Théorie de la mesure, Intégration, Probabilités
Classe sino-française, USTC

Stéphane Nonnenmacher*

Mars-Avril 2021

Résumé

Ces notes de cours présentent la théorie de la mesure et de l'intégrale de Lebesgue. On suppose connus les bases de la théorie des ensembles et de la topologie. Ce cours est destiné aux étudiants de 2e année de la classe sino-française de l'Université des Sciences et Technologies de Chine, à Hefei. On l'étend ensuite à la théorie des probabilités (non discrètes).
Ces notes reprennent le schéma du cours que Jean-François Le Gall a donné à l'ENS en 2006.

*https://www.imo.universite-paris-saclay.fr/∼nonnenma/index.html

Table des matières

Première partie

Théorie de la mesure et Intégration

1 Introduction - Motivation

La théorie de l'intégration a d'abord consisté à calculer des primitives de fonctions explicites, de façon à résoudre l'équation différentielle $u'(x) = f(x)$, avec $f : I \to \mathbb{R}$ une fonction donnée (explicite), et u la fonction inconnue (ici $I \subset \mathbb{R}$ est un intervalle borné, non réduit à un point ; on pourra prendre par exemple $I = [0, 1]$).

1.1 Rappel sur l'intégrale de Riemann

A partir du 19e siècle, les mathématiciens se sont posé la question d'intégrer des fonctions moins explicites, en généralisant de plus en plus les classes de fonctions pour lesquelles l'intégrale avait un sens. Un pas important a été accompli par B.Riemann, qui a défini l'intégrale de fonctions définies sur $\mathbb{R}$ au moyen d'approximations par des fonctions constantes par morceaux. Sa définition permet alors d'intégrer n'importe quelle fonction $f : I \to \mathbb{R}$ *continue*, et en fait n'importe quelle fonction *réglée*. Pour définir les fonctions réglées, on commencera par définir une fonction en escalier.

Définition 1.1. Soit $a < b$ deux points de $\mathbb{R}$, et soit $I = [a, b]$ l'intervalle correspondant. Une fonction $f : I \to \mathbb{R}$ est une *fonction en escalier* (ou *fonction constante par morceaux*) s'il existe une suite finie de po ;ints $a = x_0 < x_1 < x_2 \cdots < x_K = b$, telle que f est constante sur chaque sous-intervalle ouvert $\mathring{I}_k =]x_{k-1}, x_k[$, $k = 1, \dots, K$ (les valeurs prises par f aux points x_k sont arbitraires).

Pour étudier les suites de fonctions, on utilisera tout d'abord la norme de la convergence uniforme. Pour toute fonction $f : I \to \mathbb{R}$, on définit la norme sup :

$$\|f\|_\infty = \sup_{x \in I} |f(x)| .$$

Pour une suite $(f_n)_{n \geq 1}$ de fonctions définies sur I, on dit que la suite (f_n) converge *uniformément* vers f si et seulement si

$$\|f - f_n\|_\infty \xrightarrow[n \to \infty]{} 0.$$

Définition 1.2. Une fonction $f : I \to \mathbb{R}$ est une fonction réglée s'il existe une suite (f_n) de fonctions en escalier sur I, qui converge uniformément vers f lorsque $n \to \infty$.

Notons qu'en général, la partition de I en sous-intervalles *dépend* de la fonctions f_n : on notera alors $I = \bigcup_{k=1}^{K(n)} I_k^{(n)}$.

Proposition 1.3. *Toute fonctions f continue est réglée.*

Démonstration. : Supposons pour simplifier que $I = [0, 1]$. On peut par exemple, pour chaque $n > 0$, découper I en n intervalles $I = \bigcup_{k=1}^{n} \left[\frac{k-1}{n}, \frac{k}{n}\right]$ et construire la fonction en escalier

$$f_n(x) = f\left(\frac{k}{n}\right) \quad \text{si } x \in \left[\frac{k-1}{n}, \frac{k}{n}\right[, k = 1, \ldots, n \quad \text{et} \quad f_n(1) = f(1).$$

La fonction f étant continue sur l'intervalle compact I, elle y est *uniformément continue* (Théorème de Heine) :

$$\forall \epsilon > 0, \exists n = n(\epsilon) \in \mathbb{N}^*, \forall x, y \in I,\ |x - y| \leq \frac{1}{n} \Longrightarrow |f(x) - f(y)| \leq \epsilon.$$

Cette propriété implique alors que

$$\forall m \geq n(\epsilon),\ \forall x \in I, \quad |f_m(x) - f(x)| \leq \epsilon,$$

donc que les fonctions (f_n) convergent uniformément vers f lorsque $n \to \infty$. □

Affirmation 1.4. On peut montrer qu'une fonction $f : I \to \mathbb{R}$ est réglée si et seulement si elle admet, en tout point $x \in I$, une limite à droite et une limite à gauche (ces deux limites peuvent être différentes, et peuvent être toutes deux différentes de $f(x)$).

La définition de l'intégrale d'une fonction en escalier est simple. Soit f une fonction en escalier, telle que

$$f(x) = a_k \quad \text{lorsque } x \in]x_{k-1}, x_k[, \quad \text{pour } k = 1, \ldots, K.$$

Alors l'intégrale de f sur I est donnée par

$$\int_I f(x)\, dx = \sum_{k=1}^{K} a_k\, (x_k - x_{k-1})\,.$$

Proposition 1.5. *Soit une fonction réglée $f : I \to \mathbb{R}$, et soit (f_n) une suite de fonctions en escalier convergeant uniformément vers f. On montre alors que les intégrales $S_n = \int_I f_n(x)\, dx$*

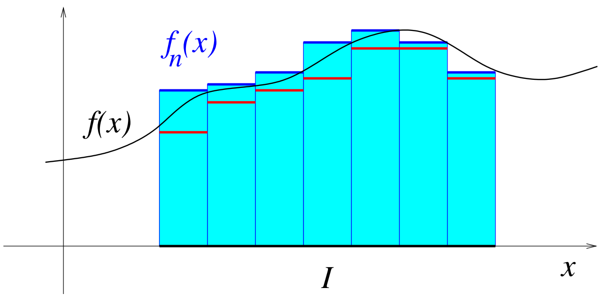

FIGURE 1.1 – Approximation d'une fonction continue f par une fonction en escalier f_n, et représentation de lintégrale $\int_I f_n(x)\,dx$ (en bleu). Une seconde fonction en escalier (en rouge) minore f.

ont une limite lorsque $n \to \infty$. Cette limite définit l'inégrale de la fonction f sur I, notée $\int_I f(x)\,dx = \int_a^b f(x)\,dx$.

Démonstration. Le fait que f_n converge vers f uniformément implique que, pour tout $\epsilon > 0$, il existe $n(\epsilon)$ tel que, si $n, m \geq n(\epsilon)$, alors $\|f_n - f_m\|_\infty \leq \epsilon$. La fonction $f_n - f_m$ est alors une fonction en escalier, associée à la partition formée par l'intersection des partitions $(I_k^{(n)})$ et $(I_j^{(m)})$. Cette fonction satisfaisant $\sup_x |f_n(x) - f_m(x)| \leq \epsilon$, son intégrale vérifie alors l'inégalité

$$|S_n - S_m| = \left| \int_I (f_n(x) - f_m(x))\,dx \right| \leq \epsilon\,(b-a) = \epsilon |I|,$$

où on a noté $|I| = b - a$ la longueur de l'intervall I. Ceci montre que (S_n) est une suite de Cauchy, donc qu'elle converge.

Il reste à vérifier que la limite $\lim_{n\to\infty} S_n$ ne dépend pas du choix de la suite (f_n) convergeant vers f. En effet, si on considère deux suites (f_n) et $(\tilde{f}_n)$ de fonctions en escalier convergeant uniformément vers f, alors la suite $f_n - \tilde{f}_n$ converge vers zéro uniformément, ce qui implique que $\int f_n(x)\,dx - \int \tilde{f}_n(x)\,dx$ tend vers zéro. □

Graphiquement, l'intégrale de Riemann d'une fonction réglée f représente l'aire algébrique entre le graphe de f au-dessus de I et l'intervalle I, voir la figure .

1.2 Problèmes de l'intégrale de Riemann

Dans la seconde moitié du 19e siècle, les mathématiciens ont introduit des familles de fonctions de plus en plus compliquées, et qu'on ne pouvait pas obtenir comme limites uniformes de fonctions en escalier, mais seulement comme limites simples de fonction en escalier (ou limites simples de fonctions continues). L'intégrale de Riemann ne permet alors pas de donner un sens à l'intégrale d'une telle fonction. Pouvait-on néanmoins définir, d'une autre manière, l'intégrale d'une telle fonction ?

Exemple 1.6. Sur l'intervalle $I = [0, 1]$, on considère la fonction caractéristique des nombres rationnels, qu'on notera $\mathbb{1}_{\mathbb{Q}\cap I}$. Cette fonction est discontinue en tout point ; elle n'admet de limite à droite ou à gauche en aucun point $x \in I$, donc ce n'est pas une fonction réglée. Cette fonction peut être obtenue comme la double limite :

$$\forall x \in I, \qquad \mathbb{1}_{\mathbb{Q}\cap I}(x) = \lim_{n\to\infty} \lim_{m\to\infty} \cos(\pi n! x)^{2m} \overset{\text{def}}{=} \lim_{n\to\infty} F_n(x),$$

mais aucune des deux limites n'est uniforme. En effet, $F_n(x) = 1$ sur les points $x \in I$ multiples de $\frac{1}{n!}$, et $F_n(x) = 0$ partout ailleurs. Les fonctions F_n sont des fonctions en escalier d'intégrale nulle. Elles convergent simplement vers $\mathbb{1}_{\mathbb{Q}\cap I}$, mais pas uniformément. Peut-on néanmoins donner un sens à l'intégrale de $\mathbb{1}_{\mathbb{Q}\cap I}$?

Intégrale de Riemann - bis

Pour toute fonction bornée $f : I \to \mathbb{R}$, on peut chercher à assouplir la définition de l'intégrale de Riemann donnée ci-dessus, en considérant uniquement les fonctions en escalier f_n qui minorent f (respectivement qui majorent f) : on peut ainsi définir

$$S_-(f) \overset{\text{def}}{=} \sup_{f_n \le f} \int_I f_n(x)\, dx,$$

où on prend le supremum sur toutes les fonctions en escalier f_n minorant f. De même, on définit

$$S_+(f) \overset{\text{def}}{=} \inf_{g_n \ge f} \int_I g_n(x)\, dx.$$

On montre facilement que $S_-(f) \le S_+(f)$. Si la fonction f est telle que $S_-(f) = S_+(f)$, cette valeur définit l'intégrale de Riemann de f ; une telle fonction f est dite Riemann-intégrable. Les fonctions Riemann-intégrables forment une classe plus large que les fonctions réglées.

Exemple 1.7. 1.On voit facilement que notre fonction $\mathbb{1}_{\mathbb{Q}\cap I}$ n'est pas Riemann-intégrable. En effet, le fait que $\mathbb{R}\setminus\mathbb{Q}$ soit dense dans $\mathbb{R}$ montre que pour toute fonction en escalier $f_n \le \mathbb{1}_{\mathbb{Q}\cap I}$, $f_n(x)$ doit prendre des valeurs négatives ou nulles sur tout intervalle de sa partition associée, de sorte que $\int_I f_n(x)\, dx \le 0$. Tandis que pour toute fonction en escalier $g_n \ge \mathbb{1}_{\mathbb{Q}\cap I}$, du fait de la densité de $\mathbb{Q}$ dans $\mathbb{R}$, $g_n(x) \ge 1$ sur tout intervalle de sa partition, de sorte que $\int_I g_n(x)\, dx \ge 1$.

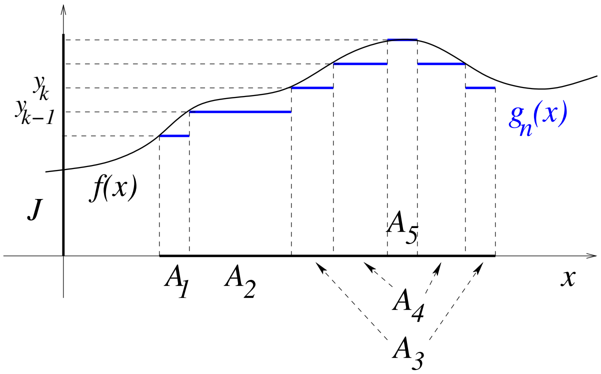

FIGURE 1.2 – Stratégie de Lebesgue : approximation de f par une fonction étagée g_n.

2. Par contre, la fonction $x \mapsto \sin(1/x)$ définie sur l'intervalle (0,1) n'est pas réglée (elle n'admet pas de limite en $x \searrow 0$), mais on peut montrer qu'elle est Riemann-intégrable.

1.3 Approche de Lebesgue de l'intégration : « il faut découper dans le sens horizontal »

L'idée formulée par H.Lebesgue au tout début de 20e siècle, est de calculer l'aire (algébrique) entre le graphe de f et l'intervalle I non pas en découpant la surface par des rectangles fins *verticaux* comme le fait Riemann (v. la Figure 1.1), mais par des « bandes horizontales ». Plutôt que de découper l'intervalle « source » I en sous-intervalles $I_k^{(n)}$ pour fabriquer une fonction en escalier f_n approximant f, on découpe l'intervalle « image » $J \subset \mathbb{R}$ en sous-intervalles $J_k^{(n)} = [y_{k-1}^{(n)}, y_k^{(n)}] \cap J$, par exemple en prenant $y_k^{(n)} = \frac{k}{n}$, et on fabrique une fonction approximante g_n, telle que

$$g_n(x) = y_{k-1}^{(n)} \quad \text{pour } x \in A_k^{(n)} \stackrel{\text{def}}{=} f^{-1}\left(\left[y_{k-1}^{(n)}, y_k^{(n)}\right[\right) \cap I. \tag{1.1}$$

La fonction g_n est appelée une *fonction étagée*. Elle ressemble à une fonction en escalier, mais la différence est que les ensembles $A_k \subset I$ peuvent être beaucoup plus compliqués que des sous-intervalles. Sur la figure, on voit déjà que certains A_k sont des sous-intervalles, d'autres des unions (finies) de sous-intervalles. L'intégrale de la fonction étagée g_n est définie par

$$\int_I g_n(x)\, dx = \sum_k y_{k-1}^{(n)} \left|A_k^{(n)}\right| \tag{1.2}$$

avec $\left|A_k^{(n)}\right|$ indiquant la longueur du sous-ensemble (« l'étage ») $A_k^{(n)}$.

Pour une fonction $f : I \to \mathbb{R}$ continue, on peut montrer que les intégrales $\int g_n(x)\,dx$ convergent vers l'intégrale de Riemann $\int f(x)\,dx$, c'est donc une seconde manière d'obtenir la même valeur.

> Que gagne-t-on avec cette second manière d'approximer f par un « découpage horizontal » ?

On voit que le calcul de l'intégrale par cette second manière impose de *mesurer* les sous-ensembles $A_k^{(n)}$, définis comme les images réciproques par f des intervalles $J_k^{(n)}$. Que signifie « mesurer » ? Lorsque $A_k^{(n)}$ est une union finie d'intervalles (comme sur la figure), sa longueur est la somme des longueurs des sous-intervalles. Mais que se passe-t-il si $A_k^{(n)}$ est un ensemble plus compliqué ? Peut-on définir de façon raisonnable sa « longueur » ?

La théorie de la Lebesgue va répondre de façon précise à ces questions. Elle aura pour objectif de

1. définir une famille de sous-ensembles $A \subset I$, auxquels on peut assigner une « longueur » ; on appellera cette longueur la *mesure de Lebesgue* de A, et un tel sous-ensemble A sera appelé « ensemble mesurable ».
2. On définira ensuite la une classe de fonctions (dites « fonctions mesurables »), pour lesquelles chaque image réciproque d'un intervalle $f^{-1}([a,b])$ est un ensemble mesurable.
3. Une fonction mesurable pourra être approchée par une famille de fonctions étagées, du type $g_n = \sum_k c_k^{(n)} \mathbb{1}_{A_k^{(n)}}$ avec $A_k^{(n)}$ des ensembles mesurables (la construction de g_n peut être différente de celle donnée en (1.1)). On peut définir l'intégrale de g_n comme en (1.2), et il faudra montrer sous quelles conditions les intégrales $\int g_n(x)dx$ ont une limite.

Ce qu'on gagne par rapport à la théorie de Riemann, c'est que la convergence des g_n vers f n'a plus besoin d'être uniforme, mais elle peut être *simple* : il suffit que

$$\forall x \in I, \quad g_n(x) \xrightarrow[n\to\infty]{} f(x).$$

Pour une telle convergence simple des g_n, la convergence des $\int g_n(x)dx$ sera assurée par des théorèmes importants de la théorie de Lebesgue : le *Théorème de convergence monotone*, et le *Théorème de convergence dominée*, ainsi que le *Lemme de Fatou*.

Un autre avantage de la théorie de Lebesgue, c'est qu'on va élargir la notion de « longueur » d'un sous-ensemble $A \subset I$, en la notion plus générale de « mesure » de cet ensemble. La mesure d'un ensemble A sera un « poids » (un nombre réel positif , parfois $+\infty$) associé à cet ensemble, on notera ce poids $\mu(A)$. On définira alors l'intégrale d'une fonction étagée, relativement à cette mesure μ, par la même formule qu'en (1.2), mais en remplaçant les longueurs $\left|A_k^{(n)}\right|$ par les poids $\mu\left(A_k^{(n)}\right)$. On aura alors les mêmes théorèmes de convergence vers une intégrale « de mesure μ », qu'on notera $\int_I f(x)\, d\mu(x)$.

Cette généralisation de la notion de « longueur » à celle de « mesure » retire à l'intégrale son caractère géométrique : l'intégrale $\int_I f(x)\, d\mu(x)$ ne sera plus reliée à l'aire (algébrique) comprise entre le graphe de f et I. Mais ce type d'intégrale aura une interprétation importante en *théorie des probabilités*, que nous étudierons dans la seconde partie de ce cours. Le poids $\mu(A)$ sera alors interprété comme la probabilité de l'ensemble A, avec comme condition que l'ensemble total (ici, l'intervalle I) est de probabilité pleine $\mu(I) = 1$; une telle mesure est appelée mesure de probabilité. Une fonction f sera alors interprétée comme une variable aléatoire, et l'intégrale $\int_I f(x)\, d\mu(x)$ comme la moyenne (ou l'espérance) de cette variable aléatoire, par rapport à la mesure de probabilité μ.

La théorie de Lebesgue permet aussi de généraliser l'espace portant la mesure, et les fonctions f : on a pris jusqu'à présent $I \subset \mathbb{R}$ un intervalle réel. On pourra considérer à la place de I un ouvert de $\mathbb{R}^d$, un espace métrique, voire un ensemble E quelconque, auquel cas la mesure d'un sous-ensemble A n'a plus rien à voir avec une « longueur » ou un « volume ».

2 Mesures positives, fonctions mesurables

Dans ce chapitre nous introduisons les éléments de base de la théorie de l'intégration de Lebesgue, sur un ensemble de départ quelconque E, et pour une mesure μ quelconque. On définira la notion de sous-ensemble mesurable, avant celle de la mesure. Les sous-ensembles mesurables forment une famille satisfaisant des propriétés axiomatiques spécifiques, qu'on appelle une *tribu*, ou *σ-algèbre*.

Cette présentation « axiomatique » de la théorie de la mesure trouvera son intérêt dans un cas particulier, celui de la mesure de Lebesgue sur l'espace euclidien $E = \mathbb{R}^d$, qui généralise la notion de volume euclidien. On construira en détail la tribu des sous-ensembles de $\mathbb{R}^d$ sur laquelle cette mesure est définie, qu'on appellera la tribu de Lebesgue.

A une tribu de E, on pourra associer une famille de fonctions à valeurs rélles sur E, les fonctions mesurables. Une fois choisie une mesure μ définie sur cette tribu, on pourra alors définir l'intégrale d'une fonction mesurable. On donnera les théorèmes fondamentaux de cette théorie de l'intégration : les théorèmes de convergence monotone et de convergence dominée.

2.1 Ensembles mesurables, tribu

On se donne un ensemble E quelconque, sur lequel seront définies nos fonctions. Une tribu est une famille de sous-ensembles de E, satisfaisant certaines propriétés axiomatiques.

Définition 2.1. Soit E un ensemble quelconque. Une tribu (aussi appelée σ-algèbre) sur E est une famille $\mathcal{A} \subset \mathscr{P}(E)$ de parties de E, telle que :

(i) l'ensemble E fait partie de la tribu $\mathcal{A}$

(ii) si $A \in \mathcal{A}$, alors son complémentaire $A^c \in \mathcal{A}$. (une tribu est stable par prise du complémentaire)

(iii) Si $(A_n)_{n\in\mathbb{N}}$ est une famille d'éléments de $\mathcal{A}$, alors leur union $\bigcup_{n\in\mathbb{N}} A_n$ appartient aussi à $\mathcal{A}$. On dit qu'une tribu est stable par *union dénombrable*.

Une fois qu'on s'est donné une tribu $\mathcal{A}$ sur E, les éléments de $\mathcal{A}$ sont appelés les *ensembles mesurables* (ou $\mathcal{A}$-mesurables). Le couple $(E, \mathcal{A})$ est appelé un *espace mesurable*. On déduit facilement les propriétés suivantes de la tribu $\mathcal{A}$:

1. Par (i) et (ii), on déduit que l'ensemble vide $\emptyset$ est dans $\mathcal{A}$.

2. Si $(A_n)_{n\in\mathbb{N}}$ est une famille d'éléments de $\mathcal{A}$, alors leur intersection $\bigcap_{n\in\mathbb{N}} A_n = \left(\bigcup_n A_n^c\right)^c$ est aussi dans $\mathcal{A}$.

3. Les A_n ne sont pas supposés distincts les uns des autres. Si on prend par exemple $A_n = \emptyset$ pour $n \geq N$, $\bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n=0}^{N} A_n$ est effectivement une union finie. Donc $\mathcal{A}$ est aussi stable par unions ou intersections finies.

Si on avait supposé les propriétés (i), (ii) et la stabilité uniquement par union finies, on aurait appelé $\mathcal{A}$ une algèbre. Le fait d'autoriser des unions *dénombrables* est indiqué par le préfixe σ.

Exemple. Les exemples les plus simples de tribus sur un ensemble E sont :

1. la tribu complète $\mathcal{A} = \mathscr{P}(E)$.

2. la tribu minimale, ou triviale $\mathcal{A} = \{\emptyset, E\}$

Ces deux tribus ne sont pas très intéressantes, dans le sens qu'elles ne pourront pas porter les mesures qui vont nous intéresser.

Proposition 2.2. *Si $\mathcal{A}$ et $\mathcal{B}$ sont deux tribus sur E, alors leur intersection $\mathcal{A} \cap \mathcal{B} = \{A \subset E\,;\, A \in \mathcal{A} \text{ et } A \in \mathcal{B}\}$ est également une tribu.*

Démonstration. Facile, à partir de la définition d'une tribu. □

Cette propriété d'intersection permet de définir une tribu à partir d'une famille quelconque de parties de E :

Définition 2.3. Soit $\mathcal{C}$ un sous-ensemble de $\mathscr{P}(E)$ (donc une famille de parties de E). Il existe alors une plus petite tribu sur E contenant $\mathcal{C}$. Cette tribu est unique, définie par

$$\sigma(\mathcal{C}) \stackrel{\text{def}}{=} \bigcap_{\mathcal{A}\supset\mathcal{C}} \mathcal{A}.$$

Dans le membre de droite on prend l'intersection sur toutes les tribus $\mathcal{A}$ sur E qui contiennent $\mathcal{C}$. Il en existe au moins une, la tribu complète $\mathscr{P}(E)$. La tribu $\sigma(\mathcal{C})$ est appelée la *tribu engendrée par $\mathcal{C}$*.

Cette notion de tribu engendrée va constituer notre outil majeur pour fabriquer des tribus intéressantes. Il faut pour cela qu'on suppose que l'ensemble E possède des structures supplémentaires. Les cas qui vont nous occuper le plus souvent :

1. E est un espace topologique. Cela signifie qu'on a défini une topologie sur E, qui consiste à définir la famille $\mathcal{O} \subset \mathscr{P}(E)$ l'ensemble des ouverts de E. La tribu engendrée par $\mathcal{O}$, $\sigma(\mathcal{O})$, est appelée **la tribu des boréliens, ou tribu borélienne de** E. On la notera généralement $\mathcal{B}(E)$. Les éléments de cette tribu sont appelée les (sous-ensembles) boréliens de E. Cette tribu est un cas particulier très important dans ce qui suit.

2. Un cas particulier du cas précédent est celui où E est un espace métrique, donc muni d'une distance $d(\bullet, \bullet)$. Dans ce cas, cette distance permet de définir une topologie sur E, à partir des boules ouvertes $B(x, r) = \{y \in E\,;\, d(x, y) < r\}$.

3. Un cas particulier du cas précédent est celui de l'espace euclidien $E = \mathbb{R}^n$, muni de la distance euclidienne $d(x, y) = \|x - y\|$.

Remarque 2.4. On rappelle que pour tout espace topologique E, l'ensemble des ouverts $\mathcal{O}$ contient toujours E et $\emptyset$, qu'il est stable par intersection finie et par union quelconque (pas forcément dénombrable). Notez les différences subtiles entre ces propriétés, et les propriétés ci-dessus vérifiée par une tribu.

> Dans la suite du cours, lorsqu'on parlera de tribu sur un espace topologique E sans préciser (par exemple pour $E = \mathbb{R}^n$), il s'agira implicitement de la tribu des boréliens $\mathcal{B}(E)$.

Exercice 2.5. Montrer que la tribu des boréliens sur l'espace euclidien $E = \mathbb{R}$ est aussi engendrée par l'ensemble des intervalles $\{]-\infty, a[\,,\ a \in \mathbb{Q}\}$.

Indication : on pourra se servir du fait que $\mathbb{Q}$ est dense dans $\mathbb{R}$, et que tout ouvert de $\mathbb{R}$ est donné par une union finie ou dénombrable d'intervalles ouverts disjoints.

Sur la complexité de la tribu des boréliens

Soit E un espace topologique. La tribu des boréliens contient des ensembles assez « compliqués ». Par définition, la tribu $\mathcal{B}(E)$ contient tous les ouverts. Elle contient aussi les intersections dénombrables d'ouverts (qui ne sont en général pas des ouverts) ; on appelle G_δ l'ensemble de ces intersections dénombrables (ici, le symbole δ signifie « intersection dénombrable ». Par complémentarité, $\mathcal{B}(E)$ contient aussi les fermés, et les unions dénombrables de fermés (qui en général ne sont pas des fermés) ; on appelle F_σ ce type de sous-ensembles. On peut aussi considérer les unions dénombrables de compacts de E, notés K_σ. A partir de ces nouveaux types d'ensembles, on peut recommencer à prendre des intersections ou union dénombrables : on obtient par exemple des ensembles de type $G_{\delta\sigma}$, $G_{\delta\sigma\delta}$, etc.

Lorsque $E = \mathbb{R}$, on montre qu'à chaque nouvelle opération dénombrable, on obtient effectivement de nouveaux sous-ensembles. Ceci montre la complexité des boréliens, même sur $E = \mathbb{R}$.

Produit de deux espaces mesurables

Il est fréquent de considérer un espace produit cartésien de deux espaces E_1, E_2. Si chacun de ces espaces est équipé d'une tribu, y a-t-il une tribu naturelle sur $E_1 \times E_2$? Oui, on l'appelle la tribu-produit.

Définition 2.6. Soit $(E_1, \mathcal{A}_1)$ et $(E_2, \mathcal{A}_2)$ deux espaces mesurables. La tribu-produit sur $E_1 \times E_2$ est la tribu définie par

$$\mathcal{A}_1 \otimes \mathcal{A}_2 \stackrel{\text{def}}{=} \sigma\left(\{A_1 \times A_2\,;\, A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2\}\right).$$

Remarquons que cette tribu-produit n'est pas seulement constituée des produits cartésiens $A_1 \times A_2$, mais c'est la plus petite tribu contenant tous ces produits.

Exemple 2.7. Pour $E_1 = E_2 = \mathbb{R}$ munis de la tribu des boréliens, la tribu-produit $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ est égale à la tribu borélienne $\mathcal{B}(\mathbb{R}^2)$.

Pour deux espaces topologiques E_1, E_2 quelconques munis de leurs tribus boréliennes, la tribu-produit $\mathcal{B}(E_1) \otimes \mathcal{B}(E_2)$ est toujours contenue dans $\mathcal{B}(E_1 \times E_2)$, mais la réciproque n'est pas toujours vraie.

2.2 Mesure positive définie sur une tribu

On a pour l'instant défini une tribu $\mathcal{A} \subset \mathscr{P}(E)$, qui contient par définition les ensembles *mesurables* de E. « Mesurable » signifie : « qui peut être mesuré ». Mesurer un sous-ensemble, c'est lui allouer un nombre positif, un « poids ». Ce poids doit satisfaire une propriété naturelle d'additivité vis-à-vis de l'union de deux sous-ensembles disjoints. La particularité de la définition ci-dessous, c'est que cette propriété d'additivité s'étend aux unions disjointes dénombrables.

Définition 2.8. Soit $(E, \mathcal{A})$ un ensemble mesurable. Une mesure positive sur cet ensemble est une application $\mu : \mathcal{A} \to [0, \infty]$ vérifiant les propriétés suivantes :

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

(ii) Pour toute famille $(A_n)_{n\in\mathbb{N}}$ de parties mesurables *disjointes deux à deux*,

$$\mu\left(\bigsqcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n) = \lim_{N\to\infty} \sum_{n=0}^{N} \mu(A_n). \qquad (\sigma\text{-additivité})$$

L'espace mesurable $(E, \mathcal{A})$ équipé d'une mesure μ est appelé un *espace mesuré* $(E, \mathcal{A}, \mu)$.

Remarquons que :

0. la mesure d'un ensemble peut valoir $+\infty$.

1. l'union du membre de gauche est bien dans $\mathcal{A}$, d'après la définition d'une tribu. On aurait pu écrire l'union $\bigcup_{n\in\mathbb{N}} A_n$, mais $\bigsqcup_{n\in\mathbb{N}}$ permet d'insister sur le caractère disjoint des A_n.

2. la limite dans le membre de droite est toujours bien définie, puisque tous ses termes sont ≥ 0, de sorte que la somme est croissante par rapport à N ; cette limite peut être finie, ou valoir $+\infty$.

3. La propriété (ii) est appelée **σ-additivité** (comme plus haut, le préfixe σ indique le caractère dénombrable de l'union et de la somme). En prenant $A_n = \emptyset$ tous vides pour $n \geq n_0$, on vérifie la propriété d'additivité vis-à-vis d'une union finie de sous-ensembles disjoints $A_0, \ldots, A_{n_0}$.

4. L'additivité ne sera pas vraie pour une union *indénombrable* d'ensembles disjoints. De toute façon, il est difficile de donner à un sens à une somme indénombrable de termes positifs ou nuls. Mais remarquons que la mesure de Lebesgue d'un intervalle $]a, b[$ avec $a < b$ n'est pas égale à la somme des mesures $\mu(\{x\}) = 0$ des points qui le constituent.

Remarque 2.9. Il est important de se souvenir que la mesure μ n'est en général pas définie sur tous les sous-ensembles de E, mais seulement sur ceux appartenant à la tribut $\mathcal{A}$. Nous montrerons plus tard que c'est le cas en particulier pour la mesure de Lebesgue sur $\mathcal{B}(\mathbb{R})$: il existe des sous-ensembles de $\mathbb{R}$ dont on ne peut pas mesurer la « longueur ».

Propriétés de la mesure :

Tous les ensembles ci-dessous sont supposés être $\mathcal{A}$-mesurables.

1. Si $A \subset B$, alors $\mu(A) \leq \mu(B)$. Si de plus $\mu(A) < \infty$, alors $\mu(B \setminus A) = \mu(B) - \mu(A)$. (le cas $\mu(A) = \infty$ implique $\mu(B) = \infty$, et on ne peut pas donner un sens à $\infty - \infty$)

2. Pour tous $A, B \in \mathcal{A}$, alors $\mu(A) + \mu(B) = \mu(A \cup B) + \mu(A \cap B)$.

3. Si les $(A_n)_{n\in\mathbb{N}}$ forment une famille croissante ($A_n \subset A_{n+1}$), alors

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \uparrow \mu(A_n),$$

où la notation $\lim \uparrow$ signifie une limite croissante.

4. Si les $(B_n)_{n\in\mathbb{N}}$ forment une famille décroissante ($B_n \supset B_{n+1}$) et si $\mu(B_0) < \infty$, alors

$$\mu\left(\bigcap_{n\in\mathbb{N}} B_n\right) = \lim_{n\to\infty} \downarrow \mu\left(B_n\right). \tag{2.1}$$

Le résultat reste vrai si $\mu(B_N) = 0$ pour un certain $N \geq 0$.

5. Pour une famille $(A_n)_{n\in\mathbb{N}}$ quelconque d'ensembles mesurables, $\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leq \sum_{n\in\mathbb{N}} \mu(A_n)$.

Démonstration. 1 : on a $B = A \sqcup (B \setminus A)$, donc $\mu(B) = \mu(A) + \mu(B \setminus A)$.

2 : On rappelle les décompositions disjointes $A = (A \setminus B) \sqcup (A \cap B)$, $B = (B \setminus A) \sqcup (A \cap B)$, $A \cup B = (A \setminus B) \sqcup (A \cup B) \sqcup (B \setminus A)$. En passant aux mesures de ces ensembles, on trouve l'égalité.

3. A partir de la famille croissante $(A_n)_{n\in\mathbb{N}}$, on fabrique des ensembles disjoints $C_0 = A_0$, $C_n = A_n \setminus A_{n-1} \in \mathcal{A}$ pour $n \geq 1$, qui vérifient $A_n = \bigsqcup_{j=0}^n C_j$, donc $\mu(A_n) = \sum_{j=0}^n \mu(C_j)$, et $\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{j=0}^\infty \mu(C_j) = \lim \uparrow \mu(A_n)$.

4. En prenant $A_n = B_0 \setminus B_n$, on on obtient une suite $(A_n)_{n\in\mathbb{N}}$ croissante. En se servant des propriétes 1. et 3., et du fait que $\mu(B_0) < \infty$, on déduit

$$\begin{aligned}\mu(B_0) - \mu\left(\bigcap_{n\in\mathbb{N}} B_n\right) &= \mu\left(B_0 \setminus \bigcap_{n\in\mathbb{N}} B_n\right) = \mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim \uparrow \mu(A_n) \\ &= \lim \uparrow (\mu(B_0) - \mu(B_n)) = \mu(B_0) - \lim \downarrow \mu(B_n).\end{aligned}$$

5. On se réduit à une famille d'ensembles disjoints en prenant $C_0 = A_0$, $C_1 = A_1 \setminus A_0$, et $C_n = A_n \setminus \left(\bigcup_{j=0}^{n-1} A_j\right)$ pour tout $n \geq 1$. Ces ensembles sont mesurables, disjoints et vérifient $\bigcup_{j=0}^n A_j = \bigsqcup_{j=0}^j C_j$, donc en prenant la limite $n \to \infty$, $\bigcup_{j=0}^\infty A_j = \bigsqcup_{j=0}^\infty C_j$, ce qui amène $\mu\left(\bigcup_{j=0}^\infty A_j\right) = \mu\left(\bigsqcup_{j=0}^\infty C_j\right) = \sum_{j=0}^\infty \mu(C_j) \leq \sum_{j=0}^\infty \mu(A_j)$. □

Après avoir énoncé les propriétés générales d'une mesure, il est temps de présenter des exemples concrets.

Exemples de mesures

1. Pour $E = \mathbb{N}$ et la tribu complète $\mathcal{A} = \mathscr{P}(\mathbb{N})$, on peut considérer la mesure de comptage $\mu(A) = \#A$. On remarque que les ensembles $B_n = \{n, n+1, \ldots\}$ forment une suite décroissante ($B_{n+1} \subset B_n$), avec $\mu(B_n) = \infty$ pour tout $n \geq 0$, mais $\bigcap_n B_n = \emptyset$. La

formule (2.1) est donc fausse dans ce cas, ce qui est dû au fait que tous les ensembles B_n sont de mesure infinie.

La mesure de comptage peut être définie pour n'importe quel ensemble mesurable de la forme $(E, \mathscr{P}(E))$, mais pour des ensembles indénombrables elle n'est pas très intéressante, puisqu'elle prend la valeur $+\infty$ sur la plupart des sous-ensembles.

2. Pour un ensemble mesurable quelconque $(E, \mathcal{A})$ et x un élément de E, on peut définir une mesure δ_x par :
$$\forall A \in \mathcal{A}, \qquad \delta_x(A) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A. \end{cases}$$
En se servant des fonctions caractéristiques des ensembles $A \in \mathcal{A}$, on peut réécrire la définition de cette mesure par $\delta_x(A) = \mathbb{1}_A(x)$.On appelle δ_x la mesure de Dirac au point x.

Note historique : Paul Dirac était un physicien théoricien travaillant sur la mécanique quantique, il a utilisé la notation δ_x pour définir une « fonction singulière » sur $\mathbb{R}^d$ qui serait nulle en tout point $y \neq x$, mais telle que son « intégrale » $\int \delta_x(y)\, dy = 1$. Les mathématiciens ont interprété sa « fonction singulière » comme une mesure définie sur la tribu des boréliens de $\mathbb{R}^d$, telle que $\delta_x(\{x\}) = 1$ mais $\delta_x(A) = 0$ si $x \notin A$.

3. Pour deux mesures μ_1, μ_2 définies sur une même tribu, et pour $\alpha_1, \alpha_2 \in [0, \infty]$, on peut définir la mesure combinaison linéaire $\mu = \alpha_1\mu_1 + \alpha_2\mu_2$.

4. Mesure de Lebesgue (par exemple sur $E = \mathbb{R}$). Nous en avons déjà parlé dans l'introduction du cours, c'est cette mesure qui permettra de faire le lien avec l'intégrale de Riemann. Cette mesure est définie sur la tribu des boréliens de $\mathbb{R}$, elle mesure la « longueur » des boréliens. On la note λ (ou parfois m, ou parfois dL, ou parfois dx). Comme on le verra dans l'Exemple 2.17, c'est l'unique mesure sur $\mathcal{B}(\mathbb{R})$ telle que, pour tout intervalle $]a, b[$, on ait $\lambda\left(]a, b[\right) = b - a$. La construction rigoureuse de cette mesure (qui semble pourtant très intuitive) va réclamer du travail, nous la ferons dans une section ultérieure. Cette mesure se généralise aux espaces euclidiens $\mathbb{R}^d$ de toute dimension.

Les mesures peuvent être classifiées en plusieurs catégories.

Définition 2.10. 1. μ est dite *finie* si $\mu(E) < \infty$. Le nombre $\mu(E)$ est appelé la masse totale de μ. En particulier, si $\mu(E) = 1$, on appelle μ une *mesure de probabilité.*

2. Une mesure μ est dite *σ-finie* s'il existe une suite croissante $(E_n)_{n\in\mathbb{N}}$ de sous-ensembles mesurables, telle que $E = \bigcup_{n\in\mathbb{N}} E_n$, et $\mu(E_n) < \infty$ pour tout $n \in \mathbb{N}$. Cette définition inclut

formellement le cas où μ est finie, mais on l'utilise surtout dans des cas où $\mu(E) = \infty$. Par exemple, la mesure de comptage sur $(\mathbb{N}, \mathscr{P}(\mathbb{N}))$ est σ-finie, ce qu'on peut voir en prenant $E_n = [\![0, n]\!]$. Par contre, si on considère $(E, \mathscr{P}(E))$ avec un ensemble E indénombrable, alors la mesure de comptage sur E n'est pas σ-finie.

3. Un point $x \in E$ est appelé un *atome* de la mesure μ si $\{x\} \in A$ et $\mu(\{x\}) > 0$.

4. Une mesure μ est dite *diffuse* si elle n'a pas d'atomes.

5. Un ensemble $B \subset E$ tel qu'il existe $A \supset B$ mesurable avec $\mu(A) = 0$, est dit *μ-négligeable*. Si la mesure μ est sous-entendue, on dira plus simplement que l'ensemble B est négligeable. On dit que la tribu $\mathcal{A}$ est complète (relativement à la mesure μ) si tous les ensembles μ-négligeables sont dans $\mathcal{A}$. Il est possible de compléter la tribu $\mathcal{A}$, de manière à obtenir une tribu $\bar{\mathcal{A}} \supset \mathcal{A}$, complète pour μ.

Exemple 2.11. La mesure de Lebesgue sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ est une mesure diffuse : pour tout $x \in \mathbb{R}$, $\lambda(\{x\}) = 0$. Par σ-additivité, pour tout sous-ensemble $A \subset \mathbb{R}^d$ dénombrable, on aura aussi $\lambda(A) = 0$. En particulier, l'ensemble des rationnels $\mathbb{Q} \subset \mathbb{R}$ est de mesure de Lebesgue nulle.

On peut aussi construire des sous-ensembles indénombrables de $\mathbb{R}$ négligeables pour la mesure de Lebesgue. Par exemple. l'ensemble de Cantor $Can \subset [0, 1]$, composé des nombres réels pouvant s'écrire $x = \sum_{n\geq 1} \frac{k_n}{3^n}$ avec tous les $k_n \in \{0, 2\}$, forme un borélien sur $\mathbb{R}$, qui est indénombrable. On peut montrer que $\lambda(Can) = 0$.

Exercice 2.12. Sur $\mathbb{R}^2$, considérons un morceau fini courbe lisse $\mathcal{C}$. Montrer qu'elle est de mesure de Lebesgue nulle.

Indication : construire des recouvrements de $\mathcal{C}$ par des unions de petits cubes.

2.3 Lemme de classe monotone : identification d'une mesure

Avant de passer à la définition de l'intégrale, nous introduisons la structure de *classe monotone*, une notion proche de celle de la tribu, qui va nous être utile pour identifier une mesure μ sur une tribu $\mathcal{A}$ (par ex. la mesure de Lebesgue sur $\mathcal{B}(\mathbb{R})$) en ne connaissant que ses valeurs sur une sous-classe $\mathcal{C} \subset \mathcal{A}$ satisfaisant certains propriétés.

Définition 2.13. Une sous-famille $\mathcal{M}$ de $\mathcal{P}(E)$ est appelée une classe monotone si :

(i) $E \in \mathcal{M}$;

(ii) Si $A, B \in \mathcal{M}$ et $A \subset B$, alors $B \setminus A \in \mathcal{M}$;

(iii) Si les $A_n \in \mathcal{M}$ forment une suite croissante ($A_n \subset A_{n+1}$), alors $\bigcup_n A_n \in \mathcal{M}$.

En comparant ces conditions avec celles d'une tribu (Def. 2.1), on voit qu'une tribu est un cas particulier de classe monotone. D'une part, (i) et (ii) impliquent que pour tout $A \in \mathcal{A}$, $A^c \in \mathcal{A}$. Par contre, la condition (iii) ci-dessus est moins contraignante que dans le cas d'une tribu, puisque qu'elle n'est valable que si on considère une suite (A_n) croissante, tandis que pour une tribu . Les classes monotones sont donc « plus nombreuses » que les tribus parmi toutes les sous-familles de $\mathscr{P}(E)$.

Lemme 2.14. *Une classe monotone est une tribu si et seulement si elle est stable par intersections finies.*

Démonstration. Supposons qu'une classe monotone $\mathcal{M}$ est invariante par intersections finies. Comme $\mathcal{M}$ est invariante par passage au complémentaire, elle sera aussi invariante par unions finies. En se servant de la propriété (iii), on montre l'invariance par union dénombrable ; en effet, pour toute suite $(B_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{M}$, on peut construire $A_0 = B_0$, $A_1 = A_0 \cup B_1$, et similairement $A_n = A_{n-1} \cup B_n$. Grâce à l'invariance par union finie, on montre alors par récurrence que $A_n \in \mathcal{M}$. De plus, on a forcément $A_n \supset A_{n-1}$, donc la suite (A_n) est croissante. Le point (iii) montre donc que $\bigcup_n A_n = \bigcup_n B_n$ est dans $\mathcal{M}$. On a donc montré que $\mathcal{M}$ est une tribu □

Si $\mathcal{M}_1$ et $\mathcal{M}_2$ sont deux classes monotones, la famille $\mathcal{M}_1 \cap \mathcal{M}_2$ l'est aussi. On peut donc, comme pour les tribus, définir la plus petite classe monotone contenant une sous-famille $\mathcal{C}$ de $\mathscr{P}(E)$; on appelle cette classe $\mathcal{M}(\mathcal{C})$ la classe monotone engendrée par $\mathcal{C}$:

$$\mathcal{M}(\mathcal{C}) \stackrel{\text{def}}{=} \bigcap_{\mathcal{M}\,\text{cl. monot.},\ \mathcal{M}\supset\mathcal{C}} \mathcal{M}.$$

Le principal résultat de cette section est le Lemme de classe monotone. Sa preuve utilise déjà les méthodes qui seront utilisées pour construire la mesure de Lebesgue, ce que nous ferons dans un chapitre ultérieur.

Théorème 2.15. *[Lemme de classe monotone] Si $\mathcal{C} \subset \mathscr{P}(E)$ est stable par intersections finies, alors la classe monotone engendrée $\mathcal{M}(\mathcal{C})$ est égale à la tribu $\sigma(\mathcal{C})$.*

Démonstration. Comme les tribus sont toutes des classes monotones, la classe monotone engendrée par $\mathcal{C}$ est incluse dans la tribu engendrée : $\mathcal{M}(\mathcal{C}) \subset \sigma(\mathcal{C})$. On aura égalité si et seulement si on montre que $\mathcal{M}(\mathcal{C})$ est une tribu. D'après le lemme 2.14, pour montrer que

$\mathcal{M}(\mathcal{C})$ est une tribu, il suffit de montrer que $\mathcal{M}(\mathcal{C})$ est stable par intersections finies (on a juste fait l'hypothèse que $\mathcal{C}$ est stable par intersections finies).

Pour un élément $A \in \mathcal{M}(\mathcal{C})$ quelconque, on définit la famille

$$\mathcal{M}_A \stackrel{\text{def}}{=} \{B \in \mathcal{M}(\mathcal{C})\,;\, A \cap B \in \mathcal{M}(\mathcal{C})\}\,. \tag{2.2}$$

Notre objectif est alors de montrer que $\mathcal{M}_A = \mathcal{M}(\mathcal{C})$.

Pour commencer, on va supposer que $A \in \mathcal{C}$. Comme $\mathcal{C}$ est stable par intersection, pour tout $B \in \mathcal{C}$ on aura $A \cap B \in \mathcal{C} \subset \mathcal{M}(\mathcal{C})$, ce qui montre que $B \in \mathcal{M}_A$; donc la famille $\mathcal{C} \subset \mathcal{M}_A$. L'astuce consiste à montrer que $\mathcal{M}_A$ est une classe monotone :

(i) Comme $A \cap E = A$, on a évidemment $E \in \mathcal{M}_A$.

(ii) Supposons $B \subset B'$ deux éléments de $\mathcal{M}_A$: on aura alors $(A \cap B) \subset (A \cap B')$ tous deux dans $\mathcal{M}(\mathcal{C})$. Par conséquent, $(A \cap B') \setminus (A \cap B)$ est aussi dans $\mathcal{M}(\mathcal{C})$. Or ce sous-ensemble s'écrit aussi $A \cap (B' \setminus B)$: on a donc montré que $A \cap (B' \setminus B) \in \mathcal{M}(\mathcal{C})$, donc que $B' \setminus B \in \mathcal{M}_A$.

(iii) Prenons maintenant une suite croissante $(B_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{M}_A$. Les ensembles $(A \cap B_n)$ sont alors dans $\mathcal{M}(\mathcal{C})$, et ils forment une suite croissante ; leur union $\bigcup_n (A \cap B_n) = A \cap (\bigcup_n B_n)$ est donc dans $\mathcal{M}(\mathcal{C})$. Par conséquent, $\bigcup_n B_n \in \mathcal{M}_A$.

On a montré que pour $A \in \mathcal{C}$, la famille $\mathcal{M}_A$ satisfait les propriétés d'une classe monotone. Comme $\mathcal{C} \subset \mathcal{M}_A$ et que $\mathcal{M}(\mathcal{C})$ est la classe monotone la plus petite contenant $\mathcal{C}$, on en déduit $\mathcal{M}_A = \mathcal{M}(\mathcal{C})$. Par conséquent, nous avons montré que :

$$\forall A \in \mathcal{C}, \forall B \in \mathcal{M}(\mathcal{C}),\ A \cap B \in \mathcal{M}(\mathcal{C}). \tag{2.3}$$

Pour conclure, il faut élargir cette propriété à tout élément $A \in \mathcal{M}(\mathcal{C})$. On reprend alors la même stratégie, en définissant, pour un $A' \in \mathcal{M}(\mathcal{C})$ donné, la famille $\mathcal{M}_{A'}$ comme ci-dessus. La propriété (2.3) que nous venons de montrée, écrite en remplaçant (A, B) par (B, A'), montre que tout élément $B \in \mathcal{C}$ est dans $\mathcal{M}_{A'}$, de sorte que $\mathcal{C} \subset \mathcal{M}_{A'}$. La même preuve que ci-dessus montre que $\mathcal{M}_{A'}$ est une classe monotone, et donc qu'elle est égale à $\mathcal{M}(\mathcal{C})$. On a donc montré que cette dernière est invariante par intersection finie, ce qui achève la preuve the théorème. □

La principale conséquence de ce lemme est de permettre d'identifier deux mesures lorsqu'elles sont égales sur une sous-classe $\mathcal{C} \subset \mathcal{A}$.

Corollaire 2.16. *Soient μ, ν deux mesures sur un espace mesurable $(E, \mathcal{A})$. Supposons qu'il* *existe un famille $\mathcal{C} \subset \mathcal{A}$ stable par intersection finie, telle que $\sigma(\mathcal{C}) = \mathcal{A}$, et telle que pour tout $A \in \mathcal{C}$, on ait $\mu(A) = \nu(A)$. Si une des conditions suivantes est vérifiée :*

1. $\mu(E) = \nu(E) < \infty$;

2. μ et ν sont σ-finies par rapport à une suite croissante $(E_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{C}$ (c'est-à-dire $E = \bigcup_n E_n$, $\mu(E_n) = \nu(E_n) < \infty$) ;

alors $\mu = \nu$.

Démonstration. 1. Commençons par le cas de μ, ν finies. Encore une fois, il faut définir un classe astucieusement. On pose $\mathcal{G} \overset{\text{def}}{=} \{A \in \mathcal{A}\,;\, \mu(A) = \nu(A)\}$. D'une part, cette classe contient $\mathcal{C}$; d'autre part, on vérifie facilement que $\mathcal{G}$ est une classe monotone :

(i) il est évident que $E \in \mathcal{G}$.

(ii) si $A \subset B$ sont dans $\mathcal{G}$, alors $\mu(B \setminus A) = \mu(B) - \mu(A)$ (car $\mu(A) < \infty$), et de même pour ν. On en déduit que $\mu(B \setminus A) = \nu(B \setminus A)$, ce qui montre que $B \setminus A \in \mathcal{G}$.

(iii) si $(A_n)_{n\geq 0}$ est une suite croissante dans $\mathcal{G}$, alors $\mu(\bigcup_n A_n) = \lim \uparrow \mu(A_n) = \lim \uparrow \nu(A_n) = \nu(\bigcup_n A_n)$, donc $(\bigcup_n A_n) \in \mathcal{G}$.

Comme $\mathcal{C}$ est stable par intersection finie, le Lemme de classe monotone montre que $\mathcal{A} = \sigma(\mathcal{C}) = \mathcal{M}(\mathcal{C})$. Comme la classe monotone $\mathcal{G}$ contient $\mathcal{G}$, elle contient $\mathcal{M}(\mathcal{C})$, donc on a forcément $\mathcal{G} = \mathcal{A}$, ce qui revient à dire que $\mu = \nu$ sur la tribu $\mathcal{A}$ entière.

2. Considérons le 2e cas. Pour tout n on définit les mesures $\mu_n(A) = \mu_n(A \cap E_n)$, $\nu_n(A) = \nu_n(A \cap E_n)$, (on appelle μ_n la restriction de μ à l'ensemble E_n). Pour tout $A \in \mathcal{C}$ l'ensemble $A \cap E_n$ est aussi dans $\mathcal{C}$ (car $\mathcal{C}$ est stable par intersection), ce qui implique que $\mu_n(A) = \nu_n(A)$. Les mesures μ_n et ν_n satisfont donc l'hypothèse 1. ; d'après la preuve ci-dessus, elle sont donc égales. On montre enfin que pour tout $A \in \mathcal{A}$,

$$\mu(A) = \lim_n \uparrow \mu_n(A) = \lim_n \uparrow \nu_n(A) = \nu(A).$$

□

Exemple 2.17. *Unicité de la mesure de Lebesgue.* Nous n'avons pas encore construit la mesure de Lebesgue sur $\mathbb{R}$, mais nous avons déjà indiqué une condition qu'elle doit satisfaire : pour tout intervalle borné $]a, b[$, $\lambda\left(]a, b[\right) = b - a$. Si $\mathcal{C}$ est la famille de ces intervalles, on a déjà expliqué que la tribu des boréliens est engendrée par cette famille $\mathcal{C}$. D'autre part, cette famille est stable par intersection finie. Enfin, la suite $(E_n =]-n, n[)_{n\geq 1}$ satisfait bien $E_n \in \mathcal{C}$, elle est croissante, et $\bigcup_n E_n = \mathbb{R}$. Le corollaire ci-dessus montre alors qu'il existe au plus une seule mesure satisfaisant ces propriétés. Nous montrerons dans un chapitre ultérieur que cette mesure existe effectivement sur toute la tribu $\mathcal{B}(\mathbb{R})$.

Un seconde application sera très utile lorsqu'on s'intéressera aux mesures de probabilité.

Exemple 2.18. *Fonction de répartition d'une mesure finie sur* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Soit μ une mesure finie sur $\mathbb{R}$. Pour identifier μ, il suffit de connaître ses valeurs sur la famille $\{]-\infty, a[\,,\ a \in \mathbb{R}\}$, ou bien ses valeurs sur la famille $\{]-\infty, a]\,,\ a \in \mathbb{R}\}$. Par convention, en théorie des probabilités on utilise plutôt la seconde famille : la fonction $a \mapsto \mu\,(]-\infty, a])$ est alors appelée la fonction de répartition de μ.

2.4 Fonctions mesurables (entre deux tribus)

Après avoir défini la classe des sous-ensembles mesurables (= la tribu), on peut définir une classe de fonctions relant deux espaces mesurables, qui sont compatibles avec cette structure mesurable. On les appelle les fonctions mesurables.

Remarque 2.19. Il faudra faire attention au mot « mesurable », qui s'applique à différents objets de natures différentes :

- un espace mesurable $(E, \mathcal{A})$, c'est-à-dire un espace muni d'une tribu

- un ensemble mesurable $A \in \mathcal{A}$, élément de la tribu dont il est question

- une fonction mesurable, notion qui fait référence à deux tribus $\mathcal{A}$ et $\mathcal{B}$.

2.4.1 Définition et premières propriétés des fonctions mesurables

Définition 2.20. Soient $(E, \mathcal{A})$ et $(F, \mathcal{B})$ deux espaces mesurables. Une application $f : E \to F$ est dite mesurable si

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

Dans le cas où E et F sont des espaces topologiques munis de leurs tribus boréliennes, on dit qu'une telle fonction f est borélienne. Très souvent l'espace image sera $\mathbb{R}$, $\mathbb{C}$ ou $\mathbb{R}^d$ (fonction à valeurs réelles, complexes ou vectorielles), dans ce cas $\mathcal{B}$ sera la tribu des boréliens.

Remarque 2.21. On ne demande pas que l'image d'un sous-ensemble mesurable soit mesurable, mais que l'*image réciproque* d'un mesurable le soit. La situation est analogue au cas d'une fonction entre deux espaces topologiques : la fonction est continue si l'image réciproque de tout ouvert est un ouvert.

Dans notre introduction, on avait rencontré les ensembles $f^{-1}([y_{k-1}, y_k[)$ dans la définition intuitive de l'intégrale de Lebesgue, avec $f : I \to \mathbb{R}$; on remarque que les intervalles $[y_{k-1}, y_k[$ sont des boréliens de $\mathbb{R}$. Il faut donc s'assurer que les $f^{-1}([y_{k-1}, y_k[)$ sont aussi des boréliens.

Proposition 2.22. *La composition de deux fonctions mesurables est encore mesurable.*

Démonstration. Immédiat si on remarque que $(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C))$. □

En général il est difficile de vérifier la propriété pour tous les éléments de la tribu (par exemple, on a vu la tribu des boréliens sur $\mathbb{R}$ contient des ensembles compliqués). Heureusement, la proposition suivante montre que pour prouver la mesurabilité de f, il suffit de se restreindre à une sous-famille de $\mathcal{B}$ qui engendre la tribu.

Proposition 2.23. *Supposons que la tribu $\mathcal{B}$ est engendrée par une sous-famille $\mathcal{C} \subset \mathcal{B}$. Pour vérifier que $f : E \to F$ est mesurable, il suffit de vérifier que $f^{-1}(B) \in \mathcal{A}$ pour tout B élément de la famille $\mathcal{C}$.*

Démonstration. On considère la sous-famille de la tribu $\mathcal{B}$ définie par $\mathcal{G} \stackrel{\text{def}}{=} \{B \in \mathcal{B} \,;\, f^{-1}(B) \in \mathcal{A}\}$. On voit que $\mathcal{G}$ a les propriétés d'une tribu :

(i) $f^{-1}(\emptyset) = \emptyset$,

(ii) $f^{-1}(B^c) = \left(f^{-1}(B)\right)^c$. On remarque que cette propriété est fausse si on remplace f^{-1} par f. Cela explique pourquoi la mesurabilité est définie à partir de l'image réciproque, et non de l'image directe.

(iii) $f^{-1}(\bigcup B_n) = \bigcup_n f^{-1}(B_n)$.

Supposons maintenant que pour tout B dans la famille $\mathcal{C}$, $f^{-1}(B) \in \mathcal{A}$. Cela implique que la sous-tribu $\mathcal{G}$ contient la famille $\mathcal{C}$. Donc, par définition, $\mathcal{G}$ contient la tribu engendrée par $\mathcal{C}$, qui n'est autre que $\mathcal{B}$ par hypothèse ; comme $\mathcal{G} \subset \mathcal{B}$, on en déduit donc $\mathcal{G} = \mathcal{B}$, autrement dit : $f^{-1}(B) \in \mathcal{A}$ pour tout $B \in \mathcal{B}$. On a montré que f est mesurable entre les tribus $\mathcal{A}$ et $\mathcal{B}$. □

Corollaire 2.24. *Soit une fonction $f : (E, \mathcal{A}) \to \mathbb{R}$; comme expliqué plus haut, on considèrera implicitement que la tribu associée à $\mathbb{R}$ est la tribu borélienne $\mathcal{B}(\mathbb{R})$. Donc on est dans le cas $F, \mathcal{B}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Comme la tribu $\mathcal{B}(\mathbb{R})$ est engendrée par les intervalles $]\alpha, +\infty[$, pour vérifier que f est mesurable (entre $\mathcal{A}$ et $\mathcal{B}(\mathbb{R})$), il suffit donc de montrer que*

$$\forall \alpha \in \mathbb{R}, \quad f^{-1}(]\alpha, +\infty[) \in \mathcal{A}.$$

On aurait pu choisir comme autre critère, par exemple $f^{-1}(]-\infty,\alpha[) \in \mathcal{A}$*, ou bien* $f^{-1}(]-\infty,\alpha]) \in \mathcal{A}$.

Ce corollaire nous fournit le résultat important suivant, qui relie les propriétés topologiques aux propriétés de mesurabilité :

Corollaire 2.25. *Si* E, F *sont deux espaces topologiques munis de leurs tribus boréliennes, alors une application* $f : E \to F$ *continue est aussi mesurable (entre* $\mathcal{B}(E)$ *et* $\mathcal{B}(F)$*).*

Démonstration. On sait que pour $f : E \to F$ continue, pour tout ouvert $V \subset F$, l'image réciproque $f^{-1}(V)$ est un ouvert de E, en particulier c'est un borélien de E. Comme (par définition de la tribu borélienne) l'ensemble des ouverts de F engendre $\mathcal{B}(F)$, la proposition 2.23 montre que f est mesurable. □

Tribu définie à partir d'une famille de fonctions

On peut procéder dans le sens inverse : pour un espace E donné, on peut partir d'une famille $\mathcal{F}$ de fonctions $f : E \to \mathbb{R}$. Y a-t-il alors une tribu $\mathcal{A}$ sur E, telle que toutes les fonctions $f \in \mathcal{F}$ sont mesurables de $\mathcal{A}$ vers $\mathcal{B}(\mathbb{R})$?

Proposition 2.26. *Si on se donne une famille* $\mathcal{F}$ *de fonctions* $f : E \to \mathbb{R}$*, il existe une tribu minimale* $\mathcal{A} = \mathcal{A}_{\mathcal{F}}$ *sur* E*, telle que toutes les fonctions* $f \in \mathcal{F}$ *sont mesurables.*

Démonstration. Considérons la famille de sous-ensembles de E définie par

$$\mathcal{R}_{\mathcal{F}} \stackrel{\text{def}}{=} \left\{ f^{-1}(]-\infty,\alpha[) \subset E,\ f \in \mathcal{F},\ a \in \mathbb{R} \right\},$$

et soit $\mathcal{A}_{\mathcal{F}} \stackrel{\text{def}}{=} \sigma(\mathcal{R}_{\mathcal{F}})$ la tribu sur E engendrée par $\mathcal{R}_{\mathcal{F}}$. Comme les intervalles $\{]-\infty,\alpha[\,, \alpha \in \mathbb{R}\}$ engendrent les boréliens de $\mathbb{R}$, et que pour toute fonction $f \in \mathcal{F}$, les images réciproques $f^{-1}(]-\infty,\alpha[)$ sont dans $\mathcal{R}_{\mathcal{F}}$, donc dans la tribu $\mathcal{A}_{\mathcal{F}}$, alors la Proposition 2.23 montre que cette fonction f est mesurable sur $\mathcal{A}_{\mathcal{F}}$. On voit aussi que la tribu $\mathcal{A}_{\mathcal{F}}$ est la plus petite ayant cette propriété : n'importe quelle tribu rendant les f mesurables doit forcément contenir $\mathcal{R}_{\mathcal{F}}$, donc contenir la tribu $\mathcal{A}_{\mathcal{F}} = \sigma(\mathcal{R}_{\mathcal{F}})$.

□

Exemple 2.27. Si $E = F = \mathbb{R}$ et $\mathcal{F} = C^0(\mathbb{R},\mathbb{R})$ la famille des fonctions continues $\mathbb{R} \to \mathbb{R}$, alors la plus petite tribu rendant ces fonctions mesurables est la tribu des boréliens de $\mathbb{R}$.

Démonstration. Pour tout intervalle ouvert $]a, b[$, il existe une fonction continue $f : \mathbb{R}^d \to \mathbb{R}$ telle que $f^{-1}(]-\infty, 0[) =]a, b[$. On peut par exemple prendre la fonction à symétrie sphérique $f(x) = \left|x - \frac{a+b}{2}\right| - \frac{b-a}{2}$. Une tribu $\mathcal{A}$ rendant les fonctions f mesurable va donc contenir tous les intervalles $]a, b[$. On a déjà vu qu'une telle tribu doit alors contenir tous les boréliens de $\mathbb{R}$. D'un autre côté, le corollaire 2.25 montre que les fonctions continues sont mesurables $\mathcal{B}(\mathbb{R}) \to \mathcal{B}(\mathbb{R})$, donc la tribu $\mathcal{A}$ minimale satisfaisant cette propriété est bien la tribu des boréliens.

□

2.4.2 Opérations sur les fonctions mesurables

A partir de deux fonctions *scalaires* (= à valeurs dans $\mathbb{R}$) $f_1, f_2 : E \to \mathbb{R}$, on obtient une fonction à valeurs vectorielles $f = (f_1, f_2) : E \to \mathbb{R}^2$. Qu'en est-il de leur mesurabilité ?

Proposition 2.28. *Soient $f_1 : (E, \mathcal{A}) \mapsto (F_1, \mathcal{B}_1)$ et $f_2 : (E, \mathcal{A}) \mapsto (F_2, \mathcal{B}_2)$ deux fonctions mesurables. Alors l'application produit $f : (E, \mathcal{A}) \mapsto (F_1 \times F_2, \mathcal{B}_1 \otimes \mathcal{B}_2)$ définie par $f(x) = (f_1(x), f_2(x))$ est aussi mesurable. La réciproque est vraie : si $f = (f_1, f_2)$ est mesurable, chacune de ses composantes l'est aussi.*

Démonstration. Par la définition 2.6, la tribu $\mathcal{B}_1 \otimes \mathcal{B}_2$ est engendrée par la famille des $\mathcal{C} = \{B_1 \times B_2,\ B_i \in \mathcal{B}_i,\ i = 1, 2\}$. D'après la Proposition 2.23, pour vérifier que f est mesurable il suffit de montrer que les ensembles $f^{-1}(B_1 \times B_2)$ sont dans $\mathcal{A}$. Or que valent ces ensembles ?

$$\begin{aligned} f^{-1}(B_1 \times B_2) &= \{x \in E\,;\ f(x) \in B_1 \times B_2\} = \{x \in E\,;\ f_1(x) \in B_1 \text{ et } f_2(x) \in B_2\} \\ &= \{x \in E\,;\ f_1(x) \in B_1\} \cap \{x \in E\,;\ f_2(x) \in B_2\} \\ &= f_1^{-1}(B_1) \cap f_2^{-1}(B_2). \end{aligned}$$

Comme f_1 et f_2 sont mesurables, les ensembles $f_i^{-1}(B_i)$ sont dans $\mathcal{A}$, donc leur intersection l'est aussi. Donc f est bien mesurable.

Inversement, si f est supposée mesurable, alors pour tout $B_1 \in \mathcal{B}_1$, $f^{-1}(B_1 \times F_2) = f_1^{-1}(B_1) \cap f_2^{-1}(F_2) = f_1^{-1}(B_1) \cap E$ est dans $\mathcal{A}$; cela montre que f_1 est mesurable. En échangeant les indices, on montre que f_2 est également mesurable. □

Ce résultat sur la mesurabilité de la fonction vectorielle $f = (f_1, f_2)$ va nous permettre des opérations sur les fonctions scalaires.

Corollaire 2.29. *Soient $f, g : (E, \mathcal{A}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ deux fonctions mesurables. Soient $\alpha, \beta \in \mathbb{R}$. Alors les fonctions*

$$\alpha f + \beta g, \quad f^{+} = \max(f, 0), \quad f^{-} = \max(-f, 0) \qquad \textit{sont mesurables.}$$

Démonstration. On se sert du Lemme précédent. On sait que la fonction $x \mapsto (f(x), g(x))$ est mesurable de $\mathcal{A}$ dans $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^2)$. D'autre part, l'application

$$\begin{aligned} \mathbb{R}^2 &\to \mathbb{R} \\ (a, b) &\mapsto \alpha a + \beta b \end{aligned}$$

est continue, donc elle est mesurable $\mathcal{B}(\mathbb{R}^2) \to \mathcal{B}(\mathbb{R})$ (cf. le corollaire 2.25). La composée de ces deux applications mesurables est donc mesurable.

Pour les deux autres énoncés, il suffit de remarquer que les applications $\mathbb{R} \to \mathbb{R}$ données par $a \mapsto \max(a, 0)$, respectivement $a \mapsto \max(-a, 0)$ sont aussi continues, donc mesurables. □

2.4.3 Suites de fonctions mesurables à valeurs dans $\overline{\mathbb{R}}$

On va maintenant considérer des suites de fonctions. Même si chaque f_n est à valeurs dans $\mathbb{R}$, il se peut qu'en certains points x, les valeurs $(f_n(x))_n$ admettent comme valeur d'adhérence $+\infty$ ou $-\infty$. On va donc étendre $\mathbb{R}$ en $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\} = [-\infty, +\infty]$, qu'on appelle $\overline{\mathbb{R}}$ la *droite réelle achevée*, et on supposera que les fonctions f_n prennent déjà leurs valeurs dans $\overline{\mathbb{R}}$. Arrêtons-nous un peu sur la topologie de cette droite achevée. $\overline{\mathbb{R}}$ est homéomorphe à l'intervalle $[-1, 1]$, par exemple à travers la fonction $x \mapsto \tanh(x)$. Ainsi, les ouverts de $\overline{\mathbb{R}}$ incluent les ouverts de $\mathbb{R}$, mais également les intervalles du type $[-\infty, a[$ ou $]a, +\infty]$. A partir de cette topologie, on construit la tribu borélienne $\mathcal{B}(\overline{\mathbb{R}})$.

Exercice 2.30. Montrer que la tribu $\mathcal{B}(\overline{\mathbb{R}})$ est engendrée par les intervalles $\{[-\infty, a[, \ a \in \mathbb{R}\}$.

On va maintenant s'intéresser aux opérations sur les suites des fonctions. Soit $(a_n)_{n \in \mathbb{N}}$ une suite d'éléments de $\overline{\mathbb{R}}$. Alors la limite supérieure (respectivement limite inférieure) de cette suite sont définies par :

$$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} \downarrow \left(\sup_{k\geq n} a_k\right) = \inf_{n\geq 0} \left(\sup_{k\geq n} a_k\right)_n,$$

$$\text{respectivement} \quad \liminf_{n\to\infty} a_n = \lim_{n\to\infty} \uparrow \left(\inf_{k\geq n} a_k\right) = \sup_{n\geq 0} \left(\inf_{k\geq n} a_k\right). \tag{2.4}$$

La lim sup et la lim inf existent pour toute suite $(a_n)_{n\geq 0}$, ce sont les plus grande et plus petite valeurs d'adhérence de la suite. La suite $(a_n)_{n\in\mathbb{N}}$admet une limite en $n \to \infty$ si et seulement si $\liminf a_n = \limsup a_n$ (la limite est alors égale à la valeur commune). Notons que ces limites supérieure et inférieure prennent leurs valeurs dans $\overline{\mathbb{R}}$.

On s'intéresse maintenant à des suites de fonctions mesurables.

Proposition 2.31. *Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables de $(E, \mathcal{A})$ vers $\big(\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}})\big)$. Alors les fonctions*

$$\sup_n f_n, \quad \inf_n f_n, \quad \limsup_n f_n, \quad \liminf_n f_n \qquad \textit{sont aussi mesurables.}$$

En particulier, si la suite $(f_n)_{n\in\mathbb{N}}$ converge simplement vers une fonction $f : E \to \overline{\mathbb{R}}$, alors f est mesurable.

Enfin, pour une suite $(f_n)_{n\in\mathbb{N}}$ quelconque, le sous-ensemble de E,

$$\{x \in E\,;\; f_n(x) \textit{ admet une limite lorsque } n \to \infty\}\,,$$

est mesurable.

Cette proposition nous montre que la notion de mesurabilité s'accomode bien de suites de fonctions admettant une limite *simple*. C'est un des grands atouts de la théorie de Lebesgue.

Démonstration. Définissons la fonction $f \overset{\text{def}}{=} \inf_n f_n : E \to \overline{\mathbb{R}}$. En adaptant le Corollaire 2.24 au cas de $\overline{\mathbb{R}}$, on voit que pour montrer que f est mesurable il suffit de montrer que pour tout $\alpha \in \mathbb{R}$, $f^{-1}\left([-\infty, \alpha[\right)$ est mesurable. Remarquons que $\inf_n f_n(x) < \alpha$ implique qu'il existe un $n_0 \in \mathbb{N}$ tel que $f_n(x) < \alpha$. Ainsi,

$$\begin{aligned} f^{-1}\left([-\infty, \alpha[\right) = \Big\{x \in E\,;\, \inf_n f_n(x) < \alpha\Big\} &= \bigcup_n \{x \in E\,;\; f_n(x) < \alpha\} \\ &= \bigcup_n f_n^{-1}\left([-\infty, \alpha[\right). \end{aligned}$$

Par hypothèse chaque $f_n^{-1}\left([-\infty, \alpha[\right)$ est mesurable, donc leur union l'est également. Pour traiter le cas de $\sup_n f_n$, on peut noter que $\sup_n f_n = -\left(\inf_n(-f_n)\right)$.

En combinant ces résultats, on obtient que $g_n \overset{\text{def}}{=} \left(\sup_{k\geq n} f_k\right)$ est mesurable pour tout $n \geq 0$, et donc que $\inf_n g_n = \limsup_k f_k$ est mesurable.

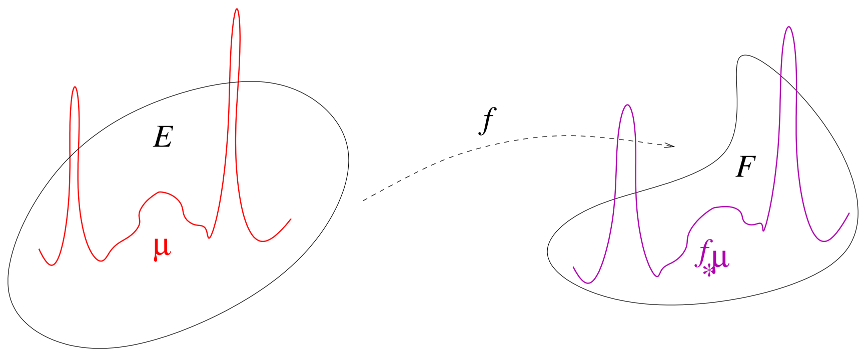

FIGURE 2.1 – Mesure image par une application $f : E \to F$. Les courbes rouge et violette schématisent des mesures sur E et sur F.

Le sous-ensemble de la dernière assertion peut être défini par l'égalité $\limsup_n f_n(x) = \liminf_n f_n(x)$. Les deux fonctions $\limsup_n f_n$ et $\liminf_n f_n$ sont mesurables, donc l'application $g(x) \stackrel{\text{def}}{=} \limsup_n f_n(x) - \liminf_n f_n(x)$ l'est aussi. L'ensemble que nous cherchons vaut $g^{-1}(\{0\})$, il est mesurable puisque $\{0\}$ est un borélien. $\square$

2.4.4 Transport d'une mesure par une application

Une façon de fabriquer une mesure sur un espace mesurable $(F, \mathcal{B})$ est de partir d'une mesure μ sur un espace mesurable $(E, \mathcal{A})$, et de la « pousser » vers F au moyen d'une application mesurable $f : (E, \mathcal{A}) \to (F, \mathcal{B})$.

Définition 2.32. Soit $f : (E, \mathcal{A}) \to (F, \mathcal{B})$ une application mesurable, et soit μ une mesure positive sur $(E, \mathcal{A})$. La mesure image de μ par f, notée $f_*\mu$ (ou parfois $f(\mu)$), est la mesure positive sur $(F, \mathcal{B})$ définie par :

$$\forall B \in \mathcal{B}, \quad f_*\mu(B) = \mu\left(f^{-1}(B)\right).$$

On vérifie facilement que μ est une mesure grâce aux propriétés suivantes :

- $f^{-1}(\emptyset) = \emptyset$, donc $f_*\mu(\emptyset) = 0$

- $f^{-}(\bigsqcup_n B_n) = \bigsqcup_n f^{-1}(B_n)$.

La mesure image $f_*\mu$ aura la même masse totale que μ. Par contre, il est possible que μ soit σ-finie, sans que $f_*\mu$ le soit.

Exercice 2.33. On prend la mesure de Lebesgue λ sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, et on considère des fonctions $f : \mathbb{R} \to \mathbb{R}$ mesurables (par rapport à la tribu borélienne sur $\mathbb{R}$). Décrire la mesure $f_*\lambda$ dans le cas des applications suivantes :

1. $f(x) = 0$ pour tout $x \in \mathbb{R}$. La mesure $f_*\mu$ est-elle σ-finie ?

2. $f(x) = 2x$ pour tout $x \in \mathbb{R}$.

3 Intégration d'une fonction par rapport à une mesure

Nous avons défini une mesure sur une tribu de sous-ensembles mesurables ; nous avons défini ce qu'est une fonction mesurable. Il ne reste plus qu'à combiner les deux notions, en définissant l'intégrale d'une fonction mesurable par rapport à cette mesure. Comme dans le cas de l'intégrale de Riemann, on commencera par définir l'intégrale sur une famille particulière de fonctions : les fonctions étagées, qui sont mesurables mais ne prennent qu'un nombre fini de valeurs. Ceci nous amènera à la définition de l'intégrale (de Lebesgue) d'une fonction mesurable *positive*. La notion de fonction intégrable nous permettra d'étendre la définition aux fonctions changeant de signe.

Ce chapitre nous permettra de présenter les principaux théorèmes de convergence pour l'intégrale de Lebesgue, par rapport à une suite de fonctions mesurables (f_n) convergeant *simplement* vers une fonction limite :

1. Le théorème de convergence monotone
2. Le Lemme de Fatou
3. Le théorème de convergence dominée, qui illustre la « puissance » de la théorie de Lebesgue sur celle de Riemann.

Tout au long de ce chapitre, la mesure d'intégration sera quelconque.

3.1 Intégration d'une fonction mesurable positive

Contrairement à la définition de l'intégrale de Riemann, l'intégrale de Lebesgue d'une présente une forme d'asymétrie, en cela qu'on cherche des fonctions approximantes par valeurs inférieures à f. C'est pourquoi on la définition concerne tout d'abord des fonctions positives.

3.1.1 Fonctions étagées et définition de l'intégrale de fonctions positives

Comme indiqué dans l'introduction, l'idée sera d'approximer f par des *fonctions étagées* (en anglais on dit « simple functions »).

Définition 3.1. Soit $(E, \mathcal{A})$ un espace mesurable. Une fonction $f : (E, \mathcal{A}) \to \mathbb{R}$ est dite étagée si elle ne prend qu'un nombre fini de valeurs distinctes $\alpha_1 < \alpha_2 < \cdots < \alpha_n$, et que pour chaque valeur α_j, l'ensemble $A_j = f^{-1}(\{\alpha_j\})$ est $\mathcal{A}$-mesurable. Autrement dit,

$$f(x) = \sum_{j=1}^{n} \alpha_j \, \mathbb{1}_{A_j}(x), \qquad A_j \in \mathcal{A}, \quad E = \bigsqcup_{j=1}^{n} A_j. \tag{3.1}$$

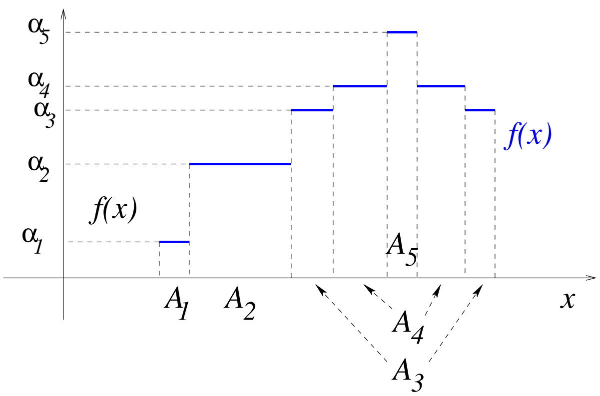

FIGURE 3.1 – Fonction étagée

Cette représentation de f est appelée son *écriture canonique* (on montre facilement qu'elle est unique).

Donnons-nous maintenant une mesure μ définie sur la tribu $\mathcal{A}$. On peut alors définir l'intégrale d'une fonction étagée, relativement à cette mesure.

Définition 3.2. Supposons la fonction étagée f à valeurs dans $\mathbb{R}_+$, définie sur l'espace mesuré $(E, \mathcal{A}, \mu)$. L'intégrale de f par rapport à μ est définie par :

$$\int f\, d\mu := \sum_{j=1}^{n} \alpha_j \mu\left(A_j\right).$$

On prend la convention que si $\alpha_j = 0$ et $\mu(A_j) = \infty$, alors $\alpha_j \mu(A_j) = 0$. A priori, l'intégrale prend ses valeurs dans $[0, \infty]$.

Ecritures non canoniques d'une fonction étagée

Plus généralement, une fonction étagée f peut être définie par une partition mesurable $E = \bigsqcup_{i=1}^{N} B_i$, et une suite de valeurs réelles $\beta_1, \ldots, \beta_N$, non nécessairement distinctes. On définit alors

$$f = \sum_{i=1}^{N} \beta_i \mathbb{1}_{B_i}.$$

Pour retrouver l'écriture canonique (3.1), il suffit de regrouper ensemble les β_i qui sont égaux à un α_j donné : les indices i correspondants forment un sous-ensemble $I(\alpha_j) \subset [\![1, N]\!]$. On a alors $A_j = \sqcup_{i \in I(\alpha_j)} B_i$, et donc $\mathbb{1}_{A_j} = \sum_{i \in I(\alpha_j)} \mathbb{1}_{B_i}$. L'additivité de la mesure μ implique alors que $\mu(A_j) = \sum_{i \in I(\alpha_j)} \mu(B_i)$, de sorte que

$$\int f\, d\mu = \sum_{i=1}^{N} \beta_i \mu(B_i).$$

Ainsi, la formule ci-dessus pour $\int f\, d\mu$ n'est pas seulement valable pour sa décomposition canonique de la fonction étagée f, mais pour n'importe quelle décomposition (finie).

Proposition 3.3. *Soient f et g deux fonctions étagées positives sur l'espace mesuré $(E, \mathcal{A}, \mu)$.*

1. Pour tous réels $a, b \geq 0, \int (af + bg)\, d\mu = a \int f\, d\mu + b \int g\, d\mu$.

2. Si $f \leq g$, alors $\int f\, d\mu \leq \int g\, d\mu$.

Démonstration. 1. La fonction $af + bg$ est une fonction étagée positive. En effet, si les fonctions f, g se décomposent en $f = \sum_{i=1}^{N} \beta_i \mathbb{1}_{B_i}$ et $g = \sum_{i=1}^{N'} \beta'_i \mathbb{1}_{B'_i}$, alors les ensembles $\left\{C_{ij} = B_i \cap B'_j, i = 1, \ldots, N,\, j = 1, \ldots, N'\right\}$ forment une partition de E, sur laquelle f et g peuvent toutes deux se décomposer (on dit que la partition $\sqcup C_{ij}$ est plus « fine » que les partitions $\sqcup B_i$ et $\sqcup B'_j$). On pourra donc écrire $f = \sum_{ij} \beta_i \mathbb{1}_{C_{ij}}$, $g = \sum_{ij} \beta'_j \mathbb{1}_{C_{ij}}$, et donc

$$af + bg = \sum_{ij} \gamma_{ij} \mathbb{1}_{C_{ij}}, \text{ avec les coefficients } \gamma_{ij} = a\beta_i + b\beta'_j.$$

Les fonctions $f, g, af + bg$ sont toutes trois définies sur la même partition $\sqcup C_{ij}$, on peut facilement relier leurs intégrales :

$$\begin{aligned}
\int (af + bg)\, d\mu &= \sum_{ij} \gamma_{ij} \mu(C_{ij}) = \sum_{ij} \left(a\beta_i + b\beta'_j\right) \mu(C_{ij}) \\
&= a \sum_i \sum_j \beta_i \mu(C_{ij}) + b \sum_j \sum_i \beta'_j \mu(C_{ij}) \\
&= a \sum_i \beta_i \mu(B_i) + b \sum_j \beta'_j \mu(B'_j).
\end{aligned}$$

2. Si $f \geq g$, la fonction $f - g$ est étagée positive, et $f = g + (f - g)$. On retombe sur le cas précédent. Comme $\int (f - g)\, d\mu \geq 0$, on obtient le résultat. □

Notations

$\mathcal{E}_+$ est l'espace des fonctions étagées positives. La proposition ci-dessus utilise l'invariance cet espace par la somme de fonctions, et par leur multiplication par un nombre réel positif.

On dira qu'une fonction f est *mesurable positive* sur $(E, \mathcal{A})$ si elle est mesurable, et qu'elle prend ses valeurs dans $\overline{\mathbb{R}}_+ = [0, \infty]$.

On peut maintenant définir l'intégrale d'une fonction mesurable positive, à partir de celle des fonctions étagées.

Définition 3.4. [Intégrale d'une fonction mesurable positive] Soit $(E, \mathcal{A}, \mu)$ un espace mesuré, et soit $f : (E, \mathcal{A}) \to \overline{\mathbb{R}}_+$ une fonction mesurable positive. L'intégrale de f est définie par

$$\int f \, d\mu := \sup_{h \in \mathcal{E}_+, h \leq f} \int h \, d\mu.$$

Selon le contexte, l'intégrale d'une fonction peut être notée indifféremment $\int f \, d\mu = \int f(x) \, d\mu(x) = \mu(f) = \langle \mu, f \rangle$ (ou parfois $\int f$ lorsque la mesure est évidente).

Proposition 3.5. *1. Si f est une fonction étagée, toute fonction étagée $h \leq f$ vérifiera $\int h \, d\mu \leq \int f \, d\mu$, tandis que l'égalité sera attente en prenant $h = f$. La définition ci-dessus est donc bien compatible avec celle d'une fonction étagée.*

2. Si $f \leq g$ sont mesurables positives, alors la définition implique directement que $\int f \, d\mu \leq \int g \, d\mu$.

3. Si f est mesurable positive et $\mu(\{x \in E \, ; \, f(x) > 0\}) = 0$, alors $\int f \, d\mu = 0$. En effet, cette hypothèse implique que $\mu(\{x \in E \, ; \, h(x) > 0\}) = 0$ pour toute fonction étagée positive $h \leq f$, ce qui implique que $\int h \, d\mu = 0$. On peut noter que cette hypothèse sur f autorise à ce que $f(x) = \infty$ sur un sous-ensemble $A \subset E$ (avec nécessairement $\mu(A) = 0$).

Remarque 3.6. La propriété 3. ci-dessus montre que les valeurs de f sur un sous-ensemble $A \subset E$ de mesure nulle ne sont par pertinentes dans le calcul de l'intégrale : ces valeurs sont « peu importantes ». Ceci nous amène à la définition d'un ensemble négligeable.

Définition 3.7. [Ensembles négligeables - propriété vraie presque partout] On dira qu'un ensemble $A \subset E$ est *μ-négligeable* s'il existe un $B \in \mathcal{A}$ tel que $A \subset B$ et $\mu(B) = 0$. Si la mesure μ est implicite, on dira que A est un ensemble négligeable. On dira qu'une propriété est vérifiée μ-presque partout (souvent abrégé par « μ-p.p. », ou « p.p. » lorsque la mesure est connue) si elle est vraie en-dehors d'un sous-ensemble négligeable. On dit aussi qu'une telle propriété est *presque sure*.

La plupart des énoncés de la théorie de l'intégration feront référence à des propriétés vraies p.p.

3.1.2 Le théorème de convergence monotone

La définition de l'intégrale se fait par approximations inférieures de f. Cette définition débouche rapidement sur le premier théorème de convergence pour des suites de fonctions.

Théorème 3.8. *[Théorème de convergence monotone] Soit $(f_n)_{n\in\mathbb{N}}$ une suite croissante de fonctions mesurables positives (à valeurs dans $[0,\infty]$), et définissons $f = \sup_n f_n = \lim_n \uparrow f_n$. Alors, pour toute mesure μ sur $\mathcal{A}$, on a :*

$$\int f\,d\mu = \lim_n \uparrow \int f_n\,d\mu.$$

Remarquons que la limite $f(x) = \lim_n \uparrow f_n(x)$ est une limite simple, pas forcément une limite uniforme.

Démonstration. On sait que déjà que $f = \sup f_n$ est mesurable et positive, donc l'intégrale $\int f\,d\mu$ est bien définie. Comme (f_n) est croissante, la suite $\left(\int f_n\,d\mu\right)_n$ est également croissante, cf. Proposition 3.5 (2). Comme $f \geq f_n$ pour tout $n \geq 0$, on a aussi $\int f\,d\mu \geq \int f_n\,d\mu$, et donc

$$\int f\,d\mu \geq \lim_n \uparrow \int f_n d\mu.$$

Pour montrer l'égalité, il nous faut montrer l'inégalité inverse. Choisissons une fonction étagée positive $h \leq f$. Fixons $0 < \epsilon < 1$. Comme $f_n(x) \to f(x)$, on définit pour tout n l'ensemble $E_n = \{x \in E : (1-\epsilon)h(x) \leq f_n(x)\}$. Comme la suite $(f_n)_{n\in\mathbb{N}}$ est croissante, la suite d'ensembles $(E_n)_n$ est également croissante, et comme $f_n(x) \to f(x)$, pour tout x on va avoir, à partir d'un certain rang, $f_n(x) \geq (1-\epsilon)h(x)$; ceci montre que $\bigcup_{n=0}^{\infty} E_n = E$.

Par définition de E_n, on a $f_n \geq (1-\epsilon)\mathbb{1}_{E_n} h$, de sorte que $\int f_n\,d\mu \geq (1-\epsilon)\int \mathbb{1}_{E_n} h\,d\mu$. Souvenons-nous que h est une fonction étagée : $h = \sum_{k=1}^{K} \alpha_k \mathbb{1}_{A_k}$, donc

$$\int \mathbb{1}_{E_n} h\,d\mu = \sum_{k=1}^{K} \alpha_k \mu(A_k \cap E_n).$$

Comme les $(E_n)_n$ recouvrent tout l'ensemble E, on a $\lim_n \uparrow \mu(A_k \cap E_n) = \mu(A_k)$. On trouve donc, à la limite $n \to \infty$:

$$\lim_n \uparrow \int f_n\,d\mu \geq (1-\epsilon) \lim_n \uparrow \sum_{k=1}^{K} \alpha_k \mu(A_k \cap E_n) = (1-\epsilon)\int h\,d\mu.$$

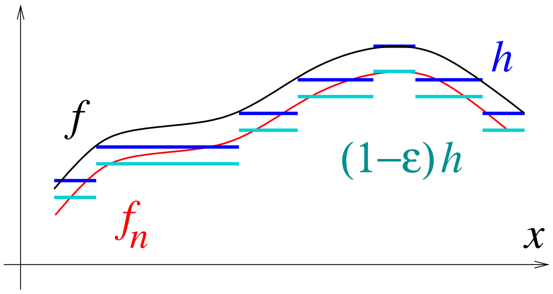

FIGURE 3.2 – Les fonctions $f, f_n, h, (1-\epsilon)h$ apppараissant dans la preuve du théorème de convergence monotone

Cette équation est vraie pour tout $\epsilon > 0$, donc on peut prendre $\epsilon \to 0$ et on trouve

$$\lim_n \uparrow \int f_n \, d\mu \geq \int h \, d\mu.$$

Cette inégalité est vraie pour toute fonction étagée h minorant f, donc on peut prendre le supremum sur toutes ces fonctions $h \leq f$, et on trouve bien le résultat. □

Exemple 3.9. Revenons à la fonction $f = \mathbb{1}_{\mathbb{Q} \cap I}$ de l'Exemple 1.6, qui n'admettait pas d'intégrale de Riemann. On avait montré que $f = \lim_n F_n$, avec $F_n = \mathbb{1}_{\frac{1}{n!}\mathbb{N} \cap I}$. La suite $(F_n)_{n \geq 1}$ est croissante, et les fonctions F_n sont étagées par rapport à la tribu de Borel, telles que $\int F_n \, d\lambda = 0$, puisque l'ensemble $\{x \in I, \, F_n(x) > 0\}$ est λ-négligeable. Le théorème ci-dessus montre donc que $\int f \, d\lambda = 0$. On obtient le même résultat en notant que l'ensemble $\{x \in I, \, f(x) > 0\} = \mathbb{Q} \cap I$ est dénombrable, donc il est borélien, et négligeable pour la mesure λ qui est diffuse.

Le théorème suivant montre qu'on peut approximer (au sens de la convergence simple, et de façon croissante) une fonction mesurable positive par des fonctions étagées.

Théorème 3.10. *[Approximation par des fonctions étagées] Soit f une fonction mesurable positive. Alors il existe une suite croissante $(f_n)_{n \in \mathbb{N}}$ de fonctions positives étagées, telles que* $\lim_n \uparrow f_n(x) = f(x)$ *pour tout $x \in E$.*

Démonstration. Notons que toute fonction étagée est nécessairement bornée, tandis que f peut valoir $+\infty$ sur un ensemble de E. On va construire les approximations f_n en tronquant leurs valeurs dans l'intervalle $[0, n]$, et en découpant cet intervalle en petits sous-intervalles

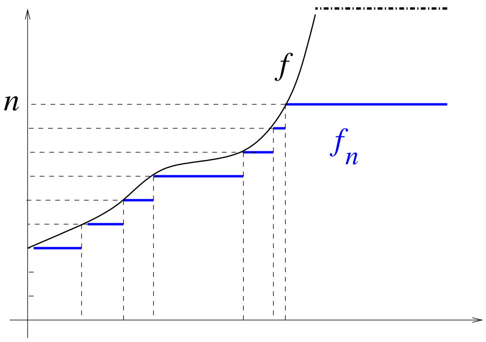

FIGURE 3.3 – Approximation d'une fonction mesurable positive par des fonctions étagées.

de longueur 2^{-n}. Par image réciproque, on définit donc les boréliens :

$$\begin{aligned} A_\infty^{(n)} &= f^{-1}\left([n,\infty[\right), \\ A_i^{(n)} &= f^{-1}\left(\left[\frac{i}{2^n}, \frac{i+1}{2^n}\right[\right), \; i = 0,\dots,n2^n-1, \end{aligned}$$

et la fonction étagée associée

$$f_n = \sum_{i=0}^{n2^n-1} \frac{i}{2^n}\,\mathbb{1}_{A_i^{(n)}} + n\mathbb{1}_{A_\infty^{(n)}}.$$

Cette fonction minore clairement f. De plus, on voit facilement que la partition en intervalles $\bigsqcup_i \left[\frac{i}{2^{n+1}}, \frac{i+1}{2^{n+1}}\right[\sqcup [n+1,\infty[$ est plus fine que la partition $\bigsqcup_i \left[\frac{i}{2^n}, \frac{i+1}{2^n}\right[\sqcup [n,\infty[$, ce qui implique que la fonction $f_{n+1} \geq f_n$.

Il reste à montrer que $f_n \nearrow f$. Sur $A^{(n)} := \bigsqcup_i A_i^{(n)} = f^{-1}\left([0,n[\right)$ on a l'encadrement $f_n(x) \leq f(x) \leq f_n(x) + \frac{1}{2^n}$. Les ensembles $A^{(n)}$ sont croissants, et leur union $\bigsqcup_n A^{(n)} = f^{-1}(\mathbb{R}_+)$, ce qui montre que pour tout $x \in f^{-1}(\mathbb{R}_+)$, $f_n(x) \nearrow f(x)$. Pour tout x sur l'ensemble restant $f^{-1}(\infty) = \bigcap_n A_\infty^{(n)}$, on aura $f_n(x) \geq n$, donc $f_n(x) \nearrow \infty = f(x)$. On a donc montré que $f_n \nearrow f$ sur tout E. □

Equipés de ce théorème d'approximation, on peut maintenant généraliser les propriétés de « linéarité positive » de l'intégrale des fonctions étagées.

Proposition 3.11. *1. Soient f, g des fonctions mesurables positives, et $a, b \geq 0$ deux réels positifs. Alors*

$$\int (af + bg)\, d\mu = a \int f\, d\mu + b \int g\, d\mu. \tag{3.2}$$

2. Si $(f_n)_{n\in\mathbb{N}}$ est une suite quelconque de fonctions mesurables positives, alors

$$\int \left(\sum_{n=0}^{\infty} f_n\right) d\mu = \sum_{n=0}^{\infty} \int f_n \, d\mu. \tag{3.3}$$

Autrement dit, pour toute suite de fonctions positives, on peut intervertir la somme et l'intégrale.

Attention : la formule (2) est valable car toutes les fonctions f_n sont positives. Elle ne sera pas vraie pour des fonctions f_n générales.

Démonstration. 1. Remarquons que la relation de linéarité (3.2) n'est pas complètement évidente à partir de la définition de l'intégrale. Grâce au théorème d'approximation, on peut approcher f (respectivement g) par une suite croissante de fonctions étagées $(f_n)_{n\in\mathbb{N}}$, et g par une suite similaire $(g_n)_{n\in\mathbb{N}}$. On aura alors automatiquement $af_n(x)+bg_n(x) \to af(x)+bg(x)$, et la suite $(af_n + bg_n)_n$ est croissante et faite de fonctions positives étagées. Le théorème de convergence monotone montre donc que

$$\int f_n \, d\mu \xrightarrow[n\to\infty]{} \int f \, d\mu, \quad \int g_n \, d\mu \xrightarrow[n\to\infty]{} \int g \, d\mu, \qquad \int (af_n + bg_n) \, d\mu \xrightarrow[n\to\infty]{} \int (af + bg) \, d\mu.$$

D'autre part, la Prop.3.3 montre que pour les fonctions étagées,

$$\int (af_n + bg_n) \, d\mu = a \int f_n d\mu + b \int g_n d\mu,$$

ce qui prouve le résultat.

2. La première partie (linéarité) montre l'égalité pour les sommes finies $F_N = \sum_{n=0}^{N} f_n$. Ensuite, il suffit de remarquer que $F_N \nearrow F_\infty = \sum_{n=0}^{\infty} f_n$, en particulier F_∞ est bien mesurable. Le théorème de convergence monotone montre donc que

$$\int F_\infty \, d\mu = \lim_N \uparrow \int F_N d\mu = \lim_N \uparrow \sum_{n=0}^{N} \int f_n d\mu = \sum_{n=0}^{\infty} \int f_n d\mu.$$

La somme de droite est composée de termes positifs, donc elle a bien une limite. □

Exercice 3.12. Retrouver le fait que, pour une suite double $(a_{n,k})_{n,k\in\mathbb{N}}$ de nombres positifs, on a l'égalité des sommes doubles

$$\sum_{k\in\mathbb{N}} \left(\sum_{n\in\mathbb{N}} a_{n,k}\right) = \sum_{n\in\mathbb{N}} \left(\sum_{k\in\mathbb{N}} a_{n,k}\right).$$

Une mesure μ peut être pondérée par une fonction positive f, pour former une nouvelle mesure $\nu = f \cdot \mu$.

Définition 3.13. [Mesure modulée par une fonction densité] Soit μ une mesure sur $(E, \mathcal{A})$ et f une fonction mesurable positive. Alors la mesure $\nu = f \cdot \mu$ est définie comme suit :

$$\forall A \in \mathcal{A}, \quad \nu(A) = \int \mathbb{1}_A f \, d\mu := \int_A f \, d\mu.$$

On peut interpréter la fonction f comme une « densité », qui vient moduler la mesure μ. Par exemple la mesure μ pourrait représenter le nombre de particules (atomes) par unité de volume, et f indiquer la masse de chaque atome, qui dépend de l'espèce chimique dont il s'agit.

Démonstration. Il faut tout de même vérifier, dans la définition ci-dessus, que ν définit bien une mesure sur la tribu $\mathcal{A}$. On a bien $\nu(\emptyset) = 0$. Si (A_n) forme une suite disjointe de parties mesurables, alors $\nu\left(\bigsqcup_n A_n\right) = \int \left(\sum_n \mathbb{1}_{A_n}\right) f \, d\mu$ peut se réécrire, grâce à la formule d'échange (3.3),

$$\nu\left(\bigsqcup_n A_n\right) = \sum_n \int \mathbb{1}_{A_n} f \, d\mu = \sum_n \nu(A_n).$$

□

Remarque 3.14. Notons que la mesure ν obtenue par cette modulation n'est pas quelconque : elle vérifier la propriété suivante :

$$\forall A \in \mathcal{A}, \qquad \text{si } \mu(A) = 0, \quad \text{alors forcément } \nu(A) = 0.$$

Autrement dit, un ensemble A négligeable pour la mesure μ sera aussi négligeable pour ν. Ceci reste vrai *si la fonction f prend la valeur $+\infty$ sur l'ensemble A*. (on rappelle la règle $0 \cdot \infty = 0$).

On a défini plus haut la notion d'ensemble négligeable pour une mesure μ, et inversement de propriété vraie μ-presque partout. Voici quelques inégalités très utiles, qui utilisent cette terminologie.

Proposition 3.15. *Soit f une fonction mesurable positive sur l'espace mesuré$(E, \mathcal{A}, \mu)$.*

1. Pour tout réel $a > 0$,

$$\mu\left(\{x \in E \,:\, f(x) \geq a\}\right) \leq \frac{1}{a} \int f \, d\mu. \tag{3.4}$$

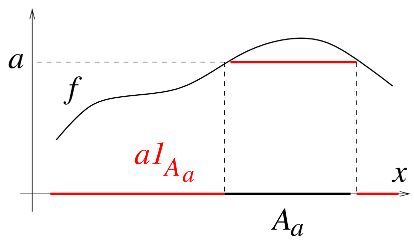

FIGURE 3.4 – Inégalité de Markov

*Cette inégalité est très utile en théorie des probabilités, où elle prend le nom d'*inégalité de Markov.

2. $\int f\,d\mu < \infty \Longrightarrow f < \infty \quad p.p.$ *(ou $\mu - p.p$).*

3. $\int f\,d\mu = 0 \Longleftrightarrow f = 0 \quad p.p.$

4. *Si g est une autre fonction mesurable positive, alors* $g = f \quad p.p. \Longrightarrow \int f\,d\mu = \int g\,d\mu$.

Cette dernière propriété montre que, dans le cadre de l'intégration, il n'est pas nécessaire de connaître une fonction en tout point (« partout ») pour déterminer ses propriétés d'intégration, mais uniquement « presque partout ». Ceci nous amènera à définir les classes d'équivalences de fonctions égales entre elles presque partout.

Démonstration. 1. Notons $A_a = f^{-1}([a, \infty]) = \{x \in E\,;\, f(x) \geq a\}$. Comme f est mesurable, A_a l'est aussi. On remarque l'inégalité évidente (« tautologie ») $f \geq a\mathbb{1}_{A_a}$. La positivité de l'intégrale implique alors

$$\int f\,d\mu \geq \int a\mathbb{1}_{A_a}\,d\mu = a\mu(A_a).$$

2. Notons, pour tout $n \geq 1$, $A_n = f^{-1}([n, \infty])$, et $A_\infty = f^{-1}(\{\infty\})$. On remarque que la suite $(A_n)_n$ est décroissante, et que $\bigcap_n A_n = A_\infty$, de sorte que

$$\mu(A_\infty) = \mu\left(\bigcap_n A_n\right) = \lim_n \downarrow \mu(A_n).$$

D'après l'inégalité (3.4), $\mu(A_n) \leq \frac{1}{n}\int f\,d\mu$, donc si f est d'intégrale finie, on aura $\lim_n \downarrow \mu(A_n) = 0$.

3. On a déjà montré l'implication $\Longleftarrow$ dans la Proposition 3.5. Pour montrer $\Longrightarrow$, on va utiliser les ensembles $B_n := f^{-1}\left(\left[\frac{1}{n}, \infty\right]\right)$. Ils forment une suite croissante, et leur union $\bigcup_{n\in\mathbb{N}} B_n = f^{-1}(]0, \infty])$. L'inégalité de Markov (3.4) implique $\mu(B_n) \leq n\int f\,d\mu$ pour tout

$n \geq 1$, donc d'après l'hypothèse, $\mu(B_n) = 0$. Comme Par sous-additivité de la mesure, $\mu\left(f^{-1}(]0,\infty])\right) = 0$, ce qui revient à dire que f est nulle presque partout.

4. Remarquons qu'on ne sait pas encore intégrer des fonctions changeant de signe, donc on ne peut pas directement essayer de calculer $\int (f-g) d\mu$. On va plutôt considérer les fonctions mesurables $f \vee g := \max(f,g)$ et $f \wedge g := \min(f,g)$. En effet, ces fonctions satisfont $f \vee g \geq f \wedge g$, de sorte que $f \vee g - f \wedge g$ est une fonction mesurable positive. D'autre part, l'hypothèse $f = g$ p.p. implique que $(f \vee g - f \wedge g) = 0$ p.p. Ceci implique que

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

Enfin, on a les inégalités $f \wedge g \leq f \leq f \vee g$ et de même en remplaçant f par g. Ces encadrements, avec l'égalité ci-dessus, montrent donc que

$$\int f \, d\mu = \int f \vee g \, d\mu \quad \text{et} \quad \int g \, d\mu = \int f \vee g \, d\mu,$$

donc ces deux intégrales sont égales. □

Le théorème de convergence monotone concernait exclusivement des suites croissantes de fonctions. Le second résultat important, le Lemme [1] de Fatou, supprime cette hypothèse de monotonie..

Théorème 3.16. *[Lemme de Fatou] Soit $(E, \mathcal{A}, \mu)$ un espace mesuré, et $(f_n)_{n\in\mathbb{N}}$ une suite quelconque de fonctions mesurables positives. Alors*

$$\int \left(\liminf_{n\to\infty} f_n \right) d\mu \leq \liminf_{n\to\infty} \int f_n \, d\mu.$$

On remarque que, d'après le théorème de convergence monotone, si la suite $(f_n)_n$ est croissante, alors $\liminf_n f_n = \lim_n f_n$, et on a égalité dans l'expression ci-dessus.

Démonstration. Par définition, $\liminf_n f_n = \lim_n \uparrow (\inf_{k\geq n} f_k)$. Les fonctions $F_n := (\inf_{k\geq n} f_k)$ sont mesurables et forment une suite croissante, donc d'après le théorème de convergence monotone,

$$\lim_n \uparrow \int F_n \, d\mu = \int \left(\lim_n \uparrow F_n \right) d\mu = \int \left(\liminf_n f_n \right) d\mu.$$

1. En général, un Lemme est un résultat intermédiaire, qui n'a pas d'intérêt en soi. Il se trouve que certains Lemmes sont devenus célèbres, car ils ont eu des conséquences inattendues sur le reste de la théorie. On devrait donc plutôt parler de « Théorème de Fatou », mais on a gardé la dénomination historique de « Lemme ».

FIGURE 3.5 – Illustration du Lemme de Fatou

D'autre part, on a forcément $F_n \leq f_n$, donc $\int F_n \, d\mu \leq \int f_n \, d\mu$; en passant à la $\liminf_n$ on obtient

$$\liminf_n \int F_n \, d\mu \leq \liminf_n \int f_n \, d\mu.$$

Enfin, on se souvient que la suite $\left(\int F_n \, d\mu\right)_n$ est croissante, de sorte que $\liminf_n \int F_n \, d\mu = \lim_n \int F_n \, d\mu$. En combinant avec l'inégalité et légalité ci-dessus, on obtient le résultat. $\square$

Exercice 3.17. 1. On se donne les fonctions $f_n := \mathbb{1}_{[n,n+1]}$ sur l'espace mesuré $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$. Calculer les deux côtés de l'inégalité dans le Lemme de Fatou, et vérifier qu'on n'a pas égalité. Dans ce cas, l'inégalité stricte est due au fait que le support de f_n « part à l'infini » lorsque $n \to \infty$.

2. On définit $f_n = f_0$ si $n$est pair, $f_n = f_1$ si n est impair, avec $f_0 = \mathbb{1}_{[0,1/2\}}$ et $f_1 = \mathbb{1}_{[1/2,0]}$. Calculer alors les deux côtés du Lemme de Fatou, et vérifier qu'on a inégalité stricte. Ici, l'inégalité stricte est due à l'« oscillation » du support de f_n.

3. On définit $f_n = n\mathbb{1}_{[0,1/n]}$. Montrer qu'on a encore une inégalité stricte dans le Lemme de Fatou.

3.2 Fonctions intégrables

Jusqu'à présent on a considéré des fonctions mesurables positives, qui pouvaient prendre la valeur $+\infty$ sur n'importe quel sous-ensemble (mesurable), et dont l'intégrale pouvait prendre la valeur $+\infty$. Afin d'étendre l'intégration aux fonctions prenant des valeurs négatives, ou aux fonctions changeant de signe, on va être obligé d'imposer une condition d'intégrabilité ; en effet, on doit éviter de se retrouver face à des expressions du type $+\infty - (+\infty)$.

3.2.1 Intégrale d'une fonction réelle intégrable

Définition 3.18. Soit $(E, \mathcal{A}, \mu)$ un espace mesuré, et soit $f : (E, \mathcal{A}) \to \mathbb{R}$ une fonction mesurable (pas forcément positive). On dit que f est *intégrable par rapport à* μ (ou μ-

intégrable, ou intégrable si la mesure est sous-entendue) si

$$\int |f| \, d\mu < \infty.$$

Dans ce cas, on définit l'intégrale de f par

$$\int f \, d\mu := \int f^+ d\mu - \int f^- d\mu, \tag{3.5}$$

où $f^+ = \max(f, 0)$ et $f_- = \max(-f, 0)$ sont appelées respectivement la partie positive et la partie négative de f. On notera que f^+ et f^- sont mesurables, et que $f = f^+ - f^-$, respectivement $|f| = f^+ + f^-$ (en particulier $|f|$ est bien mesurable et positif).

On note $\mathcal{L}^1(E, \mathcal{A}, \mu)$ l'espace des fonctions intégrables. Notons que cet espace dépend de la mesure μ, et pas uniquement de la tribu $\mathcal{A}$. On notera $\mathcal{L}^1_+(E, \mathcal{A}, \mu)$ l'espace des fonctions intégrables à valeurs positives.

Remarque 3.19. 1. On a forcément $f^+ \leq |f|$ et $f^- \leq |f|$, donc l'hypothèse d'intégrabilité impose que $\int f^+ d\mu$ et $\int f^- d\mu$ sont finies, de sorte que la différence (3.5) a bien un sens, et donne un nombre fini.

2. Si f est positive, on retrouve la définition précédente.

3. On aurait pu, dans la définition ci-dessus, considérer des fonctions f mesurables à valeurs dans $\overline{\mathbb{R}}$. Cependant, l'intégrabilité de f impose alors que les ensembles $f^{-1}(\{+\infty\})$ et $f^{-1}(\{-\infty\})$ sont de mesure nulle. Donc on peut modifier la valeur de f sur ces ensembles négligeables (par exemple poser $f(x) = 0$ sur ces ensembles), sans changer la valeur de l'intégrale de f. Après ce changement, f prend ses valeurs dans $\mathbb{R}$.

On va montrer que la plupart des propriétés satisfaites par l'intégrale des fonctions positives s'étendent aux fonctions intégrables.

Proposition 3.20. *a) Pour tout $f \in \mathcal{L}^1(E, \mathcal{A}, \mu)$, on a l'inégalité $\left|\int f \, d\mu\right| \leq \int |f| \, d\mu$. On qualifie cette inégalité de « triangulaire » : en effet, on la voit comme une généralisation de $|\sum_i a_i| \leq \sum_i |a_i|$, pour une famille $(a_1, \ldots, a_N)$ de nombre réels.*

b) $\mathcal{L}^1(E, \mathcal{A}, \mu)$ est un espace vectoriel, et l'application $f \mapsto \int f \, d\mu$ est une forme linéaire sur cet espace vectoriel. Autrement dit, la formule (3.2) s'étend aux fonctions $f, g \in \mathcal{L}^1(E, \mathcal{A}, \mu)$ et aux réels $a, b \in \mathbb{R}$ quelconques.

c) Si $f, g \in \mathcal{L}^1(E, \mathcal{A}, \mu)$ et $f \leq g$, alors $\int f \, d\mu \leq \int g \, d\mu$. On dit parfois que l'intégrale est une forme linéaire positive.

d) Si $f, g \in \mathcal{L}^1(E, \mathcal{A}, \mu)$ et $f = g$ p.p., alors $\int f\, d\mu = \int g\, d\mu$.

Démonstration. a) $\left|\int f\right| = \left|\int f^+ - \int f^-\right| \leq \left|\int f^+\right| + \left|\int f^-\right| = \int (f^+ + f^-) = \int |f|$ (on a omis d'indiquer les $d\mu$ dans les intégrales). Ici on s'est servi de l'inégalité triangulaire.

b) Bien que la linéarité semble évidente, il faut un peu « ruser » pour le montrer, car on n'a à notre disposition uniquement des résultats pour les fonctions positives. Il faut donc se ramener à des fonctions positives.

Pour $f \in \mathcal{L}^1(E, \mathcal{A}, \mu)$ et $a \in \mathbb{R}$ on a $\int |af| d\mu = |a| \int |f| d\mu < \infty$, donc $af \in \mathcal{L}^1$. Pour montrer la linéarité de l'intégrale par multiplication par le réel a, on sépare les cas $a \geq 0$ et $a \leq 0$:

- si $a \geq 0$, alors $af = (af)^+ - (af)^- = af^+ - af^-$, donc $\int (af)\, d\mu = \int af^+\, d\mu - \int af^- d\mu$. En se servant de (3.2), cette expression vaut $a \left(\int f^+ - \int f^-\right) = a \int f\, d\mu$.

- si $a < 0$, on a $(af)^+ = (-a)f^-$ et $(af)^- = (-a)f^+$. On trouve alors, par les mêmes manipulations, $\int (af)\, d\mu = (-a) \left(\int f^- - \int f^+\right) = (-a) \left(-\int f\right) = a \int f\, d\mu$.

Si f et g sont intégrables, l'inégalité triangulaire ponctuelle $|f(x) + g(x)| \leq |f(x)| + |g(x)|$ implique que $(f+g)$ est aussi intégrable. Pour montrer que $\int (f+g) = \int f + \int g$, il faut décomposer $(f+g)$ en parties positive et négative. A partir de l'identité :

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

on obtient l'identité entre fonctions positives

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

On prend l'intégrale de cette identité :

$$\int \left[(f+g)^+ + f^- + g^-\right] d\mu = \int \left[(f+g)^- + f^+ + g^+\right] d\mu,$$

et on utilise la linéarité de l'intégrale pour les fonctions positives, de façon à obtenir :

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

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

c) Supposons f, g intégrables et $f \leq g$. La fonction $(g-f)$ est intégrable et positive, donc

$0 \leq \int (g-f) d\mu < \infty$. La linéarité montrée en *b)* permet alors d'écrire

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

d) L'égalité presque sure $f = g$ p.p. implique que $f^+ = g^+$ p.p. et $f^- = g^-$ p.p. A partir du point (4) de la Proposition 3.15, on en déduit que $\int f^+ d\mu = \int g^+ d\mu$ et $\int f^- d\mu = \int g^- d\mu$, et finalement $\int f \, d\mu = \int g \, d\mu$. $\square$

Remarque 3.21. En se servant du *d)*, on peut affaiblir l'hypothèse du *c)* en l'hypothèse : « Si $f, g \in \mathcal{L}^1$ et si $f \leq g$ p.p. ». Comme expliqué plus haut, modifier les valeurs d'une fonction sur un ensemble négligeable ne va pas modifier son intégrale.

3.2.2 Fonctions à valeurs complexes

Pour l'instant on a supposé que les fonctions f prennent des valeurs réelles (ou dans $\overline{\mathbb{R}}$). On peut étendre la théorie aux fonctions à valeurs dans $\mathbb{C}$, puis dans $\mathbb{R}^d$ pour $d \geq 1$ quelconque.

On rappelle qu'une fonction $f : E \to \mathbb{C}$ est mesurable si, pour tout borélien $B \subset \mathbb{C}$, $f^{-1}(B)$ est dans la tribu $\mathcal{A}$. Comme $\mathbb{C} \equiv \mathbb{R}^2$, la Proprosition 2.28 montre que f est mesurable si et seulement si les fonctions $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ sont mesurables.

Définition 3.22. Soit $(E, \mathcal{A}, \mu)$ un espace mesuré. Une fonction $f : E \to \mathbb{C}$ mesurable est dite intégrable si la fonction $|f| : E \to \mathbb{R}_+$ est intégrable. On note alors $f \in \mathcal{L}^1_{\mathbb{C}}(E, \mathcal{A}, \mu)$. On définit alors l'intégrale de f par

$$\int f \, d\mu = \int \mathrm{Re}(f) \, d\mu + i \int \mathrm{Im}(f) \, d\mu \in \mathbb{C}. \tag{3.6}$$

Comme $|\mathrm{Re}(f)| \leq |f|$, $|\mathrm{Im}(f)| \leq |f|$ et $|f| \leq |\mathrm{Re}(f)| + |\mathrm{Re}(f)|$, l'intégrabilité de f est équivalente à l'intégrabilité simultanée de ses parties réelle et imaginaire.

Proposition 3.23. *Les propriétés a), b), d) de la Proposition 3.20 s'étendent aux fonctions intégrables à valeurs complexes.*

Démonstration. Les propriétés b) et d) sont évidentes à étendre, à partir de l'expression (3.6).

Pour montrer l'inégalité triangulaire *a)* il faut travailler un peu plus. On utilise la propriété suivante du module d'un nombre complexe :

$$\forall z \in \mathbb{C}, \quad |z| = \max_{u \in \mathbb{C}, |u|=1} \mathrm{Re}(uz). \tag{3.7}$$

On aura donc
$$\left|\int f\,d\mu\right| = \max_{u\in\mathbb{C},|u|=1} \mathrm{Re}(u\int f\,d\mu).$$
En décomposant $u = \mathrm{Re}(u) + i\,\mathrm{Im}(u)$, et en prenant la définition de l'intégrale de f, on obtient
$$\left|\int f\,d\mu\right| = \max_{u\in\mathbb{C},|u|=1} \left(\mathrm{Re}(u)\int \mathrm{Re}(f)\,d\mu - \mathrm{Im}(u)\int \mathrm{Im}(f)\,d\mu\right).$$
Par linéarité de l'intégrale des fonctions réelles, cela donne
$$\left|\int f\,d\mu\right| = \max_{u\in\mathbb{C},|u|=1} \left(\int [\mathrm{Re}(u)\,\mathrm{Re}(f) - \mathrm{Im}(u)\,\mathrm{Im}(f)]\;d\mu\right) = \max_{u\in\mathbb{C},|u|=1} \left(\int \mathrm{Re}(uf)\,d\mu\right).$$
Finalement, on a pour tout $u \in \mathbb{C}$ de norme 1, l'inégalité entre fonctions intégrables $\mathrm{Re}(uf) \leq |f|$, donc $\int \mathrm{Re}(uf)\,d\mu \leq \int |f|\,d\mu$. □

Exercice 3.24. Définir la notion d'intégrabilité pour une fonction $f : E \to \mathbb{R}^d$, définir son intégrale en termes de ses composantes $(f_j)_{j=1,\ldots,d}$, et montrer la généralisation des propriétés *a), c), d)* de la proposition 3.20 à ce contexte.

3.2.3 Le théorème de convergence dominée

Le troisième résultat de convergence de la théorie de l'intégration de Lebesgue (probablement le plus important des trois théorèmes présentés dans ce chapitre) est le théorème de convergence dominée.

Théorème 3.25. *[Théorème de convergence dominée] Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions dans $\mathcal{L}^1(E,\mathcal{A},\mu)$ (ou $\mathcal{L}^1_{\mathbb{C}}(E,\mathcal{A},\mu)$). On fait les hypothèses suivantes :*

(1) il existe une fonction f mesurable à valeurs dans $\mathbb{R}$ (ou dans $\mathbb{C}$) telle que
$$f_n(x) \xrightarrow[n\to\infty]{} f(x) \quad \mu - p.p.$$
(2) il existe une fonction $g : E \to \mathbb{R}_+$ mesurable et intégrable ($\int g\,d\mu < \infty$) telle que
$$\forall n \in \mathbb{N}, \quad |f_n(x)| \leq g(x) \quad \mu - p.p.$$
Alors la fonction $f \in \mathcal{L}^1(E,\mathcal{A},\mu)$ (ou $f \in \mathcal{L}^1_{\mathbb{C}}(E,\mathcal{A},\mu)$), et on a les convergences
$$\int f_n\,d\mu \xrightarrow[n\to\infty]{} \int f\,d\mu \tag{3.8}$$

$$\text{et} \qquad \int |f - f_n| \, d\mu \xrightarrow[n\to\infty]{} 0. \tag{3.9}$$

La dernière ligne est souvent appelée « convergence $\mathcal{L}^1$ », ou plus exactement « convergence L^1 ». Le terme « convergence dominée » provient du fait que les fonctions f_n sont toutes « dominées » (bornées) par une même fonction $g \in \mathcal{L}^1$.

Démonstration. Dans un premier temps, on va supposer que les deux hypothèses presques sures sont vraies partout :

(1)' $\lim_{n\to\infty} f_n(x) = f(x)$ pour tout $x \in E$,

(2)' pour tout $n \geq 0$ et pour tout $x \in E$, $|f_n(x)| \leq g(x)$.

Ces propriétés impliquent que $|f| \leq g$, donc que la fonction limite f est intégrable (on savait déjà que f était mesurable).

Pour montrer la limite (3.9) on va appliquer le Lemme de Fatou à une suite de fonctions bien choisie. Les propriétés $|f_n| \leq g$ et $|f| \leq g$ impliquent que $|f - f_n| \leq 2g$, donc que la fonction $(2g - |f - f_n|)$ est positive. Cette fonction converge (simplement) vers $2g$, le Lemme de Fatou nous donne donc :

$$\liminf_n \int (2g - |f - f_n|) \, d\mu \geq \int \liminf_n (2g - |f - f_n|) \, d\mu = \int 2g \, d\mu.$$

Ceci peut se réécrire, en utilisant la linéarité de l'intégrale à gauche :

$$\int 2g \, d\mu - \limsup_n \int |f - f_n| \, d\mu \geq \int 2g \, d\mu,$$

donc $\limsup_n \int |f - f_n| \, d\mu \leq 0$. Comme ces intégrales sont positives, on a donc montré $\lim_n \int |f - f_n| \, d\mu = 0$.

A partir de cette limite, l'inégalité triangulaire donne

$$\left| \int f \, d\mu - \int f_n \, d\mu \right| \leq \int |f - f_n| \, d\mu \xrightarrow[n\to\infty]{} 0,$$

ce qui montre la convergence (3.8). On a donc montré le théorème sous les hypothèses plus fortes (1)' et (2)'.

On fait maintenant les hypothèses (1) et (2). On définit les « bons ensembles » $B_n = \{x \in E\,;\, f_n(x) \leq g(x)\}$ et $B_{lim} = \{x \in E\,;\, \lim_n f_n(x) = f(x)\}$. Les hypothèses disent que leurs complémentaires B_n^c et B_{lim}^c sont négligeables, autrement dit ils sont inclus respectivement dans des ensembles N_n et N_{lim} mesurables de mesure nulle (on peut montrer que B_n et B_{lim} sont mesurables, et donc prendre $N_n = B_n^c$, $N_{lim} = B_{lim}^c$, mais on ne s'en servira pas ici). Le « bon ensemble global » $B = B_{lim} \cap (\bigcap_n B_n)$, celui des points x satisfaisant toutes les bonnes propriétés, a pour complémentaire

$$B^c = B_{lim}^c \cup \left(\bigcup_n B_n^c \right) \subset N_{lim} \cup \left(\bigcup_n N_n \right) := N.$$

L'ensemble de droite est de mesure nulle, donc B^c est négligeable. On fabrique alors les fonctions mesurables

$$\tilde{f}_n = f_n \, \mathbb{1}_{N^c}, \quad \tilde{f} = f \, \mathbb{1}_{N^c}.$$

Comme $N^c \subset B$, ces fonctions satisfont les propriétés (1)' et (2)', donc elles vérifient la conclusion du théorème. On remarque finalement que, comme N est de mesure nulle, on $f_n = \tilde{f}_n$ p.p. et $f = \tilde{f}$ p.p., donc on peut remplacer dans les intégrales les fonctions $\tilde{f}_n, \tilde{f}$ par f_n, f sans changer les valeurs des intégrales. $\square$

Exercice 3.26. Discuter les trois exemples de l'Exercice 3.17 : les fonctions $(f_n)_n$ convergent-elles simplement ? Si c'est le cas, sont-elles dominées par une fonction g intégrable ?

3.3 Intégrales dépendant d'un paramètre

On considère à présent la situation où la fonction f dépend de deux variables, qui ont des rôles différents : une variable $x \in E$, sur la quelle on va intégrer ; et une variable u appartenant à un « espace des paramètres » U, sur laquelle on ne va pas intégrer.

Le cas d'une suite de fonctions $(f_n)_{n\in\mathbb{N}}$ est déjà dans ce cadre : ici le paramètre u est l'entier $n \in \mathbb{N}$. On va généraliser au cas où l'espace des paramètres U est un espace métrique quelconque, par exemple un ouvert de l'espace euclidien $\mathbb{R}^d$. On se pose alors naturellement la question de la variation de l'intégrale $\int f(x,u)\, d\mu(x)$ par rapport au paramètre u : sous quelles conditions cette intégrale va-t-elle varier continûment par rapport à u ? Plus bas, on s'intéressera aussi à la dérivabilité de l'intégrale par rapport au paramètre $u \in \mathbb{R}^d$. Les résultats de cette section sont des corollaires du TCD.

3.3.1 Continuité d'une intégrale à paramètre

Notre espace mesuré E n'a pas forcément de structure topologique, alors que l'espace métrique en a une. Il s'agit donc de comprendre les relations entre la structure mesurable (par rapport à la variable x) et la structure topologique (continuité, par rapport à la variable u). Ce mélange de deux structures rend les hypothèses dans les théorèmes ci-dessous un peu délicates, les variables u, x ne jouant pas les mêmes rôles.

Théorème 3.27. *Soit $(E, \mathcal{A}, \mu)$ un espace mesuré, et (U, d) un espace métrique. On considère une application $f : U \times E \to \mathbb{R}$, et on fixe un point $u_0 \in U$. On fait les hypothèses suivantes :*

i) pour tout paramètre $u \in U$, l'application $x \mapsto f(u, x)$ est mesurable ;

ii) pour μ-presque tout $x \in E$, l'application $u \mapsto f(u, x)$ est continue en u_0 ;

iii) il existe une fonction $g \in \mathcal{L}^1_+(E, \mathcal{A}, \mu)$ telle que, pour tout $u \in U$, $|f(u, x)| \leq g(x)$ $\mu(dx)$-p.p.

Alors la fonction $u \mapsto \int f(u, x)\, d\mu(x)$ est bien définie en tout point $u \in U$, et elle est continue en u_0.

Remarque 3.28. [Comparaison avec le théorème de convergence dominée (TCD)] Ce théorème généralise le TCD au cas d'un espace des paramètre continu. En effet, en prenant $U = \mathbb{N}$ et sa topologie discrète, on retrouve exactement le TCD (la condition *ii.* de continuité en u_0 est l'analogue de la condition de convergence $\lim_{n\to\infty} f_n = f$). La nouveauté ici est d'autoriser des familles de fonctions indicées par un ensemble continu.

Démonstration. Les hypothèses i) et iii) impliquent que pour tout $u \in U$, la fonctions $f(u, \cdot)$ est intégrable, de sorte que son intégrale $\int f(u, x)\, d\mu(x)$ est bien définie. Si on choisit une suite $(u_n)_{n\geq 1}$ une suite de points de U convergeant vers u_0, alors les hypothèses ci-dessus montrent que les fonctions $(f(u_n, \cdot))_{n\geq 1}$ satisfont les hypothèses du théorème 3.25 de convergence dominée, avec comme fonction limite $f(u_0, \cdot)$. Le TCD montre donc que les intégrales$\int f(u_n, x)\, d\mu(x)$ convergent vers $\int f(u_0, x)\, d\mu(x)$ lorsque $n \to \infty$. □

Exemple 3.29. [μ-primitive d'une fonction intégrable] On se place sur l'espace mesuré $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$ avec μ une mesure diffuse (sans atomes) ; on prend $U = \mathbb{R}$ avec sa structure métrique usuelle. Si $\varphi \in \mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$, on considère l'intégrale

$$F(u) := \int_{]-\infty,u]} \varphi(x)\, \mu(dx) = \int_{-\infty}^{u} \varphi(x)\, d\mu(x) = \int \varphi \mathbb{1}_{]-\infty,u]} d\mu$$

(les trois écritures sont équivalentes). On montre alors que la fonction $u \mapsto F(u)$ est continue. Attention : cette expression intégrale ressemble à celle de la primitive d'une fonction φ intégrable continue, $\int_{-\infty}^{u} \varphi(x)\,dx = \int_{-\infty}^{u} \varphi(x)\,\lambda(dx)$. Mais dans notre cas, avec une mesure diffuse μ quelconque, la fonction F n'est pas forcément dérivable, même lorsque φ est continue intégrable.

Montrons la continuité de F. On considère ici la famille de fonctions $f(u,x) = \varphi(x)\mathbb{1}_{]-\infty,u]}(x)$. Pour tout $u \in U$, la fonction $f(u,\cdot)$ (autrement dit $x \mapsto f(u,x)$) est borélienne par rapport à x, et elle est dominée par la fonction $g = |\varphi| \in \mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$; on a donc bien les hypothèses *i)* et *iii)* du théorème. Pour vérifier l'hypothèse *ii)*, il faut remarquer que, la fonction $u \mapsto f(u,x)$ est constante sur $u \in]-\infty, x[$ et sur $u \in]x, \infty[$, donc son seul point de discontinuité est en $u = x$. Donc si on se fixe un point $u_0 \in \mathbb{R}$, alors $u \mapsto f(u,x)$ est continue en $u = u_0$, sauf si $x = u_0$. L'ensemble $\{u_0\} \subset \mathbb{R}$ est négligeable par rapport à la mesure μ, puisque celle-ci est diffuse ; la fonction $u \mapsto f(u,x)$ est donc bien continue en $u = u_0$ pour μ-presque tout point x, ce qui est l'hypothèse *ii)*.

On remarque facilement que ce résultat est faux en général si la mesure μ admet des atomes. Par exemple dans le cas où $\mu = \delta_{u_0}$ la mesure de Dirac en u_0, on vérifie que $F(u) = 0$ pour $u < u_0$, tandis que $F(u) = \varphi(u_0)$ pour $u \geq u_0$.

Exemple 3.30. *a)* [Transformée de Fourier] Si $\varphi \in \mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, la fonction $\hat{\varphi}(u) := \int e^{-iux}\varphi(x)\,\lambda(dx)$ est appelée la transformée de Fourier de φ. Cette fonction est continue et bornée sur $\mathbb{R}$ entier.

b) [Convolution] soit $\varphi \in \mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ et soit $h : \mathbb{R} \to \mathbb{R}$ une fonction continue et bornée. Alors la fonction

$$u \mapsto h * \varphi(u) = \int h(u-x)\varphi(x)\,\lambda(dx)$$

est continue et bornée sur $\mathbb{R}$ entier.

3.3.2 Dérivabilité d'une intégrale à paramètre

Pour pouvoir dériver par rapport à $u \in U$, il faut que U ait une structure euclidienne. On prend ici $U = I$ un intervalle ouvert de $\mathbb{R}$.

Théorème 3.31. *Soit $(E, \mathcal{A}, \mu)$ un espace mesuré, et $I \subset \mathbb{R}$ un intervalle ouvert. On considère une application $f : I \times E \to \mathbb{R}$, et on fixe un point $u_0 \in I$. Supposons que :*

i) pour tout $u \in I$, l'application $x \mapsto f(u,x)$ est dans $\mathcal{L}^1(E, \mathcal{A}, \mu)$;

ii) Pour μ-presque tout x, l'application $u \mapsto f(u,x)$ est dérivable en u_0, de dérivée notée $\frac{\partial f}{\partial u}(u_0,x)$;

iii) il existe $g \in \mathcal{L}^1_+(E,\mathcal{A},\mu)$ telle que, pour tout $u \in I$,

$$|f(u,x) - f(u_0,x)| \leq g(x)\,|u-u_0| \qquad \mu(dx) - p.p.$$

Alors la fonction $F(u) = \int f(u,x)\,\mu(dx)$ est continue et dérivable en u_0, et sa dérivée en ce point vaut

$$F'(u_0) = \int \frac{\partial f}{\partial u}(u_0,x)\,\mu(dx).$$

On remarque que la borne uniform $g(x)$ ne s'applique plus directement à la fonction $f(x)$, mais au *taux de variation* de f entre u_0 et u, $\frac{f(u,x)-f(u_0,x)}{u-u_0}$. La dérivée $\frac{\partial f}{\partial u}(u_0,x)$ n'est a priori définie que sur le complémentaire d'un ensemble négligeable, mais on peut la compléter de façon arbitraire en dehors de cet ensemble, de façon à obtenir une fonction définie partout. Il n'est pas évident que cette fonction soit mesurable, la preuve montrera qu'elle est égale p.p. à une fonction mesurable (et intégrable).

Démonstration. Considérons une suite $(u_n)_{n\in\mathbb{N}}$ dans $I \setminus \{x_0\}$, convergeant vers u_0, et définissons les fonctions $\varphi_n(x) = \frac{f(u_n,x)-f(u_0,x)}{u_n-u_0}$. Comme les fonctions $f(u,\cdot)$ sont toutes mesurables et μ-intégrables, les fonctions φ_n le sont également. De plus, l'hypothèse iii) montrer que μ-presque partout, les fonctions $|\varphi_n|$ sont bornées par la fonction intégrable g. On peut donc définir

$$\varphi_\infty(x) = \limsup_{n\to\infty} \varphi_n(x),$$

une fonction mesurable qui vérifie

$$|\varphi_\infty(x)| \leq g(x) \qquad \mu - p.p.,$$

donc $\varphi_\infty \in \mathcal{L}^1(E,\mathcal{A},\mu)$. D'autre part l'hypothèse ii) montre que

$$\text{pour } \mu\text{-presque tout } x, \qquad \varphi_\infty(x) = \lim_{n\to\infty} \varphi_n(x) = \frac{\partial f}{\partial u}(u_0,x).$$

Le théorème de convergence dominée montre alors que

$$\begin{aligned}\int \varphi_\infty(x)\,d\mu(x) &= \lim_{n\to\infty} \int \varphi_n(x)\,d\mu(x) = \lim_{n\to\infty} \int \left(\frac{f(u_n,x)-f(u_0,x)}{u_n-u_0}\right) d\mu(x) \\ &= \lim_{n\to\infty} \frac{1}{u-u_0}\left(F(u_n) - F(u_0)\right).\end{aligned}$$

Comme $\varphi_\infty(x)$ est égal à $\frac{\partial f}{\partial u}(u_0, x)$ pour presque tout x, l'intégrale du membre de gauche peut s'écrire $\int \frac{\partial f}{\partial u}(u_0, x) d\mu(x)$. Comme ce membre de gauche est indépendant de la suite (u_n) convergeant vers u_0, l'égalité ci-dessus montre que la fonction F est dérivable en u_0, de dérivée $F'(u_0) = \int \varphi_\infty(x)\, d\mu(x)$. □

Remarque 3.32. Très souvent, les hypothèses *ii)* et *iii)* sont remplacées par des hypothèses plus fortes :

ii)' pour μ-presque tout x, l'application $u \mapsto f(x, u)$ est dérivable sur I ;

iii)' il existe $g \in \mathcal{L}^1_+(E, \mathcal{A}, \mu)$ telle que pour μ-presque tout x,

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

Le théorème des accroissements finis montre que *iii)'* implique *iii)*. Ces hypothèses renforcées impliquent que F est dérivable sur I.

Exercice 3.33. Reprenons les exemples des exercices 3.30 et 3.29.

a) Supposons que $\varphi \in \mathcal{L}^1(R, \mathcal{B}(\mathbb{R}), \lambda)$ est tel que la fonction $x \mapsto x\varphi(x)$ est également dans $\mathcal{L}^1(R, \mathcal{B}(\mathbb{R}), \lambda)$. Montrer alors que la transformée de Fourier $\hat{\varphi}(u)$ est dérivable sur $\mathbb{R}$, et qu'elle vaut

$$\hat{\varphi}(u) = -i \int x e^{-iux} \varphi(x)\, \lambda(dx).$$

b) Soit $\varphi \in \mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ et soit $h : \mathbb{R} \to \mathbb{R}$ une fonction de classe C^1, bornée ainsi que sa dérivée. Alors montrer que la convolution $h * \varphi$ est dérivable sur $\mathbb{R}$, et de dérivée $(h * \varphi)' = h' * \varphi$.

c) Soit μ une mesure diffuse sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ et soit $\varphi \in \mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$ telle que $x \mapsto x\varphi(x)$ est aussi dans $\mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$. Pour tout $u \in \mathbb{R}$, on pose

$$F(u) := \int (u - x)^+ \varphi(x)\, \mu(dx).$$

Montrer que F est dérivable sur $\mathbb{R}$, de dérivée $F'(u) = \int_{]-\infty, u]} \varphi(x)\, \mu(dx)$.

4 Construction et propriétés de la mesure de Lebesgue

Nous avons pour l'instant évoqué la mesure de Lebesgue sur $\mathbb{R}$ comme étant la mesure « naturelle », celle qui associe à un intervalle sa longueur. Nous avons aussi affirmé qu'elle est bien définie sur toute la tribu des boréliens. Nous avons enfin montré qu'il ne peut exister au plus qu'une telle mesure (grâce au lemme de classe monotone). Dans ce chapitre nous allons construire la mesure de Lebesgue, satisfaisant toutes ces propriétés. On montrera, en particulier, qu'elle permet de faire le lien avec l'intégrale de Riemann sur $\mathbb{R}$, c'est-à-dire que l'intégrale de Lebesgue d'une fonction réglée est égale à l'intégrale de Riemann correspondante.

Nous nous focaliserons sur la construction de la mesure de Lebesgue, mais notre stratégie de construction pourra en fait s'appliquer à bien d'autres mesures. Cette stratégie consiste à construire tout d'abord un objet moins contraint qu'une mesure, qu'on appelle une *mesure extérieure* (ou mesure σ-sous-additive) ; cet assouplissement des contraintes permet à cette mesure extérieure d'être sur toutes les parties de E. On introduira aussi la notion d'algèbre, une notion qui anticipe celle de σ-algèbre. Notre mesure finale sera finalement obtenue comme une restriction (sur une certaine tribu) de cette mesure extérieure.

Notre construction d'une mesure μ ira à rebours : on expliquera tout d'abord comment, à partir d'une mesure extérieure μ^*, on définit une tribu $\mathcal{M}(\mu^*)$ et une mesure μ dessus. Dans un second temps, on expliquera comment fabriquer une mesure extérieure μ^*, à partir d'une mesure additive définie sur une algèbre d'ensembles $\mathscr{A}$ (qui sera incluse dans la tribu $\mathcal{M}(\mu^*)$.

4.1 Mesures extérieures

Définition 4.1. Soit E un ensemble quelconque. Une application $\mu^* : \mathscr{P}(E) \to [0, \infty]$ est appelée une mesure extérieure (ou mesure σ-sous-additive) si :

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

ii) μ^* est croissante : si $A \subset B$, alors $\mu^*(A) \leq \mu^*(B)$;

iii) μ^* est σ-sous-additive : pour toute suite $(A_k)_{k\in\mathbb{N}}$ d'éléments de $\mathscr{P}(E)$,

$$\mu^*\left(\bigcup_{k\in\mathbb{N}} A_k\right) \leq \sum_{k\in\mathbb{N}} \mu^*(A_k).$$

On pourra comparer ces propriétés avec celles d'une mesure (v. Définition 2.8). Les inégalités ci-dessus sont également vérifiées par une mesure, puisqu'elles découlent de la propriété plus forte de σ-additivité. Nous n'imposons à μ^* que ces propriétés « faibles » ; par contre, nous imposons à μ^* d'être définie sur toute partie $A \subset E$, et pas uniquement sur une tribu particulière. En fait, la construction de la tribu sur laquelle notre mesure sera définie, va se faire « dynamiquement », à partir de μ^*.

Nous montrerons dans une section ultérieure comment construire une telle mesure ultérieure. Notre objectif présent est de montrer comment, à partir d'une mesure extérieure μ^* donnée, on construit une vraie mesure μ. La première étape est de définir la tribu sur laquelle sera définie μ, ou de façon équivalente, la notion de mesurabilité par rapport à la mesure extérieure μ^*.

Définition 4.2. Une partie $B \subset E$ est dite μ^*-mesurable si,

$$\text{pour toute partie } A \subset E, \qquad \mu^*(A) = \mu^*(A \cap B) + \mu^*(A \cap B^c). \tag{4.1}$$

On note $\mathcal{M}(\mu^*)$ l'ensemble des parties μ^*-mesurables.

Remarque 4.3. La notion de « μ^*-mesurable » est encore une nouvelle utilisation du mot « mesurable » ! Elle ne signfie pas que $\mu^*(B)$ est bien définie, puisque μ^* est en fait définie sur tout $\mathscr{P}(E)$.

On remarque aussi que l'inégalité $\mu^*(A) \leq \mu^*(A \cap B) + \mu^*(A \cap B^c)$ vraie pour toutes parties B, A, par σ-sous-additivité. L'égalité, par contre, ne sera vraie que pour certaines parties ; pour vérifier l'égalité, il suffira donc de vérifier l'inégalité inverse $\mu^*(A) \geq \mu^*(A \cap B) + \mu^*(A \cap B^c)$. La mesurabilité de la partie B indique que la mesure extérieure μ^* est additive vis-à-vis de la partition de toute partie A en $A = (A \cap B) \sqcup (A \cap B^c)$.

L'intérêt de cette notion est manifeste dans le

Théorème 4.4. *1. L'ensemble $\mathcal{M}(\mu^*)$ est une tribu sur E, qui contient toutes les parties B de E telles que $\mu^*(B) = 0$.*

2. La restriction de μ^* à cette tribu $\mathcal{M}(\mu^*)$, qu'on notera μ, est une mesure.

La propriété, pour $\mathcal{M}(\mu^*)$, de contenir tous les ensembles μ^*-négligeables, fait de la mesure μ une mesure complète.

Ce théorème nous a donc fabriqué une tribu et une mesure μ définie sur cette tribu, à partir de la seule mesure extérieure μ^*.

Démonstration. 1. Pour alléger les notations, on écrira dans cette preuve $\mathcal{M} = \mathcal{M}(\mu^*)$. Montrons tout d'abord que la classe $\mathcal{M}$ contient tous les ensembles μ^*-négligeables. Si $\mu^*(B) = 0$, on a pour toute partie A, en utilisant la croissance de μ^* :

$$\mu^*(A) \geq \mu^*(A \cap B^c) \qquad \text{et} \qquad 0 = \mu^*(B) \geq \mu^*(A \cap B),$$
$$\text{d'où on déduit} \mu^*(A) \geq \mu^*(A \cap B^c) + \mu^*(A \cap B),$$

et donc l'égalité. On a donc montré que B est μ^*-mesurable.

Montrons que $\mathcal{M}$ satisfait la définition 2.1 d'une tribu. Il est évident que $B = \emptyset$ satisfait la propriété de μ^* mesurabilité, puisque $\mu^*(\emptyset) = 0$; donc l'ensemble vide est bien dans $\mathcal{M}$. D'autre part, la propriété (4.1) est symétrique par passage au complémentaire $B \to B^c$, ce qui montre que $\mathcal{M}$ est stable par passage au complémentaire. Il reste à montrer la stabilité de $\mathcal{M}$ par union dénombrable.

Montrons déjà que $\mathcal{M}$ est stable par union finie, en particulier par union de deux de ses éléments. Soient $B_1, B_2 \in \mathcal{M}$. On veut montrer que $(B_1 \cap B_2)$ est également μ^*-mesurable. Pour tout $A \in \mathscr{P}(E)$, on a :

$$\begin{aligned}
\mu^*\left(A \cap (B_1 \cup B_2)\right) &\overset{B_1 \in \mathcal{M}}{=} \mu^*\left([A \cap (B_1 \cup B_2)] \cap B_1\right) + \mu^*\left([A \cap (B_1 \cup B_2)] \cap B_1^c\right) \\
&= \mu^*\left(A \cap B_1\right) + \mu^*\left(A \cap B_2 \cap B_1^c\right),
\end{aligned}$$

donc

$$\begin{aligned}
\mu^*\left(A \cap (B_1 \cup B_2)\right) + \mu^*\left(A \cap (B_1 \cup B_2)^c\right) &= \mu^*\left(A \cap B_1\right) + \mu^*\left(A \cap B_2 \cap B_1^c\right) + \mu^*\left(A \cap B_1^c \cap B_2^c\right) \\
&\overset{B_2 \in \mathcal{M}}{=} \mu^*\left(A \cap B_1\right) + \mu^*\left(A \cap B_1^c\right) \\
&\overset{B_1 \in \mathcal{M}}{=} \mu^*(A).
\end{aligned}$$

(on a indiqué l'hypothèse utilisée sur les égalités correspondantes). Ceci montre donc que $(B_1 \cup B_2) \in \mathcal{M}$. La classe $\mathcal{M}$ est donc stable par unions finies, et donc par intersections finies. Et donc aussi par l'opération $B, B' \in \mathcal{M} \Longrightarrow B \setminus B' \in \mathcal{M}$.

Si on considère une suite $(A_k)_{k \in \mathbb{N}}$ d'éléments de $\mathcal{M}$, on peut fabriquer la suite $B_0 = A_0$, $B_1 = A_1 \setminus A_0$, $B_2 = A_2 \setminus (A_0 \cup A_1)$, etc. La suite $(B_k)_{k \in \mathbb{N}}$ obtenue sera composée d'éléments de $\mathcal{M}$ qui sont deux à deux disjoints, et tels que pour tout $n \in \mathbb{N}$, on a $\bigcup_{k=0}^n B_k = \bigcup_{k=0}^n A_k$.

En conséquence, $\bigcup_{k=0}^{\infty} B_k = \bigcup_{k=0}^{\infty} A_k$. Pour chaque n, on a la partition :

$$E = B_0 \sqcup B_1 \cdots \sqcup B_n \sqcup \left(\bigcap_{k=0}^{n} B_k^c \right).$$

Montrons par récurrence que μ^* est additive par rapport à cette partition, c'est-à-dire que

$$\forall A \in \mathscr{P}(E), \qquad \mu^*(A) = \sum_{k=0}^{n} \mu^*(A \cap B_k) + \mu^* \left(A \cap \left(\bigcap_{k=0}^{n} B_k^c \right) \right). \tag{4.2}$$

La propriété est vraie à l'ordre $n = 0$, puisque $B_0 \in \mathcal{M}$. Supposons qu'elle est vraie à l'ordre n. En se servant du fait que $B_{n+1} \in \mathcal{M}$, on décompose le second terme en :

$$\begin{aligned} \mu^* \left(A \cap \left(\bigcap_{k=0}^{n} B_k^c \right) \right) &= \mu^* \left(A \cap \left(\bigcap_{k=0}^{n} B_k^c \right) \cap B_{n+1} \right) + \mu^* \left(A \cap \left(\bigcap_{k=0}^{n} B_k^c \right) \cap B_{n+1}^c \right) \\ &= \mu^* \left(A \cap B_{n+1} \right) + \mu^* \left(A \cap \left(\bigcap_{k=0}^{n+1} B_k^c \right) \right), \end{aligned}$$

où on a utilisé le fait que $B_{n+1} \subset (\bigcap_{k=0}^{n} B_k^c)$, puisque $\bigcup_{k=0}^{n} B_k \subset B_{n+1}^c$ (les B_k sont tous disjoints). En remplaçant cette expression dans l'égalité (4.2), on obtient l'égalité à l'ordre $n+1$, ce qui montre que la propriétés est vraie à tout ordre $n \in \mathbb{N}$.

On voudrait maintenant faire tendre n vers $+\infty$. Comme l'intersection dénombrable $(\bigcap_{k=0}^{\infty} B_k^c)$ est incluse dans $(\bigcap_{k=0}^{n} B_k^c)$, la croissance de μ^*, combinée avec (4.2), montre que pour tout $n \in \mathbb{N}$,

$$\forall A \in \mathscr{P}(E), \qquad \mu^*(A) \geq \sum_{k=0}^{n} \mu^*(A \cap B_k) + \mu^* \left(A \cap \left(\bigcap_{k=0}^{\infty} B_k^c \right) \right).$$

Dans le membre de droite on peut alors faire tendre n vers $+\infty$, et on obtient

$$\begin{aligned} \forall A \in \mathscr{P}(E), \qquad \mu^*(A) &\geq \sum_{k=0}^{\infty} \mu^*(A \cap B_k) + \mu^* \left(A \cap \left(\bigcap_{k=0}^{\infty} B_k^c \right) \right) \\ &\geq \mu^* \left(\bigcup_k (A \cap B_k) \right) + \mu^* \left(A \cap \left(\bigcap_{k=0}^{\infty} B_k^c \right) \right), \end{aligned} \tag{4.3}$$

où, dans la dernière ligne, on s'est servi de la propriété de σ-sous-additivité de μ^*. On remarque alors que les deux ensembles ci-dessus peuvent s'écrire

$$\bigcup_k (A\cap B_k) = A\cap\left(\bigsqcup_k B_k\right) = A\cap\left(\bigcup_k A_k\right), \quad \text{et} \qquad \left(\bigcap_{k=0}^{\infty} B_k^c\right) = \left(\bigsqcup_k B_k\right)^c = \left(\bigcup_k A_k\right)^c.$$

On a donc montré $\mu^*(A) \geq \mu^*\left(A\cap(\bigcup_k A_k)\right) + \mu^*\left(A\cap(\bigcup_k A_k)^c\right)$; comme on a l'inégalité inverse, on a l'égalité entre les deux membres :

$$\mu^*(A) = \mu\left(A\cap\left(\bigcup_k A_k\right)\right) + \mu^*\left(A\cap\left(\bigcup_k A_k\right)^c\right), \tag{4.4}$$

ce qui montre que$\bigcup_k A_k$ est μ^*-mesurable. On a ainsi montré que la classe $\mathcal{M}$ est une tribu.

2. Notons μ la restriction de μ^* à la tribu $\mathcal{M}$. Vérifions que μ satisfait bien la Définition 2.8 d'une mesure. On sait déjà que $\mu(\emptyset) = 0$. Soit $(B_k)_{k\in\mathbb{N}}$ une suite d'éléments disjoints de $\mathcal{M}$. La preuve du 1. ci-dessus nous a montré que $\bigsqcup_k B_k \in \mathcal{M}$. D'autre part, l'inégalité (4.3) combinée avec la σ-sous-additivité, montre que cette inégalité est en fait une égalité :

$$\forall A \in \mathscr{P}(E), \qquad \mu^*(A) = \sum_{k=0}^{\infty} \mu^*(A\cap B_k) + \mu^*\left(A\cap\left(\bigsqcup_k B_k\right)^c\right).$$

Si on choisit $A = \bigsqcup_k B_k$, on a alors $A \in \mathcal{M}$, et l'égalité ci-dessus s'écrit $\mu\left(\bigsqcup_k B_k\right) = \sum_{k\in\mathbb{N}} \mu(B_k)$. On a montré la σ-additivité de μ. Ceci conclut la preuve que μ est une mesure. □

4.2 Construction de la mesure extérieure de Lebesgue λ^* sur $\mathbb{R}$

Avant d'élaborer une théorie générale, on construira la mesure extérieure λ^* sur $\mathscr{P}(\mathbb{R})$, dont la restriction donnera la mesure de Lebesgue λ. On définit λ^* par la formule variationnelle suivante :

$$\forall A \subset E, \quad \lambda^*(A) := \inf\left\{\sum_{i\in\mathbb{N}} (b_i - a_i)\,;\, A \subset \bigcup_{i\in\mathbb{N}}]a_i, b_i[\right\}. \tag{4.5}$$

A droite on prend l'infimum sur toutes les suites d'intervalles ouverts $(]a_i, b_i[)_{i\in\mathbb{N}}$ tels que $A \subset \bigcup_{i\in\mathbb{N}}]a_i, b_i[$; on inclut a possibilité d'intervalles vides $]a_i, a_i[$, autrement dit on considère aussi les unions finies d'intervalles. Notons que l'infimum peut valoir $+\infty$, si toute famille d'intervalles contenant A satisfait $\sum_{i\in\mathbb{N}}(b_i - a_i) = +\infty$.

Cette définition nous montre déjà que $\lambda^*(A) \in [0, \infty]$ pour toute partie A, et que $\lambda^*(\emptyset) = 0$.

Par contre, il n'est pas complètement évident que $\lambda^*(]a,b[) = b-a$, nous le montrerons ci-dessous.

Il faut vérifier que la fonction λ^* définit bien une mesure extérieure.

Théorème 4.5. *i) λ^* est une mesure extérieure sur $\mathbb{R}$.*

ii) La tribu $\mathcal{M}(\lambda^*)$ contient la tribu borélienne $\mathcal{B}(\mathbb{R})$.

iii) Pour tous $a \leq b$, $\lambda^*([a,b]) = \lambda^*(]a,b[) = b-a$.

La restriction de λ^* à la tribu $\mathcal{M}(\lambda^*)$ s'appelle la mesure de Lebesgue sur $\mathbb{R}$, notée λ. Comme sa tribu $\mathcal{M}(\lambda^*)$ contient tous les boréliens, elle contient tous les intervalles ouverts, et leur mesure vaut $\lambda(]a,b[) = b-a$. Comme on l'a montré dans l'Exemple 2.17, λ est l'unique mesure possédant cette propriété.

Démonstration. i) la définition montre immédiatement que $\lambda^*(\emptyset) = 0$, et que λ^* est croissante. Il reste à montrer la σ-sous-additivité.

Soit $(A_n)_{n\in\mathbb{N}}$ une suite de parties de $\mathbb{R}$. Si l'une des parties satisfait $\lambda^*(A_i) = \infty$, il n'y a rien à montrer, donc on supposera que toutes les $\lambda^*(A_i) < \infty$. Choisissons un $\epsilon > 0$. Pour tout $n \in \mathbb{N}$, la définition de $\lambda^*(A_n)$ montre qu'on peut recouvrir A_n par une union d'intervalles

$$A_n \subset \bigcup_{i\in\mathbb{N}}]a_i^{(n)}, b_i^{(n)}[, \quad \text{tels que} \quad \sum_{i\in\mathbb{N}} (b_i^{(n)} - a_i^{(n)}) \leq \lambda^*(A_n) + \frac{\epsilon}{2^n}.$$

Si on somme sur les indices i et n, on obtient une famille *dénombrable* d'intervalles $\left(]a_i^{(n)}, b_i^{(n)}[\right)_{i,n\in\mathbb{N}}$, dont l'union recouvre $\bigcup_n A_n$, et telle que

$$\sum_{i,n\in\mathbb{N}} (b_i^{(n)} - a_i^{(n)}) \leq \sum_n \left(\lambda^*(A_n) + \frac{\epsilon}{2^n}\right) = 2\epsilon + \sum_n \lambda^*(A_n).$$

Par définition, $\lambda^*\left(\bigcup_n A_n\right)$ est majoré par le membre de gauche :

$$\lambda^*\left(\bigcup_n A_n\right) \leq 2\epsilon + \sum_n \lambda^*(A_n).$$

Enfin, on peut prendre $\epsilon \searrow 0$ dans l'expression ci-dessus, ce qui montre la σ-sous-additivité de λ^*. On a donc montré que λ^* est une mesure extérieure.

ii) On sait déjà que $\mathcal{M}(\lambda^*)$ est une tribu. Pour montrer qu'elle contient les boréliens, il suffit de montrer qu'elle contient une sous-famille $\mathcal{C}$ qui engendre les boréliens. Par exemple, il suffit

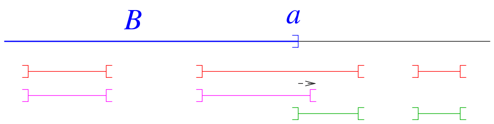

FIGURE 4.1 – Représentation schématique des intervalles $]a_i, b_i[$ (rouge) recouvrant A, et de leurs troncations recouvrant $A \cap B$ (famille $\mathcal{F}_1$, violet) et $A \cap B^c$ (famille $\mathcal{F}_2$, vert). Le prolongement par $\epsilon 2^{-i}$ est indiqué par une flèche.

de montrer qu'elle contient les intervalles de la forme $\{]-\infty, a],\ a \in \mathbb{R}\}$. Choisissons donc un tel intervalle $Soit B =]-\infty, a]$. Pour montrer qu'il appartient à $\mathcal{M}(\lambda^*)$, il faut vérifier que pour toute partie $A \subset \mathbb{R}$,

$$\lambda^*(A) \geq \lambda^*(A \cap B) + \lambda^*(A \cap B^c) \tag{4.6}$$

(l'inégalité inverse étant vraie par la sous-additivité de λ^* qu'on vient de montrer). Il faut donc comparer les recouvrements de $A, A \cap B$ et $A \cap B^c$ par des unions d'intervalles ouverts. Soit $(]a_i, b_i[)_{i \in \mathbb{N}}$ une famille d'intervalles recouvrant A, et telle que

$$\sum_{i \in \mathbb{N}} (b_i - a_i) \leq \lambda^*(A) + \epsilon. \tag{4.7}$$

En rappelant la notation $a \wedge b = \min(a, b)$, on définit alors les deux familles :

$$\mathcal{F}_1 = \big(]a_i \wedge a, (b_i \wedge a) + \epsilon 2^{-i}[\big)_{i \in \mathbb{N}}, \quad \mathcal{F}_2 = (]a_i \vee a, b_i \vee a[)_{i \in \mathbb{N}}.$$

La première famille recouvre $A \cap]-\infty, a]$, la second recouvre $A \cap]a, \infty[$. V. la figure pour visualiser ces deux familles. A partir de ces deux familles, on trouve alors les inégalités :

$$\begin{aligned}
\lambda^*(A \cap B) &\leq \sum_{i \in \mathbb{N}} \left[(b_i \wedge a) + \epsilon 2^{-i} - a_i \wedge a\right] = 2\epsilon + \sum_{i \in \mathbb{N}} \left[(b_i \wedge a) - a_i \wedge a\right], \\
\lambda^*(A \cap B^c) &\leq \sum_{i \in \mathbb{N}} \left[b_i \wedge a - a_i \wedge a\right].
\end{aligned}$$

En sommant ces deux inégalités et en remarquant que pour tout intervalle $]a_i, b_i[$, on a

$$(b_i \wedge a) - a_i \wedge a + b_i \wedge a - a_i \wedge a = b_i - a_i,$$

on trouve que

$$\lambda^*(A \cap B) + \lambda^*(A \cap B^c) \leq 2\epsilon + \sum_{i \in \mathbb{N}} \left[b_i - a_i\right].$$

En comparant avec l'inégalité (4.7), on trouve donc

$$\lambda^*(A \cap B) + \lambda^*(A \cap B^c) \leq \lambda^*(A) + 3\epsilon.$$

Comme $\epsilon > 0$ est arbitrairement petit, on a donc montré l'inégalité (4.6) qu'on voulait, montrant que l'intervalle B est λ^*-mesurable. La tribu $\mathcal{M}(\mu^*)$ contient donc les boréliens.

iii) Comme l'intervalle fermé $[a,b]$ est recouvert par $]a-\epsilon, b+\epsilon[$ pour tout $\epsilon > 0$, on a $\lambda^*([a,b]) \leq b-a+2\epsilon$, et donc $\lambda^*([a,b]) \leq b-a$. Montrons l'inégalité inverse. Supposons que $[a,b]$ est recouvert par la famille $(]a_i, b_i[)_{i\in\mathbb{N}}$. Comme $[a,b]$ est compact, le théorème de Borel-Lebesgue montre qu'il est en fait recouvert par une union finie, qu'on peut appeler $\bigcup_{i=0}^{N}]a_i, b_i[$. Vérifions la propriété intuitive que la somme des longueurs de ces intervalles doit forcément majorer celle de $[a,b]$:

$$b - a \leq \sum_{i=0}^{N} (b_i - a_i).$$

Quitte à restreindre l'union encore plus, on peut supposer que tous les $]a_i, b_i[$ intersectent $[a,b]$. Un des intervalles (on dira que c'est le premier) doit contenir le point a, donc vérifier $a_0 < a < b_0$. Si $b_0 > b$, on a donc $b_0 - a_0 > b - a$, et la preuve est finie. Si $b_0 < b$, le point b_1 doit être recouvert par un intervalle de la famille (supposons qu'il s'agit de $]a_1, b_1[$) donc tel que $a_1 < b_0 < b_1$. On a alors $b_1 - a_1 + b_0 - a_0 > b_1 - a_0$. Si $b_1 > b$ on a montré la propriété voulue. Si $b_1 < b$, on identifie un troisième intervalle de la famille contenant b_1 ; on aura alors $\sum_{i=0}^{2} (b_i - a_i) > b_2 - a_0$. On procède ainsi de suite, jusqu'à ce qu'avoir un intervalle $]a_n, b_n[$ tel que $b_n > b$, et donc tel que

$$\sum_{i=0}^{n} (b_i - a_i) > b_2 - a_0 > b - a.$$

En rajoutant les intervalles « inutiles », on a donc

$$b - a \leq \sum_{i=0}^{\infty} (b_i - a_i).$$

En prenant l'infimum sur tous les recouvrements, on trouve donc $b - a \leq \lambda^*([a,b])$, et finalement $\lambda^*([a,b]) = b - a$.

Pour calculer la mesure de l'intervalle ouvert $]a,b[$, il suffit de remarquer que les singletons $\{a\} = [a,a]$ et $\{b\} = [b,b]$ sont de mesure nulle, et que l'additivité de λ^* sur sa tribu implique

que $\lambda^*([a,b]) = \lambda^*(]a,b[) + \lambda^*(\{a\}) + \lambda^*(\{b\})$. $\square$

Avec cette construction de la mesure extérieure s'achève celle de la mesure de Lebesgue. Le point ii) du théorème ci-dessus nous dit que la tribu $\mathcal{M}(\lambda^*)$ sur laquelle λ est définie (qu'on appellera la *tribu de Lebesgue*) contient la tribu borélienne $\mathcal{B}(\mathbb{R})$, mais elle n'est pas forcément égale à celle-ci. Mais la différence entre ces deux tribus est, d'une certaine façon, peu importante. On sait aussi que cette tribu $\mathcal{M}(\lambda^*)$ contient tous les ensembles λ-négligeables, ce qui n'était pas forcément le cas de la tribu borélienne. Nous allons donner une autre caractérisation de la tribu de Lebesgue, comme la complétion de la tribu de Borel par rapport à λ. Commençons par définir la complétée d'une tribu par rapport à une de ses mesures.

Proposition 4.6. *Soit $(E, \mathcal{A}, \mu)$ un espace mesuré. La tribu complétée de $\mathcal{A}$ par rapport à μ est définie par $\bar{\mathcal{A}} = \sigma(\mathcal{A} \cup \mathcal{N})$, où $\mathcal{N}$ est l'ensemble des parties μ-négligeables de E. Il existe alors sur $(E, \bar{\mathcal{A}})$ une unique mesure qui prolonge μ.*

Démonstration. Montrons qu'une façon de construire la tribu complétée est de considérer toutes les parties de E qui sont « encadrées » par deux éléments de $\mathcal{A}$ de même mesure. On définit la classe

$$\mathcal{C} := \{A \in \mathscr{P}(E)\,;\, \exists B, B' \in \mathcal{A},\ B \subset A \subset B',\ \mu(B' \setminus B) = 0\},$$

nous allons montrer que $\mathcal{C} = \bar{\mathcal{A}}$. On montre d'abord que $\mathcal{C}$ est une tribu. Si A est dans $\mathcal{C}$, le passage au complémentaire montre que $B'^c \subset A^c \subset B^c$, avec B'^c, B^c dans $\mathcal{A}$, et $B^c \setminus B'^c = B' \setminus B$ est de mesure nulle, donc $A \in \mathcal{C}$. Si les $(A_k)_{k\in\mathbb{N}}$ sont encadrés par $(B_k)_k$ et $(B'_k)_k$, alors $\bigcup_k A_k$ est encadrée par les éléments $\bigcup_k B_k$ et $\bigcup_k B'_k$ de $\mathcal{A}$, et on vérifie que

$$\begin{aligned}\mu\left(\left(\bigcup_k B'_k\right) \setminus \left(\bigcup_j B_j\right)\right) &= \mu\left(\bigcup_k \left(B'_k \setminus \left(\bigcup_j B_j\right)\right)\right) \leq \mu\left(\bigcup_k (B'_k \setminus B_k)\right) \\ &\leq \sum_k \mu(B'_k \setminus B_k) = 0.\end{aligned}$$

Ceci montre que $\bigcup_k A_k$ est bien dans $\mathcal{C}$. Comme la classe $\mathcal{C}$ contient à la fois $\mathcal{A}$ et l'ensemble des négligeables $\mathcal{N}$, elle contient $\bar{\mathcal{A}}$. Montrons l'inclusion inverse $\mathcal{C} \subset \bar{\mathcal{A}}$. Soit un élément $A \in \mathcal{C}$; il existe alors $B, B' \in \mathcal{A}$ encadrant A, donc en particulier $A = B \cup (A \setminus B)$, et $A \setminus B \subset B' \setminus B$ est négligeable. Ceci montre que A est dans l'union $\mathcal{A} \cup \mathcal{N}$, donc dans sa tribu engendrée $\bar{\mathcal{A}}$.

Remarque : On peut voir les élements de $\mathcal{C} = \bar{\mathcal{A}}$ comme des éléments de $\mathcal{A}$ auxquels ont aurait rajouté, ou retiré, de la « poussière » de mesure nulle.

Il est alors facile de prolonger la mesure μ à cette tribu complétée : si $A \in \bar{\mathcal{A}}$ est encadrée par B, B', comme $\mu(B') = \mu(B) + \mu(B' \setminus B) = \mu(B)$, on posera naturellement $\mu(A) = \mu(B)$. Il est facile de voir que cette définition ne dépend pas du choix de la paire (B, B') encadrant A. En effet, si A est encadrée par une seconde paire $(\tilde{B}, \tilde{B}')$, on a alors $\mu(\tilde{B}') = \mu(\tilde{B})$, mais aussi $\mu(\tilde{B}') \geq \mu(B)$ puisque $B \subset A \subset \tilde{B}'$, et inversement $\mu(\tilde{B}) \leq \mu(B')$. On en déduit donc que ces quatre ensembles ont la même mesure. Montrons la σ-additivité de cette mesure prolongée. Si $(A_n)_n$ est une suite d'éléments disjoints de $\bar{\mathcal{A}}$, chacun est encadré par une paire (B_n, B'_n). Les $(B_n)_n$ forment alors une suite disjointe, et on aura alors

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

Enfin, on remarque comme ci-dessus que $(\bigcup_n A_n) \setminus (\bigcup_k B_k) = \bigcup_n (A_n \setminus B_n)$ du fait de la disjonction des $(A_n)_n$. Cette dernière union est négligeable comme union d'ensembles négligeables, de sorte que $\mu(\bigcup_n A_n) = \mu(\bigcup_n B_n) = \sum_n \mu(A_n)$. $\square$

On va pouvoir appliquer cette notion de tribu complétée pour décrire la tribu de Lebesgue à partir de celle de Borel.

Proposition 4.7. *La tribu de Lebesgue* $\mathcal{M}(\lambda^*)$ *est la complétée de la tribu de Borel* $\mathcal{B}(\mathbb{R})$ *par rapport à la mesure de Lebesgue. On notera cette tribu complétée* $\overline{\mathcal{B}(\mathbb{R})}$*, en gardant en mémoire qu'elle est complète pour la mesure de Lebesgue. O*

Démonstration. D'après le théorème 4.5, $\mathcal{M}(\lambda^*)$ contient les boréliens, et contient les ensembles λ^*-négligeables de $\mathbb{R}$, qui sont aussi les ensembles λ-négligeables. Elle contient donc $\overline{\mathcal{B}(\mathbb{R})}$, la tribu complétée par rapport à la mesure de Lebesgue.

Montrons inversement que si $A \in \mathcal{M}(\lambda^*)$, alors $A \in \overline{\mathcal{B}(\mathbb{R})}$. On notera ici λ la mesure de Lebesgue sur $\mathcal{M}(\lambda^*)$.

1. Supposons dans un premier temps que A est borné, c'est-à-dire que $A \subset]-K, K[$ pour un certain $K > 0$. Par croissance de λ on aura alors $\lambda(A) < \infty$. Pour tout $n > 0$, on peut trouver une famille d'intervalles $\left(]a_i^{(n)}, b_i^{(n)}[\right)_{i\in\mathbb{N}}$ qui recouvrent A, en « l'approximant à une erreur $\frac{1}{n}$ près », autrement dit tels que

$$\lambda(A) + \frac{1}{n} \geq \sum_i \left(b_i^{(n)} - a_i^{(n)}\right).$$

Notons $B_n = \bigcup_i]a_i^{(n)}, b_i^{(n)}[$ ce recouvrement de A ; c'est bien sûr un borélien, et il satisfaira (par σ-sous-additivité) $\lambda(A) + \frac{1}{n} \geq \lambda^*(B_n)$. Bien que les $(B_n)_n$ ne forment pas forcément une suite décroissante, on imagine qu'ils doivent se « serrer de plus en plus » autour de A lorsque $n \to \infty$. On peut considérer leur intersection $B' := \bigcap_n B_n$, qui est aussi un borélien contenant A. Cet ensemble vérifie

$$\lambda(A) + \frac{1}{n} \geq \lambda(B'), \text{ pour tout } n, \text{ donc } \quad \lambda(A) \geq \lambda(B'), \quad \text{donc} \quad \lambda(A) = \lambda(B').$$

En considérant l'ensemble $\tilde{A} =]-K, K[\setminus A$, qui est également dans la tribu $\mathcal{M}(\lambda^*)$, on fabrique de la même manière un borélien $\tilde{B}' \supset \tilde{A}$ (qu'on peut supposer inclus dans $]-K, K[$), et tel que $\lambda(\tilde{B}') = \lambda(\tilde{A})$. Son complémentaire dans cet intervalle, $B =]-K, K[\setminus \tilde{B}'$, satisfait $B \subset A$ et $\lambda(B) = \lambda(A)$ par additivité de la mesure. On a donc identifié deux boréliens (B, B') encadrant A, et tels que $\lambda(B' \setminus B) = 0$. On peut interpréter B comme un « rétrécissement » de A, et B' un « épaississement » de A, de manière à entrer dans la tribu de Borel.

2. On abandonne maintenant l'hypothèse que $A \in \mathcal{M}(\lambda^*)$ est borné. On peut néanmoins obtenir A comme l'union croissante des ensembles $A_K := A \cap]-K, K[$, qui sont tous dans $\mathcal{M}(\lambda^*)$. La preuve ci-dessus, appliquée à A_K, montre que chaque $A_K \in \overline{\mathcal{B}(\mathbb{R})}$. Comme $\overline{\mathcal{B}(\mathbb{R})}$ est une tribu, leur union $A = \bigcup_{K \geq 1} A_K \in \overline{\mathcal{B}(\mathbb{R})}$. $\square$

Proposition 4.8. *Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction borélienne. Supposons que la fonction $g : \mathbb{R} \to \mathbb{R}$ est égale à f λ-presque partout. Alors g est mesurable pour la tribu complétée $\overline{\mathcal{B}(\mathbb{R})}$.*

Démonstration. Par hypothèse, il existe un ensemble borélien $A \in \mathcal{B}(\mathbb{R})$ tel que $\lambda(A) = 0$, et tel que $f_{\restriction A^c} = g_{\restriction A^c}$. Pour tout borélien $B \in \mathcal{B}(\mathbb{R})$, l'image réciproque de B par g peut se décomposer en :

$$\begin{aligned} g^{-1}(B) &= \left(g^{-1}(B) \cap A^c\right) \cup \left(g^{-1}(B) \cap A\right) \\ &= \left(f^{-1}(B) \cap A^c\right) \cup \left(g^{-1}(B) \cap A\right). \end{aligned}$$

L'ensemble $(f^{-1}(B) \cap A^c)$ est borélien, et $(g^{-1}(B) \cap A) \subset A$ est λ-négligeable, donc leur union est bien dans la tribu complétée $\overline{\mathcal{B}(\mathbb{R})}$. $\square$

4.3 Généralisations en dimension supérieure

La mesure extérieure de Lebesgue sur $\mathbb{R}$ a été définie en (4.5) à partir des longueurs d'intervalles ouverts. En dimension $d > 1$, il est naturel de remplacer les intervalles par des pavés ouverts, qui sont des ouverts de $\mathbb{R}^d$ de la forme :

$$P = \prod_{j=1}^{d}]a_j, b_j[.$$

On peut aussi considérer les pavés fermés $\bar{P} = \prod_{j=1}^{d} [a_j, b_j]$. Le volume d'un pavé vaut, par définition, $\mathrm{vol}(P) = \prod_{j=1}^{d} (b_j - a_j)$. On peut alors définir la mesure extérieure extérieure de Lebesgue sur $\mathbb{R}^d$, par

$$\forall A \subset \mathbb{R}^d, \qquad \lambda^*(A) := \inf \left\{ \sum_{i \in \mathbb{N}} \mathrm{vol}(P_i)\, ;\, A \subset \bigcup_{i \in \mathbb{N}} P_i \right\},$$

où on prend l'infimum sur tous les recouvrements de A par une union dénombrable de pavés ouverts. On montre alors un théorème analogue au Thme 4.5 :

i) la fonction λ^* ci-dessus est une mesure extérieure

ii) sa tribu engendrée $\mathcal{M}(\lambda^*)$ contient la tribu $\mathcal{B}(\mathbb{R}^d)$. On peut pour utiliser le fait que les boréliens sur $\mathbb{R}^d$ peuvent être engendrés par les « cylindres » de la forme

$$B = \mathbb{R} \times \cdots \times] - \infty, a] \times \mathbb{R} \times \cdots \times \mathbb{R},$$

donc il suffira de montrer que ces cylindres sont λ^*-mesurables. Comme on l'a fait en dimension 1, à partir d'un recouvrement de $A \subset \mathscr{P}(\mathbb{R}^d)$ par une famille dénombrable $\mathcal{F}$ de pavés, et pour tout choix de $\epsilon > 0$, on peut fabriquer deux familles $\mathcal{F}_1, \mathcal{F}_2$ de pavés recouvrant respectivement $A \cap B$ et $A \cap B^c$, et tels que

$$\sum_{P \in \mathcal{F}_1} \mathrm{vol}(P) + \sum_{P \in \mathcal{F}_2} \mathrm{vol}(P) \leq \epsilon + \sum_{P \in \mathcal{F}} \mathrm{vol}(P).$$

Si la famille $\mathcal{F}$ recouvre A à ϵ près, on aura donc

$$\lambda^*(A \cap B) + \lambda^*(A \cap B^c) \leq \lambda^*(A) + 2\epsilon, \quad \text{donc} \quad \lambda^*(A \cap B) + \lambda^*(A \cap B^c) = \lambda^*(A),$$

et donc la mesurabilité de B.

iii) On veut montrer que la mesure du pavé fermé $\lambda^*(\bar{P}) = \mathrm{vol}(\bar{P})$. Il s'agit alors de vérifier que pour tout recouvrement de $\bar{P}$ par une union finie de pavés ouverts

$$\bar{P} \subset \bigcup_{i=1}^{n} P_i, \quad \text{on a alors} \qquad \mathrm{vol}(\bar{P}) \leq \sum_{i=1}^{n} \mathrm{vol}(P_i).$$

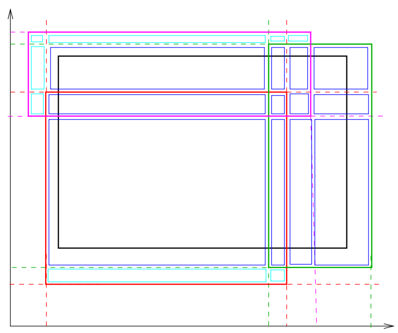

FIGURE 4.2 – Recouvrement d'un pavé P (en noir) en l'union finie de pavés P_i (en rouge, vert, rose). En prolongeant les faces des P_i, on partitionne $\bigcup_i P_i$ en « petits pavés » $P^{(k)}$ (en bleu). Certains sont utilisés pour recouvrir P (bleu foncé), d'autres seront inutiles (bleu clair). Le calcul de $\mathrm{vol}(P_i)$ et de $\mathrm{vol}(P)$ peut se faire en factorisant selon les $d = 2$ dimensions.

Cette vérification n'est pas aussi immédiate qu'en dimension 1. Une façon de procéder est de prolonger les faces de tous les P_i par des hyperplans, de façon à définir des « petits » pavés $\left(P^{(k)}\right)_{1\leq k\leq K}$; chaque pavé P_i est partitionné en une union de plusieurs pavés $P^{(k)}$: $P_i = \bigsqcup_{k\in I(i)} P^{(k)}$, certains $P^{(k)}$ pouvant apparaître dans la décomposition de plusieurs P_i. La décomposition forme un « quadrillage », de sorte que la formule

$$\mathrm{vol}(P_i) = \sum_{k\in I(i)} \mathrm{vol}(P^{(k)})$$

se montre en partitionnant chacun des intervalles formant P_i, de façon à factoriser la somme de droite sur les d dimensions. Le pavé P est aussi décomposé dans l'union disjointe d'un certain nombre de pavés, $P = \bigsqcup_{k\in I} P^{(k)}$, et on a encore $\mathrm{vol}(P) = \sum_{k\in I} \mathrm{vol}(P^{(k)})$ par factorisation sur les dimensions. Lorsqu'on compare cette somme à

$$\sum_{i=1}^{n} \mathrm{vol}(P_i) = \sum_{i=1}^{n} \sum_{k\in I(i)} \mathrm{vol}(P^{(k)}),$$

on voit que tous les $\left(P^{(k)}\right)_{1\leq k\leq K}$ apparaissent dans le membre de droite au moins une fois, et certains peuvent y apparaître plusieurs fois. Cette somme majore donce la somme $\mathrm{vol}(P) = \sum_{k\in I} \mathrm{vol}(P^{(k)})$, dans laquelle chaque $P^{(k)}$ apparaît au plus une fois.

4.4 Quelques propriétés de la mesure de Lebesgue

4.4.1 Invariance par translation

L'espace euclidien $\mathbb{R}^d$ admet une structure de groupe abélien, le groupe des translations. Cette action de groupe permet de caractériser la mesure de Lebesgue.

Théorème 4.9. *La mesure de Lebesgue λ sur $\mathbb{R}^d$ est invariante par translation : pour tout $A \in \mathcal{B}(\mathbb{R}^d)$ et pour tout vecteur $v \in \mathbb{R}^d$, $\lambda(A) = \lambda(A+v)$.*

Inversement, si μ est une mesure sur $\mathcal{B}(\mathbb{R}^d)$ localement finie, et invariante par translation, alors il existe une constante $c \geq 0$ telle que $\mu = c\lambda$.

Ainsi, à une constante multiplicative près, la mesure de Lebesgue est la seule mesure localement finie qui soit invariante par translations.

Démonstration. V. la feuille d'exercices « Week 2 », exercice 1. La première assertion est assez simple à vérifier. La formule $\lambda(A) = \lambda(A+v) := \sigma_v(\lambda)(A)$ est vraie lorsque A est un pavé, par calcul explicite. Le Lemme de classe monotone implique alors que les mesures λ et $\sigma_v(\lambda)$ coincident sur toute la tribu des boréliens.

Pour montrer la réciproque, on se place sur l'hypercube unité $C = [0,1]^d$, et on note $c = \mu(C)$. Si on décompose C en un treillis de n^d petits hypercubes $C_k^{(n)}$, on a par invariance par translation et additivité, que $\mu(C_k^{(n)}) = \frac{c}{n^d}$. Pour tout $a_1, \cdots, a_d \in]0,1]$, le pavé $P = \prod_{j=1}^d [0, a_j]$ peut être encadré par deux unions disjointes de petits cubes :

$$\bigsqcup_{k \in I^-} P^{(k)} \subset P \subset \bigsqcup_{k \in I^+} P^{(k)},$$

tels que le nombre $\#I^+ - \#I^- = \mathcal{O}(n^{d-1})$. La différence de volumen entre les deux unions est donc d'ordre $\mathcal{O}(n^{-1})$. En étudiant plus précisément les unions disjointes de droite et de gauche, et en faisant tendre n vers l'infini, on trouve que $\mu(P) = c \prod_{j=1}^d a_j = c\lambda(P)$. □

4.4.2 Régularité

Lorsqu'on a une mesure μ définie sur la tribu de Borel d'un espace topologique E, on peut étudier le comportement de μ lorsqu'on approche un borélien A par au-dessus par des ouverts, et par en-dessous par des compacts. Le bon comportement de μ vis-à-vis de ces limites définit la notion de régularité de cette mesure.

Définition 4.10. Soit E un espace topologique séparé, et μ une mesure définie sur la tribu de Borel $\mathcal{B}(E)$. Cette mesure est dite

1. extérieurement régulière si, pour tout $A \in \mathcal{B}(E)$, $\mu(A) = \inf\{\mu(O)\,;\, O \text{ ouvert de } E,\ A \subset O\}$;
2. intérieurement régulière si, pour tout $A \in \mathcal{B}(E)$, $\mu(A) = \sup\{\mu(K)\,;\, K \text{ compact de } E,\ K \subset A\}$;
3. régulière si elle est à la fois extérieurement régulière et intérieurement régulière.

Proposition 4.11. *La mesure de Lebesgue sur $\mathbb{R}^d$ est régulière.*

Démonstration. On cherche d'abord à montrer la régularité extérieure. Soit $A \in \mathcal{B}(\mathbb{R}^d)$. Si $\lambda(A) = \infty$, le résultat est évident, vu que $\lambda(A) \leq \lambda(O)$ pour tout ouvert $O \supset A$. On suppose maintenant que $\lambda(A) < \infty$. Comme $\lambda(A) = \lambda^*(A)$, on peut trouver un recouvrement de A par des pavés $(P_i)_{i\in\mathbb{N}}$, tel que $\lambda^*(A) + \epsilon \geq \sum_i \text{vol}(P_i)$. Par sous-additivité, l'ouvert $O = \bigcup_i P_i$ vérifiera $\lambda(O) \leq \sum_i \text{vol}(P_i)$, et $\lambda(O) \geq \lambda(A)$ puisqu'il recouvre A. On a donc obtenu

$$\lambda(O) \geq \lambda(A) \geq \lambda(O) - \epsilon,$$

ce qui montre bien la régularité extérieure.

Pour montrer la régularité intérieure on va d'abord supposer que A est contenu dans un hypercube fermé $C = [-n, n]^d$. L'ensemble $C \setminus A$ est un borélien, on peut donc l'approcher par un ouvert O le contenant, de telle façon que $\lambda(C \setminus A) \geq \lambda(O) - \epsilon$. En passant au complémentaire par rapport à C, on en déduit :

$$\lambda(A) \leq \lambda(C \setminus O) + \epsilon.$$

On vérifie que $C\setminus O$ est fermé borné, donc compact. On a ainsi pu approcher A par en-dessous par des compacts.

Si maintenant A n'est pas borné, on construit la suite croissante $A_n := A \cap [-n, n]^d$. On aura alors $\lambda(A) = \lim_{n\to\infty} \uparrow \lambda(A_n)$. Pour chaque n, et chaque $\epsilon > 0$, on peut trouver un compact $K_n \subset [-n, n]^d$ tel que $\lambda(K_n) \geq \lambda(A_n) - \epsilon$. Dans les deux cas où $\lambda(A) < \infty$ ou $\lambda(A) = \infty$, on vérifier donc la régularité intérieure. $\square$

4.4.3 Relation avec l'intégrale de Riemann

Dans l'introduction du cours, nous avons rappelé la définition de l'intégrale de Riemann sur un intervalle $I \subset \mathbb{R}$, d'abord pour les fonctions continues, puis les fonctions réglées, puis finalement les fonctions dites de Riemann. Dans le cadre le plus général, la définition de l'intégrale de Riemann est donnée de façon variationnelle, mais à partir de fonctions en escalier, qui sont une sous-classe des fonctions étagées. Pour toute fonction $f : I \to \mathbb{R}$ bornée, on peut définir deux expressions variationnelles :

$$S_-(f) \stackrel{\text{def}}{=} \sup \left\{ \int_I h(x)\, dx \,;\, h \text{ en escalier, } g \leq f \right\},$$
$$S_+(f) \stackrel{\text{def}}{=} \inf \left\{ \int_I \tilde{h}(x)\, dx \,;\, \tilde{h} \text{ en escalier, } \tilde{h} \geq f \right\}.$$

Pour une fonction f bornée quelconque, on a forcément $S_-(f) \leq S_+(f)$.

Définition 4.12. Une fonction bornée $f : I \to \mathbb{R}$ est dite Riemann-intégrable si $S_-(f) = S_+(f)$. Dans ce cas la valeur commune définit l'intégrale de Riemann de f, qu'on notera $S(f)$ ou $I(f)$.

Il est alors naturel de se poser la question suivante :

> Une fonction f Riemann-intégrable est-elle borélienne ? Si c'est le cas, son intégrale de Riemann coïncide-t-elle avec son intégrale par la mesure de Lebesgue ?

Théorème 4.13. *Soit une fonction $f : I \to \mathbb{R}$ Riemann-intégrable, d'intégrale de Riemann $S(f)$. Alors f est mesurable pour la tribu complétée $\overline{\mathcal{B}(I)}$, et on a l'égalité*

$$S(f) = \int_I f\, d\lambda.$$

Démonstration. Soit une suite $(h_n)_{n\in\mathbb{N}}$ de fonctions en escalier, telles que $h_n \leq f$ et $\lim_{n\to\infty} \uparrow S(h_n) = S(f)$. Notons que si on remplace h_n par $h'_n = \max(h_1, \ldots, h_n)$, la fonction h'_n est toujours en escalier, elle est toujours majorée par f, donc on aura aussi $\lim_{n\to\infty} \uparrow S(h'_n) = S(f)$. Autrement dit, on peut supposer que la suite $(h_n)_{n\in\mathbb{N}}$ est croissante.

De façon identique, on peut construire une suite $(\tilde{h}_n)_{n\in\mathbb{N}}$ décroissante de fonctions en escalier, majorant f, et telles que $\lim_{n\to\infty} \downarrow S(\tilde{h}_n) = S(f)$. Comme les suites $(h_n)_n$ et $(\tilde{h}_n)_n$ sont monotones, on peut définir les fonctions $h_\infty = \lim_n \uparrow h_n$ et $\tilde{h}_\infty = \lim_n \downarrow \tilde{h}_n$. Comme les fonctions $h_n, \tilde{h}_n$ sont en escalier, elles sont boréliennes, donc les fonctions h_∞ et $\tilde{h}_\infty$ sont aussi borélienes ; elles sont aussi bornées sur I, puisque f l'est, donc elles sont forcément intégrables par rapport à λ. Le théorème de convergence dominée montre que

$$\int h_\infty \, d\lambda = \lim_{n\to\infty} \uparrow \int h_n \, d\lambda = \lim_{n\to\infty} \uparrow S(h_n) = S_-(f),$$
$$\int \tilde{h}_\infty \, d\lambda = \lim_{n\to\infty} \downarrow \int \tilde{h}_n \, d\lambda = \lim_{n\to\infty} \downarrow S(\tilde{h}_n) = S_+(f),$$

où on a utilisé le fait que pour toute fonction en escalier h, l'intégrale $\int_I h(x) \, dx = \int_I h \, d\lambda$. Si la fonction f est Riemann-intégrable, ces deux expressions coïncident. On a trouvé une paire de fonctions boréliennes h_∞ et $\tilde{h}_\infty$ qui encadrent f, et telles que $\int h_\infty \, d\lambda = \int \tilde{h}_\infty \, d\lambda$. Ceci impose que $\tilde{h}_\infty = h_\infty = f$ λ-p.p.. La proposition 4.8 montre alors que f est $\overline{\mathcal{B}(\mathbb{R})}$-mesurable. De plus, son intégrale de Lebesgue $\int f \, d\lambda = \int h_\infty \, d\lambda = S(f)$. □

4.4.4 Il existe un ensemble non mesurable pour la mesure de Lebesgue

On a vu que les ensembles de la tribu de Lebesgue peuvent être compliqués : ensembles $G_\delta, F_\sigma, G_{\delta\sigma}, \dots$ On a en plus rajouté aux ensembles boréliens les ensembles λ-négligeables, de manière à former la tribué complétée $\overline{\mathcal{B}(\mathbb{R})} = \mathcal{M}(\lambda^*)$. Il est légitime de se demander si cette tribu complétée, ou même la tribu des boréliens sur $\mathbb{R}$, ne contient déjà pas toutes les parties de $\mathbb{R}$.

Nous montrons que ce n'est pas le cas, en exhibant un sous-ensemble de $\mathbb{R}$ qui ne peut pas faire partie de cette tribu. Cet ensemble est « compliqué » : pour le construire on doit faire appel à l'axiome du choix de la théorie des ensembles.

Considérons l'espace $\mathbb{R}/\mathbb{Q}$ des classes d'équivalence des réels modulo les rationnels : $x \sim y$ ssi $x - y \in \mathbb{Q}$. Pour chaque classe d'équivalence $a \in \mathbb{R}/\mathbb{Q}$, on peut choisir un représentant x_a, qu'on peut prendre dans l'intervalle $[0, 1]$. C'est ce choix d'un représentant pour tout a, qui nécessite l'axiome du choix. En effet, $\mathbb{R}/\mathbb{Q}$ est indénombrable. On définit alors l'ensemble de ces représentats :

$$F := \{x_a \,;\, a \in \mathbb{R}/\mathbb{Q}\} \subset [0, 1],$$

qui est un ensemble indénombrable, puisque tous les x_a sont forcément différents.

Théorème 4.14. *L'ensemble F n'est pas contenu dans la tribu de Lebesgue $\overline{\mathcal{B}(\mathbb{R})}$.*

Démonstration. Supposons au contraire que F est mesurable pour la tribu de Lebesgue. Comme la tribu de Lebesgue est invariante par translation, pour tout rationnel q l'ensemble $(q + F)$ sera aussi mesurable. Comme tout $x \in \mathbb{R}$ appartient à une classe $a \in \mathbb{R}/\mathbb{Q}$, et que chaque classe regroupe tous les éléments $x_a + \mathbb{Q}$, on vérifie que

$$\bigcup_{q\in\mathbb{Q}} (q+F) = \mathbb{R}.$$

Si $\lambda(F) = 0$, alors par invariance par translation on aura pour tout rationnel $\lambda(q+F) = 0$, et par σ-additivité $\lambda(\mathbb{R}) = 0$, ce qui est absurde. Donc on a forcément $\lambda(F) = c > 0$. Comme $F \subset [0,1]$, on doit cependant avoir $0 < c \leq 1$. L'invariance par translation de λ impose alors que pour tout rationnel q, $\lambda(q+F) = c$.

Par ailleurs, les ensembles $(q+F)$ sont tous disjoints (si $q + x_a = q' + x_{a'}$ alors $x_a \sim x_{a'}$, donc $a = a'$ et $x_a = x_{a'}$). A partir de l'inclusion

$$\bigcup_{q\in\mathbb{Q}\cap[0,1]} (q+F) \subset [0,2]$$

on déduit par σ-additivité

$$\sum_{q\in\mathbb{Q}\cap[0,1]} \lambda(q+F) \leq \lambda\left([0,2]\right) = 2,$$

ce qui est impossible puisqu le membre de gauche vaudra $\infty \cdot c = \infty$.

On a donc une contradiction au fait que F est mesurable. Donc F ne peut pas être mesurable.

□

4.5 Construction plus générale d'une mesure

Dans cette section on généralisera la construction d'une mesure extérieure μ^* sur $\mathscr{P}(E)$, à partir d'une mesure μ définie sur une classe restreinte de parties de E. Cette classe restreinte de parties de E doit former une *semi-algèbre* sur un ensemble E. Les preuves sont laissées aux lecteurs, qui pourront s'inspirer des preuves données dans le cas de la mesure de Lebesgue.

Définition 4.15. Soit E un ensemble. Une semi-algèbre $\mathscr{S}$ sur E est un ensemble de parties de E, tel que :

(i) $\emptyset \in \mathscr{S}$;

(ii) $\mathscr{S}$ est stable par intersections finies ;

(iii) $\forall A \in \mathscr{S}$, il existe $B_1, \cdots, B_n$ des éléments disjoints de $\mathscr{S}$, tels que $A^c = \bigsqcup_{i=1}^n B_i$.

Remarquons que cette dernière propriété est une version affaiblie de l'invariance par passage au complémentaire.

Exemple 4.16. La mesure de Lebesgue est au départ définie sur les intervalles ouverts et fermés de $\mathbb{R}$: $\lambda\left(]a,b[\right) = \lambda\left([a,b]\right) = b-a$ pour tous réels $a<b$. L'ensemble des intervalles (ouverts et fermés) de $\mathbb{R}$ forme bien une semi-algèbre.

Définition 4.17. Une mesure définie sur une semi-algèbre $\mathscr{S}$ est une application $\mu : \mathscr{S} \to [0,\infty]$ telle que

1. $\mu(\emptyset) = 0$;

2. μ est σ-additive sur la semi-algèbre $\mathscr{S}$: si $A \in \mathscr{S}$ se décompose en $A = \bigsqcup_{i\in\mathbb{N}} B_i$ avec les $B_i \in \mathscr{S}$, alors $\mu(A) = \sum_{i\in\mathbb{N}} \mu(B_i)$.

La semi-algèbre $\mathscr{S}$ engendre naturellement une algèbre :

Définition 4.18. Soit E un ensemble. Une algèbre $\mathscr{A}$ sur E est un ensemble de parties de E tels que :

(i) $\emptyset \in \mathscr{A}$;

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

(iii) $\mathscr{A}$ est stable par union finie (et donc par intersection finie).

Si on ajoute à une sous-algèbre $\mathscr{S}$ toutes les unions finies de ses éléments, on obtient une algèbre, qu'on appelle l'algèbre engendrée par $\mathscr{S}$.

Remarque : La seule différence entre la notion d'algèbre et celle de tribu, est l'absence de stabilité par unions *dénombrables.*

Proposition 4.19. *Soit μ une mesure définie sur une sous-algèbre $\mathscr{S}$. Alors on peut étendre μ à l'algèbre $\mathscr{A}$ engendrée par $\mathscr{S}$, en posant :*

$$\forall A \in \mathscr{A},\ si\ A = \bigsqcup_{i=1}^{n} B_i \ (avec\ B_i \in \mathscr{S}\ \ disjoints), \qquad \mu(A) := \sum_{i=1}^{n} \mu(B_i).$$

La mesure μ étendue est alors σ-additive sur l'algèbre $\mathscr{A}$ (si $A = \bigsqcup_{i\in\mathbb{N}} B_i$ avec $A, B_i \in \mathscr{A}$, alors $\mu(A) = \sum_{i\in\mathbb{N}} \mu(B_i)$).

Exemple 4.20. L'algèbre engendrée par la semi-algèbre des intervalles de $\mathbb{R}$, est celle des unions finies d'intervalles de $\mathbb{R}$.

A partir d'une mesure σ-additive sur une algèbre $\mathscr{A}$, on va pouvoir construire une mesure extérieure.

Théorème 4.21. *Soit μ une mesure σ-additive sur une algèbre $\mathscr{A} \subset \mathscr{P}(E)$. Pour tout $A \in \mathscr{P}(E)$, posons*

$$\mu^*(A) := \inf \left\{ \sum_{i=0}^{\infty} \mu(A_i)\,;\, A_i \subset \mathscr{A},\, A \subset \bigcup_{i=0}^{n} A_i \right\}.$$

Alors μ^* définit une mesure extérieure (ous σ-sous-additive) sur E, qui prolonge μ.

Exemple 4.22. La mesure extérieure λ^* avait été définie en (4.5) de façon similaire, à partir de recouvrements de A par des unions dénombrables d'intervalles ouverts. On remarque que, partant de la mesure de Lebesgue sur la semi-algèbre des intervalles ouverts ou fermés, la définition ci-dessus inclura également comme possible A_i les intervalles fermés. Comme les singletons $\{a\}$ sont de mesure de Lebesgue nulle, cela ne change pas le résultat de λ^*.

Une fois qu'on dispose d'une mesure extérieure, la construction de la tribu $\mathcal{M}(\mu^*)$ et de la mesure $\mu = \mu^*_{\restriction \mathcal{M}(\mu^*)}$ a été présentée dans la section 4.1.