Intégrale de Lebesgue

T. Deheuvels

Table des matières

Introduction

Dans ce cours, deux notions essentielles et fondamentalement liées seront abordées : la notion de mesure, et la notion d'intégration. Ces notions ont en particulier une importance capitale dans les cours de probabilités et d'analyse fonctionnelle qui suivront.

Notion de mesure

La première question abordée est celle de la mesure : étant donné un ensemble E quelconque, comment donner un sens à la mesure d'un sous-ensemble de E ? Quelle "place" prend une partie de E dans E ? Si on considère une propriété particulière $\mathscr{P}$ sur l'ensemble E, on pourrait en répondant à ces questions mesurer l'ensemble des points de E qui vérifient $\mathscr{P}$. Si cet ensemble est de mesure nulle, on pourrait alors donner un sens à "la propriété $\mathscr{P}$ est vraie presque partout dans E".

À première vue, il semble naturel de définir une mesure sur E comme une application de $\mathscr{P}(E)$ dans $\overline{\mathbb{R}}_+$ qui vérifierait $\mu(\emptyset) = 0$, ainsi qu'une propriété d'additivité telle que : la mesure d'une union (dénombrable) d'ensembles deux à deux disjoints est la somme des mesures de ces ensembles. En fait, nous verrons qu'il s'avèrera utile de définir les mesures sur un sous-ensemble $\mathscr{A}$ de $\mathscr{P}(E)$. Afin que les propriétés de la mesure aient un sens, il convient que $\mathscr{A}$ ait une structure particulière. Dans la pratique, on impose que $\mathscr{A}$ contienne $\emptyset$ et soit stable par passage au complémentaire et par union dénombrable : c'est la notion de tribu.

On remarque immédiatement qu'on connait déjà des mesures :

– Le "cardinal" d'un ensemble : si $A \subset E$, $\mu(A)$ est donné par le nombre d'éléments de A si A est un ensemble fini, et $+\infty$ sinon. On appelle cette mesure la mesure de comptage. On se rend assez vite compte que cette mesure est beaucoup trop imprécise dès qu'on considère des ensembles non dénombrables : mesurer des ensembles de $\mathbb{R}$ de cette façon n'aurait pas beaucoup d'intérêt.

– Les probabilités : il s'agit des mesures décrites plus haut auxquelles on impose de plus que $\mu(\Omega) = 1$. On dit que ces mesures sont de masse 1.

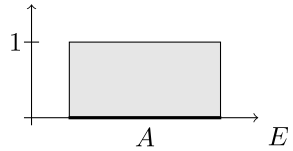
Nous pouvons bien sûr citer également le cas de $\mathbb{R}$, ou plus généralement de $\mathbb{R}^n$, pour lesquels nous souhaitons considérer la notion de longueur, ou de volume en toute généralité. Ces notions sont en général définies en partant de la longueur des intervalles dans $\mathbb{R}$, ou du volume des pavés dans $\mathbb{R}^n$. Elles ne sont toutefois pas définies *a priori* pour n'importe quel sous-ensemble de $\mathbb{R}$ ou de $\mathbb{R}^n$.

Si on s'intéresse au cas de $\mathbb{R}$, la question est de savoir s'il existe une mesure λ de $\mathscr{P}(\mathbb{R})$ dans $\overline{\mathbb{R}}_+$ telle que pour tout intervalle I d'extrémités $a, b \in \mathbb{R}$, $\lambda(I) = b - a$. Si on travaille avec les axiomes habituels de la théorie des ensembles, c'est-à-dire avec l'axiome du choix, la réponse à cette question est non ! Nous détaillerons ceci plus loin.

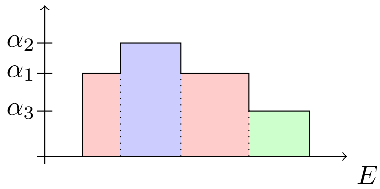
Nous allons donc définir une mesure qui vérifie ces propriétés sur une tribu incluse dans $\mathscr{P}(\mathbb{R})$. Pour assurer que nous pouvons mesurer assez de parties de $\mathbb{R}$, nous imposerons que cette tribu contienne au moins les ouverts de $\mathbb{R}$ (donc également les fermés). Nous consacrerons une partie de ce cours à construire proprement cette mesure, et nous nous interrogerons sur la tribu sur laquelle elle est définie. La construction proposée sera valable dans $\mathbb{R}^n$, ce qui définira également la notion de volume.

Lien avec l'intégration

Si on considère un ensemble E muni d'une mesure μ et un ensemble $A \in \mathscr{A}$, il semble naturel d'associer à la fonction $\mathbb{1}_A$ l'intégrale $1 \cdot \mu(A) = \mu(A)$:

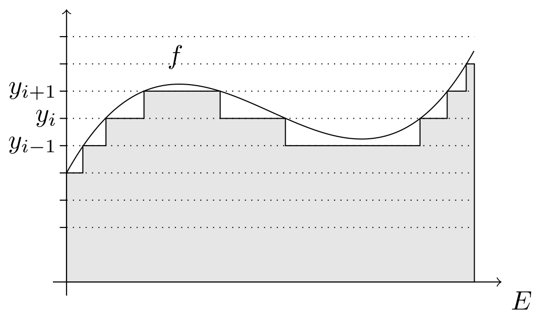


On note $\int \mathbb{1}_A \, \mathrm{d}\mu = \mu(A)$ pour préciser que l'intégration se fait par rapport à la mesure μ. Comme il est naturel de demander à l'intégrale une propriété de linéarité, on observe qu'on sait tout de suite donner un sens à l'intégrale de fonctions du type $\sum_{i=1}^N \alpha_i \mathbb{1}_{A_i}$, où $A_i \in \mathscr{A}$ pour tout i :



on doit avoir $\int \sum_{i=1}^N \alpha_i \mathbb{1}_{A_i} = \sum_{i=1}^N \alpha_i \mu(A_i)$. Les fonctions de ce type sont appelées les fonctions *étagées*. Elles jouent dans la construction de l'intégrale de Lebesgue un rôle similaire à celui des fonctions en escalier pour l'intégrale de Riemann. Il convient par ailleurs de noter que les fonctions en escalier sur un intervalle de $\mathbb{R}$ sont exactement les fonctions étagées telles que les A_i soient tous des intervalles.

Donnons maintenant une idée de la démarche employée pour donner un sens à l'intégrale d'une fonction $f : E \to \mathbb{R}$ qui n'est pas nécessairement étagée. Comme nous le verrons, sous une condition dite de *mesurabilité*, on peut approcher une fonction f par des fonctions étagées, que l'on sait intégrer. Une façon de le voir est de considérer une subdivision de l'espace d'arrivée : $\mathbb{R}$ (à comparer avec l'intégrale de Riemann, où on fait une subdivision de l'intervalle d'intégration) :



On approche alors f par la fonction étagée $\sum_i y_i \mathbb{1}_{f^{-1}([y_i, y_{i+1}[)}$, ce qui conduit à approcher l'intégrale de f par l'intégrale de cette fonction étagée. Ceci donne la valeur approchée de l'intégrale de f : $\sum_i y_i \mu(f^{-1}([y_i, y_{i+1}[))$. On note que pour que cette formule soit valide, il faut avoir $f^{-1}([y_i, y_{i+1}[) \in \mathscr{A}$ pour tout i. C'est une condition importante sur la fonction f, qui sera assurée par la propriété de mesurabilité.

Il faut noter que la théorie de l'intégration de Lebesgue généralise la théorie d'intégration de Riemann : tous les résultats propres à l'intégrale de Riemann restent vrai pour l'intégrale de Lebesgue. Nous vérifierons ce résultat après la construction de l'intégrale de Lebesgue. En outre, la théorie de Lebesgue présente de nombreux avantages :

- On peut intégrer davantage de fonctions. Par exemple, la fonction $\mathbb{1}_{\mathbb{Q}}$ dont on attendrait qu'elle soit intégrable et que son intégrale vaille 0, n'est pas Riemann-intégrable. Elle est intégrable au sens de Lebesgue, et son intégrale est nulle.
- Les théorèmes de convergence sont vrais sous des hypothèses plus faibles. Par ailleurs, ils s'énoncent plus simplement.
- L'intégrale de Lebesgue permet d'écrire les sommes comme des intégrales (par rapport à la mesure de comptage). Par conséquent, tous les résultats vrais pour les intégrales le sont immédiatement pour les séries.
- Le plus gros avantage de la théorie de Lebesgue est de permettre d'intégrer sur un ensemble différent de $\mathbb{R}$, ou $\mathbb{R}^n$. les applications seront nombreuses, notamment en analyse fonctionnelle.

CHAPITRE 1

Rappels : cardinaux et dénombrabilité

1.1 Cardinalité

Définition 1.1 – On dit que deux ensembles E et F sont *équipotents* s'il existe une bijection de E dans F. On note alors $E \simeq F$ ou
$$\text{Card } E \ = \ \text{Card } F.$$
On dit dans ce cas que E et F ont *même cardinal* (attention, il ne s'agit que d'une notation !).

Remarque 1.2 – On voit immédiatement que si E, F, G sont des ensembles

- $E \simeq E$,
- si $E \simeq F$, alors $F \simeq E$,
- si $E \simeq F$ et $F \simeq G$, alors $E \simeq G$.

$\simeq$ se comporte donc comme une relation d'équivalence (mais n'en est pas une à proprement parler puisqu'il n'existe pas d'ensemble de tous les ensembles).

Exemples –

1. $\mathbb{N} \simeq 2\mathbb{N}$,
2. $\mathscr{P}(E) \simeq \mathscr{F}(E, \{0,1\}) = \{0,1\}^E$: l'application $A \subset E \mapsto \mathbb{1}_A \in \{0,1\}^E$ est bijective.

Notation 1.3 – Soient E et F deux ensembles. S'il existe une injection de E dans F, on notera
$$E \preceq F \qquad \text{ou} \qquad \text{Card } E \leqslant \text{Card } F$$
(on prendra garde aussi à cette dernière notation). Si de plus E et F ne sont pas équipotents, on notera $E \prec F$ ou Card $E <$ Card F.
De même, lorsqu'il existe une surjection de E dans F, on notera $E \succeq F$ ou Card $E \geqslant$ Card F, et $E \succ F$ ou Card $E >$ Card F lorsque de plus E et F ne sont pas équipotents.

Propriétés 1.4 –

- *Si* Card $E \leqslant$ Card F, *alors* Card $F \geqslant$ Card E.
- *Si* Card $E \geqslant$ Card F, *alors* Card $F \leqslant$ Card E (AC).

Remarque 1.5 – La démonstration du deuxième point des propriétés ci-dessus repose sur l'axiome du choix (précisé (AC) dans l'énoncé). En fait, cet énoncé est même équivalent à l'axiome du choix.
On donne deux énoncés équivalents de l'axiome du choix :

– pour tout ensemble E non vide, il existe une fonction $f : \mathscr{P}(E) \to E$ telle que pour tout $A \subset E$ non vide, $f(A) \in A$ (une telle fonction est appelée *fonction choix*),
– pour toute famille $(E_i)_{i\in I}$ d'ensembles non vides, $\prod_{i\in I} E_i$ est non vide.

Théorème 1.6 (*Cantor-Bernstein*) – *Si E et F sont deux ensembles vérifiant* $\text{Card } E \leqslant \text{Card } F$ *et* $\text{Card } F \leqslant \text{Card } E$*, alors* $\text{Card } E = \text{Card } F$.

Remarque 1.7 – La version duale du thérème de Cantor-Bernstein ($\text{Card } E \geqslant \text{Card } F$ et $\text{Card } F \geqslant \text{Card } E \Rightarrow \text{Card } E = \text{Card } F$) peut donc se montrer en utilisant l'axiome du choix. La question de savoir si elle est équivalente à l'axiome du choix est ouverte, mais on sait qu'on ne peut pas montrer la version duale du théorème dans l'axiomatique de Zermelo-Fraenkel.

Proposition 1.8 – *Soit E est un ensemble non vide,* $\text{Card } E < \text{Card } \mathscr{P}(E)$.

Démonstration. Supposons qu'il existe une surjection φ de E dans $\mathscr{P}(E)$. Soit $A = \{x \in E,\ x \notin \varphi(x)\}$. Comme φ est surjective, il existe $a \in E$ tel que $\varphi(a) = A$. Alors :
– soit $a \in A$, alors $a \notin \varphi(a) = A$, contradiction,
– soit $a \notin A$, alors $a \in \varphi(a) = A$, contradiction.

On aboutit donc à une contradiction, et il n'existe pas de surjection de E dans $\mathscr{P}(E)$. En considérant l'injection $x \in E \mapsto \{x\} \in \mathscr{P}(E)$, on voit par ailleur que $\text{Card } E \leqslant \text{Card } \mathscr{P}(E)$, d'où le résultat. □

1.2 Ensembles infinis, ensembles dénombrables

Définition 1.9 – On dit qu'un ensemble E est infini s'il existe une injection de $\mathbb{N}$ dans E ($\text{Card } N \leqslant \text{Card } E$).

Exemples –
1. $\mathbb{R}$ est infini : $\mathbb{N} \subset \mathbb{R}$ donc $\text{Card } \mathbb{N} \leqslant \text{Card } \mathbb{R}$.
2. $\mathscr{P}(\mathbb{N})$ est infini.

Définition 1.10 – On dit qu'un ensemble E est dénombrable s'il existe une injection de E dans $\mathbb{N}$ ($\text{Card } E \leqslant \text{Card } \mathbb{N}$). Si $\text{Card } E > \text{Card } \mathbb{N}$, on dit que E est non dénombrable.

Remarque 1.11 – On a la définition équivalente : E est dénombrable si et seulement si $\text{Card } \mathbb{N} \geqslant \text{Card } E$ (sans l'axiome du choix).

Attention : on rencontre parfois une terminologie différente : on dit parfois que E est au plus dénombrable si $\text{Card } E \leqslant \text{Card } \mathbb{N}$, et que E est dénombrable si $\text{Card } E = \text{Card } \mathbb{N}$.

Exemples –
1. $2\mathbb{N}$ est dénombrable : $n \mapsto 2n$ est bijective.
2. $\mathbb{Z}$ est dénombrable : $n \in \mathbb{Z} \mapsto \begin{cases} 2n & \text{si } n \geqslant 0 \\ -2n-1 & \text{si } n < 0 \end{cases}$ est bijective.
3. $\mathbb{N}^2$ est dénombrable : $(n, m) \in \mathbb{N}^2 \mapsto \frac{(n+m)(n+m+1)}{2} + m$ est une bijection de $\mathbb{N}^2$ dans $\mathbb{N}$.
4. $\mathbb{N}^k$ est dénombrable pour tout $k \in \mathbb{N}$: par récurrence à partir du résultat pour $k = 2$.
5. Si $E_1, \ldots E_k$ sont des ensembles dénombrables, alors $E_1 \times \ldots \times E_k$ est dénombrable : si pour tout $i \in [|1, k|]$, φ_i désigne une injection de E_i dans $\mathbb{N}$, alors la fonction
$$\varphi : (x_1, \ldots, x_k) \in E_1 \times \ldots \times E_k \ \mapsto\ (\varphi_1(x_1), \ldots, \varphi_k(x_k))$$
est injective, donc $\text{Card } E_1 \times \ldots \times E_k \ \leqslant\ \text{Card } \mathbb{N}^k \ =\ \text{Card } \mathbb{N}$.
6. $\mathbb{Q}$ est dénombrable :
$$\begin{array}{rcl} \mathbb{Q} & \to & \mathbb{Z} \times \mathbb{N} \\ \frac{p}{q} \text{ avec } p \wedge q = 1 & \mapsto & (p, q) \end{array}$$
est injective, donc $\text{Card } \mathbb{Q} \leqslant \text{Card } \mathbb{Z} \times \mathbb{N} \leqslant \text{Card } \mathbb{N}$.

7. $\mathscr{P}(\mathbb{N})$ n'est pas dénombrable : Card $\mathbb{N} <$ Card $\mathscr{P}(\mathbb{N})$ par la proposition 1.8.
8. $\{0,1\}^{\mathbb{N}}$ n'est pas dénombrable : Card $\{0,1\}^{\mathbb{N}} =$ Card $\mathscr{P}(\mathbb{N})$.

PROPOSITION 1.12 – *Une union dénombrable d'ensembles dénombrables est dénombrable.*

Démonstration. Soient $(E_n)_{n\in\mathbb{N}}$ une suite d'ensembles dénombrables et $E = \bigcup_{n\in\mathbb{N}} E_n$. Pour tout $n \in \mathbb{N}$, soit $\varphi_n : \mathbb{N} \to E_n$ surjective. On introduit la fonction

$$\begin{array}{cccc} \varphi & : & \mathbb{N}^2 & \to & E \\ & & (n,m) & \mapsto & \varphi_n(m) \end{array}.$$

On voit aisément que φ est surjective donc Card $\mathbb{N} =$ Card $\mathbb{N}^2 \geqslant$ Card E, et E est dénombrable. □

THÉORÈME 1.13 – $\mathbb{R}$ *n'est pas dénombrable.*

Première démonstration. Il s'agit du célèbre procédé diagonal de Cantor. On suppose que $D := [0,1]$ est dénombrable, on peut donc écrire D comme une suite : $D = \{x_n,\ n \in \mathbb{N}\}$. Pour tout $n \in \mathbb{N}$, on note $0,x_n^1x_n^2x_n^3\ldots$ une écriture décimale de x_n (la suite $(x_n^k)_k$ peut stagner en 0 ou en 9 dans le cas où x_n est décimal). On définit alors $x = 0,x^1x^2x^3\ldots \in [0,1]$ par son écriture décimale : pour tout $k \in \mathbb{N}$,

$$x^k = \begin{cases} 2 & \text{si } x_k^k = 1, \\ 1 & \text{sinon.} \end{cases}$$

La construction de x assure alors que $x \neq x_n$ pour tout $n \in \mathbb{N}$, donc $x \in [0,1] \setminus D$, et il y a contradiction.

Deuxième démonstration. On considère :

$$\begin{array}{cccc} \varphi & : & \{0,1\}^{\mathbb{N}^\star} & \to & \mathbb{R} \\ & & (x_n)_{n\geqslant 1} & \mapsto & \sum_{n\geqslant 1} \dfrac{x_n}{3^n} \end{array}.$$

La fonction φ est injective : soient deux suites distinctes $x = (x_n)_{n\geqslant 1}$ et $y = (y_n)_{n\geqslant 1}$ à valeurs dans $\{0,1\}$, montrons que $\varphi(x) \neq \varphi(y)$. On note $n_0 = \min\{n,\ x_n \neq y_n\}$. On peut supposer que $x_{n_0} = 0$ et $y_{n_0} = 1$. On a alors :

$$\begin{aligned} |\varphi(y) - \varphi(x)| &= \left| \frac{1}{3^{n_0}} + \sum_{n>n_0} \frac{y_n - x_n}{3^n} \right| \\ &\geqslant \frac{1}{3^{n_0}} - \left| \sum_{n>n_0} \frac{y_n - x_n}{3^n} \right| \geqslant \frac{1}{3^{n_0}} - \sum_{n>n_0} \left| \frac{y_n - x_n}{3^n} \right| \\ &\geqslant \frac{1}{3^{n_0}} - \sum_{n>n_0} \frac{1}{3^n} = \frac{1}{2} \cdot \frac{1}{3^{n_0}} > 0. \end{aligned}$$

Alors, Card $\mathbb{N} <$ Card $\{0,1\}^{\mathbb{N}^\star} \leqslant$ Card $\mathbb{R}$, et $\mathbb{R}$ n'est pas dénombrable. □

EXERCICE –

- Quel est le cardinal de $\mathbb{R}^{\mathbb{N}}$?
- Quel est le cardinal de l'ensemble $\mathscr{O}(\mathbb{R})$ des ouverts de $\mathbb{R}$?
- Quel est le cardinal des fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$?

CHAPITRE 2

Tribus, tribu borélienne

2.1 Définitions

Soit E un ensemble. On appelle *classe de parties* de E tout sous-ensemble de $\mathscr{P}(E)$.

DÉFINITION 2.1 – On appelle *tribu* sur E (ou σ-*algèbre*) une classe de partie $\mathscr{A} \subset \mathscr{P}(E)$ telle que :

- $\emptyset \in \mathscr{A}$,
- si $A \in \mathscr{A}$, alors $A^c \in \mathscr{A}$ (stabilité par complémentaire),
- si $A_n \in \mathscr{A}$, $n \in \mathbb{N}$, alors $\bigcup_{n\in\mathbb{N}} A_n \in \mathscr{A}$ (stabilité par union dénombrable).

On dit alors que $(E, \mathscr{A})$ est un espace *mesurable*.

REMARQUE 2.2 – Par définition, une tribu $\mathscr{A}$ sur E est aussi stable par intersection dénombrable : si $A_n \in \mathscr{A}$, $n \in \mathbb{N}$,

$$\bigcap_{n\in\mathbb{N}} A_n = \left(\bigcup_{n\in\mathbb{N}} {A_n}^c\right)^c \in \mathscr{A}.$$

EXEMPLES –

1. $\mathscr{P}(E)$: tribu triviale,
2. $\{\emptyset, E\}$: tribu grossière,
3. Si $A \subset E$, $\{\emptyset, A, A^c, E\}$ est une tribu. C'est la plus petite tribu sur E contenant A.
4. $\mathscr{A} = \{A \subset E,\ A \text{ dénombrable ou } A^c \text{ dénomrable}\}$ est une tribu sur E.

   Démonstration. Il est clair que $\emptyset \in \mathscr{A}$ et que $\mathscr{A}$ est stable par passage au complémentaire. Maintenant, si $A_n \in \mathscr{A}$ pour tout $n \in \mathbb{N}$, alors soit tous les A_n sont dénombrables, et $\bigcup_n A_n$ est dénombrable, soit il existe n_0 tel que A_{n_0} n'est pas dénombrable, auquel cas $(A_{n_0})^c$ est dénombrable, et $(\bigcup_n A_n)^c = \bigcap_n (A_n)^c \subset (A_{n_0})^c$ est dénombrable. Dans les deux cas, $\bigcup_n A_n \in \mathscr{A}$.

REMARQUE 2.3 – Une intersection quelconque de tribus sur E est une tribu sur E.

DÉFINITION-PROPOSITION 2.4 – Soit $\mathscr{C}$ une classe de parties de E. Il existe une plus petite tribu sur E (au sens de l'inclusion) contenant les éléments de $\mathscr{C}$. On la note $\sigma(\mathscr{C})$ et on l'appelle *tribu engendrée* par $\mathscr{C}$ sur E.
En d'autres termes, pour toute tribu $\mathscr{A}$ sur E telle que $\mathscr{C} \subset \mathscr{A}$, on a $\sigma(\mathscr{C}) \subset \mathscr{A}$.

Démonstration. Il suffit de poser

$$\sigma(\mathscr{C}) = \bigcap_{\substack{\mathscr{A} \text{ tribu} \\ \mathscr{C} \subset \mathscr{A}}} \mathscr{A}.$$

$\sigma(\mathscr{C})$ est bien une tribu comme intersection de tribus. □

Exemple 2.5 – Si $A \subset E$, $\sigma(\{A\}) = \{\emptyset, A, A^c, E\}$.

Exercice – Quelle est la tribu engendrée par l'ensemble $\{\{x\},\ x \in E\}$ des singletons de E ? (voir TD)

Il arrive fréquemment lorsqu'on est confronté à une tribu engendrée par une classe de parties $\mathscr{C}$ qu'on veuille démontrer qu'une propriété vérifiée par les éléments de $\mathscr{C}$ reste vraie pour $\sigma(\mathscr{C})$ tout entière. Pour montrer une telle chose, on pourrait vouloir essayer de décrire un élément quelconque de $\sigma(\mathscr{C})$ à partir d'éléments de $\mathscr{C}$. Malheureusement, il n'y a pas de procédé général pour y parvenir : un élément de $\sigma(\mathscr{C})$ ne peut pas toujours être décrit comme une union dénombrable d'ensembles de $\mathscr{C}$, ni comme une intersection d'union de tels ensembles, etc. La remarque suivante suggère un procédé de démonstration bien commode pour montrer que $\sigma(\mathscr{C})$ continue de vérifier une propriété vraie pour $\mathscr{C}$. Nous l'utiliserons à de nombreuses reprises dans ce cours.

Remarque 2.6 – Pour montrer que tous les éléments d'une tribu $\mathscr{A}$ sur E engendrée par $\mathscr{C} \subset \mathscr{P}(E)$ vérifient une propriété (P), il suffit de montrer que :
- $\mathscr{B} := \{B \in \mathscr{A},\ B \text{ vérifie } (P)\}$ est une tribu,
- $\forall C \in \mathscr{C},\ C$ vérifie (P).

Comme $\mathscr{B} \supset \mathscr{C}$, on aura alors $\mathscr{B} \supset \sigma(\mathscr{C}) = \mathscr{A}$. On aura souvent recours à ce type de raisonnement pour monter des propriétés générales sur une tribu engendrée par une classe de parties simple.

2.2 Tribu borélienne

2.2.1 Espace topologique

On introduit ci-dessous la notion de topologie d'un ensemble E. La définition est à comparer à celle de tribu vue plus haut.

Définition 2.7 – Soit $\mathscr{T}$ une classe de parties de E. On dit que $\mathscr{T}$ est une *topologie* sur E si
- $\emptyset, E \in \mathscr{T}$,
- $\mathscr{T}$ est stable par union quelconque,
- $\mathscr{T}$ est stable par intersection finie.

On appelle alors *ouverts* de E des éléments de la topologie $\mathscr{T}$, et *fermés* les complémentaires des ensembles de $\mathscr{T}$. On dit alors que $(E, \mathscr{T})$ est un espace topologique.

Exemple 2.8 – $(\mathbb{R}, \mathscr{O}(\mathbb{R}))$ est un espace topologique.

Remarque 2.9 – Une intersection quelconque de topologies sur E est encore une topologie sur E. On en déduit que si $\mathscr{C}$ est une classe de parties de E, il existe une plus petite topologie sur E contenant $\mathscr{C}$: il s'agit tout simplement de l'intersection des topologies sur E qui contiennent $\mathscr{C}$. On l'appelle *topologie engendrée par* $\mathscr{C}$.

2.2.2 Tribu borélienne

Définition 2.10 – Soit $(E, \mathscr{T})$ un espace topologique. On appelle *tribu borélienne* et on note $\mathscr{B}(E)$ la tribu engendrée par les ouverts de E :

$$\mathscr{B}(E) = \sigma(\mathscr{T}).$$

Les éléments de $\mathscr{B}(E)$ sont appelés les *boréliens* de E.

Remarque 2.11 – ⋄ Par définition, $\mathscr{B}(E)$ est aussi la tribu engendrée par les fermés de E.

◇ En général, $\mathscr{B}(E) \neq \mathscr{P}(E)$. C'est par exemple le cas pour $E = \mathbb{R}$. Nous verrons en effet plus loin qu'il est facile de construire des parties de $\mathbb{R}$ qui ne sont pas boréliennes, en ayant recours à l'axiome du choix. Il est en fait possible d'exhiber de telles parties sans utiliser l'axiome du choix, mais leur construction n'est pas triviale.

En fait, on peut montrer (sans axiome du choix) le résultat plus général :

$$\text{Card } \mathscr{B}(\mathbb{R}) \;=\; \text{Card } \mathbb{R}.$$

Ce résultat est assez difficile, et sa démonstration repose sur une récurrence sur les ordinaux. On déduit donc que Card $\mathscr{B}(\mathbb{R}) <$ Card $\mathscr{P}(\mathbb{R})$. Autrement dit, les boréliens occupent une place minuscule parmi les parties de $\mathbb{R}$. Dans la pratique, il sera pourtant tout à fait satisfaisant de se limiter à travailler dans les boréliens : en témoigne la difficulté de construire des ensembles non boréliens !

Boréliens de $\mathbb{R}$, de $\mathbb{R}^d$

Nous allons donner d'autres classes de parties de $\mathbb{R}$ ou de $\mathbb{R}^d$ qui engendrent les boréliens. Leur intérêt est que leurs éléments sont plus simples que les ouverts en général. Ce point peut s'avérer très utile, en particulier dans l'esprit de la remarque 2.6.

Proposition 2.12 –
$$\begin{aligned} \mathscr{B}(\mathbb{R}) &= \sigma\big(\{]\alpha,\beta[,\ \alpha<\beta,\ \alpha,\beta\in\mathbb{Q}\}\big) \\ &= \sigma\big(\{]-\infty,a[,\ a\in\mathbb{Q}\}\big) \;=\; \sigma\big(\{]a,+\infty[,\ a\in\mathbb{Q}\}\big) \\ &= \sigma\big(\{]-\infty,a],\ a\in\mathbb{Q}\}\big) \;=\; \sigma\big(\{[a,+\infty[,\ a\in\mathbb{Q}\}\big). \end{aligned}$$

Démonstration. — Si $\alpha,\beta\in\mathbb{Q}$ et $\alpha<\beta$, $]\alpha,\beta[\in\mathscr{O}(\mathbb{R})$, donc $\sigma\big(\{]\alpha,\beta[,\ \alpha<\beta,\ \alpha,\beta\in\mathbb{Q}\}\big)\subset\mathscr{B}(\mathbb{R})$.

Réciproquement, on remarque que si $\Omega\in\mathscr{O}(\mathbb{R})$, on a $\Omega=\bigcup\{]p,q[,\ p,q\in\mathbb{Q},\ ]p,q[\subset\Omega\}$. On a alors $\Omega\in\sigma\big(\{]\alpha,\beta[,\ \alpha<\beta,\ \alpha,\beta\in\mathbb{Q}\}\big)$, ce qui fournit la deuxième inclusion.

— Si $a\in\mathbb{Q}$, $]-\infty,a[\in\mathscr{O}(\mathbb{R})$ donc $\sigma\big(\{]-\infty,a[,\ a\in\mathbb{Q}\}\big)\subset\mathscr{B}(\mathbb{R})$.

D'autre part, si $\alpha,\beta\in\mathbb{Q}$,

$$]\alpha,\beta[\, = \,]-\infty,\beta[\setminus]-\infty,\alpha] \;=\,]-\infty,\beta[\setminus\bigcap_{n\in\mathbb{N}^\star}\Big]-\infty,\alpha+\frac{1}{n}\Big[$$

donc $\mathscr{B}(\mathbb{R})=\sigma\big(\{]\alpha,\beta[,\ \alpha<\beta,\ \alpha,\beta\in\mathbb{Q}\}\big)\subset\sigma\big(\{]-\infty,a[,\ a\in\mathbb{Q}\}\big)$.

Les trois derniers cas se traitent de la même façon. □

Remarque 2.13 – On montre de manière similaire que, plus généralement, si $d\in\mathbb{N}^\star$,

$$\mathscr{B}(\mathbb{R}^d) \;=\; \sigma\big(\{]a_1,b_1[\times\ldots\times]a_d,b_d[,\ a_i,b_i\in\mathbb{Q},\ i=1,\ldots,d\}\big).$$

La propriété de stabilité par translation suivante sera utile pour définir proprement la mesure de Lebesgue.

Proposition 2.14 – *Si $B\in\mathscr{B}(\mathbb{R}^d)$ et $a\in\mathbb{R}^d$, alors $B+a:=\{x+a,\ x\in B\}$ est un borélien de $\mathbb{R}^d$.*

Démonstration. Soit $a\in\mathbb{R}^d$, on pose $\mathscr{A}=\{B\in\mathscr{B}(\mathbb{R}^d),\ B+a\in\mathscr{B}(\mathbb{R}^d)\}$. Il est clair que $\mathscr{A}$ contient $\mathscr{O}(\mathbb{R}^d)$. Par ailleurs, $\mathscr{A}$ est une tribu sur $\mathbb{R}^d$:

◇ $\emptyset\in\mathscr{A}$,
◇ si $B\in\mathscr{A}$, alors $B^c+a=(B+a)^c\in\mathscr{B}(\mathbb{R}^d)$,
◇ si $B_n\in\mathscr{A}$, $n\in\mathbb{N}$, alors $\big(\bigcup_n B_n\big)+a=\bigcup_n(B_n+a)\in\mathscr{B}(\mathbb{R}^d)$.

On en déduit que $\mathscr{A}\supset\sigma(\mathscr{O}(\mathbb{R}^d))=\mathscr{B}(\mathbb{R}^d)$, d'où le résultat. □

Nous verrons plus loin une manière plus directe de montrer ce résultat, à l'aide de la notion de fonction mesurable.

Boréliens de $\overline{\mathbb{R}}$

On commence par introduire deux nouveaux éléments $-\infty$ et $+\infty$, et on note

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

On prolonge ensuite l'ordre total $\leqslant$ sur $\mathbb{R}$ en posant

$$-\infty \ \leqslant \ x \ \leqslant \ +\infty, \quad \forall x \in \overline{\mathbb{R}}.$$

On munit ensuite $\overline{\mathbb{R}}$ de la topologie engendrée par $\mathscr{O}(\mathbb{R}) \cup \{[-\infty, a[, \ a \in \mathbb{R}\} \cup \{]b, +\infty], \ b \in \mathbb{R}\}$. On peut alors montrer le résultat suivant.

PROPOSITION 2.15 – $$\begin{aligned} \mathscr{B}(\overline{\mathbb{R}}) &= \sigma\big(\{[-\infty, a[, \ a \in \mathbb{Q}\}\big) = \sigma\big(\{]a, +\infty], \ a \in \mathbb{Q}\}\big) \\ &= \sigma\big(\{[-\infty, a], \ a \in \mathbb{Q}\}\big) = \sigma\big(\{[a, +\infty], \ a \in \mathbb{Q}\}\big). \end{aligned}$$

2.3 Tribu image réciproque, tribu image

Soient E et F deux ensembles, et $f : E \to F$. Pour toute classe de paties $\mathscr{C}$ de F, on notera

$$f^{-1}(\mathscr{C}) = \{f^{-1}(C), \ C \in \mathscr{C}\}.$$

2.3.1 Tribu image réciproque

DÉFINITION-PROPOSITION 2.16 – Si $\mathscr{B}$ est une tribu sur F, la classe de parties $f^{-1}(\mathscr{B})$ est une tribu sur E. On l'appelle tribu image réciproque de $\mathscr{B}$ par f.

Démonstration. — $\emptyset = f^{-1}(\emptyset) \in f^{-1}(\mathscr{B})$,

— si $A \in \mathscr{A}$, alors $A = f^{-1}(B)$ avec $B \in \mathscr{B}$, et $A^c = f^{-1}(B^c) \ \in f^{-1}(\mathscr{B})$,

— si $A_n \in \mathscr{A}$, $n \in \mathbb{N}$, alors $A_n = f^{-1}(B_n)$ avec $B_n \in \mathscr{B}$, et $\bigcup_n A_n = f^{-1}\big(\bigcup_n B_n\big) \in f^{-1}(\mathscr{B})$.

Donc $\mathscr{A}$ est bien une tribu sur E. □

EXEMPLE 2.17 – Soient $\mathscr{B}$ une tribu sur E et $A \subset E$. On considère l'injection canonique $i : x \in A \mapsto x$ de A dans E. Alors, la classe de parties

$$\{i^{-1}(B), \ B \in \mathscr{B}\} \ = \ \{A \cap B, \ B \in \mathscr{B}\}$$

est une tribu sur A. On l'appelle *tribu trace* de $\mathscr{B}$ sur A.

2.3.2 Tribu image

En général, si $\mathscr{A}$ est une tribu sur E, la classe de parties $\{f(A), \ A \in \mathscr{A}\}$ n'est pas une tribu sur F. Par exemple, si $E = F = \{0, 1\}$, $\mathscr{A} = \{\emptyset, E\}$ et $f : x \in E \mapsto 0$, on a $f(\mathscr{A}) = \{\emptyset, \{0\}\}$, qui n'est pas une tribu. On montre facilement que $\mathscr{B} = \{B \subset F, \ f^{-1}(B) \in \mathscr{A}\}$ est une tribu sur F. En effet,

- $f^{-1}(\emptyset) = \emptyset \in \mathscr{A}$, donc $\emptyset \in \mathscr{B}$,
- si $B \in \mathscr{B}$, alors $f^{-1}(B^c) = (f^{-1}(B))^c \in \mathscr{A}$, donc $B^c \in \mathscr{B}$,
- si $B_n \in \mathscr{B}$ pour tout $n \in \mathbb{N}$, alors $f^{-1}(\bigcup_n B_n) = \bigcup_n f^{-1}(B_n) \in \mathscr{A}$, donc $\bigcup_n B_n \in \mathscr{B}$.

DÉFINITION 2.18 – Si $\mathscr{A}$ est une tribu sur E, on appelle tribu image de $\mathscr{A}$ par f la tribu sur F

$$\{B \subset F, \ f^{-1}(B) \in \mathscr{A}\}.$$

2.3.3 Lemme de transport

LEMME 2.19 – *Soit $\mathscr{C}$ une classe de parties de F, on a*

$$\sigma(f^{-1}(\mathscr{C})) \;=\; f^{-1}(\sigma(\mathscr{C})). \tag{2.1}$$

REMARQUE 2.20 – Les deux classes de parties de (2.1) sont des tribus sur E : $\sigma(f^{-1}(\mathscr{C}))$ est la tribu sur E engendrée par la classe de parties $f^{-1}(\mathscr{C})$, et $f^{-1}(\sigma(\mathscr{C}))$ est la tribu image réciproque de la tribu $\sigma(\mathscr{C})$ par f.

Démonstration. — On remarque d'abord que $f^{-1}(\mathscr{C}) \subset f^{-1}(\sigma(\mathscr{C}))$. Comme $f^{-1}(\sigma(\mathscr{C}))$ est une tribu contenant la classe de parties $f^{-1}(\mathscr{C}) \subset$, elle contient la tribu qu'elle engendre, *ie* $\sigma(f^{-1}(\mathscr{C})) \subset f^{-1}(\sigma(\mathscr{C}))$.

— Montrons la deuxième inclusion. Il suffit de montrer que pour tout $B \in \sigma(\mathscr{C})$, $f^{-1}(B) \in \sigma(f^{-1}(\mathscr{C}))$. On note $\mathscr{B} = \{B \subset F,\ f^{-1}(B) \in \sigma(f^{-1}(\mathscr{C}))\}$. $\mathscr{B}$ est la tribu image par f de la tribu $\sigma(f^{-1}(\mathscr{C}))$. C'est donc une tribu, et on remarque immédiatement qu'elle contient $\mathscr{C}$, elle contient donc $\sigma(\mathscr{C})$, ce qui achève la démonstration. □

CHAPITRE 3

Mesures

3.1 Définitions et exemples

Soit $(E, \mathscr{A})$ un espace mesurable.

DÉFINITION 3.1 – On appelle *mesure* sur $(E, \mathscr{A})$ une application $\mu : \mathscr{A} \to \overline{\mathbb{R}}_+$ telle que :

- $\mu(\emptyset) = 0$,
- si $A_n \in \mathscr{A}$, $n \in \mathbb{N}$ deux à deux disjoints, alors $\mu\left(\bigsqcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} \mu(A_n)$.

On dit alors que $(E, \mathscr{A}, \mu)$ est un *espace mesuré*.

Voici quelques exemples de mesures classiques.

- Mesure *nulle* : μ définie par $\mu(A) = 0$ pour tout $A \in \mathscr{A}$ est une mesure sur $(E, \mathscr{A})$ appelée mesure nulle.
- Mesure *grossière* : μ définie par $\mu(A) = \infty$ pour tout $A \in \mathscr{A} \setminus \emptyset$ est une mesure sur $(E, \mathscr{A})$ appelée mesure grossière.
- Mesure *de comptage* : m définie par

$$m(A) = \begin{cases} \text{Card } A & \text{si } A \text{ est fini,} \\ +\infty & \text{sinon,} \end{cases}$$

  pour tout $A \subset E$ est une mesure sur $(E, \mathscr{P}(E))$, appelée mesure de comptage.
- Mesure de *Dirac* : si $x \in E$, δ_x définie par

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

  pour tout $A \in \mathscr{A}$ est une mesure sur $(E, \mathscr{A})$ appelée mesure de Dirac au point $x \in E$.
- Si μ est une mesure sur $(E, \mathscr{A})$ et $B \in \mathscr{A}$, alors $\nu : \mathscr{A} \to \overline{\mathbb{R}}_+$ définie par $\nu(A) = \mu(A \cap B)$ pour tout $A \in \mathscr{A}$ est aussi une mesure sur $(E, \mathscr{A})$.

DÉFINITION 3.2 (*Mesures finies, mesures σ-finies*) –

- On dit qu'une mesure μ est une mesure *finie* si $\mu(E) < \infty$. On appelle alors $\mu(E)$ la *masse* de μ. Si μ est de masse 1, on dit que μ est une *mesure de probabilité* sur E, et on dit que $(E, \mathscr{A}, \mu)$ est un espace probabilisé.

– On dit qu'une mesure μ est σ-finie s'il existe une suite $(E_n)_n$ de $\mathscr{A}$ telle que $E = \bigcup_n E_n$ et $\mu(E_n) < \infty$ pour tout n.

Exemple 3.3 – La mesure de comptage m sur $(\mathbb{N}, \mathscr{P}(\mathbb{N}))$ est σ-finie. En effet, pour tout $n \in \mathbb{N}$, $m([|0,n|]) = n+1 < \infty$, et on a bien sûr $\mathbb{N} = \bigcup_n [|0,n|]$.

Remarque 3.4 – Si pour tout $n \in \mathbb{N}$, μ_n est une mesure sur $(E, \mathscr{A})$, alors $\sum_n \mu_n$ est aussi une mesure sur $(E, \mathscr{A})$ (exercice).
Par exemple, si $(\alpha_n)_{n\in\mathbb{N}} \in \mathbb{R}_+^{\mathbb{N}}$ et $(x_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$, alors $\sum_n \alpha_n \delta_{x_n}$ est une mesure sur $(E, \mathscr{A})$. On note ainsi que si X est une variable aléatoire réelle discrète sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$ ayant pour valeurs les réels $x_n, n \in \mathbb{N}$, alors on peut écrire la loi de X sous la forme $\mathbb{P}_X = \sum_n \alpha_n \delta_{x_n}$, où $\alpha_n = \mathbb{P}(X = x_n)$ pour tout $n \in \mathbb{N}$.

3.2 Propriétés des mesures

Soit μ une mesure sur l'espace mesurable $(E, \mathscr{A})$.

Proposition 3.5 (*Croissance*) – *Si $A, B \in \mathscr{A}$ avec $A \subset B$, alors $\mu(A) \leqslant \mu(B)$ (μ est croissante). Si de plus $\mu(A) < \infty$, alors*

$$\mu(B \setminus A) = \mu(B) - \mu(A).$$

Démonstration. A et $B \setminus A$ appartiennent à $\mathscr{A}$ et sont disjoints, donc $\mu(B) = \mu(A) + \mu(B \setminus A) \geqslant \mu(A)$. Si $\mu(A) < \infty$, alors $\mu(B) - \mu(A) = \mu(B \setminus A)$. □

Proposition 3.6 (σ*-sous-additivité*) – *Pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathscr{A}$,*

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leqslant \sum_{n\in\mathbb{N}} \mu(A_n).$$

Démonstration. Pour tout n, on pose $B_n = A_n \setminus \bigcup_{k=0}^{n-1} A_k$. On note que pour tout n, $B_n \in \mathscr{A}$, et que $B_n \subset A_n$.

- Les B_n sont deux à deux disjoints : si $m < n$, $B_m \subset A_m$ et $B_n \cap A_m = \emptyset$, d'où $B_m \cap B_n = \emptyset$.
- $\bigsqcup_n B_n = \bigcup_n A_n$: si $x \in \bigcup_n A_n$, soit $n_0 = \min\{n,\ x \in A_n\}$, alors $x \in A_{n_0} \setminus \bigcup_{k=0}^{n_0-1} A_k = B_{n_0}$. La deuxième inclusion est évidente.

En conséquence,

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \mu\left(\bigsqcup_{n\in\mathbb{N}} B_n\right) = \sum_{n\in\mathbb{N}} \mu(B_n) \leqslant \sum_{n\in\mathbb{N}} \mu(A_n).$$ □

Proposition 3.7 (*Continuité*) – *Soit $(A_n)_n$ une suite d'éléments de $\mathscr{A}$.*

1. *Si la suite (A_n) est croissante pour l'inclusion (i.e. $A_n \subset A_{n+1}$ pour tout $n \in \mathbb{N}$), alors*
$$\lim_{n\to\infty} \mu(A_n) = \mu\left(\bigcup_n A_n\right).$$
On dit alors que μ est "continue à gauche".
2. *Si la suite (A_n) est décroissante pour l'inclusion (i.e. $A_n \subset A_{n+1}$ pour tout $n \in \mathbb{N}$) et il existe n_0 tel que $\mu(A_{n_0}) < \infty$, alors*
$$\lim_{n\to\infty} \mu(A_n) = \mu\left(\bigcap_n A_n\right).$$
On dit alors que μ est "continue à droite".

Démonstration.

1. Notons d'abord que s'il existe n_0 tel que $\mu(A_{n_0}) = \infty$, alors $\mu(\bigcup_n A_n) = \infty$, et $\mu(A_n) = \infty$ à partir d'un certain rang, d'où le résultat.

   Sinon, on introduit la suite (B_n) définie par $B_0 = A_0$, et pour tout $n \geqslant 1$, $B_n = A_n \setminus A_{n-1}$. Il s'agit en fait de la même suite que celle construite dans la démonstration précédente. En particulier, les B_n sont deux à deux disjoints, et $\bigcup_n A_n = \bigsqcup_n B_n$. Par conséquent,

$$\mu\Big(\bigcup_n A_n\Big) = \mu\Big(\bigsqcup_n B_n\Big) = \sum_{n=0}^{\infty} \mu(B_n) = \lim_{n\to\infty}\Big(\sum_{k=1}^{n} \mu(A_k) - \mu(A_{k-1})\Big) + \mu(A_0) = \lim_{n\to\infty} \mu(A_n).$$

2. Quitte à considérer la suite $(A_{n_0+n})_{n\in\mathbb{N}}$, on peut supposer que $\mu(A_0) < \infty$. Comme $(A_0 \setminus A_n)_n$ est une suite de $\mathscr{A}$ croissante pour l'inclusion, on peut lui appliquer le résultat précédent. On obtient alors

$$\lim_{n\to\infty} \mu(A_0 \setminus A_n) \;=\; \mu\Big(\bigcup_n A_0 \setminus A_n\Big) \;=\; \mu\Big(A_0 \setminus \bigcap_n A_n\Big).$$

   Comme $\mu(A_0) < \infty$, on en déduit que $\lim_n \mu(A_0) - \mu(A_n) = \mu(A_0) - \mu(\bigcap_n A_n)$, ce qui donne le résultat. □

Remarque 3.8 – L'existence de n_0 avec $\mu(A_{n_0}) < \infty$ est nécessaire : si on considère les ensembles $A_n = [\![n, \infty[\![$ pour $n \in \mathbb{N}$, la suite $(A_n)_n$ est décroissante, et on a $\bigcap_n A_n = \emptyset$. Si on munit $(\mathbb{N}, \mathscr{P}(\mathbb{N}))$ de la mesure de comptage m sur $\mathbb{N}$, $m(\bigcap_n A_n) = 0$, mais pour tout n, $m(A_n) = \infty$.
On peut construire un exemple semblable avec la mesure de Lebesgue en posant $A_n = [n, \infty[$.

Exercice (*voir TD*) – Montrer qu'une application $\mu : \mathscr{A} \to \overline{\mathbb{R}}_+$ est une mesure sur l'espace mesurable $(E, \mathscr{A})$ si et seulement si :

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

(*ii*) μ est finiment additive : si $A, B \in \mathscr{A}$ et $A \cap B = \emptyset$, alors $\mu(A \sqcup B) = \mu(A) + \mu(B)$,

(*iii*) μ est continue à gauche.

3.3 Mesure de Lebesgue sur $\mathbb{R}^d$

La mesure de Lebesgue sur $\mathbb{R}$ a vocation à donner un sens à la notion de longueur pour les boréliens de $\mathbb{R}$, c'est-à-dire à généraliser la notion de longueur définie pour les intervalles : si $I \subset \mathbb{R}$ est un intervalle d'extrémités $a, b \in \mathbb{R}$, on note sa longueur

$$\ell(I) \;=\; b - a.$$

Par ailleurs, si I est un intervalle non borné de $\mathbb{R}$, on pose $\ell(I) = \infty$.
On cherchera plus généralement à donner un sens à la notion de volume pour tout borélien de $\mathbb{R}^d$, où d est un entier quelconque. Pour ce faire, on souhaitera généraliser la notion de volume d'un pavé de $\mathbb{R}^d$: si $P = I_1 \times \ldots \times I_k$ est un pavé de $\mathbb{R}^d$, où $I_1, \ldots, I_d$ sont des intervalles de $\mathbb{R}$, on note son volume

$$\mathscr{V}(P) \;=\; \prod_{k=1}^{d} \ell(I_k),$$

avec la convention $\mathscr{V}(P) = 0$ dès qu'il existe $k \in [\![1, d]\!]$ tel que $\ell(I_k) = 0$.

Nous allons en fait introduire la mesure de Lebesgue sur $\mathbb{R}^d$ comme l'unique mesure invariante par translation telle que la mesure du cube unité soit 1. Nous montrerons dans la proposition 3.10 que ceci entraîne que la mesure de Lebesgue généralise la notion de volume. Nous laissons temporairement de côté la démonstration du théorème qui suit, le chapitre 6 étant consacré aux résultats généraux qui permettent de l'établir.

Théorème 3.9 – *Soit $d \in \mathbb{N}^\star$. Il existe une unique mesure λ_d sur $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ telle que :*

– $\lambda_d([0,1]^d) = 1$,

– λ_d *est invariante par translation, i.e.*

$$\forall x \in \mathbb{R},\ \forall B \in \mathscr{B}(\mathbb{R}^d), \quad \lambda_d(B+x) = \lambda_d(B).$$

On appelle λ_d *la mesure de Lebesgue d-dimensionnelle. Dans le cas* $d = 1$*, on appelle* λ_1 *la mesure de Lebesgue, et on la note plus simplement* λ*.*

Les propriétés de la mesure de Lebesgue permettent de déduire directement que la mesure de Lebesgue de tout intervalle correspond à sa longueur. C'est l'objet de la proposition suivante.

Proposition 3.10 – *Pour tout intervalle* I *de* $\mathbb{R}$*,* $\lambda(I) = \ell(I)$*. En particulier,* $\lambda(\{x\}) = 0$ *pour tout* $x \in \mathbb{R}$*.*

Démonstration. On commence par remarquer que l'invariance par translation donne que pour tout $x \in \mathbb{R}$, $\lambda(\{x\}) = \lambda(\{0\})$. Par ailleurs, si $n \in \mathbb{N}^\star$, la croissance de la mesure donne $\lambda(\{1, \frac{1}{2}, \dots, \frac{1}{n}\}) \leqslant \lambda([0,1]) = 1$, donc $n\lambda(\{0\}) \leqslant 1$ par additivité, et $\lambda(\{0\}) \leqslant \frac{1}{n}$. En laissant tendre n vers ∞, on a donc $\lambda(\{0\}) = 0$.

Soit $q \in \mathbb{N}^\star$. On remarque que $\lambda\left(\bigsqcup_{k=1}^q [\frac{k-1}{q}, \frac{k}{q}[\right) = \lambda([0,1]) = 1$. Par conséquent, $q\lambda([0, \frac{1}{q}[) = 1$, et $\lambda([0, \frac{1}{q}[) = \frac{1}{q}$. L'additivité donne alors que pour tout $p \in \mathbb{N}$, $\lambda([0, \frac{p}{q}[) = \frac{p}{q}$. On en déduit que si $a, b \in \mathbb{Q}$, alors l'invariance de λ par translation implique que $\lambda([a,b[) = \lambda([0, b-a[) = b-a$.

Soient maintenant a, b deux réels distincts. Il existe une suite décroissante (a_n) et une suite croissante (b_n) de rationnels telles que $\lim_n a_n = a$ et $\lim_n b_n = b$. Ainsi, par continuité à gauche de λ, on a $\lambda(]a,b[) = \lambda(\bigcup_n]a_n, b_n[) = \lim_n \lambda(]a_n, b_n[) = \lim_n b_n - a_n = b - a$.

Si maintenant I et un intervalle non borné, il existe une suite (x_n) dans I telle que $\lim_n |x_n| = \infty$. On peut supposer sans perte de généralité que $\lim_n x_n = \infty$ et (x_n) est croissante. Alors pour tout n, $\lambda(I) \geqslant \lambda([x_0, x_n]) = x_n - x_0 \underset{n\to\infty}{\longrightarrow} \infty$. □

Remarque 3.11 – Une conséquence immédiate est que tout ensemble dénombrable D de $\mathbb{R}$ est de mesure de Lebesgue nulle : tout d'abord, D est un borélien comme union dénombrable de singletons, puis

$$\lambda(D) \;=\; \lambda\biggl(\bigsqcup_{x\in D} \{x\}\biggr) \;=\; \sum_{x\in D} \lambda(\{x\}) \;=\; 0.$$

Attention toutefois, la réciproque est fausse : il existe des boréliens de $\mathbb{R}$ de mesure nulle qui ne sont pas dénombrables. C'est le cas par exemple de l'ensemble de Cantor K, défini par $K = \bigcap_{n\in\mathbb{N}} K_n$, où $K_0 = [0,1]$, et $K_{n+1} = \frac{K_n}{3} \cup \frac{K_n+2}{3}$ pour tout $n \in \mathbb{N}$ (le vérifier en exercice !).

Plus généralement, si P est un pavé de $\mathbb{R}^d$, alors $\lambda_d(P) = \mathscr{V}(P)$. Pour le voir il suffirait d'adapter la preuve ci-dessus dans le cas $d = 1$. Nous verrons par ailleurs plus loin que λ_d peut s'interpréter comme une mesure produit, ce qui simplifie beaucoup la preuve de ce résultat.

Régularité de la mesure de Lebesgue

Nous avons vu que bien que les ouverts engendrent les boréliens, il n'est pas aisé du tout d'exprimer un borélien quelconque à partir d'ouverts ou de fermés. Pourtant, nous allons voir que pour n'importe quel borélien B de $\mathbb{R}^d$, on peut trouver un ouvert contenant B et un fermé inclus dans B dont la mesure est arbitrairement proche de celle de B. Il s'agit de la propriété de régularité de la mesure de Lebesgue sur $\mathbb{R}^d$.

La propriété proprement dite est énoncée dans le théorème 3.13 ci-dessous, et repose sur le lemme ci-dessous.

Lemme 3.12 – *Soient* $B \in \mathscr{B}(\mathbb{R}^d)$ *et* $\varepsilon > 0$*. Il existe un fermé* F *et un ouvert* Ω *de* $\mathbb{R}^d$ *tels que* $F \subset B \subset \Omega$ *et* $\lambda_d(\Omega \setminus F) < \varepsilon$*.*

Démonstration. Soient $n \in \mathbb{N}$ et λ_d^n la mesure sur $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ définie par $\lambda_d^n(B) = \lambda_d(B \cap]-n, n[^d)$ pour tout $B \in \mathscr{B}(\mathbb{R}^d)$. Nous allons d'abord montrer le résultat pour la mesure λ_d^n. Pour ce faire, on introduit la classe de parties

$$\mathscr{B} = \{B \in \mathscr{B}(\mathbb{R}^d),\ \forall \varepsilon > 0,\ \exists F \text{ fermé},\ \exists \Omega \text{ ouvert},\ F \subset B \subset \Omega,\ \lambda_d^n(\Omega \setminus F) < \varepsilon\}.$$

Nous allons montrer que $\mathscr{B}$ est une tribu contenant les pavés ouverts bornés de $\mathbb{R}^d$. Comme cette classe de parties engendre $\mathscr{B}(\mathbb{R}^d)$, ceci montrera que $\mathscr{B} = \mathscr{B}(\mathbb{R}^d)$, et donnera le résultat pour la mesure λ_d^n.

- On remarque d'abord que $\emptyset \in \mathscr{B}$: il suffit de prendre $F = \Omega = \emptyset$.
- Ensuite, si $B \in \mathscr{B}$, alors $B^c \in \mathscr{B}$: si F et Ω sont les ensembles associés à B pour un $\varepsilon > 0$, alors $\Omega^c \subset B^c \subset F^c$. Comme Ω^c est fermé, F^c est ouvert, et $\lambda_d^n(F^c \setminus \Omega^c) = \lambda_d^n(\Omega \setminus F) < \varepsilon$.
- Pour finir, si $(B_k)_{k\in\mathbb{N}}$ est une suite d'éléments de $\mathscr{B}$ et $\varepsilon > 0$, on sait qu'on peut construire une suite de fermés $(F_k)_{k\in\mathbb{N}}$ et une suite d'ouverts $(\Omega_k)_{k\in\mathbb{N}}$ telles que pour tout $k \in \mathbb{N}$, $F_k \subset B_k \subset \Omega_k$ et $\lambda_d^n(\Omega_k \setminus F_k) < \frac{\varepsilon}{2^{k+2}}$.

  On introduit ensuite l'ouvert $\Omega = \bigcup_{k\in\mathbb{N}} \Omega_k$ et les fermés $F^K = \bigcup_{k=0}^K F_k$ pour tout $K \in \mathbb{N}$. On remarque qu'on a alors

$$\lambda_d^n\Big(\Omega \setminus \bigcup_K F_K\Big) \;=\; \lambda_d^n\Big(\Omega \setminus \bigcup_k F_k\Big) \;\leqslant\; \lambda_d^n\Big(\bigcup_k \Omega_k \setminus F_k\Big) \;\leqslant\; \sum_{k\in\mathbb{N}} \lambda_d^n(\Omega_k \setminus F_k) \;<\; \frac{\varepsilon}{2}.$$

  On observe ensuite que, comme la suite $(\Omega \setminus F_K)_{K\in\mathbb{N}}$ est décroissante et la mesure λ_d^n est finie, on a

$$\lim_{K\to\infty} \lambda_d^n(\Omega \setminus F_K) \;=\; \lambda_d^n\Big(\bigcap_{K\in\mathbb{N}} \Omega \setminus F^K\Big) \;=\; \lambda_d^n\Big(\Omega \setminus \bigcup_{K\in\mathbb{N}} F^K\Big) \;<\; \frac{\varepsilon}{2}.$$

  Par conséquent, il existe $K \in \mathbb{N}$ tel que $\lambda_d^n(\Omega \setminus F^K) < \varepsilon$. Comme on a $F^K \subset \bigcup_k B_k \subset \Omega$, on a bien $\bigcup_k B_k \in \mathscr{B}$.

Nous venons de montrer que $\mathscr{B}$ est une tribu. Pour voir qu'elle contient les pavés ouverts bornés, il suffit de se donner un tel pavé $P = \prod_{i=1}^d]a_i, b_i[$ et un réel $\varepsilon > 0$, et de choisir $\Omega = P$, et $F = \prod_{i=1}^d [a_i + \eta, b_i + \eta]$, où $\eta \in]0, \min(\frac{b_i - a_i}{2})[$ est choisi assez petit pour avoir $\prod_{i=1}^d (b_i - a_i) - \prod_{i=1}^d (b_i - a_i - 2\eta) < \varepsilon$.

Nous allons maintenant en déduire le résultat pour la mesure λ_d. Donnons-nous $B \in \mathscr{B}(\mathbb{R}^d)$ et $\varepsilon > 0$, et commençons par montrer qu'il existe un ouvert Ω de $\mathbb{R}^d$ tel que $\lambda_d(\Omega \setminus B) < \frac{\varepsilon}{2}$. Pour $n \in \mathbb{N}$, on pose $B_n = B \cap]-n, n[^d$, de sorte que $B = \bigcup_n B_n$. Le résultat précédent permet d'affirmer que pour tout n, il existe un ouvert Ω_n tel que $B_n \subset \Omega_n$ et $\lambda_d^n(\Omega_n \setminus B_n) < \frac{\varepsilon}{2^{n+2}}$. Quitte à remplacer Ω_n par $\Omega_n \cap]-n, n[$, on peut supposer que $\Omega_n \subset]-n, n[$, ce qui donne en fait $\lambda(\Omega_n \setminus B_n) < \frac{\varepsilon}{2^{n+2}}$. En introduisant l'ouvert $\Omega = \bigcup_n \Omega_n$, on a donc $B \subset \Omega$, et

$$\lambda(\Omega \setminus B) \;\leqslant\; \lambda\Big(\bigcup_n \Omega_n \setminus B_n\Big) \;\leqslant\; \sum_n \lambda(\Omega_n \setminus B_n) \;<\; \frac{\varepsilon}{2},$$

ce qui prouve l'assertion.

On peut ensuite appliquer ce dernier résultat au borélien B^c pour obtenir un ouvert ω tel que $B^c \setminus \omega$ et $\lambda_d(\omega \setminus B^c) < \frac{\varepsilon}{2}$. En considérant le fermé $F = \omega^c$, on a alors $\lambda_d(\Omega \setminus F) = \lambda_d(\Omega \setminus B) + \lambda_d(B \setminus F) < \varepsilon$ et $F \subset B \subset \Omega$, ce qui achève la preuve. □

Théorème 3.13 – *La mesure λ_d est régulière, c'est-à-dire que pour tout $B \in \mathscr{B}(\mathbb{R}^d)$,*

i. $\lambda(B) \;=\; \inf\{\lambda(\Omega),\ \Omega \supset B,\ \Omega$ *est ouvert*$\}$,

ii. $\lambda(B) \;=\; \sup\{\lambda(K),\ K \subset B,\ K$ *est compact*$\}$.

Démonstration. Si $B \in \mathscr{B}(\mathbb{R}^d)$ est tel que $\lambda(B) < \infty$, alors pour tout ε, le lemme 3.12 affirme qu'il existe un ouvert $\Omega \supset B$ tel que $\lambda(B) \leqslant \lambda(\Omega) \leqslant \lambda(B) + \varepsilon$. En passant à l'infimum sur les ouverts contenant B, et en laissant tendre ε vers 0, on obtient *i.* Par ailleurs, si $\lambda(B) = \infty$, alors *i.* est trivialement vérifié.

De même que ci-dessus, le lemme 3.12 montre que $\lambda(B) = \sup\{\lambda(F),\ F \subset B,\ K \text{ est fermé}\}$. Or, si F est un fermé vérifiant $F \subset B$ et $K_n = F \cap [-n, n]^d$ pour tout $n \in \mathbb{N}$, alors $\lim_{n\to\infty} \lambda_d(K_n) = \lambda_d(F)$, ce qui montre que $\sup\{\lambda(F),\ F \subset B,\ K \text{ est fermé}\} = \sup\{\lambda(K),\ K \subset B,\ K \text{ est compact}\}$. □

Nous verrons au chapitre 6 une autre preuve de ce résultat, basée sur la définition de la mesure de Lebesgue.

3.4 Ensembles négligeables

Définition 3.14 – Si $(E, \mathscr{A}, \mu)$ est un espace mesuré, on dit qu'un ensemble $N \subset E$ est μ-négligeable (ou négligeable) s'il existe $A \in \mathscr{A}$ tel que

$$\begin{cases} N \subset A, \\ \mu(A) = 0. \end{cases}$$

On dit qu'une propriété $\mathscr{P}$ sur E est vraie μ-presque partout sur E (on note μ-p.p.) si

$$\{x \in E,\ x \text{ ne vérifie pas la propriété } \mathscr{P}\} \text{ est } \mu\text{-négligeable.}$$

Lorsque le contexte est clair, on dira aussi que $\mathscr{P}$ est vraie presque partout (on note p.p.). À titre d'exemples :

- la fonction $\mathbb{1}_{\mathbb{Q}}$ est nulle λ-presque partout,
- la fonction $\mathbb{1}_{[0,\infty[} : \mathbb{R} \to \mathbb{R}$ est continue λ-presque partout,
- la suite de fonction $\left(\mathbb{1}_{[0,\frac{1}{n}]}\right)_{n \in \mathbb{N}^\star}$ converge λ-presque partout vers 0.

Remarque 3.15 – Revenons sur les deux premiers exemples ci-dessus : les fonctions $\mathbb{1}_{\mathbb{Q}}$ et $\mathbb{1}_{\mathbb{R}_+}$ sont toutes deux continues λ-presque partout, mais seule la première est égale presque partout à une fonction continue (la fonction nulle). Il faudra prendre garde à ne pas confondre ces deux propriétés.

Si $(E, \mathscr{A}, \mu)$ est un espace mesuré et $A \in \mathscr{A}$ est tel que $\mu(A) = 0$, la propriété de croissance de la mesure semble suggérer qu'on sait donner un sens à la mesure de tout sous-ensemble de A, et que cette mesure devrait être nulle. En d'autres termes, tout ensemble μ-négligeable devrait avoir une mesure nulle également. Il peut arriver cependant qu'un tel sous-ensemble n'appartienne pas à la tribu $\mathscr{A}$, et ne puisse donc être associé à une mesure. On dit alors que l'espace mesuré n'est pas complet. Nous verrons plus loin que c'est le cas par exemple de l'espace mesuré $(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$.

Toutefois, nous verrons au paragraphe 6.3 qu'on peut toujours introduire une tribu $\widetilde{\mathscr{A}}$ contenant $\mathscr{A}$, et prolonger la mesure μ en une mesure $\tilde{\mu}$ sur $(E, \widetilde{\mathscr{A}})$ de manière à obtenir un espace mesuré complet. De cette façon, nous prolongerons la tribu $\mathscr{B}(\mathbb{R})$ en une tribu $\widetilde{\mathscr{B}}(\mathbb{R})$ qui vérifiera $\text{Card } \widetilde{\mathscr{B}}(\mathbb{R}) = \text{Card } \mathscr{P}(\mathbb{R})$. On appelle $\widetilde{\mathscr{B}}(\mathbb{R})$ la tribu de Lebesgue.

CHAPITRE 4

Fonctions mesurables

Dans l'introduction, nous avons évoqué une condition pour pouvoir définir l'intégrale d'une fonction $f : E \to \mathbb{R}$ avec l'approche de la théorie de Lebesgue, où $(E, \mathscr{A})$ est un espace mesurable. La condition que nous avons rencontrée peut se formuler de la manière suivante :

$$f^{-1}([a,b[) \in \mathscr{A} \quad \text{pour tous} \quad a, b \in \mathbb{R}.$$

Il s'agit précisément de la propriété de mesurabilité de la fonction f, comme nous allons le voir. Cette condition sera cruciale pour pouvoir associer à f une intégrale sur E.

Plus généralement, dans ce chapitre, $(E, \mathscr{A})$ et $(F, \mathscr{B})$ désigneront deux espaces mesurables, et f une fonction de E dans F. La condition de mesurabilité sur f impose que $f^{-1}(B) \in \mathscr{A}$ pour tout $B \in \mathscr{B}$. Nous allons voir comment montrer qu'une fonction donnée est mesurable, ainsi que les propriétés de ces fonctions.

4.1 Mesurabilité

Définition 4.1 – Une fonction $f : E \to F$ est dite $(\mathscr{A}, \mathscr{B})$-*mesurable*, ou plus simplement *mesurable* ssi $f^{-1}(\mathscr{B}) \subset \mathscr{A}$. Autrement dit,

$$\forall B \in \mathscr{B},\ f^{-1}(B) \in \mathscr{A}.$$

Si les ensembles E et F sont des espaces topologiques munis de leurs tribus boréliennes respectives, on appelle les fonctions mesurables des fonctions *boréliennes*.

Souvent, l'espace d'arrivée F est $\mathbb{R}$, $\mathbb{R}_+$, $\overline{\mathbb{R}}$, $\mathbb{C}$, ou $\mathbb{R}^d$. Dans ce cas, sauf mention contraire, il sera sous-entendu que F est muni de sa tribu borélienne.

Exemples

- Les fonctions constantes $f : (E, \mathscr{A}) \to (F, \mathscr{B})$ sont mesurables : si $f \equiv c \in F$ sur E tout entier, alors pour tout $B \in \mathscr{B}$, $f^{-1}(B) = \begin{cases} \emptyset & \text{si } c \notin B \\ E & \text{si } c \in B \end{cases}$. Dans tous les cas, $f^{-1}(B) \in \mathscr{A}$.
- Soit $A \subset E$, alors la fonction $\mathbb{1}_A : (E, \mathscr{A}) \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ est mesurable si et seulement si $A \in \mathscr{A}$.

Remarque 4.2 – Dans le cas où $\mathscr{A} = \mathscr{P}(E)$, toutes les fonctions de $(E, \mathscr{A})$ dans $(F, \mathscr{B})$ sont mesurables. C'est le cas par exemple des suites de réels, qui peuvent être vues comme des fonctions de $(\mathbb{N}, \mathscr{P}(\mathbb{N}))$ dans $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$.

Mesure image

Définition-Proposition 4.3 – Si $f : (E, \mathscr{A}) \to (F, \mathscr{B})$ mesurable, alors μ_f définie par

$$\mu_f(B) = \mu(f^{-1}(B)), \quad \text{pour tout } B \in \mathscr{B}$$

définit une mesure sur $(F, \mathscr{B})$, qu'on appelle *mesure image* par f de μ.

Démonstration. On remarque tout d'abord que μ_f est bien définie du fait que f est mesurable, et définit bien une application de $\mathscr{B}$ dans $\overline{\mathbb{R}}_+$. Par ailleurs, $\mu(f^{-1}(\emptyset)) = \mu(\emptyset) = 0$, et si $(B_n)_n$ est une suite d'ensembles deux à deux disjoints de $\mathscr{B}$, alors $\mu(f^{-1}(\bigsqcup_n B_n)) = \mu(\bigsqcup_n f^{-1}(B_n)) = \sum_n \mu_f(B_n)$, d'où le résultat. □

4.2 Montrer la mesurabilité

Dans la pratique, il est difficile de montrer directement qu'une fonction $f : (E, \mathscr{A}) \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ (ou plus généralement $f : (E, \mathscr{A}) \to (F, \sigma(\mathscr{C}))$ avec $\mathscr{C} \subset \mathscr{P}(F)$) est mesurable, du fait qu'on ne peut pas décrire simplement les éléments d'une tribu à partir des éléments d'une tribu qui l'engendre. Une telle description est possible en ayant recours à un procédé de récurrence transfinie. Ce procédé est beaucoup trop compliqué pour une telle tâche, et n'est pas l'objet de ce cours.
La proposition suivante montre qu'on peut se restreindre à montrer la propriété sur une classe de parties qui engendre la tribu d'arrivée.

Proposition 4.4 – *On suppose que $\mathscr{B} = \sigma(\mathscr{C})$ pour $\mathscr{C} \subset \mathscr{P}(F)$. Alors, $f : (E, \mathscr{A}) \to (F, \mathscr{B})$ est mesurable si et seulement si $f^{-1}(\mathscr{C}) \subset \mathscr{A}$.*

Démonstration. Il est clair que si $f^{-1}(\mathscr{B}) \subset \mathscr{A}$ alors $f^{-1}(\mathscr{C}) \subset f^{-1}(\mathscr{B}) \subset \mathscr{A}$. D'autre part, si $f^{-1}(\mathscr{C}) \subset \mathscr{A}$, alors $\sigma(f^{-1}(\mathscr{C})) \subset \mathscr{A}$ puisque $\mathscr{A}$ est une tribu contenant $f^{-1}(\mathscr{C})$. On déduit alors du le lemme 2.19 que $f^{-1}(\sigma(\mathscr{C})) \subset \mathscr{A}$, soit $f^{-1}(\mathscr{B}) \subset \mathscr{A}$. □

Application 4.5 – Toute fonction $f : \mathbb{R} \to \mathbb{R}$ monotone est mesurable.

Démonstration. Il suffit de remarquer que si I est un intervalle de $\mathbb{R}$, alors $f^{-1}(I)$ est un intervalle de $\mathbb{R}$, donc un borélien. Comme les boréliens de $\mathbb{R}$ sont engendrés par les intervalles, f est mesurable. □

Dans le cas où $F = \mathbb{R}$, on utilisera les notations suivantes :

$$\{f < a\} := f^{-1}(]-\infty, a[) \qquad \{f \leqslant a\} := f^{-1}(]-\infty, a]),$$
$$\{f > a\} := f^{-1}(]a, +\infty[) \qquad \{f \geqslant a\} := f^{-1}([a, +\infty[),$$

D'après la proposition 2.12, $f : (E, \mathscr{A}) \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ est mesurable si et seulement si l'une des propriétés suivantes est vérifiée :

$$\forall a \in \mathbb{R}, \{f < a\} \in \mathscr{A}, \qquad \forall a \in \mathbb{R}, \{f \leqslant a\} \in \mathscr{A},$$
$$\forall a \in \mathbb{R}, \{f > a\} \in \mathscr{A}, \qquad \forall a \in \mathbb{R}, \{f \geqslant a\} \in \mathscr{A}.$$

La même remarque vaut pour le cas où l'espace d'arrivée est $\overline{\mathbb{R}}$: pour montrer que $f : (E, \mathscr{A}) \to (\overline{\mathbb{R}}, \mathscr{B}(\overline{\mathbb{R}}))$ est mesurable, il suffit par exemple de vérifier que $\{f \leqslant a\} := f^{-1}([-\infty, a]) \in \mathscr{A}$ pour tout $a \in \mathbb{R}$.

Continuité et mesurabilité

On considère deux espaces topologiques $(E, \mathscr{T})$ et $(F, \mathscr{S})$, que l'on munit de leurs tribus boréliennes respectives $\mathscr{B}(E)$ et $\mathscr{B}(F)$.

Proposition 4.6 – *Une fonction $f : E \to F$ continue est borélienne (i.e.$(\mathscr{B}(E), \mathscr{B}(F))$-mesurable).*

Démonstration. Comme f est continue, on sait que si Ω est un ouvert de F, alors $f^{-1}(\Omega)$ est un ouvert de E, donc un borélien de E. Ainsi, $f^{-1}(\mathscr{S}) \subset \mathscr{B}(E)$, et f est borélienne d'après la proposition 4.4. □

Application 4.7 – La proposition 4.6 fournit une nouvelle démonstration de la proposition 2.14, qui dit que la tribu $\mathscr{B}(\mathbb{R}^d)$ est stable par translation. En effet, si $a \in \mathbb{R}^d$ et $f : x \in \mathbb{R}^d \mapsto x - a$, alors $B + a = f^{-1}(B) \in \mathscr{B}(\mathbb{R}^d)$, du fait que f est continue donc borélienne.

On dit qu'une fonction $f : E \to \mathbb{R}$ est semi-continue inférieurement (s.c.i) si pour tout $a \in \mathbb{R}$, l'ensemble $\{f \leqslant a\}$ est fermé. On dit par ailleurs que f est *semi-continue supérieurement* (s.c.s) si pour tout $a \in \mathbb{R}$, l'ensemble $\{f \geqslant a\}$ est fermé.

Proposition 4.8 – *Les fonctions s.c.i. et les fonctions s.c.s. sont des fonctions boréliennes.*

Démonstration. C'est immédiat : pour tout $a \in \mathbb{R}$, $\{f \leqslant a\}$ est un borélien, donc f est mesurable. □

4.3 Propriétés des fonctions mesurables

Proposition 4.9 (*Composition des fonctions mesurables*) – *Soient $(E, \mathscr{A})$, $(F, \mathscr{B})$ et $(G, \mathscr{C})$ trois espaces mesurables. Si $f : E \to F$ et $g : F \to G$ sont mesurables, alors $g \circ f$ est mesurable.*

Démonstration. Immédiat : $(g \circ f)^{-1}(\mathscr{C}) = f^{-1}(g^{-1}(\mathscr{C})) \subset \mathscr{A}$ car $g^{-1}(\mathscr{C}) \subset \mathscr{B}$. □

Application 4.10 – En combinant la proposition 4.6 avec la proposition 4.9, on obtient par exemple :
- si $f : E \to \mathbb{R}$ est mesurable, alors $|f|$ est mesurable,
- si $f : E \to \mathbb{R}^\star$ est mesurable, alors $\frac{1}{f}$ est mesurable, etc.

Proposition 4.11 (*Couple de fonctions mesurables*) – *Soit $f : x \in E \mapsto (f_1(x), f_2(x)) \in \mathbb{R}^2$. La fonction f est mesurable si et seulement si les fonctions f_1 et f_2 le sont.*

Démonstration. — Supposons que f est mesurable. On a $f_i = \pi_i \circ f$, où $\pi_i : (x_1, x_2) \in \mathbb{R}^2 \mapsto x_i$ $(i = 1, 2)$. Comme les fonctions π_i sont continues et f est mesurable, les f_i sont mesurables.

— Supposons que f_1 et f_2 sont mesurables. On sait que $\mathscr{B}(\mathbb{R}^2) = \sigma\big(\{I_1 \times I_2,\ I_1, I_2 \text{ intervalles de } \mathbb{R}\}\big)$. On remarque que
$$f^{-1}(I_1 \times I_2) \;=\; {f_1}^{-1}(I_1) \cap {f_2}^{-1}(I_2) \;\in\; \mathscr{A}$$
car les f_i sont mesurables. La proposition 4.4 implique donc que f est mesurable. □

Proposition 4.12 – *Si f et g sont des fonctions mesurables de $(E, \mathscr{A})$ dans $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, alors*

1. *$\alpha f + g$ est mesurable, pour tout $\alpha \in \mathbb{R}$,*
2. *fg est mesurable.*

En d'autres termes, les fonctions mesurables de $(E, \mathscr{A})$ dans $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ forment une $\mathbb{R}$-algèbre.

Démonstration.
1. On peut écrire $\alpha f + g = \psi \circ \varphi$, où $\varphi(x) = (f(x), g(x)) \in \mathbb{R}^2$ pour tout $x \in \mathbb{R}$, et $\psi(x, y) = \alpha x + y$ pour tout $(x, y) \in \mathbb{R}^2$.
La fonction ψ est continue de $\mathbb{R}^2$ dans $\mathbb{R}$, donc elle est $(\mathscr{B}(\mathbb{R}^2), \mathscr{B}(\mathbb{R}))$-mesurable. Comme f et g sont mesurables, la proposition 4.11 implique que la fonction φ est mesurable. Par conséquent, la fonction $\alpha f + g$ est mesurable comme composée de fonctions mesurables.
2. On raisonne de même en prenant $\psi : (x, y) \mapsto xy$, qui est aussi continue. □

Exemple 4.13 – Les fonctions de la forme $\sum_{i=1}^N \alpha_i \mathbb{1}_{A_i}$, où $\alpha_i \in \mathbb{R}$ et $A_i \in \mathscr{A}$ pour tout $i = 1, \ldots, N$, sont des fonctions mesurables, comme combinaisons linéaires de fonctions mesurables. On les appelle des fonctions *étagées*.

Proposition 4.14 – *Si $f : (E, \mathscr{A}) \to (\mathbb{C}, \mathscr{B}(\mathbb{C}))$, f est mesurable si et seulement si $\mathfrak{Re}\ f$ et $\mathfrak{Im}\ f$ le sont, et les points 1. (avec $\alpha \in \mathbb{C}$) et 2. de la proposition 4.12 restent vrais.*

Démonstration. Si $\mathfrak{Re}\ f$ et $\mathfrak{Im}\ f$ sont mesurables, il suffit de remarquer que $\mathfrak{Re}\ f + i\mathfrak{Im}\ f = \psi \circ \varphi$, où $\varphi(x) = (\mathfrak{Re}\ f(x), \mathfrak{Im}\ f(x))$ pour tout $x \in \mathbb{R}$, et $\psi(x, y) = x + iy$ pour tout $(x, y) \in \mathbb{R}^2$. Comme φ est mesurable, et ψ est continue donc mesurable, on obtient que f est mesurable.
Si f est mesurable, il suffit de voir que $\mathfrak{Re}\ f = \mathfrak{Re}\ \circ f$ et $\mathfrak{Im}\ f = \mathfrak{Im}\ \circ f$. Comme $\mathfrak{Re}$ et $\mathfrak{Im}$ sont des fonctions continues de $\mathbb{C}$ dans $\mathbb{R}$, on obtient bien que $\mathfrak{Re}\ f$ et $\mathfrak{Im}\ f$ sont mesurables.
Les points 1. et 2. de la proposition 4.12 dans le cas présent s'obtiennent avec la même démonstration. □

4.4 Suites de fonctions mesurables à valeurs réelles

PROPOSITION 4.15 – *Soit $(f_n)_n$ une suite de fonctions mesurables de $(E, \mathscr{A})$ dans $(\overline{\mathbb{R}}, \mathscr{B}(\overline{\mathbb{R}}))$. Alors*

1. $\inf_n f_n : x \in E \mapsto \inf_n f_n(x)$ *et* $\sup_n f_n : x \in E \mapsto \sup_n f_n(x)$ *sont mesurables,*
2. $\liminf_n f_n : x \in E \mapsto \liminf_n f_n(x)$ *et* $\limsup_n f_n : x \in E \mapsto \limsup_n f_n(x)$ *sont mesurables,*
3. *Si la suite $(f_n)_n$ converge simplement vers f sur E, alors f est mesurable.*

Rappel. Si $(a_n)_{n\in\mathbb{N}}$ est une suite de $\overline{\mathbb{R}}$, on remarque que $(\inf_{k\geqslant n} a_k)_n$ est une suite croissante, et que $(\sup_{k\geqslant n} a_k)_n$ est une suite décroissante. On note

$$\liminf_n a_n := \lim_{n\to\infty} \inf_{k\geqslant n} a_k = \sup_{n\in\mathbb{N}} \inf_{k\geqslant n} a_k,$$

$$\limsup_n a_n := \lim_{n\to\infty} \sup_{k\geqslant n} a_k = \inf_{n\in\mathbb{N}} \sup_{k\geqslant n} a_k.$$

On note aussi $\underline{\lim}$ et $\overline{\lim}$ pour $\liminf_n$ et $\limsup_n$.
On peut montrer que $\liminf_n a_n$ est la plus petite valeur d'adhérence de $(a_n)_n$ dans $\overline{\mathbb{R}}$, et que $\limsup_n a_n$ est la plus grande valeur d'adhérence de $(a_n)_n$ dans $\overline{\mathbb{R}}$.
Exemple. $\liminf_n(-1)^n = -1$, et $\limsup_n(-1)^n = 1$.
On observe également que si $\liminf_n a_n = \limsup_n a_n =: a \in \overline{\mathbb{R}}$, alors la suite $(a_n)_n$ converge, et $\lim\limits_{n\to\infty} a_n = a$.

Démonstration. 1. Soit $a \in \mathbb{R}$, on a :

— $\{\inf_n f_n < a\} = \bigcup_n \{f_n < a\} \in \mathscr{A}$ car les f_n sont mesurables, donc $\inf_n f_n$ est mesurable,

— $\{\sup_n f_n \leqslant a\} = \bigcap_n \{f_n \leqslant a\} \in \mathscr{A}$ car les f_n sont mesurables, donc $\sup_n f_n$ est mesurable.

2. Il suffit de remarquer que $\liminf_n f_n = \sup_n(\inf_{k\geqslant n} f_k)$. Par le point précédent, les fonctions $g_n := \inf_{k\geqslant n} f_k$ sont mesurables, donc $\sup_n g_n$ est mesurable.
De même, on remarque que $\limsup_n f_n = \inf_n(\sup_{k\geqslant n} f_k)$ et on conclut de la même manière.
3. Il suffit de remarquer que si la suite $(f_n)_n$ converge simplement vers f, alors $f = \liminf_n f_n$, et le point précédent permet de conclure. □

REMARQUE 4.16 – On peut en fait montrer un résultat plus général dans le cadre des fonctions à valeurs dans un espace métrique F : si $(f_n)_n$ est une suite de fonctions mesurables de E dans F qui converge simplement vers une fonction f, alors f est mesurable. Notons par ailleurs que le résultat devient faux dans le cadre plus général des fonctions à valeurs dans un espace topologique quelconque.

4.5 Approximation des fonctions mesurables à valeurs réelles

Dans ce paragraphe, nous allons voir que toute fonction $f : E \to \mathbb{R}$ mesurable peut être approchée par une suite de fonctions étagées, au sens de la convergence simple. Dans le cas où f est positive, la suite de fonctions étagées peut être choisie croissante. Ce résultat sera fondamental dans le chapitre suivant : couplé au théorème de convergence monotone, il fournira un moyen très simple d'étendre certaines propriétés vérifiées par les fonctions étagées aux fonctions mesurables.

DÉFINITION 4.17 – On appelle fonction étagée sur E toute fonction mesurable $\varphi : E \to \mathbb{R}$ qui ne prend qu'un nombre fini de valeurs.

Notons que si $\varphi : E \to \mathbb{R}$ est une fonction étagée, et si on note $\alpha_1, \ldots, \alpha_N$ ses valeurs distinctes, alors φ s'écrit

$$\sum_{i=1}^{N} \alpha_i \mathbb{1}_{\{\varphi=\alpha_i\}}.$$

Plus généralement, on voit immédiatement que les fonctions étagées sur E sont exactement les fonctions de E dans $\mathbb{R}$ qui s'écrivent sous la forme

$$\sum_{i=1}^{N} \alpha_i \mathbb{1}_{A_i}, \quad \text{où } \alpha_i \in \mathbb{R} \text{ et } A_i \in \mathscr{A} \text{ pour tout } i \in [\![1, N]\!].$$

Quitte à poser $A_i = \{\varphi = \alpha_i\}$, on peut par ailleurs toujours supposer que les A_i forment une partition de E dans cette écriture. Dans la suite, on fait toujours cette hypothèse lorsqu'on écrit une fonction étagée sous cette forme.

On note $\mathscr{E}$ l'ensemble des fonctions étagées de E dans $\mathbb{R}$, et $\mathscr{E}_+$ le sous-ensemble de $\mathscr{E}$ formé des fonctions à valeurs positives.

Proposition 4.18 – *Soit $f : E \to \mathbb{R}$ une fonction mesurable positive. Il existe une suite $(\varphi_n)_{n\in\mathbb{N}}$ de fonctions de $\mathscr{E}_+$ telles que*

- *$(\varphi_n)_n$ est croissante,*
- *pour tout $x \in E$, $\varphi_n(x) \underset{n\to\infty}{\longrightarrow} f(x)$.*

De plus, si f est bornée, alors la convergence de $(\varphi_n)_{n\in\mathbb{N}}$ vers f est uniforme.

Démonstration. On pose, pour $n \in \mathbb{N}$,

$$\varphi_n = \sum_{k=0}^{n2^n-1} \frac{k}{2^n} \mathbb{1}_{\{\frac{k}{2^n} \leqslant f < \frac{k+1}{2^n}\}} + n\mathbb{1}_{\{f \geqslant n\}}.$$

On note d'abord que $\varphi_n \in \mathscr{E}_+$ pour tout n : les fonctions φ_n ont un nombre fini de valeurs, et sont mesurables car f l'est.

- La suite (φ_n) converge simplement vers f : si $x \in E$ est tel que $f(x) = \infty$, alors $\varphi_n(x) = n$ pour tout n, donc $\varphi_n(x) \to \infty = f(x)$. Si $f(x) < \infty$, alors pour n assez grand, il existe $k \leqslant n2^n - 1$ tel que $k \leqslant 2^n f(x) < k + 1$, donc $|f(x) - \varphi_n(x)| = |f(x) - \frac{k}{2^n}| \leqslant \frac{1}{2^n} \to 0$.
- Montrons maintenant que la suite est croissante. Soit $n \in \mathbb{N}$,
  - si $f(x) \geqslant n + 1$, alors $\varphi_n(x) = n < n + 1 = \varphi_{n+1}(x)$,
  - si $f(x) < n$, on note k l'entier tel que $\frac{k}{2^n} \leqslant f(x) < \frac{k+1}{2^n}$, on a alors $\varphi_n(x) = \frac{k}{2^n}$. Comme $f(x) \geqslant \frac{2k}{2^{n+1}}$, on en déduit que $\varphi_{n+1}(x) \geqslant \frac{2k}{2^{n+1}} = \varphi_n(x)$.
  - si $n \leqslant f(x) < n + 1$, alors $\varphi_n(x) = n$, et il existe $k \geqslant n2^{n+1}$ tel que $\frac{k}{2^{n+1}} \leqslant f(x) < \frac{k+1}{2^{n+1}}$, donc $\varphi_{n+1}(x) \geqslant n = \varphi_n(x)$.

Si f est bornée et $N \geqslant \|f\|_\infty$, alors pour tout $n \geqslant N$, on a $\|f - \varphi_n\|_\infty \leqslant \frac{1}{2^n}$, ce qui montre que (φ_n) converge uniformément vers f. □

Remarque 4.19 – Dans le cas où f est seulement supposée mesurable, il existe une suite de fonctions étagées qui converge simplement vers f. Si la fonction f est bornée, la convergence est uniforme.

Pour voir ce résultat, il suffit de modifier la construction de la suite (φ_n) dans la preuve de la proposition 4.18, en posant

$$\varphi_n = \sum_{k=-n2^n}^{n2^n-1} \frac{k}{2^n} \mathbb{1}_{\{\frac{k}{2^n} \leqslant f < \frac{k+1}{2^n}\}} - n\mathbb{1}_{\{f < -n\}} + n\mathbb{1}_{\{f \geqslant n\}}.$$

Parmi les applications de ce résultat, on donne une démonstration du théorème de Lusin qui précise le lien entre fonctions mesurables et fonctions continues. Comme nous l'avons vu au paragraphe 4.2, une fonction continue de $\mathbb{R}$ dans $\mathbb{R}$ est toujours mesurable pour la tribu borélienne. Il va sans dire que la réciproque est fausse : nous avons déjà rencontré des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ mesurables et discontinues. Néanmoins, le théorème suivant permet de mettre en avant un fort lien entre continuité et mesurabilité.

Théorème 4.20 (*Théorème de Lusin*) – *Soit $f : \mathbb{R} \to \mathbb{R}$ mesurable. Pour tout $\varepsilon > 0$, il existe un fermé F de $\mathbb{R}$ tel que $\lambda(F^c) < \varepsilon$ et $f_{|F}$ est continue.*

Démonstration. On commence par prouver le résultat dans le cas où $f = \mathbb{1}_A$, où A est un borélien de $\mathbb{R}$. Le lemme 3.12 donne l'existence d'un fermé F_1 et d'un ouvert Ω tels que $F_1 \subset A \subset \Omega$ et $\lambda(\Omega \setminus F_1) < \varepsilon$. On pose $F_2 = \Omega^c$, et on introduit le fermé $F = F_1 \cup F_2$.

Nous allons montrer que f est continue sur F. On considère $x \in F$ et une suite (x_n) de F qui converge vers x. Si $x \in F_1$, alors $x_n \in F_1$ à partir d'un certain rang : sinon, (x_n) admet une sous-suite dans F_2 dont la limite x n'est pas dans F_2, ce qui contredit le caractère fermé de F_2. Ainsi, à partir d'un certain rang, $f(x_n) = 1 = f(x)$. De même, si $x \in F_2$, alors $f(x_n) = 0 = f(x)$ à partir d'un certain rang.

On en déduit que si f est étagée, alors le théorème est vrai. En effet, si $f = \sum_{i=1}^N \alpha_i \mathbb{1}_{A_i}$, où les A_i sont boréliens, alors il existe des fermés $F_1, \ldots, F_N$ tels que $\lambda({F_i}^c) < \frac{\varepsilon}{N}$ et f_i est continue sur F_i pour tout i. Si on pose $F = \bigcap_{i=1}^N F_i$, alors f est continue sur F, et $\lambda(F^c) \leqslant \sum_{i=1}^N \lambda({F_i}^c) < \varepsilon$.

Supposons maintenant que f est mesurable bornée. On sait alors qu'il existe une suite $(f_n)_{n\in\mathbb{N}^\star}$ de fonctions étagées qui converge uniformément vers f. Pour tout $n \in \mathbb{N}^\star$, il existe un fermé F_n tel que $\lambda({F_n}^c) < \frac{\varepsilon}{2^n}$ et f_n est continue sur F_n. En posant $F = \bigcap_n F_n$, on a f continue sur F car les f_n sont continues et convergent uniformément vers f sur F. D'autre part, $\lambda(F^c) \leqslant \sum_{n\in\mathbb{N}^\star} \lambda({F_n}^c) < \varepsilon$.

Maintenant, si f n'est pas bornée, on pose $g = \arctan$. On remarque que $g \circ f$ est mesurable et bornée, donc il existe un fermé F tel que $g \circ f$ est continue sur F et $\lambda(F^c) < \varepsilon$. Comme $f = \tan(g \circ f)$, on en déduit que f est continue sur F. □

Ce résultat se généralise à des espaces topologiques séparés localement compacts.

CHAPITRE 5

Intégrale de Lebesgue

Dans tout le chapitre, on se place sur un espace mesuré $(E, \mathscr{A}, \mu)$, et $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$. Nous allons définir l'intégrale d'une fonction mesurable de E dans $\mathbb{C}$, lorsqu'elle existe. Dans ce but, nous commencerons par définir l'intégrale des fonctions étagées positives, puis des fonctions mesurables positives. La définition de l'intégrale des fonctions intégrables à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$ s'en déduira.

5.1 Intégration des fonctions étagées positives

Comme nous l'avons vu dans le chapitre précédent, toute fonction étagée $\varphi \in \mathscr{E}$ est décrite de manière unique par son écriture

$$\sum_{i=1}^{N} \alpha_i \mathbb{1}_{A_i},$$

où $\alpha_1, \ldots, \alpha_N$ désignent les valeurs (deux à deux distinctes) de φ. Notons que les ensembles $A_i = \{\varphi = \alpha_i\}$ sont bien des éléments de $\mathscr{A}$, du fait que φ est mesurable. Nous allons définir l'intégrale des fonctions de $\mathscr{E}_+$ à partir de cette écriture.

Définition 5.1 – Soit $\varphi = \sum_{i=1}^{N} \alpha_i \mathbb{1}_{A_i}$ une fonction de $\mathscr{E}_+$, où $A_1, \ldots, A_N$ sont des éléments de $\mathscr{A}$ qui forment une partition de E. On définit l'intégrale de φ par rapport à la mesure μ par

$$\int_E \varphi \, \mathrm{d}\mu \ := \ \sum_{i=1}^{N} \alpha_i \mu(A_i),$$

avec la convention $\alpha_i \mu(A_i) = 0$ si $\alpha_i = 0$ et $\mu(A_i) = \infty$. On note aussi cette quantité $\int_E \varphi(x) \, \mathrm{d}\mu(x)$, ou$\int_E \varphi(x) \, \mu(\,\mathrm{d}x)$. On omet parfois de mentionner E lorsque le contexte est clair.

Remarque 5.2 – La définition de $\int \varphi \, \mathrm{d}\mu$ ne dépend pas de l'écriture $\varphi = \sum_{i=1}^{N} \alpha_i \mathbb{1}_{A_i}$ choisie, où les A_i sont dans $\mathscr{A}$ et forment une partition de E.

Démonstration. Soient $\alpha_1, \ldots, \alpha_N$ les valeurs de φ, et $A_i = \{\varphi = \alpha_i\}$. Si $\varphi = \sum_{j=1}^{M} \beta_j \mathbb{1}_{B_j}$, où les B_j sont dans $\mathscr{A}$ et forment une partition de E, alors

$$\sum_{j=1}^{M} \beta_j \mu(B_j) \ = \ \sum_{i=1}^{N} \sum_{j,\ \beta_j = \alpha_i} \beta_j \mu(B_j) \ = \ \sum_{i=1}^{N} \alpha_i \sum_{j,\ \beta_j = \alpha_i} \mu(B_j)$$

Or pour tout $i \in [\![1, n]\!]$, $\displaystyle\bigsqcup_{j,\ \beta_j = \alpha_i} B_j = \{\varphi = \alpha_i\} = A_i$, donc $\displaystyle\sum_{j,\ \beta_j = \alpha_i} \mu(B_j) = \mu(A_i)$. □

Remarque 5.3 – On note que si φ est étagée positive, alors $\int \varphi \,\mathrm{d}\mu \geqslant 0$.

Exemple 5.4 – 1. *Mesure de Dirac* : Soit $x \in E$, pour toute fonction étagée positive φ,

$$\int_E \varphi \,\mathrm{d}\delta_x = \varphi(x).$$

En effet, si $\alpha_1, \ldots, \alpha_N$ sont les valeurs distinctes de φ, et si $\varphi(x) = \alpha_{i_0}$, on a

$$\int \varphi \,\mathrm{d}\delta_x = \sum_{i=1}^N \alpha_i \underbrace{\delta_x(\{\varphi = \alpha_i\})}_{=\begin{cases} 1 & \text{si } i = i_0 \\ 0 & \text{si } i \neq i_0 \end{cases}} = \varphi(x).$$

2. *Mesure de comptage sur* $\mathbb{N}$: Soit m la mesure de comptage sur $(\mathbb{N}, \mathscr{P}(\mathbb{N}))$, si φ est étagée positive sur $\mathbb{N}$, alors

$$\int_{\mathbb{N}} \varphi \,\mathrm{d}m = \sum_{n\in\mathbb{N}} \varphi(n).$$

En effet, si φ a pour valeurs distinctes $\alpha_1, \ldots, \alpha_N$,

$$\begin{aligned} \sum_{n\in\mathbb{N}} \varphi(n) &= \sum_{i=1}^N \sum_{\substack{n\in\mathbb{N} \\ \varphi(n)=\alpha_i}} \varphi(n) = \sum_{i=1}^N \alpha_i \sum_{\substack{n\in\mathbb{N} \\ \varphi(n)=\alpha_i}} 1 \\ &= \sum_{i=1}^N \alpha_i m(\{\varphi = \alpha_i\}) = \int \varphi \,\mathrm{d}m. \end{aligned}$$

Proposition 5.5 – *Si φ et ψ sont deux fonctions étagées positives sur l'espace mesuré $(E, \mathscr{A}, \mu)$,*

1. *$\varphi + \psi \in \mathscr{E}_+$, et $\int(\varphi + \psi) \,\mathrm{d}\mu = \int \varphi \,\mathrm{d}\mu + \int \psi \,\mathrm{d}\mu$,*
2. *si $\alpha \in \mathbb{R}_+$, alors $\alpha\varphi \in \mathscr{E}_+$ et $\int \alpha\varphi \,\mathrm{d}\mu = \alpha \int \varphi \,\mathrm{d}\mu$,*
3. *si $\varphi \leqslant \psi$, alors $\int \varphi \,\mathrm{d}\mu \leqslant \int \psi \,\mathrm{d}\mu$.*

Démonstration. 1. On écrit $\varphi = \sum_{i=1}^N \alpha_i \mathbb{1}_{A_i}$ et $\psi = \sum_{j=1}^M \beta_j \mathbb{1}_{B_j}$ où les A_i et (*resp.* les B_j) forment une partition de E. Alors $\varphi + \psi = \sum_{i=1}^N \sum_{j=1}^M (\alpha_i + \beta_j) \mathbb{1}_{A_i \cap B_j}$,

$$\begin{aligned} \int (\varphi + \psi) \,\mathrm{d}\mu &= \sum_{i=1}^N \sum_{j=1}^M (\alpha_i + \beta_j) \mu(A_i \cap B_j) = \sum_{i=1}^N \alpha_i \sum_{j=1}^M \mu(A_i \cap B_j) + \sum_{j=1}^M \beta_j \sum_{i=1}^N \mu(A_i \cap B_j) \\ &= \sum_{i=1}^N \alpha_i \mu\big(\textstyle\bigsqcup_{j=1}^M A_i \cap B_j\big) + \displaystyle\sum_{j=1}^M \beta_j \mu\big(\textstyle\bigsqcup_{i=1}^N A_i \cap B_j\big) \\ &= \sum_{i=1}^N \alpha_i \mu(A_i) + \sum_{j=1}^M \beta_j \mu(B_j) = \int \varphi \,\mathrm{d}\mu + \int \psi \,\mathrm{d}\mu. \end{aligned}$$

2. Si $\varphi = \sum_{i=1}^N \alpha_i \mathbb{1}_{A_i}$ est étagée positive, $\alpha\varphi$ est étagée positive et

$$\int \alpha\varphi \,\mathrm{d}\mu = \sum_{i=1}^N \alpha\alpha_i \mu(A_i) = \alpha \sum_{i=1}^N \alpha_i \mu(A_i) = \alpha \int \varphi \,\mathrm{d}\mu. \qquad \square$$

3. Comme $\varphi \leqslant \psi$, la fonction $\psi - \varphi$ est étagée positive. On a alors par le point 1,

$$\int \psi \,\mathrm{d}\mu = \int (\varphi + (\psi - \varphi)) \,\mathrm{d}\mu = \int \varphi \,\mathrm{d}\mu + \int (\psi - \varphi) \,\mathrm{d}\mu \geqslant \int \varphi \,\mathrm{d}\mu.$$

Notation. Si φ est une fonction étagée positive et $A \in \mathscr{A}$, on note

$$\int_A \varphi \, \mathrm{d}\mu = \int_E \varphi \mathbb{1}_A \, \mathrm{d}\mu.$$

Cette dernière intégrale est bien définie : on observe que $\varphi \mathbb{1}_A \in \mathscr{E}_+$. En effet, si $\varphi = \sum_{i=1}^N \alpha_i \mathbb{1}_{A_i}$ où les A_i sont dans $\mathscr{A}$, alors $\varphi \mathbb{1}_A = \sum_{i=1}^N \alpha_i \mathbb{1}_{A \cap A_i}$.

LEMME 5.6 – *Si $(E_n)_{n \in \mathbb{N}}$ est une suite croissante de $\mathscr{A}$ vérifiant $E = \bigcup_{n \in \mathbb{N}} E_n$, alors pour toute fonction $\varphi \in \mathscr{E}_+$,*

$$\int_{E_n} \varphi \, \mathrm{d}\mu \underset{n \to \infty}{\longrightarrow} \int_E \varphi \, \mathrm{d}\mu.$$

Démonstration. On écrit $\varphi = \sum_{i=1}^N \alpha_i \mathbb{1}_{A_i}$, on a alors

$$\int_{E_n} \varphi \, \mathrm{d}\mu \; = \; \sum_{i=1}^N \alpha_i \mu(A_i \cap E_n).$$

On observe que pour tout $i \in [\![1, N]\!]$, on a $\mu(A_i \cap E_n) \underset{n \to \infty}{\longrightarrow} \mu(\bigcup_n A_i \cap E_n) = \mu(A_i)$, du fait que $(A_i \cap E_n)_n$ est une suite croissante de $\mathscr{A}$. Le résultat en découle. □

5.2 Intégration des fonctions mesurables positives

5.2.1 Définition

Dans toute la suite, on note $\mathscr{M}_+$ l'ensembles des fonctions mesurables de $(E, \mathscr{A})$ dans $(\overline{\mathbb{R}}_+, \mathscr{B}(\overline{\mathbb{R}}_+))$.

DÉFINITION 5.7 – Si $f \in \mathscr{M}_+$, on définit l'intégrale de f par rapport à la mesure μ comme

$$\int_E f \, \mathrm{d}\mu \; = \; \sup \left\{ \int_E \varphi \, \mathrm{d}\mu, \; \varphi \in \mathscr{E}_+, \; \varphi \leqslant f \right\} \; \in \; \overline{\mathbb{R}}_+.$$

Si $\int_E f \, \mathrm{d}\mu < \infty$, on dit que f est μ-intégrable (ou intégrable).
Si $A \in \mathscr{A}$, on sait que $f \mathbb{1}_A \in \mathscr{M}_+$, et on étend donc la notation du paragraphe précédent :

$$\int_A f \, \mathrm{d}\mu \; = \; \int_E f \mathbb{1}_A \, \mathrm{d}\mu.$$

On note que, par croissance de l'intégrale des fonctions de $\mathscr{E}_+$, la définition 5.7 est compatible avec la définition 5.1.

REMARQUE 5.8 – Les propriétés suivantes se déduisent immédiatement de la définition de l'intégrale.

1. (Positivité) Si $f \in \mathscr{M}_+$, alors $\int_E f \, \mathrm{d}\mu \geqslant 0$: si $\varphi \equiv 0$, alors $\varphi \in \mathscr{E}_+$, et $\varphi \leqslant f$.
2. (Croissance) Si $f, g \in \mathscr{M}_+$ et $f \leqslant g$, alors $\int_E f \, \mathrm{d}\mu \leqslant \int_E g \, \mathrm{d}\mu$. En effet, si $\varphi \in \mathscr{E}_+$ et $\varphi \leqslant f$, alors $\varphi \leqslant g$.

EXEMPLES –

1. *Mesure de Dirac* : si $x \in E$, alors pour toute fonction $f \in \mathscr{M}_+$, on a

$$\int_E f \, \mathrm{d}\delta_x \; = \; f(x).$$

En effet, $\int f \, \mathrm{d}\delta_x = \sup\{\varphi(x), \; \varphi \in \mathscr{E}_+, \; \varphi \leqslant f\} = f(x)$: d'une part, il est évident que $\int f \, \mathrm{d}\delta_x \leqslant f(x)$, d'autre part, $\varphi = f(x)\mathbb{1}_{\{f \geqslant f(x)\}} \in \mathscr{E}_+$ vérifiant bien $\varphi \leqslant f$, on obtient l'autre inégalité.

2. *Mesure de comptage* sur $(\mathbb{N}, \mathscr{P}(\mathbb{N}))$: pour toute fonction $f : \mathbb{N} \to \mathbb{N}$, on a

$$\int_{\mathbb{N}} f \, \mathrm{d}m \; = \; \sum_{n \in \mathbb{N}} f(n).$$

En effet,

– pour toute fonction $\varphi \in \mathscr{E}_+$ telle que $\varphi \leqslant f$, on a $\int \varphi \, \mathrm{d}m = \sum_n \varphi(n) \leqslant \sum_n f(n)$, donc $\int f \, \mathrm{d}m \leqslant \sum_n f(n)$,
– pour tout $n \in \mathbb{N}$, la fonction $\varphi_n = f \mathbb{1}_{[|0,n|]}$ est étagée positive, et vérifie $\varphi_n \leqslant f$, donc $\int \varphi_n \, \mathrm{d}m = \sum_{k=0}^n f(k) \leqslant \int f \, \mathrm{d}m$ pour tout n, d'où le résultat.

5.2.2 Théorème de Beppo Levi et conséquences

Théorème 5.9 (*Beppo Levi*) – *Soit $(f_n)_{n\in\mathbb{N}}$ une suite croissante de fonctions mesurables positives : pour tout $n \in \mathbb{N}$, $0 \leqslant f_n \leqslant f_{n+1}$. Alors, $f := \lim_{n\to\infty} f_n$ est mesurable positive, et*

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

Le théorème de Beppo Levi est aussi appelé théorème de convergence monotone.

Démonstration. Par la proposition 4.15, $f \in \mathscr{M}_+$.

◇ La suite $(f_n)_n$ étant croissante, on a $f_n \leqslant f$ pour tout $n \in \mathbb{N}$, donc par croissance, $\int_E f_n \, \mathrm{d}\mu \leqslant \int_E f \, \mathrm{d}\mu$, et en passant à la limite,

$$\lim_{n\to\infty} \int_E f_n \, \mathrm{d}\mu \leqslant \int_E f \, \mathrm{d}\mu.$$

◇ Montrons l'autre inégalité. Soient $\varphi \in \mathscr{E}_+$ telle que $\varphi \leqslant f$, et $\alpha \in]0,1[$. Posons

$$E_n = \{x \in E, \ f_n(x) \geqslant \alpha\varphi(x)\}.$$

On a

– pour tout n, $E_n \subset E_{n+1}$,
– $E = \bigcup_n E_n$: si $x \in E$, comme $f_n(x) \underset{n\to\infty}{\longrightarrow} f(x) > \alpha\varphi(x)$, il existe $n \in \mathbb{N}$ tel que $\alpha\varphi(x) \leqslant f_n(x)$.

Alors, par le lemme 5.6, on a

$$\int_E \alpha\varphi \, \mathrm{d}\mu = \lim_{n\to\infty} \int_{E_n} \alpha\varphi \, \mathrm{d}\mu.$$

Pour tout $n \in \mathbb{N}$, on note que $\alpha\varphi\mathbb{1}_{E_n} \leqslant f_n$ et $\alpha\varphi\mathbb{1}_{E_n} \in \mathscr{E}_+$, ce qui implique que $\int_E \alpha\varphi \, \mathrm{d}\mu \leqslant \int_E f_n \, \mathrm{d}\mu$. On en déduit que

$$\alpha \int_E \varphi \, \mathrm{d}\mu \leqslant \lim_{n\to\infty} \int_E f_n \, \mathrm{d}\mu.$$

En laissant tendre α vers 1, puis en passant au sup sur les fonctions φ de $\mathscr{E}_+$ telles que $\varphi \leqslant f$, on obtient l'inégalité souhaitée. □

Proposition 5.10 – *Si f, g deux fonctions mesurables positives, alors les propriétés suivantes sont vérifiées.*

1. *(Additivité)* $\int_E (f+g) \, \mathrm{d}\mu = \int f \, \mathrm{d}\mu + \int g \, \mathrm{d}\mu$.
2. *(Homogénéité positive) Si $\alpha \geqslant 0$, alors* $\int \alpha f \, \mathrm{d}\mu = \alpha \int f \, \mathrm{d}\mu$.

Démonstration. 1. On sait qu'il existe des suites croissantes $(\varphi_n)_n$ et $(\psi_n)_n$ de fonctions de $\mathscr{M}_+$ telles que $f = \lim_{n\to\infty} \varphi_n = f$ et $g = \lim_{n\to\infty} \psi_n = f$. Alors, $(\varphi_n + \psi_n)_n$ est une suite croissante de $\mathscr{M}_+$ qui converge vers $f+g$, et par Beppo Levi,

$$\begin{array}{ccc} \int_E (\varphi_n + \psi_n) \, \mathrm{d}\mu & \underset{n\to\infty}{\longrightarrow} & \int_E (f+g) \, \mathrm{d}\mu, \\ \| & & \\ \int_E \varphi_n \, \mathrm{d}\mu + \int_E \psi_n \, \mathrm{d}\mu & \underset{n\to\infty}{\longrightarrow} & \int_E f \, \mathrm{d}\mu + \int_E g \, \mathrm{d}\mu. \end{array}$$

2. On utilise les même notations que dans le point 1. La suite $(\alpha\varphi_n)_n$ est une suite croissante de $\mathscr{M}_+$ qui converge vers αf, donc par Beppo Levi,

$$\begin{array}{ccc} \int_E \alpha\varphi_n \,\mathrm{d}\mu & \underset{n\to\infty}{\longrightarrow} & \int_E \alpha f \,\mathrm{d}\mu, \\ \| & & \\ \alpha \int_E \varphi_n \,\mathrm{d}\mu & \underset{n\to\infty}{\longrightarrow} & \alpha \int_E f \,\mathrm{d}\mu. \end{array}$$

□

Corollaire 5.11 (*Relation de Chasles*) – *Si $f \in \mathscr{M}_+$ et $A, B \in \mathscr{A}$ avec $A \cap B = \emptyset$, alors*

$$\int_{A\sqcup B} f \,\mathrm{d}\mu = \int_A f \,\mathrm{d}\mu + \int_B f \,\mathrm{d}\mu.$$

Démonstration. On note que comme $A \cap B = \emptyset$, on a $f\mathbb{1}_{A\sqcup B} = f\mathbb{1}_A + f\mathbb{1}_B$. L'additivité de l'intégrale donne le résultat. □

Remarque 5.12 – Si $f, g \in \mathscr{M}_+$ avec $f \leqslant g$ et f μ-intégrable, alors $\int_E (g - f) \,\mathrm{d}\mu = \int_E g \,\mathrm{d}\mu - \int_E f \,\mathrm{d}\mu$. En effet, on remarque que $g = f + (g - f)$, et le résultat découle de l'additivité de l'intégrale sur $\mathscr{M}_+$.

Remarque 5.13 – Le théorème de Beppo Levi devient faux si la suite de fonctions est décroissante. Par exemple, si $f_n = \mathbb{1}_{[n,\infty[}$, on a $\int_\mathbb{R} f_n d\lambda = \infty$ pour tout $n \in \mathbb{N}$, mais $\int_\mathbb{R} \lim_n f_n \,\mathrm{d}\lambda = 0$.
En revanche, si on suppose en plus qu'il existe une fonction g μ-intégrable telle que $f_n \leqslant g$ pour tout $n \in \mathbb{N}$, alors la conclusion de Beppo Levi reste vraie. En effet, on peut appliquer le théorème de Beppo Levi à $(g - f_n)_n$ qui est bien une suite croissante de fonctions de $\mathscr{M}_+$. On obtient

$$\int_E \lim_n (g - f_n) \,\mathrm{d}\mu \;=\; \lim_n \int_E (g - f_n) \,\mathrm{d}\mu,$$

ce qui donne $\int_E g \,\mathrm{d}\mu - \int_E \lim_n f_n \,\mathrm{d}\mu = \int_E g \,\mathrm{d}\mu - \lim_n \int_E f_n \,\mathrm{d}\mu$.

Le résultat suivant, qui porte sur l'intégration des séries de fonctions de $\mathscr{M}_+$, est équivalent au théorème de Beppo Levi.

Proposition 5.14 – *Si $(f_n)_{n\in\mathbb{N}}$ est une suite de fonctions mesurables positives, alors*

$$\int_E \sum_{n\in\mathbb{N}} f_n \,\mathrm{d}\mu \;=\; \sum_{n\in\mathbb{N}} \int_E f_n \,\mathrm{d}\mu.$$

Démonstration. On note que la suite $\left(\sum_{k=0}^n f_k\right)_n$ est une suite croissante de $\mathscr{M}_+$, donc par le théorème de Beppo Levi,

$$\begin{array}{ccc} \displaystyle\int_E \sum_{k=0}^n f_k \,\mathrm{d}\mu & \underset{n\to\infty}{\longrightarrow} & \displaystyle\int_E \sum_{k\in\mathbb{N}} f_k \,\mathrm{d}\mu, \\ = & & \\ \displaystyle\sum_{k=0}^n \int_E f_k \,\mathrm{d}\mu & \underset{n\to\infty}{\longrightarrow} & \displaystyle\sum_{k\in\mathbb{N}} \int_E f_k \,\mathrm{d}\mu, \end{array}$$

où on a utilisé l'additivité de l'intégrale des fonctions de $\mathscr{M}_+$. □

Mesures à densité

Proposition 5.15 – *Si $g \in \mathscr{M}_+$ et pour tout $A \in \mathscr{A}$,*

$$\nu(A) \;=\; \int_A g \,\mathrm{d}\mu,$$

alors ν définit une mesure sur $(E, \mathscr{A})$. On dit que ν est la mesure à densité g par rapport à la mesure μ. Par ailleurs, si $f \in \mathscr{M}_+$, alors

$$\int_E f \,\mathrm{d}\nu \;=\; \int_E fg \,\mathrm{d}\mu. \tag{5.1}$$

Démonstration. On a bien sûr $\nu(\emptyset) = 0$. Par ailleurs, si $(A_n)_n$ est une suite d'éléments de $\mathscr{A}$ distincts deux à deux, alors

$$\nu\Big(\bigsqcup_n A_n\Big) \;=\; \int g \sum_{n\in\mathbb{N}} \mathbb{1}_{A_n}\,\mathrm{d}\mu \;=\; \sum_{n\in\mathbb{N}} \int g\mathbb{1}_{A_n}\,\mathrm{d}\mu \;=\; \sum_{n\in\mathbb{N}} \nu(A_n),$$

par Beppo Levi, ce qui prouve que ν est bien une mesure.

On note que, par définition de ν, l'égalité (5.1) est vraie si f est une indicatrice d'élément de $\mathscr{A}$. Par additivité de l'intégrale, elle reste donc vraie si $f \in \mathscr{E}_+$. Pour finir, si $f \in \mathscr{M}_+$, alors il existe une suite croissante $(g_n)_n$ de fonctions de $\mathscr{E}_+$ qui converge simplement vers f. Comme $(f_n g)_n$ est aussi une suite croissante de fonctions mesurables, on a par Beppo Levi

$$\begin{array}{ccc} \displaystyle\int f_n\,\mathrm{d}\nu & \underset{n\to\infty}{\longrightarrow} & \displaystyle\int f\,\mathrm{d}\nu \\ \| & & \\ \displaystyle\int f_n g\,\mathrm{d}\mu & \underset{n\to\infty}{\longrightarrow} & \displaystyle\int fg\,\mathrm{d}\mu, \end{array}$$

d'où le résultat. □

Exemple 5.16 – Une variable aléatoire réelle X sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$ suit la loi normale $\mathscr{N}(\mu, \sigma)$ si sa loi $\mathbb{P}_X$ est la mesure qui a pour densité la fonction $x \mapsto \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2}}$, ce qui donne pour tout $A \in \mathscr{B}(\mathbb{R})$,

$$\mathbb{P}_X(A) \;=\; \int_A \frac{e^{-\frac{(x-\mu)^2}{2}}}{\sigma\sqrt{2\pi}}\,\mathrm{d}\lambda(x).$$

5.2.3 Inégalité de Markov et conséquences

Lemme 5.17 (*Inégalité de Markov*) – *Soient $f \in \mathscr{M}_+$ et $a \in \mathbb{R}_+^\star$, on a*

$$\mu(\{f \geqslant a\}) \;\leqslant\; \frac{1}{a}\int_E f\,\mathrm{d}\mu.$$

Démonstration. Comme $f \geqslant a\mathbb{1}_{\{f\geqslant a\}}$, la croissance de l'intégrale donne

$$\int_E f\,\mathrm{d}\mu \geqslant \int_E a\;\mathbb{1}_{\{f\geqslant a\}}\,\mathrm{d}\mu = a\;\mu(\{f \geqslant a\}),$$

d'où le résultat. □

Corollaire 5.18 – *Si $f \in \mathscr{M}_+$ et $\int_E f\,\mathrm{d}\mu < \infty$, alors f est finie μ-presque partout (ie $\mu(\{f = \infty\}) = 0$).*

Démonstration. Pour tout $n \in \mathbb{N}^\star$, on a

$$\mu(\{f = \infty\}) \;\leqslant\; \mu(\{f \geqslant n\}) \;\leqslant\; \frac{1}{n}\int_E f\,\mathrm{d}\mu \underset{n\to\infty}{\longrightarrow} 0,$$

par l'inégalité de Markov, d'où $\mu(\{f = \infty\}) = 0$. □

Corollaire 5.19 –
1. *Si $f \in \mathscr{M}_+$, alors $\int_E f\,\mathrm{d}\mu = 0$ si et seulement si $f = 0$ μ-p.p.*
2. *Si $f, g \in \mathscr{M}_+$ et $f = g$ μ-p.p., alors $\int_E f\,\mathrm{d}\mu = \int_E g\,\mathrm{d}\mu$.*

Remarque 5.20 – Si $f \in \mathscr{M}_+$ et $A \in \mathscr{A}$ avec $\mu(A) = 0$, alors $\int_A f\,\mathrm{d}\mu = 0$ (car $f\mathbb{1}_A = 0$ μ-p.p.).

Démonstration. 1. — Supposons que $f \in \mathscr{E}_+$ et $f = 0$ μ-p.p. On peut alors écrire $f = \sum_{i=1}^N \alpha_i \mathbb{1}_{A_i}$, où $\mu(A_i) = 0$ dès que $\alpha_i > 0$. Par conséquent, $\int_E f\,\mathrm{d}\mu = \sum_{i=1}^N \alpha_i\mu(A_i) = 0$.
Si maintenant $f \in \mathscr{M}_+$ et $f = 0$ μ-p.p., alors pour toute fonction $\varphi \in \mathscr{E}_+$ telle que $\varphi \leqslant f$, $\varphi = 0$ μ-p.p., ce qui implique que $\int_E \varphi\,\mathrm{d}\mu = 0$. On a donc bien montré que $\int_E f\,\mathrm{d}\mu = 0$.

— Si maintenant $\int_E f \, \mathrm{d}\mu = 0$, alors pour tout $n \in \mathbb{N}^\star$, $\mu(\{f \geqslant \frac{1}{n}\}) \leqslant n \int f \, \mathrm{d}\mu = 0$ par l'inégalité de Markov. Par conséquent,

$$\mu(\{f \neq 0\}) \;=\; \mu\Bigg(\bigcup_{n \in \mathbb{N}^\star} \left\{ f \geqslant \frac{1}{n} \right\} \Bigg) \;\leqslant\; \sum_{n \in \mathbb{N}^\star} \mu\left(\left\{ f \geqslant \frac{1}{n} \right\} \right) \;=\; 0.$$

2. Si $f = g$ μ-p.p., alors $f \mathbb{1}_{\{f \neq g\}}$ et $g \mathbb{1}_{\{f \neq g\}}$ sont nulles μ-p.p., ce qui implique par 1. que $\int_{\{f \neq g\}} f \, \mathrm{d}\mu = \int_{\{f \neq g\}} g \, \mathrm{d}\mu = 0$. On en déduit alors que

$$\int_E f \, \mathrm{d}\mu \;=\; \int_{\{f=g\}} f \, \mathrm{d}\mu \;=\; \int_{\{f=g\}} g \, \mathrm{d}\mu \;=\; \int_E g \, \mathrm{d}\mu.$$ □

5.2.4 Lemme de Fatou

Proposition 5.21 (*Lemme de Fatou*) – *Si $(f_n)_{n \in \mathbb{N}}$ est une suite de fonctions mesurables positives, alors*

$$\int_E \liminf_n f_n \, \mathrm{d}\mu \;\leqslant\; \liminf_n \int_E f_n \, \mathrm{d}\mu.$$

Démonstration. Soit $g_n = \inf_{k \geqslant n} f_k$. La suite $(g_k)_k$ est une suite croissante de $\mathscr{M}_+$, donc par le théorème de Beppo Levi, $\displaystyle\int_E \lim_{n \to \infty} g_n \, \mathrm{d}\mu = \lim_{n \to \infty} \int_E g_n \, \mathrm{d}\mu$, soit

$$\int_E \liminf_n f_n \;=\; \lim_{n \to \infty} \int_E \inf_{k \geqslant n} f_k \, \mathrm{d}\mu.$$

Il suffit donc de montrer qu'on a $\int_E \inf_{k \geqslant n} f_k \, \mathrm{d}\mu \leqslant \inf_{k \geqslant n} \int_E f_k \, \mathrm{d}\mu$. Pour tout $k_0 \geqslant n$, on a $\inf_{k \geqslant n} f_k \leqslant f_{k_0}$, ce qui implique que $\int_E \inf_{k \geqslant n} f_k \leqslant \int_E f_{k_0}$. Par passage à l'infimum pour $k_0 \geqslant n$, on obtient bien le résultat. □

Remarque 5.22 – Attention, le résultat n'est valable que pour des fonctions positives. On verra plus loin que si $f_n = \mathbb{1}_{[0,1]} - \mathbb{1}_{[n,n+1]}$, alors pour tout n, $\int_\mathbb{R} f_n \, \mathrm{d}\lambda = 0$, mais $\int_\mathbb{R} \liminf_n f_n \, \mathrm{d}\lambda = 1$.

Remarque 5.23 – On n'a ni $\int_E \limsup_n f_n \, \mathrm{d}\mu \leqslant \limsup_n \int_E f_n \, \mathrm{d}\mu$, ni $\limsup_n \int_E f_n \, \mathrm{d}\mu \leqslant \int_E \limsup_n f_n \, \mathrm{d}\mu$ en général.

En revanche, on a le résultat suivant : si $(f_n)_n$ est une suite de $\mathscr{M}_+$ telle que pour tout $n \in \mathbb{N}$, $f_n \leqslant g$, où $g \in \mathscr{M}_+$ est μ-intégrable, alors

$$\int_E \limsup_n f_n \, \mathrm{d}\mu \;\geqslant\; \limsup_n \int_E f_n \, \mathrm{d}\mu.$$

En effet, on peut appliquer le lemme de Fatou à la suite de fonctions $g - f_n$ qui appartiennent à $\mathscr{M}_+$:

$$\begin{aligned} \int_E \liminf_n (g - f_n) \, \mathrm{d}\mu \;&\leqslant\; \liminf_n \int_E (g - f_n) \, \mathrm{d}\mu \\ \int_E g \, \mathrm{d}\mu - \int_E \limsup_n f_n \, \mathrm{d}\mu \;&\leqslant\; \int_E g \, \mathrm{d}\mu - \limsup_n \int_E f_n \, \mathrm{d}\mu, \end{aligned}$$

d'où l'inégalité.

5.3 Fonctions intégrables

5.3.1 L'espace $\mathscr{L}^1(E, \mathscr{A}, \mu)$

Définition 5.24 – Une fonction $f : (E, \mathscr{A}) \to (\mathbb{K}, \mathscr{B}(\mathbb{K}))$ est dite μ-intégrable (ou intégrable) si elle est mesurable et $\int_E |f| \, \mathrm{d}\mu < \infty$. On note

$$\mathscr{L}^1(E, \mathscr{A}, \mu) \;=\; \{f : (E, \mathscr{A}) \to (\mathbb{K}, \mathscr{B}(\mathbb{K})),\; f \text{ est } \mu\text{-intégrable}\}.$$

$\mathscr{L}^1(E,\mathscr{A},\mu)$ est parfois seulement noté $\mathscr{L}^1_\mu(E)$, voire $\mathscr{L}^1(E)$ lorsqu'il n'y a pas d'ambiguïté.

On étend la définition aux fonctions à valeurs dans $\overline{\mathbb{R}}$, ou un sous-ensemble de $\mathbb{K}$ (muni de la tribu des boréliens).

Notation. Si $f : E \to \overline{\mathbb{R}}$ est mesurable, alors on appelle respectivement partie positive et partie négative de f les fonctions f^+ et f^- données par

$$f^+ = \max(f,0) \text{ et } f^- = -\min(f,0).$$

On remarque que f^+ etf^- sont positives, et mesurables comme composées d'une fonction continue et d'une fonction mesurable. Par ailleurs, on a clairement :

$$\begin{aligned} f &= f^+ - f^-, \\ |f| &= f^+ + f^-. \end{aligned}$$

On déduit de cette écriture que f est mesurable si et seulement si f^+ et f^- le sont.

Proposition 5.25 –

1. *Si $f : E \to \overline{\mathbb{R}}$, alors*
$$f \in \mathscr{L}^1_\mu(E) \;\Leftrightarrow\; f^+, f^- \in \mathscr{L}^1_\mu(E).$$
2. *Si $f : E \to \mathbb{C}$, alors*
$$f \in \mathscr{L}^1_\mu(E) \;\Leftrightarrow\; \mathfrak{Re}\, f, \mathfrak{Im}\, f \in \mathscr{L}^1_\mu(E).$$

Démonstration. 1. Comme $|f| = f^+ + f^-$, on a

$$\int_E f^\pm \,\mathrm{d}\mu \;\leqslant\; \int_E |f|\,\mathrm{d}\mu \;=\; \int_E f^+\,\mathrm{d}\mu + \int_E f^-\,\mathrm{d}\mu$$

ce qui implique que $\int_E |f|\,\mathrm{d}\mu < \infty$ si et seulement si $\int_E f^+\,\mathrm{d}\mu < \infty$ et $\int_E f^-\,\mathrm{d}\mu < \infty$.

2. On sait déjà que $f : E \to \mathbb{C}$ est mesurable si et seulement si $\mathfrak{Re}\, f$ et $\mathfrak{Im}\, f$ le sont. Comme $|f| = \sqrt{\mathfrak{Re}\, f^2 + \mathfrak{Im}\, f^2}$, on a

$$\left.\begin{array}{l} \int_E |\mathfrak{Re}\, f|\,\mathrm{d}\mu \\ \int_E |\mathfrak{Im}\, f|\,\mathrm{d}\mu \end{array}\right\} \;\leqslant\; \int_E |f|\,\mathrm{d}\mu \;\leqslant\; \int_E |\mathfrak{Re}\, f|\,\mathrm{d}\mu + \int_E |\mathfrak{Im}\, f|\,\mathrm{d}\mu,$$

d'où le résultat. □

Définition 5.26 (*Intégrale de Lebesgue*) –

– Si $f : (E,\mathscr{A}) \to (\overline{\mathbb{R}}, \mathscr{B}(\overline{\mathbb{R}}))$ est μ-intégrable, on pose

$$\int_E f\,\mathrm{d}\mu \;:=\; \int_E f^+\,\mathrm{d}\mu - \int_E f^-\,\mathrm{d}\mu.$$

– Si $f : (E,\mathscr{A}) \to (\mathbb{C}, \mathscr{B}(\mathbb{C}))$ est μ-intégrable, on pose

$$\int_E f\,\mathrm{d}\mu \;:=\; \int_E \mathfrak{Re}\, f\,\mathrm{d}\mu + i\int_E \mathfrak{Im}\, f\,\mathrm{d}\mu.$$

Si $f \in \mathscr{L}^1_\mu(E)$ et $A \in \mathscr{A}$, alors $f\mathbb{1}_A \in \mathscr{L}^1_\mu(E)$ (du fait que $\int_A |f|\,\mathrm{d}\mu \leqslant \int_E |f|\,\mathrm{d}\mu < \infty$). On note alors comme précédemment

$$\int_A f\,\mathrm{d}\mu \;=\; \int_E f\mathbb{1}_A\,\mathrm{d}\mu.$$

Proposition 5.27 – *Si $f, g : E \to \mathbb{K}$ sont mesurables et $f = g$ $\mu-$p.p., alors*

$$f \in \mathscr{L}^1_\mu(E) \;\Leftrightarrow\; g \in \mathscr{L}^1\mu(E),$$

et dans ce cas, $\displaystyle\int_E f\,\mathrm{d}\mu = \int_E g\,\mathrm{d}\mu.$

Démonstration. Cas où $f, g : E \to \overline{\mathbb{R}}$: comme $\{f^+ \neq g^+\} \subset \{f \neq g\}$ et $\{f^- \neq g^-\} \subset \{f \neq g\}$, $f^+ = g^+ \mu$–p.p. et $f^- = g^- \mu$–p.p., d'où f^+ et f^- sont μ-intégrables ssi g^+ et g^- le sont, et dans ce cas $\int_E f \,\mathrm{d}\mu = \int_E g \,\mathrm{d}\mu$.
Cas où $f, g : E \to \mathbb{C}$: on note que $\{\mathfrak{Re}\, f \neq \mathfrak{Re}\, g\} \subset \{f \neq g\}$ et $\{\mathfrak{Im}\, f \neq \mathfrak{Im}\, g\} \subset \{f \neq g\}$, donc $\mathfrak{Re}\, f$ et $\mathfrak{Im}\, f$ sont μ-intégrables ssi $\mathfrak{Re}\, g$ et $\mathfrak{Im}\, f$ le sont, et dans ce cas $\int_E f \,\mathrm{d}\mu = \int_E g \,\mathrm{d}\mu$. □

En vue de la proposition 5.27, on s'autorisera à parler de l'intégrale de fonctions définies μ-presque partout, si elles sont égales presque partout à une fonction intégrable.

Remarque 5.28 – Si $f : E \to \overline{\mathbb{R}}$ est intégrable, alors $\mu(\{f = \pm\infty\}) = \mu(\{|f| = \infty\}) = 0$, donc f est finie μ-presque partout. Dans la suite, on se contentera donc de considérer l'intégrale de fonctions intégrables à valeurs dans $\mathbb{R}$ (quitte à remplacer f par $f\mathbb{1}_{\{|f|<\infty\}}$).

5.3.2 Propriétés des fonctions intégrables

Proposition 5.29 (*Linéarité de l'intégrale*) –

- *Si $f, g \in \mathscr{L}^1_\mu(E)$, alors $f + g \in \mathscr{L}^1_\mu(E)$ et $\int_E (f+g) \,\mathrm{d}\mu = \int_E f \,\mathrm{d}\mu + \int_E g \,\mathrm{d}\mu$.*
- *Si $f \in \mathscr{L}^1_\mu(E)$ et $\alpha \in \mathbb{K}$, alors $\alpha f \in \mathscr{L}^1_\mu(E)$ et $\int_E \alpha f \,\mathrm{d}\mu = \alpha \int_E f \,\mathrm{d}\mu$.*

Démonstration.

- Tout d'abord, si $f, g \in \mathscr{L}^1_\mu(E)$, on a $\int_E |f+g| \,\mathrm{d}\mu \leqslant \int_E |f| \,\mathrm{d}\mu + \int_E |g| \,\mathrm{d}\mu < \infty$, du fait que $|f+g| \leqslant |f| + |g|$.
  On suppose d'abord $\mathbb{K} = \mathbb{R}$. On note que

$$\begin{aligned} f+g &= (f+g)^+ - (f+g)^- \\ f+g &= f^+ - f^- + g^+ - g^- \end{aligned}$$

  On en déduit que $(f+g)^+ + f^- + g^- = (f+g)^- + f^+ + g^+$. Comme les fonctions des membres de gauche et de droite sont mesurables positives, on a

$$\int_E (f+g)^+ \,\mathrm{d}\mu + \int_E f^- \,\mathrm{d}\mu + \int_E g^- \,\mathrm{d}\mu = \int_E (f+g)^- \,\mathrm{d}\mu + \int_E f^+ \,\mathrm{d}\mu + \int_E g^+ \,\mathrm{d}\mu.$$

  Donc $\int_E (f+g)^+ \,\mathrm{d}\mu - \int_E (f+g)^- \,\mathrm{d}\mu = \underbrace{\int_E f^+ \,\mathrm{d}\mu - \int_E f^- \,\mathrm{d}\mu}_{\int_E f \,\mathrm{d}\mu} + \underbrace{\int_E g^+ \,\mathrm{d}\mu - \int_E g^- \,\mathrm{d}\mu}_{\int_E g \,\mathrm{d}\mu}.$

  Si $\mathbb{K} = \mathbb{C}$, on a $\int_E (f+g) \,\mathrm{d}\mu = \int_E (\mathfrak{Re}\, f + \mathfrak{Re}\, g) \,\mathrm{d}\mu + i \int_E (\mathfrak{Im}\, f + \mathfrak{Im}\, g) \,\mathrm{d}\mu$, et le résultat se déduit du point précédent.
- On note d'abord que si $f \in \mathscr{L}^1_\mu(E)$ et $\alpha \in \mathbb{K}$, alors on a $\int_E |\alpha f| \,\mathrm{d}\mu = |\alpha| \int_E |f| \,\mathrm{d}\mu < \infty$.
  Supposons $\mathbb{K} = \mathbb{R}$.
  - Si $\alpha \geqslant 0$, alors $(\alpha f)^+ = \alpha f^+$ et $(\alpha f)^- = \alpha f^-$, d'où le résultat.
  - Si $\alpha < 0$, alors $(\alpha f)^+ = -\alpha f^-$ et $(\alpha f)^- = -\alpha f^+$, d'où le résultat.

  Si $\mathbb{K} = \mathbb{C}$ et $\alpha = a + ib$, alors

$$\begin{aligned} \int_E \alpha f \,\mathrm{d}\mu &= \int_E (a\mathfrak{Re}\, f - b\mathfrak{Im}\, f) \,\mathrm{d}\mu + i \int_E (a\mathfrak{Im}\, f + b\mathfrak{Re}\, f) \,\mathrm{d}\mu \\ &= (a+ib) \int_E \mathfrak{Re}\, f \,\mathrm{d}\mu + i(a+ib) \int_E \mathfrak{Im}\, f \,\mathrm{d}\mu = \alpha \int_E f \,\mathrm{d}\mu. \end{aligned}$$ □

Corollaire 5.30 (*Relation de Chasles*) – *Si $A, B \in \mathscr{A}$ et $A \cap B = \emptyset$, alors*

$$\int_{A \sqcup B} f \,\mathrm{d}\mu = \int_A f \,\mathrm{d}\mu + \int_B f \,\mathrm{d}\mu.$$

Démonstration. On déduit le résultat de l'égalité $f\mathbb{1}_{A \sqcup B} = f\mathbb{1}_A + f\mathbb{1}_B$ et de la linéarité de l'intégrale. □

PROPOSITION 5.31 (*Croissance*) – *Si $f, g : E \to \mathbb{R}$ sont intégrables et $f \leqslant g$, alors $\int_E f \, \mathrm{d}\mu \leqslant \int_E g \, \mathrm{d}\mu$.*

Démonstration. On a $g - f \in \mathscr{M}_+$, donc $\int_E (g-f) \, \mathrm{d}\mu \geqslant 0$. On conclut par la linéarité de l'intégrale. □

REMARQUE 5.32 – Pour avoir la propriété de croissance, il suffit en fait de supposer que $f \leqslant g$ μ–p.p. : il suffit de considérer $f\mathbb{1}_{\{f\leqslant g\}}$ et $g\mathbb{1}_{\{f\leqslant g\}}$ au lieu de f et g.

PROPOSITION 5.33 (*Inégalité triangulaire*) – *Si $f \in \mathscr{L}^1_\mu(E)$, alors*

$$\left| \int_E f \, \mathrm{d}\mu \right| \leqslant \int_E |f| \, \mathrm{d}\mu.$$

Démonstration. – Si $\mathbb{K} = \mathbb{R}$, on a

$$\left| \int f \, \mathrm{d}\mu \right| = \left| \int f^+ \, \mathrm{d}\mu - \int f^- \, \mathrm{d}\mu \right| \leqslant \int f^+ \, \mathrm{d}\mu + \int f^- \, \mathrm{d}\mu = \int |f| \, \mathrm{d}\mu.$$

– Si $\mathbb{K} = \mathbb{C}$, on commence par remarquer que si $\int_E f \, \mathrm{d}\mu = 0$, l'inégalité est vérifiée. Sinon, on remarque que $\left| \int_E f \, \mathrm{d}\mu \right| = \int_E f \frac{|\int_E f \, \mathrm{d}\mu|}{\int_E f \, \mathrm{d}\mu} \, \mathrm{d}\mu$. En posant $\alpha = \frac{|\int_E f \, \mathrm{d}\mu|}{\int_E f \, \mathrm{d}\mu}$, on a

$$\begin{aligned}
\left| \int_E f \, \mathrm{d}\mu \right| &= \int_E (\alpha f) \, \mathrm{d}\mu = \int_E \mathfrak{Re}\,(\alpha f) \, \mathrm{d}\mu && \text{car } \left| \textstyle\int_E f \, \mathrm{d}\mu \right| \in \mathbb{R} \\
&\leqslant \int_E |\mathfrak{Re}\,(\alpha f)| \, \mathrm{d}\mu && \text{par le cas } \mathbb{K} = \mathbb{R} \\
&\leqslant \int_E |\alpha f| \, \mathrm{d}\mu = \int_E |f| \, \mathrm{d}\mu.
\end{aligned}$$

□

5.3.3 Formule de transfert

Soient $(E, \mathscr{A})$ et $(F, \mathscr{B})$ et φ une fonction mesurable de E dans F. On rappelle que μ_φ définie par

$$\mu_\varphi(B) = \mu(\varphi^{-1}(B)), \quad B \in \mathscr{B},$$

est une mesure sur $(F, \mathscr{B})$ appelée mesure image de μ par φ.
Le lemme suivant peut-être vu comme un résultat de changement de variable.

PROPOSITION 5.34 (*Formule de transfert*) – *Soit f une fonction mesurable de F dans $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$). La fonction f est μ_φ-intégrable sur F si et seulement si la fonction $f \circ \varphi$ est μ-intégrable sur E, et dans ce cas on a*

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

Démonstration. Voir TD. □

APPLICATION 5.35 – On peut déjà obtenir une première formule de changement de variable très simple : si $\varphi : x \in \mathbb{R} \mapsto x + a$, pù $a \in \mathbb{R}$, alors $\int_{\mathbb{R}} f \, \mathrm{d}\lambda = \int_{\mathbb{R}} f(x+a) \, \mathrm{d}\lambda(x)$, pour toute fonction $f \in \mathscr{L}^1(\mathbb{R})$.
En effet, pour tout $B \in \mathscr{B}(\mathbb{R})$, $\lambda_\varphi(B) = \lambda(B - a) = \lambda(B)$, donc $\lambda_\varphi = \lambda$, et la formule de transfert donne le résultat.

APPLICATION 5.36 (*Calcul de l'espérance d'une variable aléatoire*) – Soit X une variable aléatoire réelle sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$. L'espérance de X est définie par $\mathbb{E}(X) = \int_\Omega X \, \mathrm{d}\mathbb{P}$. Le lemme de transfert implique alors que

$$\mathbb{E}(X) = \int_{\mathbb{R}} x \, \mathrm{d}\mathbb{P}_X(x),$$

où on a remplacé μ par $\mathbb{P}$, φ par X et f par $\mathrm{Id} : \mathbb{R} \to \mathbb{R}$ dans la formule de transfert.
N.B. : Dans le cas particulier où X est à valeurs dans $\mathbb{N}$, on a bien la formule classique de l'espérance :

$$\mathbb{E}(X) = \int_E X \, \mathrm{d}\mathbb{P} = \sum_{k\in\mathbb{N}} \int_{\{X=k\}} X \, \mathrm{d}\mathbb{P} = \sum_{k\in\mathbb{N}} k \, \mathbb{P}(\{X = k\}).$$

5.4 Théorème de convergence dominée et applications

5.4.1 Théorème de convergence dominée

Théorème 5.37 – *Soient f_n, $n \in \mathbb{N}$, et f des fonctions mesurables de E dans $\mathbb{K}$ telles que*

(i) *la suite $(f_n)_n$ converge vers f μ-p.p. (i.e. $f_n(x) \underset{n\to\infty}{\longrightarrow} f(x)$ pour μ-presque tout $x \in E$),*

(ii) *il existe $g \in \mathscr{L}^1_\mu(E)$ telle que pour tout $n \in \mathbb{N}$, $|f_n| \leqslant g$ μ-p.p.*

Alors, les fonctions f_n, f sont intégrables, et

$$\int_E |f - f_n| \, \mathrm{d}\mu \underset{n\to\infty}{\longrightarrow} 0.$$

En particulier, $\int_E f_n \, \mathrm{d}\mu \underset{n\to\infty}{\longrightarrow} \int_E f \, \mathrm{d}\mu$.

La démonstration repose sur l'idée donnée dans la remarque 5.23 : on note qu'il suffit de montrer l'inégalité

$$\int_E \limsup_n |f - f_n| \, \mathrm{d}\mu \;\geqslant\; \limsup_n \int_E |f - f_n| \, \mathrm{d}\mu$$

pour obtenir le résultat. D'après la remarque 5.23, il suffit de "dominer" la suite $|f - f_n|$ par une fonction intégrable pour obtenir l'inégalité. On va voir que les hypothèses du lemme de Fatou permettent de le faire.

Démonstration. Soit $A = \{x \in E,\ f_n(x) \underset{n\to\infty}{\longrightarrow} f(x) \text{ et } \forall n \in \mathbb{N},\ |f_n(x)| \leqslant g(x)\}$. On a

$$\mu(A^c) \;=\; \mu\Big(\{f_n \underset{n\to\infty}{\not\longrightarrow} f\} \cup \bigcup_{n\in\mathbb{N}} \{|f_n| > g\}\Big) \;\leqslant\; \mu(\{f_n \underset{n\to\infty}{\not\longrightarrow} f\}) + \sum_{n\in\mathbb{N}} \mu(\{|f_n| > g\}) \;=\; 0.$$

On remarque que si $x \in A$, alors $|f(x) - f_n(x)| \leqslant |f(x)| + |f_n(x)| \leqslant 2g(x)$, ce qui implique que $\mathbb{1}_A(2g - |f - f_n|) \in \mathscr{M}_+$. Par le lemme de Fatou, on a

$$\underbrace{\int_A \liminf_n \,(2g - |f - f_n|) \, \mathrm{d}\mu}_{=\int_A 2g \, \mathrm{d}\mu} \;\leqslant\; \underbrace{\liminf_n \int_A (2g - |f - f_n|) \, \mathrm{d}\mu}_{=\int_A 2g \, \mathrm{d}\mu - \limsup_n \int_A |f - f_n| \, \mathrm{d}\mu},$$

du fait que $\liminf\limits_n \,(2g - |f - f_n|) = 2g$ sur A. On obtient alors

$$\limsup_n \int_A |f - f_n| \, \mathrm{d}\mu \leqslant 0$$

et $\liminf_n \int_A |f - f_n| \, \mathrm{d}\mu = \limsup_n \int_A |f - f_n| \, \mathrm{d}\mu = 0$, donc $\lim_{n\to\infty} \int_E |f - f_n| \, \mathrm{d}\mu = 0$. □

Remarque 5.38 – Il existe des situations où $\int \lim_n f_n \, \mathrm{d}\mu = \lim_n \int f_n \, \mathrm{d}\mu$, mais la suite $(f_n)_n$ n'est dominée par aucune fonction intégrable (*i.e.* (*ii*) n'est pas vraie dans le théorème de convergence dominée). *Exemple :* $f_n = n\mathbb{1}_{[\frac{1}{n+1}, \frac{1}{n}[}$ dans $(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$.

Corollaire 5.39 – *Soit $(f_n)_n$ une suite de fonctions mesurables de E dans $\mathbb{K}$ vérifiant $\sum_n \int |f_n| \, \mathrm{d}\mu < \infty$. Alors, la fonction $\sum_n f_n$ est définie presque partout, μ-intégrable, et*

$$\sum_{n\in\mathbb{N}} \int_E f_n \, \mathrm{d}\mu \;=\; \int_E \sum_{n\in\mathbb{N}} f_n \, \mathrm{d}\mu.$$

Démonstration. On observe d'abord que comme $\sum_n \int |f_n| \, \mathrm{d}\mu < \infty$, les fonctions f_n sont en fait intégrables. Par ailleurs, par Beppo Levi, on a

$$\int_E \sum_{n\in\mathbb{N}} |f_n| \, \mathrm{d}\mu \;=\; \sum_{n\in\mathbb{N}} \int_E |f_n| \, \mathrm{d}\mu \;<\; \infty. \tag{5.2}$$

Par conséquent, la fonction $\sum_n |f_n|$ est finie μ-presque partout, donc $\sum_n f_n(x)$ converge absolument pour μ-presque tout $x \in E$. On en déduit que la fonction $\sum_n f_n$ est bien définie μ-presque partout.
On pose par exemple $\sum_n f_n(x) = 0$ lorsque $x \in A := \{\sum_n |f_n| = \infty\} \in \mathscr{A}$. La fonction $\sum_n f_n$ est alors mesurable comme limite simple de la suite de fonctions mesurables $(\sum_{k=0}^n f_k \mathbb{1}_A)_n$. Par ailleurs, $\int |\sum_n f_n| \,\mathrm{d}\mu \leqslant \int \sum_n |f_n| < \infty$, donc $\sum_n f_n$ est intégrable.
On applique le théorème de convergence dominée à la suite de fonctions $(g_n)_n$, où $g_n = \sum_{k=0}^n f_k$. En effet, on a

- $g_n \underset{n\to\infty}{\longrightarrow} \sum_n f_n$ μ-p.p.,
- $|g_n| \leqslant \sum_n |f_n|$ μ-p.p.

On en déduit que

$$\begin{array}{ccc} \displaystyle\int_E \sum_{k=0}^n f_k \,\mathrm{d}\mu & \underset{n\to\infty}{\longrightarrow} & \displaystyle\int_E \sum_{n\in\mathbb{N}} f_n \,\mathrm{d}\mu \\ = & & = \\ \displaystyle\sum_{k=0}^n \int_E f_k \,\mathrm{d}\mu & \underset{n\to\infty}{\longrightarrow} & \displaystyle\sum_{n\in\mathbb{N}} \int_E f_n \,\mathrm{d}\mu, \end{array}$$

d'où le résultat. □

5.4.2 Intégrales dépendant d'un paramètre

Limite et continuité sous l'intégrale

On considère un espace métrique (Y, d) et une fonction $f : E \times Y \to \mathbb{K}$. On s'intéresse à des propriétés de limite et de continuité de la fonction

$$F : y \in Y \mapsto \int_E f(x, y) \,\mathrm{d}\mu$$

lorsqu'elle est bien définie.

Proposition 5.40 – *Soit $\bar{y} \in Y$. On suppose qu'il existe une fonction mesurable $\ell : E \to \mathbb{K}$ telle que*

(i) pour tout $y \in Y$, $x \mapsto f(x, y)$ est mesurable,
(ii) pour μ-presque tout $x \in E$, $f(x, y) \underset{y\to\bar{y}}{\longrightarrow} \ell(x)$,
(iii) il existe $g \in \mathscr{L}^1_\mu(E)$ telle que pour tout $y \in Y$,

$$|(f(x, y)| \leqslant g(x) \quad p.p.\ x.$$

Alors F est bien définie sur Y entier, et $F(y) \underset{y\to\bar{y}}{\longrightarrow} \int_E \ell \,\mathrm{d}\mu$.

Démonstration. Soit $(y_k)_{k\in\mathbb{N}}$ une suite de Y telle que $\lim_{k\to\infty} y_k = \bar{y}$. Pour tout $k \in \mathbb{N}$, on pose

$$\varphi_k(x) = f(x, y_k).$$

Par hypothèse, les fonctions φ_k sont mesurables, la suite (φ_k) converge μ-presque partout vers ℓ, et pour tout $k \in \mathbb{N}$, $|\varphi_k| \leqslant g$ μ-presque partout. Comme $g \in \mathscr{L}^1_\mu(E)$, on peut appliquer le théorème de convergence dominée à la suite de fonctions (φ_k), et on obtient que

$$F(y) \underset{y\to\bar{y}}{\longrightarrow} \int_E \ell \,\mathrm{d}\mu.$$

□

Corollaire 5.41 (*Continuité sous l'intégrale*) – *Soit $\bar{y} \in Y$. On suppose :*

(i) pour tout $y \in Y$, $x \mapsto f(x, y)$ est mesurable,
(ii) pour presque tout $x \in E$, $y \mapsto f(x, y)$ est continue en $\bar{y} \in Y$,
(iii) il existe $g \in \mathscr{L}^1_\mu(E)$ telle que pour tout $y \in Y$,

$$|f(x, y)| \leqslant g(x) \quad p.p.\ x.$$

Alors la fonction $F : y \in Y \mapsto \int_E f(x, y)\, \mathrm{d}\mu$ *est bien définie sur* Y *tout entier, et est continue en* $\bar{y}$*.*

Démonstration. On pose $\ell : x \in E \mapsto f(x, \bar{y})$. Par hypothèse, ℓ est une fonction mesurable. On peut donc appliquer le théorème précédent, qui donne

$$F(y) \underset{y \to \bar{y}}{\longrightarrow} \int_E \ell\, \mathrm{d}\mu \;=\; F(\bar{y}). \qquad \square$$

Dérivation sous l'intégrale

Soient I un intervalle non vide de $\mathbb{R}$ et une fonction $f : E \times I \to \mathbb{R}$. On s'interroge maintenant sur le lien entre la dérivabilité de la fonction $y \mapsto f(x, y)$ pour presque tout $x \in E$ et celle de la fonction $y \mapsto \int_E f(x, y)\, \mathrm{d}\mu(x)$ lorsqu'elle existe.

PROPOSITION 5.42 (*Dérivation sous l'intégrale*) – *On suppose :*

(*i*) *pour tout* $y \in I$*,* $x \mapsto f(x, y)$ *est intégrable,*

(*ii*) *pour presque tout* $x \in E$*,* $y \mapsto f(x, y)$ *est dérivable sur* I*,*

(*iii*) *il existe* $g \in \mathscr{L}^1_\mu(E)$ *telle que pour presque tout* $x \in E$*,*

$$|\partial_y(x, y)| \leqslant g(x) \quad \forall y \in I,$$

où $\frac{\partial f}{\partial y}(x, \cdot)$ *désigne la dérivée de la fonction* $y \mapsto f(x, y)$*. Alors la fonction* $F : y \in I \mapsto \int_E f(x, y)\, \mathrm{d}\mu$ *est dérivable sur* I*, et pour tout* $y \in I$*,*

$$F'(y) = \int_E \partial_y(x, y)\, \mathrm{d}\mu.$$

Démonstration. Soients $\bar{y} \in I$ et une suite $(y_n)_{n \in \mathbb{N}}$ de I telle que $\lim_{n\to\infty} y_n$ et $y_n \neq y$ pour tout n.On considère la suite de fonctions φ_n définie par

$$\varphi_n(x) \;=\; \frac{f(x, y_n) - f(x, \bar{y})}{y_n - \bar{y}}.$$

On note que par hypothèse, les fonctions φ_n sont intégrables, et $\varphi_n(x) \underset{n\to\infty}{\longrightarrow} \partial_y f(x, \bar{y})$ μ-p.p. x. Par ailleurs, pour μ-presque tout $x \in E$, le théorème des accroissements finis donne

$$|\varphi_n(x)| \;\leqslant\; \sup_{y \in [y_n, \bar{y}]} |\partial_y f(x, y)| \;\leqslant\; |g(x)|.$$

On peut alors appliquer le théorème de convergence dominée, qui donne

$$\frac{F(y_n) - F(\bar{y})}{y_n - \bar{y}} \;=\; \int_E \varphi_n(x)\, \mathrm{d}\mu(x) \underset{n\to\infty}{\longrightarrow} \int_E \partial_y f(x, y)\, \mathrm{d}\mu(x)$$

et qui conclut la preuve. $\square$

REMARQUE 5.43 – On note que si on remplace dérivable par dérivable par C^1 dans les hypothèses de la proposition 5.42, on obtient le résultat tient bien sûr toujours, et on obtient en plus que F est de classe C^1 sur I. Il suffit en effet d'appliquer le théorème de continuité sous l'intégrale à la fonction $y \mapsto \int_E \partial_y f(x, y)\, \mathrm{d}\mu(x)$.

5.5 Lien entre l'intégrale de Riemann et l'intégrale de Lebesgue

5.5.1 Intégrale de Riemann

On se place sur un intervalle $[a, b] \subset \mathbb{R}$. On rappelle que $\varphi : [a, b] \to \mathbb{R}$ est dite *en escalier* si elle est de la forme $\varphi = \sum_{i=1}^N \alpha_i \mathbb{1}_{I_i}$, où $I_1, \ldots, I_N$ sont des intervalles de $[a, b]$. On note $\mathscr{E}sc$ l'ensemble des fonctions en escalier définies sur $[a, b]$. Remarquons que toute fonction en escalier sur $[a, b]$ est en particulier une fonction étagée sur $[a, b]$, mais l'inverse n'est pas vrai.

On définit alors naturellement l'intégrale de $f = \sum_{i=1}^{N} \alpha_i \mathbb{1}_{I_i} \in \mathscr{E}sc$ sur $[a,b]$ au sens de Riemann par

$$\int_a^b f(x)\,\mathrm{d}x \;=\; \sum_{i=1}^{N} \alpha_i \ell(I_i) \;=\; \int_{[a,b]} f\,\mathrm{d}\lambda.$$

Définition 5.44 – Si $f : [a,b] \to \mathbb{R}$ est une fonction bornée, on note

$$\overline{\int_a^b} f(x)\,\mathrm{d}x \;:=\; \sup\left\{\int_a^b \varphi(x)\,\mathrm{d}x,\ \varphi \leqslant f,\ \varphi \in \mathscr{E}sc\right\},$$

$$\underline{\int_a^b} f(x)\,\mathrm{d}x \;:=\; \inf\left\{\int_a^b \varphi(x)\,\mathrm{d}x,\ \varphi \geqslant f,\ \varphi \in \mathscr{E}sc\right\}.$$

Si $\overline{\int_a^b} f(x)\,\mathrm{d}x = \underline{\int_a^b} f(x)\,\mathrm{d}x$, on dit que f est Riemann-intégrable sur $[a,b]$, et on note

$$\int_a^b f(x)\,\mathrm{d}x = \overline{\int_a^b} f(x)\,\mathrm{d}x = \underline{\int_a^b} f(x)\,\mathrm{d}x.$$

Remarque 5.45 – On voit aisément que la définition ci-dessus est équivalente à la définition classique de l'intégrale de Riemann par sommes de Riemann, ou encore à la définition par sommes de Darboux.

Théorème 5.46 – *Si $(f_n)_{n\in\mathbb{N}}$ est une suite de fonctions Riemann-intégrables de $[a,b]$ dans $\mathbb{R}$ qui converge uniformément vers f, alors f est Riemann-intégrable, et*

$$\int_a^b f(x)\,\mathrm{d}x \;=\; \lim_{n\to\infty} \int_a^b f_n(x)\,\mathrm{d}x.$$

Définition 5.47 – On appelle fonction réglée sur $[a,b]$ toute fonction $f : [a,b] \to \mathbb{R}$ qui est limite uniforme d'une suite de fonctions en escalier.

On peut montrer que les fonctions réglées sur $[a,b]$ sont exactement les fonctions qui admettent une limite à gauche et à droite en tout point. Il est donc aisé de voir que toutes les fonctions continues, les fonctions continues par morceaux, les fonctions monotones sont des fonctions réglées sur $[a,b]$. On peut aussi montrer que les fonctions réglées ont un nombre dénombrable de discontinuités.

Une conséquence immédiate du théorème 5.46 est que toute fonction réglée est Riemann-intégrable. Il est alors naturel de se demander si ce sont les seules. Le théorème suivant répond par la négative à cette question, et caractérise les fonctions Riemann-intégrables sur un segment $[a,b]$.

Théorème 5.48 (*Critère de Lebesgue*) – *Une fonction $f : [a,b] \to \mathbb{R}$ bornée est Riemann-intégrable si et seulement si elle est continue λ-presque partout.*

Le critère de Lebesgue permet de voir immédiatement qu'il existe des fonctions Riemann-intégrables qui ne sont pas réglées. On peut citer l'exemple de l'indicatrice de l'ensemble de Cantor C sur $[0,1]$: la fonction est discontinue sur C qui n'est pas dénombrable, donc $\mathbb{1}_C$ n'est pas réglée. En revanche, $\mathbb{1}_C$ est Riemann-intégrable car l'ensemble de ses points de discontinuité est C, qui est de mesure nulle. Un autre exemple est a fonction f définie sur $[0,1]$ par $f(x) = \sin\frac{1}{x}$ si $x > 0$, et $f(0) = 0$. Elle n'est pas réglée puisqu'elle n'admet pas de limite à droite en 0, mais elle est Riemann-intégrable puisqu'elle est continue sur $]0,1]$ qui est de mesure pleine dans $[0,1]$.

5.5.2 Fonctions Riemann-intégrables et fonctions Lebesgue-intégrables

Théorème 5.49 – *Si $f : [a,b] \to \mathbb{R}$ est Riemann-intégrable, alors il existe une fonction $g \in \mathscr{L}^1_\lambda([a,b])$ telle que $f = g$ λ-presque partout, et*

$$\int_a^b f(x)\,\mathrm{d}x \;=\; \int_{[a,b]} g\,\mathrm{d}\lambda.$$

Pour toute fonction $g \in \mathscr{L}^1_\lambda([a,b])$, on notera alors sans amiguité $\int_{[a,b]} g \, \mathrm{d}\lambda = \int_a^b g(x) \, \mathrm{d}x$.

Démonstration. Soit $(\varphi_n)_{n\in\mathbb{N}}$ une suite de $\mathscr{E}sc$ telle que pour tout $n \in \mathbb{N}$, $\varphi_n \leqslant f$, et

$$\int_{[a,b]} \varphi_n \, \mathrm{d}\lambda \underset{n\to\infty}{\longrightarrow} \int_a^b f(x) \, \mathrm{d}x.$$

On remarque que, quitte à remplacer φ_n par $\max(\varphi_0, \ldots, \varphi_n)$, on peut supposer que la suite $(\varphi_n)_n$ est croissante. On introduit alors sa limite simple g, qui est mesurable comme limite simple d'une suite de fonctions mesurables, et qui vérifie donc $g \leqslant f$.
Comme la suite $(\varphi_n - \varphi_0)_n$ est une suite croissante de $\mathscr{M}_+$, on peut lui appliquer le théorème de Beppo Levi. Il vient alors

$$\int_{[a,b]} (\varphi_n - \varphi_0) \, \mathrm{d}\lambda \underset{n\to\infty}{\longrightarrow} \int_{[a,b]} (g - \varphi_0) \, \mathrm{d}\lambda.$$

Comme $\varphi_0 \in \mathscr{L}^1_\lambda([a,b])$, on en déduit que $\int_{[a,b]} \varphi_n \, \mathrm{d}\lambda \underset{n\to\infty}{\longrightarrow} \int_{[a,b]} g \, \mathrm{d}\lambda$, et donc $\int_{[a,b]} g \, \mathrm{d}\lambda = \int_a^b f(x) \, \mathrm{d}x$.

On considère maintenant une suite $(\psi_n)_{n\in\mathbb{N}}$ de $\mathscr{E}sc$ telle que pour tout $n \in \mathbb{N}$, $\psi_n \geqslant f$, et

$$\int_{[a,b]} \psi_n \, \mathrm{d}\lambda \underset{n\to\infty}{\longrightarrow} \int_a^b f(x) \, \mathrm{d}x.$$

Quitte à remplacer ψ_n par $\min(\psi_0, \ldots, \psi_n)$, on peut supposer que la suite $(\psi_n)_n$ est décroissante. On introduit alors sa limite simple h, qui est mesurable et vérifie $h \geqslant f$. On applique cette fois Beppo Levi à la suite $(\psi_0 - \psi_n)_n$, et on obtient, du fait que les fonctions h, ψ_n, ψ_0 sont intégrables ($h \leqslant \psi_n \leqslant \psi_0$ pour tout n),

$$\int_{[a,b]} \psi_n \, \mathrm{d}\lambda \underset{n\to\infty}{\longrightarrow} \int_{[a,b]} h \, \mathrm{d}\lambda.$$

On en déduit alors que $\int_{[a,b]} h \, \mathrm{d}\lambda = \int_a^b f(x) \, \mathrm{d}x$.

On remarque alors que $h - g \geqslant 0$, et $\int_{[a,b]} (h - g) \, \mathrm{d}\lambda = 0$, donc $g = h$ λ-presque partout. De ce fait, $f = g$ λ-presque partout, d'où le résultat. $\square$

Remarque 5.50 – On ne peut pas dire en général que si f est Riemann-intégrable, alors elle est intégrable au sens de Lebesgue : il faudrait pour cela qu'elle soit mesurable, ce qui n'est pas garanti. Il peut en effet arriver qu'une fonction soit presque partout égale à une fonction mesurable, mais ne soit pas mesurable elle-même. Bien sûr, si on ajoute dans les hypothèses du théorème 5.49 que f est mesurable, alors la conclusion devient sur $f \in \mathscr{L}^1([a,b])$, et $\int_a^b f(x) \, \mathrm{d}x = \int_{[a,b]} f \, \mathrm{d}\lambda$.

Une conséquence directe du théorème 5.49 est que tous les résultats connus pour l'intégrale de Riemann s'appliquent à l'intégrale de Lebesgue des fonctions Riemann-intégrables mesurables.

Par ailleurs, à partir de maintenant, lorsque nous travaillerons avec une fonction $f : [a,b] \to \mathbb{R}$ Riemann-intégrable et mesurable, nous ne ferons plus de distinction entre les écritures $\int_a^b f(x) \, \mathrm{d}x$, $\int_{[a,b]} f \, \mathrm{d}\lambda$, et on s'autorisera par exemple à écrire $\int_a^b f \, \mathrm{d}\lambda$.

Cas des intégrales généralisées

Théorème 5.51 – *On suppose que f est une fonction mesurable de $]a,b[$ dans $\mathbb{R}$ localement Riemann-intégrable et d'intégrale absolument convergente. Alors, f est Lebesgue-intégrable, et*

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

Démonstration. Soit $(a_n)_{n\in\mathbb{N}}$ une suite décroissante qui converge vers a, et $b_n)_{n\in\mathbb{N}}$ une suite croissante qui converge vers b. On note que pour tout $n \in \mathbb{N}$, on a

$$\int_{[a_n,b_n]} |f| \, \mathrm{d}\lambda = \int_{a_n}^{b_n} |f(x)| \, \mathrm{d}x.$$

Comme la suite $|f|\mathbb{1}_{[a_n,b_n]}$ est croissante, on peut appliquer Beppo Levi, et la convergence absolue de l'intégrale de Riemann de f sur $]a,b[$ donne alors

$$\int_{[a,b]} f \, \mathrm{d}\lambda \;=\; \int_a^b |f(x)| \, \mathrm{d}x \;<\; \infty.$$

On en déduit que $f \in \mathscr{L}^1([a,b])$. On peut donc appliquer le théorème de convergence dominée à la suite $\big(|f|\mathbb{1}_{[a_n,b_n]}\big)_n$, ce qui donne

$$\int_{[a_n,b_n]} f \, \mathrm{d}\lambda \underset{n\to\infty}{\longrightarrow} \int_{[a,b]} f \, \mathrm{d}\lambda.$$

La convergence de l'intégrale de Riemann de f sur $]a,b[$ fournit alors le résultat. $\square$

CHAPITRE 6

Construction de mesures - Unicité

6.1 Construction de mesures

Cas de la mesure de Lebesgue

On cherche à définir une mesure λ sur $\mathbb{R}$ qui soit conforme à la notion de longueur des intervalles de $\mathbb{R}$. En d'autres termes, si $\mathscr{I}$ désigne l'ensemble des intervalles ouverts bornés de $\mathbb{R}$, on souhaite étendre l'application

$$\begin{array}{rccc} \lambda & : & \mathscr{I} & \to & \mathbb{R}_+ \\ & &]a,b[& \mapsto & b-a \end{array}$$

à une classe plus large de parties de $\mathbb{R}$, tout en imposant que le prolongement obtenu soit une mesure sur une tribu $\mathscr{A}$ de $\mathbb{R}$.

Plus précisément, comme détaillé dans le théorème 3.9 énoncé plus haut, on souhaite définir une mesure λ sur $\mathbb{R}$ telle que :

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

(ii) λ est invariante par translation.

Le résultat suivant montre qu'il est impossible de définir une telle mesure sur $\mathscr{P}(\mathbb{R})$.

Théorème 6.1 – *Il n'existe pas de mesure sur* $(\mathbb{R}, \mathscr{P}(\mathbb{R}))$ *vérifiant (i) et (ii).*

La démonstration (voir TD) repose sur l'axiome du choix et implique en particulier l'existence de parties de $\mathbb{R}$ dites "non mesurables", c'est-à-dire qui n'appartiennent à aucune tribu de $\mathbb{R}$ sur laquelle on peut définir une mesure vérifiant (i) et (ii).

Nous allons donc construire une mesure λ sur $\mathbb{R}$ vérifiant les propriétés (i) et (ii) sur une tribu de $\mathbb{R}$ différente de $\mathscr{P}(\mathbb{R})$. Pour assurer que nous pourrons tout de même mesurer assez d'ensembles de $\mathbb{R}$, nous chercherons à construire λ sur une tribu assez large : nous imposerons qu'elle contienne les ouverts. En fait, nous allons même la construire sur la plus petite tribu qui contient les ouverts : la tribu borélienne. Nous verrons ensuite que cette définition se généralise à des tribus plus grandes. Toutefois, dans la pratique, il est largement suffisant de travailler dans $\mathscr{B}(\mathbb{R})$.

6.1.1 Mesures extérieures

On commence par définir la notion de mesure extérieure, plus faible que la notion de mesure. Une mesure extérieure sera toujours définie sur toutes les parties d'un ensemble E.

Définition 6.2 – On appelle *mesure extérieure* sur E une application $\mu^\star : \mathscr{P}(E) \to \overline{\mathbb{R}}_+$ telle que :

- $\mu^\star(\emptyset) = 0$,
- $\mu^\star$ est croissante : si $A \subset B \subset E$, $\mu^\star(A) \leqslant \mu^\star(B)$,
- $\mu^\star$ est σ-sous-additive : si pour tout $n \in \mathbb{N}$, $A_n \subset E$, alors $\mu^\star\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leqslant \sum_{n\in\mathbb{N}} \mu^\star(A_n)$.

On note que la propriété de σ-sous-additivité vérifiée par les mesures extérieures est plus faible que la propriété de σ-additivité des mesures. Par conséquent, toute mesure sur $(E, \mathscr{P}(E))$ est aussi une mesure extérieure sur E (les mesures extérieures sont toujours définies sur l'ensemble des parties).
Les propositions qui suivent permettent de voir qu'on peut restreindre toute mesure extérieure à une tribu sur laquelle elle est σ-additive, ce qui fait d'elle une mesure sur cette tribu.

Définition 6.3 – On dit qu'un ensemble $A \subset E$ est une partie $\mu^\star$-*mesurable* ssi pour tout $B \subset E$,

$$\mu^\star(B) = \mu^\star(B \cap A) + \mu^\star(B \setminus A).$$

On note $\mathscr{M}_{\mu^\star}$ l'ensemble des parties $\mu^\star$-mesurables de E.

On peut voir les élements de $\mathscr{M}_{\mu^\star}$ comme les parties de E pour lesquelles $\mu^\star$ “se comporte comme une mesure”, c'est-à-dire a une propriété d'additivité. Le théorème suivant donne un sens à cette remarque.

Théorème 6.4 –

1. $\mathscr{M}_{\mu^\star}$ est une tribu sur E.

2. $\mu^\star$ est une mesure sur $(E, \mathscr{M}_{\mu^\star})$.

Démonstration. 1. Il est clair que $\emptyset \in \mathscr{M}_{\mu^\star}$. D'autre part, si $A \in \mathscr{M}_{\mu^\star}$, alors pour tout $B \subset E$

$$\mu^\star(B) = \mu^\star(B \cap A) + \mu^\star(B \setminus A) = \mu^\star(B \setminus A^c) + \mu^\star(B \cap A^c),$$

donc $A^c \in \mathscr{M}_{\mu^\star}$. Il reste à vérifier le dernier point. Soient $(A_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathscr{M}_{\mu^\star}$, et $B \subset E$. Par définition de la mesure extérieure, on a $\mu^\star(B) \leqslant \mu^\star(B \cap \bigcup_n A_n) + \mu^\star(B\setminus \bigcup_n A_n)$. Montrons l'autre inégalité : on observe que

$$\begin{aligned}
\mu^\star(B) &= \mu^\star(B \cap A_0) + \mu^\star(B\setminus A_0)\\
&= \mu^\star(B \cap A_0) + \mu^\star((B\setminus A_0) \cap A_1) + \mu^\star(B\setminus(A_0 \cup A_1))\\
&= \mu^\star(B \cap A_0) + \mu^\star((B\setminus A_0) \cap A_1) + \mu^\star(B\setminus(A_0 \cup A_1) \cap A_2) + \mu^\star(B\setminus(A_0 \cup A_1 \cup A_2))\\
&= \ldots\\
&= \sum_{n=0}^{N} \mu^\star\left(\left(B\setminus \bigcup_{k=0}^{n-1} A_k\right) \cap A_n\right) + \mu^\star\left(B\setminus \bigcup_{n=0}^{N} A_n\right)
\end{aligned}$$

pour tout $N \in \mathbb{N}$. Alors

$$\mu^\star(B) \geqslant \sum_{n=0}^{N} \mu^\star\left(\left(B\setminus \bigcup_{k=0}^{n-1} A_k\right) \cap A_n\right) + \mu^\star\left(B\setminus \bigcup_{n\in\mathbb{N}} A_n\right)$$

par croissance de $\mu^\star$. En laissant tendre Nvers ∞, on obtient :

$$\begin{aligned}
\mu^\star(B) &\geqslant \sum_{n\in\mathbb{N}} \mu^\star\left(\left(B\setminus \bigcup_{k=0}^{n-1} A_k\right) \cap A_n\right) + \mu^\star\left(B\setminus \bigcup_{n\in\mathbb{N}} A_n\right)\\
&\geqslant \mu^\star\left(\bigcup_{n\in\mathbb{N}}\left(B\setminus \bigcup_{k=0}^{n-1} A_k\right) \cap A_n\right) + \mu^\star\left(B\setminus \bigcup_{n\in\mathbb{N}} A_n\right).
\end{aligned}$$

Comme $\bigcup_n \left(B\setminus \bigcup_{k=0}^{n-1} A_k\right) \cap A_n = \bigcup_n B \cap \left(A_n \setminus \bigcup_{k=0}^{n-1} A_k\right) = B \cap \bigcup_n A_n$, la deuxième inégalité est démontrée.

2. Seule la σ-additivité est à démontrer. Soient A_n, $n \in \mathbb{N}$ des éléments de $\mathscr{M}_{\mu^\star}$ deux à deux disjoints. Il suffit de montrer que $\mu^\star\left(\bigsqcup_{n\in\mathbb{N}} A_n\right) \geqslant \sum_{n\in\mathbb{N}} \mu^\star(A_n)$, l'autre inégalité étant vérifiée par définition de $\mu^\star$. Pour tout $m \in \mathbb{N}$, on a :

$$\mu^\star\left(\bigsqcup_{n\in\mathbb{N}} A_n\right) \geqslant \mu^\star\left(\bigsqcup_{n=0}^{m} A_n\right) = \mu^\star(A_m) + \mu^\star\left(\bigsqcup_{n=0}^{m-1} A_n\right) = \ldots = \sum_{k=0}^{m} \mu^\star(A_n).$$

En laissant tendre m vers ∞, on a bien l'inégalité désirée. □

6.1.2 La mesure de Lebesgue

Comme on l'a dit précédemment, on cherche à étendre la notion de longueur pour les intervalles de $\mathbb{R}$. On remarque dans un premier temps que, par croissance de la mesure qu'on souhaite définir, si un ensemble $A \subset \mathbb{R}$ vérifie $A \subset \bigcup_{n\in\mathbb{N}} I_n$, où $I_n \in \mathscr{I}$, alors la mesure de A sera majorée par $\sum_{n\in\mathbb{N}} \ell(I_n)$, où $\ell(I)$ désigne la longueur d'un intervalle I.

Il semble donc naturel de poser

$$\lambda^\star(A) := \inf\Big\{\sum_{n\in\mathbb{N}} \ell(I_n),\ A \subset \bigcup_{n\in\mathbb{N}} I_n,\ I_n \in \mathscr{I}\ \forall n \in \mathbb{N}\Big\},$$

pour tout $A \subset \mathbb{R}$.

Remarque 6.5 – On ne peut pas faire la même construction en prenant des intervalles inclus dans A et en prenant le sup. Une telle construction donnerait par exemple une mesure nulle à $\mathbb{R}\setminus\mathbb{Q}$ qui ne contient aucun intervalle non vide.

Remarque 6.6 – On peut remplacer $\mathscr{I}$ dans la définition de la mesure de Lebesgue par l'ensemble des intervalles bornés, ou encore par l'ensemble des intervalles.

Proposition 6.7 – *$\lambda^\star$ est une mesure extérieure sur $\mathbb{R}$.*

Démonstration. Tout d'abord, il est clair que $\lambda^\star(\emptyset) = 0$. D'autre part, si $A \subset B$, alors si $I_n \in \mathscr{I}$, $n \in \mathbb{N}$,

$$B \subset \bigcup_{n\in\mathbb{N}} I_n \ \Rightarrow\ A \subset \bigcup_{n\in\mathbb{N}} I_n,$$

donc $\lambda^\star(B) \geqslant \lambda^\star(A)$.
Soit maintenant $(A_k)_{k\in\mathbb{N}}$ une suite de parties de $\mathbb{R}$. Posons $\varepsilon > 0$, et pour chaque $k \in \mathbb{N}$, prenons une suite d'intervalles $(I_n^k)_n$ de $\mathscr{I}$ telle que

$$\sum_{n\in\mathbb{N}} \ell(I_n^k) \ \leqslant\ \lambda^\star(A_k) + \frac{\varepsilon}{2^{k+1}}.$$

Alors $(I_n^k)_{(k,n)\in\mathbb{N}^2}$ est un recouvrement dénombrable de $\bigcup_{k\in\mathbb{N}} A_k$ par des éléments de $\mathscr{I}$, et

$$\lambda^\star\left(\bigcup_{k\in\mathbb{N}} A_k\right) \ \leqslant\ \sum_{(k,n)\in\mathbb{N}^2} \ell(I_n^k) \ \leqslant\ \sum_{k\in\mathbb{N}}\Big(\lambda^\star(A_k) + \frac{\varepsilon}{2^{k+1}}\Big) \ =\ \sum_{k\in\mathbb{N}} \lambda^\star(A_k) + \varepsilon.$$

Comme l'inégalité tient pour tout $\varepsilon > 0$, on obtient bien l'inégalité souhaitée en laissant tendre ε vers 0. □

Proposition 6.8 –

1. *$\lambda^\star([0,1]) = 1$,*
2. *$\lambda^\star$ est invariante par translation.*

Démonstration.

1. Pour commencer, $\lambda^\star([0,1]) \leqslant 1$: pour tout $\varepsilon > 0$, $[0,1] \subset]-\varepsilon, 1+\varepsilon[$ donc $\lambda^\star([0,1]) \leqslant 1 + 2\varepsilon$ pour tout $\varepsilon > 0$, on obtient l'inégalité en laissant tendre ε vers 0.
   Montrons l'autre inégalité. Supposons que $[0,1] \subset \bigcup_{i\in\mathbb{N}}]a_i, b_i[$. Par compacité, il existe un entier n tel que $[0,1] \subset \bigcup_{i=0}^n]a_i, b_i[$.
   Soit i_1 tel que $0 \in]a_{i_1}, b_{i_1}[$. Si $b_{i_1} \in [0,1]$, il existe i_2 tel que $b_{i_1} \in]a_{i_2}, b_{i_2}[$. On répète ensuite l'opération pour construire les i_k suivants avec $b_{i_{k-1}} \in]a_{i_k}, b_{i_k}[$, jusqu'à obtenir $b_{i_N} > 1$ (le fait qu'il y a un nombre fini d'intervalles $]a_i, b_i[$ assure l'existence de N). Par conséquent,

$$\sum_{i\in\mathbb{N}} b_i - a_i \ \geqslant\ \sum_{k=1}^{N} b_{i_k} - a_{i_k} \ \geqslant\ \sum_{k=2}^{N}(b_{i_k} - b_{i_{k-1}}) + b_{i_1} - a_{i_1} \ =\ b_{i_N} - a_{i_1} \ \geqslant\ 1.$$

2. Il est clair que $\lambda^\star$ est invariante par translation : si $A \subset \mathbb{R}$ et $x \in \mathbb{R}$,

$$\lambda^\star(A) = \inf_{\substack{(I_n)\in\mathscr{I}^{\mathbb{N}} \\ A\subset\bigcup_n I_n}} \sum_{n\in\mathbb{N}} \ell(I_n) = \inf_{\substack{(J_n)\in\mathscr{I}^{\mathbb{N}} \\ A+x\subset\bigcup_n J_n}} \sum_{n\in\mathbb{N}} \ell(J_n - x) = \inf_{\substack{(J_n)\in\mathscr{I}^{\mathbb{N}} \\ A+x\subset\bigcup_n J_n}} \sum_{n\in\mathbb{N}} \ell(J_n) = \lambda^\star(A+x). \quad \square$$

Par le paragraphe précédent, le Théorème 6.7 montre alors que $\lambda^\star$ est une mesure sur la tribu $\mathscr{M}_{\lambda^\star}$ de ses ensembles mesurables, qui vérifie bien les propriétés (i) et (ii). Il reste à vérifier que la tribu $\mathscr{M}_{\lambda^\star}$ contient les ensembles qui nous intéressent, c'est-à-dire les boréliens de $\mathbb{R}$.

Proposition 6.9 – *La tribu $\mathscr{M}_{\lambda^\star}$ contient $\mathscr{B}(\mathbb{R})$.*

Démonstration. Comme $\mathscr{M}_{\lambda^\star}$ est une tribu, il suffit de montrer qu'elle contient une classe qui engendre les boréliens. Par exemple, il suffit de montrer que pour tout $a \in \mathbb{R}$, $]-\infty, a] \in \mathscr{M}_{\lambda^\star}$. Soit donc $I =]-\infty, a]$, on souhaite montrer que pour tout $B \subset \mathbb{R}$, $\lambda^\star(B) = \lambda^\star(B \cap I) + \lambda^\star(B \setminus I)$. On sait par sous-additivité de $\lambda^\star$ que

$$\lambda^\star(B) \leqslant \lambda^\star(B \cap I) + \lambda^\star(B \setminus I),$$

montrons la deuxième inégalité. Pour ce faire, on considère une suite $(I_n) = (]a_n, b_n[)$ de $\mathscr{I}$ telle que $B \subset \bigcup_n I_n$, et on pose $\varepsilon > 0$. On note que les intervalles $]\min(a_i, a), \min(b_i, a) + \frac{\varepsilon}{2^n}[$ recouvrent $B \cap I$, et que les intervalles $]\max(a_i, a), \max(b_i, a)[$ recouvrent $B \setminus I$. Ainsi,

$$\begin{aligned}
\lambda^\star(B \cap I) + \lambda^\star(B \setminus I) &\leqslant \sum_{n\in\mathbb{N}} (\min(b_n, a) - \min(a_n, a)) + 2\varepsilon + \sum_{n\in\mathbb{N}} (\max(b_n, a) - \max(a_n, a)) \\
&= \sum_{n\in\mathbb{N}} (\min(b_n, a) + \max(b_n, a)) - (\max(a_n, a) + \min(a_n, a)) + 2\varepsilon \\
&= \sum_{n\in\mathbb{N}} (b_n - a_n) + 2\varepsilon
\end{aligned}$$

Finalement, en laissant tendre ε vers 0, on obtient que $\lambda^\star(B \cap I) + \lambda^\star(B \setminus I) \leqslant \sum_{n\in\mathbb{N}} (b_n - a_n)$, et on obtient le résultat en passant à l'inf sur les recouvrements dénombrables de B par des éléments de $\mathscr{I}$. $\square$

Remarque 6.10 – On peut définir de manière similaire la mesure de Lebesgue de dimension $d \geqslant 2$, en définissant d'abord la mesure extérieure $\lambda_d^\star$ par

$$\lambda_d^\star(A) = \inf \Big\{ \sum_{n\in\mathbb{N}} \mathscr{V}(P_n),\ A \subset \bigcup_{n\in\mathbb{N}} P_n,\ (P_n)_n \in \mathscr{P}^{\mathbb{N}} \Big\},$$

pour tout $A \subset \mathbb{R}^d$, où $\mathscr{P}$ désigne l'ensemble des pavés ouverts bornés de $\mathbb{R}^d$. On remarque que la définition reste la même si $\mathscr{P}$ désigne l'ensemble des pavés ouverts de $\mathbb{R}^d$, ou même plus généralement l'ensemble des pavés de $\mathbb{R}^d$.

On peut ensuite montrer comme ci-dessus que $\lambda_d^\star$ définit une mesure λ_d sur $\mathscr{B}(\mathbb{R}^d)$, qui vérifie les propriétés souhaitées :

(i) $\lambda_d([0,1]^d) = 1$,

(ii) λ_d est invariante par translation.

Il convient de remarquer qu'on peut aussi construire les mesures de Lebesgue d-dimensionnelles comme des mesures produits, à partir de la mesure de Lebesgue sur $\mathbb{R}$. Nous détaillerons cette construction au prochain chapitre.

Nous pouvons utiliser la définition de la mesure extérieure de Lebesgue pour proposer une nouvelle preuve de la régularité de la mesure de Lebesgue. Le résultat ci-dessous concerne $\lambda_d^\star$ et est valable pour tout élément de $\mathscr{M}_{\lambda_d^\star}$. Il est donc plus fort que le résultat de régularité du chapitre 3.

Lemme 6.11 – *Si $A \in \mathscr{M}_{\lambda_d^\star}$ et $\varepsilon > 0$, il existe un fermé F et un ouvert Ω de $\mathbb{R}^d$ tels que $F \subset A \subset \Omega$ et $\lambda_d(\Omega \setminus F) < \varepsilon$.*

Démonstration. On remarque d'abord qu'il suffit de montrer que pour tout $\varepsilon > 0$, on peut trouver un ouvert Ω de $\mathbb{R}^d$ tel que $A \subset \Omega$ et $\lambda_d^\star(\Omega \setminus A) < \frac{\varepsilon}{2}$. En effet, on pourra alors appliquer le résultat à A^c pour obtenir un ouvert ω tel que $A^c \subset \omega$ et $\lambda_d^\star(\omega \setminus A^c) < \frac{\varepsilon}{2}$. Ainsi, en introduisant le fermé $F = \omega^c$, on aura bien $F \subset A \subset \Omega$, et $\lambda_d(\Omega \setminus F) \leqslant \lambda_d^\star(\Omega \setminus A) + \lambda_d^\star(A \setminus F) < \varepsilon$.

On suppose d'abord que $\lambda_d^\star(A) < \infty$. La définition de $\lambda_d^\star$ donne l'existence d'une suite $(P_n)_n$ de $\mathscr{P}$ telle que $A \subset \bigcup_n P_n$ et $\sum_n \mathscr{V}(P_n) < \lambda_d^\star(A) + \frac{\varepsilon}{2^{n+2}}$. Si on considère l'ouvert $\Omega = \bigcup_n P_n$, on a alors $A \subset \Omega$, et

$$\lambda_d(\Omega) \;\leqslant\; \sum_n \lambda_d(P_n) \;<\; \lambda_d^\star(A) + \frac{\varepsilon}{2}.$$

Comme $A \in \mathscr{M}_{\lambda_d^\star}$, on a $\lambda_d^\star(\Omega \setminus A) = \lambda_d^\star(\Omega) - \lambda_d^\star(\Omega \cap A) = \lambda_d^\star(\Omega) - \lambda_d^\star(A) < \frac{\varepsilon}{2}$.

Si maintenant $\lambda_d^\star(A) = \infty$, on pose $A_n = A \cap [-n, n]^d$ pour tout $n \in \mathbb{N}$. Il existe alors une suite d'ouverts $(\Omega_n)_n$ telle que $A_n \subset \Omega_n$ et $\lambda_d^\star(\Omega_n \setminus A_n) < \frac{\varepsilon}{2^{n+2}}$ pour tout n. On a donc $A \subset \Omega := \bigcup_n \Omega_n$, et

$$\lambda_d^\star(\Omega \setminus A) \;\leqslant\; \lambda_d^\star\Big(\bigcup_n \Omega_n \setminus A_n\Big) \;\leqslant\; \sum_n \lambda_d^\star(\Omega_n \setminus A_n) \;<\; \frac{\varepsilon}{2},$$

ce qui conclut. □

Remarque 6.12 – On peut en fait montrer que la réciproque du lemme 6.1.2 est vraie : une partie A de $\mathbb{R}^d$ est Lebesgue-mesurable si et seulement si pour tout $\varepsilon > 0$, il existe un ouvert Ω et un fermé F tels que $F \subset A \subset \Omega$ et $\lambda_d(\Omega \setminus F) < \varepsilon$.

> *Démonstration.* Il suffit de montrer la réciproque du lemme 6.1.2. On se ramène donc à prouver que si $B \subset \mathbb{R}^d$, alors $\lambda_d^\star(B) \geqslant \lambda_d^\star(B \cap A) + \lambda_d^\star B \setminus A)$. On se donne $\varepsilon > 0$, et un ouvert Ω de $\mathbb{R}^d$ tel que $A \subset \Omega$ et $\lambda_d^\star(\Omega \setminus A) < \varepsilon$. Ainsi,
>
> $$\lambda_d^\star(B \cap A) + \lambda_d^\star(B \setminus A) \;\leqslant\; \lambda_d^\star(B \cap \Omega) + \lambda_d^\star(B \setminus \Omega) + \lambda_d^\star(\Omega \setminus A) \;\leqslant\; \lambda_d^\star(B) + \varepsilon,$$
>
> du fait que $\Omega \in \mathscr{M}_{\lambda_d^\star}$. En laissant tendre ε vers 0, on obtient le résultat. □

On en déduit du lemme la régularité de la mesure extérieure de Lebesgue de la même manière qu'on a déduit le théorème 3.13 du lemme 3.12.

Théorème 6.13 – *Pour tout $A \in \mathscr{M}_{\lambda_d^\star}$,*

1. $\lambda_d^\star(A) = \inf\{\lambda_d(\Omega),\ \Omega \text{ ouvert},\ A \subset \Omega\}$,
2. $\lambda_d^\star(A) = \sup\{\lambda_d(K),\ K \text{ compact},\ K \subset A\}$.

Remarque 6.14 – On déduit du théorème 6.13 que pour toute partie $A \in \mathscr{M}_{\lambda_d^\star}$, il existe des boréliens $B_1, B_2 \in \mathscr{B}(\mathbb{R}^d)$ tels que $B_1 \subset A \subset B_2$ et $\lambda_d(B_1) = \lambda_d^\star(A) = \lambda_d(B_2)$.

6.2 Unicité des mesures

La question de la caractérisation des mesures se pose naturellement : lorsqu'on fixe une mesure sur une classe de parties $\mathscr{C}$ d'un ensemble E, définit-on une unique mesure sur $\sigma(\mathscr{C})$?

On constate aisément que la réponse est non. Par exemple, si $E = \{0, 1\}$ et $\mathscr{C} = \{\{0\}\}$, on note que $\sigma(\mathscr{C}) = \mathscr{P}(E)$. Si μ est la mesure nulle sur $(E, \mathscr{P}(E))$ et ν la seule mesure sur $(E, \mathscr{P}(E))$ telle que $\nu(\{0\}) = 0$ et $\nu(\{1\}) = 1$, alors μ et ν coïncident sur $\mathscr{C}$, mais ne sont pas égales. En fait, on peut même construire des exemples où les deux mesures ont même masse : si $E = \{0, 1, 2, 3\}$ et $\mathscr{C} = \{\{0, 1\}, \{0, 2\}\}$, alors on a toujours $\sigma(\mathscr{C}) = \mathscr{P}(E)$. On note qu'il suffit de définir la mesure des singletons de E pour définir de manière unique une mesure sur $(E, \mathscr{P}(E))$. Si μ et ν sont les mesures vérifiant

A	$\{0\}$	$\{1\}$	$\{2\}$	$\{3\}$
$\mu(A)$	1	0	0	1
$\nu(A)$	0	1	1	0

alors $\mu(E) = \nu(E) = 2$, μ et ν coïncident sur $\mathscr{C}$, mais μ et ν ne sont pas égales.

Toutefois, le résultat devient vrai si $\mathscr{C}$ est stable par intersection finie, et si l'on fixe la masse de la mesure, comme nous le verrons dans le théorème 6.19 d'unicité des mesures.

La preuve de ce théorème fait appel à un raisonnement par *classe monotone*, c'est donc l'occasion d'introduire ici ce type de raisonnement, qu'il est utile de connaître. Lorsqu'on souhaite démontrer qu'une classe de parties est une tribu, il est parfois plus aisé de montrer que cette classe de partie est une classe monotone. Comme nous le verrons, la structure de classe monotone est comparable à celle de tribu, mais elle n'impose la stabilité par union dénombrable que pour les unions croissantes. Il reste alors à se demander sous quelles hypothèses une classe monotone est en fait une tribu. Ou plutôt, sous quelles hypothèses sur une classe de parties $\mathscr{C}$ la classe monotone engendrée par $\mathscr{C}$ est une tribu : c'est l'objet du lemme des classes monotones.

Dans le cadre du théorème d'unicité des mesures, on souhaite naturellement montrer que, sous les bonnes hypothèses sur $\mathscr{C}$, la classe de parties

$$\mathscr{M} = \{A \in \mathscr{A},\ \mu(A) = \nu(A)\}$$

est une tribu (elle contiendra alors $\sigma(\mathscr{C}) = \mathscr{A}$). On remarque qu'il n'est pas évident de montrer que $\mathscr{M}$ est stable par union dénombrable, mais qu'il est très aisé de montrer que $\mu(\bigcup_n A_n) = \nu(\bigcup_n A_n)$ si la suite $(A_n)_n$ d'éléments de $\mathscr{M}$ est supposée croissante. Cette observation permet de montrer que $\mathscr{M}$ a une structure de classe monotone, et le lemme des classes monotones permet de conclure.

6.2.1 Lemme des classes monotones

Définition 6.15 – On appelle classe monotone sur E une classe de parties $\mathscr{M}$ telle que

— $E \in \mathscr{M}$,

— si $A, B \in \mathscr{M}$ et $A \subset B$, alors $B \setminus A \in \mathscr{M}$ (*stabilité par différence*),

— si $(A_n)_{n\in\mathbb{N}}$ est une suite croissante de $\mathscr{M}$, alors $\bigcup_n A_n \in \mathscr{M}$ (*stabilité par union dénombrable croissante*).

Remarque 6.16 – Toute tribu sur E est une classe monotone sur E.

Lemme 6.17 – *Une classe monotone $\mathscr{M}$ sur E stable par intersection finie (i.e. $A, B \in \mathscr{M} \Rightarrow A \cap B \in \mathscr{M}$) est une tribu sur E.*

Démonstration. On commence par remarquer que $\emptyset = E \setminus E$ donc $\emptyset \in \mathscr{M}$, et si $A \in \mathscr{M}$, alors $A^c = E \setminus A \in \mathscr{M}$. On considère maintenant une suite $(A_n)_{n\in\mathbb{N}}$ de $\mathscr{M}$, et on pose $B_n = \bigcup_{k=0}^n A_k$. Comme $\mathscr{M}$ est stable par intersection finie et par passage au complémentaire, $\mathscr{M}$ est aussi stable par union finie, donc pour tout $n \in \mathbb{N}$, $B_n \in \mathscr{M}$. Comme la suite $(B_n)_n$ est croissante, on a $\bigcup_n A_n = \bigcup_n B_n \in \mathscr{M}$. On a bien montré que $\mathscr{M}$ est une tribu. □

On remarque qu'une intersection quelconque de classes monotones sur E est aussi une classe monotone sur E. Pour toute classe de parties $\mathscr{C}$ de E, il existe alors une plus petite classe monotone sur E contenant $\mathscr{C}$. On l'appelle classe monotone engendrée par $\mathscr{C}$, et on la note $m(\mathscr{C})$:

$$m(\mathscr{C}) = \bigcap \{\mathscr{M},\ \mathscr{M} \text{ classe monotone sur } E,\ \mathscr{C} \subset \mathscr{M}\}.$$

Lemme 6.18 (*Lemme des classes monotones*) – *Soit $\mathscr{C}$ une classe de parties de E stable par intersection finie. On a alors*

$$\sigma(\mathscr{C}) = m(\mathscr{C}).$$

Démonstration. Comme $\sigma(\mathscr{C})$ est une classe monotone contenant $\mathscr{C}$, on a immédiatement $\sigma(\mathscr{C}) \supset m(\mathscr{C})$. Pour obtenir l'autre inclusion, il suffit de montrer que $m(\mathscr{C})$ est une tribu. Le lemme 6.17 assure qu'il suffit de montrer que $m(\mathscr{C})$ est stable par intersection finie. Commençons par montrer que

$$\forall C \in \mathscr{C},\ \forall B \in m(\mathscr{C}), C \cap B \in m(\mathscr{C}). \tag{6.1}$$

Il suffit de montrer que $\mathscr{M}_1 := \{B \in m(\mathscr{C}),\ \forall C \in \mathscr{C}, C \cap B \in m(\mathscr{C})\}$ est une classe monotone. Comme $\mathscr{C} \subset \mathscr{M}_1$ par hypothèse, on aura alors $\mathscr{M}_1 = m(\mathscr{C})$, ce qui donne (6.1).

— Pour tout $C \in \mathscr{C}$, on a $C \cap E = E = C \in m(\mathscr{C})$ donc $E \in \mathscr{M}_1$.

— Si $B_1, B_2 \in \mathscr{M}_1$ avec $B_1 \subset B_2$, alors pour tout $C \in \mathscr{C}$, $C \cap (B_2 \setminus B_1) = (C \cap B_2) \setminus (C \cap B_1) \in m(\mathscr{C})$ donc $B_2 \setminus B_1 \in \mathscr{M}_1$.

— Si $(B_n)_n$ est une suite croissante de $\mathscr{M}_1$, alors pour tout $C \in \mathscr{C}$, $C \cap B_n$ est une suite croissante de $m(\mathscr{C})$ donc $\bigcup_n C \cap B_n = C \cap \bigcup_n B_n \in m(\mathscr{C})$, et $\bigcup_n B_n \in \mathscr{M}_1$.

Pour montrer que $m(\mathscr{C})$ est stable par intersection finie, c'est-à-dire

$$\forall A \in m(\mathscr{C}),\ \forall B \in m(\mathscr{C}), A \cap B \in m(\mathscr{C}),$$

il suffit de montrer que $\mathscr{M}_2 := \{A \in m(\mathscr{C}),\ \forall B \in m(\mathscr{C}), A \cap B \in m(\mathscr{C})\}$ est une classe monotone contenant $\mathscr{C}$. On montre aisément comme ci-dessus que $\mathscr{M}_2$ est une classe monotone. Par ailleurs, (6.1) exprime que $\mathscr{M}_2$ contient $\mathscr{C}$, d'où le résultat. □

6.2.2 Théorème d'unicité des mesures

Théorème 6.19 (*Théorème d'unicité des mesures*) – *Soient μ et ν deux mesures sur un espace mesurable $(E, \mathscr{A})$. Supposons que μ et ν coïncident sur une classe $\mathscr{C} \subset \mathscr{A}$ stable par intersection finie contenant E, et qui engendre $\mathscr{A}$: $\sigma(\mathscr{C}) = \mathscr{A}$. Alors :*

1. *si μ et ν sont finies, on a $\mu = \nu$,*
2. *s'il existe une suite croissante de parties $E_n \in \mathscr{C}$ telles que $\mu(E_n) < \infty$ et $E = \bigcup_n E_n$, on a $\mu = \nu$.*

Remarque 6.20 – On note que si la propriété 2 du Théorème d'unicité est vérifiée, alors les mesures μ et ν sont en particulier σ-finies.

Démonstration. 1. On pose

$$\mathscr{M} := \{A \in \mathscr{A},\ \mu(A) = \nu(A)\}.$$

On remarque que $\mathscr{M}$ est une classe monotone :

— $\mu(E) = \nu(E)$ car $E \in \mathscr{C}$, donc $E \in \mathscr{M}$,

— si $A, B \in \mathscr{M}$ avec $A \subset B$, alors $\mu(B \setminus A) = \mu(B) - \mu(A) = \nu(B) - \nu(A) = \nu(B \setminus A)$, du fait que μ et ν sont finies, donc $B \setminus A \in \mathscr{M}$,

— si $(A_n)_n$ est une suite croissante de $\mathscr{M}$ alors la continuité à gauche de μ et ν implique que $\mu\big(\bigcup_n A_n\big) = \nu\big(\bigcup_n A_n\big)$, donc $\bigcup_n A_n \in \mathscr{M}$.

Comme $\mathscr{M}$ contient $\mathscr{C}$ par hypothèse, on a donc $\mathscr{M} \supset m(\mathscr{C}) = \sigma(\mathscr{C}) = \mathscr{A}$, par le lemme des classes monotones, ce qui prouve que $\mu = \nu$ sur $\mathscr{A}$.

2. Pour tout $n \in \mathbb{N}$, on définit les mesures μ_n et ν_n sur $(E, \mathscr{A})$ par

$$\mu_n(A) := \mu(A \cap E_n), \quad \nu_n(A) := \nu(A \cap E_n).$$

Pour tout $n \in \mathbb{N}$, les mesures μ_n et ν_n sont donc des mesures finies. Par ailleurs, si $C \in \mathscr{C}$, on a alors $C \cap E_n \in \mathscr{C}$, et $\mu_n(C) = \nu_n(C)$. On déduit alors du point 1 que $\mu_n = \nu_n$ sur $\mathscr{A}$. Si $A \in \mathscr{A}$, comme la suite $(A \cap E_n)_n$ est croissante et $\bigcup_n E_n = E$, la continuité à gauche de μ et ν implique que

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

ce qui achève la preuve. □

6.2.3 Unicité de la mesure de Lebesgue

Comme nous l'avons vu au chapitre 3, si μ est une mesure invariante par translation sur $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ et telle que $\mu([0,1]) = 1$, alors pour tout intervalle I de $\mathbb{R}$,

$$\mu(I) = \ell(I).$$

On déduit alors qu'il existe au plus une mesure λ sur $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ invariante par translation et telle que $\lambda([0,1]) = 1$. En effet, si λ et λ' satisfont ces propriétés, alors elles coïncident sur la classe $\mathscr{I}$ des intervalles. Or, la classe $\mathscr{I}$ est stable par intersection finie, elle contient $\mathbb{R}$ tout entier, et elle engendre la tribu $\mathscr{B}(\mathbb{R})$. On peut donc appliquer le point 2 du théorème d'unicité en prenant $E_n =]-n, n[$.

On montre l'unicité de la mesure de Lebesgue d-dimensionnelle λ_d sur $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ de manière analogue, en montrant que λ_d fixe la mesure des ensembles de la classe $\mathscr{P}$ des pavés de $\mathbb{R}^d$. Comme $\mathscr{P}$ est stable par intersection finie et contient $\mathbb{R}^d$, on peut raisonner de la même manière en posant $E_n =]-n, n[^d$.

6.3 Tribu complétée, mesure complétée

6.3.1 Généralités

Soit $(E, \mathscr{A}, \mu)$ un espace mesuré. On rappelle qu'un ensemble $N \subset E$ est dit *μ-négligeable* s'il existe $A \in \mathscr{A}$ tel que $N \subset A$ et $\mu(A) = 0$. On note

$$\mathscr{N}_\mu = \{N \subset E,\ E \text{ est } \mu\text{-négligeable}\}.$$

Remarque 6.21 – On remarque que $\mathscr{N}_\mu$ est stable par union dénombrable. En effet, si $(N_n)_{n\in\mathbb{N}}$ est une suite de $\mathscr{N}_\mu$, on sait que pour tout $n \in \mathbb{N}$, il existe $A_n \in \mathscr{A}$ tel que $\mu(A_n) = 0$ et $N_n \subset A_n$. Ainsi, $\bigcup_n N_n \subset \bigcup_n A_n$, et $\mu(\bigcup_n A_n) \leqslant \sum_n \mu(A_n) = 0$.

Définition 6.22 – Si $\mathscr{N}_\mu \subset \mathscr{A}$, on dit que l'espace mesuré $(E, \mathscr{A}, \mu)$ est complet.

Définition-Proposition 6.23 – On appelle *tribu complétée* de $\mathscr{A}$ la tribu

$$\bar{\mathscr{A}} = \{B \cup N,\ B \in \mathscr{A},\ N \in \mathscr{N}_\mu\}.$$

Démonstration. Montrons que $\bar{\mathscr{A}}$ est une tribu sur E :

— $\emptyset \in \mathscr{A} \cap \mathscr{N}_\mu$ donc $\emptyset \in \bar{\mathscr{A}}$.

— Supposons $B \cup N \in \bar{\mathscr{A}}$. On sait alors qu'il existe $A \in \mathscr{A}$ tel que $N \subset A$ et $\mu(A) = 0$, d'où $(B \cup N)^c = (B^c \setminus A) \cup ((A \setminus N) \cap B^c)$. Comme $A \setminus N \cap B^c \subset A$, on a bien $(B \cup N)^c \in \bar{\mathscr{A}}$.

— Si $B_n \cup N_n \in \bar{\mathscr{A}}$ pour tout n, alors $\bigcup_n (B_n \cup N_n) = \bigcup_n B_n \cup \bigcup_n N_n \in \bar{\mathscr{A}}$ car $\bigcup_n N_n \in \mathscr{N}_\mu$. □

Si $B \cup N \in \bar{\mathscr{A}}$, où $B \in \mathscr{A}$ et $N \in \mathscr{N}_\mu$, on pose

$$\bar{\mu}(B \cup N) = \mu(B).$$

La définition de $\bar{\mu}$ ne dépend pas de l'écriture $B \cup N$, $B \in \mathscr{A}$, $N \in \mathscr{N}_\mu$ choisie : si $B \cup N = B' \cup N' \in \bar{\mathscr{A}}$, alors on a

$$\begin{array}{ccccc} B & \subset & B \cup N & \subset & B \cup A, \\ B' & \subset & B' \cup N' & \subset & B' \cup A', \end{array}$$

avec $A, A' \in \mathscr{A}$ et $\mu(A) = \mu(A') = 0$. On en déduit que $\mu(B) \leqslant \mu(B' \cup A') = \mu(B')$, et de même $\mu(B') \leqslant \mu(B)$, d'où $\bar{\mu}(B \cup N) = \bar{\mu}(B' \cup N')$.

Proposition 6.24 –

1. *$\bar{\mu}$ est une mesure sur $\bar{\mathscr{A}}$, qui prolonge la mesure μ.*
2. *L'espace mesuré $(E, \bar{\mathscr{A}}, \bar{\mu})$ est complet.*

Démonstration. 1. Tout d'abord, il est clair que $\bar{\mu}$ prolonge la mesure μ : si $B \in \mathscr{A}$, $\bar{\mu}(B) = \bar{\mu}(B \cup \emptyset) = \mu(B)$. En particulier, $\bar{\mu}(\emptyset) = 0$. Pour montrer que $\bar{\mu}$ est une mesure, il reste à montrer la σ-additivité : si pour tout $n \in \mathbb{N}$, $B_n \cup N_n \in \bar{\mathscr{A}}$ avec $B_n \in \mathscr{A}$, $N_n \in \mathscr{N}_\mu$, et les $B_n \cup N_n$ sont deux à deux disjoints, alors

$$\bar{\mu}\big(\textstyle\bigsqcup_n (B_n \cup N_n)\big) = \bar{\mu}\big(\bigsqcup_n B_n \cup \bigsqcup_n N_n\big) = \mu\big(\bigsqcup_n B_n\big) = \sum_n \mu(B_n) = \sum_n \bar{\mu}(B_n \cup N_n),$$

car $\bigsqcup_n N_n \in \mathscr{N}_\mu$.

2. Il suffit de montrer que $\mathscr{N}_{\bar{\mu}} \subset \mathscr{N}_\mu$. En effet, on aura alors $\mathscr{N}_{\bar{\mu}} \subset \mathscr{N}_\mu \subset \bar{\mathscr{A}}$, ce qui montre que $(E, \bar{\mathscr{A}}, \bar{\mu})$ est complet. Pour voir ça, on remarque que si $N \in \mathscr{N}_{\bar{\mu}}$, alors $N \subset B \cup N'$ où $B \in \mathscr{A}$, $N' \in \mathscr{N}_\mu$ et $\bar{\mu}(B \cup N') = 0$. Il existe donc $A' \in \mathscr{A}$ tel que $N' \subset A'$ et $\mu(A') = 0$. On a donc
$$N \subset B \cup N' \subset B \cup A'.$$
Comme $0 = \bar{\mu}(B \cup N') = \mu(B)$, on a $\mu(B \cup A') = 0$, ce qui implique que $N \in \mathscr{N}_\mu$. □

6.3.2 Complétion de $\mathscr{B}(\mathbb{R})$

Définition 6.25 – On appelle *tribu de Lebesgue* la tribu $\overline{\mathscr{B}(\mathbb{R})}$ complétée de la tribu $\mathscr{B}(\mathbb{R})$ pour la mesure de Lebesgue. On la note $\mathscr{L}(\mathbb{R})$.

On appelle encore mesure de Lebesgue la mesure $\bar{\lambda}$ complétée de la mesure de Lebesgue λ. On la note généralement aussi λ.

Théorème 6.26 –
1. *La tribu $\mathscr{M}_{\lambda^\star}$ des parties mesurables pour la mesure extérieure $\lambda^\star$ coïncide avec la tribu $\mathscr{L}(\mathbb{R})$.*
2. *La mesure $\bar{\lambda}$ coïncide avec $\lambda^\star$ sur $\mathscr{M}_{\lambda^\star} = \mathscr{L}(\mathbb{R})$.*

Démonstration.
1. Commençons par montrer que si $N \in \mathscr{N}_\lambda$, alors N est λ^*-mesurable. On sait qu'il existe $A \in \mathscr{A}$ tel que $N \subset A$ et $\lambda(A) = 0$, donc si $B \subset \mathbb{R}$, alors $\lambda^\star(B \cap N) \leqslant \lambda^\star(A) = 0$ et $\lambda^\star(B \setminus N) \leqslant \lambda^\star(B)$. On en déduit que
$$\lambda^\star(B \cap N) + \lambda^\star(B \setminus N) \leqslant \lambda^\star(B).$$
L'autre inégalité découle de la σ-sous-additivité de $\lambda^\star$, donc on a bien montré que $\mathscr{N}_\lambda \subset \mathscr{M}_{\lambda^\star}$. Comme $\mathscr{B}(\mathbb{R}) \subset \mathscr{M}_{\lambda^\star}$, on obtient immédiatement que $\mathscr{L}(\mathbb{R}) \subset \mathscr{M}_{\lambda^\star}$.

   Supposons maintenant que $A \in \mathscr{M}_{\lambda^\star}$. Par la régularité de la mesure de Lebesgue, on sait qu'il existe un borélien B de $\mathbb{R}$ tel que $B \subset A$ et $\lambda(B) = \lambda^\star(A)$. Comme $B \in \mathscr{M}_{\lambda^\star}$, on a $\lambda^\star(A) = \lambda^\star(B) + \lambda^\star(A \setminus B)$, ce qui donne $\lambda^\star(A \setminus B) = 0$. Toujours par régularité de la mesure de Lebesgue, il existe un borélien C tel que $A \setminus B$ et $\lambda(C) = \lambda^\star(B \setminus A) = 0$. On en déduit donc que $A \setminus B \in \mathscr{N}_\lambda$, et $A = B \cup (A \setminus B) \in \mathscr{L}(\mathbb{R})$.
2. Soient $B \in \mathscr{B}(\mathbb{R})$ et $N \in \mathscr{N}_\lambda$. On a
$$\lambda^\star(B) \leqslant \lambda^\star(B \cup N) \leqslant \lambda^\star(B) + \lambda^\star(N).$$
On a $\lambda^\star(N) = 0$ car $N \subset A$ avec $A \in \mathscr{B}(\mathbb{R})$ et $\lambda(A) = 0$, donc $\lambda^\star(N) \leqslant \lambda^\star(A) = 0$. On en déduit que
$$\lambda^\star(B \cup N) = \lambda^\star(B) = \lambda(B) = \bar{\lambda}(B \cup N),$$
d'où $\lambda^\star = \bar{\lambda}$ sur $\mathscr{L}(\mathbb{R}) = \mathscr{M}_{\lambda^\star}$. □

Proposition 6.27 – *La tribu de Lebesgue $\mathscr{L}(\mathbb{R})$ a même cardinal que $\mathscr{P}(\mathbb{R})$.*

Démonstration. L'ensemble de Cantor $C \subset [0, 1]$ est compact donc borélien, il vérifie $\lambda(C) = 0$, et il est équipotent à $\mathbb{R}$. Par conséquent, Card $\mathscr{P}(C) =$ Card $\mathscr{P}(\mathbb{R})$. Comme $\mathscr{P}(C) \subset \mathscr{N}_\lambda \subset \mathscr{L}(\mathbb{R})$, on a
$$\text{Card } \mathscr{P}(\mathbb{R}) = \text{Card } \mathscr{P}(C) \leqslant \text{Card } \mathscr{L}(\mathbb{R}),$$
donc comme $\mathscr{L}(\mathbb{R}) \subset \mathscr{P}(\mathbb{R})$, Card $\mathscr{P}(\mathbb{R}) =$ Card $\mathscr{L}(\mathbb{R})$. □

Remarque 6.28 – Ce résultat est à comparer avec celui de la remarque 2.11 : la tribu $\mathscr{B}(\mathbb{R})$ est a même cardinal que celui de $\mathbb{R}$. On voit donc par un argument de cardinal qu'il existe des ensembles Lebesgue-mesurables qui ne sont pas des boréliens. Il en existe même énormément, malgré la difficulté à en expliciter.