue

On va voir que la définition de l'intégrale de Lebesgue d'une fonction indicatrice d'une partie mesurable est naturelle. On cherche donc, pour définir plus généralement l'intégrale des fonctions mesurables positives, à les approcher par des « combinaisons linéaires » d'indicatrices de parties mesurables. Ce point de vue motive l'introduction des fonctions simples.

Définition 4.1. Soit X un ensemble. Une fonction $s : X \rightarrow \mathbf{R}$ est une *fonction simple* si s ne prend qu'un nombre fini de valeurs (i.e. $s(X)$ est une partie finie de $\mathbf{R}$).

Proposition 4.2. Si s est simple, alors il existe un unique $n \in \mathbf{N}_{>0}$, $(\alpha_1, \dots, \alpha_n) \in \mathbf{R}^n$, $(A_i)_{1 \leq i \leq n} \in \mathcal{P}(X)$ forment une partition de X tels que

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

Démonstration. Soit n le nombre de valeurs et notons $\alpha_1, \dots, \alpha_n$ ces n valeurs distinctes. Notons pour $1 \leq i \leq n$, $A_i := s^{-1}(\{\alpha_i\})$. Ainsi définit, les $(A_i)_i$ partitionnent X . Pour tout $i = 1, \dots, n$, $x \in A_i$, on a $s(x) = \alpha_i$ et $\sum_{j=1}^n \alpha_j \mathbf{1}_{A_j}(x) = \alpha_i$. Ainsi, l'égalité de la proposition est vraie. $\square$

Définition 4.3. On appelle *écriture canonique* d'une fonction simple la décomposition

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

Proposition 4.4. Si X est un ensemble et $(B_i)_{i=1, \dots, n} \in \mathcal{P}(X)$, $(\beta_i)_{i=1, \dots, n} \in \mathbf{R}$, alors

$$s = \sum_{i=1}^n \beta_i \mathbf{1}_{B_i}$$

est simple.

Démonstration. La fonction s prend un nombre fini de valeurs, incluses dans

$$\left\{ \sum_{i \in I} \beta_i \mid I \subset \llbracket 1, n \rrbracket \right\}$$

qui est fini. $\square$

L'écriture $\sum_{i=1}^n \beta_i \mathbf{1}_{B_i}$ n'est pas, en général, la décomposition canonique.

Exemple 4.5. On a $\mathbf{1}_{[0, +\infty[} + \mathbf{1}_{]-\infty, 0[} = \mathbf{1}_{\mathbf{R}}$. À gauche, l'écriture est non canonique. Plusieurs écritures sont possibles mais une seule est canonique.

Proposition 4.6. La somme et le produit de deux fonctions simples sont simples.

Démonstration. Si s et t prennent un nombre fini de valeurs, disons respectivement $\{\alpha_i; i = 1, \dots, n\}$ et $\{\beta_j; j = 1, \dots, m\}$, alors $s + t$ et st prennent des valeurs qui sont incluses respectivement dans $\{\alpha_i + \beta_j; (1 \leq i \leq n) (1 \leq j \leq m)\}$ et $\{\alpha_i \beta_j; (1 \leq i \leq n) (1 \leq j \leq m)\}$. $\square$

De plus, si $s = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i}$ et $t = \sum_{j=1}^m \beta_j \mathbf{1}_{B_j}$ sont les décompositions canoniques de s et t , alors

$$s + t = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} (\alpha_i + \beta_j) \mathbf{1}_{A_i \cap B_j} \quad \text{et} \quad st = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} \alpha_i \beta_j \mathbf{1}_{A_i \cap B_j}.$$

On peut même remarquer que

$$\max(s, t) = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} \max(\alpha_i, \beta_j) \mathbf{1}_{A_i \cap B_j} \quad \text{et} \quad \min(s, t) = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} \min(\alpha_i, \beta_j) \mathbf{1}_{A_i \cap B_j}.$$

Ces propriétés de stabilités par $+$, $\times$, $\max$, $\min$, se généralisent au cas d'un nombre fini de fonctions simples :

Exemple 4.7. Si pour tout $k = 1, \dots, K$,

$$s^{(k)} = \sum_{i=1}^{n_k} \alpha_i^{(k)} \mathbf{1}_{A_i^{(k)}}$$

est une écriture canonique, alors

$$\sum_{k=1}^K s^{(k)} = \sum_{k=1}^K \sum_{i=1}^{n_k} \alpha_i^{(k)} \mathbf{1}_{\bigcap_{1 \leq i \leq K} A_i^{(k)}}.$$

Proposition 4.8. Soit $(X, \mathcal{A})$ un espace mesurable et $s: X \rightarrow \mathbf{R}$ une fonction simple. La fonction s est $(\mathcal{A}, \mathcal{B}(\mathbf{R}))$ -mesurable si et seulement si il existe $n \in \mathbf{N}_{>0}$, des parties deux à deux disjointes $A_i \in \mathcal{A}$ et des réels α_i ($1 \leq i \leq n$) tels que l'on ait $s = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i}$.

Remarque 4.9. Soit X un ensemble à deux éléments : $X = \{a, b\}$ muni de la tribu grossière : $\mathcal{A} = \{\emptyset, X\}$. On voit que $s = \mathbf{1}_{\{a\}} + \mathbf{1}_{\{b\}}$ est mesurable alors que ni $\{a\}$, ni $\{b\}$ ne l'est. Ceci est dû au fait que ce n'est pas la décomposition canonique, qui est ici $s = \mathbf{1}_X$.

Démonstration. • Supposons que s est mesurable. Comme les singletons sont des fermés de $\mathbf{R}$, en particulier les $\{\alpha_i\}$ et donc leurs préimages $A_i = s^{-1}(\{\alpha_i\})$ sont dans $\mathcal{A}$.

• Réciproquement, si les $(A_i)_i$ sont dans $\mathcal{A}$, alors pour tout $I \subset \{1, \dots, n\}$, $\bigcup_{i \in I} A_i \in \mathcal{A}$ (stable par union dénombrable de $\mathcal{A}$). Or l'ensemble $\mathcal{P}$ des préimages par s est justement

$$\mathcal{P} = \left\{ \bigcup_{i \in I} A_i \mid I \subset \{1, \dots, n\} \right\}.$$

Ainsi, $\mathcal{P} \subset \mathcal{A}$. □

Définition 4.10. Soit $(X, \mathcal{A}, \mu)$ un espace mesuré et $E \in \mathcal{A}$. Si $s: X \rightarrow \mathbf{R}_{\geq 0}$ est une fonction mesurable simple écrite sous la forme

$$s = \sum_{i=1}^r \alpha_i \mathbf{1}_{A_i}$$

avec $\alpha_i \in \mathbf{R}_{\geq 0}$ et $A_i \in \mathcal{A}$, pour tout i , alors l'intégrale de Lebesgue de s sur E est définie comme étant :

$$\int_E s \, d\mu := \sum_{i=1}^r \alpha_i \mu(A_i \cap E).$$

Lemme 4.11. Soit $(X, \mathcal{A}, \mu)$ un espace mesuré et soit $s, t: X \rightarrow \mathbf{R}_{\geq 0}$ des fonctions simples mesurables. On a les propriétés suivantes.

(i) Pour tout $E \in \mathcal{A}$,

$$\int_E (s + t) \, d\mu = \int_E s \, d\mu + \int_E t \, d\mu.$$

(ii) Soit $(E_i)_{i \geq 1}$ une suite d'éléments deux à deux disjointes de $\mathcal{A}$. On a alors

$$\int_E s \, d\mu = \sum_{n=1}^{+\infty} \int_{E_n} s \, d\mu$$

avec $E := \bigcup_{k=1}^{+\infty} E_k$.

Démonstration. (i) On a $s + t = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} (\alpha_i + \beta_j) \mathbf{1}_{A_i \cap B_j}$ et donc

$$\begin{aligned} \int_E (s + t) \, d\mu &= \sum_{i,j} (\alpha_i + \beta_j) \mu(E \cap (A_i \cap B_j)) \\ &= \sum_{i,j} \alpha_i \mu(E \cap A_i \cap B_j) + \sum_{i,j} \beta_j \mu(E \cap A_i \cap B_j). \end{aligned}$$

Mais $\sum_{i,j} \alpha_i \mu(E \cap A_i \cap B_j) = \sum_{i,j} \alpha_i (\sum_j \mu(E \cap A_i \cap B_j)) = \sum_{i,j} \alpha_i \mu(\bigsqcup_j E \cap A_i \cap B_j)$. De plus, $\bigsqcup_j E \cap A_i \cap B_j = E \cap A_i \cap \bigsqcup_j B_j = E \cap A_i$.

Ainsi,

$$\sum_i \alpha_i \mu(E \cap A_i) = \int_E s \, d\mu.$$

On fait de même avec la deuxième somme et on a

$$\sum_j \beta_j \mu(E \cap B_j) = \int_E t \, d\mu$$

d'où l'égalité annoncée.

(ii) On a :

$$\begin{aligned} \int_{\bigcup_{n \in \mathbf{N}} E_n} s \, d\mu &= \sum_i \alpha_i \mu\left(\left(\bigcup_{n \in \mathbf{N}} E_n\right) \cap A_i\right) \\ &= \sum_i \alpha_i \sum_{n \in \mathbf{N}} \mu(E_n \cap A_i) \\ &= \sum_{n \in \mathbf{N}} \int_{E_n} s \, d\mu. \end{aligned}$$

□

5 Intégrale de Lebesgue de fonctions mesurables positives

On va maintenant définir l'intégrale des fonctions mesurables positives à partir de la définition de l'intégrale des fonctions mesurables positives simples.

Définition 5.1. Soit $(X, \mathcal{A}, \mu)$ un espace mesuré et $E \in \mathcal{A}$. Soit $f: X \rightarrow [0, +\infty]$ une fonction mesurable (on munit $[0, +\infty]$ de sa tribu borélienne). L'intégrale de Lebesgue de f sur E est définie comme étant :

$$\int_E f \, d\mu := \sup_{s \leq f} \int_E s \, d\mu$$

où le sup est pris sur les fonctions mesurables simples positives $s: X \rightarrow \mathbf{R}_{\geq 0}$ telles que $s(x) \leq f(x)$ pour tout $x \in X$.

On peut aisément vérifier que la définition de $\int_E s \, d\mu$ ne dépend que de s et pas du choix des réels α_i ou des parties mesurables A_i .

Remarque 5.2. En conservant les notations de la définition, on se donne $E \in \mathcal{A}$. Soit f une application définie sur E (mais pas, *a priori*, sur X) et à valeurs dans $[0, +\infty]$. On peut encore définir l'intégrale de f sur E : on note $\mathcal{A}|_E = \{A \in \mathcal{A}; A \subset E\}$ (dont on peut vérifier, en exercice, que c'est une tribu sur E) et $\mu|_E: \mathcal{A}|_E \rightarrow [0, +\infty]$ la restriction de μ à $\mathcal{A}|_E$. On a alors que $(E, \mathcal{A}|_E, \mu|_E)$ est un espace mesuré et si f est $(\mathcal{A}|_E, \mathcal{B}([0, +\infty]))$ -mesurable, alors on définit

$$\int_E f \, d\mu|_E.$$

Si maintenant, $g: X \rightarrow [0, +\infty]$ est $(\mathcal{A}, \mathcal{B}([0, +\infty]))$ -mesurable, alors $g|_E$ est $(\mathcal{A}|_E, \mathcal{B}([0, +\infty]))$ -mesurable (où $g|_E$ désigne la restriction de g à E) et on a

$$\int_E g \, d\mu = \int_E g|_E \, d\mu|_E.$$

On souhaite que la définition ci-dessus d'intégrale satisfasse un certain nombre de propriétés « intuitives », dont on sait par exemple qu'elles sont vraies dans le cas de l'intégrale de Riemann. On établit ci-dessous quelques unes de ces propriétés.

Théorème 5.3. Soit $(X, \mathcal{A}, \mu)$ un espace mesuré. Soit $f, g: X \rightarrow [0, +\infty]$ des fonctions mesurables et soit $E \in \mathcal{A}$.

(i) Si $f(x) \leq g(x)$ pour tout $x \in E$, alors

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

(ii) On a,

$$\int_E f \, d\mu = \int_X (f \times \mathbf{1}_E) \, d\mu.$$

(iii) Si $\mu(E) = 0$, alors

$$\int_E f \, d\mu = 0.$$

(iv) S'il existe des parties $E_0 \subset X$ et $E_1 \in \mathcal{A}$ telles que $E_0 \subset E_1$, $\mu(E_1) = 0$ et pour tout $x \in E \setminus E_0$, $f(x) = 0$, alors $\int_E f \, d\mu = 0$.

(v) Si $A \in \mathcal{A}$ et $A \subset E$, alors

$$\int_A f \, d\mu \leq \int_E f \, d\mu.$$

(vi) Pour tout $c \geq 0$, on a

$$\int_E (c \cdot f) \, d\mu = c \int_E f \, d\mu.$$

Démonstration. (i) Soit s simple positive mesurable telle que $s \leq f$ sur X , alors $s \cdot \mathbf{1}_E \leq g$ sur X et comme $s \cdot \mathbf{1}_E$ est simple, positive, mesurable, on déduit

$$\int_E s \, d\mu = \int_E s \cdot \mathbf{1}_E \, d\mu \leq \int_E g \, d\mu.$$

On conclut en passant au sup sur les fonctions s .

(ii) On a la série d'égalités :

$$\int_E f \, d\mu = \sup_{s \leq f} \int_E s \, d\mu = \sup_{s \leq f} \int_X s \cdot \mathbf{1}_E \, d\mu = \sup_{t \leq f \cdot \mathbf{1}_E} \int_X t \, d\mu = \int_X f \cdot \mathbf{1}_E \, d\mu.$$

(iii) Soit s positive, mesurable, simple telle que $s(x) \leq f(x)$ pour tout $x \in X$. La définition de l'intégrale des fonctions simples montre que $\int_E s \, d\mu = 0$ puisque $\mu(E) = 0$. On conclut par passage au sup sur s .

(iv) Soit $s: X \rightarrow \mathbf{R}_{\geq 0}$ simple mesurable telle que $s(x) \leq f(x)$ pour tout $x \in X$. Par hypothèse sur f , on a donc $s(x) = 0$ pour tout $x \in E \setminus E_0$. Dans l'écriture $s = \sum_{i=1}^r \alpha_i \mathbf{1}_{A_i}$, on peut prendre $\alpha_1 = 0$ (quitte à renuméroter les α_i). On a donc $E \setminus E_0 \subset A_1$ et $A_i \subset E_0 \subset E_1$ pour tout $i \geq 2$. En particulier, $A_i = A_i \cap E_1$ vérifie $\mu(A_i) = 0$, pour tout $i \geq 2$. Par croissance de μ , on a donc $\mu(A_i \cap E) = 0$ dès que $i \geq 2$ et on conclut :

$$\int_E s \, d\mu = \sum_{i=1}^r \alpha_i \mu(A_i \cap E) = 0 + \sum_{i=2}^r \alpha_i \mu(A_i \cap E) = 0.$$

On obtient l'assertion souhaitée en passant au sup sur les fonctions $s \leq f$.

Les assertions suivantes sont laissées en exercice. □

Dans le point (iv) du théorème apparaît la notion de partie négligeable, sur laquelle on reviendra longuement dans le chapitre suivant. On donne dès à présent la définition de telles parties de X .

Définition 5.4 (PARTIE NÉGLIGEABLE, FONCTION NULLE PRESQUE PARTOUT). Soit $(X, \mathcal{A}, \mu)$ un espace mesuré et soit $f: X \rightarrow \mathbf{R}$ une fonction mesurable.

(i) Une partie $E_0 \subset X$ est dite μ -négligeable s'il existe $E_1 \in \mathcal{A}$ tel que $E_0 \subset E_1$ et $\mu(E_1) = 0$.

(ii) On dit que f est *nulle μ -presque partout* s'il existe une partie μ -négligeable $E_0 \subset X$ telle que $f(x) = 0$ pour tout $x \in X \setminus E_0$.

En utilisant le vocabulaire introduit dans cette définition, on peut déduire du point (iv) du théorème que l'intégrale sur X d'une fonction mesurable positive et nulle presque partout vaut 0.

On souhaite aussi bien sûr que l'intégrale ainsi construite soit linéaire (*i.e.* que « l'intégrale de $f + g$ est la somme de l'intégrale de f et de l'intégrale de g »).

On va voir dans la section suivante que ce point n'est pas évident à établir et nécessite le recours à un résultat important du cours : le théorème de convergence monotone, ainsi que le résultat suivant qui précise l'idée suivant laquelle on peut approcher les fonctions mesurables positives par des fonctions simples mesurables.

Théorème 5.5 (APPROXIMATION DES FONCTIONS MESURABLES POSITIVES). Soit $(X, \mathcal{A})$ un espace mesurable et soit $f: X \rightarrow [0, +\infty]$. On munit $[0, +\infty]$ de sa tribu de Borel. La fonction f est mesurable si et seulement si il existe une suite $(s_i)_{i \geq 1}$ de fonctions simples mesurables sur X vérifiant,

$$(\forall n \in \mathbf{N}) \quad (\forall x \in X) \quad 0 \leq s_1(x) \leq s_2(x) \leq \dots \leq s_n(x) \leq f(x) \quad \text{et} \quad f(x) = \lim_{k \rightarrow +\infty} s_k(x).$$

Démonstration. • Pour le sens réciproque, fixons $x \in X$, la suite $(s_n(x))_{n \geq 1}$ croît donc converge dans $[0, +\infty]$. D'après le théorème 3.16 du chapitre 3, la fonction $f: X \rightarrow [0, +\infty]$ définie pour tout $x \in X$ par

$$f(x) := \lim_{n \rightarrow +\infty} s_n(x)$$

est mesurable (comme limite ponctuelle de fonctions mesurables).

• Pour le sens direct, on définit pour tout $n \in \mathbf{N}_{>0}$:

$$\phi_n = \left(\sum_{k=0}^{n2^n-1} \frac{k}{2^n} \cdot \mathbf{1}_{\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right]} \right) + n \cdot \mathbf{1}_{[n, +\infty]}.$$

D'après la proposition précédente, chaque $\phi_n: [0, +\infty] \rightarrow \mathbf{R}$ est une fonction simple mesurable (relativement aux tribus de Borel de $[0, +\infty]$ et $\mathbf{R}$), positive et l'on vérifie que la suite de terme général $s_n = \phi_n \circ f: X \rightarrow \mathbf{R}$ convient. □

6 Théorème de convergence monotone de Lebesgue

Un ingrédient permettant d'établir la linéarité de l'intégrale des fonctions mesurables positives, qui est aussi l'un des résultats les plus importants de ce cours est le théorème suivant.

Théorème 6.1 (CONVERGENCE MONOTONE DE LEBESGUE). Soit $(X, \mathcal{A}, \mu)$ un espace mesuré et soit $f_n : X \rightarrow [0, +\infty]$ le terme général d'une suite de fonctions mesurables telles que,

$$(\forall n \geq 1) (\forall x \in X) \quad f_n(x) \leq f_{n+1}(x)$$

alors la fonction

$$\begin{aligned} f : X &\rightarrow [0, +\infty] \\ x &\mapsto \lim_{n \rightarrow +\infty} f_n(x) \end{aligned}$$

est mesurable. De plus, la suite $(\int_X f_n d\mu)_{n \geq 1}$ converge et

$$\lim_{n \rightarrow +\infty} \int_X f_n d\mu = \int_X f d\mu.$$

Démonstration. Notons

$$I := \int_X f d\mu$$

et pour tout $n \in \mathbf{N}$,

$$I_n := \int_X f_n d\mu.$$

Comme la suite de fonction (f_n) est croissante et majorée par f , la suite $(I_n)_n$ (qui est à valeurs dans $[0, +\infty]$) est croissante, donc sa limite (qui est éventuellement infinie) existe et cette suite est majorée par I , de sorte que l'on a

$$\lim_{n \rightarrow +\infty} I_n \leq I.$$

Pour démontrer l'inégalité dans l'autre sens, considérons une fonction simple mesurable s qui vérifie $0 \leq s \leq f$. Soit c une constante strictement inférieure à 1. Posons pour tout $n \in \mathbf{N}$,

$$X_n := \{x \in X; f_n(x) \leq c \cdot s(x)\}.$$

Les ensembles X_n sont mesurables, car les fonctions $(f_n)_n$ et s le sont et pour tout $n \in \mathbf{N}$, l'ensemble X_n peut donc se réécrire comme la préimage d'un borélien par une application mesurable. De plus, la suite $(X_n)_n$ est croissante car la suite $(f_n)_n$ est croissante. Enfin, $\bigcup_{n \in \mathbf{N}} X_n = X$. Pour tout $x \in X$, la suite $(f_n(x))_n$ converge simplement vers $f(x)$ et $f(x) \geq s(x) > c \cdot s(x)$. En utilisant la définition de l'intégrale d'une fonction simple, on en déduit

$$\lim_{n \rightarrow +\infty} \int_{X_n} s d\mu = \int_X s d\mu.$$

Par ailleurs, on a

$$I_n \geq \int_{X_n} f_n d\mu \geq c \int_{X_n} s d\mu.$$

En prenant la limite des deux côtés, on en déduit

$$\lim_{n \rightarrow +\infty} I_n \geq c \int_X s d\mu.$$

En faisant tendre c vers 1, on en déduit que

$$\lim_{n \rightarrow +\infty} I_n \geq \int_X s d\mu.$$

Comme cette inégalité est vraie pour toute fonction simple positive mesurable s majorée par f , on obtient bien que

$$\lim_{n \rightarrow +\infty} I_n \geq I.$$

□

Le théorème de convergence monotone combiné au théorème 5.5 du chapitre 5 et au lemme 3.16 du chapitre 4 permet d'obtenir finalement les propriétés générales de linéarité souhaitées.

Théorème 6.2 (LINÉARITÉ DE L'INTÉGRALE DES FONCTIONS POSITIVES). Soit $(X, \mathcal{A}, \mu)$ un espace mesuré. On a les propriétés suivantes :

(i) Si $f, g: X \rightarrow [0, +\infty]$ sont mesurables, alors pour tout $E \in \mathcal{A}$,

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

(ii) Soit $f_n: X \rightarrow [0, +\infty]$ le terme général d'une suite de fonctions mesurables. La fonction

$$\begin{aligned} f: X &\rightarrow [0, +\infty] \\ x &\mapsto \sum_{n=1}^{+\infty} f_n(x) \end{aligned}$$

est mesurable et pour tout $E \in \mathcal{A}$, on a

$$\int_E f d\mu = \sum_{n=1}^{+\infty} \left(\int_E f_n d\mu \right).$$

(iii) Si $f: X \rightarrow [0, +\infty]$ est mesurable et si $(E_i)_{i \geq 1}$ est une suite d'éléments deux à deux disjoints de $\mathcal{A}$, alors

$$\int_E f d\mu = \sum_{n=1}^{+\infty} \int_{E_n} f d\mu$$

avec $E := \bigcup_{k=1}^{\infty} E_k$.

Démonstration. (i) On a vu qu'il existe une suite croissante $(s_n)_n$ de fonctions simples mesurable qui converge simplement vers f et de même, une suite de fonctions simples mesurables qui converge simplement vers g , alors $(s_n + t_n)_n$ est une suite de fonctions simples mesurables qui converge simplement vers $f + g$. On a que pour tout n ,

$$\int_X (s_n + t_n) d\mu = \int_X s_n d\mu + \int_X t_n d\mu.$$

On applique le théorème de convergence monotone aux 3 intégrales pour arriver à

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

(ii) On applique le théorème de convergence monotone à la suite des sommes partielles $s_N(x) = \sum_{n=0}^N f_n(x)$. Cela donne

$$\lim_N \int_X s_N d\mu = \int_X (\lim_N s_N) d\mu$$

soit

$$\lim_N \sum_{n=0}^N \int_X f_n d\mu = \int_X \left(\sum_{n=0}^{+\infty} f_n \right) d\mu$$

en étendant par récurrence la linéarité de l'intégrale à un nombre finie arbitraire de fonctions. Ainsi,

$$\sum_{n \geq 0} \int_X f_n d\mu = \int_X \left(\sum_{n \geq 0} f_n \right) d\mu.$$

(iii) Il suffit d'appliquer le théorème de convergence monotone à $f_N := \sum_{n=0}^N \mathbf{1}_{A_n}$. □

Théorème 6.3 (LEMME DE FATOU). Soit $(X, \mathcal{A}, \mu)$ un espace mesuré et $f_n: X \rightarrow [0, +\infty]$ le terme général d'une suite de fonctions mesurables. On a alors

$$\int_X \liminf_{n \rightarrow +\infty} f_n d\mu \leq \liminf_{n \rightarrow +\infty} \int_X f_n d\mu.$$

Démonstration. Notons $g_n := \inf_{k \geq n} f_k$. C'est une suite croissante de fonction mesurables, donc par le théorème de convergence monotone,

$$\lim_n \int_X g_n \, d\mu = \int_X \lim_n g_n \, d\mu.$$

On a pour tout n , $g_n \leq f_n$, donc $\int_X g_n \, d\mu \leq \int_X f_n \, d\mu$. On en déduit donc de l'inégalité précédente que

$$\liminf_n \int_X g_n \, d\mu \leq \liminf_n \int_X f_n \, d\mu.$$

Ici, comme la suite $(g_n)_n$ est croissante, $\lim_n \int_X g_n \, d\mu$ existe, donc

$$\liminf_n \int_X g_n \, d\mu = \lim_n \int_X g_n \, d\mu$$

d'où le résultat. □

Remarque 6.4. On peut en déduire une autre inégalité avec la limite sup, mais le sens de l'inégalité change !

7 Théorème de convergence dominée de Lebesgue

On fixe un espace mesuré $(X, \mathcal{A}, \mu)$.

Définition 7.1. Une fonction $f: X \rightarrow \mathbf{R}$ est μ -intégrable (on dira simplement intégrable lorsqu'il n'y a pas d'ambiguïté sur l'espace mesuré ambiant) si

(i) f est $(\mathcal{A}, \mathcal{B}(\mathbf{R}))$ -mesurable,

(ii) $\int_X |f| d\mu < +\infty$.

On note

$$\mathcal{L}^1(\mu) := \mathcal{L}^1(X, \mathcal{A}, \mu) := \{f: X \rightarrow \mathbf{R}; f \text{ est } \mu\text{-intégrable}\}.$$

Pour $E \in \mathcal{A}$, on définit pour tout $f \in \mathcal{L}^1(\mu)$,

$$\int_E f d\mu := \int_E f^+ d\mu - \int_E f^- d\mu$$

où $f^+(x) = \max\{f(x), 0\}$ et $f^-(x) = \max\{-f(x), 0\}$ pour tout $x \in X$.

On observe que pour toute fonction f , on a les décompositions suivantes :

$$f = f^+ - f^- \quad |f| = f^+ + f^-$$

et les propriétés :

$$f^+ \geq 0 \quad f^- \geq 0 \quad \text{supp } f^+ \cap \text{supp } f^- = \emptyset.$$

Remarque 7.2. En utilisant les points (i) et (v) du théorème 5.3 du chapitre 5, on voit que

$$\int_E f^+ d\mu \leq \int_X |f| d\mu < +\infty \quad \text{et} \quad \int_E f^- d\mu \leq \int_X |f| d\mu < +\infty$$

puisque $f \in \mathcal{L}^1(\mu)$. Cela permet de justifier que $\int_E f d\mu$ est définie ci-dessus comme la différence de deux quantités finies.

Théorème 7.3 (PROPRIÉTÉS DE L'INTÉGRALE DE LEBESGUE). On conserve les notations ci-dessus.

(i) L'ensemble $\mathcal{L}^1(\mu)$ est un $\mathbf{R}$ -espace vectoriel et pour tout $E \in \mathcal{A}$, l'application

$$\begin{aligned} \mathcal{L}^1(\mu) &\rightarrow \mathbf{R} \\ f &\mapsto \int_E f d\mu \end{aligned}$$

est linéaire. En d'autre termes, pour tout $f, g \in \mathcal{L}^1(\mu)$ et tout $c \in \mathbf{R}$, on a :

$$\int_E (f + g) d\mu = \int_E f d\mu + \int_E g d\mu \quad \text{et} \quad \int_E (cf) d\mu = c \int_E f d\mu.$$

(ii) Pour tout $f, g \in \mathcal{L}^1(\mu)$ et tout $E \in \mathcal{A}$:

$$(\forall x \in E) f(x) \leq g(x) \Rightarrow \int_E f d\mu \leq \int_E g d\mu.$$

(iii) Pour tout $f \in \mathcal{L}^1(\mu)$ et tout $E \in \mathcal{A}$:

$$\left| \int_E f d\mu \right| \leq \int_E |f| d\mu$$

(iv) Pour tout $f \in \mathcal{L}^1(\mu)$ et toute suite $(E_i)_{i \geq 1}$ d'éléments de $\mathcal{A}$ deux à deux disjoints :

$$\int_E f d\mu = \sum_{k \geq 1} \int_{E_k} f d\mu \quad \text{avec} \quad E = \bigcup_{k \geq 1} E_k.$$

(v) Pour tout $E \in \mathcal{A}$ et tout $f \in \mathcal{L}^1(\mu)$:

$$\int_E f d\mu = \int_X f \cdot \mathbf{1}_E d\mu.$$

(vi) Soit $E \in \mathcal{A}$. Pour tout $f \in \mathcal{L}^1(\mu)$, si l'une des deux conditions suivantes est vérifiée :

- $\mu(E) = 0$;
 - $f(x) = 0$ pour tout $x \in E \setminus E_0$ où $E_0 \subset X$ vérifie $E_0 \subset E_1$ pour un $E_1 \in \mathcal{A}$ tel que $\mu(E_1) = 0$;
- alors on a :

$$\int_E f d\mu = 0.$$

On peut maintenant énoncer le résultat qui est probablement le plus important de tout le cours : le théorème de convergence dominée de Lebesgue.

Théorème 7.4 (CONVERGENCE DOMINÉE DE LEBESGUE). Soit $(X, \mathcal{A}, \mu)$ un espace mesuré et soit

$$g: X \rightarrow [0, +\infty[, \quad f_n: X \rightarrow \mathbf{R}$$

avec $n \geq 1$ des fonctions μ -intégrables satisfaisant les deux hypothèses suivantes :

- *hypothèse de domination* : $(\forall x \in X) |f_n(x)| \leq g(x)$,
- *hypothèse de convergence simple* : la suite $(f_n(x))_{n \geq 1}$ converge simplement vers une limite $f(x)$ pour tout $x \in X$.

On a alors,

(i) la fonction $f: X \rightarrow \mathbf{R}$ ainsi définie et les $(f_n)_{n \geq 1}$ sont dans $\mathcal{L}^1(\mu)$.

(ii) lorsque $n \rightarrow +\infty$,

$$\int_X |f_n - f| d\mu \rightarrow 0.$$

(iii) lorsque $n \rightarrow +\infty$,

$$\int_X f_n d\mu \rightarrow \int_X f d\mu.$$

Démonstration. (i) La fonction f en tant que limite simple de fonctions mesurables est mesurable. De plus, en passant à la limite dans l'hypothèse de domination, on obtient pour tout $x \in X$, $|f(x)| \leq g(x)$. En intégrant, cela donne

$$\int_X |f| d\mu \leq \int_X g d\mu$$

et donc f est dans $\mathcal{L}^1(\mu)$.

(ii) On considère la suite de fonctions auxiliaires $h_n := 2g - |f_n - f|$, qui est une suite de fonctions mesurables positives, par l'inégalité triangulaire. Par le lemme de Fatou, on a donc

$$\int_X \liminf_{n \rightarrow +\infty} h_n d\mu \leq \liminf_{n \rightarrow +\infty} \int_X h_n d\mu.$$

De plus, avec l'hypothèse de convergence simple, on a

$$\liminf_{n \rightarrow +\infty} h_n = 2g$$

et donc

$$\int_X \liminf_{n \rightarrow +\infty} h_n d\mu = \int_X 2g d\mu.$$

D'autre part,

$$\liminf_{n \rightarrow +\infty} \int_X h_n d\mu = \int_X 2g d\mu - \limsup_{n \rightarrow +\infty} \int_X |f_n - f| d\mu.$$

On a donc

$$\limsup_{n \rightarrow +\infty} \int_X |f_n - f| d\mu \leq 0.$$

Puisque $|f_n - f|$ est positive, cela implique que lorsque $n \rightarrow +\infty$, on a

$$\int_X |f_n - f| d\mu \rightarrow 0.$$

(iii) Par l'inégalité triangulaire, on a

$$\left| \int_X f_n d\mu - \int_X f d\mu \right| \leq \int_X |f_n - f| d\mu$$

ce qui implique que lorsque $n \rightarrow +\infty$,

$$\int_X f_n d\mu \rightarrow \int_X f d\mu.$$

□

8 Ensembles négligeables et espace $L^1(\mu)$

On fixe un espace mesuré $(X, \mathcal{A}, \mu)$. On revient dans cette section sur la définition qui va nous permettre de définir un quotient de l'espace $\mathcal{L}^1(\mu)$ qui a la propriété d'être un $\mathbf{R}$ -espace normé complet (autrement dit un espace de Banach).

Définition 8.1 (PROPRIÉTÉ VRAIE μ -PRESQUE PARTOUT). Soit $\mathcal{P}$ une propriété relative aux éléments $x \in X$ (penser par exemple à une inégalité $f(x) \geq g(x)$ entre fonctions mesurables, à une convergence en x d'une suite de fonctions mesurables, etc...). On dit que $\mathcal{P}$ est vérifiée μ -presque partout sur X s'il existe $N \in \mathcal{A}$ tel que

- $\mu(N) = 0$,
- $\mathcal{P}$ est vérifiée en tout $x \in X \setminus N$.

Exemple 8.2. La fonction indicatrice des rationnels $\mathbf{1}_{\mathbf{Q}}$ est nulle λ -presque partout, où λ est la mesure de Lebesgue sur $\mathbf{R}$.

On notera que l'on n'exige pas dans cette définition que $\{x \in X; \mathcal{P} \text{ est vérifiée en } x\}$ soit un élément de $\mathcal{A}$.

Comme illustration importante des notions ci-dessus, on a la propriété suivante de finitude presque partout des fonctions mesurables positives.

Lemme 8.3. Soit $f: X \rightarrow [0, +\infty]$ une fonction mesurable telle que $\int f d\mu < +\infty$. Si $N := \{x \in X; f(x) = +\infty\}$, alors $\mu(N) = 0$.

Observons ensuite que deux fonctions positives mesurables coïncidant μ -presque partout sur X ont même intégrale sur tout $E \in \mathcal{A}$.

Lemme 8.4. Soit $f, g: X \rightarrow [0, +\infty]$ des fonctions mesurables (resp. $f, g: X \rightarrow \mathbf{R}$ sont éléments de $\mathcal{L}^1(\mu)$). Soit

$$N := \{x \in X; f(x) \neq g(x)\}.$$

Si $\mu(N) = 0$, alors pour tout $E \in \mathcal{A}$,

$$\int_E f d\mu = \int_E g d\mu.$$

La réciproque est fautive en général pour les fonctions mesurables positives. On peut en revanche prouver une forme de réciproque (le cas $g = 0$) à la fois pour les fonctions intégrables et pour les fonctions mesurables positives. C'est l'objet du lemme suivant :

Lemme 8.5. Si $f: X \rightarrow [0, +\infty]$ est une fonction mesurable ou si $f: X \rightarrow \mathbf{R}$ est un élément de $\mathcal{L}^1(\mu)$, alors les assertions suivantes sont équivalentes.

- (i) $f(x) = 0$ μ -presque partout ;
- (ii) $(\forall E \in \mathcal{A}) \int_E f d\mu = 0$;
- (iii) on a $\int_X |f| d\mu = 0$.

Démonstration. (i) $\Rightarrow$ (ii). Évident.

(ii) $\Rightarrow$ (iii). On utilise (ii) avec les ensembles

$$E^+ = f^{-1}([0, +\infty]) \quad \text{et} \quad E^- = f^{-1}(]-\infty, 0])$$

qui sont bien dans $\mathcal{A}$ comme préimage par des fonctions mesurables de boréliens. On observe que $f^\pm = f \times \mathbf{1}_{A^\pm}$ et on décompose (iii) pour conclure.

(iii) $\Rightarrow$ (i). On pose $g = f$ si f est mesurable positive et $g = |f|$ si $f \in \mathcal{L}^1(\mu)$. On définit pour $n \geq 1$:

$$A_n := \{x \in X; g(x) > 2^{-n}\}.$$

On a alors l'inégalité :

$$2^{-n} \mu(A_n) \leq \int_{A_n} g d\mu \leq \int_X g d\mu$$

et donc comme d'une part $\mu(A_n) \geq 0$ et que d'autre part on a supposé (iii), $\mu(A_n) = 0$ pour tout n . Notons

$$N := \{x \in X; g(x) \neq 0\}.$$

On a $N = \bigcup_{n \in \mathbf{N}} A_n$ et on en déduit $\mu(N) = 0$. On a donc bien (i). □

La notion de « propriété vraie μ -presque partout » permet souvent de renforcer des énoncés donnant des propriétés relatives à l'intégrale de fonctions plutôt qu'aux fonctions elles-mêmes. Pour illustrer ce principe, on donne la « forme forte » suivante :

Théorème 8.6 (CONVERGENCE DOMINÉE, FORME FORTE). Soit $(X, \mathcal{A}, \mu)$ un espace mesuré et soit

$$g: X \rightarrow \mathbf{R}_{>0}, \quad f_n: X \rightarrow \mathbf{R}, \quad (n \geq 1)$$

des fonctions μ -intégrables satisfaisant

(i) pour μ -presque tout $x \in X$, $|f_n(x)| \leq g(x)$,

(ii) la suite $(f_n(x))_{n \geq 1}$ admet une limite pour tout $x \in X \setminus N$, où $N \in \mathcal{A}$ vérifie $\mu(N) = 0$.

Si on définit $f: X \rightarrow \mathbf{R}$ par :

$$f(x) := \begin{cases} \lim_{n \rightarrow +\infty} f_n(x) & \text{si } x \notin N \\ 0 & \text{si } x \in N \end{cases}$$

alors $f \in \mathcal{L}^1(\mu)$ et pour tout $E \in \mathcal{A}$ on a :

$$\int_E f \, d\mu = \lim_{n \rightarrow +\infty} \int_E f_n \, d\mu.$$

Donnons maintenant la construction de l'espace $L^1(\mu)$ à partir de $\mathcal{L}^1(\mu)$. On conserve les notations ci-dessus et on définit sur $\mathcal{L}^1(\mu)$ la relation d'équivalence $\sim_\mu$:

$$(f \sim_\mu g) \Leftrightarrow (f = g \text{ } \mu\text{-presque partout}).$$

D'après le lemme 8.4, si $f \in \mathcal{L}^1(\mu)$ et $f \sim_\mu g$, alors $g \in \mathcal{L}^1(\mu)$. De plus :

$$V_0(\mu) = \{f \in \mathcal{L}^1(\mu); f \sim_\mu 0\} = \{f \in \mathcal{L}^1(\mu); f = 0 \text{ } \mu\text{-presque partout}\}$$

est un sous- $\mathbf{R}$ -espace vectoriel de $\mathcal{L}^1(\mu)$. La relation d'équivalence $\sim_\mu$ permet de définir l'espace $L^1(\mu)$.

Définition 8.7. Soit $(X, \mathcal{A}, \mu)$ un espace mesuré. On note $L^1(\mu)$ l'ensemble des classes d'équivalences de $\mathcal{L}^1(\mu)$ pour la relation $\sim_\mu$.

De manière équivalente, $L^1(\mu)$ est l'espace vectoriel quotient $\mathcal{L}^1(\mu)/V_0(\mu)$ et une classe d'équivalence de $\mathcal{L}^1(\mu)$ pour la relation $\sim_\mu$ (*i.e.* un élément de $L^1(\mu)$) est une partie de $\mathcal{L}^1(\mu)$ du type

$$f + V_0(\mu) := \{f + g; g \in V_0(\mu)\}$$

pour un $f \in \mathcal{L}^1(\mu)$. On notera alors $\bar{f} = f + V_0(\mu)$ l'élément de $L^1(\mu)$, dont f est un représentant dans $\mathcal{L}^1(\mu)$. Le $\mathbf{R}$ -espace vectoriel quotient $L^1(\mu)$ a des propriétés remarquables. Tout d'abord, il est naturellement muni d'une norme.

Proposition 8.8. L'application

$$\begin{aligned} L^1(\mu) &\rightarrow \mathbf{R}_{\geq 0} \\ f &\mapsto \|f\|_1 := \int_X |f| \, d\mu \end{aligned}$$

est une norme sur $L^1(\mu)$.

Démonstration. On a déjà vu que l'on a, même sur $\mathcal{L}^1$,

- $\|f + g\|_{L^1} \leq \|f\|_{L^1} + \|g\|_{L^1}$;
- $(\forall \lambda \in \mathbf{R}) \|\lambda f\|_{L^1} = |\lambda| \cdot \|f\|_{L^1}$.

De plus, si $\|f\|_{L^1} = 0$ *i.e.* $\int_X |f| \, d\mu = 0$, alors $f = 0$ p.p. *i.e.* $f \sim 0$, soit $f = 0$ dans L^1 . □

Proposition 8.9. Soit $(X, \mathcal{A}, \mu)$ un espace mesuré et soit $(f_n)_{n \geq 1}$ une suite de $\mathcal{L}^1(\mu)$; si

$$\sum_{n \geq 1} \int_X |f_n| \, d\mu < +\infty$$

alors il existe une partie $N \subset X$ de mesure nulle et une fonction $f \in \mathcal{L}^1(\mu)$ tels que

(i) $(\forall x \in X \setminus N) |f_n(x)| < +\infty$ et $f(x) = \sum_{n \geq 1} f_n(x)$;

(ii) $(\forall A \in \mathcal{A}) \int_A f \, d\mu = \sum_{n \geq 1} \int_A f_n \, d\mu$;

(iii) On a $\lim_{n \rightarrow +\infty} \int_X \left| f - \sum_{k=1}^n f_k \right| \, d\mu = 0$.

Démonstration. (i) Pour $x \in X$, on définit la fonction $\phi: X \rightarrow [0, +\infty]$ par

$$\phi(x) := \sum_{k=1}^{+\infty} |f_k(x)|.$$

La fonction ϕ est mesurable comme limite d'une suite de fonctions mesurables (la suite des sommes partielles associées à la suite $(f_n)_{n \geq 1}$). Par positivité de $|f_k|$, le théorème de convergence monotone s'applique et donne

$$\int_X \phi \, d\mu = \lim_{n \rightarrow +\infty} \int_X \sum_{k=1}^n |f_k| \, d\mu = \lim_{n \rightarrow +\infty} \sum_{k=1}^n \int_X |f_k| \, d\mu = \sum_{k=1}^{+\infty} \int_X |f_k| \, d\mu < +\infty$$

où la dernière inégalité provient de l'hypothèse faite dans la proposition. Ainsi, $\phi \in \mathcal{L}^1(\mu)$ et donc l'ensemble $N := \{x \in X; \phi(x) = +\infty\}$ est un élément de $\mathcal{A}$ de mesure nulle. Définissons $f: X \rightarrow \mathbf{R}$ par

$$f(x) = \begin{cases} 0 & \text{si } x \in N \\ \sum_{k=1}^{+\infty} f_k(x) & \text{si } x \in X \setminus N \end{cases}$$

L'inégalité triangulaire montre que $f \in \mathcal{L}^1(\mu)$, puisque $\phi \in \mathcal{L}^1(\mu)$.

(ii) Définissons pour cela la fonction $g: X \rightarrow \mathbf{R}$ et les fonctions $g_n: X \rightarrow \mathbf{R}$ pour $n \geq 1$ par

$$g := \phi \mathbf{1}_{X \setminus N}, \quad g_n = \left(\sum_{k=1}^n f_k \right) \mathbf{1}_{X \setminus N}.$$

Ces fonctions sont $(\mathcal{A}, \mathcal{B}(\mathbf{R}))$ -mesurables, comme sommes de produits de fonctions mesurables. Aussi, g est à valeurs positives et comme $\mu(N) = 0$, on a

$$\int_X g \, d\mu = \int_X \phi \, d\mu < +\infty$$

de sorte que $g \in \mathcal{L}^1(\mu)$. Comme $(g_n)_{n \geq 1}$ converge simplement vers f sur X , le théorème de convergence dominée affirme alors que $f \in \mathcal{L}^1(\mu)$ et que pour tout $A \in \mathcal{A}$,

$$\int_A f \, d\mu = \lim_{n \rightarrow +\infty} \int_A g_n \, d\mu = \lim_{n \rightarrow +\infty} \int_A \sum_{k=1}^n f_k \, d\mu = \sum_{n=1}^{+\infty} \int_A f_n \, d\mu$$

où la seconde égalité utilise à nouveau le fait que $\mu(N) = 0$.

(iii) Notons que l'égalité

$$f - \sum_{k=1}^n f_k = f - g_n$$

est vraie pour μ -presque tout $x \in X$. Aussi, $f(x) - g_n(x) \rightarrow 0$ lorsque n tend vers $+\infty$ pour tout $x \in X$. Enfin, $f - g_n \in \mathcal{L}^1(\mu)$ puisque $|f - g_n| \leq |f| + g$ et que cette dernière somme de fonctions est intégrable. L'application du théorème de convergence dominée donne alors :

$$\lim_{n \rightarrow +\infty} \int_X \left| f - \sum_{k=1}^n f_k \right| \, d\mu = \lim_{n \rightarrow +\infty} \int_X |f - g_n| \, d\mu = 0.$$

□

On conclut en donnant une propriété très importante de l'espace vectoriel normé $(L^1(\mu), \|\cdot\|_1)$.

Théorème 8.10. Si $(X, \mathcal{A}, \mu)$ est un espace mesuré et si $(f_n)_{n \geq 1}$ est une suite de Cauchy de $\mathcal{L}^1(\mu)$, alors :

- (i) il existe $f \in \mathcal{L}^1(\mu)$ et une suite extraite $(f_{\varphi(n)})_{n \geq 1}$ de (f_n) qui converge μ -presque partout vers f .
- (ii) la suite $(f_n)_{n \geq 1}$ (toute entière) converge vers f dans $\mathcal{L}^1(\mu)$, i.e. lorsque n tend vers $+\infty$:

$$\|f_n - f\|_1 := \int_X |f_n - f| \, d\mu \rightarrow 0.$$

- (iii) le $\mathbf{R}$ -espace vectoriel normé $(L^1(\mu), \|\cdot\|_1)$ est un espace de Banach.

Démonstration. (i) Comme la suite $(f_n)_{n \geq 1}$ est de Cauchy pour $\|\cdot\|_1$, il existe une application strictement croissante $\varphi: \mathbf{N}_{>0} \rightarrow \mathbf{N}_{>0}$ telle que pour tout $i \geq 1$:

$$\|f_{\varphi(i+1)} - f_{\varphi(i)}\|_1 < 2^{-i}.$$

Considérons alors la suite de terme général $g_n = f_{\varphi(n+1)} - f_{\varphi(n)}$, pour $n \geq 1$. C'est une suite de $\mathcal{L}^1(\mu)$ satisfaisant l'hypothèse de la proposition 8.9 d'après ce qui précède. En appliquant la proposition, on en déduit l'existence de $g \in \mathcal{L}^1(\mu)$ coïncidant μ -presque partout avec

$$\sum_{i \geq 1} f_{\varphi(i+1)} - f_{\varphi(i)}.$$

On pose alors $f := f_{\varphi(1)} + g$ et on voit donc que $f \in \mathcal{L}^1(\mu)$, car g et $f_{\varphi(1)} \in \mathcal{L}^1(\mu)$ et que la suite $(f_{\varphi(n)})_{n \geq 1}$ converge μ -presque partout vers f .

(ii) Soit $\varepsilon > 0$. D'après le point (iii) de la proposition 8.9, on a :

$$0 = \lim_{n \rightarrow +\infty} \left\| g - \sum_{i=1}^n f_{\varphi(i+1)} - f_{\varphi(i)} \right\|_1 = \lim_{n \rightarrow +\infty} \|g - f_{\varphi(n+1)} - f_{\varphi(n)}\|_1$$

Il existe donc un entier ℓ tel que pour tout $k \geq \ell$:

$$\|f_{\varphi(k)} - f\|_1 \leq \frac{\varepsilon}{2}.$$

Par ailleurs, quitte à augmenter la taille de ℓ , pour $n, m \geq \varphi(\ell)$ on a

$$\|f_n - f_m\|_1 \leq \frac{\varepsilon}{2}.$$

Par l'inégalité triangulaire, on en déduit pour tout $n \geq \varphi(\ell)$,

$$\|f_n - f\|_1 \leq \|f_n - f_{\varphi(\ell)}\|_1 + \|f_{\varphi(\ell)} - f\|_1 \leq \varepsilon.$$

Ainsi $\|f_n - f\|_1 \rightarrow 0$ lorsque n tend vers $+\infty$.

(iii) D'après la proposition 8.8, l'espace $(L^1(\mu), \|\cdot\|_1)$ est $\mathbf{R}$ -espace vectoriel normé et on vient de voir qu'il est complet. C'est donc un espace de Banach. $\square$

9 Lien avec l'intégrale de Riemann

Cette section explique succinctement en quoi l'intégration de Lebesgue est une généralisation de l'intégrale au sens de Riemann. Rappelons le cadre classique dans lequel la notion d'intégrale au sens de Riemann est en général développée.

Proposition-Définition 9.1. Fixons un intervalle $[a, b]$ de $\mathbf{R}$.

- (i) Une fonction $f: [a, b] \rightarrow \mathbf{R}$ est *continue par morceaux* si elle est bornée sur $[a, b]$ et si elle admet un nombre fini de discontinuités sur $[a, b]$.
- (ii) Pour tout $n \geq 1$ et tout $i \in \{0, \dots, n\}$, on note

$$a_i = a + i \frac{b-a}{n}.$$

Si $f: [a, b] \rightarrow \mathbf{R}$ est continue par morceaux alors la suite de terme général

$$\frac{b-a}{n} \sum_{i=0}^{n-1} f(a_i)$$

converge et l'intégrale de Riemann de f sur $[a, b]$ est par définition la limite de cette suite.

On montre que toute fonction continue par morceaux sur un intervalle $[a, b]$ est aussi intégrable au sens de Lebesgue et que les notions d'intégrales de Riemann et Lebesgue coïncident alors.

Proposition 9.2. Si $f: [a, b] \rightarrow \mathbf{R}$ est une fonction continue par morceaux, alors $f \in \mathcal{L}^1(\lambda_{[a,b]})$, où $\lambda_{[a,b]}$ est la restriction à $[a, b]$ de la mesure de Lebesgue λ sur $\mathbf{R}$. De plus les notions d'intégrale de Riemann et Lebesgue coïncident, *i.e.*

$$\int_{[a,b]} f \, d\lambda_{[a,b]} = \int_a^b f(x) \, dx.$$

Démonstration. • La mesurabilité de f relativement aux tribus boréliennes de $[a, b]$ et $\mathbf{R}$ est claire, puisque à un nombre fini de points près, on peut partitionner $[a, b]$ en un nombre fini d'intervalles ouverts, sur lesquels f se restreint en une fonction continue. L'intégrabilité de f sur $[a, b]$ s'en déduit du fait que f est bornée sur l'intervalle compact $[a, b]$.

• Soit $n \geq 1$. On note $(a_i)_{0 \leq i \leq n}$ la subdivision de $[a, b]$ définie dans la définition-proposition précédente. On considère la suite de fonctions de terme général $g_n: [a, b] \rightarrow \mathbf{R}$:

$$g_n(x) = \sum_{i=0}^{n-1} f(a_i) \mathbf{1}_{[a_i, a_{i+1}[}(x) + f(b) \mathbf{1}_{\{b\}}(x).$$

Chaque g_n est simple et la continuité par morceaux de f sur $[a, b]$ assure que $g_n(x) \rightarrow f(x)$ pour λ -presque tout $x \in [a, b]$. Enfin on a pour tout $n \geq 1$ et pour tout $x \in [a, b]$, la majoration :

$$|g_n(x)| \leq \sup_{[a,b]} |f|.$$

Du fait que f est continue par morceaux sur $[a, b]$, le membre de droite est fini et donc λ -intégrable sur $[a, b]$ qui est un intervalle fermé borné. Le théorème de convergence dominée de Lebesgue permet alors de conclure que $f \in \mathcal{L}^1(\lambda_{[a,b]})$ et que

$$\int_{[a,b]} g_n(x) \, d\lambda(x) \xrightarrow{n \rightarrow +\infty} \int_{[a,b]} f \, d\lambda.$$

Comme g_n est simple pour tout n , on voit que le membre de gauche vaut $\frac{b-a}{n} \sum_{i=0}^{n-1} f(a_i)$, ce qui termine la preuve. $\square$

10 Mesure produit, théorème de Fubini et Tonelli

On n'expose la théorie des *espaces mesurés produits* que pour la classe des espaces mesurés σ -finis, que l'on définit maintenant.

Définition 10.1. Soit $(X, \mathcal{A}, \mu)$ un espace mesuré. On dit que $(X, \mathcal{A}, \mu)$ est σ -fini s'il existe une suite $(A_n)_{n \in \mathbb{N}}$ de $\mathcal{A}$ telle que $\mu(A_n) < +\infty$ pour tout n et $X = \bigcup_{n \geq 0} A_n$.

Cette hypothèse est par exemple vérifiée quand $\mu(X) < +\infty$ (donc en particulier quand μ est une mesure de probabilité, i.e. $\mu(X) = 1$), quand $X = \mathbb{N}$ muni de la tribu discrète et de la mesure de comptage, ou quand $X = \mathbb{R}^n$ muni de la mesure de Lebesgue. La méthode de base pour calculer une intégrable d'une fonction de deux variables est de se ramener à des intégrales de fonctions d'une variable. Pour cela il nous faut d'abord expliquer comment on peut munir $X \times Y$ d'une structure d'espace mesuré quand X, Y sont tous les deux munis d'une telle structure.

Proposition-Définition 10.2. (i) Soit $(X, \mathcal{A}_1, \mu_1)$ et $(Y, \mathcal{A}_2, \mu_2)$ deux espaces mesurés σ -finis. On note $\mathcal{A}_1 \otimes \mathcal{A}_2$ la tribu sur $X \times Y$ engendrée par les parties de la forme $A \times B$, où $A \in \mathcal{A}_1$, $B \in \mathcal{A}_2$; on l'appelle *tribu produit* des tribus $\mathcal{A}_1$ et $\mathcal{A}_2$.

(ii) Il existe alors une unique mesure ν sur $\mathcal{A}_1 \otimes \mathcal{A}_2$ vérifiant $\nu(A \times B) = \mu_1(A)\mu_2(B)$ pour tout $A \in \mathcal{A}_1$ et tout $B \in \mathcal{A}_2$. Cette mesure est notée $\mu_1 \otimes \mu_2$.

(iii) L'espace mesuré $(X \times Y, \mathcal{A}_1 \otimes \mathcal{A}_2, \mu_1 \otimes \mu_2)$ est σ -fini.

On n'essaiera pas de rentrer dans le détail de la construction de cette mesure produit. Notons que, pour les tribus boréliennes des espaces $\mathbb{R}^d$, on a $\mathcal{B}(\mathbb{R}^{n+m}) = \mathcal{B}(\mathbb{R}^n) \otimes \mathcal{B}(\mathbb{R}^m)$ et que, si λ_d désigne la mesure de Lebesgue sur $\mathbb{R}^d$, alors on a toujours $\lambda_{n+m} = \lambda_n \otimes \lambda_m$. La mesure produit $\mu_1 \otimes \mu_2$ étant définie à partir de μ_1 et μ_2 , on s'attend à ce qu'il en aille de même de l'intégrale d'une fonction mesurable positive relativement à $\mu_1 \otimes \mu_2$. C'est effectivement le contenu des théorèmes de Fubini (cas des fonctions intégrables) et Tonelli (cas des fonctions mesurables positives).

Théorème 10.3 (TONELLI). Soit $(X_1, \mathcal{A}_1, \mu_1)$ et $(X_2, \mathcal{A}_2, \mu_2)$ deux espaces mesurés σ -finis. Si $f: X_1 \times X_2 \rightarrow [0, +\infty]$ est une fonction $\mathcal{A}_1 \otimes \mathcal{A}_2$ -mesurable, alors :

(i) $y \mapsto f(x, y)$ est une fonction $(\mathcal{A}_2, \mathcal{B}([0, +\infty]))$ -mesurable pour tout $x \in X_1$. De plus $x \mapsto \int_{X_2} f(x, y) d\mu_2(y)$ est une fonction $(\mathcal{A}_1, \mathcal{B}([0, +\infty]))$ -mesurable.

(ii) $x \mapsto f(x, y)$ est une fonction $(\mathcal{A}_1, \mathcal{B}([0, +\infty]))$ -mesurable pour tout $y \in X_2$. De plus $y \mapsto \int_{X_1} f(x, y) d\mu_1(x)$ est une fonction $(\mathcal{A}_2, \mathcal{B}([0, +\infty]))$ -mesurable.

(iii) On a :

$$\begin{aligned} \int_{X_1 \times X_2} f(x, y) d(\mu_1 \otimes \mu_2)(x, y) &= \int_{X_1} \left(\int_{X_2} f(x, y) d\mu_2(y) \right) d\mu_1(x) \\ &= \int_{X_2} \left(\int_{X_1} f(x, y) d\mu_1(x) \right) d\mu_2(y). \end{aligned}$$

Dans le cas particulier des fonctions boréliennes, on obtient :

Corollaire 10.4 (TONELLI POUR LES FONCTIONS BORÉLIENNES). Soit m, p deux entiers ≥ 1 , $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^p$ des boréliens et $f: U \times V \rightarrow [0, +\infty]$ une fonction borélienne. On a alors :

(i) les fonctions $x \mapsto f(x, y)$ et $y \mapsto \int_U f(x, y) d\lambda_m$, définies respectivement sur U et V , sont boréliennes.

(ii) les fonctions $y \mapsto f(x, y)$ et $x \mapsto \int_V f(x, y) d\lambda_p$, définies respectivement sur V et U , sont boréliennes.

(iii) on a :

$$\begin{aligned} \int_{U \times V} f d\lambda_{m+p} &= \int_{x \in U} \left(\int_{y \in V} f(x, y) d\lambda_p(y) \right) d\lambda_m(x) \\ &= \int_{y \in V} \left(\int_{x \in U} f(x, y) d\lambda_m(x) \right) d\lambda_p(y). \end{aligned}$$

En pratique dans les exercices, m et p vaudront le plus souvent 1 ou 2.

Exemple 10.5. Calculons l'aire du disque unité $D = \{(x, y) \in \mathbb{R}^2; x^2 + y^2 \leq 1\}$. Par définition, l'aire d'une

partie D est l'intégrale de la fonction caractéristique de D sur $\mathbf{R}^2$, on a donc

$$\begin{aligned} \text{aire}(D) &= \int_{\mathbf{R}^2} \mathbf{1}_D(x, y) \, d\lambda_2(x, y) \\ &= \int_{x \in [-1, 1]} \left(\int_{y \in [-\sqrt{1-x^2}, \sqrt{1-x^2}]} d\lambda(y) \right) d\lambda(x) \\ &= 2 \int_{x \in [-1, 1]} \sqrt{1-x^2} \, dx \\ &= 2 \int_{-\pi/2}^{\pi/2} \cos^2(t) \, dt = \pi \end{aligned}$$

où la dernière étape utilise la coïncidence des intégrales au sens de Riemann et de Lebesgue pour les fonctions continues bornées sur des intervalles bornés, puis un changement de variable au sens de l'intégrale de Riemann.

Dans le cas où f n'est pas à valeurs positives, on a d'abord besoin de s'assurer que f est intégrable, ce qui peut être fait en appliquant le théorème de Tonelli. On peut alors appliquer le théorème suivant valable pour toutes les fonctions intégrables.

Théorème 10.6 (FUBINI). Soit $(X_1, \mathcal{A}_1, \mu_1)$ et $(X_2, \mathcal{A}_2, \mu_2)$ deux espaces mesurés σ -finis. Si $f: X_1 \times X_2 \rightarrow \mathbf{R}$ est une fonction de $\mathcal{L}^1(\mu_1 \otimes \mu_2)$, alors :

- (i) $y \mapsto f(x, y)$ est une fonction de $\mathcal{L}^1(\mu_2)$ pour presque tout $x \in X_1$, et $x \mapsto \int_{X_2} f(x, y) \, d\mu_2(y)$ définit un élément de $L^1(\mu_1)$.
- (ii) $x \mapsto f(x, y)$ est une fonction de $\mathcal{L}^1(\mu_1)$ pour presque tout $y \in X_2$, et $y \mapsto \int_{X_1} f(x, y) \, d\mu_1(x)$ définit un élément de $L^1(\mu_2)$.
- (iii) On a :

$$\begin{aligned} \int_{X_1 \times X_2} f(x, y) \, d(\mu_1 \otimes \mu_2)(x, y) &= \int_{X_1} \left(\int_{X_2} f(x, y) \, d\mu_2(y) \right) d\mu_1(x) \\ &= \int_{X_2} \left(\int_{X_1} f(x, y) \, d\mu_1(x) \right) d\mu_2(y). \end{aligned}$$

Pour les fonctions boréliennes intégrables, on obtient :

Corollaire 10.7 (FUBINI POUR LES FONCTIONS BORÉLIENNES). Soit m, p deux entiers ≥ 1 , $U \subset \mathbf{R}^m, V \subset \mathbf{R}^p$ des boréliens, et $f: U \times V \rightarrow \mathbf{R}$ une fonction borélienne λ_{m+p} -intégrable. On a alors :

- (i) les fonctions $x \mapsto f(x, y)$ et $y \mapsto \int_U f(x, y) \, d\lambda_m$, définissent respectivement des éléments de $\mathcal{L}^1(U, \lambda_m)$ et $L^1(V, \lambda_p)$.
- (ii) les fonctions $y \mapsto f(x, y)$ et $x \mapsto \int_V f(x, y) \, d\lambda_p$, définissent respectivement des éléments de $\mathcal{L}^1(V, \lambda_p)$ et $L^1(U, \lambda_m)$.
- (iii) on a :

$$\begin{aligned} \int_{U \times V} f \, d\lambda_{m+p} &= \int_{x \in U} \left(\int_{y \in V} f(x, y) \, d\lambda_p(y) \right) d\lambda_m(x) \\ &= \int_{y \in V} \left(\int_{x \in U} f(x, y) \, d\lambda_m(x) \right) d\lambda_p(y). \end{aligned}$$

Remarque 10.8. Dans le cas particulier d'une fonction continue sur un fermé borné $U \times V$ de $\mathbf{R}^{m+p}$, on sait qu'elle est toujours intégrable sur $U \times V$; le théorème de Fubini s'applique alors automatiquement. En revanche il faut bien s'assurer que la fonction f est continue et surtout que le domaine d'intégration est fermé borné...

En pratique, les théorèmes de Tonelli et Fubini sont utilisés conjointement; dans le cas d'une fonction mesurable de plusieurs variables à valeurs réelles, Tonelli permet de montrer l'intégrabilité de la fonction, condition nécessaire à l'application de Fubini pour le calcul effectif de l'intégrale multiple.

11 Théorème fondamental de l'analyse et changement de variable

Dans les cours d'introduction à l'intégration au sens de Riemann, on voit en général très tôt le lien entre intégration et calcul de primitives. On va voir que ce lien se généralise au cadre de l'intégration au sens de Lebesgue.

Théorème 11.1 (FONDAMENTAL DE L'ANALYSE). Soit $f: [a, b] \rightarrow \mathbf{R}$ une fonction λ -intégrable, où λ désigne la mesure de Lebesgue sur $\mathbf{R}$. Considérons $F: [a, b] \rightarrow \mathbf{R}$ définie par

$$F(x) = \int_{[a,b]} f(t) \mathbf{1}_{[a,x]}(t) d\lambda(t)$$

alors F est dérivable λ -presque partout sur $[a, b]$ et pour λ -presque tout $x \in [a, b]$, on a

$$F'(x) = f(x).$$

Démonstration. Dans le cas où l'on suppose f continue, on peut raisonner comme suit. Notons $g(x, t) = f(t) \mathbf{1}_{[a,x]}(t)$ qui est une fonction mesurable (comme produit de fonctions mesurables) définie sur $[a, b] \times [a, b]$. Fixons un $x \in [a, b]$ et $\varepsilon > 0$. Par continuité de f en x , il existe $\alpha > 0$ tel que pour tout $h \in [0, \alpha[$, on a $|f(x+h) - f(x)| \leq \varepsilon$. En conservant ces notations on a donc :

$$\begin{aligned} \left| \frac{1}{h} (F(x+h) - F(x) - hf(x)) \right| &= \left| \frac{1}{h} \left(\int_{[a,b]} f(t) (\mathbf{1}_{[a,x+h]}(t) - \mathbf{1}_{[a,x]}(t)) d\lambda(t) - hf(x) \right) \right| \\ &= \left| \frac{1}{h} \int_{[a,b]} (f(t) - f(x)) \mathbf{1}_{[x,x+h]}(t) d\lambda(t) \right| \\ &\leq \left(\sup_{x \leq t \leq x+h} |f(t) - f(x)| \right) \frac{1}{h} \int_{[a,b]} \mathbf{1}_{[x,x+h]}(t) d\lambda(t) \leq \varepsilon. \end{aligned}$$

On a donc montré que F est dérivable en tout $x \in [a, b]$ et que $F'(x) = f(x)$.

Le cas général se démontre par un argument de densité, on peut en effet montrer que les fonctions continues sur $[a, b]$ sont denses dans $\mathcal{L}^1([a, b])$ pour la norme $\|\cdot\|_1$. $\square$

Théorème 11.2 (DIFFÉRENTIATION DE LEBESGUE). Si $\mathbf{R}^n$ est munit de la mesure de Lebesgue λ_n et $f \in \mathcal{L}^1(\mathbf{R}^n)$, alors presque partout en x on a :

$$\frac{1}{\lambda_n[B(x, r)]} \int_{B(x, r)} |f(y) - f(x)| d\lambda_n(y) \xrightarrow{r \rightarrow 0} 0.$$

« L'écart moyen au sens L^1 à la fonction tend vers 0 presque partout. »

On donne maintenant un énoncé du théorème de changement de variable dans le cadre de l'intégration de Lebesgue, qui généralise l'énoncé que l'on rappelle ci-dessous pour l'intégrale de Riemann en dimension 1.

Théorème 11.3. Soit $a < b$ des nombres réels et soit $\varphi: [a, b] \rightarrow \varphi([a, b])$ une fonction de classe $\mathcal{C}^1$ dont la dérivée ne s'annule pas sur $[a, b]$. Si f est une fonction définie sur $\varphi([a, b])$ et n'ayant qu'un nombre fini de discontinuités sur cet intervalle, alors

$$\int_{\varphi(a)}^{\varphi(b)} f(t) dt = \int_a^b (f \circ \varphi(x)) \varphi'(x) dx.$$

On a bien sûr $\varphi([a, b]) = [\varphi(a), \varphi(b)]$ (resp. $\varphi([a, b]) = [\varphi(b), \varphi(a)]$) si $\varphi'(x) > 0$ (resp. $\varphi'(x) < 0$) pour tout $x \in [a, b]$.

L'énoncé ci-dessous généralise le théorème précédent de deux manières : il est valable pour l'intégrale au sens de Lebesgue et il est énoncé pour toute dimension $d \geq 1$.

Théorème 11.4 (FORMULE DU CHANGEMENT DE VARIABLE). Soit $d \geq 1$ et soit U, V des ouverts de $\mathbf{R}^d$. Soit $\varphi: U \rightarrow V$ un difféomorphisme de classe $\mathcal{C}^1$ (i.e. une bijection différentiable de classe $\mathcal{C}^1$ dont la réciproque est différentiable de classe $\mathcal{C}^1$). Notons pour tout $x = (x_1, \dots, x_d) \in U$:

$$\varphi(x) = (\varphi_1(x_1, \dots, x_d), \dots, \varphi_d(x_1, \dots, x_d)).$$

Le déterminant jacobien de φ en $x \in U$ est :

$$\det(\text{Jac}(\varphi)) = \det \begin{pmatrix} \partial_{x_1} \varphi_1 & \cdots & \partial_{x_d} \varphi_1 \\ \vdots & & \vdots \\ \partial_{x_1} \varphi_m & \cdots & \partial_{x_d} \varphi_m \end{pmatrix}.$$

Si $f \in \mathcal{L}^1(V)$, alors la composée

$$f \circ \varphi: U \xrightarrow{\varphi} V \xrightarrow{f} \mathbf{R}$$

est éléments de $\mathcal{L}^1(U)$ et on a

$$\int_V f \, d\lambda_d = \int_U f \circ \varphi(x) |\det(\text{Jac}(\varphi))| \, d\lambda_d(x).$$

Exemple 11.5. Exemple fondamental : l'intégral de Gauss. Soit

$$I = \int_{\mathbf{R}} e^{-x^2} \, dx$$

on a alors

$$I^2 = \left(\int_{\mathbf{R}} e^{-x^2} \, dx \right) \left(\int_{\mathbf{R}} e^{-y^2} \, dy \right) = \int_{\mathbf{R}^2} e^{-(x^2+y^2)} \, dx \, dy \quad (\text{par Fubini-Tonelli}).$$

On va utiliser un changement de coordonnées, des coordonnées euclidiennes vers les coordonnées polaires : $x = r \cos(\theta)$, $y = r \sin(\theta)$. Posons $\Omega = \mathbf{R}^2$ et $\tilde{\Omega} = \mathbf{R}_+ \times [0, 2\pi[$. Soit

$$\begin{aligned} \varphi: \tilde{\Omega} &\rightarrow \Omega \\ (r, \theta) &\mapsto (r \cos(\theta), r \sin(\theta)) \end{aligned}$$

c'est une bijection d'inverse

$$(x, y) \mapsto (x^2 + y^2, \arctan\left(\frac{y}{x}\right))$$

On a

$$\text{Jac}(\varphi) = \begin{pmatrix} \cos(\theta) & -r \sin(\theta) \\ \sin(\theta) & r \cos(\theta) \end{pmatrix}$$

donc $\det(\text{Jac}(\varphi)) = r \cos^2(\theta) + r \sin^2(\theta) = r \geq 0$. On obtient que

$$I^2 = \int_{\tilde{\Omega}} r e^{-r^2} \, dr \, d\theta = \int_0^{+\infty} \left(\int_0^{2\pi} r e^{-r^2} \, d\theta \right) \, dr \quad (\text{Fubini-Tonelli}).$$

On a donc

$$I^2 = 2\pi \int_0^{+\infty} r e^{-r^2} \, dr.$$

Or, $-2r e^{-r^2} = \frac{d}{dr}(e^{-r^2}) \Leftrightarrow r e^{-r^2} = -\frac{1}{2} \frac{d}{dr}(e^{-r^2})$, d'où

$$\int_0^{+\infty} r e^{-r^2} \, dr = \left[-\frac{1}{2} e^{-r^2} \right]_0^{+\infty} = \frac{1}{2}.$$

On a donc $I^2 = 2\pi \times \frac{1}{2} = \pi$ et I étant positive, on en déduit que $I = \sqrt{\pi}$.

12 Intégrales à paramètre

Le cas d'étude typique auquel cette section s'intéresse est le suivant : soit E un borélien de $\mathbf{R}^n$ et soit I un intervalle de $\mathbf{R}$. On considère une application

$$\begin{aligned} f: E \times I &\rightarrow \mathbf{R} \\ (x, t) &\mapsto f(x, t) \end{aligned}$$

et on s'intéresse aux propriétés (continuité, dérivabilité, ...) de la fonction

$$\begin{aligned} F: I &\rightarrow \mathbf{R} \\ t &\mapsto \int_E f(x, t) d\lambda_n(x) \end{aligned}$$

(Pour que cette fonction soit bien définie, il faut pour tout $t \in I$, que $x \mapsto f(x, t)$ soit dans $\mathcal{L}^1(E)$, ou bien soit positive mesurable - auquel cas on doit en général autoriser des valeurs infinies pour F). Le *paramètre* dont il est question dans le titre de cette section est ici t .

On se concentre d'abord sur un résultat assurant la continuité de la fonction F ci-dessus (localement en un $t_0 \in I$, puis globalement en I).

Théorème 12.1 (CONTINUITÉ SOUS LE SYMBOLE $\int$). Soit $(X, \mathcal{A}, \mu)$ un espace mesuré, soit $E \in \mathcal{A}$ et soit $I \subset \mathbf{R}$ un intervalle. On considère une application $f: E \times I \rightarrow \mathbf{R}$. Soit $t_0 \in I$. Supposons que :

- (i) pour $t \neq t_0$ dans un voisinage de t_0 , la fonction définie par $x \mapsto f(x, t)$ est éléments de $\mathcal{L}^1(\mu|_E)$, où $\mu|_E$ désigne la restriction de la mesure μ à la partie mesurable E ;
- (ii) $\lim_{t \rightarrow t_0} f(x, t) = f(x, t_0)$ pour μ -presque tout $x \in E$;
- (iii) il existe $g: E \rightarrow \mathbf{R}_{\geq 0}$, élément de $\mathcal{L}^1(\mu|_E)$, $N \in \mathcal{A}$ satisfaisant $\mu(N) = 0$ et $\varepsilon > 0$ tels que

$$(\forall t \in I \setminus \{t_0\}) (|t - t_0| < \varepsilon \Rightarrow (|f(x, t)| \leq g(x); \forall x \in E \setminus N)).$$

On a alors que l'application définie par $x \mapsto f(x, t_0)$ définit un élément de $\mathcal{L}^1(\mu|_E)$ et on a la formule

$$\lim_{t \rightarrow t_0} \int_E f(x, t) d\mu(x) = \int_E f(x, t_0) d\mu(x).$$

Démonstration. Il s'agit d'une application du théorème de convergence dominée. On utilise l'interprétation séquentielle de la continuité : on introduit une suite $(t_n)_{n \geq 1}$ de I dont tous les termes sont distincts de t_0 et qui converge vers t_0 . On considère alors la suite dont le terme général est défini par $f_n(x) = f(x, t_n)$. $\square$

La conséquence suivante est évidente. Le résultat traduit le fait qu'une fonction est par définition, continue sur un intervalle si elle est continue en chacun de ses points.

Corollaire 12.2. On conserve les notations du théorème précédent. On suppose que :

- (i) pour $t \in I$, la fonction définie par $x \mapsto f(x, t)$ est élément de $\mathcal{L}^1(\mu|_E)$, où $\mu|_E$ désigne la restriction de la mesure μ à la partie mesurable E ;
- (ii) pour μ -presque tout $x \in E$, l'application $I \rightarrow \mathbf{R}$ définie par $t \mapsto f(x, t)$ est continue ;
- (iii) il existe $g: E \rightarrow \mathbf{R}_{\geq 0}$, élément de $\mathcal{L}^1(\mu|_E)$, telle que pour tout $t \in I$ et pour μ -presque tout $x \in E$, on a

$$|f(x, t)| \leq g(x).$$

On a alors que l'application définie sur I par $t \mapsto \int_E f(x, t) d\mu(x)$ est continue sur I .

On souhaite maintenant étudier la dérivabilité de la fonction F . L'analogue du théorème 12.1 est le suivant.

Théorème 12.3 (DÉRIVABILITÉ SOUS LE SYMBOLE $\int$). Soit $(X, \mathcal{A}, \mu)$ un espace mesuré, soit $E \in \mathcal{A}$ et soit $I \subset \mathbf{R}$ un intervalle ouvert. On considère une application $f: E \times I \rightarrow \mathbf{R}$. Soit $t_0 \in I$. Supposons que :

- (i) il existe $\varepsilon > 0$ tel que pour $t \in]t_0 - \varepsilon, t_0 + \varepsilon[\subset I$, la fonction définie par $x \mapsto f(x, t)$ est élément de $\mathcal{L}^1(\mu|_E)$; ainsi on a bien notre fonction $F:]t_0 - \varepsilon, t_0 + \varepsilon[\rightarrow \mathbf{R}$;
- (ii) pour $\varepsilon > 0$ comme dans (i), l'application $t \mapsto f(x, t)$ est dérivable sur $]t_0 - \varepsilon, t_0 + \varepsilon[$ pour μ -presque tout $x \in E$; on note $\frac{\partial f}{\partial t}(x, t)$ cette dérivée ;
- (iii) il existe $g: E \rightarrow \mathbf{R}_{\geq 0}$, élément de $\mathcal{L}^1(\mu|_E)$, $N \in \mathcal{A}$ satisfaisant $\mu(N) = 0$ et $\varepsilon > 0$ comme dans (i) et (ii) :

$$(\forall t \in I \setminus \{t_0\}) \quad |t - t_0| < \varepsilon \implies \left(\left| \frac{\partial f}{\partial t}(x, t) \right| \leq g(x) \quad (\forall x \in E \setminus N) \right).$$

On a alors que l'application donnée pour $t \in]t_0 - \varepsilon, t_0 + \varepsilon[$ définie par $x \mapsto \frac{\partial f}{\partial t}(x, t)$ définit un élément de $L^1(\mu|_E)$ et l'application F de (i) est dérivable en t_0 et

$$F'(t_0) = \int_E \frac{\partial f}{\partial t}(x, t_0) d\mu(x).$$

Démonstration. On procède de manière analogue à la preuve du théorème 12.1. □

De même que pour la continuité, on a la forme « globale » suivante du théorème précédent.

Corollaire 12.4. On conserve les notations du théorème précédent. On suppose que :

- (i) pour $t \in I$, la fonction définie par $x \mapsto f(x, t)$ est élément de $\mathcal{L}^1(\mu|_E)$, où $\mu|_E$ désigne la restriction de la mesure μ à la partie mesurable E ,
- (ii) l'application $t \mapsto f(x, t)$ est dérivable sur I pour μ -presque tout $x \in E$; on note $\frac{\partial f}{\partial t}(x, t)$ cette dérivée,
- (iii) il existe $g : E \rightarrow \mathbf{R}_{\geq 0}$, élément de $\mathcal{L}^1(\mu|_E)$, telle que pour tout $t \in I$ et pour μ -presque tout $x \in E$, on a

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

On a alors que l'application définie pour tout $t \in I$ par $x \mapsto \frac{\partial f}{\partial t}(x, t)$ est élément de $L^1(\mu|_E)$. De plus, l'application F de (i) est dérivable sur I et pour tout $t_0 \in I$,

$$F'(t_0) = \int_E \frac{\partial f}{\partial t}(x, t_0) d\mu(x).$$

Exemple 12.5. Application à l'étude de la fonction Gamma d'Euler :

$$(t > 0) \quad \Gamma(t) = \int_0^{+\infty} x^{t-1} e^{-x} dx.$$

Posons $f(x, t) = x^{t-1} e^{-x}$. On commence par observer que :

$$(\forall t > 0) \quad f(\cdot, t) \in L^1([0, +\infty[, \mathcal{B}([0, +\infty[, \lambda_1]).$$

De plus, $f(x, t) = e^{(t-1)\ln(x)-x}$ et pour tout $x \in]0, +\infty[$,

$$\frac{\partial f}{\partial t}(x, t) = (\ln(x)) e^{(t-1)\ln(x)-x} = \ln(x) x^{t-1} e^{-x}$$

À cause de la dégénérescence de $\frac{\partial f}{\partial t}$ en $t = 0$ (le terme en $\frac{1}{x}$, qui empêche d'isoler une fonction de domination $g_1 \in L^1$) on introduit $\delta > 0$ tel que $I_\delta =]\delta, +\infty[$ et $g_1(x) = |\ln(x)| x^{\delta-1} e^{-x}$ qui est bien L^1 et qui domine $\frac{\partial f}{\partial t}$ pour $t \in I_\delta$. Ainsi, par dérivation sous l'intégrale, on a

$$\Gamma'(t) = \int_0^{+\infty} \ln(x) x^{t-1} e^{-x} dx \quad (\forall t \in]\delta, +\infty[) \quad (\forall \delta > 0)$$

donc pour tout $t > 0$. En observant que pour tout $k \in \mathbf{N}_{>0}$, on a :

$$\frac{\partial^k f}{\partial t^k}(x, t) = (\ln(x))^k x^{t-1} e^{-x}$$

et donc, en posant

$$g_k(x) = (\ln(x))^k x^{\delta-1} e^{-x}$$

on a que $g_k \in L^1$ et que

$$\left| \frac{\partial^k f}{\partial t^k}(x, t) \right| \leq g_k(x) \quad (\forall t \in I_\delta).$$

Ainsi, la fonction Γ est $\mathcal{C}^\infty$ sur $\mathbf{R}_{>0}$ avec

$$(\forall t > 0) \quad (\forall k \in \mathbf{N}) \quad \Gamma^k(t) = \int_0^{+\infty} (\ln(x))^k x^{t-1} e^{-x} dx.$$