

# Mesure, Intégration, Fourier

Benjamin Charlier, Philippe Castillon

7 décembre 2023

# Table des matières

<b>1</b>	<b>Quelques Rappels et Motivations</b>	<b>7</b>
1.1	Rudiments de théorie des ensembles	7
1.1.1	Opérations sur $\mathcal{P}(X)$	7
1.1.2	Images directes et réciproques	8
1.1.3	Dénombrabilité	8
1.2	L'ordre sur $\mathbb{R}$	9
1.2.1	La droite réelle achevée	9
1.2.2	Limites inférieures et limites supérieures	10
1.3	Motivations	10
1.3.1	En géométrie	11
1.3.2	En analyse	12
1.3.3	En probabilité	12
<b>2</b>	<b>Théorie générale de la mesure</b>	<b>13</b>
2.1	Espaces mesurables	13
2.1.1	Tribu, tribu engendrée	13
2.1.2	Les boréliens	14
2.2	Applications mesurables	16
2.2.1	Définition	16
2.2.2	Propriétés	19
2.3	Espaces mesurés	20
2.3.1	Mesure positive	20
2.3.2	Propriétés	21
2.4	Exemples de mesures	25
2.4.1	Mesure image	25
2.4.2	Mesure de Lebesgue	25
2.5	Espaces de probabilité	26
2.6	Le "presque partout"	27
2.6.1	Ensembles négligeables	27
2.6.2	Ensemble de Cantor	28
<b>3</b>	<b>Théorie générale de l'intégration</b>	<b>31</b>
3.1	Fonctions étagées	31
3.1.1	Définition et propriétés	31
3.1.2	Intégrale des fonctions étagées	32
3.2	Intégrale des fonctions positives	34
3.2.1	Définition	34
3.2.2	Propriétés	36
3.2.3	Théorème de convergence monotone	36
3.2.4	Propriétés et lemme de Fatou	38
3.3	Intégrales des fonctions réelles et complexes	39
3.3.1	Définitions	39
3.3.2	Théorème de convergence dominée	42
3.4	Intégrales dépendant d'un paramètre	44
3.4.1	Continuité	45

3.4.2	Dérivabilité . . . . .	45
<b>4</b>	<b>Intégration pour quelques mesures particulières</b>	<b>49</b>
4.1	Mesure de Dirac et mesure de comptage sur $\mathbb{N}$ . . . . .	49
4.1.1	Mesure de Dirac . . . . .	49
4.1.2	Mesure de comptage sur $\mathbb{N}$ . . . . .	50
4.2	Riemann vs. Lebesgue . . . . .	51
4.2.1	Sur un segment . . . . .	51
4.2.2	Le cas des intégrales impropres . . . . .	52
4.3	Le théorème de transfert . . . . .	54
4.3.1	Intégration et mesure image . . . . .	54
4.3.2	Un peu de probabilités . . . . .	55
4.4	Intégration sur les espaces produits . . . . .	57
4.4.1	Tribu et mesure produit . . . . .	57
4.4.2	Théorèmes de Tonelli et Fubini . . . . .	61
4.5	Un exemple à connaître : la fonction $\Gamma$ . . . . .	64
<b>5</b>	<b>Les espaces <math>L^p</math> et la mesure de Lebesgue sur <math>\mathbb{R}^n</math></b>	<b>67</b>
5.1	Espaces $\mathcal{L}^p$ et $L^p$ . . . . .	67
5.1.1	Les semi-normes $\ \cdot\ _p$ . . . . .	67
5.1.2	Les espaces $L^p$ . . . . .	70
5.2	La mesure de Lebesgue sur $\mathbb{R}^n$ . . . . .	75
5.2.1	Invariance par translation . . . . .	76
5.2.2	Changement de variable affine . . . . .	76
5.2.3	Changement de variable . . . . .	77
5.3	Produit de convolution et théorèmes de densité . . . . .	80
5.3.1	Densité des fonctions continues à support compact . . . . .	81
5.3.2	Le produit de convolution . . . . .	85
5.3.3	Densité des fonctions $C^\infty$ à support compact . . . . .	88
<b>6</b>	<b>Analyse de Fourier sur <math>\mathbb{R}</math></b>	<b>93</b>
6.1	Définition et propriétés . . . . .	93
6.1.1	Transformée de Fourier et dérivation . . . . .	95
6.1.2	Produit de convolution et formule d'inversion . . . . .	97
6.2	Le cadre fonctionnel . . . . .	99
6.2.1	Espace de Schwartz . . . . .	99
6.2.2	Transformée de Fourier $L^2$ . . . . .	100
<b>7</b>	<b>Existence et unicité de la mesure de Lebesgue</b>	<b>103</b>
7.1	Mesures extérieures . . . . .	103
7.1.1	Existence de la mesure de Lebesgue . . . . .	105
7.1.2	Complétion de mesure . . . . .	107
7.2	Classes monotones . . . . .	108
7.2.1	Caractérisation de la mesure de Lebesgue et autres conséquences . . . . .	110



# Avant Propos

Ce document a été initialement écrit par P. Castillon (Univ. Montpellier) et grandement inspiré du polycopié [2] de T. Gallay (Univ. Grenoble). Ce sont ajoutées diverses sources au fil des rédactions : le très complet livre [3] de T. Gallouët et R. Herbin, le plus avancé (et un peu daté) livre [4] de H.L. Royden, les très rigoureuses (et duales) notes [1] de F. De Marçay/J. Merker (Univ. Paris Saclay).



# Chapitre 1

## Quelques Rappels et Motivations

### 1.1 Rudiments de théorie des ensembles

Si  $X$  est un ensemble, une **partie** (ou un **sous-ensemble**) de  $X$  est un ensemble  $A$  vérifiant  $A \subset X$ , c'est à dire  $\forall x \in A \ x \in X$ . Il existe un ensemble noté  $\mathcal{P}(X)$  dont les éléments sont exactement les parties de  $X$  :

$$\mathcal{P}(X) = \{A \mid A \subset X\}.$$

En particulier, on a toujours  $\emptyset \in \mathcal{P}(X)$ ,  $X \in \mathcal{P}(X)$ , et  $\emptyset \subset A \subset X$  pour tout  $A \in \mathcal{P}(X)$ .

Si  $X$  et  $I$  sont des ensembles, une **famille d'éléments** de  $X$  indexée par  $I$  est une application de  $I$  dans  $X$  notée  $(x_i)_{i \in I}$ , où  $x_i$  est l'image de  $i \in I$  par cette application. Par exemple :

- un  $n$ -uplet  $(x_1, \dots, x_n)$  de réels, avec  $n \in \mathbb{N}^*$ , est une famille de  $\mathbb{R}$  indexée par  $\{1, \dots, n\}$  ;
- une suite réelle  $(u_n)_{n \in \mathbb{N}}$  est une famille de  $\mathbb{R}$  indexée par  $\mathbb{N}$  ;
- $([x, x + x^2])_{x \in \mathbb{R}}$  est une famille de  $\mathcal{P}(\mathbb{R})$  indexée par  $\mathbb{R}$ .

De façon générale, si l'ensemble d'indexation est  $I = \mathbb{N}$  on parlera indifféremment de famille de  $X$  indexée par  $\mathbb{N}$  ou de **suite** de  $X$ .

Dans la suite, on s'intéressera beaucoup à des parties de  $\mathcal{P}(X)$ , c'est à dire à des ensembles dont les éléments sont des parties de  $X$ . On appellera **classe** sur  $X$  une partie de  $\mathcal{P}(X)$ . Ainsi, si  $\mathcal{C} \subset \mathcal{P}(X)$  est une classe sur  $X$  et  $A \in \mathcal{C}$ , alors  $A \subset X$ .

#### 1.1.1 Opérations sur $\mathcal{P}(X)$

L'ensemble  $\mathcal{P}(X)$  est muni de plusieurs opérations :

- le **complémentaire** d'un élément  $A \in \mathcal{P}(X)$  est  $A^c$  défini par

$$\forall x \in X \quad x \in A^c \Leftrightarrow x \notin A.$$

- la **réunion** d'une famille  $(A_i)_{i \in I}$  de  $\mathcal{P}(X)$  est la partie

$$\bigcup_{i \in I} A_i = \{x \in X \mid \exists i \in I, x \in A_i\} \in \mathcal{P}(X).$$

Pour une famille finie  $(A_1, A_2, \dots, A_n)$ , on note  $A_1 \cup A_2 \cup \dots \cup A_n$  sa réunion.

- l'**intersection** d'une famille  $(A_i)_{i \in I}$  de  $\mathcal{P}(X)$  est la partie

$$\bigcap_{i \in I} A_i = \{x \in X \mid \forall i \in I, x \in A_i\} \in \mathcal{P}(X).$$

Pour une famille finie  $(A_1, A_2, \dots, A_n)$ , on note  $A_1 \cap A_2 \cap \dots \cap A_n$  son intersection.

- la **différence** de deux parties  $A, B \in \mathcal{P}(X)$  est

$$A \setminus B = A \cap B^c = \{x \in A \mid x \notin B\}.$$

- la **différence symétrique** de deux parties  $A, B \in \mathcal{P}(X)$  est

$$A \triangle B = (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B).$$

Parmi les nombreuses propriétés de ces opérations, on utilisera souvent le fait que le passage au complémentaire “échange” les opérations de réunion et intersection : pour toute famille  $(A_i)_{i \in I}$  de  $\mathcal{P}(X)$  on a

$$\left( \bigcup_{i \in I} A_i \right)^c = \bigcap_{i \in I} A_i^c \quad \text{et} \quad \left( \bigcap_{i \in I} A_i \right)^c = \bigcup_{i \in I} A_i^c.$$

L’inclusion est une relation d’ordre (partielle) sur  $\mathcal{P}(X)$ , on peut donc parler de **suites monotones** d’ensemble. Une suite  $(A_k)_{k \in \mathbb{N}}$  de  $\mathcal{P}(X)$  est

- croissante** si  $\forall k \in \mathbb{N} \quad A_k \subset A_{k+1}$ ,
- décroissante** si  $\forall k \in \mathbb{N} \quad A_{k+1} \subset A_k$ .

### 1.1.2 Images directes et réciproques

Soient  $X, Y$  deux ensembles et  $f : X \rightarrow Y$  une application. Si  $A \in \mathcal{P}(X)$ , l’**image directe** de  $A$  par  $f$  est

$$f(A) = \{y \in Y \mid \exists x \in A, f(x) = y\} = \{f(x) \mid x \in A\} \in \mathcal{P}(Y).$$

Si  $B \in \mathcal{P}(Y)$ , l’**image réciproque** de  $B$  par  $f$  est

$$f^{-1}(B) = \{x \in X \mid f(x) \in B\} \in \mathcal{P}(X).$$

**Remarque 1.1.1.** Ces notations peuvent prêter à confusion...

- On note de la même façon l’image par  $f$  d’un élément de  $X$  et l’image directe d’une partie de  $X$ , alors que ce sont deux choses différentes. En effet, étant donné  $x \in X$  on a  $f(x) \in Y$  alors que  $f(\{x\}) = \{f(x)\} \in \mathcal{P}(Y)$ . Il convient donc d’être très attentif à l’argument entre parenthèse pour savoir de quel objet on parle...
- La même observation est valable pour l’image réciproque, avec cette remarque supplémentaire : on ne suppose rien sur  $f$  pour définir l’image réciproque d’une partie, en particulier, si  $f$  n’est pas bijective sa fonction réciproque n’est même pas définie, alors que l’image réciproque d’une partie l’est bel et bien.

L’image réciproque se comporte bien avec les opérations sur  $\mathcal{P}(X)$  et  $\mathcal{P}(Y)$ . Pour toute famille  $(B_i)_{i \in I}$  de  $\mathcal{P}(Y)$  on a

$$f^{-1}\left(\bigcup_{i \in I} B_i\right) = \bigcup_{i \in I} f^{-1}(B_i) \quad \text{et} \quad f^{-1}\left(\bigcap_{i \in I} B_i\right) = \bigcap_{i \in I} f^{-1}(B_i).$$

De plus, pour tout  $B \in \mathcal{P}(Y)$  on a également  $f^{-1}(B^c) = f^{-1}(B)^c$ .

Avec l’image directe il faut faire attention ! Pour toute famille  $(A_i)_{i \in I}$  de  $\mathcal{P}(X)$  on a

$$f\left(\bigcup_{i \in I} A_i\right) = \bigcup_{i \in I} f(A_i) \quad \text{et} \quad f\left(\bigcap_{i \in I} A_i\right) \subset \bigcap_{i \in I} f(A_i)$$

et on ne peut rien dire a priori sur l’image directe du complémentaire d’une partie de  $X$ .

### 1.1.3 Dénombrabilité

Deux ensembles  $X$  et  $Y$  sont **équipotents** s’il existe une bijection  $f : X \rightarrow Y$ . Un ensemble  $X$  non vide est **dénombrable** s’il est équipotent à une partie de  $\mathbb{N}$ . Les trois assertions suivantes sont équivalentes :

- $X$  est dénombrable,
- il existe une application injective  $f : X \rightarrow \mathbb{N}$ ,



- il existe une application surjective  $f : \mathbb{N} \rightarrow X$ .

En utilisant la définition et ces propriétés, on montre que

1. toute partie d'un ensemble dénombrable est dénombrable,
2. une réunion finie d'ensembles dénombrables est dénombrable,
3. si  $X$  et  $Y$  sont dénombrables, alors  $X \times Y$  est dénombrable,
4. une réunion dénombrable d'ensembles dénombrables est dénombrable : si  $I$  est dénombrable et  $(X_i)_{i \in I}$  est une famille d'ensembles dénombrables, alors  $\bigcup_{i \in I} X_i$  est dénombrable.

En particulier,  $\mathbb{N}$ ,  $\mathbb{Z}$  et  $\mathbb{Q}$  sont dénombrables. Par contre  $\mathbb{R}$  et  $\mathcal{P}(\mathbb{N})$  ne sont pas dénombrables (de façon générale,  $X$  et  $\mathcal{P}(X)$  ne sont pas équipotents).

## 1.2 L'ordre sur $\mathbb{R}$

Un réel  $M$  est un majorant d'une partie  $A \in \mathcal{P}(\mathbb{R})$  si  $\forall x \in A, x \leq M$ . Le réel  $M$  est le plus petit des majorants de  $A$  si  $M$  est un majorant de  $A$  et si pour tout majorant  $P$  de  $A$  on a  $M \leq P$ . S'il existe, ce plus petit majorant est unique, on l'appelle la **borne supérieure** (ou **supremum**) de  $A$  et on le note  $\sup(A)$ .

On définit de façon analogue un minorant, le plus grand minorant d'une partie  $A \in \mathcal{P}(\mathbb{R})$  et la **borne inférieure** (ou **infimum**) de  $A$ .

L'ensemble des nombres réels a la propriété fondamentale suivante : *toute partie de  $\mathbb{R}$  non vide et majorée (resp. minorée) possède une borne supérieure (resp. une borne inférieure).*

**Remarque 1.2.1.** Ce résultat est une conséquence d'une propriété topologique de  $\mathbb{R}$  (la complétude). Par exemple, il est faux dans  $\mathbb{Q}$  où l'ensemble  $\{q \in \mathbb{Q} | q^2 < 2\}$  n'admet pas de supremum.

Il a comme conséquence le résultat suivant.

Si  $(x_k)_{k \in \mathbb{N}}$  est une suite réelle croissante, alors on a l'alternative suivante :

- soit la suite est bornée et alors elle converge,  $\lim_{k \rightarrow +\infty} x_k = \sup_{k \in \mathbb{N}} (x_k)$  ;
- soit la suite n'est pas bornée et alors  $\lim_{k \rightarrow +\infty} x_k = +\infty$ .

De façon analogue, une suite décroissante  $(x_k)_{k \in \mathbb{N}}$  converge vers  $\inf_{k \in \mathbb{N}} (x_k)$  si elle est minorée et tend vers  $-\infty$  sinon.

### 1.2.1 La droite réelle achevée

Pour unifier les deux issues des alternatives ci-dessus, on rajoute deux éléments à  $\mathbb{R}$  : la **droite réelle achevée** est

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

L'ordre usuel sur  $\mathbb{R}$  est étendu à  $\overline{\mathbb{R}}$  par  $\forall x \in \overline{\mathbb{R}} \quad -\infty \leq x \leq +\infty$ . La notion de convergence d'une suite  $(x_k)_{k \in \mathbb{N}}$  est étendue par les définition usuelles de  $\lim_{k \rightarrow +\infty} x_k = +\infty$  et  $\lim_{k \rightarrow +\infty} x_k = -\infty$ . Dans la mesure où  $+\infty \in \overline{\mathbb{R}}$ , on peut bien dire que la suite converge vers  $+\infty$ .

Compte tenu de la définition de l'ordre dans  $\overline{\mathbb{R}}$ , toute partie non vide  $A \subset \overline{\mathbb{R}}$  possède un majorant  $(+\infty)$  et un minorant  $(-\infty)$ . La propriété fondamentale devient alors : *toute partie non vide de  $\overline{\mathbb{R}}$  possède une borne supérieure et une borne inférieure.*

Pour une suite, il n'y a plus d'alternative : *toute suite monotone de  $\overline{\mathbb{R}}$  converge.* En particulier, si  $(a_k)_{k \in \mathbb{N}}$  est une suite de réels positifs, alors les sommes partielles de la série  $\sum_{k \in \mathbb{N}} a_k$  forment une suite croissante qui est donc convergente dans  $\overline{\mathbb{R}}_+$ . La somme de la série est bien définie :  $\sum_{k=0}^{+\infty} a_k \in \overline{\mathbb{R}}_+$ . Par ailleurs, la valeur de la somme ne dépend pas de l'ordre de sommation, et si  $(a_{i,j})_{(i,j) \in \mathbb{N}^2}$  est une famille indexée par  $\mathbb{N}^2$  de réels positifs, on a

$$\sum_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} a_{i,j} = \sum_{j \in \mathbb{N}} \sum_{i \in \mathbb{N}} a_{i,j}.$$

Si on a pu étendre l'ordre de façon satisfaisante à  $\overline{\mathbb{R}}$ , on ne peut cependant pas étendre les opérations usuelles de  $\mathbb{R}$  à  $\overline{\mathbb{R}}$ . En particulier, la somme  $(+\infty) + (-\infty)$  n'est pas définie... Dans  $\overline{\mathbb{R}}_+ = \mathbb{R}_+ \cup \{+\infty\}$ , on conviendra que

- $\forall a \in \overline{\mathbb{R}}_+, a + (+\infty) = +\infty + a = +\infty$ ;
- $\forall a \in \overline{\mathbb{R}}_+ \setminus \{0\}, a \cdot (+\infty) = +\infty$ ;

Mais attention : les éléments  $+\infty$  et  $-\infty$  ne sont pas opposés l'un de l'autre pour l'addition. Ainsi, si  $a, b \in \overline{\mathbb{R}}_+$  vérifient  $a + \infty = b + \infty$ , on ne peut pas en conclure que  $a = b$  (de même que pour  $x, y \in \mathbb{R}$ , le fait que  $x \cdot 0 = y \cdot 0$  ne permet pas de conclure que  $x = y$ ...)

Comme vous avez appris à ne pas diviser par 0, il faut prendre la bonne habitude de ne jamais retrancher  $+\infty$ , et de ne jamais diviser par  $+\infty$ .

## 1.2.2 Limites inférieures et limites supérieures

Soit  $(x_k)_{k \in \mathbb{N}}$  une suite  $\mathbb{R}$ . Pour tout  $n \in \mathbb{N}$  on note

$$a_n = \inf\{x_k \mid k \geq n\} \quad \text{et} \quad b_n = \sup\{x_k \mid k \geq n\}.$$

Comme  $\{x_k \mid k \geq n+1\} \subset \{x_k \mid k \geq n\}$ , on a  $a_{n+1} \geq a_n$  et  $b_{n+1} \leq b_n$ . Les suites  $(a_n)_{n \in \mathbb{N}}$  et  $(b_n)_{n \in \mathbb{N}}$  étant monotones, elles convergent dans  $\overline{\mathbb{R}}$ .

La **limite inférieure** de  $(x_k)_{k \in \mathbb{N}}$  est

$$\liminf_{k \rightarrow +\infty} (x_k) = \lim_{n \rightarrow +\infty} a_n = \sup_{n \in \mathbb{N}} \left( \inf_{k \geq n} (x_k) \right) \in \overline{\mathbb{R}}$$

La **limite supérieure** de  $(x_k)_{k \in \mathbb{N}}$  est

$$\limsup_{k \rightarrow +\infty} (x_k) = \lim_{n \rightarrow +\infty} b_n = \inf_{n \in \mathbb{N}} \left( \sup_{k \geq n} (x_k) \right) \in \overline{\mathbb{R}}$$

**Remarque 1.2.2.** Lorsqu'il n'y a pas d'ambiguïté, et suivant les auteurs, on omet les indices  $\limsup_{k \rightarrow +\infty} x_k = \limsup_k x_k = \lim x_k$ . De même avec  $\liminf$ . On trouve aussi la notation  $\overline{\lim}$  pour la limite supérieure et  $\underline{\lim}$  pour la limite inférieure.

Comme  $a_n \leq x_n \leq b_n$  pour tout  $n \in \mathbb{N}$ , on a par passage à la limite,  $\liminf_{k \rightarrow +\infty} (x_k) \leq \limsup_k (x_k)$ . De plus, on a  $\liminf_k (x_k) = \limsup_k (x_k)$  si et seulement si la suite  $(x_k)_{k \in \mathbb{N}}$  converge dans  $\mathbb{R}$ , et alors  $\lim_k x_k = \liminf_k (x_k) = \limsup_k (x_k)$ .

Les limites supérieures et inférieures existent pour toutes les suites, convergentes comme non-convergentes. On peut étendre ces notions aux suites de fonctions en considérant des limites ponctuelles. Ainsi, une suite de fonction  $f_k : X \rightarrow \mathbb{R}$  indexée par  $k \in \mathbb{N}$  **converge simplement** vers  $f : X \rightarrow \mathbb{R}$  si

$$\forall x \in X, f(x) = \lim_{k \rightarrow +\infty} f_k(x) = \limsup_{k \rightarrow +\infty} f_k(x) = \liminf_{k \rightarrow +\infty} f_k(x).$$

Par ailleurs on notera  $\inf_k f_k$ ,  $\sup_k f_k$ ,  $\liminf_k f_k$  et  $\limsup_k f_k$  les fonctions définies sur  $X$  par

- $\forall x \in X, (\inf_k f_k)(x) = \inf_k (f_k(x))$ ;
- $\forall x \in X, (\sup_k f_k)(x) = \sup_k (f_k(x))$ ;
- $\forall x \in X, (\liminf_k f_k)(x) = \liminf_k (f_k(x))$ ;
- $\forall x \in X, (\limsup_k f_k)(x) = \limsup_k (f_k(x))$ .

## 1.3 Motivations

Intégrer et sommer sont deux concepts unifiés par la théorie de Lebesgue. Avant de commencer ce cours, on pourra lire avec intérêt l'histoire du signe intégrale : [https://fr.wikipedia.org/wiki/S\\_long](https://fr.wikipedia.org/wiki/S_long).

### 1.3.1 En géométrie

**Mesurer des aires et des volumes** L'idée est de définir rigoureusement un objet permettant de mesurer des ensembles de points :

- des longueurs dans  $\mathbb{R}$  ;
- des aires dans  $\mathbb{R}^2$  ;
- des volumes dans  $\mathbb{R}^3, \mathbb{R}^4, \dots$

Il s'agit alors de définir une mesure qui est une **fonction d'ensemble**, c'est à dire une application  $\mu : \mathcal{P}(\mathbb{R}^n) \rightarrow \mathbb{R}_+$ . On souhaite que  $\mu$  satisfasse les propriétés suivantes :

- (i) **Le vide est de mesure nulle** :  $\mu(\emptyset) = 0$ .
- (ii) **Volume des pavés** : si  $P = [a_1, b_1] \times \dots \times [a_n, b_n]$ , alors  $\mu(P) = \prod_{k=1}^n (b_k - a_k)$ .  
Par exemple, dans  $\mathbb{R}^2$  ces pavés sont des rectangles ayant des côtés parallèles aux axes, et leur mesure est alors le produit des longueurs de leurs côtés.
- (iii) **Invariance par isométrie** : si  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$  est une isométrie et  $A \in \mathcal{P}(\mathbb{R}^n)$ , alors  $\mu(\phi(A)) = \mu(A)$ .  
Dans  $\mathbb{R}^2$ , cela permet d'étendre la propriété précédente à tous les rectangles, y compris ceux dont les côtés ne sont pas parallèles aux axes.
- (iv) **Additivité** : soient  $A, B \in \mathcal{P}(\mathbb{R}^n)$ , si  $A \cap B = \emptyset$ , alors  $\mu(A \cup B) = \mu(A) + \mu(B)$ . De façon générale, si  $A$  et  $B$  ne sont pas disjointes,  $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$ .  
Cette propriété permet de montrer que l'aire d'un triangle est  $\frac{1}{2} \cdot \text{Base} \cdot \text{Hauteur}$ , puis par découpage, de calculer l'aire d'un polygone.
- (v) **Additivité dénombrable** : si  $(A_k)_{k \in \mathbb{N}}$  est une famille dénombrable de parties de  $\mathbb{R}^n$  deux à deux disjointes, alors  $\mu(\bigcup_{k \in \mathbb{N}} A_k) = \sum_{k \in \mathbb{N}} \mu(A_k)$ .  
Par une méthode d'approximation, on peut déterminer l'aire de domaine plus généraux que les polygones, par exemple l'aire d'un disque.

Ces propriétés sont raisonnables et permettent de retrouver ce qu'on sait faire avec la notion intuitive de volume, mais il y a un (très gros) problème : une telle fonction n'existe pas.

**Paradoxe de Banach-Tarski** Le paradoxe de Banach-Tarski est un résultat qui s'énonce de la façon suivante : *Dans  $\mathbb{R}^3$ , la boule unité est équidécomposable à deux boules unité.*

Soient  $B = \{x \in \mathbb{R}^3 \mid \|x\| < 1\}$  la boule unité et  $C_1 = \{x \in \mathbb{R}^3 \mid \|x - c_1\| < 1\}$ ,  $C_2 = \{x \in \mathbb{R}^3 \mid \|x - c_2\| < 1\}$  deux autres boules de rayon 1 disjointes. Le résultat de Banach-Tarski donne l'existence de  $A_1, \dots, A_n \in \mathcal{P}(\mathbb{R}^3)$  deux à deux disjointes,  $\phi_1, \dots, \phi_n$  isométries de  $\mathbb{R}^3$ , et  $p$  compris entre 1 et  $n - 1$  qui sont tels que

- $B = A_1 \cup \dots \cup A_n$  ;
- $\phi_1(A_1), \dots, \phi_n(A_n)$  sont deux à deux disjointes ;
- $C_1 = \phi_1(A_1) \cup \dots \cup \phi_p(A_p)$  et  $C_2 = \phi_{p+1}(A_{p+1}) \cup \dots \cup \phi_n(A_n)$ .

**Remarque 1.3.1.** Le lecteur intéressé pourra consulter les références suivantes :

- [https://en.wikipedia.org/wiki/Banach%E2%80%93Tarski\\_paradox](https://en.wikipedia.org/wiki/Banach%E2%80%93Tarski_paradox)
- <https://www.youtube.com/watch?v=s86-Z-CbaHA>

**Mesurabilité** Supposons maintenant qu'une mesure  $\mu$  existe et satisfasse aux points (i)–(v) ci-dessus. On a alors  $0 < \mu(B) = \mu(C_1) = \mu(C_2)$  car  $C_1$  et  $C_2$  sont des translatés de  $B$  et contiennent un pavé. Par ailleurs, on a

$$\begin{aligned}
 \mu(B) &= \mu(A_1) + \dots + \mu(A_n) \\
 &= \mu(\phi_1(A_1)) + \dots + \mu(\phi_n(A_n)) \\
 &= \left( \mu(\phi_1(A_1)) + \dots + \mu(\phi_p(A_p)) \right) + \left( \mu(\phi_{p+1}(A_{p+1})) + \dots + \mu(\phi_n(A_n)) \right) \\
 &= \mu(C_1) + \mu(C_2) \\
 &= 2\mu(B)
 \end{aligned}$$

Ce qui est impossible, car  $\mu(B) > 0$ .

L'approche adoptée pour échapper à ce paradoxe consiste à renoncer à mesurer toutes les parties de  $\mathbb{R}^n$  (en particulier, les parties  $A_1, \dots, A_n$  données par le théorème de Banach-Tarski n'ont pas de mesure). Ainsi,

- on commence par dire quelles sont les **parties mesurables**, c'est l'objet de la Section 2.1 ;
- on définit ensuite ce qu'est une **mesure** sur l'ensemble des parties mesurables, c'est l'objet de la Section 2.3.

### 1.3.2 En analyse

On va introduire la notion de mesure dans un cadre très général qui aura d'autres implications que simplement donner un sens à la notion de volume dans  $\mathbb{R}^n$ . En particulier, une mesure sert à **intégrer** une fonction sur son domaine de définition. Si l'on choisit bien l'espace fonctionnel, l'intégrale de la valeur absolue est une norme. Formellement, et lorsque cela à un sens, pour une fonction  $f : \mathbb{R} \rightarrow \mathbb{R}$  on pose

$$N_p(f) = \left( \int |f(x)|^p dx \right)^{\frac{1}{p}}, \quad p \geq 1.$$

On cherche alors à définir des espaces vectoriels **normés** qui sont **complets** (dit **espace de Banach**). Les Chapitres 5 et 6 présentent les prémices de ce programme avec les espaces  $L^p$  et la théorie de Fourier. L'approche **duale** sera systématisée en théorie des distributions (niveau Master).

### 1.3.3 En probabilité

*Last but not least...* La théorie moderne des probabilité s'appuie sur l'axiomatique de la théorie de ma mesure. Dans ce cas, un **évènement** est un **ensemble**, une **variable aléatoire** est une **fonction mesurable** et une **probabilité** est une **mesure finie positive**. Ce développement est l'objet du cours HAX506X en seconde partie de ce semestre.

Notons qu'à notre niveau, les problèmes de mesurabilité soulevés par le paradoxe de Banach-Tarski n'ont pas une incidence très grande en géométrie et en analyse : la totalité des fonctions que nous manipulons sont mesurables pour les mesures usuelles. En probabilité cependant, considérer une classe de parties mesurables plutôt qu'une autre relève d'un choix de modélisation ( *e.g.* dans le cadre des processus temporels, cela contient l'information disponible à un temps donné et est modélisée par les filtrations).

## Chapitre 2

# Théorie générale de la mesure

### 2.1 Espaces mesurables

Pour un ensemble  $X$ , on rappelle qu'une classe sur  $X$  est une partie de  $\mathcal{P}(X)$ .

#### 2.1.1 Tribu, tribu engendrée

Pour réaliser le programme esquissé à la fin du chapitre précédent, il faut définir précisément le “domaine de définition” d'une mesure. Cette classe doit être stable par un certains nombre d'opérations.

**Définition 2.1.1.** Une **tribu** (aussi appelée une  **$\sigma$ -algèbre**) sur un ensemble  $X$  est une classe  $\mathcal{A} \subset \mathcal{P}(X)$  vérifiant :

- (i)  $X \in \mathcal{A}$  ;
- (ii) si  $A \in \mathcal{A}$  alors  $A^c \in \mathcal{A}$  ;
- (iii) si  $(A_k)_{k \in \mathbb{N}}$  est une famille dénombrable de  $\mathcal{A}$ , alors  $\bigcup_{k \in \mathbb{N}} A_k \in \mathcal{A}$ .

Un **espace mesurable**  $(X, \mathcal{A})$  est un ensemble muni d'une tribu  $\mathcal{A}$ . Les éléments de  $\mathcal{A}$  sont les **parties mesurables** de  $X$ .

**Remarque 2.1.1.** Soit  $(X, \mathcal{A})$  un espace mesurable.

- Comme  $X \in \mathcal{A}$ , on a  $\emptyset = X^c \in \mathcal{A}$  ;
- Si  $(A_k)_{k \in \mathbb{N}}$  est une famille dénombrable de parties mesurables, alors  $\left(\bigcap_{k \in \mathbb{N}} A_k\right)^c = \bigcup_{k \in \mathbb{N}} A_k^c \in \mathcal{A}$ ,  
donc  $\bigcap_{k \in \mathbb{N}} A_k \in \mathcal{A}$  ;
- Si  $A, B \in \mathcal{A}$  alors  $A \setminus B = A \cap B^c \in \mathcal{A}$  et  $A \Delta B = (A \setminus B) \cup (B \setminus A) \in \mathcal{A}$ .

**Exemple 2.1.1.** Soit  $\Omega$  un ensemble.

1. La tribu grossière ou triviale :  $\mathcal{F} = \{\emptyset, \Omega\}$  ;
2. La tribu pleine :  $\mathcal{F} = \mathcal{P}(\Omega)$ , où  $\mathcal{P}(\Omega)$  désigne l'ensemble des parties de  $\Omega$  ;
3. La tribu engendrée par une partie  $A$  de  $\Omega$  :  $\mathcal{F} = \{A, A^c, \emptyset, \Omega\}$ .

**Proposition 2.1.1.** Si  $(\mathcal{A}_i)_{i \in I}$  est une famille de tribus sur  $X$ , alors  $\mathcal{A} = \bigcap_{i \in I} \mathcal{A}_i$  est une tribu sur  $X$ .

*Démonstration.* (i) On a  $\forall i \in I$   $X \in \mathcal{A}_i$ , donc  $X \in \bigcap_{i \in I} \mathcal{A}_i = \mathcal{A}$ .

(ii) Soit  $A \in \mathcal{A}$ . Pour tout  $i \in I$ , on a  $A \in \mathcal{A}_i$  et  $\mathcal{A}_i$  est une tribu, donc  $A^c \in \mathcal{A}_i$ . On en déduit que  $A^c \in \bigcap_{i \in I} \mathcal{A}_i = \mathcal{A}$ .

(iii) Soit  $(A_k)_{k \in \mathbb{N}}$  une suite d'éléments de  $\mathcal{A}$ . Pour tout  $i \in I$ , on a  $\mathcal{A} \subset \mathcal{A}_i$ , donc  $(A_k)_{k \in \mathbb{N}}$  une suite d'éléments de  $\mathcal{A}_i$  et, comme  $\mathcal{A}_i$  est une tribu, on a  $\bigcup_{k \in \mathbb{N}} A_k \in \mathcal{A}_i$ . On en déduit que

$$\bigcup_{k \in \mathbb{N}} A_k \in \bigcap_{i \in I} \mathcal{A}_i = \mathcal{A}.$$

Les trois points de la définition d'une tribu sont vérifiés, donc  $\mathcal{A}$  est une tribu sur  $X$ .  $\square$

**Remarque 2.1.2.** La réunion de deux tribus n'est, en général, pas une tribu. Soit  $\Omega = \{a, b, c\}$ ,  $\mathcal{T}_1 = \{\emptyset, \Omega, \{a\}, \{b, c\}\}$ ,  $\mathcal{T}_2 = \{\emptyset, \Omega, \{b\}, \{a, c\}\}$ . Alors  $\mathcal{T}_1$  et  $\mathcal{T}_2$  sont des tribus, mais pas  $\mathcal{T}_1 \cup \mathcal{T}_2$  qui n'est pas stable par union finie car contient  $\{a\}$  et  $\{b\}$  mais pas leur union  $\{a, b\}$ .

En général, il est délicat de décrire tous les éléments d'une tribu... De plus, on considère une tribu qui contient une classe d'ensembles intéressants et/ou facile à manipuler (on parle de **générateurs**). Pourquoi dans ce cas, se priverait-on de prendre la plus "grossière" :

**Définition 2.1.2.** Soit  $\mathcal{C}$  une classe sur  $X$ . La **tribu engendrée** par  $\mathcal{C}$  est

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

**Remarque 2.1.3.** Soit  $\mathcal{C}$  une classe sur  $X$ . Compte tenu de la proposition précédente,  $\sigma(\mathcal{C})$  est bien une tribu sur  $X$ . C'est la plus petite tribu sur  $X$  contenant  $\mathcal{C}$ . En effet, on a :

- $\mathcal{C} \subset \sigma(\mathcal{C})$
- si  $\mathcal{A}$  est une tribu et  $\mathcal{C} \subset \mathcal{A}$ , alors  $\sigma(\mathcal{C}) \subset \mathcal{A}$ .

## 2.1.2 Les boréliens

**Topologie et tribu borelienne** On munit  $\mathbb{R}^n$  d'une norme  $\|\cdot\|$  et on note  $B(x, r) = \{y \in \mathbb{R}^n \mid \|y - x\| < r\}$  la boule ouverte de rayon  $r > 0$  et de centre  $x \in \mathbb{R}^n$ . On rappelle de plus que les toutes les normes sur  $\mathbb{R}^n$  sont équivalentes. La topologie de  $\mathbb{R}^n$  est la classe de ses parties ouvertes :

$$\mathcal{O}_{\mathbb{R}^n} = \left\{ U \in \mathcal{P}(\mathbb{R}^n) \mid \forall x \in U, \exists \varepsilon > 0, B(x, \varepsilon) \subset U \right\}.$$

Si  $X \subset \mathbb{R}^n$ , sa topologie est la classe des **ouverts induits** :

$$\mathcal{O}_X = \left\{ X \cap U \mid U \in \mathcal{O}_{\mathbb{R}^n} \right\}.$$

**Définition 2.1.3.** Soit  $X \subset \mathbb{R}^n$ . La **tribu borélienne** sur  $X$  est  $\mathcal{B}(X) = \sigma(\mathcal{O}_X)$ .

De façon plus générale, dès qu'un espace est muni d'une topologie (c'est à dire d'une classe qui définit ses parties ouvertes), sa tribu borélienne est la tribu engendrée par cette topologie. C'est en particulier le cas des espaces métriques (*c.f.* le cours **topologie des espaces métriques** du second semestre). Sur  $\mathbb{R}^n$ , toutes les normes définissent la même topologie  $\mathcal{O}_{\mathbb{R}^n}$ , donc la tribu borélienne ne dépend pas de la norme considérée.

**Exemple 2.1.2.** La droite réelle achevée  $\overline{\mathbb{R}}$  possède également une topologie naturelle  $\mathcal{O}_{\overline{\mathbb{R}}} : U \subset \overline{\mathbb{R}}$  est un ouvert si, pour tout  $x \in U$  on a

- si  $x \in \mathbb{R}$ , alors il existe  $\varepsilon > 0$  tel que  $]x - \varepsilon, x + \varepsilon[ \subset U$  ;
- si  $x = -\infty$ , alors il existe  $a \in \mathbb{R}$  tel que  $[-\infty, a[ \subset U$  ;
- si  $x = +\infty$ , alors il existe  $a \in \mathbb{R}$  tel que  $]a, +\infty[ \subset U$ .

On notera  $\mathcal{B}(\overline{\mathbb{R}})$  la tribu borélienne de  $\overline{\mathbb{R}}$ .

**Les boréliens de  $\mathbb{R}$**  Les tribus  $\mathcal{B}(\mathbb{R}^n)$  sont en fait engendrées par des classes beaucoup plus restreintes que les ouverts, ce qui sert souvent dans la pratique. Pour illustrer ce fait sur  $\mathcal{B}(\mathbb{R})$ , on commence par un lemme sur les ouverts de  $\mathbb{R}$  :

**Lemme 2.1.1.** Pour tout ouvert  $U \subset \mathbb{R}$  non vide, il existe une famille dénombrable  $(]a_k, b_k[)_{k \in \mathbb{N}}$  d'intervalles ouverts bornés tels que  $U = \bigcup_{k \in \mathbb{N}} ]a_k, b_k[$

*Démonstration.* Soit donc  $U$  un ouvert de  $\mathbb{R}$  tel que  $U \neq \emptyset$ . On pose

$$A = \{(\beta, \gamma) \in \mathbb{Q}^2 \mid \beta < \gamma, ]\beta, \gamma[ \subset U\},$$

l'ensemble des couples de rationnels extrémités d'intervalle inclus dans  $U$ . On remarque que  $A$  est non vide ( $\mathbb{Q}$  est dense dans  $U$ ) et au plus dénombrable (car  $\mathbb{Q}$  l'est).

- On a par construction  $\bigcup_{(\beta, \gamma) \in A} ]\beta, \gamma[ \subset U$ .
- Réciproquement, soit  $x \in U$ , il existe  $\varepsilon_x > 0$  tel que  $]x - \varepsilon_x, x + \varepsilon_x[ \subset U$ . L'ensemble  $\mathbb{Q}$  étant dense dans  $\mathbb{R}$ , on peut considérer

$$\beta_x \in ]x - \varepsilon_x, x[ \quad \text{et} \quad \gamma_x \in ]x, x + \varepsilon_x[.$$

On a donc  $x \in ]\beta_x, \gamma_x[ \subset U$  et donc  $(\beta_x, \gamma_x) \in A$ . On en déduit que

$$x \in ]\beta_x, \gamma_x[ \subset \bigcup_{(\beta, \gamma) \in A} ]\beta, \gamma[.$$

On a bien montré que  $U \subset \bigcup_{(\beta, \gamma) \in A} ]\beta, \gamma[$ .

Par double inclusion, on a  $U = \bigcup_{(\beta, \gamma) \in A} ]\beta, \gamma[$  qui est au plus dénombrable et le lemme est démontré.  $\square$

**Proposition 2.1.2.** Soient  $\mathcal{I} = \{]a, b[ \mid a, b \in \mathbb{R} \text{ et } a < b\}$  la classe des intervalles ouverts, et  $\mathcal{K} = \{]a, +\infty[ \mid a \in \mathbb{R}\}$ . On a

$$\sigma(\mathcal{I}) = \sigma(\mathcal{K}) = \mathcal{B}(\mathbb{R}).$$

*Démonstration.* • On montre d'abord l'égalité  $\sigma(\mathcal{I}) = \mathcal{B}(\mathbb{R})$  par double inclusion. Comme  $\mathcal{I} \subset \mathcal{O}_{\mathbb{R}} \subset \mathcal{B}(\mathbb{R})$  et  $\mathcal{B}(\mathbb{R})$  est une tribu, on a  $\sigma(\mathcal{I}) \subset \mathcal{B}(\mathbb{R})$ .

Soit  $U \in \mathcal{O}_{\mathbb{R}}$  un ouvert de  $\mathbb{R}$ . D'après le lemme,  $U$  est une réunion dénombrable d'éléments de  $\mathcal{I}$ , donc de  $\sigma(\mathcal{I})$ . On a donc  $U \in \sigma(\mathcal{I})$  et  $\mathcal{O}_{\mathbb{R}} \subset \sigma(\mathcal{I})$ . Comme  $\sigma(\mathcal{I})$  est une tribu qui contient  $\mathcal{O}_{\mathbb{R}}$ , elle contient aussi la tribu engendrée par  $\mathcal{O}_{\mathbb{R}}$ , d'où  $\mathcal{B}(\mathbb{R}) \subset \sigma(\mathcal{I})$ . On a montré par double inclusion que  $\sigma(\mathcal{I}) = \mathcal{B}(\mathbb{R})$ .

- De même, on montre  $\sigma(\mathcal{K}) = \mathcal{B}(\mathbb{R})$ . Les éléments de  $\mathcal{K}$  sont des ouverts, on a donc  $\mathcal{K} \subset \mathcal{B}(\mathbb{R})$  et  $\sigma(\mathcal{K}) \subset \mathcal{B}(\mathbb{R})$ .

Soit  $a \in \mathbb{R}$ . Les tribu étant stables par intersection dénombrable, on a

$$]a, +\infty[ = \bigcap_{k \in \mathbb{N}^*} ]a - \frac{1}{k}, +\infty[ \in \sigma(\mathcal{K}).$$

Par passage au complémentaire on obtient

$$]-\infty, a[ = ]a, +\infty[^c \in \sigma(\mathcal{K}).$$

Soit  $]a, b[ \in \mathcal{I}$ , on a alors

$$]a, b[ = ]-\infty, b[ \cap ]a, +\infty[ \in \sigma(\mathcal{K}),$$

donc  $\mathcal{I} \subset \sigma(\mathcal{K})$  et  $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{I}) \subset \sigma(\mathcal{K})$ .

On a montré par double inclusion que  $\sigma(\mathcal{K}) = \mathcal{B}(\mathbb{R})$ .  $\square$

Par des arguments similaires, on peut montrer que  $\mathcal{B}(\mathbb{R})$  est engendrée par différentes classes assez simples :

- $\{]a, b[ \mid a, b \in \mathbb{R} \text{ et } a < b\}$
- $\{[a, b] \mid a, b \in \mathbb{R} \text{ et } a \leq b\}$
- $\{]a, b[ \mid a, b \in \mathbb{R} \text{ et } a < b\}$
- $\{[a, b[ \mid a, b \in \mathbb{R} \text{ et } a < b\}$
- $\{]a, +\infty[ \mid a \in \mathbb{R}\}$
- $\{[a, +\infty[ \mid a \in \mathbb{R}\}$
- $\{]-\infty, a[ \mid a \in \mathbb{R}\}$
- $\{]-\infty, a] \mid a \in \mathbb{R}\}$

**Remarque 2.1.4.** La base de ces résultats est le lemme qui permet d'écrire tout ouvert comme réunion d'intervalles dont les bornes sont rationnelles. La proposition reste donc vraie si on suppose, pour chacune des classes considérées, que les bornes sont dans  $\mathbb{Q}$ .

Ainsi, la tribu  $\mathcal{B}(\mathbb{R})$  est engendrée par une classe **dénombrable**. On peut en particulier montrer que  $\mathcal{B}(\mathbb{R})$  est *équipotente* à  $\mathbb{R}$ , d'où on déduit que  $\mathcal{B}(\mathbb{R}) \neq \mathcal{P}(\mathbb{R})$ . Il existe des parties non boréliennes dans  $\mathbb{R}$  (et il en existe même beaucoup).

**Exemple 2.1.3.** La tribu Borélienne de  $\overline{\mathbb{R}}$  vérifie  $\mathcal{B}(\overline{\mathbb{R}}) = \sigma(\{]a, +\infty[ \mid a \in \mathbb{R}\}) = \sigma(\{]-\infty, a[ \mid a \in \mathbb{R}\})$ .

**Les boréliens de  $\mathbb{R}^n$**  Pour les boréliens de  $\mathbb{R}^n$ , on a des résultats similaires. On donnera plus de détails au Chapitre 7.

**Proposition 2.1.3.** La tribu  $\mathcal{B}(\mathbb{R}^n)$  est engendrée par chacune des classes suivantes :

1.  $\{B(a, r) \mid a \in \mathbb{R}^n, r > 0\}$ , la classe des boules ouvertes ;
2.  $\{]a_1, b_1[ \times \cdots \times ]a_n, b_n[ \mid a_1, b_1, \dots, a_n, b_n \in \mathbb{R}\}$ .

*Démonstration.* Proposition admise. □

Sur  $\mathbb{R}^n$ , toutes les normes définissent la même topologie  $\mathcal{O}_{\mathbb{R}^n}$ , donc la tribu borélienne ne dépend pas de la norme. Par contre, les boules ouvertes ne sont pas les mêmes pour toutes les normes, et le premier point de la proposition ci-dessus est valable pour toutes les normes de  $\mathbb{R}^n$ . En particulier, le plan complexe  $\mathbb{C}$  est muni naturellement d'une tribu borélienne engendrée par les ouverts, par les boules euclidiennes ou par les rectangles.

## 2.2 Applications mesurables

### 2.2.1 Définition

**Définition 2.2.1.** Soient  $(X, \mathcal{A})$  et  $(Y, \mathcal{B})$  deux espaces mesurables. Une application  $f : X \rightarrow Y$  est  **$\mathcal{A}$ - $\mathcal{B}$ -mesurable** si

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

Une fonction mesurable sur  $(X, \mathcal{A})$  est une application mesurable à valeurs dans  $\mathbb{R}$  ou  $\mathbb{C}$  muni de sa tribu borélienne.

Si les tribus sont implicites dans le contexte, on parlera d'application mesurables sans préciser les tribus. Lorsque les ensembles  $X$  et  $Y$  sont des parties de  $\mathbb{R}^n$  munies de leurs tribu boréliennes, on parle parfois d'**applications boréliennes** au lieu de mesurables.

**Remarque 2.2.1.** Si on ne précise rien, les tribus utilisées par défaut sont :

- $X$  est un ensemble fini ou dénombrable : ensemble des parties  $\mathcal{P}(X)$  (tribu pleine) ;
- $X$  est un  $\mathbb{R}^n$ ,  $n \in \mathbb{N}^*$  : les boréliens  $\mathcal{B}(X)$ .

**Tribu image et image réciproque** La mesurabilité d'une application  $f : X \rightarrow Y$  dépend des tribus  $\mathcal{A}$  de  $X$  et  $\mathcal{B}$  de  $Y$ . Si on a une application  $f : (X, \mathcal{A}) \rightarrow Y$  (resp.  $f : X \rightarrow (Y, \mathcal{B})$ ), on peut se demander quelle tribu sur  $Y$  (resp.  $X$ ) considérer pour que cette application soit mesurable.

Pour toute classe  $\mathcal{E} \subset \mathcal{P}(X)$  et  $\mathcal{F} \subset \mathcal{P}(Y)$  on pose alors

$$f_*(\mathcal{E}) = \{B \subset Y \mid f^{-1}(B) \in \mathcal{E}\} \subset \mathcal{P}(Y)$$



et

$$f^{-1}(\mathcal{F}) = \{A \subset X \mid A = f^{-1}(B), B \in \mathcal{F}\} \subset \mathcal{P}(X).$$

**Lemme 2.2.1.** Soit  $f : X \rightarrow Y$  une application. Pour toute classe  $\mathcal{E}, \mathcal{E}' \subset \mathcal{P}(X)$  et  $\mathcal{F}, \mathcal{F}' \subset \mathcal{P}(Y)$  on a toujours :

1.  $f^{-1}(f_*(\mathcal{E})) \subset \mathcal{E}$ ;
2.  $\mathcal{F} \subset f_*(f^{-1}(\mathcal{F}))$ ;
3. si  $\mathcal{E} \subset \mathcal{E}'$  alors  $f_*(\mathcal{E}) \subset f_*(\mathcal{E}')$ ;
4. si  $\mathcal{F} \subset \mathcal{F}'$  alors  $f^{-1}(\mathcal{F}) \subset f^{-1}(\mathcal{F}')$ .

*Démonstration.* 1. On a en effet,

$$\begin{aligned} f^{-1}(f_*(\mathcal{E})) &= \{A \subset X \mid A = f^{-1}(B), B \in f_*(\mathcal{E})\} \\ &= \{A \subset X \mid A = f^{-1}(B), f^{-1}(B) \in \mathcal{E}\} \subset \mathcal{E} \end{aligned}$$

2. De même

$$\begin{aligned} f_*(f^{-1}(\mathcal{F})) &= \{B \subset Y \mid f^{-1}(B) \in f^{-1}(\mathcal{F})\} \\ &= \{B \subset Y \mid \exists C \in \mathcal{F}, f^{-1}(B) = f^{-1}(C)\} \supset \mathcal{F} \end{aligned}$$

□

**Remarque 2.2.2.** Les inclusions ne sont pas nécessairement des égalités comme on peut le voir avec les exemples suivants.

1.  $E = \{a_1, a_2\}$ ,  $F = \{b\}$ ,  $f(a_1) = f(a_2) = b$ ,  $\mathcal{E} = \{\{a_1\}, \{a_1, a_2\}\}$ , de sorte que  $f_*(\mathcal{E}) = \{\{b\}\}$  et donc  $f^{-1}(f_*(\mathcal{E})) = \{\{a_1, a_2\}\}$ .
2.  $E = \{a\}$ ,  $F = \{b_1, b_2, b_3\}$ ,  $f(a) = b_1$ ,  $\mathcal{F} = \{\{b_2\}\}$ , de sorte que  $f^{-1}(\mathcal{F}) = \{\emptyset\}$  et donc  $f_*(f^{-1}(\mathcal{F})) = \{\emptyset, \{b_2\}, \{b_3\}, \{b_2, b_3\}\}$ .

**Définition - Proposition 2.2.1.** Soit  $(X, \mathcal{A})$  un espace mesurable et  $f : X \rightarrow Y$  une application. On a alors

1.  $f_*(\mathcal{A}) = \{B \subset Y \mid f^{-1}(B) \in \mathcal{A}\}$  est une tribu sur  $Y$ . C'est la **tribu image** de  $\mathcal{A}$  par  $f$ .
2. L'application  $f$  est  $\mathcal{A}$ - $f_*(\mathcal{A})$ -mesurable, et  $f_*(\mathcal{A})$  est la **plus grande** tribu sur  $Y$  qui rend  $f$  mesurable.

*Démonstration.* 1. On vérifie les conditions de la Définition 2.1.1

- (i) Comme  $f^{-1}(Y) = X \in \mathcal{A}$ , on a bien  $Y \in f_*(\mathcal{A})$ .
- (ii) Soit  $B \in f_*(\mathcal{A})$ . On a  $f^{-1}(B) \in \mathcal{A}$  et  $f^{-1}(B^c) = f^{-1}(B)^c \in \mathcal{A}$ , donc  $B^c \in f_*(\mathcal{A})$ .
- (iii) Soit  $(B_k)_{k \in \mathbb{N}}$  une suite de  $f_*(\mathcal{A})$ . Pour tout  $k \in \mathbb{N}$  on a  $f^{-1}(B_k) \in \mathcal{A}$ , donc  $f^{-1}(\bigcup_{k \in \mathbb{N}} B_k) =$

$$\bigcup_{k \in \mathbb{N}} f^{-1}(B_k) \in \mathcal{A} \text{ et } \bigcup_{k \in \mathbb{N}} B_k \in f_*(\mathcal{A}).$$

2. Soit  $B \in f_*(\mathcal{A})$ . Par définition de cette tribu, on a  $f^{-1}(B) \in \mathcal{A}$ . Autrement dit,  $f$  est  $\mathcal{A}$ - $f_*(\mathcal{A})$ -mesurable.

Soit  $\mathcal{B}$  une tribu sur  $Y$  telle que  $f$  soit  $\mathcal{A}$ - $\mathcal{B}$ -mesurable, et soit  $B \in \mathcal{B}$ . Par mesurabilité de  $f$ , on a  $f^{-1}(B) \in \mathcal{A}$ , donc  $B \in f_*(\mathcal{A})$  et  $\mathcal{B} \subset f_*(\mathcal{A})$ .

□

**Remarque 2.2.3.** Attention  $f_*(\mathcal{A})$  ne désigne pas  $f(\mathcal{A}) = \{B \in Y \mid f(A) = B, A \in \mathcal{A}\}$  qui n'est pas une tribu en général.

De même, on peut tirer en arrière une tribu

**Définition - Proposition 2.2.2.** Soit  $(Y, \mathcal{B})$  un espace mesurable et  $f : X \rightarrow Y$  une application. On a alors

- $f^{-1}(\mathcal{B}) = \{A \subset X \mid A = f^{-1}(B), B \in \mathcal{B}\}$  est une tribu sur  $X$ . On l'appelle **tribu image réciproque**.
- c'est la plus petite tribu sur  $X$  tel que  $f$  est  $f^{-1}(\mathcal{B})$ - $\mathcal{B}$ -mesurable, i.e. si  $\tilde{\mathcal{A}}$  est une tribu sur  $X$  tel que  $f$  est  $\tilde{\mathcal{A}}$ - $\mathcal{B}$ -mesurable, alors  $f^{-1}(\mathcal{B}) \subset \tilde{\mathcal{A}}$ .

*Démonstration.* En exercice □

L'image réciproque de la tribu engendrée par une classe  $\mathcal{C}$  dans l'espace d'arrivée est la tribu dans l'espace de départ engendrée par l'image réciproque de  $\mathcal{C}$  :

**Lemme 2.2.2.** Soit  $f : X \rightarrow Y$  une application de  $X$  dans  $Y$  et  $\mathcal{C}$  un ensemble de parties de  $Y$ . Alors

$$f^{-1}(\sigma(\mathcal{C})) = \sigma(f^{-1}(\mathcal{C})).$$

*Démonstration.* On procède par double inclusion.

- Comme  $\mathcal{C} \subset \sigma(\mathcal{C})$ , on a  $f^{-1}(\mathcal{C}) \subset f^{-1}(\sigma(\mathcal{C}))$ . Comme  $f^{-1}(\sigma(\mathcal{C}))$  est une tribu, elle contient  $\sigma(f^{-1}(\mathcal{C}))$ , autrement dit on a bien  $\sigma(f^{-1}(\mathcal{C})) \subset f^{-1}(\sigma(\mathcal{C}))$ .
- Montrons que  $f^{-1}(\sigma(\mathcal{C})) \subset \sigma(f^{-1}(\mathcal{C}))$ . On considère la tribu image

$$\mathcal{B} = f_*(\sigma(f^{-1}(\mathcal{C}))).$$

D'après le Lemme 2.2.1 points 2 et 3 on a  $\mathcal{C} \subset f_*(f^{-1}(\mathcal{C})) \subset f_*(\sigma(f^{-1}(\mathcal{C}))) = \mathcal{B}$ . Ce qui donne  $\sigma(\mathcal{C}) \subset \mathcal{B}$  car  $\mathcal{B}$  est une tribu. Enfin, d'après le Lemme 2.2.1 points 4 et 1 on a

$$f^{-1}(\sigma(\mathcal{C})) \subset f^{-1}(\mathcal{B}) = f^{-1}(f_*(\sigma(f^{-1}(\mathcal{C})))) \subset \sigma(f^{-1}(\mathcal{C})). \quad \square$$

**Montrer qu'une application est mesurable** Pour cela, et si l'on suit à la lettre la Définition 2.2.1, il faut vérifier que l'image réciproque de la tribu d'arrivée  $\mathcal{B}$  est dans  $\mathcal{A}$ . Mais en fait, il suffit de vérifier que l'image réciproque d'une classe génératrice de la tribu d'arrivée est dans la tribu de départ.

**Proposition 2.2.1.** Soient  $(X, \mathcal{A})$  et  $(Y, \mathcal{B})$  deux espaces mesurables, et soit  $\mathcal{C}$  une classe sur  $Y$  telle que  $\sigma(\mathcal{C}) = \mathcal{B}$ . Alors une application  $f : X \rightarrow Y$  est  $\mathcal{A}$ - $\mathcal{B}$ -mesurable si et seulement si

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

*Démonstration.* On procède par double implication

- Supposons que  $f$  est  $\mathcal{A}$ - $\mathcal{B}$ -mesurable. On a alors  $f^{-1}(B) \in \mathcal{A}$  pour tout  $B \in \mathcal{B}$ , et comme  $\mathcal{C} \subset \mathcal{B}$ , on a bien  $\forall B \in \mathcal{C}, f^{-1}(B) \in \mathcal{A}$ .
- Réciproquement, on suppose que  $\forall B \in \mathcal{C}, f^{-1}(B) \in \mathcal{A}$ . On a alors

$$\mathcal{C} \subset \{B \subset Y \mid f^{-1}(B) \in \mathcal{A}\} = f_*(\mathcal{A}).$$

Comme  $f_*(\mathcal{A})$  est une tribu elle contient aussi  $\sigma(\mathcal{C})$  et on a donc  $\mathcal{B} = \sigma(\mathcal{C}) \subset f_*(\mathcal{A})$ . Autrement dit, pour tout  $B \in \mathcal{B}$  on a alors  $B \in f_*(\mathcal{A})$ , donc  $f^{-1}(B) \in \mathcal{A}$ . Donc  $f$  est  $\mathcal{A}$ - $\mathcal{B}$ -mesurable. □

Une conséquence importante de cette proposition est que la mesurabilité d'une fonction est d'autant plus facile à montrer que la tribu de l'espace d'arrivée est engendrée par une petite partie.

**Remarque 2.2.4.** Pour montrer que  $f : (X, \mathcal{A}) \rightarrow \mathbb{R}$  est mesurable, il suffit de montrer que les ensembles de niveau de  $f$  sont mesurable :

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

Une autre application importante est la proposition suivante avec  $n, p \in \mathbb{N}^*$  :

**Proposition 2.2.2.** Soient  $X \subset \mathbb{R}^p$  et  $\mathbb{R}^n$  munis de leurs tribus boréliennes respectives  $\mathcal{B}(X)$  et  $\mathcal{B}(\mathbb{R}^n)$ . Si  $f : X \rightarrow \mathbb{R}^n$  est continue, alors  $f$  est mesurable.

*Démonstration.* Comme  $\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{O}_{\mathbb{R}^n})$ , il suffit de montrer que  $f^{-1}(U) \in \mathcal{B}(X)$  pour tout  $U \in \mathcal{O}_{\mathbb{R}^n}$ . Or,  $f$  étant continue, pour tout  $U \in \mathcal{O}_{\mathbb{R}^n}$  on a  $f^{-1}(U) \in \mathcal{O}_X \subset \mathcal{B}(X)$ . Donc  $f$  est mesurable.  $\square$

## 2.2.2 Propriétés

**Proposition 2.2.3.** Si  $f : (X, \mathcal{A}) \rightarrow (Y, \mathcal{B})$  et  $g : (Y, \mathcal{B}) \rightarrow (Z, \mathcal{C})$  sont mesurables, alors  $g \circ f : (X, \mathcal{A}) \rightarrow (Z, \mathcal{C})$  est mesurable.

*Démonstration.* Soit  $C \in \mathcal{C}$ . Comme  $g$  est mesurable, on a  $g^{-1}(C) \in \mathcal{B}$ , et comme  $f$  est mesurable, on a  $f^{-1}(g^{-1}(C)) \in \mathcal{A}$ . On obtient donc  $(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C)) \in \mathcal{A}$ , et  $g \circ f$  est mesurable.  $\square$

**Proposition 2.2.4.** Si les fonctions  $f, g : (X, \mathcal{A}) \rightarrow \mathbb{R}$  sont mesurables, alors les fonctions  $f + g$ ,  $f \cdot g$ ,  $\max(f, g)$  et  $\min(f, g)$  sont aussi mesurables. Si de plus  $g$  ne s'annule pas, alors  $\frac{f}{g}$  est mesurable.

*Démonstration.* Les opérations en jeu (addition, multiplication, maximum, minimum, division) sont des *opérateurs binaires*. L'idée est de passer dans  $\mathbb{R}^2$  puis d'utiliser une *opération de réduction* bien choisie.

- On considère d'abord  $F : X \rightarrow \mathbb{R}^2$  définie par  $F(x) = (f(x), g(x))$ . Soient  $I = ]a, b[$  et  $J = ]c, d[$  deux intervalles ouverts de  $\mathbb{R}$ . On a

$$\begin{aligned} x \in F^{-1}(I \times J) &\Leftrightarrow F(x) \in I \times J \\ &\Leftrightarrow f(x) \in I \text{ et } g(x) \in J \\ &\Leftrightarrow x \in f^{-1}(I) \cap g^{-1}(J) \end{aligned}$$

On a donc  $F^{-1}(I \times J) = f^{-1}(I) \cap g^{-1}(J)$ , et comme  $f$  et  $g$  sont mesurables on a  $F^{-1}(I \times J) \in \mathcal{A}$  pour tout pavé ouvert  $I \times J$ . Comme la tribu  $\mathcal{B}(\mathbb{R}^2)$  est engendrée par les pavés ouverts, on en déduit par la Proposition 2.2.1 que  $F$  est mesurable.

- Pour l'addition : Soit  $s : \mathbb{R}^2 \rightarrow \mathbb{R}$  définie par  $s(u, v) = u + v$ . C'est une fonction continue, donc mesurable. Par ailleurs, pour tout  $x \in X$  on a  $f(x) + g(x) = s(f(x), g(x)) = (s \circ F)(x)$ , donc  $f + g = s \circ F$ . Comme  $s$  et  $F$  sont mesurables,  $f + g$  est mesurable.
- Pour les autres opérations, on raisonne de même en considérant, à la place de  $s$ , les fonctions  $(u, v) \mapsto u \cdot v$ ,  $(u, v) \mapsto \max(u, v)$ ,  $(u, v) \mapsto \min(u, v)$  et  $(u, v) \mapsto \frac{u}{v}$  qui sont continues, donc mesurables.  $\square$

En particulier, les fonctions  $f_+ = \max(f, 0)$  et  $f_- = \max(-f, 0)$  sont mesurables, de même que  $|f| = f_+ + f_-$ .

**Proposition 2.2.5.** Soit  $(f_n)_{n \in \mathbb{N}}$  une suite de fonctions mesurables sur  $(X, \mathcal{A})$  à valeurs dans  $\overline{\mathbb{R}}$ . Les fonctions  $\inf_{n \in \mathbb{N}}(f_n)$ ,  $\sup_{n \in \mathbb{N}}(f_n)$ ,  $\liminf(f_n)$  et  $\limsup(f_n)$  sont mesurables sur  $(X, \mathcal{A})$ . En particulier, si la suite  $(f_n)_{n \in \mathbb{N}}$  converge simplement vers  $f$ , alors  $f$  est mesurable.

*Démonstration.* • Soit  $g = \sup_{n \in \mathbb{N}}(f_n)$ . D'après la remarque 2.2.4, il suffit de montrer que  $g^{-1}(]a, +\infty]) \in \mathcal{A}$ , car  $\mathcal{B}(\overline{\mathbb{R}})$  est engendrée par  $\{]a, +\infty] \mid a \in \mathbb{R}\}$ . On a

$$\begin{aligned} x \in g^{-1}(]a, +\infty]) &\Leftrightarrow g(x) > a \\ &\Leftrightarrow \exists n \in \mathbb{N}, f_n(x) > a \\ &\Leftrightarrow x \in \bigcup_{n \in \mathbb{N}} f_n^{-1}(]a, +\infty]) \end{aligned}$$

Comme les fonctions  $f_n$  sont mesurables, on a  $f_n^{-1}(]a, +\infty]) \in \mathcal{A}$  pour tout  $n$  et  $g^{-1}(]a, +\infty]) = \bigcup_{n \in \mathbb{N}} f_n^{-1}(]a, +\infty]) \in \mathcal{A}$ .

- Pour la fonction  $\inf_{n \in \mathbb{N}}(f_n)$  le raisonnement est identique en remarquant que  $\mathcal{B}(\overline{\mathbb{R}})$  est engendrée par  $\{[-\infty, a[ \mid a \in \mathbb{R}\}$ .
- Pour tout  $n \in \mathbb{N}$ , posons  $g_n = \sup_{k \geq n}(f_k)$ . On a alors  $\limsup(f_n) = \inf_{n \in \mathbb{N}}(g_n)$ . Comme les fonctions  $(f_k)_{k \geq n}$  sont mesurables, le point précédent implique que les fonctions  $(g_n)_{n \in \mathbb{N}}$  sont mesurables, ainsi que la fonction  $\limsup(f_n)$ . Pour la fonction  $\liminf(f_n)$  le raisonnement est similaire.
- Si la suite de fonctions  $(f_n)_{n \in \mathbb{N}}$  converge simplement vers  $f$ , on a alors  $f = \limsup(f_n) = \liminf(f_n)$  et  $f$  est mesurable par ce qui précède.

□

**Remarque 2.2.5.** Si la suite  $(f_n)_{n \in \mathbb{N}}$  ne converge pas simplement, alors l'ensemble des points où elle converge est mesurable.

Soient  $C = \{x \in X \mid \lim(f_n(x)) \text{ existe dans } \overline{\mathbb{R}}\}$  et  $F : X \rightarrow \overline{\mathbb{R}}^2$  définie par

$$F(x) = (\liminf(f_n(x)), \limsup(f_n(x))).$$

Cette fonction est mesurable par ce qui précède et on a

$$C = \{x \in X \mid \liminf(f_n(x)) = \limsup(f_n(x))\} = F^{-1}(\Delta)$$

où  $\Delta = \{(u, u) \mid u \in \overline{\mathbb{R}}\}$ . Comme  $\Delta$  est fermé dans  $\overline{\mathbb{R}}^2$ , il est mesurable et  $C$  est donc mesurable.

**Remarque 2.2.6.** Les fonctions continues par morceaux sur un segment  $[a, b] \subset \mathbb{R}$  sont mesurables. C'est une conséquence de la Proposition 2.2.2 et de l'exercice 6 de la feuille de TD II. Si  $f : [a, b] \rightarrow \mathbb{R}$  est continue par morceaux, alors il existe une subdivision  $a = x_0 < x_1 < \dots < x_n < x_{n+1} = b$

du segment  $[a, b]$  telle que  $f|_{]x_k, x_{k+1}[}$  soit continue pour tout  $k \in \llbracket 0, n \rrbracket$ . On a alors  $f = \sum_{k=0}^n f_k +$

$\sum_{k=0}^{n+1} f(x_k) \mathbb{1}_{\{x_k\}}$ , où  $f_k(x) = \begin{cases} f(x) & \text{si } x \in ]x_k, x_{k+1}[ \\ 0 & \text{sinon.} \end{cases}$ . La fonction  $f_k$  est mesurable par continuité de  $f|_{]x_k, x_{k+1}[}$  et par le résultat démontré à la question 1. de l'exercice 6.

Cette remarque sera utile lors de la comparaison des intégrales de Riemann et de Lebesgue.

## 2.3 Espaces mesurés

Un peu de terminologie sur les familles de parties d'un ensemble  $X$  :

- La suite  $(A_k)_{k \in \mathbb{N}}$  est une **suite croissante de parties** de  $X$  si elle est croissante pour l'inclusion (qui est une relation d'ordre sur  $\mathcal{P}(X)$ ), c'est à dire si elle vérifie

$$\forall k \in \mathbb{N} \quad A_k \subset A_{k+1}.$$

De même on parle de suite **décroissante d'ensembles** si  $\forall k \in \mathbb{N} \quad A_{k+1} \subset A_k$ .

- La famille  $(A_i)_{i \in I}$  est une famille de parties **deux à deux disjointes** d'un ensemble  $X$  si elle vérifie

$$\forall i, j \in I \quad i \neq j \Rightarrow A_i \cap A_j = \emptyset.$$

Cela implique en particulier que  $\bigcap_{i \in I} A_i = \emptyset$  (à condition que  $\text{Card}(I) \geq 2$ ), mais c'est une condition strictement plus forte.

### 2.3.1 Mesure positive

**Définition 2.3.1.** Soit  $(X, \mathcal{A})$  un espace mesurable. Une **mesure** sur  $(X, \mathcal{A})$  est une application  $\mu : \mathcal{A} \rightarrow \overline{\mathbb{R}}_+$  vérifiant :

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

(ii) pour toute famille dénombrable  $(A_k)_{k \in \mathbb{N}}$  de parties mesurables deux à deux disjointes on a

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

Un **espace mesuré**  $(X, \mathcal{A}, \mu)$  est un espace mesurable  $(X, \mathcal{A})$  muni d'une mesure  $\mu$ .

**Remarque 2.3.1.** • Les parties de mesure nulle sont appelées les **négligeables**.

- Une partie mesurable peut être de mesure  $+\infty$ .
- La série  $\sum_{k \in \mathbb{N}} \mu(A_k)$  est à termes positifs, donc la somme est bien définie dans  $\overline{\mathbb{R}}_+$ , et elle ne dépend pas de l'ordre des termes.
- Soient  $A, B \in \mathcal{A}$  tels que  $A \cap B = \emptyset$ . Le point (ii) de la définition appliqué à la famille  $(A, B, \emptyset, \emptyset, \dots)$  donne

$$\mu(A \cup B) = \mu(A) + \mu(B) + \mu(\emptyset) + \dots = \mu(A) + \mu(B).$$

Donc le point (ii) est bien valable pour les familles finies.

- Il est important de se restreindre aux familles dénombrables.

**Définition - Proposition 2.3.1.** Soit  $(X, \mathcal{A})$  un espace mesurable et  $a \in X$ . L'application  $\delta_a : \mathcal{A} \rightarrow \overline{\mathbb{R}}_+$  définie par

$$\delta_a(A) = \begin{cases} 1 & \text{si } a \in A \\ 0 & \text{sinon} \end{cases}$$

est une mesure sur  $(X, \mathcal{A})$ . C'est la **mesure de Dirac** en  $a$ .

*Démonstration.* On vérifie les trois hypothèses de la Définition 2.3.1

(i) Comme  $a \notin \emptyset$ , on a  $\delta_a(\emptyset) = 0$ .

(ii) Soit  $(A_k)_{k \in \mathbb{N}}$  une famille de parties mesurables deux à deux disjointes. On a l'alternative suivante :

- Si  $\forall k \in \mathbb{N} \quad a \notin A_k$ . Dans ce cas,  $a \notin \bigcup_{k \in \mathbb{N}} A_k$  et  $\delta_a(A_k) = 0$  pour tout  $k \in \mathbb{N}$ , on a donc

$$\delta_a\left(\bigcup_{k \in \mathbb{N}} A_k\right) = 0 = \sum_{k \in \mathbb{N}} \delta_a(A_k).$$

- Si  $\exists k_0 \in \mathbb{N}, a \in A_{k_0}$ . Dans ce cas,  $a \in \bigcup_{k \in \mathbb{N}} A_k$ ,  $\delta_a(A_{k_0}) = 1$  et  $\delta_a(A_k) = 0$  pour tout  $k \neq k_0$  car les parties sont deux à deux disjointes. On a donc

$$\sum_{k \in \mathbb{N}} \delta_a(A_k) = \delta_a(A_{k_0}) = 1 = \delta_a\left(\bigcup_{k \in \mathbb{N}} A_k\right). \quad \square$$

En particulier, la mesure de Dirac ne prend que deux valeurs, et  $\delta_a(X) = 1$ . Il n'y a pas de partie de mesure infinie.

**Définition 2.3.2.** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré.

- La **masse totale** de  $\mu$  est  $\mu(X)$ . La mesure  $\mu$  est **finie** si sa masse totale est finie :  $\mu(X) < +\infty$ .
- La mesure  $\mu$  est **une mesure de probabilité** si  $\mu(X) = 1$ .
- La mesure  $\mu$  est  **$\sigma$ -finie** s'il existe une suite croissante  $(A_k)_{k \in \mathbb{N}}$  de parties mesurables telle que  $X = \bigcup_{k \in \mathbb{N}} A_k$  et  $\forall k \in \mathbb{N}, \mu(A_k) < +\infty$ .

### 2.3.2 Propriétés

Les combinaisons linéaires de mesures sont des mesures.

**Proposition 2.3.1.** Soit  $(X, \mathcal{A})$  un espace mesurable.

1. Si  $\mu$  est une mesure sur  $(X, \mathcal{A})$  et  $\alpha \in \mathbb{R}_+$ , alors l'application

$$\begin{aligned} \alpha\mu : \mathcal{A} &\rightarrow \overline{\mathbb{R}}_+ \\ A &\mapsto \alpha\mu(A) \end{aligned}$$

est une mesure sur  $(X, \mathcal{A})$ .

2. Si  $(\mu_k)_{k \in \mathbb{N}}$  est une famille de mesures sur  $(X, \mathcal{A})$ , alors l'application

$$\begin{aligned} \mu : \mathcal{A} &\rightarrow \overline{\mathbb{R}}_+ \\ A &\mapsto \sum_{k \in \mathbb{N}} \mu_k(A) \end{aligned}$$

est une mesure sur  $(X, \mathcal{A})$ . On notera  $\mu = \sum_{k \in \mathbb{N}} \mu_k$

*Démonstration.* On ne traite que le point 2. On vérifie les hypothèses de la Définition 2.3.1.

- (i) On a  $\mu(\emptyset) = \sum_{k \in \mathbb{N}} \mu_k(\emptyset)$ , et comme  $\mu_k(\emptyset) = 0$  pour tout  $k \in \mathbb{N}$ , on obtient  $\mu(\emptyset) = 0$ .

- (ii) Soit  $(A_i)_{i \in \mathbb{N}}$  une famille de parties mesurables deux à deux disjointes. On a

$$\mu\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \sum_{k \in \mathbb{N}} \mu_k\left(\bigcup_{i \in \mathbb{N}} A_i\right).$$

Comme  $\mu_k$  est une mesure pour tout  $k \in \mathbb{N}$  on obtient

$$\mu\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \sum_{k \in \mathbb{N}} \sum_{i \in \mathbb{N}} \mu_k(A_i).$$

Les termes  $\mu_k(A_i)$  étant positifs pour tout  $k \in \mathbb{N}$  et tout  $i \in \mathbb{N}$ , on peut permuter les sommations, et la définition de  $\mu$  donne

$$\mu\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \sum_{i \in \mathbb{N}} \sum_{k \in \mathbb{N}} \mu_k(A_i) = \sum_{i \in \mathbb{N}} \mu(A_i).$$

L'application  $\mu$  est bien une mesure sur  $(X, \mathcal{A})$ .  $\square$

**Remarque 2.3.2.** Cette proposition et la définition des mesures de Dirac permet de construire beaucoup d'exemples de mesures : si  $(X, \mathcal{A})$  est un espace mesurable,  $(a_k)_{k \in \mathbb{N}} \in X^{\mathbb{N}}$  et  $(\alpha_k)_{k \in \mathbb{N}} \in (\mathbb{R}_+)^{\mathbb{N}}$ , alors  $\mu = \sum_{k \in \mathbb{N}} \alpha_k \delta_{a_k}$  est une mesure sur  $(X, \mathcal{A})$ . Ces combinaisons linéaires de mesures de Dirac s'appellent des **mesures discrètes**.

**Définition 2.3.3.** La **mesure de comptage** sur  $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$  est définie par  $\chi = \sum_{n \in \mathbb{N}} \delta_n$ .

**Remarque 2.3.3.** • Dans cette définition,  $\delta_n$  désigne la mesure de Dirac en  $n$ . Compte tenu de la proposition précédente,  $\chi$  ainsi définie est bien une mesure.

- On peut aussi mettre des poids devant les masses de Dirac en se donnant une suite  $(p_n)_{n \in \mathbb{N}}$  (que l'on peut penser comme une fonction  $p : \mathbb{N} \rightarrow \mathbb{R}_+$  aussi appelée **densité**). Par exemple, si  $p \in ]0, 1[$  est fixé et si  $p_n = (1-p)^{n-1}p$  alors  $\mu = \sum_{n \in \mathbb{N}} p_n \delta_n$  est la loi géométrique de paramètre  $p$ .

**Proposition 2.3.2.** Pour tout  $A \in \mathcal{P}(\mathbb{N})$  on a  $\chi(A) = \text{Card}(A)$  et la mesure de comptage est  $\sigma$ -finie.

*Démonstration.* Soit  $A \in \mathcal{P}(\mathbb{N})$ . En utilisant que  $\delta_n(A) = 0$  si  $n \notin A$ , puis que  $\delta_n(A) = 1$  si  $n \in A$ , on obtient

$$\chi(A) = \sum_{n \in \mathbb{N}} \delta_n(A) = \sum_{n \in A} \delta_n(A) = \sum_{n \in A} 1 = \text{Card}(A).$$

On a en particulier  $\chi(\mathbb{N}) = +\infty$ , donc la mesure de comptage n'est pas finie. Mais en considérant  $A_k = \llbracket 0, k \rrbracket$  pour tout  $k \in \mathbb{N}$ , on a  $\bigcup_{k \in \mathbb{N}} A_k = \mathbb{N}$  et  $\chi(A_k) < +\infty$  pour tout  $k \in \mathbb{N}$ . La mesure de comptage est donc  $\sigma$ -finie.  $\square$

**Proposition 2.3.3.** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré.

1. Monotonie : soient  $A, B \in \mathcal{A}$ , si  $A \subset B$  alors  $\mu(A) \leq \mu(B)$ .

2. Sous-additivité : pour toute famille  $(A_k)_{k \in \mathbb{N}}$  de parties mesurables on a  $\mu\left(\bigcup_{k \in \mathbb{N}} A_k\right) \leq \sum_{k \in \mathbb{N}} \mu(A_k)$ .

3. Limite croissante : si  $(A_k)_{k \in \mathbb{N}}$  est une suite croissante de parties mesurables, alors

$$\mu\left(\bigcup_{k \in \mathbb{N}} A_k\right) = \lim_{k \rightarrow +\infty} \mu(A_k) = \sup_{k \in \mathbb{N}} (\mu(A_k)).$$

4. Limite décroissante : si  $(A_k)_{k \in \mathbb{N}}$  est une suite décroissante de parties mesurables et si  $\mu(A_0) < +\infty$ , alors

$$\mu\left(\bigcap_{k \in \mathbb{N}} A_k\right) = \lim_{k \rightarrow +\infty} \mu(A_k) = \inf_{k \in \mathbb{N}} (\mu(A_k)).$$

*Démonstration.* 1. Comme  $A \subset B$ , on a  $B \cap A = A$ , donc

$$B = (B \cap A) \sqcup (B \cap A^c) = A \sqcup (B \setminus A)$$

où  $\sqcup$  désigne une union disjointe. Les propriétés de la mesure donnent

$$\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A).$$

2. Soit  $(A_k)_{k \in \mathbb{N}}$  une famille de  $\mathcal{A}$ , et pour tout  $k \in \mathbb{N}$ , soit  $B_k = A_k \setminus \left(\bigcup_{i=0}^{k-1} A_i\right)$ , en particulier,

$B_0 = A_0$ . La famille  $(B_k)_{k \in \mathbb{N}}$  a les propriétés suivantes :

- (i) Pour tout  $k \in \mathbb{N}$  on a  $B_k \in \mathcal{A}$  et  $B_k \subset A_k$ .
- (ii) Les  $B_k$  sont deux à deux disjoints. Soit  $j \neq k$ , et supposons que  $j < k$ . On a alors  $A_j \subset \bigcup_{i=0}^{k-1} A_i$  car  $j \in \llbracket 0, k-1 \rrbracket$ . Comme  $B_k = A_k \cap \left(\bigcup_{i=0}^{k-1} A_i\right)^c \subset \left(\bigcup_{i=0}^{k-1} A_i\right)^c$ , on a  $B_j \cap B_k = \emptyset$ .
- (iii)  $\forall k \in \mathbb{N}$ ,  $\bigcup_{i=0}^k B_i = \bigcup_{i=0}^k A_i$ . Cette propriété est vraie pour  $k = 0$ , et si on la suppose vraie pour  $k \in \mathbb{N}$  donné on obtient

$$\begin{aligned} \bigcup_{i=0}^{k+1} B_i &= B_{k+1} \cup \bigcup_{i=0}^k B_i \\ &= A_{k+1} \setminus \left(\bigcup_{i=0}^k A_i\right) \cup \bigcup_{i=0}^k A_i \quad \text{par hypothèse de récurrence} \\ &= A_{k+1} \cup \bigcup_{i=0}^k A_i = \bigcup_{i=0}^{k+1} A_i \end{aligned}$$

et la propriété est démontrée par récurrence. On en déduit en particulier que  $\bigcup_{k \in \mathbb{N}} A_k = \bigcup_{k \in \mathbb{N}} B_k$ .

De ces trois point on déduit

$$\begin{aligned} \mu\left(\bigcup_{k \in \mathbb{N}} A_k\right) &= \mu\left(\bigcup_{k \in \mathbb{N}} B_k\right) \quad \text{par (iii)} \\ &= \sum_{k \in \mathbb{N}} \mu(B_k) \quad \text{par (ii)} \\ &\leq \sum_{k \in \mathbb{N}} \mu(A_k) \quad \text{par (i)} \end{aligned}$$

3. Soit  $(A_k)_{k \in \mathbb{N}}$  une suite croissante de  $\mathcal{A}$ , et soit  $(B_k)_{k \in \mathbb{N}}$  définie comme au point 2 ci-dessus.

Comme la suite  $(A_k)_{k \in \mathbb{N}}$  est croissante, on a  $A_k = \bigcup_{i=0}^k A_i = \bigcup_{i=0}^k B_i$  et

$$\begin{aligned} \mu\left(\bigcup_{k \in \mathbb{N}} A_k\right) &= \mu\left(\bigcup_{k \in \mathbb{N}} B_k\right) \quad \text{par (iii)} \\ &= \sum_{k \in \mathbb{N}} \mu(B_k) \quad \text{par (ii)} \\ &= \lim_{k \rightarrow +\infty} \sum_{i=0}^k \mu(B_i) \quad \text{par définition de la somme d'une série} \\ &= \lim_{k \rightarrow +\infty} \mu\left(\bigcup_{i=0}^k B_i\right) \quad \text{par (ii)} \\ &= \lim_{k \rightarrow +\infty} \mu(A_k). \end{aligned}$$

Comme  $\mu(A_k) \leq \mu(A_{k+1})$  pour tout  $k \in \mathbb{N}$ , la limite est la borne supérieure.

4. Soit  $(A_k)_{k \in \mathbb{N}}$  une suite décroissante de  $\mathcal{A}$  telle que  $\mu(A_0) < +\infty$ . Pour tout  $k \in \mathbb{N}$ , notons  $C_k = A_0 \setminus A_k$ . Comme  $A_k \subset A_0$ , on a  $\mu(A_k) < +\infty$  et  $\mu(A_0) = \mu(A_k) + \mu(C_k)$  (égalité entre nombre finis).

La suite  $(A_k)_{k \in \mathbb{N}}$  étant décroissante, la suite  $(C_k)_{k \in \mathbb{N}}$  est croissante. On déduit de la monotonie que la suite  $(\mu(A_k))_{k \in \mathbb{N}}$  est décroissante et minorée par 0, alors que la suite  $(\mu(C_k))_{k \in \mathbb{N}}$  est croissante et majorée par  $\mu(A_0)$ . Par passage à la limite, on obtient

$$\lim_{k \rightarrow +\infty} \mu(A_k) = \mu(A_0) - \lim_{k \rightarrow +\infty} \mu(C_k).$$

Par ailleurs, la suite  $(C_k)_{k \in \mathbb{N}}$  étant croissante, la propriété de limite croissante donne

$$\lim_{k \rightarrow +\infty} \mu(C_k) = \mu\left(\bigcup_{k \in \mathbb{N}} C_k\right) \text{ et on a}$$

$$\bigcup_{k \in \mathbb{N}} C_k = \bigcup_{k \in \mathbb{N}} A_0 \cap A_k^c = A_0 \cap \bigcup_{k \in \mathbb{N}} A_k^c = A_0 \cap \left(\bigcap_{k \in \mathbb{N}} A_k\right)^c = A_0 \setminus \left(\bigcap_{k \in \mathbb{N}} A_k\right).$$

On en déduit que  $\mu\left(\bigcup_{k \in \mathbb{N}} C_k\right) = \mu(A_0) - \mu\left(\bigcap_{k \in \mathbb{N}} A_k\right)$  et

$$\lim_{k \rightarrow +\infty} \mu(A_k) = \mu\left(\bigcap_{k \in \mathbb{N}} A_k\right).$$

La suite  $(\mu(A_k))_{k \in \mathbb{N}}$  étant décroissante on a  $\lim_{k \rightarrow +\infty} \mu(A_k) = \inf_{k \in \mathbb{N}} (\mu(A_k))$ . □



**Remarque 2.3.4.** Pour le dernier point, l'hypothèse  $\mu(A_0) < +\infty$  est nécessaire. Par exemple, sur  $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \chi)$  on peut considérer  $A_k = \llbracket k, +\infty \llbracket$  pour tout  $k \in \mathbb{N}$ . On a alors  $\bigcap_{k \in \mathbb{N}} A_k = \emptyset$ , donc  $\chi(\bigcap_{k \in \mathbb{N}} A_k) = 0$ , mais  $\chi(A_k) = +\infty$  pour tout  $k \in \mathbb{N}$ , donc  $\lim_{k \rightarrow +\infty} \chi(A_k) = +\infty$ .

## 2.4 Exemples de mesures

### 2.4.1 Mesure image

**Définition - Proposition 2.4.1.** Soient  $(X, \mathcal{A}, \mu)$  un espace mesuré,  $(Y, \mathcal{B})$  un espace mesurable et  $F : X \rightarrow Y$  une application mesurable. L'application

$$\begin{aligned} F_*\mu : \mathcal{B} &\rightarrow \overline{\mathbb{R}}_+ \\ B &\mapsto \mu(F^{-1}(B)) \end{aligned}$$

est une mesure sur  $(Y, \mathcal{B})$ . Elle s'appelle la **mesure image** de  $\mu$  par  $F$ .

*Démonstration.* On remarque d'abord que l'application  $F_*\mu$  est bien définie : comme  $F$  est mesurable et  $B \in \mathcal{B}$ , on a  $F^{-1}(B) \in \mathcal{A}$  et  $\mu(F^{-1}(B))$  est bien définie. Reste à voir que c'est une mesure en vérifiant les hypothèses de la Définition 2.3.1.

- (i) Comme  $F^{-1}(\emptyset) = \emptyset$ , on a  $(F_*\mu)(\emptyset) = \mu(\emptyset) = 0$ .
- (ii) Soit  $(B_k)_{k \in \mathbb{N}}$  une famille de parties mesurables de  $Y$  deux à deux disjointes. Comme  $F^{-1}(\bigcup_{k \in \mathbb{N}} B_k) = \bigcup_{k \in \mathbb{N}} F^{-1}(B_k)$  on a

$$(F_*\mu)\left(\bigcup_{k \in \mathbb{N}} B_k\right) = \mu\left(\bigcup_{k \in \mathbb{N}} F^{-1}(B_k)\right).$$

Par ailleurs, pour  $i \neq k$  on a  $B_i \cap B_k = \emptyset$  et

$$F^{-1}(B_i) \cap F^{-1}(B_k) = F^{-1}(B_i \cap B_k) = \emptyset,$$

donc  $(F^{-1}(B_k))_{k \in \mathbb{N}}$  est une famille de parties mesurables deux à deux disjointes de  $(X, \mathcal{A})$  et, comme  $\mu$  est une mesure, on obtient

$$(F_*\mu)\left(\bigcup_{k \in \mathbb{N}} B_k\right) = \sum_{k \in \mathbb{N}} \mu(F^{-1}(B_k)) = \sum_{k \in \mathbb{N}} (F_*\mu)(B_k).$$

□

Si  $(X, \mathcal{A}, \mu)$  est un espace mesuré et  $F : X \rightarrow X$  est mesurable, alors  $F_*\mu$  est une mesure sur  $(X, \mathcal{A})$ . Si  $F_*\mu = \mu$ , on dit que l'application  $F$  **préserve la mesure**  $\mu$ , ou que la mesure  $\mu$  est **invariante** par  $F$ .

### 2.4.2 Mesure de Lebesgue

**Théorème 2.4.1.** Il existe une unique mesure sur  $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ , notée  $\lambda_n$ , telle que pour tout pavé  $P = [a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$  on ait

$$\lambda_n(P) = \prod_{i=1}^n (b_i - a_i).$$

*Démonstration.* Ce théorème est admis dans un premier temps, sa preuve fera l'objet du chapitre 7. □

Ce théorème est crucial : quand on parle de la mesure de Lebesgue, on ne parle pas dans le vide (existence), et on parle d'un objet bien identifié (unicité), il n'y a pas d'ambiguïté.

La mesure de Lebesgue n'est pas donnée par une "formule", comme les mesures de Dirac ou la mesure de comptage. Mais comme  $\mathcal{B}(\mathbb{R}^n)$  est engendrée par les pavés et comme  $\lambda_n$  est caractérisée par ses valeurs prises sur les pavés, cela permet de montrer les propriétés de  $\lambda_n$ . On verra en particulier que la mesure de Lebesgue est invariante par les isométries de  $\mathbb{R}^n$ .

Pour calculer la mesure d'un borélien donné, il faut se ramener aux pavés en utilisant les propriétés générales des mesures et les propriétés particulières de la mesure de Lebesgue. En utilisant son invariance par les translations et les rotations de  $\mathbb{R}^2$ , on peut calculer la mesure d'un segment quelconque du plan, la mesure d'un triangle rectangle, puis d'un triangle quelconque. En utilisant une union dénombrable de polygones, on peut calculer la mesure de la boule unité.

Ce procédé de construction de mesures peut être systématisé. On prescrit les valeurs prises par la mesure sur une classe bien choisie d'ensembles et on **complète** (on dit aussi **étend**) la mesure sur la tribu engendrée par cette classe (voir par exemple le théorème de Carathéodory).

## 2.5 Espaces de probabilité

Le formalisme de la théorie des probabilités utilise les outils de la théorie de la mesure en adoptant un vocabulaire spécifique aux probabilités.

**Définition 2.5.1.** Un **espace probabilisable**  $(\Omega, \mathcal{F})$  est un ensemble  $\Omega$  muni d'une tribu  $\mathcal{F}$ .

Un **espace de probabilité**  $(\Omega, \mathcal{F}, \mathbb{P})$  est une espace probabilisable muni d'une mesure de probabilité  $\mathbb{P}$ .

Dans ce formalisme :

- $\Omega$  est un ensemble dont les éléments, appelés **événements élémentaires**, correspondent aux issues possibles d'une expérience aléatoire.
- $\mathcal{F}$  est une tribu sur  $\Omega$  dont les éléments  $A \in \mathcal{F}$  sont appelées des **événements**.
- $\mathbb{P}$  est une mesure positive telle que  $\mathbb{P}(\Omega) = 1$ . Si  $A \in \mathcal{F}$  est un événement, alors  $\mathbb{P}(A)$  mesure la probabilité que l'événement  $A$  soit observé à l'issue de l'expérience aléatoire.

**Exemple 2.5.1.** On jette un dé à 6 faces. L'ensemble des issues possibles de l'expérience est

$$\Omega = \{\square, \dots, \begin{smallmatrix} \blacksquare \\ \blacksquare \\ \blacksquare \end{smallmatrix}\}.$$

L'ensemble étant fini on prend  $\mathcal{F} = \mathcal{P}(\Omega)$  comme tribu des événements.

L'événement "le lancer donne un nombre pair" est  $\{\square, \begin{smallmatrix} \blacksquare \\ \blacksquare \end{smallmatrix}, \begin{smallmatrix} \blacksquare \\ \blacksquare \\ \blacksquare \end{smallmatrix}\}$ .

Si le dé est équilibré la probabilité est

$$\mathbb{P} = \frac{1}{6}(\delta_{\square} + \dots + \delta_{\begin{smallmatrix} \blacksquare \\ \blacksquare \\ \blacksquare \end{smallmatrix}})$$

et les issues de l'expérience sont équiprobables. La probabilité que le lancer donne un nombre pair est  $\mathbb{P}(A) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$ .

Si le dé est pipé, on a par exemple

$$\mathbb{P} = \frac{1}{10}(\delta_{\square} + \dots + \delta_{\begin{smallmatrix} \blacksquare \\ \blacksquare \end{smallmatrix}}) + \frac{1}{2}\delta_{\begin{smallmatrix} \blacksquare \\ \blacksquare \\ \blacksquare \end{smallmatrix}},$$

de sorte que la face 6 sort une fois sur 2... La probabilité que le lancer donne un nombre pair est alors  $\mathbb{P}(A) = \frac{1}{10} + \frac{1}{10} + \frac{1}{2} = \frac{7}{10}$ .

On peut être amené à s'intéresser à des quantités calculées à partir des issues de l'expérience, c'est à dire aux images d'une application  $X : \Omega \rightarrow \mathbb{R}$ . En particulier on peut se demander quelle est la probabilité que  $X(\omega) \in B$  pour une certaine partie  $B \subset \mathbb{R}$ , c'est à dire à la probabilité de l'événement  $\{\omega \in \Omega \mid X(\omega) \in B\} = X^{-1}(B)$ .

**Définition 2.5.2.** Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace de probabilité et  $(E, \mathcal{E})$  un espace mesurable. Une **variable aléatoire**  $X$  à valeur dans  $E$  est une application mesurable  $X : \Omega \rightarrow E$ . La **loi** de la variable aléatoire  $X$  est la mesure  $\mathbb{P}_X = X_*\mathbb{P}$  sur  $(E, \mathcal{E})$ .

Souvent, l'ensemble  $E$  est  $\mathbb{R}$ ,  $\mathbb{C}$  ou  $\mathbb{R}^n$  muni de sa tribu borélienne. On parle alors de variables aléatoires réelles, complexes ou vectorielles.

Si  $X : \Omega \rightarrow \mathbb{R}$  est une variable aléatoire, sa loi  $\mathbb{P}_X$  est la mesure image par  $X$  de la mesure  $\mathbb{P}$ . Si  $B \in \mathcal{B}(\mathbb{R})$  alors  $\mathbb{P}_X(B) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) \in B\})$  est la probabilité que  $X(\omega) \in B$ . On la note de façon abrégée  $\mathbb{P}(X \in B)$ .

**Exemple 2.5.2.** On lance deux dés (par exemple un noir et un bleu), et on s'intéresse à la somme des deux nombres obtenus. On modélise cette expérience par

$$\{\square, \dots, \begin{smallmatrix} \blacksquare \\ \blacksquare \\ \blacksquare \end{smallmatrix}\} \times \{\begin{smallmatrix} \blacksquare \\ \blacksquare \\ \blacksquare \end{smallmatrix}, \dots, \square\} \simeq \{1, 2, 3, 4, 5, 6\}^2 = \Omega.$$

(Remarquer que l'on a utilisé le concept de variable aléatoire en associant  $\square$  à 1,  $\begin{smallmatrix} \blacksquare \\ \blacksquare \\ \blacksquare \end{smallmatrix}$  à 2, etc...). L'ensemble  $\Omega$  est muni de la tribu de ses parties  $\mathcal{F} = \mathcal{P}(\Omega)$  et de la loi uniforme

$$\mathbb{P} = \sum_{(i,j) \in \Omega} \frac{1}{36} \delta_{(i,j)},$$

car les deux dés sont équilibrés. Une issue de l'expérience est un couple  $(p, q)$  où  $p$  désigne le résultat du premier dé, et  $q$  celui du second.

On considère ensuite  $X : \Omega \rightarrow \mathbb{R}$  définie par  $X(i, j) = i + j$ . Les probabilités liées à la somme des résultats d'un lancer sont données par la loi  $\mathbb{P}_X$  de  $X$ . Par exemple, la probabilité de faire au moins 10 est  $\mathbb{P}(X \geq 10) = \mathbb{P}(X^{-1}(\{10, 11, 12\})) = \mathbb{P}_X(\{10, 11, 12\})$ .

Le même espace de probabilité permet également de s'intéresser, par exemple, au plus grand des deux nombre d'un lancer en considérant la variable aléatoire  $Y : \Omega \rightarrow \mathbb{R}$  définie par  $Y(i, j) = \max(i, j)$  et en calculant sa loi  $\mathbb{P}_Y$ .

## 2.6 Le “presque partout”

### 2.6.1 Ensembles négligeables

**Définition 2.6.1.** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré.

1. Une partie mesurable  $A \in \mathcal{A}$  est **négligeable** pour  $\mu$  (ou  $\mu$ -négligeable) si  $\mu(A) = 0$ .
2. Une propriété  $P(x)$  paramétrée par  $x \in X$  est **vraie  $\mu$ -presque partout** (noté  $\mu$ -pp) si il existe une partie  $A \in \mathcal{A}$  négligeable telle que  $\forall x \in X \setminus A, P(x)$ .

**Exemple 2.6.1.** 1. Si  $\chi$  est la mesure de comptage sur  $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$  et  $A \in \mathcal{P}(\mathbb{N})$ , alors

$$\chi(A) = 0 \Leftrightarrow \text{Card}(A) = 0 \Leftrightarrow A = \emptyset.$$

Donc le vide est la seule partie négligeable de  $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \chi)$ .

2. Si  $\lambda_1$  est la mesure de Lebesgue sur  $\mathbb{R}$ , alors pour tout  $a \in \mathbb{R}$  on a  $\{a\} = [a, a]$ , donc  $\lambda_1(\{a\}) = a - a = 0$  : les singletons sont négligeables dans  $\mathbb{R}$ .

Comme  $\mathbb{Q} = \bigcup_{a \in \mathbb{Q}} \{a\}$  est une union dénombrable de singletons deux à deux disjoints, on a

$\lambda_1(\mathbb{Q}) = \sum_{a \in \mathbb{Q}} \lambda_1(\{a\}) = 0$ , donc  $\mathbb{Q}$  est  $\lambda_1$ -négligeable dans  $\mathbb{R}$ . Il en est de même pour toute partie dénombrable de  $\mathbb{R}$ .

Par ailleurs, la fonction caractéristique de  $\mathbb{Q}$  vérifie  $\forall x \in \mathbb{R} \setminus \mathbb{Q}, \mathbf{1}_{\mathbb{Q}}(x) = 0$ . Comme  $\mathbb{Q}$  est négligeable, on a  $\mathbf{1}_{\mathbb{Q}} = 0$ ,  $\lambda_1$ -pp.



FIGURE 2.1 – Construction de l'ensemble de Cantor

3. Dans  $\mathbb{R}^2$ , soit  $D$  la droite d'équation  $y = 0$ . Pour tout  $n \in \mathbb{N}$ , le segment  $S_n = [-n, n] \times \{0\}$  vérifie  $\lambda_2(S_n) = 2n \cdot 0 = 0$ . Comme  $D = \bigcup_{n \in \mathbb{N}} S_n$  est une union croissante de parties mesurables on a  $\lambda_2(D) = \lim_{n \rightarrow +\infty} \lambda_2(S_n) = 0$ . La mesure de Lebesgue étant invariante par translation et rotation, toutes les droites du plan sont  $\lambda_2$ -négligeables.
- Exercice :** dans  $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \lambda_n)$ , montrer que tous les sous-espaces affines de dimension au plus  $n - 1$  sont  $\lambda_n$ -négligeables.
4. On considère  $\mathbb{R}^n$  muni de la norme  $\|\cdot\|_\infty$ . Soit  $U \subset \mathbb{R}^n$  un ouvert non vide et soit  $a = (a_1, \dots, a_n) \in U$ . Il existe  $\varepsilon > 0$  tel que  $B(a, \varepsilon) \subset U$ . Comme  $B(a, \varepsilon) = ]a_1 - \varepsilon, a_1 + \varepsilon[ \times \dots \times ]a_n - \varepsilon, a_n + \varepsilon[$  on a  $[a_1 - \frac{\varepsilon}{2}, a_1 + \frac{\varepsilon}{2}] \times \dots \times [a_n - \frac{\varepsilon}{2}, a_n + \frac{\varepsilon}{2}] \subset U$  et  $\lambda_n(U) \geq \varepsilon^n > 0$ . Dans  $\mathbb{R}^n$ , un borélien d'intérieur non vide n'est jamais négligeable.
5. Soit  $\delta_0$  la mesure de Dirac en 0 sur  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ . Comme  $0 \notin \mathbb{R}^*$ , on a  $\delta_0(\mathbb{R}^*) = 0$  et  $\mathbb{R}^*$  est  $\delta_0$ -négligeable. Par contre  $\delta_0(\{0\}) = 1$  et le singleton  $\{0\}$  n'est pas  $\delta_0$ -négligeable.

## 2.6.2 Ensemble de Cantor

On construit une suite décroissante de parties fermées de  $\mathbb{R}$  qui sont contenues dans  $[0, 1]$  de la façon suivante :

$$\begin{aligned} F_0 &= [0, 1] \\ F_1 &= [0, \frac{1}{3}] \cup [\frac{2}{3}, 1] \\ F_2 &= [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1] \\ &\vdots \end{aligned}$$

La Figure 2.1 illustre le procédé. Pour  $n \in \mathbb{N}$ , le fermé  $F_n$  est une réunion disjointe de  $2^n$  intervalles fermés de longueur  $\frac{1}{3^n}$ . Pour les décrire, on peut les numéroter de 0 à  $2^n - 1$  en respectant leur ordre et en écrivant leur numéro en base 2 :

$$\begin{aligned} F_1 &= [0, \frac{1}{3}] \cup [\frac{2}{3}, 1] = G_1^0 \cup G_1^1 \\ F_2 &= [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1] = G_2^{00} \cup G_2^{01} \cup G_2^{10} \cup G_2^{11} \\ &\vdots \end{aligned}$$

Ainsi, on a

$$F_n = \bigcup_{(a_1, \dots, a_n) \in \{0, 1\}^n} G_n^{a_1 a_2 \dots a_n}.$$

Pour tout  $n \in \mathbb{N}^*$  et tout  $(a_1, a_2, \dots, a_n) \in \{0, 1\}^n$ , l'ensemble  $G_n^{a_1 a_2 \dots a_n}$  est un segment de longueur  $\frac{1}{3^n}$ .

L'ensemble de Cantor est

$$K = \bigcap_{n \in \mathbb{N}} F_n.$$

Cet ensemble a les propriétés suivantes :

1.  $K$  est une intersection décroissante de compacts non vides, donc  $K$  est lui-même un compact non vide (c’est un théorème de topologie).
2. Pour tout  $n$ ,  $F_n$  est une réunions disjointe de  $2^n$  segments de longueur  $\frac{1}{3^n}$ , on a donc  $\lambda_1(F_n) = \frac{2^n}{3^n}$ , et comme  $\lambda_1(F_0) < +\infty$  on obtient

$$\lambda_1(K) = \lim_{n \rightarrow +\infty} \lambda_1(F_n) = 0.$$

Donc  $K$  est  $\lambda_1$ -négligeable.

3. Soit  $a = (a_k)_{k \in \mathbb{N}^*} \in \{0, 1\}^{\mathbb{N}^*}$  une suite de  $\{0, 1\}$ , et soit  $H_a = \bigcap_{n \in \mathbb{N}^*} G_n^{a_1 a_2 \dots a_n}$ . L’ensemble

$H_a$  est une intersection décroissante d’intervalles compacts non vides, donc  $H_a$  est lui-même non vide. Par ailleurs, si  $x, y \in H_a$ , alors, pour tout  $n \in \mathbb{N}^*$  on a  $x, y \in G_n^{a_1 a_2 \dots a_n}$ , donc  $|x - y| \leq \frac{1}{3^n}$ . On en déduit que  $x = y$ , et  $H_a$  est un singleton.

Pour tout  $a = (a_k)_{k \in \mathbb{N}^*} \in \{0, 1\}^{\mathbb{N}^*}$ , notons  $\phi(a)$  l’unique élément de  $H_a$ . On définit ainsi une application  $\phi : \{0, 1\}^{\mathbb{N}^*} \rightarrow K$ . Soient  $a = (a_k)_{k \in \mathbb{N}^*}$  et  $b = (b_k)_{k \in \mathbb{N}^*}$ . Si  $a \neq b$ , alors  $p = \min(\{k \in \mathbb{N}^* \mid a_k \neq b_k\}) \in \mathbb{N}^*$  est bien défini, et on a  $\phi(a) \in G_p^{a_1 \dots a_p}$ ,  $\phi(b) \in G_p^{b_1 \dots b_p}$ , et  $G_p^{a_1 \dots a_p} \cap G_p^{b_1 \dots b_p} = \emptyset$ . Donc  $\phi(a) \neq \phi(b)$  et  $\phi$  est injective.

On en déduit que  $K$  n’est pas dénombrable.

Il existe donc des parties de  $\mathbb{R}$  qui sont négligeables et non-dénombrables. De plus,  $K$  est équipotent à  $\mathbb{R}$ , donc  $\mathcal{P}(K)$  est équipotent à  $\mathcal{P}(\mathbb{R})$ . Comme  $\mathcal{B}(\mathbb{R})$  est équipotent à  $\mathbb{R}$ , il existe donc des parties de  $K$  qui ne sont pas boréliennes.

**Remarque 2.6.1.** Le lecteur intéressé pourra aller voir [https://en.wikipedia.org/wiki/Cantor\\_function](https://en.wikipedia.org/wiki/Cantor_function).



## Chapitre 3

# Théorie générale de l'intégration

### 3.1 Fonctions étagées

#### 3.1.1 Définition et propriétés

**Définition 3.1.1.** Soit  $(X, \mathcal{A})$  un espace mesurable.

- Une fonction mesurable  $f : X \rightarrow \mathbb{R}$  est **étagée** si  $f$  ne prend qu'un nombre fini de valeurs.
- On note  $\mathcal{E}(X, \mathcal{A})$  (ou  $\mathcal{E}(X)$ ) l'espace des fonctions étagées et  $\mathcal{E}_+(X, \mathcal{A})$  (ou  $\mathcal{E}_+(X)$ ) l'espace des fonctions étagées positives.

**Remarque 3.1.1.** • Soit  $f : X \rightarrow \mathbb{R}$  une fonction étagée et soit  $\{\alpha_1, \dots, \alpha_n\}$  les valeurs prises par  $f$ . Pour  $k \in \llbracket 1, n \rrbracket$ , soit  $A_k = f^{-1}(\{\alpha_k\})$ . On a alors

- les réels  $\alpha_1, \dots, \alpha_n$  sont deux à deux disjoints.
- $A_1, \dots, A_n \in \mathcal{A}$  car  $f$  est mesurable et  $(A_1, \dots, A_n)$  est une partition finie de  $X$  (*i.e.* recouvrent  $X$  et sont deux à deux disjoints).
- On a  $f = \sum_{k=1}^n \alpha_k \mathbf{1}_{A_k}$ .
- Réciproquement, si  $(A_1, \dots, A_n)$  est une partition finie de  $X$  par des parties mesurables, et si  $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ , alors  $\sum_{k=1}^n \alpha_k \mathbf{1}_{A_k}$  est une fonction étagée. Mais attention : on supposera toujours que les  $A_i$  sont disjoints deux à deux, mais on ne supposera pas que les valeurs  $\alpha_1, \dots, \alpha_n$  sont nécessairement deux à deux distinctes. En particulier, il existe plusieurs partitions de  $X$  permettant l'écriture d'une fonction étagée sous la forme d'une combinaison linéaire d'indicatrices.

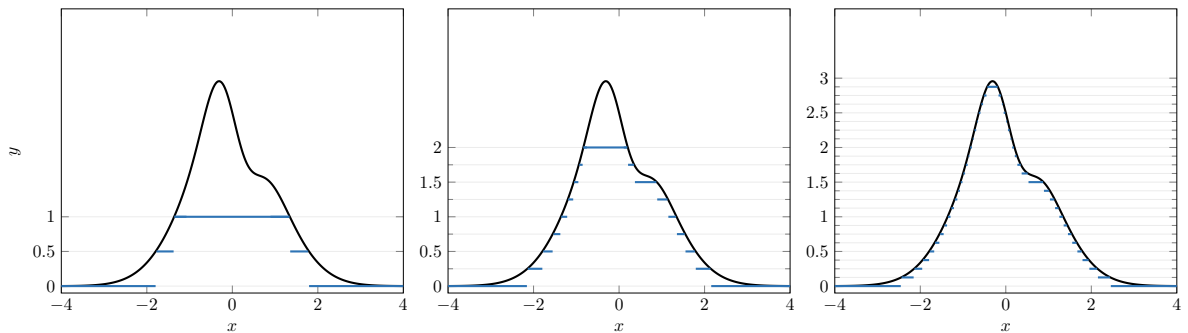


FIGURE 3.1 – Approximation d'une fonction réelle mesurable par les trois premiers termes de la suite croissante de fonctions étagées donnée par (3.1).

**Proposition 3.1.1.** Soit  $(X, \mathcal{A})$  un espace mesurable et  $f : X \rightarrow \overline{\mathbb{R}}_+$  une fonction mesurable positive. Alors il existe une suite croissante  $(f_n)_{n \in \mathbb{N}}$  de  $\mathcal{E}_+(X)$  telle que  $\forall x \in X, \lim_{n \rightarrow +\infty} f_n(x) = f(x)$ .

*Démonstration.* Soit donc  $f : X \rightarrow \overline{\mathbb{R}}_+$  mesurable, on pose pour tout  $n \in \mathbb{N}^*$ ,

$$\begin{aligned} f_n(x) &= \begin{cases} \frac{p}{2^n} & \text{si } f(x) \in [\frac{p}{2^n}, \frac{p+1}{2^n}[ \text{ avec } p \in \llbracket 0, n2^n - 1 \rrbracket \\ n & \text{si } f(x) \geq n. \end{cases} \\ &= n \mathbb{1}_{\{x \in E, f(x) \geq n\}} + \sum_{p=0}^{n2^n-1} \frac{p}{2^n} \mathbb{1}_{\{x \in E, f(x) \in [\frac{p}{2^n}, \frac{p+1}{2^n}[ \}}. \end{aligned} \quad (3.1)$$

On montre dans la suite que la suite  $(f_n)_{n \in \mathbb{N}}$  réalise le résultat annoncé.

- Elle est dans  $\mathcal{E}_+(X)$ . Comme  $f$  est mesurable, les ensembles  $\{x \in E, f(x) \geq n\}$  et  $\{x \in E, f(x) \in [\frac{p}{2^n}, \frac{p+1}{2^n}[ \}$  sont dans  $\mathcal{A}$  pour tout  $n$  et tout  $p$ . Les  $f_n$  étant positifs, il vient  $f_n \in \mathcal{E}_+(X)$  pour tout  $n \in \mathbb{N}$ .
- Elle converge simplement vers  $f$ . Soit  $x \in X$ . De deux choses l'une :
  1. si  $f(x) < +\infty$  alors dès lors que  $n \geq f(x)$ , on a  $|f(x) - f_n(x)| \leq \frac{1}{2^n}$  ;
  2. si  $f(x) = +\infty$  alors  $f_n(x) = n$  pour tout  $n \in \mathbb{N}^*$ .

Dans les deux cas,  $f_n(x) \rightarrow f(x)$  quand  $n \rightarrow +\infty$ .

- Elle est croissante. Soit  $x \in X$  et  $n \in \mathbb{N}^*$ , on a trois régimes :
  1. Si  $f(x) \geq n+1$ . On a alors  $f_{n+1}(x) = n+1 > n = f_n(x)$ .
  2. Si  $n \leq f(x) < n+1$ . Il existe alors  $i \in \llbracket n2^{n+1}, (n+1)2^{n+1} - 1 \rrbracket$  tel que  $f_{n+1}(x) \in [\frac{i}{2^{n+1}}, \frac{i+1}{2^{n+1}}[$ . On a alors  $f_n(x) = n \leq \frac{i}{2^{n+1}} = f_{n+1}(x)$ .
  3. Si  $f(x) < n$ . Il existe alors  $p \in \llbracket 0, n2^n - 1 \rrbracket$  tel que  $f(x) \in [\frac{p}{2^n}, \frac{p+1}{2^n}[ = [\frac{2p}{2^{n+1}}, \frac{2p+2}{2^{n+1}}[$ . Deux cas se présentent alors :
    - (a) si  $f_n(x) \in [\frac{2p}{2^{n+1}}, \frac{2p+1}{2^{n+1}}[$ , on a  $f_n(x) = \frac{p}{2^n} = \frac{2p}{2^{n+1}} = f_{n+1}(x)$  ;
    - (b) si  $f_n(x) \in [\frac{2p+1}{2^{n+1}}, \frac{2p+2}{2^{n+1}}[$ , on a  $f_n(x) = \frac{p}{2^n} < \frac{2p+1}{2^{n+1}} = f_{n+1}(x)$ .

Ainsi, on a toujours  $f_n(x) \leq f_{n+1}(x)$ . □

**Remarque 3.1.2.** En fait la réciproque de la proposition précédente est vrai : une fonction est mesurable positive si et seulement si elle est limite simple d'une suite croissante de fonctions étagées positives. Voir Proposition 3.17 de [3].

### 3.1.2 Intégrale des fonctions étagées

**Définition 3.1.2.** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré et soit  $f : X \rightarrow \mathbb{R}_+$  une fonction étagée positive qui s'écrit  $f = \sum_{k=1}^n \alpha_k \mathbb{1}_{A_k}$ , où  $(A_1, \dots, A_n) \in \mathcal{A}$  est une partition de  $X$ . L'intégrale de  $f$  pour la mesure  $\mu$  est

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

**Lemme 3.1.1.** L'écriture d'une fonction étagée comme combinaison linéaire d'indicatrices n'est pas unique (c.f. remarque 3.1.1), mais la valeur de son intégrale ne dépend pas de la partition choisie.

*Démonstration.* Soit  $f \in \mathcal{E}_+(X)$  et supposons que  $f$  s'écrive  $f = \sum_{j=1}^n \alpha_j \mathbb{1}_{A_j} = \sum_{k=1}^p \beta_k \mathbb{1}_{B_k}$ , où  $(A_1, \dots, A_n)$  et  $(B_1, \dots, B_p)$  sont des partitions mesurables de  $X$ . En particulier, pour tout  $j \in \llbracket 1, n \rrbracket$



on a  $A_j = A_j \cap \bigsqcup_{k=1}^p B_k = \bigsqcup_{k=1}^p (A_j \cap B_k)$ . On obtient donc

$$\begin{aligned} \sum_{j=1}^n \alpha_j \mu(A_j) &= \sum_{j=1}^n \alpha_j \sum_{k=1}^p \mu(A_j \cap B_k) \\ &= \sum_{k=1}^p \sum_{j=1}^n \alpha_j \mu(A_j \cap B_k). \end{aligned}$$

Pour  $j \in \llbracket 1, n \rrbracket$  et  $k \in \llbracket 1, p \rrbracket$  fixés, on a

- soit  $A_j \cap B_k \neq \emptyset$ , et pour  $x \in A_j \cap B_k$  on obtient  $\alpha_j = f(x) = \beta_k$ , donc  $\alpha_j \mu(A_j \cap B_k) = \beta_k \mu(A_j \cap B_k)$ ;
- soit  $A_j \cap B_k = \emptyset$ , et  $\alpha_j \mu(A_j \cap B_k) = 0 = \beta_k \mu(A_j \cap B_k)$ .

D'où

$$\begin{aligned} \sum_{j=1}^n \alpha_j \mu(A_j) &= \sum_{k=1}^p \sum_{j=1}^n \beta_k \mu(A_j \cap B_k) \\ &= \sum_{k=1}^p \beta_k \sum_{j=1}^n \mu(A_j \cap B_k) \\ &= \sum_{k=1}^p \beta_k \mu(B_k \cap \bigsqcup_{j=1}^n A_j) \\ &= \sum_{k=1}^p \beta_k \mu(B_k). \end{aligned} \quad \square$$

**Proposition 3.1.2.** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré. Pour toutes fonctions  $f, g \in \mathcal{E}_+(X)$  et tout  $a \in \mathbb{R}_+$  on a

1.  $f + g \in \mathcal{E}_+(X)$  et  $\int_X (f + g) d\mu = \int_X f d\mu + \int_X g d\mu$ .
2.  $af \in \mathcal{E}_+(X)$  et  $\int_X af d\mu = a \int_X f d\mu$ .
3. si  $f \leq g$  alors  $\int_X f d\mu \leq \int_X g d\mu$ .

*Démonstration.* 1. Il existe  $(A_1, \dots, A_n)$  et  $(B_1, \dots, B_p)$  partitions mesurables de  $X$  et  $\alpha_1, \dots, \alpha_n$ ,  $\beta_1, \dots, \beta_p \in \mathbb{R}_+$  tels que  $f = \sum_{j=1}^n \alpha_j \mathbb{1}_{A_j}$  et  $g = \sum_{k=1}^p \beta_k \mathbb{1}_{B_k}$ .

La famille  $(A_j \cap B_k)_{(j,k) \in \llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket}$  est encore une partition mesurable de  $X$ , et comme  $A_j = \bigsqcup_{k=1}^p (A_j \cap B_k)$  pour tout  $j \in \llbracket 1, n \rrbracket$ , on a  $f = \sum_{j=1}^n \sum_{k=1}^p \alpha_j \mathbb{1}_{A_j \cap B_k}$ . De même, la fonction

$g$  s'écrit  $g = \sum_{j=1}^n \sum_{k=1}^p \beta_k \mathbb{1}_{A_j \cap B_k}$ . On a en particulier  $\int_X f d\mu = \sum_{j=1}^n \sum_{k=1}^p \alpha_j \mu(A_j \cap B_k)$  et

$$\int_X g d\mu = \sum_{j=1}^n \sum_{k=1}^p \beta_k \mu(A_j \cap B_k).$$

Par ailleurs, on obtient  $f + g = \sum_{j=1}^n \sum_{k=1}^p (\alpha_j + \beta_k) \mathbb{1}_{A_j \cap B_k}$ , donc  $f + g \in \mathcal{E}_+(X)$  et

$$\begin{aligned} \int_X (f + g) d\mu &= \sum_{j=1}^n \sum_{k=1}^p (\alpha_j + \beta_k) \mu(A_j \cap B_k) \\ &= \sum_{j=1}^n \sum_{k=1}^p \alpha_j \mu(A_j \cap B_k) + \sum_{j=1}^n \sum_{k=1}^p \beta_k \mu(A_j \cap B_k) \\ &= \int_X f d\mu + \int_X g d\mu. \end{aligned}$$

2. Il existe  $(A_1, \dots, A_n)$  partition mesurable de  $X$  et  $\alpha_1, \dots, \alpha_n \in \mathbb{R}_+$  tels que  $f = \sum_{j=1}^n \alpha_j \mathbb{1}_{A_j}$ ,

et on a  $af = \sum_{j=1}^n a\alpha_j \mathbb{1}_{A_j}$ . Comme  $a \geq 0$ , on a  $af \in \mathcal{E}_+(X)$  et

$$\int_X af d\mu = \sum_{j=1}^n a\alpha_j \mu(A_j) = a \sum_{j=1}^n \alpha_j \mu(A_j) = a \int_X f d\mu.$$

3. Par un raisonnement analogue à celui de la preuve du point 1, on montre que la fonction  $g - f$  est étagée, et comme  $f \leq g$  on a  $g - f \in \mathcal{E}_+(X)$ . En particulier, il résulte de la définition de l'intégrale que  $\int_X (g - f) d\mu \geq 0$ . On a alors, en utilisant le point 1,

$$\int_X g d\mu = \int_X (f + (g - f)) d\mu = \int_X f d\mu + \int_X (g - f) d\mu \geq \int_X f d\mu. \quad \square$$

## 3.2 Intégrale des fonctions positives

### 3.2.1 Définition

On démontre un lemme qui va permettre de définir l'intégrale des fonctions mesurables positives.

**Lemme 3.2.1.** Soient  $(X, \mathcal{A}, \mu)$  un espace mesuré,  $(f_n)_{n \in \mathbb{N}} \subset \mathcal{E}_+(X)$ , et  $g \in \mathcal{E}_+(X)$ , tels que :

- (i)  $f_{n+1}(x) \geq f_n(x)$ , pour tout  $n \in \mathbb{N}$ , et tout  $x \in X$ ,
- (ii)  $\lim_{n \rightarrow +\infty} f_n(x) \geq g(x)$ , pour tout  $x \in X$ ,

alors

$$\lim_{n \rightarrow +\infty} \int_X f_n d\mu \geq \int_X g d\mu.$$

*Démonstration.* Noter que la suite  $(\int_X f_n d\mu)_{n \in \mathbb{N}}$  est une suite croissante de  $\overline{\mathbb{R}}_+$ , donc sa limite existe dans  $\overline{\mathbb{R}}_+$ . Pour  $x \in X$ , on pose  $f(x) = \lim_{n \rightarrow +\infty} f_n(x)$  (cette limite existe et appartient à  $\overline{\mathbb{R}}_+$ ). Il se peut que  $f \notin \mathcal{E}_+(X)$ , mais on a toujours  $f$  mesurable et les hypothèses du lemme donnent  $f_n \rightarrow f$  quand  $n \rightarrow +\infty$  de manière croissante.

Soit  $\alpha \in ]0, 1[$ , on définit, pour  $n \in \mathbb{N}$  :

$$A_n = \{x \in X \mid \alpha g(x) \leq f_n(x)\} = (f_n - \alpha g)^{-1}([0, +\infty[) \in \mathcal{A}.$$

Cette suite d'ensembles est :

- **Monotone.** On a  $A_n \subset A_{n+1}$  car  $f_n \leq f_{n+1}$ .
- **Recouvre  $X$ .** Autrement dit  $X = \bigcup_{n \in \mathbb{N}} A_n$ . En effet, si  $x \in X$ , on distingue deux cas :
  1. Si  $g(x) = 0$ , alors  $x \in A_n$  pour tout  $n \in \mathbb{N}$  donc  $x \in \bigcup_{n \in \mathbb{N}} A_n$ ,
  2. Si  $g(x) > 0$ , on a alors

$$\alpha g(x) < g(x) \leq \lim_{n \rightarrow +\infty} f_n(x).$$

Il existe donc  $n_x \in \mathbb{N}$  (dépendant de  $x$ ) tel que  $x \in A_n$  pour  $n \geq n_x$ . Donc,  $x \in \bigcup_{n \in \mathbb{N}} A_n$ .

Comme  $A_n \subset A_{n+1}$ , pour tout  $n \in \mathbb{N}$ , on peut aussi remarquer que la suite de fonctions  $(\alpha g \mathbb{1}_{A_n})_{n \in \mathbb{N}}$  converge simplement et en croissant vers la fonction  $\alpha g$ .

On remarque maintenant que pour tout  $n \in \mathbb{N}$ , on a  $\alpha g \mathbb{1}_{A_n} \in \mathcal{E}_+(X)$ ,  $f_n \in \mathcal{E}_+(X)$  et que, grâce à la définition de  $A_n$ , on a  $\alpha g \mathbb{1}_{A_n} \leq f_n$ . La monotonie de l'intégrale sur  $\mathcal{E}_+(X)$  donne donc :

$$\int_X \alpha g \mathbb{1}_{A_n} d\mu \leq \int_X f_n d\mu. \quad (3.2)$$

Il existe alors un nombre  $p \in \mathbb{N}$  fini de réels  $0 < b_1 < \dots < b_p$ , et une partition mesurable  $(B_1, \dots, B_p)$  de  $X$  tels que  $g = \sum_{i=1}^p b_i \mathbb{1}_{B_i}$ . On a donc

$$\alpha g \mathbb{1}_{A_n} = \sum_{i=1}^p \alpha b_i \mathbb{1}_{B_i \cap A_n}$$

et donc :

$$\int_X \alpha g \mathbb{1}_{A_n} d\mu = \sum_{i=1}^p \alpha b_i \mu(B_i \cap A_n).$$

Comme  $\bigcup_{n \in \mathbb{N}} (B_i \cap A_n) = B_i \cap (\bigcup_{n \in \mathbb{N}} A_n) = B_i \cap X = B_i$ , la continuité croissante de  $\mu$  donne  $\mu(B_i \cap A_n) \rightarrow \mu(B_i)$ , quand  $n \rightarrow +\infty$ . On en déduit :

$$\lim_{n \rightarrow +\infty} \int_X \alpha g \mathbb{1}_{A_n} d\mu = \lim_{n \rightarrow +\infty} \sum_{i=1}^p \alpha b_i \mu(B_i \cap A_n) = \sum_{i=1}^p \alpha b_i \mu(B_i) = \int_X \alpha g d\mu.$$

On peut donc passer à la limite, quand  $n \rightarrow +\infty$ , dans (3.2) et obtenir :

$$\int_X \alpha g d\mu \leq \lim_{n \rightarrow +\infty} \int_X f_n d\mu.$$

Enfin, la linéarité positive de l'intégrale sur  $\mathcal{E}_+(X)$  donne  $\int_X \alpha g d\mu = \alpha \int_X g d\mu$ . On conclut la démonstration du lemme en faisant tendre  $\alpha$  vers 1.  $\square$

**Lemme 3.2.2.** Soient  $(X, \mathcal{A}, \mu)$  un espace mesuré et  $f : X \rightarrow \overline{\mathbb{R}}_X$  mesurable. Soient deux suites  $(f_n)_{n \in \mathbb{N}}$  et  $(g_n)_{n \in \mathbb{N}}$  à valeurs dans  $\mathcal{E}_+(X)$  convergeant simplement et en croissant vers  $f$ . On a alors :

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

*Démonstration.* On applique le Lemme 3.2.1 avec  $g = g_p$ , à  $p \in \mathbb{N}$  fixé. On obtient  $\int_X g_p d\mu \leq \lim_{n \rightarrow +\infty} \int_X f_n d\mu$ . Puis, en passant à la limite quand  $p \rightarrow +\infty$ , on obtient :

$$\lim_{p \rightarrow +\infty} \int_X g_p d\mu \leq \lim_{n \rightarrow +\infty} \int_X f_n d\mu.$$

On obtient enfin l'égalité demandé en changeant les rôles de  $f_n$  et  $g_p$ .  $\square$

Ce dernier résultat permet donc de définir l'intégrale sur des fonctions mesurables positives de la manière suivante :

**Définition - Proposition 3.2.1 Intégrale des fonctions mesurables positives.** Soient  $(X, \mathcal{A}, \mu)$  un espace mesuré, et  $f : X \rightarrow \overline{\mathbb{R}}$  une application mesurable et positive. Il existe une suite croissante  $(f_n)_{n \in \mathbb{N}} \subset \mathcal{E}_+(X)$  telle que  $f_n \rightarrow f$  quand  $n \rightarrow +\infty$ , c'est-à-dire :

(i) Pour tout  $x \in X$ ,  $f_n(x) \rightarrow f(x)$ , quand  $n \rightarrow +\infty$ ,

(ii)  $f_{n+1}(x) \geq f_n(x)$ , pour tout  $x \in X$ , et tout  $n \in \mathbb{N}$ .

On définit alors l'intégrale de  $f$  en posant :

$$\int_X f d\mu = \lim_{n \rightarrow +\infty} \int_X f_n d\mu \in \overline{\mathbb{R}}_+. \quad (3.3)$$

Si  $A \in \mathcal{A}$ , on note  $\int_A f d\mu = \int_X f \mathbb{1}_A d\mu$ .

*Démonstration.* L'existence de la suite croissante  $(f_n)_{n \in \mathbb{N}}$  est donnée par la Proposition 3.1.1. Le fait que la formule (3.3) soit bien définie dans  $\overline{\mathbb{R}}$  de manière unique (ne dépendant donc pas du choix de la suite  $(f_n)_{n \in \mathbb{N}}$ ) est une conséquence immédiate du Lemme 3.2.2.  $\square$

On a aussi la caractérisation suivante, parfois bien utile car ne dépendant pas d'une suite approximante, de l'intégrale d'une fonction mesurable positive à partir d'intégrales de fonctions étagées positives :

**Proposition 3.2.1.** Soient  $(X, \mathcal{A}, \mu)$  un espace mesuré et  $f : X \rightarrow \overline{\mathbb{R}}$  une application mesurable positive. Alors

$$\int_X f d\mu = \sup \left\{ \int_X v d\mu \mid v \in \mathcal{E}_+(X) \text{ et } v \leq f \right\}.$$

*Démonstration.* Laissé en exercice ou voir le Lemme 4.9 de [3].  $\square$

**Remarque 3.2.1.** Si  $f \in \mathcal{E}_+(X)$ , alors cette caractérisation de la définition de l'intégrale de  $f$  sur  $X$  coïncide avec la Définition 3.1.2.

### 3.2.2 Propriétés

**Proposition 3.2.2.** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré et soit  $f, g : X \rightarrow \overline{\mathbb{R}}_+$  mesurables.

1. Si  $f \leq g$  alors  $\int_X f d\mu \leq \int_X g d\mu$ .
2. Soient  $A, B \in \mathcal{A}$ , si  $A \subset B$  alors  $\int_A f d\mu \leq \int_B f d\mu$ .

*Démonstration.* 1. Notons  $E_f = \{v \in \mathcal{E}_+(X) \mid v \leq f\}$  et  $E_g = \{v \in \mathcal{E}_+(X) \mid v \leq g\}$ . Si  $f \leq g$  on a  $E_f \subset E_g$  et  $\int_X f d\mu = \sup_{v \in E_f} \left( \int_X v d\mu \right) \leq \sup_{v \in E_g} \left( \int_X v d\mu \right) = \int_X g d\mu$ .

2. Si  $A \subset B$ , alors  $\mathbb{1}_A \leq \mathbb{1}_B$  et, comme  $f \geq 0$  on a aussi  $f \mathbb{1}_A \leq f \mathbb{1}_B$ , donc  $\int_A f d\mu \leq \int_B f d\mu$ .  $\square$

### 3.2.3 Théorème de convergence monotone

Pour définir l'intégrale d'une fonction  $f$  mesurable positive, on a utilisé un processus limite : on approxime l'objet d'intérêt par une suite de fonctions simples (étagées) et on définit l'intégrale de  $f$  comme la limite des intégrales le long de la suite. Mais la suite approximant  $f$  n'est pas choisie n'importe comment... Elle se doit d'être croissante.

**Exemple 3.2.1 La bosse fuyante.** Pour tout  $n \in \mathbb{N}$ , soit  $f_n = \mathbb{1}_{[n, n+1]}$ . On a alors  $\lim_{n \rightarrow +\infty} f_n = 0$ , et  $\int_{\mathbb{R}} f_n d\lambda_1 = 1$  pour tout  $n \in \mathbb{N}$ , d'où  $0 = \int_{\mathbb{R}} \lim_{n \rightarrow +\infty} f_n d\lambda_1 < \lim_{n \rightarrow +\infty} \left( \int_{\mathbb{R}} f_n d\lambda_1 \right) = 1$ . En effet la suite  $(f_n)_{n \in \mathbb{N}}$  n'est pas croissante.

De plus, l'additivité de l'intégrale ne se déduit pas immédiatement de la définition, et nécessite le théorème de convergence suivant, qui est par ailleurs utile dans d'autres circonstances.

**Théorème 3.2.1 Convergence monotone (a.k.a Beppo-Levi).** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré et soit  $(f_n)_{n \in \mathbb{N}}$  une suite croissante de fonctions  $X \rightarrow \overline{\mathbb{R}}_+$  mesurables positives sur  $X$ , i.e et  $f_{n+1}(x) \geq f_n(x)$ , pour tout  $n \in \mathbb{N}$  et tout  $x \in X$ . On pose  $f(x) = \lim_{n \rightarrow +\infty} f_n(x) \in \overline{\mathbb{R}}$  la limite

simple de  $(f_n)_{n \in \mathbb{N}}$ . On a alors

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

*Démonstration.* On note d'abord que si  $f_n \in \mathcal{E}_+(X)$  pour tout  $n \in \mathbb{N}$ , alors le fait que  $\int_X f_n d\mu = \int_X f d\mu$  lorsque  $n \rightarrow +\infty$ , est donné par la Définition 3.2.1. L'enjeu de la preuve est de travailler avec des  $f_n$  mesurables et non seulement étagées.

La fonction  $f$  est limite simple (croissante) de la suite de fonctions mesurables  $(f_n)_{n \in \mathbb{N}}$  et est donc bien mesurable (c'est la Proposition 2.2.5). On montre alors l'égalité annoncée en deux temps.

- La première inégalité est immédiate : c'est la monotonie de l'intégrale donnée par la Proposition 3.2.2 qui s'écrit

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

- Pour montrer l'inégalité inverse, on va construire une suite croissante de fonctions étagées  $(g_p)_{p \in \mathbb{N}}$  de  $\mathcal{E}_+(X)$  qui converge simplement vers  $f$  et telle que  $g_p \leq f_p$  pour tout  $p \in \mathbb{N}$ . Pour cela, on note qu'à  $n \in \mathbb{N}$  fixé, il existe une suite croissante de fonctions étagées  $(f_{n,p})_{p \in \mathbb{N}}$  qui converge simplement vers  $f_n$  (Proposition 3.1.1). On définit alors

$$g_p = \sup_{n \leq p} f_{n,p}.$$

On a les propriétés suivantes pour la suite  $(g_p)_{p \in \mathbb{N}}$  :

1. Suite de fonctions étagées. On a pour tout  $p \in \mathbb{N}$ ,  $g_p \in \mathcal{E}_+(X)$  car  $g_p$  est le supremum d'un nombre fini d'éléments de  $\mathcal{E}_+(X)$ .
2. Suite croissante. En effet, comme  $f_{n,p+1} \geq f_{n,p}$  pour tout  $n, p \in \mathbb{N}$ , on a

$$g_{p+1} = \sup \{f_{p+1,p+1}, \sup_{n \leq p} f_{n,p+1}\} \geq \sup_{n \leq p} f_{n,p+1} \geq \sup_{n \leq p} f_{n,p} = g_p.$$

3. Suite simplement convergente vers  $f$ . Pour tout  $x \in X$ , les suites  $(g_p(x))_{p \in \mathbb{N}}$  sont donc croissantes et convergent vers  $g(x) \in \overline{\mathbb{R}}_+$ . On note donc  $g$  la limite et on montre que  $g = f$ . En effet, on remarque que  $g \geq f_n$  pour tout  $n \in \mathbb{N}$  et en passant à la limite on a  $g \geq f$ . D'autre part, on a par construction  $f_{n,p} \leq f_n \leq f$  pour tout  $n, p \in \mathbb{N}$ . On a donc  $g_p \leq f$  pour tout  $p \in \mathbb{N}$  et en passant à la limite on a bien  $g \leq f$ . Ce qui montre que  $g = f$ .
4. Suite majorée. On a  $g_p \leq f_p$  pour tout  $p \in \mathbb{N}$ . En effet,  $f_{n,p} \leq f_n \leq f_p$  si  $n \leq p$ . On a donc  $g_p = \sup_{n \leq p} f_{n,p} \leq f_p$ .

Les points 1, 2 et 3 permettent d'appliquer la Définition 3.2.1 et on a

$$\int_X f d\mu = \lim_{p \rightarrow +\infty} \int_X g_p d\mu.$$

La monotonie de l'intégrale et le point 4 impliquent  $\int_X g_p d\mu \leq \int_X f_p d\mu$  pour tout  $p \in \mathbb{N}$  et on a

$$\int_X f d\mu = \lim_{p \rightarrow +\infty} \int_X g_p d\mu \leq \lim_{p \rightarrow +\infty} \int_X f_p d\mu.$$

Finalement, on obtient bien  $\int_X f d\mu = \lim_{n \rightarrow +\infty} \int_X f_n d\mu$ . □

**Corollaire 3.2.1.** Soient  $(X, \mathcal{A}, \mu)$  un espace mesuré,  $f, g : X \rightarrow \overline{\mathbb{R}}_+$  mesurables, et  $a \in \mathbb{R}_+$ .

On a alors  $\int_X (f + g) d\mu = \int_X f d\mu + \int_X g d\mu$  et  $\int_X a f d\mu = a \int_X f d\mu$ .

*Démonstration.* • Soient  $(u_n)_{n \in \mathbb{N}}$  et  $(v_n)_{n \in \mathbb{N}}$  deux suites croissantes de  $\mathcal{E}_+(X)$  telles que  $\lim_{n \rightarrow +\infty} u_n = f$  et  $\lim_{n \rightarrow +\infty} v_n = g$  (on sait que de telles suites existent par la Proposition 3.1.1). On a alors  $\lim_{n \rightarrow +\infty} (u_n + v_n) = f + g$ , et  $(u_n + v_n)_{n \in \mathbb{N}}$  est une suite croissante de fonctions

positives. Le Théorème de convergence monotone et la linéarité de l'intégrale pour les fonctions étagées donnent

$$\begin{aligned}\int_X (f + g) d\mu &= \lim_{n \rightarrow +\infty} \int_X (u_n + v_n) d\mu \\ &= \lim_{n \rightarrow +\infty} \int_X u_n d\mu + \lim_{n \rightarrow +\infty} \int_X v_n d\mu \\ &= \int_X f d\mu + \int_X g d\mu.\end{aligned}$$

• De même, on a

$$\int_X (af) d\mu = \lim_{n \rightarrow +\infty} \int_X (au_n) d\mu = a \lim_{n \rightarrow +\infty} \int_X u_n d\mu = a \int_X f d\mu. \quad \square$$

**Corollaire 3.2.2.** Soient  $(X, \mathcal{A}, \mu)$  un espace mesuré et  $(f_n)_{n \in \mathbb{N}}$  une suite de fonctions mesurables positives sur  $X$ . Alors la série  $\sum_{n \in \mathbb{N}} f_n$  est mesurable positive et  $\int_X \left( \sum_{n \in \mathbb{N}} f_n \right) d\mu = \sum_{n \in \mathbb{N}} \int_X f_n d\mu$ .

*Démonstration.* Comme, pour tout  $n \in \mathbb{N}$ , les fonctions  $f_n$  sont positives, les sommes partielles  $F_n = \sum_{k=0}^n f_k$  forment une suite croissante de fonctions positives qui converge vers  $\sum_{k \in \mathbb{N}} f_k$ . Par

ailleurs, par linéarité de l'intégrale, on a  $\int_X F_n d\mu = \sum_{k=0}^n \int_X f_k d\mu$  pour tout  $n \in \mathbb{N}$ . Le Théorème de convergence monotone donne alors  $\int_X \left( \sum_{k \in \mathbb{N}} f_k \right) d\mu = \sum_{k \in \mathbb{N}} \int_X f_k d\mu$ .  $\square$

### 3.2.4 Propriétés et lemme de Fatou

Si une suite de fonctions  $(f_n)_{n \in \mathbb{N}}$  n'est pas monotone, elle ne converge pas forcément, mais on peut considérer les limites inférieures. On conserve alors une inégalité entre intégrale de la limite et limite de l'intégrale.

**Théorème 3.2.2 Lemme de Fatou.** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré, soit  $(f_n)_{n \in \mathbb{N}}$  une suite de fonctions  $X \rightarrow \overline{\mathbb{R}}_+$  mesurables positives définies sur  $X$ . On a

$$\int_X \liminf_{n \rightarrow +\infty} (f_n) d\mu \leq \liminf_{n \rightarrow +\infty} \left( \int_X f_n d\mu \right).$$

*Démonstration.* Pour tout  $n \in \mathbb{N}$ , soit  $g_n = \inf_{k \geq n} f_k$ . La suite  $(g_n)_{n \in \mathbb{N}}$  est une suite croissante de fonctions positives, et on a, par définition de la limite inférieure,  $\lim_{n \rightarrow +\infty} g_n = \liminf_{n \rightarrow +\infty} f_n$ . Le Théorème

de convergence monotone donne donc  $\int_X \liminf_{n \rightarrow +\infty} f_n d\mu = \lim_{n \rightarrow +\infty} \int_X g_n d\mu$ .

Pour tout  $k \geq n$ , on a  $g_n \leq f_k$ , donc  $\int_X g_n d\mu \leq \int_X f_k d\mu$  et  $\int_X g_n d\mu \leq \inf_{k \geq n} \left( \int_X f_k d\mu \right)$ . Par passage à la limite on obtient  $\int_X \liminf_{n \rightarrow +\infty} (f_n) d\mu \leq \lim_{n \rightarrow +\infty} \inf_{k \geq n} \left( \int_X f_k d\mu \right) = \liminf_{n \rightarrow +\infty} \left( \int_X f_n d\mu \right)$ .  $\square$

Le “Lemme” de Fatou est primordial : si l'on calcule l'intégrale le long d'une suite de fonctions alors le seul comportement pathologique possible (*i.e.* l'obstruction à l'interversion entre limite et intégrale) est une *perte de masse*. *L'intégrale de la limite est plus petite ou égale que la limite des intégrales.*

**Exemple 3.2.2 La bosse évanescence.** En général, il n'y a pas égalité. Pour tout  $n \in \mathbb{N}^*$ , soit

$f_n = \frac{1}{n} \mathbb{1}_{[0, n]}$ . On a alors  $\liminf_{n \rightarrow +\infty} f_n = \lim_{n \rightarrow +\infty} f_n = 0$ , et  $\int_{\mathbb{R}} f_n d\lambda_1 = 1$  pour tout  $n \in \mathbb{N}$ , d'où  $0 = \int_{\mathbb{R}} \liminf_{n \rightarrow +\infty} f_n d\lambda_1 < \liminf_{n \rightarrow +\infty} \left( \int_{\mathbb{R}} f_n d\lambda_1 \right) = 1$ .

**Remarque 3.2.2 Utilisation du Lemme de Fatou.** Le Lemme de Fatou est souvent utilisé avec des suites  $(f_n)_{n \in \mathbb{N}}$  mesurables positives telles que la suite  $(\int_X f_n d\mu)_{n \in \mathbb{N}}$  des intégrales est bornée et la suite  $(f_n(x))_{n \in \mathbb{N}}$  est convergente pour presque tout  $x \in X$ . Il permet alors de montrer que la limite (au sens de la convergence p.p.) de la suite  $(f_n)_{n \in \mathbb{N}}$  est intégrable.

**Proposition 3.2.3.** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré et soit  $f, g : X \rightarrow \overline{\mathbb{R}}_+$  mesurables et positives.

1. On a  $\int_X f d\mu = 0$  si et seulement si  $f = 0$   $\mu$ -pp.
2. Si  $\int_X f d\mu < +\infty$  alors  $f < +\infty$   $\mu$ -pp.
3. Si  $f = g$   $\mu$ -pp, alors  $\int_X f d\mu = \int_X g d\mu$ .

*Démonstration.* 1. On procède par double implication :

- Si  $f = 0$   $\mu$ -pp et  $v \in \mathcal{E}_+(X)$  vérifie  $0 \leq v \leq f$ , alors la fonction  $v$  est nulle  $\mu$ -pp et  $A_0 = v^{-1}(\{0\})$  vérifie  $\mu(A_0^c) = 0$ . La fonction  $v$  peut s'écrire  $v = 0\mathbb{1}_{A_0} + \sum_{k=1}^n \alpha_k \mathbb{1}_{A_k}$ , où  $\alpha_1, \dots, \alpha_n \in \mathbb{R}_+^*$  sont les valeurs non nulles prises par la fonction  $v$ . On a alors  $A_k \subset A_0^c$  et donc  $\mu(A_k) = 0$ , pour  $k = 1, \dots, n$ . On obtient donc  $\int_X v d\mu = 0\mu(A_0) + \sum_{k=1}^n \alpha_k \mu(A_k) = 0$ . Par définition de l'intégrale de  $f$  on en déduit que  $\int_X f d\mu = 0$ .

- Supposons que  $\int_X f d\mu = 0$ . Pour tout  $n \in \mathbb{N}^*$ , soit  $A_n = \{x \in X \mid f(x) \geq \frac{1}{n}\}$ . On a  $\frac{1}{n}\mathbb{1}_{A_n} \leq f$ , donc  $\frac{1}{n}\mu(A_n) = \int_X \frac{1}{n}\mathbb{1}_{A_n} d\mu \leq \int_X f d\mu = 0$  et  $\mu(A_n) = 0$  pour tout  $n \in \mathbb{N}^*$ . Par ailleurs, on a  $\bigcup_{n \in \mathbb{N}^*} A_n = \{x \in X \mid f(x) > 0\} = f^{-1}(\{0\})^c$ , d'où  $\mu(f^{-1}(\{0\})^c) = \mu(\bigcup_{n \in \mathbb{N}^*} A_n) \leq \sum_{n \in \mathbb{N}^*} \mu(A_n) = 0$ , et  $f$  est nulle  $\mu$ -pp.

2. Par contraposition : supposons que  $A = \{x \in X \mid f(x) = +\infty\}$  vérifie  $\mu(A) > 0$ . On a alors, pour tout  $n \in \mathbb{N}$ ,  $n\mathbb{1}_A \leq f$  et  $n\mathbb{1}_A \in \mathcal{E}_+(X)$ , donc  $\int_X f d\mu \geq \int_X n\mathbb{1}_A d\mu = n\mu(A)$ . Comme  $\mu(A) > 0$ , un passage à la limite pour  $n \rightarrow +\infty$  donne  $\int_X f d\mu = +\infty$ .
3. Soit  $h = \min(f, g)$  et  $k = \max(f, g)$ . On a alors  $h \leq f \leq k$  et  $h \leq g \leq k$ . De plus, pour tout  $x \in X$  vérifiant  $f(x) = g(x)$  on a aussi  $h(x) = k(x)$ , donc  $h = k$   $\mu$ -pp. Comme  $k - h \geq 0$  et  $k - h = 0$   $\mu$ -pp, on peut écrire

$$\int_X k d\mu = \int_X (h + (k - h)) d\mu = \int_X h d\mu + \int_X (k - h) d\mu = \int_X h d\mu.$$

Par monotonie de l'intégrale on a  $\int_X h d\mu \leq \int_X f d\mu \leq \int_X k d\mu$ , donc  $\int_X f d\mu = \int_X k d\mu$ . De même pour la fonction  $g$ , donc  $\int_X f d\mu = \int_X g d\mu$ .  $\square$

## 3.3 Intégrales des fonctions réelles et complexes

### 3.3.1 Définitions

Soit  $X$  un ensemble. On rappelle que si  $f : X \rightarrow \overline{\mathbb{R}}$  alors  $f_+ = \max\{0, f\}$  et  $f_- = \max\{0, -f\}$  de sorte que  $f = f_+ - f_-$ .

**Définition 3.3.1.** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré. Une fonction mesurable  $f : X \rightarrow \mathbb{R}$  est **intégrable** si  $\int_X |f| d\mu < +\infty$ , et dans ce cas, son intégrale est

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

On note  $\mathcal{L}^1(X, \mathcal{A}, \mu)$  ou  $\mathcal{L}^1(X)$  l'espace des fonctions intégrables sur  $X$ .

**Remarque 3.3.1.** • Si  $f : X \rightarrow \mathbb{R}$  est mesurable, alors les fonctions  $f_+ = \max(f, 0)$ ,  $f_- = \max(-f, 0)$  et  $|f|$  sont mesurables positives, donc leurs intégrales sont bien définies.

- Comme  $|f| = f_+ + f_-$ , on a  $f_- \leq |f|$  donc  $\int_X f_- d\mu < +\infty$  et on peut soustraire cette intégrale. L'hypothèse d'intégrabilité est donc nécessaire pour écrire l'intégrale de  $f$ . De même, on a  $f_+ \leq |f|$  donc  $\int_X f_+ d\mu < +\infty$ , donc  $\int_X f d\mu \in \mathbb{R}$ .
- Si  $f \geq 0$  alors  $f_- = 0$  et  $f_+ = f$ , donc cette définition est bien cohérente avec la définition précédente.

**Proposition 3.3.1.** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré.

1. L'espace  $\mathcal{L}^1(X)$  est un  $\mathbb{R}$ -espace vectoriel et l'application  $f \mapsto \int_X f d\mu$  est une forme linéaire sur  $\mathcal{L}^1(X)$ .
2. Pour tout  $f \in \mathcal{L}^1(X)$  on a  $\left| \int_X f d\mu \right| \leq \int_X |f| d\mu$ .
3. Soient  $f, g \in \mathcal{L}^1(X)$ . Si  $f \leq g$   $\mu$ -pp, alors  $\int_X f d\mu \leq \int_X g d\mu$ , et si  $f = g$   $\mu$ -pp, alors  $\int_X f d\mu = \int_X g d\mu$ .

*Démonstration.* 1. L'espace  $\mathcal{L}^1(X)$  est bien un espace vectoriel. En effet, soient  $f, g \in \mathcal{L}^1(X)$  et  $\alpha \in \mathbb{R}$ . Comme  $|\alpha f + g| \leq |\alpha| |f| + |g|$ , on a  $\int_X |\alpha f + g| d\mu \leq |\alpha| \int_X |f| d\mu + \int_X |g| d\mu < +\infty$ , donc  $\alpha f + g \in \mathcal{L}^1(X)$ .

Pour montrer que l'intégrale est une forme linéaire, on procède en deux temps :

- Les décompositions en parties positives et négatives des fonctions donnent

$$(f + g)_+ - (f + g)_- = f + g = f_+ - f_- + g_+ - g_-.$$

Toutes ces fonctions étant d'intégrale finies sur  $X$ , elles sont finies  $\mu$ -pp et on a

$$(f + g)_+ + f_- + g_- = (f + g)_- + f_+ + g_+ \quad \mu\text{-pp}.$$

Par linéarité de l'intégrale des fonctions positives on obtient

$$\int_X (f + g)_+ d\mu + \int_X f_- d\mu + \int_X g_- d\mu = \int_X (f + g)_- d\mu + \int_X f_+ d\mu + \int_X g_+ d\mu,$$

et comme tous les termes de cette égalité sont finis on a

$$\int_X (f + g)_+ d\mu - \int_X (f + g)_- d\mu = \int_X f_+ d\mu - \int_X f_- d\mu + \int_X g_+ d\mu - \int_X g_- d\mu,$$

soit

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



- Si  $\alpha \geq 0$ , alors  $(\alpha f)_+ = \alpha f_+$  et  $(\alpha f)_- = \alpha f_-$ , d'où

$$\int_X (\alpha f) d\mu = \int_X \alpha f_+ d\mu - \int_X \alpha f_- d\mu = \alpha \int_X f_+ d\mu - \alpha \int_X f_- d\mu = \alpha \int_X f d\mu.$$

Si  $\alpha < 0$ , alors  $(\alpha f)_+ = -\alpha f_-$  et  $(\alpha f)_- = -\alpha f_+$ , et comme  $-\alpha > 0$ , la linéarité de l'intégrale des fonctions positives donne

$$\int_X (\alpha f) d\mu = \int_X (-\alpha) f_- d\mu - \int_X (-\alpha) f_+ d\mu = (-\alpha) \int_X f_- d\mu - (-\alpha) \int_X f_+ d\mu = \alpha \int_X f d\mu.$$

2. Par inégalité triangulaire, et comme  $f_+ + f_- = |f|$ , on obtient

$$\begin{aligned} \left| \int_X f d\mu \right| &= \left| \int_X f_+ d\mu - \int_X f_- d\mu \right| \\ &\leq \left| \int_X f_+ d\mu \right| + \left| \int_X f_- d\mu \right| \\ &\leq \int_X f_+ d\mu + \int_X f_- d\mu \\ &\leq \int_X (f_+ + f_-) d\mu \\ &\leq \int_X |f| d\mu. \end{aligned}$$

3. Pour une fonction  $h \geq 0$   $\mu$ -pp, on a  $h_- = 0$   $\mu$ -pp et  $h_+ = h$   $\mu$ -pp, donc  $\int_X h d\mu = \int_X h_+ d\mu \geq 0$ .

Si  $f \leq g$   $\mu$ -pp on a donc  $g - f \geq 0$   $\mu$ -pp et  $0 \leq \int_X (g - f) d\mu = \int_X g d\mu - \int_X f d\mu$ .

Si  $f = g$   $\mu$ -pp, alors  $f_+ = g_+$   $\mu$ -pp et  $f_- = g_-$   $\mu$ -pp. Le cas des fonctions mesurables positives étant connu, on a  $\int_X f_+ d\mu = \int_X g_+ d\mu$  et  $\int_X f_- d\mu = \int_X g_- d\mu$ , d'où  $\int_X f d\mu = \int_X g d\mu$ .  $\square$

**Définition 3.3.2.** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré. Une fonction mesurable à valeurs complexes  $f : X \rightarrow \mathbb{C}$  est **intégrable** si  $\int_X |f| d\mu < +\infty$ , et dans ce cas, son intégrale est

$$\int_X f d\mu \doteq \int_X \operatorname{Re}(f) d\mu + i \int_X \operatorname{Im}(f) d\mu.$$

On note  $\mathcal{L}_{\mathbb{C}}^1(X, \mathcal{A}, \mu)$  ou  $\mathcal{L}_{\mathbb{C}}^1(X)$  l'espace des fonctions intégrables sur  $X$ .

**Remarque 3.3.2.** Comme  $|f|^2 = \operatorname{Re}(f)^2 + \operatorname{Im}(f)^2$ , on a  $|\operatorname{Re}(f)| \leq |f|$  et  $|\operatorname{Im}(f)| \leq |f|$ . L'hypothèse d'intégrabilité de  $f$  implique donc que  $\operatorname{Re}(f), \operatorname{Im}(f) \in \mathcal{L}^1(X)$  et l'intégrale de  $f$  est bien définie.

**Proposition 3.3.2.** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré.

1. L'espace  $\mathcal{L}_{\mathbb{C}}^1(X)$  est un  $\mathbb{C}$ -espace vectoriel et l'application  $f \mapsto \int_X f d\mu$  est une forme linéaire sur  $\mathcal{L}_{\mathbb{C}}^1(X)$ .
2. Pour tout  $f \in \mathcal{L}_{\mathbb{C}}^1(X)$  on a  $\left| \int_X f d\mu \right| \leq \int_X |f| d\mu$ .
3. Soient  $f, g \in \mathcal{L}_{\mathbb{C}}^1(X)$ . Si  $f = g$   $\mu$ -pp, alors  $\int_X f d\mu = \int_X g d\mu$ .

*Démonstration.* 1. Laissé en exercice.

2. Soit  $\theta \in \mathbb{R}$  tel que  $\int_X f d\mu = \left| \int_X f d\mu \right| e^{i\theta}$ . On a alors

$$\left| \int_X f d\mu \right| = e^{-i\theta} \int_X f d\mu = \int_X e^{-i\theta} f d\mu = \int_X \operatorname{Re}(e^{-i\theta} f) d\mu + i \int_X \operatorname{Im}(e^{-i\theta} f) d\mu.$$

Comme  $\left| \int_X f d\mu \right|$  est un réel positif, sa partie imaginaire est nulle et sa partie réelle est positive, on a donc

$$\left| \int_X f d\mu \right| = \left| \int_X \operatorname{Re}(e^{-i\theta} f) d\mu \right| \leq \int_X \left| \operatorname{Re}(e^{-i\theta} f) \right| d\mu \leq \int_X |e^{-i\theta} f| d\mu \leq \int_X |f| d\mu.$$

3. Laissé en exercice. □

**Remarque 3.3.3.** On peut, de façon analogue, définir l'intégrale d'une fonction mesurable à valeurs vectorielles : la fonction  $f : X \rightarrow \mathbb{R}^n$  est intégrable si  $\int_X \|f\| d\mu < +\infty$ , où  $\|\cdot\|$  est une norme sur  $\mathbb{R}^n$  (par exemple la norme euclidienne).

Pour une telle fonction, on a  $f(x) = (f_1(x), \dots, f_n(x))$  pour tout  $x \in X$ , et les fonctions  $f_k$  sont mesurables et vérifient  $|f_k| \leq \|f\|$ . Si  $f$  est intégrable, on définit alors

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

### 3.3.2 Théorème de convergence dominée

Le théorème suivant est le résultat principal permettant de permuter limite et intégrale. Comme on l'a vu avec le Lemme de Fatou, il est nécessaire d'ajouter une hypothèse concernant la suite de fonctions.

**Théorème 3.3.1 Théorème de convergence dominée de Lebesgue (CDL).** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré et  $(f_n)_{n \in \mathbb{N}}$  une suite de fonctions intégrables sur  $X$ . On suppose que

- (i) **Convergence simple.** Il existe  $f : X \rightarrow \mathbb{C}$  telle  $\forall x \in X, \lim_{n \rightarrow +\infty} f_n(x) = f(x)$ .
- (ii) **Domination.** Il existe  $g : X \rightarrow \mathbb{R}_+$  intégrable telle que  $\forall n \in \mathbb{N}, \forall x \in X, |f_n(x)| \leq g(x)$ .

Alors,  $f$  est intégrable sur  $X$  et on a  $\int_X f d\mu = \lim_{n \rightarrow +\infty} \int_X f_n d\mu$  et  $\lim_{n \rightarrow +\infty} \int_X |f_n - f| d\mu = 0$ .

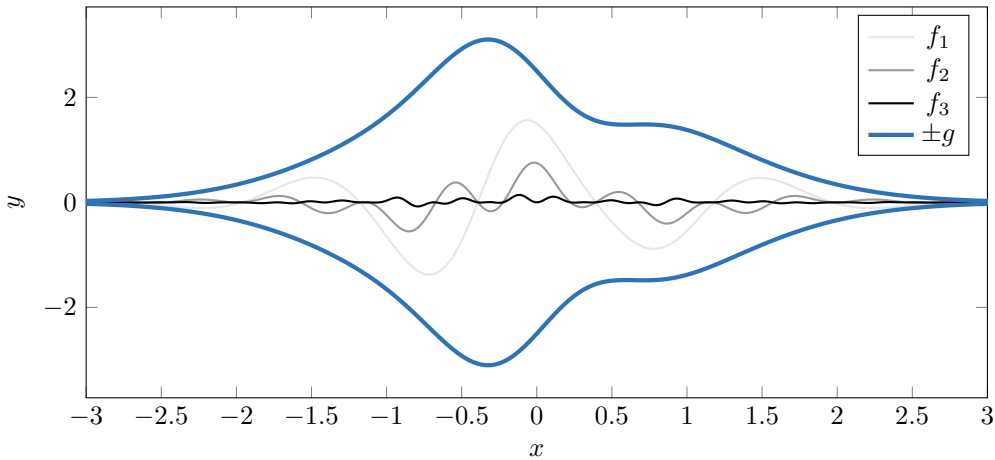


FIGURE 3.2 – L'hypothèse de majoration (uniforme en  $n$ ) dans le Théorème de Convergence dominée permet d'éviter la "perte de masse" pour la fonction limite : les graphes des  $f_n$  sont "coincés" entre celui de  $g$  et de  $-g$ .

*Démonstration.* La démonstration repose sur la constations suivante :

$$\begin{cases} -g \leq f_n \leq g \\ -g \leq f \leq g \end{cases} \Rightarrow \begin{cases} -g \leq -f_n \leq g \\ -g \leq f \leq g \end{cases} \Rightarrow |f - f_n| \leq 2g.$$

Posons donc  $g_n = 2g - |f - f_n|$  pour tout  $n \in \mathbb{N}$ . La suite  $(g_n)_{n \in \mathbb{N}}$  est une suite de fonctions mesurables **positives** qui converge simplement vers  $2g$ . On peut appliquer le Lemme de Fatou à  $(g_n)_{n \in \mathbb{N}}$  :

$$0 \leq \int_X 2gd\mu = \int_X \liminf_{n \rightarrow +\infty} g_n d\mu \leq \liminf_{n \rightarrow +\infty} \int_X g_n d\mu = \lim_{n \rightarrow +\infty} \inf_{p \geq n} \left( \int_X 2g - |f - f_p| d\mu \right). \quad (3.4)$$

La linéarité de l'intégrale donne  $\int_X 2g - |f - f_p| d\mu = \int_X 2g - \int_X |f - f_p| d\mu$  et on a

$$\inf_{p \geq n} \left( \int_X 2g - |f - f_p| d\mu \right) = \int_X 2gd\mu - \sup_{p \geq n} \left( \int_X |f - f_p| d\mu \right)$$

qui donne

$$0 \leq \liminf_{n \rightarrow +\infty} \int_X g_n d\mu = \int_X 2gd\mu - \limsup_{n \rightarrow +\infty} \left( \int_X |f - f_n| d\mu \right).$$

Puis que  $\int_X 2gd\mu < +\infty$  on peut "simplifier" par cette quantité dans (3.4) on a

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

Autrement dit  $\limsup_{n \rightarrow +\infty} \left( \int_X |f - f_n| d\mu \right) = 0$ . Enfin, en appliquant le point 2 de la Proposition 3.3.1 on a

$$\left| \int_X f d\mu - \int_X f_n d\mu \right| = \left| \int_X f - f_n d\mu \right| \leq \int_X |f - f_n| d\mu, \text{ pour tout } n \in \mathbb{N},$$

qui permet de conclure en passant à la limite.  $\square$

On déduit de ce résultat une version souvent utile où les hypothèses de convergence simple et de domination sont satisfaites seulement  $\mu$ -pp.

**Corollaire 3.3.1.** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré et  $(f_n)_{n \in \mathbb{N}}$  une suite de fonctions intégrables sur  $X$ . On suppose que

- (i) **Convergence p.p.** Il existe  $f : X \rightarrow \mathbb{C}$  telle  $\lim_{n \rightarrow +\infty} f_n(x) = f(x)$  pour  $\mu$ -pt  $x \in X$ .
- (ii) **Domination.** Il existe  $g : X \rightarrow \mathbb{R}_+$  intégrable telle que  $|f_n(x)| \leq g(x)$  pour  $\mu$ -pt  $x \in X$ .

Alors,  $f$  est intégrable sur  $X$  et on a  $\int_X f d\mu = \lim_{n \rightarrow +\infty} \int_X f_n d\mu$  et  $\lim_{n \rightarrow +\infty} \int_X |f_n - f| d\mu = 0$ .

*Démonstration.* On applique simplement le Théorème de convergence dominée à des fonctions bien choisies.

- Comme  $f_n \rightarrow f$   $\mu$ -pp quand  $n \rightarrow +\infty$ , il existe  $A \in \mathcal{A}$  tel que  $\mu(A) = 0$  et  $f_n(x) \rightarrow f(x)$  pour tout  $x \in A^c$  quand  $n \rightarrow +\infty$ .
- Comme, pour tout  $n \in \mathbb{N}$ , on a  $|f_n| \leq g$   $\mu$ -pp alors il existe  $B_n \in \mathcal{A}$  tel que  $\mu(B_n) = 0$  et  $|f_n| \leq f$ .

On pose alors  $C = A \cup \left( \bigcup_{n \in \mathbb{N}} B_n \right)$  et on remarque au passage que  $C$  est de mesure nulle comme réunion dénombrable d'ensembles négligeables. On peut appliquer le Théorème de convergence dominée aux fonctions

$$f_n \mathbb{1}_{C^c} = f_n \mu\text{-pp pour tout } n \in \mathbb{N}, \quad f \mathbb{1}_{C^c} = f \mu\text{-pp} \quad \text{et} \quad g \mathbb{1}_{C^c} = g \mu\text{-pp}$$

On conclut en remarquant que  $\lim_{n \rightarrow +\infty} \int_X |f_n \mathbb{1}_{C^c} - f \mathbb{1}_{C^c}| d\mu = \int_X |f_n - f| d\mu = 0$  et que

$$\lim_{n \rightarrow +\infty} \int_X f_n d\mu = \lim_{n \rightarrow +\infty} \int_X f_n \mathbb{1}_{C^c} d\mu = \int_X f \mathbb{1}_{C^c} d\mu = \int_X f d\mu. \quad \square$$

**Exemple 3.3.1.** L'hypothèse de domination est nécessaire. Sur  $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda_1)$  on considère la suite de fonctions  $(\mathbb{1}_{[n, n+1]})_{n \in \mathbb{N}}$ . La plus petite fonction qui majore toutes les fonctions de cette suite est  $g = \mathbb{1}_{[0, +\infty[}$  qui n'est pas intégrable et on a  $\lim_{n \rightarrow +\infty} \mathbb{1}_{[n, n+1]} = 0$  alors que  $\int_X \mathbb{1}_{[n, n+1]} d\lambda_1 = 1$  pour tout  $n \in \mathbb{N}$ .

**Corollaire 3.3.2.** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré et  $(f_n)_{n \in \mathbb{N}} \in (\mathcal{L}_{\mathbb{C}}^1(X))^{\mathbb{N}}$  une suite de fonctions intégrables sur  $X$  telle que  $\sum_{n \in \mathbb{N}} \int_X |f_n| d\mu < +\infty$ .

Alors la série  $\sum_{n \in \mathbb{N}} f_n$  converge  $\mu$ -pp vers une fonction intégrable  $F : X \rightarrow \mathbb{C}$  et  $\int_X F d\mu = \sum_{n \in \mathbb{N}} \int_X f_n d\mu$ .

*Démonstration.* Soit  $m : X \rightarrow \overline{\mathbb{R}}_+$  définie par  $m(x) = \sum_{n \in \mathbb{N}} |f_n(x)|$ . Par le théorème de convergence

monotone on a  $\int_X m d\mu = \sum_{n \in \mathbb{N}} \int_X |f_n| d\mu < +\infty$ , donc la fonction  $m$  est finie  $\mu$ -pp.

Soit  $E = \{x \in X \mid m(x) < +\infty\}$ , on a  $\mu(E^c) = 0$  et, pour tout  $x \in E$  la série numérique  $\sum_{n \in \mathbb{N}} f_n(x)$  est absolument convergente, donc convergente. Notons  $F : X \rightarrow \mathbb{C}$  la fonction définie par

$$F(x) = \begin{cases} \sum_{n \in \mathbb{N}} f_n(x) & \text{si } x \in E \\ 0 & \text{sinon.} \end{cases}$$

La série de fonction  $\sum_{n \in \mathbb{N}} f_n$  converge  $\mu$ -pp vers cette fonction  $F$ .

Pour tout  $n \in \mathbb{N}$ , soit  $F_n : X \rightarrow \mathbb{C}$  définie par  $F_n(x) = \sum_{k=0}^n f_k(x)$ . Pour tout  $x \in E$  on a  $\lim_{n \rightarrow +\infty} F_n(x) = F(x)$ , donc  $\lim_{n \rightarrow +\infty} F_n = F$   $\mu$ -pp. Par ailleurs, pour tout  $n \in \mathbb{N}$  et tout  $x \in X$  on a

$$|F_n(x)| \leq \sum_{k=0}^n |f_k(x)| \leq m(x).$$

Comme  $m$  est intégrable, le Théorème de convergence dominée donne

$$\int_X F d\mu = \lim_{n \rightarrow +\infty} \int_X \sum_{k=0}^n f_k d\mu = \lim_{n \rightarrow +\infty} \sum_{k=0}^n \int_X f_k d\mu = \sum_{n \in \mathbb{N}} \int_X f_n d\mu. \quad \square$$

### 3.4 Intégrales dépendant d'un paramètre

Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré,  $T$  un ensemble de paramètres, et  $f : X \times T \rightarrow \mathbb{C}$  une fonction telle que  $x \mapsto f(x, t)$  soit intégrable pour tout  $t \in T$ . On peut alors considérer la fonction  $F : T \rightarrow \mathbb{C}$  définie par  $F(t) = \int_X f(x, t) d\mu(x)$ .

Question générale : la fonction  $F$  hérite-t-elle des propriétés de  $f$  vue comme fonction de  $t$ ? On va considérer deux cas.

- (i) Si  $T$  est une partie de  $\mathbb{R}^n$  et  $t \mapsto f(x, t)$  est continue pour tout  $x$ , la fonction  $F$  est-elle continue?
- (ii) Si  $T$  est un intervalle de  $\mathbb{R}$  et  $t \mapsto f(x, t)$  est dérivable pour tout  $x$ , la fonction  $F$  est-elle dérivable?

### 3.4.1 Continuité

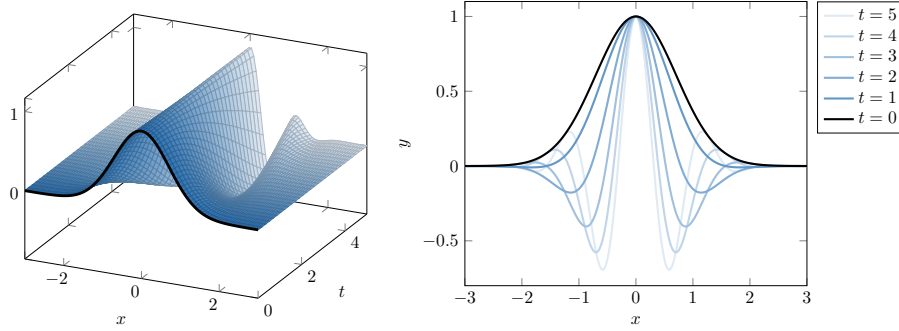


FIGURE 3.3 – Graphe de la fonction  $f(x, t) = \exp(-x^2) \cos(xt)$  pour différentes valeurs du paramètre  $t$ . Cette fonction est utilisée pour calculer l'intégrale de Gauss  $\int_{\mathbb{R}} f(\cdot, 0) d\lambda_1$ .

**Théorème 3.4.1 Théorème de continuité.** Soient  $(X, \mathcal{A}, \mu)$  un espace mesuré,  $U \subset \mathbb{R}^n$  un ouvert et  $f : X \times U \rightarrow \mathbb{C}$  vérifiant les hypothèses

- (i) Pour tout  $t \in U$  l'application  $x \mapsto f(x, t)$  est intégrable sur  $X$ .
- (ii) Pour  $\mu$ -pt  $x \in X$  l'application  $t \mapsto f(x, t)$  est continue sur  $U$ .
- (iii) Il existe  $g : X \rightarrow \mathbb{R}_+$  intégrable telle que, pour tout  $t \in U$ , on ait  $|f(x, t)| \leq g(x)$   $\mu$ -pp.

Alors l'application  $F : U \rightarrow \mathbb{C}$  définie par  $F(t) = \int_X f(x, t) d\mu(x)$  est continue sur  $U$ .

*Démonstration.* L'hypothèse (i) implique que la fonction  $F$  est bien définie. Soit  $t \in U$ , pour montrer que  $F$  est continue en  $t$ , il suffit de montrer que  $\lim_{n \rightarrow +\infty} F(t_n) = F(t)$  pour toute suite  $(t_n)_{n \in \mathbb{N}}$  à valeurs dans  $U$  et convergeant vers  $t$ .

Soit donc  $(t_n)_{n \in \mathbb{N}} \in U$  convergeant vers  $t$ , et pour tout  $n \in \mathbb{N}$ , soit  $u_n : X \rightarrow \mathbb{C}$  définie par  $u_n(x) = f(x, t_n)$ .

- La continuité de  $f$  par rapport à la variable  $t$  (hypothèses (ii)) donne  $\lim_{n \rightarrow +\infty} u_n(x) = \lim_{n \rightarrow +\infty} f(x, t_n) = f(x, t)$  pour  $\mu$ -pt  $x \in X$ .
- Par ailleurs, pour tout  $n \in \mathbb{N}$ , on a (par l'hypothèse (iii))  $|u_n(x)| = |f(x, t_n)| \leq g(x)$  pour  $\mu$ -pt  $x \in X$ . Comme  $g$  est intégrable.

Les hypothèses du Théorème de convergence dominée sont alors satisfaites et on a

$$\lim_{n \rightarrow +\infty} F(t_n) = \lim_{n \rightarrow +\infty} \int_X u_n d\mu = \int_X \lim_{n \rightarrow +\infty} u_n d\mu = \int_X f(x, t) d\mu(x) = F(t).$$

□

### 3.4.2 Dérivabilité

Le résultat suivant étend aux intégrales à paramètre le fait que *la dérivée de la somme, est la somme des dérivées*.

**Théorème 3.4.2 Théorème de dérivabilité.** Soient  $(X, \mathcal{A}, \mu)$  un espace mesuré,  $U \subset \mathbb{R}$  un ouvert, et  $f : X \times U \rightarrow \mathbb{C}$  vérifiant les hypothèses

- (i) Pour tout  $t \in U$  l'application  $x \mapsto f(x, t)$  est intégrable sur  $X$ .
- (ii) Pour  $\mu$ -pt  $x \in X$  l'application  $t \mapsto f(x, t)$  est dérivable sur  $U$ .
- (iii) Il existe  $g : X \rightarrow \mathbb{R}_+$  intégrable telle que, pour  $\mu$ -pt  $x \in X$ , on ait  $\forall t \in U, \left| \frac{\partial f}{\partial t}(x, t) \right| \leq g(x)$ .

Alors l'application  $F : U \rightarrow \mathbb{C}$  définie par  $F(t) = \int_X f(x, t) d\mu(x)$  est dérivable sur  $U$  et on a

$$F'(t) = \int_X \frac{\partial f}{\partial t}(x, t) d\mu(x). \quad (3.5)$$

De plus, si pour  $\mu$ -pt  $x \in X$  l'application  $t \mapsto f(x, t)$  est  $C^1$  sur  $U$ , alors  $F'$  est continue.

*Démonstration.* L'hypothèse (i) implique que la fonction  $F$  est bien définie. Fixons maintenant  $\tau \in U$  et montrons que la relation (3.5) est vérifiée en  $t = \tau$ . Comme  $U$  est un ouvert de  $\mathbb{R}$ , il existe  $\varepsilon > 0$  et un intervalle  $I = ]\tau - \varepsilon, \tau + \varepsilon[ \subset U$  qui contient une suite  $(t_n)_{n \in \mathbb{N}}$  convergeant vers  $\tau$  et avec  $t_n \neq \tau$  pour tout  $n \in \mathbb{N}$ . On calcule des taux d'accroissement en introduisant les fonctions  $h_n : X \rightarrow \mathbb{C}$  définies pour tout  $x \in X$  par

$$h_n(x) = \frac{f(x, t_n) - f(x, \tau)}{t_n - \tau}, \text{ pour tout } n \in \mathbb{N}.$$

- L'hypothèse (ii) implique que pour  $\mu$ -pt  $x \in X$ , la suite  $h_n(x)$  converge vers

$$h(x) = \frac{\partial f}{\partial t}(x, \tau).$$

On note donc  $A \in \mathcal{A}$ , l'ensemble des points  $x \in X$  pour lesquels  $(h_n)_{n \in \mathbb{N}}$  converge simplement vers  $h$ . L'ensemble  $A$  dépend de  $\tau$  et satisfait  $\mu(A^c) = 0$ .

- Reste à voir que l'hypothèse (iii) implique que les fonctions  $(h_n)_{n \in \mathbb{N}}$  sont bien majorées  $\mu$ -p.p. par la fonction  $g$  (pour pouvoir intervertir limite et intégrale). En effet, comme on a supposé que  $\tau, t_n \in I$  pour tout  $n \in \mathbb{N}$ , le Théorème des accroissements finis implique, qu'à  $x \in A$  fixé, il existe un  $\theta_{n,x} \in ]0, 1[$  tel que

$$|h_n(x)| = \left| \frac{f(x, t_n) - f(x, \tau)}{t_n - \tau} \right| = \left| \frac{\partial f}{\partial t}(x, \theta_{n,x}t_n + (1 - \theta_{n,x})\tau) \right| \leq g(x).$$

Ce qui est la majoration voulue.

On conclut en appliquant le Théorème de convergence dominée :

$$\lim_{n \rightarrow +\infty} \int_X h_n d\mu = \int_X h d\mu,$$

ce qui revient à dire que pour tout

$$F'(\tau) = \lim_{n \rightarrow +\infty} \frac{F(t_n) - F(\tau)}{t_n - \tau} = \int_X \frac{\partial f}{\partial t}(x, \tau) d\mu(x).$$

La dernière remarque du théorème se déduit en appliquant le Théorème 3.4.1 de continuité à l'intégrale à paramètre donnée par  $F'$ .  $\square$

Pour démontrer le caractère  $C^k$  ou  $C^\infty$ , on peut utiliser la proposition suivante qui se déduit du théorème de dérivation par une récurrence.

**Corollaire 3.4.1.** Soient  $(X, \mathcal{A}, \mu)$  un espace mesuré,  $U \subset \mathbb{R}$  un ouvert,  $n \in \mathbb{N}$ , et  $f : X \times U \rightarrow \mathbb{C}$  vérifiant les hypothèses

- (i) Pour tout  $t \in U$  l'application  $x \mapsto f(x, t)$  est intégrable sur  $X$ .
- (ii) Pour  $\mu$ -pt  $x \in X$  l'application  $t \mapsto f(x, t)$  est  $n$  fois dérivable sur  $U$ .
- (iii) Il existe des fonctions  $g_k : X \rightarrow \mathbb{R}_+$  pour  $k \in \llbracket 1, n \rrbracket$  toutes intégrables et telles que, pour  $\mu$ -pt  $x \in X$ , on ait  $\forall t \in U, \forall k \in \llbracket 1, n \rrbracket, \left| \frac{\partial^k f}{\partial t^k}(x, t) \right| \leq g_k(x)$ .

Alors l'application  $F : U \rightarrow \mathbb{C}$  définie par  $F(t) = \int_X f(x, t) d\mu(x)$  est  $n$  fois dérivable sur  $U$  et on a

$$F^{(k)}(t) = \int_X \frac{\partial^k f}{\partial t^k}(x, t) d\mu(x) \text{ pour tout } k \in \llbracket 1, n \rrbracket.$$

De plus, si pour  $\mu$ -pt  $x \in X$  l'application  $t \mapsto f(x, t)$  est  $C^n$  sur  $U$ , alors  $F^{(n)}$  est continue.

*Démonstration.* Laissé en exercice.  $\square$

Comme conséquence immédiate, si  $t \mapsto f(x, t)$  est  $C^\infty$  pour  $\mu$ -pt  $x \in X$  et si la domination  $\left| \frac{\partial^k f}{\partial t^k}(x, t) \right| \leq g_k(x)$  par une fonction intégrable  $g_k$  est satisfaite pour tout  $k \in \mathbb{N}$ , alors la fonction  $F$  est  $C^\infty$ .

Enfin, on peut considérer des dérivées partielles si le paramètre  $s = (s_1, \dots, s_n)$  appartient à un ouvert de  $\mathbb{R}^n$ .

**Corollaire 3.4.2.** Soient  $(X, \mathcal{A}, \mu)$  un espace mesuré,  $U \subset \mathbb{R}^n$  un ouvert, et  $f : X \times U \rightarrow \mathbb{C}$  vérifiant les hypothèses

- (i) Pour tout  $s \in U$  l'application  $x \mapsto f(x, s)$  est intégrable sur  $X$ .
- (ii) Pour  $\mu$ -pt  $x \in X$  l'application  $s \mapsto f(x, s)$  est de classe  $C^1$  sur  $U$ .
- (iii) Il existe des fonctions  $g_k : X \rightarrow \mathbb{R}_+$  pour  $k \in \llbracket 1, n \rrbracket$  toutes intégrables et telles que, pour  $\mu$ -pt  $x \in X$ , on ait  $\forall s \in U, \forall k \in \llbracket 1, n \rrbracket, \left| \frac{\partial f}{\partial s_k}(x, s) \right| \leq g_k(x)$ .

Alors l'application  $F : U \rightarrow \mathbb{C}$  définie par  $F(s) = \int_X f(x, s) d\mu(x)$  est  $C^1$  sur  $U$  et on a

$$\frac{\partial F}{\partial s_k}(s) = \int_X \frac{\partial f}{\partial s_k}(x, s) d\mu(x) \text{ pour tout } k \in \llbracket 1, n \rrbracket.$$

*Démonstration.* On rappelle ici que les fonctions dérivées partielles sont les fonctions dérivées des applications partielles : en notant  $e_1 = (1, 0, \dots, 0)$ ,  $e_2 = (0, 1, 0, \dots, 0)$ ,  $\dots$ ,  $e_n = (0, \dots, 0, 1)$  on a

$$\frac{\partial f}{\partial s_k}(x, s) = \lim_{\varepsilon \rightarrow 0} \frac{f(x, s + \varepsilon e_k) - f(x, s)}{\varepsilon}.$$

Le reste de la preuve revient à appliquer le Théorème 3.4.2 aux fonctions partielles  $t \mapsto f(x, s + te_k)$ ,  $k = \{1, \dots, n\}$  et est laissé au lecteur.  $\square$





## Chapitre 4

# Intégration pour quelques mesures particulières

### 4.1 Mesure de Dirac et mesure de comptage sur $\mathbb{N}$

#### 4.1.1 Mesure de Dirac

Soit  $(X, \mathcal{A})$  un espace mesurable, la mesure de Dirac en  $a \in X$  est l'application  $\delta_a : \mathcal{A} \rightarrow \overline{\mathbb{R}}_+$  définie par

$$\delta_a(A) = \begin{cases} 1 & \text{si } a \in A \\ 0 & \text{sinon} \end{cases}$$

**Proposition 4.1.1.** Soit  $(X, \mathcal{A})$  un espace mesurable et  $a \in X$ . Pour toute fonction mesurable  $f : X \rightarrow \mathbb{C}$  on a  $f \in \mathcal{L}^1(X, \delta_a)$  et  $\int_X f d\delta_a = f(a)$ .

*Démonstration.* La preuve repose sur la constatation que pour tout  $A \in \mathcal{A}$ , on a

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

par conséquent, on a  $\delta_a(A) = \mathbb{1}_A(a)$ .

Pour démontrer la proposition, on suit les étapes de la construction de l'intégrale.

Étape 1. Soit  $v = \sum_{k=1}^n \alpha_k \mathbb{1}_{A_k} \in \mathcal{E}_+(X)$ . Par définition de l'intégrale des fonctions étagées positives et en utilisant la remarque préliminaire, on a

$$\int_X v d\delta_a = \sum_{k=1}^n \alpha_k \delta_a(A_k) = \sum_{k=1}^n \alpha_k \mathbb{1}_{A_k}(a) = v(a).$$

Étape 2. Soit  $f : X \rightarrow \overline{\mathbb{R}}_+$  une fonction mesurable positive. On sait qu'il existe une suite croissante  $(v_n)_{n \in \mathbb{N}} \in \mathcal{E}_+(X)^{\mathbb{N}}$  telle que  $\lim_{n \rightarrow +\infty} v_n(x) = f(x)$  pour tout  $x \in X$ . Par le Théorème de convergence monotone on a

$$\int_X f d\delta_a = \lim_{n \rightarrow +\infty} \int_X v_n d\delta_a = \lim_{n \rightarrow +\infty} v_n(a) = f(a).$$

Étape 3. Soit  $f : X \rightarrow \mathbb{R}$  mesurable. La fonction  $|f|$  est mesurable positive, donc  $\int_X |f| d\delta_a = |f(a)| \in \mathbb{R}$ , donc  $f$  est intégrable. De plus, comme les fonctions  $f_+$  et  $f_-$  sont mesurables positives, on a

$$\int_X f d\delta_a = \int_X f_+ d\delta_a - \int_X f_- d\delta_a = f_+(a) - f_-(a) = f(a).$$

Si  $f : X \rightarrow \mathbb{C}$  est mesurable, alors  $\int_X |f| d\delta_a = |f(a)| \in \mathbb{R}$ , donc  $f$  est intégrable. Comme  $\operatorname{Re}(f)$  et  $\operatorname{Im}(f)$  sont à valeurs réelles on a

$$\int_X f d\delta_a = \int_X \operatorname{Re}(f) d\delta_a + i \int_X \operatorname{Im}(f) d\delta_a = \operatorname{Re}(f(a)) + i \operatorname{Im}(f(a)) = f(a). \quad \square$$

#### 4.1.2 Mesure de comptage sur $\mathbb{N}$

**Proposition 4.1.2.** Soit  $f : \mathbb{N} \rightarrow \mathbb{C}$ . On a  $f \in \mathcal{L}^1(\mathbb{N}, \mathcal{P}(\mathbb{N}), \chi)$  si et seulement si la série  $\sum_{k \in \mathbb{N}} |f(k)|$  converge dans  $\mathbb{R}$ , et on a alors  $\int_{\mathbb{N}} f d\chi = \sum_{k \in \mathbb{N}} f(k)$ .

*Démonstration.* La preuve suit les étapes de la construction de l'intégrale.

Étape 1. Soit  $v = 0 \cdot \mathbb{1}_{A_0} + \sum_{k=1}^n \alpha_k \mathbb{1}_{A_k} \in \mathcal{E}_+(\mathbb{N})$ , où  $\alpha_1, \dots, \alpha_n \in \mathbb{R}_+^*$  et  $(A_0, \dots, A_n)$  est une partition de  $\mathbb{N}$ . Si les ensembles  $A_1, \dots, A_n$  sont finis, alors on a

$$\begin{aligned} \int_{\mathbb{N}} v d\chi &= 0 \cdot \chi(A_0) + \sum_{k=1}^n \alpha_k \chi(A_k) \\ &= \sum_{k=1}^n \alpha_k \operatorname{Card}(A_k) \\ &= \sum_{k=1}^n \sum_{p \in A_k} \alpha_k \\ &= \sum_{k=1}^n \sum_{p \in A_k} v(p) \\ &= \sum_{p \in \cup_{k=1}^n A_k} v(p) \\ &= \sum_{p \in \mathbb{N}} v(p). \end{aligned}$$

S'il existe  $k_0 \in \llbracket 1, n \rrbracket$  tel que  $A_{k_0}$  soit infini, alors on a  $\int_{\mathbb{N}} v d\chi \geq \alpha_{k_0} \operatorname{Card}(A_{k_0}) = +\infty$  et  $\sum_{p \in \mathbb{N}} v(p) \geq \sum_{p \in A_{k_0}} v(p) = +\infty$ , donc  $\int_{\mathbb{N}} v d\chi = +\infty = \sum_{p \in \mathbb{N}} v(p)$ .

Étape 2. Soit  $f : \mathbb{N} \rightarrow \overline{\mathbb{R}}_+$  une fonction mesurable positive, et pour tout  $n \in \mathbb{N}$ , soit  $v_n = f \mathbb{1}_{\llbracket 0, n \rrbracket}$ . La suite  $(v_n)_{n \in \mathbb{N}}$  est une suite de  $\mathcal{E}_+(\mathbb{N})$ , comme  $f \geq 0$  cette suite est croissante, et on a  $\lim_{n \rightarrow +\infty} v_n(k) = f(k)$  pour tout  $k \in \mathbb{N}$ . Le Théorème de convergence monotone donne alors

$$\int_{\mathbb{N}} f d\chi = \lim_{n \rightarrow +\infty} \int_{\mathbb{N}} v_n d\chi = \lim_{n \rightarrow +\infty} \sum_{k=0}^n f(k) = \sum_{k=0}^{+\infty} f(k).$$

Étape 3. Soit  $f : \mathbb{N} \rightarrow \mathbb{R}$ . La fonction  $|f|$  est intégrable si et seulement si  $\sum_{k=0}^{+\infty} |f(k)| < +\infty$ , c'est à dire si et seulement si la série  $\sum f(k)$  est absolument convergente. Les séries  $\sum f_+(k)$  et

$\sum f_-(k)$  sont alors aussi absolument convergentes et on a

$$\begin{aligned} \int_{\mathbb{N}} f d\chi &= \int_{\mathbb{N}} f_+ d\chi - \int_{\mathbb{N}} f_- d\chi \\ &= \sum_{k=0}^{+\infty} f_+(k) - \sum_{k=0}^{+\infty} f_-(k) \\ &= \sum_{k=0}^{+\infty} f_+(k) - f_-(k) \\ &= \sum_{k=0}^{+\infty} f(k). \end{aligned}$$

La preuve est analogue pour les fonctions à valeurs complexes en utilisant les parties réelles et imaginaires.  $\square$

**Remarque 4.1.1.** Les fonctions intégrables sont celles pour lesquelles la série  $\sum f(n)$  est absolument convergente. Par exemple, la fonction  $f : \mathbb{N} \rightarrow \mathbb{R}$  définie par  $f(n) = \frac{(-1)^n}{n+1}$  n'est pas intégrable car  $\sum_{n=0}^{+\infty} \frac{1}{n+1} = +\infty$ , même si la série  $\sum_{n=0}^{+\infty} \frac{(-1)^n}{n+1} = \ln(2)$  converge.

**Remarque 4.1.2.** Comme les sommes de séries sont des intégrales pour la mesure de comptage, les séries de fonctions peuvent être vues comme des intégrales dépendant d'un paramètre, et les théorèmes de continuité et de dérivabilité s'appliquent.

## 4.2 Riemann vs. Lebesgue

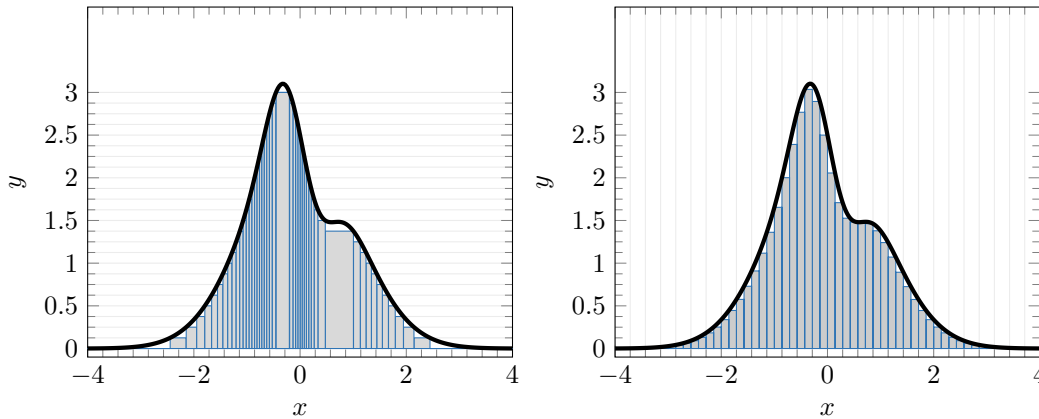


FIGURE 4.1 – Calcul d'intégrale par une approximation de minorantes étagées : à la Lebesgue et à la Riemann. Voir aussi à ce propos <https://www.geogebra.org/m/ygdfg9b7>.

### 4.2.1 Sur un segment

Soit  $[a, b]$  un segment de  $\mathbb{R}$  et  $f : [a, b] \rightarrow \mathbb{R}$  une application continue par morceaux. La fonction  $f$  admet une intégrale au sens de Riemann qu'on notera  $\int_a^b f(t)dt$ . En particulier, cette intégrale s'obtient comme limite des sommes de Riemann. Voir <https://www.geogebra.org/m/A9ka5Rbs>.

**Proposition 4.2.1.** Si  $f : [a, b] \rightarrow \mathbb{R}$  est une fonction continue par morceaux, alors  $f$  est intégrable pour la mesure de Lebesgue et  $\int_a^b f(t)dt = \int_{[a,b]} f d\lambda_1$ .

*Démonstration.* • La fonction  $f$  est Lebesgue intégrable. Comme  $f$  est continue par morceau, c'est une fonction borélienne (c'est l'objet de la Remarque 2.2.6). Elle est de plus bornée (car définie sur un compact) et est donc Lebesgue intégrable.

- L'intégrale de Riemann et l'intégrale de Lebesgue coïncident. Si  $f$  est une fonction étagée, le résultat est clair. Dans le cas général, on sait qu'il existe deux suites de fonctions étagées  $(\varphi_n)_{n \in \mathbb{N}}$  et  $(\psi_n)_{n \in \mathbb{N}}$  (e.g. les sommes de Darboux) telles que : pour tout  $n \in \mathbb{N}$  on ait  $\varphi_n \leq f \leq \psi_n$  et

$$\lim_{n \rightarrow +\infty} \int_a^b \varphi_n(x) dx = \lim_{n \rightarrow +\infty} \int_a^b \psi_n(x) dx = \int_a^b f(x) dx.$$

Par monotonie de l'intégration, on a pour tout  $n \geq 0$  :

$$\int_{[a,b]} \varphi_n(x) d\lambda_1 \leq \int_{[a,b]} f(x) d\lambda_1 \leq \int_{[a,b]} \psi_n(x) d\lambda_1.$$

Mais puisque les intégrales au sens de Lebesgue et de Riemann coïncident pour les fonctions étagées, on a donc :

$$\lim_{n \rightarrow +\infty} \int_a^b \varphi_n(x) dx = \lim_{n \rightarrow +\infty} \int_a^b \psi_n(x) dx = \int_{[a,b]} f d\lambda_1,$$

et par conséquent

$$\int_{[a,b]} f d\lambda_1 = \int_a^b f(x) dx. \quad \square$$

**Remarque 4.2.1.** • Il existe une notion de **fonction Riemann-intégrables** : ce sont les fonctions bornées sur  $[a, b]$  dont les sommes de Darboux inférieures et supérieures coïncident. Les fonctions continues par morceaux sont des exemples particuliers de fonctions Riemann-intégrables.

- Le résultat précédent est encore vrai pour les fonctions Riemann-intégrable, mais la question de la mesurabilité est délicate et nécessite l'introduction d'une nouvelle tribu qui contient la tribu borélienne.
- Dans la pratique, les fonctions Riemann-intégrables que nous rencontrerons seront en fait continues par morceau.

**Remarque 4.2.2.** Il y a plusieurs enseignements à tirer de ce résultat :

1. La réciproque est fausse : la fonction  $\mathbb{1}_{\mathbb{Q} \cap [a,b]}$  est Lebesgue-intégrable, mais pas Riemann-intégrable.
2. Dès lors qu'une fonction est continue par morceaux, on peut utiliser les outils usuels (issue de l'intégration au sens de Riemann) pour calculer son intégrale de Lebesgue : calcul de primitives, intégration par parties, changement de variable ...
3. La théorie générale de l'intégration de Lebesgue nous donne de nouveaux outils pour traiter des intégrales de Riemann : théorèmes de convergence monotone et dominée, théorèmes de continuité et de dérivabilité.

Dans la suite, lorsque  $f : [a, b] \rightarrow \mathbb{R}$  sera une fonction continue par morceaux, on notera indifféremment son intégrale  $\int_{[a,b]} f d\lambda_1$  ou  $\int_a^b f(t) dt$  ou  $\int_{[a,b]} f(t) dt$ .

## 4.2.2 Le cas des intégrales impropres

Si  $]a, b[$  est un intervalle non compact de  $\mathbb{R}$  (où  $a, b \in \overline{\mathbb{R}}$ ), l'intégrale de Riemann permet de définir, pour certaines fonctions  $f : ]a, b[ \rightarrow \mathbb{R}$  une intégrale sur  $]a, b[$ . Pour cela, on considère les intégrales de  $f$  sur des segments contenus dans  $]a, b[$  et on fait tendre (séparément) les bornes des segments vers les bornes de  $]a, b[$ .

Soit  $c \in ]a, b[$ . L'intégrale de  $f$  est convergente en  $a$  (resp. convergente en  $b$ ) si  $\lim_{\substack{x \rightarrow a \\ x > a}} \int_x^c f(t)dt$  existe et est finie (resp. si  $\lim_{\substack{x \rightarrow b \\ x < b}} \int_c^x f(t)dt$  existe et est finie). Si l'intégrale de  $f$  est convergente en  $a$  et en  $b$ , alors on note  $\int_a^b f(t)dt$  la somme de ces deux limites. Le résultat obtenu est indépendant du choix de  $c \in ]a, b[$  grâce à la relation de Chasle.

**Proposition 4.2.2.** Soient  $a, b \in \overline{\mathbb{R}}$  et  $f : ]a, b[ \rightarrow \mathbb{C}$  continue par morceaux. Si l'intégrale de Riemann de  $|f|$  est convergente en  $a$  et en  $b$ , alors on a  $f \in \mathcal{L}^1(]a, b[, \lambda_1)$  et  $\int_a^b f(t)dt = \int_{]a, b[} f d\lambda_1$ .

*Démonstration.* On suppose  $I = [0, +\infty[$ , les autres cas se traitant de la même façon.

- On considère la suite de fonctions  $(g_n)_{n \geq 0}$  définies par :

$$g_n = |f| \mathbb{1}_{[0, n]}.$$

Puisque l'intégrale de  $|f|$  est finie sur  $[0, n]$ , la suite  $(g_n)_{n \geq 0}$  est croissante, à valeurs dans l'espace des fonctions mesurables positives et convergeant simplement vers  $|f|$ . Par conséquent  $|f|$  est mesurable et par le Théorème de convergence monotone :

$$\lim_{n \rightarrow +\infty} \int_I g_n d\lambda_1 = \int_I |f| d\lambda_1.$$

Or d'après le résultat précédent :

$$\int_I g_n d\lambda_1 = \int_I \mathbb{1}_{[0, n]} |f| d\lambda_1 = \int_{[0, n]} |f| d\lambda_1 = \int_0^n |f(x)| dx$$

et par définition de l'intégrale de Riemann généralisée :

$$\lim_{n \rightarrow +\infty} \int_0^n |f(x)| dx = \int_0^{+\infty} |f(x)| dx.$$

On en déduit que :

$$\int_0^{+\infty} |f(x)| dx = \int_I |f| d\lambda_1.$$

- Considérons maintenant la suite de fonctions  $(f_n)_{n \geq 0}$  définies par :

$$f_n = f \mathbb{1}_{[0, n]}.$$

$(f_n)_{n \geq 0}$  est une suite de fonctions mesurables convergeant simplement vers  $f$ , dominées par la fonction intégrable  $|f|$ , donc on peut appliquer le Théorème de convergence dominée :

$$\lim_{n \rightarrow +\infty} \int_I f_n d\lambda_1 = \int_I f d\lambda_1.$$

Mais, d'après la Proposition 4.2.2 on sait que pour tout  $n \geq 0$  :

$$\int_I f_n d\lambda_1 = \int_0^n f(x) dx$$

et par définition de l'intégrale de Riemann généralisée :

$$\lim_{n \rightarrow +\infty} \int_0^n f(x) dx = \int_0^{+\infty} f(x) dx.$$

On en déduit que :

$$\int_0^{+\infty} f(x) dx = \int_I f d\lambda_1.$$

□

**Remarque 4.2.3.** Attention : il existe des fonctions dont l'intégrale de Riemann est convergente sur un intervalle, mais qui ne sont pas Lebesgue-intégrables sur cet intervalle. Un exemple est la fonction  $t \mapsto \frac{\sin t}{t}$  dont l'intégrale de Riemann est convergente sur  $\mathbb{R}_+$  :  $\int_{\mathbb{R}_+} \frac{\sin t}{t} dt = \frac{\pi}{2}$ . Mais cette fonction n'est pas intégrable car  $\int_{\mathbb{R}_+} \frac{|\sin t|}{t} dt = +\infty$ .

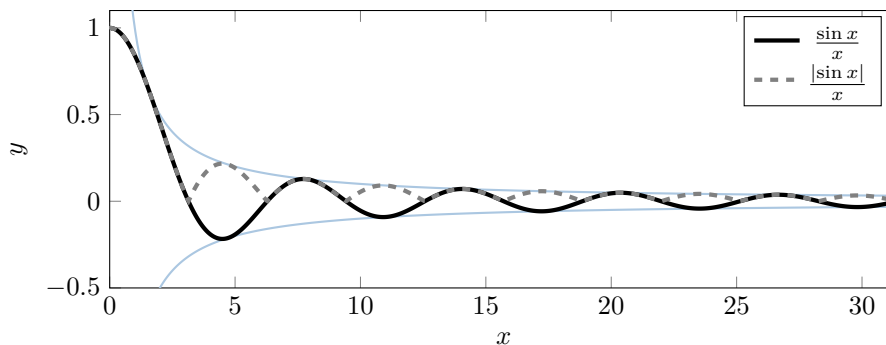


FIGURE 4.2 – La fonction sinus cardinal n'est pas Lebesgue intégrable sur  $\mathbb{R}_+$  mais possède une intégrale de Riemann convergente.

**Remarque 4.2.4.** Pour l'étude des intégrales impropres de fonctions continues par morceaux, il y a donc deux étapes :

1. Finitude de  $\int_I |f| d\lambda_1$  : tous les outils permettant de montrer que des intégrales impropres de fonctions positives sont convergentes peuvent être utilisés pour montrer que ces fonctions sont intégrables au sens de Lebesgue : inégalités, comparaison à des fonctions de référence...
2. Si  $\int_I |f| d\lambda_1 < +\infty$  la suite de l'étude peut se faire avec les méthodes usuelles : intégration par parties, changement de variable, calcul de primitives, théorèmes de convergence dominée, théorèmes de continuité et de dérivabilité.

## 4.3 Le théorème de transfert

### 4.3.1 Intégration et mesure image

Si  $(X, \mathcal{A}, \mu)$  est un espace mesuré,  $(Y, \mathcal{B})$  un espace mesurable et  $T : X \rightarrow Y$  une application mesurable, la mesure image  $T_*\mu$  est donnée par la Définition 2.4.1. L'intégration des fonctions pour la mesure image est donnée par le Théorème de transfert :

**Théorème 4.3.1.** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré,  $(Y, \mathcal{B})$  un espace mesurable et  $T : X \rightarrow Y$  une application mesurable. Pour toute fonction mesurable  $f : Y \rightarrow \mathbb{C}$  on a  $f \in \mathcal{L}^1_c(Y, T_*\mu)$  si et seulement si  $f \circ T \in \mathcal{L}^1_c(X, \mu)$ , et dans ce cas on a  $\int_Y f d(T_*\mu) = \int_X (f \circ T) d\mu$ .

*Démonstration.* Remarque préliminaire : soit  $B \in \mathcal{B}$  et  $x \in X$ , on a

$$\mathbb{1}_{T^{-1}(B)}(x) = 1 \Leftrightarrow x \in T^{-1}(B) \Leftrightarrow T(x) \in B \Leftrightarrow \mathbb{1}_B(T(x)) = 1$$

donc  $\mathbb{1}_{T^{-1}(B)} = \mathbb{1}_B \circ T$ .

Étape 1. Soit  $u \in \mathcal{E}_+(Y)$ ,  $u = \sum_{k=1}^n \beta_k \mathbf{1}_{B_k}$ . On a

$$\begin{aligned} \int_Y u d(T_*\mu) &= \sum_{k=1}^n \beta_k T_*\mu(B_k) \\ &= \sum_{k=1}^n \beta_k \mu(T^{-1}(B_k)) = \int_X v d\mu \end{aligned}$$

où  $v = \sum_{k=1}^n \beta_k \mathbf{1}_{T^{-1}(B_k)}$ . Compte tenu de la remarque préliminaire, on a  $v = \sum_{k=1}^n \beta_k \mathbf{1}_{B_k} \circ T =$

$$u \circ T \text{ et } \int_Y u d(T_*\mu) = \int_X (u \circ T) d\mu$$

Étape 2. Soit  $f : Y \rightarrow \overline{\mathbb{R}}_+$  mesurable et  $(u_n)_{n \in \mathbb{N}} \in \mathcal{E}_+(Y)^{\mathbb{N}}$  une suite croissante qui converge simplement vers  $f$ . Pour tout  $n \in \mathbb{N}$  on a  $u_n \circ T \in \mathcal{E}_+(X)$  et  $u_{n+1} \circ T \geq u_n \circ T$ . De plus la suite  $(u_n \circ T)_{n \in \mathbb{N}}$  converge simplement vers  $f \circ T$ . Par le Théorème de convergence monotone on a

$$\int_Y f d(T_*\mu) = \lim_{n \rightarrow +\infty} \int_Y u_n d(T_*\mu) = \lim_{n \rightarrow +\infty} \int_X (u_n \circ T) d\mu = \int_X (f \circ T) d\mu.$$

Étape 3. Soit  $f : Y \rightarrow \mathbb{C}$  mesurable. On a

$$\begin{aligned} f \in \mathcal{L}^1(Y, T_*\mu) &\Leftrightarrow \int_Y |f| d(T_*\mu) < +\infty \\ &\Leftrightarrow \int_X |f \circ T| d\mu < +\infty \\ &\Leftrightarrow f \circ T \in \mathcal{L}^1(X, \mu). \end{aligned}$$

Si  $f$  est à valeurs réelles, pour tout  $x \in X$  on a

$$f_+(T(x)) = \max \{f(T(x)), 0\} = \max \{(f \circ T)(x), 0\} = (f \circ T)_+(x)$$

donc  $f_+ \circ T = (f \circ T)_+$ . On a de même  $f_- \circ T = (f \circ T)_-$ . Si  $f \in \mathcal{L}^1(Y, T_*\mu)$  alors

$$\begin{aligned} \int_Y f d(T_*\mu) &= \int_Y f_+ d(T_*\mu) - \int_Y f_- d(T_*\mu) \\ &= \int_X (f_+ \circ T) d\mu - \int_X (f_- \circ T) d\mu \\ &= \int_X (f \circ T)_+ d\mu - \int_X (f \circ T)_- d\mu \\ &= \int_X (f \circ T) d\mu. \end{aligned}$$

Si  $f$  est à valeurs complexes, on a  $f = \operatorname{Re}(f) + i \operatorname{Im}(f)$  et  $f \circ T = \operatorname{Re}(f) \circ T + i \operatorname{Im}(f) \circ T$  en appliquant ce qui précède à  $\operatorname{Re}(f)$  et  $\operatorname{Im}(f)$  on obtient  $\int_Y f d(T_*\mu) = \int_X (f \circ T) d\mu$ .  $\square$

### 4.3.2 Un peu de probabilités

**Définition 4.3.1.** Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace probabilisé et  $X : \Omega \rightarrow \mathbb{R}$  une variable aléatoire. Si  $X \in \mathcal{L}^1(\Omega, \mathbb{P})$ , ou si  $X \geq 0$ , alors **l'espérance de  $X$**  est  $\mathbb{E}(X) = \int_{\Omega} X d\mathbb{P}$ .

L'espérance hérite de toutes les propriétés de l'intégrale : linéarité, monotonie, etc. On définit également, de façon similaire, l'espérance d'une variable aléatoire à valeurs complexes ou vectorielles.

**Exemple 4.3.1.** L'espérance est en fait la moyenne des valeurs prises par  $X$ . Dans le cas d'une variable aléatoire ne prenant qu'un nombre fini de valeurs, son espérance est bien la moyenne des ses valeurs pondérées par leurs probabilités. En effet, dans ce cas, la variable aléatoire est une fonction

étagée,  $X = \sum_{k=1}^n \alpha_k \mathbb{1}_{\{X=\alpha_k\}}$  et l'intégrale des fonctions étagées donne  $\mathbb{E}(X) = \sum_{k=1}^n \alpha_k \mathbb{P}(X = \alpha_k)$ .

Par exemple, si  $X$  prend ses  $n$  valeurs de façon équiprobables on obtient  $\mathbb{E}(X) = \frac{\alpha_1 + \dots + \alpha_n}{n}$ .

**Proposition 4.3.1 Formule de transfert.** Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace probabilisé et  $X : \Omega \rightarrow \mathbb{R}$  une variable aléatoire réelle et  $f : \mathbb{R} \rightarrow \mathbb{R}$  mesurable. Alors  $f(X) \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$  si et seulement si  $f \in \mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_X)$  et dans ce cas on a

$$\mathbb{E}(f(X)) = \int_{\Omega} f(X) d\mathbb{P} = \int_{\mathbb{R}} f d\mathbb{P}_X.$$

*Démonstration.* C'est une re-formulation du Théorème de transfert. □

En particulier, si  $f = \text{Id}_{\mathbb{R}}$  on obtient  $\mathbb{E}(X) = \int_{\Omega} X d\mathbb{P} = \int_{\mathbb{R}} x d\mathbb{P}_X(x)$ .

**Exemple 4.3.2.** 1. On lance un dés équilibré à 6 faces. On modélise cela par une variable aléatoire  $X$  de loi uniforme sur  $\{1, \dots, 6\}$  et  $\mathbb{E}(X) = \frac{1}{6}(1 + \dots + 6) = \frac{21}{6} = 3.5$ . L'espérance de  $Y = X^2$  est alors  $\mathbb{E}(Y) = \mathbb{E}(X^2) = \frac{1}{6}(1^2 + 2^2 + \dots + 6^2) = \frac{91}{6} = 15.16666\dots$   
 2. Soit  $X$  une variable aléatoire de loi uniforme sur  $\{-3, -2, \dots, 3\}$ . On a bien sûr  $\mathbb{E}(X) = 0$ . Pour calculer  $\mathbb{E}(X^2)$  on peut déterminer la loi de  $X^2$  :

$$\mathbb{P}(X^2 = k) = \begin{cases} \frac{1}{7}, & \text{si } k = 0 \\ \frac{2}{7}, & \text{si } k = 1, 4, 9 \\ 0, & \text{sinon.} \end{cases}$$

On a alors  $\mathbb{E}(X^2) = \frac{0}{7} + \frac{2}{7}(1 + 4 + 9) = 4$ . Mais on aurait pu utiliser directement le Théorème de transfert car on a  $\mathbb{E}(X^2) = \frac{1}{7}((-3)^2 + (-2)^2 + \dots + 3^2) = 4$ .

Donc la loi de  $X$  détermine les espérances de toutes les variables aléatoires de la forme  $f(X)$ . La réciproque est vraie : soient  $X, Y$  deux variables aléatoires réelles sur  $(\Omega, \mathcal{F}, \mathbb{P})$  telles que  $\mathbb{E}(f(X)) = \mathbb{E}(f(Y))$  pour toute fonction mesurable positive  $f : \mathbb{R} \rightarrow \mathbb{R}_+$ . Soit  $A \in \mathcal{B}(\mathbb{R})$ , on a alors

$$\mathbb{P}_X(A) = \int_{\mathbb{R}} \mathbb{1}_A d\mathbb{P}_X = \mathbb{E}(\mathbb{1}_A(X)) = \mathbb{E}(\mathbb{1}_A(Y)) = \int_{\mathbb{R}} \mathbb{1}_A d\mathbb{P}_Y = \mathbb{P}_Y(A).$$

Donc  $\mathbb{P}_X = \mathbb{P}_Y$  et les espérances des variables aléatoires de la forme  $f(X)$  déterminent la loi. De façon plus générale, on peut se demander quelles familles de fonctions déterminent la loi d'une variable aléatoire.

### Exemples de fonctions associées à une variable aléatoire

Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace probabilisé et  $X : \Omega \rightarrow \mathbb{R}$  une variable aléatoire.

- La **fonction de répartition** de  $X$  est  $F_X : \mathbb{R} \rightarrow \mathbb{R}$  définie par

$$F_X(t) = \mathbb{P}(X \leq t) = \mathbb{P}(X^{-1}(\cdot] - \infty, t]) = \mathbb{E}(\mathbb{1}_{\cdot] - \infty, t]}(X)) = \int_{\mathbb{R}} \mathbb{1}_{\cdot] - \infty, t]} d\mathbb{P}_X.$$

La fonction de répartition est déterminée par la loi de  $X$  et la réciproque est vraie : si deux variables aléatoires vérifient  $F_X = F_Y$ , alors  $\mathbb{P}_X = \mathbb{P}_Y$ .

- La **fonction génératrice des moments** de  $X$  est  $M_X : \mathbb{R} \rightarrow \mathbb{R}$  définie par

$$M_X(t) = \mathbb{E}(e^{tX}) = \int_{\mathbb{R}} e^{tx} d\mathbb{P}_X(x).$$

La fonction génératrice est déterminée par la loi de  $X$  et la réciproque est vraie : si deux variables aléatoires vérifient  $M_X = M_Y$ , alors  $\mathbb{P}_X = \mathbb{P}_Y$ .



- La **fonction caractéristique** de  $X$  est  $\varphi_X : \mathbb{R} \rightarrow \mathbb{C}$  définie par

$$\varphi_X(\xi) = \mathbb{E}(e^{i\xi X}) = \int_{\mathbb{R}} e^{i\xi x} d\mathbb{P}_X(x).$$

La fonction de caractéristique est déterminée par la loi de  $X$  et la réciproque est vraie : si deux variables aléatoires vérifient  $\varphi_X = \varphi_Y$ , alors  $\mathbb{P}_X = \mathbb{P}_Y$ .

## 4.4 Intégration sur les espaces produits

### 4.4.1 Tribu et mesure produit

Soient  $(X, \mathcal{A}, \mu)$  et  $(Y, \mathcal{B}, \nu)$  deux espaces mesurés, y a-t-il un moyen naturel de munir  $X \times Y$  d'une tribu et d'une mesure ?

#### Tribu produit

**Définition 4.4.1.** Soient  $(X, \mathcal{A})$  et  $(Y, \mathcal{B})$  des espaces mesurables. La **tribu produit** sur  $X \times Y$  est

$$\mathcal{A} \otimes \mathcal{B} = \sigma\left(\{A \times B \mid A \in \mathcal{A}, B \in \mathcal{B}\}\right).$$

**Remarque 4.4.1.** Soit  $X \times Y$  muni de la tribu produit.

- Si  $A \subset X$  et  $B \subset Y$  sont mesurables, alors  $A \times B \subset X \times Y$  est mesurable.
- Par contre, il existe des parties mesurables de  $X \times Y$  qui ne sont pas le produit de deux parties mesurables de  $X$  et  $Y$ . En effet,  $\{A \times B \mid A \in \mathcal{A}, B \in \mathcal{B}\}$  n'est pas une tribu (ce n'est pas stable par réunion).

À l'espace produit  $X \times Y$  on associe les deux projections canoniques

$$\begin{array}{ccc} \pi_1 : X \times Y & \rightarrow & X \\ (x, y) & \mapsto & x \end{array} \quad \text{et} \quad \begin{array}{ccc} \pi_2 : X \times Y & \rightarrow & Y \\ (x, y) & \mapsto & y \end{array}$$

**Proposition 4.4.1.** Soient  $(X, \mathcal{A})$  et  $(Y, \mathcal{B})$  des espaces mesurables.

1.  $\mathcal{A} \otimes \mathcal{B}$  est la plus petite tribu pour laquelle les projections  $\pi_1$  et  $\pi_2$  sont mesurables.
2. Soit  $(Z, \mathcal{C})$  et  $h : Z \rightarrow X \times Y$ . L'application  $h = (h_1, h_2)$  est mesurable si et seulement si les **applications coordonnées**

$$h_1 = \pi_1 \circ h : Z \rightarrow X \quad \text{et} \quad h_2 = \pi_2 \circ h : Z \rightarrow Y$$

sont mesurables.

*Démonstration.* 1. Pour tout  $A \in \mathcal{A}$  on a

$$\pi_1^{-1}(A) = \{(x, y) \in X \times Y \mid x \in A\} = A \times Y.$$

Comme  $A \in \mathcal{A}$  et  $Y \in \mathcal{B}$ , on a  $A \times Y \in \mathcal{A} \otimes \mathcal{B}$  et  $\pi_1$  est mesurable. On montre de même que  $\pi_2$  est mesurable.

Soit  $\mathcal{C}$  une tribu sur  $X \times Y$  telle que  $\pi_1$  et  $\pi_2$  soient mesurables. Pour tout  $A \in \mathcal{A}$  et  $B \in \mathcal{B}$ , et par mesurabilité de  $\pi_1$  et  $\pi_2$ , on a  $A \times Y = \pi_1^{-1}(A) \in \mathcal{C}$  et  $X \times B = \pi_2^{-1}(B) \in \mathcal{C}$ . On a donc  $A \times B = (A \times Y) \cap (X \times B) \in \mathcal{C}$  et  $\{A \times B \mid A \in \mathcal{A}, B \in \mathcal{B}\} \subset \mathcal{C}$ . Comme  $\mathcal{C}$  est une tribu, on a

$$\mathcal{A} \otimes \mathcal{B} = \sigma\left(\{A \times B \mid A \in \mathcal{A}, B \in \mathcal{B}\}\right) \subset \mathcal{C},$$

donc  $\mathcal{A} \otimes \mathcal{B}$  est la plus petite tribu pour laquelle les projections  $\pi_1$  et  $\pi_2$  sont mesurables.

2. Si  $h$  est mesurable, alors, compte tenu de ce qui précède, les applications  $h_1 = \pi_1 \circ h$  et  $h_2 = \pi_2 \circ h$  sont mesurables comme composées d'applications mesurables.

Réciproquement, supposons que  $h_1$  (resp.  $h_2$ ) soit  $\mathcal{C}$ - $\mathcal{A}$ -mesurable (resp.  $\mathcal{C}$ - $\mathcal{B}$ -mesurable). Comme  $\mathcal{A} \otimes \mathcal{B}$  est engendrée par  $\{A \times B \mid A \in \mathcal{A}, B \in \mathcal{B}\}$ , il suffit de montrer que  $h^{-1}(A \times B)$  est mesurable pour tout  $A \in \mathcal{A}$  et tout  $B \in \mathcal{B}$  (Proposition 2.2.1). Soit  $z \in Z$ , comme  $h(z) = (h_1(z), h_2(z))$  on a

$$\begin{aligned} z \in h^{-1}(A \times B) &\Leftrightarrow h(z) \in A \times B \\ &\Leftrightarrow h_1(z) \in A \text{ et } h_2(z) \in B \\ &\Leftrightarrow z \in h_1^{-1}(A) \text{ et } z \in h_2^{-1}(B) \\ &\Leftrightarrow z \in h_1^{-1}(A) \cap h_2^{-1}(B) \end{aligned}$$

donc  $h^{-1}(A \times B) = h_1^{-1}(A) \cap h_2^{-1}(B)$ . Comme  $h_1$  et  $h_2$  sont mesurables, on a  $h_1^{-1}(A) \in \mathcal{C}$  et  $h_2^{-1}(B) \in \mathcal{C}$ , donc  $h^{-1}(A \times B) \in \mathcal{C}$  et  $h$  est mesurable.  $\square$

**Proposition 4.4.2.** Si  $X \subset \mathbb{R}^n$  et  $Y \subset \mathbb{R}^p$  alors  $X \times Y \subset \mathbb{R}^{n+p}$  et on a

$$\mathcal{B}(X \times Y) = \mathcal{B}(X) \otimes \mathcal{B}(Y).$$

*Démonstration.* On se contente d'écrire la démonstration dans le cas où  $X = Y = \mathbb{R}$ .

1. Première inclusion :  $\mathcal{B}(\mathbb{R}^2) \subset \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ . On montre que tout ouvert de  $\mathbb{R}^2$  est réunion au plus dénombrable de produits d'intervalles ouverts de  $\mathbb{R}$ . La preuve est similaire à celle du Lemme 2.1.1.

Soit  $O$  un ouvert de  $\mathbb{R}^2$ . Pour tout  $x = (x_1, x_2) \in O$ , il existe  $r > 0$  tel que  $]x_1 - r, x_1 + r[ \times ]x_2 - r, x_2 + r[ \subset O$ . Comme les rationnels sont denses dans  $\mathbb{R}$ , on peut trouver  $y_1 \in \mathbb{Q} \cap ]x_1 - r, x_1[$ ,  $z_1 \in \mathbb{Q} \cap ]x_1, x_1 + r[$ ,  $y_2 \in \mathbb{Q} \cap ]x_2 - r, x_2[$  et  $z_2 \in \mathbb{Q} \cap ]x_2, x_2 + r[$ . On a donc  $x \in ]y_1, z_1[ \times ]y_2, z_2[ \subset O$ .

On note alors  $I = \{(y_1, z_1, y_2, z_2) \in \mathbb{Q}^4 \mid ]y_1, z_1[ \times ]y_2, z_2[ \subset O\}$ . Pour tout  $x \in O$ , il existe donc  $(y_1, z_1, y_2, z_2) \in I$  tel que  $x \in ]y_1, z_1[ \times ]y_2, z_2[$ . On en déduit que

$$O = \bigcup_{(y_1, z_1, y_2, z_2) \in I} ]y_1, z_1[ \times ]y_2, z_2[.$$

Comme  $I$  est au plus dénombrable (car  $\mathbb{Q}^4$  est dénombrable), on en déduit que  $O \in \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ . On a ainsi montré que  $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$  est une tribu contenant tous les ouverts de  $\mathbb{R}^2$ , et donc contenant la tribu engendrée par les ouverts de  $\mathbb{R}^2$  (c'est-à-dire  $\mathcal{B}(\mathbb{R}^2)$ ). Donc,  $\mathcal{B}(\mathbb{R}^2) \subset \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ .

2. Réciproquement, on montre que  $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) \subset \mathcal{B}(\mathbb{R}^2)$ . On considère

$$\mathcal{T} = \{A \in \mathcal{B}(\mathbb{R}) \mid A \times \mathbb{R} \in \mathcal{B}(\mathbb{R}^2)\}.$$

Notons que si  $A$  est un ouvert de  $\mathbb{R}$ . Alors  $A \times \mathbb{R}$  est un ouvert de  $\mathbb{R}^2$  et donc un borélien de  $\mathbb{R}^2$ . Cela montre que  $\mathcal{T}$  contient tous les ouverts de  $\mathbb{R}$ . On montre maintenant que  $\mathcal{T}$  est une tribu.

- $\emptyset \in \mathcal{T}$  car  $\emptyset \times \mathbb{R} = \emptyset \in \mathcal{B}(\mathbb{R}^2)$ .
- On montre ici que  $\mathcal{T}$  est stable par passage au complémentaire. Soit  $A \in \mathcal{T}$ , on a  $A^c = \mathbb{R} \setminus A \in \mathcal{B}(\mathbb{R})$  et

$$A^c \times \mathbb{R} = \mathbb{R}^2 \setminus (A \times \mathbb{R}) \in \mathcal{B}(\mathbb{R}^2)$$

car  $A \times \mathbb{R} \in \mathcal{B}(\mathbb{R}^2)$ . Ce qui prouve que  $A^c \in \mathcal{T}$ .

- Enfin,  $\mathcal{T}$  est stable par union dénombrable. En effet, si  $(A_n)_{n \in \mathbb{N}} \subset \mathcal{T}$ , on a

$$\left( \bigcup_{n \in \mathbb{N}} A_n \right) \times \mathbb{R} = \bigcup_{n \in \mathbb{N}} (A_n \times \mathbb{R}) \in \mathcal{B}(\mathbb{R}^2)$$

car  $A_n \times \mathbb{R} \in \mathcal{B}(\mathbb{R}^2)$  pour tout  $n \in \mathbb{N}$ . Donc,  $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{T}$ .

Ainsi  $\mathcal{T}$  est une tribu sur  $\mathbb{R}$  contenant les ouverts de  $\mathbb{R}$ , ce qui prouve que  $\mathcal{T} \supset \mathcal{B}(\mathbb{R})$ . Cela montre que  $A \times \mathbb{R} \in \mathcal{B}(\mathbb{R}^2)$  pour tout  $A \in \mathcal{B}(\mathbb{R})$ . On montre de façon analogue que  $\mathbb{R} \times B \in \mathcal{B}(\mathbb{R}^2)$  pour tout  $B \in \mathcal{B}(\mathbb{R})$ .

Soit  $A, B \in \mathcal{B}(\mathbb{R})$  alors

$$A \times B = (A \times \mathbb{R}) \cap (\mathbb{R} \times B) \in \mathcal{B}(\mathbb{R}^2).$$

Ainsi,  $\mathcal{B}(\mathbb{R}^2)$  est une tribu contenant toutes les parties de la forme  $A \times B$  avec  $A, B \in \mathcal{B}(\mathbb{R})$ . Elle contient donc la tribu engendrée qui n'est autre que  $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ .

□

**Remarque 4.4.2.** En particulier, la proposition donne  $\mathcal{B}(\mathbb{R}^2) = \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ , de même que  $\mathcal{B}(\mathbb{R}^n) = \mathcal{B}(\mathbb{R}^p) \otimes \mathcal{B}(\mathbb{R}^{n-p})$  pour tout  $1 \leq p < n$ .

La preuve utilise l'existence d'une partie dénombrable dense de  $X$  et de  $Y$ . Cette propriété n'est pas vérifiée par tous les espaces métriques, et la proposition se généralise uniquement aux espaces métriques possédant une partie dénombrable dense.

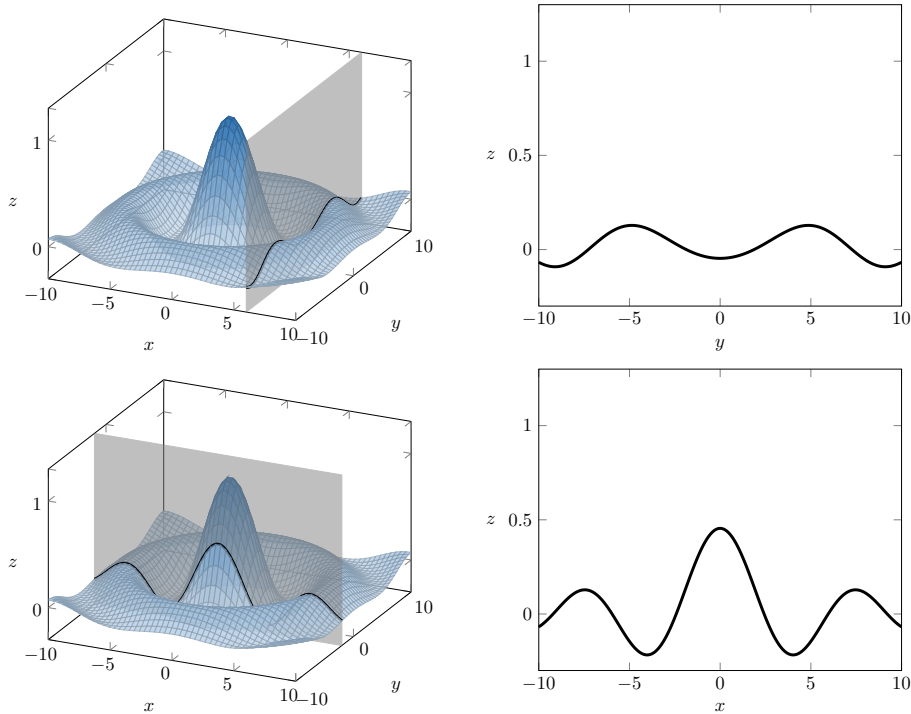


FIGURE 4.3 – Applications partielles de  $f(x, y) = \frac{\sin(\sqrt{x^2 + y^2})}{\sqrt{x^2 + y^2}}$ . En haut :  $f(6, \cdot)$ . En bas :  $f(\cdot, -2)$ .

**Définition 4.4.2.** Soit  $f : X \times Y \rightarrow Z$  une application et  $(a, b) \in X \times Y$ . On notera  $f(a, \cdot)$  et  $f(\cdot, b)$  les **applications partielles** :

$$\begin{array}{ccc} f(a, \cdot) : Y & \rightarrow & Z \\ y & \mapsto & f(a, y) \end{array} \quad \text{et} \quad \begin{array}{ccc} f(\cdot, b) : X & \rightarrow & Z \\ x & \mapsto & f(x, b) \end{array}$$

De même, pour toute partie  $E \subset X \times Y$  on définit les **sections** de  $E$  par

$$E_{a, \cdot} = \{y \in Y \mid (a, y) \in E\} \quad \text{et} \quad E_{\cdot, b} = \{x \in X \mid (x, b) \in E\}.$$

**Remarque 4.4.3.** Avec ces notations on a  $\mathbb{1}_E(a, \cdot) = \mathbb{1}_{E_{a, \cdot}}$  et  $\mathbb{1}_E(\cdot, b) = \mathbb{1}_{E_{\cdot, b}}$ .

**Proposition 4.4.3.** Soient  $(X, \mathcal{A})$  et  $(Y, \mathcal{B})$  des espaces mesurables.

1. Pour tout  $E \in \mathcal{A} \otimes \mathcal{B}$ , et tout  $(a, b) \in X \times Y$  on a  $E_{a, \cdot} \in \mathcal{B}$  et  $E_{\cdot, b} \in \mathcal{A}$ .
2. Pour toute application mesurable  $f : X \times Y \rightarrow (Z, \mathcal{C})$  et tout  $(a, b) \in X \times Y$ , l'application  $f(a, \cdot) : Y \rightarrow Z$  est  $\mathcal{B}$ - $\mathcal{C}$ -mesurable et l'application  $f(\cdot, b) : X \rightarrow Z$  est  $\mathcal{A}$ - $\mathcal{C}$ -mesurable.

*Démonstration.* 1. On fait la preuve pour  $E_{a, \cdot}$ , la preuve pour  $E_{\cdot, b}$  étant similaire.

Soit  $a \in X$  et  $\mathcal{C}_a = \{E \in \mathcal{A} \otimes \mathcal{B} \mid E_{a, \cdot} \in \mathcal{B}\}$ . On veut montrer que  $\mathcal{C}_a = \mathcal{A} \otimes \mathcal{B}$ , on a  $\mathcal{C}_a \subset \mathcal{A} \otimes \mathcal{B}$  par définition, et on prouve l'inclusion inverse en montrant que  $\mathcal{C}_a$  est une tribu qui contient  $\{A \times B \mid A \in \mathcal{A}, B \in \mathcal{B}\}$ .

- Soient  $A \in \mathcal{A}$  et  $B \in \mathcal{B}$ , on a

$$(A \times B)_{a, \cdot} = \{y \in Y \mid (a, y) \in A \times B\} = \begin{cases} \emptyset & \text{si } a \notin A \\ B & \text{si } a \in A. \end{cases}$$

Dans les deux cas on a  $(A \times B)_{a, \cdot} \in \mathcal{B}$ , donc  $A \times B \in \mathcal{C}_a$ . En particulier,  $X \times Y \in \mathcal{C}_a$ .

- Soit  $E \in \mathcal{C}_a$ , on a  $E_{a, \cdot} \in \mathcal{B}$  et, pour tout  $y \in Y$

$$y \in (E^c)_{a, \cdot} \Leftrightarrow (a, y) \in E^c \Leftrightarrow (a, y) \notin E \Leftrightarrow y \notin E_{a, \cdot} \Leftrightarrow y \in (E_{a, \cdot})^c.$$

On a donc  $(E^c)_{a, \cdot} = (E_{a, \cdot})^c$ , d'où  $(E^c)_{a, \cdot} \in \mathcal{B}$  et  $E^c \in \mathcal{C}_a$ .

- Soit  $(E_k)_{k \in \mathbb{N}} \in \mathcal{C}_a^{\mathbb{N}}$ , on a

$$\begin{aligned} y \in \left( \bigcup_{k \in \mathbb{N}} E_k \right)_{a, \cdot} &\Leftrightarrow (a, y) \in \bigcup_{k \in \mathbb{N}} E_k \\ &\Leftrightarrow \exists k \in \mathbb{N}, (a, y) \in E_k \\ &\Leftrightarrow \exists k \in \mathbb{N}, y \in (E_k)_{a, \cdot} \\ &\Leftrightarrow y \in \bigcup_{k \in \mathbb{N}} (E_k)_{a, \cdot}. \end{aligned}$$

On a donc  $\left( \bigcup_{k \in \mathbb{N}} E_k \right)_{a, \cdot} = \bigcup_{k \in \mathbb{N}} (E_k)_{a, \cdot}$ , d'où  $\left( \bigcup_{k \in \mathbb{N}} E_k \right)_{a, \cdot} \in \mathcal{B}$  et  $\bigcup_{k \in \mathbb{N}} E_k \in \mathcal{C}_a$ .

On a montré que  $\mathcal{C}_a$  est une tribu qui contient  $\{A \times B \mid A \in \mathcal{A}, B \in \mathcal{B}\}$ , donc  $\mathcal{A} \otimes \mathcal{B} \subset \mathcal{C}_a$  et  $\mathcal{C}_a = \mathcal{A} \otimes \mathcal{B}$ .

2. Soit  $a \in X$ . Pour tout  $C \in \mathcal{C}$  et  $y \in Y$  on a

$$y \in \left( f(a, \cdot) \right)^{-1}(C) \Leftrightarrow f(a, y) \in C \Leftrightarrow (a, y) \in f^{-1}(C) \Leftrightarrow y \in (f^{-1}(C))_{a, \cdot}.$$

Or  $f$  est mesurable, donc  $f^{-1}(C) \in \mathcal{A} \otimes \mathcal{B}$  et  $(f^{-1}(C))_{a, \cdot} \in \mathcal{B}$  par ce qui précède. L'application  $f(a, \cdot)$  est donc mesurable.

□

### Mesure produit.

Une façon naturelle de mettre une mesure sur  $(X \times Y, \mathcal{A} \otimes \mathcal{B})$  est de demander que la mesure d'un produit soit le produit des mesures de chacun des facteurs, comme pour les rectangles de  $\mathbb{R}^2$ . Cette propriété a le bon goût de caractériser la mesure produit.

**Théorème 4.4.1 Existence et caractérisation de la mesure produit.** Soient  $(X, \mathcal{A}, \mu)$  et  $(Y, \mathcal{B}, \nu)$  deux espaces mesurés  $\sigma$ -finis.

1. il existe une unique mesure, notée  $\mu \otimes \nu$ , sur  $(X \times Y, \mathcal{A} \otimes \mathcal{B})$  telle que

$$\forall A \in \mathcal{A}, \forall B \in \mathcal{B}, \quad \mu \otimes \nu(A \times B) = \mu(A)\nu(B).$$

2. Pour tout  $E \in \mathcal{A} \otimes \mathcal{B}$ , les fonctions  $x \mapsto \nu(E_x, \cdot)$  et  $y \mapsto \mu(E, \cdot, y)$  sont mesurables et on a

$$\mu \otimes \nu(E) = \int_X \nu(E_x, \cdot) d\mu(x) = \int_Y \mu(E, \cdot, y) d\nu(y)$$

*Démonstration.* Admis dans un premier temps. Ce théorème sera démontré au chapitre 7.  $\square$

**Remarque 4.4.4.** Cet énoncé est assez proche du théorème d'existence et de caractérisation de la mesure de Lebesgue sur  $\mathbb{R}^2$ , on remplace simplement les produits de segments par des produits de mesurables. Il n'est donc pas surprenant que la mesure de Lebesgue soit une mesure produit. Sur  $\mathbb{R}^{n+p} = \mathbb{R}^n \times \mathbb{R}^p$ , pour tout pavé  $P$  on a

$$P = [a_1, b_1] \times \cdots \times [a_n, b_n] \times [a_{n+1}, b_{n+1}] \times \cdots \times [a_{n+p}, b_{n+p}] = A \times B$$

où

$$A = [a_1, b_1] \times \cdots \times [a_n, b_n] \quad \text{et} \quad B = [a_{n+1}, b_{n+1}] \times \cdots \times [a_{n+p}, b_{n+p}].$$

la mesure  $\lambda_n \otimes \lambda_p$  vérifie donc

$$\begin{aligned} \lambda_n \otimes \lambda_p(P) &= \lambda_n(A) \lambda_p(B) \\ &= \prod_{k=1}^n (b_k - a_k) \cdot \prod_{k=n+1}^{n+p} (b_k - a_k) \\ &= \prod_{k=1}^{n+p} (b_k - a_k) \\ &= \lambda_{n+p}(P). \end{aligned}$$

La caractérisation de la mesure de Lebesgue donne donc  $\lambda_{n+p} = \lambda_n \otimes \lambda_p$

**Remarque 4.4.5.** Les résultats précédents s'étendent à des produits multiples. Si  $(X_1, \mathcal{A}_1, \mu_1), \dots, (X_n, \mathcal{A}_n, \mu_n)$  sont des espaces mesurés  $\sigma$ -finis alors

- la tribu produit sur  $X_1 \times \cdots \times X_n$  est

$$\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n = \sigma\left(\left\{A_1 \times \cdots \times A_n \mid \forall k \in \llbracket 1, n \rrbracket A_k \in \mathcal{A}_k\right\}\right).$$

- il existe une unique mesure  $\mu_1 \otimes \cdots \otimes \mu_n$  sur  $(X_1 \times \cdots \times X_n, \mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n)$  telle que

$$\mu_1 \otimes \cdots \otimes \mu_n(A_1 \times \cdots \times A_n) = \mu_1(A_1) \cdots \mu_n(A_n).$$

Tous les résultats précédents s'étendent aux produits multiples : les projections canoniques sont mesurables, et si  $f : X_1 \times \cdots \times X_n \rightarrow Z$  est mesurable, alors toutes ses fonctions partielles sont mesurables.

En particulier, sur  $\mathbb{R}^n$  on a  $\mathcal{B}(\mathbb{R}^n) = \mathcal{B}(\mathbb{R}) \otimes \cdots \otimes \mathcal{B}(\mathbb{R})$ , et  $\lambda_n = \lambda_1 \otimes \cdots \otimes \lambda_1$ .

#### 4.4.2 Théorèmes de Tonelli et Fubini

**Théorème 4.4.2 Théorème de Tonelli.** Soient  $(X, \mathcal{A}, \mu)$  et  $(Y, \mathcal{B}, \nu)$  deux espaces mesurés  $\sigma$ -finis, et soit  $f : X \times Y \rightarrow \overline{\mathbb{R}}_+$  une fonction  $\mathcal{A} \otimes \mathcal{B}$ -mesurable.

1. Les fonctions

$$\begin{aligned} \varphi_f : X &\rightarrow \overline{\mathbb{R}}_+ & \psi_f : Y &\rightarrow \overline{\mathbb{R}}_+ \\ x &\mapsto \int_Y f(x, y) d\nu(y) & y &\mapsto \int_X f(x, y) d\mu(x) \end{aligned} \quad \text{et}$$

sont mesurables.

2. On a

$$\int_{X \times Y} f d\mu \otimes \nu = \int_X \left( \int_Y f(x, y) d\nu(y) \right) d\mu(x) = \int_Y \left( \int_X f(x, y) d\mu(x) \right) d\nu(y) \quad (4.1)$$

*Démonstration.* La preuve se fait en trois temps. On montre les conclusions du Théorème pour les indicatrices, puis pour les fonctions étagées positives et enfin pour les applications mesurables.

Étape 1. Soit  $E \in \mathcal{A} \otimes \mathcal{B}$  et  $f = \mathbb{1}_E$ . La Proposition 4.4.3 implique que les fonctions partielles  $f(x, \cdot) = \mathbb{1}_{E_{x, \cdot}}$ , sont mesurables pour tout  $x \in X$ . De plus, le point 2 du Théorème 4.4.1 implique que la fonction définie par

$$\varphi_{\mathbb{1}_E}(x) = \int_Y \mathbb{1}_E(x, \cdot) d\nu = \int_Y \mathbb{1}_{E_{x, \cdot}} d\nu = \nu(E_{x, \cdot}), \text{ pour tout } x \in X$$

est  $\mathcal{A}$ - $\mathcal{B}(\overline{\mathbb{R}})$ -mesurable et que l'on a bien

$$\begin{aligned} \int_{X \times Y} \mathbb{1}_E d\mu \otimes \nu &= \mu \otimes \nu(E) \\ &= \int_X \nu(E_{x, \cdot}) d\mu(x) = \int_X \left( \int_Y \mathbb{1}_E(x, y) d\nu(y) \right) d\mu(x). \end{aligned}$$

Le même raisonnement fonctionne en considérant  $f(\cdot, y) = \mathbb{1}_{E_{\cdot, y}}$ , la fonction  $\psi_f$  est mesurable et on a bien la relation (4.1) pour les indicatrices.

Étape 2. Soit maintenant  $f : X \times Y \rightarrow \overline{\mathbb{R}}_+$  une fonction étagée. Il existe donc  $e_1, \dots, e_n \in \overline{\mathbb{R}}_+$  et une partition  $(E_1, \dots, E_n)$  de  $X$  composée d'éléments de  $\mathcal{A} \otimes \mathcal{B}$  tels que

$$f = \sum_{i=1}^n e_i \mathbb{1}_{E_i}.$$

On a alors pour tout  $x \in X$ ,  $f(x, \cdot) = \sum_{i=1}^n e_i \mathbb{1}_{E_i}(x, \cdot) : Y \rightarrow \overline{\mathbb{R}}_+$  mesurable comme combinaison linéaire de fonctions indicatrices mesurables (par la Proposition 4.4.3). De même, la fonction

$$\varphi_f(x) = \int_Y \left( \sum_{i=1}^n e_i \mathbb{1}_{E_i}(x, \cdot) \right) d\nu = \sum_{i=1}^n e_i \varphi_{\mathbb{1}_{E_i}}(x), \text{ pour tout } x \in X$$

est mesurable comme combinaison linéaire de fonctions mesurables (d'après l'étape précédente). Enfin, on a

$$\begin{aligned} \int_{X \times Y} f d\mu \otimes \nu &= \sum_{i=1}^n e_i \mu \otimes \nu(E_i) \\ &= \sum_{i=1}^n e_i \int_X \varphi_{\mathbb{1}_{E_i}} d\mu = \int_X \left( \sum_{i=1}^n e_i \varphi_{\mathbb{1}_{E_i}} \right) d\mu \\ &= \int_X \left( \sum_{i=1}^n e_i \int_Y \mathbb{1}_{E_i}(x, y) d\nu(y) \right) d\mu(x) = \int_X \left( \int_Y f(x, y) d\nu(y) \right) d\mu(x) \end{aligned}$$

Le même raisonnement fonctionne si l'on considère la seconde fonction partielle  $f(\cdot, y) : X \rightarrow \overline{\mathbb{R}}$  et cela donne bien la relation (4.1) pour les fonction dans  $\mathcal{E}_+(X \times Y)$ .

Étape 3. Soit maintenant  $f : X \times Y \rightarrow \overline{\mathbb{R}}_+$  une fonction mesurable positive. Il existe donc une suite croissante  $(f_n)_{n \in \mathbb{N}}$  dans  $\mathcal{E}_+(X \times Y)$  qui converge simplement vers  $f$ . Pour tout  $x \in X$ , la suite  $(f_n(x, \cdot))_{n \in \mathbb{N}}$  des applications partielles est elle aussi croissante et converge simplement vers  $f(x, \cdot)$ . Cela donne, au passage, la mesurabilité de l'application partielle  $f(x, \cdot)$ .

On peut alors appliquer le Théorème de convergence monotone pour la mesure  $\nu$ . Ainsi, pour tout  $x \in X$ ,

$$\lim_{n \rightarrow +\infty} \varphi_{f_n}(x) = \lim_{n \rightarrow +\infty} \int_Y f_n(x, \cdot) d\nu = \int_Y f(x, \cdot) d\nu = \varphi_f(x).$$

Par monotonie de l'intégrale, on constate que la suite  $\varphi_{f_n}$  est croissante et converge bien vers  $\varphi_f : X \rightarrow \overline{\mathbb{R}}$  qui est donc  $\mathcal{A}$ - $\mathcal{B}(\overline{\mathbb{R}})$ -mesurable.

On peut appliquer le Théorème de convergence monotone pour la mesure  $\mu$  et la mesure  $\mu \otimes \nu$  :

$$\lim_{n \rightarrow +\infty} \int_X \varphi_{f_n} d\mu = \int_X \varphi_f d\mu \quad \text{et} \quad \lim_{n \rightarrow +\infty} \int_{X \times Y} f_n d\mu \otimes \nu = \int_{X \times Y} f d\mu \otimes \nu.$$

Or, l'Étape 2. donne  $\int_{X \times Y} f_n d\mu \otimes \nu = \int_X \varphi_{f_n} d\mu$  pour tout  $n \in \mathbb{N}$ . On en déduit donc que  $\int_{X \times Y} f d\mu \otimes \nu = \int_X \varphi_f d\mu$  qui n'est autre que la première égalité dans la formule (4.1).

On conclut la preuve du théorème en intervertissant, encore une fois, les rôles des applications partielles et des mesures  $\mu$  et  $\nu$ .

□

**Remarque 4.4.6.** • Les égalités dans la formule (4.1) s'entendent dans  $\overline{\mathbb{R}}$ .

- Un cas d'usage classique du Théorème de Tonelli : pour montrer qu'une fonction  $X \times Y \rightarrow \overline{\mathbb{R}}_+$  est intégrable sur  $X \times Y$  (i.e.  $f \in \mathcal{L}^1(X \times Y, \mathcal{A} \otimes \mathcal{B}, \mu \otimes \nu)$ ), il suffit de montrer que  $\varphi_f$  (resp.  $\psi_f$ ) est intégrable sur  $X$  (resp. sur  $Y$ ) et donc dans  $\mathcal{L}^1(X, \mathcal{A}, \mu)$  (resp.  $\mathcal{L}^1(Y, \mathcal{B}, \nu)$ ).

**Théorème 4.4.3 Théorème de Fubini.** Soient  $(X, \mathcal{A}, \mu)$  et  $(Y, \mathcal{B}, \nu)$  deux espaces mesurés  $\sigma$ -finis, et soit  $f : X \times Y \rightarrow \mathbb{C}$  une fonction intégrable.

1. Pour  $\mu$ -pt  $x \in X$ , la fonction  $f(x, \cdot)$  est dans  $\mathcal{L}^1(Y, \nu)$  et la fonction  $\varphi_f : x \mapsto \int_Y f(x, y) d\nu(y)$  est intégrable sur  $(X, \mathcal{A}, \mu)$ .

Pour  $\nu$ -pt  $y \in Y$ , la fonction  $f(\cdot, y)$  est dans  $\mathcal{L}^1(X, \mu)$  et la fonction  $\psi_f : y \mapsto \int_X f(x, y) d\mu(x)$  est intégrable sur  $(Y, \mathcal{B}, \nu)$ .

2. On a

$$\int_{X \times Y} f d\mu \otimes \nu = \int_X \left( \int_Y f(x, y) d\nu(y) \right) d\mu(x) = \int_Y \left( \int_X f(x, y) d\mu(x) \right) d\nu(y) \quad (4.2)$$

*Démonstration.* On se contente d'écrire la preuve dans le cas où  $f : X \times Y \rightarrow \mathbb{R}$  (pour le cas complexe, il suffit de considérer la partie réelle et la partie imaginaire). Sous les hypothèses, on a  $f_-$  et  $f_+$  mesurables et on peut appliquer le Théorème de Tonelli :

$$\int_{X \times Y} f_- d\mu \otimes \nu = \int_X \varphi_{f_-} d\mu \quad \text{et} \quad \int_{X \times Y} f_+ d\mu \otimes \nu = \int_X \varphi_{f_+} d\mu.$$

- Les fonctions partielles  $f(x, \cdot)$  sont intégrables. Comme  $f \in \mathcal{L}^1(X \times Y, \mathcal{A} \otimes \mathcal{B}, \mu \otimes \nu)$  alors  $\int_{X \times Y} f_- d\mu \otimes \nu < +\infty$  et  $\int_{X \times Y} f_+ d\mu \otimes \nu < +\infty$ . Ainsi, les fonctions  $\varphi_{f_-} < +\infty$  et  $\varphi_{f_+} < +\infty$   $\mu$ -pp (Proposition 3.2.3) et les applications partielles  $f_-(x, \cdot), f_+(x, \cdot) \in \mathcal{L}^1(Y, \mathcal{B}, \nu)$  pour  $\mu$ -pt  $x \in X$ . On en déduit donc que  $f(x, \cdot) = f_+(x, \cdot) - f_-(x, \cdot) \in \mathcal{L}^1(Y, \mathcal{B}, \nu)$  pour  $\mu$ -pt  $x \in X$ .
- La fonction  $\varphi_f$  est intégrable. La fonction  $\varphi_f : X \rightarrow \mathbb{R}$  est donc définie  $\mu$ -pp et on a  $\varphi_f(x) = \varphi_{f_+}(x) - \varphi_{f_-}(x)$  en tout point  $x \in X$  tel que  $\varphi_{f_-}(x), \varphi_{f_+}(x) < +\infty$ . Comme  $\varphi_{f_-}, \varphi_{f_+} < +\infty$   $\mu$ -pp, on peut trouver un ensemble  $A \in \mathcal{A}$  tel que  $\mu(A) = 0$  avec  $\varphi_{f_-}, \varphi_{f_+} < +\infty$  sur  $A^c$ . En posant

$$g(x) = \begin{cases} \varphi_{f_+}(x) - \varphi_{f_-}(x) & \text{si } x \in A^c \\ 0 & \text{sinon} \end{cases}$$

on a  $g = \varphi_f$   $\mu$ -pp et  $g \in \mathcal{L}^1(X, \mathcal{A}, \mu)$  car  $\int_X |g| d\mu \leq \int_X \varphi_{f_+} d\mu + \int_X \varphi_{f_-} d\mu < +\infty$ .

- La formule (4.2) est vérifiée. La première égalité de la formule (4.2) est vraie car on a

$$\begin{aligned} \int_X \varphi_f d\mu &= \int_X g d\mu = \int_X \varphi_{f_+} d\mu - \int_X \varphi_{f_-} d\mu \\ &\stackrel{\text{Tonelli}}{=} \int_{X \times Y} f_+ d\mu \otimes \nu - \int_{X \times Y} f_- d\mu \otimes \nu = \int_{X \times Y} f d\mu \otimes \nu. \end{aligned}$$

Comme pour le Théorème de Tonelli, la relation (4.2) s'obtient en considérant l'application partielle  $f(\cdot, y)$  et en échangeant les rôles entre  $\mu$  et  $\nu$ .

□

**Remarque 4.4.7.** Si  $f : X \times Y \rightarrow \mathbb{C}$  est de la forme  $f(x, y) = g(x)h(y)$  et satisfait aux hypothèses du Théorème de Tonelli, ou aux hypothèses du Théorème de Fubini, alors on a

$$\begin{aligned} \int_{X \times Y} f d\mu \otimes \nu &= \int_X \left( \int_Y g(x)h(y) d\nu(y) \right) d\mu(x) \\ &= \int_X g(x) \left( \int_Y h(y) d\nu(y) \right) d\mu(x) \\ &= \left( \int_Y h(y) d\nu(y) \right) \left( \int_X g(x) d\mu(x) \right) \end{aligned}$$

**Exemple 4.4.1.** Dans l'énoncé du Théorème de Fubini, l'hypothèse d'intégrabilité est nécessaire. Par exemple, la fonction  $(x, y) \mapsto \frac{x^2 - y^2}{(x^2 + y^2)^2}$  n'est pas intégrable sur  $]0, 1[ \times ]0, 1[$ , ses intégrales itérées sont bien définies, mais différentes :

$$\int_0^1 \left( \int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dy \right) dx = \frac{\pi}{4} \quad \text{et} \quad \int_0^1 \left( \int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dx \right) dy = -\frac{\pi}{4}$$

Pour démontrer ces résultats, on peut poser  $v(y) = x^2 + y^2$  et chercher  $y \mapsto u(y)$  telle que

$$\frac{u'(y)v(y) - u(y)v'(y)}{v^2(y)} = \frac{x^2 - y^2}{(x^2 + y^2)^2}.$$

En procédant par identification, on trouve  $u(y) = y$ . Ce qui donne

$$\int_0^1 \left( \int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dy \right) dx = \int_0^1 \left[ \frac{y}{x^2 + y^2} \right]_{y=0}^1 dx = \int_0^1 \frac{1}{1 + x^2} dx = \frac{\pi}{4}.$$

## 4.5 Un exemple à connaître : la fonction $\Gamma$

**Proposition 4.5.1.** Pour tout  $x \in \mathbb{R}_+^*$  la fonction  $t \mapsto t^{x-1}e^{-t}$  est intégrable sur  $\mathbb{R}_+^*$  et la fonction

$$\begin{aligned} \Gamma : \mathbb{R}_+^* &\longrightarrow \mathbb{R} \\ x &\longmapsto \int_0^{+\infty} t^{x-1}e^{-t} dt \end{aligned}$$

a les propriétés suivantes :

1.  $\Gamma$  est de classe  $C^\infty$  sur  $\mathbb{R}_+^*$  et pour tout  $k \in \mathbb{N}$  on a  $\Gamma^{(k)}(x) = \int_0^{+\infty} (\ln t)^k t^{x-1} e^{-t} dt$ .
2. Pour tout  $x \in \mathbb{R}_+^*$  on a  $\Gamma(x+1) = x\Gamma(x)$ . En particulier,  $\Gamma(n+1) = n!$  pour tout  $n \in \mathbb{N}$ .

*Démonstration.* Dans toute la preuve, on note

$$\begin{aligned} f : \mathbb{R}_+^* \times \mathbb{R}_+^* &\longrightarrow \mathbb{R} \\ (t, x) &\longmapsto t^{x-1}e^{-t} . \end{aligned}$$

Étape 1. La fonction  $\Gamma$  est bien définie sur  $\mathbb{R}_+^*$ . Soit  $x \in \mathbb{R}_+^*$ . Pour tout  $t \in \mathbb{R}_+^*$  on a  $f(t, x) \geq 0$ . Donc, pour montrer que  $t \mapsto f(t, x)$  est intégrable sur  $\mathbb{R}_+^*$ , il suffit de montrer que  $\int_0^{+\infty} t^{x-1}e^{-t} dt$  est convergente en 0 et en  $+\infty$ .

- En  $x \rightarrow 0$  on a  $t^{x-1}e^{-t} \sim_0 t^{x-1}$  et  $\int_0^1 t^{x-1} dt$  est convergente, donc  $\int_0^{+\infty} t^{x-1}e^{-t} dt$  est convergente en 0.



- En  $x \rightarrow +\infty$  on a  $t^{x-1}e^{-t} = t^{x-1}e^{-t/2}e^{-t/2} = o(e^{-t/2})$  car  $\lim_{t \rightarrow +\infty} t^{x-1}e^{-t/2} = 0$ . Comme  $\int_0^{+\infty} e^{-t/2} dt$  est convergente,  $\int_0^{+\infty} t^{x-1}e^{-t} dt$  est aussi convergente en  $+\infty$  et la fonction  $\Gamma$  est bien définie.

Étape 2. La fonction  $\Gamma$  est continue sur  $\mathbb{R}_+^*$ . Soient  $a, b \in \mathbb{R}_+^*$  tels que  $0 < a < 1 < b$ , montrons que  $\Gamma$  est continue sur  $]a, b[$  en utilisant le Théorème de continuité.

- Pour tout  $x \in ]a, b[$ , la fonction  $t \mapsto f(t, x)$  est intégrable sur  $\mathbb{R}_+^*$  (cf. étape 1).
- Pour tout  $t \in \mathbb{R}_+^*$ , la fonction  $x \mapsto f(t, x)$  est continue sur  $]a, b[$ .
- Pour tout  $t \in \mathbb{R}_+^*$  et tout  $x \in ]a, b[$  on a
  - si  $t \geq 1$  alors  $x < b \Rightarrow t^{x-1}e^{-t} \leq t^{b-1}e^{-t} = u(t)$  ;
  - si  $t < 1$  alors  $x > a \Rightarrow t^{x-1}e^{-t} \leq t^{a-1}e^{-t} = v(t)$ .

Donc, pour tout  $(t, x) \in \mathbb{R}_+^* \times ]a, b[$  on a  $|f(t, x)| \leq g(t)$ , où  $g = u\mathbb{1}_{[1, +\infty[} + v\mathbb{1}_{]0, 1[}$ . De plus, le raisonnement de l'étape 1 montre que les intégrales  $\int_1^{+\infty} t^{b-1}e^{-t} dt$  et  $\int_0^1 t^{a-1}e^{-t} dt$  sont convergentes, donc les fonctions  $u\mathbb{1}_{[1, +\infty[}$  et  $v\mathbb{1}_{]0, 1[}$  sont intégrables, et  $g$  est aussi intégrable.

Les trois hypothèses du Théorème de continuité étant satisfaites, la fonction  $\Gamma$  est continue sur  $]a, b[$ , et la continuité étant une propriété locale, la fonction  $\Gamma$  est continue sur  $\bigcup_{0 < a < 1 < b} ]a, b[ = \mathbb{R}_+^*$ .

Étape 3. La fonction  $\Gamma$  est  $C^\infty$  sur  $\mathbb{R}_+^*$  et  $\Gamma^{(k)}(x) = \int_0^{+\infty} (\ln t)^k t^{x-1} e^{-t} dt$  pour tout  $k \in \mathbb{N}$ . Pour tout  $k \in \mathbb{N}$ , notons

$$\begin{aligned} f_k : \mathbb{R}_+^* \times \mathbb{R}_+^* &\longrightarrow \mathbb{R} \\ (t, x) &\longmapsto (\ln t)^k t^{x-1} e^{-t} \end{aligned}$$

Pour tout  $k \in \mathbb{N}$  et tout  $t \in \mathbb{R}_+^*$  la fonction  $x \mapsto f(t, x)$  est  $k$  fois dérivable et  $\frac{\partial^k f}{\partial x^k} = f_k$ .

Soient  $a, b \in \mathbb{R}_+^*$  tels que  $0 < a < b$ , pour tout  $(t, x) \in \mathbb{R}_+^* \times ]a, b[$  et tout  $k \in \mathbb{N}$ , on a les majorations suivantes :

- si  $t < 1$  alors  $x > a$  implique

$$|f_k(t, x)| \leq |\ln t|^k t^{a-1} e^{-t} \leq |\ln t|^k t^{a/2} t^{a/2-1} e^{-t} = v_k(t)$$

et  $v_k(t) = o(t^{a/2-1}e^{-t})$ , car  $\lim_{t \rightarrow 0} |\ln t|^k t^{a/2} = 0$ .

- si  $t \geq 1$  alors  $0 \leq \ln t \leq t$  et  $x < b \Rightarrow |f_k(t, x)| \leq t^{k+b-1}e^{-t} = u_k(t)$  ;

Pour tout  $k \in \mathbb{N}$  et tout  $(t, x) \in \mathbb{R}_+^* \times ]a, b[$  on a la majoration  $|f_k(t, x)| \leq g_k(t)$  où  $g_k = u_k\mathbb{1}_{[1, +\infty[} + v_k\mathbb{1}_{]0, 1[}$ . De plus, le raisonnement de l'étape 1 montre que les intégrales  $\int_1^{+\infty} t^{k+b-1}e^{-t} dt$  et  $\int_0^1 t^{a/2-1}e^{-t} dt$  sont convergentes, donc les fonctions  $u_k\mathbb{1}_{[1, +\infty[}$  et  $v_k\mathbb{1}_{]0, 1[}$  sont intégrables, et  $g_k$  est aussi intégrable.

De la Proposition 3.4.1 on déduit que la fonction  $\Gamma$  est  $C^\infty$  sur  $]a, b[$  et  $\Gamma^{(k)}(x) = \int_0^{+\infty} (\ln t)^k t^{x-1} e^{-t} dt$ .

On a montré que la fonction  $\Gamma$  est  $C^\infty$  sur  $]a, b[$ , et la dérivabilité étant une propriété locale, la fonction  $\Gamma$  est  $C^\infty$  sur  $\bigcup_{0 < a < b} ]a, b[ = \mathbb{R}_+^*$ .

Étape 4. Pour tout  $n \in \mathbb{N}$  on a  $\Gamma(n+1) = n!$ . Soit  $x \in \mathbb{R}_+^*$ , une intégration par partie donne

$$\Gamma(x+1) = \int_0^{+\infty} t^x e^{-t} dt = \left[ -t^x e^{-t} \right]_0^{+\infty} + \int_0^{+\infty} x t^{x-1} e^{-t} dt = x\Gamma(x)$$

Par ailleurs, on a  $\Gamma(1) = \int_0^{+\infty} e^{-t} dt = \left[ -e^{-t} \right]_0^{+\infty} = 1$ .

On a donc  $\Gamma(0+1) = 0!$  et  $\Gamma(n+1) = n\Gamma(n)$  pour tout  $n \in \mathbb{N}$ . Une récurrence immédiate donne  $\Gamma(n+1) = n!$  pour tout  $n \in \mathbb{N}$ .

□



## Chapitre 5

# Les espaces $L^p$ et la mesure de Lebesgue sur $\mathbb{R}^n$

Dans ce chapitre,  $(X, \mathcal{A}, \mu)$  désigne un espace mesuré, et  $\mathbb{K}$  est le corps  $\mathbb{R}$  ou  $\mathbb{C}$ .

Étant donné  $x \in \mathbb{R}^n$ , on note  $\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$  et  $B(x, r) = \{y \in \mathbb{R}^n \mid \|y - x\| < r\}$  la boule euclidienne ouverte centrée en  $x$  et de rayon  $r > 0$ .

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

L'idée est d'étendre les définitions des différentes normes rencontrées dans les espaces de dimension finies.

#### 5.1.1 Les semi-normes $\|\cdot\|_p$

##### L'espace $\mathcal{L}^\infty$ des fonctions essentiellement bornées

Soit  $f : X \rightarrow \mathbb{R}$  une fonction mesurable. On rappelle qu'un majorant de  $f$  est un nombre  $a \in \mathbb{R}$  vérifiant  $f(x) \leq a$  pour tout  $x \in X$ . Autrement dit, on a

$$\{f > a\} = f^{-1}(]a, +\infty[) = \emptyset.$$

Comme on travaille avec des espaces mesurés, on peut relaxer légèrement cette définition et considérer les **majorants essentiels**. C'est à dire les  $a \in \mathbb{R}$  tels que

$$\mu(\{f > a\}) = \mu(f^{-1}(]a, +\infty[)) = 0.$$

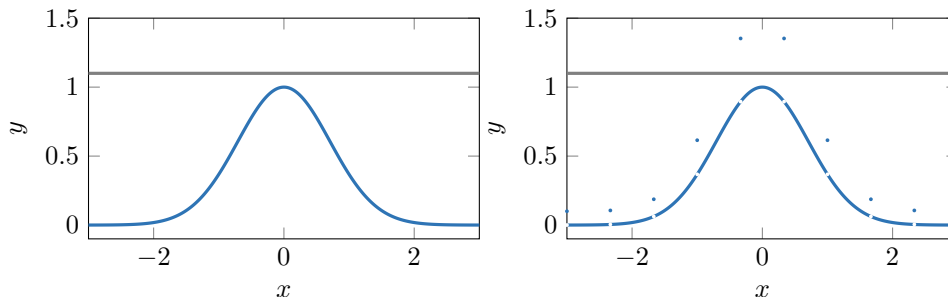


FIGURE 5.1 – A gauche, la fonction est bien majorée par  $a = 1.1$ . A droite, la fonction est majorée essentiellement par  $a = 1.1$

On arrive alors à cette définition :

**Définition 5.1.1.** Une fonction  $f : X \rightarrow \mathbb{K}$  mesurable est **essentiellement bornée** s'il existe  $a \in \mathbb{R}$  tel que  $|f| \leq a$   $\mu$ -pp, le réel  $a$  est une **borne essentielle** de  $f$ . On définit

$$\|f\|_\infty = \begin{cases} \inf \{a \in \mathbb{R}_+ \mid |f| \leq a \text{ } \mu\text{-pp}\} & \text{si } f \text{ est essentiellement bornée} \\ +\infty & \text{sinon.} \end{cases}$$

On note  $\mathcal{L}^\infty(X, \mathbb{K})$  ou  $\mathcal{L}^\infty(X)$  l'ensemble des fonctions essentiellement bornées.

**Exemple 5.1.1.** Une fonction peut être essentiellement bornée sans être bornée. Par exemple,  $f(x) = x \mathbf{1}_{\mathbb{N}}(x)$  n'est pas bornée, mais  $f = 0$   $\lambda_1$ -pp sur  $\mathbb{R}$ , elle est donc essentiellement bornée et  $\|f\|_\infty = 0$ .

**Lemme 5.1.1.** Pour toute fonction  $f : X \rightarrow \mathbb{K}$  on a  $|f(x)| \leq \|f\|_\infty$  pour  $\mu$ -pt  $x \in X$ .

En particulier, si  $f \in \mathcal{L}^\infty(X)$ , alors  $\|f\|_\infty$  est un majorant essentiel de  $f$ , et la définition implique que c'est le plus petit des majorants essentiels.

*Démonstration.* • Si  $\|f\|_\infty = +\infty$  on a bien sûr  $f \leq \|f\|_\infty$   $\mu$ -pp.

- Si  $\|f\|_\infty < +\infty$ . Par définition de la borne inférieure, il existe  $(C_n)_{n \in \mathbb{N}}$  suite décroissante de  $\{M \in \mathbb{R}_+ \mid |f| \leq M \text{ } \mu\text{-pp}\}$  qui converge vers  $\|f\|_\infty$  quand  $n \rightarrow +\infty$ . De plus, pour chaque  $n \in \mathbb{N}$ , il existe  $A_n \in \mathcal{A}$  tel que  $\mu(A_n) = 0$  et  $|f| < C_n$  sur  $A_n^c$ . En posant  $A = \bigcup_{n \in \mathbb{N}} A_n$ , on a  $\mu(A) = 0$  et  $|f(x)| \leq C_n$  pour tout  $n \in \mathbb{N}$  dès lors que  $x \in A^c$ . En passant à la limite, on a que  $|f(x)| \leq \|f\|_\infty$  pour tout  $x \in A^c$  et donc que  $|f| \leq \|f\|_\infty$   $\mu$ -pp. □

**Exemple 5.1.2.** Soit  $f : (\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda) \rightarrow \mathbb{R}$  une fonction continue, alors  $\sup_{x \in \mathbb{R}} \{|f(x)|\} = \|f\|_\infty$ . Voir Proposition 6.18 de [3]

**Proposition 5.1.1.** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré. L'espace  $\mathcal{L}^\infty(X)$  est un espace vectoriel.

*Démonstration.* • Soit  $\alpha \in \mathbb{R}$  et  $f \in \mathcal{L}^\infty(X)$ . On a  $\alpha f : X \rightarrow \mathbb{K}$  mesurable et  $|\alpha f(x)| = |\alpha| |f(x)| \leq |\alpha| \|f\|_\infty < +\infty$  pour  $\mu$ -pt  $x \in X$ . Et donc  $\alpha f \in \mathcal{L}^\infty(X)$ .

- Soit  $f, g \in \mathcal{L}^\infty(X)$ . On veut montrer que  $f + g \in \mathcal{L}^\infty(X)$ . La mesurabilité de  $f + g$  découle du fait que l'espace des fonctions mesurables est lui même un espace vectoriel. De plus,

$$\begin{aligned} |f(x) + g(x)| &\leq |f(x)| + |g(x)|, \quad \text{pour tout } x \in X \\ &\leq \|f\|_\infty + \|g\|_\infty, \quad \text{pour } \mu\text{-pt } x \in X \end{aligned}$$

ce qui montre que  $f + g \in \mathcal{L}^\infty(X)$ . □

### Les espaces $\mathcal{L}^p$ ds fonctions de puissance $p$ intégrables

**Définition 5.1.2.** Soit  $p \in [1, +\infty[$ . Si  $f : X \rightarrow \mathbb{K}$  est mesurable, on note

$$\|f\|_p \doteq \left( \int_X |f|^p d\mu \right)^{\frac{1}{p}} \in \overline{\mathbb{R}}_+,$$

et on note

$$\mathcal{L}^p(X) = \mathcal{L}^p(X, \mathbb{K}) \doteq \{f : X \rightarrow \mathbb{K} \mid f \text{ mesurable et } \|f\|_p < +\infty\}.$$

**Remarque 5.1.1.** •  $\|f\|_p$  est bien définie car la fonction  $|f|^p$  est mesurable positive.

- Si  $p = 1$  on retrouve l'espace des fonctions intégrables.
- $\|f\|_p$  est notée comme une norme, mais ce n'est pas une norme : si  $f = 0$   $\mu$ -pp, alors  $\|f\|_p = 0$ , mais  $f$  n'est pas forcément la fonction nulle.

**Proposition 5.1.2.** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré et  $1 \leq p < +\infty$ . L'espace  $\mathcal{L}^p(X)$  est un espace vectoriel.

*Démonstration.* • Soit  $\alpha \in \mathbb{R}$  et  $f \in \mathcal{L}^p(X)$ . On a  $\alpha f : X \rightarrow \mathbb{K}$  mesurable et  $\int_X |\alpha f|^p d\mu = |\alpha|^p \int_X |f|^p d\mu < +\infty$ . Et donc  $\alpha f \in \mathcal{L}^p(X)$ .  
 • Soit  $f, g \in \mathcal{L}^p(X)$ . On veut montrer que  $f + g \in \mathcal{L}^p(X)$ . La mesurabilité de  $f + g$  découle du fait que l'espace des fonctions mesurables est lui même un espace vectoriel. Comme  $p > 1$ , on utilise la convexité de  $t \mapsto t^p$  qui donne pour tout  $a, b \in \mathbb{R}$ ,  $\left|\frac{a+b}{2}\right|^p \leq \frac{|a|^p + |b|^p}{2} \leq |a|^p + |b|^p$ . Ainsi,

$$|f(x) + g(x)|^p \leq 2^p |f|^p + 2^p |g|^p$$

et donc

$$\int_X |f(x) + g(x)|^p d\mu \leq 2^p \int_X |f|^p d\mu + 2^p \int_X |g|^p d\mu < +\infty,$$

ce qui montre que  $f + g \in \mathcal{L}^p(X)$ . □

### Inégalités entre semi-normes

**Définition 5.1.3.** Deux éléments  $p, q \in [1, +\infty]$  sont des **exposants conjugués** si

$$\frac{1}{p} + \frac{1}{q} = 1.$$

On conviendra que  $\frac{1}{\infty} = 0$ , de sorte que 1 et  $\infty$  sont des exposants conjugués. Pour tout  $p \in ]1, +\infty[$ , l'exposant conjugué de  $p$  est  $\frac{p}{p-1}$ . En particulier, 2 est son propre conjugué.

**Lemme 5.1.2 Inégalités de Young.** Soient  $a, b \in \mathbb{R}_+$  et  $p, q \in ]1, +\infty[$  deux exposants conjugués. Alors

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

*Démonstration.* La fonction exponentielle  $\theta \mapsto e^\theta$  est convexe (de  $\mathbb{R}$  dans  $\mathbb{R}$ ). On a donc, pour tout  $\theta_1, \theta_2 \in \mathbb{R}$  et tout  $t \in [0, 1]$ ,

$$e^{t\theta_1 + (1-t)\theta_2} \leq te^{\theta_1} + (1-t)e^{\theta_2}.$$

Soit  $a, b > 0$  (les autres cas sont triviaux). On prend  $t = \frac{1}{p}$  (de sorte que  $(1-t) = \frac{1}{q}$ ),  $\theta_1 = p \ln(a)$  et  $\theta_2 = q \ln(b)$ . On obtient bien  $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$ . □

**Proposition 5.1.3 Inégalité de Hölder.** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré et  $p, q \in [1, +\infty]$  des exposants conjugués. Si  $f \in \mathcal{L}^p(X, \mathbb{K})$  et  $g \in \mathcal{L}^q(X, \mathbb{K})$  alors  $fg \in \mathcal{L}^1(X, \mathbb{K})$  et

$$\|fg\|_1 \leq \|f\|_p \|g\|_q. \quad (5.1)$$

*Démonstration.* Le produit  $fg$  est mesurable car  $f$  et  $g$  le sont.

- Si  $p = 1$  et  $q = +\infty$ , alors  $\int_X |f(x)g(x)| d\mu \leq \|g\|_\infty \int_X |f| d\mu = \|f\|_1 \|g\|_\infty$  et (5.1) est vérifiée.
- Si  $p \in ]1, +\infty[$  alors l'inégalité de Young donne  $|f(x)g(x)| \leq \frac{|f(x)|^p}{p} + \frac{|g(x)|^q}{q}$  pour tout  $x \in X$ . Les propriétés de l'intégrale permettent d'écrire :

$$\int_X |fg| d\mu \leq \frac{1}{p} \int_X |f|^p d\mu + \frac{1}{q} \int_X |g|^q d\mu < +\infty. \quad (5.2)$$

Donc,  $fg \in \mathcal{L}^1(X)$ . Pour montrer l'inégalité (5.1), on distingue maintenant 3 cas :

Cas 1.  $\|f\|_p = 0$  ou  $\|g\|_q = 0$ . On a alors  $f = 0$   $\mu$ -pp ou  $g = 0$   $\mu$ -pp. On en déduit  $fg = 0$   $\mu$ -pp, donc  $\|fg\|_1 = 0$  et (5.1) est vérifié.

Cas 2.  $\|f\|_p = 1$  et  $\|g\|_q = 1$ . On a alors, avec (5.2),

$$\|fg\|_1 = \int_X |fg| d\mu \leq \frac{1}{p} + \frac{1}{q} = 1 = \|f\|_p \|g\|_q.$$

L'inégalité (5.1) est donc vraie.

Cas 3.  $\|f\|_p > 0$  et  $\|g\|_q > 0$ . On pose alors  $f_1 = \frac{f}{\|f\|_p}$  et  $g_1 = \frac{g}{\|g\|_q}$ , de sorte que  $\|f_1\|_p = \|g_1\|_q = 1$ . Le Cas 2. donne alors

$$\frac{\|fg\|_1}{\|f\|_p \|g\|_q} = \|f_1 g_1\|_1 \leq 1.$$

Ce qui démontre l'inégalité (5.1). □

**Proposition 5.1.4 Inégalité de Minkowski.** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré et  $p \in [1, +\infty]$ . Pour toutes fonctions  $f, g \in \mathcal{L}^p(X, \mathbb{K})$  alors  $f + g \in \mathcal{L}^p(X, \mathbb{K})$  et on a

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p. \quad (5.3)$$

*Démonstration.* On remarque tout d'abord que les Propositions 5.1.1 et 5.1.2 impliquent que  $f + g$  est bien dans  $\mathcal{L}^p(X)$ .

- Si  $p = \infty$ , il suffit de remarquer que  $|f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f\|_\infty + \|g\|_\infty$ , pour  $\mu$ -pt  $x \in X$  et retrouver l'inégalité (5.3).
- Si  $p = 1$ , il suffit de remarquer que  $|f + g| \leq |f| + |g|$  et d'utiliser les propriétés de monotonie et de linéarité de l'intégrale pour déduire l'inégalité (5.3).
- Si  $p > 1$ , on peut supposer que  $\|f + g\|_p \neq 0$  sinon (5.3) est trivial. On remarque alors que

$$|f + g|^p = |f + g| |f + g|^{p-1} \leq FH + GH, \quad \text{avec} \quad \begin{cases} F = |f|, \\ G = |g|, \\ H = |f + g|^{p-1}. \end{cases} \quad (5.4)$$

On pose considère l'exposant  $q = \frac{p}{p-1}$  conjugué de  $p$ . On a  $F, G \in \mathcal{L}^p(X)$  par hypothèse et  $H \in \mathcal{L}^q(X)$  par construction. On peut alors appliquer l'inégalité de Hölder (5.1) :

$$\|FH\|_1 \leq \|F\|_p \|H\|_q, \quad \text{et} \quad \|GH\|_1 \leq \|G\|_p \|H\|_q,$$

on réinjecte cela dans (5.4) qui donne

$$\int_X |f + g|^p d\mu \leq (\|f\|_p + \|g\|_p) \left( \int_X |f + g|^p d\mu \right)^{1-\frac{1}{p}}.$$

On en déduit immédiatement l'inégalité (5.3). □

**Remarque 5.1.2.** L'inégalité de Minkowski implique que  $\|\cdot\|_p$  est une **semi-norme** sur  $\mathcal{L}^p(X)$ . Pour tout  $f, g \in \mathcal{L}^p(X)$ , on a

- $\|f\|_p \geq 0$  et  $\|0\|_p = 0$ .
- si  $\alpha \in \mathbb{K}$  alors  $\|\alpha f\|_p = |\alpha| \|f\|_p$ .
- $\|f + g\|_p \leq \|f\|_p + \|g\|_p$ .

Mais  $\|f\|_p = 0$  implique  $f = 0$   $\mu$ -pp mais n'implique pas  $f = 0$ .

### 5.1.2 Les espaces $L^p$

L'égalité  $\mu$ -pp permet de définir une relation d'équivalence  $\mathcal{R}$  sur  $\mathcal{L}^p(X)$  :  $f \mathcal{R} g \Leftrightarrow f = g$   $\mu$ -pp. Cette relation a un lien avec ce qui précède : pour tout  $p \in [1, +\infty]$  on a  $f \mathcal{R} g \Leftrightarrow \|f - g\|_p = 0$ . Le fait que  $\|\cdot\|_p$  soit une semi-norme permet de montrer que  $\mathcal{R}$  est bien une relation d'équivalence, et la construction ci-dessous est très générale et n'utilise que la propriété de semi-norme de  $\|\cdot\|_p$ .

**Définition 5.1.4.** Si  $(X, \mathcal{A}, \mu)$  est une espace mesuré, alors

$$L^p(X) = \mathcal{L}^p(X)/\mathcal{R} = \{[f] \mid f \in \mathcal{L}^p(X)\}$$

où  $[f]$  désigne la classe d'équivalence de  $f$  pour la relation  $\mathcal{R}$ .

Un élément de  $L^p(X)$  désigne donc une **partie** de  $\mathcal{L}^p(X)$ , autrement dit un ensemble de fonction de  $\mathcal{L}^p(X)$ .

**Proposition 5.1.5.** Les opérations suivantes sont bien définies sur  $L^p(X)$  :

- l'addition définie par  $[f] + [g] = [f + g]$  pour tout  $[f], [g] \in L^p(X)$  ;
- la multiplication externe définie par  $\alpha[f] = [\alpha f]$  pour tout  $[f] \in L^p(X)$  et tout  $\alpha \in \mathbb{K}$  ;
- la norme définie par  $\|[f]\|_p = \|f\|_p$  pour tout  $[f] \in L^p(X)$ .

Muni de ces opérations,  $L^p(X)$  est un  $\mathbb{K}$ -espaces vectoriel normé.

*Démonstration.* • La classe d'équivalence  $[f + g]$  ne dépend pas du choix de  $f$  dans  $[f]$  et  $g$  dans  $[g]$ . En effet,  $f = f_1$   $\mu$ -pp (resp.  $g = g_1$   $\mu$ -pp) implique qu'il existe  $A \in \mathcal{A}$  (resp.  $B \in \mathcal{A}$ ) tel que  $\mu(A) = 0$  (resp.  $\mu(B) = 0$ ) et  $f(x) = f_1(x)$  (resp.  $g(x) = g_1(x)$ ) pour tout  $x \in A^c$  (resp. tout  $x \in B^c$ ). Ainsi, pour tout  $x \in (A \cup B)^c$  on a  $f(x) + g(x) = f_1(x) + g_1(x)$  et donc  $f + g = f_1 + g_1$   $\mu$ -pp.

• La classe d'équivalence  $[\alpha f]$  ne dépend pas du choix de  $f$  dans  $[f]$  car  $f = f_1$   $\mu$ -pp implique  $\alpha f = \alpha f_1$   $\mu$ -pp.

• Le fait que la norme  $\|[f]\|_p$  ne dépende pas du représentant est une conséquence directe du point 3 la Proposition 3.2.3.

□

**Remarque 5.1.3.** • Dans la pratique, on oublie les  $[\cdot]$  : on désigne les classes d'équivalence par l'un de leurs représentants, quand on écrit  $f = g$  dans  $L^p(X)$ , on écrit en fait une égalité  $\mu$ -pp.

- Attention : il faut quand même faire la différence entre les fonctions dans  $\mathcal{L}^p(X)$  et les éléments de  $L^p(X)$ . Généralement, une propriété portant sur un élément de  $L^p(X)$  est décrite en utilisant un représentant dans la classe d'équivalence. Il faut alors s'assurer que tout ce qui est dit ne dépend pas du représentant choisi.
- Par exemple, si  $f \in L^p(\mathbb{R}, \lambda_1)$ , alors l'évaluation de  $f$  en  $x_0 \in \mathbb{R}$ , c'est à dire considérer  $f(x_0)$  n'a pas de sens car  $f$  est seulement définie aux ensembles négligeables près. Et comme le singleton  $\{x_0\}$  est  $\lambda_1$ -négligeable et on peut trouver deux représentants de  $f$  dont les valeurs en  $x_0$  diffèrent.

### Un peu de topologie

Sur un espace vectoriel normé  $(E, \|\cdot\|)$  on a une notion de convergence de suites similaire à celle qu'on a sur  $\mathbb{Q}$ ,  $\mathbb{R}$  ou  $\mathbb{C}$ . Il suffit de remplacer la valeur absolue (ou le module) par la norme : soient  $(u_n)_{n \in \mathbb{N}} \in E^{\mathbb{N}}$  et  $u \in E$ , on a  $\lim_{n \rightarrow +\infty} u_n = u$  si

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, \|u_n - u\| < \varepsilon,$$

ce qui revient à demander  $\lim_{n \rightarrow +\infty} \|u_n - u\| = 0$ . De même, une suite  $(u_n)_{n \in \mathbb{N}} \in E^{\mathbb{N}}$  est une **suite de Cauchy** si

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall p, q \geq N, \|u_p - u_q\| < \varepsilon.$$

Le lien entre suites convergentes et suites de Cauchy est le même que pour les suites de nombres : *si une suite est convergente, alors elle est une suite de Cauchy*, mais la réciproque est fausse en général.

Une différence fondamentale entre  $\mathbb{Q}$  et  $\mathbb{R}$  est liée à ces propriétés : il existe des suites de rationnels qui sont de Cauchy, mais ne convergent pas dans  $\mathbb{Q}$  (par exemple la suite définie par  $u_n = \sum_{k=0}^n \frac{1}{k!}$

et voir [https://en.wikipedia.org/wiki/Proof\\_that\\_e\\_is\\_irrational](https://en.wikipedia.org/wiki/Proof_that_e_is_irrational)), alors que sur  $\mathbb{R}$  toute suite de Cauchy est convergente. Pour les espaces vectoriels normés, il en va de même. Il existe des espaces sur lesquels les suites de Cauchy ne convergent pas nécessairement.

**Définition 5.1.5.** Un espace vectoriel normé  $(E, \|\cdot\|)$  est **complet** si toute suite de Cauchy de  $E^{\mathbb{N}}$  est convergente. On appelle **espace de Banach** les espaces vectoriels normés complets.

Les espaces vectoriels normés de dimension finie sur  $\mathbb{R}$  sont complets (c'est par exemple le cas de  $\mathbb{R}^n$  ou  $\mathbb{C}$ ), mais il existe des espaces de dimension infinie qui ne le sont pas. Les espaces complets ont un intérêt majeur : pour y montrer qu'une suite est convergente, il suffit de montrer qu'elle est de Cauchy (c'est ce qu'on appelle **appliquer le critère de Cauchy**). C'est par exemple ce que l'on fait pour montrer qu'une série numérique absolument convergente est bien convergente.

Cet intérêt est suffisamment important pour que, en présence d'un espace vectoriel normé, on cherche à savoir s'il est complet ou non.

**Théorème 5.1.1 Théorème de Riesz-Fischer.** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré et  $p \in [1, +\infty]$ . L'espace  $L^p(X)$  muni de la norme  $\|\cdot\|_p$  est un espace de Banach, c'est à dire un espace vectoriel normé complet.

*Démonstration.* Pour démontrer le théorème de Riesz-Fischer, la stratégie consiste à montrer que toute suite de Cauchy  $(f_n)_{n \in \mathbb{N}} \in L^p(X, \mu)^{\mathbb{N}}$  possède une sous-suite qui converge  $\mu$ -pp vers une certaine fonction  $f$ . On montre ensuite que  $f \in L^p(X, \mu)$  et que la suite  $(f_n)_{n \in \mathbb{N}}$  converge vers  $f$  en norme  $\|\cdot\|_p$ .

- Si  $1 \leq p < \infty$ . Par récurrence, on construit une suite extraite  $(f_{n_k})_{k \in \mathbb{N}}$  de  $(f_n)_{n \in \mathbb{N}}$  telle que pour tout  $k \geq 0$ ,  $\|f_{n_{k+1}} - f_{n_k}\|_p < 2^{-k}$ . Le problème est donc de construire une limite à cette suite. On note pour tout  $N \in \mathbb{N}$ ,

$$g_N(x) = \sum_{k=0}^N |f_{n_{k+1}}(x) - f_{n_k}(x)|, \quad x \in X.$$

Alors,  $g_N \in \mathcal{L}^p(X)$ , et par l'inégalité de Minkowski :

$$\|g_N\|_p \leq \sum_{k=0}^N \|f_{n_{k+1}} - f_{n_k}\|_p \leq \sum_{k=0}^N 2^{-k} < 2, \quad \forall N \in \mathbb{N}.$$

Ainsi, la suite  $(g_N^p)_{N \in \mathbb{N}}$  est une suite croissante de fonctions intégrables et positives et qui converge simplement vers  $x \mapsto \sup_{N \in \mathbb{N}} g_N(x) \in \overline{\mathbb{R}}_+$ . Par le Théorème de convergence monotone,  $\int_X \sup_{N \in \mathbb{N}} g_N(x) d\mu(x) = \sup_{N \in \mathbb{N}} \int_X g_N^p d\mu \leq 2^p < +\infty$  et qui implique que  $\sup_{N \in \mathbb{N}} g_N(x) < +\infty$  pour  $\mu$ -pt  $x \in X$  (Proposition 3.2.3 point 2). Plus précisément, il existe  $E \in \mathcal{A}$  tel que  $\mu(E) = 0$  et  $\sum_{k=0}^{\infty} |f_{n_{k+1}}(x) - f_{n_k}(x)| < +\infty$  pour tout  $x \in E^c$ . En outre, si  $g : X \rightarrow \mathbb{R}_+$  est définie par :

$$g(x) = \begin{cases} \sum_{k=0}^{\infty} |f_{n_{k+1}}(x) - f_{n_k}(x)|, & \text{si } x \notin E \\ 0, & \text{si } x \in E \end{cases}$$

alors  $g \in \mathcal{L}^p(X)$  et  $\|g\|_p \leq 2$ . D'autre part, on a évidemment :

$$f_{n_k}(x) = f_{n_0}(x) + \sum_{i=0}^{k-1} (f_{n_{i+1}}(x) - f_{n_i}(x)), \quad \forall x \in X.$$

On définit

$$f(x) = \begin{cases} \lim_{k \rightarrow \infty} f_{n_k}(x), & \text{si } x \notin E \\ 0, & \text{si } x \in E \end{cases}$$

Comme  $f_{n_k}(x) \xrightarrow[k \rightarrow \infty]{} f(x)$  pour  $\mu$ -pt  $x \in X$ , et que  $|f_{n_k}| \leq |f_{n_0}| + g \in \mathcal{L}^p(X)$ , le Théorème de convergence dominée montre que  $f \in \mathcal{L}^p(X)$ . Enfin,  $|f_{n_k}(x) - f(x)| \xrightarrow[k \rightarrow +\infty]{} 0$  presque



partout, et  $|f_{n_k} - f| \leq |f_{n_0} - f| + g \in \mathcal{L}^p(X)$ , donc par le même argument,  $\|f_{n_k} - f\|_p \xrightarrow{k \rightarrow \infty} 0$ . Enfin, comme  $(f_n)_{n \in \mathbb{N}}$  est une suite de Cauchy, on conclut que  $\|f_n - f\|_p \xrightarrow{n \rightarrow \infty} 0$ .<sup>1</sup>

- Si  $p = \infty$ . Pour  $n, m \in \mathbb{N}$ , on définit :

$$\begin{aligned} A_n &= \{x \in X \mid |f_n(x)| > \|f_n\|_\infty\}, \\ B_{n,m} &= \{x \in X \mid |f_n(x) - f_m(x)| > \|f_n - f_m\|_\infty\}, \\ E &= \left( \bigcup_{n \in \mathbb{N}} A_n \right) \cup \left( \bigcup_{\substack{n \in \mathbb{N} \\ m \in \mathbb{N}}} B_{n,m} \right). \end{aligned}$$

On a par définition  $\mu(A_n) = 0$ ,  $\mu(B_{n,m}) = 0$  qui donne  $\mu(E) = 0$ . Si  $x \notin E$ , on a

$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\|_\infty \xrightarrow{n, m \rightarrow \infty} 0.$$

On définit donc

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

Alors  $f$  est mesurable,  $|f(x)| \leq \liminf_{n \rightarrow +\infty} \|f_n\|_\infty$  pour tout  $x \in X$  et donc  $f \in \mathcal{L}^\infty(X)$  avec  $\|f\|_\infty \leq \liminf_{n \rightarrow +\infty} \|f_n\|_\infty < +\infty$ . De même,

$$|f_n(x) - f(x)| \leq \liminf_{m \rightarrow +\infty} \|f_n - f_m\|_\infty, \quad \text{pour tout } x \in X$$

$$\text{donc } \|f_n - f\|_\infty \leq \liminf_{n \rightarrow +\infty} \|f_n - f_m\|_\infty \xrightarrow{n \rightarrow +\infty} 0.$$

□

**Remarque 5.1.4.** Comme conséquence de cette preuve on a le résultat suivant : si la suite  $(f_n)_{n \in \mathbb{N}} \in L^p(X, \mu)^\mathbb{N}$  converge vers  $f \in L^p(X, \mu)$  en norme  $\|\cdot\|_p$  (c'est à dire si  $\lim_{n \rightarrow +\infty} \|f_n - f\|_p = 0$ ), alors il existe une sous suite  $(f_{n_k})_{k \in \mathbb{N}}$  telle que  $\lim_{k \rightarrow +\infty} f_{n_k} = f$   $\mu$ -pp.

**Exemple 5.1.3.** Sur  $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \chi)$ , la seule partie négligeable est l'ensemble vide. Pour  $a, b : \mathbb{N} \rightarrow \mathbb{C}$  on a donc  $a = b$   $\chi$ -pp et seulement si  $\forall k \in \mathbb{N} \ a_k = b_k$ . Dans ce cas très particulier, on a  $L^p(\mathbb{N}, \chi) = \mathcal{L}^p(\mathbb{N}, \chi)$ .

Pour  $p \in [1, +\infty]$ , on note

$$\ell^p = L^p(\mathbb{N}, \chi) = \left\{ (a_k)_{k \in \mathbb{N}} \in \mathbb{K}^\mathbb{N} \mid \sum_{k \in \mathbb{N}} |a_k|^p < +\infty \right\}.$$

Pour  $a \in \ell^p$  on a  $\|a\|_p = \left( \sum_{k \in \mathbb{N}} |a_k|^p \right)^{\frac{1}{p}}$  et  $(\ell^p, \|\cdot\|_p)$  est un espace de Banach.

**Remarque 5.1.5 Espace de Hilbert.** Lorsque  $p = 2$ , la norme  $\|\cdot\|_2$  sur l'espace  $L^2(X, \mu)$  est associée à un produit scalaire.

Si  $\mathbb{K} = \mathbb{R}$ , pour  $f, g \in L^2(X, \mu)$  l'inégalité de Hölder donne  $\int_X |fg| d\mu \leq \|f\|_2 \|g\|_2 < +\infty$ . La fonction  $fg$  est donc intégrable et on peut définir

$$\begin{aligned} \langle \cdot, \cdot \rangle : L^2(X, \mu) \times L^2(X, \mu) &\rightarrow \mathbb{R} \\ (f, g) &\mapsto \langle f, g \rangle = \int_X fg d\mu. \end{aligned} \tag{5.5}$$

---

1. L'argument montre que toute suite de Cauchy dans  $\mathcal{L}^p(X)$  possède une sous-suite qui converge point par point presque partout.

L'application  $\langle \cdot, \cdot \rangle$  est bilinéaire, symétrique et définie-positive : c'est un produit scalaire sur l'espace vectoriel  $L^2(X, \mu)$ . La norme associée à ce produit scalaire est

$$\langle f, f \rangle^{\frac{1}{2}} = \left( \int_X f^2 d\mu \right)^{\frac{1}{2}} = \|f\|_2.$$

Si  $\mathbb{K} = \mathbb{C}$ , on définit un produit sesquilinéaire par  $\langle f, g \rangle = \int_X f \bar{g} d\mu$ . La norme associée à ce produit est bien la norme  $\|\cdot\|_2$ . Un espace vectoriel normé complet dont la norme provient d'un produit scalaire s'appelle un **espace de Hilbert**.

### Inclusions des espaces $L^p$

Il est toujours utile d'avoir en tête les relations d'inclusion entre les espaces  $L^p$ . Étant donné un espace mesuré  $(X, \mathcal{A}, \mu)$ , on cherche ici à décrire les conditions assurant qu'une fonction dans  $L^p(X)$  appartienne aussi dans  $L^{p'}(X)$  pour  $p, p' \in [1, +\infty]$ . En général, il n'y a pas de règle simple d'inclusion sauf dans quelques cas particuliers. On en décrit deux dans ce cours : les espaces de mesure finie (penser aux espaces de probabilité) et ceux ne contenant aucun négligeable à part l'ensemble vide.

**Exemple 5.1.4.** On considère des fonctions  $(\mathbb{R}, \mathcal{B}, \lambda_1) \rightarrow \mathbb{R}$ .

1. La fonction  $x \mapsto e^{-x} \mathbb{1}_{\mathbb{R}_+}(x)$  appartient à  $\mathcal{L}^p(\mathbb{R})$  pour tout  $p \in [1, +\infty]$ .
2. La fonction  $x \mapsto \begin{cases} \frac{1}{x}, & \text{si } x \geq 1 \\ 0, & \text{sinon} \end{cases}$  appartient à  $\mathcal{L}^p(\mathbb{R})$  pour tout  $p \in ]1, +\infty]$ .
3. La fonction  $x \mapsto \begin{cases} \frac{1}{\sqrt{x}}, & \text{si } x \geq 1 \\ 0, & \text{sinon} \end{cases}$  appartient à  $\mathcal{L}^p(\mathbb{R})$  pour tout  $p \in ]2, +\infty]$ .
4. La fonction  $x \mapsto \begin{cases} \frac{1}{\sqrt{x}}, & \text{si } 0 < x \leq 1 \\ 0, & \text{sinon} \end{cases}$  appartient à  $\mathcal{L}^p(\mathbb{R})$  pour tout  $p \in [1, 2[$ .

Les deux propositions suivantes sont valables avec les espaces  $\mathcal{L}^p$  ou  $L^p$ .

**Proposition 5.1.6.** On suppose que  $\mu(X) < +\infty$ . Si  $f \in \mathcal{L}^p(X, \mathcal{A}, \mu)$  alors :

1. pour tout  $p' \in [1, p]$ , on a  $f \in \mathcal{L}^{p'}(X, \mathcal{A}, \mu)$  et on a

$$\|f\|_{p'} \leq C(X) \|f\|_p, \quad \text{où } C(X) = \mu(X)^{\frac{1}{p'} - \frac{1}{p}}.$$

2. Si  $f \in \mathcal{L}^\infty(X, \mathcal{A}, \mu)$  alors

$$\|f\|_\infty = \lim_{p \rightarrow \infty} \|f\|_p. \quad (5.6)$$

Dans ce cas on a l'inclusion :

$$L^\infty(X) \subset \dots \subset L^2(X) \subset \dots \subset L^1(X).$$

*Démonstration.* 1. C'est l'objet de l'exercice 2 du TD V.

2. Pour montrer (5.6), on peut supposer que  $\|f\|_\infty > 0$ . Soit  $\varepsilon \in ]0, 1[$ . Il existe  $E \in \mathcal{A}$  avec  $\mu(E) > 0$  tel que

$$|f(x)| \geq \|f\|_\infty (1 - \varepsilon), \quad \text{pour tout } x \in E.$$

Soit  $q \geq 1$  assez grand pour que  $\mu(E)^{\frac{1}{q}} \geq 1 - \varepsilon$  et que  $\mu(X)^{\frac{1}{q}} \leq 1 + \varepsilon$ . Si  $p \geq q$ , on a donc :

$$\|f\|_p \leq \|f\|_\infty \mu(X)^{\frac{1}{p}} \leq \|f\|_\infty (1 + \varepsilon).$$

D'autre part,

$$\|f\|_p^p = \int_X |f|^p d\mu \geq \int_E \|f\|_\infty^p (1 - \varepsilon)^p d\mu = \|f\|_\infty^p (1 - \varepsilon)^p \mu(E).$$

Donc

$$\|f\|_p \geq \|f\|_\infty (1 - \varepsilon) \mu(E)^{\frac{1}{p}} \geq \|f\|_\infty (1 - \varepsilon)^2.$$

Ainsi, pour tout  $p \geq q$  on a

$$(1 - \varepsilon)^2 \leq \frac{\|f\|_p}{\|f\|_\infty} \leq 1 + \varepsilon,$$

ce qui donne le résultat.  $\square$

**Exemple 5.1.5.** Dans les espaces de probabilité, une variable aléatoire qui admet une variance (*i.e.* qui est dans  $L^2(\Omega, \mathbb{P})$ ) admet nécessairement une moyenne (*i.e.* appartient à  $L^1(\Omega, \mathbb{P})$ )

**Proposition 5.1.7.** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré tel que  $\mu_0(X) > 0$ , où

$$\mu_0(X) = \inf\{\mu(E) \mid E \in \mathcal{A}, \mu(E) > 0\}.$$

Soient  $p, q \in [1, +\infty]$  et  $p \geq q$ . Alors, si  $f \in \mathcal{L}^q(X, \mathcal{A}, \mu)$ , on a  $f \in \mathcal{L}^p(X, \mathcal{A}, \mu)$  et

$$\|f\|_p \leq C(X) \|f\|_q, \quad \text{où } C(X) = \mu_0(X)^{\frac{1}{p} - \frac{1}{q}}.$$

Dans ce cas, on a l'emboîtement suivant :

$$L^1(X) \subset \dots \subset L^2(X) \subset \dots \subset L^\infty(X).$$

*Démonstration.* Soit  $f \in \mathcal{L}^q(X, \mathcal{A}, \mu)$  pour  $q < +\infty$ . S'il existe  $A > 0$  et  $E \in \mathcal{A}$  avec  $\mu(E) > 0$  tel que  $|f(x)| \geq A$  pour tout  $x \in E$ , alors :

$$\int_X |f|^q d\mu \geq \int_E A^q d\mu = A^q \mu(E) \geq A^q \mu_0(X).$$

Donc

$$A \leq \|f\|_q \left( \frac{1}{\mu_0(X)} \right)^{\frac{1}{q}}.$$

On a donc  $f \in \mathcal{L}^\infty(X, \mathcal{A}, \mu)$  et  $\|f\|_\infty \leq \|f\|_q \left( \frac{1}{\mu_0(X)} \right)^{\frac{1}{q}}$ . Si  $1 \leq q < p < +\infty$ , alors on a :

$$\int_X |f|^p d\mu = \int_X |f|^{p-q} |f|^q d\mu \leq \|f\|_\infty^{p-q} \|f\|_q^q \leq \|f\|_q^p \left( \frac{1}{\mu_0(X)} \right)^{\frac{p-q}{q}}$$

Donc  $\|f\|_p \leq \|f\|_q \left( \frac{1}{\mu_0(X)} \right)^{\frac{1}{q} - \frac{1}{p}}$ .  $\square$

**Exemple 5.1.6.** C'est le cas de l'espace mesuré  $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \chi)$  comme décrit dans l'Exemple 5.1.3. On a alors

$$\ell^1(\mathbb{N}) \subset \dots \subset \ell^2(\mathbb{N}) \subset \dots \subset \ell^\infty(\mathbb{N}).$$

## 5.2 La mesure de Lebesgue sur $\mathbb{R}^n$

La mesure de Lebesgue  $\lambda_n$  sur  $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$  est caractérisée par sa valeur sur les pavés :

$$\lambda_n([a_1, b_1] \times \dots \times [a_n, b_n]) = \prod_{i=1}^n (b_i - a_i).$$

La mesure de Lebesgue est finie sur les parties bornées : si  $A \in \mathcal{B}(\mathbb{R}^n)$  est bornée, alors il existe  $R > 0$  tel que  $A \subset [-R, R]^n$  et  $\lambda_n(A) \leq (2R)^n < +\infty$ . La mesure de Lebesgue est  $\sigma$ -finie :  $\mathbb{R}^n = \bigcup_{k \in \mathbb{N}} [-k, k]^n$ .

Par ailleurs,  $\mathbb{R}^n$  est un ensemble produit :  $\mathbb{R}^n = \mathbb{R} \times \dots \times \mathbb{R} = \mathbb{R}^p \times \mathbb{R}^{n-p}$  pour tout  $1 \leq p < n$ . La séparabilité donne  $\mathcal{B}(\mathbb{R}^n) = \mathcal{B}(\mathbb{R}) \otimes \dots \otimes \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^p) \otimes \mathcal{B}(\mathbb{R}^{n-p})$ , et les caractérisations des mesures de Lebesgue et des mesures produits donnent

$$\lambda_n = \lambda_1 \otimes \dots \otimes \lambda_1 = \lambda_p \otimes \lambda_{n-p}.$$

### 5.2.1 Invariance par translation

**Proposition 5.2.1.** Pour tout  $A \in \mathcal{B}(\mathbb{R}^n)$  et tout  $v \in \mathbb{R}^n$  on a  $\lambda_n(A + v) = \lambda_n(A)$ , où  $A + v = \{x + v \mid x \in A\}$

*Démonstration.* Soit  $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$  définie par  $T(x) = x - v$ , la translation de vecteur  $-v$ . Pour tout  $A \in \mathcal{B}(\mathbb{R}^n)$  on a  $T^{-1}(A) = A + v$  et

$$\lambda_n(A + v) = \lambda_n(T^{-1}(A)) = (T_*\lambda_n)(A).$$

On veut donc montrer que  $T_*\lambda_n = \lambda_n$ .

Si  $v = (v_1, \dots, v_n)$ , alors, pour tout pavé  $[a_1, b_1] \times \dots \times [a_n, b_n]$  on a  $T^{-1}([a_1, b_1] \times \dots \times [a_n, b_n]) = [a_1 + v_1, b_1 + v_1] \times \dots \times [a_n + v_n, b_n + v_n]$  et

$$\begin{aligned} (T_*\lambda_n)([a_1, b_1] \times \dots \times [a_n, b_n]) &= \lambda_n(T^{-1}([a_1, b_1] \times \dots \times [a_n, b_n])) \\ &= \lambda_n([a_1 + v_1, b_1 + v_1] \times \dots \times [a_n + v_n, b_n + v_n]) \\ &= \prod_{i=1}^n (b_i + v_i - (a_i + v_i)) \\ &= \prod_{i=1}^n (b_i - a_i) \end{aligned}$$

et par caractérisation de la mesure de Lebesgue on a  $T_*\lambda_n = \lambda_n$ . □

Cette invariance par translation caractérise la mesure de Lebesgue à une constante près :

**Proposition 5.2.2.** a Si  $\mu$  est une mesure sur  $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$  invariante par translations et finie sur les parties bornées, alors il existe une constante  $c \in \mathbb{R}_+$  telle que  $\mu = c\lambda_n$ .

*Démonstration.* C'est l'objet de l'exercice 14 du TD II. □

### 5.2.2 Changement de variable affine

Soit  $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$  une application linéaire. On note  $\det(L)$  le déterminant de l'endomorphisme  $L$ . Ce nombre réel peut se calculer en pratique en considérant une matrice de  $L$  écrite dans n'importe quelles bases de  $\mathbb{R}^n$ . À ce stade du cours, il peut être bon de regarder <https://www.youtube.com/watch?v=kYB8IZa5AuE> et <https://www.youtube.com/watch?v=Ip3X9L0h2dk> et de revoir les propriétés élémentaires décrites <https://en.wikipedia.org/wiki/Determinant> ou dans tout autre cours d'algèbre linéaire.

**Proposition 5.2.3.** Soit  $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$  une application linéaire. Si  $A \in \mathcal{B}(\mathbb{R}^n)$ , alors  $\lambda_n(L(A)) = |\det(L)| \lambda_n(A)$ .

*Démonstration.* Si  $L$  n'est pas injective, on a  $\det(L) = 0$ . De plus  $\lambda_n(L(A)) = 0$  pour tout borélien  $A$  car l'ensemble  $L(A)$  est alors contenu dans un sous-espace vectoriel strict de  $\mathbb{R}^n$  qui est  $\lambda_n$ -négligeable. La proposition est donc vérifiée dans ce cas.

On suppose maintenant que  $L$  est inversible. Pour tout borélien  $A$  de  $\mathbb{R}^n$ , on pose :

$$\mu(A) = \lambda_n(L(A)).$$

$\mu$  est une mesure positive et invariante par translation (grâce à la linéarité de  $L$ ). De plus, si  $A$  est borné, alors  $L(A)$  aussi car  $L$  est un endomorphisme de  $\mathbb{R}^n$ . La Proposition 5.2.2 permet de conclure que  $\mu = c\lambda_n$ . Il s'agit donc de voir que  $c = \det(L)$ .

- Soit  $L_1, L_2 : \mathbb{R}^n \rightarrow \mathbb{R}^n$  deux applications linéaires. Notons que la constante  $c$  (qui dépend de l'application linéaire) vérifie :

$$c(L_1 \circ L_2) = c(L_1)c(L_2).$$

car

$$c(L_1 \circ L_2)\lambda_n(A) = \lambda_n(L_1 \circ L_2(A)) = \lambda_n(L_1(L_2(A))) = c(L_1)\lambda_n(L_2(A)) \quad (5.7)$$

$$= c(L_1)c(L_2)\lambda_n(A). \quad (5.8)$$

Elle vérifie aussi  $c(\text{Id}_{\mathbb{R}^n}) = 1$ .

- Si  $L = D$  est un opérateur diagonal, c'est-à-dire que sa matrice, dans la base canonique, est de la forme :

$$\begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix},$$

alors  $D([0, 1]^n)$  est un pavé dont les côtés sont de longueur  $|d_1|, \dots, |d_n|$ . Donc :

$$c(D) = |d_1| |d_2| \cdots |d_n| = |d_1 d_2 \cdots d_n| = |\det(D)|.$$

- Si  $L = U$  est une transformation orthogonale, on a  $U^t U = U U^t = \text{Id}_{\mathbb{R}^n}$  (et donc  $U^t = U^{-1}$ ).  $U$  est alors une isométrie linéaire, et on a  $U(\text{B}(0, 1)) = \text{B}(0, 1)$ . Donc :

$$\lambda_n(\text{B}(0, 1)) = \lambda_n(U(\text{B}(0, 1))) = c(U)\lambda_n(\text{B}(0, 1)),$$

ce qui entraîne (car  $\lambda_n(\text{B}(0, 1)) > 0$ ) que  $c(U) = 1$  (rappelons que tout ouvert non vide  $\Omega$  de  $\mathbb{R}^n$  contient un pavé ouvert non vide et donc  $\lambda_n(\Omega) > 0$ ).

- Le théorème s'en déduit car toute application linéaire  $L$  peut se décomposer en :

$$L = U_1 D U_2,$$

où  $D$  est diagonal et  $U_1, U_2$  sont orthogonales.

□

**Remarque 5.2.1.** 1. Si  $L$  est linéaire non inversible, alors  $\det(L) = 0$  et  $L(A)$  est, pour tout  $A \in \mathcal{B}(\mathbb{R}^n)$ , contenu dans un hyperplan. En particulier,  $\lambda_n(L(\mathbb{R}^n)) = 0 = |\det(L)| \lambda_n(\mathbb{R}^n)$  ce qui montre que les sous-espaces vectoriels stricts de  $\mathbb{R}^n$  ont une mesure (de Lebesgue  $n$ -dimensionnelle) nulle.

2. Si  $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$  est une transformation affine, alors il existe  $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$  linéaire et  $v \in \mathbb{R}^n$  tels que  $T(x) = L(x) + v$  pour tout  $x \in \mathbb{R}^n$ . De plus, pour tout  $A \in \mathcal{B}(\mathbb{R}^n)$ , on a  $\lambda_n(T(A)) = |\det(L)| \lambda_n(A)$ .

3. Si  $H : \mathbb{R}^n \rightarrow \mathbb{R}^n$  est l'homothétie de centre  $v$  et de rapport  $r$ , alors

$$H(x) = v + r(x - v) = rx + (1 - r)v,$$

donc  $\lambda_n(H(A)) = |r|^n \lambda_n(A)$  pour tout  $A \in \mathcal{B}(\mathbb{R}^n)$ .

### 5.2.3 Changement de variable

On souhaite étendre les résultats précédents à des applications qui ne soient pas linéaires ou affines. On considère un difféomorphisme de classe  $C^1$ ,  $\varphi : U \rightarrow V$  où  $U$  et  $V$  sont des ouverts de  $\mathbb{R}^n$ . Pour tout  $x \in U$ , le jacobien de  $\varphi$  en  $x$  est le déterminant de la matrice jacobienne de  $\varphi$  au point  $x$ . Si on note  $J_\varphi(x)$  cette matrice jacobienne, l'application  $\det(J_\varphi) : U \rightarrow \mathbb{R}$  ainsi définie est continue (et donc mesurable).

**Proposition 5.2.4.** Soient  $U, V$  deux ouverts de  $\mathbb{R}^n$ , et  $\varphi : U \rightarrow V$  un difféomorphisme de classe  $C^1$ . Pour tout  $A \in \mathcal{B}(U)$  on a  $\varphi(A) \in \mathcal{B}(V)$  et

$$\lambda_n(\varphi(A)) = \int_A |\det(J_\varphi(x))| d\lambda_n(x).$$

*Démonstration.* Admis dans un premier temps. On pourra consulter <https://www.geogebra.org/m/V9PZC5SU> pour développer son intuition à propos de ce résultat. □

De cette proposition on déduit un théorème de changement de variable dans  $\mathbb{R}^n$ .

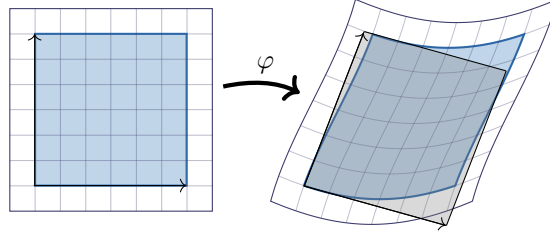


FIGURE 5.2 – Une fonction régulière  $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  envoie un élément de surface carré (en bleu) sur un élément de surface courbe (en bleu). La meilleure approximation linéaire de  $\varphi$  en  $x_0 \in \mathbb{R}^2$  est donnée par la différentielle (dont la matrice est la Jacobienne  $J_\varphi(x_0)$ ) qui envoie le carré (en bleu) sur le rectangle (en gris). Le déterminant  $\det(J_\varphi(x_0))$  donne alors le rapport entre la surface du carré bleu et la surface du rectangle gris. [https://commons.wikimedia.org/wiki/File:Jacobian\\_determinant\\_and\\_distortion.svg](https://commons.wikimedia.org/wiki/File:Jacobian_determinant_and_distortion.svg)

**Théorème 5.2.1 Changement de variable.** Soient  $U, V$  deux ouverts de  $\mathbb{R}^n$ , et  $\varphi : U \rightarrow V$  un difféomorphisme de classe  $C^1$ .

Si  $f : V \rightarrow \mathbb{R}_+$  est une fonction mesurable, alors  $(f \circ \varphi) |\det(J_\varphi)| : U \rightarrow \mathbb{R}_+$  est mesurable et

$$\int_V f(y) d\lambda_n(y) = \int_U f(\varphi(x)) |\det(J_\varphi(x))| d\lambda_n(x).$$

Si  $f : V \rightarrow \mathbb{K}$  est une fonction intégrable, alors  $(f \circ \varphi) |\det(J_\varphi)| : U \rightarrow \mathbb{K}$  est intégrable et

$$\int_V f(y) d\lambda_n(y) = \int_U f(\varphi(x)) |\det(J_\varphi(x))| d\lambda_n(x).$$

*Démonstration.* La preuve suit un plan similaire à celle du Théorème de Tonelli : on démontre la propriété pour les fonctions indicatrices, puis par linéarité pour les fonction étagées et enfin par densité pour les fonctions mesurables.

Étape 1. Pour  $f = \mathbb{1}_B$  où  $B \subset V$  est un borélien de  $\mathbb{R}^n$ , on a  $f \circ \varphi = \mathbb{1}_B \circ \varphi = \mathbb{1}_{\varphi^{-1}(B)} = \mathbb{1}_A$  avec  $A \subset U$  un borélien lui aussi (car  $\varphi$  est de classe  $C^1$ ). La formule de changement de variable a été établie dans ce cas à la Proposition 5.2.4.

Étape 2. Soit maintenant  $f \in \mathcal{E}_+(V)$  une fonction étagée. Il existe alors  $\alpha_1, \dots, \alpha_p \in \mathbb{R}_+$  et une partition mesurable  $(B_1, \dots, B_p)$  de  $V$  tels que  $f(y) = \sum_{i=1}^p \alpha_i \mathbb{1}_{B_i}$ . Comme  $\varphi$  est un difféomorphisme, la classe  $(A_1, \dots, A_p) = (\varphi^{-1}(B_1), \dots, \varphi^{-1}(B_p))$  est une partition mesurable de  $U$  et on a

$$f(y) = \sum_{i=1}^p \alpha_i \mathbb{1}_{\varphi(A_i)}(y), \quad y \in V \quad \text{et} \quad f \circ \varphi(x) = \sum_{i=1}^p \alpha_i \mathbb{1}_{A_i}(x), \quad x \in U.$$

On a alors

$$\begin{aligned} \int_V f d\lambda_n(y) &= \int_V \left( \sum_{i=1}^p \alpha_i \mathbb{1}_{\varphi(A_i)}(y) \right) d\lambda_n(y) \\ &= \sum_{i=1}^p \alpha_i \lambda_n(\varphi(A_i)) \\ &= \sum_{i=1}^p \alpha_i \int_U \mathbb{1}_{A_i}(x) |\det(J_\varphi(x))| d\lambda_n(x) \\ &= \int_U \left( \sum_{i=1}^p \alpha_i \mathbb{1}_{A_i}(x) \right) |\det(J_\varphi(x))| d\lambda_n(x) \\ &= \int_U f \circ \varphi(x) |\det(J_\varphi(x))| d\lambda_n(x). \end{aligned}$$

Cela montre la formule pour les fonctions étagées.

Étape 3. Maintenant, si  $f : V \rightarrow \overline{\mathbb{R}}_+$  est une application mesurable, on peut trouver une suite croissante  $(e_k)_{k \in \mathbb{N}}$  de fonctions de  $\mathcal{E}_+(V)$  qui converge vers  $f$ . On a alors, d'après le théorème de convergence monotone (TCM) appliqué deux fois,

$$\begin{aligned} \int_V f(y) d\lambda_n(y) &= \int_V \lim_{k \rightarrow +\infty} e_k(y) d\lambda_n(y) \\ &\stackrel{\text{TCM}}{=} \lim_{k \rightarrow +\infty} \int_V e_k(y) d\lambda_n(y) \\ &= \lim_{k \rightarrow +\infty} \int_U e_k \circ \varphi(x) |\det(J_\varphi(x))| d\lambda_n(x) \\ &\stackrel{\text{TCM}}{=} \int_U \lim_{k \rightarrow +\infty} e_k \circ \varphi(x) |\det(J_\varphi(x))| d\lambda_n(x) \\ &= \int_U f \circ \varphi(x) |\det(J_\varphi(x))| d\lambda_n(x). \end{aligned}$$

On a donc montré la formule pour les fonctions mesurables positives.

Étape 4. Enfin, si  $f : V \rightarrow \mathbb{K}$  est une fonction mesurable. Dans le cas où  $\mathbb{K} = \mathbb{R}$ , on utilise la décomposition  $f = f_+ - f_-$  pour démontrer la formule de changement de variables. Si  $\mathbb{K} = \mathbb{C}$  on applique le résultat précédent à  $\operatorname{Re}(f)$  et  $\operatorname{Im}(f)$  pour conclure la démonstration.  $\square$

**Remarque 5.2.2.** • La présence d'une valeur absolue dans cette formule de changement de variables en dimension quelconque  $n \geq 1$  provient du fait que les mesures de Lebesgue  $d\lambda_n(x) = d\lambda_1(x_1) \otimes \cdots \otimes d\lambda_1(x_n)$  sur le  $\mathbb{R}^n$ -source et  $d\lambda_n(y) = d\lambda_1(y_1) \otimes \cdots \otimes d\lambda_1(y_n)$  sur le  $\mathbb{R}^n$ -image ont été définies comme positives (à la physicienne), contrairement au  $dx$  riemannien sur  $\mathbb{R}$  en dimension 1, lequel possède un signe ; d'ailleurs, même en dimension  $n = 1$ , la théorie de Lebesgue requiert une valeur absolue :

$$\int_{\varphi^{-1}([a,b])} f d\lambda_1 = \int_{[a,b]} f \circ \varphi |\varphi'| d\lambda_1(x)$$

puisque si la fonction  $\varphi \in C^1(\mathbb{R}^n)$  du Théorème de changement de variable a une dérivée négative  $\varphi' < 0$ , donc décroît, si  $I = [a, b]$  alors  $J = [\varphi(b), \varphi(a)]$  – noter l'inversion. Grâce à la théorie de Riemann, on réalise alors pourquoi, en théorie de Lebesgue, on est forcé d'insérer une valeur absolue :

$$\begin{aligned} \int_{\varphi^{-1}([a,b])} f d\lambda_1 &= \int_{\varphi(b)}^{\varphi(a)} f(y) dy \\ &= - \int_{\varphi(a)}^{\varphi(b)} f(y) dy \\ &= \int_a^b f \circ \varphi(x) (-\varphi'(x)) dx \\ &= \int_{[a,b]} f \circ \varphi |\varphi'| d\lambda_1. \end{aligned}$$

Pour pouvoir parler de mesures de Lebesgue orientées, souvent notée  $dx_1 \wedge \cdots \wedge dx_d$  et qui généraliseraient le  $dx$  riemannien orienté sur  $\mathbb{R}$  il faudrait entrer dans la théorie des formes différentielles extérieures, ce que nous ne pourrions pas faire dans ce cours. On pourra tout de même voir [https://www.youtube.com/watch?v=xRf9-hdxB0w&list=PL9\\_jI1bdZmz0hIrNCMQW1YmZysAiIYSSS&index=6](https://www.youtube.com/watch?v=xRf9-hdxB0w&list=PL9_jI1bdZmz0hIrNCMQW1YmZysAiIYSSS&index=6).

### Coordonnées polaires sur $\mathbb{R}^2$

Le changement de variable en coordonnées polaires envoie des disques sur des rectangles. Voir absolument <https://www.geogebra.org/m/dN7Bjw96>. On note  $U = \mathbb{R}_+^* \times ]-\pi, \pi[ \subset \mathbb{R}^2$  et on

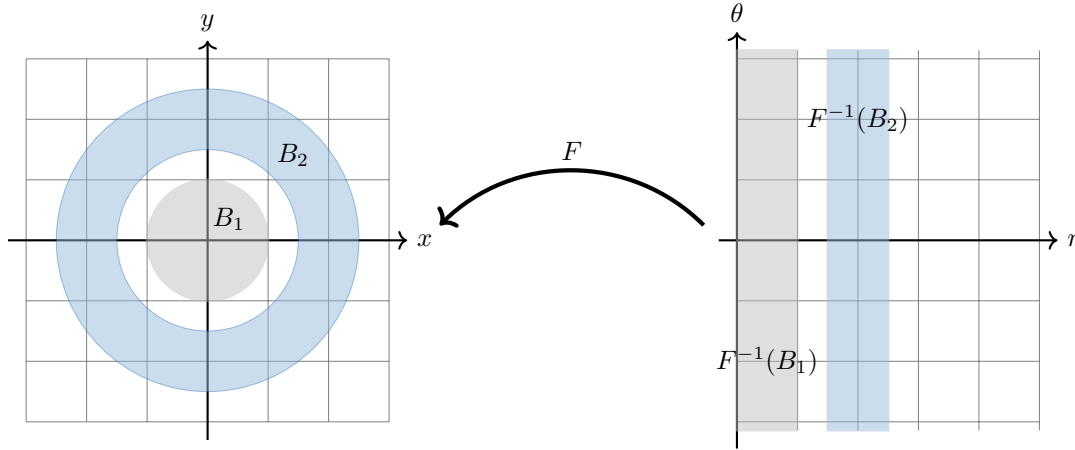


FIGURE 5.3 – L'aire du rectangle  $F^{-1}(B_1)$  et du rectangle  $F^{-1}(B_2)$  dans le plan  $(r, \theta)$  sont égales. Pour retrouver les aires des couronnes  $B_1$  et  $B_2$  correspondantes dans le plan  $(x, y)$  il faut multiplier par un facteur correctif  $|\det(J_F(r, \theta))| = r$ . En effet,  $\text{Aire}(B_1) = \int_{B_1} d\lambda_2 = \int_{F^{-1}(B_1)} r d\lambda_2(r, \theta) = \int_0^1 r dr \int_{-\pi}^{\pi} d\theta = \pi$  et  $\text{Aire}(B_2) = \int_{B_2} d\lambda_2 = \int_{F^{-1}(B_2)} r d\lambda_2(r, \theta) = \int_{\frac{5}{2}}^{\frac{7}{2}} r dr \int_{-\pi}^{\pi} d\theta = 4\pi$ .

considère

$$\begin{aligned} F : U &\longrightarrow \mathbb{R}^2 \\ (r, \theta) &\longmapsto (r \cos \theta, r \sin \theta) \end{aligned}$$

On a alors  $F(U) = \mathbb{R}^2 \setminus (\mathbb{R}_- \times \{0\})$  et  $F : U \rightarrow F(U)$  est un difféomorphisme. Pour toute fonction  $f : \mathbb{R}^2 \rightarrow \mathbb{K}$  intégrable ou positive on a

$$\int_{\mathbb{R}^2} f(x, y) d\lambda_2(x, y) = \int_U f(r \cos \theta, r \sin \theta) r d\lambda_2(r, \theta) = \int_{-\pi}^{\pi} \int_0^{+\infty} f(r \cos \theta, r \sin \theta) r dr d\theta$$

### Coordonnées sphériques sur $\mathbb{R}^3$

De façon analogue, le changement de coordonnées sphériques envoie des sphères sur des pavés. Attention ici aux différentes conventions : latitude, colatitude, longitude, *etc.* On note  $V = \mathbb{R}_+^* \times ]-\pi, \pi[ \times ]0, \pi[ \subset \mathbb{R}^3$  et on considère

$$\begin{aligned} G : V &\longrightarrow \mathbb{R}^3 \\ (r, \theta, \varphi) &\longmapsto (r \cos \theta \sin \varphi, r \sin \theta \sin \varphi, r \cos \varphi). \end{aligned}$$

On a alors  $G(V) = \mathbb{R}^3 \setminus (\mathbb{R}_- \times \{0\} \times \mathbb{R})$  et  $G : V \rightarrow G(V)$  est un difféomorphisme. Pour toute fonction  $f : \mathbb{R}^3 \rightarrow \mathbb{K}$  intégrable ou positive on a

$$\begin{aligned} \int_{\mathbb{R}^3} f(x, y, z) d\lambda_3(x, y, z) &= \int_V f(G(r, \theta, \varphi)) r^2 \sin \varphi d\lambda_3(r, \theta, \varphi) \\ &= \int_0^{\pi} \int_{-\pi}^{\pi} \int_0^{+\infty} f(G(r, \theta, \varphi)) r^2 \sin \varphi dr d\theta d\varphi \end{aligned}$$

Voir la Figure 5.4 pour une illustration.

## 5.3 Produit de convolution et théorèmes de densité

La propriété de complétude des espaces  $L^p(\mathbb{R}^n)$  est cruciale dans beaucoup de situations, dès qu'on utilise la norme  $\|\cdot\|_p$ . Mais on a plus l'habitude de travailler avec des fonctions continues, voire différentiables, voire de classe  $C^\infty$ .

Les théorèmes de densité montrent que toute fonction  $f \in L^p(\mathbb{R}^n)$  est limite d'une suite de fonctions continues, ou d'une suite de fonctions  $C^\infty$ .



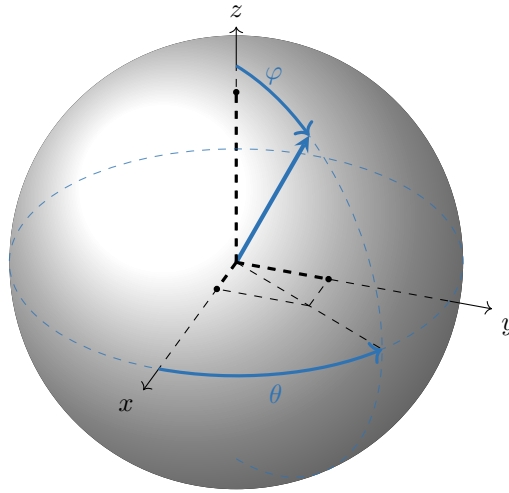


FIGURE 5.4 – Coordonnées sphériques.

**Définition 5.3.1.** Un sous-espace vectoriel  $E \subset L^p(\mathbb{R}^n)$  est dense dans  $L^p(\mathbb{R}^n)$  si, pour tout  $f \in L^p(\mathbb{R}^n)$  il existe une suite  $(u_n)_{n \in \mathbb{N}} \in E^{\mathbb{N}}$  telle que  $\lim_{n \rightarrow +\infty} \|f - u_n\|_p = 0$ .

**Remarque 5.3.1.** Cette propriété peut aussi s'écrire

$$\forall f \in L^p(\mathbb{R}^n), \forall \varepsilon > 0, \exists u \in E, \|f - u\|_p < \varepsilon.$$

Cette notion de densité est analogue à la densité de  $\mathbb{Q}$  dans  $\mathbb{R}$ .

**Définition 5.3.2.** Une fonction  $f : \mathbb{R}^n \rightarrow \mathbb{C}$  est **à support compact** s'il existe un compact  $K \subset \mathbb{R}^n$  tel que

$$\forall x \in \mathbb{R}^n \setminus K, f(x) = 0.$$

On note  $C_c^0(\mathbb{R}^n)$  (resp.  $C_c^{0,1}(\mathbb{R}^n)$ , resp.  $C_c^\infty(\mathbb{R}^n)$ ) l'espace vectoriel des fonctions continues (resp. lipschitziennes, resp. de classe  $C^\infty$ ) à support compact.

**Remarque 5.3.2.** Les parties compactes de  $\mathbb{R}^n$  étant les parties fermées et bornées, une fonction  $f : \mathbb{R}^n \rightarrow \mathbb{C}$  est à support compact si et seulement si il existe  $R > 0$  tel que pour tout  $x \in \mathbb{R}^n \setminus B(0, R)$  on ait  $f(x) = 0$ . Dans la suite, on utilisera souvent cette caractérisation.

La somme de deux fonctions à support compact est encore à support compact, de même que le produit d'une fonction à support compact par un scalaire. Les espaces  $C_c^0(\mathbb{R}^n)$ ,  $C_c^{0,1}(\mathbb{R}^n)$  et  $C_c^\infty(\mathbb{R}^n)$  sont donc bien des espaces vectoriels. Par ailleurs, les fonctions continues sur un compact étant bornées, pour toute fonction  $f \in C_c^0(\mathbb{R}^n)$  il existe  $R > 0$  tel que

$$\int_{\mathbb{R}^n} |f|^p d\lambda_n = \int_{B(0, R)} |f|^p d\lambda_n \leq M^p \lambda_n(B(0, R)) < +\infty$$

où  $M = \sup_{B(0, R)}(f)$ . Donc  $f \in L^p(\mathbb{R}^n)$  et on a les inclusions suivantes

$$C_c^\infty(\mathbb{R}^n) \subset C_c^{0,1}(\mathbb{R}^n) \subset C_c^0(\mathbb{R}^n) \subset L^p(\mathbb{R}^n).$$

### 5.3.1 Densité des fonctions continues à support compact

**Théorème 5.3.1 Régularité de la mesure de Lebesgue.** On se place dans  $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \lambda_n)$ . Les compacts de  $\mathbb{R}^n$  sont des boréliens de mesure finis, c'est-à-dire que  $\lambda_n(K) < +\infty$  pour tout compact  $K$  de  $\mathbb{R}^n$ . Alors, pour tout  $A \in \mathcal{B}(\mathbb{R}^n)$  et tout  $\varepsilon > 0$ , il existe un ouvert  $O$  et un fermé

$F$  tel que  $F \subset A \subset O$  et  $\lambda_n(O \setminus F) \leq \varepsilon$ .

*Démonstration.* La preuve suit la stratégie suivante : on appelle  $\mathcal{T}$  l'ensemble des  $A \in \mathcal{B}(\mathbb{R}^n)$  tel que pour tout  $\varepsilon > 0$ , il existe  $O$  ouvert et  $F$  fermé vérifiant  $F \subset A \subset O$  et  $\lambda(O \setminus F) \leq \varepsilon$ . On montre que  $\mathcal{T}$  est une tribu contenant l'ensemble  $C$  des pavés ouverts de  $\mathbb{R}^n$ . Comme  $C$  engendre  $\mathcal{B}(\mathbb{R}^n)$ , ceci donne  $\mathcal{T} = \mathcal{B}(\mathbb{R}^n)$ .

Elle est détaillée dans [3] Théorème 2.43 (pour le cas  $n = 1$  et avec des mesures finies sur tout compact).  $\square$

**Théorème 5.3.2.** L'espace  $C_c^{0,1}(\mathbb{R}^n)$  des fonctions lipschitziennes à support compact dans  $\mathbb{R}^n$  est dense dans  $L^p(\mathbb{R}^n, \lambda_n)$ , pour tout  $p \in [1, +\infty[$ .

*Démonstration.* On rédige la preuve dans le cas  $n = 1$ . Le cas général étant très similaire. On a déjà vu que  $C_c^{0,1}(\mathbb{R}) \subset L^1(\mathbb{R})$ . En toute rigueur, on a plutôt  $C_c^{0,1}(\mathbb{R}) \subset \mathcal{L}^1(\mathbb{R}) = \mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda_1)$ . L'objectif est donc de montrer que pour tout  $f \in \mathcal{L}^1(\mathbb{R})$  et tout  $\varepsilon > 0$ , il existe  $\varphi \in C_c^{0,1}(\mathbb{R})$  telle que  $\|f - \varphi\|_1 \leq \varepsilon$ . On va raisonner une nouvelle fois en plusieurs étapes : fonctions caractéristiques, fonctions étagées, fonctions mesurables positives et enfin dans  $\mathcal{L}^1(\mathbb{R})$ .

**Étape 1.** On suppose ici que  $f = \mathbf{1}_A$  avec  $A \in \mathcal{B}(\mathbb{R})$  et  $\lambda_1(A) < +\infty$ . Soit  $\varepsilon > 0$ . Comme  $\lambda_1$  est une mesure régulière (Théorème 5.3.1), il existe un ouvert  $O$  et un fermé  $F$  tels que  $F \subset A \subset O$  et  $\lambda_1(O \setminus F) \leq \varepsilon$ . Pour  $k \in \mathbb{N}^*$ , on pose  $F_k = F \cap [-k, k]$ , de sorte que  $F_k$  est compact pour tout  $k \in \mathbb{N}$ , et  $F = \bigcup_{k \in \mathbb{N}^*} F_k$ . La continuité croissante de  $\lambda_1$  donne alors  $\lambda_1(F_k) \rightarrow \lambda_1(F)$ , quand  $k \rightarrow +\infty$ . Comme  $\lambda_1(F) \leq \lambda_1(A) < +\infty$ , on a aussi  $\lambda_1(F \setminus F_k) = \lambda_1(F) - \lambda_1(F_k) \rightarrow 0$  quand  $k \rightarrow +\infty$ . Il existe donc  $k_0 \in \mathbb{N}$  tel que  $\lambda_1(F \setminus F_{k_0}) \leq \varepsilon$ . On pose  $K = F_{k_0}$  et on obtient donc  $K \subset F \subset A \subset O$ , ce qui donne

$$\lambda_1(O \setminus K) \leq \lambda_1(O \setminus F) + \lambda_1(F \setminus K) \leq 2\varepsilon.$$

On a donc trouvé un compact  $K$  et un ouvert  $O$  tels que  $K \subset A \subset O$  et  $\lambda_1(O \setminus K) \leq 2\varepsilon$ . Ceci va nous permettre de construire  $\varphi \in C_c^{0,1}(\mathbb{R})$  telle que  $\|f - \varphi\|_1 \leq 2\varepsilon$ . On pose

$$d = d(K, O^c) = \inf \{d(x, y), x \in K, y \in O^c\}.$$

Montrons que  $d > 0$  : par caractérisation de l'infimum, il existe des suites  $(x_j)_{j \in \mathbb{N}} \subset K$  et  $(y_j)_{j \in \mathbb{N}} \subset O^c$  telles que

$$d(x_j, y_j) = |x_j - y_j| \rightarrow d \text{ quand } j \rightarrow +\infty.$$

Par compacité de  $K$ , on peut supposer (après extraction éventuelle d'une sous-suite) que  $x_j \rightarrow x$ , quand  $j \rightarrow +\infty$ . Si  $d = 0$ , on a alors aussi  $y_j \rightarrow x$  quand  $j \rightarrow +\infty$  et donc  $x \in O^c \cap K$  (car  $K$  et  $O^c$  sont fermés), ce qui est impossible puisque  $K \subset A \subset O$  et donc  $O^c \cap K = \emptyset$ . On a donc bien montré  $d > 0$ .

On définit maintenant la fonction<sup>2</sup>  $\varphi : \mathbb{R} \rightarrow \mathbb{R}_+$  par

$$\varphi(x) = \frac{1}{d}(d - d(x, K))_+, \quad \text{avec } d(x, K) = \inf \{d(x, y), y \in K\}.$$

La fonction  $\varphi$  est continue car  $x \mapsto d(x, K)$  est continue (cette fonction est même lipschitzienne, on peut montrer que  $|d(x, K) - d(y, K)| \leq |x - y|$ ). Elle est à support compact car il existe  $A > 0$  tel que  $K \subset [-A, A]$  et on remarque alors que  $\varphi = 0$  sur  $[-A - d, A + d]^c$ . On a donc  $\varphi \in C_c^{0,1}(\mathbb{R})$ . Enfin, on remarque que  $\varphi = 1$  sur  $K$ ,  $\varphi = 0$  sur  $O^c$  et  $0 \leq \varphi \leq 1$ . On en déduit que  $f - \varphi = 0$  sur  $K \cup O^c$  et  $0 \leq |f - \varphi| \leq 1$ , ce qui donne

$$\|f - \varphi\|_1 \leq \lambda_1(O \setminus K) \leq 2\varepsilon,$$

qui termine la principale étape.

2. Ce résultat est connue sous le nom de Lemme d'Urysohn

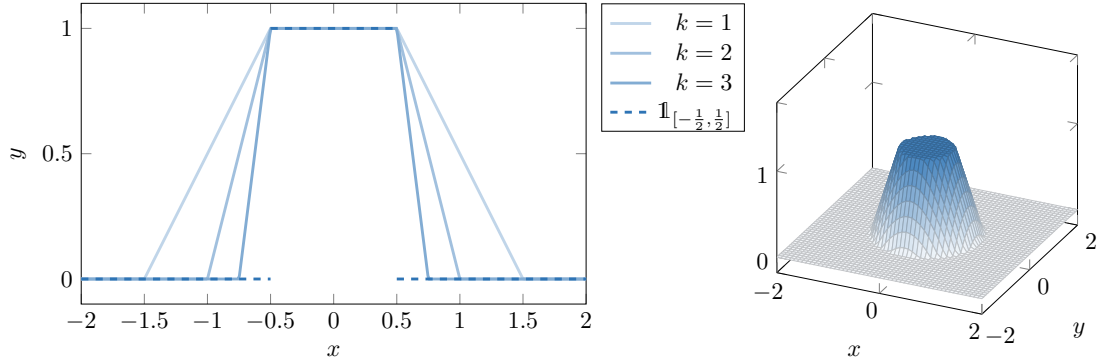


FIGURE 5.5 – Approximation d'une indicatrice de  $B(0, \frac{1}{2})$  dans  $\mathbb{R}^n$  par une suite de fonction Lipschitziennes à support compact. À gauche : pour  $n = 1$  et différentes valeurs du paramètre  $k \in \mathbb{N}^*$ . À droite : pour  $n = 2$  et  $k = 2$ .

Étape 2. On suppose ici que  $f \in \mathcal{E}_+(\mathbb{R}) \cap \mathcal{L}^1(\mathbb{R})$ . Il existe donc  $a_1, \dots, a_p > 0$  et une partition  $(A_1, \dots, A_p)$  mesurable tels que  $f = \sum_{i=1}^p a_i \mathbf{1}_{A_i}$ . Comme  $f \in \mathcal{L}^1(\mathbb{R})$ , on a aussi  $\lambda_1(A_i) < +\infty$  pour tout  $i = 1, \dots, p$ . Soit  $\varepsilon > 0$ , l'Étape 1. donne, pour tout  $i = 1, \dots, p$ , l'existence de  $\varphi_i \in C_c^{0,1}(\mathbb{R})$  telle que  $\|\mathbf{1}_{A_i} - \varphi_i\|_1 \leq \varepsilon$ . On pose

$$\varphi = \sum_{i=1}^p a_i \varphi_i \in C_c^{0,1}(\mathbb{R})$$

et on obtient

$$\|f - \varphi\|_1 \leq \left( \sum_{i=1}^p a_i \right) \varepsilon$$

(ce qui est bien arbitrairement petit).

Étape 3. On suppose ici que  $f$  est mesurable, positive et dans  $\mathcal{L}^1(\mathbb{R})$ . Soit  $\varepsilon > 0$ . D'après la définition de l'intégrale des fonction mesurable positive (Proposition-Définition 3.2.1), il existe  $g \in \mathcal{E}_+(\mathbb{R})$  telle que  $g \leq f$  et

$$\int_{\mathbb{R}} f d\lambda_1 - \varepsilon \leq \int_{\mathbb{R}} g d\lambda_1 \leq \int_{\mathbb{R}} f d\lambda_1,$$

de sorte que

$$\|g - f\|_1 = \int_{\mathbb{R}} (f - g) d\lambda_1 \leq \varepsilon,$$

L'Étape 2. donne alors l'existence de  $\varphi \in C_c^{0,1}(\mathbb{R})$  telle que  $\|g - \varphi\|_1 \leq \varepsilon$ . D'où l'on déduit

$$\|f - \varphi\|_1 \leq 2\varepsilon,$$

ce qui termine l'Étape 3..

Étape 4. On suppose enfin que  $f \in \mathcal{L}^1(\mathbb{R})$  et on fixe  $\varepsilon > 0$ . Comme  $f_{\pm}$  sont mesurables, positives et dans  $\mathcal{L}^1(\mathbb{R})$ , l'Étape 3. implique qu'il existe  $\varphi_1, \varphi_2 \in C_c^{0,1}(\mathbb{R})$  telle que

$$\|f_+ - \varphi_1\|_1 \leq \varepsilon \quad \text{et} \quad \|f_- - \varphi_2\|_1 \leq \varepsilon.$$

On pose alors  $\varphi = \varphi_1 - \varphi_2$ . On a  $\varphi \in C_c^{0,1}(\mathbb{R})$  et  $\|f - \varphi\|_1 \leq 2\varepsilon$ , ce qui prouve bien la densité de  $C_c^{0,1}(\mathbb{R})$  dans  $\mathcal{L}^1(\mathbb{R})$ . □

**Remarque 5.3.3.** • Ce résultat est faux si  $p = +\infty$ . Un contre-exemple est donné par la fonction de Heavyside  $\mathbf{1}_{\mathbb{R}_+}$ .

- Les fonctions lipschitziennes étant continues, l'espace des fonctions continues à support compact est aussi dense dans  $L^p(\mathbb{R}^n)$ .

- Le résultat de densité que nous venons de voir n'est pas limité à la mesure de Lebesgue. Il est vrai pour toute mesure sur  $\mathcal{B}(\mathbb{R}^n)$ , finie sur les compacts.

Ce résultat permet, dans certaines situations, d'étendre à  $L^p(\mathbb{R}^n)$  des propriétés vérifiées par les fonctions continues ou Lipschitziennes. Un exemple typique (et très utile pour la suite) est la propriété suivante qui concerne la translatée d'une fonction  $f : \mathbb{R}^n \rightarrow \mathbb{C}$ .

Pour une fonction  $f : \mathbb{R}^n \rightarrow \mathbb{K}$  et  $y \in \mathbb{R}^n$ , la **translatée** de  $f$  par  $y$  est la fonction  $\tau_y f : \mathbb{R}^n \rightarrow \mathbb{K}$  définie par

$$(\tau_y f)(x) = f(x - y), \text{ pour tout } x \in \mathbb{R}^n.$$

**Proposition 5.3.1 Continuité en moyenne.** Soit  $p \in [1, +\infty[$ , pour tout  $f \in L^p(\mathbb{R}^n)$  on a  $\lim_{y \rightarrow 0} \|\tau_y f - f\|_p = 0$ .

*Démonstration.* Soient  $f \in L^p(\mathbb{R}^n)$  et  $\varepsilon > 0$ . Par densité des fonctions Lipschitziennes à support compact, il existe  $u \in C_c^{0,1}(\mathbb{R}^n)$  telle que  $\|f - u\|_p < \frac{\varepsilon}{3}$ . Soit  $y \in \mathbb{R}^n$ . On a  $\tau_y f - f = (\tau_y f - \tau_y u) + (\tau_y u - u) + (u - f)$ , donc, par inégalité triangulaire

$$\|\tau_y f - f\|_p \leq \|\tau_y f - \tau_y u\|_p + \|\tau_y u - u\|_p + \|u - f\|_p.$$

Montrons que, pour  $\|y\|$  assez petit, chacun de ces trois termes est inférieur à  $\frac{\varepsilon}{3}$ .

- $\|u - f\|_p < \frac{\varepsilon}{3}$  par choix de  $u$ .
- $\|\tau_y f - \tau_y u\|_p^p = \int_{\mathbb{R}^n} |f(x - y) - u(x - y)|^p d\lambda_n(x)$ .

Le changement de variable  $z = x - y$  donne  $x = z + y = F(z)$  et  $J_F(z) = \text{Id}_n$ . Donc  $|\det(J_F(z))| = 1$  et on obtient

$$\|\tau_y f - \tau_y u\|_p^p = \int_{\mathbb{R}^n} |f(z) - u(z)|^p 1 d\lambda(z) = \|u - f\|_p^p.$$

Donc  $\|\tau_y f - \tau_y u\|_p = \|u - f\|_p < \frac{\varepsilon}{3}$ .

- $\|\tau_y u - u\|_p^p = \int_{\mathbb{R}^n} |u(x - y) - u(x)|^p d\lambda_n(x)$ . Remarquons que
  - comme  $u$  est à support compact, il existe  $R > 0$  tel que  $u(x) = 0$  pour tout  $x \in \mathbb{R}^n \setminus \mathbf{B}(0, R)$ . Ainsi, pour tout  $x \in \mathbb{R}^n$  vérifiant  $\|x\| > R + \|y\|$ , on a  $\|x\| > R$  et  $\|x - y\| \geq \|x\| - \|y\| > R$  par inégalité triangulaire, d'où  $u(x) = u(x - y) = 0$ .
  - comme  $u$  est Lipschitzienne, il existe  $C > 0$  telle que  $|u(x - y) - u(x)| \leq C \|x - y - x\| = C\|y\|$  pour tout  $x, y \in \mathbb{R}^n$ .

On a donc

$$\|\tau_y u - u\|_p^p = \int_{\mathbf{B}(0, R + \|y\|)} |u(x - y) - u(x)|^p dx \leq C^p \|y\|^p \lambda_n(\mathbf{B}(0, R + \|y\|)).$$

Soit  $\alpha = \min\left(1, \frac{\varepsilon}{3C\lambda_n(\mathbf{B}(0, R + 1))^{\frac{1}{p}}}\right)$ . Si  $\|y\| < \alpha$ , on a  $\lambda_n(\mathbf{B}(0, R + \|y\|)) \leq \lambda_n(\mathbf{B}(0, R + 1))$  et

$$\|\tau_y u - u\|_p \leq C\|y\| \lambda_n(\mathbf{B}(0, R + 1))^{\frac{1}{p}} \leq C\alpha \lambda_n(\mathbf{B}(0, R + 1))^{\frac{1}{p}} \leq \frac{\varepsilon}{3}.$$

Pour tout  $y \in \mathbb{R}^n$ , si  $\|y\| < \alpha$  on a donc

$$\|\tau_y f - f\|_p \leq \|\tau_y f - \tau_y u\|_p + \|\tau_y u - u\|_p + \|u - f\|_p < \varepsilon,$$

ce qui montre que  $\lim_{y \rightarrow 0} \|\tau_y f - f\|_p = 0$ . □

**Remarque 5.3.4.** Pour ce résultat encore, l'hypothèse  $p < +\infty$  est nécessaire. Si  $p = +\infty$ , un contre-exemple est encore donné par la fonction de Heavyside.

### 5.3.2 Le produit de convolution

On pourra voir cette vidéo <https://www.youtube.com/watch?v=KuXjwB4LzSA> en guise d'introduction.

**Définition 5.3.3.** Soient  $f, g : \mathbb{R}^n \rightarrow \mathbb{K}$  deux fonctions mesurables. Pour tout  $x \in \mathbb{R}^n$  tel que  $y \mapsto f(y)g(x-y)$  soit intégrable, le **produit de convolution** de  $f$  et  $g$  en  $x$  est défini par

$$f * g(x) = \int_{\mathbb{R}^n} f(y)g(x-y)d\lambda_n(y).$$

**Remarque 5.3.5.** À priori, le produit de convolution  $f * g$  ne définit pas une fonction sur tout  $\mathbb{R}^n$ , ni même presque partout. Par contre, s'il est défini en  $x$ , alors le changement de variable  $z = x - y$  donne  $f * g(x) = g * f(x)$  : le produit de convolution est commutatif.

**Exemple 5.3.1 Probabilité.** Utile pour calculer la loi de la somme de deux variables aléatoires indépendantes. L'expérience aléatoire d'un lancé de deux dés  $X$  et  $Y$  est modélisé par la mesure (produit) sur  $\{1, \dots, 6\}^2$  de deux loi uniforme sur  $\{1, \dots, 6\}$ . La loi de la somme  $S = X + Y$  est alors pour tout  $k = 2, \dots, 12$

$$\mathbb{P}(S = k) = \sum_{i=1}^6 \mathbb{P}(X = i)\mathbb{P}(Y = k - i)$$

avec la convention  $\{X = \ell\} = \{Y = \ell\} = \emptyset$  dès lors que  $\ell \notin \{1, \dots, 6\}$ .

**Exemple 5.3.2 Traitement du signal.** 1. Moyenne mobile :

2. Flou gaussien : <https://docs.gimp.org/2.10/nl/gimp-filter-gaussian-blur.html>

3. Filtre de Sobel : pour la détection de contours [https://en.wikipedia.org/wiki/Sobel\\_operator](https://en.wikipedia.org/wiki/Sobel_operator)

**Exemple 5.3.3 Apprentissage machine.** Les réseaux de neurones convolutifs : [https://www.youtube.com/watch?v=aircAruvnKk&list=PLZHQObOWTQDNU6R1\\_67000Dx\\_ZCJB-3pi](https://www.youtube.com/watch?v=aircAruvnKk&list=PLZHQObOWTQDNU6R1_67000Dx_ZCJB-3pi)

Question : sous quelles hypothèses portant sur  $f$  et  $g$  le produit de convolution  $f * g$  définit-il une fonction sur  $\mathbb{R}^n$  ?

**Théorème 5.3.3.** Soient  $p, q, r \in [1, +\infty]$  tels que  $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$ . Si  $f \in L^p(\mathbb{R}^n)$  et  $g \in L^q(\mathbb{R}^n)$  on a :

1. si  $r = +\infty$  (i.e.  $p$  et  $q$  sont des exposants conjugués), alors  $f * g(x)$  est défini pour tout  $x \in \mathbb{R}^n$ ,  $|f * g(x)| \leq \|f\|_p \|g\|_q$  et  $f * g$  est continue sur  $\mathbb{R}^n$ .
2. si  $r < +\infty$ , alors  $f * g(x)$  est bien défini pour  $\lambda_n$  presque tout  $x \in \mathbb{R}^n$  et  $f * g \in L^r(\mathbb{R}^n)$ . De plus on a  $\|f * g\|_r \leq \|f\|_p \|g\|_q$ .

**Remarque 5.3.6.** On a  $1 \leq p, q \leq r$  nécessairement, car si on avait par exemple  $p > r$  et donc  $\frac{1}{p} < \frac{1}{r}$ , alors à l'aide de  $0 \leq \frac{1}{q} \leq 1$  il viendrait par addition  $\frac{1}{p} + \frac{1}{q} < \frac{1}{r} + 1$ , ce qui contredirait l'hypothèse  $\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1$ , et le même argument s'applique pour justifier  $q \leq r$ .

**Démonstration.** 1. Comme les exposants  $p$  et  $q$  sont conjugués, l'inégalité de Hölder et le changement de variable  $z = x - y$  donnent

$$\begin{aligned} \int_{\mathbb{R}^n} |f(y)g(x-y)| d\lambda_n(y) &\leq \left( \int_{\mathbb{R}^n} |f(y)|^p d\lambda_n(y) \right)^{\frac{1}{p}} \left( \int_{\mathbb{R}^n} |g(x-y)|^q d\lambda_n(y) \right)^{\frac{1}{q}} \\ &= \left( \int_{\mathbb{R}^n} |f(y)|^p d\lambda_n(y) \right)^{\frac{1}{p}} \left( \int_{\mathbb{R}^n} |g(z)|^q d\lambda(z) \right)^{\frac{1}{q}} \\ &= \|f\|_p \|g\|_q < +\infty \end{aligned}$$

Pour tout  $x \in \mathbb{R}^n$  la fonction  $y \mapsto f(y)g(x-y)$  est donc intégrable et  $f * g(x)$  est bien définie.

Pour tout  $x, z \in \mathbb{R}^n$  on a

$$\begin{aligned} f * g(x) - f * g(x - z) &= \int_{\mathbb{R}^n} f(y)(g(x - y) - g(x - z - y)) d\lambda_n(y) \\ &= \int_{\mathbb{R}^n} f(y)(g(x - y) - \tau_z g(x - y)) d\lambda_n(y) \end{aligned}$$

Par inégalité triangulaire et inégalité de Hölder on obtient alors

$$|f * g(x) - f * g(x - z)| \leq \int_{\mathbb{R}^n} |f(y)| |g(x - y) - \tau_z g(x - y)| d\lambda_n(y) \leq \|f\|_p \|g - \tau_z g\|_q.$$

- Si  $q < +\infty$  alors la Proposition 5.3.1 donne  $\lim_{z \rightarrow 0} \|g - \tau_z g\|_q = 0$ , donc  $\lim_{z \rightarrow 0} |f * g(x) - f * g(x - z)| = 0$  et  $f * g$  est continue en  $x$ .
- Si  $q = +\infty$ , alors  $p = 1$  et la commutativité du produit de convolution et un calcul similaire au précédent donnent

$$|f * g(x) - f * g(x - z)| = |g * f(x) - g * f(x - z)| \leq \|g\|_q \|f - \tau_z f\|_p.$$

On conclut de la même façon, car  $p < +\infty$ .

2. On traite tout d'abord le cas  $p = q = r = 1$  en montrant que

$$\|f * g\|_1 \leq \|f\|_1 \|g\|_1. \quad (5.9)$$

En effet, on observe que l'on peut appliquer le Théorème de Fubini-Tonelli car la fonction  $(x, y) \mapsto |f(y)g(x - y)|$  est positive :

$$\begin{aligned} \int_{\mathbb{R}^n} \left( \int_{\mathbb{R}^n} |f(y)g(x - y)| d\lambda_n(y) \right) d\lambda_n(x) &= \int_{\mathbb{R}^n} |f(y)| \underbrace{\left( \int_{\mathbb{R}^n} |g(x - y)| d\lambda_n(x) \right)}_{=\|g\|_1} d\lambda_n(y) \\ &= \|g\|_1 \int_{\mathbb{R}^n} |f| d\lambda_n = \|f\|_1 \|g\|_1 < +\infty \end{aligned}$$

Ainsi, l'application  $x \mapsto \int_{\mathbb{R}^n} |f(y)g(x - y)| d\lambda_n(y)$  prend des valeurs finies  $\lambda_n$ -pp car son intégrale est finie. On en déduit par l'inégalité triangulaire que  $f * g$  est elle aussi bien définie  $\lambda_n$ -pp. Pour avoir le contrôle sur la norme, on utilise directement l'inégalité triangulaire et la majoration que l'on vient de montrer :

$$\|f * g\|_1 = \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} f(y)g(x - y) d\lambda_n(y) \right| d\lambda_n(x) \leq \int_{\mathbb{R}^n} \left( \int_{\mathbb{R}^n} |f(y)g(x - y)| d\lambda_n(y) \right) d\lambda_n(x),$$

qui permet de montrer l'inégalité (5.9).

Pour démontrer le cas général, on va admettre provisoirement que l'inégalité suivante

$$((|f| * |g|)(x))^r \leq \underbrace{\|f\|_p^{r-p} \|g\|_q^{r-q}}_{\text{constantes}} (|f|^p * |g|^q)(x) \quad (5.10)$$

est satisfaite pour presque tout  $x \in \mathbb{R}^n$ .

Dans le second membre de l'inégalité (5.10), on a par hypothèse  $|f|^p \in L^1(\mathbb{R}^n)$  et  $|g|^q \in L^1(\mathbb{R}^n)$ , et puisque  $L^1(\mathbb{R}^n) * L^1(\mathbb{R}^n) \subset L^1(\mathbb{R}^n)$ , on voit que le membre à droite de l'inégalité est une fonction intégrable par rapport à  $x$ , d'où l'on déduit que le membre à gauche est dans  $L^1(\mathbb{R}^n)$ , ce qui revient à dire que  $|f| * |g|$  appartient à  $L^r(\mathbb{R}^n)$ , en d'autres termes :

$$\int_{\mathbb{R}^n} \left( \int_{\mathbb{R}^n} |f(x - y)| |g(y)| d\lambda_n(y) \right)^r d\lambda_n(x) < +\infty.$$

Il découle en particulier de cette inégalité que la fonction  $y \mapsto f(x - y)g(y)$  est intégrable pour presque tout  $x \in \mathbb{R}^n$  fixé, et donc que la convolée  $f * g(x)$  est définie pour presque tout  $x \in \mathbb{R}^n$ .

On peut estimer la norme  $L^r(\mathbb{R}^n)$  de la convolée (à la puissance  $r$ ) :

$$\begin{aligned} \|f * g\|_r^r &= \int_{\mathbb{R}^n} |(f * g)(x)|^r d\lambda_n(x) = \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} f(x-y)g(y) d\lambda_n(y) \right|^r d\lambda_n(x) \\ &\leq \int_{\mathbb{R}^n} \underbrace{\left( \int_{\mathbb{R}^n} |f(x-y)g(y)| d\lambda_n(y) \right)^r}_{|f| * |g|(x)} d\lambda_n(x) \\ &= \| |f| * |g| \|_r^r. \end{aligned}$$

Grace à l'inégalité (5.10) on a alors

$$\|f * g\|_r^r \leq \|f\|_p^{r-p} \|g\|_q^{r-q} \underbrace{\| |f|^p \|_1}_{\in L^1} * \underbrace{\| |g|^q \|_1}_{\in L^1}.$$

Puis on peut appliquer l'inégalité (5.9) qui donne

$$\begin{aligned} \|f * g\|_r^r &\leq \|f\|_p^{r-p} \|g\|_q^{r-q} \| |f|^p \|_1 \| |g|^q \|_1 \\ &= \|f\|_p^{r-p} \|g\|_q^{r-q} \|f\|_p^p \|f\|_q^q \\ &= \|f\|_p^r \|f\|_q^r, \end{aligned}$$

et en prenant finalement la racine  $r$ -ème de cette dernière inégalité, on obtient la majoration annoncée dans le théorème.

Il reste maintenant à établir (5.10). On note  $p'$  et  $q'$  les deux exposants conjugués de  $p$  et de  $q$ . De la relation  $\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1$  qui existe par hypothèse entre  $p$ ,  $q$  et  $r$ , on remarque que

$$1 = \frac{1}{p'} + \frac{1}{q'} + \frac{1}{r}$$

Or il existe une inégalité de Hölder généralisée à trois termes. Pour tout triplet d'exposants réels  $1 < s_1, s_2, s_3 < +\infty$  tels que  $1 = \frac{1}{s_1} + \frac{1}{s_2} + \frac{1}{s_3}$ , et tout triplet de fonctions mesurables  $f_1 \in L^{s_1}(\mathbb{R}^n)$ ,  $f_2 \in L^{s_2}(\mathbb{R}^n)$  et  $f_3 \in L^{s_3}(\mathbb{R}^n)$  le produit  $f_1 f_2 f_3$  appartient à  $L^1(\mathbb{R}^n)$ , et on a

$$\|f_1 f_2 f_3\|_1 \leq \|f_1\|_{s_1} \|f_2\|_{s_2} \|f_3\|_{s_3}.$$

L'idée est de décomposer  $(x, y) \mapsto |f(x-y)| |g(y)|$  en trois facteurs rentrant dans les hypothèses de l'inégalité de Hölder généralisée

$$|f(x-y)| |g(y)| = \underbrace{(|f(x-y)|^p)^{\frac{1}{q'}}}_{=h_1(x,y)} \underbrace{(|g(y)|^q)^{\frac{1}{p'}}}_{=h_2(x,y)} \underbrace{(|f(x-y)|^p |g(y)|^q)^{\frac{1}{r}}}_{=h_3(x,y)}.$$

Ainsi, en appliquant l'inégalité de Hölder à trois termes, on obtient :

$$\begin{aligned} (|f| * |g|)(x) &= \int_{\mathbb{R}^n} |h_1(x, y) h_2(x, y) h_3(x, y)| d\lambda_n(x) \\ &\leq \left( \int_{\mathbb{R}^n} |f(x-y)|^p d\lambda_n(y) \right)^{\frac{1}{q'}} \left( \int_{\mathbb{R}^n} |g(y)|^q d\lambda_n(y) \right)^{\frac{1}{p'}} \left( \int_{\mathbb{R}^n} |f(x-y)|^p |g(y)|^q d\lambda_n(y) \right)^{\frac{1}{r}} \\ &= \|f\|_p^{\frac{p}{q'}} \|g\|_q^{\frac{q}{p'}} ((|f|^p * |g|^q)(x))^{\frac{1}{r}}. \end{aligned} \tag{5.11}$$

Remarquons alors que

$$\frac{rp}{q'} = rp \left( 1 - \frac{1}{p'} - \frac{1}{r} \right) = rp \left( 1 - \frac{p-1}{p} - \frac{1}{r} \right) = r - p$$

et de même que  $\frac{rq}{p'} = r - q$ . En prenant alors la puissance  $r$ -ème dans (5.11), on a :

$$\begin{aligned} ((|f| * |g|)(x))^r &\leq \|f\|_p^{\frac{rp}{q'}} \|g\|_q^{\frac{rq}{p'}} (|f|^p * |g|^q)(x) \\ &= \|f\|_p^{r-p} \|g\|_q^{r-q} (|f|^p * |g|^q)(x), \end{aligned}$$

ce qui est bien l'inégalité (5.10). □

### 5.3.3 Densité des fonctions $C^\infty$ à support compact

Le produit de convolution est utile pour ses propriétés de régularisation. Ainsi, une fonction juste continue et intégrable, après convolution par une application  $C^\infty$ , peut devenir  $C^\infty$ .

Dans l'énoncé ci-dessous, si  $f : \mathbb{R}^n \rightarrow \mathbb{K}$  est  $C^\infty$  et  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$  est un multi-indice, on note

$$\frac{\partial^\alpha f}{\partial x^\alpha}(x_1, \dots, x_n) = \frac{\partial^{\alpha_1 + \dots + \alpha_n} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}(x_1, \dots, x_n)$$

On note de plus  $|\alpha| = \alpha_1 + \dots + \alpha_n$ .

**Proposition 5.3.2.** Soit  $p \in [1, +\infty[$ . Si  $f \in L^p(\mathbb{R}^n)$  et  $g \in C_c^\infty(\mathbb{R}^n)$  alors

1.  $f * g$  est bien définie.
2.  $f * g$  est de classe  $C^\infty(\mathbb{R}^n)$  et  $\frac{\partial^\alpha (f * g)}{\partial x^\alpha} = f * \left( \frac{\partial^\alpha g}{\partial x^\alpha} \right)$ .

*Démonstration.* 1. Idée : Pour la bonne définition de  $f * g$  on peut noter que  $g \in C_c^\infty(\mathbb{R}^n) \subset L^q(\mathbb{R}^n)$  avec  $p$  et  $q$  conjugué. Le Théorème 5.3.3 point 1 donne le résultat. Mais la majoration suivante sera utile dans la suite de la preuve.

On note que pour tout  $x \in \mathbb{R}^n$  l'application  $f(\cdot)g(x - \cdot)$  est mesurable car  $g$  est continue (donc borélienne) et  $f$  est mesurable.

Comme  $g$  est à support compact, il existe  $\delta > 0$  tel que  $g(x - \cdot) = 0$  sur  $B(x, \delta)^c$ . On note  $K_x = B(x, \delta)$  et on a pour tout  $x \in \mathbb{R}^n$

$$|f(\cdot)g(x - \cdot)| \leq \|g\|_\infty |f(\cdot)\mathbf{1}_{K_x}(\cdot)|$$

qui est intégrable car

$$\int_{\mathbb{R}^n} |f| \mathbf{1}_{K_x} d\lambda_n \leq \|f\|_p \|\mathbf{1}_{K_x}\|_q = \|f\|_p \lambda_n(B(0, \delta))^{\frac{1}{q}} < +\infty$$

pour  $\frac{1}{p} + \frac{1}{q} = 1$  par l'inégalité de Hölder. La fonction  $f * g$  est donc bien définie sur tout  $\mathbb{R}^n$ .

2. On procède en trois temps :

Étape 1. On fixe  $x_0 \in \mathbb{R}^n$  arbitraire et on applique le Théorème de continuité sur  $B(x_0, 1)$ . En effet,  $x \mapsto f(y)g(x - y)$  est continue sur  $B(x_0, 1)$  pour presque tout  $y \in \mathbb{R}^n$  et on a

$$|f(y)g(x - y)| \leq \|g\|_\infty |f(y)\mathbf{1}_{B(x_0, \delta+1)}(y)|$$

qui est bien intégrable (appliquer l'inégalité de Hölder comme au point précédent). La fonction  $f * g$  est donc continue en tout  $x_0 \in \mathbb{R}^n$ .

Étape 2. Pour  $|\alpha| = 1$ , on fixe  $x_0 \in \mathbb{R}^n$  arbitraire et on applique le Théorème de dérivation sur  $B(x_0, 1)$ . En effet, on observe que  $x \mapsto f(y)g(x - y)$  est  $C^1(B(x_0, 1))$  pour presque tout  $y$  et, pour tout  $j = 1, \dots, n$ ,

$$\left| \frac{\partial}{\partial x_j} f(y)g(x - y) \right| \leq \left\| \frac{\partial g}{\partial x_j} \right\|_\infty |f(y)\mathbf{1}_{B(x_0, \delta+1)}(y)|.$$

qui est intégrable sur  $\mathbb{R}^n$  (encore une fois, appliquer l'inégalité de Hölder). La fonction  $x \mapsto f * g(x)$  est donc  $C^1$  en tout  $x_0 \in \mathbb{R}^n$  et  $\frac{\partial f * g}{\partial x_j} = f * \frac{\partial g}{\partial x_j}$ .

Étape 3. Pour  $|\alpha| \geq 2$ , on procède par récurrence : L'Étape 1. donne l'initialisation ; on suppose ensuite la propriété vérifiée au rang  $k \in \mathbb{N}$  : l'Étape 2. permet de conclure que  $f * g$  est dérivable  $k + 1$  fois et l'Étape 1. donne la continuité de cette dérivée et donc la propriété au rang  $k + 1$ .

□

### Régularisation par convolution



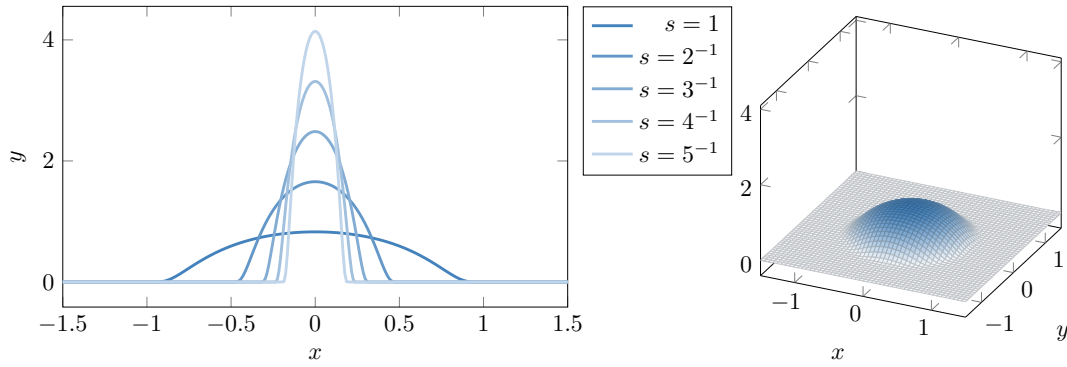


FIGURE 5.6 – La fonction  $\rho_s : \mathbb{R}^n \rightarrow \mathbb{R}$  définie avec la formule (5.12). À gauche : pour  $n = 1$  et différentes valeurs du paramètre  $s > 0$ . À droite : pour  $n = 2$  et  $s = 1$ .

**Définition 5.3.4 Suite régularisante.** Soit  $\rho : \mathbb{R}^n \rightarrow \mathbb{R}_+$  une fonction de  $C_c^\infty(\mathbb{R}^n)$  telle que  $\rho \geq 0$ ,  $\{x \in \mathbb{R}^n \mid \rho(x) \neq 0\} \subset B(0, 1)$  et  $\int_{\mathbb{R}^n} \rho d\lambda_n = 1$ . On appelle **famille régularisante** les fonctions définies pour tout  $s > 0$  par

$$\rho_s : x \mapsto \frac{1}{s^n} \rho\left(\frac{x}{s}\right).$$

La famille de fonction  $(\rho_s)_{s>0}$  est parfois appelée **unité approchée** : les  $f_s$  se “concentrent” autour de l’origine tout en restant d’intégrale 1.

**Remarque 5.3.7.** • En utilisant le changement de variable (linéaire)  $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^n$  définie par  $x \mapsto \frac{x}{s}$  on a  $|\det(J_\varphi(x))| = s^{-n}$  et

$$\int_{\mathbb{R}^n} \rho_s d\lambda_n = \int_{\mathbb{R}^n} \rho\left(\frac{x}{s}\right) \frac{1}{s^n} d\lambda_n(x) = \int_{\mathbb{R}^n} \rho(\varphi(x)) |\det(J_\varphi(x))| d\lambda_n(x) = \int_{\mathbb{R}^n} \rho(y) d\lambda_n(y) = 1$$

- La famille régularisante est à support compact et on a  $\{x \in \mathbb{R}^n \mid \rho_s(x) \neq 0\} \subset B(0, s)$ .

**Exemple 5.3.4.** Soit  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}_+$  définie par

$$\phi(x) = \begin{cases} \exp\left(-\frac{1}{1-\|x\|^2}\right) & \text{si } \|x\| < 1 \\ 0 & \text{sinon.} \end{cases}$$

Cette fonction est bien de classe  $C^\infty(\mathbb{R}^n)$  comme composée de fonctions  $C^\infty(\mathbb{R}^n)$ , et à support compact. Dans la suite, on note  $\rho : \mathbb{R}^n \rightarrow \mathbb{R}_+$  définie par

$$\rho(x) = \frac{\phi(x)}{\int_{\mathbb{R}^n} \phi d\lambda_n} \quad (5.12)$$

On a alors bien  $\rho \in C_c^\infty(\mathbb{R}^n)$ ,  $\rho \geq 0$ ,  $\int_{\mathbb{R}^n} \rho d\lambda_n = 1$  et  $\rho$  a son support dans la boule de rayon 1.

**Lemme 5.3.1 Régularisation par convolution.** Soient  $(\rho_s)_{s>0}$  une famille régularisante et  $f \in L^p(\mathbb{R}^n)$ . On pose alors  $f_k = f * \rho_{\frac{1}{k}}$  pour tout  $k \in \mathbb{N}^*$ . On a

1.  $f_k$  bien définie partout sur  $\mathbb{R}^n$ .
2.  $f_k \in C^\infty(\mathbb{R}^n)$ .
3. s’il existe  $a > 0$  tel que  $f = 0$   $\lambda_n$ -pp sur  $B(0, a)^c$ , on a alors  $f * \rho_s = 0$  sur  $B(0, a + s)^c$ . Autrement dit, si  $f$  est à support compact alors  $f * \rho_s$  est aussi à support compact.

*Démonstration.* 1. C’est une conséquence de la Proposition 5.3.2.

2. C’est une conséquence de la Proposition 5.3.2.

3. Soit  $x \in \mathbf{B}(0, a+s)^c$ , on va montrer que  $f(\cdot)\rho_s(x-\cdot) = 0$   $\lambda_n$ -pp et donc que  $f * \rho_s(x) = 0$ . Comme  $f = 0$   $\lambda_n$ -pp sur  $\mathbf{B}(0, a)^c$ , on a aussi  $f(\cdot)\rho_s(x-\cdot) = 0$   $\lambda_n$ -pp sur  $\mathbf{B}(0, a)^c$ . Comme  $\rho_s = 0$  sur  $\mathbf{B}(0, s)^c$ , on a  $f(\cdot)\rho_s(x-\cdot) = 0$  sur  $\mathbf{B}(x, s)^c$  (car  $y \in \mathbf{B}(x, s)^c \Leftrightarrow (x-y) \in \mathbf{B}(0, s)^c$ ). On a donc :

$$f(\cdot)\rho_s(x-\cdot) = 0 \text{ } \lambda_n\text{-pp sur } x \in \mathbf{B}(0, a)^c \cup \mathbf{B}(x, s)^c. \quad (5.13)$$

Or  $\mathbf{B}(0, a) \cap \mathbf{B}(x, s) = \emptyset$  car  $y \in \mathbf{B}(0, a) \cap \mathbf{B}(x, s)$  implique  $\|y\| < a$  et  $\|x-y\| < s$  et donc  $\|x\| = \|x-y+y\| < a+s$ , ce qui contredit  $x \in \mathbf{B}(0, a+s)^c$ . On a donc  $\mathbf{B}(0, a)^c \cup \mathbf{B}(x, s)^c = (\mathbf{B}(0, a) \cap \mathbf{B}(x, s))^c = \mathbb{R}^n$  et (5.13) donne alors  $f(\cdot)\rho_s(x-\cdot) = 0$   $\lambda_n$ -pp et donc  $f * \rho_s$  est définie au point  $x \in \mathbf{B}(0, a+s)^c$  et  $f * \rho_s(x) = 0$ . □

Si on convole une fonction  $f$  par une suite de fonctions régularisante, on peut retrouver la fonction  $f$  à la limite.

**Proposition 5.3.3 Densité de  $C^\infty(\mathbb{R}^n)$  dans  $L^p(\mathbb{R}^n)$ .** Soit  $p \in [1, +\infty[$ , pour tout  $f \in L^p(\mathbb{R}^n)$  on a  $\lim_{k \rightarrow \infty} \|f * \rho_{\frac{1}{k}} - f\|_p = 0$ .

*Démonstration.* La démonstration est une conséquence de la continuité en moyenne (Proposition 5.3.1). On choisit un représentant de  $f \in L^p(\mathbb{R}^n)$ , encore noté  $f \in \mathcal{L}^p(\mathbb{R}^n)$ . Pour  $k \in \mathbb{N}^*$ , on pose  $f_k = f * \rho_{\frac{1}{k}}$ . Pour  $x \in \mathbb{R}^n$ , on a :

$$\begin{aligned} f_k(x) - f(x) &= \int_{\mathbb{R}^n} f(x-y)\rho_{\frac{1}{k}}(y)d\lambda_n(y) - f(x) \int_{\mathbb{R}^n} \rho_{\frac{1}{k}}(y)d\lambda_n(y) \\ &= \int_{\mathbb{R}^n} (f(x-y) - f(x))\rho_{\frac{1}{k}}(y)d\lambda_n(y), \end{aligned}$$

et donc :

$$|f_k(x) - f(x)|^p \leq \left( \int_{\mathbb{R}^n} |f(x-y) - f(x)| \rho_{\frac{1}{k}}(y)d\lambda_n(y) \right)^p$$

Pour  $p > 1$ , on utilise l'inégalité de Höder en écrivant  $\rho_{\frac{1}{k}} = \rho_{\frac{1}{k}}^{\frac{1}{p}} \rho_{\frac{1}{k}}^{\frac{1}{q}}$  (avec  $q = \frac{p}{p-1}$ ) et on obtient (ce qui est aussi immédiatement vrai pour  $p = 1$ ) :

$$\begin{aligned} |f_k(x) - f(x)|^p &\leq \int_{\mathbb{R}^n} |f(x-y) - f(x)|^p \rho_{\frac{1}{k}}(y)d\lambda_n(y) \underbrace{\left( \int_{\mathbb{R}^n} \rho_{\frac{1}{k}}(y)d\lambda_n(y) \right)^{\frac{p}{q}}}_{=1} \\ &\leq \int_{\mathbb{R}^n} |f(x-y) - f(x)|^p \rho_{\frac{1}{k}}(y)d\lambda_n(y). \end{aligned} \quad (5.14)$$

On remarque maintenant que pour tout  $k \in \mathbb{N}^*$ , l'application  $(x, y) \mapsto |f(x-y) - f(x)|^p \rho_{\frac{1}{k}}(y)$  est mesurable et positive de  $\mathbb{R}^2$  dans  $\mathbb{R}_+$ . On peut donc appliquer le théorème de Fubini-Tonelli pour déduire de (5.14) que pour  $\delta > 0$  on a :

$$\begin{aligned} \int_{\mathbb{R}^n} |f_k(x) - f(x)|^p d\lambda_n(x) &\leq \int_{\mathbb{R}^n} \left( \int_{\mathbb{R}^n} |f(x-y) - f(x)|^p \rho_{\frac{1}{k}}(y)d\lambda_n(y) \right) d\lambda_n(x) \\ &= \int_{\mathbb{R}^n} \left( \int_{\mathbb{R}^n} |f(x-y) - f(x)|^p d\lambda_n(x) \right) \rho_{\frac{1}{k}}(y)d\lambda_n(y) \\ &= \int_{\mathbf{B}(0, \frac{1}{k})} \|\tau_y f - f\|_p^p \rho_{\frac{1}{k}}(y)d\lambda_n(y). \end{aligned}$$

Soit maintenant  $\varepsilon > 0$  fixé. En appliquant la Proposition 5.3.1, on peut choisir  $\eta > 0$  suffisamment petit tel que pour tout  $h \in \mathbb{R}^n$ ,  $|h| < \eta$  implique  $\|\tau_h f - f\|_p \leq \varepsilon$ . Ainsi, si  $\frac{1}{k} \leq \eta$  on a  $\|f_k - f\|_p \leq \varepsilon$ , ce qui démontre le résultat. □

La proposition précédente est le point clé de la preuve du théorème de densité suivant :

**Théorème 5.3.4 Densité de  $C_c^\infty(\mathbb{R}^n)$  dans  $L^p(\mathbb{R}^n)$ .** Sur  $\mathbb{R}^n$  muni de la mesure de Lebesgue  $\lambda_n$ , l'espace  $C_c^\infty(\mathbb{R}^n)$  des fonctions  $C^\infty$  à support compact est dense dans  $L^p(\mathbb{R}^n)$  pour tout  $p \in [1, +\infty[$ .

*Démonstration.* La démonstration proposée ici utilise une méthode dite de **troncature et régularisation**. On note  $A$  l'ensemble des fonctions  $f \in L^p(\mathbb{R}^n)$  à support compact.

Étape 1. On montre dans cette étape que  $A$  est dense dans  $L^p(\mathbb{R}^n)$ . Soit donc  $f \in L^p(\mathbb{R}^n)$  et pour  $k \in \mathbb{N}$ , on pose

$$f_k = f \mathbb{1}_{B(0,k)}.$$

C'est l'étape de troncature. Comme  $f_k \rightarrow f$   $\lambda_n$ -pp quand  $k \rightarrow +\infty$  et  $|f_k| \leq |f|$   $\lambda_n$ -pp (pour tout  $k \in \mathbb{N}$ ), on peut appliquer le Théorème de convergence dominée dans  $L^p(\mathbb{R}^n)$  (on utilise ici le fait que  $p < \infty$ ). Il donne que  $f_k \rightarrow f$  dans  $L^p(\mathbb{R}^n)$  quand  $k \rightarrow +\infty$ . Comme  $(f_k)_{k \in \mathbb{N}}$  est une suite de  $A$ , on a bien montré la densité de  $A$  dans  $L^p(\mathbb{R}^n)$ .

Étape 2. Soit maintenant  $f \in A$ . Pour conclure la démonstration, il suffit de montrer qu'il existe une suite  $(f_k)_{k \in \mathbb{N}}$  de  $C_c^\infty(\mathbb{R}^n)$  telle que  $f_k \rightarrow f$  dans  $L^p(\mathbb{R}^n)$  quand  $k \rightarrow +\infty$ . Or cette suite est donnée par

$$f_k = f * \rho_{\frac{1}{k}}, \text{ pour } k \in \mathbb{N}^*,$$

où  $(\rho_s)_{s>0}$  une suite régularisante à support compact définie à l'aide de la formule (5.12). En effet, le Lemme 5.3.1 permet d'affirmer que la suite  $(f_k)_{k \in \mathbb{N}}$  est dans  $C_c^\infty(\mathbb{R}^n)$  et la Proposition 5.3.3 assure que  $f_k \rightarrow f$  dans  $L^p(\mathbb{R}^n)$  quand  $k \rightarrow +\infty$ .

□



## Chapitre 6

# Analyse de Fourier sur $\mathbb{R}$

En guise d'introduction, on pourra regarder la vidéo suivante : <https://www.youtube.com/watch?v=spUNpyF58BY>.

### 6.1 Définition et propriétés

La transformée de Fourier est un exemple d'**opérateur linéaire intégral** communément appelé **opérateur à noyau** : une fonction linéaire définie à l'aide d'une intégrale paramétrique sur certains espaces vectoriels de fonctions.

**Définition 6.1.1.** Si  $f \in L^1(\mathbb{R})$  la transformée de Fourier de  $f$  est la fonction  $\hat{f} : \mathbb{R} \rightarrow \mathbb{C}$  définie par

$$\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{-ix\xi} dx.$$

On peut noter l'analogie d'écriture entre opérateurs à noyaux et opérateurs matriciels dans les espaces de dimension fini. Étant donnée une matrice  $M \in \mathbb{R}^{n \times n}$ , l'image  $\hat{f} \in \mathbb{R}^n$  d'un vecteur  $f = (f_1, \dots, f_n) \in \mathbb{R}^n$  s'écrit avec une multiplication matricielle

$$\begin{pmatrix} \hat{f}_1 \\ \vdots \\ \hat{f}_n \end{pmatrix} = Mf = \sum_{\ell=1}^n M_{k\ell} f_{\ell} \quad \Leftrightarrow \quad \hat{f}_k = \sum_{\ell=1}^n M_{k\ell} f_{\ell}, \quad \text{pour tout } k = 1, \dots, n.$$

On remarque que le rôle de  $M = (M_{k\ell})_{k,\ell=1}^n$  est joué par le noyau  $\mathbb{R}^2 \ni (x, \xi) \mapsto \frac{1}{\sqrt{2\pi}} e^{-ix\xi}$  et le signe  $\sum$  remplacé par le signe  $\int$ . Contrairement à la dimension finie où les espaces de départ et d'arrivées sont facilement identifiés, en dimension infinie, il faut les préciser rigoureusement. Par exemple, la Définition 6.1.1 affirme que la transformée de Fourier est définie pour  $f \in L^1(\mathbb{R})$  et la Proposition 6.1.1 affirme que  $\hat{f}$  est "seulement" dans  $C_0^0(\mathbb{R})$  l'espace des fonctions continues qui s'annulent à l'infinie (mais  $\hat{f}$  n'est pas nécessairement dans  $L^1(\mathbb{R})$ ).

**Remarque 6.1.1.** • Comme  $|f(x)e^{-ix\xi}| = |f(x)|$  et  $f \in L^1(\mathbb{R})$ , la fonction  $x \mapsto f(x)e^{-ix\xi}$  est intégrable et  $\hat{f}(\xi)$  est bien défini pour tout  $\xi \in \mathbb{R}$ .

- La transformée de Fourier se définit aussi pour une fonction  $f \in L^1(\mathbb{R}^n)$  par

$$\hat{f}(\xi) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} f(x) e^{-i\langle x, \xi \rangle} dx.$$

où  $\langle x, \xi \rangle = \sum_{k=1}^n x_k \xi_k$  est le produit scalaire canonique sur  $\mathbb{R}^n$ . Les propriétés (et leurs preuves) sont similaires à celle de la transformée de Fourier sur  $\mathbb{R}$ . Dans la suite on se contentera de traiter le cas de  $\mathbb{R}$ .

**Proposition 6.1.1.** La transformée de Fourier a les propriétés suivantes :

- Linéarité :  $\widehat{\alpha f + g} = \alpha \hat{f} + \hat{g}$
- $\hat{f} : \mathbb{R} \rightarrow \mathbb{R}$  est bien définie en tout point, continue et bornée. En particulier,  $\|\hat{f}\|_\infty \leq \frac{1}{\sqrt{2\pi}}\|f\|_1$  et  $f \mapsto \hat{f}$  est une application linéaire continue de  $L^1(\mathbb{R})$  dans  $L^\infty(\mathbb{R})$ .
- Lemme de Riemann-Lebesgue :  $\lim_{|\xi| \rightarrow +\infty} \hat{f}(\xi) = 0$ .

*Démonstration.* • La linéarité est immédiate.

- Pour tout  $x \in \mathbb{R}$ , l'application  $\xi \mapsto f(x)e^{-ix\xi}$  est continue et on a  $|f(x)e^{-ix\xi}| = |f(x)|$  pour tout  $x$  et tout  $\xi$ . Comme  $f \in L^1(\mathbb{R})$ , le théorème de continuité s'applique et  $\hat{f}$  est continue sur  $\mathbb{R}$ . De plus, l'inégalité triangulaire donne

$$|\hat{f}(\xi)| \leq \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} |f(x)e^{-ix\xi}| dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} |f(x)| dx = \frac{1}{\sqrt{2\pi}} \|f\|_1,$$

donc  $\hat{f}$  est bornée sur  $\mathbb{R}$  et  $\|\hat{f}\|_\infty \leq \frac{1}{\sqrt{2\pi}}\|f\|_1$ .

- Soit  $\xi \neq 0$ , comme  $e^{-i\pi} = -1$ , on a

$$\begin{aligned} \hat{f}(\xi) &= \frac{e^{-i\pi}}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x)e^{-ix\xi} dx \\ &= \frac{-1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x)e^{-i\xi(x+\frac{\pi}{\xi})} dx \\ &= \frac{-1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(y - \frac{\pi}{\xi})e^{-iy\xi} dy \end{aligned}$$

où on a effectué le changement de variable  $y = x + \frac{\pi}{\xi}$ . En ajoutant à cette égalité l'expression usuelle de  $\hat{f}(\xi)$  on obtient  $2\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} (f(x) - f(x - \frac{\pi}{\xi}))e^{-ix\xi} dx$  et

$$|\hat{f}(\xi)| \leq \frac{1}{2\sqrt{2\pi}} \int_{\mathbb{R}} |f(x) - f(x - \frac{\pi}{\xi})| dx = \frac{1}{2\sqrt{2\pi}} \|\tau_{\frac{\pi}{\xi}} f - f\|_1.$$

Comme  $f \in L^1(\mathbb{R})$  et  $\lim_{|\xi| \rightarrow +\infty} \frac{\pi}{\xi} = 0$ , on a  $\lim_{|\xi| \rightarrow +\infty} \|\tau_{\frac{\pi}{\xi}} f - f\|_1 = 0$  par la Proposition 5.3.1. □

**Exemple 6.1.1 Fonction caractéristique de la Loi de Laplace.** Pour tout  $a > 0$ , soit  $v_a : \mathbb{R} \rightarrow \mathbb{R}$  définie par  $v_a(x) = e^{-a|x|}$ . Cette fonction vérifie  $x \mapsto x^n v_a(x)$  intégrable, pour tout  $n \in \mathbb{N}$ , donc sa transformée de Fourier est  $C^\infty$ . Elle se calcule :

$$\begin{aligned} \hat{v}_a(\xi) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-a|x|} e^{-ix\xi} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{x(a-i\xi)} dx + \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} e^{-x(a+i\xi)} dx \\ &= \frac{1}{\sqrt{2\pi}} \left[ \frac{e^{x(a-i\xi)}}{a-i\xi} \right]_{x \rightarrow -\infty}^{x=0} + \frac{1}{\sqrt{2\pi}} \left[ \frac{e^{-x(a+i\xi)}}{-a-i\xi} \right]_{x=0}^{x \rightarrow +\infty} \\ &= \frac{1}{\sqrt{2\pi}} \left[ \frac{1}{a-i\xi} - 0 \right] + \frac{1}{\sqrt{2\pi}} \left[ 0 - \frac{1}{-a-i\xi} \right] \\ &= \frac{1}{\sqrt{2\pi}} \frac{a+i\xi}{a^2+\xi^2} + \frac{1}{\sqrt{2\pi}} \frac{a-i\xi}{a^2+\xi^2} \\ &= \frac{1}{\sqrt{2\pi}} \frac{2a}{a^2+\xi^2} \end{aligned}$$

où on a utilisé que  $x \mapsto e^{-ix\xi}$  est bornée sur  $\mathbb{R}$  pour les passages à la limite. En probabilité, on utilise des conventions légèrement différentes. Soit  $X$  une variable aléatoire de loi de Laplace de paramètre  $\alpha > 0$ , alors la densité de  $X$  est  $\frac{1}{2}v_{\frac{1}{\alpha}}$  et la fonction caractéristique est

$$\varphi_X(\xi) = \mathbb{E}(e^{iX\xi}) = \frac{1}{2\alpha} \int_{\mathbb{R}} v_{\frac{1}{\alpha}} e^{ix\xi} dx = \frac{\sqrt{2\pi}}{2\alpha} \hat{v}_{\frac{1}{\alpha}}(-\xi) = \frac{1}{1 + \alpha^2 \xi^2}.$$

**Proposition 6.1.2.** Pour  $f, g \in L^1(\mathbb{R})$  on a  $\int_{\mathbb{R}} \hat{f}g d\lambda_1 = \int_{\mathbb{R}} f\hat{g} d\lambda_1$ .

*Démonstration.* C'est l'Exercice 5 du TD VI. □

Si  $f : \mathbb{R}^n \rightarrow \mathbb{C}$ , pour  $a \in \mathbb{R}^n$  et  $\lambda \in \mathbb{R} \setminus \{0\}$ , on rappelle la définition de la translatée de  $f$  par  $a$ , et on définit la **dilatée** de  $f$  par  $\lambda$  comme

$$\begin{aligned} (\tau_a f)(x) &= f(x - a), \\ (\delta_\lambda f)(x) &= f\left(\frac{x}{\lambda}\right), \end{aligned}$$

pour tout  $x \in \mathbb{R}^n$ .

**Lemme 6.1.1.** Si  $f \in L^1(\mathbb{R})$ ,  $a \in \mathbb{R}$  et  $\lambda \in \mathbb{R} \setminus \{0\}$ , on a

$$\begin{aligned} \widehat{\tau_a f}(\xi) &= e^{-ia\xi} \hat{f}(\xi), \\ \widehat{\delta_\lambda f}(\xi) &= |\lambda| \delta_{\frac{1}{\lambda}} \hat{f}(\xi), \end{aligned}$$

pour tout  $\xi \in \mathbb{R}$ .

*Démonstration.* On a, en utilisant un changement de variable,

$$\begin{aligned} \widehat{\tau_a f}(\xi) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x - a) e^{-ix\xi} dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(y) e^{-i(y+a)\xi} dy \\ &= \frac{e^{-ia\xi}}{\sqrt{2\pi}} \int_{\mathbb{R}} f(y) e^{-iy\xi} dy = e^{-ia\xi} \hat{f}(\xi). \end{aligned}$$

De même,

$$\begin{aligned} \widehat{\delta_\lambda f}(\xi) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f\left(\frac{x}{\lambda}\right) e^{-ix\xi} dx = \frac{|\lambda|}{\sqrt{2\pi}} \int_{\mathbb{R}} f(y) e^{-i(\lambda y)\xi} dy \\ &= \frac{|\lambda|}{\sqrt{2\pi}} \int_{\mathbb{R}} f(y) e^{-iy(\lambda\xi)} dy = |\lambda| \hat{f}(\lambda\xi) \\ &= |\lambda| \delta_{\frac{1}{\lambda}} \hat{f}(\xi). \end{aligned}$$

□

### 6.1.1 Transformée de Fourier et dérivation

**Proposition 6.1.3.** Soit  $f \in L^1(\mathbb{R})$  et  $n \in \mathbb{N}$ . Si  $x \mapsto x^n f(x)$  est intégrable, alors  $\hat{f}$  est  $C^n$  et

$$\hat{f}^{(k)}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} (-ix)^k f(x) e^{-ix\xi} dx$$

pour tout  $k \in \llbracket 0, n \rrbracket$ .

**Remarque 6.1.2.** On remarque en particulier que  $\hat{f}^{(k)}$  est la transformée de Fourier de  $x \mapsto (-ix)^k f(x)$ , donc elle hérite de toutes les propriétés des transformées de Fourier.

*Démonstration.* Soit  $F : \mathbb{R}^2 \rightarrow \mathbb{C}$  définie par  $F(x, \xi) = f(x) e^{-ix\xi}$ . On vérifie les hypothèses du théorème de dérivation sous le signe intégral (Corollaire 3.4.1) :

- Comme  $f \in L^1(\mathbb{R})$ ,  $x \mapsto F(x, \xi)$  est intégrable pour tout  $\xi \in \mathbb{R}$ .
- Pour tout  $k \in \llbracket 0, n \rrbracket$  et tout  $x \in \mathbb{R}$ , l'application  $\xi \mapsto F(x, \xi)$  est de classe  $C^k$  et  $\frac{\partial^k F}{\partial \xi^k}(x, \xi) = (-ix)^k f(x) e^{-ix\xi}$ .
- Pour tout  $k \in \llbracket 0, n \rrbracket$  et tout  $(x, \xi) \in \mathbb{R}^2$  on a

$$\left| \frac{\partial^k F}{\partial \xi^k}(x, \xi) \right| = |x^k f(x)| \leq |f(x)| \mathbb{1}_{[-1, 1]}(x) + |x^n f(x)| \mathbb{1}_{\mathbb{R} \setminus [-1, 1]}(x).$$

Comme  $f$  et  $x \mapsto x^n f(x)$  sont intégrables,  $\frac{\partial^k F}{\partial \xi^k}$  est dominée par une fonction intégrable.

On en déduit que  $\hat{f}$  est de classe  $C^n(\mathbb{R})$  et, pour tout  $k \in \llbracket 0, n \rrbracket$  on a

$$\hat{f}^{(k)}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \frac{\partial^k F}{\partial \xi^k} dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} (-ix)^k f(x) e^{-ix\xi} dx.$$

□

**Proposition 6.1.4.** Soit  $f \in L^1(\mathbb{R}) \cap C^1(\mathbb{R})$  telle que  $f'$  soit intégrable. Alors

$$\widehat{f'}(\xi) = i\xi \hat{f}(\xi).$$

Si  $f$  est  $C^n(\mathbb{R})$  et  $f^{(k)}$  est intégrable pour tout  $k \in \llbracket 0, n \rrbracket$ , alors

$$\widehat{f^{(k)}}(\xi) = (i\xi)^k \hat{f}(\xi)$$

pour tout  $k \in \llbracket 0, n \rrbracket$ .

*Démonstration.* Comme  $f$  est  $C^1(\mathbb{R})$ , le Théorème fondamental de l'analyse donne  $f(x) = f(0) + \int_0^x f'(t) dt$ , et comme  $f'$  est intégrable sur  $\mathbb{R}$  on peut passer à la limite et on obtient

$$\lim_{x \rightarrow +\infty} f(x) = f(0) + \int_0^{+\infty} f'(t) dt \in \mathbb{C}.$$

De même pour la limite en  $-\infty$ . Comme  $f$  est intégrable et possède des limites finies en  $+\infty$  et  $-\infty$ , ces limites sont nécessairement nulles :  $\lim_{|x| \rightarrow +\infty} f(x) = 0$ .

La définition de  $\widehat{f'}(\xi)$  et une intégration par parties donnent

$$\begin{aligned} \widehat{f'}(\xi) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f'(x) e^{-ix\xi} dx \\ &= \frac{1}{\sqrt{2\pi}} \left[ f(x) e^{-ix\xi} \right]_{x \rightarrow -\infty}^{x \rightarrow +\infty} - \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) (-i\xi) e^{-ix\xi} dx \\ &= 0 + i\xi \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{-ix\xi} dx \\ &= i\xi \hat{f}(\xi), \end{aligned}$$

où on a utilisé que  $x \mapsto e^{-ix\xi}$  est bornée sur  $\mathbb{R}$  et  $\lim_{|x| \rightarrow +\infty} f(x) = 0$ .

Si  $f$  est de classe  $C^n(\mathbb{R})$  et  $f^{(k)}$  est intégrable pour tout  $k \in \llbracket 0, n \rrbracket$ , la preuve se fait par récurrence sur  $k$ . L'initialisation pour  $k = 0$  est immédiate et l'hérédité s'obtient en appliquant le calcul précédent à la fonction  $f^{(k)}$ . □

**Remarque 6.1.3.** • la transformée de Fourier transforme la dérivation en une multiplication, d'où son intérêt dans le cadre des équations différentielles.



- Si  $f$  est  $C^\infty(\mathbb{R})$  et  $f^{(k)}$  est intégrable pour tout  $k \in \mathbb{N}$ , la proposition précédente montre que  $\xi \mapsto (i\xi)^k \hat{f}(\xi)$  est la transformée de Fourier de  $f^{(k)}$ , donc le Lemme de Riemann-Lebesgue donne  $\lim_{|\xi| \rightarrow +\infty} |\xi|^k |\hat{f}(\xi)| = 0$  et  $\hat{f}(\xi) = o\left(\frac{1}{|\xi|^k}\right)$  en  $\pm\infty$ . Si une fonction est très régulière, sa transformée de Fourier décroît très vite à l'infini (plus vite que n'importe quelle puissance de  $\xi$ ).

**Exemple 6.1.2.** On peut utiliser la transformée de Fourier pour trouver une solution particulière à l'équation différentielle

$$u''(x) - u(x) = f(x)$$

où  $f \in L^1(\mathbb{R}) \cap C^0(\mathbb{R})$  est donnée. Si on cherche des solutions intégrables et de dérivées intégrables, on peut passer en Fourier. L'équation devient pour tout  $\xi \in \mathbb{R}$ ,

$$\begin{aligned} \widehat{u''}(\xi) - \hat{u}(\xi) &= \hat{f}(\xi) \quad \Leftrightarrow \quad -\xi^2 \hat{u}(\xi) - \hat{u}(\xi) = \hat{f}(\xi) \\ &\Leftrightarrow \quad \hat{u}(\xi) = \frac{-1}{1+\xi^2} \hat{f}(\xi). \end{aligned}$$

Une façon de terminer le calcul est la suivante :

1. trouver une fonction  $g$  telle que  $\hat{g}(\xi) = \frac{-1}{1+\xi^2}$ . L'exemple précédent montre que  $g : \mathbb{R} \rightarrow \mathbb{R}$  définie par  $g(x) = -\sqrt{\frac{\pi}{2}} e^{-|x|}$  vérifie  $\hat{g}(\xi) = -\sqrt{\frac{\pi}{2}} \hat{v}_1(\xi) = \frac{-1}{1+\xi^2}$ . Notre équation devient alors  $\hat{u} = \hat{g}\hat{f}$ .
2. déterminer une fonction  $h$  telle que  $\hat{h} = \hat{g}\hat{f}$  (de façon générale, quels sont les liens entre transformée de Fourier et produit). L'équation devient  $\hat{u} = \hat{h}$
3. pouvoir “inverser” la transformée de Fourier pour en conclure que  $u = h$

La suite donne les résultats permettant de réaliser les étapes 2 et 3.

### 6.1.2 Produit de convolution et formule d'inversion

**Proposition 6.1.5.** Pour  $f, g \in L^1(\mathbb{R})$  on a  $\widehat{f * g}(\xi) = \sqrt{2\pi} \hat{f}(\xi) \hat{g}(\xi)$ .

*Démonstration.* Comme  $f, g \in L^1(\mathbb{R})$ ,  $f * g(x)$  est bien défini presque partout et  $f * g \in L^1(\mathbb{R})$ . Donc  $\widehat{f * g}(\xi)$  est bien défini pour tout  $\xi$  et

$$\widehat{f * g}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left( \int_{\mathbb{R}} f(x-y)g(y)dy \right) e^{-ix\xi} dx.$$

Par le théorème de Tonelli on obtient

$$\begin{aligned} \int_{\mathbb{R}^2} |f(x-y)g(y)e^{-ix\xi}| dx dy &= \int_{\mathbb{R}} \left( \int_{\mathbb{R}} |f(x-y)||g(y)| dx \right) dy \\ &= \int_{\mathbb{R}} |g(y)| \left( \int_{\mathbb{R}} |f(x-y)| dx \right) dy \\ &= \|f\|_1 \int_{\mathbb{R}} |g(y)| dy \\ &= \|f\|_1 \|g\|_1 < +\infty. \end{aligned}$$

Donc  $(x, y) \mapsto f(x-y)g(y)e^{-ix\xi}$  est intégrable sur  $\mathbb{R}^2$  et le théorème de Fubini donne

$$\widehat{f * g}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g(y) \left( \int_{\mathbb{R}} f(x-y)e^{-ix\xi} dx \right) dy.$$

Pour  $y$  fixé, le changement de variable  $z = x - y$  donne

$$\begin{aligned} \widehat{f * g}(\xi) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g(y) \left( \int_{\mathbb{R}} f(z)e^{-i(z+y)\xi} dz \right) dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g(y)e^{-iy\xi} \left( \int_{\mathbb{R}} f(z)e^{-iz\xi} dz \right) dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g(y)e^{-iy\xi} \sqrt{2\pi} \hat{f}(\xi) dy \\ &= \sqrt{2\pi} \hat{f}(\xi) \hat{g}(\xi). \end{aligned}$$

□

**Exemple 6.1.3 (suite).** Pour suivre l'exemple 6.1.2, en prenant  $g : \mathbb{R} \rightarrow \mathbb{R}$  définie par  $g(x) = -\sqrt{\frac{\pi}{2}}e^{-|x|}$ , on a  $\hat{g}(\xi) = \frac{-1}{1+\xi^2}$ , et

$$\hat{u}(\xi) = \hat{g}(\xi)\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}}\widehat{g * f}(\xi),$$

et si on sait inverser la transformée de Fourier on peut en conclure que  $u = \frac{1}{\sqrt{2\pi}}g * f$ .

**Théorème 6.1.1 Formule d'inversion.** Soit  $f \in L^1(\mathbb{R})$  telle que  $\hat{f} \in L^1(\mathbb{R})$ . La fonction  $g : \mathbb{R} \rightarrow \mathbb{R}$  définie par

$$g(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(\xi) e^{i\xi x} d\xi.$$

vérifie que  $f(x) = g(x)$  pour presque tout  $x \in \mathbb{R}$ .

*Démonstration.* Idée : On applique deux fois la transformée de Fourier à  $f$  et on montre que  $\hat{\hat{f}}(\cdot) = f(-\cdot)$   $\lambda_1$ -pp (autrement dit, on retrouve quasiment  $f$  quand on applique successivement deux fois la transformée de Fourier). Pour arriver à cela, on part de la définition

$$\hat{\hat{f}}(x) = \int_{\mathbb{R}} \hat{f}(\xi) e^{-i\xi x} d\xi.$$

Malheureusement, il n'est pas possible d'appliquer la Proposition 6.1.2 car  $\xi \mapsto e^{-i\xi x}$  n'est pas intégrable. La preuve consiste à perturber le problème avec une suite de fonctions bien choisie et qui tends vers 1...

Soit  $(s_n)_{n \in \mathbb{N}}$  une suite de réels telle que  $s_n > 0$  pour tout  $n \in \mathbb{N}$  et  $\lim_{n \rightarrow +\infty} s_n = 0$ . On remarque que la famille de fonctions définie pour tout  $x \in \mathbb{R}$  par

$$\rho_{s_n}(x) = \frac{1}{\pi} \frac{s_n}{s_n^2 + x^2}$$

est une suite régularisante au sens de la Définition 5.3.4. On a donc, d'après la Proposition 5.3.3

$$\lim_{n \rightarrow \infty} \|f * \rho_{s_n} - f\|_1 = 0. \quad (6.1)$$

De plus, d'après l'Exemple 6.1.1 on a  $\widehat{e^{|\cdot|s_n}}(\xi) = \sqrt{2\pi}\rho_{s_n}(\xi)$ ,

$$\begin{aligned} f * \rho_{s_n}(x) &= \int_{\mathbb{R}} f(x-y)\rho_{s_n}(y)dy \\ &\stackrel{\text{Ex. 6.1.1}}{=} \int_{\mathbb{R}} \tau_{-x}f(-y) \frac{\widehat{e^{|\cdot|s_n}}}{\sqrt{2\pi}}(y)dy \\ &\stackrel{\text{Lem. 6.1.1, Prop. 6.1.2}}{=} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \widehat{\tau_{-x}f(-y)} e^{|y|s_n} dy \\ &\stackrel{\text{Lem. 6.1.1}}{=} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(-y) e^{-ixy} e^{|y|s_n} dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(\xi) e^{ix\xi} e^{|\xi|s_n} d\xi. \end{aligned}$$

On remarque enfin que  $|\hat{f}(\xi)e^{ix\xi}e^{|\xi|s_n}| \leq |\hat{f}(\xi)|$  qui est bien intégrable par hypothèse. On peut alors appliquer le Théorème de convergence dominée :

$$\lim_{n \rightarrow +\infty} f * \rho_{s_n}(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \lim_{n \rightarrow +\infty} \hat{f}(\xi) e^{ix\xi} e^{|\xi|s_n} d\xi = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(\xi) e^{ix\xi} d\xi = \hat{f}(-x).$$

Enfin, la relation (6.1) implique que  $f * \rho_{s_n} \rightarrow f$  dans  $L^1(\mathbb{R})$  quand  $n \rightarrow +\infty$ . Cela permet d'appliquer la remarque 5.1.4 (Théorème de Riesz-Fischer) qui implique, qu'à extraction de sous suite près, on a  $f * \rho_{s_n} \rightarrow f$   $\lambda_1$ -pp quand  $n \rightarrow +\infty$ .

L'unicité de la limite (pour les suites dans  $\mathbb{R}$ ) permet alors de conclure que pour  $\lambda_1$ -pt  $x \in \mathbb{R}$  on a  $\hat{f}(-x) = f(x)$ .  $\square$

**Remarque 6.1.4.** La formule d'inversion est très proche de la formule donnant la transformée de Fourier. Si  $\hat{f} \in L^1(\mathbb{R})$ , la fonction  $g$  donnée par la formule d'inversion possède des propriétés analogues (qui se démontrent de la même façon) :  $g$  est continue sur  $\mathbb{R}$ , bornée par  $\|\hat{f}\|_1$  et  $\lim_{|x| \rightarrow +\infty} g(x) = 0$ .

**Exemple 6.1.4.** Il existe des fonctions  $f \in L^1(\mathbb{R})$  telles que  $\hat{f} \notin L^1(\mathbb{R})$  : c'est le cas dès que  $f$  n'est pas égale  $\lambda_1$ -pp à une fonction continue, ou dès que  $f$  ne tends pas vers 0 à l'infini. Par exemple, si  $f : x \mapsto e^{-x} \mathbb{1}_{\mathbb{R}_+}(x)$ , alors  $\hat{f} : \xi \mapsto \frac{1}{\sqrt{2\pi}(1+i\xi)}$  qui n'est pas dans  $L^1(\mathbb{R})$ .

**Corollaire 6.1.1.** Soient  $f, g \in L^1(\mathbb{R})$ . Si  $\hat{f} = \hat{g}$  alors  $f = g$ . La transformée de Fourier est injective sur  $L^1(\mathbb{R})$ .

*Démonstration.* Soit  $f, g \in L^1(\mathbb{R})$  telles que  $\hat{f} = \hat{g}$ . La linéarité de la transformée de Fourier implique que  $\widehat{f-g} = 0$  et donc que  $\widehat{f-g} \in L^1(\mathbb{R})$ . En appliquant le Théorème 6.1.1 d'inversion on a  $f = g$   $\lambda_1$ -pp.  $\square$

**Exemple 6.1.5 (suite et fin).** Si  $u$  est une solution intégrable et de dérivées intégrables de l'équation différentielle

$$u''(x) - u(x) = f(x)$$

alors on a montré que  $\hat{u}(\xi) = \hat{g}(\xi)\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}}\widehat{g*f}(\xi)$ , avec  $g(x) = -\sqrt{\frac{\pi}{2}}e^{-|x|}$ . L'injectivité de la transformée de Fourier donne

$$u(x) = \frac{1}{\sqrt{2\pi}}g * f(x) = \frac{1}{\sqrt{2\pi}}\left(-\sqrt{\frac{\pi}{2}}\right) \int_{\mathbb{R}} f(y)e^{-|x-y|}dy = -\frac{1}{2} \int_{\mathbb{R}} f(y)e^{-|x-y|}dy.$$

On obtient donc, en fonction de  $f$  une solution particulière de l'équation différentielle.

## 6.2 Le cadre fonctionnel

La transformée de Fourier transforme les fonctions intégrables en fonctions continues, bornées qui tendent vers 0 à l'infini. La formule d'inversion montre qu'elle est injective, mais elle n'est pas surjective. Il existe des fonctions continues bornées tendant vers 0 à l'infini et qui ne sont pas des transformées de Fourier de fonctions intégrables.

Le but ici est de construire un espace de fonctions  $E$  stable par la transformée de Fourier et tel que

$$\begin{array}{ccc} \mathcal{F} : E & \longrightarrow & E \\ f & \longmapsto & \hat{f} \end{array}.$$

soit un isomorphisme. On se base pour cela sur les liens que fait la transformée de Fourier entre dérivabilité et décroissance à l'infini.

### 6.2.1 Espace de Schwartz

**Définition 6.2.1.** Une fonction  $f : \mathbb{R} \rightarrow \mathbb{C}$  est à **décroissance rapide** si  $f$  est  $C^\infty(\mathbb{R})$  et, pour tout  $j, k \in \mathbb{N}$ , on a  $\sup_{x \in \mathbb{R}} |x^k f^{(j)}(x)| < +\infty$ . L'**espace de Schwartz** est l'ensemble des fonctions à décroissance rapide :

$$\mathcal{S}(\mathbb{R}) = \{f \in C^\infty(\mathbb{R}) \mid \forall j, k \in \mathbb{N} \ x \mapsto x^k f^{(j)}(x) \text{ est bornée} \}.$$

**Remarque 6.2.1.** L'espace de Schwartz contient les fonctions régulières dont les dérivées décroissent plus vite que tout polynôme. Il contient en particulier  $C_c^\infty(\mathbb{R})$  et est dense dans  $L^p(\mathbb{R})$  pour tout  $p \in [1, +\infty[$ .

**Proposition 6.2.1.** L'application  $\mathcal{F} : \mathcal{S}(\mathbb{R}) \rightarrow \mathcal{S}(\mathbb{R})$  est une bijection dont la réciproque est donnée par la formule d'inversion :

$$\mathcal{F}^{-1}(g)(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g(\xi) e^{i\xi x} d\xi.$$

*Démonstration.* • On commence par remarquer que  $\mathcal{S}(\mathbb{R}) \subset L^1(\mathbb{R})$ . En effet, si  $f \in \mathcal{S}(\mathbb{R})$ , en prenant des choix convenables de  $k$  et  $j$  dans la définition de  $\mathcal{S}(\mathbb{R})$ , on remarque qu'il existe  $C \in \mathbb{R}_+$  tel que  $(1 + |x|^2) |f(x)| \leq C$ . On en déduit alors que  $f \in L^1(\mathbb{R})$ . Plus généralement, si  $f \in \mathcal{S}(\mathbb{R})$ , on remarque que  $x \mapsto x^k f^{(j)}(x)$  est dans  $L^1(\mathbb{R})$  pour tout  $k, j \in \mathbb{N}$ .

- Les Propositions 6.1.3 et 6.1.4 impliquent alors que si  $f \in \mathcal{S}(\mathbb{R})$ , on a pour tout  $k, j \in \mathbb{N}$  :

$$\begin{aligned} \mathcal{F} \left[ ((-i \cdot)^k f)^{(j)} \right] &= (i \cdot)^j \mathcal{F} [(-i \cdot)^k f] = (i \cdot)^j \hat{f}^{(k)}, \\ \mathcal{F} [(-i \cdot)^k f^{(j)}] &= \left( \widehat{f^{(j)}} \right)^{(k)} = \left[ (i \cdot)^j \hat{f} \right]^{(k)}. \end{aligned}$$

- En utilisant le point précédent, on montre que  $\mathcal{F}$  envoie  $\mathcal{S}(\mathbb{R})$  dans  $\mathcal{S}(\mathbb{R})$ . Le théorème 6.1.1 d'inversion donne alors que  $\mathcal{F}$  est injective et que  $f = \hat{\hat{f}}(-\cdot)$  pour tout  $f \in \mathcal{S}(\mathbb{R})$ . De cette dernière formule, on déduit que  $\mathcal{F}$  est surjective (et donc bijective) de  $\mathcal{S}(\mathbb{R})$  dans  $\mathcal{S}(\mathbb{R})$ .  $\square$

## 6.2.2 Transformée de Fourier $L^2$

L'espace de Schwartz peut être muni d'une multitude de normes. La norme la plus naturelle à considérer serait la norme  $\|\cdot\|_\infty$ . Mais on choisit la norme  $\|\cdot\|_2$  car elle est préservée par transformation de Fourier (pour les fonctions de carré intégrable). Le produit hermitien associé a été défini par la formule (5.5) à la Remarque 5.1.5.

**Proposition 6.2.2.** Si  $f, g \in \mathcal{S}(\mathbb{R})$  (on a donc  $f, g \in L^2(\mathbb{R})$ ), alors  $\hat{f}, \hat{g} \in L^2(\mathbb{R})$  et  $\langle f, g \rangle = \langle \hat{f}, \hat{g} \rangle$ .

*Démonstration.* Soient  $f, g \in \mathcal{S}(\mathbb{R})$ . On a  $\hat{f}, \hat{g} \in \mathcal{S}(\mathbb{R})$  par la Proposition 6.2.1. Ainsi  $f, g, \hat{f}, \hat{g} \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ . Reste à voir que  $\langle f, g \rangle = \langle \hat{f}, \hat{g} \rangle$ .

On peut appliquer la formule d'inversion au produit scalaire de  $f$  avec  $g$ . On a

$$\langle f, g \rangle = \int_{\mathbb{R}} \hat{f}(-x) \bar{g}(x) dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left( \int_{\mathbb{R}} e^{i\xi x} \hat{f}(\xi) d\xi \right) \bar{g}(x) dx.$$

On peut maintenant utiliser le Théorème de Fubini car  $\hat{f}, \bar{g} \in L^1(\mathbb{R})$  et on obtient

$$\langle f, g \rangle = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left( \int_{\mathbb{R}} e^{i\xi x} \bar{g}(x) dx \right) \hat{f}(\xi) d\xi = \langle \hat{f}, \hat{g} \rangle.$$

$\square$

**Remarque 6.2.2.** En particulier,  $\mathcal{F}$  est une isométrie de l'espace  $(\mathcal{S}(\mathbb{R}), \|\cdot\|_2)$ .

L'espace  $(\mathcal{S}(\mathbb{R}), \|\cdot\|_2)$  n'est pas complet, mais  $\mathcal{F}$  s'étend naturellement en une isométrie sur tout  $L^2(\mathbb{R})$ . Le théorème suivant est un cas particulier d'un résultat très général d'extension des applications uniformément continues définies sur un sous-espace dense (cf. cours de topologie).

**Théorème 6.2.1 Plancherel.** Il existe une application linéaire continue  $\tilde{\mathcal{F}}$  de  $L^2(\mathbb{R})$  dans  $L^2(\mathbb{R})$  telle que :

1. Si  $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ , on a alors  $\tilde{\mathcal{F}}(f) = \hat{f}$   $\lambda_1$ -pp.
2. Pour tout  $f, g \in L^2(\mathbb{R})$ , on a  $\langle f, g \rangle = \langle \tilde{\mathcal{F}}(f), \tilde{\mathcal{F}}(g) \rangle$ .
3. Pour tout  $f \in L^2(\mathbb{R})$ , on a  $f = \tilde{\mathcal{F}}(\tilde{\mathcal{F}}(f))(-\cdot)$ .

4.  $\tilde{\mathcal{F}}$  est une bijection de  $L^2(\mathbb{R})$  dans  $L^2(\mathbb{R})$ .

Pour  $f \in L^2(\mathbb{R})$ , la fonction  $\tilde{\mathcal{F}}(f)$  s'appelle la **transformée de Fourier-Plancherel** de  $f$ . Compte tenu du premier item, on notera en général,  $\hat{f}$  la transformée de Fourier de  $f$  si  $f \in L^1(\mathbb{R})$  (et alors  $\hat{f} = \mathcal{F}(f) \in C_0(\mathbb{R})$ ) ou si  $f \in L^2(\mathbb{R})$  (et alors  $\hat{f} = \tilde{\mathcal{F}}(f) \in L^2(\mathbb{R})$ ).

*Démonstration.* L'application  $\mathcal{F} : \mathcal{S}(\mathbb{R}) \rightarrow \mathcal{S}(\mathbb{R})$  définie par  $f \mapsto \hat{f}$  est linéaire et continue pour la norme de  $L^2(\mathbb{R})$ , et définie sur  $\mathcal{S}(\mathbb{R})$  qui est dense dans  $L^2(\mathbb{R})$  (car  $\mathcal{S}(\mathbb{R})$  contient  $C_c^\infty(\mathbb{R})$  qui est dense dans  $L^2(\mathbb{R})$  d'après le Théorème 5.3.4). On décrit comment prolonger  $\mathcal{F}$  par  $\tilde{\mathcal{F}} : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$  en conservant le caractère linéaire et continue.

Soit donc  $f \in L^2(\mathbb{R})$ , il existe  $(f_n)_{n \in \mathbb{N}}$  suite de  $\mathcal{S}(\mathbb{R})$  telle que  $f_n \rightarrow f$  dans  $L^2(\mathbb{R})$  quand  $n \rightarrow +\infty$ . La suite  $(f_n)_{n \in \mathbb{N}}$  est donc de Cauchy dans  $L^2(\mathbb{R})$ . La Proposition 6.2.2 implique que la suite  $(\hat{f}_n)_{n \in \mathbb{N}}$  est aussi de Cauchy dans  $L^2(\mathbb{R})$  et converge donc dans  $L^2(\mathbb{R})$ . On aimerait définir  $\tilde{\mathcal{F}}(f)$  comme étant la limite (dans  $L^2(\mathbb{R})$ ) de la suite  $(\hat{f}_n)_{n \in \mathbb{N}}$ . Ceci est possible à condition que cette limite ne dépende que de  $f$  et non du choix de la suite  $(f_n)_{n \in \mathbb{N}}$  convergeant dans  $L^2(\mathbb{R})$  vers  $f$ . Or, ce dernier point est facile car si  $(g_n)_{n \in \mathbb{N}}$  est une autre suite convergeant dans  $L^2(\mathbb{R})$  vers  $f$ , on a  $\|\hat{f}_n - \hat{g}_n\|_2 = \|f_n - g_n\|_2 \rightarrow 0$  quand  $n \rightarrow +\infty$ . On a ainsi défini  $\tilde{\mathcal{F}}$  de  $L^2(\mathbb{R})$  dans lui-même.

La linéarité de  $\tilde{\mathcal{F}}$  découle immédiatement du fait que l'application  $f \mapsto \hat{f}$  est linéaire de  $\mathcal{S}(\mathbb{R})$  dans  $L^2(\mathbb{R})$ . Enfin, soit  $f \in L^2(\mathbb{R})$  et  $(f_n)_{n \in \mathbb{N}} \subset \mathcal{S}(\mathbb{R})$  tels que  $f_n \rightarrow f$  dans  $L^2(\mathbb{R})$ . La Proposition 6.2.2 donne que  $\|\hat{f}_n\|_2 = \|f_n\|_2$  pour tout  $n \in \mathbb{N}$ . En passant à la limite quand  $n \rightarrow +\infty$ , on en déduit  $\|\tilde{\mathcal{F}}(f)\|_2 = \|f\|_2$ . Ce qui prouve la continuité de  $\tilde{\mathcal{F}}$  de  $L^2(\mathbb{R})$  dans  $L^2(\mathbb{R})$ .

On montre maintenant les 4 items du théorème.

1. Soit  $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ . En reprenant la démonstration du Théorème 5.3.4, il est facile de voir qu'il existe une suite  $(f_n)_{n \in \mathbb{N}} \subset C_c^\infty(\mathbb{R})$  telle que  $f_n \rightarrow f$  dans  $L^2(\mathbb{R})$  et  $L^1(\mathbb{R})$  lorsque  $n \rightarrow +\infty$  (car dans la démonstration du Théorème 5.3.4, la suite construite pour converger vers  $f$  dans  $L^p$  ne dépend pas de  $p$ ). On en déduit que  $\hat{f}_n \rightarrow \hat{f}$  uniformément sur  $\mathbb{R}$  lorsque  $n \rightarrow +\infty$  (car  $f_n \rightarrow f$  dans  $L^1(\mathbb{R})$ ) et que  $\hat{f}_n \rightarrow \tilde{\mathcal{F}}(f)$  dans  $L^2(\mathbb{R})$  lorsque  $n \rightarrow +\infty$  (car  $f_n \rightarrow f$  dans  $L^2(\mathbb{R})$ ) et donc que, après extraction éventuelle d'une sous-suite, on peut supposer que  $\hat{f}_n \rightarrow \tilde{\mathcal{F}}(f)$   $\lambda_1$ -pp quand  $n \rightarrow +\infty$ . On en déduit bien que  $\hat{f} = \tilde{\mathcal{F}}(f)$   $\lambda_1$ -pp.
2. Soit  $f, g \in L^2(\mathbb{R})$ . Il existe deux suites  $(f_n)_{n \in \mathbb{N}}$  et  $(g_n)_{n \in \mathbb{N}}$  de  $\mathcal{S}(\mathbb{R})$  telles que  $f_n \rightarrow f$  et  $g_n \rightarrow g$  dans  $L^2(\mathbb{R})$ . La Proposition 6.2.2 implique que  $\langle \hat{f}_n, \hat{g}_n \rangle = \langle f_n, g_n \rangle$  pour tout  $n \in \mathbb{N}$ . En passant à la limite quand  $n \rightarrow +\infty$  on obtient bien  $\langle \tilde{\mathcal{F}}(f), \tilde{\mathcal{F}}(g) \rangle = \langle f, g \rangle$ .
3. Soit  $f \in L^2(\mathbb{R})$  et  $(f_n)_{n \in \mathbb{N}}$  suite de  $\mathcal{S}(\mathbb{R})$  telles que  $f_n \rightarrow f$  dans  $L^2(\mathbb{R})$  quand  $n \rightarrow +\infty$ . On a donc  $\hat{f}_n \rightarrow \tilde{\mathcal{F}}(f)$  dans  $L^2(\mathbb{R})$ , ce qui donne aussi  $\hat{f}_n(-\cdot) \rightarrow \tilde{\mathcal{F}}(f)(-\cdot)$  dans  $L^2(\mathbb{R})$  et donc  $\widehat{\hat{f}_n}(-\cdot) \rightarrow \widehat{\tilde{\mathcal{F}}(f)(-\cdot)}$  dans  $L^2(\mathbb{R})$  quand  $n \rightarrow +\infty$ . La Proposition 6.2.1 implique que  $\widehat{\hat{f}_n} = \widehat{\widehat{f}_n}(-\cdot)$  pour tout  $n \in \mathbb{N}$ . On en déduit donc (par unicité de la limite dans  $L^2$ ) que  $f = \tilde{\mathcal{F}}(\tilde{\mathcal{F}}(f))(-\cdot)$   $\lambda_1$ -pp.
4. L'injectivité de  $\tilde{\mathcal{F}}$  (de  $L^2(\mathbb{R})$  dans  $L^2(\mathbb{R})$ ) découle du fait que  $\|\tilde{\mathcal{F}}(f)\|_2 = \|f\|_2$  et que  $\tilde{\mathcal{F}}$  est linéaire. La surjectivité est une conséquence immédiate du troisième item.

□

**Remarque 6.2.3.** • Attention : si  $f \in L^2(\mathbb{R})$  mais  $f \notin L^1(\mathbb{R})$  et pour  $\xi \in \mathbb{R}$  on ne peut pas calculer  $\mathcal{F}(f)(\xi)$  avec la formule  $\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{-ix\xi} dx$  car elle n'a pas de sens.

- Par ailleurs, il existe des fonctions  $f \in L^1(\mathbb{R})$  telles que  $f \notin L^2(\mathbb{R})$ . Pour ces fonctions, la transformée de Fourier  $\hat{f}$  est bien définie, mais  $\mathcal{F}(f)$  n'est pas définie, car  $f \notin L^2(\mathbb{R})$ .



## Chapitre 7

# Existence et unicité de la mesure de Lebesgue

### 7.1 Mesures extérieures

Remember Banach-Tarski : la boule unité de  $\mathbb{R}^n$  est équi-décomposable à l'union disjointe de deux boules unités. Si on veut mesurer toutes les parties de  $\mathbb{R}^n$  de sorte que toutes les boules unités aient la même mesure, il faut accepter que la somme des mesures des parties soit strictement supérieure à la mesure de leur union. Il existe des partitions de la boule unité qui "rajoute de la mesure". En particulier, on ne peut pas espérer qu'une telle mesure soit additive, on peut demander au mieux qu'elle soit sous-additive.

**Définition 7.1.1.** Une **mesure extérieure** sur  $X$  est une application  $\mu^* : \mathcal{P}(X) \rightarrow \overline{\mathbb{R}}_+$  telle que :

1.  $\mu^*(\emptyset) = 0$ .
2.  $\mu^*$  est croissante : si  $A \subset B$  alors  $\mu^*(A) \leq \mu^*(B)$ .
3.  $\mu^*$  est sous-additive : pour toute famille  $(A_k)_{k \in \mathbb{N}}$  de  $\mathcal{P}(X)$  on a  $\mu^*\left(\bigcup_{k \in \mathbb{N}} A_k\right) \leq \sum_{k \in \mathbb{N}} \mu^*(A_k)$ .

Ces trois propriétés sont satisfaites par les mesures, donc une mesure sur  $\mathcal{P}(X)$  est une mesure extérieure. On remarque que la propriété 3 n'implique pas que la mesure extérieure d'une union disjointe de parties est la somme des mesures extérieures des parties. Pour combler ce manque, on considère la définition suivante

**Définition 7.1.2.** Soit  $X$  un ensemble muni d'une mesure extérieure  $\mu^*$ . Une partie  $B \in \mathcal{P}(X)$  est  $\mu^*$ -régulière si

$$\forall A \in \mathcal{P}(X), \quad \mu^*(A) = \mu^*(A \cap B) + \mu^*(A \cap B^c).$$

On note  $\mathcal{R}(\mu^*)$  l'ensemble des parties régulières.

Une partie  $B$  est régulière si elle permet de découper toute partie  $A$  en deux morceaux disjoints ( $A \cap B$  et  $A \cap B^c$ ) sans rajouter de mesure.

**Remarque 7.1.1.** Pour tout  $A \in \mathcal{P}(X)$ , comme  $A = (A \cap B) \cup (A \cap B^c)$ , la sous-additivité de la mesure extérieure donne

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

Pour montrer que  $B$  est une partie  $\mu^*$ -régulière, il suffit de montrer que pour tout  $A \in \mathcal{P}(X)$ ,

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

**Proposition 7.1.1.** Si  $\mu^*$  est une mesure extérieure sur  $X$ . Alors

1. si  $\mu^*(B) = 0$  alors  $B \in \mathcal{R}(\mu^*)$
2.  $\mathcal{R}(\mu^*)$  est une tribu sur  $X$

3.  $\mu^* : \mathcal{R}(\mu^*) \rightarrow \overline{\mathbb{R}}_+$  est une mesure sur  $(X, \mathcal{R}(\mu^*))$ .

*Démonstration.* 1. Soit  $B \subset X$  telle que  $\mu^*(B) = 0$ . Alors, pour toute partie  $A$  de  $X$ ,

$$\mu^*(A \cap B) \leq \mu^*(B) = 0 \quad \text{par croissance.}$$

Donc

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

et  $B \in \mathcal{R}(\mu^*)$ .

2. Il est clair que  $\emptyset \in \mathcal{R}(\mu^*)$  et que  $B \in \mathcal{R}(\mu^*)$  si et seulement si  $B^c \in \mathcal{R}(\mu^*)$ . Pour montrer que  $\mathcal{R}(\mu^*)$  est une tribu, il reste donc à voir que  $\mathcal{R}(\mu^*)$  est stable par réunion dénombrable. Commençons par l'établir pour une réunion finie. Soient  $B_1, B_2 \in \mathcal{R}(\mu^*)$  et  $A$  une partie de  $X$ . Comme  $B_1 \in \mathcal{R}(\mu^*)$ ,

$$\begin{aligned} \mu^*(A \cap (B_1 \cup B_2)) &= \mu^*(A \cap (B_1 \cup B_2) \cap B_1) + \mu^*(A \cap (B_1 \cup B_2) \cap B_1^c) \\ &= \mu^*(A \cap B_1) + \mu^*(A \cap B_2 \cap B_1^c). \end{aligned}$$

Donc, comme  $B_1, B_2 \in \mathcal{R}(\mu^*)$ ,

$$\begin{aligned} \mu^*(A \cap (B_1 \cup B_2)) + \mu^*(A \cap (B_1 \cup B_2)^c) &= \mu^*(A \cap B_1) + \mu^*(A \cap B_1^c \cap B_2) + \mu^*(A \cap B_1^c \cap B_2^c) \\ &= \mu^*(A \cap B_1) + \mu^*(A \cap B_1^c) \\ &= \mu^*(A) \end{aligned}$$

On en déduit que  $B_1 \cup B_2 \in \mathcal{R}(\mu^*)$ .

Pour terminer la preuve, considérons une suite  $(B_n)_{n \in \mathbb{N}}$  d'éléments de  $\mathcal{R}(\mu^*)$ . On peut supposer, sans perte de généralité, les  $B_n$  deux à deux disjoints. On veut montrer que

$$B = \bigcup_{n \in \mathbb{N}} B_n \in \mathcal{R}(\mu^*) \quad \text{et} \quad \mu^*(B) = \sum_{n \in \mathbb{N}} \mu^*(B_n).$$

Notons  $E_n = \bigcup_{k=0}^n B_k$ . Pour toute partie  $A \subset X$  et tout  $n \in \mathbb{N}$ , on va montrer par récurrence que

$$\mu^*(A \cap E_n) = \sum_{k=0}^n \mu^*(A \cap B_k). \quad (7.1)$$

La formule (7.1) est évidemment vraie au rang  $n = 0$  et si l'on suppose la propriété vraie au rang  $n - 1$ , on a

$$\begin{aligned} \mu^*(A \cap E_n) &= \mu^*(A \cap \underbrace{E_n \cap B_n}_{=B_n}) + \mu^*(A \cap \underbrace{E_n \cap B_n^c}_{=E_{n-1}}) \\ &= \mu^*(A \cap B_n) + \mu^*(A \cap E_{n-1}). \end{aligned}$$

Ce qui donne la propriété au rang  $n$ . En prenant la limite lorsque  $n$  tend vers  $+\infty$ , et en se souvenant que  $E_n \subset B$ , on trouve

$$\lim_{n \rightarrow \infty} \mu^*(A \cap E_n) = \sum_{k=0}^{\infty} \mu^*(A \cap B_k) \leq \mu^*(A \cap B).$$

Par ailleurs, d'après la remarque 7.1.1 et avec la sous-additivité, on a bien

$$\mu^*(A \cap B) = \mu^*\left(\bigcup_{n \in \mathbb{N}} (A \cap B_n)\right) = \sum_{n \in \mathbb{N}} \mu^*(A \cap B_n)$$



En particulier, si  $A = X$  cela donne

$$\mu^*(B) = \sum_{n \in \mathbb{N}} \mu^*(B_n).$$

Enfin, comme  $\mathcal{R}(\mu^*) \ni E_n \subset B$ , on a pour tout  $n \in \mathbb{N}$  et pour toute partie  $A \subset X$ ,

$$\mu^*(A) = \mu^*(A \cap E_n) + \mu^*(A \cap E_n^c) \geq \mu^*(A \cap E_n) + \mu^*(A \cap B^c)$$

En prenant la limite quand  $n \rightarrow +\infty$ , on obtient  $\mu^*(A) \geq \mu^*(A \cap B) + \mu^*(A \cap B^c)$  donc  $B \in \mathcal{R}(\mu^*)$ . □

### 7.1.1 Existence de la mesure de Lebesgue

Un pavé de  $\mathbb{R}^n$  est une partie  $P \in \mathcal{P}(\mathbb{R}^n)$  de la forme

$$P = I_1 \times I_2 \times \cdots \times I_n,$$

où  $I_1, \dots, I_n$  sont des intervalles bornés de  $\mathbb{R}$ .

Pour chaque intervalle  $I \subset \mathbb{R}$  on note  $\ell(I) = \sup(I) - \inf(I) \in \overline{\mathbb{R}}_+$ , et on définit la mesure d'un pavé  $P = I_1 \times I_2 \times \cdots \times I_n$  par

$$m(P) = \prod_{k=1}^n \ell(I_k).$$

**Définition 7.1.3.**  $\lambda^* : \mathcal{P}(\mathbb{R}^n) \rightarrow \overline{\mathbb{R}}$  est définie par

$$\lambda^*(A) = \inf \left\{ \sum_{k=0}^{+\infty} m(P_k) \mid (P_k)_{k \in \mathbb{N}} \text{ suite de pavés telle que } A \subset \bigcup_{k \in \mathbb{N}} P_k \right\}$$

pour tout  $A \in \mathcal{P}(\mathbb{R}^n)$ .

**Théorème 7.1.1.** L'application  $\lambda^* : \mathcal{P}(\mathbb{R}^n) \rightarrow \overline{\mathbb{R}}$  a les propriétés suivantes :

1.  $\lambda^*$  est une mesure extérieure sur  $\mathbb{R}^n$ .
2. La tribu  $\mathcal{L}(\mathbb{R}^n) = \mathcal{R}(\lambda^*)$  est appelée **tribu de Lebesgue** et contient les boréliens :  $\mathcal{B}(\mathbb{R}^n) \subset \mathcal{R}(\lambda^*)$ .
3. Pour tout pavé  $P$  de  $\mathbb{R}^n$  on a  $\lambda^*(P) = m(P)$ .

**Remarque 7.1.2.** La tribu de Lebesgue est strictement plus grosse que la tribu de Lebesgue : on peut montrer que  $\mathcal{B}(\mathbb{R}^n)$  est équipotent à  $\mathbb{R}^n$  alors que  $\mathcal{L}(\mathbb{R}^n)$  est équipotent à  $\mathcal{P}(\mathbb{R}^n)$ .

*Démonstration.* 1. On montre que  $\lambda^*$  est une mesure extérieure. Il est clair que  $\lambda^*(\emptyset) = 0$  et que  $\lambda^*$  est croissante. Montrons la sous-additivité : soit  $(A_k)_{k \in \mathbb{N}}$  une suite de sous-ensembles de  $\mathbb{R}^n$ . On peut supposer que  $\lambda^*(A_k) < +\infty$ , pour tout  $k \in \mathbb{N}$  (sans quoi il n'y a rien à montrer). Soit alors  $\varepsilon > 0$ . Pour tout  $k \in \mathbb{N}$ , il existe une suite  $(P_i^k)_{i \in \mathbb{N}}$  de pavés ouverts tels que

$$A_k \subset \bigcup_{i \in \mathbb{N}} P_i^k \quad \text{et} \quad \sum_{i \in \mathbb{N}} m(P_i^k) \leq \lambda^*(A_k) + \frac{\varepsilon}{2^k}.$$

Alors

$$\bigcup_{k \in \mathbb{N}} A_k = \bigcup_{k \in \mathbb{N}} \bigcup_{i \in \mathbb{N}} P_i^k$$

et

$$\sum_{k \in \mathbb{N}} \sum_{i \in \mathbb{N}} m(P_i^k) \leq \sum_{k \in \mathbb{N}} \left( \lambda^*(A_k) + \frac{\varepsilon}{2^k} \right) = \sum_{k \in \mathbb{N}} \lambda^*(A_k) + 2\varepsilon.$$

Comme  $\varepsilon > 0$  peut être pris arbitrairement petit, on conclut que

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

2. Montrons que  $\mathcal{R}(\lambda^*)$  contient  $\mathcal{B}(\mathbb{R}^n)$ . Comme  $\mathcal{R}(\lambda^*)$  est une tribu de  $\mathbb{R}^n$  (Proposition 7.1.1), il suffit de montrer qu'elle contient un ensemble générateur de  $\mathcal{B}(\mathbb{R}^n)$  pour conclure. Considérons pour cela les demi-espaces ouverts définis à l'aide de  $j \in \{1, \dots, n\}$  et  $\alpha \in \mathbb{R}$  :

$$B_j(\alpha) = \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R} \times \underbrace{]\alpha, +\infty[}_{j\text{-ème position}} \times \mathbb{R} \times \cdots \times \mathbb{R}.$$

La famille  $\{B_j(\alpha)\}_{j \in \{1, \dots, n\}, \alpha \in \mathbb{R}}$  engendre bien la tribu des boréliens  $\mathcal{B}(\mathbb{R}^n)$ . Il suffit donc de vérifier que  $B_j(\alpha) \in \mathcal{R}(\lambda^*)$  pour tout  $j \in \{1, \dots, n\}$  et  $\alpha \in \mathbb{R}$ .

Soit  $A$  une partie de  $\mathbb{R}^n$  et soit  $\varepsilon > 0$ . Il existe une suite de pavés ouverts  $(P_i)_{i \in \mathbb{N}}$  tels que  $A \subset \bigcup_{i \in \mathbb{N}} P_i$  et  $\sum_{i \in \mathbb{N}} m(P_i) \leq \lambda^*(A) + \varepsilon$ . Alors,

$$A \cap B_j(\alpha) \subset \bigcup_{i \in \mathbb{N}} (P_i \cap B_j(\alpha))$$

où  $P_i \cap B_j(\alpha)$  est un pavé ouvert comme intersection d'un pavé ouvert et d'un demi-espace ouvert, et

$$A \cap B_j(\alpha)^c \subset \bigcup_{i \in \mathbb{N}} (P_i \cap B_j(\alpha)^c)$$

où  $P_i \cap B_j(\alpha)^c$  est un pavé quelconque. Or,

$$\forall i \in \mathbb{N}, \quad m(P_i) = m(P_i \cap B_j(\alpha)) + m(P_i \cap B_j(\alpha)^c)$$

Ainsi,

$$\begin{aligned} \lambda^*(A \cap B_j(\alpha)) + \lambda^*(A \cap B_j(\alpha)^c) &\leq \sum_{i \in \mathbb{N}} m(P_i \cap B_j(\alpha)) + \sum_{i \in \mathbb{N}} m(P_i \cap B_j(\alpha)^c) \\ &= \sum_{i \in \mathbb{N}} m(P_i) \\ &\leq \lambda^*(A) + \varepsilon \end{aligned}$$

Et comme  $\varepsilon$  était arbitraire, on obtient

$$\lambda^*(A \cap B_j(\alpha)) + \lambda^*(A \cap B_j(\alpha)^c) \leq \lambda^*(A)$$

donc  $B_j(\alpha) \in \mathcal{R}(\lambda^*)$ .

3. Pour tout pavé  $P \subset \mathbb{R}^n$ , on a évidemment

$$\lambda^*(P) \leq m(P).$$

Inversement, soit  $P \subset \mathbb{R}^n$  un pavé fermé et soit  $(P_i)_{i \in \mathbb{N}}$  un recouvrement de  $P$  par des pavés ouverts :

$$P \subset \bigcup_{i \in \mathbb{N}} P_i$$

Par compacité de  $P$ , il existe un entier  $N \in \mathbb{N}$  tel que  $P \subset \bigcup_{i=1}^N P_i$ . Par le Lemme 7.1.1 on a

$$m(P) \leq \sum_{i=0}^N m(P_i) \leq \sum_{n=0}^{\infty} m(P_i)$$

Ainsi,

$$\lambda^*(P) \geq m(P).$$

Enfin, si  $P$  est un pavé quelconque, on note  $\mathring{P}$  (resp.  $\overline{P}$ ) le pavé ouvert (resp. fermé) appelé aussi l'intérieur (resp. fermeture) de  $P$ . On a

$$\lambda^*(\mathring{P}) \leq m(\mathring{P}) \leq m(\overline{P}) \leq \lambda^*(\overline{P}).$$

Par ailleurs,  $\overline{P} \setminus \mathring{P} = \partial P$  est une union finie de pavés fermés de  $m$ -mesure nulle, donc  $\lambda^*(\partial P) = 0$  par sous-additivité. Ainsi,

$$\lambda^*(\overline{P}) \leq \lambda^*(\mathring{P}) + \lambda^*(\partial P) = \lambda^*(\mathring{P}).$$

On conclut que

$$\lambda^*(P) = \lambda^*(\mathring{P}) = \lambda^*(\overline{P}) = m(P).$$

□

**Lemme 7.1.1.** Soit  $N \in \mathbb{N}^*$ . Si  $P, P_1, \dots, P_N$  sont des pavés de  $\mathbb{R}^n$  avec  $P \subset \bigcup_{i=1}^N P_i$  alors

$$m(P) \leq \sum_{i=1}^N m(P_i).$$

*Démonstration.* Voir [2] Lemme 3.8.

□

**Définition 7.1.4.** La **mesure de Lebesgue** sur  $\mathbb{R}^n$  est la restriction de  $\lambda^*$  aux boréliens de  $\mathbb{R}^n$ . Autrement dit, c'est l'application

$$\begin{aligned} \lambda_n : \mathcal{B}(\mathbb{R}^n) &\rightarrow \overline{\mathbb{R}}_+ \\ A &\mapsto \lambda^*(A) \end{aligned}$$

## 7.1.2 Complétion de mesure

**Définition 7.1.5.** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré. La mesure  $\mu$  est **complète** si tout sous-ensemble d'une partie négligeable est mesurable.

**Remarque 7.1.3.** Certains sous-ensembles de l'ensemble de Cantor n'étant pas mesurable, la mesure de  $\lambda_1$  n'est pas complète sur  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ .

**Proposition 7.1.2.** Soit  $(X, \mathcal{A}, \mu)$  un espace mesuré et soit

$$\mathcal{A}^\sim = \left\{ E \in \mathcal{P}(X) \mid \exists A, B \in \mathcal{A}, A \subset E \subset B \text{ et } \mu(B \setminus A) = 0 \right\}.$$

Si  $E \in \mathcal{A}^\sim$  alors on note  $\bar{\mu}(E) = \mu(A)$ , où  $A \in \mathcal{A}$  est telle qu'il existe  $B \in \mathcal{A}$  vérifiant  $A \subset E \subset B$  et  $\mu(B \setminus A) = 0$ . On a alors

- $\mathcal{A}^\sim$  est une tribu et  $\mathcal{A} \subset \mathcal{A}^\sim$ .
- pour tout  $A \in \mathcal{A}$  on a  $\bar{\mu}(A) = \mu(A)$ .
- $\bar{\mu}$  est une mesure complète sur  $(X, \mathcal{A}^\sim)$ .

*Démonstration. Admise.* La preuve n'est pas difficile, vous pouvez la faire en exercice. Il faut penser à tout démontrer :

- $\bar{\mu}$  est bien définie ( $\bar{\mu}(E)$  ne dépend pas du choix de  $A$  et  $B$ )
- $\mathcal{A}^\sim$  est une tribu contenant  $\mathcal{A}$ .

- $\bar{\mu}$  est une mesure et elle coïncide avec  $\mu$  sur  $\mathcal{A}$ .
- cette mesure est complète.

□

Si  $(X, \mathcal{A}, \mu)$  est un espace mesuré, cette proposition implique qu'on peut toujours compléter la tribu et étendre la mesure pour en faire une mesure complète.

Dans le cas de la mesure de Lebesgue sur  $\mathbb{R}^n$ , on a :

- $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \lambda_n)$  n'est pas complet : Il existe des parties de l'ensemble de Cantor qui ne sont pas boréliennes.
- $(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n), \lambda_n)$  est complet : c'est une conséquence des propriétés de la tribu  $\mathcal{R}(\lambda^*)$ .
- On a

$$\mathcal{B}(\mathbb{R}^n) \subsetneq \mathcal{L}(\mathbb{R}^n) \subsetneq \mathcal{P}(\mathbb{R}^n).$$

Ces deux inclusions sont strictes.

**Attention :** en particulier, la tribu  $\mathcal{L}(\mathbb{R}^n)$  n'est pas engendrée par les ouverts, ni par les pavés, ... Certains résultats démontrés sur les boréliens, ou les fonctions boréliennes, utilisent le fait que  $\mathcal{B}(\mathbb{R}^n)$  est engendré par une classe particulière. Ces résultats ne sont plus vrais si on remplace la tribu borélienne par la tribu de Lebesgue.

## 7.2 Classes monotones

Soit  $(X, \mathcal{A})$  un espace mesurable, pour montrer que la mesure de Lebesgue est caractérisée par sa valeur sur les pavés, on a besoin de se poser la question suivante : Si  $\mathcal{A} = \sigma(\mathcal{F})$  et si deux mesures  $\mu$  et  $\nu$  sont égales sur  $\mathcal{F}$ , sont-elles égales sur  $\mathcal{A}$  ? En général, c'est faux, il faut en dire un peu plus sur  $\mathcal{F}$ .

Le résultat central utilise un outil et un lemme.

**Définition 7.2.1.** Soit  $X$  un ensemble,  $\mathcal{M} \subset \mathcal{P}(X)$  est une **classe monotone** sur  $X$  si

1.  $X \in \mathcal{M}$
2. Si  $A, B \in \mathcal{M}$  et  $A \subset B$  alors  $B \setminus A \in \mathcal{M}$
3. Si  $(A_k)_{k \in \mathbb{N}}$  est une suite de  $\mathcal{M}$  telle que  $\forall k \in \mathbb{N} \ A_k \subset A_{k+1}$ , alors  $\bigcup_{k \in \mathbb{N}} A_k \in \mathcal{M}$ .

**Exemple 7.2.1.** Soit  $(X, \mathcal{A})$  un espace probabilisable. Si  $\mathbb{P}$  et  $\mathbb{Q}$  sont deux lois de probabilité sur  $(X, \mathcal{A})$ . Alors

$$\mathcal{F} = \{A \in \mathcal{A} \mid \mathbb{P}(A) = \mathbb{Q}(A)\}$$

est une classe monotone. Voir la preuve l'Étape 1. de la preuve du Théorème 7.2.1.

**Remarque 7.2.1.** • Si  $A \in \mathcal{M}$  alors  $A^c = X \setminus A \in \mathcal{M}$ . En particulier,  $\emptyset \in \mathcal{M}$ .

- Toute tribu sur  $X$  est une classe monotone.
- Une intersection de classes monotones est une classe monotone. Si  $\mathcal{F} \subset \mathcal{P}(X)$ , on note

$$\rho(\mathcal{F}) = \bigcap_{\substack{\mathcal{M} \text{ classe monotone} \\ \mathcal{F} \subset \mathcal{M}}} \mathcal{M}$$

la classe monotone engendrée par  $\mathcal{F}$ . Comme les tribus sont des classes monotones, on a en particulier

$$\mathcal{F} \subset \rho(\mathcal{F}) \subset \sigma(\mathcal{F}).$$

**Lemme 7.2.1.** Si  $\mathcal{M}$  est une classe monotone sur  $X$  et si  $\mathcal{M}$  est stable par intersection finie, alors  $\mathcal{M}$  est une tribu.

*Démonstration.* En effet toute classe monotone  $\mathcal{M}$  contient  $X$  et est stable par passage au complémentaire (puisque stable par différence et contenant  $X$ ). Si de plus  $\mathcal{M}$  est stable par intersection, les Lois de De Morgan entraîne qu'elle est stable par réunion :

$$\begin{aligned} \{A, B \in \mathcal{M}\} &\Rightarrow \{A^c, B^c \in \mathcal{M}\} \\ &\Rightarrow \{A^c \cap B^c \in \mathcal{M}\} \\ &\Rightarrow \{(A^c \cap B^c)^c (= A \cup B) \in \mathcal{M}\}. \end{aligned}$$

Par une récurrence évidente, la réunion d'une famille finie d'éléments de  $\mathcal{M}$  est encore un élément de  $\mathcal{M}$ . Soit alors une famille dénombrable d'éléments de  $\mathcal{M}$  notée  $(A_n)_{n \geq 1}$ . Posons  $B_n = \bigcup_{1 \leq k \leq n} A_k$ . Alors  $(B_n)_{n \geq 1}$  est une suite croissante d'éléments de  $\mathcal{M}$ , donc  $\bigcup_{n \geq 1} B_n \in \mathcal{M}$ , mais par ailleurs, on peut démontrer, par exemple par double inclusion, que  $\bigcup_{n \geq 1} B_n = \bigcup_{n \geq 1} A_n$ .  $\square$

**Lemme 7.2.2 Lemme des classes monotones.** Si  $\mathcal{F} \subset \mathcal{P}(X)$  est stable par intersections finies, alors  $\rho(\mathcal{F}) = \sigma(\mathcal{F})$ .

*Démonstration.* Il suffit de montrer que  $\rho(\mathcal{F})$  est stable par intersections finies d'après la Proposition 7.2.2. Soit  $A \in \mathcal{F}$ . Notons

$$\mathcal{N}_1 = \{B \in \rho(\mathcal{F}) \mid B \cap A \in \rho(\mathcal{F})\}.$$

Par hypothèse,  $\mathcal{N}_1$  contient  $\mathcal{F}$ . En outre, c'est une classe monotone. En effet,

- $X \in \mathcal{N}_1$ , puisque  $X \cap A = A \in \rho(\mathcal{F})$
- Si  $B_1, B_2 \in \mathcal{N}_1$ , avec  $B_1 \subset B_2$ , alors  $B_2 \setminus B_1 \in \rho(\mathcal{F})$  et

$$(B_2 \setminus B_1) \cap A = (B_2 \cap A) \setminus (B_1 \cap A) \in \rho(\mathcal{F})$$

Ainsi,  $B_2 \setminus B_1 \in \mathcal{N}_1$ .

- Si  $B_k \in \rho(\mathcal{F})$  pour tout  $k \in \mathbb{N}$ , avec  $B_k \subset B_{k+1}$ , alors  $B = \bigcup_{k \in \mathbb{N}} B_k \in \rho(\mathcal{F})$  et

$$B \cap A = \bigcup_{k \in \mathbb{N}} (B_k \cap A) \in \rho(\mathcal{F})$$

Ainsi,  $B \in \mathcal{N}_1$ .

La classe monotone  $\mathcal{N}_1$  contient donc  $\rho(\mathcal{F})$ , mais comme on a l'inclusion inverse par définition de  $\mathcal{N}_1$ , on déduit que  $\mathcal{N}_1 = \rho(\mathcal{F})$ .

On a montré que si  $A \in \mathcal{F}$  et  $B \in \rho(\mathcal{F})$  alors  $A \cap B \in \rho(\mathcal{F})$ . Soit maintenant  $B \in \rho(\mathcal{F})$ . Notons

$$\mathcal{N}_2 = \{A \in \rho(\mathcal{F}) \mid A \cap B \in \rho(\mathcal{F})\}.$$

Par la première étape, on a que  $\mathcal{N}_2$  contient  $\mathcal{F}$ . On vérifie comme précédemment que  $\mathcal{N}_2$  est une classe monotone, et on a donc  $\mathcal{N}_2 = \rho(\mathcal{F})$ . Ainsi, si  $A, B \in \rho(\mathcal{F})$  alors  $A \cap B \in \rho(\mathcal{F})$ .  $\square$

**Exemple 7.2.2.** L'intersection de deux pavés de  $\mathbb{R}^n$  est un pavé de  $\mathbb{R}^n$ . Si  $\mathcal{F} \subset \mathcal{P}(\mathbb{R}^n)$  est la classe des pavés, alors  $\rho(\mathcal{F}) = \sigma(\mathcal{F}) = \mathcal{B}(\mathbb{R}^n)$ .

**Théorème 7.2.1.** Soit  $(X, \mathcal{A})$  un espace mesurable, et  $\mu, \nu$  deux mesures sur  $(X, \mathcal{A})$ . S'il existe  $\mathcal{F} \subset \mathcal{P}(X)$  tel que

1.  $\mathcal{F}$  est stable par intersection finie, et  $\sigma(\mathcal{F}) = \mathcal{A}$ ,
2. pour tout  $A \in \mathcal{F}$  on a  $\mu(A) = \nu(A)$ ,
3. il existe une suite croissante  $(E_k)_{k \in \mathbb{N}}$  de  $\mathcal{F}$  telle que  $X = \bigcup_{k \in \mathbb{N}} E_k$  et  $\mu(E_k) = \nu(E_k) < +\infty$

pour tout  $k \in \mathbb{N}$ ,

alors  $\mu = \nu$ .

*Démonstration.* La preuve procède en deux étapes :

Étape 1. Supposons, dans un premier temps que  $\mu(X) < +\infty$ . Soit  $\mathcal{N} = \{A \in \mathcal{A} \mid \mu(A) = \nu(A)\} \subset \mathcal{A}$ . Par hypothèse,  $\mathcal{N} \supset \mathcal{F}$ . Montrons que  $\mathcal{N}$  est une classe monotone.

- $X \in \mathcal{N}$  par l'hypothèse 3.
- Si  $A, B \in \mathcal{N}$  avec  $A \subset B$ , alors

$$\mu(B \setminus A) \underset{\mu \text{ finie}}{=} \mu(B) - \mu(A) = \nu(B) - \nu(A) \underset{\nu \text{ finie}}{=} \nu(B \setminus A)$$

et  $B \setminus A \in \mathcal{N}$ .

- Si  $A_i \in \mathcal{N}$  pour tout  $i \in \mathbb{N}$  et que  $A_i \subset A_{i+1}$ , alors

$$\mu\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \lim_{i \rightarrow \infty} \mu(A_i) = \lim_{i \rightarrow \infty} \nu(A_i) = \nu\left(\bigcup_{i \in \mathbb{N}} A_i\right).$$

Cela implique que  $\bigcup_{i \in \mathbb{N}} A_i \in \mathcal{N}$ .

On a bien montré que  $\mathcal{N}$  est une classe monotone. Ainsi, par le Lemme des classes monotones,

$$\mathcal{N} \supset \mathcal{N}(\mathcal{F}) = \sigma(\mathcal{F}) = \mathcal{A} \supset \mathcal{N}$$

Donc  $\mathcal{N} = \mathcal{A}$  et les mesures  $\mu$  et  $\nu$  sont les mêmes.

Étape 2. On ne suppose plus que  $X$  soit de mesure finie. Pour tout  $k \in \mathbb{N}$ , on restreint  $\mu$  et  $\nu$  à  $E_k$  en posant :

$$\forall A \in \mathcal{A}, \quad \mu_k(A) = \mu(A \cap E_k) \quad \text{et} \quad \nu_k(A) = \nu(A \cap E_k).$$

Comme  $E_k \in \mathcal{F}$ , on a  $\mu_k(A) = \nu_k(A)$  pour tout  $A \in \mathcal{F}$ . En appliquant l'Étape 1., on trouve que

$$\mu_k(A) = \nu_k(A) \quad \forall A \in \mathcal{A}.$$

En prenant la limite quand  $k$  tend vers  $+\infty$ , on trouve

$$\mu(A) = \lim_{k \rightarrow +\infty} \mu_k(A) = \lim_{k \rightarrow +\infty} \nu_k(A) = \nu(A)$$

pour tout  $A \in \mathcal{A}$  et  $\nu = \mu$ . □

### 7.2.1 Caractérisation de la mesure de Lebesgue et autres conséquences

Les trois résultats suivants utilisent directement le Théorème 7.2.1 ou le Lemme 7.2.2.

La caractérisation de la mesure de Lebesgue est une conséquence directe du théorème 7.2.1 appliquée à la classe des pavés de  $\mathbb{R}^n$ .

**Théorème 7.2.2.** Soit  $\mu$  une mesure sur  $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ . Si, pour tout  $P = \prod_{k=1}^n [a_k, b_k]$  on a

$$\mu(P) = \prod_{k=1}^n (b_k - a_k), \text{ alors } \mu = \lambda_n.$$

*Démonstration.* Voir la Proposition 5.2.2. □

### Retour sur les mesures produit

Le théorème d'existence et de caractérisation de la mesure produit a été admis au chapitre 5, on le montre ici. Sa preuve utilise le théorème 7.2.1 pour la partie caractérisation, et le Lemme 7.2.2 pour montrer la mesurabilité des fonctions  $x \mapsto \nu(E_{x,\cdot})$  et  $y \mapsto \mu(E_{\cdot,y})$ .

**Théorème 7.2.3.** Soient  $(X, \mathcal{A}, \mu)$  et  $(Y, \mathcal{B}, \nu)$  deux espaces mesurés  $\sigma$ -finis.

1. il existe une unique mesure, notée  $\mu \otimes \nu$ , sur  $(X \times Y, \mathcal{A} \otimes \mathcal{B})$  telle que

$$\forall A \in \mathcal{A}, \forall B \in \mathcal{B}, \quad \mu \otimes \nu(A \times B) = \mu(A)\nu(B). \quad (7.2)$$

2. Pour tout  $E \in \mathcal{A} \otimes \mathcal{B}$ , les fonctions  $x \mapsto \nu(E_{x,\cdot})$  et  $y \mapsto \mu(E_{\cdot,y})$  sont mesurables et on a

$$\mu \otimes \nu(E) = \int_X \nu(E_{x,\cdot}) d\mu(x) = \int_Y \mu(E_{\cdot,y}) d\nu(y)$$

*Démonstration.* 1. **Unicité.** Par hypothèse, il existe  $(A_k)_{k \in \mathbb{N}}$  et  $(B_k)_{k \in \mathbb{N}}$  telles que  $X = \bigcup_{k \in \mathbb{N}} A_k$ ,  $Y = \bigcup_{k \in \mathbb{N}} B_k$  et pour tout  $k \in \mathbb{N}$ ,  $A_k \in \mathcal{A}$ ,  $B_k \in \mathcal{B}$ ,  $A_k \subset A_{k+1}$ ,  $B_k \subset B_{k+1}$ ,  $\mu(A_k) < +\infty$  et  $\nu(B_k) < +\infty$ . Alors,  $X \times Y = \bigcup_{k \in \mathbb{N}} E_k$  avec  $E_k = A_k \times B_k$  pour tout  $k \in \mathbb{N}$ . Soient  $m_1, m_2$  deux mesures sur  $(X \times Y, \mathcal{A} \otimes \mathcal{B})$  qui vérifient (7.2). Alors,

- $m_1$  et  $m_2$  coïncident sur la famille des rectangles mesurables, qui est stable par intersection finie et engendre la tribu  $\mathcal{A} \otimes \mathcal{B}$ .
- $m_1(E_k) = \mu(A_k)\nu(B_k) = m_2(E_k) < +\infty$  pour tout  $k \in \mathbb{N}$ .

Par le Théorème 7.2.1 on a  $m_1 = m_2$ . Notons cette mesure (unique)  $\mu \otimes \nu$ . Cet argument montre que toute mesure vérifiant (7.2) est  $\sigma$ -finie.

2. **Existence.** L'idée est de définir  $\mu \otimes \nu$  par la formule

$$\mu \otimes \nu(E) = \int_X \nu(E_x) d\mu_x \quad \text{pour } E \in \mathcal{A} \otimes \mathcal{B}.$$

De même, on aurait pu utiliser la formule  $\mu \otimes \nu(E) = \int_Y \mu(E^y) d\nu_y$  et l'unicité permet de voir qu'elle définit bien le même objet.

Remarquons que  $\nu(E_x)$  est bien définie pour tout  $x \in E$  d'après la Proposition 4.4.3. Pour vérifier que la formule (7.2) a bien un sens, il faut vérifier que pour tout  $E \in \mathcal{A} \otimes \mathcal{B}$ , l'application  $x \mapsto \nu(E_x)$  est  $\mathcal{A}$ -mesurable.

- (a) Supposons tout d'abord que  $\nu$  est une mesure finie. Notons

$$\mathcal{M} = \{E \in \mathcal{A} \otimes \mathcal{B} \mid x \mapsto \nu(E_x) \text{ est } \mathcal{A}\text{-mesurable}\} \subset \mathcal{A} \otimes \mathcal{B}.$$

Alors,

- $\mathcal{M}$  contient les rectangles mesurables : si  $E = A \times B$ ,  $A \in \mathcal{A}$ ,  $B \in \mathcal{B}$ , on a  $\nu(E_x) = \mathbf{1}_A(x)\nu(B)$ . De plus, la classe des rectangles mesurables est stable par intersections finies.
- on vérifie aisément que  $\mathcal{M}$  est une classe monotone :
  - (i)  $X \times Y \in \mathcal{M}$  car  $\nu((X \times Y)_x) = \mathbf{1}_X(x)\nu(Y) = \nu(Y)$ .
  - (ii) Si  $E, F \in \mathcal{M}$  et  $E \subset F$ , on a :

$$\nu((F \setminus E)_x) = \nu(F_x \setminus E_x) \underset{\nu \text{ finie}}{=} \nu(F_x) - \nu(E_x)$$

qui est une fonction mesurable, donc  $F \setminus E \in \mathcal{M}$ .

- (iii) Si  $(E_n)_{n \in \mathbb{N}}$  est une famille croissante d'éléments de  $\mathcal{M}$ , alors :

$$\nu\left(\left(\bigcup_{n \in \mathbb{N}} E_n\right)_x\right) = \nu\left(\bigcup_{n \in \mathbb{N}} (E_n)_x\right) \underset{\text{suite croissante}}{=} \lim_{n \rightarrow +\infty} \underbrace{\nu((E_n)_x)}_{\text{mesurable}}$$

est une limite de fonctions mesurables, donc est mesurable et  $\bigcup_{n \in \mathbb{N}} E_n \in \mathcal{M}$ .

Par le Lemme des classes monotones, on a  $\mathcal{M} = \mathcal{A} \otimes \mathcal{B}$ . Ainsi,  $x \mapsto \nu(E_x)$  est mesurable pour tout  $E \in \mathcal{A} \otimes \mathcal{B}$ .

- (b) Si  $\nu$  est seulement  $\sigma$ -finie, on écrit comme avant que  $Y = \bigcup_{n \in \mathbb{N}} B_n$ , et on applique le résultat de (a) à la mesure  $\nu_n$  définie par  $\nu_n(B) = \nu(B \cap B_n)$  pour tout  $B \in \mathcal{B}$ . Comme  $\nu(E_x) = \lim_{n \rightarrow +\infty} \nu_n(E_x)$ , on en déduit que  $x \mapsto \nu(E_x)$  est mesurable pour tout  $E \in \mathcal{A} \otimes \mathcal{B}$ .

On peut donc définir

$$\mu \otimes \nu(E) = \int_X \nu(E_x) d\mu_x \quad \forall E \in \mathcal{A} \otimes \mathcal{B}.$$

Reste à vérifier que  $\mu \otimes \nu$  est bien une mesure. Par construction,  $\mu \otimes \nu(\emptyset) = 0$ , et  $\mu \otimes \nu(A \times B) = \mu(A)\nu(B)$  pour tout  $A \in \mathcal{A}$  et  $B \in \mathcal{B}$ . Si  $(E_n)_{n \in \mathbb{N}}$  est une famille disjointe de  $\mathcal{A} \otimes \mathcal{B}$ , les  $\{(E_n)_x\}_{n \in \mathbb{N}}$  sont aussi disjoints pour tout  $x \in X$ , et on a

$$\begin{aligned} \mu \otimes \nu\left(\bigcup_{n \in \mathbb{N}} E_n\right) &= \int_X \nu\left(\bigcup_{n \in \mathbb{N}} (E_n)_x\right) d\mu_x \\ &= \int_X \sum_{n \in \mathbb{N}} \nu((E_n)_x) d\mu_x \\ &= \sum_{n \in \mathbb{N}} \int_X \nu((E_n)_x) d\mu_x \\ &= \sum_{n \in \mathbb{N}} \mu \otimes \nu(E_n) \end{aligned}$$

□

### Retour sur le théorème du changement de variable

Le résultat clé pour démontrer la formule du changement de variables dans  $\mathbb{R}^n$  est la proposition suivante (admise au chapitre 5) :

**Proposition 7.2.1.** Soient  $U, V$  deux ouverts de  $\mathbb{R}^n$ , et  $F : U \rightarrow V$  un difféomorphisme de classe  $C^1$ . Pour tout  $A \in \mathcal{B}(U)$  on a  $F(A) \in \mathcal{B}(V)$  et

$$\lambda_n(F(A)) = \int_A |\det(J_F(x))| d\lambda_n(x).$$

*Démonstration.* On en trouvera une preuve (très) complète dans [1]. Elle repose sur le Théorème 7.2.1 et consiste à montrer que les applications  $A \mapsto \lambda_n(F(A))$  et  $A \mapsto \int_A |\det(J_F(x))| d\lambda_n(x)$  sont des mesures sur  $(U, \mathcal{B}(U))$  qui coïncident sur les pavés de  $\mathbb{R}^n$  contenus dans  $U$ . La classe des pavés contenus dans  $U$  est stable par intersection finie et engendre  $\mathcal{B}(U)$ , donc le Théorème 7.2.1 permet de conclure que ces mesures sont égales sur  $\mathcal{B}(U)$ . □



# Bibliographie

- [1] F. De Marçay. Notes en théorie de la mesure et de l'intégration, Non datées ; <https://www.imo.universite-paris-saclay.fr/~joel.merker/>.
- [2] T. Gallay. Théorie de la mesure et de l'intégration, 2009 ; <https://www-fourier.ujf-grenoble.fr/~edumas/integration.pdf>.
- [3] T. Gallouët and R. Herbin. *Mesure, Integration, Probabilités*. Ellipses Edition Marketing, 2013, <https://hal.science/hal-01283567v2/file/totoro2.pdf>.
- [4] H.L. Royden. *Real Analysis*. Mathematics and statistics. Macmillan, 1988.