Intégration

Cours magistral, Semestre 5

Yulia Kuznetsova

30 novembre 2020

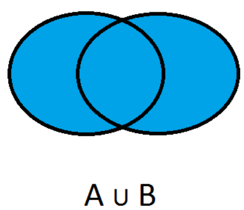
Table des matières

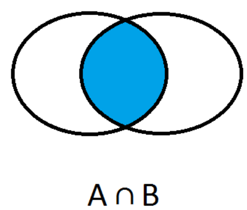
Chapitre 1

Rappels de la théorie des ensembles ; espaces mesurés

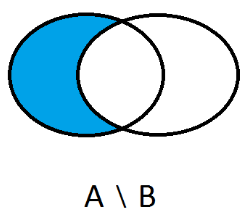
1.1 Opérations sur les ensembles

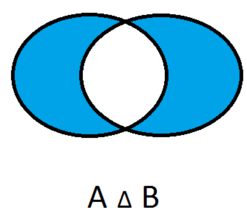
On connait bien les opérations arithmétiques sur des ensembles : la réunion $A \cup B$, l'intersection $A \cap B$,





la différence $A \setminus B$, la différence symétrique $A \Delta B = (A \setminus B) \cup (B \setminus A)$.



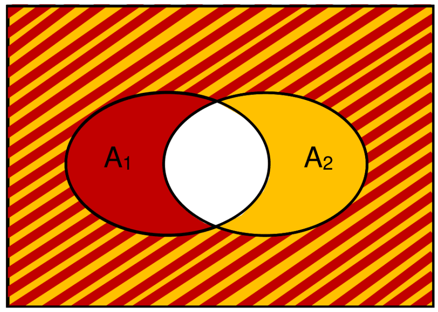


Si on fixe un ensemble «englobant» X et l'on considère seulement des sous-ensembles de X, on note par $A^c = X \setminus A$ le complément de A dans X.

Proposition 1.1.1. *Soit $(A_i)_{i \in I}$ une famille de sous-ensembles de X, alors :*

$$\left(\bigcup_{i \in I} A_i \right)^c = \bigcap_{i \in I} A_i^c; \qquad \left(\bigcap_{i \in I} A_i \right)^c = \bigcup_{i \in I} A_i^c.$$

Démonstration. On peut voir le cas de deux ensembles sur le dessin : A_1^c est présenté en jaune, A_2^c en rouge, et leur intersection est exactement le complément de $A_1 \cup A_2$.



Dans le cas général, la logique est la même : si $x \in \Big(\cup_{i\in I} A_i\Big)^c = X \setminus \cup_{i\in I} A_i$, alors $x \notin A_i$ pour tout i, c'est-à-dire $x \in A_i^c$; cela vaut pour tout i donc $x \in \cap_{i\in I} A_i^c$. Reciproquement, si $x \in \cap_{i\in I} A_i^c$ alors $x \notin A_i$ pour tout i, donc x n'appartient pas à la réunion $\cup_{i\in I} A_i$.

La démonstration de la deuxième égalité est laissée aux TDs. □

Cette propriété sera utilisée plusieurs fois dans nos démonstrations.

Exemple 1.1.2. $]\,0, +\infty) = \cup_n \big(\frac{1}{n}, +\infty\big)$ donc $(-\infty, 0\,] = \cap_n \big(-\infty, \frac{1}{n}\big)$.

Exemple 1.1.3. Pour toute fonction f à valeurs réelles,

$$\{x : f(x) \neq 0\} = \bigcup_{n\in\mathbb{N}^*} \{x : |f(x)| > 1/n\}$$

donc $\{x : f(x) = 0\} = \cap\{x : |f(x)| < 1/n\}$.

1.2 Cardinalité

Définition 1.2.1. Un ensemble A est dit **dénombrable** s'il existe une *bijection* $\varphi : \mathbb{N} \to A$. Dans ce cas-là, on peut énumérer les éléments de A, en posant $a_n = \varphi(n)$: $A = \{a_0, a_1, \alpha_2, \dots\}$.

Exemples :

1. L'ensemble $\mathbb{Z}$ des entiers :

$$\varphi(n) = \begin{cases} \frac{n}{2}, & n \text{ est pair;} \\ -\frac{n+1}{2}, & n \text{ est impair.} \end{cases}$$

2. L'ensemble $\mathbb{N}^2$ de paires de nombres naturels, qui est (par le précédent) en bijection avec l'ensemble $\mathbb{Z} \times \mathbb{N}$ de paires (m, n) où m est entier et n naturel. On construit $\varphi : \mathbb{N} \to \mathbb{N}^2$ comme suit : $\varphi(0) = (0, 0)$, $\varphi(1) = (0, 1)$, et ensuite on suit l'ordre indiqué ci-dessous : $\varphi(2) = (1, 0)$, $\varphi(3) = (2, 0)$ et ainsi de suite. Avec un certain effort, on aurait pû écrire une formule pour φ. Mais même sans formule il est évident que chaque paire

sera l'image d'un n naturel, et d'un seul.

$$\begin{tikzcd}
(0,0) \arrow[r] & (0,1) \arrow[dl] & (0,2) \arrow[r] & (0,3) \arrow[dl] & \cdots \\
(1,0) \arrow[d] & (1,1) \arrow[ur] & (1,2) \arrow[dl] & \cdots & \\
(2,0) \arrow[ur] & (2,1) \arrow[dl] & \cdots & & \\
(3,0) & \cdots & & & \\
\vdots & & & &
\end{tikzcd}$$

On peut parler d'ensembles quelconques :

Définition 1.2.2. Deux ensembles A et B sont dits **équipotents** s'il existe une bijection $\varphi : A \to B$. On le note $A \sim B$.

C'est une relatoin d'équivalence : elle est

- ⋆ reflexive : $A \sim A$ (l'application identité $\varphi : A \to A$ est bijective) ;
- ⋆ symétrique : si $A \sim B$ alors $B \sim A$ (si $\varphi : A \to B$ est bijective alors $\varphi^{-1} : B \to A$ l'est aussi) ;
- ⋆ transitive : si $A \sim B$ et $B \sim C$, alors $A \sim C$ (si $\varphi : A \to B$, $\psi : B \to C$ sont des bijections, alors $\psi \circ \varphi : A \to C$ l'est aussi).

On dit que les ensembles de la même classe d'équivalence ont la même *cardinalité*. On note $|A|$ la cardinalité de A, ce qui signifie, strictement dit, la classe d'équivalence de A. Pour un ensemble fini, $|A|$ est le nombre d'éléments dans A.

On peut comparer les cardinalités :

Définition 1.2.3. $|A| \leqslant |B|$ s'il existe une application **injective** $\varphi : A \to B$ (c'est-à-dire une bijection entre A et une partie de B).

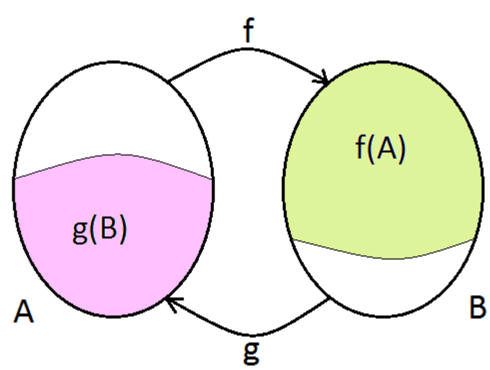
On peut démontrer que deux ensembles sont toujours comparables : soit $|A| \leqslant |B|$, soit $|B| \leqslant |A|$ (ou tous les deux, ce qui implique $|A| = |B|$, mais c'est un théorème qu'on va démontrer).

Exemple 1.2.4. $|\mathbb{Q}| \leqslant |\mathbb{R}|$ (évidemment). On montrera que l'inverse n'est pas vrai.

Théorème 1.2.5 (Cantor-Bernstein)**.** *Si* $|A| \leqslant |B|$ *et* $|B| \leqslant |A|$, *alors* $A \sim B$.

Démonstration. Par l'hypothèse, il existent des applications injectives $f : A \to B$ et $g : B \to A$.

À partir d'elles, on va construire une bijection $\varphi : A \to B$. Il est évident que sur $A \setminus g(B)$, il faut poser $\varphi(a) = f(a)$. Mais on ne peut pas définir $\varphi(a)$ par $g^{-1}(a)$ sur $g(B)$ car $b = g^{-1}(a)$ pourrait appartenir à $f\big(A \setminus g(B)\big)$, ce qui nous ferait perdre l'injectivité.

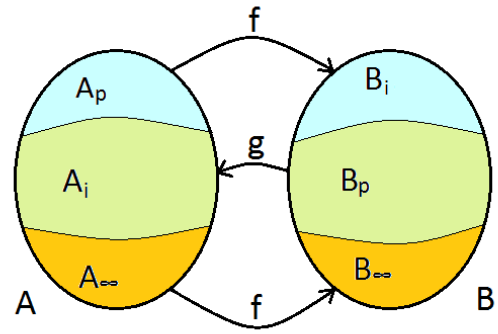


$$\varphi(a) = \begin{cases} f(a), & a \in A \setminus g(B) \\ ?, & a \in g(B) \end{cases}$$

Il faudra donc diviser A et B plus finement. Disons que $x \in A \cup B$ est un antécédent de $a \in A$ si $a = g \circ f \dots g \circ \dots (x)$, avec un nombre fini de fonctions f et g alternées. On définit de la même façon des antécédents pour $b \in B$.

Soit $N(a)$ le nombre d'antécédents de $a \in A$; autrement dit, c'est le nombre maximal de fonctions f ou g dans des représentations $a = g \circ f \dots g \circ \dots (x)$. Si $a \notin g(B)$, il n'a pas d'antécédents et on pose $N(a) = 0$. Si a en possède un nombre infini, on pose $N(a) = \infty$. Pour $b \in B$, le nombre $N(b)$ est défini de la même façon.

On peut donc diviser A en trois parties : $A_\infty = \{a \in A : N(a) = \infty\}$, $A_i = \{a \in A : N(a) \text{ est impair}\}$ et $A_p = \{a \in A : N(a) \text{ est pair}\}$; de même, on divise B en B_∞, B_p et B_i. On peut noter que l'ensemble $A \backslash g(B)$ où $N(a) = 0$ est contenu dans A_p.



Chaque application de f ou g augmente de 1 le nombre $N(a)$, respectivement $N(b)$: on voit donc que $N(f(a)) = N(a) + 1$ et $f(A_p) \subset B_i$. Reciproquement, si $b \in B_i$, alors $N(b)$ est impair donc non-nul, ce qui fait qu'il existe $a \in A$ tel que $f(a) = b$. Comme $N(b) = N(a)+1$, $N(a)$ est pair et $a \in A_p$. Il en suit $f(A_p) = B_i$; en échangeant A et B, on peut obtenir de la même manière $g(B_p) = A_i$. Si $N(a) = \infty$, on aura également $N(f(a)) = \infty$, et reciproquement, alors $f(A_\infty) = B_\infty$.

Les fonctions f, g étant injectives, ces égalités d'images impliquent que f est bijective sur $A_p \cup A_\infty$, et le même est vrai pour g^{-1} sur A_i. On peut alors compléter la définition de la fonction φ, et conclure qu'elle est bijective :

$$\varphi(a) = \begin{cases} f(a), & a \in A_\infty \cup A_p; \\ g^{-1}(a), & a \in A_i. \end{cases}$$

□

Exemple 1.2.6. Il est évident que $|\mathbb{N}| \leqslant |\mathbb{Q}|$, car on a au moins l'injection naturelle de $\mathbb{N}$ dans $\mathbb{Q}$. Au même temps, on peut représenter chaque nombre rationnel en forme $q = \dfrac{m}{n}$ avec $m \in \mathbb{Z}$ et $n \in \mathbb{N}$, en choisissant par exemple les nombres m, n de sorte que la fraction soit irréductible. Cela définit une application injective $\varphi : \mathbb{Q} \to \mathbb{Z} \times \mathbb{N}$. On avait montré que ce dernier ensemble est dénombrable, donc qu'il existe une bijection $\psi : \mathbb{Z} \times \mathbb{N} \to \mathbb{N}$. La composition $\psi \circ \varphi : \mathbb{Q} \to \mathbb{N}$ sera injective, donc $|\mathbb{Q}| \leqslant |\mathbb{N}|$. Par le théorème de Cantor–Bernstein, on conclut que $\mathbb{Q} \sim \mathbb{N}$, c'est-à-dire $\mathbb{Q}$ est dénombrable.

——————— (fin du cours 1) ———————

Notation 1.2.7. La cardinalité de $\mathbb{N}$ est notée $\aleph_0$ (aleph-0). On sait donc que $|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}| = \aleph_0$.

Formulons deux simple mais utiles propriétés :

Proposition 1.2.8. *Si A est un ensemble infini, alors $|A| \geqslant |\mathbb{N}|$.*

Démonstration. Par définition, il faut construire une application $\varphi : \mathbb{N} \to A$ injective. Par l'hypothèse, A est non-vide. Soit $a_0 \in A$ quelconque, posons $\varphi(0) = a_0$. Comme A est infini, $A \setminus \{a_0\}$ est nonvide ; on choisit $a_1 \in A \setminus \{a_0\}$ quelconque et on pose $\varphi(1) = a_1$. Si on a choisi de cette façon $\varphi(k)$, $k < n$, il est toujours possible de choisir $\varphi(n) \in A \setminus \{a_k : k < n\}$. L'application $\varphi : \mathbb{N} \to A$ obtenue par récurrence est injective par construction. □

Proposition 1.2.9. *Si A est dénombrable, alors toute partie infinie $B \subset A$ est dénombrable.*

Démonstration. On a $|B| \geqslant |\mathbb{N}|$ par la proposition précédente. Au même temps, $|B| \leqslant |A| = |\mathbb{N}|$. Par le théorème de Cantor-Bernstein, $|B| = |\mathbb{N}|$. □

Exemple 1.2.10. $\mathbb{R}$ n'est pas dénombrable, c'est-à-dire $|\mathbb{R}| > |\mathbb{Q}|$.

Démonstration. Rappelons que tout nombre réel peut être écrit en forme d'une fraction décimale, finie ou infinie. On continue les fractions finies par des zéros pour avoir des représentations infinie pour tous les nombres. (On sait que les rationnels sont exactement les nombres dont les écritures sont périodiques à partir d'un certain moment.)

Il faut noter aussi que tout nombre finissant par "$\ldots 999 \ldots$" a une autre écriture qui fint par "$\ldots 0000 \ldots$, mais ce n'est pas important pour la prevue.

Supossons maintenant que $\mathbb{R}$ est dénombrable. Sa partie $[0,1[$, qui est évidemment infinie, est donc aussi dénombrable. Il existe donc $\varphi : \mathbb{N} \to \mathbb{R}$ bijective. En passant à l'écriture décimale, on peut écrire $\varphi(n) = 0, a_{n0}a_{n1}\ldots$. On forme le tableau suivant :

$$\begin{array}{ccccc} a_{00} & a_{01} & a_{02} & a_{03} & \cdots \\ a_{10} & a_{11} & a_{12} & a_{13} & \cdots \\ a_{20} & a_{21} & a_{22} & a_{23} & \cdots \\ a_{30} & a_{31} & a_{32} & a_{33} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{array}$$

La n-ème ligne contient donc les chiffres décimaux de $\varphi(n)$, sans compter le 0 initial.

Pour tout $n \in \mathbb{N}$, soit maintenant $b_n \in \{0, 1, \ldots, 9\}$ quelconque, mais différent de a_{nn} et, pour éviter le cas de deux écritures pour le même nombre, inégal à 0 et 9. Le nombre $b = 0, b_0b_1b_2\ldots$ appartient à $[0,1[$ (on l'a choisi strictement inférieur à 1), alors il est égal à $\varphi(n)$ pour un certain n et doit être présent dans le tableau. Mais son n-ème chiffre est $b_n \neq a_{nn}$, et il n'a pas de "queue" de 9s ou de 0s, donc $b \neq \varphi(n)$ pour tout n. Cette contradiction démontre le théorème.

La preuve porte le nom de la méthode diagonale de Cantor. □

On note $\mathfrak{c} = |\mathbb{R}|$ (la cardinalité du continu).

En TDs, il sera démontré ce que $|\mathbb{R} \setminus \mathbb{Q}| = \mathfrak{c}$.

Notation 1.2.11. 2^X est l'ensemble de toutes les parties d'un ensemble X. (La notation s'explique si on considère X fini.)

On peut montrer que $|\mathbb{R}| = |2^{\mathbb{N}}|$ et, en généralisant le fait démontré ci-dessus, que $|2^X| > |X|$ pour tout X.

Définition 1.2.12. On dit qu'un ensemble A est *au plus dénombrable* s'il est fini ou dénombrable. C'est équivalent au fait que $|A| \leqslant |\mathbb{N}|$.

Proposition 1.2.13. *Soient $(A_n)_{n\in\mathbb{N}}$ des ensembles au plus dénombrables, alors $A = \cup_{n\in\mathbb{N}} A_n$ est au plus dénombrable.*

Démonstration. Nous avions déjà montré que $\mathbb{N}^2$, qu'on peut voir comme la réunion $\mathbb{N}^2 = \cup_n \{n\} \times \mathbb{N}$, est dénombrable. Le cas actuel est très similaire, il faudra seulement tenir compte du fait que A_n peuvent avoir d'éléments en commun.

Remplaçons-les par des ensembles deux à deux disjoints. Soit $B_0 = A_0$, et pour tout $n > 0$ posons $B_n = A_n \backslash \cup_{k<n} A_k$. On a $\cup_n B_n = \cup_n A_n$, et les ensembles

(B_n) sont deux à deux disjoints (si $n < m$ et $x \in B_n \cap B_m$, on aurait eu $x \in A_n$ et à la fois $x \notin \cup_{k<m} A_k \supset A_n$). Ils sont soit finis, soit dénombrables. Soient $\varphi_n : B_n \to \mathbb{N}$ injectives, alors on peut poser $\varphi(a) = (n, \varphi_n(a))$ pour $b \in B_n$. Cette application est bien définie car $a \in A$ appartient à un et un seul B_n.

Si $\varphi(a) = \varphi(b)$, alors a et b appartiennent au même B_n et $\varphi_n(a) = \varphi_n(b)$. L'application φ_n étant injective, il en suit $a = b$. On conclut que φ est injective, donc $|A| \leqslant \aleph_0$, c'est-à-dire A est au plus dénombrable. □

Corollaire 1.2.14. *Soient* $(A_n)_{n\in\mathbb{N}}$ *des ensembles dénombrables, alors* $A = \cup_{n\in\mathbb{N}} A_n$ *est dénombrable.*

Démonstration. On a déjà montré que A est au plus dénombrable. De plus, A_0 est contenu dans A et par l'hypothèse est infini, donc $|A| = \aleph_0$. □

1.3 Espaces mesurables

Définition 1.3.1. Une *tribu* (ou une σ-algèbre) sur un ensemble X est une partie $\mathcal{F} \subset 2^X$ telle que :

- ⋆ $X \in \mathcal{F}$;
- ⋆ si $A, B \in \mathcal{F}$ alors $A \setminus B \in \mathcal{F}$;
- ⋆ si $A_n \in \mathcal{F}$, $n \in \mathbb{N}$, alors $\bigcup\limits_{n=0}^{\infty} A_n \in \mathcal{F}$.

On remarque tout de suite qu'il en suit $\emptyset = X \setminus X \in \mathcal{F}$, et que $\mathcal{F}$ contient les réunions finies de ses éléments.

Exercice 1.3.2. Montrer que si $A, B \in \mathcal{F}$ alors $A^c = X \setminus A \in \mathcal{F}$ et $A \cap B \in \mathcal{F}$. (Utiliser l'identité $(A \cap B)^c = A^c \cup B^c$.)

Toujours avec les formules du complément, on peut montrer que $\mathcal{F}$ contient les intersections dénombrables de ses éléments.

Exemple 1.3.3. $\mathcal{F} = (\emptyset, X)$ est une tribu.

Exemple 1.3.4. La famille $\mathcal{G}$ de tous les ouverts de $\mathbb{R}$ n'est pas une tribu. La deuxième propriété n'est pas vérifiée : on a $]0, +\infty[\in \mathcal{G}$ mais $\mathbb{R} \setminus]0, +\infty[=]-\infty, 0] \notin \mathcal{G}$. La troisième propriété n'est pas vérifiée non plus : on a $]-1/n, 1/n[\in \mathcal{G}$ mais

$$\bigcap_{n\in\mathbb{N}^*} \left] \frac{1}{n}, \frac{1}{n} \right[= \{0\} = [0, 0] \notin \mathcal{G}.$$

Définition 1.3.5. Un *espace mesurable* est une paire $(X, \mathcal{F})$ d'un ensemble X et d'une tribu $\mathcal{F}$ sur X.

Exemple 1.3.6. On peut toujours poser $\mathcal{F} = 2^X$, donc $(X, 2^X)$ est un espace mesurable.

C'est sur des tribus qu'on définit des mesures, et les éléments de $\mathcal{F}$ sont des ensembles mesurables.

Exemple 1.3.7. Soit $\mathcal{F} = \{A \subset \mathbb{R} : A \text{ ou } A^c \text{ est au plus dénombrable}\}$. Alors $\mathcal{F}$ est une tribu.

Démonstration. On a $\mathbb{R}^c = \emptyset$ fini donc au plus dénombrable, alors $\mathbb{R} \in \mathcal{F}$.

Si $A, B \setminus \mathcal{F}$ et A est au plus dénombrable, alors $A \setminus B \subset A$ aussi. Si B^c est au plus dénombrable, alors $A \setminus B = A \cap B^c$ l'est aussi. Enfin, soit A^c et B au plus dénombrables. On a $(A \setminus B)^c = B \cup A^c$ au plus dénombrable, donc $(A \setminus B)^c \in \mathcal{F}$. Dans tous les trois cas, on conclut que $A \setminus B \in \mathcal{F}$.

Si $A_n \in \mathcal{F}$ pour $n \in \mathbb{N}$ et A_m^c est au plus dénombrable pour certain m, alors $(\cup_n A_n)^c = \cap_n A_n^c \subset A_m^c$ est au plus dénombrable donc $\cup_n A_n \in \mathcal{F}$. Si A_n est au plus dénombrable pour tout n, alors le même est vrai pour $\cup_n A_n$ donc $\cup_n A_n \in \mathcal{F}$. □

Exemple 1.3.8. En général, $\mathcal{F} = \{A \subset X : |A| \leqslant \aleph_0 \text{ ou } |A^c| \leqslant \aleph_0\}$ est une tribu, quel que soit X. Dans le cas $X = \mathbb{R}$, cette tribu est différente de 2^X (quelle partie de $\mathbb{R}$ elle ne contient pas ?).

——————— (fin du cours 2) ———————

Proposition 1.3.9. *L'intersection d'une famille quelconque de tribus est une tribu.*

Démonstration. $\mathcal{F} = \bigcap\limits_{i \in I} \mathcal{F}_i$. Les trois propriétés sont faciles à vérifier. □

Définition 1.3.10. La tribu *engendrée* par une famille $S \subset 2^X$ est l'intersection de toutes les tribus sur X qui contiennent S.

Comme $2^X \supset S$ est une tribu et en vertu de la proposition 1.3.9, la tribu engendrée existe toujours.

Cela nous donne l'une des tribus les plus importantes :

Définition 1.3.11. Soit X un espace métrique. La *tribu borélienne* est la tribu engendrée par la famille de tous les ouverts de X. Elle est notée $\mathcal{B}(X)$.

Exemple 1.3.12. En particulier, $\mathcal{B}(\mathbb{R})$ contient : tous les ouverts ; tous les fermés car ils sont les compléments des ouverts ; tous les intervales $[a, b[= [a, b] \setminus [b, +\infty[$ ou $]a, b]$; toutes les parties dénombrables de $\mathbb{R}$ car un seul point $\{a\}$ est un fermé et $A = \{a_0, a_1, \ldots\} = \cup_{n=0}^{\infty} \{a_n\}$.

1.4 Espaces mesurés

En simple mots, une mesure est une fonction additive de l'ensemble, une généralisation des notions de surface ou de volume. Pour $X = \mathbb{R}^n$, on voudrait avoir une mesure telle que

$$\mu([a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]) = (b_1 - a_1) \cdots (b_n - a_n).$$

Simple qu'elle paraisse, sa construction est assez longue et fera une partie importante de notre cours.

On voit tout de suite qu'il faut autoriser des valeurs infinies à μ car évidemment $\mu(\mathbb{R}) = \infty$. On introduit pour cela

Notation 1.4.1. On note $\mathbb{R}_+ = [0, +\infty[$ et $\overline{\mathbb{R}}_+ = [0, +\infty] = [0, +\infty[\cup \{+\infty\}$. On admet les règles formelles suivantes : $x + \infty = \infty + x = \infty + \infty = \infty$ pour tout $x \in \mathbb{R}_+$.

Définition 1.4.2. Soit $(X, \mathcal{F})$ un espace mesurable. Une *mesure* (σ-additive) est une application $\mu : \mathcal{F} \to \overline{\mathbb{R}}_+$ telle que

- $\star$ $\mu(\emptyset) = 0$;
- $\star$ pour toute famille $(A_n)_{n \in \mathbb{N}}$ d'ensembles $A_n \subset \mathcal{F}$ deux à deux disjoints,

$$\mu(\bigcup_{n=0}^{\infty} A_n) = \sum_{n=0}^{\infty} \mu(A_n),$$

Le triple $(X, \mathcal{F}, \mu)$ est appelé dans ce cas un *espace mesuré*.

(On pourrait considérer aussi des mesures finiment additives, ne permettant que des sommes finies d'ensembles.)

Remarque 1.4.3. Une mesure σ-additive est toujours finiment additive. Si on a une famille finie $A_0, \dots, A_n$ d'ensembles deux à deux disjoints, on peut poser $A_k = \emptyset$, $k > n$, et obtenir une famille infinie $(A_n)_{n \in \mathbb{N}}$ d'ensembles qui sont toujours deux à deux disjoints. On aura donc

$$\mu(\bigcup_{k=0}^{n} A_k) = \mu(\bigcup_{k=0}^{\infty} A_k) = \sum_{k=0}^{\infty} \mu(A_k) = \sum_{k=0}^{n} \mu(A_k).$$

Exemple 1.4.4. Soit $\mathcal{F} = 2^X$. Posons $\mu(A) = |A|$ pour tout A fini et $\mu(A) = \infty$ pour A infini. C'est une mesure σ-additive, appelée la mesure de comptage.

Exemple 1.4.5. Soit $\mathcal{F} = 2^X$ et $a \in X$. La *mesure de Dirac* concentrée en a est définie par

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

Remarque 1.4.6. La condition $\mu(\emptyset) = 0$ peut être remplacée par la condition que $\mu(A) < \infty$ pour au moins un ensemble $A \in \mathcal{F}$. En effet, dans ce cas on a $\mu(A) = \mu(A \cup \emptyset) = \mu(A) + \mu(\emptyset)$, car A et $\emptyset$ sont disjoints : $A \cap \emptyset = \emptyset$. Mais sans aucune de ses conditions on pourrait avoir une mesure qui est infinie partout, et qui est impossible de travailler avec.

On peut déduire les propriétés suivantes des mesures.

Proposition 1.4.7 (Monotonie)**.** *Soit $(X, \mathcal{F}, \mu)$ un espace mesuré. Si $A, B \in \mathcal{F}$ et $A \subset B$, alors $\mu(A) \leqslant \mu(B)$.*

Démonstration. On a $B = A \cup (B \setminus A)$, donc $\mu(B) = \mu(A) + \mu(B \setminus A)$. Les trois nombres étant finis ou infinis, il en suit $\mu(A) \leqslant \mu(B)$. $\square$

Proposition 1.4.8 (Continuité de la mesure)**.** *Soit $(X, \mathcal{F}, \mu)$ un espace mesuré. Soient $A_n \in \mathcal{F}$, $n \in \mathbb{N}$, tels que $A_n \subset A_{n+1}$ pour tout n (autrement dit, la famille $(A_n)_{n \in \mathbb{N}}$ est croissante). Alors*

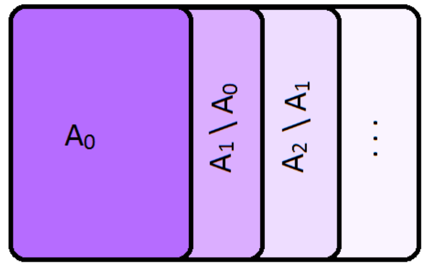
$$\mu(\bigcup_{n=0}^{\infty} A_n) = \lim_{n \to \infty} \mu(A_n).$$

Démonstration. Par le précédent, $\mu(A_n) \leqslant \mu(A_{n+1})$ et $\mu(A_n) \leqslant \mu(\cup_{n=0}^{\infty} A_n)$ pour tout n. La suite croissante $(\mu(A_n))$ tend vers sa limite $L = \sup_{n \in \mathbb{N}} \mu(A_n)$, finie si la suite est bornée ou infinie sinon. $\mu(\cup_{n=0}^{\infty} A_n)$ est un majorant de cette suite et il est donc $\geqslant L$. Dans le cas où $L = \infty$ cela termine la preuve.

Supposons alors que $L < \infty$. Pour tout n, on peut représenter A_n en tant que réunion disjointe

$$A_n = A_0 \cup (A_1 \setminus A_0) \cup (A_2 \setminus A_1) \cup \cdots \cup (A_n \setminus A_{n-1}).$$

De même,



$$\bigcup_{n=0}^{\infty} A_n = A_0 \cup \bigcup_{n=1}^{\infty} (A_n \setminus A_{n-1})$$

où les ensembles sont toujours deux à deux disjoints. Par la σ-additivité de μ,

$$\begin{aligned}
\mu\Big(\bigcup_{n=0}^{\infty} A_n\Big) &= \mu(A_0) + \sum_{n=1}^{\infty} \mu(A_n \setminus A_{n-1}) \\
&= \mu(A_0) + \lim_{m\to\infty} \sum_{n=1}^{m} \mu(A_n \setminus A_{n-1}) \\
&= \lim_{m\to\infty} \mu(A_m).
\end{aligned}$$

□

La reciproque est vraie aussi, dans la forme suivante :

Théorème 1.4.9. *Soit $(X, \mathcal{F})$ un espace mesurable. Soit $\mu : \mathcal{F} \to \overline{\mathbb{R}}_+$ une application telle que*

- $\mu(\emptyset) = 0$ *;*
- $\mu(A \cup B) = \mu(A) + \mu(B)$ *pour $A, B \in \mathcal{F}$ disjoints ;*
- *pour toute famille croissante $(A_n)_{n\in\mathbb{N}}$ d'ensembles $A_n \subset \mathcal{F}$*

$$\mu\big(\bigcup_{n=0}^{\infty} A_n\big) = \lim_{n\to\infty} \mu(A_n).$$

Alors μ est une mesure σ-additive.

Chapitre 2

Construction de la mesure de Lebesgue sur $\mathbb{R}^d$

2.1 Anneaux d'ensembles. La mesure des intervalles

La notion suivante est plus faible mais plus maniable que celle d'une tribu.

Définition 2.1.1. Un anneau d'ensembles est une famille $\mathcal{R} \subset 2^X$ de parties d'un ensemble X, telle que
- $\star$ $\emptyset \in \mathcal{R}$;
- $\star$ si $A, B \in \mathcal{R}$, alors $A \cap B \in \mathcal{R}$;
- $\star$ si $A, B \in \mathcal{R}$, alors $A \triangle B \in \mathcal{R}$.

Proposition 2.1.2. *Une tribu est un anneau d'ensembles.*

Proposition 2.1.3. *Dans $X = \mathbb{R}^d$, soit*

$$\mathcal{I} = \{ \rrbracket a_1, b_1 \llbracket \times \rrbracket a_2, b_2 \llbracket \times \cdots \times \rrbracket a_d, b_d \llbracket : k_k, b_k \in \mathbb{R} \},$$

où $\llbracket$ peut signifier l'un des symboles $]$ ou $[$, et soit

$$\mathcal{R} = \{ \cup_{k=1}^n A_k : A_k \in \mathcal{I}, \text{ deux à deux disjoints} \}.$$

Alors $\mathcal{R}$ est un anneau d'ensembles sur $\mathbb{R}^d$.

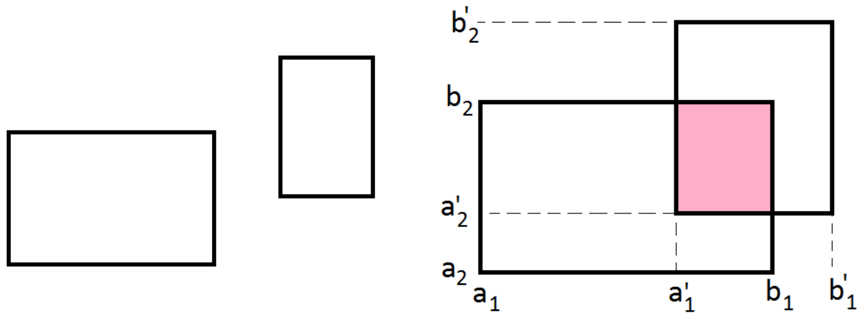
Démonstration. Il est évident que $\emptyset \in \mathcal{I} \subset \mathcal{R}$: il suffit de prendre $b_k < a_k$. Si $A = \bigsqcup_{k=1}^n A_k$ et $B = \bigsqcup_{l=1}^m B_l$, alors

$$A \cap B = \bigcup_{\substack{1 \leqslant k \leqslant n \\ 1 \leqslant l \leqslant m}} A_k \cap B_l.$$

Cette réunion reste disjointe [en détail : si $x \in A_k \cap B_l$ et $x \in A_{k'} \cap B_{l'}$, alors $x \in A_k \cap A_{k'}$ donc $k = k'$, et $x \in B_l \cap B_{l'}$ donc $l = l'$].

Pour chaque k, l on a $A_k \cap B_l \in \mathcal{I}$. Dans le cas $d = 2$ on le voit sur le dessin ci-dessous : soit les intervalles n'intersectent pas, soit

$$[a_1, b_1] \times [a_2, b_2] \cap [a_1', b_1'] \times [a_2', b_2'] = [\max(a_1, a_1'), \min(b_1, b_1')] \times [\max(a_2, a_2'), \min(b_2, b_2')].$$



Le cas $d \neq 2$ est similaire.

——————————— (fin du cours 3) ———————————

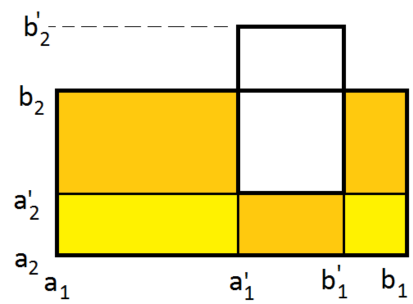
Ensuite, pour les mêmes A, B

$$A \triangle B = (A \setminus B) \sqcup (B \setminus A).$$

Si on montre que $A \setminus B \in \mathcal{R}$, alors de la même manière on aura $B \setminus A \in \mathcal{R}$, et comme ces deux ensembles sont dijsoints, il en suivra $A \triangle B \in \mathcal{R}$. Considérons alors

$$A \setminus B = \bigcup_{k=1}^{n} \Big(A_k \setminus \bigcup_{l=1}^{m} B_l \Big) = \bigcup_{k=1}^{n} \bigcap_{l=1}^{m} \Big(A_k \setminus B_l \Big).$$

Pour chaque k, l on a $A_k \setminus B_l \in \mathcal{R}$. Soit de nouveau $d = 2$: la différence $D = [a_1, b_1] \times [a_2, b_2] \setminus [a'_1, b'_1] \times [a'_2, b'_2]$ se décompose en réunion disjointe d'intervalles, dont le nombre dépend des positions relatives de a_j et b_j. Dans le cas présenté sur le dessin, on a $D = [a_1, a'_1] \times [a_2, a'_2] \sqcup]a'_1, b'_1] \times [a_2, a'_2[\sqcup \dots$ (5 intervalles au total).



Comme montré plus haut, l'intersection par l les laisse dans $\mathcal{R}$, et enfin la réunion par k est toujours disjointe (car A_k l'étaient). On conclut que $A \setminus B \in \mathcal{R}$. □

Définition 2.1.4. Pour $A =]\!]a_1, b_1]\!] \times]\!]a_2, b_2]\!] \times \cdots \times]\!]a_d, b_d]\!] \in \mathcal{I}$, posons

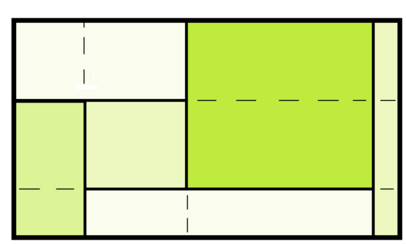
$$\lambda(A) = \prod_{k=1}^{d} (b_k - a_k).$$

Pour $A = \sqcup_{k=1}^{n} A_k \in \mathcal{R}$, posons $\lambda(A) = \sum_{k=1}^{n} \lambda(A_k)$.

Il faut bien évidemment vérifier que $\lambda(A)$ est bien défini, c'est-à-dire que si on peut décomposer

$$A = \sqcup_{k=1}^{n} A_k = \sqcup_{l=1}^{m} B_l$$

de deux façons différentes en tant que réunion d'intervalles, alors $\sum_k \lambda(A_k) = \sum_l \lambda(B_l)$. Supposons alors qu'on a deux telles décompositions. Les ensembles $A_k \cap B_l$ sont des intervalles et sont deux à deux disjoints ; pour chaque k, on a



$$A_k = \sqcup_{l=1}^{m} A_k \cap B_l$$

(voir le dessin). Le volume des intervalles est additif (on va le supposer connu, ou bien facile à vérifier), donc $\lambda(A_k) = \sum_l \lambda(A_k \cap B_l)$, d'où

$$\sum_k \lambda(A_k) = \sum_{k,l} \lambda(A_k \cap B_l).$$

Par symétrie, cette somme est aussi égale à $\sum_l \lambda(B_l)$, ce qui était à démontrer.

Par définition, $\lambda(A) \geqslant 0$ pour tout $A \in \mathcal{R}$.

Proposition 2.1.5. *λ est σ-additive, c'est-à-dire si $A_n \in \mathcal{R}$, $n \in \mathbb{N}$ sont deux-à-deux disjoints et $A = \bigsqcup_{n=1}^{\infty} A_n \in \mathcal{R}$, alors*

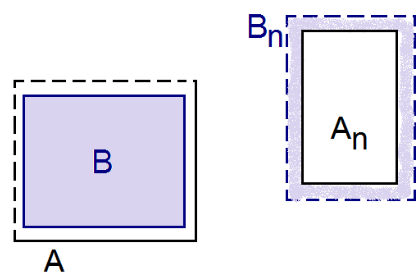
$$\lambda(A) = \sum_{n=1}^{\infty} \lambda(A_n).$$

Démonstration. 1) Montrons d'abord que λ est additive, c.à.d. la formule est vérifiée si le nombre de $A_n \in \mathcal{R}$ est fini. Pour chaque n, on a $A_n = \cup_{k=1}^{K_n} A_{nk}$ où A_{nk} sont des intervalles ; on peut noter qu'ils sont deux à deux disjoints. Par définition, comme $A = \sqcup A_{nk}$, on a

$$\lambda(A) = \sum_{n,k} \lambda(A_{nk}) = \sum_n \Big(\sum_k \lambda(A_{nk}) \Big) = \sum_n \lambda(A_n).$$

De l'additivité, il en suit aussi que si $A \subset \cup_{n=1}^{N} A_n$ alors $\lambda(A) \leqslant \sum_{n=1}^{N} \lambda(A_n)$ (il faut remplacer A_n par $A_n \cap A$ et).

2) En supposant $A = \bigsqcup_{n=1}^{\infty} A_n$, démontrons que $\lambda(A) \leqslant \sum_{n=1}^{\infty} \lambda(A_n)$. Pour tout $\varepsilon > 0$, on peut choisir $B \subset A$ fermé tel que $B \in \mathcal{R}$ et $\lambda(B) > \lambda(A) - \varepsilon$. (Dans chaque intervalle composant A, peut-être pas fermé, on inscrit un intervalle fermé en peu plus petit.) Pour tout n, on peut également choisir $B_n \supset A_n$ ouvert dans $\mathcal{R}$ tel que $\lambda(B_n) < \lambda(A_n) + \varepsilon/2^n$ (en entourant chaque intervalle d'un intervalle ouvert). On aura



$$B \subset \bigcup_{n \in \mathbb{N}} B_n$$

un recouvrement ouvert d'un fermé borné donc d'un compact ; on peut en choisir un sous-recouvrement fini : $B \subset \bigcup_{1 \leqslant n \leqslant N} B_n$. Par 1),

$$\lambda(B) \leqslant \sum_{n \in \mathbb{N}} \lambda(B_n),$$

et par leur choix

$$\lambda(A) < \lambda(B) + \varepsilon \leqslant \sum_{n=1}^{\infty} \Big(\lambda(A_n) + \frac{\varepsilon}{2^n} \Big) + \varepsilon = \sum_{n=1}^{\infty} \lambda(A_n) + 2\varepsilon.$$

Comme cela est vérifié avec tout $\varepsilon > 0$, l'inégalité est démontrée.

3) Pour tout N,

$$\bigsqcup_{n=1}^{N} A_n \subset A,$$

donc

$$\lambda\Big(\bigsqcup_{n=1}^{N} A_n \Big) = \sum_{n=1}^{N} \lambda(A_n) \leqslant \lambda(A),$$

d'où $\lambda(A) \geqslant \sum_{n \in \mathbb{N}} \lambda(A_n)$. Cela termine la preuve. $\square$

2.2 Théorème de Carathéodory

Définition 2.2.1. Soit $\mathcal{R}$ un anneau de sous-ensembles de X, et soit $\mu : \mathcal{R} \to \mathbb{R}_+$ une application σ-additive. Pour tout $E \subset X$, définissons sa *mesure extérieure* :

$$\mu^*(E) = \inf \Big\{ \sum_{n=1}^{\infty} \mu(A_n) : A_n \in \mathcal{R} \text{ et } E \subset \bigcup_{n=1}^{\infty} A_n \Big\}.$$

En général, μ^* n'est pas additive, mais on a la majoration suivante :

Proposition 2.2.2. *Si* $E \subset \bigcup_{n=1}^{\infty} E_n$*, alors*

$$\mu^*(E) \leqslant \sum_{n=1}^{\infty} \mu^*(E_n).$$

Démonstration. Le seul cas nontrivial est celui où $\mu^*(E_n) < \infty$ pour tout n. Soit $\varepsilon > 0$ fixe. Pour tout n, on peut recouvrir E_n par $\cup_k A_{nk}$ où $A_{nk} \in \mathcal{R}$ et

$$\sum_k \mu(A_{nk}) < \frac{\varepsilon}{2^n} + \mu^*(A_n).$$

On aura alors $E \subset \cup_{n,k} A_{n,k}$, ce qui est une famille dénombrable, et

$$\sum_{n,k} \mu(A_{nk}) < \sum_n \Big(\frac{\varepsilon}{2^n} + \mu^*(A_n) \Big) < \varepsilon + \sum_n \mu^*(A_n),$$

ce qui démontre que $\mu^*(E)$ est majorée par $\sum_n \mu^*(A_n)$. $\square$

Il en suit aussi que μ^* est monotone.

Définition 2.2.3. Soit $\mathcal{R}$ un anneau de sous-ensembles de X muni d'une application σ-additive $\mu : \mathcal{R} \to \mathbb{R}_+$. Un ensemble $E \subset X$ est dit *mesurable* si pour tout $\varepsilon > 0$ il existe $A \in \mathcal{R}$ tel que $\mu^*(E \triangle A) < \varepsilon$.

Evidemment, tout $A \in \mathcal{R}$ est mesurable ($E = A$). Le but est de montrer que les ensembles mesurables forment une tribu.

La construction qui suit se fait d'abord dans le cas d'une mesure finie. Dans le cas de $\mathbb{R}^d$, on considère maintenant $X = [0,1]^d$ ou tout autre cube fixe.

Proposition 2.2.4. *Soit $\mathcal{R}$ un anneau de sous-ensembles de X muni d'une application σ-additive $\mu : \mathcal{R} \to \mathbb{R}_+$. Supposons que $X \in \mathcal{R}$. Alors la famille $\mathcal{F}$ de tous les ensembles mesurables est une tribu sur X.*

Démonstration. Par l'assomption, $X \in \mathcal{R} \subset \mathcal{F}$.

Soient $E, F \in \mathcal{F}$. Pour $\varepsilon > 0$ donné, soient $A, B \in \mathcal{R}$ tels que $\mu^*(E \triangle A) < \varepsilon$ et $\mu^*(F \triangle B) < \varepsilon$. On cherche à montrer que $E \setminus F$ est "presque égal" à $A \setminus B$. On a $E \subset A \cup (E \setminus A)$, alors

$$E \setminus F \subset \big(A \cup (E \setminus A)\big) \setminus F \subset (A \setminus F) \cup (E \setminus A).$$

Ensuite,

$$A \setminus F \subset \big(B \cup (A \setminus B)\big) \setminus F \subset (B \setminus F) \cup (A \setminus B).$$

Donc

$$E \setminus F \subset (A \setminus B) \cup (B \setminus F) \cup (E \setminus A) = (A \setminus B) \cup C$$

avec $\mu^*(C) < 2\varepsilon$. De même (en échangeant A avec E et B avec F),

$$A \setminus B \subset (E \setminus F) \cup D$$

avec $\mu^*(D) < 2\varepsilon$. Finalement,

$$(E \setminus F) \triangle (A \setminus B) \subset C \cup D$$

où $\mu^*(C \cup D) < 4\varepsilon$, ce qui prouve que $E \setminus F$ est mesurable.

De façon similaire, on démontre que si $E = \cup_n E_n$ et pour tout $\varepsilon > 0$ on choisit $A_n \in \mathcal{R}$ tels que $\mu^*(E_n \triangle A_n) < \varepsilon/2^n$, alors $\mu^*(E \triangle (\cup_n A_n)) < 4\varepsilon$, donc E est mesurable. (Démontré en TDs ?) □

Théorème 2.2.5. *Soit $\mathcal{R}$ un anneau de sous-ensembles de X muni d'une application σ-additive $\mu : \mathcal{R} \to \mathbb{R}_+$. Supposons que $X \in \mathcal{R}$. Alors la famille $\mathcal{F}$ de tous les ensembles mesurables est une tribu sur X et μ^* est une mesure (σ-additive) sur $\mathcal{F}$ qui prolonge μ.*

——————————— (fin du cours 4) ———————————

Démonstration. On a déjà vérifié ce que μ^* prolonge μ et que $\mathcal{F}$ est une tribu. Il reste alors à vérifier ce que μ^* est une mesure.

Par définition, $\mu^*(E) \geqslant 0$ pour tout E, et comme μ^* prolonge μ, $\mu^*(\emptyset) = 0$. Montrons d'abord que μ^* est additive dans les cas de deux ensembles. Soient $E, F \in \mathcal{F}$, $E \cap F = \emptyset$; pour $\varepsilon > 0$, soient $A, B \in \mathcal{R}$ tels que $\mu^*(E \triangle A) < \varepsilon$ et $\mu^*(F \triangle B) < \varepsilon$.

On a toujours $\mu^*(E \cup F) \leqslant \mu^*(E) + \mu^*(F)$, la difficulté est dans l'inégalité inverse. Comme $E \subset A \cup (A \setminus E) \subset A \cup (E \triangle A)$, on a par le choix de A

$$\mu^*(E) \leqslant \mu^*(A) + \mu^*(E \triangle A) < \mu(A) + \varepsilon,$$

et de même $\mu^*(F) < \mu(B) + \varepsilon$. Alors

$$\mu^*(E) + \mu^*(F) < \mu(A) + \mu(B) + 2\varepsilon = \mu(A \cup B) + \mu(A \cap B) + 2\varepsilon.$$

On a $A \cup B \subset E \cup (A \setminus E) \cup F \cup (B \setminus F)$, donc

$$\mu(A \cup B) \leqslant \mu^*(E \cup F) + \mu^*(A \setminus E) + \mu^*(B \setminus F) < \mu^*(E \cup F) + 2\varepsilon.$$

Ensuite, comme les ensembles E, F sont disjoints, tout $x \in X$ est dans le complément soit de E, soit de F, donc

$$A \cap B \subset \big((A \cap B) \setminus E\big) \cup \big((A \cap B) \setminus F\big) \subset (A \setminus E) \cup (B \setminus F),$$

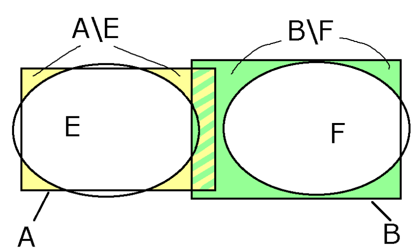
d'où $\mu^*(A \cap B) < 2\varepsilon$.

Finalement, nous arrivons à

$$\mu^*(E) + \mu^*(F) < \mu^*(E \cup F) + 6\varepsilon,$$

et comme ε est arbitrairement petit, on conclut que $\mu^*(E \cup F) \geqslant \mu^*(E) + \mu^*(F)$.

Par récurrence, on peut montrer l'additivité pour un nombre fini quelconque d'ensembles disjoints.



Soient maintenant $E = \bigsqcup_{n=1}^{\infty} E_n$, où $E_n, E \in \mathcal{F}$. On a vu que $\mu^*(E) \leqslant \sum_n \mu^*(E_n)$. Comme μ^* est monotone,

$$\mu^*(E) \geqslant \mu^*\Big(\bigcup_{n=1}^{N} E_n\Big) = \sum_{n=1}^{N} \mu^*(E_n)$$

pour tout $N \geqslant 1$ (l'additivité finie est déjà démontrée). Donc

$$\mu^*(E) \geqslant \sum_{n=1}^{\infty} \mu^*(E_n)$$

et finalement $\mu^*(E) = \sum_{n=1}^{\infty} \mu^*(E_n)$. $\square$

Définition 2.2.6. La mesure sur la tribu des ensembles mesurables dans $\mathbb{R}^d$ qui prolonge la mesure standard des intervalles s'appelle *la mesure de Lebesgue* sur $\mathbb{R}^d$. On la note λ.

Exemple 2.2.7. Soit $A = \cup_{n=1}^{\infty} [n, n+2^{-n}]$. Alors A est mesurable est $\lambda(A) = \sum_n 2^{-n} = 1$.

Remarque 2.2.8. On peut montrer que toutes les parties de $\mathbb{R}^n$ ne sont pas mesurables.

Exemple 2.2.9 (Vitali)**.** Un ensemble non-mesurable peut être construit comme suit. Pour $x, y \in [0,1]$, disons que $x \sim y$ si $x - y \in \mathbb{Q}$. On vérifie que c'est une relation d'équivalence, et $[0,1]$ se décompose alors en réunion disjointe de classes d'équivalence. Soit V un ensemble qui contient exactement un représentant de toute classe d'équivalence. [L'existence de V suit de l'axiome du choix, l'un des axiomes fondamentaux de la théorie des ensembles.]

Pour tout $x \in [0,1]$, par le choix il existe $y \in E$ et $q \in \mathbb{Q}$ tels que $x = q + y$, alors

$$[0,1] \subset \bigcup_{q \in \mathbb{Q}} (V+q) = \bigcup_{q \in \mathbb{Q}} \{y + q : y \in V\}.$$

Par la construction de la mesure de Lebesque, la translation d'un ensemble mesurable est encore mesurable et de même mesure. Si on suppose que V est mesurable, on a alors $\lambda(V+q) = \lambda(V)$ pour tout $q \in \mathbb{Q}$. La réunion est dénombrable, alors

$$1 = \lambda([0,1]) \leqslant \sum_{q \in \mathbb{Q}} \lambda(V),$$

d'où $\lambda(V) > 0$. Au même temps,

$$[0,2] \supset \sum_{q \in \mathbb{Q} \cap [0,1]} (q+V),$$

donc

$$2 = \lambda([0,2]) \geqslant \sum_{q \in \mathbb{Q} \cap [0,1]} \lambda(V).$$

Mais cette série a les termes identiques strictement positifs donc elle diverge et sa somme ne peut pas être majorée par 2. Cette contradiction montre que V n'est pas mesurable.

Chapitre 3

Fonctions mesurables. Convergence presque partout

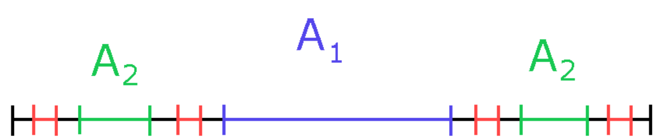
3.1 Ensembles négligeables

Définition 3.1.1. Soit $(X, \mathcal{F}, \mu)$ un espace mesuré. Un ensemble $A \subset X$ est dit *négligeable* s'il est contenu dans un ensemble de mesure nulle : il existe $B \in \mathcal{F}$ tel que $A \subset B$ et $\mu(B) = 0$.

En particulier, tout ensemble de mesure nulle est négligeable.

Exemple 3.1.2 (L'ensemble de Cantor)**.** Pour tout $n \geqslant 1$, soit A_n la réunion de 2^{n-1} intervalles de longueur 3^{-n} (voir dessin) :

$$\begin{aligned} A_1 &=]\frac{1}{3}, \frac{2}{3}[\\ A_2 &=]\frac{1}{9}, \frac{2}{9}[\cup]\frac{7}{9}, \frac{8}{9}[\\ &\dots \end{aligned}$$



Alors $A = \cup A_n$ est mesurable de mesure

$$\lambda(A) = \sum_{n=1}^{\infty} 2^{n-1} 3^{-n} = \frac{1}{3} \sum_{k=0}^{\infty} 2^k 3^{-k} = \frac{1}{3} \cdot \left(1 - \frac{2}{3}\right)^{-1} = 1.$$

Son complément $C = [0,1] \setminus A$ est appelé l'ensemble de Cantor et a la mesure nulle : $\lambda(C) = \lambda([0,1]) - \lambda(A) = 0$. Il est donc négligeable. On sait aussi que C est fermé, non-vide (contient toutes les bornes des intervalles de A_n) et possède plusieurs autres propriétés intéressantes. Il est en particulier non-dénombrable.

Proposition 3.1.3. *1. Si A est négligeable et $B \subset A$, alors B est négligeable.*

2. Si A_n est négligeable, $n \in \mathbb{N}$, alors $A = \cup_n A_n$ est négligeable.

Démonstration. Pour le deuxième, on a $B_n \in \mathcal{F}$, $\mu(B_n) = 0$; donc $B = \cup_n B_n \in \mathcal{F}$, $A \subset B$ et $\mu(B) \leqslant \sum_n \mu(B_n) = 0$. $\square$

Tout ensemble négligeable n'est pas mesurable, cela dépend de la tribu :

Exemple 3.1.4. Soit $X = \mathbb{R}$, $\mathcal{F} = \{\emptyset, \mathbb{R}, [0,1], \mathbb{R} \setminus [0,1]\}$. C'est une tribu ; on pose $\mu(\mathbb{R}) = \mu([0,1]) = 1$ et sinon 0. L'ensemble $[-1, 0[$ est alors négligeable mais non mesurable.

Définition 3.1.5. Soit $(X, \mathcal{F}, \mu)$ un espace mesuré. La mesure μ est dite *complète* si tout ensemble négligeable est mesurable.

Proposition 3.1.6. *Si la mesure $\mu = \nu^*$ sur $\mathcal{F}$ est construite par l'extension de Carathéodory d'une mesure ν (définie sur un anneau d'ensembles $\mathcal{R}$), alors μ est complète. C'est en particulier le cas de la mesure de Lebesgue sur $\mathbb{R}^n$.*

Démonstration. Effectivement, si $A \subset B$ et $\mu(B) = \nu^*(B) = 0$, alors $\nu^*(A) = 0$ car une mesure extérieure est monotone. Alors pour tout $\varepsilon > 0$ on a $\nu^*(A \triangle \emptyset) = \nu^*(A) = 0$ avec $\emptyset \in \mathcal{R}$ donc A est mesurable, et $\mu(A) = \nu^*(A) = 0$. $\square$

Chaque mesure peut être *complétée* : on élargit la tribu $\mathcal{F}$ en y ajoutant tous les ensembles négligeables et leurs compléments. De façon similaire à l'exemple 1.3.8, on montre que c'est bien une tribu.

On peut montrer qu'il existent des ensembles négligeables non-boréliens dans $\mathbb{R}^n$.

——————————— (fin du cours 5) ———————————

3.2 Fonctions mesurables

Rappel : si f est une fonction sur X, alors $f^{-1}(A) = \{x \in X : f(x) \in A\}$.

La continuité est d'habitude définie par des limites, mais on peut montrer l'équivalence suivante :

Proposition 3.2.1. *Une fonction $f : \mathbb{R}^n \to \mathbb{R}$ est continue si et seulement si $f^{-1}(U)$ est ouvert pout tout ouvert $U \subset \mathbb{R}$; et si et seulement si*

$$f^{-1}\big(]a, b[\big) = \{x : a < f(x) < b\}$$

est ouvert pour tous $a, b \in \mathbb{R}$.

Démonstration. Soit $f : \mathbb{R}^n \to \mathbb{R}$ continue. Pour un ouvert $U \subset \mathbb{R}$ et $x \in f^{-1}(U)$ on a $y = f(x) \in U$, alors il existe un intervalle ouvert autour de y contenu dans U ; on peut donc trouver $\varepsilon > 0$ tel que $\mathcal{O}_\varepsilon(y) =]y - \varepsilon, y + \varepsilon[\subset U$. Par continuité, il existe $\delta > 0$ tel que $f(\mathcal{O}_\delta(x)) \subset \mathcal{O}_\varepsilon(y) \subset U$ (il serait plus familier de l'écrire à l'aide des inégalités : si $\|x - x'\| < \delta$, alors $|y - f(x')| < \varepsilon$). On a alors $\mathcal{O}_\delta(x) \subset f^{-1}(U)$, ce qui montre que $f^{-1}(U)$ contient une boule ouverte autour de chaqun de ses points alors il est ouvert.

Reciproquement, soit f telle que $f^{-1}(U)$ est ouvert pour tout $U \subset \mathbb{R}$ ouvert. Pour $x \in \mathbb{R}^n$ soit $y = f(x)$. Pour chaque $\varepsilon > 0$ le voisinage $U = \mathcal{O}_\varepsilon(y)$ est ouvert, alors $f^{-1}(U)$ est ouvert et contient une boule ouverte $\mathcal{O}_\delta(x)$ avec $\delta > 0$. On a $f(\mathcal{O}_\delta(x)) \subset \mathcal{O}_\varepsilon(y)$; c'est la définition de la continuité en x.

Enfin, si cette propriété de f n'est vérifiée que pour des intervalles $]a, b[$, on peut rappeler que tout ouvert $U \subset \mathbb{R}$ est une réunion d'intervalles ouverts : $U = \cup_\alpha]a_\alpha, b_\alpha[$. L'égalité

$$f^{-1}\Big(\bigcup_a]a_\alpha, b_\alpha[\Big) = \bigcup_\alpha f^{-1}\big(]a_\alpha, b_\alpha[\big)$$

montre que $f^{-1}(U)$ est ouvert, ce qui termine la preuve. $\square$

On verra alors bien la logique de la définition suivante :

Définition 3.2.2. Soit $(X, \mathcal{F})$ un espace mesurable. Une fonction $f : X \to \mathbb{R}$ est dite *mesurable* si $f^{-1}(U) \in \mathcal{F}$ (est mesurable) pour tout ouvert $U \subset \mathbb{R}$.

En fait, on peut à la place considérer les images inverses de tous les boréliens :

Proposition 3.2.3. *Soit $(X, \mathcal{F})$ un espace mesurable. Une fonction $f : X \to \mathbb{R}$ est mesurable si et seulement si $f^{-1}(A) \in \mathcal{F}$ pour tout borélien $A \subset \mathbb{R}$.*

Démonstration. La seule implication nontriviale est de supposer que f est mesurable et de montrer la mesurabilité de $f^{-1}(A)$ pour tout borélien A. Soit

$$\mathcal{G} = \{A \subset \mathbb{R} : f^{-1}(A) \in \mathcal{F}\}.$$

Par l'hypothèse, $\mathcal{G}$ contients tous les ouverts. On vérifie que c'est une tribu : $\mathbb{R} \in \mathcal{G}$ car il est ouvert ; si $A, B \in \mathcal{G}$ alors $f^{-1}(A \setminus B) = f^{-1}(A) \setminus f^{-1}(B) \in \mathcal{F}$ donc $A \setminus B \in \mathcal{G}$, et enfin si $A_n \in \mathcal{G}$, $n \in \mathbb{N}$, alors $f^{-1}(\cup A_n) = \cup_n f^{-1}(A_n) \in \mathcal{F}$ et $\cup A_n \in \mathcal{G}$.

Par définition, $\mathcal{B}(\mathbb{R})$ est la plus petite tribu qui contient les ouverts, donc $\mathcal{G} \supset \mathcal{B}(\mathbb{R})$, ce qui était à démontrer. □

On pourrait considérer, plus généralement, des applications mesurables :

Définition 3.2.4. Soient $(X, \mathcal{F})$ et $(Y, \mathcal{G})$ deux espaces mesurables. On dit que $f : X \to Y$ est mesurable si $f^{-1}(A) \in \mathcal{F}$ pour tout $A \in \mathcal{G}$.

La définition en haut est donc le cas particulier où on considère $Y = \mathbb{R}$ et $\mathcal{G} = \mathcal{B}(\mathbb{R})$. Un autre cas souvent considéré est celui de $Y = \mathbb{C}$:

Définition 3.2.5. On identifie $\mathbb{C}$ avec $\mathbb{R}^2$ de façon standard : $J : \mathbb{C} \to \mathbb{R}^2$, $J(z) = (\mathrm{Re} z, \mathrm{Im} z)$. On définit la tribu borélienne sur $\mathbb{C}$ comme

$$\mathcal{B}(\mathbb{C}) = \{J^{-1}(A) : A \in \mathcal{B}(\mathbb{R}^2)\}.$$

Une fonction $f : X \to \mathbb{C}$ est dite *mesurable* si $f^{-1}(A) \in \mathcal{F}$ pour tout ensemble borélien $A \in \mathcal{B}(\mathbb{C})$.

La plupart des propositions de cette section sont valables pour $\mathbb{C}$ aussi.

Proposition 3.2.6. *Soit $(X, \mathcal{F})$ un espace mesurable. Pour qu'une fonction $f : X \to \mathbb{R}$ soit mesurable, il est nécessaire et suffisant que*

$$f^{-1}(]-\infty, a[) = \{x \in X : f(x) < a\} \in \mathcal{F}$$

pour tout ouvert $a \in \mathbb{R}$.

Démonstration. Nécessité est évidente. Supposons que $f^{-1}(]-\infty, a[)$ est mesurable pour tout $a \in \mathbb{R}$. Comme $]a, b[=]-\infty, b[\setminus]-\infty, a]$ et $]-\infty, a] = \cap_{n \in \mathbb{N}^*}]-\infty, a + 1/n[$, on voit que $f^{-1}(]a, b[)$ est mesurable pour tous $a, b \in \mathbb{R}$. Enfin, tout ouvert $U \subset \mathbb{R}$ est la réunion d'intervalles ouverts dont on peut supposer les bornes rationnelles, donc $f^{-1}(U)$ est mesurable. □

Proposition 3.2.7. *Si $\mathcal{F}$ contient la tribu borélienne, alors toute fonction continue est mesurable.*

Démonstration. Pour tout ouvert $U \subset \mathbb{R}$, $f^{-1}(U)$ est ouvert donc dans $\mathcal{F}$. □

Exemple 3.2.8. On peut donner une multitude d'exemple de fonctions mesurables sur $\mathbb{R}$: $f(x) = x^2$, $g(x) = e^x$,...

Exemple 3.2.9. Soit C l'ensemble de Cantor et soit

$$f(x) = \begin{cases} 1, & x \in C \\ 0, & \text{sinon.} \end{cases}$$

Alors f est mesurable (car $f^{-1}(A)$ est toujours l'un des ensembles C, $\mathbb{R} \setminus C$, $\emptyset$ ou $\mathbb{R}$), mais discontinue.

Exemple 3.2.10. En général, soit I_E la fonction indicatrice d'un ensemble $E \subset X$:

$$I_E(x) = \begin{cases} 1, & x \in E \\ 0, & \text{sinon.} \end{cases}$$

Alors I_E est mesurable si et seulement si E est mesurable.

On peut vérifier que

$$f^{-1}(]-\infty, a[) = \begin{cases} \emptyset, & a \leqslant 0 \\ X \setminus E, & 0 < a \leqslant 1 \\ X, & a > 1. \end{cases}$$

Les fonctions mesurables forment une classe large. Certaines propriétés sont les mêmes que des fonctions continues :

Proposition 3.2.11. *Soient $(X, \mathcal{F}, \mu)$, $(Y, \mathcal{G}, \nu)$ et $(Z, \mathcal{H}, \lambda)$ des espaces mesurés. Si $f : X \to Y$ et $g : Y \to Z$ sont mesurables, alors $g \circ f : X \to Z$ est mesurable.*

Démonstration. Pour $A \in \mathcal{H}$, on a $g^{-1}(A) \in \mathcal{G}$ et donc

$$(g \circ f)^{-1}(A) = f^{-1}\big(g^{-1}(A)\big) \in \mathcal{F}.$$

$\square$

Proposition 3.2.12. *Soit $(X, \mathcal{F}, \mu)$ un espace mesuré. Si $f, g : X \to \mathbb{R}$ sont mesurables, alors :*

1. *cf est mesurable pour tout $c \in \mathbb{R}$;*
2. *$f + g$, $f - g$ sont mesurables ;*
3. *fg est mesurable ;*
4. *si $g(x) \neq 0$ pour tout x, alors f/g est mesurable.*

Démonstration.

1. Pour $c = 0$, on a $cf = 0$ est le fait est évident. Si $c \neq 0$, on a $(cf)^{-1}(A) = f^{-1}(A/c) \in \mathcal{F}$ car la famille des ouverts est stable sous dilatations, donc la tribu borélienne aussi.
2. Pour $a \in \mathbb{R}$,

$$E_a = \{x : f(x) + g(x) < a\} = \{x : f(x) < a - g(x)\}.$$

Pour tout $x \in E_a$, il existe $q \in \mathbb{Q}$ tel que $f(x) < q < a - g(x)$, et à l'inverse, si un tel q existe, alors $x \in E_a$. Donc

$$E_a = \bigcup_{q \in \mathbb{Q}} \{x : f(x) < q\} \cap \{x : q < a - g(x)\}$$

Pour chaque q, les deux ensembles à droite sont mesurables (le deuxième est $g^{-1}(]-\infty, a - q[)$), alors leur réunion (dénombrable !) est mesurable aussi.

Ensuite, $f - g = f + (-g)$.

3. On peut écrire $fg = \frac{1}{4}\big((f+g)^2 - (f-g)^2\big)$, où $(f \pm g)^2$ sont mesurables par composition ; le reste suit des propriétés déjà démontrées.
4. Comme $\frac{f}{g} = f \cdot \frac{1}{g}$, il suffit de montrer que $\frac{1}{g}$ est mesurable. Mais c'est la composée de $x \mapsto 1/x$, $\mathbb{R} \setminus \{0\} \to \mathbb{R}$, avec g.

□

Le fait suivant montre la différence importante entre les fonctions mesurables et continues :

Proposition 3.2.13. *Soient $f_n, n \in \mathbb{N}$, des fonctions mesurables, telles que pour tout $x \in X$ il existe $\lim_{n\to\infty} f_n(x) =: f(x)$. Alors f est mesurable.*

Démonstration. Pour $a \in \mathbb{R}$, $f(x) < a$ si et seulement s'il existe $q < a$ (qu'on peut supposer rationnel) et $n \in \mathbb{N}$ tels que $f_k(x) \leqslant q$ pour $k \geqslant n$. Alors

$$\{x : f(x) < a\} = \bigcup_{q\in\mathbb{Q}:q<a} \bigcup_{n\in\mathbb{N}} \bigcap_{k\geqslant n} \{x : f_k(x) < q\}.$$

Pour chaque q, k l'ensemble $\{x : f_k(x) < q\}$ est mesurable, donc leurs intersections et les réunions sont mesurables aussi. □

Exemple 3.2.14. Soient $f_n(x) = x^n$, $x \in [0,1]$. Alors

$$\lim_{n\to\infty} f_n(x) = \begin{cases} 0, & x \in [0,1[; \\ 1, & x = 1. \end{cases}$$

C'est une fonction discontinue mais mesurable.

3.3 Convergence presque partout

Définition 3.3.1. Soient $f, g : X \to \mathbb{R}$. On dit qu'elles sont égales presque partout, notation : $f \sim g$ ou $f \underset{p.p.}{=} g$ si l'ensemble $\{X : f(x) \neq g(x)\}$ est négligeable.

——— (fin du cours 6) ———

C'est une relation d'équivalence. On dit également “presque partout” pour “partout sur X sauf peut-être un ensemble négligeable”.

Exemple 3.3.2. La fonction indicatrice $I_{\mathbb{Q}}$ de l'ensemble des rationnels est presque partout nulle.

Le théorème de convergence peut être renforcé comme suit :

Théorème 3.3.3. *Soit $(X, \mathcal{F}, \mu)$ un espace mesuré. Soient $f_n, n \in \mathbb{N}$, des fonctions mesurables, f une fonction sur X, telle que pour presque tout $x \in X$ il existe $\lim_{n\to\infty} f_n(x) = f(x)$. On dit que la suite (f_n) converge presque partout vers f. Alors f est équivalente à une fonction mesurable. Si μ est complète, on peut conclure que f est mesurable.*

Démonstration. Soit

$$A = \{x \in X : \lim_{n\to\infty} f_n(x) \text{ existe et vaut } f(x)\}.$$

Par l'hypothèse, A^c est négligeable donc il existe $B \subset X$ de mesure nulle tel que $A^c \subset B$. Sur A (qui est un espace mesuré avec la mesure héritée) $f_n \to f$ partout, donc $f|_A$ est mesurable. On pose $f|_B := 0$. Pour $0 \leqslant a \in \mathbb{R}$ on a

$$\{x \in X : f(x) < a\} = \{x \in A : f(x) < a\},$$

et pour $a > 0$

$$\{x \in X : f(x) < a\} = B \cup \{x \in A : f(x) < a\}.$$

On conclut que f changée ainsi est mesurable, et elle est équivalente à f initiale. Si μ est complète, f est mesurable sans changement car dans la formule

$$\{x \in X : f(x) < a\} = \{x \in A : f(x) < a\} \cup \{x \in A^c : f(x) < a\}$$

le deuxième ensemble est négligeable donc de mesure nulle. □

Théorème 3.3.4 (Yegorov)**.** *Soit* (f_n) *une suite de fonctions mesurables sur* $(X, \mathcal{F}, \mu)$ *qui converge presque partout vers une fonction* f. *Alors pour tout* $\delta > 0$ *il existe* $E_\delta \subset X$ *mesurable tel que*

- ⋆ $\mu(X \setminus E_\delta) < \delta$;
- ⋆ $f_n \to f$ *uniformément sur* E_δ.

Démonstration. Démontré en TDs. □

3.4 Fonctions étagées

Supposons que $(X, \mathcal{F}, \mu)$ est un espace mesuré.

Définition 3.4.1. Une fonction $f : X \to \mathbb{R}$ est dite *étagée* si elle est mesurable et l'ensemble de ses valeurs est fini.

Autrement dit, il existent $n \in \mathbb{N}$, $y_1, \ldots, y_n \in \mathbb{R}$, $A_1, \ldots, A_n \in \mathcal{F}$ tels que $f|_{A_k} = y_k$ et $X = \sqcup_k A_k$. Si on note par I_E la fonction indicatrice d'un ensemble $E \subset X$, alors f s'écrit

$$f = \sum_{k=1}^{n} y_k I_{A_k}, \qquad \sqcup_k A_k = X. \tag{3.1}$$

(N'oublions pas que A_k sont disjoints.) Évidemment f est mesurable si A_k sont mesurables. Pour l'implication inverse, il faut supposer que y_k sont distincts ; dans ce cas-là,

$$A_k = f^{-1}\big(\{y_k\}\big),$$

ce qui oblige A_k à être mesurables. On a montré donc que f est étagée si et seulement si elle admet une représentation (3.1).

Proposition 3.4.2. *Une fonction* $f : X \to \mathbb{R}$ *est mesurable si et seulement s'il existe une suite* (f_n) *de fonctions étagées qui converge vers* f *partout. Si* f *est bornée,* (f_n) *peuvent être choisies de sorte qu'elles convergent uniformément.*

Démonstration. Comme toute fonction étagée est mesurable, "seulement si" est évident. Supposons alors que f est mesurable. Pour $n \in \mathbb{N}$, $-n^2 < m < n^2$ soit

$$A_{n,m} = \{x \in X : \frac{m}{n} \leqslant f(x) < \frac{m+1}{n}\};$$

ensuite,

$$A_{n,-n^2} = \{x \in X : f(x) < -n\}, \qquad A_{n,n^2} = \{x \in X : f(x) \geqslant n\}.$$

Ce sont des ensembles mesurables. On pose

$$f_n(x) = \sum_{m=-n^2}^{n^2} \frac{m}{n} I_{A_{n,m}}.$$

Si n_0 est tel que $|f(x)| < n_0$, on a alors pour tout $n \geqslant n_0$

$$|f(x) - f_n(x)| \leqslant \frac{1}{n}.$$

Il en suit que $f_n(x) \to f(x)$ pour tout x, est si f est bornée, la convergence est uniforme. $\square$

Remarque 3.4.3. Si f est positive, la suite $\varphi_k \to f$ de fonctions étagées peut être choisie croissante. En effet, soit $\varphi_k = f_{2^k}$ comme définie dans la preuve précédente, avec $n = 2^k$. Les ensembles $A_{n,-n^2}$ seront vides car $f \geqslant 0$. Si $x \in A_{n,n^2}$, alors par construction $\varphi_k(x) = f_{2^k}(x) = 2^k > 2^{k-1} \geqslant f_{2k-1}(x) = \varphi_{k-1}(x)$. Pour $x \in A_{n,m}$ on a $\frac{m}{2^k} \leqslant f(x) < \frac{m+1}{2^k}$; si m est pair, alors

$$\frac{m/2}{2^{k-1}} \leqslant f(x) < \frac{(m/2)+1}{2^{k-1}} \text{ et } f_{k-1}(x) = \frac{m/2}{2^{k-1}} = \frac{m}{2^k} = f_{2^k}(x).$$

Si m est impair, alors

$$\frac{(m+1)/2-1}{2^{k-1}} \leqslant f(x) < \frac{(m+1)/2}{2^{k-1}} \text{ et } f_{k-1}(x) = \frac{(m-1)/2}{2^{k-1}} < \frac{m}{2^k} = f_{2^k}(x).$$

Dans tous les deux cas, nous avons $\varphi_{k-1}(x) = f_{2^{k+1}}(x) \leqslant f_{2^k}(x) = \varphi_k(x)$, ce qui était à démontrer.

Chapitre 4

Intégrale de Lebesgue

4.1 L'intégrale des fonctions étagées

Supposons que $(X, \mathcal{F}, \mu)$ est un espace mesuré avec une mesure complète finie μ (c.à.d. $\mu(X) < \infty$). On définit d'abord l'intégrale des fonctions étagées :

Définition 4.1.1. Soit $f = \sum_{k=1}^{n} y_k I_{A_k}$ une fonction étagée, $X = \sqcup_k A_k$. L'intégrale de Lebesgue de f est définie comme

$$\int_X f d\mu := \sum_{k=1}^{n} y_k \mu(A_k).$$

Proposition 4.1.2. *L'intégrale de dépend pas de décomposition de f en somme* (3.1).

Démonstration. Supposons que f admet deux représentations :

$$f = \sum_{k=1}^{n} y_k I_{A_k} = \sum_{l=1}^{m} z_l I_{B_l}.$$

En passant à des réunions de A_k et B_l si nécessaire, on peut admettre que $\{y_k\}$ sont distincts (et respectivement $\{z_l\}$). Effectivement, soient $k_1, \dots, k_r$ tel que $\{y_k : 1 \leqslant k \leqslant n\} = \{y_{k_j} : 1 \leqslant j \leqslant r\}$. On a donc

$$f = \sum_{j=1}^{r} y_{k_j} \sum_{k : y_k = y_{k_j}} I_{A_k} = \sum_{j=1}^{r} y_{k_j} I_{\cup\{A_k : y_k = y_{k_j}\}},$$

et

$$\int_X f d\mu = \sum_{j=1}^{r} y_{k_j} \sum_{k : y_k = y_{k_j}} \mu(A_k) = \sum_{j=1}^{r} y_{k_j} \mu\big(I_{\cup\{A_k : y_k = y_{k_j}\}}\big).$$

Maintenant, si $A_k \cap B_l \neq \emptyset$, alors la valeur de f y est $y_k = b_l$, donc pour chaque k c'est possible pour au plus un seul l ; et pour au moins un car $A_k \subset X = \cup B_l$. Il en suit que pour chaque k il existe l_k tel que $A_k \subset B_{l_k}$. En appliquant un raisonnement symétrique, on déduit qu'en fait $A_k = B_{l_k}$; dans ce cas évidemment $y_k = z_{l_k}$, et donc

$$\sum_{k=1}^{n} y_k \mu(A_k) = \sum_{l=1}^{m} z_l \mu(B_l).$$

□

Proposition 4.1.3. *Si f, g sont des fonctions étagées, alors*

1. $\int_X cf d\mu = c \int_X f d\mu$ *pour tout* $c \in \mathbb{R}$ *;*
2. $\int_X (f \pm g) d\mu = \int_X f d\mu \pm \int_X g d\mu$ *;*
3. *Si* $f \geqslant 0$*, alors* $\int_X f d\mu \geqslant 0$.
4. *Si* $|f| \leqslant M$*, alors* $\left| \int_X f d\mu \right| \leqslant M\mu(X)$.

Démonstration. 1. Si $f = \sum_{k=1}^n y_k I_{A_k}$, alors $cf = \sum_{k=1}^n c y_k I_{A_k}$, d'où l'égalité.

2. On peut toujours choisir des ensembles C_k de telle façon que f et g se décompose en
$$f = \sum_{k=1}^n y_k I_{C_k}, \qquad g = \sum_{k=1}^n z_k I_{C_k}.$$
(passer à $A_j \cap B_l$). Ensuite,
$$f \pm g = \sum_{k=1}^n (y_k \pm z_k) I_{C_k}$$
et
$$\int_X (f \pm g) d\mu = \sum_{k=1}^n (y_k \pm z_k)\mu(C_k) = \int_X f d\mu \pm \int_X g d\mu.$$
3. Si $f \geqslant 0$, alors
$$\sum_{k=1}^n y_k \mu(A_k) \geqslant 0.$$
4. Si $|f| \leqslant M$, alors
$$\Big| \sum_{k=1}^n y_k \mu(A_k) \Big| \leqslant \sum_{k=1}^n M\mu(A_k) = M\mu(X).$$

□

Exemple 4.1.4. $\int_X 1 d\mu = 1 \cdot \mu(X) = \mu(X)$.

Exemple 4.1.5. $\int_X I_A d\mu = \mu(A)$.

——————————— (fin du cours 7) ———————————

4.2 Définition de l'intégrale

Nous supposons toujours que $(X, \mathcal{F}, \mu)$ est un espace mesuré avec une mesure complète finie μ (c.à.d. $\mu(X) < \infty$).

Définition 4.2.1. Soit $f : X \to \mathbb{R}_+$ une fonction mesurable positive. On pose
$$\int_X f d\mu = \sup \Big\{ \int_X g d\mu : \ g \text{ étagée, } 0 \leqslant g \leqslant f \Big\}.$$
On dit que f est intégrable si $\int_X f d\mu < +\infty$.

On voit immédiatement que si f est étagée, les deux définitions coïncident.

Définition 4.2.2. Pour toute fonction $f : X \to \mathbb{R}$, on définit ses parties positive et négative f_+, f_- :

$$f_+(x) = \begin{cases} f(x), & f(x) \geqslant 0; \\ 0, & \text{sinon.} \end{cases}$$

On pose $f_- = f_+ - f$, en obtenant $f = f_+ - f_-$. Toutes les deux fonctions f_+, f_- sont positives, et

$$f_-(x) = \begin{cases} -f(x), & f(x) < 0; \\ 0, & \text{sinon.} \end{cases}$$

Si f est mesurable, f_+ et f_- le sont aussi. (Il suffit de remarquer que $f_+ = f \cdot I_{\{x:f(x)\geqslant 0\}}$.)

Définition 4.2.3. Soit $f : X \to \mathbb{R}$ mesurable. On dit que f est intégrable si toutes les deux f_+, f_- sont intégrables, et dans ce cas on pose

$$\int_X f d\mu = \int_X f_+ d\mu - \int_X f_- d\mu.$$

Remarque 4.2.4. Dans l'intégrale de Lebesgue, il n'y a pas de fonctions intégrables qui ne soient pas absolument intégrables, c'est la différence avec l'intégrale de Riemann impropre (exemple : $f(x) = \sin(x)/x$ sur $[0, +\infty)$, ou, dans le cas de mesure finie, $f(x) = \frac{1}{x}\sin(\frac{1}{x})$ sur $]0, 1]$).

Remarque 4.2.5. Si f est bornée positive, disons $0 \leqslant f \leqslant M$, alors pour toute fonction étagée $0 \leqslant g \leqslant f$ on aura aussi $0 \leqslant g \leqslant M$; par les propriétés de l'intégrale des fonctions étagées, on a $0 \leqslant \int_X g d\mu \leqslant M\mu(X)$. Il en suit que l'intégrale de f, qui est par définition la borne supérieure de $\int_X g d\mu$, est aussi bornée par $M\mu(X)$, donc f est intégrable. En décomposant f arbitraire *bornée* en sa partie positive et négative, on conclut que **toute fonction mesurable bornée est intégrable** (sur un espace de mesure finie).

On peut rappeler que sur un segment dans $\mathbb{R}$, une fonction est Riemann intégrable si et seulement si l'ensemble de ses points de discontinuité est négligeable (c'est un résultat de Lebesgue) ; autrement dit, elle est presque partout continue. La fonction $I_{\mathbb{Q}}$ n'est pas intégrable sur $[0, 1]$ par exemple. Dans le sens de Lebesgue, elle l'est ; son intégrale est $\int_{[0,1]} I_{\mathbb{Q}} d\lambda = \lambda(\mathbb{Q} \cap [0, 1]) = 0$.

Proposition 4.2.6. *Soient f et g des fonctions mesurables.*

1. *Si $0 \leqslant f \leqslant M$, où $M \geqslant 0$ est une constante, alors $0 \leqslant \int_X f d\mu \leqslant M\mu(X)$.*
2. *Si $f \underset{p.p.}{=} 0$, alors $\int_X f d\mu = 0$.*
3. *Si $f \underset{p.p.}{\geqslant} 0$, alors $\int_X f d\mu \geqslant 0$.*
4. *Si $0 \leqslant f \underset{p.p.}{\leqslant} g$ et g est intégrable, alors f est intégrable et $\int_X f d\mu \leqslant \int_X g d\mu$.*
5. *Si $f \underset{p.p.}{=} g$, alors f et g sont intégrables ou pas simultanément, et $\int_X f d\mu = \int_X g d\mu$.*

Démonstration. 1. Déjà démontré.

2. Dans la décomposition $f = f_+ - f_-$, on a $f_+ \underset{p.p.}{=} 0$ et $f_- \underset{p.p.}{=} 0$, toutes les deux étant positives. L'ensemble $A_+ = \{x : f_+(x) > 0\}$ est donc de mesure nulle. Si g est une fonction étagée et $0 \leqslant g \leqslant f_+$, alors g peut être non-nulle sur l'ensemble A_+ seulement (dans sa décomposition en combinaison linéaire de fonctions charactéristiques, la valeur 0 correspond à une partie dont le complément a la mesure nulle), donc son intégrale est nulle. Cela implique que $\int_X f_+ d\mu = 0$ comme la borne supérieure des intégrales de g étagées. De même, $\int_X f_- d\mu = 0$, d'où $\int_X f d\mu = 0$.
3. La condition $f \underset{p.p.}{\geqslant} 0$ implique que $f \underset{p.p.}{=} f_+$ et $f_- \underset{p.p.}{=} 0$. On a $\int_X f_+ d\mu \geqslant 0$ par 1, et $\int_X f_- d\mu = 0$ par 2, donc $\int_X f d\mu \geqslant 0$.
4. Soit $0 \leqslant h \leqslant f$ une fonction étagée. On peut la changer si nécéssaire sur l'ensemble $A = \{x : f(x) > g(x)\}$ qui a la mesure nulle, pour obtenir une fonction $\tilde{h}$ étagée telle que $0 \leqslant \tilde{h} \leqslant g$; on peut par exemple poser $\tilde{h}|_A = 0$. On peut remarquer que
$$\int_X h d\mu = \int_X \tilde{h} d\mu \leqslant \int_X g d\mu;$$
ceci étant vrai pour toute h, on conclut que $\int_X f d\mu \leqslant \int_X g d\mu$. En particulier, si g est intégrable, c'est-à-dire son intégrale est finie, alors f est intégrable aussi.
5. On a $f_\pm \underset{p.p.}{=} g_\pm$, il suffit alors de considérer des fonctions positives. Par symétrie, il suffit de supposer que g est intégrable. Par 4., f est alors intégrable et $\int_X f d\mu \leqslant \int_X g d\mu$. En appliquant 4. à $f \geqslant g$, on obtient $\int_X f d\mu \geqslant \int_X g d\mu$, donc les intégrales sont égales.

□

Corollaire 4.2.7. *Si on change une fonction mesurable sur un ensemble négligeable, cela ne change pas le fait d'être intégrable ou pas, ni son intégrale.*

4.3 Lemme de Fatou

Lemme 4.3.1 (Fatou)**.** *Soient $f_n \geqslant 0$ mesurables, $f \underset{p.p.}{=} \varliminf_{n\to\infty} f_n$. Alors*
$$\int_X f d\mu \leqslant \varliminf_{n\to\infty} \int_X f_n d\mu,$$
la limite étant finie ou infinie.

Démonstration. Si on change f sur l'ensemble négligeable $\{x : f(x) \neq \varliminf_{n\to\infty} f_n(x)\}$, cela ne changera pas son intégrale ni l'intégrabilité ; on peut alors supposer que $f = \varliminf_{n\to\infty} f_n$ partout. En particulier, on a $f \geqslant 0$.

On ne fait pas la preuve du fait que f est mesurable, mais cette preuve serait similaire au cas de la limite simple.

Soit g une fonction étagée, telle que $0 \leqslant g \leqslant f$. Fixons $\varepsilon > 0$. Pour tout $x \in X$, on a par l'hypothèse
$$g(x) \leqslant f(x) \leqslant \varliminf_{n\to\infty} f_n(x).$$

Par conséquence, il existe $m(x) \in \mathbb{N}$ tel que $f_n(x) > g(x) - \varepsilon$ pour tout $n \geqslant m(x)$. Pour tout m naturel, soit $B_m = \{x : m(x) \leqslant m\}$, alors $X = \cup_m B_m$ et les ensembles (B_m) sont croissants. De plus, ils sont mesurables car

$$B_m = \bigcap_{n \geqslant m} \{x : f_n(x) - g(x) > -\varepsilon\}$$

et les fonctions $f_n - g$ sont mesurables. On définit la fonction suivante :

$$h_m(x) = \begin{cases} g(x) - \varepsilon, & x \in B_m \\ 0, & x \notin B_m \end{cases}$$

Elle est étagée (car g l'est, et h_m a peut-être une valeur de plus par rapport à g), et par construction $h_m \leqslant f_n$ si $n \geqslant m$. Il en suit que

$$\int_X h_m d\mu \leqslant \inf_{n \geqslant m} \int_X f_n d\mu \leqslant \varliminf_{n \to \infty} \int f_n d\mu.$$

Ensuite, sur B_m, on a $g(x) \leqslant h_m(x) + \varepsilon$, et sur $X \setminus B_m$ la fonction g est majorée par $\|g\|_\infty$, alors

$$g \leqslant h_m + \varepsilon + \|g\|_\infty I_{X \setminus B_m}.$$

En passant aux intégrales, on obtient

$$\int_X g d\mu \leqslant \int_X h_m d\mu + \varepsilon\mu(X) + \|g\|_\infty \mu(X \setminus B_m),$$

donc

$$\int_X g d\mu \leqslant \varliminf_{n \to \infty} \int f_n d\mu + \varepsilon\mu(X) + \|g\|_\infty \mu(X \setminus B_m), \tag{4.1}$$

et ceci pour tout m. On a remarqué que $X = \cup_m B_m$ et les ensembles (B_m) sont croissants, ce qu'on peut reformuler en disant que les ensembles $(X \setminus B_m)$ sont décroissants de l'intersection vide. Par la continuité de la mesure, il en suit que

$$0 = \mu(\emptyset) = \lim_{m \to \infty} \mu(X \setminus B_m).$$

En passant à la limite par m dans la formule (4.1), on conclut que

$$\int_X g d\mu \leqslant \varliminf_{n \to \infty} \int f_n d\mu + \varepsilon\mu(X).$$

En choisissant ε arbitrairement petit, on en déduit $\int_X g d\mu \leqslant \varliminf_{n \to \infty} \int f_n d\mu$. La partie à droite est donc un majorant pour toutes les intégrales de g, alors elle majore aussi leur borne supérieure, et $\int_X f d\mu \leqslant \varliminf_{n \to \infty} \int f_n d\mu$. $\square$

Remarque 4.3.2. On peut noter que f n'est pas forcément intégrable, mais elle pourrait être non-intégrable seulement dans le cas où les intégrales de f_n tendent vers $+\infty$.

Théorème 4.3.3 (Convergence majorée, cas positif)**.** *Soient $f_n \geqslant 0$ mesurables telles que $f_n \leqslant f$ pour tout n et $f_n \underset{p.p.}{\to} f$. Alors*

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

Démonstration. Par le lemme de Fatou,

$$\int_X f d\mu \leqslant \varliminf_{n\to\infty} \int_X f_n d\mu.$$

Au même temps, $f \geqslant f_n$ implique $\int_X f d\mu \geqslant \int_X f_n d\mu$ pour tout n (Proposition 4.2.6), alors

$$\int_X f d\mu \geqslant \varlimsup_{n\to\infty} \int_X f_n d\mu.$$

Cela implique que la suite des intégrales admet une limite et elle est égale à l'intégrale de f. On peut noter que toutes les deux peuvent être finies ou infinies. $\square$

On n'a pas encore démontré l'additivité de l'intégrale, ce sera dans la section 4.4 plus bas. Mais le corollaire suivant, qui l'utilise, a sa meilleure place ici ; la section 4.4 ne l'utilise évidemment pas.

Corollaire 4.3.4. *Soient $f_n \geqslant 0$ des fonctions mesurables, telles que la série $\sum\limits_{n\in\mathbb{N}} f_n$ converge presque partout vers la fonction f. Alors*

$$\int_X f d\mu = \sum_{n=0}^{\infty} \int_X f_n d\mu,$$

la somme étant finie ou infinie.

Démonstration. On peut admettre que $f \geqslant 0$ (partout et non seulement presque partout). Pour $N \in \mathbb{N}$, on pose

$$g_N(x) = \sum_{n=0}^{N} f_n(x)$$

si $f(x) = \sum f_n(x)$ et $g_N(x) = 0$ sinon. Ces fonctions sont mesurables, $0 \leqslant g_N \leqslant f$ et $f \underset{p.p.}{=} \lim_{N\to\infty} g_N$. Par le Théorème 4.3.3,

$$\int_X f d\mu = \lim_{N\to\infty} \int_X g_N d\mu = \lim_{N\to\infty} \sum_{n=0}^{N} \int_X f_n d\mu = \sum_{n=0}^{\infty} \int_X f_n d\mu,$$

la somme étant finie ou infinie. $\square$

——————— (fin du cours 8) ———————

4.4 Propriétés de l'intégrale I

Théorème 4.4.1. *Une fonction mesurable $f \geqslant 0$ est intégrable si et seulement s'il existe une suite (f_n) de fonctions étagées telle que $0 \leqslant f_n \leqslant f$, $f_n \to f$ et*

$$\sup_{n\in\mathbb{N}} \int_X f_n d\mu < \infty. \tag{4.2}$$

Dans ce cas, $\displaystyle\int_X f d\mu = \lim_{n\to\infty} \int_X f_n d\mu.$

Démonstration. Si f est mesurable, il existe une suite croissante de fonctions étagées f_n qui converge vers f. Comme $0 \leqslant f_n \leqslant f$, on a $\int_X f_n d\mu \leqslant \int_X f d\mu$, ce qui démontre (4.2).

Reciproquement, si $f_n \to f$ vérifient les assomptions du théorème, alors en particulier $f = \lim f_n \leqslant \underline{\lim} f_n$ et on peut appliquer le lemme de Fatou :

$$\int_X f d\mu \leqslant \underline{\lim}_{n\to\infty} \int_X f_n d\mu \leqslant \sup_n \int_X f_n d\mu < \infty,$$

alors f est intégrable. Comme $f_n \leqslant f$ pour chaque n, on a aussi $\sup_n \int_X f_n d\mu \leqslant \int_X f d\mu$. La suite $a_n = \int_X f_n d\mu$ est alors majorée par $a = \int_X f d\mu$, mais au même temps $\underline{\lim}_{n\to\infty} a_n \geqslant a$; cela signifie qu'en réalité il existe $\lim_{n\to\infty} a_n = a$, ce qui termine la preuve. □

Proposition 4.4.2.

1. *Si f est intégrable, alors pour tout $c \in \mathbb{R}$ la fonction cf est intégrable, et*
$$\int_X cf d\mu = c \int_X f d\mu.$$
2. *Si f, g sont* **positives** *et intégrables, alors $f + g$ est intégrable et*
$$\int_X (f+g) d\mu = \int_X f d\mu + \int_X g d\mu.$$
3. *Si $0 \leqslant |f| \underset{p.p.}{\leqslant} g$ et g est intégrable, alors f est intégrable et $\left| \int_X f d\mu \right| \leqslant \int_X g d\mu$.*
4. *Si f est intégrable, alors $|f|$ aussi, et $\left| \int_X f d\mu \right| \leqslant \int_X |f| d\mu$.*
5. *Si f est intégrable et $\int_X |f| d\mu = 0$, alors $f \underset{p.p.}{=} 0$.*

Démonstration.

1. Si $c = 0$ le fait est évident. On peut admettre $c > 0$, le cas $c < 0$ se démontre de façon similaire. On a $(cf)_\pm = cf_\pm$. Par Théorème 4.4.1, il existe une suite (f_n) de fonctions étagées telles que $0 \leqslant f_n \leqslant f_+$ et $f_n \to 0$. Il en suit $0 \leqslant cf_n \leqslant cf_+$ et $cf_n \to cf_+$, donc cf_+ est intégrable (par le même théorème) et
$$\int_X cf_+ d\mu = \lim_{n\to\infty} \int_X cf_n d\mu = c \lim_{n\to\infty} \int_X f_n d\mu = c \int_X f_+ d\mu.$$
De même, on obtient
$$\int_X cf_- d\mu = c \int_X f_- d\mu$$
d'où le résultat.
2. On approche f et g par des suites (f_n), (g_n) de fonctions étagées selon le Théorème 4.4.1. On a $0 \leqslant f_n + g_n \leqslant f + g$ et, en passant à la limite,
$$\int_X (f+g) d\mu = \int_X f d\mu + \int_X g d\mu.$$
3. Ce fait a été démontré dans la Proposition 4.2.6 dans le cas $f \geqslant 0$. Dans le cas général, on représente $f = f_+ - f_-$, ce qui implique $|f| = f_+ + f_-$. Par l'assomption $0 \leqslant |f| \leqslant g$, alors $|f|$ est intégrable et son intégrale est majorée par $\int_X g d\mu$. Comme $0 \leqslant f_\pm \leqslant |f|$, les fonctions f_+, f_- sont intégrables elles aussi, et $\int_X f d\mu = \int_X f_+ d\mu - \int_X f_- d\mu$. Les intégrales de $f_\pm$ sont positives, alors
$$\left| \int_X f d\mu \right| \leqslant \int_X f_+ d\mu + \int_X f_- d\mu = \int_X |f| d\mu \leqslant \int_X g d\mu.$$

4. Suit de la fin de la preuve de 3.
5. De $0 \leqslant f_+ \leqslant |f|$ il suit que f_+ est intégrable et $0 \leqslant \int_X f_+ d\mu \leqslant \int_X |f| d\mu = 0$. Si $0 \leqslant f_n \leqslant f_+$ sont des fonctions étagées convergeant vers f_+, alors de même $\int_X f_n d\mu = 0$. Cette intégrale peut s'écrire en détail comme

$$\sum_{k=1}^{K_n} y_{n,k}\mu(A_{n,k}) = 0$$

où $y_{n,k} \geqslant 0$ (car $f_n \geqslant 0$), donc $\mu(A_{n,k}) = 0$ pour k tels que $y_{n,k} \neq 0$. On conclut que $f_n \underset{p.p.}{=} 0$ pour tout n, donc $f_+ = \lim f_n \underset{p.p.}{=} 0$. On démontre de la même façon que $f_- \underset{p.p.}{=} 0$ et enfin $f \underset{p.p.}{=} 0$.

□

On peut remarquer qu'il n'y a toujours pas d'additivité de l'intégrale pour des fonctions non positives. Avant de le faire, il nous faudra discuter des intégrales sur des sous-ensembles, dans la section prochaine.

4.5 Mesures à densité

Soit toujours $(X, \mathcal{F}, \mu)$ est un espace mesuré avec une mesure complète finie μ.

Définition 4.5.1. Soit f une fonction mesurable sur X. Pour $A \in \mathcal{F}$ on dit que f est intégrable sur A si $f \cdot I_A$ est intégrable sur X, et dans ce cas on pose

$$\int_A f d\mu = \int_X f\, I_A d\mu.$$

Remarque 4.5.2. On peut vérifier que $(A, \mathcal{F}|_A, \mu_A = \mu|_A)$ est un espace mesuré, et $\int_A f d\mu = \int_A f\, d\mu_A$, l'intégrabilité étant la même dans tous les deux sens.

Remarque 4.5.3. Si $\mu(A) = 0$, alors toute f est intégrable sur A et $\int_A f d\mu = 0$ (on remarque que $fI_A \underset{p.p.}{=} 0$).

Proposition 4.5.4. *Soient $A_n \in \mathcal{F}$ deux à deux disjoints, $n \in \mathbb{N}$, et soit $f : X \to \mathbb{R}$ intégrable sur $A = \cup_{n\in\mathbb{N}} A_n$. Alors f est intégrable sur chaque A_n et*

$$\int_A f d\mu = \sum_{n\in\mathbb{N}} \int_{A_n} f d\mu.$$

Démonstration. Pour tout n on a $(fI_{A_n})_\pm = f_\pm I_{A_n} \leqslant f_\pm$ donc f est intégrable sur A_n. Il est clair que pour la suite, il suffit de considérer f positive.

Si on pose $f_n = fI_{A_n}$, on aura $0 \leqslant f_n \leqslant fI_A$ et

$$fI_A = \sum_{n=0}^{\infty} f_n.$$

Par le Corollaire 4.3.4,

$$\int_A f d\mu = \int_X fI_A = \sum_{n=0}^{\infty} \int_X f_n d\mu = \sum_{n=0}^{\infty} \int_{A_n} f d\mu.$$

□

Corollaire 4.5.5. *Si f est intégrable sur $A \subset X$, alors f est intégrable sur chaque $B \subset A$ mesurable.*

Corollaire 4.5.6. *Soit $f \geqslant 0$ une fonction intégrable sur X. L'application $\mu_f : \mathcal{F} \to \mathbb{R}_+$, $A \mapsto \int_A f d\mu$ est une mesure sur $\mathcal{F}$.*

Démonstration. On a vérifié que cette application a des valeurs dans $\mathbb{R}_+$ et qu'est σ-additive ; il ne reste qu'à remarquer qu'elle vérifie $\mu_f(\emptyset) = 0$. □

Définition 4.5.7. Les mesures obtenues de cette façon sont appelées des mesures à densité.

Exemple 4.5.8. Mesures à densité sont souvent utilisées dans la théorie des probabilités, comme par exemple : $X = \mathbb{R}$, μ est la mesure gaussienne :

$$\mu(A) = \int_A e^{-x^2}\, d\lambda(x),$$

où λ est la mesure de Lebesgue (l'intégrale dans le cas d'une mesure infinie est définie dans la section qui suit). Un autre exemple sur $X = [-1, 1]$ est donné par la mesure semi-circulaire :

$$\mu(A) = \frac{2}{\pi} \int_A \sqrt{1 - x^2}\, d\lambda(x).$$

4.6 Propriétés de l'intégrale II

Il nous reste à montrer la linéarité de l'intégrale dans le cas général (non positif).

Proposition 4.6.1. *Si f, g sont des fonctions intégrables, alors $f + g$ est intégrable et*

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

Démonstration. Comme $0 \leqslant (f+g)_\pm \leqslant f_\pm + g_\pm$, on voit que $f+g$ est intégrable. Soit

$$\begin{aligned}
A_{++} &= \{x : f(x) \geqslant 0,\ g(x) \geqslant 0\},\\
A_{+-} &= \{x : f(x) + g(x) \geqslant 0,\ g(x) < 0\},\\
A_{-+} &= \{x : f(x) + g(x) \geqslant 0,\ f(x) < 0\},
\end{aligned}$$

et $A = A_{++} \cup A_{+-} \cup A_{-+}$, alors

$$(f+g)_+ = (f_+ + g_+) I_{A_{++}} + (f_+ - g_-) I_{A_{+-}} + (g_+ - f_-) I_{A_{-+}},$$

tous les trois termes étant positifs, donc

$$\int_X (f+g)_+ d\mu = \int_{A_{++}} (f_+ + g_+) d\mu + \int_{A_{+-}} (f_+ - g_-) d\mu + \int_{A_{-+}} (g_+ - f_-) d\mu.$$

Sur A_{+-}, on a $f_+ - g_- \geqslant 0$, donc

$$\int_{A_{+-}} f_+ d\mu = \int_{A_{+-}} (f_+ - g_- + g_-) d\mu = \int_{A_{+-}} (f_+ - g_-) d\mu + \int_{A_{+-}} g_- d\mu,$$

d'où

$$\int_{A_{+-}} (f_+ - g_-)d\mu = \int_{A_{+-}} f_+ d\mu - \int_{A_{+-}} g_- d\mu.$$

De même on décompose la troisième intégrale, pour arriver à

$$\begin{aligned}\int_X (f+g)_+ d\mu &= \int_{A_{++}\cup A_{+-}} f_+\, d\mu + \int_{A_{++}\cup A_{-+}} g_+\, d\mu - \int_{A_{+-}} g_- d\mu - \int_{A_{-+}} f_- d\mu \\ &= \int_A f_+\, d\mu + \int_A g_+\, d\mu - \int_A g_- d\mu - \int_A f_- d\mu = \int_A f d\mu + \int_A g d\mu.\end{aligned}$$

Avec un raisonnement similaire, on obtient l'additivité sur $B = \{x : f(x) + g(x) < 0\}$, ce qui démontre la proposition. $\square$

Corollaire 4.6.2. *L'intégrale est une fonctionnelle linéaire sur l'espace des fonctions intégrables.*

Chapitre 5

Théorèmes de convergence

5.1 Théorème de Lebesgue

Le théorème suivant est fondamental et le plus souvent utilisé dans la théorie de l'intégration.

Théorème 5.1.1 (de Lebesgue sur la convergence dominée)**.** *Soit* (f_n) *une suite de fonctions mesurables telle que* $f_n \underset{p.p.}{\to} f$*. Supposons qu'il existe* $g \geqslant 0$ *intégrable telle que* $|f_n| \leqslant g$ *pour tout* n*. Alors* f *est intégrable et*

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

Démonstration. La majoration $|f_n| \leqslant g$ implique que f_n est intégrable pour tout n. Comme $g - f_n \geqslant 0$, par le lemme de Fatou

$$\int_X (g-f) d\mu \leqslant \varliminf_{n\to\infty} \int_X (g-f_n) d\mu = \int_X g d\mu - \varlimsup_{n\to\infty} \int_X f_n d\mu,$$

donc

$$\int_X f d\mu \geqslant \varlimsup_{n\to\infty} \int_X f_n d\mu.$$

En appliquant le lemme de Fatou à $g + f_n \geqslant 0$, on obtient

$$\int_X f d\mu \leqslant \varliminf_{n\to\infty} \int_X f_n d\mu.$$

La suite des intégrales donc converge, et sa limite est $\int_X f d\mu$. □

5.2 Inégalité de Tchebychev. Théorème de Lévi

Théorème 5.2.1 (Inégalité de Tchebychev)**.** *Soit* f *intégrable sur* X*. Alors pour tout* $\delta > 0$

$$\mu(\{x : |f(x)| \geqslant \delta\}) \leqslant \frac{1}{\delta} \int_X |f| d\mu.$$

Démonstration. On rappelle que par définition, f est intégrable si et seulement si $|f| = f_+ + f_-$ est intégrable. Pour $\delta > 0$ soit $A_\delta = \{x : |f(x)| \geqslant \delta\}$. On a $|f| I_A \geqslant \delta I_A$, alors

$$\delta \int_X I_A d\mu = \delta\mu(A) \leqslant \int_A |f| d\mu \leqslant \int_X |f| d\mu.$$

Il ne reste qu'à diviser par δ. ☐

(fin du cours 9)

Théorème 5.2.2 (de Lévi sur la convergence monotone)**.** *Soit (f_n) une suite croissante de fonctions mesurables : $f_n \leqslant f_{n+1}$ pour tout n. Supposons que chaque f_n est intégrable et*

$$\sup_n \int_X f_n d\mu = K < \infty.$$

Alors presque partout sur X il existe la limite finie

$$f(x) = \lim_{n\to\infty} f_n(x),$$

la fonction f définie ainsi (en posant $f(x) = 0$ dans les points x où la limite est infinie) est intégrable, et

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

Démonstration. En passant si besoin à la suite $g_n = f_n - f_0$, on peut supposer que $f_n \geqslant 0$. Par la monotonie, la suite $(f_n(x))$ converge en tout point x vers $\sup_n f_n(x)$, finie ou infinie. Soit

$$C = \{x : \lim_{n\to\infty} f_n(x) = \infty\}.$$

Si on pose $C_{N,n} = \{x : f_n(x) \geqslant N\}$, alors $C_{N,n} \subset C_{N,n+1}$ pour tous N, n et

$$C = \bigcap_{N\in\mathbb{N}} \bigcup_{n\geqslant N} C_{N,n}.$$

Par l'inégalité de Tchebychev,

$$\mu(C_{N,n}) \leqslant \frac{1}{N} \int_X f_n d\mu \leqslant \frac{K}{N},$$

alors

$$\mu\Big(\bigcup_{n\geqslant N} C_{N,n}\Big) = \lim_{n\to\infty} \mu(C_{N,n}) \leqslant \frac{K}{N},$$

donc $\mu(C) \leqslant K/N$ pour tout N. Il en suit que $\mu(C) = 0$, ce qui démontre la convergence de (f_n) vers f presque partout.

Maintenant, comme $f_n \underset{p.p.}{\leqslant} f$, par le théorème 4.3.3

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

La suite des intégrales est croissante et bornée, la limite est donc finie, et f est intégrable. ☐

Résumé des théorèmes de convergence : on peut conclure que

$$\int_X f d\mu = \lim_{n\to\infty} \int_X f_n d\mu$$

si :

- $\star$ $f_n \underset{p.p.}{\to} f$, $|f_n| \leqslant g$, g intégrable (convergence dominée, Lebesgue, 5.1.1) ;
- $\star$ (f_n) croissante, $\sup_n \int f_n < \infty$ (convergence monotone, Lévi, 5.2.2).

Et on peut toujours majorer, par le lemme de Fatou,

$$\int_X \varliminf_{n\to\infty} f_n d\mu \leqslant \varliminf_{n\to\infty} \int_X f_n d\mu.$$

Avec ces théorèmes, on obtient de puissants outils pour le calcul des intégrales.

5.3 Intégrale dans le cas d'une mesure infinie

On peut définir l'intégrale dans le cas suivant, pour une classe particulière de mesures infinies :

Définition 5.3.1. Une mesure μ sur un espace mesurable $(X, \mathcal{F})$ est dite σ-finie s'il existent $X_n \in \mathcal{F}$ disjoints tels que $X = \cup_n X_n$ et $\mu(X_n) < \infty$ pour tout n.

Exemple 5.3.2. La mesure de Lebesgue sur $\mathbb{R}^n$ est σ-finie.

Définition 5.3.3. Soit $(X, \mathcal{F}, \mu)$ un espace mesuré où μ est σ-finie. Soit (X_n) une décomposition de X en ensembles disjoints de mesure finie. Un fonction mesurable $f : X \to \mathbb{R}$ est dite intégrable si elle est intégrable sur X_n pour chaque n et

$$\sum_n \int_{X_n} |f| d\mu < \infty.$$

Dans ce cas, on pose

$$\int_X f d\mu = \sum_n \int_{X_n} f d\mu.$$

(Il est évident que la série des intégrales converge.)

Proposition 5.3.4. *L'intégrabilité de f et son intégrale de dépendent pas du choix de (X_n).*

Démonstration. Soit f intégrable, alors pour tout $A \subset X$

$$\int_A |f| d\mu = \sum_n \int_{A\cap X_n} |f| d\mu \leqslant \sum_n \int_{X_n} |f| d\mu = \int_X |f| d\mu < \infty.$$

Soit (Y_m) une autre décomposition de X avec les mêmes propriétés. On peut noter $Z_M = \cup_{m\leqslant M} Y_m$, alors pour tout $M \in \mathbb{N}$

$$\sum_{m=0}^{M} \int_{Y_m} |f| d\mu = \int_{Z_M} |f| d\mu \leqslant \int_X |f| d\mu,$$

donc la série $\sum_m \int_{Y_m} |f|d\mu$ converge et f est intégrable par rapport à (Y_m). Ensuite, la double série convergeant absolument,

$$\sum_m \int_{Y_m} f d\mu = \sum_m \sum_n \int_{Y_m \cap X_n} f d\mu = \sum_n \int_{X_n} f d\mu,$$

ce qui termine la preuve. □

Exemple 5.3.5. La fonction $f(x) = \sin(x)/x$ n'est pas intégrable sur $\mathbb{R}$; la fonction $f(x) = x^p$ est intégrable sur $[1, +\infty[$ ssi $p < -1$.

5.4 Définition de l'espace L^1

On suppose que $(X, \mathcal{F}, \mu)$ est un espace mesuré.

Rappelons que deux fonctions mesurables f, g sont dites équivalentes si $f \underset{p.p.}{=} g$. On a vu que f et g sont intégrables ou non au même temps, et dans ce cas leurs intégrales sont égales.

Définition 5.4.1. On note $L^1(X, \mu)$ l'ensemble de classes d'équivalence de fonctions intégrables. Si μ est fixe, on peut le noter aussi $L^1(X)$. Si $f : X \to \mathbb{R}$ est une fonction intégrable, on peut utiliser la notation f aussi pour sa classe d'équivalence, en écrivant donc $f \in L^1(X)$.
Pour $f \in L^1(X)$, on pose

$$\|f\|_1 = \int_X |f| d\mu.$$

Par les propriétés de l'intégrale,

1. $\|f\|_1 \geqslant 0$; si $\|f\|_1 = 0$ alors $f \underset{p.p.}{=} 0$ donc $f = 0$ dans $L^1(X)$;
2. $\|cf\|_1 = |c| \, \|f\|_1$ pour $c \in \mathbb{R}$;
3. $\|f + g\|_1 = \int_X |f + g|_1 d\mu \leqslant \int_X |f| d\mu + \int_X |g| d\mu = \|f\|_1 + \|g\|_1$.

On conclut que $\|\cdot\|_1$ est une norme sur $L^1(X)$ (qui est évidemment un espace vectoriel).

Proposition 5.4.2. *L'espace des fonctions étagées est dense dans $L^1(X, \mu)$.*

Démonstration. Supposons d'abord que la mesure de X est finie. Soit $f \in L^1(X)$. Par la définition de l'intégrale, pour tout $\varepsilon > 0$ il existent des fonctions étagées g_+, g_- telles que $0 \leqslant g_\pm \leqslant f_\pm$ et

$$\int_X g_\pm d\mu \leqslant \int_X f_\pm d\mu < \int_X g_\pm d\mu + \frac{\varepsilon}{2}.$$

Il en suit que

$$\int_X |f - (g_+ - g_-)| d\mu \leqslant \int_X (f_+ - g_+) d\mu + \int_X (f_- - g_-) d\mu < \varepsilon,$$

ce qui démontre la densité.

Dans le cas où μ est seulement σ-finie, soit $X = \sqcup_n X_n$ une décomposition de X en ensembles de mesure finie. Pour $f \in L^1(X)$, la série suivante converge :

$$\sum_{n=1}^{\infty} \int_{X_n} |f| d\mu < \infty.$$

Pour tout $\varepsilon > 0$, on peut alors choisir $N \in \mathbb{N}$ tel que le reste de la série est inférieur à ε : $\sum_{n=N+1}^{\infty} \int_{X_n} |f| d\mu < \varepsilon$; il suffira alors d'approcher f sur X_n avec $n \leqslant N$ seulement.

Pour tout $n \leqslant N$, par ce qui précède il existe une fonction étagée g_n sur X_n telle que $\int_{X_n} |f - g_n| d\mu < \varepsilon/N$. Soit

$$g(x) = \begin{cases} g_n(x), & x \in X_n,\ n \leqslant N \\ 0, & \text{sinon} \end{cases}$$

C'est bien une fonction étagée, et

$$\begin{aligned} \int_X |f - g| d\mu = \sum_{n=1}^{\infty} \int_{X_n} |f - g| d\mu &= \sum_{n=1}^{N} \int_{X_n} |f - g_n| d\mu + \sum_{n=N+1}^{\infty} \int_{X_n} |f| d\mu \\ &< N \cdot \frac{\varepsilon}{N} + \varepsilon = 2\varepsilon. \end{aligned}$$

□

Dans $\mathbb{R}$, on peut démontrer plus précisément :

Proposition 5.4.3. *Dans $L^1(\mathbb{R}, \lambda)$ ou dans $L^1([a,b], \lambda)$ les fonctions continues à support compact ou les fonctions lisses (c.à.d. C^∞) à support compact sont des sous-espaces denses.*

Démonstration. Se démontre en approchant une fonction indicatrice d'un intervalle par une fonction continue ou lisse. (A préciser en TDs.) □

5.5 Convergence en norme

Définition 5.5.1. Une suite de fonctions (f_n) est dit de converger *en norme* vers f si $\|f_n - f\|_1 \to 0$, $n \to \infty$.

Théorème 5.5.2. *$L^1(X, \mu)$ est un espace de Banach avec la norme $\|\cdot\|_1$. Autrement dit, il est un espace métrique complet.*

Démonstration. Soit $(f_n) \subset L^1(X, \mu)$ une suite de Cauchy ; montrons qu'elle converge en norme vers $f \in L^1(X, \mu)$. Pour tout k, il existe N_k tel que $\|f_n - f_m\|_1 < 2^{-k}$ si $n, m \geqslant N_k$. On peut supposer que $N_{k+1} > N_k$ pour tout k. Si on pose $g_k = f_{N_k}$, on aura alors $\|g_k - g_{k+1}\|_1 < 2^{-k}$. Cela implique

$$\sum_{k=0}^{\infty} \int_X |g_{k+1} - g_k| d\mu \leqslant 1$$

et donc par le théorème de Lévi la série à termes positifs

$$\sum_{k=0}^{\infty} |g_{k+1} - g_k| \underset{p.p.}{=} g$$

converge presque partout vers une fonction intégrable g. Il en suit que la série

$$g_0 + \sum_{k=0}^{\infty} (g_{k+1} - g_k) = \lim_{n \to \infty} g_n \underset{p.p.}{=} f$$

converge presque partout aussi, tout en restant majorée par g. Par le théorème de Lebesgue, f est intégrable. Par le choix de (g_k), on a pour tout $l > k$

$$\int_X |g_k - g_l| d\mu < 2^{-k}.$$

Par le lemme de Fatou, comme $g_k - f \underset{p.p.}{=} \lim_{l\to\infty}(g_k - g_l)$,

$$\int_X |g_k - f| d\mu \leqslant \varliminf_{l\to\infty} \int_X |g_k - g_l| d\mu \leqslant 2^{-k},$$

donc $g_k \to f$ en norme. Il reste à noter que pour $n > N_k$, $\|g_k - f_n\|_1 < 2^{-k}$, donc $\|f_n - f\|_1 \to 0$. $\square$

Dans la preuve, on a démontré en particulier :

Corollaire 5.5.3. *Toute suite convergente dans $L^1(X)$ contient une sous-suite convergente presque partout.*

La convergence presque partout ne suffit par pour la convergence en norme, un exemple simple est donné par $f_n(x) = nx^n$ sur $[0, 1]$.

Il existe encore un type de convergence, surtout utile dans la théorie des probabilités :

Définition 5.5.4. Une suite (f_n) de fonction mesurables est dite *converger en mesure vers f* si pour tout $\delta > 0$

$$\mu\Big(\big\{x \in X : |f(x) - f_n(x)| > \delta\big\}\Big) \to 0, \quad n \to \infty.$$

Cette convergence sera considérée en TDs.

(fin du cours 10)

Chapitre 6

L'intégrale de Riemann vs. l'intégrale de Lebesgue. Lien avec les dérivées

6.1 Toute fonction Riemann intégrable est Lebesgue intégrable

Théorème 6.1.1. *Soit f Riemann intégrable sur le segment fini $[a, b]$. Alors f est Lebesgue intégrable sur $[a, b]$ et*

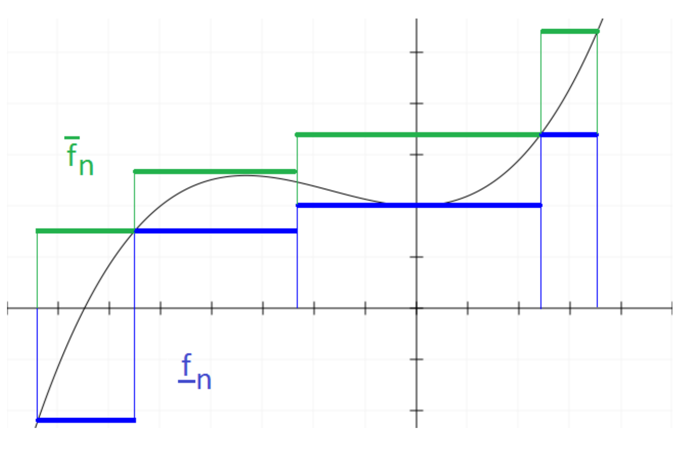
$$(R)\int_a^b f(x)dx = (L)\int_{[a,b]} f d\lambda.$$

Démonstration. L'intégrabilité au sens de Riemann est définie par la convergence de sommes suivantes : si $a = x_0 < x_1 < \cdots < x_n = b$, on définit les sommes inférieure $\underline{S}(f)$ et supérieure $\overline{S}(f)$ comme

$$\underline{S}(f) = \sum_{k=1}^{n}(x_k - x_{k-1}) \inf_{[x_{k-1},x_k]} f, \qquad \overline{S}(f) = \sum_{k=1}^{n}(x_k - x_{k-1}) \sup_{[x_{k-1},x_k]} f.$$

f est alors intégrable s'il existe la limite $\lim_{\Delta\to 0} \underline{S}(f) = \lim_{\Delta\to 0} \overline{S}(f)$ par rapport à $\Delta = \max\{x_k - x_{k-1} : 1 \leqslant k \leqslant n\}$. On remarque que f est en fait approchée par des fonctions étagées

$$\underline{f}_{n,\bar{x}} = \sum_{k=1}^{n} I_{[x_{k-1},x_k]} \inf_{[x_{k-1},x_k]} f, \qquad \overline{f}_{n,\bar{x}} = \sum_{k=1}^{n} I_{[x_{k-1},x_k]} \sup_{[x_{k-1},x_k]} f,$$



et $\underline{S}(f) = (L)\int_{[a,b]} \underline{f}_{n,\bar{x}} d\lambda$, $\overline{S}(f) = (L)\int_{[a,b]} \overline{f}_{n,\bar{x}} d\lambda$. En particulier, $\underline{f}_{n,\bar{x}} \leqslant f \leqslant \overline{f}_{n,\bar{x}}$ quels que soient les points $\{x_k\}$.

Maintenant pour chaque n choisissons des intervalles de longueur identique : $x_k = a + \frac{k}{n}(b-a)$. On aura $\Delta_n = (b-a)/n \to 0$. Par le lemme de Fatou (à rappeler, il est applicable à des fonctions positives, mais il est facile à voir que le lemme est aussi valable dans le cas où toutes les fonctions sont minorées par la même constante),

$$\int_{[a,b]} \varliminf_{n\to\infty} \overline{f}_n d\lambda \leqslant \varliminf_{n\to\infty} \int_{[a,b]} \overline{f}_n d\lambda = \lim_{n\to\infty} \int_{[a,b]} \overline{f}_n d\lambda = (R)\int_a^b f(x)dx;$$

à noter, l'intégrale à gauche est dans le sens de Lebesgue, et les intégrales de $\overline{f}_n$ peuvent être comprises dans l'un ou l'autre sens car ces fonctions sont étagées. Ensuite, on peut appliquer le lemme de Fatou à la suite $(-f_n)$ qui est bornée aussi, et obtenir

$$\int_{[a,b]} \varliminf_{n\to\infty} (-\underline{f}_n) d\lambda \leqslant \varliminf_{n\to\infty} \int_{[a,b]} (-\underline{f}_n) d\lambda.$$

Changement de signe inverse les limites supérieure et inférieure, alors

$$\int_{[a,b]} \varlimsup_{n\to\infty} \underline{f}_n d\lambda \geqslant \varlimsup_{n\to\infty} \int_{[a,b]} \underline{f}_n d\lambda = \lim_{n\to\infty} \int_{[a,b]} \underline{f}_n d\lambda = (R)\int_a^b f(x)dx.$$

Comme on a $\underline{f}_n \leqslant f \leqslant \overline{f}_n$,

$$\overline{f} := \varliminf_{n\to\infty} \overline{f}_n \geqslant f \geqslant \varlimsup_{n\to\infty} \underline{f}_n =: \underline{f}$$

et $\overline{f} - \underline{f} \geqslant 0$. Au même temps, par les inégalités ci-dessus

$$\int_{[a,b]} \overline{f} d\lambda \leqslant (R)\int_a^b f(x)dx \leqslant \int_{[a,b]} \underline{f} d\lambda$$

donc en fait l'intégrale de $\overline{f} - \underline{f}$ est nulle. Il en suit que $\overline{f} - \underline{f} \underset{p.p.}{=} 0$; donc f est aussi égale à chacune de ses fonctions presque partout et est donc mesurable. Son intégrale de Lebesgue est minorée par $\int_{[a,b]} \underline{f} d\lambda$ et majorée par $\int_{[a,b]} \overline{f} d\lambda$, et donc elle est égale à ces deux intégrales (égales entre elles) et à $(R)\int_a^b f(x)dx$. $\square$

A partir de ce moment, on écrit $\int_a^b f d\lambda$ pour une fonction mesurable, dans le sens de Lebesgue aussi.

Comme il a été déjà dit, dans le cas d'une fonction non-bornée ou d'un intervalle non-borné (si f vérifie les conditions où l'intégrale de Riemann impropre est définie) l'intégrale de Lebesgue existe si et seulement si $|f|$ est intégrable.

Nous pourrons aussi calculer les intégrales de Lebesgue à l'aide de primitives si elles existent, tout comme avec les intégrales de Riemann. On y aura même plus de liberté, comme on verra plus tard.

6.2 Continuité absolue et l'intégrale indéfinie

Théorème 6.2.1 (Continuité absolue de l'intégrale)**.** *Soit f intégrable sur X. Alors pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que si $\mu(A) < \delta$, alors*

$$\Big|\int_A f d\mu\Big| \leqslant \int_A |f| d\mu < \varepsilon.$$

Démonstration. En utilisant l'inégalité $\left|\int f d\mu\right| \leqslant \int |f| d\mu$, on peut supposer f positive. Selon la définition de l'intégrale, il existe une fonction étagée g telle que $0 \leqslant g \leqslant f$ et

$$\int_X g d\mu > \int_X f d\mu - \frac{\varepsilon}{2},$$

alors

$$0 \leqslant \int_X (f-g) d\mu < \frac{\varepsilon}{2}.$$

On peut supposer f non-nulle (sinon l'assertion est triviale), alors on pourra toujours choisir g non-nulle. Comme g est bornée, pour tout $A \in \mathcal{F}$

$$\int_A g d\mu \leqslant \mu(A)\|g\|_\infty \neq 0.$$

On pose $\delta = \dfrac{\varepsilon}{2\|g\|_\infty}$, alors pour tout ensemble A de mesure $\mu(A) < \delta$ on obtient

$$0 \leqslant \int_A f d\mu = \int_A g d\mu + \int_A (f-g) d\mu < \mu(A)\|g\|_\infty + \frac{\varepsilon}{2} < \varepsilon.$$

□

Définition 6.2.2. Soit $(X, \mathcal{F}, \mu)$ un espace mesuré. Une autre mesure ν sur $(X, \mathcal{F})$ est dite *absolument continue* par rapport à μ si pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que $\mu(A) < \delta$ implique $|\nu(A)| < \varepsilon$.

Remarque 6.2.3. Le théorème dit que toute mesure à densité est absolument continue par rapport à μ. On peut montrer la reciproque : si ν est absolument continue, alors il existe f telle que $\nu = \mu_f$ est une mesure à densité f. Ce résultat s'appelle le théorème de Radon-Nikodym mais il ne fait pas partie du cours.

> Soit f intégrable sur $[a, b]$. Son intégrale indéfinie est une fonction F sur $[a, b]$,
> $$F(x) = \int_a^x f d\lambda.$$

On va montrer que F est continue, et même dans un sens fort suivant.

Définition 6.2.4. On dit que $F : [a, b] \to \mathbb{R}$ est *absolument continue* si pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que pour toute famille d'intervalles disjoints $\sqcup_k (a_k, b_k) \subset [a, b]$

$$\sum_k (b_k - a_k) < \delta \quad \text{implique} \quad \sum_k |F(b_k) - F(a_k)| < \varepsilon.$$

Remarque 6.2.5. Toute fonction absolument continue est continue en tout point $x \in [a, b]$: si ε, δ sont comme dans la définition et $|y - x| < \delta$, alors $|F(y) - F(x)| < \varepsilon$.

Proposition 6.2.6. *Si F est l'intégrale indéfinie d'une fonction f intégrable sur $[a, b]$, alors F est absolument continue.*

Démonstration. Suit directement du Théorème 6.2.1 : en posant $I = \sqcup_k (a_k, b_k)$, on a $\lambda(I) = \sum\limits_k (b_k - a_k) < \delta$, alors

$$\begin{aligned}\sum_k |F(b_k) - F(a_k)| &= \sum_k \left| \int_a^{b_k} f d\lambda - \int_a^{a_k} f d\lambda \right| \\ &= \sum_k \left| \int_{[a_k, b_k]} f d\lambda \right| \leqslant \sum_k \int_{[a_k, b_k]} |f| d\lambda = \int_I |f| d\lambda < \varepsilon.\end{aligned}$$

□

6.3 La dérivée de l'intégrale

Soit f intégrable sur $[a,b]$. On peut définir son intégrale indéfinie qui est une fonction F sur $[a,b]$:

$$F(x) = \int_a^x f d\lambda.$$

On ne peut pas s'attendre à ce que F soit dérivable partout : il suffit de considérer $[a,b] = [-1,1]$ et

$$f(x) = \begin{cases} 1, & x \geqslant 0 \\ 0, & x < 0. \end{cases}$$

Mais on verra que F est dérivable *presque partout*.

Il est clair que F est la différence de deux fonctions monotones :

$$F(x) = \int_a^x f_+ d\lambda - \int_a^x f_- d\lambda.$$

Théorème 6.3.1 (Lebesgue)**.** *Toute fonction monotone est dérivable presque partout.*

La preuve de ce théorème est longue mais pas très compliquée (à consulter par exemple [5]). En l'admettant, on peut démontrer la propriété suivante de l'intégrale :

Théorème 6.3.2. *Soit f intégrable sur $[a,b]$, et soit F son intégrale indéfinie. Alors $F'(x)$ existe presque partout et $F' \underset{p.p.}{=} f$.*

Démonstration. L'existence presque partout suit du théorème de Lebesgue, et le reste de la preuve est pour montrer l'égalité $F' \underset{p.p.}{=} f$. On va démontrer d'abord que $f \underset{p.p.}{\geqslant} F'$.

Si x est tel que $f(x) < F'(x)$, alors ils existent $q, r \in \mathbb{Q}$ tels que $f(x) < q < r < F'(x)$. Soit

$$E_{q,r} = \{x : f(x) < q < r < F'(x)\}.$$

On vérifie qu'en fait, l'ensemble $E = \{x : f(x) < F'(x)\}$ est la réunion

$$E = \bigcup_{q,r \in \mathbb{Q}} E_{q,r}.$$

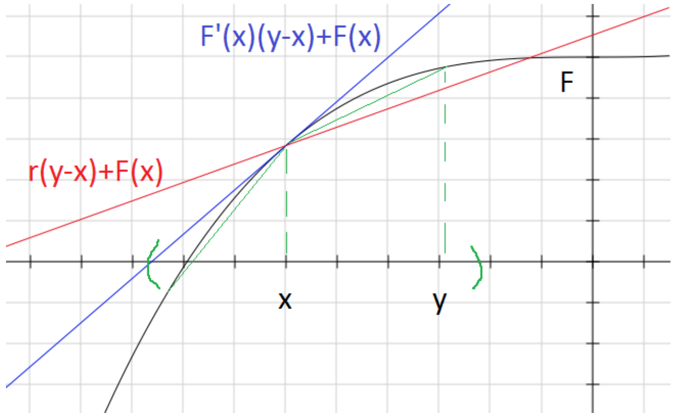
Le but est de montrer que $\lambda(E_{q,r}) = 0$ pour tout q, r ; comme la réunion est dénombrable, il en suivra $\lambda(E) = 0$ et donc $f \underset{p.p.}{\geqslant} F'$.

Avec q, r fixes, on peut ajouter une constante à f — ce qui ajoute la même constante à F' — et admettre $q \geqslant 0$.

Fixons $\varepsilon > 0$. Par la continuité absolue, il existe $\delta > 0$ tel que $|\int_I f d\lambda| < \varepsilon$ pour tout intervalle I de longueur $|I| = \lambda(I) < \delta$. On peut choisir $\delta < \varepsilon/r$. Par la construction de la mesure de Lebesgue (à rappeler, la mesure extérieure est la borne inférieure par des recouvrements par des intervalles, et ceux-là on peut choisir ouverts) il existe un ouvert V tel que $E_{qr} \subset V$ et $\lambda(V) < \lambda(E_{qr}) + \delta$.

Comme $F'(x) > r$ pour $x \in E_{qr}$, il existe un voisinage assez petit $]a_x, b_x[$ de x (qu'on peut supposer contenu dans V) tel que dans ce voisinage,

$$\frac{F(y) - F(x)}{y - x} > r. \tag{6.1}$$



Si on note $G(y) = F(y) - ry$, on remarque, en transformant (6.1), que

$$G(y) > G(x) \text{ si } y > x \text{ et } G(y) < G(x) \text{ si } y < x, \tag{6.2}$$

pour $y \in]a_x, b_x[$.

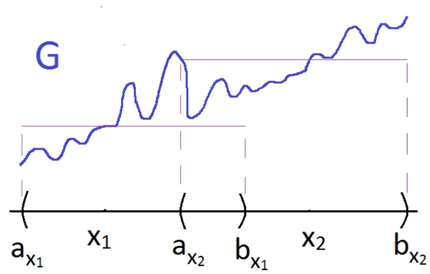
L'ensemble $U = \bigcup\limits_{x \in E_{qr}}]a_x, b_x[$ est ouvert, alors il est égal à une réunion dénombrable d'intervalles ouverts disjoints : $U = \bigcup\limits_n]\alpha_n, \beta_n[$.

On peut vérifier que $G(\alpha_n) \leqslant G(\beta_n)$ pour tout n. [Supposons au contraire que $G(\alpha_n) > G(\beta_n)$. On sait que G est continue, alors il existe $\delta_n > 0$ tel que $G(y_1) > \frac{1}{2}\big(G(\alpha_n) + G(\beta_n)\big) > G(y_2)$ pour $y_1 \in [\alpha_n, \alpha_n + \delta_n]$, $y_2 \in [\beta_n - \delta_n, \beta_n]$. Le segment $[\alpha_n + \delta_n, \beta_n - \delta_n]$ est compact et recouvert par les intervalles ouverts $]a_x, b_x[$, $x \in E_{qr}$ (les intervalles $]a_x, b_x[$ qui rencontrent $]\alpha_n, \beta_n[$ y sont contenus, par le fait que $]\alpha_n, \beta_n[$ sont ouverts et disjoints). Il existe alors un sous-recouvrement fini :

$$[\alpha_n + \delta_n, \beta_n - \delta_n] \subset \bigcup_{k=1}^{m}]a_{x_k}, b_{x_k}[,$$

où l'on peut supposer que $\alpha_n + \delta_n \in]a_{x_1}, b_{x_1}[$ et $\beta_n - \delta_n \in]a_{x_m}, b_{x_m}[$.

Pour tout k, on a par la construction $G(a_{x_k}) < G(x_k) < G(b_{x_k})$ (voir le dessin), ainsi que $G(\alpha_n + \delta_n) < G(x_1)$ et $G(x_2) < G(\beta_n - \delta_n)$. Il est facile à voir que cela implique $G(a_{x_1}) < G(b_{x_m})$, ce qui est une contradiction car $a_{x_1} \in [\alpha_n, \alpha_n + \delta_n]$ et $b_{x_m} \in [\beta_n - \delta_n, \beta_n]$; alors en fait $G(\alpha_n) < G(\beta_n)$.]



L'inégalité obtenue $G(\alpha_n) \leqslant G(\beta_n)$ est équivalente à

$$F(\beta_n) - F(\alpha_n) \geqslant r(\beta_n - \alpha_n)$$

pour tout n.

On aura alors :

$$\int_U f d\lambda = \sum_{k=0}^{\infty} \int_{\alpha_k}^{\beta_k} f d\lambda = \sum_{k=0}^{\infty} [F(\beta_k) - F(\alpha_k)] \geqslant r \sum_{k=0}^{\infty} (\beta_k - \alpha_k) = r\lambda(U),$$

et au même temps

$$\int_U f d\lambda = \int_{E_{qr}} f d\lambda + \int_{U \setminus E_{qr}} f d\lambda < q\lambda(E_{qr}) + \varepsilon$$

car $\lambda(U \setminus E_{qr}) \leqslant \lambda(V \setminus E_{qr}) < \delta$. On obtient $q\lambda(E_{qr}) + \varepsilon > r\lambda(U) \geqslant r\lambda(E_{qr})$, et

$$\lambda(E_{qr}) < \frac{\varepsilon}{r - q}.$$

En choisissant ε arbitrairement petit, on montre alors que $\lambda(E_{qr}) = 0$. Cela implique $f \underset{p.p.}{\geqslant} F'(x)$.

En passant à $-f$ et donc à $-F$, on montre que $-f \underset{p.p.}{\geqslant} -F'(x)$ et $f \underset{p.p.}{\leqslant} F'(x)$, ce qui termine la preuve. □

6.4 Théorème fondamental de l'analyse

On appelle ainsi la formule de Newton–Leibnitz :

$$F(b) - F(a) = \int_a^b F'(x)dx.$$

Il faut préciser que même dans le sens de Lebesgue, l'égalité n'est pas toujours vérifié : on considérera en TDs l'exemple de l'«escalier de Cantor» F où on a $F(1) - F(0) = 1$ mais $F' \underset{p.p.}{=} 0$. La formule est vraie pour les fonctions absolument continues ; on a vu que c'est exactement la classe des intégrales indéfinies de fonctions intégrables.

Théorème 6.4.1. *Si F est absolument continue sur $[a, b]$, alors elle est presque partout dérivable et*

$$F(b) - F(a) = \int_a^b F'(x)dx.$$

La preuve sera peut-être considérée au dernier cours car elle est assez longue. On peut aussi consulter [5]. Pour le moment, on démontre le cas particulier suivant :

Théorème 6.4.2. *Si F est continue et dérivable sur $[a, b]$, avec la dérivée F' bornée sur $[a, b]$, alors F' est intégrable et*

$$F(b) - F(a) = \int_a^b F'(x)dx.$$

Démonstration. On peut prolonger F sur $[b, b+1]$ en posant $F(x+b) = F(b) + F'(b)x$. Pour $n \in \mathbb{N}$, $n > 0$, on peut alors définir sur $[a, b]$ les fonctions

$$F_n(x) = \frac{F(x + \frac{1}{n}) - F(x)}{1/n} = n\big[F(x + \frac{1}{n}) - F(x)\big],$$

$x \in [a, b]$. Par l'hypothèse, $F_n \to F'$ sur $[a, b]$. Il en suit déjà que F' est mesurable, et en étant bornée elle est intégrable. Par le théorème d'accroissement finis pour chaque $x \in [a, b]$ et chaque n il existe $y \in [x, x + \frac{1}{n}]$ tel que $F_n(x) = F'(y)$, donc

$$|F_n(x)| \leqslant C = \sup\{|F'(x)| : x \in [a, b]\}.$$

Par le théorème de convergence dominée,

$$\int_a^b F'd\lambda = \lim_{n\to\infty} \int_a^b F_n d\lambda = \lim_{n\to\infty} \int_a^b n\big[F(x + \frac{1}{n}) - F(x)\big]d\lambda(x) =$$
$$= \lim_{n\to\infty} n\Big[\int_b^{b+1/n} Fd\lambda - \int_a^{a+1/n} Fd\lambda\Big]$$

(On peut faire le changement de variable $x \mapsto x - 1/n$ car les fonctions sous les intégrales sont continues, alors on peut calculer les intégrales au sens de Riemann.) Encore par le théorème d'accroissement finis mais appliqué à l'intégrale indéfinie de F, ces dernières intégrales sont égales à $F(b_n)/n$ et $F(a_n)/n$ respectivement, où $b_n \in [b, b+1/n]$ et $a_n \in [a, a+1/n]$. On conclut que

$$\int_a^b F' d\lambda = \lim_{n\to\infty} \big[F(b_n) - F(a_n)\big] = F(b) - F(a).$$

$\square$

(fin du cours 11)

Chapitre 7

Théorème de Fubini

7.1 Produits de mesures

Soient $(X_k, \mathcal{F}_k, \mu_k)$ des espaces mesurés, $k = 1, 2$. Notre but est de construire une mesure μ sur (une tribu sur) $X_1 \times X_2$ de sorte que $\mu(A \times B) = \mu_1(A)\mu_2(B)$. On l'appelle le produit des mesures μ_1, μ_2 et note $\mu = \mu_1 \times \mu_2$. C'est possible également pour un produit de n ensembles, $n \geqslant 2$.

Sur $X_1 \times X_2$, il est nécessaire de commencer d'anneaux car la tribu produit ne se construit pas directement. Rappelons que $\mathcal{R}$ est un anneau d'ensembles sur X si $\emptyset \in \mathcal{R}$, et $A, B \in \mathcal{R}$ implique $A \cap B$, $A \triangle B \in \mathcal{R}$. Soit

$$\mathcal{R} = \{ \bigsqcup_{j=1}^{m} A_j \times B_j : \ A_j \in \mathcal{F}_1, \ B_j \in \mathcal{F}_2 \}.$$

Alors $\mathcal{R}$ est un anneau d'ensembles sur $X_1 \times X_2$. La preuve est identique à celle dans le cas des intervalles dans $\mathbb{R}^2$, voir chapitre 2.

On définit l'application $\mu : \mathcal{R} \to \mathbb{R}_+$ en posant

$$\mu\big(\bigsqcup_{j=1}^{m} A_j \times B_j \big) = \sum_{j=1}^{m} \mu_1(A_j)\mu_2(B_j).$$

Exactement comme dans le cas de $\mathbb{R}^n$, on montre que μ est bien définie (ne dépend pas de la décomposition en A_j, B_j) et additive.

Plus généralement, on peut considérer n ensembles $X_1, \ldots, X_n$ munis de tribus (ou même d'anneaux) $\mathcal{F}_1, \ldots, \mathcal{F}_n$, et construire de la même façon un anneau $\mathcal{R}$ sur $X_1 \times \cdots \times X_n$ et l'application $\mu : \mathcal{R} \to \mathbb{R}_+$.

Pour montrer la σ-additivité dans $\mathbb{R}^n$, nous avons utilisé l'approximation par des ouverts, ce qui n'est pas possible dans le cas abstrait. On doit alors faire autrement :

Proposition 7.1.1. *La mesure de produit μ est σ-additive sur $\mathcal{R}$.*

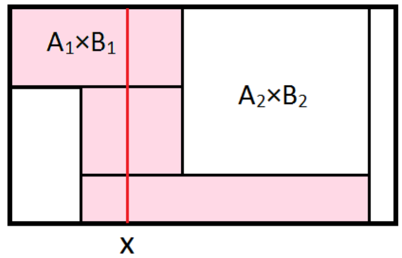
Démonstration. Il faut démontrer que si $C_m \in \mathcal{R}$, $m \in \mathbb{N}$ sont disjoints et $C = \sqcup_m C_m \in \mathcal{R}$, alors $\mu(C) = \sum_m \mu(C_m)$. On fait la preuve pour le cas $n = 2$, $X = X_1 \times X_2$. En changeant des indices si nécessaire, on peut supposer $C_m = A_m \times B_m$ pour chaque m (et non pas une réunion finie d'ensembles de ce type). Ensuite, a priori

$$C = \bigsqcup_{j=1}^{s} \tilde{A}_j \times \tilde{B}_j,$$

mais en remplaçant si besoin A_m, B_m par $A_m \cap \tilde{A}_j$, $B_m \cap \tilde{B}_j$ on peut supposer que $C = A \times B$ avec $A \in \mathcal{F}_1$, $B \in \mathcal{F}_2$, et le cas général en suivra.

On définit pour chaque $m \in \mathbb{N}$ la fonction suivante (mesurable) sur X_1 :

$$f_m(x) = \begin{cases} \mu_2(B_m), & x \in A_m; \\ 0, & \text{sinon,} \end{cases}$$



autrement dit $f_m = \mu_2(B_m) I_{A_m}$. Par exemple, sur le dessin $f_1(x) = \mu_2(B_1)$ et $f_2(x) = 0$. On a

$$\int_{X_1} f_m d\mu_1 = \mu_1(A_m)\mu_2(B_m) = \mu(C_m),$$

et pour $x \in X_1$,

$$\sum_m f_m(x) = \sum_{m: x \in A_m} \mu_2(B_m) = \begin{cases} \mu_2(B), & x \in A \\ 0, & \text{sinon} \end{cases} \tag{7.1}$$

(chaque terme est nul si $x \notin A$; si $x \in A$, alors l'égalité

$$A \times B = \bigsqcup_m A_m \times B_m$$

implique que $\{x\} \times B = \bigsqcup_{m: x \in A_m} \{x\} \times B_m$, d'où (par la σ-additivité de μ_2) suit l'égalité (7.1).)

Dans la série aux termes positifs $\sum f_m$, les intégrales des sommes partielles sont bornées :

$$\int_{X_1} \sum_{m=1}^{M} f_m d\mu_1 = \sum_{m=1}^{M} \mu_1(A_m)\mu_2(B_m) = \mu\Big(\bigsqcup_{m=1}^{M} A_m \times B_m\Big) \leqslant \mu(C),$$

car μ est finiment additive et monotone. Par le théorème de Lévi (plus précisément, par son corollaire pour les séries),

$$\begin{aligned} \sum_m \int_{X_1} f_m d\mu_1 &= \sum_m \mu_2(B_m)\mu_1(A_m) = \int_{X_1} \Big(\sum_m f_m\Big) d\mu_1 \\ &= \int_{X_1} \mu_2(B_m) I_{A_m} d\mu_1 = \mu_1(A)\mu_2(B). \end{aligned}$$

On y retrouve donc l'égalité $\sum_m \mu(C_m) = \mu(C)$. □

Enfin, le théorème de Carathéodory implique que μ s'étend à une tribu $\mathcal{F}$ d'ensembles mesurables, et la mesure obtenue est σ-additive et complète.

On aurait pû définir la mesure de Lebesgue dans $\mathbb{R}^n$ de cette façon, en commencánt par la mesure dans la dimension 1. Cela nous ferait pourtant attendre sa définition jusqu'à la fin du cours, après tous les théorèmes de convergence.

On sait bien qu'en général, il est difficile de décrire les ensembles mesurables. Pour la mesure produit, on a la description «approximative» suivante, par des intersections des réunions d'éléments de $\mathcal{R}$:

Lemma 7.1.2. Soit A mesurable par rapport à $\mu = \mu_1 \times \mu_2$. Il existe alors $B \in \mathcal{F}$ tel que $A \subset B$, $\mu(A) = \mu(B)$ et

$$B = \bigcap_n B_n, \text{ où } B_n \supset B_{n+1} \; \forall n, \tag{7.2}$$

chaque B_n se représente comme

$$B_n = \bigcup_k C_{nk}, \text{ où } C_{nk} \subset C_{n,k+1} \; \forall k, \tag{7.3}$$

et $C_{nk} \in \mathcal{R}$ pour tout n, k.

Démonstration. Par la définition de la mesure extérieure, pour chaque $n \in \mathbb{N}^*$ il existe $D_n \in \mathcal{R}$ tel que $A \subset D_n$ et $\mu(D_n) < \mu(A) + 1/n$. On pose $B_n = \bigcap_{k=1}^{n} D_k$ et $B = \cap_n B_n$. Avec ce choix on a $B_n \supset B_{n+1}$ pour tout n, $A \subset B$ et

$$\mu(A) \leqslant \mu(B) = \lim_{n\to\infty} \mu(B_n) \leqslant \mu(A)$$

donc $\mu(A) = \mu(B)$. Comme $B_n \in \mathcal{R}$, il existent des ensembles produits $E_{nk} = E_{nk}^{(1)} \times E_{nk}^{(2)} \in \mathcal{R}$ tels que $B_n = \sqcup_k E_{nk}$. En posant $C_{nk} = \cup_{j=1}^{k} E_{nj}$, on obtient $C_{nk} \in \mathcal{R}$, $C_{nk} \subset C_{n,k+1}$ pour tout n, k et $B_n = \cup_k C_{nk}$. □

7.2 La mesure d'un ensemble en tant qu'intégrale double

Le théorème de Fubini se réduit au cas particulier de $f = I_A$, l'indicatrice d'un ensemble $A \subset X \times Y$. Nous démontrons ce cas dans cette section.

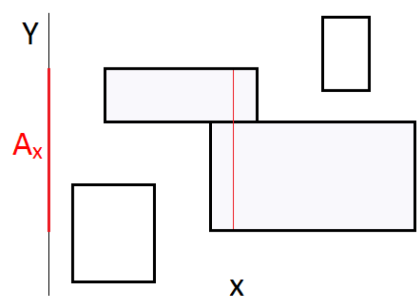
Soient alors $(X, \mathcal{F}_X, \mu_X)$ et $(Y, \mathcal{F}_Y, \mu_Y)$ des ensembles mesurés, avec des mesures finies complètes μ_X et μ_Y. Soit $\mu = \mu_X \times \mu_Y$. Pour $A \subset X \times Y$ on note

$$A_x = \{y \in Y : (x, y) \in A\}, \quad x \in X;$$
$$A_y = \{x \in X : (x, y) \in A\}, \quad y \in Y.$$

Théorème 7.2.1. *Dans les hypothèses ci-dessus, pour tout ensemble μ-mesurable $A \subset X \times Y$ on a :*

1. *les ensembles A_x, A_y sont mesurables pour presque tous $x \in X$ (respectivement $y \in Y$) ;*
2. *les fonctions $x \mapsto \mu_Y(A_x)$, $y \mapsto \mu_X(A_y)$ (égales à 0 en points où A_x ou A_y n'est pas mesurable) sont mesurables ;*
3. $$\mu(A) = \int_X \mu_Y(A_x) d\mu_X = \int_Y \mu_X(A_y) d\mu_Y. \qquad (7.4)$$

Démonstration. Par la symétrie entre X_1 et X_2, il nous suffit de démontrer seulement les assertions sur A_x, $\mu_Y(A_x)$ et l'égalité $\mu(A) = \int_X \mu_Y(A_x) d\mu_X$.



Pour $A = \bigsqcup_{j=1}^{m} B_j \times C_j \in \mathcal{R}$, le théorème est trivial : $A_x = \bigcup_{x \in B_j} C_j$ est mesurable, la fonction $x \mapsto \mu_Y(A_x)$ est égale à $\sum_j \mu_Y(C_j) I_{B_j}$ (comme dans la Proposition 7.1.1), et enfin

$$\mu(A) = \sum_j \mu_X(B_j)\mu_Y(C_j)$$
$$= \sum_j \int_X \mu_Y(C_j) I_{B_j} d\mu_X = \int_X \mu_Y(A_x) d\mu_X.$$

Pour un ensemble $A = B_n$ de forme (7.3) avec $C_{nk} \in \mathcal{R}$, on a $A_x = \bigcup_k (C_{nk})_x$, c'est donc un ensemble mesurable pour tout x. La fonction

$$x \mapsto \mu_Y(A_x) = \lim_{k\to\infty} \mu_Y\big((C_{nk})_x\big)$$

est limite de fonctions mesurables donc mesurable, et

$$\int_X \mu_Y(A_x)d\mu_X = \lim_{k\to\infty}\int_X \mu_Y\big((C_{nk})_x\big)d\mu_X = \lim_{k\to\infty}\mu(C_{nk}) = \mu(A).$$

Nous avons successivement utilisé le théorème de Lebesgue, l'égalité (7.4) pour C_{nk} et la continuité de la mesure μ.

Pour un ensemble B de forme (7.2) la preuve est exactement pareille.

Soit maintenant A seulement approché par un tel ensemble B selon le Lemme 7.1.2. En particulier, $A_x \subset B_x$ pour chaque $x \in X$. Si $\mu(A) = 0$, alors $\mu(B) = \int_Y \mu_Y(B_x)d\mu_Y = 0$ d'où $\mu_Y(B_x) \underset{p.p.}{=} 0$. Les ensembles $A_x \subset B_x$ sont donc négligeables pour presque tout x ; on suppose μ_Y complète, alors tout ensemble négligeable est mesurable, et la fonction $x \mapsto \mu_Y(A_x)$ est mesurable. Son intégrale est évidemment nulle.

Enfin, soit $A \in \mathcal{F}$ quelconque. Soit B son ensemble approchant construit selon Lemme 7.1.2, et soit $C = B \setminus A$. Les ensembles $C_x = B_x \setminus A_x$ sont mesurables pour presque tout x donc A_x aussi ; la fonction $x \mapsto \mu_Y(C_x) = \mu_Y(B_x) - \mu_Y(A_x)$ est mesurable alors $x \mapsto \mu_Y(A_x)$ est mesurable, et

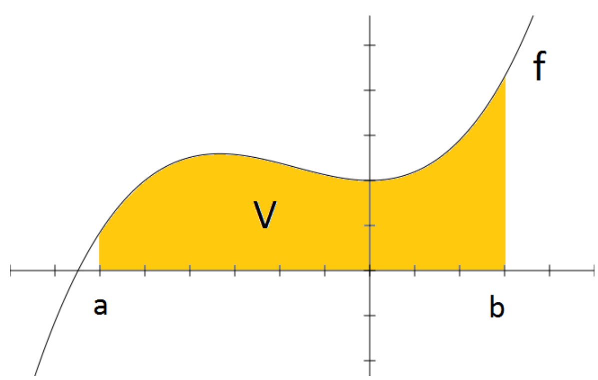
$$\mu(A) = \mu(B) - \mu(C) = \int_X \mu_Y(B_x)d\mu_X - \int_X \mu_Y(C_x)d\mu_X = \int_X \mu_Y(A_x)d\mu_X.$$

□

7.3 Intégrale en tant que la surface sous le graphe

Si f est une fonction positive continue sur $[a,b]$, alors comme il est connu $\int_a^b f(x)dx$ est égale à la surface du domaine

$$V = \{(x,y) \in \mathbb{R}^2 : x \in [a,b],\ 0 \leqslant y \leqslant f(x)\}.$$



Ce fait est vrai pour l'intégrale de Lebesgue sur un ensemble arbitraire.

Soit $(X, \mathcal{F}_X, \mu_X)$ un espace mesuré, f positive et intégrable sur X. Soit $Y = \mathbb{R}$, qu'on considère avec la mesure de Lebesgue, et soit

$$V(f) = \{(x,y) \in X \times \mathbb{R} :\ 0 \leqslant y \leqslant f(x)\}.$$

Si f est étagée, $f = \sum_k y_k I_{A_k}$, alors $V(f) = \cup_k A_k \times [0, y_k]$, ce qui est un ensemble mesurable. Dans le cas général, $f = \lim f_n$ où les fonctions (f_n) sont étagées et $0 \leqslant f_n \leqslant f_{n+1} \leqslant f$ pour tout n. On remarque que $V(f) = \cup_n V(f_n)$, alors $V(f)$ est mesurable.

Pour $x \in X$, l'ensemble $V_x = \{y \in \mathbb{R} : (x,y) \in V(f)\}$ est juste $[0, f(x)]$, car f est positive. Du Théorème 7.2.1 il suit maintenant :

$$\mu\big(V(f)\big) = \int_X \mu_{\mathbb{R}}\big([0, f(x)]\big)d\mu_X(x) = \int_X f d\mu_X. \tag{7.5}$$

En fait, on peut montrer aussi la mesurabilité reciproque : si $V = V(f)$ est mesurable et $f \geqslant 0$, alors f est mesurable. Pour $y \in \mathbb{R}$, on peut remarquer l'égalité

$$V_y = \{x \in X : (x, y) \in V\} = \{x \in X : f(x) \geqslant y\} = f^{-1}\big([y, +\infty[\big).$$

Cet ensemble est mesurable pour presque tout $y \in \mathbb{R}$. Maintenant, pour tout $y \in \mathbb{R}$ on peut choisir une suite **croissante** $y_n \to y$ telle que V_{y_n} est mesurable, alors $f(x) \geqslant y$ si et seulement si $f(x) \geqslant f(y_n)$ pour tout n, et on conclut que $V_y = \cap_n V_{y_n}$ est mesurable aussi.

7.4 Théorème de Fubini

On suppose toujours que $(X, \mathcal{F}_X, \mu_X)$ et $(Y, \mathcal{F}_Y, \mu_Y)$ sont des ensembles mesurés, avec des mesures finies complètes μ_X et μ_Y.

Théorème 7.4.1 (Fubini)**.** *Soit f intégrable sur $X \times Y$ par rapport à la mesure produit $\mu = \mu_X \times \mu_Y$, alors*

$$\begin{aligned}\int_{X\times Y} f d\mu &= \int_X \Big(\int_Y f(x,y) d\mu_Y(y)\Big) d\mu_X(x)\\ &= \int_Y \Big(\int_X f(x,y) d\mu_X(x)\Big) d\mu_Y(y).\end{aligned}$$

Démonstration. L'énoncé sous-entend l'implication que les fonctions $f_x(y) = f(x, y)$ et $f_y(x) = f(x, y)$ sont mesurables pour presque tous x et y respectivement.

La représentation $f = f_+ - f_-$ de la fonction f réduit le théorème au cas $f \geqslant 0$, ce que nous admettons dans la suite. Soit

$$V = \{(x, y, z) \in X \times Y \times \mathbb{R} : \ 0 \leqslant z \leqslant f(x, y)\}.$$

On peut considérer la mesure produit $\tilde{\mu} = \mu_X \times \mu_Y \times \lambda$ sur $X \times Y \times \mathbb{R}$. Par la formule (7.5),

$$\tilde{\mu}(V) = \int_{X\times Y} f d\mu.$$

Mais par Théorème 7.2.1, on a aussi, en représentant $\tilde{\mu} = \mu_X \times (\mu_Y \times \lambda)$,

$$\tilde{\mu}(V) = \int_X (\mu_Y \times \lambda)(V_x) d\mu_X.$$

Pour $x \in X$,

$$V_x = \{(y, z) \in Y \times \mathbb{R} : \ (x, y, z) \in V\} = \{(y, z) \in Y \times \mathbb{R} : \ 0 \leqslant z \leqslant f_x(y)\},$$

donc

$$(\mu_Y \times \lambda)(V_x) = \int_Y f_x d\mu_Y.$$

Il faut noter que les fonctions f_x sont mesurables pour presque tout x, selon la preuve en fin de la section 7.3. La deuxième égalité dans l'énoncé se démontre de la même façon. $\square$

Théorème 7.4.2 (Fubini, condition suffisante)**.** *Soit f mesurable sur $X \times Y$ par rapport à la mesure produit $\mu = \mu_X \times \mu_Y$. Si l'une des intégrales suivantes converge,*

$$I_1 = \int_X \Big(\int_Y |f(x,y)| d\mu_Y(y) \Big) d\mu_X(x),$$
$$I_2 = \int_Y \Big(\int_X |f(x,y)| d\mu_X(x) \Big) d\mu_Y(y),$$

alors f est intégrable sur $X \times Y$.

Démonstration. Sans perte de généralité, on peut admettre que $I_1 < \infty$. Pour $n \in \mathbb{N}$, soit $f_n = \min(|f|, n)$. Cette fonction est mesurable, $0 \leqslant f_n \leqslant |f|$ et $f_n \to |f|$, $n \to \infty$. Par le théorème de Fubini et par l'hypothèse,

$$\int_{X\times Y} f_n d\mu = \int_X \Big(\int_Y f_n(x,y) d\mu_Y(y) \Big) d\mu_X(x) \leqslant I_1.$$

On peut donc appliquer le théorème de Lévi, qui nous dit que $|f| = \lim f_n$ est intégrable. Par définition de l'intégrale, f est donc intégrable. □

Exemple 7.4.3. L'égalité

$$\int_X \Big(\int_Y f(x,y) d\mu_Y(y) \Big) d\mu_X(x) = \int_Y \Big(\int_X f(x,y) d\mu_X(x) \Big) d\mu_Y(y)$$

n'implique pas, en général, que f est intégrable ! Par exemple, soit $X = Y = [-1, 1]$ et

$$f(x,y) = \begin{cases} \dfrac{xy}{(x^2+y^2)^2}, & (x,y) \neq (0,0) \\ 0, & (x,y) = (0,0) \end{cases}.$$

Cette fonction est impaire, alors

$$\int_{-1}^1 \Big(\int_{-1}^1 f(x,y) dx \Big) dy = \int_{-1}^1 \Big(\int_{-1}^1 f(x,y) dy \Big) dx = 0,$$

mais si f était intégrable, l'intégrale suivante serait finie (à noter le changement de variable $z = x^2 + y^2$) :

$$\begin{aligned}\int_{-1}^1 \Big(\int_{-1}^1 |f(x,y)| dx \Big) dy &= 4 \int_0^1 \int_0^1 \frac{xy}{(x^2+y^2)^2} dx dy \\ &= 2 \int_0^1 \int_{y^2}^{y^2+1} \frac{y}{z^2} dz dy = 2 \int_0^1 \Big[-\frac{y}{z} \Big]_{y^2}^{y^2+1} dy \\ &= 2 \int_0^1 \Big(-\frac{y}{y^2+1} + \frac{1}{y} \Big) dy = 2 \int_0^1 \frac{1}{y(y^2+1)} dy,\end{aligned}$$

ce qui n'est pas le cas.

(fin du cours 12)

Chapitre 8

Changement de variables

On connait la formule de changement de variable :

$$\int_{a=\varphi(\alpha)}^{b=\varphi(\beta)} f(x)dx = \int_{\alpha}^{\beta} f\big(\varphi(t)\big)\varphi'(t)dt.$$

On peut utiliser cette formule sans problème si f est continue et φ est C^1.

Pour pouvoir trouver les nouvelles bornes α, β, on suppose souvent l'existence de la fonction inverse φ^{-1} sur $[a, b]$. C'est équivalent à ce que φ est strictement monotone et donc $\varphi'(t)$ est partout non-nulle ; on pose $\alpha = \varphi^{-1}(a)$, $\beta = \varphi^{-1}(b)$.

Notre but est de démontrer la formule analogue dans $\mathbb{R}^n$. L'une des conditions sera l'inversibilité de φ.

8.1 Mesure image

Soit $(X, \mathcal{F}, \mu)$ un espace mesuré, $(Y, \mathcal{G})$ un espace mesurable, et $\varphi : X \to Y$ mesurable. On sait que les images inverses d'une tribu forment une tribu :

$$\varphi^{-1}(\mathcal{G}) = \{\varphi^{-1}(G) : \ G \in \mathcal{G}\}.$$

(À noter, ce n'est pas forcément vrai dans le sens direct.) On peut poser alors

Définition 8.1.1. L'image de la mesure μ est la mesure $\varphi \circ \mu$ sur $(Y, \mathcal{G})$ définie par $(\varphi \circ \mu)(G) = \mu\big(\varphi^{-1}(G)\big)$. On peut la noter aussi $\varphi(\mu)$.

Il faut vérifier que c'est bien une mesure : elle est positive, σ-additive, et $(\varphi \circ \mu)(\emptyset) = 0$.

On peut démontrer le théorème général de changement de variables. Les hypothèses sur X, Y, f sont les mêmes qu'au début de la section.

Théorème 8.1.2. *Une fonction $f : Y \to \mathbb{R}$ est $\varphi \circ \mu$-intégrable si et seulement si $f \circ \varphi$ est μ-intégrable, et dans ce cas*

$$\int_Y f \, d\varphi(\mu) = \int_X (f \circ \varphi) \, d\mu.$$

Démonstration. Soit d'abord $f = I_A$ l'indicatrice d'un ensemble $A \in \mathcal{G}$. On peut vérifier que $f\big(\varphi(x)\big) = 1$ si et seulement si $\varphi(x) \in A$, c'est-à-dire $x \in \varphi^{-1}(A)$, donc $f \circ \varphi = I_{\varphi^{-1}(A)}$. La fonction f est intégrable si et seulement si

$(\varphi \circ \mu)(A) = \mu\big(\varphi^{-1}(A)\big)$ est finie, et donc si et seulement si $f \circ \varphi$ est intégrable. Les intégrales sont égales :

$$\int_Y f\, d(\varphi \circ \mu) = (\varphi \circ \mu)(A) = \mu\big(\varphi^{-1}(A)\big) = \int_X I_{\varphi^{-1}(A)}\, d\mu = \int_X (f \circ \varphi)\, d\mu.$$

Par linéarité, l'égalité vaut aussi pour toute fonction étagée. Ensuite, soit $f \geqslant 0$ mesurable. Il existe une suite (f_n) croissante de fonctions étagées convergente vers f, alors $(f_n \circ \varphi)$ est croissante aussi et converge vers $f \circ \varphi$. Par Théorème 4.3.3 (qui suit du Lemme de Fatou),

$$\int_Y f\, d(\varphi \circ \mu) = \lim_{n\to\infty} \int_Y f_n\, d(\varphi \circ \mu) = \lim_{n\to\infty} \int_X (f_n \circ \varphi)\, d\mu = \int_X (f \circ \varphi)\, d\mu,$$

et on peut noter que les deux intégrales sont finies ou infinies simultanément. Le cas d'une fonction quelconque en suit immédiatement. □

8.2 L'unicité et les translations

Dans la construction de la mesure de Lebesgue, on a implicitement obtenu la propriété suivante :

Proposition 8.2.1. *Soit $A \subset \mathbb{R}^n$ un ensemble mesurable de mesure finie. Alors pour tout $\varepsilon > 0$ il existent un ouvert U et un fermé F tels que $F \subset A \subset U$ et $\lambda(U) < \lambda(A) + \varepsilon$, $\lambda(F) > \lambda(A) - \varepsilon$.*

Démonstration. Par la construction de la mesure extérieure, pour tout $\varepsilon > 0$ il existe une réunion dénombrable $(I_k)_{k\geqslant 1}$ d'intervalles telle que $A \subset \cup_k I_k$ et $\sum_k \lambda(I_k) < \lambda(A) + \varepsilon/2$. Si l'intervalle I_k n'est pas ouvert, ou peut l'agrandir : choisir un intervalle ouvert $J_k \supset I_k$ tel que $\lambda(J_k) < \lambda(I_k) + \varepsilon/2^{k+1}$ (si I_k est ouvert, on pose $J_k = I_k$). La réunion $U = \cup_k J_k$ est ouverte, $A \subset U$ et

$$\lambda(U) \leqslant \sum_{k=1}^{\infty} \lambda(J_k) < \sum_{k=1}^{\infty} \Big(\lambda(I_k) + \frac{\varepsilon}{2^{k+1}}\Big) < \lambda(A) + \varepsilon.$$

Pour l'autre inclusion, il faut faire attention au cas où A ne serait pas contenu dans un fermé de mesure finie (c'est possible, il suffit de considérer $\mathbb{Q}^n$ qui est de mesure nulle mais dense dans $\mathbb{R}^n$).

On peut représenter $\mathbb{R}^n = \cup_k B_k$ comme la réunion de cubes fermés de côté 1 qui n'ont en commun que leurs faces ; la réunion Φ de cettes faces a la mesure nulle. Pour tout k, l'ensemble $B_k \setminus A$ est mesurable de mesure finie, alors il existe un ouvert V_k qui le contient et dont la mesure est $\lambda(V_k) < \lambda(B_k \setminus A) + \varepsilon/2^k$. L'ensemble $F_k = B_k \setminus V_k = B_k \cap (\mathbb{R}^n \setminus V_k)$ est fermé, contenu dans A et

$$\begin{aligned}\lambda(F_k) &= \lambda(B_k) - \lambda(B_k \cap V_k) \geqslant \lambda(B_k) - \lambda(V_k)\\ &> \lambda(B_k) - \lambda(B_k \setminus A) - \frac{\varepsilon}{2^k} = \lambda(B_k \cap A) - \frac{\varepsilon}{2^k}.\end{aligned}$$

La réunion $F = \cup_k F_k$ est fermée : ce n'est pas toujours le cas d'une réunion de fermés, mais si une suite $(x_m) \subset F$ converge, elle est bornée et est contenue en fait dans une réunion finie de F_k (qui sont contenus chacun dans un cube borné). On a aussi $F \subset A$ et

$$\lambda(F) = \lambda(\cup_k F_k) \geqslant \sum_k \lambda(F_k) - \lambda(\Phi) > \sum_k \lambda(B_k \cap A) - \frac{\varepsilon}{2^k} = \lambda(A) - \varepsilon.$$

□

Corollaire 8.2.2 (Unicité de la mesure de Lebesgue)**.** *Si μ est une mesure sur la tribu $\mathcal{F}(\mathbb{R}^n)$ des ensembles mesurables et μ coïncide avec λ sur les ouverts, alors $\mu = \lambda$.*
La même conclusion vaut si μ coïncide avec λ sur les intervalles.

Démonstration. Supposons μ et λ coIncident sur les ouverts. Pour un ensemble $A \in \mathcal{F}(\mathbb{R}^n)$ de mesure finie, on choisit pour tout $\varepsilon > 0$ un ouvert U et un fermé F tels que $F \subset A \subset U$ et $\lambda(F) + \varepsilon > \lambda(A) > \lambda(U) - \varepsilon$.

Par l'hypothèse, $\mu(U) = \lambda(U)$; comme $U \setminus F$ est ouvert, on a aussi $\mu(U \setminus F) = \lambda(U \setminus F)$ et donc $\mu(F) = \lambda(F)$. Ensuite,

$$\mu(A) \leqslant \mu(U) = \lambda(U) < \lambda(A) + \varepsilon < \lambda(F) + 2\varepsilon = \mu(F) + 2\varepsilon \leqslant \mu(A) + 2\varepsilon,$$

ce qui implique (comme ε est arbitraire) $\mu(A) = \lambda(A)$.

Pour le deuxième cas, il suffit de remarquer que tout ouvert se représente comme la réunion d'une famille dénombrable de cubes ouverts (C_k), et qu'on peut en passer à une famille d'intervalles disjoints par le procédé standard (considérer $C_k \setminus \cup_{j<k} C_j$ puis décomposer en intervalles si besoin). $\square$

On peut en déduire que λ est invariante par les translations. Soit $a \in \mathbb{R}^n$ et $\mu(A) = \lambda(A - a)$; à rappeler, on définit

$$A - a = \{x - a : \ x \in A\}.$$

Il est évident que μ est une mesure sur la même tribu, mais on sait que aussi la mesure d'un intervalle ne change pas sous les translations, donc $\mu = \lambda$.

8.3 Cas d'un opérateur linéaire

Nous allons calculer la mesure image $\Phi(\lambda)$ où Φ est un opérateur linéaire sur $\mathbb{R}^n$ et λ est la mesure de Lebesgue. On peut représenter Φ comme une matrice dans $M_n(\mathbb{R})$.

Théorème 8.3.1. *Soit Φ un opérateur linéaire inversible sur $\mathbb{R}^n$ et λ est la mesure de Lebesgue, alors*

$$\Phi(\lambda) = |\det \Phi|^{-1} \lambda. \tag{8.1}$$

Démonstration. On peut déjà remarquer que toute application mesurable inversible envoie la tribu des ensembles mesurables dans elle-même. Dans la définition de la mesure intervient l'application inverse $\Psi = \Phi^{-1}$; c'est aussi un opérateur linéaire, dont la matrice est l'inverse de celle de Φ. Selon la définition de la mesure image et les résultats de la Section 8.2, il suffit de vérifier que

$$\lambda\big(\Psi(I)\big) = |\det \Psi| \lambda(I) = |\det \Phi|^{-1} \lambda(I) \tag{8.2}$$

pour tout intervalle I.

Une matrice inversible se décompose en produit de matrices élémentaires :

- ⋆ transpositions T_{kl}, $k \neq l$: $T_{kl}(e_k) = e_l$, $T_{kl}(e_l) = e_k$, $T_{kl}(e_j) = e_j$, $j \notin \{k, l\}$;
- ⋆ dilatations $D_{k,r}$, $r \in \mathbb{R}^*$: $D_{k,r}(e_k) = re_k$, $D_{k,r}(e_j) = e_j$, $j \neq k$;
- ⋆ sommes $S_{k,l}$, $k \neq l$: $S_{k,l}(e_l) = e_l + e_k$, $S_{k,l}(e_j) = e_j$, $j \neq l$.

Ce fait revient à la méthode de Gauss de résolution de systèmes des équations linéaires.

Si la formule (8.2) est démontrée pour chaque matrice élémentaire, alors on en déduira facilement (8.2) pour Ψ (le déterminant d'un produit de matrices est le produit des déterminants).

Si $\Psi = T_{kl}$, la mesure de chaque intervalle ne change pas, alors par l'unicité (Section 8.2) $\Phi \circ \lambda = \lambda$. Au même temps, le determinant de la matrice T_{kl} est 1 ou -1, le résultat correspond donc à la formule (8.2).

Soit $\Psi = D_{k,r}$. Un intervalle $I =]\!]a_1, b_1]\!] \times]\!]a_2, b_2]\!] \times \cdots \times]\!]a_d, b_d]\!]$ est envoyé dans l'intervalle

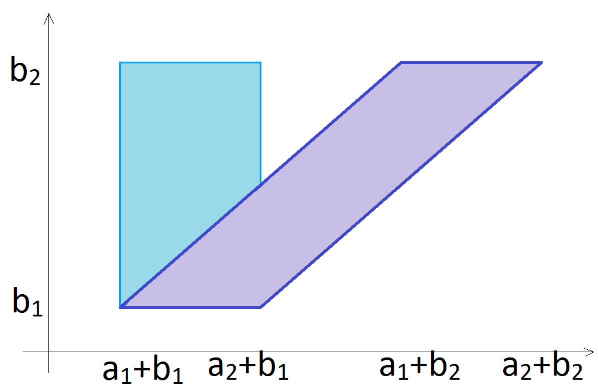
$$D_{k,r}(I) =]\!]a_1, b_1]\!] \times]\!]a_2, b_2]\!] \times \cdots \times]\!]ra_k, rb_k]\!] \times \cdots \times]\!]a_d, b_d]\!],$$

dont la mesure est (attention à la possibilité $r < 0$)

$$\lambda\big(D_{k,r}(I)\big) = |rb_k - ra_k| \prod_{j \neq k} (b_j - a_j) = |r|\lambda(I).$$

Comme $\det \Psi = r$, l'égalité (8.2) et vérifiée.

Enfin, soit $\Psi = S_{k,l}$. En coordonnées de $\mathbb{R}^n$, sa formule est $S_{kl}(x_1, \ldots, x_n) = (x_1, \ldots, x_{k-1}, x_k + x_l, x_{k+1}, \ldots, x_n)$. Pour comprendre la transformation d'un intervalle I, il suffit de considérer le plan $\langle e_k, e_l \rangle$, et on peut admettre $k = 1$, $l = 2$ sans perte de généralité. Le rectangle $]\!]a_1, b_1]\!] \times]\!]a_2, b_2]\!]$ est envoyé dans le parallèlogramme aux sommets $(a_i + b_j, b_j)$, $i, j = 1, 2$ (voir dessin).



L'aire d'un parallèlogramme est le produit de sa base et son hauteur, elle ne change donc pas sous cette transformation ; ceci montre que le volume de I ne change pas, et $\Psi \circ \lambda = \lambda$. On a $\det \Psi = 1$, donc la formule (8.2) est vérifiée, ce qui termine la preuve. □

8.4 La formule générale

Théorème 8.4.1 (Changement de variables)**.** *Soient $U \subset \mathbb{R}^n$ un ouvert, $\varphi : U \to \mathbb{R}^n$ injective de classe C^1, telle que $d\varphi(x)$ est inversible pour tout $x \in U$. Alors pour tout ensemble mesurable $E \subset U$ et toute f intégrable sur $\varphi(E)$*

$$\int_{\varphi(E)} f(y)dy = \int_E f \circ \varphi(x)|J_\varphi(x)|dx,$$

où dx, dy signifie la mesure de Lebesgue dans $\mathbb{R}^n$ et $J_\varphi(x) = \det d\varphi(x)$.

Démonstration. Si on note par μ la mesure à densité $|J_\varphi(x)|$, l'énoncé prend la forme
$$\int_{\varphi(E)} f d\lambda = \int_E (f\circ\varphi)\, d\mu.$$
Supposons qu'on a montré que $\varphi(\mu) = \lambda$; l'égalité suivra alors du Théorème 8.1.2, avec $X = E$, $Y = \varphi(E)$.

Pour démontrer l'égalité $\varphi(\mu) = \lambda$, il suffit de montrer que
$$\mu\big(\varphi^{-1}(U)\big) \equiv \int_{\varphi^{-1}(U)} |J_\varphi(x)| d\lambda(x) = \lambda(U)$$
pour un ouvert U. Par la continuité de φ, c'est équivalent à montrer que
$$\int_I |J_\varphi(x)| d\lambda(x) = \lambda\big(\varphi(I)\big)$$
pour un ouvert I, et par la σ-additivité des deux termes il suffit de le démontrer pour un intervalle fermé I.

Nous allons le faire d'abord localement. Pour la suite, il nous sera plus pratique de supposer que la norme de $\mathbb{R}^n$ est $\|x\|_\infty = \max\{|x_1|, \dots, |x_n|\}$. Supposons I borné, et soit
$$\ell = \max_{x\in I} \|d\varphi(x)^{-1}\|;$$
ce nombre est fini par continuité de $d\varphi$ et par la compacité de I.

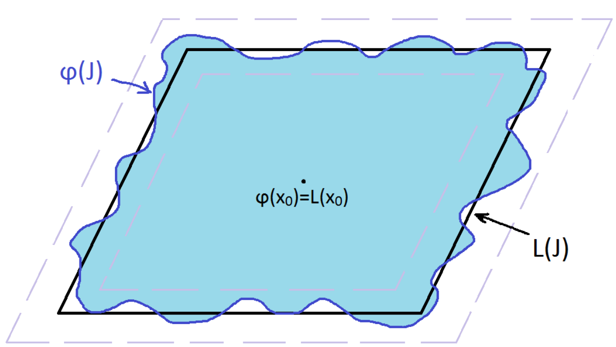
Lemme 8.4.2. *Soit I un intervalle fermé borné. Pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que pour tout intervalle $J \subset I$ centré en $x_0 \in I$ de côté inférieur à δ,*
$$\Big|\lambda\big(\varphi(J)\big) - |J_\varphi(x_1)|\lambda(J)\Big| < \varepsilon\ell|\lambda(J)|.$$

Démonstration. Par la formule de Taylor, pour $x_0, x \in I$ on a
$$\varphi(x) = \varphi(x_0) + d\varphi(x_0)(x - x_0) + \alpha_{x_0}(x),$$
où $\alpha_{x_0}(x) = o(\|x - x_0\|)$ quand $x \to x_0$. Pour $\varepsilon > 0$, soit $\delta > 0$ tel que $\|x - x_0\| < \delta$ implique $|\alpha_{x_0}(x)| < \varepsilon\|x - x_0\|$.

On peut montrer — ceci appartient plutôt au cours du calcul différentiel — qu'un tel δ peut être choisi le même pour tous $x_0 \in I$ (ceci suit des applications répétées n fois du théorème des accroissements finis, de la continuité des dérivées partielles et de la compacité de I).

Soit J un intervalle de côté δ centré en x_0 ; il est alors contenu dans la boule $\{x : \|x - x_0\| \leqslant \delta/2\}$ et contient x_0 dans son intérieur.

Notons $L_0 = d\varphi(x_0)$; à noter, $\|L_0^{-1}\| \leqslant \ell$. L'application affine $L(x) = \varphi(x_0) + L_0(x - x_0)$ transforme J dans un *paralléloèdre* dont le volume on peut



calculer par Théorème 8.3.1, en utilisant l'invariance de λ par translations :

$$\lambda\big(L(J)\big) = \lambda\big(L_0(J - x_0)\big) = |\det L_0|\lambda(J) = |J_\varphi(x_0)|\lambda(J).$$

L'image de J sous l'application φ est proche de $L(J)$, dans le sens suivant. Pour $x \in J$, on a $L(x) - L(x_0) = L_0(x - x_0)$ et alors $x - x_0 = L_0^{-1}\big(L(x) - L(x_0)\big)$. On peut majorer

$$\begin{aligned}\|\varphi(x) - \varphi(x_0)\| = \|L(x) - L(x_0) + \alpha_{x_0}(x)\| &< \|L(x) - L(x_0)\| + \varepsilon\|x - x_0\| \\ &\leqslant (1 + \varepsilon\ell)\|L(x) - L(x_0)\|,\end{aligned}$$

alors $\varphi(J) \subset (1 + \varepsilon\ell)\big(L(J) - L(x_0)\big)$ et

$$\lambda\big(\varphi(J)\big) \leqslant (1 + \varepsilon\ell)\lambda\big(L(J)\big) = (1 + \varepsilon\ell)\,|J_\varphi(x_0)|\,\lambda(J).$$

Au même temps, on a $\|\varphi(x) - \varphi(x_0)\| > (1 - \varepsilon\ell)\|L(x) - L(x_0)\|$. Cela ne montre pas encore que $\varphi(J) - \varphi(x_0)$ contient $U = (1 - \varepsilon\ell)\big(L(J) - L(x_0)\big)$, mais montre au moins que l'image de la frontière de J n'intersecte pas S. Pour montrer que $S \subset \varphi(J)$, il faut des arguments topologiques sur la connexité [6, Théorème 8.24] que nous admettons ; on obtient alors

$$\lambda\big(\varphi(J)\big) > (1 - \varepsilon\ell)\lambda\big(L(x) - L(x_0)\big) = (1 - \varepsilon\ell)|J_\varphi(x_0)|\lambda(J).$$

□

La suite de la démonstration du théorème. Pour $\varepsilon > 0$, soit $\delta > 0$ comme dans le Lemme. On peut diviser I en réunion disjointe $I = \sqcup_k J_k$ d'intervalles de côté inférieur à δ. Grâce à bijectivité de φ, on a également $\varphi(I) = \sqcup_k \varphi(J_k)$. En utilisant le Lemme, on peut minorer

$$\lambda\big(\varphi(I)\big) = \sum_k \lambda\big(\varphi(J_k)\big) > (1 - \varepsilon\ell)\sum_k |J_\varphi(x_k)|\lambda(J_k).$$

À droite nous retrouvons une somme intégrale (dans le sens de Riemann), qui converge vers $\displaystyle\int_I |J_\varphi(x)|dx$ quand $\delta \to +0$, donc $\lambda\big(\varphi(I)\big) \geqslant \displaystyle\int_I |J_\varphi(x)|dx$. De la même façon, on obtient $\lambda\big(\varphi(I)\big) \leqslant \int_I |J_\varphi(x)|dx$, ce qui termine la preuve. □

——— (fin du cours 13) ———

Bibliographie

[1] J.-P. Ansel, Y. Ducel, Exercices corrigés en théorie de la mesure et de l'intégration, Ellipses, 2015.

[2] M. Briane, G. Pagès, Théorie de l'intégration, Vuibert, 2012.

[3] J.B. Bruckner, A.M. Bruckner, B.S. Thomson, Real Analysis, CreateSpace, 2008, téléchargeable depuis le site des auteurs : `http://classicalrealanalysis.info/Free-Downloads.php`

[4] H. Bauer, Measure and Integration Theory, de Gruyter, 2001.

[5] A. Kolmogorov, S. Fomine, Eléments de la théorie de fonctions et de l'analyse fonctionnelle, Ellipses, 1994.

[6] W. Rudin, Real and complex analysis, 1966.