2009

Théorie de la mesure et de l'intégration.

Thierry Gallay

Transcrit par Tancrède Lepoint

2009

Université Joseph Fourier, Grenoble

TABLE DES MATIÈRES

AVANT-PROPOS

Ce polycopié est le support du cours de *Théorie de la mesure et de l'intégration* enseigné à l'université Joseph Fourier de Grenoble en troisième année de licence de mathématiques fondamentales par Thierry Gallay [1]. Il a été transcrit tout au long de l'année et ne saurait en aucun cas remplacer le cours.

Ce document est très proche du cours enseigné, et excepté quelques infimes modifications (et l'annexe), il retranscrit le cours tel qu'il a été donné à tous les étudiants. En conséquence de quoi, il n'est pas un approfondissement du cours, au contraire des livres disponibles dans la bibliographie. L'annexe présente la mesure de Hausdorff, née une quinzaine d'année après celle de Lebesgue, qui permet notamment la mesure d'objets de dimension inférieures, et n'a pas été traitée en cours. Elle nécessite de connaître les chapitres 1 et 3.

Pour toute remarque, suggestion ou correction concernant ce document, merci de me contacter pour que je puisse modifier et corriger ce polycopié.

Tancrède Lepoint.
http ://www.kilomaths.com/ tanc/
Tancrede.Lepoint@e.ujf-grenoble.fr

BIBLIOGRAPHIE SOMMAIRE

N. Bourbaki, *Éléments de mathématiques, livre VI : Intégration*, Chapitres 1-9, Hermann, Paris, 1952-1969.
M. Briane et G. Pagès, *Théorie de l'intégration*, Vuibert, Paris, 2000.
D. L. Cohn, *Measure Theory*, Birkhäuser, Boston, 1980.
J. L. Doob, *Measure Theory*, Graduate Texts in Mathematics **143**, Springer, New-York, 1994.
R. M. Dudley, /em Real analysis and probability, Cambridge Studies in Advanced Mathematics **74**, Cambridge University Press, 2002.
P. R. Halmos, *Measure Theory*, Graduate Texts in Mathematics **18**, Springer, 1974.
E. H. Lieb et M. Loss, *Analysis*, Graduate Studies in Mathematics **14**, AMS, Providence, 1997.
W. Rudin, *Analyse réelle et complexe*, Masson, Paris, 1980.
W. Rudin, Real and complex analysis (3ème éd.), McGraw-Hill, New York, 1987.
D. W. Stroock, *A concise introduction to the theory of integration* (3ème éd.), Birkhäuser, Boston, 1999.
J. Yeh, *Real analysis. Theory of measure and integration* (2ème éd.), World Scientific, Hackensack, 2006.

1. Thierry Gallay - *http ://www-fourier.ujf-grenoble.fr/ gallay/ - Thierry.Gallay@ujf-grenoble.fr*

INTRODUCTION

Une **théorie de l'intégration** est un procédé qui associe à toute fonction f (dans une certaine classe) un nombre $I(f)$, appelé **intégrale de** f et qui vérifie certaines propriétés (linéarité, positivité, ...).

Exemple (fondamental). On retrouve pour la première fois l'exemple suivant (actuellement connu comme l'intégrale de Riemann) dans le cours de Cauchy en 1820.
Soit $\mathcal{C}^0([a,b],\mathbb{R})$ l'espace des fonctions continues sur un intervalle $[a,b]$ à valeurs dans $\mathbb{R}$. Pour tout $f \in \mathcal{C}^0([a,b],\mathbb{R})$, la limite suivante existe :

$$I(f) = \lim_{N\to+\infty} \frac{b-a}{N} \sum_{i=0}^{N-1} f(a + i\frac{b-a}{N}) \tag{1}$$

La correspondance $f \mapsto I(f)$ est :
- **linéraire** :
  - $I(f_1+f_2) = I(f_1) + I(f_2), \forall f_1, f_2 \in \mathcal{C}^0([a,b],\mathbb{R})$
  - $I(\lambda f) = \lambda I(f), \forall f \in \mathcal{C}^0([a,b],\mathbb{R}), \forall \lambda \in \mathbb{R}$
- **positive** : Si $f \geqslant 0$, alors $I(f) \geqslant 0$. Dans ce cas, $I(f)$ a une interprétation graphique (figure 1) : c'est l'aire sous le graphe de f.

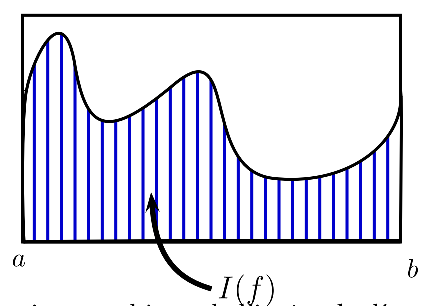


FIGURE 1 – Interprétation graphique de l'intégrale d'une fonction positive

Remarque. Il n'est pas nécessaire d'utiliser une subdivision régulière pour calculer $I(f)$.
Pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que pour toute subdivision $a = x_0 < x_1 < \cdots < x_N = b$ de l'intervalle $[a,b]$ tel que $\max_{i=1,\ldots,N}(x_i - x_{i-1}) \leqslant \delta$ et pour tous points $\xi_i \in [x_{i-1}, x_i]$, $i = 1,\ldots,N$ on a :

$$\left| I(f) - \sum_{i=1}^{N} f(\xi_i)(x_i - x_{i-1}) \right| \leqslant \varepsilon \tag{2}$$

Mais pourquoi appelle-t-on cela l'intégrale de Riemann ? Riemann s'est demandé pour quelles classes de fonctions les procédés (1) et (2) permettent de définir l'intégrale. Est-il nécessaire de se limiter aux fonctions continues ? Il a remarqué en 1854 que l'on pouvait utiliser ces procédés pour une certaine classe de fonctions non continues. Les fonctions $f\colon\ [a,b] \to \mathbb{R}$ pour lesquelles on peut définir $I(f)$ par (1) et (2) sont appelées des fonctions **intégrables au sens de Riemann**. L'intégrale $I(f)$ est souvent notée $I(f) = \displaystyle\int_a^b f(x)\mathrm{d}x$.

Limitations de l'intégrale de Riemann.

– Toute fonction $f\colon\ [a,b] \to \mathbb{R}$ intégrable au sens de Riemann est *bornée*.
  En pratique, ce n'est pas très gênant, on peut généraliser un peu le procédé pour intégrer certaines fonctions non bornées.

  Exemple. $\displaystyle\int_0^1 \frac{1}{\sqrt{x}}\mathrm{d}x$ peut être définie comme $\displaystyle\lim_{\varepsilon\to 0}\int_\varepsilon^1 \frac{1}{\sqrt{x}}\mathrm{d}x = 2$

– Soit $f\colon\ [0,1] \to \mathbb{R}$ définie par $f(x) = \left\{\begin{array}{ll} 1 & \text{si } x \in \mathbb{Q} \\ 0 & \text{sinon} \end{array}\right.$. C'est la fonction indicatrice des rationnels, et f n'est *pas intégrable* au sens de Riemann. En effet, dans le procédé (2) on peut choisir tous les ξ_i rationnels (respectivement irrationnels) et on en déduit que $I(f)$ vaudrait 1 (resp. 0)...
  Henri Lebesgue (mathématicien français du début du XX$^{\text{ème}}$ siècle) a montré qu'une fonction bornée $f\colon\ [a,b] \to \mathbb{R}$ est intégrable au sens de Riemann si et seulement si l'ensemble de ses points de discontinuité est négligeable.
  Un ensemble $A \subset \mathbb{R}$ est dit négligeable si $\forall\varepsilon > 0$, A peut être recouvert par une union dénombrable d'intervalles dont la somme des longueurs est $\leqslant \varepsilon$.

– *Intégrales et primitives.* Soit $F \in \mathcal{C}^0([a,b]\,,\mathbb{R})$. Si F est dérivable sur $]a,b[$ et si $f = F'$, a-t-on nécessairement le théorème fondamental du calcul intégral $\displaystyle\int_a^b f(x)\mathrm{d}x = F(b) - F(a)$? La réponse est non, il n'y a aucune raison que f soit intégrable au sens de Riemann.

– *Limites simples.* Soit $f_n\colon\ [a,b] \to \mathbb{R}$ une suite de fonctions intégrables au sens de Riemann. On suppose que $\forall x \in [a,b]\,,\ \displaystyle\lim_{n\to+\infty} f_n(x) = f(x)$.

  A-t-on que $\displaystyle\int_a^b f(x)\mathrm{d}x = \lim_{n\to+\infty}\int_a^b f_n(x)\mathrm{d}x$? Si il y a convergence uniforme, on sait que cela est vrai, sinon un contre exemple est facile à trouver. En fait, la question est mal posée : même si l'on suppose que $|f_n(x)| \leqslant M, \forall n \in \mathbb{N}, \forall x \in [a,b]$, la limite f n'est en général pas intégrable au sens de Riemann.

  Remarque. Par contre, on peut montrer que, sous ces hypothèses, la limite des intégrales est toujours définie.

Remarque. Tout cet exemple fondamental se généralise au cas des fonctions $f\colon \mathbb{R}^d \to \mathbb{R}$. Il faut remplacer les intervalles $[a,b]$ par des **pavés** de la forme $\Pi = [a_1,b_1] \times \cdots \times [a_d,b_d]$

Une **théorie de la mesure** est un procédé qui associe à tout ensemble A (dans une certaine classe) un nombre positif $\mu(A)$, appelé **mesure de** A, et qui vérifie certaines propriétés (monotonie, additivité, ...). En dimension 1, la mesure correspond à la longueur, à l'aire en dimension 2 et au volume au dimension 3, d'où la généralisation.
Intégration et mesure sont étroitement liées. Si $A \subset \mathbb{R}^d$, on définit **la fonction indicatrice** de A par $\mathbb{1}_A(x) = \left\{\begin{array}{ll} 1 & \text{si } x \in A \\ 0 & \text{sinon} \end{array}\right.$, $\mathbb{1}_A : \mathbb{R}^d \to \mathbb{R}$.
Si $\mathbb{1}_A$ est intégrable (pour un certain procédé d'intégration), on peut définir la mesure de A par $\mu(A) = \displaystyle\int \mathbb{1}_A(x)\mathrm{d}x$. On obtient ainsi une application μ qui vérifie les propriétés suivantes :

– **monotonie** : Si $A \subset B$, alors $\mu(A) \leqslant \mu(B)$ (car $\mathbb{1}_A \leqslant \mathbb{1}_B$).
– **additivité** : Si $A \cap B = \varnothing$ alors $\mu(A \cup B) = \mu(A) + \mu(B)$ (car $\mathbb{1}_{A\cup B} = \mathbb{1}_A + \mathbb{1}_B$).

Remarque. **Quels sont les sous-ensembles de $\mathbb{R}^d$ qui sont intégrables au sens de Riemann ?** On dit que $A \subset \mathbb{R}^d$ est **quarrable** si $\mathbb{1}_A$ est intégrable au sens de Riemann. On peut montrer que $A \subset \mathbb{R}^d$ est quarrable si et seulement si A est bornée et si ∂A est négligeable.
Attention ! Un espace compact de $\mathbb{R}^d$ n'est pas nécessairement quarrable.

Mais alors, peut-on définir une mesure sur tout sous-ensemble de $\mathbb{R}^d$? On cherche à construire une application $\mu : \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$ qui vérifie :
– μ est additive
– μ est invariante par translation et rotation
– $\mu([0,1]^d) = 1$.
Attention ! Une telle mesure n'existe pas si $d \geqslant 3$. Une démonstration de ceci utilise l'axiome du choix. C'est **le paradoxe de Banach-Tarski** (1923) : *Il est possible de découper la boule unité de $\mathbb{R}^3$ en un nombre fini de morceaux et de les réarranger (après translations et rotations) de façon à obtenir deux copies de la boule unité.*
Moralité : On ne peut pas mesurer tous les sous-ensembles de $\mathbb{R}^d$. Il faut se restreindre à une sous-famille $\mathcal{M}$ qui possède au moins les propriétés suivantes :
– $\varnothing \in \mathcal{M}$
– Si $A, B \in \mathcal{M}$, alors $A \cup B \in \mathcal{M}$ et $A \cap B \in \mathcal{M}$
– Si $A, B \in \mathcal{M}$ alors $A \setminus B \in \mathcal{M}$
$\mathcal{M}$ est un **anneau booléen**. Par exemple, les ensembles quarrables forment un anneau.

Si on a une théorie de la mesure, on peut en déduire une théorie de l'intégration.

Idée de la construction de l'intégrale à partir de la mesure (Henri Lebesgue, 1901).
Soit $f : \mathbb{R}^d \to [0,1]$. Contrairement à la théorie de l'intégration de Riemann qui subdivise l'espace de départ de la fonction, ici on utilise une subdivision régulière de l'espace d'arrivée de la fonction $[0,1]$. On peut approcher f par une fonction *étagée* de la forme

$$f_N = \sum_{i=1}^{N} \frac{i}{N} \mathbb{1}_{\{x / \frac{i-1}{N} < f(x) \leqslant \frac{i}{N}\}} \geqslant f$$

Remarque. On peut aussi écrire $f_N = \displaystyle\sum_{i=1}^{N} \frac{1}{N} \mathbb{1}_{\{x/f(x) > \frac{i-1}{N}\}}$.

Pour chaque N, on définit

$$\begin{aligned}\int f_N \mathrm{d}\mu &= \sum_{i=1}^{N} \frac{i}{N} \mu(\{x / \frac{i-1}{N} < f(x) \leqslant \frac{i}{N}\}) \\ &= \sum_{i=1}^{N} \frac{1}{N} \mu(\{x / f(x) > \frac{i-1}{N}\})\end{aligned}$$

En passant à la limite quand N tend vers $+\infty$, on trouve formellement

$$\int f \mathrm{d}\mu = \underbrace{\int_0^1 \mu(\{x / f(x) \geqslant t\}) \mathrm{d}t}_{\text{Intégrale de Riemann}}$$

x

★

CHAPITRE 1

THÉORIE GÉNÉRALE DE LA MESURE

1.1 ESPACES MESURABLES

Définition 1.1. Soit X un ensemble. On appelle **tribu** ou **σ-algèbre** sur X une famille $\mathcal{M}$ de parties de X possédant les propriétés suivantes :

i) $X \in \mathcal{M}$

ii) Si $A \in \mathcal{M}$, alors $A^{\complement} \in \mathcal{M}$ (où $A^{\complement} = X \setminus A$ est le complémentaire de A dans X)

iii) Si $A_n \in \mathcal{M}, \forall n \in \mathbb{N}$, alors $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{M}$

Les éléments de $\mathcal{M}$ sont appelés les **parties mesurables** de X. On dit que $(X, \mathcal{M})$ est un **espace mesurable.**

Remarque. Si au lieu de iii) on demande seulement

iii') Si $A, B \in \mathcal{M}$, alors $A \cup B \in \mathcal{M}$

c'est-à-dire la stabilité de $\mathcal{M}$ par intersection finie, alors on obtient un **anneau booléen**.

Conséquences: – *$\varnothing \in \mathcal{M}$ car $X \in \mathcal{M}$*

– *Si $A_n \in \mathcal{M}, \forall n \in \mathbb{N}$ alors $\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{M}$ (car $\left(\bigcap_{n\in\mathbb{N}} A_n\right)^{\complement} = \bigcup_{n\in\mathbb{N}} A_n^{\complement}$)*

– *M est stable par intersection ou union finie.*

– *Si A et B sont mesurables, alors la différence non symétrique* [1] $A \setminus B = A \cap B^{\complement} \in \mathcal{M}$

Evidemment, tout ensemble X possède des tribus, par exemple :

– $\mathcal{M} = \{\varnothing, X\}$ la plus petite

– $\mathcal{M} = \mathcal{P}(X)$ la plus grande

Pour construire des tribus "intéressantes" sur X, on utilise souvent le résultat suivant :

Lemme 1.2. *Soit $\{\mathcal{M}_i\}_{i\in I}$ une famille quelconque de tribus sur X. Alors $\mathcal{M} = \bigcap_{i\in I} \mathcal{M}_i$ est encore une tribu sur X.*

Démonstration. La vérification est immédiate. □

1. Dans ce cours, on utilise la différence non symétrique. On peut aussi définir la différence symétrique

$$A \vartriangle B = (A \setminus B) \cup (B \setminus A) = (A \cap B)^{\complement} \cap (A \cup B)$$

mais elle ne sera pas utilisée ici.

Définition 1.3. Soit F une famille de parties de X. On note

$$\sigma(F) = \bigcap_{\mathcal{M} \text{ tribu sur } X, \mathcal{M} \supset F} \mathcal{M}$$

Alors, $\sigma(F)$ est une tribu sur X appelée **tribu engendrée par** F. C'est la plus petite tribu sur X qui contient F.

Définition 1.4 (Rappel). Une **topologie** sur X est une famille $\mathcal{T}$ de parties de X telles que :
– $\varnothing \in \mathcal{T}, X \in \mathcal{T}$
– Si $O_1, \ldots, O_n \in \mathcal{T}$, alors $\bigcap_{i=1}^n O_i \in \mathcal{T}$
– Si $\{O_i\}_{i \in I}$ est une famille quelconque d'éléments de $\mathcal{T}$ alors $\bigcup_{i \in I} O_i \in \mathcal{T}$
Les éléments de $\mathcal{T}$ s'appellent **les ouverts** de X. On dit que $(X, \mathcal{T})$ est un **espace topologique**.

Définition 1.5 (Tribu de Borel). Soit $(X, \mathcal{T})$ un espace topologique. On appelle **tribu de Borel sur** X la tribu engendrée par les ouverts de X : $\mathcal{M} = \sigma(\mathcal{T})$.

Considérons plus en détail le cas de la tribu de Borel sur $\mathbb{R}$ notée $\mathcal{B}(\mathbb{R})$
– $\mathcal{B}(\mathbb{R})$ contient tous les ouverts et tous les fermés de $\mathbb{R}$
– $\mathcal{B}(\mathbb{R})$ contient les unions dénombrables de fermés (ensembles F_σ)
– $\mathcal{B}(\mathbb{R})$ contient les intersections dénombrables d'ouverts (ensembles G_δ)
– ... et bien plus encore.
On peut montrer que la tribu $\mathcal{B}(\mathbb{R})$ a la puissance du continu. En conséquence, $\mathcal{B}(\mathbb{R}) \neq \mathcal{P}(\mathbb{R})$.

Proposition 1.6. *La tribu $\mathcal{B}(\mathbb{R})$ est engendrée par les intervalles $]a, +\infty[$ pour $a \in \mathbb{R}$*

Démonstration. Soit $\mathcal{M}$ la tribu engendrée par les intervalles $]a, +\infty[$ où $a \in \mathbb{R}$. Par construction, $\mathcal{M} \subset \mathcal{B}(\mathbb{R})$.
D'autre part, $\forall a \in \mathbb{R}$ on a $[a, +\infty[= \bigcap_{n \in \mathbb{N}} \left] a - \frac{1}{n}, +\infty \right[\in \mathcal{M}$. Par complémentaire, $]-\infty, a[= [a, +\infty[^{\complement} \in \mathcal{M}$. Par intersection, si $a < b$, $]a, b[=]-\infty, b[\cup]a, +\infty[\in \mathcal{M}$.
On sait que tout ouvert de $\mathbb{R}$ est une réunion au plus dénombrable d'intervalles de la forme $]-\infty, a[$, $]a, b[$, $]a, +\infty[$, donc $\mathcal{M}$ contient tous les ouverts de $\mathbb{R}$ et $\mathcal{M} \supset \mathcal{B}(\mathbb{R})$. $\square$

Remarque. $\mathcal{B}(\mathbb{R})$ est engendré par les intervalles $]a, +\infty[$, $a \in \mathbb{Q}$. En effet, pour tout $a \in \mathbb{R}$, il existe une suite de rationnels $(a_n)_{n \in \mathbb{N}}$ décroissant vers a et $]a, +\infty[= \bigcup_{n=1}^{\infty}]a_n, +\infty[$

Exercice (Pavés dyadiques dans $\mathbb{R}^d$). Dans $\mathbb{R}^d$, on note $\forall k \in \mathbb{N}, P_k$ l'ensemble des pavés de la forme $\left[0, 2^{-k}\right]^d + n2^{-k}, \forall n \in \mathbb{Z}^d$. On note $P = \bigcup_{k=0}^{\infty} P_k$.

a. Montrer que tout ouvert de $\mathbb{R}^d$ est une réunion dénombrable de pavés dyadiques.
b. En déduire que la tribu de Borel $\mathcal{B}(\mathbb{R}^d)$ est engendrée par P.

1.2 Définition et exemples de mesures

Définition 1.7. Soit (X, M) un espace mesurable.
On appelle **mesure positive sur** X une application $\mu : \mathcal{M} \to [0, +\infty]$ vérifiant :

1. $\mu(\varnothing) = 0$
2. **Additivité dénombrable** : si $\{A_n\}_{n\in\mathbb{N}}$ est une famille dénombrable d'ensembles mesurables deux à deux disjoints alors
$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n)$$
On dit que $(X, \mathcal{M}, \mu)$ est **un espace mesuré**.

Commentaires. – On dira souvent "mesure" au lieu de "mesure positive"
– La condition $\mu(\varnothing) = 0$ est nécessaire pour éviter des situations triviales.
En effet, $\forall A \in \mathcal{M}$, on a $A = A \cup \varnothing \cup \varnothing \cup \cdots$, donc $\mu(A) = \mu(A) + \sum_{n\in\mathbb{N}} \mu(\varnothing)$.

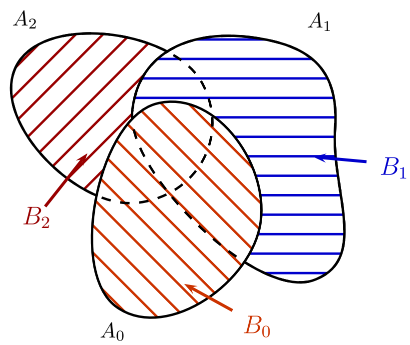
Proposition 1.8 (propriétés élémentaires d'une mesure positive).

1. ***Monotonie*** *: Si $A, B \in \mathcal{M}$ et $A \subset B$, alors $\mu(A) \leqslant \mu(B)$*
2. ***Sous-additivité*** *: Si $A_n \in \mathcal{M}, \forall n \in \mathbb{N}$ alors*
$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leqslant \sum_{n\in\mathbb{N}} \mu(A_n)$$
3. *Si $A_n \in \mathcal{M}, \forall n \in \mathbb{N}$ et si $A_n \subset A_{n+1}, \forall n \in \mathbb{N}$, alors*
$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \mu(A_n)$$
4. *Si $A_n \in \mathcal{M}, \forall n \in \mathbb{N}$ et si $A_n \supset A_{n+1}, \forall n \in \mathbb{N}$, $\underline{\text{avec } \mu(A_0) < \infty}$ alors*
$$\mu(\bigcap_{n\in\mathbb{N}} A_n) = \lim_{n\to\infty} \mu(A_n)$$

Démonstration. 1. On a $B = A \cup (B \setminus A)$, union disjointe d'éléments de $\mathcal{M}$ donc $\mu(B) = \mu(A) + \mu(B \setminus A) \geqslant \mu(A)$.
Si $\mu(A) < \infty$, on déduit que $\mu(B \setminus A) = \mu(B) - \mu(A)$

2. Posons $B_0 = A_0$, et $\forall n \geqslant 1, B_n = A_n \setminus \bigcup_{k=0}^{n-1} A_k$. Pour tout $n \in \mathbb{N}$, on a $\displaystyle\bigcup_{k=0}^{n} A_k = \bigcup_{k=0}^{n} B_k$. En outre, les $\{B_n\}_{n\in\mathbb{N}}$ sont deux à deux disjoints et $\mu(B_n) \leqslant \mu(A_n)$, alors
$$\mu(\bigcup_{n\in\mathbb{N}} A_n) = \mu\left(\bigcup_{n\in\mathbb{N}} B_n\right) = \sum_{n\in\mathbb{N}} \mu(B_n) \leqslant \sum_{n\in\mathbb{N}} \mu(A_n)$$

3. Posons $B_0 = A_0$ et $\forall n \geqslant 1, B_n = A_n \setminus A_{n-1}$. Alors les B_n sont 2 à 2 disjoints et $\forall n \in \mathbb{N}, A_n = \displaystyle\bigcup_{k=0}^{n} B_k$.



Ainsi,

$$\mu(A_n) = \sum_{k=0}^{n} \mu(B_k) \underset{n\to+\infty}{\longrightarrow} \sum_{k\in\mathbb{N}} \mu(B_k) = \mu\left(\bigcup_{k\in\mathbb{N}} B_k\right) = \mu\left(\bigcup_{n\in\mathbb{N}} A_n\right)$$

4. Posons $B_n = A_0 \setminus A_n, \forall n \in \mathbb{N}$. Alors la suite $\{B_n\}_{n\in\mathbb{N}}$ est croissante avec $\bigcup_{n\in\mathbb{N}} B_n = A_0 \setminus \bigcap_{n\in\mathbb{N}} A_n$.
En outre, $\mu(B_n) = \mu(A_0) - \mu(A_n)$ car $\mu(A_0) < \infty$.
Par 3), on a donc

$$\mu(A_0) - \mu(\bigcap_{n\in\mathbb{N}} A_n) = \mu\left(\bigcup_{n\in\mathbb{N}} B_n\right) = \lim_{n\to\infty} \mu(B_n) = \mu(A_0) - \lim_{n\to\infty} \mu(A_n)$$

□

Exemples (Exemples élémentaires de mesures).

1. **Mesure de comptage** sur un ensemble X
Sur $(X, \mathcal{P}(X))$, on définit la mesure de comptage $\mu(A), A \subset X$ par

$$\mu(A) = \begin{cases} \mathrm{card}(A) & \text{si } A \text{ fini} \\ \infty & \text{sinon} \end{cases}$$

Cette mesure est surtout utilisée sur des ensembles "discrets" $(\mathbb{N}, \mathbb{Z}, \mathbb{Z}^d, \ldots)$

2. **Mesure de Dirac** en un point
Soit $(X, \mathcal{M})$ un ensemble mesurable, avec $X \neq \varnothing$ et soit $x \in X$. On définit $\mu\colon \mathcal{M} \to [0, +\infty]$ par

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

On note souvent $\mu = \delta_x$

Remarque. La condition $\mu(A_0) < \infty$ du (4) de la proposition 1.8 est nécessaire. En effet, considérons $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ muni de la mesure de comptage et considérons $A_n = \{n, n+1, n+2, \ldots\}$, alors $A_n \supset A_{n+1}$ et $\bigcap_{n\in\mathbb{N}} A_n = \varnothing$, mais $\forall n \in \mathbb{N}, \mu(A_n) = +\infty$.

Théorème 1.9 (Mesure de Lebesgue sur $\mathbb{R}$). *Il existe une unique mesure positive sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, notée λ, telle que*

$$\lambda(]a, b[) = b - a, \forall a, b \in \mathbb{R}, a < b$$

Remarque. La mesure de Lebesgue est **diffuse** : $\lambda(\{x\}) = 0, \forall x \in \mathbb{R}$.
En effet, $\{x\} = \bigcap_{n=1}^{\infty}]x - \frac{1}{n}, x + \frac{1}{n}[$, donc, par la proposition 1.8, on a :

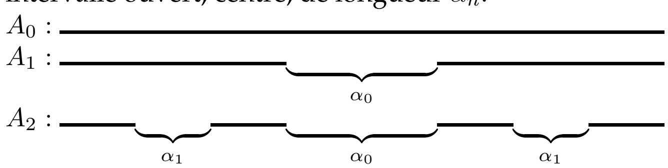
$$\lambda(\{x\}) = \lim_{n\to+\infty} \lambda(]x - \frac{1}{n}, x + \frac{1}{n}[) = \lim_{n\to+\infty} \frac{2}{n} = 0$$

On dit aussi qu'elle **ne charge pas les points** (au contraire de la mesure de Dirac).

Conséquences: – $\lambda(]a,b[) = \lambda([a,b[) = \lambda(]a,b]) = \lambda([a,b]) = b - a$ *si* $a \leqslant b$
– *Les ensembles dénombrables sont de mesure nulle.*

1.3 Exemple : l'ensemble de Cantor

Soit $(\alpha_n)_{n\in\mathbb{N}}$ une suite de réels > 0 tels que $\sum_{n=0}^{\infty} 2^n \alpha_n \leqslant 1$. On construit une suite décroissante de compacts inclus dans $[0,1]$ de la façon suivante :
– $A_0 = [0,1]$
– $A_1 = \left[0, \frac{1-\alpha_0}{2}\right] \cup \left[\frac{1+\alpha_0}{2}, 1\right] \ldots$
Pour tout $n \in \mathbb{N}$, A_{n+1} s'obtient de A_n en retranchant de chacun des 2^n intervalles le composant un intervalle ouvert, centré, de longueur α_n.



Par construction, $A_{n+1} \subset A_n$ et A_n est constitué de 2^n intervalles compacts disjoints de longueurs égales.
En outre, $\lambda(A_n) = 1 - \sum_{k=0}^{n-1} 2^k \alpha_k > 0$. On définit

$$K = \bigcap_{n\in\mathbb{N}} A_n$$

Alors K est un compact non vide (car intersection d'une suite décroissante de compacts non vides). On a $\mathring{K} = \varnothing$. En effet, $\forall n \in \mathbb{N}, K \subset A_n$ et A_n ne contient aucun intervalle de longueur supérieure à 2^{-n}. On en déduit que $K = \partial K$ (où ∂K est la frontière de K). De plus, on a

$$\lambda(K) = \lim_{n\to+\infty} \lambda(A_n) = 1 - \sum_{n=0}^{\infty} 2^n \alpha_n \geqslant 0$$

Si l'on choisit les (α_n) tels que $\sum_{n=0}^{\infty} 2^n \alpha_n = 1$, alors c'est un **Cantor maigre** (de mesure nulle).
Si l'on choisit les (α_n) tels que $\sum_{n=0}^{\infty} 2^n \alpha_n < 1$, alors c'est un **Cantor gras** (de mesure strictement positive).
Enfin, K est équipotent à $2^{\mathbb{N}}$ (chaque point du Cantor est définit de manière unique avec une suite de 0 et de 1 selon qu'il se trouve dans l'intervalle de gauche ou de droite à chaque nouveau découpage du Cantor), et a donc la puissance du continu.

Remarque. K contient nécessairement des sous-ensembles *non boréliens*, même lorsque $\lambda(K) = 0$. En particulier, il existe des sous-ensembles d'ensemble de mesure nulle qui ne sont pas mesurables !

1.4 Complétion des mesures

Définition 1.10. Soit $(X, \mathcal{M}, \mu)$ un espace mesuré.

1. On dit que $A \subset X$ est **négligeable** (pour la mesure μ) si $A \in \mathcal{M}$ et $\mu(A) = 0$.
2. On dit que la mesure μ **est complète** si tout sous-ensemble d'un ensemble négligeable est encore négligeable.

Remarque. On a vu qu'un Cantor de mesure nulle contenait des sous-ensembles non mesurables, donc non négligeables. On en déduit donc que la mesure de Lebesgue λ sur $\mathcal{B}(\mathbb{R})$ n'est pas complète !

Proposition 1.11 (complétion des mesures). *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré. Soit $\mathcal{M}^*$ l'ensemble de toutes les parties E de X telles qu'il existe $A, B \in \mathcal{M}$ avec $A \subset E \subset B$ et $\mu(B \setminus A) = 0$. On définit alors $\mu^*(E) = \mu(A)$. Ainsi, $\mathcal{M}^*$ est une tribu sur X et μ^* une mesure complète sur $\mathcal{M}$ qui prolonge μ.*

Remarque. Si $E \in \mathcal{M}$, et comme $E \subset E \subset E$, alors $E \in \mathcal{M}^*$ et $\mu^*(E) = \mu(E)$.

Démonstration. – Il faut vérifier que μ^* est bien définie.

Supposons que $\begin{cases} A_1 \subset E \subset B_1, & \mu(B_1 \setminus A_1) = 0 \\ A_2 \subset E \subset B_2, & \mu(B_2 \setminus A_2) = 0 \end{cases}$, $A_i, B_i \in \mathcal{M}, i = 1, 2$. Alors $A_1 \subset B_2$, donc $\mu(A_1) \leqslant \mu(B_2) = \mu(A_2)$. En répétant l'argument dans l'autre sens, on montre aussi que $\mu(A_2) \leqslant \mu(A_1)$, donc μ^* est bien définie.

– Il faut vérifier que $\mathcal{M}^*$ est une tribu sur X.

  – $X \in \mathcal{M}^*$
  – Si $A \subset E \subset B$, par passage au complémentaire on a $B^{\complement} \subset E^{\complement} \subset A^{\complement}$ avec $\mu(\overbrace{A^{\complement} \setminus B^{\complement}}^{=B\setminus A}) = \mu(B \setminus A) = 0$. On a donc $E \in \mathcal{M}^* \implies E^{\complement} \in \mathcal{M}^*$
  – Si $A_n \subset E_n \subset B_n$ avec $\mu(B_n \setminus A_n) = 0$ alors $\underbrace{\bigcup_{n\in\mathbb{N}} A_n}_{A} \subset \underbrace{\bigcup_{n\in\mathbb{N}} E_n}_{E} \subset \underbrace{\bigcup_{n\in\mathbb{N}} B_n}_{B}$.

    Or $B \setminus A = \bigcup_{n\in\mathbb{N}} B_n \setminus A \subset \bigcup_{n\in\mathbb{N}} B_n \setminus A_n$ donc $\mu(B \setminus A) \leqslant \sum_{n\in\mathbb{N}} \mu(B_n \setminus A_n) = 0$.

    Ainsi, $E_n \in \mathcal{M}^*, \forall n \in \mathbb{N} \implies \bigcup_{n\in\mathbb{N}} E_n \in \mathcal{M}^*$

  On en déduit que $\mathcal{M}^*$ est une tribu sur X.

– Il reste à vérifier que μ^* est bien une mesure. Clairement, $\mu^*(\varnothing) = \mu(\varnothing) = 0$ car $\varnothing \in \mathcal{M}$. Par ailleurs, si $\{E_n\}_{n\in\mathbb{N}}$ est une famille disjointe dans $\mathcal{M}^*$, alors dans l'étape précédente les A_n le sont aussi et on a donc :

$$\mu^*(\bigcup_{n\in\mathbb{N}} E_n) = \mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n) = \sum_{n\in\mathbb{N}} \mu^*(E_n)$$

□

Définition 1.12. On appelle **tribu de Lebesgue** sur $\mathbb{R}$, et on note $\mathcal{L}(\mathbb{R})$, la tribu qui complète la tribu de Borel $\mathcal{B}(\mathbb{R})$ pour la mesure de Lebesgue λ.
On appelle encore **mesure de Lebesgue** la mesure complétée $\lambda\colon\ \mathcal{L}(\mathbb{R}) \to [0, +\infty]$

Remarque. La tribu de Lebesgue n'est pas équipotente à $\mathbb{R}$ mais à $\mathcal{P}(\mathbb{R})$. En effet, tout sous-ensemble d'un ensemble de Cantor de mesure nulle est mesurable pour la mesure de Lebesgue.

Question : A-t-on $\mathcal{L}(\mathbb{R}) = \mathcal{P}(\mathbb{R})$? La réponse est non.

Exemple (d'un ensemble non mesurable). Sur $\mathbb{R}$ on définit une relation d'équivalence : $x \sim y \iff x - y \in \mathbb{Q}$. On note $X = \mathbb{R}/_\sim$ le quotient de $\mathbb{R}$ par cette relation d'équivalence. Pour tout $x \in X$, on choisit[2] un représentant $e_x \in [0,1]$. On note $E = \{e_x \mid x \in X\} \subset [0,1]$. Alors E n'est pas mesurable pour la mesure de Lebesgue ($E \notin \mathcal{L}(\mathbb{R})$). Supposons en effet que $E \in \mathcal{L}(\mathbb{R})$. On a :

$$\mathbb{R} = \bigcup_{q \in \mathbb{Q}} q + E \text{ (union disjointe)}$$

D'où

$$\lambda(\mathbb{R}) = +\infty = \sum_{q \in \mathbb{Q}} \lambda(q+E) = \sum_{q \in \mathbb{Q}} \lambda(E) \Longrightarrow \lambda(E) > 0$$

D'autre part, $\displaystyle\bigcup_{q \in \mathbb{Q} \cap [0,1]} q + E \subset [0,2]$ donc

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

On obtient une absurdité.

On a aussi un théorème d'existence de la mesure de Lebesgue sur $\mathbb{R}^d$

Théorème 1.13 (Mesure de Lebesgue sur $\mathbb{R}^d$). *Il existe une unique mesure positive sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, notée λ, telle que pour tout pavé $P = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_d, b_d] \subset \mathbb{R}^d$, on ait :*

$$\lambda(P) = \prod_{i=1}^{d} (b_i - a_i)$$

Comme précédemment, on peut compléter la tribu $\mathcal{B}(\mathbb{R}^d)$ et étendre la mesure λ à $\mathcal{L}(\mathbb{R}^d)$. La mesure de Lebesgue (sur $\mathcal{B}(\mathbb{R}^d)$ ou $\mathcal{L}(\mathbb{R}^d)$) possède en outre les propriétés suivantes :

a) λ est **invariante par translation et rotation**

b) λ est **régulière**, i.e. $\forall E \subset \mathbb{R}^d$ mesurable, on a
   - $\lambda(E) = \sup \{\lambda(K) \mid K \text{ compact}, K \subset E\}$. C'est la régularité intérieure.
   - $\lambda(E) = \inf \{\lambda(V) \mid V \text{ ouvert}, V \supset E\}$. C'est la régularité extérieure.

2. Ici, on se sert de l'axiome du choix !

CHAPITRE 2

THÉORIE GÉNÉRALE DE L'INTÉGRATION

Maintenant que l'on a étudié la théorie générale de la mesure dans le chapitre 1, et en admettant l'existence et l'unicité de la mesure de Lebesgue sur $\mathbb{R}^d$, on va pouvoir définir, comme conséquence, une théorie de l'intégration. On construira donc une intégrale à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$, puis on comparera avec l'intégrale de Riemann. Enfin, on donnera des théorèmes de continuité et dérivabilité sur les intégrales dépendant d'un paramètre avant d'étudier la fonction Γ d'Euler.

2.1 FONCTIONS MESURABLES

2.1.1 Définitions et généralités

Définition 2.1. Soit $(X, \mathcal{M})$ et $(Y, \mathcal{N})$ deux espaces mesurables. On dit qu'une application $f : X \to Y$ est **mesurable** (pour les tribus $\mathcal{M}$ et $\mathcal{N}$) si

$$f^{-1}(B) \in \mathcal{M}, \forall B \in \mathcal{N}$$

Cela rappelle la notion de fonction continue dans les espaces topologiques.

Définition 2.2 (Rappel). Soit $(X, \mathcal{T})$ et $(Y, \mathcal{S})$ deux espaces topologiques. On dit qu'une application $f : X \to Y$ est **continue** si

$$f^{-1}(B) \in \mathcal{T}, \forall B \in \mathcal{S}$$

Les fonctions mesurables sont aux espaces mesurables ce que les fonctions continues sont aux espaces topologiques.

Remarques.

1. Si $(X, \mathcal{M})$ est un espace mesurable, si Y est un ensemble quelconque et $f : X \to Y$, alors on peut toujours munir Y d'une tribu $\mathcal{N}$ telle que f soit mesurable.
   Evidemment, on peut prendre $\mathcal{N} = \{\varnothing, Y\}$ (et c'est la plus petite tribu possible). Un meilleur choix est de poser $\mathcal{N} = \left\{B \subset Y \mid f^{-1}(B) \in \mathcal{M}\right\}$. Alors $\mathcal{N}$ est une tribu (**exercice**), et c'est la plus grande tribu sur Y qui rende f mesurable.
   On dit que $\mathcal{N}$ est la **tribu image** de $\mathcal{M}$ par f.
2. Si X est un ensemble quelconque, si $(Y, \mathcal{N})$ est un espace mesurable, et $f : X \to Y$, alors on peut toujours munir X d'une tribu $\mathcal{M}$ telle que f soit mesurable.
   Evidemment, on peut prendre $\mathcal{M} = \mathcal{P}(X)$ (et c'est la plus grande tribu possible). Un meilleur choix est de poser $\mathcal{M} = \left\{f^{-1}(B) \mid B \in \mathcal{N}\right\}$. Alors $\mathcal{M}$ est une tribu (**exercice**), et c'est la plus petite tribu sur X qui rende f mesurable.
   On dit que $\mathcal{M}$ est la **tribu engendrée** par f.

Le critère suivant est fort utile :

Lemme 2.3. *Soient $(X, \mathcal{M})$ et $(Y, \mathcal{N})$ deux espaces mesurables, et $f : X \to Y$. On suppose que $\mathcal{N}$ est engendrée par une famille $\mathcal{F}$ de parties de Y : $\mathcal{N} = \sigma(\mathcal{F})$.*
Alors f est mesurable si et seulement si $f^{-1}(B) \in \mathcal{M}, \forall B \in \mathcal{F}$

Démonstration. La condition est évidemment nécessaire. Supposons donc que $f^{-1}(B) \in \mathcal{M}, \forall B \in \mathcal{F}$, et considérons la tribu image de $\mathcal{M}$ par f :

$$\widetilde{\mathcal{N}} = \left\{B \subset Y \mid f^{-1}(B) \in \mathcal{M}\right\}$$

Alors $\widetilde{\mathcal{N}}$ contient $\mathcal{F}$, donc $\widetilde{\mathcal{N}}$ contient $\sigma(\mathcal{F}) = \mathcal{N}$, et en particulier, f est mesurable. □

Cas particulier. Si X et Y sont deux espaces topologiques munis de leurs tribus boréliennes, une application $f : X \to Y$ mesurable est appelée **borélienne**.
Par le lemme, $f : X \to Y$ est borélienne si et seulement si, pour tout ouvert $V \subset Y$, $f^{-1}(V)$ est borélien.
En particulier, si $f : X \to Y$ est continue, alors f est borélienne.
Si $Y = \mathbb{R}$ muni de la tribu de Borel $\mathcal{B}(\mathbb{R})$, alors $f : X \to Y$ est borélienne si et seulement si $f^{-1}(]a, +\infty[)$ est borélien pour tout $a \in \mathbb{R}$.

Exemple. Soit $(X, \mathcal{M})$ un espace mesurable, et soit $A \subset X$. On définit la **fonction indicatrice de** A par

$$\begin{array}{rccl} \mathbb{1}_A : & X & \to & \mathbb{R} \\ & x & \mapsto & \mathbb{1}_A(x) = \left\{ \begin{array}{ll} 1 & \text{si } x \in A \\ 0 & \text{sinon} \end{array} \right. \end{array}$$

alors

$$\forall a \in \mathbb{R}, (\mathbb{1}_A^{-1})(]a, +\infty[) = \left\{ \begin{array}{ll} \varnothing & \text{si } a \geqslant 1 \\ A & \text{si } 0 \leqslant a < 1 \\ X & \text{si } a < 0 \end{array} \right.$$

Ainsi, $\mathbb{1}_A$ est mesurable si et seulement si A est mesurable ($A \in \mathcal{M}$).

2.1.2 Stabilité de la classe des fonctions mesurables

Lemme 2.4. *Si $f\colon (X_1, \mathcal{M}_1) \to (X_2, \mathcal{M}_2)$ et $g\colon (X_2, \mathcal{M}_2) \to (X_3, \mathcal{M}_3)$ sont mesurables, alors $g \circ f : (X_1, \mathcal{M}_1) \to (X_3, \mathcal{M}_3)$ est mesurable.*

Démonstration. Evident car $(g \circ f)^{-1}(A) = f^{-1}(g^{-1}(A)), \forall A \subset X_3$. □

Proposition 2.5. *Soit $(X, \mathcal{M})$ un espace mesurable, $(Y, \mathcal{T})$ un espace topologique, $f_1, f_2\colon X \to \mathbb{R}$ des applications mesurables, et $\Phi : \mathbb{R}^2 \to Y$ une application continue.*
Pour tout $x \in X$, on note $h(x) = \Phi(f_1(x), f_2(x))$. Alors $h\colon X \to Y$ est mesurable.

On pourrait formuler cela comme : "Une combinaison continue de deux applications mesurables est mesurable".

Démonstration. Notons $F(x) = (f_1(x), f_2(x))$ de sorte que $F\colon X \to \mathbb{R}^2$. Alors $h = \Phi \circ F$. Comme Φ est borélienne (car continue), il suffit de vérifier que F est mesurable.
Soient $I_1, I_2 \subset \mathbb{R}$ deux intervalles, et soit $R = I_1 \times I_2 \subset \mathbb{R}^2$. Alors $F^{-1}(R) = f_1^{-1}(I_1) \cap f_2^{-1}(I_2) \in \mathcal{M}$. Comme la tribu de Borel sur $\mathbb{R}^2$ est engendrée par les rectangles de la forme ci-dessus, on a que F est mesurable. □

Corollaire 2.6. *Soient $f, g\colon X \to \mathbb{R}$ deux fonctions mesurables. Alors $f+g$, fg, $\min(f,g)$ et $\max(f,g)$ sont mesurables.*

Démonstration. Il suffit d'appliquer la proposition avec $Y = \mathbb{R}$ et $\Phi(x,y) = x+y$, xy, $\min(x,y)$, $\max(x,y)$. □

Corollaire 2.7. *Si $f\colon X \to \mathbb{R}$ est mesurable, alors $f_+ = \max(f,0)$, $f_- = \max(-f,0)$, et $|f| = f_+ + f_-$ sont mesurables.*

Corollaire 2.8. *Si $f\colon X \to \mathbb{R}$ est mesurable, et si $f(x) \neq 0, \forall x \in X$, alors g définie par $g(x) = \dfrac{1}{f(x)}$ est mesurable.*

Démonstration. Soit $Y = \mathbb{R} \setminus \{0\}$ muni de la tribu borélienne, et soit $\varphi : Y \to Y$ définie par $\varphi(y) = \dfrac{1}{y}$.
Comme $f\colon X \to Y$ est mesurable et φ est continue alors $g = \varphi \circ f\colon X \to Y$ (ou $\mathbb{R}$) est mesurable. □

Corollaire 2.9.
1. *Une fonction $f\colon X \to \mathbb{C}$ est mesurable si et seulement si $\Re(f)\colon X \to \mathbb{R}$ et $\Im(f)\colon X \to \mathbb{R}$ sont mesurables.*
2. *Si $f, g\colon X \to \mathbb{C}$ sont mesurables, il en va de même de $f+g$, fg et $|f|$*
3. *Si $f\colon X \to \mathbb{C}$ est mesurable, il existe une fonction $\alpha\colon X \to \mathbb{C}$ mesurable telle que $\forall x, |\alpha(x)| = 1$ et $f = \alpha\,|f|$*

Démonstration (Démonstration du 3). Soit $A = \{x \in X \mid f(x) = 0\} = f^{-1}(\{0\}) \in \mathcal{M}$. Soit $Y = \mathbb{C} \setminus \{0\}$ et $\varphi(z) = \dfrac{z}{|z|}, \forall z \in Y$. On pose $\alpha\colon X \to \mathbb{C}$ une application mesurable (comme composition de deux fonctions mesurables) définie par

$$\alpha(x) = \begin{cases} 1 & \text{si } x \in A \\ \dfrac{f(x)}{|f(x)|} & \text{si } x \notin A \end{cases}$$

La fonction α ainsi définie convient. □

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

1. **relation d'ordre** : On munit $\overline{\mathbb{R}}$ de la relation d'ordre sur $\mathbb{R}$, complétée de

$$\forall a \in \mathbb{R}, -\infty < a < +\infty$$

   $\overline{\mathbb{R}}$ est donc *totalement ordonné*.

2. **topologie** : Les ouverts de $\overline{\mathbb{R}}$ sont les unions d'intervalles de la forme

$$[-\infty, a[\,,]a, b[\,,]b, +\infty]\,, \forall a, b \in \mathbb{R}$$

$\overline{\mathbb{R}}$ est un espace topologique compact.

Remarque. $\overline{\mathbb{R}}$ est homéomorphe à $[-1, 1]$. Un homéomorphisme est donné par $f(x) = \dfrac{2}{\pi}\arctan(x)$

3. **structure borélienne** : la tribu de Borel sur $\overline{\mathbb{R}}$ est engendrée par les intervalles $]a, +\infty]\,, a \in \mathbb{R}$

Attention ! Les opérations algébriques du corps $\mathbb{R}$ ne s'étendent pas à $\overline{\mathbb{R}}$.

Exemple. $a + b$ n'est pas défini si $a = +\infty$ et $b = -\infty$ (ou vice-versa)
ab n'est pas défini si $a = 0$ et $b = \pm\infty$ (ou vice versa)

Remarque. On peut étendre les opérations algébriques sur $\overline{\mathbb{R}_+} = [0, +\infty]$ en posant :
– $a + \infty = \infty + a = \infty$ si $0 \leqslant a \leqslant \infty$
– $a \cdot \infty = \infty \cdot a = \begin{cases} 0 & \text{si } a = 0 \\ \infty & \text{si } 0 < a \leqslant \infty \end{cases}$

Attention !
$a + b = a + c \Longrightarrow b = c$ si $0 \leqslant a < \infty$
$a \cdot b = a \cdot c \Longrightarrow b = c$ si $0 < a < \infty$

Notation. $(X, \mathcal{M})$ un espace mesurable, et $\{f_n\}_n$ une suite de fonctions de X dans $\overline{\mathbb{R}}$. On note :
– $(\sup\limits_n f_n)(x) = \sup\limits_n f_n(x)$
– $(\limsup\limits_{n\to\infty} f_n)(x) = \lim\limits_{n\to\infty} \sup\limits_{k\geqslant n} f_k(x)$
On définit de même $\inf\limits_n f_n$ et $\liminf\limits_{n\to\infty} f_n$

Proposition 2.10 (Stabilité des fonctions mesurables par limite ponctuelle). *Soit $(X, \mathcal{M})$ un espace mesurable et $f_n : X \to \overline{\mathbb{R}}$ une suite de fonction mesurables, alors*

$$\sup_n f_n, \inf_n f_n, \limsup_{n\to\infty} f_n, \liminf_{n\to\infty} f_n : X \to \overline{\mathbb{R}}$$

sont mesurables.
En particulier, si $f(x) = \lim_{n\to\infty} f_n(x)$ existe $\forall x \in X$, alors $f : X \to \overline{\mathbb{R}}$ est mesurable.
Plus généralement, l'ensemble $\{x \in X \mid \lim_{n\to\infty} f_n(x) \text{ existe}\}$ est mesurable.

Démonstration. Si $g = \sup_n f_n$, on a $\forall a \in \mathbb{R}$,

$$g^{-1}(]a, +\infty]) = \bigcup_{n\in\mathbb{N}} f_n^{-1}(]a, +\infty]) \in \mathcal{M}$$

Ainsi, g est mesurable. Il en va de même de $\inf_n f_n = -\sup_n(-f_n)$. En conséquence, $\limsup\limits_{n\to\infty} f_n = \inf\limits_n \sup\limits_{k\geqslant n} f_k$ est mesurable, et il en va de même de $\liminf\limits_{n\to\infty} f_n$.
Pour montrer la dernière affirmation, on pose $F = (\liminf\limits_{n\to\infty} f_n, \limsup\limits_{n\to\infty} f_n)$ et on remarque que

$$\left\{x \in X \mid \liminf_{n\to\infty} f_n(x) = \limsup_{n\to\infty} f_n(x)\right\} = F^{-1}(\Delta)$$

où $\Delta = \left\{(x, x) \mid x \in \overline{\mathbb{R}}\right\} \subset \overline{\mathbb{R}}^2$ Cet ensemble est mesurable car Δ est fermé et F est mesurable. $\square$

2.2 LES FONCTIONS ÉTAGÉES

2.2.1 Généralités

Définition 2.11. Soit $(X, \mathcal{M})$ un espace mesurable. On dit qu'une application mesurable $f : X \to \mathbb{R}$ est **étagée** si f ne prend qu'un nombre fini de valeurs.
En notant $\alpha_1, \ldots, \alpha_n$ les valeurs de f et $A_i = f^{-1}(\alpha_i)$ pour $i = 1, \ldots, n$, on a donc

$$f = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{A_i} \tag{2.1}$$

Exemple. La fonction indicatrice des rationnels $\mathbb{1}_{\mathbb{Q}}$ est une fonction étagée.

L'écriture (2.1) est unique aux renumérotations près. Les $\alpha_1, \ldots, \alpha_n$ sont deux à deux distincts et les $A_1, \ldots, A_n$ sont deux à deux disjoints. Si f prend la valeur 0, on peut omettre le terme correspondant dans la somme (2.1).
Les fonctions (mesurables) étagées sont exactement les combinaisons linéaires finies de fonctions indicatrices d'ensembles mesurables. Si

$$f = \sum_{k=1}^{N} \beta_k \mathbb{1}_{B_k}$$

avec $\beta_1, \ldots, \beta_N \in \mathbb{R}$ non nécessairement distincts et $B_1, \ldots, B_N \in \mathcal{M}$ non nécessairement disjoints, alors f ne prend qu'un nombre fini de valeurs et possède donc aussi une écriture canonique de la forme (2.1).

Proposition 2.12. *Soit $f : (X, \mathcal{M}) \to [0, +\infty]$ une fonction mesurable. Alors il existe une suite croissante de fonctions (mesurables) étagées qui converge ponctuellement vers f.*

Démonstration. Pour tout $n \in \mathbb{N}$, on définit $\varphi_n \colon [0, +\infty] \to [0, +\infty[$ par

$$\varphi_n(t) = \begin{cases} 2^{-n} E(2^n t) & \text{si } 0 \leqslant t < n \\ n & \text{si } t \geqslant n \end{cases}$$

où $E(x)$ désigne la partie entière de x.
Il est clair que φ_n est étagée telle que $0 \leqslant \varphi_n(t) \leqslant \varphi_{n+1}(t), \forall t \in [0, +\infty]$. On pose $f_n = \varphi_n \circ f, n \in \mathbb{N}$, alors f_n est étagée, $\forall n \in \mathbb{N}, f_n \leqslant f_{n+1}$ et $f_n(x) \underset{n\to\infty}{\longrightarrow} f(x), \forall x \in X$.
En effet, si $f(x) = +\infty$, on a $\forall n, f_n(x) = n$. D'autre part, si $f(x) < \infty$ et $n \geqslant f(x)$ alors $f(x) - 2^{-n} \leqslant f_n(x) \leqslant f(x)$. $\square$

2.2.2 Définition de l'intégrale d'une fonction étagée positive

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

Définition 2.13. Soit $f \colon (X, \mathcal{M}, \mu) \to [0, +\infty[$ une fonction (mesurable) étagée. On appelle **intégrale de** f (pour la mesure positive μ) la quantité

$$\int f \mathrm{d}\mu = \underbrace{\sum_{i=1}^{n} \alpha_i \mu(A_i)}_{\in [0, +\infty]}$$

où $f = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{A_i}$ est l'écriture canonique de f comme dans (2.1)

Lemme 2.14. *Si $\beta_1, \dots, \beta_N \in [0, +\infty[$ et $B_1, \dots, B_n \in \mathcal{M}$, et si $f = \displaystyle\sum_{k=1}^{N} \beta_k \mathbb{1}_{B_k}$, alors*

$$\int f \mathrm{d}\mu = \sum_{k=1}^{N} \beta_k \mu(B_k)$$

Démonstration. On suppose $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_N \neq 0$.

1er cas : Les $B_1, \dots, B_N$ sont 2 à 2 disjoints. Dans ce cas là, les $\beta_1, \dots, \beta_N$ parcourent $\{\alpha_1, \dots, \alpha_n\}$, et on a $A_i = \displaystyle\bigcup_{k/B_k \subset A_i} B_k$ et $B_k \subset A_i \iff \beta_k = \alpha_i$. Ainsi,

$$\begin{aligned}
\sum_{i=1}^{n} \alpha_i \mu(A_i) &= \sum_{i=1}^{n} \alpha_i \left(\sum_{k/B_k \subset A_i} \mu(B_k) \right) \\
&= \sum_{i=1}^{n} \sum_{k/B_k \subset A_i} \beta_k \mu(B_k) \\
&= \sum_{k=1}^{N} \beta_k \mu(B_k)
\end{aligned}$$

2e cas (cas général) : La σ-algèbre engendrée par les $B_1, \dots, B_N$ est également engendrée par les ensembles $C_1, \dots, C_m$ deux à deux disjoints. On définit

$$\gamma_j = \sum_{k/B_k \supset C_j} \beta_k, j = 1, \dots, M$$

□

alors $f = \displaystyle\sum_{j=1}^{m} \gamma_j \mathbb{1}_{C_j}$ avec $C_1, \dots, C_m$ deux à deux disjoints. En outre,

$$\begin{aligned}
\sum_{k=1}^{N} \beta_k \mu(B_k) &= \sum_{k=1}^{N} \beta_k \left(\sum_{j/B_k \supset C_j} \mu(C_j) \right) \\
&= \sum_{j=1}^{m} \mu(C_j) \left(\sum_{k/B_k \supset C_j} \beta_k \right) \\
&= \sum_{j=1}^{m} \gamma_j \mu(C_j)
\end{aligned}$$

Notons $\mathcal{E}_+$ l'ensemble des fonctions (mesurables) étagées sur X à valeurs dans $[0, +\infty[$. L'application $\begin{array}{rccc} I\colon & \mathcal{E}_+ & \longrightarrow & [0,+\infty] \\ & f & \longmapsto & \int f \mathrm{d}\mu \end{array}$ possède les propriétés suivantes :

i) **Additivité :**

$$\int (f+g)\mathrm{d}\mu = \int f\mathrm{d}\mu + \int g\mathrm{d}\mu, \forall f, g \in \mathcal{E}_+$$

ii) **Homogénéité :**

$$\int \lambda f\mathrm{d}\mu = \lambda \int f\mathrm{d}\mu, \forall f \in \mathcal{E}_+, \forall \lambda \in \mathbb{R}_+$$

iii) **Monotonie :** Si f et $g \in \mathcal{E}_+$ et si $f \leqslant g$, alors

$$\int f\mathrm{d}\mu \leqslant \int g\mathrm{d}\mu$$

En effet,

i) si $f = \sum_{i=1}^n \alpha_i \mathbb{1}_{A_i}$ et $g = \sum_{j=1}^m \beta_j \mathbb{1}_{B_j}$, alors $f + g = \sum_{i=1}^n \alpha_i \mathbb{1}_{A_i} + \sum_{j=1}^m \beta_j \mathbb{1}_{B_j}$, et donc

$$\begin{aligned} \int (f+g)\mathrm{d}\mu &\underset{\text{lemme}}{=} \sum_{i=1}^n \alpha_i \mu(A_i) + \sum_{j=1}^m \beta_j \mu(B_j) \\ &= \int f\mathrm{d}\mu + \int g\mathrm{d}\mu \end{aligned}$$

ii) C'est évident par définition de l'intégrale.

iii) Suit de $i)$ car

$$\int g\mathrm{d}\mu = \int f\mathrm{d}\mu + \underbrace{\int (g-f)\mathrm{d}\mu}_{\geqslant 0} \geqslant \int f\mathrm{d}\mu$$

2.3 Intégration des fonctions mesurables positives.

Dans toute la suite, $(X, \mathcal{M}, \mu)$ désigne un espace mesuré.

2.3.1 Définitions et théorème de convergence monotone

Définition 2.15. Soit $f : X \to [0, +\infty]$ une fonction mesurable. On appelle **intégrale de** f (sur X, pour la mesure μ), la quantité :

$$\int f\mathrm{d}\mu = \sup \left\{ \int h\mathrm{d}\mu \mid h \in \mathcal{E}_+, h \leqslant f \right\} \in [0, +\infty]$$

Si $E \subset X$ est une partie mesurable, on note aussi

$$\int_E f\mathrm{d}\mu = \int f \mathbb{1}_E \mathrm{d}\mu$$

Si $f \in \mathcal{E}_+$, on retrouve bien la définition précédente de l'intégrale.

Cette intégrale possède la propriété de la monotonie : $\int f\mathrm{d}\mu \leqslant \int g\mathrm{d}\mu$ si $f \leqslant g$.

Théorème 2.16 (de la convergence monotone – Beppo-Levi). *Soit $f_n \colon X \to [0,+\infty]$ une suite croissante de fonctions mesurables positives, et soit $f = \lim_{n\to\infty} f_n$ la limite ponctuelle des f_n. Alors, f est mesurable et*

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

Démonstration. Comme $\displaystyle\int f_n \mathrm{d}\mu \leqslant \int f_{n+1} \mathrm{d}\mu$ du fait de la croissance de (f_n), la limite $\alpha = \displaystyle\lim_{n\to\infty} \int f_n \mathrm{d}\mu \in [0,+\infty]$ existe. Soit $f = \lim_{n\to\infty} f_n$, alors f est mesurable, et comme $\forall n, f_n \leqslant f$, on a

$$\int f_n \mathrm{d}\mu \leqslant \int f \mathrm{d}\mu, \forall n \Longrightarrow \alpha \leqslant \int f \mathrm{d}\mu$$

Par ailleurs, soit $h \in \mathcal{E}_+$ telle que $h \leqslant f$ et soit $c \in]0,1[$. On définit :

$$A_n = \{x \in X \mid f_n(x) \geqslant ch(x)\}$$

alors $A_n \in \mathcal{M}$ (comme image inverse de $[0,+\infty]$ par $f_n - ch$ mesurable), et d'autre part $A_n \subset A_{n+1}$ et $\bigcup_{n\in\mathbb{N}} A_n = X$. Or,

$$\int f \mathrm{d}\mu \geqslant \int_{A_n} f_n \mathrm{d}\mu \geqslant \int_{A_n} ch \mathrm{d}\mu = c \int_{A_n} h \mathrm{d}\mu$$

De plus, si $h = \sum_{j=1}^m \beta_j \mathbb{1}_{B_j}$, on a

$$\int_{A_n} h \mathrm{d}\mu = \sum_{j=1}^m \beta_j \mu(B_j \cap A_n)$$

Comme on a une somme finie, on peut passer à la limite quand n tend vers $+\infty$ et on a

$$\int_{A_n} h \mathrm{d}\mu \underset{n\to\infty}{\longrightarrow} \sum_{j=1}^m \beta_j \mu(B_j) = \int h \mathrm{d}\mu$$

Ainsi, $\alpha \geqslant c \displaystyle\int h \mathrm{d}\mu$ et ce $\forall c \in]0,1[, \forall h \in \mathcal{E}_+$ tel que $h \leqslant f$. On prend d'abord le sup sur $c \in]0,1[$, puis le sup sur $h \in \mathcal{E}_+, h \leqslant f$ et on trouve $\alpha \geqslant \displaystyle\int f \mathrm{d}\mu$. □

Remarques. 1) Si f_n est une suite **décroissante** de fonctions mesurables positives, et si $\int f_0 \mathrm{d}\mu < \infty$, alors on a

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

où $f = \lim_{n\to\infty} f_n$

Démonstration. Appliquer le théorème de Beppo-Levi à la suite $g_n = f_0 - f_n$. □

2) On rappelle que si $f \colon X \to [0,+\infty]$ est mesurable, il existe une suite croissante $f_n \leqslant f_{n+1}, f_n \in \mathcal{E}_+$ telle que $f_n(x) \underset{n\to\infty}{\longrightarrow} f(x), \forall x \in X$. Alors,

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

L'intégrale des fonctions mesurables positives vérifie :

i) **Additivité :**
$$\int (f+g)\mathrm{d}\mu = \int f\mathrm{d}\mu + \int g\mathrm{d}\mu, \forall f,g \in \mathcal{E}_+$$
En effet, soient (f_n) et (g_n) deux suites croissantes dans $\mathcal{E}_+$ telles que $f_n \underset{n}{\longrightarrow} f$ et $g_n \underset{n}{\longrightarrow} g$. Alors $f_n + g_n \in \mathcal{E}_+$ et $f_n + g_n \underset{n\to\infty}{\longrightarrow} f+g$. Or, $\forall n$, $\displaystyle\int (f_n+g_n)\mathrm{d}\mu = \int f_n\mathrm{d}\mu + \int g_n\mathrm{d}\mu$, on obtient donc le résultat en passant à la limite grâce au théorème de convergence monotone.

ii) **Monotonie :** Si $f \leqslant g$, alors
$$\int f\mathrm{d}\mu \leqslant \int g\mathrm{d}\mu$$

Corollaire 2.17. *Soit (f_n) une suite de fonctions mesurables positives et soit $f = \sum_{n\in\mathbb{N}} f_n$, $f\colon X \to [0,+\infty]$. Alors f est mesurable et*
$$\int f\mathrm{d}\mu = \sum_{n\in\mathbb{N}} \int f_n\mathrm{d}\mu$$
C'est la propriété "d'additivité dénombrable".

Démonstration. Soit $F_N = \sum_{n=0}^{N} f_n$, alors $(F_N)_N$ est une suite croissante de fonctions mesurables positives, et $\forall N \in \mathbb{N}$, on a
$$\int F_N\mathrm{d}\mu = \sum_{n=0}^{N} \int f_n\mathrm{d}\mu$$
En prenant la limite quand $N \longrightarrow \infty$, et en utilisant le théorème de convergence monotone, on obtient le résultat. $\square$

2.3.2 Propriétés de l'intégrale

Définition 2.18 (Terminologie). Dans un espace mesuré $(X, \mathcal{M}, \mu)$, on dit qu'une propriété $P(x)(x \in X)$ est vraie **presque partout** (ou μ-presque partout) si elle est vraie *en dehors d'un ensemble négligeable.*

Exemple. Si $f,g\colon X \to \mathbb{R}$ ou $\mathbb{C}$ sont mesurables, alors $\{x \mid f(x) \neq g(x)\} \in \mathcal{M}$ et donc
$$f = g \text{ presque partout } \iff \mu\left(\{x \mid f(x) \neq g(x)\}\right) = 0$$

Remarque. μ n'est pas nécessairement complète dans l'exemple précédent.
Ainsi, si μ est complète ou si f, g sont mesurables, alors
$$f = g \text{ presque partout } \iff \mu\left(\{x \in X \mid f(x) \neq g(x)\}\right) = 0$$

Proposition 2.19. *Soit $f\colon X \to [0,+\infty]$ une fonction mesurable.*

1) $\forall a > 0, \mu\left(\{x \in X \mid f(x) \geqslant a\}\right) \leqslant \dfrac{1}{a}\displaystyle\int f\mathrm{d}\mu.$

2) $\displaystyle\int f\mathrm{d}\mu = 0 \iff f = 0$ *presque partout.*

3) Si $\int f \mathrm{d}\mu < \infty$*, alors* $f < \infty$ *presque partout.*

4) Si f *et* $g\colon X \to [0, +\infty]$ *sont mesurables, alors*

$$f = g \text{ presque partout} \Longrightarrow \int f \mathrm{d}\mu = \int g \mathrm{d}\mu$$

Démonstration. 1) Soit $A = \{x \in X \mid f(x) \geqslant a\} = f^{-1}([a, +\infty]) \in \mathcal{M}$. On a $f \geqslant a\mathbb{1}_A$ et ainsi,

$$\int f \mathrm{d}\mu \geqslant a\mu(A)$$

2) Si $f = 0$ presque partout, alors si $h \in \mathcal{E}_+$ tel que $h \leqslant f$, on a $h = 0$ presque partout. En utilisant la définition de l'intégrale dans $\mathcal{E}_+$, on en déduit que $\int h \mathrm{d}\mu = 0$ d'où $\int f \mathrm{d}\mu = 0$.

Inversement, supposons que $\int f \mathrm{d}\mu = 0$. Pour tout entier $n \in \mathbb{N}^*$, on note

$$A_n = \left\{x \in X \mid f(x) \geqslant \frac{1}{n}\right\}$$

Alors A_n est mesurable (image réciproque de $\left[\frac{1}{n}, +\infty\right]$ par f), $A_n \subset A_{n+1}$ et $\bigcup_{n \in \mathbb{N}^*} A_n = \{x \in X \mid f(x) > 0\}$. Par ailleurs,

$$\forall n \in \mathbb{N}^*, \mu(A_n) \leqslant n \int f \mathrm{d}\mu = 0 \qquad \text{par 1)}$$

Ainsi, on en déduit que

$$\mu\left(\{x \in X \mid f(x) > 0\}\right) = \lim_{n \to \infty} \mu(A_n) = 0$$

3) Supposons que $f(x) = +\infty, \forall x \in A$ avec $A \in \mathcal{M}$ et $\mu(A) > 0$. Alors, $\forall n \in \mathbb{N}, f \geqslant n\mathbb{1}_A$ et donc

$$\int f \mathrm{d}\mu \geqslant n\mu(A), \forall n \in \mathbb{N}$$

donc $\int f \mathrm{d}\mu = +\infty$.

4) Supposons que $f = g$ presque partout, et notons $h_+ = \max(f, g)$ et $h_- = \min(f, g)$ (point par point) alors h_+ et h_- sont mesurables, $h_+ = h_-$ presque partout et $h_- \leqslant f, g \leqslant h_+$. Or,

$$\int h_+ \mathrm{d}\mu = \int h_- \mathrm{d}\mu + \underbrace{\int (h_+ - h_-) \mathrm{d}\mu}_{=0 \text{ avec 2)}} = \int h_- \mathrm{d}\mu$$

Par monotonie,

$$\int f \mathrm{d}\mu = \int g \mathrm{d}\mu = \int h_+ \mathrm{d}\mu = \int h_- \mathrm{d}\mu$$

□

Lemme 2.20 (de Fatou - Corollaire du théorème de convergence monotone). *Soit* $(f_n)_n$, $f_n\colon X \to [0,+\infty]$ *une suite de fonctions mesurables. Alors*

$$\int (\liminf_{n\to\infty} f_n)\mathrm{d}\mu \leqslant \liminf_{n\to\infty} \int f_n \mathrm{d}\mu$$

Démonstration. On rappelle que $\liminf_{n\to\infty} f_n = \lim_{k\to\infty}(\inf_{n\geqslant k} f_n)$, et c'est une limite croissante, donc par le théorème de convergence monotone,

$$\int (\liminf_{n\to\infty} f_n)\mathrm{d}\mu = \lim_{k\to\infty} \int (\inf_{n\geqslant k} f_n)\mathrm{d}\mu$$

D'autre part, si $p \geqslant k$, on a $\inf_{n\geqslant k} f_n \leqslant f_p$, donc

$$\int (\inf_{k\geqslant n} f_n)\mathrm{d}\mu \leqslant \int f_p \mathrm{d}\mu$$

On passe à l'inf sur $p \geqslant k$, et on a

$$\int (\inf_{k\geqslant n} f_n)\mathrm{d}\mu \leqslant \inf_{p\geqslant k} \int f_p \mathrm{d}\mu$$

Ainsi,

$$\int (\liminf_{n\to\infty} f_n)\mathrm{d}\mu \leqslant \lim_{k\to\infty} \inf_{p\geqslant k} \int f_p \mathrm{d}\mu = \liminf_{n\to\infty} \int f_n \mathrm{d}\mu$$

□

2.3.3 Application : Mesures à densité

Soit $(X, \mathcal{M}, \mu)$ un espace mesuré, et $f\colon X \to [0,+\infty]$ une fonction mesurable.
On définit une application $\nu\colon \mathcal{M} \to [0,+\infty]$ par

$$\nu(A) = \int_A f \mathrm{d}\mu = \int f \mathbb{1}_A \mathrm{d}\mu$$

Alors, ν est une mesure sur $(X, \mathcal{M})$ appelée **mesure de densité** f **par rapport à** μ.
En effet, $\nu(\varnothing) = 0$, et si $(A_n)_{n\in\mathbb{N}}$ est une suite d'élements de $\mathcal{M}$ deux à deux disjoints, on a :

$$\begin{aligned}
\nu\left(\bigcup_{n\in\mathbb{N}} A_n\right) &= \int f \mathbb{1}_{\bigcup_n A_n} \mathrm{d}\mu \\
&= \int f \sum_{n\in\mathbb{N}} \mathbb{1}_{A_n} \mathrm{d}\mu \text{ car les } A_n \text{ sont 2 à 2 disjoints} \\
&= \sum_{n\in\mathbb{N}} \int f \mathbb{1}_{A_n} \mathrm{d}\mu \text{ interversion somme/intégrale} \\
& \quad \text{licite quand tout est positif} \\
&= \sum_{n\in\mathbb{N}} \nu(A_n)
\end{aligned}$$

Remarque. Si $A \in \mathcal{M}$ vérifie que $\mu(A) = 0$ alors $\nu(A) = 0$. On dit que ν **est absolument continue par rapport à** μ.

Exercice. Montrer que, si $g\colon X \to [0, +\infty]$ est mesurable, alors

$$\int g \mathrm{d}\nu = \int g f \mathrm{d}\mu$$

Démonstration (Solution). Soit $g\colon X \to [0, +\infty]$ mesurable.

– Si $g = \mathbb{1}_A$ où $A \in \mathcal{M}$, on a

$$\int g \mathrm{d}\nu = \int \mathbb{1}_A \mathrm{d}\nu = \nu(A) = \int_A f \mathrm{d}\mu = \int f \mathbb{1}_A d\mu = \int f g \mathrm{d}\mu$$

– Si g est une fonction étagée positive, le résultat est vrai par additivité et homogénéité positive de l'intégrale.
– Cas général : il existe une suite croissante $(g_n)_{n\in\mathbb{N}}$ de fonctions étagées positives convergeant simplement vers g (proposition 2.12). D'après ce qui précède, on a pour tout $n \in \mathbb{N}$,

$$\int g_n \mathrm{d}\nu = \int g_n f \mathrm{d}\mu$$

Le théorème de convergence monotone (2.16) entraîne donc que

$$\int g \mathrm{d}\nu = \lim_{n\to\infty} \int g_n \mathrm{d}\nu = \lim_{n\to\infty} \int g_n f \mathrm{d}\mu$$

D'autre part, la suite $(g_n f)_n$ est croissante, positive et converge simplement vers fg donc toujours par le théorème de convergence monotone,

$$\lim_{n\to\infty} \int g_n f \mathrm{d}\mu = \int f g \mathrm{d}\mu$$

Finalement,

$$\int g \mathrm{d}\nu = \int f g \mathrm{d}\mu$$

□

2.4 Fonctions intégrables à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$

Dans toute cette partie, $(X, \mathcal{M}, \mu)$ est un espace mesuré quelconque.

2.4.1 Cas de $\mathbb{R}$

Définition 2.21. Soit $f\colon X \to \mathbb{R}$ une fonction mesurable. On dit que f **est intégrable** (ou **sommable**) par rapport à μ si

$$\int |f| \, \mathrm{d}\mu < \infty$$

Dans ce cas, on pose

$$\int f \mathrm{d}\mu = \int f_+ \mathrm{d}\mu - \int f_- \mathrm{d}\mu \tag{2.2}$$

où $f_+ = \max(f, 0)$ et $f_- = \max(-f, 0)$.
On note $\mathcal{L}^1(X, \mathcal{M}, \mu)$ l'espace des fonctions intégrables sur X.

Remarques. 1) Si $f \geqslant 0$, on retrouve la définition précédente.

2) Si $\displaystyle\int |f| \,\mathrm{d}\mu < \infty$, alors comme $f_+, f_- \leqslant |f|$, on a aussi les intégrales de f_+ et f_- qui sont finies, et la définition (2.2) fait sens.

Proposition 2.22.

a) $\mathcal{L}^1(X, \mathcal{M}, \mu)$ *est un espace vectoriel sur* $\mathbb{R}$, *et l'application* $f \mapsto \displaystyle\int f \mathrm{d}\mu$ *est linéaire.*

b) $\displaystyle\left|\int f \mathrm{d}\mu\right| \leqslant \int |f| \,\mathrm{d}\mu, \forall f \in \mathcal{L}^1(X, \mathcal{M}, \mu)$

c) Si $f, g \in \mathcal{L}^1(X, \mathcal{M}, \mu)$ *et* $f \leqslant g$ *alors* $\displaystyle\int f \mathrm{d}\mu \leqslant \int g \mathrm{d}\mu$.

d) Si $f, g \in \mathcal{L}^1(X, \mathcal{M}, \mu)$ *et si* $f = g$ *presque partout, alors* $\displaystyle\int f \mathrm{d}\mu = \int g \mathrm{d}\mu$

Démonstration. a) Si $f, g \in \mathcal{L}^1(X, \mathcal{M}, \mu)$, alors $f + g$ est mesurable et $|f + g| \leqslant |f| + |g|$ donc $f + g \in \mathcal{L}^1(X, \mathcal{M}, \mu)$.
En outre,

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

donc

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

Ainsi,

$$\int (f+g)_+ \mathrm{d}\mu + \int f_- \mathrm{d}\mu + \int g_- \mathrm{d}\mu = \int (f+g)_- \mathrm{d}\mu + \int f_+ \mathrm{d}\mu + \int g_+ \mathrm{d}\mu$$

Ce sont des intégrales finies donc

$$\begin{aligned}
\int (f+g) \mathrm{d}\mu &= \int (f+g)_+ \mathrm{d}\mu - \int (f+g)_- \mathrm{d}\mu \\
&= \left(\int f_+ \mathrm{d}\mu - \int f_- \mathrm{d}\mu\right) + \left(\int g_+ \mathrm{d}\mu - \int g_- \mathrm{d}\mu\right) \\
&= \int f \mathrm{d}\mu + \int g \mathrm{d}\mu
\end{aligned}$$

D'autre part, si $f \in \mathcal{L}^1(X, \mathcal{M}, \mu)$ et $\lambda \in \mathbb{R}$, alors λf est mesurable et $\displaystyle\int |\lambda f| \,\mathrm{d}\mu < +\infty$ donc $\lambda f \in \mathcal{L}^1(X, \mathcal{M}, \mu)$.
– si $\lambda \geqslant 0$,

$$\begin{aligned}
\int (\lambda f) \mathrm{d}\mu &= \int (\lambda f)_+ \mathrm{d}\mu + \int (\lambda f)_- \mathrm{d}\mu \\
&= \lambda \int f_+ \mathrm{d}\mu - \lambda \int f_- \mathrm{d}\mu \\
&= \lambda \int f \mathrm{d}\mu
\end{aligned}$$

– si $\lambda \leqslant 0$,

$$\begin{aligned}\int (\lambda f) \mathrm{d}\mu &= \int (\lambda f)_+ \mathrm{d}\mu + \int (\lambda f)_- \mathrm{d}\mu \\ &= (-\lambda) \int f_- \mathrm{d}\mu - (-\lambda) \int f_+ \mathrm{d}\mu \\ &= \lambda \int f \mathrm{d}\mu\end{aligned}$$

b)

$$\begin{aligned}\left|\int f \mathrm{d}\mu\right| &= \left|\int f_+ \mathrm{d}\mu - \int f_- \mathrm{d}\mu\right| \\ &\leqslant \left|\int f_+ \mathrm{d}\mu\right| + \left|\int f_- \mathrm{d}\mu\right| \\ &\leqslant \int |f| \, \mathrm{d}\mu\end{aligned}$$

car $|f| = f_+ + f_-$

c) Comme pour les fonctions mesurables positives.

d) Si $f = g$ presque partout, alors $f_+ = g_+$ presque partout et $f_- = g_-$ presque partout, donc $\int f \mathrm{d}\mu = \int g \mathrm{d}\mu$ □

2.4.2 Cas de $\mathbb{C}$

Définition 2.23. Soit $f \colon X \to \mathbb{C}$ une fonction mesurable. On dit que f **est intégrable** (ou **sommable**) par rapport à μ si

$$\int |f| \, \mathrm{d}\mu < \infty$$

où $|\cdot|$ est le module.
Dans ce cas, on pose

$$\int f \mathrm{d}\mu = \int \Re(f) \mathrm{d}\mu + i \int \Im(f) \mathrm{d}\mu \tag{2.3}$$

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

Remarques. 1) Si $\Im(f) = 0$, on retrouve la définition précédente.

2) Si $\int |f| \, \mathrm{d}\mu < \infty$, alors comme $|\Re(f)|, |\Im(f)| \leqslant |f|$, on a aussi les intégrales de $\Re(f)$ et $\Im(f)$ qui sont finies, et la définition (2.3) fait sens.

Proposition 2.24.

a) $\mathcal{L}^1_{\mathbb{C}}(X, \mathcal{M}, \mu)$ *est un espace vectoriel sur* $\mathbb{C}$*, et l'application* $f \mapsto \int f \mathrm{d}\mu$ *est linéaire.*

b) $\left|\int f \mathrm{d}\mu\right| \leqslant \int |f| \, \mathrm{d}\mu, \forall f \in \mathcal{L}^1_{\mathbb{C}}(X, \mathcal{M}, \mu)$

c) Si $f, g \in \mathcal{L}^1_{\mathbb{C}}(X, \mathcal{M}, \mu)$ et si $f = g$ presque partout, alors $\int f \mathrm{d}\mu = \int g \mathrm{d}\mu$

Démonstration. (partielle)

a) Si $f = f_1 + if_2 \in \mathcal{L}^1_{\mathbb{C}}(X, \mathcal{M}, \mu)$ et $\lambda = \lambda_1 + i\lambda_2 \in \mathbb{C}$, on a :

$$\begin{aligned}
\int (\lambda f) \mathrm{d}\mu &= \int \Re(\lambda f) \mathrm{d}\mu + i \int \Im(\lambda f) \mathrm{d}\mu \\
&= \int (\lambda_1 f_1 - \lambda_2 f_2) \mathrm{d}\mu + i \int (\lambda_1 f_2 + \lambda_2 f_1) \mathrm{d}\mu \\
&= \lambda_1 \left(\int f_1 \mathrm{d}\mu + i \int f_2 \mathrm{d}\mu \right) + i\lambda_2 \left(\int f_1 \mathrm{d}\mu + i \int f_2 \mathrm{d}\mu \right) \\
&= \lambda \int f \mathrm{d}\mu
\end{aligned}$$

b) Soit $f \in L^1_{\mathbb{C}}(X, \mathcal{M}, \mu)$. Il existe [1] $\lambda \in \mathbb{C}$ avec $|\lambda| = 1$ tel que

$$\left| \int f \mathrm{d}\mu \right| = \lambda \int f \mathrm{d}\mu = \int \lambda f \mathrm{d}\mu$$

Soit $g = \Re(\lambda f)$, alors $|g| \leqslant |\lambda f| \leqslant |f|$, donc

$$\begin{aligned}
\underbrace{\left| \int f \mathrm{d}\mu \right|}_{\in \mathbb{R}} &= \underbrace{\int g \mathrm{d}\mu}_{\in \mathbb{R}} + i \underbrace{\int \Im(\lambda f) \mathrm{d}\mu}_{=0} \\
&= \left| \int g \mathrm{d}\mu \right| \leqslant \int |g| \, \mathrm{d}\mu \\
&\leqslant \int |f| \, \mathrm{d}\mu
\end{aligned}$$

c) Comme avant. □

2.4.3 Théorème de la convergence dominée

Théorème 2.25 (de la convergence dominée). *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré, et $f_n \colon X \to \mathbb{C}$ une suite de fonctions mesurables [2]. On suppose que :*

1) la limite $f(x) = \lim_{n\to\infty} f_n(x)$ existe $\forall x \in X$.

2) il existe $g \colon X \to [0, +\infty[$ intégrable telle que $|f_n(x)| \leqslant g(x) \quad \forall n \in \mathbb{N}, \forall x \in X$

Alors $f \colon X \to \mathbb{C}$ est intégrable, et on a :

$$\int f \mathrm{d}\mu = \lim_{n\to\infty} \int f_n \mathrm{d}\mu \quad \text{et} \quad \lim_{n\to\infty} \int |f_n - f| \, \mathrm{d}\mu = 0$$

1. on peut prendre $\lambda = \frac{|\int f \mathrm{d}\mu|}{\int f \mathrm{d}\mu}$ si $\int f \mathrm{d}\mu \neq 0$, et $\lambda = 1$ sinon.

Démonstration. On sait que f est mesurable, et comme $|f| \leqslant g$, il suit que f est intégrable. Etant donné que $|f - f_n| \leqslant 2g$, on peut appliquer le lemme de Fatou (2.20) à la suite $2g - |f - f_n| \geqslant 0$.

$$\begin{aligned}
\int 2g \mathrm{d}\mu = \int \liminf_{n\to\infty}(2g - |f - f_n|)\mathrm{d}\mu \quad &\leqslant \quad \liminf_{n\to\infty} \int 2g - |f - f_n| \,\mathrm{d}\mu \\
&\underset{\text{linéarité}}{=} \quad \liminf_{n\to\infty} \left(\int 2g\mathrm{d}\mu - \int |f - f_n| \,\mathrm{d}\mu \right) \\
&= \quad \int 2g\mathrm{d}\mu - \limsup_{n\to\infty} \int |f - f_n| \,\mathrm{d}\mu
\end{aligned}$$

On en déduit que $\limsup_{n\to\infty} \int |f - f_n| \,\mathrm{d}\mu \leqslant 0$ donc $\lim_{n\to\infty} \int |f - f_n| \,\mathrm{d}\mu = 0$

Comme

$$\left| \int f_n \mathrm{d}\mu - \int f \mathrm{d}\mu \right| \leqslant \int |f_n - f| \,\mathrm{d}\mu$$

On conclut que $\lim_{n\to\infty} \int f_n \mathrm{d}\mu = \int f \mathrm{d}\mu$ □

Variante du théorème de la convergence dominée

Théorème 2.26. *Soit $f_n \colon X \to \mathbb{C}$ une suite de fonctions mesurables telle que :*

1) il existe $f \colon X \to \mathbb{C}$ mesurable telle que $f(x) = \lim_{n\to\infty} f_n(x)$ pour presque tout $x \in X$.

*2) il existe $g \colon X \to [0, +\infty[$ **intégrable** telle que, pour chaque $n \in \mathbb{N}$, on ait $|f_n(x)| \leqslant g(x)$ pour presque tout $x \in X$.*

Alors f est intégrable et

$$\int f \mathrm{d}\mu = \lim_{n\to\infty} \int f_n \mathrm{d}\mu \quad \text{et} \quad \lim_{n\to\infty} \int |f - f_n| \,\mathrm{d}\mu = 0$$

Démonstration. Introduisons les ensemble mesurables suivants :

$$A = \left\{ x \in X \mid \lim_{n\to\infty} f_n(x) \text{ existe } (= f(x)) \right\}. \qquad \text{On a } \mu(A^{\complement}) = 0$$

$$B_n = \{x \in X \mid |f_n(x)| \leqslant g(x)\}, n \in \mathbb{N}. \qquad \text{On a } \mu(B_n^{\complement}) = 0$$

Soit $E = A \cap (\bigcap_{n\in\mathbb{N}} B_n)$. Alors $\mu(E^{\complement}) = 0$ car $E^{\complement} = A^{\complement} \cup \left(\bigcup_{n\in\mathbb{N}} B_n^{\complement} \right)$.

On définit $\widetilde{f}_n = f_n \mathbb{1}_E, \widetilde{f} = f \mathbb{1}_E, \widetilde{g} = g \mathbb{1}_E$.

Alors le théorème de convergence dominée s'applique aux fonctions $\widetilde{f}_n$, $\widetilde{f}$ et $\widetilde{g}$. Comme

$$\int \widetilde{f}_n \mathrm{d}\mu = \int f_n \mathrm{d}\mu \quad \text{et} \quad \int \widetilde{f} \mathrm{d}\mu = \int f \mathrm{d}\mu$$

on obtient la conclusion désirée. □

Corollaire 2.27. *Soit $f_n \colon X \to \mathbb{C}$ une suite de fonctions **intégrables** telles que*

$$\sum_{n\in\mathbb{N}} \int |f_n| \,\mathrm{d}\mu < \infty$$

Alors la série $\sum_{n\in\mathbb{N}} f_n(x)$ converge (absolument) pour presque tout $x \in X$ vers une fonction f intégrable, et on a

$$\int f \mathrm{d}\mu = \sum_{n\in\mathbb{N}} \int f_n \mathrm{d}\mu$$

Démonstration. Soit $\varphi\colon X \to [0, +\infty]$ la fonction définie par

$$\varphi(x) = \sum_{n\in\mathbb{N}} |f_n(x)|$$

Alors, par le corollaire (2.17) du théorème de convergence monotone

$$\int \varphi \mathrm{d}\mu = \sum_{n\in\mathbb{N}} \int |f_n| \mathrm{d}\mu < \infty$$

Donc φ est intégrable, et on déduit qu'il existe $E \in \mathcal{M}$ tel que $\mu(E^{\complement}) = 0$ et que l'on ait $\varphi(x) < \infty, \forall x \in E$. Ainsi, la série $\sum_{n\in\mathbb{N}} f_n(x)$ converge absolument pour tout x de E, et on peut définir

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

En applicant le théorème de convergence dominée à la suite $(\sum_{n=0}^{N} f_n)_N$, dominée par $\varphi \mathbb{1}_E$, on obtient

$$\int f \mathrm{d}\mu = \lim_{N\to\infty} \int \sum_{n=0}^{N} f_n \mathrm{d}\mu = \lim_{N\to\infty} \sum_{n=0}^{N} \int f_n \mathrm{d}\mu = \sum_{n\in\mathbb{N}} \int f_n \mathrm{d}\mu$$

□

2.5 Applications

2.5.1 Comparaison avec l'intégrale de Riemann

On suppose ici que $(X, \mathcal{M}, \mu) = (\mathbb{R}^d, \mathcal{L}(\mathbb{R}^d), \lambda)$ où $\mathcal{L}(\mathbb{R}^d)$ est la tribu de Lebesgue et λ la mesure de Lebesgue.

Définition 2.28.

- Si $I_1, I_2, \ldots, I_d \subset \mathbb{R}$ sont des intervalles bornés de $\mathbb{R}$, on appelle **pavé de** $\mathbb{R}^d$ l'ensemble

$$P = I_1 \times I_2 \times \cdots \times I_d \subset \mathbb{R}^d$$

  On note $mes(P) = \ell(I_1) \cdot \ldots \cdot \ell(I_d)$, où $\ell(I_i)$ est la longueur du segment I_i, la **mesure** du pavé P.
- On dit qu'une fonction $f\colon \mathbb{R}^d \to \mathbb{C}$ est **en escalier** si f est combinaison linéaire finie de fonctions indicatrices de pavés (en particulier, f est donc étagée) :

$$f = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{P_i} \qquad \alpha_i \in \mathbb{C}, P_i \subset \mathbb{R}^d \text{ pavés}$$

  Si f est en escalier, on définit :

$$\int f \mathrm{d}x = \sum_{i=1}^{n} \alpha_i mes(P_i) \quad \in \mathbb{C}$$

  Ici comme pour les fonctions étagées, la définition de l'intégrale ne dépend que de f et pas de sa décomposition en fonctions indicatrices de pavés.

- On dit qu'une fonction $f: \mathbb{R}^d \to \mathbb{C}$ est **intégrable au sens de Riemann** s'il existe deux suites $(f_n)_n$ et $(\varphi_n)_n$ de fonctions en escalier, $\varphi_n \geqslant 0$, telles que

$$|f - f_n| \leqslant \varphi_n \quad \forall n \in \mathbb{N} \qquad \text{et} \qquad \int \varphi_n \mathrm{d}x \underset{n\to\infty}{\longrightarrow} 0$$

On définit alors

$$\int f \mathrm{d}x = \lim_{n\to\infty} \int f_n \mathrm{d}x \tag{2.4}$$

Remarques. 1) La limite (2.4) existe, car $(\int f_n \mathrm{d}x)_n$ est une suite de Cauchy. En effet,

$$|f_n - f_m| \leqslant |f_n - f| + |f - f_m| \leqslant \varphi_n + \varphi_M$$

donc

$$\left| \int f_n \mathrm{d}x - \int f_m \mathrm{d}x \right| \leqslant \int |f_n - f_m| \, \mathrm{d}x \leqslant \int \varphi_n \mathrm{d}x + \int \varphi_m \mathrm{d}x \quad \underset{n,m\to\infty}{\longrightarrow} 0$$

2) La limite (2.4) ne dépend que de f et non des suites (f_n) et (φ_n). Supposons en effet l'existence de deux suites de fonctions en escaliers $(\widetilde{f_n})_n$ et $(\widetilde{\varphi_n})_n$ telles que

$$\left|f - \widetilde{f_n}\right| \leqslant \widetilde{\varphi_n} \quad \forall n \in \mathbb{N} \qquad \text{et} \qquad \int \widetilde{\varphi_n} \mathrm{d}x \underset{n\to\infty}{\longrightarrow} 0$$

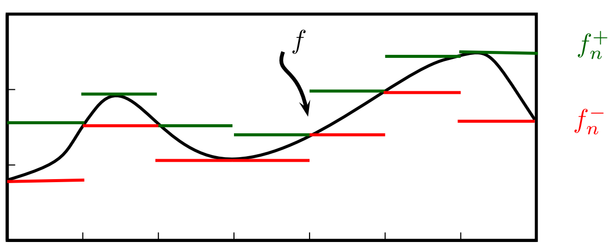
On a alors

$$\left|f_n - \widetilde{f_m}\right| \leqslant |f_n - f| + \left|f - \widetilde{f_m}\right| \leqslant \varphi_n + \widetilde{\varphi_m}$$

donc

$$\left| \int f_n \mathrm{d}x - \int \widetilde{f_m} \mathrm{d}x \right| \leqslant \int \left|f_n - \widetilde{f_m}\right| \mathrm{d}x \leqslant \int \varphi_n \mathrm{d}x + \int \widetilde{\varphi_m} \mathrm{d}x \quad \underset{n,m\to\infty}{\longrightarrow} 0$$

3) La définition (2.4) coïncide avec la définition "habituelle" à l'aide des sommes de Riemann. Considérons en effet le cas particulier où $d = 1$, et $f: \mathbb{R} \to \mathbb{R}$ sur la figure suivante.



$\int f_n^- \mathrm{d}x$ est appelée la somme de Darboux inférieure et $\int f_n^+ \mathrm{d}x$ est appelée la somme de Darboux supérieure. De plus, on a : $\forall n \in \mathbb{N} \quad f_n^- \leqslant f \leqslant f_n^+$. Ainsi, f est intégrable au sens de Riemann si $\int (f_n^+ - f_n^-) \mathrm{d}x \underset{n\to\infty}{\longrightarrow} 0$.

Lien avec ci-dessus : $f_n^+ = f_n + \varphi_n$ et $f_n^- = f_n - \varphi_n$.

Proposition 2.29. *Soit $f: \mathbb{R}^d \to \mathbb{C}$ une fonction intégrable au sens de Riemann. Alors f est mesurable pour la tribu de Lebesgue $\mathcal{L}(\mathbb{R}^d)$, f est intégrable pour la mesure de Lebesgue λ, et*

$$\underset{\text{Lebesgue}}{\int f \mathrm{d}\lambda} = \underset{\text{Riemann}}{\int f \mathrm{d}x}$$

Démonstration. Soient $(f_n)_n, (\varphi_n)_n$ deux suites de fonctions en escalier telles que

$$|f - f_n| \leqslant \varphi_n \quad \forall n \in \mathbb{N} \qquad \text{et} \qquad \int \varphi_n \mathrm{d}x \underset{n\to\infty}{\longrightarrow} 0$$

Quitte à extraire une sous-suite, on peut supposer que

$$\int \varphi_n \mathrm{d}x \leqslant 2^{-n} \quad \forall n \in \mathbb{N}$$

Notons que $\varphi_n \in \mathcal{E}_+(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ est une fonction étagée pour la tribu de Borel[3], et

$$\int \varphi_n \mathrm{d}x = \int \varphi_n \mathrm{d}\lambda \quad \forall n \in \mathbb{N}$$

Comme $\sum_{n\in\mathbb{N}} \int \varphi_n \mathrm{d}x < \infty$, alors on sait que $\sum_{n=0}^{\infty} \varphi_n(x) < \infty$ pour presque tout $x \in \mathbb{R}^d$, i.e $\forall x \in E$ où $E \in \mathcal{B}(\mathbb{R}^d)$ tel que $\lambda(E^{\complement}) = 0$.
En particulier, $\forall x \in E, \lim_{n\to\infty} \varphi_n(x) = 0$ donc on en déduit que

$$\forall x \in E, \quad \lim_{n\to\infty} f_n(x) = f(x)$$

On en déduit que f est mesurable pour la tribu de Lebesgue $\mathcal{L}(\mathbb{R}^d)$. En effet, si l'on suppose pour simplifier que f est à valeurs réelles[4] alors, $\forall a \in \mathbb{R}$

$$\left\{x \in \mathbb{R}^d \mid f(x) > a\right\} = \underbrace{\left\{x \in E \mid f(x) > a\right\}}_{\in \mathcal{B}(\mathbb{R}^d) \text{ car } f\mathbb{1}_E \text{ est borélienne}} \cup \underbrace{\left\{x \in E^{\complement} \mid f(x) > a\right\}}_{\subset E^{\complement}} \quad \in \mathcal{L}(\mathbb{R}^d)$$

Par ailleurs, comme $|f - f_0| \leqslant \varphi_0$, f est une fonction bornée et nulle en dehors d'un ensemble borné. Il existe donc $M > 0$ et un pavé $P \subset \mathbb{R}^d$ tel que $|f| \leqslant M\mathbb{1}_P$. f est donc intégrable pour la mesure λ.
Quitte à remplacer f_n par $(\chi_M \circ f_n)\mathbb{1}_P$ avec $\chi_M = \begin{cases} z & \text{si } |z| \leqslant M \\ M\dfrac{z}{|z|} & \text{si } |z| > M \end{cases}$, on peut supposer que $|f_n| \leqslant M\mathbb{1}_P, \forall n \in \mathbb{N}$ (et on a toujours que $|f - f_n| \leqslant \varphi_n$).
En appliquant le théorème de convergence dominée, on trouve donc :

$$\int f \mathrm{d}\lambda \underset{\text{convergence dominée}}{=} \lim_{n\to\infty} \int f_n \mathrm{d}\lambda = \lim_{n\to\infty} \int f_n \mathrm{d}x \underset{\text{définition}}{=} \int f \mathrm{d}x$$

□

Proposition 2.30. *Une fonction $f : \mathbb{R}^d \to \mathbb{C}$ est intégrable au sens de Riemann si et seulement si :*

a) f est bornée, et nulle en dehors d'un ensemble borné, ***et,***

b) l'ensemble des points de discontinuité de f est négligeable (pour la mesure de Lebesgue sur $\mathcal{L}(\mathbb{R}^d)$).

Démonstration. 1) Supposons que f est intégrable au sens de Riemann, et soient $(f_n)_{n\in\mathbb{N}}, (\varphi_n)_{n\in\mathbb{N}}$ des suites comme ci-dessus, telles que

$$\int \varphi_n \mathrm{d}x \leqslant 2^{-n} \quad \forall n \in \mathbb{N}$$

3. On verra dans la démonstration qu'il est nécessaire de compléter la tribu pour avoir une tribu de Lebesgue
4. Le cas complexe tombe tout de suite en séparant parties réelles et imaginaires.

On sait déjà que f est bornée et nulle en dehors d'un borné.
Pour tout $n \in \mathbb{N}$, on note $A_n \subset \mathbb{R}^d$ l'ensemble des points de discontinuité de f_n ou φ_n. A_n est une union finie de pavés de mesure nulle donc $\lambda(A_n) = 0, \forall n \in \mathbb{N}$.
Soit $A = \bigcup_{n\in\mathbb{N}} A_n$. On a encore $\lambda(A) \leqslant \sum_{n\in\mathbb{N}} \lambda(A_n) = 0$.
Par ailleurs, soit $B \subset \mathbb{R}^d$ l'ensemble des points où $\varphi_n(x)$ ne converge pas vers 0 quand $n \longrightarrow \infty$. On a vu que $\lambda(B) = 0$.
On prétend que f est continue en tout point $x_0 \in A^\complement \cap B^\complement$. En effet, comme $x_0 \notin A$, les fonctions f_n et φ_n sont constantes dans un voisinage de taille $\delta_n > 0$ de x_0. Ainsi, si $|x - x_0| < \delta_n$, on a :

$$\begin{aligned} |f(x) - f(x_0)| &\leqslant |f(x) - f_n(x)| + \underbrace{|f_n(x) - f_n(x_0)|}_{=0 \text{ car } f_n \text{ constante}} + |f_n(x_0) - f(x_0)| \\ &\leqslant \varphi_n(x) + \varphi_n(x_0) = 2\varphi_n(x_0) \underset{n\to\infty}{\longrightarrow} 0 \end{aligned}$$

Donc f est continue en x_0.

2) Soit $f \colon \mathbb{R}^d \to \mathbb{R}$ vérifiant les conditions $a)$ et $b)$ ci-dessus (le cas complexe se traite de même).
On peut supposer que $|f| \leqslant M\mathbb{1}_P$ avec $M > 0$ et $P \subset \mathbb{R}^d$ un pavé. Pour tout $n \in \mathbb{N}$, soit σ_n la subdivision de P en 2^{nd} pavés de mesures égales. On note A_n la frontière de tous ces pavés, et $A = \bigcup_{n\in\mathbb{N}} A_n$. On a encore $\lambda(A) = 0$.
On définit deux suites $(f_n^+)_n$ et $(f_n^-)_n$ de fonctions en escalier par

$$f_n^+(x) = \begin{cases} 0 & \text{si } x \notin P \\ M & \text{si } x \in A_n \\ \sup_{y\in p}(f(y)) & \text{si } x \in \mathring{p} \\ & \text{ avec } p \in \sigma_n \end{cases} \qquad f_n^-(x) = \begin{cases} 0 & \text{si } x \notin P \\ -M & \text{si } x \in A_n \\ \inf_{y\in p}(f(y)) & \text{si } x \in \mathring{p} \\ & \text{ avec } p \in \sigma_n \end{cases}$$

Alors

$$f_n^-(x) \leqslant f(x) \leqslant f_n^+(x) \qquad \forall x \in \mathbb{R}^d, \forall n \in \mathbb{N}$$

Soit $B \subset \mathbb{R}^d$ l'ensemble des points de discontinuité de f. On a $\lambda(B) = 0$ par hypothèse. Soit $x_0 \in A^\complement \cap B^\complement$, on prétend que

$$f_n^+(x_0) - f_n^-(x_0) \underset{n\to\infty}{\longrightarrow} 0$$

C'est évident si $x_0 \notin P$. Si $x_0 \in P$, alors pour chaque $n \in \mathbb{N}$, $x_0 \in \mathring{p}$, $p \in \sigma_n$ avec $\delta_n = diam(p) \underset{n\to\infty}{\longrightarrow} 0$.
Comme f est continue en x_0, on a :

$$\sup_{\|y - x_0\| < \delta_n} |f(y) - f(x_0)| \underset{n\to\infty}{\longrightarrow} 0$$

Donc

$$f_n^+(x_0) - f_n^-(x_0) \underset{n\to\infty}{\longrightarrow} 0$$

Par le théorème de convergence dominée, comme $|f_n^+(x) - f_n^-(x)| \leqslant 2M\mathbb{1}_P$,

$$\int (f_n^+(x) - f_n^-(x))\mathrm{d}x \underset{n\to\infty}{\longrightarrow} 0$$

□

Remarque. Une fonction intégrable au sens de Riemann n'est pas nécessairement borélienne.
Si E est un sous-ensemble non borélien d'un ensemble de Cantor K de mesure nulle, alors $\mathbb{1}_E$ est intégrable-Riemann, mais non borélienne.

Définition 2.31. On dit que $E \subset \mathbb{R}^d$ est **quarrable** si $\mathbb{1}_E$ est intégrable-Riemann.

Par la proposition précédente, $E \subset \mathbb{R}^d$ est quarrable $\iff$ E est borné et $\lambda(\partial E) = 0$.
Par exemple un ensemble de Cantor K est quarrable $\iff$ $\lambda(K) = 0$ (car $K = \partial K$).

Remarque sur les intégrales "semi-convergentes"

1) Soient $X = \mathbb{N}, \mathcal{M} = \mathcal{P}(\mathbb{N})$ et μ la mesure de comptage. Si $f = (f_n)_n$, $f \colon \mathbb{N} \to \mathbb{C}$, alors f est intégrable si et seulement si $\sum_{n\in\mathbb{N}} |f_n| < \infty$, et, dans ce cas, $\displaystyle\int f \mathrm{d}\mu = \sum_{n\in\mathbb{N}} f_n$.

   Contre exemple : si $f_n = \dfrac{(-1)^n}{n+1}$, alors $\sum_{n\in\mathbb{N}} |f_n| = +\infty$, donc f n'est pas intégrable. Mais

$$\lim_{N\to\infty} \sum_{n=0}^{N} f_n = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots = \ln(2)$$

2) Soit $X = [0, +\infty[$, $\mathcal{M} = \mathcal{B}(X)$ et λ la mesure de Lebesgue. Si $f \colon X \to \mathbb{R}$ est définie par

$$f(x) = \begin{cases} \dfrac{\sin x}{x} & \text{si } x > 0 \\ 1 & \text{sinon} \end{cases}$$

   alors f est continue [5] sur X (donc mesurable), et

$$\int_0^\infty |f| \,\mathrm{d}\lambda = \int_0^\infty \frac{|\sin x|}{x} \mathrm{d}x = +\infty$$

   Ainsi, f n'est pas intégrable au sens de Lebesgue. Cependant,

$$\lim_{n\to\infty} \int_0^n \frac{\sin x}{x} \mathrm{d}x = \frac{\pi}{2}$$

   Ces exemples montrent que le cadre de Lebesgue n'est pas "aussi" général qu'on pourrait le croire. Les intégrales généralisées n'y appartiennent pas.

2.5.2 Intégrales dépendant d'un paramètre

Soit $(X, \mathcal{M}, \mu)$ un espace mesuré, et soit (Λ, d) un espace métrique.
On considère une fonction

$$f \colon \begin{array}{ccc} X \times \Lambda & \longrightarrow & \mathbb{C} \\ (x, \lambda) & \longmapsto & f(x, \lambda) \end{array}$$

et on suppose que la fonction $x \mapsto f(x, \lambda)$ est intégrable sur X (pour la mesure μ).
On peut donc définir la fonction $F \colon \Lambda \longrightarrow \mathbb{C}$ par

$$F(\lambda) = \int_X f(x, \lambda) \mathrm{d}\mu_x \equiv \int_X f(x, \lambda) \mathrm{d}x \qquad \forall \lambda \in \Lambda$$

5. Exercice classique. Voir le taux de variation de la fonction sinus en 0.

Afin d'harmoniser la notation des intégrales, on notera, pour une fonction $g\colon X \to \mathbb{C}$, de façon équivalente

$$\int g\mathrm{d}\mu = \int g(x)\mathrm{d}\mu_x = \int g(x)\mathrm{d}x$$

On veut étudier la continuité et la dérivabilité de F.

Théorème 2.32 (Théorème de continuité). *On suppose :*

i) $\forall \lambda \in \Lambda$, *la fonction* $x \mapsto f(x,\lambda)$ *est intégrable sur* X

ii) Pour μ*-presque tout* $x \in X$, *la fonction* $\lambda \mapsto f(x,\lambda)$ *est continue sur* Λ

iii) Il existe $g\colon X \to \mathbb{R}_+$ ***intégrable*** *telle que,* $\forall \lambda \in \Lambda$, *on ait*

$$|f(x,\lambda)| \leqslant g(x) \quad \text{pour } \mu\text{-presque tout } x \in X$$

Alors $F\colon \Lambda \to \mathbb{C}$ *définie par* $F(\lambda) = \int_X f(x,\lambda)\mathrm{d}x$ *est continue sur* Λ.

Remarques. 1) Dans $i)$ il suffit de demander que $x \mapsto f(x,\lambda)$ soit mesurable, la condition $iii)$ impliquant l'intégrabilité.

2) Si dans $ii)$ on suppose seulement que $\lambda \mapsto f(x,\lambda)$ est continue en un point $\lambda_0 \in \Lambda$, on obtient que F est continue en λ_0.

Démonstration. L'hypothèse $i)$ garantit que $F\colon \Lambda \to \mathbb{C}$ est bien définie. Soit $\lambda_0 \in \Lambda$, et soit $(\lambda_n)_{n\geqslant 1}$ une suite dans Λ telle que $d(\lambda_n,\lambda_0) \underset{n\to\infty}{\longrightarrow} 0$.

Par $ii)$

$$f(x,\lambda_n) \underset{n\to\infty}{\longrightarrow} f(x,\lambda_0) \qquad \text{pour } \mu\text{-presque tout } x \in X$$

Par $iii)$,

$$\forall n \in \mathbb{N}^*, |f(x,\lambda_n)| \leqslant g(x) \qquad \text{pour } \mu\text{-presque tout } x \in X$$

Par le théorème de converge dominée, on conclut :

$$f(\lambda_n) = \int_X f(x,\lambda_n)\mathrm{d}x \underset{n\to\infty}{\longrightarrow} \int_X f(x,\lambda_0)\mathrm{d}x = F(\lambda_0)$$

Comme λ_0 et la suite $(\lambda_n)_n$ étaient arbitraires, on a le résultat voulu. □

On suppose désormais que $\Lambda \subset \mathbb{R}$ est un intervalle d'intérieur non vide pour parler de dérivabilité.

Théorème 2.33 (Théorème de dérivabilité). *On suppose*

i) Pour tout $\lambda \in \Lambda$, *la fonction* $x \mapsto f(x,\lambda)$ *est intégrable sur* X

ii) Pour μ*-presque tout* $x \in X$, *la fonction* $\lambda \mapsto f(x,\lambda)$ *est dérivable sur* Λ

iii) Il existe $g\colon X \to \mathbb{R}_+$ *intégrable telle que pour* μ*-presque tout* $x \in X$ *on ait*

$$\left|\frac{\partial}{\partial\lambda} f(x,\lambda)\right| \leqslant g(x) \qquad \forall \lambda \in \Lambda$$

Alors la fonction $\Lambda: \to \mathbb{C}$ *définie par* $F(\lambda) = \int_X f(x, \lambda) \mathrm{d}x$ *est dérivable sur* Λ *et*

$$F'(\lambda) = \int_X \frac{\partial}{\partial \lambda} f(x, \lambda) \mathrm{d}x$$

Remarques. 1) $\frac{\partial}{\partial \lambda} f(x, \lambda)$ est définie presque partout en x, et là où elle n'est pas définie, on lui met la valeur 0.

2) Si $\Lambda = [a, b]$, "dérivable sur Λ" signifie : dérivable sur $]a, b[$, dérivable à droite en a et dérivable à gauche en b.

3) Même si on ne souhaite la dérivabilité de F qu'en un point $\lambda_0 \in \Lambda$, il faut quand même supposer $iii)$ pour tout $\lambda \in \Lambda$.

Démonstration. Soit $\lambda_0 \in \Lambda$, et soit $(\lambda_n)_{n \geqslant 1}$ une suite dans Λ telle que $\lambda_n \neq \lambda_0, \forall n \geqslant 1$ et $\lambda_n \underset{n \to \infty}{\longrightarrow} \lambda_0$. Pour tout $n \in \mathbb{N}^*$, on définit :

$$h_n(x) = \frac{1}{\lambda_n - \lambda_0} (f(x, \lambda_n) - f(x, \lambda_0)), \quad x \in X$$

Par $ii)$, il existe un ensemble $A \in \mathcal{M}$, avec $\mu(A^{\complement}) = 0$ tel que $\forall x \in A, \lambda \mapsto f(x, \lambda)$ est dérivable. Donc, $\forall x \in A, h_n(x) \underset{n \to \infty}{\longrightarrow} \frac{\partial}{\partial \lambda} f(x, \lambda_0)$. On définit

$$h(x) = \begin{cases} \frac{\partial}{\partial \lambda} f(x, \lambda_0) & \text{si } x \in A \\ 0 & \text{si } x \notin A \end{cases}$$

Par $iii)$ il existe $B \in \mathcal{M}, B \subset A$ avec $\mu(B^{\complement}) = 0$ tel que

$$\forall x \in B, \quad |h_n(x)| \leqslant \sup_{\lambda \in \Lambda} \left| \frac{\partial}{\partial \lambda} f(x, \lambda) \right| \leqslant g(x) \qquad \forall n \in \mathbb{N}$$

Par le théorème de la convergence dominée, on a donc

$$\frac{F(\lambda_n) - F(\lambda_0)}{\lambda_n - \lambda_0} = \int_X h_n(x) \mathrm{d}x \underset{n \to \infty}{\longrightarrow} \int_X h(x) \mathrm{d}x$$

Comme la suite $(\lambda_n)_{n \geqslant 1}$ est arbitraire, on conclut que F est dérivable en λ_0 et

$$F'(\lambda_0) = \int_X h(x) \mathrm{d}x$$

□

2.5.3 Application : la fonction Γ d'Euler

Pour tout $\alpha > 0$, on définit

$$\boxed{\Gamma(\alpha) = \int_0^{\infty} x^{\alpha - 1} e^{-x} \mathrm{d}x}$$

C'est une intégrale dépendant d'un paramètre, on va ainsi appliquer les théorèmes de la partie précédente.

Proposition 2.34. *Pour tout $\alpha > 0$, $\Gamma(\alpha) \in \,]0, +\infty[$ est bien défini, et la fonction $\Gamma\colon \,]0, +\infty[\, \to \,]0, +\infty[$ est continue.*

Remarque. Pour se ramener au cas général du cours, on peut choisir
- soit $X = \,]0, +\infty[\,, \mathcal{M} = \mathcal{B}(X)$ et μ la mesure de Lebesgue
- soit $X = \mathbb{R}, \mathcal{M} = \mathcal{B}(\mathbb{R})$ et μ la mesure de Lebesgue, et on prolonge toutes les fonctions que l'on considère par zéro sur $]-\infty, 0]$.

Ces deux points de vue sont possibles et sont strictement équivalents. Nous choisirons le premier dans la suite.

Démonstration. Pour tout $\alpha > 0$, la fonction $x \mapsto x^{\alpha-1}e^{-x}$ est continue sur $X = \,]0, +\infty[$, donc mesurable, et à valeurs positives. On peut donc définir

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}\mathrm{d}x \in [0, +\infty]$$

Cette quantité est-elle finie ou infinie ?
On va voir que cette quantité est finie. On utilise la linéarité de l'intégrale pour faire les calculs sur $[0, 1]$ et sur $]1, +\infty[$.
Par le théorème de convergence monotone,

$$\int_1^\infty x^{\alpha-1}e^{-x}\mathrm{d}x = \lim_{N\to\infty}\int_1^N x^{\alpha-1}e^{-x}\mathrm{d}x \equiv \lim_{N\to\infty}\int_0^\infty \underbrace{x^{\alpha-1}e^{-x}\mathbb{1}_{[1,N]}(x)}_{\text{suite croissante de fonctions}}\ \mathrm{d}x$$

or

$$\int_1^N x^{\alpha-1}e^{-x}\mathrm{d}x = \int_1^N (x^{\alpha-1}e^{-\frac{x}{2}})e^{-\frac{x}{2}}\mathrm{d}x$$

Comme $x \mapsto x^{\alpha-1}e^{-\frac{x}{2}}$ est une fonction continue tendant vers 0 en $+\infty$, elle est bornée. Il existe une constante C_α telle que $\forall x \in \,]1, \infty[\,, x^{\alpha-1}e^{-\frac{x}{2}} \leqslant C_\alpha$. Ainsi,

$$\begin{aligned}\int_1^N x^{\alpha-1}e^{-x}\mathrm{d}x &\leqslant \int_1^N C_\alpha e^{-\frac{x}{2}}\mathrm{d}x \\ &= C_\alpha(-2)\left[e^{-\frac{x}{2}}\right]_{x=1}^{x=N} \\ &= 2C_\alpha(e^{-\frac{1}{2}} - e^{-\frac{N}{2}}) \\ &\leqslant 2C_\alpha\end{aligned}$$

On en déduit donc que

$$\int_1^\infty x^{\alpha-1}e^{-x}\mathrm{d}x < \infty$$

Par le même théorème que précédemment,

$$\int_0^1 x^{\alpha-1}e^{-x}\mathrm{d}x = \lim_{\varepsilon\to0}\int_\varepsilon^1 x^{\alpha-1}e^{-x}\mathrm{d}x$$

or

$$\begin{aligned}\int_\varepsilon^1 x^{\alpha-1}e^{-x}\mathrm{d}x &\leqslant \int_\varepsilon^1 x^{\alpha-1}\mathrm{d}x\\ &= \frac{1}{\alpha}\left[x^\alpha\right]_{x=\varepsilon}^{x=1}\\ &= \frac{1}{\alpha}(1-\varepsilon^\alpha)\\ &\leqslant \frac{1}{\alpha}\end{aligned}$$

On en déduit donc que

$$\int_0^1 x^{\alpha-1}e^{-x}\mathrm{d}x < \infty$$

Finalement, par linéarité de l'intégrale

$$\int_0^\infty x^{\alpha-1}e^{-x}\mathrm{d}x < \infty$$

(Si $\alpha \leqslant 0$, l'intégrale vaudrait $+\infty$)
On voudrait dominer uniformément par rapport à α, mais c'est impossible vu les dominations obtenues.
Pour montrer la continuité de Γ, on fixe $0 < \alpha_0 < \alpha_1 < \infty$ et on pose $\Lambda = [\alpha_0, \alpha_1]$.
Pour $\alpha \in \Lambda$, la fonction $x \mapsto x^{\alpha-1}e^{-x}$ est intégrable sur $]0, +\infty[$ et pour tout $x > 0$, la fonction $\alpha \mapsto x^{\alpha-1}e^{-x}$ est continue (et même $\mathcal{C}^\infty$).
$\forall \alpha \in \Lambda, \forall x \in]0, +\infty[$, on a

$$0 \leqslant x^{\alpha-1}e^{-x} \leqslant g(x) = \begin{cases} C_{\alpha_1}e^{-\frac{x}{2}} & \text{si } x \geqslant 1\\ x^{\alpha_0-1} & \text{si } 0 < x < 1\end{cases}$$

g est intégrable sur $]0, +\infty[$.
Par le théorème de continuité 2.32, la fonction Γ est continue sur $\Lambda = [\alpha_0, \alpha_1]$, et comme α_0 et α_1 étaient arbitraires, Γ est continue sur $]0, +\infty[$. □

Proposition 2.35. $\Gamma \in \mathcal{C}^\infty(]0, +\infty[)$ *et pour tout k entier naturel, on a*

$$\Gamma^{(k)}(\alpha) = \int_0^\infty (\ln x)^k x^{\alpha-1}e^{-x}\mathrm{d}x \qquad \forall \alpha > 0$$

Corollaire 2.36. $\Gamma\colon\ ]0, +\infty[\ \to\ ]0, +\infty[$ *est strictement convexe.*

Démonstration. En effet,

$$\Gamma''(\alpha) = \int_0^\infty \underbrace{(\ln x)^2 x^{\alpha-1}e^{-x}}_{>0}\mathrm{d}x > 0$$

□

Remarque. Pour tout $\varepsilon > 0$, il existe une constante $C_\varepsilon > 0$ telle que

$$\left|\begin{array}{lll} x^\varepsilon\,|\ln x| & \leqslant C_\varepsilon & \forall x \in]0, 1]\\ x^{-\varepsilon}\,|\ln x| & \leqslant C_\varepsilon & \forall x \geqslant 1\end{array}\right.$$

Démonstration (Démonstration de la proposition 2.35). Pour tout $k \in \mathbb{N}$, on définit $f_k\colon\ ]0,+\infty[^2 \to \mathbb{R}$ par

$$f_k(x,\alpha) = (\ln x)^k x^{\alpha-1} e^{-x}$$

f_k est une fonction de classe $\mathcal{C}^\infty$ par rapport à x et α, et

$$\frac{\partial}{\partial\alpha} f_k(x,\alpha) = f_{k+1}(x,\alpha) \qquad \forall x,\alpha > 0$$

Lemme 2.37. *Soit $0 < \alpha_0 < 1 < \alpha_1$. Pour tout $k \in \mathbb{N}$, il existe $h_k\colon\ ]0,+\infty[\to \mathbb{R}_+$ intégrable telle que*

$$\sup_{\alpha_0 \leqslant \alpha \leqslant \alpha_1} |f_k(x,\alpha)| \leqslant h_k(x) \qquad \forall x > 0$$

Démonstration (Démonstration du lemme). Choisissons $\varepsilon > 0$ tel que $\varepsilon k < \alpha_0$. Soit $\alpha \in [\alpha_0,\alpha_1]$.

– Si $0 < x \leqslant 1$, on a :

$$\left|(\ln x)^k x^{\alpha-1} e^{-x}\right| \leqslant (C_\varepsilon x^{-\varepsilon})^k x^{\alpha_0-1} = C_\varepsilon^k x^{\overbrace{\alpha_0 - \varepsilon k}^{>0} -1}$$

On a majoré par une fonction intégrable sur $]0,1]$.

– Si $x \geqslant 1$, on a :

$$\left|(\ln x)^k x^{\alpha-1} e^{-x}\right| \leqslant (C_\varepsilon x^{\varepsilon})^k x^{\alpha_1-1} e^{-x} = C_\varepsilon^k x^{\alpha_1+\varepsilon k-1} e^{-x}$$

On a majoré par une fonction intégrable sur $]1,+\infty[$.

Il suffit donc de prendre

$$h_k(x) = C_\varepsilon^k \left(x^{\alpha_0-\varepsilon k-1} \mathbb{1}_{]0,1]}(x) + x^{\alpha_1+\varepsilon k-1} e^{-x} \mathbb{1}_{]1,+\infty[}(x)\right)$$ □

Grâce au lemme, on peut donc définir pour tout $k \in \mathbb{N}$

$$F_k(\alpha) = \int_0^\infty f_k(x,\alpha)\mathrm{d}x, \qquad \alpha \in [\alpha_0,\alpha_1]$$

En outre, pour tout $\alpha \in [\alpha_0,\alpha_1]$ on a

$$\left|\frac{\partial}{\partial\alpha} f(x,\alpha)\right| = |f_{k+1}(x,\alpha)| \leqslant h_{k+1}(x) \qquad h_{k+1} \text{ intégrable}$$

Par le théorème de dérivation 2.33, F_k est dérivable sur $[\alpha_0,\alpha_1]$ et

$$F_k'(\alpha) = \int_0^\infty \frac{\partial}{\partial\alpha} f_k(x,\alpha)\mathrm{d}x = \int_0^\infty f_{k+1}(x,\alpha)\mathrm{d}x = F_{k+1}(\alpha)$$

Comme α_0 et α_1 étaient arbitraires, ceci montre que F_k est de classe $\mathcal{C}^1$ sur $]0,+\infty[$ et $F_k' = F_{k+1}$. Comme $F_0 = \Gamma$, on a le résultat. □

Proposition 2.38.

$$\Gamma(\alpha+1) = \alpha\Gamma(\alpha) \qquad \forall\alpha > 0$$

Démonstration. Soit $0 < \varepsilon < M$. En repassant aux intégrales de Riemann, on a

$$\begin{aligned}\int_\varepsilon^M x^\alpha e^{-x}\mathrm{d}x &= \left[-x^\alpha e^{-x}\right]_{x=\varepsilon}^{x=M} + \int_\varepsilon^M \alpha x^\alpha e^{-x}\mathrm{d}x \\ &= \varepsilon^\alpha e^{-\varepsilon} - M^\alpha e^{-M} + \alpha\int_\varepsilon^M x^{\alpha-1}e^{-x}\mathrm{d}x\end{aligned}$$

En prenant la limite $\varepsilon \to 0, M \to \infty$, on trouve :

$$\Gamma(\alpha+1) = \int_0^\infty x^\alpha e^{-x}\mathrm{d}x = \alpha\int_0^\infty x^{\alpha-1}e^{-x}\mathrm{d}x = \alpha\Gamma(\alpha)$$

□

Il y aura un développement ultérieur dans le cours pour traiter ceci plus rapidement, sans se ramener aux intégrales de Riemann.

Conséquences: *Valeurs particulières sur les entiers et demi-entiers*

– *Calculons la valeur de Γ sur les entiers naturels non nuls.*

$$\Gamma(1) = \int_0^\infty e^{-x}\mathrm{d}x = 1, \quad \Gamma(2) = 1\Gamma(1) = 1, \quad \Gamma(3) = 2\Gamma(2) = 2, \quad \Gamma(4) = 3\Gamma(3) = 6$$

Par récurrence, on a donc

$$\boxed{\Gamma(n+1) = n! \quad \forall n \in \mathbb{N}}$$

– *Par ailleurs,*

$$\begin{aligned}\Gamma(\frac{1}{2}) &= \int_0^\infty \frac{1}{\sqrt{x}}e^{-x}\mathrm{d}x \\ &\overset{x=y^2}{=} \int_0^\infty \frac{1}{y}e^{-y^2}2y\mathrm{d}y \\ &= \int_\mathbb{R} e^{-y^2}\mathrm{d}y \\ &= \sqrt{\pi}\end{aligned}$$

C'est l'intégrale de Gauss. Ce calcul est justifié dans la partie 4.4.2, page 64.
Il suffit ensuite d'appliquer la formule de la proposition précédente et

$$\Gamma(\frac{3}{2}) = \frac{1}{2}\sqrt{\pi}, \qquad \ldots$$

Comportement aux limites

- Si $\alpha \to 0$, alors

$$\Gamma(\alpha) = \frac{\Gamma(\alpha+1)}{\alpha} \sim \frac{1}{\alpha}$$

- Si $\alpha \to \infty$, alors

$$\begin{aligned}\Gamma(\alpha+1) &= \int_0^\infty x^\alpha e^{-x}\mathrm{d}x \\ &\overset{x=\alpha+y}{=} \int_{-\alpha}^\infty (\alpha+y)^\alpha e^{-(\alpha+y)}\mathrm{d}y \\ &= \alpha^\alpha e^{-\alpha}\int_{-\alpha}^\infty (1+\frac{y}{\alpha})^\alpha e^{-y}\mathrm{d}y \\ &\overset{y=\sqrt{\alpha}z}{=} \alpha^\alpha e^{-\alpha}\int_{-\sqrt{\alpha}}^\infty (1+\frac{z}{\sqrt{\alpha}})^\alpha e^{-z\sqrt{\alpha}}\sqrt{\alpha}\mathrm{d}z\end{aligned}$$

Ainsi,

$$\Gamma(\alpha+1) = \sqrt{\alpha}e^{-\alpha}\alpha^{\alpha}\Delta(\alpha)$$

où

$$\Delta(\alpha) = \int_{-\sqrt{\alpha}}^{\infty}(1+\frac{z}{\sqrt{\alpha}})^{\alpha}e^{-z\sqrt{\alpha}}\mathrm{d}z$$

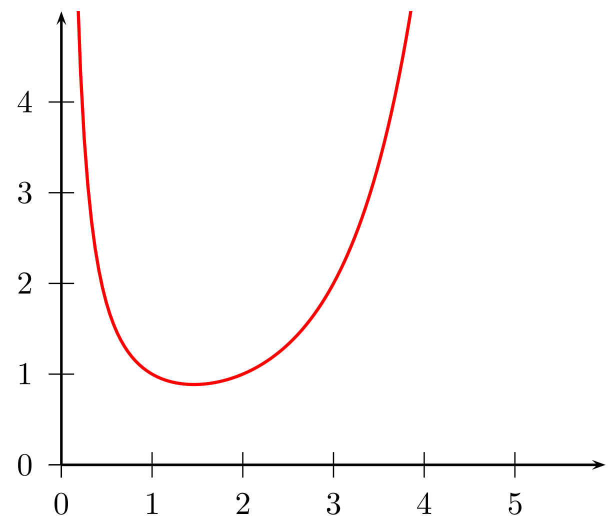


FIGURE 2.1 – La fonction Γ d'Euler

Proposition 2.39.

$$\Delta(\alpha) \underset{n\to\infty}{\longrightarrow} \sqrt{2\pi}$$

Corollaire 2.40. *On obtient la* ***formule de Stirling***

$$\Gamma(\alpha+1) \sim \sqrt{2\pi\alpha}\alpha^{\alpha}e^{-\alpha} \qquad \text{quand } \alpha\to\infty$$

Démonstration (Démonstration de la proposition 2.39). A $z\in\mathbb{R}$ fixé, on a

$$\left(1+\frac{z}{\sqrt{\alpha}}\right)^{\alpha}e^{-z\sqrt{\alpha}} = \exp\left(\alpha\ln\left(1+\frac{z}{\sqrt{\alpha}}\right)-z\sqrt{\alpha}\right)$$

On fait un développement limité [6] de $\ln(1+\frac{z}{\sqrt{\alpha}})$ quand α tend vers l'infini

$$\alpha\ln\left(1+\frac{z}{\sqrt{\alpha}}\right)-z\sqrt{\alpha} = \alpha\left(\frac{z}{\sqrt{\alpha}}-\frac{1}{2}\frac{z^2}{\alpha}+\mathcal{O}\left(\frac{|z|^3}{\alpha^{\frac{3}{2}}}\right)-\frac{z}{\sqrt{\alpha}}\right)$$

6. **Rappel :** $\ln(1+x) = x-\frac{1}{2}x^2+\mathcal{O}(x^3)$ quand $x\to 0$

Donc

$$\left(1+\frac{z}{\sqrt{\alpha}}\right)^{\alpha} \underset{\alpha\to\infty}{\longrightarrow} e^{-\frac{z^2}{2}}$$

On veut appliquer le théorème de convergence dominée pour calculer $\Delta(\alpha)$. On a besoin du lemme suivant pour la domination.

Lemme 2.41. *On a la propriété suivante*

$$\ln(1+x) \leqslant \begin{cases} \dfrac{3x}{4} & \text{si } x \geqslant 1 \\ x - \dfrac{x^2}{4} & \text{si } |x| < 1 \end{cases}$$

Démonstration (Démonstration du lemme). En effet,

– Si $x \geqslant 1$, on a par concavité de $t \mapsto \ln(1+t)$

$$\ln(1+x) \leqslant \ln(2)x \leqslant \frac{3x}{4}$$

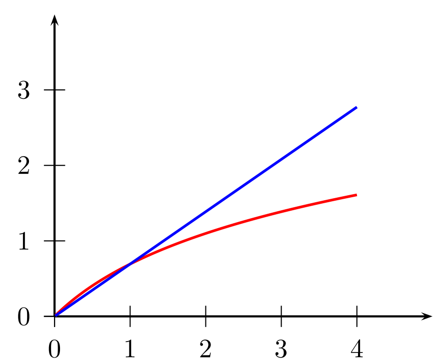


FIGURE 2.2 – Représentations de $y = \ln(1+x)$ et $y = (\ln 2)x$

– Si $0 \leqslant x \leqslant 1$, on considère $g(x) = x - \dfrac{x^2}{4} - \ln(1+x)$. On a $g(0) = 0$, et

$$\forall x \in [0,1]\,, g'(x) = 1 - \frac{x}{2} - \frac{1}{1+x} = \frac{x(1-x)}{2(1+x)} \geqslant 0$$

Ainsi, $\forall x \in [0,1]\,, g(x) \geqslant 0$.
N.B : $g(1) = \frac{3}{4} - \ln 2 \geqslant 0$ et on retrouve l'inégalité du premier cas.

– Si $-1 < x < 0$, on a

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots$$

C'est une série à termes négatifs, donc majorée par les sommes partielles et

$$\ln(1+x) \leqslant x - \frac{x^2}{2} \leqslant x - \frac{x^2}{4}$$ □

Notons $g(z,\alpha) = \left(1+\frac{z}{\sqrt{\alpha}}\right)^{\alpha} e^{-z\sqrt{\alpha}}\mathbb{1}_{[-\sqrt{\alpha},+\infty[}(z)$. On obtient donc

$$\Delta(\alpha) = \int_{\mathbb{R}} g(z,\alpha)\mathrm{d}z$$

On sait déjà que

$$g(z,\alpha) \underset{\alpha\to\infty}{\longrightarrow} e^{-\frac{z^2}{2}} \qquad \forall z \in \mathbb{R}$$

D'autre part,
- Si $z \geqslant \sqrt{\alpha}$

$$\begin{aligned} g(z,\alpha) = \exp\left(\alpha \ln\left(1+\frac{z}{\sqrt{\alpha}}\right) - z\sqrt{\alpha}\right) &\leqslant \exp\left(\alpha\frac{3z}{4\sqrt{\alpha}} - \sqrt{\alpha}z\right) \\ &= \exp\left(-\frac{1}{4}\sqrt{\alpha}z\right) \\ &\leqslant \exp\left(-\frac{|z|}{4}\right) \qquad \text{si } \alpha \geqslant 1 \end{aligned}$$

- Si $|z| \leqslant \sqrt{\alpha}$,

$$\begin{aligned} g(z,\alpha) = \exp\left(\alpha \ln\left(1+\frac{z}{\sqrt{\alpha}}\right) - z\sqrt{\alpha}\right) &\leqslant \exp\left(\alpha\left(\frac{z}{\sqrt{\alpha}} - \frac{z^2}{4\alpha}\right) - z\sqrt{\alpha}\right) \\ &= \exp\left(-\frac{z^2}{4}\right) \end{aligned}$$

Conclusion : Si $\alpha \geqslant 1$,

$$g(z,\alpha) \leqslant h(z) = \max\left(e^{-\frac{|z|}{4}}, e^{-\frac{z^2}{4}}\right)$$

Ainsi,

$$\Delta(\alpha) = \int_{\mathbb{R}} g(z,\alpha)\mathrm{d}z \underset{\alpha\to\infty}{\longrightarrow} \int_{\mathbb{R}} e^{-\frac{z^2}{2}}\mathrm{d}z = \sqrt{2\pi}$$

□

CHAPITRE 3

MESURE DE LEBESGUE SUR $\mathbb{R}^d$

Dans le chapitre 1, nous avons parlé de la théorie abstraire de la mesure, et donné l'existence (et l'unicité) de la mesure de Lebesgue, ce qui nous a permis de faire toute une théorie de l'intégration (chapitre 2). Il reste donc à construire la mesure de Lebesgue, et c'est le sujet de ce chapitre 3.

3.1 MESURES EXTÉRIEURES

Définition 3.1. Soit X un ensemble quelconque. On appelle **mesure extérieure** sur X une application $\mu^* : \mathcal{P}(X) \longrightarrow [0, +\infty]$ telle que

i) $\mu^*(\varnothing) = 0$

ii) μ^* est croissante :

$$\mu^*(A) \leqslant \mu^*(B) \qquad \text{si } A \subset B$$

iii) μ^* est sous-additive : si $\{A_n\}_{n\in\mathbb{N}}$ est une famille de parties de X, alors

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

Les propriétés d'une mesure extérieure sont moins contraignantes que celles d'une mesure. EN particulier, une mesure extérieure est définie sur toutes les parties de l'ensemble X alors que les mesures sont définies sur des tribus.
Dans cette partie, nous allons voir comment à partir d'une mesure extérieure μ^* on construit une mesure sur une tribu qui dépend de μ^*. Et c'est ainsi qu'on construira la mesure de Lebesgue.

Définition 3.2. Soit X un ensemble muni d'une mesure extérieure μ^*. On dit qu'une partie $B \subset X$ est μ^***-régulière** si, pour toute partie A de X, on a

$$\mu^*(A) = \mu^*(A \cap B) + \mu^*(A \cap B^{\complement})$$

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

Remarque. L'inégalité $\mu^*(A) \leqslant \mu^*(A \cap B) + \mu^*(A \cap B^{\complement})$ provient trivialement de la sous-additivité de la mesure extérieure. Pour vérifier que B est μ^*-régulière, c'est l'inégalité inverse qu'il importe de vérifier.

Proposition 3.3. *$\mathcal{M}(\mu^*)$ est une tribu sur X contenant toutes les parties $B \subset X$ telles que $\mu^*(B) = 0$, et la restriction de μ^* à $\mathcal{M}(\mu^*)$ est une mesure.*

Démonstration. On notera $\mathcal{M} = \mathcal{M}(\mu^*)$ pour simplifier.

- Soit $B \subset X$ telle que $\mu^*(B) = 0$. Alors, pour toute partie A de X,
$$\mu^*(A \cap B) \leqslant \mu^*(B) = 0 \qquad \text{par croissance}$$
Donc
$$\mu^*(A) \geqslant \mu^*(A \cap B^\complement) = \mu^*(A \cap B) + \mu^*(A \cap B^\complement)$$
- Il est clair que $\varnothing \in \mathcal{M}$ et que $B \in \mathcal{M}$ si et seulement si $B^\complement \in \mathcal{M}$. Pour montrer que $\mathcal{M}$ est une tribu, il reste donc à voir que $\mathcal{M}$ est stable par réunion dénombrable. Commençons par l'établir pour une réunion finie. Soient $B_1, B_2 \in \mathcal{M}$. Soit A une partie de X.
Comme $B_1 \in \mathcal{M}$,
$$\begin{aligned} \mu^*(A \cap (B_1 \cup B_2)) &= \mu^*(A \cap (B_1 \cup B_2) \cap B_1) + \mu^*(A \cap (B_1 \cup B_2) \cap B_1^\complement) \\ &= \mu^*(A \cap B_1) + \mu^*(A \cap B_2 \cap B_1^\complement) \end{aligned}$$
Donc, comme $B_1, B_2 \in \mathcal{M}$,
$$\begin{aligned} & \mu^*(A \cap (B_1 \cup B_2)) + \mu^*(A \cap (B_1 \cup B_2)^\complement) \\ = \;& \mu^*(A \cap B_1) + \mu^*(A \cap B_1^\complement \cap B_2) + \mu^*(A \cap B_1^\complement \cap B_2^\complement) \\ = \;& \mu^*(A \cap B_1) + \mu^*(A \cap B_1^\complement) \\ = \;& \mu^*(A) \end{aligned}$$
On en déduit que $B_1 \cup B_2 \in \mathcal{M}$.
Pour terminer la preuve, considérons une famille $\{B_n\}_{n\in\mathbb{N}}$ d'éléments de $\mathcal{M}$ deux à deux disjoints [1].
On veut montrer que
$$B = \bigcup_{n\in\mathbb{N}} B_n \quad \in \mathcal{M} \qquad \text{et} \qquad \mu^*(B) = \sum_{n\in\mathbb{N}} \mu^*(B_n)$$
Notons $E_n = \bigcup_{k=0}^n B_k$. Pour toute partie $A \subset X$ et tout $n \in \mathbb{N}$, on prétend que
$$\mu^*(A \cap E_n) = \sum_{k=0}^n \mu^*(A \cap B_k) \tag{3.1}$$
Par récurrence, (3.1) est vraie au rang $n = 0$ et si l'on suppose la propriété vraie au rang $n-1$, on a
$$\begin{aligned} \mu^*(A \cap E_n) &= \mu^*(A \cap \underbrace{E_n \cap B_n}_{=B_n}) + \mu^*(A \cap \underbrace{E_n \cap B_n^\complement}_{=E_{n-1}}) \\ &= \mu^*(A \cap B_n) + \mu^*(A \cap E_{n-1}) \end{aligned}$$
En prenant la limite lorsque n tend vers $+\infty$, et en se souvenant que $E_n \subset B$, on trouve
$$\lim_{n\to\infty} \mu^*(A \cap E_n) = \sum_{k=0}^\infty \mu^*(A \cap B_k) \leqslant \mu^*(A \cap B)$$
Par ailleurs, par sous-additivité,
$$\mu^*(A \cap B) = \mu^*\left(\bigcup_{n\in\mathbb{N}} (A \cap B_n)\right) \leqslant \sum_{n\in\mathbb{N}} \mu^*(A \cap B_n)$$

1. Si $\{A_n\}_n$ est une famille d'éléments de $\mathcal{M}$, on peut toujours écrire $\bigcup_n A_n = \bigcup_n B_n$, avec les éléments B_n deux à deux disjoints (cf. la démonstration de la proposition 1.8 page 3).

Ainsi, on a montré que

$$\lim_{n\to\infty} \mu^*(A \cap E_n) = \sum_{n=0}^{\infty} \mu^*(A \cap B_n) = \mu^*(A \cap B)$$

En particulier, si $A = X$,

$$\mu^*(B) = \sum_{n\in\mathbb{N}} \mu^*(B_n)$$

Enfin, comme $\mathcal{M} \ni E_n \subset B$, on a pour tout $n \in \mathbb{N}$ et pour toute partie $A \subset X$,

$$\mu^*(A) = \mu^*(A \cap E_n) + \mu^*(A \cap E_n^{\complement}) \geqslant \mu^*(A \cap E_n) + \mu^*(A \cap B^{\complement})$$

En prenant la limite quand n tend vers $+\infty$, on obtient

$$\mu^*(A) \geqslant \mu^*(A \cap B) + \mu^*(A \cap B^{\complement})$$

donc $B \in \mathcal{M}$. □

Passons enfin à la construction de la mesure de Lebesgue, en définissant une mesure extérieure que l'on restreindra.

3.2 La mesure de Lebesgue

On suppose à présent que $X = \mathbb{R}^d$.

Définition 3.4 (Rappel). Un pavé P de $\mathbb{R}^d$ est un produit d'intervalles bornés

$$P = I_1 \times I_2 \times \cdots \times I_d \qquad I_j \subset \mathbb{R} \text{ intervalle borné}$$

On note $mes(P) = \ell(I_1) \cdot \ldots \cdot \ell(I_d)$, où $\ell(I_j)$ est la longueur du segment I_j, la **mesure** du pavé P.

Définition 3.5. Pour toute partie A de $\mathbb{R}^d$, on définit

$$\lambda^*(A) = \inf\left\{ \sum_{i\in\mathbb{N}} mes(P_i) \mid A \subset \bigcup_{i\in\mathbb{N}} P_i, P_i \text{ pavé ouvert de } \mathbb{R}^d \right\}$$

L'infimum est pris sur tous les recouvrements dénombrables de A par des pavés ouverts (évidemment, il existe toujours de tels recouvrements).

Remarque. On obtient la même définition si on travaille avec des pavés quelconques. En effet, si $Q_i \subset \mathbb{R}^d$ est un pavé quelconque et si $\varepsilon > 0$, il existe P_i un pavé ouvert tel que $Q_i \subset P_i$ et $mes(P_i) \leqslant mes(Q_i) + \frac{\varepsilon}{2^i}$.

Théorème 3.6. *On a les assertions suivantes*

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

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

iii) $\lambda^*(P) = mes(P), \quad$ *pour tout pavé* $P \subset \mathbb{R}^d$.

Définition 3.7. On appelle **mesure de Lebesgue** sur $\mathbb{R}^d$ la restriction, notée λ, de la mesure extérieure λ^* à $\mathcal{B}(\mathbb{R}^d)$ ou à $\mathcal{M}(\lambda^*)$.

Démonstration. i) λ^* **est une mesure extérieure.**

Il est clair que $\lambda^*(\varnothing) = 0$ et que λ^* est croissante. Montrons la sous-additivité : soit $\{A_n\}_{n\in\mathbb{N}}$ une famille de sous-ensembles de $\mathbb{R}^d$. On peut supposer que $\lambda^*(A_n) < \infty$, $\forall n \in \mathbb{N}$ (sans quoi il n'y a rien à montrer). Soit $\varepsilon > 0$. Pour tout $n \in \mathbb{N}$, il existe une famille $\{P_i^n\}_{i\in\mathbb{N}}$ de pavés ouverts tels que

$$A_n \subset \bigcup_{i\in\mathbb{N}} P_i^n \qquad \text{et} \qquad \sum_{i\in\mathbb{N}} mes(P_i^n) \leqslant \lambda^*(A_n) + \frac{\varepsilon}{2^n}$$

Alors

$$\bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n\in\mathbb{N}} \bigcup_{i\in\mathbb{N}} P_i^n$$

et

$$\sum_{n\in\mathbb{N}} \sum_{i\in\mathbb{N}} mes(P_i^n) \leqslant \sum_{n\in\mathbb{N}} \left(\lambda^*(A_n) + \frac{\varepsilon}{2^n}\right) = \sum_{n\in\mathbb{N}} \lambda^*(A_n) + 2\varepsilon$$

Comme $\varepsilon > 0$ était arbitraire, on conclut que

$$\lambda^*\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leqslant \sum_{n\in\mathbb{N}} \lambda^*(A_n)$$

ii) $\mathcal{M}(\lambda^*)$ **est une tribu de** $\mathbb{R}^d$ **qui contient** $\mathcal{B}(\mathbb{R}^d)$**.**

Soient $j \in \{1, \ldots, d\}$ et $\alpha \in \mathbb{R}$. On définit

$$B_j(\alpha) = \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R} \times \underbrace{]\alpha, +\infty[}_{j\text{-ème position}} \times \mathbb{R} \times \cdots \times \mathbb{R}$$

C'est un demi-espace ouvert. La famille $\{B_j(\alpha)\}_{j\in\{1,\ldots,d\},\alpha\in\mathbb{R}}$ engendre la tribu des boréliens $\mathcal{B}(\mathbb{R}^d)$. Il suffit donc de vérifier que

$$B_j(\alpha) \in \mathcal{M}(\lambda^*) \qquad \forall j \in \{1, \ldots, d\}, \forall \alpha \in \mathbb{R}$$

Soit A une partie de $\mathbb{R}^d$ et soit $\varepsilon > 0$. Il existe des pavés ouverts $\{P_i\}_{i\in\mathbb{N}}$ tels que $A \subset \bigcup_{i\in\mathbb{N}} P_i$ et $\sum_{i\in\mathbb{N}} mes(P_i) \leqslant \lambda^*(A) + \varepsilon$. 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

$$A \cap B_j(\alpha)^\complement \subset \bigcup_{i\in\mathbb{N}} \left(P_i \cap B_j(\alpha)^\complement\right)$$

où $P_i \cap B_j(\alpha)^\complement$ est un pavé quelconque. Or,

$$\forall i \in \mathbb{N}, \qquad mes(P_i) = mes(P_i \cap B_j(\alpha)) + mes(P_i \cap B_j(\alpha)^\complement)$$

Ainsi,

$$\begin{aligned}\lambda^*(A \cap B_j(\alpha)) + \lambda^*(A \cap B_j(\alpha)^\complement) &\leqslant \sum_{i\in\mathbb{N}} mes(P_i \cap B_j(\alpha)) + \sum_{i\in\mathbb{N}} mes(P_i \cap B_j(\alpha)^\complement) \\ &= \sum_{i\in\mathbb{N}} mes(P_i) \\ &\leqslant \lambda^*(A) + \varepsilon\end{aligned}$$

Et comme ε était arbitraire, on obtient

$$\lambda^*(A \cap B_j(\alpha)) + \lambda^*(A \cap B_j(\alpha)^\complement) \leqslant \lambda^*(A)$$

donc $B_j(\alpha) \in \mathcal{M}(\lambda^*)$.

iii) Pour tout pavé $P \subset \mathbb{R}^d$, on a évidemment

$$\lambda^*(P) \leqslant mes(P)$$

Inversement, soit $P \subset \mathbb{R}^d$ 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 3.8 ci-dessous, on a

$$mes(P) \leqslant \sum_{i=0}^{N} mes(P_i) \leqslant \sum_{n=0}^{\infty} mes(P_i)$$

Ainsi,

$$\lambda^*(P) \geqslant mes(P)$$

Enfin, si P est un pavé quelconque, on a

$$\lambda^*(\mathring{P}) \leqslant mes(\mathring{P}) \leqslant mes(\overline{P}) \leqslant \lambda^*(\overline{P})$$

Par ailleurs, $\overline{P}\backslash\mathring{P} = \partial P$ est une union finie de pavés fermés de mesure (mes) nulle, donc $\lambda^*(\partial P) = 0$ par sous-additivité. Ainsi,

$$\lambda^*(\overline{P}) \leqslant \lambda^*(\mathring{P}) + \lambda^*(\partial P) = \lambda^*(\mathring{P})$$

On conclut que

$$\lambda^*(P) = \lambda^*(\mathring{P}) = \lambda^*(\overline{P}) = mes(P)$$ □

On a besoin d'introduire le lemme suivant pour terminer la preuve, étant donné que la notion de mesure de pavage n'a pas été définie.

Lemme 3.8. *Si $P, P_1, \dots, P_N$ sont des pavés de $\mathbb{R}^d$ avec $P \subset \bigcup_{i=1}^N P_i$ alors*

$$mes(P) \leqslant \sum_{i=1}^{N} mes(P_i)$$

Démonstration. On peut supposer sans perte de généralité que les pavés $P, P_1, \dots, P_N$ sont fermés.

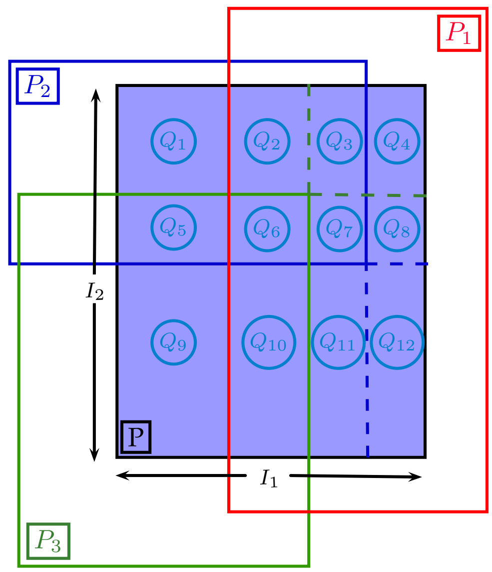


FIGURE 3.1 – Représentation de l'idée de la démonstration du lemme dans le cas général

- Le cas $N = 1$ est clair. Si $P = I_1 \times I_2 \times \cdots \times I_d$ et $Q = J_1 \times J_2 \times \cdots \times J_d$, alors

$$P \subset Q \iff \forall i \in \{1, \ldots, d\}, \quad I_i \subset J_i$$

et dans ce cas,

$$mes(P) = \ell(I_1) \cdot \ldots \cdot \ell(I_d) \leqslant \ell(J_1) \cdot \ldots \cdot \ell(J_d) = mes(Q)$$

- Considérons le cas général. L'idée de la démonstration est présentée sur la figure 3.1. Notons $P = I_1 \times I_2 \times \cdots \times I_d$ et $P_i = I_{i,1} \times I_{i,2} \times \cdots \times I_{i,d}$. Pour chaque $j \in \{1, \ldots, d\}$, on subdivise l'intervalle I_j de la façon suivante :

$$I_j = \left[a_1^j, a_2^j\right] \cup \left[a_2^j, a_3^j\right] \cup \cdots \cup \left[a_{n_j}^j, a_{n_j+1}^j\right]$$

où $\{a_k^j\}_{k \in \{1, \ldots, n_j+1\}}$ est l'ensemble des extrémités des intervalles $I_j \cap I_{i,j}$ pour i parcourant $1, \ldots, N$. Notons

$$E = \left\{k = (k_1, \ldots, k_d) \in \mathbb{N}^d \mid 1 \leqslant k_1 \leqslant n_1, \ldots, 1 \leqslant k_d \leqslant n_d\right\}$$

Pour $k \in E$, on note $Q_k = [a_{k_1}^1, a_{k_1+1}^1] \times [a_{k_2}^2, a_{k_2+1}^2] \times \cdots \times [a_{k_d}^d, a_{k_d+1}^d]$. Avec ces notations, on a que

$$P = \bigcup_{k \in E} Q_k$$

et que

$$mes(P) = \prod_{j=1}^{d} \ell(I_j) = \prod_{j=1}^{d} \sum_{k_j=1}^{n_j} \ell\left(\left[a_{k_j}^j, a_{k_j+1}^j\right]\right) = \sum_{k \in E} \ell\left([a_{k_1}^1, a_{k_1+1}^1]\right) \cdot \ldots \cdot \ell\left([a_{k_d}^d, a_{k_d+1}^d]\right) = \sum_{k \in E} mes(Q_k) \tag{3.2}$$

□

Or,

$$\begin{aligned}
\sum_{k\in E} mes(Q_k) &\leqslant \sum_{i=1}^{N}\left(\sum_{k \mid Q_k\subset P_i} mes(Q_k)\right) \\
&\underset{(3.2)\text{ pour } P_i\cap P}{=} \sum_{i=1}^{N} mes(P_i\cap P) \\
&\leqslant \sum_{i=1}^{n} mes(P_i)
\end{aligned}$$

Nous y sommes ! L'existence de la mesure de Lebesgue est maintenant un problème résolu. Il reste donc à prouver son unicité. Pour cela, d'une manière analogue à l'introduction d'une mesure extérieure – notion plus générale que la mesure sur une tribu – on va introduire la notion de classes monotones – notion plus générale que la notion de tribu.

3.3 Classes monotones

Définition 3.9. Un sous-ensemble $\mathcal{N} \subset \mathcal{P}(X)$ est appelé une **classe monotone** si :

1) $X \in \mathcal{N}$
2) Si $A, B \in \mathcal{N}$ et $A \subset B$, alors $B \setminus A \in \mathcal{N}$
3) Si $A_n \in \mathcal{N}$ pour tout $n \in \mathbb{N}$, et que $A_n \subset A_{n+1}$, alors

$$\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{N}$$

Remarques. 1) Si $A \in \mathcal{N}$, alors $A^{\complement} = X \setminus A \in \mathcal{N}$

2) Toute tribu est une classe monotone.
3) Une classe monotone est une tribu si et seulement si elle est **stable par intersection finie**, i.e. si $A_1, A_2 \in \mathcal{N}$, alors $A_1 \cap A_2 \in \mathcal{N}$. En effet, la classe monotone devient ainsi une algèbre, et donc une tribu par 3).

Comme dans le cas des tribus, toute intersection de classes monotones est encore une classe monotone. Si F est une famille de parties de X, on peut définir

$$\mathcal{N}(F) = \bigcap_{\mathcal{N}\text{ classe monotone sur } X, \mathcal{N}\supset F} \mathcal{N}$$

Alors, $\mathcal{N}(F)$ est une classe monotone sur X appelée la **classe monotone engendrée par** F. C'est la plus petite classe monotone sur X qui contient F.
On a toujours $\mathcal{N}(F) \subset \sigma(F)$ (car l'ensemble des classes monotones contient l'ensemble des tribus).

Lemme 3.10 (des classes monotones). *Si $F \subset \mathcal{P}(X)$ est une famille de partie de X* ***stable par intersections finies****, alors*

$$\mathcal{N}(F) = \sigma(F)$$

Remarque. Il suffit de montrer que $\mathcal{N}(F)$ est stable par intersections finies.

Démonstration. Soit $A \in F$. Notons

$$\mathcal{N}_1 = \{B \in \mathcal{N}(F) \mid B \cap A \in \mathcal{N}(F)\}$$

Par hypothèse, $\mathcal{N}_1$ contient F. En outre, c'est une classe monotone. En effet,
- $X \in \mathcal{N}_1$, puisque $X \cap A = A \in \mathcal{N}(F)$
- Si $B_1, B_2 \in \mathcal{N}_1$, avec $B_1 \subset B_2$, alors $B_2 \setminus B_1 \in \mathcal{N}(F)$ et

$$(B_2 \setminus B_1) \cap A = (B_2 \cap A) \setminus (B_1 \cap A) \in \mathcal{N}(F)$$

  Ainsi, $B_2 \setminus B_1 \in \mathcal{N}_1$.
- Si $B_n \in \mathcal{N}(F) \quad \forall n$, avec $B_n \subset B_{n+1}$, alors $B = \bigcup_{n\in\mathbb{N}} B_n \in \mathcal{N}(F)$ et

$$B \cap A = \bigcup_{n\in\mathbb{N}} (B_n \cap A) \in \mathcal{N}(F)$$

  Ainsi, $B \in \mathcal{N}_1$.

Ainsi, $\mathcal{N}_1$ contient $\mathcal{N}(F)$, mais comme on a l'inclusion inverse par définition de $\mathcal{N}_1$, on déduit que $\mathcal{N}_1 = \mathcal{N}(F)$.
On a montré que si $A \in F$ et $B \in \mathcal{N}(F)$ alors $A \cap B \in \mathcal{N}(F)$.
Soit maintenant $B \in \mathcal{N}(F)$. Notons

$$\mathcal{N}_2 = \{A \in \mathcal{N}(F) \mid A \cap B \in \mathcal{N}(F)\}$$

Par la première étape, on a que $\mathcal{N}_2$ contient F. On vérifie comme précédemment que $\mathcal{N}_2$ est une classe monotone, et on a donc $\mathcal{N}_2 = \mathcal{N}(F)$.
Ainsi, si $A, B \in \mathcal{N}(F)$ alors $A \cap B \in \mathcal{N}(F)$. □

Corollaire 3.11. *Soit $(X, \mathcal{M})$ un espace mesurable muni de deux mesures μ et ν. Supposons qu'il existe une famille F de parties de $\mathcal{M}$ telle que*

i) F est stable par intersection finie, et $\sigma(F) = \mathcal{M}$

ii) $\mu(A) = \nu(A) \qquad \forall A \in F$

On suppose en outre

a) soit que $\mu(X) = \nu(X) < \infty$

b) soit qu'il existe une famille $\{E_n\}_{n\in\mathbb{N}}$ d'éléments de F, telle que $E_n \subset E_{n+1}$, $\bigcup_{n\in\mathbb{N}} E_n = X$ et $\mu(E_n) = \nu(E_n) < \infty$ pour tout $n \in \mathbb{N}$.

Alors $\mu = \nu \qquad$ (i.e. $\mu(A) = \nu(A), \forall A \in \mathcal{M}$).

Démonstration. a) Soit $\mathcal{N} = \{A \in \mathcal{M} \mid \mu(A) = \nu(A)\} \subset \mathcal{M}$. Par hypothèse, $\mathcal{N} \supset F$. Montrons que $\mathcal{N}$ est une classe monotone.

1) $X \in \mathcal{N}$ par l'hypothèse a).

2) Si $A, B \in \mathcal{N}$ avec $A \subset B$, alors

$$\mu(B \setminus A) \underset{\mu \text{ finie}}{=} \mu(B) - \mu(A) = \nu(B) - \nu(A) \underset{\nu \text{ finie}}{=} \nu(B \setminus A)$$

Ainsi, $B \setminus A \in \mathcal{N}$.

3) Si $A_n \in \mathcal{N}$ pour tout $n \in \mathbb{N}$ et que $A_n \subset A_{n+1}$, alors

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

Ainsi, $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{N}$.

On a bien montré que $\mathcal{N}$ est une classe monotone. Ainsi, par le lemme des classes monotones,

$$\mathcal{N} \supset \mathcal{N}(F) = \sigma(F) = \mathcal{M} \supset \mathcal{N}$$

Donc $\mathcal{N} = \mathcal{M}$.

b) Pour tout $n \in \mathbb{N}$, on restreint μ et ν à E_n en posant :

$$\forall A \in \mathcal{M}, \qquad \mu_n(A) = \mu(A \cap E_n) \quad \text{et} \quad \nu_n(A) = \nu(A \cap E_n)$$

Comme $E_n \in F$, on a $\mu_n(A) = \nu_n(A) \quad \forall A \in F$. En appliquant le résultat de a), on trouve que

$$\mu_n(A) = \nu_n(A) \qquad \forall A \in \mathcal{M}$$

En prenant la limite quand n tend vers $+\infty$, on trouve

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

et ce, pour tout $A \in \mathcal{M}$. □

Application (Unicité de la mesure de Lebesgue). On prend $X = \mathbb{R}^d$, $\mathcal{M} = \mathcal{B}(\mathbb{R}^d)$, F la famille des pavés ouverts et $E_n =]-n, n[^d$. En appliquant le b) du corollaire ci-dessus, on voit qu'une mesure borélienne sur $\mathbb{R}^d$ finie sur les bornés est entièrement déterminée par ses valeurs sur les pavés ouverts. Ceci montre donc l'unicité de **la** mesure de Lebesgue sur $\mathbb{R}^d$.

3.4 Propriétés de la mesure de Lebesgue

On a montré l'existence d'une unique mesure $\lambda\colon \mathcal{B}(\mathbb{R}^d) \to [0, +\infty]$ telle que $\lambda(P) = mes(P)$ pour tout pavé P de $\mathbb{R}^d$. On sait aussi :

1) que λ se prolonge en une mesure sur $\mathcal{L}(\mathbb{R}^d)$.
Rappel : $\mathcal{L}(\mathbb{R}^d) = \sigma(\mathcal{B}(\mathbb{R}^d), N)$ où $N = \left\{A \subset \mathbb{R}^d \mid \exists B \in \mathcal{B}(\mathbb{R}^d) \text{ avec } A \subset B \text{ et } \lambda(B) = 0\right\}$ est l'ensemble des parties négligeables.

2) que λ se prolonge à la tribu $\mathcal{M}(\lambda^*) \supset \mathcal{B}(\mathbb{R}^d)$.
Rappel : $\lambda^*(A) = \inf\left\{\sum_{i\in\mathbb{N}} mes(P_i) \mid A \subset \bigcup_{i\in\mathbb{N}} P_i, P_i \text{ pavés ouverts.}\right\}$ et
$\mathcal{M}(\lambda^*) = \left\{B \subset \mathbb{R}^d \mid \lambda^*(A) = \lambda^*(A \cap B) + \lambda^*(A \cap B^{\complement}) \quad \forall A \subset \mathbb{R}^d\right\}$.

On peut se demander si $\mathcal{M}(\lambda^*)$ est beaucoup plus grande que la tribu $\mathcal{B}(\mathbb{R}^d)$, et si elle a un lien avec la tribu complétée $\mathcal{L}(\mathbb{R}^d)$.

Proposition 3.12. *On a* $\mathcal{M}(\lambda^*) = \mathcal{L}(\mathbb{R}^d)$.

Démonstration. L'inclusion $\mathcal{L}(\mathbb{R}^d) \subset \mathcal{M}(\lambda^*)$ est immédiate. En effet, on sait que $\mathcal{M}(\lambda^*) \supset \mathcal{B}(\mathbb{R}^d)$ et que $\mathcal{M}(\lambda^*) \supset N$ puisque si $A \in N$, il existe $B \in \mathcal{B}(\mathbb{R}^d)$ tel que $\lambda(B) = 0$ et $A \subset B$, et alors $\lambda^*(A) \leqslant \lambda^*(B) = \lambda(B) = 0$.
Inversement, soit $A \in \mathcal{M}(\lambda^*)$. Comme

$$A = \bigcup_{n\in\mathbb{N}} A \cap]-n, n[^d$$

il suffit de vérifier que $A \cap]-n, n[^d$ est λ^*-régulier et dans $\mathcal{L}(\mathbb{R}^d)$ pour tout $n \in \mathbb{N}$. On peut donc supposer sans perte de généralité que $A \subset P$ où P est un pavé ouvert. Pour tout $n \geqslant 1$, il existe une famille $\{P_i^n\}_{i\in\mathbb{N}}$ de pavés ouverts contenus dans P telle que

$$A \subset \bigcup_{i\in\mathbb{N}} P_i^n \qquad \text{et} \qquad \sum_{i\in\mathbb{N}} mes(P_i^n) \leqslant \lambda^*(A) + \frac{1}{n}$$

Notons

$$B_n = \bigcup_{i\in\mathbb{N}} P_i^n \qquad \text{et} \qquad B = \bigcap_{n\in\mathbb{N}} B_n$$

Alors $B \in \mathcal{B}(\mathbb{R}^d)$, on a $A \subset B \subset P$, et

$$\lambda^*(B) \leqslant \lambda^*(B_n) \leqslant \lambda^*(A) + \frac{1}{n} \qquad \forall n \geqslant 1$$

ce qui implique que $\lambda(B) = \lambda^*(B) = \lambda^*(A)$. On applique le même argument à $\widetilde{A} = P \setminus A$. On obtient ainsi $\widetilde{B} \in \mathcal{B}(\mathbb{R}^d)$ avec $\widetilde{A} \subset \widetilde{B} \subset P$ et $\lambda(\widetilde{B}) = \lambda^*(\widetilde{A})$. Posons $B' = P \setminus \widetilde{B} \subset P \setminus \widetilde{A} = A$. On a $B' \in \mathcal{B}(\mathbb{R}^d)$ et, comme $A \in \mathcal{M}(\lambda^*)$,

$$\lambda^*(P) = \lambda^*(P \cap A) + \lambda^*(P \cap A^{\complement}) = \lambda^*(A) + \lambda^*(\widetilde{A})$$

donc

$$\lambda(B') = \lambda(P) - \lambda(\widetilde{B}) = \lambda^*(P) - \lambda^*(\widetilde{A}) = \lambda^*(A)$$

Finalement, on a trouvé deux boréliens B et B' tels que $B' \subset A \subset B$ et $\lambda(B' \setminus B) = \lambda(B) - \lambda(B') = 0$. Ainsi, $A \in \mathcal{L}(\mathbb{R}^d)$. □

Montrons maintenant que la mesure de Lebesgue est invariante par translation. Pour cela, si $x \in \mathbb{R}^d$ et $A \in \mathcal{P}(\mathbb{R}^d)$, on notera $x + A = \{x + a \mid a \in A\}$.

Proposition 3.13. *Soit $\mathcal{M} = \mathcal{B}(\mathbb{R}^d)$ ou $\mathcal{L}(\mathbb{R}^d)$. La mesure de Lebesgue est invariante par translation, au sens où pour tout $A \in \mathcal{M}$ et $x \in \mathbb{R}^d$, on a $x + A \in \mathcal{M}$ et $\lambda(x + A) = \lambda(A)$.*

Démonstration. Tout suit naturellement du fait que si $P \subset \mathbb{R}^d$ est un pavé, alors $x + P$ est encore un pavé et que $mes(x + P) = mes(P)$. En particulier, la mesure extérieure

$$\lambda^*(A) = \inf \left\{ \sum_{i\in\mathbb{N}} mes(P_i) \mid A \subset \bigcup_{i\in\mathbb{N}} P_i, P_i \text{ pavé ouvert de } \mathbb{R}^d \right\}$$

est invariante par translation. De ce fait, la tribu

$$\mathcal{M}(\lambda^*) = \left\{ B \subset \mathbb{R}^d \mid \lambda^*(A) = \lambda^*(A \cap B) + \lambda^*(A \cap B^{\complement}), \forall A \subset \mathbb{R}^d \right\} = \mathcal{L}(\mathbb{R}^d)$$

est également invariante par translation (car cette dernière commute avec les opérations ensemblistes). D'autre part, pour tout $x \in \mathbb{R}^d$, $x + \mathcal{B}(\mathbb{R}^d) = \{x + A \mid A \in \mathcal{B}(\mathbb{R}^d)\}$ est une tribu qui contient les ouverts de $\mathbb{R}^d$, donc $x + \mathcal{B}(\mathbb{R}^d) \supset \mathcal{B}(\mathbb{R}^d)$. On en déduit également que $\mathcal{B}(\mathbb{R}^d) \supset \mathcal{B}(\mathbb{R}^d) - x$ et comme ceci est valable pour tout $x \in \mathbb{R}^d$, on a

$$x + \mathcal{B}(\mathbb{R}^d) = \mathcal{B}(\mathbb{R}^d) \qquad \forall x \in \mathbb{R}^d$$

□

Un résultat spectaculaire est la réciproque de cette proposition :

Proposition 3.14. *Si* $\mu\colon \mathcal{B}(\mathbb{R}^d) \to [0, +\infty]$ *est une mesure invariante par translation, et finie sur les bornés, alors il existe une constante* $c \geqslant 0$ *telle que* $\mu = c\lambda$*.*

Démonstration. Soit $c = \mu\left([0,1[^d\right) \geqslant 0$. Pour tout $n \in \mathbb{N}^*$, $[0,1[^d$ s'écrit comme une union disjointe de n^d pavés qui sont des translatés de $\left[0, \dfrac{1}{n}\right[^d$.

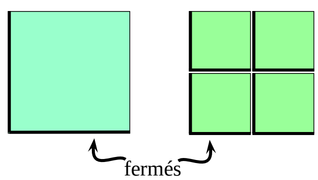


FIGURE 3.2 – Représentation du découpage effectué dans $\mathbb{R}^2$ pour $n = 2$

On déduit de l'invariance par translation de μ que

$$\mu\left(\left[0, \frac{1}{n}\right[^d\right) = \frac{c}{n^d} \qquad \forall n \in \mathbb{N}^* \qquad (*)$$

Soit maintenant $P = \prod_{i=1}^d [0, a_i[\subset \mathbb{R}^d$, où $a_1, \ldots, a_d \geqslant 0$. On a, pour tout $n \in \mathbb{N}$,

$$\prod_{i=1}^d \left[0, \frac{E(na_i)}{n}\right[\subset P \subset \prod_{i=1}^d \left[0, \frac{E(na_i)+1}{n}\right[$$

où $E(x)$ désigne la partie entière du réel $x \geqslant 0$. En utilisant le résultat précédent $(*)$ et l'invariance par translation, on trouve

$$c\prod_{i=1}^d \frac{E(na_i)}{n} \leqslant \mu(P) \leqslant c\prod_{i=1}^d \frac{E(na_i)+1}{n} \qquad \forall n \in \mathbb{N}^*$$

En prenant la limite quand $n \to \infty$, on trouve que

$$\mu(P) = c\prod_{i=1}^d a_i = c\lambda(P)$$

En prenant des unions (resp. intersections) dénombrables de translatés de tels pavés, on trouve que

$$\mu(P) = c\lambda(P) \qquad \forall P \in \mathcal{P}(\mathbb{R}^d) \text{ pavé ouvert (resp. fermé)}$$

Comme les pavés ouverts engendrent les boréliens, et par le corollaire 3.11 d'unicité des mesures, on a $\mu = c\lambda$ sur les boréliens. □

La dernière propriété concerne la régularité de la mesure de Lebesgue.

Proposition 3.15 (régularité de la mesure de Lebesgue). *Pour tout $A \in \mathcal{L}(\mathbb{R}^d)$, on a :*

$$\begin{aligned} \lambda(A) &= \inf\{\lambda(U) \mid U \textit{ ouvert}, U \supset A\} \qquad \textit{(C'est la régularité extérieure.)} \\ \lambda(A) &= \sup\{\lambda(K) \mid K \textit{ compact}, K \subset A\} \qquad \textit{(C'est la régularité intérieure.)} \end{aligned}$$

Démonstration. Il faut montrer que $\lambda(A) \geqslant \inf\{\lambda(U) \mid U \text{ ouvert}, U \supset A\}$, l'autre inégalité étant immédiate. On peut supposer que $\lambda(A) < \infty$, sans quoi il n'y a rien à montrer. Pour tout $\varepsilon > 0$, il existe une famille $\{P_i\}_{i\in\mathbb{N}}$ de pavés ouverts tels que

$$A \subset \bigcup_{i\in\mathbb{N}} P_i \qquad \text{et} \qquad \sum_{i\in\mathbb{N}} mes(P_i) \leqslant \lambda(A) + \varepsilon$$

Notons $U = \bigcup_{i\in\mathbb{N}} P_i$, alors U est un ouvert contenant A et

$$\lambda(U) \leqslant \sum_{i\in\mathbb{N}} \lambda(P_i) = \sum_{i\in\mathbb{N}} mes(P_i) \leqslant \lambda(A) + \varepsilon$$

et on a l'inégalité voulue.
De même, il faut montrer que $\lambda(A) \geqslant \sup\{\lambda(K) \mid K \text{ compact}, K \subset A\}$. Comme $\lambda(A) = \lim_{n\to\infty} \lambda\left(A \cap [-n,n]^d\right)$, on peut supposer sans perte de généralité que A est inclus dans un pavé compact P. Par régularité extérieure, pour tout $\varepsilon > 0$, il existe $U \subset \mathbb{R}^d$ ouvert, $U \supset P \setminus A$ tel que $\lambda(U) \leqslant \lambda(P \setminus A) + \varepsilon$. Alors, $K = P \setminus U$ est un fermé (donc compact), tel que $K \subset A$ et $\lambda(K) = \lambda(P) - \lambda(P \cap U) \geqslant \lambda(A) - \lambda(U) \geqslant \lambda(A) - \varepsilon$. $\square$

Il reste l'invariance de la mesure de Lebesgue par rotation à montrer. Cependant, si on applique une rotation à un pavé, on n'obtient pas nécessairement un pavé, ni même un pavage. On déduira ce résultat de la proposition 3.14 dans le chapitre des changements de variables.

3.5 Théorème de représentation de Riesz

Ce théorème fournit une autre démonstration de l'existence de la mesure de Lebesgue. Considérons un espace topologique X, muni de la tribu borélienne $\mathcal{B}(X)$, et soit $\mu \colon \mathcal{B}(X) \to [0, +\infty]$ une mesure *finie sur les compacts*. Une telle mesure est appelée une **mesure de Radon**.
On note $\mathcal{C}_c^0(X)$ l'espace des fonctions *continues à support compact* de X dans $\mathbb{R}$ ou $\mathbb{C}$. On rappelle que le support d'une fonction f est $supp(f) = \overline{\{x \in X \mid f(x) \neq 0\}}$.
On associe à la mesure μ une forme linéaire positive J sur $\mathcal{C}_c^0(X)$ définie par

$$J(f) = \int_X f \mathrm{d}\mu \qquad \in \mathbb{R} \text{ ou } \mathbb{C}$$

L'intégrale est bien définie, puisque $|f| \leqslant C\mathbb{1}_K$, avec $K = supp(f)$ compact, $C = \sup_{x\in X} |f(x)|$ et $\mu(K) < \infty$. J est positive dans le sens où $f \geqslant 0 \Rightarrow J(f) \geqslant 0$.
Le théorème de représentation de Riesz est une réciproque de ce qui précède. Sous certaines hypothèses, toutes les formes linéaires positives s'écrivent comme une intégrale. On adopte ici un point de vue différent sur la théorie de la mesure comme dual d'un espace de fonctions continues.

Théorème 3.16 (de représentation de Riesz). *Soit (X, d) un espace métrique, localement compact, séparable, et soit J une forme linéaire positive sur $\mathcal{C}_c^0(X)$. Alors, il existe une unique mesure borélienne μ, finie sur les compacts, telle que*

$$J(f) = \int_X f \mathrm{d}\mu \qquad \forall f \in \mathcal{C}_c^0(X) \tag{3.3}$$

En outre,

a) La mesure μ est régulière (extérieurement et intérieurement)

b) Pour tout ouvert $U \subset X$,

$$\mu(U) = \sup\left\{J(f) \mid f \in \mathcal{C}_c^0(X), 0 \leqslant f \leqslant 1, supp(f) \subset U\right\} \tag{3.4}$$

c) Pour tout compact $K \subset X$,

$$\mu(K) = \inf\left\{J(f) \mid f \in \mathcal{C}_c^0(X), 0 \leqslant f \leqslant 1, f \geqslant 1_K\right\} \tag{3.5}$$

Remarques. 1) Un espace topologique est localement compact si tout point x possède un voisinage ouvert d'adhérence compacte, et est séparable s'il possède une partie dénombrable dense.

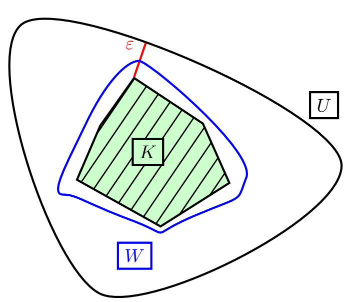
2) Si X est un espace métrique localement compact, alors X est séparable si et seulement si X est σ-compact (i.e. $X = \bigcup_{n\in\mathbb{N}} K_n$, $K_n \subset K_{n+1}$, K_n compacts).

L'ingrédient essentiel de la démonstration est le lemme suivant :

Lemme 3.17 (de Urysohn). *Soit X un espace métrique localement compact. Si $U \subset X$ est un ouvert, et $K \subset U$ un compact, il existe $f \in \mathcal{C}_c^0(X)$ telle que $supp(f) \subset U$, $f = 1$ sur K et $0 \leqslant f \leqslant 1$.*

Démonstration. Par hypothèse, tout $x \in K$ admet un voisinage ouvert V_x tel que $\overline{V_x}$ est compact. Ainsi, $K \subset \bigcup_{x\in K} V_x$, et par compacité de K, il existe $x_1, \ldots, x_N \in K$ tels que $K \subset \bigcup_{i=1}^N V_{x_i} = V$. Par construction, $\overline{V} = \bigcup_{i=1}^N \overline{V_{x_i}}$ est compact. Soit $\varepsilon = d(K, U^{\complement}) > 0$. On définit l'ouvert

$$W = V \cap \left\{x \in X \mid d(x, K) < \frac{\varepsilon}{2}\right\}$$



Alors, $K \subset W \subset \overline{W} \subset U$, et $\overline{W}$ est compact. On pose alors

$$f(x) = \frac{d(x, W^{\complement})}{d(x, K) + d(x, W^{\complement})}$$

et on vérifie que f a les propriétés requises. (f est bien définie, car le dénominateur ne s'annule pas : K et $W^{\complement}$ sont disjoints séparés d'une distance strictement positive.) □

A propos de la démonstration du théorème de Riesz

Unicité

Montrons que (3.3) ⇒ (3.4). Il suffit pour cela de montrer que

$$\mu(U) \leqslant \sup\left\{J(f) \mid f \in \mathcal{C}_c^0(X), 0 \leqslant f \leqslant 1, supp(f) \subset U\right\}$$

pour tout $U \subset X$ ouvert. L'autre inégalité est évidente car $f \leqslant \mathbb{1}_U$.

a) Supposons que $\overline{U}$ est compact dans X. Pour tout $n \in \mathbb{N}^*$, notons

$$K_n = \left\{x \in X \mid d(x, U^{\complement}) \geqslant \frac{1}{n}\right\}$$

Alors K_n est compact, $K_n \subset K_{n+1}$ et $\bigcup_{n\in\mathbb{N}^*} K_n = U$. Par le lemme d'Urysohn, il existe une application $f \in \mathcal{C}_c^0(X)$ telle que $0 \leqslant f_n \leqslant 1$, $f_n = 1$ sur K_n et $supp(f_n) \subset U$. On a donc $|f_n(x)| \leqslant \mathbb{1}_U$ pour tout $x \in X$ et pour tout $n \in \mathbb{N}^*$, et $\lim_{n\to\infty} f_n(x) = \mathbb{1}_U(x)$ pour tout $x \in X$. Par le théorème de convergence dominée,

$$J(f_n) = \int_X f_n \mathrm{d}\mu \underset{n\to\infty}{\longrightarrow} \int_X \mathbb{1}_U \mathrm{d}\mu = \mu(U)$$

b) Cas général. Comme X est σ-compact, il existe des ouverts $V_n \subset X$ tels que $V_n \subset V_{n+1}$, $\overline{V}_n$ compact et $X = \bigcup_{n\in\mathbb{N}} V_n$. Par a), on a pour tout $n \in \mathbb{N}$:

$$\begin{aligned} \sup\left\{J(f) \mid f \in \mathcal{C}_c^0(X), 0 \leqslant f \leqslant 1, supp(f) \subset U\right\} &\geqslant \sup\left\{J(f) \mid f \in \mathcal{C}_c^0(X), 0 \leqslant f \leqslant 1, supp(f) \subset U \cap V_n\right\} \\ &\geqslant \mu(U \cap V_n) \end{aligned}$$

En passant à la limite quand $n \longrightarrow \infty$ dans le membre de droite, on a l'inégalité voulue.

On procède de même pour montrer (3.5) à partir de (3.3). Cela montre l'unicité, puisque la mesure est définie d'une manière unique sur les ouverts, qui engendrent les boréliens.

Existence

On ne donnera qu'une esquisse de la démonstration.

- Etant donné une forme linéaire positive J, on définit pour tout $U \subset X$ ouvert :

$$\mu(U) = \sup\left\{J(f) \mid f \in \mathcal{C}_c^0(X), 0 \leqslant f \leqslant 1, supp(f) \subset U\right\}$$

  Pour prolonger μ en une mesure borélienne, on définit une mesure extérieure. Pour tout $A \subset X$, on pose :

$$\mu^*(A) = \inf\left\{\mu(U) \mid U \text{ ouvert}, U \supset A\right\}$$

  On vérifie facilement que μ^* définit bien une mesure extérieure sur X.
- Soit $\mathcal{M} \subset \mathcal{P}(X)$ la tribu des ensembles μ^*-réguliers. On vérifie que $\mathcal{M}$ contient les ouverts de X donc les boréliens sur X. On note encore μ la restriction de μ^* à $\mathcal{M}$ et on vérifie que μ est finie sur les compacts.
- On vérifie que (3.3) est vraie pour tout $f \in \mathcal{C}_c^0(X)$, en approchant f par des fonctions étagées, et on en déduit (3.5).
- On vérifie que μ est régulière et complète.

Application

Soit $X = \mathbb{R}^d$ et soit $J : \mathcal{C}_c^0(\mathbb{R}^d) \longrightarrow \mathbb{R}$ la forme linéaire positive définie par :

$$J(f) = \underbrace{\int_{\mathbb{R}^d} f(x)\mathrm{d}x}_{\text{intégrale de Riemann}} \qquad \forall f \in \mathcal{C}_c^0(\mathbb{R}^d)$$

Par le théorème 3.16 de représentation de Riesz, il existe une unique mesure borélienne μ telle que

$$J(f) = \int_{\mathbb{R}^d} f\mathrm{d}\mu \qquad \forall f \in \mathcal{C}_c^0(\mathbb{R}^d)$$

donc μ est la mesure de Lebesgue, d'après (3.4) et les valeurs que prend μ sur les pavés. Il y a deux

approches dans la littérature pour prouver l'existence de la mesure de Lebesgue. La première (qui est adoptée dans ce cours), consiste en la construction de la mesure de Lebesgue indépendamment du théorème de Riesz ; la seconde se focalise sur le théorème de Riesz et déduit la construction de la mesure de Lebesgue en supposant construite l'intégrale de Riemann pour les fonctions continues à support compact.

CHAPITRE 4

INTÉGRATION SUR LES ESPACES PRODUITS

Dans les trois premiers chapitres, nous avons défini une intégrale sur $\mathbb{R}^d$ et la mesure de Lebesgue associée à cette intégrale, qui permet un calcul théorique des intégrales de fonctions mesurables. Il convient maintenant dans les deux prochaines chapitres de donner des outils afin de calculer explicitement la valeur de ces intégrales.

4.1 PRODUIT D'ESPACES MESURABLES

Etant donné deux espaces mesurables, on veut définir une tribu sur le produit cartésien de ces espaces. Soient $(X_1, \mathcal{M}_1)$ et $(X_2, \mathcal{M}_2)$ deux espaces mesurables. On munit le produit $X_1 \times X_2$ de la **tribu produit** $\mathcal{M}_1 \otimes \mathcal{M}_2$ définie par :

$$\mathcal{M}_1 \otimes \mathcal{M}_2 = \sigma\left(A_1 \times A_2 \mid A_1 \in \mathcal{M}_1, A_2 \in \mathcal{M}_2\right)$$

$A_1 \times A_2, A_1 \in \mathcal{M}_1, A_2 \in \mathcal{M}_2$ est appelé un **rectangle mesurable**, et l'ensemble des rectangles mesurables n'est pas une tribu en général.
Montrons quelques propriétés de la tribu produit :

1) *La tribu $\mathcal{M}_1 \otimes \mathcal{M}_2$ est la plus petite tribu sur $X_1 \times X_2$ qui rende mesurable les deux projections canoniques $\pi_1\colon X_1 \times X_2 \to X_1$ et $\pi_2\colon X_1 \times X_2 \to X_2$.*
   En effet, si $\mathcal{M}$ est une tribu sur $X_1 \times X_2$ qui rend π_1 et π_2 mesurables, alors
   $$\forall A_1 \in \mathcal{M}_1, \forall A_2 \in \mathcal{M}_2, \quad \pi_1^{-1}(A_1) = A_1 \times X_2 \in \mathcal{M} \quad \pi_2^{-1}(A_2) = X_1 \times A_2 \in \mathcal{M}$$
   Donc $A_1 \times A_2 = (A_1 \times X_2) \cap (X_1 \times A_2) \in \mathcal{M}$ donc $\mathcal{M}_1 \otimes \mathcal{M}_2 \subset \mathcal{M}$.
2) *Soit $(X, \mathcal{M})$ un autre espace mesurable, et soit $f = (f_1, f_2)$ une application de X dans $X_1 \times X_2$. Alors $f\colon (X, \mathcal{M}) \to (X_1 \times X_2, \mathcal{M}_1 \otimes \mathcal{M}_2)$ est mesurable si et seulement si $f_i\colon (X, \mathcal{M}) \to (X_i, \mathcal{M}_i)$ est mesurable pour $i = 1, 2$.*
   En effet, si f est mesurable, alors $f_1 = \pi_1 \circ f$ et $f_2 = \pi_2 \circ f$ le sont aussi comme composition de fonctions mesurables. Inversement, si f_1 et f_2 sont mesurables, alors pour tout $A_1 \in \mathcal{M}_1, A_2 \in \mathcal{M}_2$, on a
   $$f^{-1}(A_1 \times A_2) = f_1^{-1}(A_1) \cap f_2^{-1}(A_2) \in \mathcal{M}$$
   donc f est mesurable.

On étend facilement la définition de la tribu produit pour un nombre fini d'espaces mesurables $(X_1, \mathcal{M}_1), \ldots, (X_n, \mathcal{M}_n)$ en posant

$$\mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \cdots \otimes \mathcal{M}_n = \sigma\left(A_1 \times A_2 \times \cdots \times A_n \mid A_i \in \mathcal{M}_i, \forall i \in \{1, 2, \ldots, n\}\right)$$

et on a la propriété d'"associativité" suivante :

Lemme 4.1. $(\mathcal{M}_1 \otimes \mathcal{M}_2) \otimes \mathcal{M}_3 = \mathcal{M}_1 \otimes (\mathcal{M}_2 \otimes \mathcal{M}_3) = \mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3$

Démonstration. Laissée en exercice □

Rappel : Soient $(X_1, \mathcal{T}_1), (X_2, \mathcal{T}_2)$ deux espaces topologiques. On munit $X_1 \times X_2$ de la topologie produit $\mathcal{T}_1 \otimes \mathcal{T}_2$ engendrée par les rectangles ouverts $U_1 \times U_2$, avec $U_1 \in \mathcal{T}_1$ et $U_2 \in \mathcal{T}_2$. Les éléments de $\mathcal{T}_1 \otimes \mathcal{T}_2$ sont donc les unions de rectangles ouverts [1]. La topologie $\mathcal{T}_1 \otimes \mathcal{T}_2$ est la topologie la moins fine sur $X_1 \times X_2$ qui rende continues les projections canoniques π_1 et π_2.
On veut maintenant comparer les tribus produit des tribus de Borel sur des espaces topologiques. Pour cela, on a la proposition suivante :

Proposition 4.2. *Soient X_1 et X_2 deux espaces topologiques, munis de leur tribu borélienne. Alors*

a) $\mathcal{B}(X_1 \times X_2) \supset \mathcal{B}(X_1) \otimes \mathcal{B}(X_2)$.

b) Si X_1 et X_2 sont des espaces métriques séparables, alors on a :

$$\mathcal{B}(X_1 \times X_2) = \mathcal{B}(X_1) \otimes \mathcal{B}(X_2)$$

Démonstration. a) On munit $X_1 \times X_2$ de la topologie produit. Alors les projections π_1 et π_2 sont continues, donc mesurables pour les tribus boréliennes. Par la propriété 1) des tribus produit, on a $\mathcal{B}(X_1 \times X_2) \supset \mathcal{B}(X_1) \otimes \mathcal{B}(X_2)$.

b) On suppose que X_1 et X_2 sont des espaces métriques séparables [2]. Alors, X_1 et X_2 contiennent des bases dénombrables d'ouverts $U = \{U_n\}_{n\in\mathbb{N}}$ et $V = \{V_m\}_{m\in\mathbb{N}}$ (par exemple, si on prend une suite $(x_k)_{k\in\mathbb{N}}$ dense dans X_1, on peut prendre pour U l'ensemble des boules ouverts de rayon rationnel centrées sur l'un des x_k). Il s'ensuit que la famille $W = \{U_n \times V_m \mid n, m \in \mathbb{N}\}$ est une base dénombrable de $X_1 \times X_2$ (en effet, si $O \subset X_1 \times X_2$ est ouvert, et si $z = (x, y) \in O$, alors il existe $(O_1, O_2) \in X_1 \times X_2$ un couple d'ouverts avec $z \in O_1 \times O_2$ tel que $O_1 \times O_2 \subset O$). Ainsi, chaque ouvert de $X_1 \times X_2$ appartient à la tribu $\mathcal{B}(X_1) \otimes \mathcal{B}(X_2)$ et on a $\mathcal{B}(X_1 \times X_2) \subset \mathcal{B}(X_1) \otimes \mathcal{B}(X_2)$. Le a) permet de conclure. □

Introduisons les notations suivantes. Si $(X, \mathcal{M})$ et $(Y, \mathcal{N})$ sont deux espaces mesurables quelconques et si $E \subset X \times Y$, on note :

$$E_x = \{y \in Y \mid (x, y) \in E\} \qquad \text{pour } x \in X$$
$$E^y = \{x \in X \mid (x, y) \in E\} \qquad \text{pour } y \in Y$$

Et si $f : X \times Y \to Z$ (espace quelconque), on définit :

$$f_x : Y \to Z \quad \text{par } f_x(y) = f(x, y) \qquad \text{pour } x \in X$$
$$f^y : X \to Z \quad \text{par } f^y(x) = f(x, y) \qquad \text{pour } y \in Y$$

Proposition 4.3. *Soient $(X, \mathcal{M})$ et $(Y, \mathcal{N})$ deux espaces mesurables.*

i) Si $E \in \mathcal{M} \otimes \mathcal{N}$, alors $E_x \in \mathcal{N}, \forall x \in X$ et $E^y \in \mathcal{M}, \forall y \in Y$.

1. Ici, on peut décrire tous les éléments. Ce n'est pas le cas dans les espaces mesurés : on ne peut pas décrire les éléments de la tribu produit aussi facilement.

2. La notion de séparabilité s'explique par le fait qu'une tribu est stable par union dénombrable, alors que la notion d'ouvert est stable par union quelconque.

ii) Si $f: X \times Y \to Z$ (espace mesurable quelconque) est mesurable pour la tribu produit $\mathcal{M} \otimes \mathcal{N}$, alors

$$\begin{array}{lll} \forall x \in X & f_x: Y \to Z & \text{est mesurable} \\ \forall y \in Y & f^y: X \to Z & \text{est mesurable} \end{array}$$

Démonstration. i) Pour $x \in X$, notons $\mathcal{N}_x = \{E \in \mathcal{M} \otimes \mathcal{N} \mid E_x \in \mathcal{N}\}$. Alors, $\mathcal{N}_x$ contient les rectangles mesurables : si $E = A \times B$, $A \in \mathcal{M}$, $B \in \mathcal{N}$, alors

$$(A \times B)_x = \begin{cases} B & \text{si } x \in A \\ \varnothing & \text{sinon} \end{cases}$$

Par ailleurs, on vérifie aisément que $\mathcal{N}_x$ est une tribu. Ainsi, $\mathcal{N}_x = \mathcal{M} \otimes \mathcal{N}$. On fait de même pour $y \in Y$.

ii) Soit $f: X \times Y \to Z$ mesurable pour la tribu produit $\mathcal{M} \otimes \mathcal{N}$. Alors si $E \subset Z$ est mesurable, on a pour tout $x \in X$:

$$\begin{aligned} f_x^{-1}(E) &= \{y \in Y \mid f(x,y) \in E\} \\ &= \left\{y \in Y \mid (x,y) \in f^{-1}(E)\right\} \\ &= \left(\underbrace{f^{-1}(E)}_{\in \mathcal{M} \otimes \mathcal{N}}\right)_x \qquad \in \mathcal{N} \text{ (par i))} \end{aligned}$$

Ainsi, f_x est mesurable pour tout $x \in X$. On fait de même avec $y \in Y$. □

4.2 Mesure produit

On veut maintenant construire une mesure sur l'espace produit $X \times Y$ où X et Y sont des espaces mesurés.

Définition 4.4. On dit qu'un espace mesuré $(X, \mathcal{M}, \mu)$ est σ**-fini** s'il existe une suite croissante de parties mesurables E_n telle que $X = \bigcup_{n \in \mathbb{N}} E_n$ et $\mu(E_n) < \infty$ pour tout $n \in \mathbb{N}$.

Théorème 4.5. *Soient $(X, \mathcal{M}, \mu)$ et $(Y, \mathcal{N}, \nu)$ deux espaces mesurés σ-finis. Alors :*

i) Il existe une unique mesure m sur $(X \times Y, \mathcal{M} \otimes \mathcal{N})$ telle que

$$m(A \times B) = \mu(A)\nu(B) \qquad \forall A \in \mathcal{M}, \forall B \in \mathcal{N} \tag{4.1}$$

ii) Pour tout $E \in \mathcal{M} \otimes \mathcal{N}$, on a :

$$m(E) = \int_X \nu(E_x) \mathrm{d}\mu_x = \int_Y \mu(E^y) \mathrm{d}\nu_y$$

Démonstration. a) **Unicité.** Par hypothèse, il existe $\{A_n\}_{n \in \mathbb{N}}$ et $\{B_n\}_{n \in \mathbb{N}}$ telles que $X = \bigcup_{n \in \mathbb{N}} A_n$, $Y = \bigcup_{n \in \mathbb{N}} B_n$ et pour tout $n \in \mathbb{N}$, $A_n \in \mathcal{M}$, $B_n \in \mathcal{N}$, $A_n \subset A_{n+1}$, $B_n \subset B_{n+1}$, $\mu(A_n) < \infty$ et $\nu(B_n) < \infty$. Alors, $X \times Y = \bigcup_{n \in \mathbb{N}} E_n$ avec $E_n = A_n \times B_n$ pour tout $n \in \mathbb{N}$.

Soient m_1, m_2 deux mesures sur $(X \times Y, \mathcal{M} \otimes \mathcal{N})$ qui vérifient (4.1). Alors,

- m_1 et m_2 coïncident sur la famille des rectangles mesurables, qui est stable par intersection finie et engendre la tribu $\mathcal{M} \otimes \mathcal{N}$.
- $m_1(E_n) = \mu(A_n)\nu(B_n) = m_2(E_n) < \infty$ pour tout $n \in \mathbb{N}$.

Par le corollaire 3.11 (page 46) du lemme des classes monotones, on a $m_1 = m_2$. Cet argument montre que toute mesure m vérifiant (4.1) est σ-finie.

b) **Existence.** L'idée est de définir m par la formule

$$m(E) = \int_X \nu(E_x) \mathrm{d}\mu_x \qquad \text{pour } E \in \mathcal{M} \otimes \mathcal{N} \tag{4.2}$$

De même, on peut utiliser

$$\widetilde{m}(E) = \int_Y \mu(E^y) \mathrm{d}\nu_y$$

et l'unicité montrera que $m = \widetilde{m}$.

Remarquons que $\nu(E_x)$ est bien définie pour tout $x \in E$ d'après la proposition 4.3 précédente. Pour vérifier que la formule (4.2) a bien un sens, il faut vérifier que pour tout $E \in \mathcal{M} \otimes \mathcal{N}$, l'application $x \mapsto \nu(E_x)$ est $\mathcal{M}$-mesurable.

α) Supposons tout d'abord que ν est une mesure finie. Notons

$$\mathbb{M} = \{E \in \mathcal{M} \otimes \mathcal{N} \mid x \mapsto \nu(E_x) \text{ est } \mathcal{M}\text{-mesurable}\} \subset \mathcal{M} \otimes \mathcal{N}$$

Alors,

- $\mathbb{M}$ contient les rectangles mesurables : si $E = A \times B, A \in \mathcal{M}, B \in \mathcal{N}$, on a $\nu(E_x) = \mathbb{1}_A(x)\nu(B)$. De plus, la classe des rectangles mesurables est stable par intersections finies.
- on vérifie aisément que $\mathbb{M}$ est une classe monotone :
  1) $X \times Y \in \mathbb{M}$ car $\nu\left((X \times Y)_x\right) = \mathbb{1}_X(x)\nu(Y) = \nu(Y)$.
  2) Si $E, F \in \mathbb{M}$ et $E \subset F$, on a :

  $$\nu\left((F \setminus E)_x\right) = \nu(F_x \setminus E_x) \underset{\nu \text{ finie}}{=} \nu(F_x) - \nu(E_x)$$

  qui est une fonction mesurable, donc $F \setminus E \in \mathbb{M}$.
  3) Si $\{E_n\}_{n\in\mathbb{N}}$ est une famille croissante d'éléments de $\mathbb{M}$, alors :

  $$\nu\left(\left(\bigcup_{n\in\mathbb{N}} E_n\right)_x\right) = \nu\left(\bigcup_{n\in\mathbb{N}} (E_n)_x\right) \underset{(E_n)_x \text{ est une suite croissante}}{=} \lim_{n\to\infty} \underbrace{\nu\left((E_n)_x\right)}_{\text{mesurable}}$$

  est une limite de fonctions mesurables, donc est mesurable et $\bigcup_{n\in\mathbb{N}} E_n \in \mathbb{M}$.

Par le lemme des classes monotones (lemme 3.10, page 45), on a $\mathbb{M} = \mathcal{M} \otimes \mathcal{N}$. Ainsi, $x \mapsto \nu(E_x)$ est mesurable pour tout $E \in \mathcal{M} \otimes \mathcal{N}$.

β) Si ν est seulement σ-finie, on écrit comme avant que $Y = \bigcup_{n\in\mathbb{N}} B_n$, et on applique le résultat de α) à la mesure ν_n définie par $\nu_n(B) = \nu(B \cap B_n)$ pour tout $B \in \mathcal{N}$.

Comme $\nu(E_x) = \lim_{n\to\infty} \nu_n(E_x)$, on en déduit que $x \mapsto \nu(E_x)$ est mesurable pour tout $E \in \mathcal{M} \otimes \mathcal{N}$.

On peut donc définir

$$m(E) = \int_X \nu(E_x)\mathrm{d}\mu_x \qquad \forall E \in \mathcal{M} \otimes \mathcal{N}$$

□

Reste à vérifier que m est bien une mesure. Par construction, $m(\varnothing) = 0$, et $m(A \times B) = \mu(A)\nu(B)$ pour tout $A \in \mathcal{M}$ et $B \in \mathcal{N}$. Si $\{E_n\}_{n\in\mathbb{N}}$ est une famille disjointe de $\mathcal{M}\otimes\mathcal{N}$, les $\{(E_n)_x\}_{n\in\mathbb{N}}$ sont aussi disjoints pour tout $x \in X$, et on a

$$\begin{aligned} m\left(\bigcup_{n\in\mathbb{N}} E_n\right) &= \int_X \nu\left(\bigcup_{n\in\mathbb{N}} (E_n)_x\right)\mathrm{d}\mu_x \\ &= \int_X \sum_{n\in\mathbb{N}} \nu\left((E_n)_x\right)\mathrm{d}\mu_x \\ &= \sum_{n\in\mathbb{N}} \int_X \nu\left((E_n)_x\right)\mathrm{d}\mu_x \\ &= \sum_{n\in\mathbb{N}} m(E_n) \end{aligned}$$

Remarques. 1) Par le même procédé, on peut définir le produit de n mesures σ-finies $\mu_1, \ldots, \mu_n$ en posant par exemple

$$\mu_1 \otimes \cdots \otimes \mu_n = \mu_1 \otimes (\mu_2 \otimes (\cdots \otimes \mu_n))$$

L'ordre des parenthèses n'a pas d'importance, car la mesure produit est entièrement définie par sa valeur sur les pavés mesurables :

$$(\mu_1 \otimes \cdots \otimes \mu_n)(A_1 \times \cdots \times A_n) = \mu_1(A_1)\cdots\mu_n(A_n)$$

2) Si $(X, \mathcal{M}, \mu) = (Y, \mathcal{N}, \nu) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, alors $X \times Y = \mathbb{R}^2$, $\mathcal{M} \otimes \mathcal{N} = \mathcal{B}(\mathbb{R}^2)$ et $\mu \otimes \nu = \lambda_2$ la mesure de Lebesgue sur $\mathbb{R}^2$. En effet, la dernière égalité provient du fait que les mesures prennent les mêmes valeurs sur les pavés mesurables.

   De même, $\lambda_d = \lambda_1 \otimes \cdots \otimes \lambda_1 \quad (d$ fois). Ainsi, on aurait pu se contenter de montrer l'existence de la mesure de Lebesgue sur $\mathbb{R}$ et généraliser sur $\mathbb{R}^d$ avec cette remarque.

3) L'hypothèse de σ-finitude est nécessaire. En effet, soit $(X, \mathcal{M}) = (Y, \mathcal{N}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, $\mu = \lambda$ la mesure de Lebesgue et ν la mesure de comptage (non σ-finie). Soit $E = \{(x, y) \in \mathbb{R}^2 \mid x = y\} \in \mathcal{M} \otimes \mathcal{N} = \mathcal{B}(\mathbb{R}^2)$. On a $\nu(E_x) = 1, \forall x \in X$ et $\mu(E^y) = 0, \forall y \in Y$. Or,

$$\infty = \int_X \nu(E_x)\mathrm{d}\mu_x \neq \int_Y \mu(E^y)\mathrm{d}\nu_y = 0$$

4.3 Théorèmes de Fubini

4.3.1 Enoncés des théorèmes

Ces deux théorèmes relient l'intégrale sur la mesure produit et les intégrales itérées. Il y a deux versions, la première concerne les fonctions à valeurs positives, et la deuxième les fonctions réelles ou complexes intégrables.

Théorème 4.6 (Fubini-Tonelli). *Soient $(X, \mathcal{M}, \mu)$ et $(Y, \mathcal{N}, \nu)$ deux espaces mesurés σ-finis, et soit $f : X \times Y \to [0, +\infty]$ une fonction $\mathcal{M} \otimes \mathcal{N}$-mesurable. Alors :*

i) les fonctions $\begin{array}{ccc}(X,\mathcal{M}) & \to & [0,+\infty] \\ x & \mapsto & \int_Y f(x,y)\mathrm{d}\nu_y\end{array}$ *et* $\begin{array}{ccc}(Y,\mathcal{N}) & \to & [0,+\infty] \\ y & \mapsto & \int_X f(x,y)\mathrm{d}\mu_x\end{array}$ *sont mesurables* [3].

ii) On a les égalités suivantes :

$$\int_{X\times Y} f\mathrm{d}(\mu\otimes\nu) = \int_X \left(\int_Y f(x,y)\mathrm{d}\nu_y\right)\mathrm{d}\mu_x = \int_Y \left(\int_X f(x,y)\mathrm{d}\mu_x\right)\mathrm{d}\nu_y \quad \in [0,+\infty]$$

De même que pour le théorème de Beppo-Levi, quand on a des fonctions mesurables positives, il n'y a aucune précaution à prendre. Ici, on peut échanger les deux intégrales sans problème.

Démonstration. Soit $E \in \mathcal{M}\otimes\mathcal{N}$, et soit $f = \mathbb{1}_E$. Alors,

$$\int_Y f(x,y)\mathrm{d}\nu_y = \nu(E_x), \qquad \int_X f(x,y)\mathrm{d}\mu_x = \mu(E^y)$$

donc les conclusions de i) et ii) ont été établies au paragraphe précédent, lors de la démonstration du théorème 4.5. Par linéarité, le résultat reste vrai si $f\colon X\times Y \to [0,+\infty]$ est une fonction étagée mesurable. Enfin, si $f\colon X\times Y \to [0,+\infty]$ est mesurable, il existe une suite croissante $(f_n)_{n\in\mathbb{N}}$ de fonctions étagées positives telles que $f_n(x,y) \underset{n\to\infty}{\longrightarrow} f(x,y)$ pour tout $x\in X$ et $y \in Y$. Par le théorème 2.16 de la convergence monotone, on a par exemple :

$$\int_Y f(x,y)\mathrm{d}\nu_y = \lim_{n\to\infty}\int_Y f_n(x,y)\mathrm{d}\nu_y$$

donc i) est vraie. Par ailleurs, l'égalité ii) est vraie pour f_n, donc pour f par passage à la limite croissante en appliquant deux fois le théorème de la convergence monotone. □

On passe maintenant au cas de fonctions réelles ou complexes.

Théorème 4.7 (Fubini-Lebesgue). *Soient $(X,\mathcal{M},\mu)$ et $(Y,\mathcal{N},\nu)$ deux espaces mesurés σ-finis, et soit $f\colon X\times Y\to\mathbb{R}$ ou $\mathbb{C}$ une fonction* ***intégrable****, i.e. $f\in\mathcal{L}^1(X\times Y,\mathcal{M}\otimes\mathcal{N},\mu\otimes\nu)$. Alors :*

i) Pour presque tout $x\in X$, la fonction $y\mapsto f(x,y)$ est dans $\mathcal{L}^1(Y,\mathcal{N},\nu)$.
Pour presque tout $y\in Y$, la fonction $x\mapsto f(x,y)$ est dans $\mathcal{L}^1(X,\mathcal{M},\mu)$.

ii) La fonction $x\mapsto\int_Y f(x,y)\mathrm{d}\nu_y$ (définie μ-presque partout) est dans $\mathcal{L}^1(X,\mathcal{M},\mu)$.
La fonction $y\mapsto\int_X f(x,y)\mathrm{d}\mu_x$ (définie ν-presque partout) est dans $\mathcal{L}^1(Y,\mathcal{N},\nu)$.

iii) On a les égalités suivantes :

$$\int_{X\times Y} f\mathrm{d}(\mu\otimes\nu) = \int_X \left(\int_Y f(x,y)\mathrm{d}\nu_y\right)\mathrm{d}\mu_x = \int_Y \left(\int_X f(x,y)\mathrm{d}\mu_x\right)\mathrm{d}\nu_y$$

Démonstration. Il suffit de considérer le cas où f est à valeurs réelles [4].

i) On a déjà montré que pour tout $x\in X$, la fonction $y\mapsto f(x,y)$ est $\mathcal{N}$-mesurable. En appliquant le théorème 4.6 de Fubini-Tonelli à la fonction $|f|$, on trouve que :

$$\int_X\left(\int_Y |f(x,y)|\,\mathrm{d}\nu_y\right)\mathrm{d}\mu_x = \int_{X\times Y}|f|\,\mathrm{d}(\mu\otimes\nu) < \infty$$

Donc, $\int_Y |f(x,y)|\,\mathrm{d}\nu_y < \infty$ pour presque tout $x\in X$.

4. Si f est à valeurs complexes, on applique le théorème à $\Re(f)$ et $\Im(f)$.

ii) Ecrivons $f = f_+ - f_-$, et définissions

$$g(x) = \begin{cases} \int_Y f_+(x,y)\mathrm{d}\nu_y - \int_Y f_-(x,y)\mathrm{d}\nu_y & \text{si } \int_Y |f(x,y)|\,\mathrm{d}\nu_y < \infty \\ 0 & \text{sinon} \end{cases}$$

Alors, g est mesurable (par Fubini-Tonelli appliqué à f_+ et f_-), et

$$\int_X |g|\,\mathrm{d}\mu \leqslant \int_X \left(\int_Y |f(x,y)|\,\mathrm{d}\nu_y\right)\mathrm{d}\mu_x < \infty$$

et on en déduit que $g \in \mathcal{L}^1(X, \mathcal{M}, \mu)$.

iii) Enfin, toujours par Fubini-Tonelli,

$$\int_{X\times Y} f_+\mathrm{d}(\mu\otimes\nu) = \int_X \left(\int_Y f_+(x,y)\mathrm{d}\nu_y\right)\mathrm{d}\mu_x$$

et

$$\int_{X\times Y} f_-\mathrm{d}(\mu\otimes\nu) = \int_X \left(\int_Y f_-(x,y)\mathrm{d}\nu_y\right)\mathrm{d}\mu_x$$

En faisant la différence de ces intégrales, on trouve le résultat. □

Si f vérifie les hypothèses du théorème de Fubini-Tonelli, ou de Fubini-Lebesgue, et si $f(x,y) = g(x)h(y)$ (f se factorise), alors on a :

$$\int_{X\times Y} f\mathrm{d}(\mu\otimes\nu) = \left(\int_X g\mathrm{d}\mu\right)\left(\int_Y h\mathrm{d}\nu\right)$$

4.3.2 Discussions sur les théorèmes

Nous allons voir sur un exemple pourquoi il est nécessaire d'avoir deux théorèmes de Fubini, et l'importance de l'hypothèse d'intégrabilité de la fonction que l'on veut intégrer dans le théorème de Fubini-Lebesgue.

Exemple. Soit $X =]0, +\infty[$, muni de la tribu de Borel et de la mesure de Lebesgue, notée $\mathrm{d}x$. Soit $f\colon X \times X \to \mathbb{R}$ la fonction définie par

$$f(x,y) = \frac{e^{-y}\sin(xy)}{x\sqrt{1+x^2}} \quad \forall x, y > 0$$

Montrer que f est intégrable et calculer son intégrale.

Démonstration (Solution). f est continue, donc mesurable. On se trouve face à une fonction pouvant prendre des valeurs négatives, on va donc devoir utiliser le théorème de Fubini-Lebesgue. Cependant, il est nécessaire pour l'appliquer de montrer l'intégrabilité de la fonction : pour cela, on utilise le théorème de Fubini-Tonelli sur $|f|$!

1) On rappelle que $\forall x \in \mathbb{R}, |\sin(x)| \leqslant \min(|x|, 1)$.
   - Si $0 < x \leqslant 1$, on écrit :

$$|f(x,y)| \leqslant \frac{e^{-y}xy}{x\sqrt{1+x^2}} = \frac{ye^{-y}}{\sqrt{1+x^2}}$$

   Et ainsi,

$$\int_0^\infty |f(x,y)|\,\mathrm{d}y \leqslant \frac{1}{\sqrt{1+x^2}}$$

– Si $x \geqslant 1$, on écrit :

$$|f(x,y)| \leqslant \frac{e^{-y}1}{x\sqrt{1+x^2}}$$

Et ainsi,

$$\int_0^\infty |f(x,y)|\,\mathrm{d}y \leqslant \frac{1}{x\sqrt{1+x^2}}$$

On a finalement

$$\int_0^\infty |f(x,y)|\,\mathrm{d}y \leqslant \mathbb{1}_{]0,1[}(x)\frac{1}{\sqrt{1+x^2}} + \mathbb{1}_{[1,+\infty[}(x)\frac{1}{x\sqrt{1+x^2}}$$

d'où

$$\int_0^\infty \left(\int_0^\infty |f(x,y)|\,\mathrm{d}y\right)\mathrm{d}x \leqslant \int_0^1 \mathrm{d}x + \int_1^\infty \frac{1}{x^2}\mathrm{d}x = 2 < \infty$$

□

Ainsi, f est intégrable, et $\int |f|\,\mathrm{d}x\mathrm{d}y \leqslant 2$.

2) On applique donc le théorème de Fubini-Lebesgue, et on a

$$\begin{aligned} \int_0^\infty e^{-y}\sin(xy)\mathrm{d}y &= \Im\left(\int_0^\infty e^{-y}e^{ixy}\mathrm{d}y\right) \\ &= \Im\left(\frac{1}{1-ix}\right) = \frac{x}{1+x^2} \end{aligned}$$

Donc

$$\begin{aligned} \int_{X\times X} f\mathrm{d}x\mathrm{d}y &= \int_0^\infty \left(\frac{x}{1+x^2} \times \frac{1}{x\sqrt{1+x^2}}\right)\mathrm{d}x \\ &= \int_0^\infty \frac{1}{(1+x^2)^{\frac{3}{2}}}\mathrm{d}x \\ &\overset{x=\mathrm{sh}(t)}{=} \int_0^\infty \frac{1}{(1+\mathrm{sh}(t)^2)^{\frac{3}{2}}}\,\mathrm{ch}(t)\mathrm{d}t \\ &= \int_0^\infty \frac{1}{\mathrm{ch}(t)^2}\mathrm{d}t \\ &= [\mathrm{th}(t)]_{t=0}^{t=\infty} \\ &= 1 \end{aligned}$$

Attention ! Il est indispensable lors de l'application du théorème de Fubini-Lebesgue, de vérifier l'intégrabilité avant de calculer les intégrales itérées !

Exemple. Soit $X =]0,1[$, et $f\colon X \times X \to \mathbb{R}$ définie par $f(x,y) = \dfrac{x^2-y^2}{(x^2+y^2)^2}$.
Pour $0 < x < 1$, on a :

$$\int_0^1 \frac{x^2-y^2}{(x^2+y^2)^2}\mathrm{d}y = \left[\frac{y}{x^2+y^2}\right]_{y=0}^{y=1} = \frac{1}{1+x^2}$$

On en déduit que

$$\int_0^1 \left(\int_0^1 f(x,y)\mathrm{d}y\right)\mathrm{d}x = \int_0^1 \frac{1}{1+x^2}\mathrm{d}x = [\arctan(x)]_{x=0}^{x=1} = \frac{\pi}{4}$$

Or, x et y jouent un rôle antisymétrique, donc

$$\int_0^1 \left(\int_0^1 f(x,y)\mathrm{d}x \right) \mathrm{d}y = -\frac{\pi}{4}$$

En fait, **f n'est pas intégrable sur** $]\mathbf{0},\mathbf{1}[^{\mathbf{2}}$. En effet,

$$\int_{]0,1[^2} f_+ \mathrm{d}x\mathrm{d}y = \int_0^1 \left(\int_0^x \frac{x^2-y^2}{(x^2+y^2)^2} \mathrm{d}y \right) \mathrm{d}x = \int_0^1 \left[\frac{y}{x^2+y^2} \right]_{y=0}^{y=x} \mathrm{d}x = \int_0^1 \frac{1}{2x} \mathrm{d}x = +\infty$$

Attention ! Il est possible que les intégrales itérées donnent les mêmes valeurs, mais que la fonction ne soit pas non plus intégrable.

En pratique, il faut donc se souvenir que l'application du théorème de Fubini-Lebesgue est toujours justifiée pour les fonctions mesurables positives, et que dans le cas des fonctions à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$, il faut s'assurer que

$$\int |f| \,\mathrm{d}(\mu \otimes \nu) < \infty$$

ce que l'on fait le plus souvent en appliquant le théorème de Fubini-Tonelli à propos des fonctions positives.

4.4 Applications et exemples

4.4.1 Intégration par parties dans $\mathbb{R}$

Soient $I \subset \mathbb{R}$ un intervalle, et $f, g \colon I \to \mathbb{R}$ deux fonctions **localement intégrables** (i.e. boréliennes et intégrables sur tout compact $K \subset I$). Etant donnés $\alpha, \beta \in I$, on définit les fonctions $F, G \colon I \to \mathbb{R}$ par :

$$f(x) = \int_\alpha^x f(y)\mathrm{d}y = \begin{cases} \int_{[\alpha,x]} f(y)\mathrm{d}y & \text{si } \alpha < x \\ -\int_{[x,\alpha]} f(y)\mathrm{d}y & \text{si } \alpha > x \end{cases}$$

et

$$G(x) = \int_\beta^x g(y)\mathrm{d}y$$

Alors, F et G sont continues, et pour tous $a, b \in I$, $a < b$, on a :

$$\int_a^b f(x)G(x) + \int_a^b F(x)g(x) = F(b)G(b) - F(a)G(a) \tag{4.3}$$

En effet,

i) F et G sont continues grâce au théorème de continuité des intégrales dépendant d'un paramètre [5].

ii) Par la relation de Chasles, on a par définition :

$$\int_a^b f(x)\mathrm{d}x = F(b) - F(a), \qquad \int_a^b g(x)\mathrm{d}x = G(b) - G(a)$$

Par conséquent, (4.3) est inchangée si on ajoute une contante à F ou à G. On supposera donc désormais que $\alpha = \beta = a$.

5. En fait, on ne peut pas montrer dans le cas général que ces fonctions sont $\mathcal{C}^1$, mais on peut montrer qu'elles sont *absolument continues* et presque partout dérivables.

iii) Soit $h\colon\ [a,b]^2 \to \mathbb{R}$ définie par $h(x,y) = f(x)g(y)\mathbb{1}_{\{y \leqslant x\}}$. Alors, h est mesurable (comme produit de fonctions mesurables), et

$$\int_{[a,b]^2} |h(x,y)|\,\mathrm{d}x\mathrm{d}y \leqslant \left(\int_a^b |f(x)|\,\mathrm{d}x\right)\left(\int_a^b |g(y)|\,\mathrm{d}y\right) < \infty$$

Ainsi, h est intégrable sur $[a,b]^2$, et on a par Fubini-Lebesgue :

$$\begin{aligned}\int_a^b \left(\int_a^b h(x,y)\mathrm{d}y\right)\mathrm{d}x &= \int_a^b f(x)\left(\int_a^x g(y)\mathrm{d}y\right)\mathrm{d}x \\ &= \int_a^b f(x)G(x)\mathrm{d}x\end{aligned}$$

et

$$\begin{aligned}\int_a^b \left(\int_a^b h(x,y)\mathrm{d}x\right)\mathrm{d}y &= \int_a^b g(y)\left(\int_y^b f(x)\mathrm{d}x\right)\mathrm{d}y \\ &= \int_a^b g(y)\left(F(b) - F(y)\right)\mathrm{d}y \\ &= F(b)G(b) - \int_a^b g(y)F(y)\mathrm{d}y\end{aligned}$$

ce qui donne le résultat car $F(a) = G(a) = 0$ dans ce cas.

4.4.2 Calcul de l'intégrale de Gauss

Soit $f(x,y) = e^{-x(1+y^2)}$ pour $x, y \geqslant 0$. f est mesurable (car continue) et positive. On a :

$$\int_0^\infty e^{-x(1+y^2)}\mathrm{d}x = \frac{1}{1+y^2}$$

donc par le théorème de Fubini-Tonelli, on a

$$\int_0^\infty \left(\int_0^\infty f(x,y)\mathrm{d}x\right)\mathrm{d}y = \int_0^\infty \frac{1}{1+y^2}\mathrm{d}y = \frac{\pi}{2}$$

En intégrant cette fois-ci selon y en premier, on a pour $x > 0$:

$$\int_0^\infty e^{-x(1+y)^2}\mathrm{d}y = e^{-x}\int_0^\infty e^{-xy^2}\mathrm{d}y \overset{y=z/\sqrt{x}}{=} e^{-x}\int_0^\infty e^{-z^2}\frac{1}{\sqrt{x}}\mathrm{d}z = \frac{e^{-x}}{\sqrt{x}}\left(\int_0^\infty e^{-z^2}\mathrm{d}z\right)$$

D'où, en intégrant selon x :

$$\int_0^\infty \left(\int_0^\infty e^{-x(1+y)^2}\mathrm{d}y\right)\mathrm{d}x = \left(\int_0^\infty e^{-z^2}\mathrm{d}z\right)\left(\int_0^\infty \frac{e^{-x}}{\sqrt{x}}\mathrm{d}x\right) \overset{x=z^2}{=} 2\left(\int_0^\infty e^{-z^2}\mathrm{d}z\right)^2$$

Ainsi, on peut déterminer la valeur de l'intégrale de Gauss :

$$\int_0^\infty e^{-z^2}\mathrm{d}z = \frac{\sqrt{\pi}}{2}, \qquad \int_{\mathbb{R}} e^{-z^2}\mathrm{d}z = \sqrt{\pi}$$

4.4.3 Mesure de la boule unité dans $\mathbb{R}^d$

On se place dans le cadre général $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda)$, et on pose

$$\omega_d = \lambda\left(\left\{x \in \mathbb{R}^d \mid \|x\| \leqslant 1\right\}\right)$$

où $\|.\|$ est la norme euclidienne sur $\mathbb{R}^d$, i.e. $\forall x = (x_1, \dots, x_d) \in \mathbb{R}^d$, $\|x\| = \sqrt{x_1^2 + \cdots + x_d^2}$.
On connaît les valeurs de ω_d pour $d = 1, 2$ et 3. $\omega_1 = 2$ est la longueur du segment $[-1, 1]$, $\omega_2 = \pi$ est l'aire du disque unité et $\omega_3 = \frac{4}{3}\pi$ est le volume de la sphère unité.

Remarque. On observe que $\lambda\left(\left\{x \in \mathbb{R}^d \mid \|x\| \leqslant r\right\}\right) = \omega_d r^d$ pour tout $r \geqslant 0$. On obtient ainsi le volume pour toutes les boules par *dilatation*. Plus généralement, si $A \in \mathcal{B}(\mathbb{R}^d)$, alors le *dilaté* $rA = \{rx \mid x \in A\} \in \mathcal{B}(\mathbb{R}^d), \forall r > 0$ et $\lambda(rA) = r^d\lambda(A)$ (il suffit de le montrer pour les pavés, et comme la mesure extérieure vérifie cette propriété, on en déduira qu'elle est vraie pour tous les boréliens).

Dans la suite, on identifiera $\mathbb{R}^{d+1}$ avec $\mathbb{R}^d \times \mathbb{R}$ pour ne pas s'encombrer d'écritures inutiles, et si $x \in \mathbb{R}^d$, on notera $x_1, \dots, x_d$ ses coordonnées.
Soit $f\colon \mathbb{R}^{d+1} \to \mathbb{R}$ la fonction définie par

$$f(x,t) = e^{-t}\mathbb{1}_{\{\|x\|^2 \leqslant t\}}, \qquad x \in \mathbb{R}^d, t \in \mathbb{R}$$

Alors, f est mesurable pour la tribu $\mathcal{B}(\mathbb{R}^{d+1}) = \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathbb{R})$ et f est positive, donc on a :

$$\int_{\mathbb{R}} f(x,t)\mathrm{d}t = \int_0^{\infty} f(x,t)\mathrm{d}t = \int_{\|x\|^2}^{\infty} e^{-t}\mathrm{d}t = e^{-\|x\|^2}$$

Ainsi,

$$\begin{aligned} \int_{\mathbb{R}^{d+1}} f\mathrm{d}x\mathrm{d}t &= \int_{\mathbb{R}^d} e^{-(x_1^2+\cdots+x_d^2)}\mathrm{d}x \\ &= \left(\int_{\mathbb{R}} e^{-x_1^2}\mathrm{d}x_1\right)\cdots\left(\int_{\mathbb{R}} e^{-x_d^2}\mathrm{d}x_d\right) \\ &= \pi^{\frac{d}{2}} < \infty \end{aligned}$$

Intégrons selon x cette fois-ci. Pour $t > 0$, on a :

$$\begin{aligned} \int_{\mathbb{R}^d} f(x,t)\mathrm{d}x &= e^{-t}\lambda\left(\left\{x \in \mathbb{R}^d \mid \|x\| \leqslant \sqrt{t}\right\}\right) \\ &= e^{-t}\omega_d t^{\frac{d}{2}} \end{aligned}$$

Donc, on en déduit que

$$\int_{\mathbb{R}^{d+1}} f(x,t)\mathrm{d}x\mathrm{d}t = \int_0^{\infty} e^{-t}\omega_d t^{\frac{d}{2}}\mathrm{d}t = \omega_d\Gamma\left(\frac{d}{2}+1\right)$$

où Γ désigne la fonction Γ d'Euler, que l'on a étudiée dans la partie 2.5.3.
Ainsi, on obtient

$$\omega_d = \frac{\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}+1\right)}$$

On retrouve les bons résultats pour ω_1, ω_2 et ω_3. Calculons maintenant la mesure de la boule unité dans $\mathbb{R}^4$:

$$w_4 = \frac{\pi^2}{\Gamma(3)} = \frac{\pi^2}{2}$$

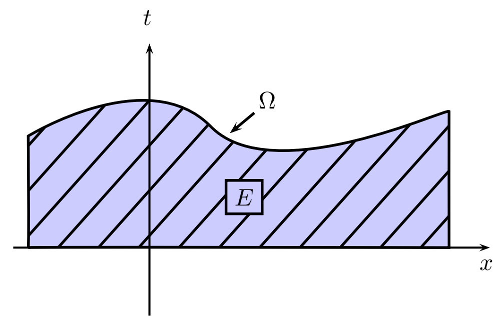


FIGURE 4.1 – Epigraphe d'une fonction réelle à valeurs positives.

4.4.4 Epigraphe d'une fonction mesurable

Ici encore, on identifiera $\mathbb{R}^{d+1}$ avec $\mathbb{R}^d \times \mathbb{R}$ et si $x \in \mathbb{R}^d$, on notera $x_1, \ldots, x_d$ ses coordonnées.

Proposition 4.8. *Soit $f\colon \mathbb{R}^d \to [0, +\infty]$. On définit $E \subset \mathbb{R}^{d+1}$ par*

$$E = \left\{(x,t) \in \mathbb{R}^{d+1} \mid 0 \leqslant t \leqslant f(x)\right\}$$

*E est appelé **l'épigraphe de** f.*
Alors f est mesurable si et seulement si E est mesurable, et dans ce cas $\lambda(E) = \int_{\mathbb{R}^d} f(x)\mathrm{d}x$.

Démonstration. i) Si f est mesurable, on définit $F\colon \mathbb{R}^{d+1} \to \mathbb{R}$ par $F(x,t) = t - f(x)$. Alors F est mesurable, et $E = F^{-1}\left(]-\infty, 0]\right) \cap \left\{(x,t) \in \mathbb{R}^{d+1} \mid t \geqslant 0\right\}$ est mesurable.

ii) Si E est mesurable, on applique le théorème de Fubini-Tonelli à $\mathbb{1}_E$, et on a :

$$\int_{\mathbb{R}} \mathbb{1}_E(x,t)\mathrm{d}t = \int_0^{f(x)} \mathrm{d}t = f(x)$$

donc f est mesurable et $\lambda(E) = \int_{\mathbb{R}^d} f(x)\mathrm{d}x$. $\square$

Remarques. a) Si on applique Fubini-Tonelli dans l'autre sens, on trouve :

$$\forall t \geqslant 0, \qquad \int_{\mathbb{R}^d} \mathbb{1}_E(x,t)\mathrm{d}x = \lambda\left(\left\{x \in \mathbb{R}^d \mid f(x) \geqslant t\right\}\right) := M(t)$$

Ainsi,

$$\int_{\mathbb{R}^d} f(x)\mathrm{d}x = \int_0^\infty M(t)\mathrm{d}t = \int_0^\infty \lambda\left(\left\{x \in \mathbb{R}^d \mid f(x) \geqslant t\right\}\right)\mathrm{d}t$$

b) Plus généralement, soit $g\colon [0, +\infty[\to [0, +\infty[$ continue et notons $G(t) = \int_0^t g(s)\mathrm{d}s$. Appliquons le théorème de Fubini-Tonelli à la fonction $(x,t) \mapsto \mathbb{1}_E(x,t)g(t)$:

- $\int_0^\infty \mathbb{1}_E(x,t)g(t)\mathrm{d}t = \int_0^{f(x)} g(t)\mathrm{d}t = G(f(x))$ pour tout $x \geqslant 0$
- $\int_{\mathbb{R}^d} \mathbb{1}_E(x,t)g(t)\mathrm{d}x = g(t)M(t)$ pour tout $t \geqslant 0$

Ainsi,

$$\int_{\mathbb{R}^d} G(f(x))\mathrm{d}x = \int_0^\infty g(t)M(t)\mathrm{d}t$$

D'où le résultat suivant

$$\int_{\mathbb{R}^d} G(f(x))\mathrm{d}x = \int_0^\infty G'(t)\lambda\left(\left\{x \in \mathbb{R}^d \mid f(x) \geqslant t\right\}\right)\mathrm{d}t$$

c) Soit $\Omega = \left\{x \in \mathbb{R}^d \mid f(x) = t\right\}$ le graphe de f. Si f est mesurable, alors Ω est mesurable, et par Fubini-Tonelli :

$$\begin{aligned}
\lambda_{d+1}(\Omega) &= \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}} \mathbb{1}_\Omega(x,t)\mathrm{d}t\right)\mathrm{d}x = 0 \\
&= \int_0^\infty \left(\int_{\mathbb{R}^d} \mathbb{1}_\Omega(x,t)\mathrm{d}x\right)\mathrm{d}t \\
&= \int_0^\infty \lambda\left(\left\{x \in \mathbb{R}^d \mid f(x) = t\right\}\right)\mathrm{d}t
\end{aligned}$$

Ainsi, pour presque tout $t > 0$, on a $\lambda\left(\left\{x \in \mathbb{R}^d \mid f(x) = t\right\}\right) = 0$.

4.5 Complétion des mesures produit

Soient $(X, \mathcal{M}, \mu)$ et $(Y, \mathcal{N}, \nu)$ deux espaces mesurés, σ-finis et complets (i.e. la mesure associée est complète). En général, la mesure $\mu \otimes \nu$ sur $X \times Y$ n'est pas complète !

Exemple. Si $(X, \mathcal{M}, \mu) = (Y, \mathcal{N}, \nu) = (\mathbb{R}, \mathcal{L}(\mathbb{R}^d), \lambda)$, alors la mesure $\lambda_2 = \lambda \otimes \lambda$ sur $\mathcal{L}(\mathbb{R}) \otimes \mathcal{L}(\mathbb{R})$ n'est pas complète. En particulier, $\mathcal{L}(\mathbb{R}) \otimes \mathcal{L}(\mathbb{R}) \subsetneq \mathcal{L}(\mathbb{R}^2)$. En effet, soit $x \in \mathbb{R}$ et $B \subset \mathbb{R}$ non mesurable[6]. Soit $E = \{x\} \times B$. Alors, avec la proposition 4.3, $E \notin \mathcal{L}(\mathbb{R}) \otimes \mathcal{L}(\mathbb{R})$ car $E_x = B \notin \mathcal{L}(\mathbb{R})$, mais $E \subset \{x\} \times \mathbb{R} \in \mathcal{L}(\mathbb{R}^2)$ et $\lambda(\{x\} \times \mathbb{R}) = 0$.

Heureusement, on a la proposition suivante :

Proposition 4.9. *Soit $d \geqslant 2$ et k, l des entiers tels que $d = k + l$. Alors, la mesure de Lebesgue λ_d sur $\mathcal{L}(\mathbb{R}^d)$ est la complétée de la mesure produit $\lambda_k \otimes \lambda_l$ sur $\mathcal{L}(\mathbb{R}^k) \otimes \mathcal{L}(\mathbb{R}^l)$.*

Ceci nous permet donc d'énoncer une généralisation (mineure) du théorème de Fubini en supposant que les espaces mesurés sont σ-finis et complets.

Proposition 4.10. *Soient $(X, \mathcal{M}, \mu)$ et $(Y, \mathcal{N}, \nu)$ deux espaces mesurés, σ-finis et complets. Soit $f : X \times Y \to \mathbb{R}$ ou $\mathbb{C}$ une fonction mesurable pour la tribu produit complétée $(\mathcal{M} \otimes \mathcal{N})^*$. Alors :*

i) Le théorème de Fubini-Tonelli s'applique si $f \geqslant 0$, à la différence près que les fonction $f_x : y \mapsto f(x, y)$ et $f^y : x \mapsto f(x, y)$ sont simplement mesurables pour presque tous les x ou y.

ii) Le théorème de Fubini-Lebesgue s'applique tel qu'il a été énoncé.

6. On a exhibé un tel ensemble dans la partie 1.4, page 6.

CHAPITRE 5

CHANGEMENTS DE VARIABLES DANS $\mathbb{R}^d$

Ce chapitre permet de présenter un autre outil pour calculer les intégrales multiples dans le cadre "concret" de $\mathbb{R}^d$. Le but est donc de calculer l'image de la mesure de Lebesgue sur $\mathbb{R}^d$ (éventuellement restreint à un ouvert) par un difféomorphisme.
Si $\varphi\colon \mathbb{R}^d \to \mathbb{R}^d$ est une fonction de classe $\mathcal{C}^1$, on notera $D_u\varphi$ la dérivée de $\varphi = (\varphi_1, \dots, \varphi_d)$ en $u = (u_1, \dots, u_d)$, que l'on peut identifier à la matrice des dérivées partielles :

$$(D_u\varphi)_{ij} = \left(\frac{\partial \varphi_i(u)}{\partial u_j}\right)$$

Le Jacobien de φ en u est le déterminant de cette matrice, on le notera $J_\varphi(u)$.

5.1 LA FORMULE DE CHANGEMENT DE VARIABLES

On commence par traiter le cas particulier où le difféomorphisme est une application affine.

Proposition 5.1. *Soit $f\colon \mathbb{R}^d \to \mathbb{R}^d$ définie par $f(x) = Mx + b$, où $M \in GL_d(\mathbb{R})$ est une matrice $d \times d$ inversible, et $b \in \mathbb{R}^d$. Si $A \in \mathcal{B}(\mathbb{R}^d)$, alors $f(A) \in \mathcal{B}(\mathbb{R}^d)$ et*

$$\lambda(f(A)) = |\det(M)|\,\lambda(A)$$

Remarque. La proposition reste valable pour une matrice M non inversible, puisque $f(A) \subset f(\mathbb{R}^d)$ est contenu dans un hyperplan, qui est de mesure de Lebesgue nulle !

On retrouve comme cas particulier de cette proposition l'invariance par translation de λ (proposition 3.13, page 48), l'invariance par rotation qu'il nous restait à montrer, et la propriété de dilatation énoncée lors de la section 4.4.3.

Démonstration. Soit $g = f^{-1}$, i.e. $g(y) = M^{-1}(y - b)$. Si $A \in \mathcal{B}(\mathbb{R}^d)$, alors $f(A) = g^{-1}(A) \in \mathcal{B}(\mathbb{R}^d)$, car $\mathbb{1}_{g^{-1}(A)} = \mathbb{1}_A \circ g$ est mesurable. Il reste à calculer $\lambda(f(A))$. On peut supposer que $b = 0$, car on sait que λ est invariante par translation. Soit $\mu\colon \mathcal{B}(\mathbb{R}^d) \to [0, +\infty]$ définie par $\mu(A) = \lambda(f(A))$ pour tout $A \in \mathcal{B}(\mathbb{R}^d)$. Alors, μ est clairement une mesure invariante par translation (car f est linéaire), et finie sur les bornés. Par la proposition 3.14, il existe une constante $c(M) \geqslant 0$ telle que $\mu = c(M)\lambda$. Il reste à voir que $c(M) = |\det(M)|$. On procède par étapes :

i) Si M est orthogonale (i.e. $M^{-1} = {}^tM$), on note $B = \left\{x \in \mathbb{R}^d \mid \|x\| \leqslant 1\right\}$. Alors, $MB = \{Mx \mid x \in B\} = B$, donc $\mu(B) = \lambda(MB) = \lambda(B) > 0$, et finalement $c(M) = 1$.

ii) Si M est une matrice symétrique définie positive, il existe R une matrice orthogonale telle que

$$ {}^tRMR = D = diag(\alpha_1, \ldots, \alpha_d), \qquad \alpha_1, \ldots, \alpha_d > 0 $$

Soit $P = [0,1]^d$, et $A = RP$. Alors,

$$ \lambda(A) = \lambda(RP) = \lambda(P) = 1 $$

et

$$ \mu(A) = \lambda(MA) = \lambda(RDP) = \lambda(\underbrace{DP}_{=\prod_{i=1}^{n}[0,\alpha_i]}) = \det(M) $$

Dans ce cas, on obtient $c(M) = \det(M)$.

iii) Dans le cas général, on écrit la décomposition polaire de $M = RS$, avec $S = ({}^tMM)^{\frac{1}{2}}$ et $R = MS^{-1}$. S est symétrique définie positive et R est orthogonale. Ainsi, pour tout $A \in \mathcal{B}(\mathbb{R}^d)$, on a :

$$ \mu(A) = \lambda(RSA) = \lambda(SA) = \det(S)\lambda(A) $$

Et on déduit que $c(M) = \det(S) = |\det(M)|$. □

Rappelons ce qu'est un difféomorphisme.

Définition 5.2. Soient U et V deux ouverts (non vides) de $\mathbb{R}^d$. On dit que $\varphi \colon U \to V$ est un difféomorphisme de classe $\mathcal{C}^1$ si :

i) $\varphi \colon U \to V$ est bijective

ii) φ est de classe $\mathcal{C}^1$ sur U

iii) φ^{-1} est de classe $\mathcal{C}^1$ sur V

Remarque. Si $\varphi \colon U \to V$ vérifie i) et ii), alors φ vérifie iii) si et seulement si $J_\varphi(u) \neq 0$ pour tout $u \in U$. C'est le théorème d'inversion locale (le caractère global provenant de l'injectivité de φ sur U).

Proposition 5.3. *Soit $\varphi \colon U \to V$ un difféomorphisme de classe $\mathcal{C}^1$. Si $A \in \mathcal{B}(\mathbb{R}^d)$, $A \subset U$, alors $\varphi(A) \in \mathcal{B}(\mathbb{R}^d)$, et*

$$ \lambda(\varphi(A)) = \int_A |J_\varphi| \,\mathrm{d}\lambda = \int_A |J_\varphi(u)| \,\mathrm{d}u $$

Pour démontrer cette proposition, on va avoir besoin du lemme suivant pour localement se ramener au cas où φ est affine.

Lemme 5.4. *Soit $K \subset U$ un compact, et $\varepsilon > 0$. Il existe $\delta > 0$ tel que pour tous $u_0 \in K$ et $r \in]0, \delta]$, on ait, en notant $P_{u_0,r} = u_0 + [-r, r[^d$,*

$$ (1-\varepsilon)\,|J_\varphi(u_0)|\,\lambda(P_{u_0,r}) \leqslant \lambda(\varphi(P_{u_0,r})) \leqslant (1+\varepsilon)\,|J_\varphi(u_0)|\,\lambda(P_{u_0,r}) $$

Démonstration. Dans cette preuve, on note $\|u\| = \max(|u_1|, \ldots, |u_d|)$ si $u \in \mathbb{R}^d$. Ainsi, $\overline{P_{u_0,r}} = \left\{u \in \mathbb{R}^d \mid \|u - u_0\| \leqslant r\right\}$. De même, si M est une matrice de taille $d \times d$, on note $\|M\| = \sup\left\{\|Mu\| \mid u \in \mathbb{R}^d, \|u\| \leqslant 1\right\}$.

- Comme $K \subset U$ est compact, on a $d(K, U^{\complement}) > 0$, et $D_u\varphi$ est uniformément continue sur K (car continue sur un compact). Etant donné $\varepsilon > 0$, on peut choisir $\delta > 0$ tel que $\delta < d(K, U^{\complement})$, et

$$\|\varphi(u) - \varphi(u_0) - (D_{u_0}\varphi)(u - u_0)\| \leqslant \varepsilon\|u - u_0\|, \quad \forall u_0 \in K, u \in U \mid \|u - u_0\| \leqslant \delta \tag{5.1}$$

- Ecrivons

$$\varphi(u) = \varphi(u_0) + (D_{u_0}\varphi)(u - u_0 + f(u, u_0)) \tag{5.2}$$

où $f(u, u_0) = (D_{u_0}\varphi)^{-1}(\varphi(u) - \varphi(u_0) - D_{u_0}\varphi(u - u_0))$. Alors, par (5.1), on a

$$\|f(u, u_0)\| \leqslant a\varepsilon\|u - u_0\|, \quad \forall u_0 \in K, \forall u \in U \mid \|u - u_0\| \leqslant \delta \tag{5.3}$$

où $a = \sup_{u_0 \in K} \left\|(D_{u_0}\varphi)^{-1}\right\|$.

- Soit $r > 0$ tel que $r \leqslant \delta$. Il suit de (5.2) et (5.3) que :

$$\varphi(P_{u_0,r}) \subset \varphi(u_0) + (D_{u_0}\varphi)\left((1 + a\varepsilon)P_{0,r}\right)$$

donc $\lambda(\varphi(P_{u_0,r})) \leqslant |J_\varphi(u_0)|\,(1 + a\varepsilon)^d\lambda(P_{0,r})$, et c'est la borne supérieure recherchée.
De même, si $a\varepsilon < 1$, on a :

$$\varphi(P_{u_0,r}) \supset \varphi(u_0) + (D_{u_0}\varphi)\left((1 - a\varepsilon)P_{0,r}\right) \qquad \square$$

donc $\lambda(\varphi(P_{u_0,r})) \geqslant |J_\varphi(u_0)|\,(1 - a\varepsilon)^d\lambda(P_{0,r})$, et c'est la borne inférieure recherchée.

Démonstration (Démonstration de la proposition 5.3). Pour tout entier n, on note Q_n l'ensemble des pavés élémentaires d'ordre n de la forme

$$P = \prod_{i=1}^{d} \left[2^{-n}k_i, 2^{-n}(k_i + 1)\right[, \quad k = (k_1, \ldots, k_d) \in \mathbb{Z}^d$$

On note $z_P = 2^{-n}\left(k_1 + \frac{1}{2}, k_2 + \frac{1}{2}, \ldots, k_d + \frac{1}{2}\right)$ le centre d'un tel pavé P.
Soit P_0 un pavé élémentaire d'ordre n_0 tel que $\overline{P_0} \subset U$ et soit $\varepsilon > 0$. On choisit un $n \geqslant n_0$ assez grand pour que :

1) le lemme 5.4 ci-dessus soit vrai avec $K = \overline{P_0}$ et $\delta = 2^{-(n+1)}$.
2) Pour tous $u, v \in \overline{P_0} = K$, avec $\|u - v\| \leqslant \delta$, on ait :

$$(1 - \varepsilon)\,|J_\varphi(u)| \leqslant |J_\varphi(v)| \leqslant (1 + \varepsilon)\,|J_\varphi(u)| \tag{5.4}$$

Comme $P_0 = \bigcup_{\substack{P \in Q_n \\ P \subset P_0}} P$, on a :

$$\begin{aligned}
\lambda(\varphi(P_0)) &= \sum_{\substack{P \in Q_n \\ P \subset P_0}} \lambda(\varphi(P)), \qquad \text{or } P = P_{z_P, 2^{-(n+1)}} \\
&\underset{\text{lemme 5.4}}{\leqslant} (1 + \varepsilon) \sum_{\substack{P \in Q_n \\ P \subset P_0}} |J_\varphi(z_P)|\,\lambda(P) \\
&\underset{(5.4)}{\leqslant} (1 + \varepsilon)^2 \sum_{\substack{P \in Q_n \\ P \subset P_0}} \int_P |J_\varphi(u)|\,\mathrm{d}u \\
&= (1 + \varepsilon)^2 \int_{P_0} |J_\varphi(u)|\,\mathrm{d}u
\end{aligned}$$

De même, on a
$$\lambda(\varphi(P_0)) \geqslant (1-\varepsilon)^2 \int_{P_0} |J_\varphi(u)|\,\mathrm{d}u, \quad \text{si } \varepsilon < 1$$
Comme $\varepsilon > 0$ était arbitraire, on a
$$\lambda(\varphi(P_0)) = \int_{P_0} |J_\varphi(u)|\,\mathrm{d}u$$
Considérons à présent les mesures boréliennes μ et ν sur U définies par :
$$\mu(A) = \lambda(\varphi(A)), \qquad \nu(A) = \int_A |J_\varphi(u)|\,\mathrm{d}u$$
On sait que μ et ν coïncident sur la famille des pavés élémentaires dont l'adhérence est un compact de U. Cette famille est stable par intersection finie, et engendre la tribu borélienne $\mathcal{B}(U)$. Par ailleurs, si $\mathcal{U}_n$ désigne l'union des pavés élémentaires d'ordre n dont l'adhérence est incluse dans $U \cap \{x \in \mathbb{R}^d \mid \|x\| \leqslant n\}$, on a $\mathcal{U}_n \subset \mathcal{U}_{n+1}$, $\bigcup_{n\in\mathbb{N}} \mathcal{U}_n = U$ et $\mu(\mathcal{U}_n) = \nu(\mathcal{U}_n) < \infty$ pour tout $n \in \mathbb{N}$. Par le corollaire 3.11 du lemme des classes monotones, on conclut que $\mu = \nu$ sur $\mathcal{B}(U)$. □

On peut à présent énoncer le théorème de changement de variable dans $\mathbb{R}^d$.

Théorème 5.5 (du changement de variable dans $\mathbb{R}^d$). *Soient U et V deux ouverts de $\mathbb{R}^d$, et $\varphi\colon U \to V$ un difféomorphisme de classe $\mathcal{C}^1$.*

1) Si $f\colon V \to [0,+\infty]$ est une application mesurable, alors $(f\circ\varphi)\,|J_\varphi| : U \to [0,+\infty]$ est mesurable, et
$$\int_V f(v)\mathrm{d}v = \int_U f(\varphi(u))\,|J_\varphi(u)|\,\mathrm{d}u \tag{5.5}$$
2) Si $f\colon V \to \mathbb{R}$ ou $\mathbb{C}$ est une application intégrable, alors $(f\circ\varphi)\,|J_\varphi| : U \to \mathbb{R}$ ou $\mathbb{C}$ est intégrable, et
$$\int_V f(v)\mathrm{d}v = \int_U f(\varphi(u))\,|J_\varphi(u)|\,\mathrm{d}u \tag{5.6}$$

Comme dans le théorème de Fubini, le théorème se décompose en deux parties. La première, moins générale, aide à trouver les hypothèses de la seconde afin de pouvoir l'appliquer.

Démonstration. Soit $B \subset V$ un borélien, et soit $A = \varphi^{-1}(B) \subset U$. Alors, si $f = \mathbb{1}_B$, on a $f\circ\varphi = \mathbb{1}_B \circ \varphi = \mathbb{1}_{\varphi^{-1}(B)} = \mathbb{1}_A$, et la formule (5.5) devient simplement
$$\lambda(\varphi(A)) = \int_A |J_\varphi(u)|\,\mathrm{d}u$$
qui a été démontré dans la proposition précédente. Ainsi, par linéarité, (5.5) est vraie pour les fonctions étagées positives, et par passage à la limite croissante, pour toute fonction f mesurable positive.
Si f est intégrable et réelle, on obtient (5.6) en appliquant (5.5) à f_+ et f_-.
Si f est intégrable et complexe, on applique (5.6) à $\Re(f)$ et $\Im(f)$. □

Remarques. 1) Le théorème est vrai si l'on travaille avec $\mathcal{L}(\mathbb{R}^d)$ au lieu de $\mathcal{B}(\mathbb{R}^d)$. En effet, si $A \in \mathcal{L}(\mathbb{R}^d)$, il existe $A_1, A_2 \in \mathcal{B}(\mathbb{R}^d)$ avec $A_1 \subset A \subset A_2$ et $\lambda(A_2 \setminus A_1) = 0$. Si $A \subset U$, alors on peut prendre $A_1, A_2 \subset U$, et on a $\varphi(A_1) \subset \varphi(A) \subset \varphi(A_2)$ avec :
$$\lambda(\varphi(A_2)\setminus\varphi(A_1)) = \lambda(\varphi(A_2\setminus A_1)) = \int_{A_2\setminus A_1} |J_\varphi(u)|\,\mathrm{d}u = 0$$

Ainsi, $\varphi(A) \in \mathcal{L}(\mathbb{R}^d)$ et

$$\lambda(\varphi(A)) = \lambda(\varphi(A_1)) = \int_{A_1} |J_\varphi(u)| \,\mathrm{d}u = \int_A |J_\varphi(u)| \,\mathrm{d}u$$

2) On peut parfois généraliser un peu la classe des applications φ qui interviennent dans la formule du changement de variables. Par exemple, pour $d = 1$, si $f \colon \mathbb{R} \to \mathbb{R}$ est continue et $\varphi \colon \mathbb{R} \to \mathbb{R}$ est de classe $\mathcal{C}^1$, alors pour tous $a, b \in \mathbb{R}$, on a :

$$\int_{\varphi(a)}^{\varphi(b)} f(y)\mathrm{d}y = \int_a^b f(\varphi(x))\varphi'(x)\mathrm{d}x$$

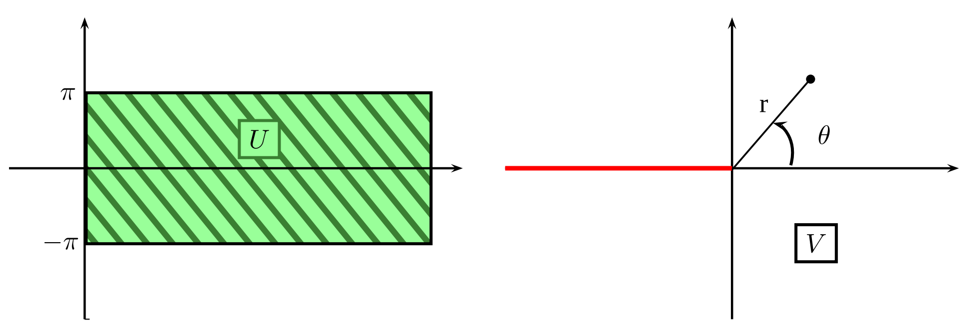
Ici, on n'a aucune hypothèse sur la bijectivité de φ, et pourtant la formule fonctionne encore.

5.2 Applications

Sont présentées ici quelques applications très classiques à connaître absolument.

5.2.1 Coordonnées polaires dans le plan $\mathbb{R}^2$

Soient $U = \{(r, \theta) \mid 0 < r < \infty, -\pi < \theta < \pi\}$ et $V = \mathbb{R}^2 \setminus \{(x, 0) \mid x \leqslant 0\}$.



Soit $\varphi \colon U \to V$ définie par $\varphi(r, \theta) = (r\cos\theta, r\sin\theta)$. Alors φ est bijective, de classe $\mathcal{C}^\infty$ et

$$J_\varphi(r, \theta) = \det \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} = r > 0$$

Ainsi, φ est un difféomorphisme de classe $\mathcal{C}^\infty$, et en appliquant le théorème, on obtient le corollaire suivant :

Corollaire 5.6. *Si f est une fonction mesurable positive ou intégrable définie sur $\mathbb{R}^2$, alors*

$$\int_{\mathbb{R}^2} f(x, y)\mathrm{d}x\mathrm{d}y = \int_U f(r\cos\theta, r\sin\theta) r\mathrm{d}r\mathrm{d}\theta$$

Remarque. En fait, $\int_{\mathbb{R}^2} f\mathrm{d}\lambda = \int_V f\mathrm{d}\lambda$ car $\lambda(\mathbb{R}^2 \setminus V) = 0$.

En particulier, si f est radiale (i.e. s'il existe f_0 telle que $f(x,y) = f_0(\sqrt{x^2+y^2})$, c'est-à-dire qu'elle est invariante par rotation), on obtient à l'aide de Fubini,

$$\int_{\mathbb{R}^2} f(x,y)\mathrm{d}x\mathrm{d}y = \int_0^\infty \underbrace{(2\pi r)}_{\text{la longueur du cercle de rayon } r} f_0(r)\mathrm{d}r$$

Application : Soit $f\colon \mathbb{R}^2 \to [0,+\infty]$ définie par

$$f(x,y) = \frac{1}{(x^2+y^2)^{\frac{\alpha}{2}}}, \quad \alpha > 0$$

Alors[1], f est intégrable sur $B = \{(x,y) \mid x^2+y^2 \leqslant 1\}$ si et seulement si $\alpha < 2$.
f est intégrable sur $B^\complement$ si et seulement si $\alpha > 2$.

En effet,

$$\int_B f(x,y)\mathrm{d}x\mathrm{d}y = \int_0^\infty \mathbb{1}_{\{r\leqslant 1\}}(2\pi r)\frac{1}{r^\alpha}\mathrm{d}r = 2\pi\int_0^1 r^{1-\alpha}\mathrm{d}r = \begin{cases} \frac{2\pi}{2-\alpha} & \text{si } \alpha < 2 \\ +\infty & \text{si } \alpha \geqslant 2 \end{cases}$$

et

$$\int_{B^\complement} f(x,y)\mathrm{d}x\mathrm{d}y = 2\pi\int_1^\infty r^{1-\alpha}\mathrm{d}r = \begin{cases} \frac{2\pi}{\alpha-2} & \text{si } \alpha > 2 \\ +\infty & \text{si } \alpha \leqslant 2 \end{cases}$$

Application (intégrale de Gauss) : Par le théorème de Fubini-Tonelli, on a que

$$\int_0^\infty e^{-x^2}\mathrm{d}x = \left(\int_0^\infty e^{-x^2}\mathrm{d}x \times \int_0^\infty e^{-y^2}\mathrm{d}y\right)^{\frac{1}{2}} = \left(\int_{]0,+\infty[^2} e^{-(x^2+y^2)}\mathrm{d}x\mathrm{d}y\right)^{\frac{1}{2}}$$

D'autre part, il vient par passage en coordonnées polaires,

$$\int_{]0,+\infty[^2} e^{-(x^2+y^2)}\mathrm{d}x\mathrm{d}y \quad = \quad \int_0^{\frac{\pi}{2}}\int_0^\infty e^{-r^2}r\mathrm{d}r\mathrm{d}\theta = \frac{\pi}{2}\int_0^\infty e^{-r^2}r\mathrm{d}r = \frac{\pi}{4}$$

On retrouve ainsi la valeur de l'intégrale de Gauss.

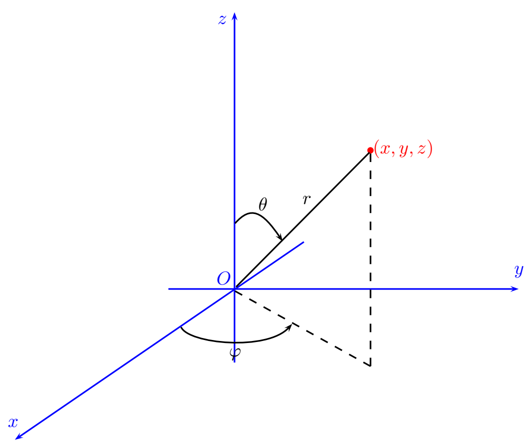
5.2.2 Coordonnées sphériques dans l'espace $\mathbb{R}^3$

Soient $U = \{(r,\theta,\varphi) \mid 0 < r < \infty, 0 < \theta < \pi, -\pi < \varphi < \pi\}$ et $V = \mathbb{R}^3 \setminus \{(x,0,z) \mid x \leqslant 0, z \in \mathbb{R}\}$.
Soit $\psi\colon U \to V$ définie par $\psi(r,\theta,\varphi) = (r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta)$.
Alors ψ est bijective, de classe $\mathcal{C}^\infty$ et

$$J_\psi(r,\theta,\varphi) = \det\begin{pmatrix} \sin\theta\cos\varphi & r\cos\theta\cos\varphi & -r\sin\theta\sin\varphi \\ \sin\theta\sin\varphi & r\cos\theta\sin\varphi & r\sin\theta\cos\varphi \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix} = r^2\sin\theta > 0$$

Ainsi, ψ est un difféomorphisme de classe $\mathcal{C}^\infty$, et en appliquant le théorème, on obtient le corollaire suivant :

1. On pourrait prendre un voisinage de 0 à la place de la boule unité.



Corollaire 5.7. *Si f est une fonction mesurable positive ou intégrable définie sur $\mathbb{R}^3$, alors*

$$\int_{\mathbb{R}^3} f(x, y, z)\mathrm{d}x\mathrm{d}y\mathrm{d}z = \int_U f(r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta)r^2\sin\theta\mathrm{d}r\mathrm{d}\theta\mathrm{d}\varphi$$

Remarque. En fait, $\int_{\mathbb{R}^3} f\mathrm{d}\lambda = \int_V f\mathrm{d}\lambda$ car $\lambda(\mathbb{R}^3 \setminus V) = 0$.

En particulier, si f est radiale (i.e. s'il existe f_0 telle que $f(x, y, z) = f_0(\sqrt{x^2 + y^2 + z^2})$), on obtient à l'aide de Fubini,

$$\begin{aligned}
\int_{\mathbb{R}^3} f(x, y, z)\mathrm{d}x\mathrm{d}y\mathrm{d}z &= \int_U f_0(r)r^2\sin\theta\mathrm{d}r\mathrm{d}\theta\mathrm{d}\varphi \\
&= \int_0^\infty f_0(r)r^2\left(\int_0^\pi \sin\theta\left(\int_{-\pi}^\pi \mathrm{d}\varphi\right)\mathrm{d}\theta\right)\mathrm{d}r \\
&= \int_0^\infty (4\pi r^2) f_0(r)\mathrm{d}r
\end{aligned}$$

l'aire de la sphère de rayon r

Application : Soit $f : \mathbb{R}^3 \to [0, +\infty]$ définie par

$$f(x, y) = \frac{1}{(x^2 + y^2 + z^2)^{\frac{\alpha}{2}}}, \quad \alpha > 0$$

Alors [2], f est intégrable sur $B = \{(x, y, z) \mid x^2 + y^2 + z^2 \leqslant 1\}$ si et seulement si $\alpha < 3$.
f est intégrable sur $B^{\complement}$ si et seulement si $\alpha > 3$.

En effet,

$$\int_B f(x, y, z)\mathrm{d}x\mathrm{d}y\mathrm{d}z = \int_0^\infty \mathbb{1}_{\{r \leqslant 1\}}(4\pi r^2)\frac{1}{r^\alpha}\mathrm{d}r = 4\pi\int_0^1 r^{2-\alpha}\mathrm{d}r = \begin{cases} \frac{4\pi}{3-\alpha} & \text{si } \alpha < 3 \\ +\infty & \text{si } \alpha \geqslant 3 \end{cases}$$

et

$$\int_{B^{\complement}} f(x, y, z)\mathrm{d}x\mathrm{d}y\mathrm{d}z = 4\pi\int_1^\infty r^{2-\alpha}\mathrm{d}r = \begin{cases} \frac{4\pi}{\alpha-3} & \text{si } \alpha > 3 \\ +\infty & \text{si } \alpha \leqslant 3 \end{cases}$$

2. On pourrait prendre un voisinage de 0 à la place de la boule unité.

5.2.3 Généralisation à $\mathbb{R}^d$

Définition 5.8. Soit $d \in \mathbb{N}^*$. On note $S^{d-1} = \{x \in \mathbb{R}^d \mid \|x\| = 1\}$, où $\|\cdot\|$ représente la norme euclidienne. Etant donné $A \subset S^{d-1}$, on note $\gamma(A)$ le secteur :

$$\gamma(A) = \left\{rx \in \mathbb{R}^d \mid x \in A, 0 < r \leqslant 1\right\}$$

On appelle **mesure euclidienne de surface** la mesure borélienne sur S^{d-1} définie par :

$$\sigma_d(A) = d\lambda_d(\gamma(A)), \qquad \forall A \in \mathcal{B}(S^{d-1})$$

On vérifie facilement que l'on a bien défini une mesure.

Remarque. Calculons la valeur de la mesure de la sphère unité :

$$\sigma_d(S^{d-1}) = d\lambda_d(\{x \mid \|x\| \leqslant 1\}) = d\frac{\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}+1\right)} = \frac{2\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)}$$

En particulier, $\sigma_1(S^0) = 2$ (c'est la mesure de comptage sur $S^0 = \{-1, 1\}$), $\sigma_2(S^1) = 2\pi$ (circonférence du cercle unité dans le plan), $\sigma_3(S^2) = 4\pi$ (aire de la sphère unité dans l'espace), ...

Proposition 5.9. *Si f est une fonction mesurable positive, ou intégrable, définie sur $\mathbb{R}^d$, alors*

$$\int_{\mathbb{R}^d} f \mathrm{d}\lambda = \int_0^\infty \int_{S^{d-1}} f(r, \omega) r^{d-1} \mathrm{d}r \mathrm{d}\sigma_\omega \tag{5.7}$$

où $\lambda = \lambda_d$ et $\sigma = \sigma_d$.

En particulier, si f est radiale (i.e. s'il existe f_0 telle que $f(x) = f_0(\|x\|)$), alors :

$$\int_{\mathbb{R}^d} f \mathrm{d}\lambda = C_d \int_0^\infty f_0(r) r^{d-1} \mathrm{d}r, \qquad \text{où } C_d = \frac{2\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)}$$

Application : Soit $f \colon \mathbb{R}^d \to [0, +\infty]$ définie par

$$f(x) = \frac{1}{\|x\|^\alpha}, \quad \alpha > 0$$

Alors, comme précédemment,

f est intégrable sur un voisinage borné de l'origine si et seulement si $\alpha < d$.
f est intégrable sur le complémentaire d'un tel voisinage si et seulement si $\alpha > d$.

Démonstration (Démonstration de la proposition 5.9). Il suffit de vérifier (5.7) pour $f = \mathbb{1}_B$ avec $B \in \mathcal{B}(U) = \mathcal{B}(\mathbb{R}^d \setminus \{0\})$. Définissons $\mu \colon \mathcal{B}(U) \to [0, +\infty]$ par :

$$\mu(B) = \int_0^\infty \int_{S^{d-1}} \mathbb{1}_B(r\omega) r^{d-1} \mathrm{d}r \mathrm{d}\sigma_\omega$$

Alors, on vérifie que μ est une mesure borélienne sur $\mathbb{R}^d \setminus \{0\}$. On veut montrer que $\mu = \lambda$. D'après le corollaire 3.11 au lemme des classes monotones, il suffit de vérifier que $\mu(B) = \lambda(B)$ pour un B de la forme :

$$B = \left\{ x \in U \ \mid \ \frac{x}{\|x\|} \in A, a < \|x\| \leqslant b \right\}, \qquad \text{avec } A \in \mathcal{B}(S^{d-1}) \text{ et } 0 < a < b$$

En effet, une telle famille vérifie les hypothèses voulues (stabilité par intersection finie, et U est l'union dénombrables de telles parties). On a donc :

•

$$\mu(B) = \int_0^{\infty} \int_{S^{d-1}} \mathbb{1}_{]a,b]} \mathbb{1}_A(\omega) r^{d-1} \mathrm{d}r \mathrm{d}\sigma_\omega = \sigma(A) \int_a^b r^{d-1} \mathrm{d}r = \sigma(A) \frac{b^d - a^d}{d}$$

• Par ailleurs, on note $\gamma_n(A) = \left\{ rx \in \mathbb{R}^d \mid \alpha^{n+1} < r \leqslant \alpha^n, x \in A \right\}$ où $\alpha = \frac{a}{b} \in]0,1[$ et $A \in S^{d-1}$. Alors, $\gamma_n(A) \cap \gamma_m(A) = \varnothing$ si $n \neq m$, et $\gamma(A) = \bigcup_{n \in \mathbb{N}} \gamma_n(A)$. Par définition, on a $B = b\gamma_0(A)$. L'invariance par dilatation de λ donne que $\lambda(\gamma_n(A)) = \alpha^{dn} \lambda(\gamma_0(A))$. Ainsi,

$$\lambda(\gamma(A)) = \sum_{n \in \mathbb{N}} \lambda(\gamma_n(A)) = \frac{\lambda(\gamma_0(A))}{1 - \alpha^d}$$

et

$$\lambda(B) = b^d \lambda(\gamma_0(A))$$

Donc,

$$\lambda(B) = b^d (1 - \alpha^d) \lambda(\gamma(A)) = \frac{b^d - a^d}{d} \sigma(A) = \mu(B)$$

□

CHAPITRE 6

ESPACES $\mathcal{L}^p$ ET L^p

Dans tous les chapitres précédents, nous avons considéré les fonctions "individuellement". Dans ce chapitre, nous étudierons des familles entières de fonctions, c'est-à-dire des "espaces de fonctions", ou **espaces fonctionnels**.
Dans tout ce chapitre, on considère un espace mesuré $(X, \mathcal{M}, \mu)$, et on notera $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

6.1 GÉNÉRALITÉS SUR LES ESPACES $\mathcal{L}^p$ ET L^p

Définition 6.1. Soit $p \in [1, +\infty[$. On définit :

$$\mathcal{L}^p(X, \mathcal{M}, \mu) = \left\{ f\colon X \to \mathbb{K} \text{ mesurable} \mid \int_X |f|^p \,\mathrm{d}\mu < \infty \right\}$$

Pour $f \in \mathcal{L}^p(X, \mathcal{M}, \mu)$, on note

$$\|f\|_p = \left(\int_X |f|^p \,\mathrm{d}\mu \right)^{\frac{1}{p}}$$

Les cas les plus importances sont $p = 1$ (i.e. l'ensemble $\mathcal{L}^1(X, \mathcal{M}, \mu)$ des fonctions intégrables, déjà défini), et $p = 2$ (i.e. l'ensemble des fonctions de carré intégrable).

Définition 6.2. On dit qu'une fonction $f\colon X \to \mathbb{K}$ est **essentiellement bornée** s'il existe un réel $A \geqslant 0$ tel que $|f| \leqslant A$ μ-presque partout, i.e.

$$\mu\left(\{x \in X \mid |f(x)| > A\}\right) = 0$$

On note $\mathcal{L}^\infty(X, \mathcal{M}, \mu)$ l'ensemble des fonctions essentiellement bornées sur X.
Pour $f \in \mathcal{L}^\infty(X, \mathcal{M}, \mu)$, on note

$$\|f\|_\infty = \sup_{x \in X} \operatorname{ess} |f(x)| = \inf\left(\{A \geqslant 0 \mid |f| \leqslant A \ \mu\text{-presque partout}\}\right)$$

Remarque. En fait, $\|f\|_\infty = \min\{A \geqslant 0 \mid |f| \leqslant A \ \mu\text{-presque partout}\}$, de sorte que

$$|f(x)| \leqslant \|f\|_\infty, \quad \text{pour } \mu\text{-presque tout } x \in X$$

En effet, soient $E_n = \left\{x \in X \mid |f(x)| > \|f\|_\infty + \frac{1}{n}\right\}$ et $E = \{x \in X \mid |f(x)| > \|f\|_\infty\}$. Alors, $E = \bigcup_{n \in \mathbb{N}} E_n$ et pour tout $n \in \mathbb{N}$, $\mu(E_n) = 0$ par définition de $\|f\|_\infty$, donc $\mu(E) = 0$.

Définition 6.3. On dit que $p, q \in [1, +\infty]$ sont des **exposants conjugués** si $\frac{1}{p} + \frac{1}{q} = 1$ (avec la convention $\frac{1}{\infty} = 0$).

Proposition 6.4 (Inégalité de Hölder). *Soient $p, q \in [1, +\infty]$ des exposants conjugués. Si $f, g\colon X \to \mathbb{K}$ sont des applications mesurables, alors*

$$\|fg\|_1 \leqslant \|f\|_p \|g\|_q \tag{6.1}$$

En particulier, si $f \in \mathcal{L}^p(X, \mathcal{M}, \mu)$ et $g \in \mathcal{L}^q(X, \mathcal{M}, \mu)$, alors $fg \in \mathcal{L}^1(X, \mathcal{M}, \mu)$ et l'inégalité (6.1) est vraie.

Remarque. Si $p, q \in]1, +\infty[$, l'inégalité (6.1) devient :

$$\int_X |fg| \,\mathrm{d}\mu \leqslant \left(\int_X |f|^p \,\mathrm{d}\mu\right)^{\frac{1}{p}} \left(\int_X |g|^q \,\mathrm{d}\mu\right)^{\frac{1}{q}} \tag{6.2}$$

et si $p = \infty$ et $q = 1$, on a :

$$\int_X |fg| \,\mathrm{d}\mu \leqslant \left(\sup_{x \in X} \text{ess} \, |f(x)|\right) \left(\int_X |g| \,\mathrm{d}\mu\right) \tag{6.3}$$

Cette dernière inégalité est évidente car $|fg| \leqslant \|f\|_\infty \, |g|$ μ-presque partout.

La démonstration de l'inégalité de Hölder repose sur le lemme suivant :

Lemme 6.5. *Soient $p, q \in]1, +\infty[$ deux exposants conjugués. Si $a, b \geqslant 0$, on a :*

$$ab \leqslant \frac{1}{p} a^p + \frac{1}{q} b^q$$

avec égalité si et seulement si $a^p = b^q$.

Démonstration. On peut supposer que $a, b > 0$ sans quoi c'est évident. Notons $x = p \ln(a)$ et $y = q \ln(b)$, de sorte que $a = e^{x/p}$ et $b = e^{y/q}$. Comme la fonction $t \mapsto \exp(t)$ est strictement convexe, on a

$$ab = \exp\left(\frac{1}{p} x + \frac{1}{q} y\right) \leqslant \frac{1}{p} \exp(x) + \frac{1}{q} \exp(y) = \frac{1}{p} a^p + \frac{1}{q} b^q$$

avec égalité si et seulement si $x = y$, i.e. $a^p = b^q$. $\square$

Démonstration (Démonstration de la proposition 6.4). Si f et g sont deux fonctions mesurables, le produit fg l'est aussi. Il suffit donc de montrer (6.1). Les cas $(p = 1, q = \infty)$ et $(p = \infty, q = 1)$ sont faciles. On suppose donc que $1 < p, q < \infty$.
Si $f = 0$ (ou $g = 0$) μ-presque partout, alors $fg = 0$ μ-presque partout et les deux membres de (6.2) sont nuls. On suppose donc dorénavant que $\|f\|_p > 0$ et $\|g\|_q > 0$.
Dans ce cas, on peut supposer que $\|f\|_p < \infty$ et $\|g\|_q < \infty$, sinon le membre de droite de (6.2) vaut $+\infty$ et il n'y a rien à montrer.
Supposons donc que $1 < p, q < \infty$ et que $0 < \|f\|_p, \|g\|_q < \infty$. Définissons les fonctions suivantes :

$$F(x) = \frac{|f(x)|}{\|f\|_p}, \qquad G(x) = \frac{|g(x)|}{\|g\|_q}$$

Alors, $F \in \mathcal{L}^p(X, \mathcal{M}, \mu)$ et $G \in \mathcal{L}^q(X, \mathcal{M}, \mu)$ avec $\|F\|_p = \|G\|_q = 1$.
En outre, par le lemme,

$$F(x)G(x) \leqslant \frac{1}{p}F(x)^p + \frac{1}{q}G(x)^q$$

En intégrant les deux membres, on trouve :

$$\frac{1}{\|f\|_p\|g\|_q}\left(\int_X |fg| \, \mathrm{d}\mu\right) \leqslant \frac{1}{p}\|F\|_p^p + \frac{1}{q}\|G\|_q^q = \frac{1}{p} + \frac{1}{q} = 1$$

□

Remarque. Si $1 < p, q < \infty$, on a égalité dans (6.2) si et seulement si

$$\alpha\,|f(x)|^p = \beta\,|g(x)|^q \qquad \text{avec } \alpha, \beta > 0, \alpha + \beta > 0$$

Si $p = \infty, q = 1$, on a égalité dans (6.3) si et seulement si

$$|g|\,(\|f\|_\infty - |f|) = 0 \qquad \mu\text{-presque partout}$$

Justifions maintenant que les applications $f \mapsto \|f\|_p$ introduites pour $p \in [1, +\infty]$ sont des semi-normes, ce qui justifiera la définition des espaces L^p pour qu'elles deviennent des normes. Nous allons tout d'abord nous intéresser à l'inégalité triangulaire, dite de Minkowski.

Proposition 6.6 (Inégalité de Minkowski). *Soit un réel $p \in [1, +\infty]$. Si $f, g\colon X \to \mathbb{K}$ sont deux applications mesurables, alors*

$$\|f + g\|_p \leqslant \|f\|_p + \|g\|_p \tag{6.4}$$

En particulier, si $f, g \in \mathcal{L}^p(X, \mathcal{M}, \mu)$, alors $f + g \in \mathcal{L}^p(X, \mathcal{M}, \mu)$ et l'inégalité (6.4) est vraie.

Démonstration. Comme $|f + g| \leqslant |f| + |g|$, alors le résultat est évident pour $p = 1$ et $p = \infty$. On suppose donc que $1 < p < \infty$, et que $f, g \in \mathcal{L}^p(X, \mathcal{M}, \mu)$ avec $f \geqslant 0$ et $g \geqslant 0$ (sinon, on remplace f par $|f|$ et g par $|g|$). La fonction $t \mapsto t^p$ étant convexe sur $[0, +\infty[$, on a donc pour tous $a, b \geqslant 0$:

$$\left(\frac{1}{2}a + \frac{1}{2}b\right)^p \leqslant \frac{1}{2}a^p + \frac{1}{2}b^p$$

donc, on en déduit immédiatement que

$$(a + b)^p \leqslant 2^{p-1}(a^p + b^p)$$

Ainsi,

$$(f(x) + g(x))^p \leqslant 2^{p-1}\left(f(x)^p + g(x)^p\right)$$

et on en déduit que $f + g \in \mathcal{L}^p(X, \mathcal{M}, \mu)$.
Par ailleurs, posons $h = f + g \geqslant 0$. Alors on a

$$h^p = (f + g)h^{p-1} = fh^{p-1} + gh^{p-1}$$

Soit $q = \frac{p}{p-1}$ l'exposant conjugué de p, on a alors

$$\|h^{p-1}\|_q = \left(\int_X h^p \mathrm{d}\mu\right)^{\frac{1}{q}} = \|h\|_p^{\frac{p}{q}} = \|h\|_p^{p-1} < \infty$$

Par l'inégalité de Hölder, on en déduit donc que

$$\|h\|_p^p \leqslant \|f\|_p\|h^{p-1}\|_q + \|g\|_p\|h^{p-1}\|_q = \|h\|_p^{p-1}\left(\|f\|_p + \|g\|_p\right)$$

Si $\|h\|_p > 0$, on obtient (6.4) en divisant les deux nombres par $\|h\|_p^{p-1}$, et si $\|h\|_p = 0$, alors $f = 0$ μ-presque partout et $g = 0$ μ-presque partout, donc les deux membres de (6.4) sont nuls. □

Soit $p \in [1, +\infty]$. Il suit de l'inégalité de Minkowski que $\mathcal{L}^p(X, \mathcal{M}, \mu)$ est un espace vectoriel, et que l'application $\| \cdot \|_p \colon f \mapsto \|f\|_p$ est une semi-norme sur $\mathcal{L}^p(X, \mathcal{M}, \mu)$. En effet,
– $\|f\|_p \geqslant 0$ pour toute application $f \in \mathcal{L}^p(X, \mathcal{M}, \mu)$
– $\|\lambda f\|_p = |\lambda| \, \|f\|_p$ pour toute application $f \in \mathcal{L}^p(X, \mathcal{M}, \mu)$ et tout $\lambda \in \mathbb{K}$
– $\|f + g\|_p \leqslant \|f\|_p + \|g\|_p$ pour toutes applications $f, g \in \mathcal{L}^p(X, \mathcal{M}, \mu)$

En revanche, $\|f\|_p = 0 \iff f(x) = 0$ **pour μ-presque tout** $x \in X$, donc $\| \cdot \|_p$ n'est pas une norme si la tribu $\mathcal{M}$ contient des ensembles non vides de mesure nulle [1].
Considérons alors $\mathcal{R}$ la relation d'équivalence sur $\mathcal{L}^p(X, \mathcal{M}, \mu)$ définie par

$$f \mathcal{R} g \text{ si et seulement si } f(x) = g(x) \text{ pour } \mu\text{-presque tout } x \in X, \text{ i.e. } \|f - g\|_p = 0$$

Etant donnée une application $f \in \mathcal{L}^p(X, \mathcal{M}, \mu)$, on note $[f]$ la classe d'équivalence de f pour la relation $\mathcal{R}$, c'est à dire que

$$[f] = \{g \in \mathcal{L}^p(X, \mathcal{M}, \mu) \mid g(x) = f(x) \ \mu\text{-presque partout}\}$$

Définition 6.7. Soit $p \in [1, +\infty]$. On note $L^p(X, \mathcal{M}, \mu)$ le quotient de $\mathcal{L}^p(X, \mathcal{M}, \mu)$ par la relation d'équivalence $\mathcal{R}$, i.e.

$$L^p(X, \mathcal{M}, \mu) = \{[f] \mid f \in \mathcal{L}^p(X, \mathcal{M}, \mu)\}$$

Par construction, $L^p(X, \mathcal{M}, \mu)$ est un espace vectoriel normé, avec les opérations d'espace vectoriel

$$[f] + [g] = [f + g], \lambda[f] = [\lambda f], \qquad \forall \lambda \in \mathbb{K}, \forall f, g \in \mathcal{L}^p(X, \mathcal{M}, \mu)$$

et la norme [2] :

$$\|[f]\|_p = \|f\|_p, \qquad \forall f \in \mathcal{L}^p(X, \mathcal{M}, \mu)$$

Attention ! En fait, lorsque l'on travaille avec des éléments de $L^p(X, \mathcal{M}, \mu)$, on raisonne avec des classes d'équivalence. Mais dans toute la suite du cours, on identifiera la classe d'équivalence avec l'une de ses fonctions représentatives. Cependant, **on ne pourra pas évaluer une fonction de $L^p(X, \mathcal{M}, \mu)$ en un point**. En effet, deux fonctions dans la même classe d'équivalence peuvent différer sur un ensemble de points de mesure nulle.

L'analyse étant basée pour une grande part sur des procédés de limite et d'approximation, on n'étudie d'ordinaire les espaces vectoriels normés que s'ils sont complets, i.e. toute suite de Cauchy converge. Le théorème suivant assure la complétude des espaces de Lebesgue.

Théorème 6.8 (Riesz-Fisher). *L'espace de Lebesgue $L^p(X, \mathcal{M}, \mu)$ est un espace vectoriel normé complet (i.e. un espace de Banach) pour tout $p \in [1, +\infty]$.*

Démonstration. Soit $(f_n)_{n \in \mathbb{N}}$ une suite de Cauchy dans $L^p(X, \mathcal{M}, \mu)$. On choisit des représentants quelconques $f_n \in \mathcal{L}^p(X, \mathcal{M}, \mu)$, et on veut montrer qu'il existe une fonction $f \in \mathcal{L}^p(X, \mathcal{M}, \mu)$ telle que $\|f - f_n\|_p \underset{n \to \infty}{\longrightarrow} 0$.

1. ... ce qui est, en particulier, le cas de la tribu de Lebesgue.
2. Il faut bien entendu vérifier que les lois de composition et que la norme ne dépendent pas du représentant choisi.

a) Si $1 \leqslant p < \infty$. 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 $k \geqslant 0$, $\|f_{n_{k+1}} - f_{n_k}\|_p < 2^{-k}$. Le problème est donc de construire une limite à cette suite. On note pour tout $N \in \mathbb{N}$,

$$g_N(x) = \sum_{k=0}^{N} \left|f_{n_{k+1}}(x) - f_{n_k}(x)\right|, \qquad x \in X$$

Alors, $g_N \in \mathcal{L}^p(X, \mathcal{M}, \mu)$, et par l'inégalité de Minkowski :

$$\|g_N\|_p \leqslant \sum_{k=0}^{N} \|f_{n_{k+1}} - f_{n_k}\|_p \leqslant \sum_{k=0}^{N} 2^{-k} < 2, \quad \forall N \in \mathbb{N}$$

Ainsi, la suite $(g_N^p)_{N\in\mathbb{N}}$ est une suite croissante de fonctions intégrables, positives et telles que $\sup_{N\in\mathbb{N}} \int g_N^p \mathrm{d}\mu \leqslant 2^p < \infty$. Par le théorème 2.16 de la convergence monotone, $\sup_{N\in\mathbb{N}} g_N(x) < \infty$ pour presque tout $x \in X$. Plus précisément, il existe $E \in \mathcal{M}$ tel que $\mu(E) = 0$ et $\sum_{k=0}^{\infty} \left|f_{n_{k+1}}(x) - f_{n_k}(x)\right| < \infty$ pour tout $x \in X \setminus E$.
En outre, si $g\colon X \to \mathbb{R}_+$ est définie par :

$$g(x) = \begin{cases} \sum_{k=0}^{\infty} \left|f_{n_{k+1}}(x) - f_{n_k}(x)\right| & \text{si } x \in X \setminus E \\ 0 & \text{si } x \in E \end{cases}$$

alors $g \in \mathcal{L}^p(X, \mathcal{M}, \mu)$ et $\|g\|_p \leqslant 2$.
D'autre part, on a évidemment :

$$f_{n_k}(x) = f_{n_0}(x) + \sum_{i=0}^{k-1} \left(f_{n_{i+1}}(x) - f_{n_i}(x)\right), \qquad \forall x \in X$$

On définit

$$f(x) = \begin{cases} \lim_{k\to\infty} f_{n_k}(x) & \text{si } x \in X \setminus E \\ 0 & \text{si } x \in E \end{cases}$$

Comme $f_{n_k}(x) \underset{k\to\infty}{\longrightarrow} f(x)$ presque partout, et que $|f_{n_k}| \leqslant |f_{n_0}| + g \in \mathcal{L}^p(X, \mathcal{M}, \mu)$, le théorème de la convergence dominée 2.25 montre que $f \in \mathcal{L}^p(X, \mathcal{M}, \mu)$. Enfin, $|f_{n_k}(x) - f(x)| \underset{k\to\infty}{\longrightarrow} 0$ presque partout, et $|f_{n_k} - f| \leqslant |f_{n_0} - f| + g \in \mathcal{L}^p(X, \mathcal{M}, \mu)$, donc par le même argument, $\|f_{n_k} - f\|_p \underset{k\to\infty}{\longrightarrow} 0$.
Enfin, comme $(f_n)_{n\in\mathbb{N}}$ est une suite de Cauchy, on conclut que $\|f_n - f\|_p \underset{n\to\infty}{\longrightarrow} 0$.

Remarque. L'argument montre que toute suite de Cauchy dans $\mathcal{L}^p(X, \mathcal{M}, \mu)$ possède une sous-suite qui converge point par point presque partout.

b) Si $p = \infty$. Pour $n, m \in \mathbb{N}$, on définit :

$$A_n = \{x \in X \mid |f_n(x)| > \|f_n\|_\infty\}, \qquad \mu(A_n) = 0$$

$$B_{n,m} = \{x \in X \mid |f_n(x) - f_m(x)| > \|f_n - f_m\|_\infty\}, \qquad \mu(B_{n,m}) = 0$$

$$E = \left(\bigcup_{n\in\mathbb{N}} A_n\right) \cup \left(\bigcup_{\substack{n\in\mathbb{N}\\ m\in\mathbb{N}}} B_{n,m}\right), \qquad \mu(E) = 0$$

Si $x \notin E$, on a

$$|f_n(x) - f_m(x)| \leqslant \|f_n - f_m\|_\infty \underset{n,m\to\infty}{\longrightarrow} 0$$

On définit donc

$$f(x) = \begin{cases} \lim_{n\to\infty} f_n(x) & \text{si } x \notin E \\ 0 & \text{sinon} \end{cases}$$

Alors f est mesurable, et $|f(x)| \leqslant \liminf_{n\to\infty} \|f_n\|_\infty$ pour tout $x \in X$, donc $f \in \mathcal{L}^\infty(X, \mathcal{M}, \mu)$ et $\|f\|_\infty \leqslant \liminf_{n\to\infty} \|f_n\|_\infty < \infty$. De même,

$$|f_n(x) - f(x)| \leqslant \liminf_{m\to\infty} \|f_n - f_m\|_\infty, \qquad \forall x \in X \setminus E$$

□

donc $\|f_n - f\|_\infty \leqslant \liminf_{m\to\infty} \|f_n - f_m\|_\infty \underset{n\to\infty}{\longrightarrow} 0$.

Corollaire 6.9 ($p = 2$). *L'espace $L^2(X, \mathcal{M}, \mu)$ est un espace de Hilbert, muni du produit scalaire :*

$$(f|g) = \int_X f\overline{g}\mathrm{d}\mu, \qquad f, g \in L^2(X, \mathcal{M}, \mu)$$

On vérifie que $(\cdot|\cdot)$ est bien un produit scalaire[3] (i.e. une forme bilinéaire symétrique définie positive sur $\mathbb{R}$, ou une forme sesquilinéaire symétrique hermitienne définie positive sur $\mathbb{C}$) sur $L^2(X, \mathcal{M}, \mu)$. En particulier, l'inégalité de Hölder (dans le cas $p = 2$) donne l'**inégalité de Cauchy-Schwarz** :

$$|(f|g)| \leqslant \|f\|_2\|g\|_2$$

Pour $p \neq 2$, aucun $L^p(X, \mathcal{M}, \mu)$ n'est un espace de Hilbert.

6.2 Inclusions des espaces $\mathcal{L}^p$ ou L^p

Il est souvent crucial de garder en tête les relations d'inclusion entre espaces de Lebesgue. Si l'espace X est de mesure finie, alors les espaces de Lebesgue $\mathcal{L}^p(X, \mathcal{M}, \mu)$ sont emboîtés :

Proposition 6.10. *On suppose que $\mu(X) < \infty$. Si $f \in \mathcal{L}^p(X, \mathcal{M}, \mu)$, alors pour tout $q \in [1, p]$, $f \in \mathcal{L}^q(X, \mathcal{M}, \mu)$ et on a :*

$$\|f\|_q \leqslant \|f\|_p \mu(X)^{\frac{1}{q}-\frac{1}{p}} \tag{6.5}$$

En outre, si $f \in \mathcal{L}^\infty(X, \mathcal{M}, \mu)$, alors

$$\|f\|_\infty = \lim_{p\to\infty} \|f\|_p \tag{6.6}$$

Dans ce cas, on a l'emboîtement suivant :

$$\mathcal{L}^\infty(X, \mathcal{M}, \mu) \subset \mathcal{L}^p(X, \mathcal{M}, \mu) \subset \mathcal{L}^q(X, \mathcal{M}, \mu) \subset \mathcal{L}^1(X, \mathcal{M}, \mu), \qquad p \geqslant q$$

Démonstration. Dans (6.5), on peut supposer que $q < p$.

3. qui ne dépend pas du représentant choisi, car l'intégrale ne change pas de valeur pour deux fonctions différentes sur un ensemble de mesure nulle.

1) Si $p = \infty$, on a $|f(x)| \leqslant \|f\|_\infty$ presque partout, donc pour tout $q < \infty$,

$$\int_X |f|^q \, \mathrm{d}\mu \leqslant \|f\|_\infty^q \mu(X)$$

En mettant à la puissance $\frac{1}{q}$, on a l'inégalité voulue.

2) Si $p < \infty$ et $1 \leqslant q < p$, alors on a

$$\int_X |f|^q \, \mathrm{d}\mu = \int_X |f|^q \cdot 1 \mathrm{d}\mu \leqslant \left(\int_X |f|^p \, \mathrm{d}\mu \right)^{\frac{q}{p}} \mu(X)^{1-\frac{q}{p}}$$

En mettant à la puissance $\frac{1}{q}$, on a l'inégalité voulue.

3) Pour montrer (6.6), on peut supposer que $\|f\|_\infty > 0$. Soit $\varepsilon \in]0,1[$. Il existe $E \in \mathcal{M}$ avec $\mu(E) > 0$ tel que

$$|f(x)| \geqslant \|f\|_\infty (1-\varepsilon), \qquad \forall x \in E$$

Soit $q \geqslant 1$ assez grand pour que $\mu(E)^{\frac{1}{q}} \geqslant 1-\varepsilon$ et que $\mu(X)^{\frac{1}{q}} \leqslant 1+\varepsilon$. Si $p \geqslant q$, on a donc :

$$\|f\|_p \leqslant \|f\|_\infty \mu(X)^{\frac{1}{p}} \leqslant \|f\|_\infty (1+\varepsilon)$$

D'autre part,

$$\|f\|_p^p = \int_X |f|^p \, \mathrm{d}\mu \geqslant \int_E \|f\|_\infty^p (1-\varepsilon)^p \mathrm{d}\mu = \|f\|_\infty^p (1-\varepsilon)^p \mu(E)$$

Donc

$$\|f\|_p \geqslant \|f\|_\infty (1-\varepsilon) \mu(E)^{\frac{1}{p}} \geqslant \|f\|_\infty (1-\varepsilon)^2$$

Ainsi,

$$(1-\varepsilon)^2 \leqslant \frac{\|f\|_p}{\|f\|_\infty} \leqslant 1+\varepsilon, \qquad p \geqslant q$$

□

Dans le cas général, les espaces de Lebesgue ne sont pas emboîtés, et il n'y a pas de règle générale. On peut d'ailleurs trouver des situations où l'emboîtement a lieu, mais dans le sens opposé à celui que nous venons de décrire.

Proposition 6.11. *On suppose que $\mu_0(X) > 0$, où*

$$\mu_0(X) = \inf\{\mu(E) \mid E \in \mathcal{M}, \mu(E) > 0\}$$

Soient $p, q \in [1, +\infty]$ et $p \geqslant q$. Alors, si $f \in \mathcal{L}^q(X, \mathcal{M}, \mu)$, on a $f \in \mathcal{L}^p(X, \mathcal{M}, \mu)$ et

$$\|f\|_p \leqslant \|f\|_q \left(\frac{1}{\mu_0(X)} \right)^{\frac{1}{q}-\frac{1}{p}} \tag{6.7}$$

Dans ce cas, on a l'emboîtement suivant :

$$\mathcal{L}^1(X, \mathcal{M}, \mu) \subset \mathcal{L}^q(X, \mathcal{M}, \mu) \subset \mathcal{L}^p(X, \mathcal{M}, \mu) \subset \mathcal{L}^\infty(X, \mathcal{M}, \mu), \qquad p \geqslant q$$

Démonstration. Soit $f \in \mathcal{L}^q(X, \mathcal{M}, \mu)$ pour $q < \infty$. S'il existe $A > 0$ et $E \in \mathcal{M}$ avec $\mu(E) > 0$ tel que $|f(x)| \geqslant A$ pour tout $x \in E$, alors :

$$\int_X |f|^q \, \mathrm{d}\mu \geqslant \int_E A^q \mathrm{d}\mu = A^q \mu(E) \geqslant A^q \mu_0(X)$$

Donc

$$A \leqslant \|f\|_q \left(\frac{1}{\mu_0(X)}\right)^{\frac{1}{q}}$$

Ainsi, $f \in \mathcal{L}^\infty(X, \mathcal{M}, \mu)$ et $\|f\|_\infty \leqslant \|f\|_q \left(\frac{1}{\mu_0(X)}\right)^{\frac{1}{q}}$.
Si $1 \leqslant q < p < \infty$, alors on a :

$$\int_X |f|^p \, \mathrm{d}\mu = \int_X |f|^{p-q} \, |f|^q \, \mathrm{d}\mu \leqslant \|f\|_\infty^{p-q} \|f\|_q^q \leqslant \|f\|_q^p \left(\frac{1}{\mu_0(X)}\right)^{\frac{p-q}{q}}$$

Donc $\|f\|_p \leqslant \|f\|_q \left(\frac{1}{\mu_0(X)}\right)^{\frac{1}{q}-\frac{1}{p}}$. □

Exemple. On pourrait se demander si l'hypothèse de la proposition ci-dessus est vérifiée... Si on considère $X = \mathbb{N}$, $\mathcal{M} = \mathcal{P}(\mathbb{N})$ et μ la mesure de comptage, on note pour $p \in [1, +\infty]$

$$\ell^p(\mathbb{N}) = \mathcal{L}^p(X, \mathcal{M}, \mu) =^{4} L^p(X, \mathcal{M}, \mu)$$

Autrement dit,

$$\ell^p(\mathbb{N}) = \{(a_n)_{n \in \mathbb{N}} \mid \|a\|_p < \infty\}$$

avec

$$a = (a_n)_{n \in \mathbb{N}}, \|a\|_p = \left(\sum_{n=0}^{\infty} |a_n|^p\right)^{\frac{1}{p}} \text{ si } p < \infty \text{, et } \|a\|_\infty = \sup_{n \in \mathbb{N}} |a_n|$$

On a dans ce cas là $\mu_0(X) = 1$, donc on a l'emboîtement suivant :

$$\ell^1(\mathbb{N}) \subset \ell^q(\mathbb{N}) \subset \ell^p(\mathbb{N}) \subset \ell^\infty(\mathbb{N}), \qquad p \geqslant q$$

En outre, l'application $p \mapsto \|a\|_p$ est décroissante.

Proposition 6.12. *Soient $p, q \in [1, +\infty]$ avec $q < p$. Si $f \in \mathcal{L}^p(X, \mathcal{M}, \mu) \cap \mathcal{L}^q(X, \mathcal{M}, \mu)$, alors $f \in \mathcal{L}^r(X, \mathcal{M}, \mu)$ pour tout $r \in [q, p]$, et on a :*

$$\|f\|_r \leqslant \|f\|_q^\alpha \, \|f\|_p^{1-\alpha}$$

où $\alpha \in [0, 1]$ est défini par $\dfrac{1}{r} = \dfrac{\alpha}{q} + \dfrac{1-\alpha}{p}$.

Corollaire 6.13.

$$\bigcap_{1 \leqslant p \leqslant \infty} \mathcal{L}^p(X, \mathcal{M}, \mu) = \mathcal{L}^1(X, \mathcal{M}, \mu) \cap \mathcal{L}^\infty(X, \mathcal{M}, \mu)$$

4. Dans la tribu, il n'y a pas d'ensemble de mesure nulle non trivial, du coup en quotientant par la relation d'équivalence $\mathcal{R}$, on obtient un espace isomorphe.

Démonstration. Il n'y a rien à montrer si $r = q$ (i.e. $\alpha = 1$), ou $r = p$ (i.e. $\alpha = 0$). Supposons donc que $q < r < p$. On écrit

$$|f|^r = |f|^{r\alpha} |f|^{r(1-\alpha)}$$

et on applique l'inégalité de Hölder avec les exposants conjugués $\frac{q}{r\alpha}$ et $\frac{p}{r(1-\alpha)}$. Si $p < \infty$, on trouve :

$$\|f\|_r^r = \int_X |f|^r \,\mathrm{d}\mu \leqslant \left(\int_X |f|^q \,\mathrm{d}\mu\right)^{\frac{r\alpha}{q}} \left(\int_X |f|^p \,\mathrm{d}\mu\right)^{\frac{r(1-\alpha)}{p}} = \|f\|_q^{r\alpha} \|f\|_p^{r(1-\alpha)}$$

□

Exemple (important). Soit $X = \mathbb{R}^d$, $\mathcal{M} = \mathcal{B}(\mathbb{R}^d)$ (la tribu de Borel) ou $\mathcal{L}(\mathbb{R}^d)$ (la tribu de Lebesgue) et $\mu = \lambda$ (la mesure de Lebesgue). Si $p \in [1, +\infty]$, on note

$$L^p(\mathbb{R}^d) = L^p(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda) \equiv L^p(\mathbb{R}^d, \mathcal{L}(\mathbb{R}^d), \lambda)$$

De plus, si $p, q, r \in [1, +\infty[$ avec $q < r < p$, alors

- si $f(x) = \dfrac{1}{(1 + \|x\|)^{d/r}}$, alors $f \in \mathcal{L}^p(\mathbb{R}^d)$ et $f \notin \mathcal{L}^q(\mathbb{R}^d)$. En effet, la puissance q-ième de la fonction ne sera pas intégrable à l'infini, d'après la section 5.2.3, page 76.
- si $g(x) = \dfrac{1}{\|x\|^{d/r}} e^{-\|x\|}$, alors $g \in \mathcal{L}^q(\mathbb{R}^d)$ et $g \notin \mathcal{L}^p(\mathbb{R}^d)$. En effet, la puissance p-ième de la fonction ne sera pas intégrable en 0, d'après la section 5.2.3, page 76.

Cet exemple montre qu'il n'y a pas d'inclusion entre $\mathcal{L}^p(\mathbb{R}^d)$ et $\mathcal{L}^q(\mathbb{R}^d)$ si $p \neq q$.

Exemple (fonctions indicatrices). Soit $A \in \mathcal{M}$. On montre facilement que

$$\mathbb{1}_A \in \mathcal{L}^1(X, \mathcal{M}, \mu) \iff \mathbb{1}_A \in \mathcal{L}^p(X, \mathcal{M}, \mu), \forall p \in [1, +\infty[$$

6.3 Théorèmes de densité

Théorème 6.14. *Soit (X, d) un espace métrique localement compact et séparable, et μ une mesure borélienne sur X, finie sur les compacts (i.e. une mesure de Radon). Alors, l'espace des fonctions lipschitziennes sur X à support compact est dense dans $L^p(X, \mathcal{B}(X), \mu)$ si $1 \leqslant p < \infty$.*

Remarques. 1) Il n'y a pas densité si $p = \infty$. Par exemple, l'adhérence dans $L^\infty(\mathbb{R}^d)$ de l'espace des fonctions continues à support compact est l'espace des fonctions continues qui s'annulent à l'infini.

2) **Rappel :** Tout espace métrique localement compact séparable est σ-compact : on peut écrire $X = \bigcup_{n\in\mathbb{N}} E_n$, où E_n est ouvert, $\overline{E_n} \subset E_{n+1}$ et $\overline{E_n}$ compact pour tout $n \in \mathbb{N}$.

3) **Rappel :** Toute mesure de Radon sur un tel espace est extérieurement régulière, i.e :

$$\mu(A) = \inf \{\mu(U) \mid U \text{ ouvert}, U \supset A\}$$

Démonstration. Soient $p \in [1, +\infty[$, $f \in \mathcal{L}^p(X, \mathcal{B}(X), \mu)$, et $\varepsilon > 0$. On cherche une fonction g lipschitzienne à support compact telle que $\|f - g\|_p \leqslant \varepsilon$.

Étape 1 : Il suffit de montrer le résultat pour $f = \mathbb{1}_A$ avec $A \in \mathcal{B}(X)$ et $\mu(A) < \infty$. En effet, supposons le résultat vrai dans ce cas, et soit $f \in \mathcal{L}^p(X, \mathcal{B}(X), \mu)$ positive. Il existe une suite croissante $(f_n)_{n\in\mathbb{N}}$ de fonctions étagées positives telle que $f_n(x) \underset{n\to\infty}{\longrightarrow} f(x)$ pour tout $x \in X$. Comme $|f - f_n|^p \leqslant$

$f^p \in \mathcal{L}^p(X, \mathcal{B}(x), \mu)$, le théorème de la convergence dominée 2.25 montre[5] que $\|f_n - f\|_p \underset{n\to\infty}{\longrightarrow} 0$. On choisit n assez grand pour que $\|f - f_n\|_p \leqslant \varepsilon/2$. On peut écrire $f_n = \sum_{j=1}^m \alpha_j \mathbb{1}_{A_j}$, où $\alpha_j > 0$, $A_j \in \mathcal{B}(X)$ tels que $\mu(A_j) < \infty$. Par hypothèse, pour tout $j \in \{1, \dots, m\}$, il existe une fonction φ_j lipschitzienne à support compact telle que

$$\left\|\mathbb{1}_{A_j} - \varphi_j\right\|_p \leqslant \frac{\varepsilon}{2m\alpha_j}$$

Notons $\varphi = \sum_{j=1}^m \alpha_j \varphi_j$. Alors φ est lipschitzienne à support compact, et

$$\|f_n - \varphi\|_p = \left\|\sum_{j=1}^m \alpha_j(\mathbb{1}_{A_j} - \varphi_j)\right\|_p \leqslant \sum_{j=1}^m \alpha_j \left\|\mathbb{1}_{A_j} - \varphi_j\right\|_p \leqslant \frac{\varepsilon}{2}$$

Ainsi, $\|f - \varphi\|_p \leqslant \|f - f_n\|_p + \|f_n - \varphi\|_p \leqslant \varepsilon$. Une fois le résultat établi pour f positive, on l'étend à f réelle (écrire $f = f_+ - f_-$), puis à f complexe (écrire $f = \Re(f) + i\Im(f)$).

Étape 2 : Soient $f = \mathbb{1}_A$, avec $A \in \mathcal{B}(X)$ et $\mu(A) < \infty$, et $\varepsilon > 0$. Par régularité extérieure, il existe un ouvert U de X contenant A tel que $\mu(U \setminus A) \leqslant \left(\frac{\varepsilon}{3}\right)^p$. Alors,

$$\|\mathbb{1}_A - \mathbb{1}_U\|_p = \left\|\mathbb{1}_{U\setminus A}\right\|_p = \mu(U \setminus A)^{1/p} \leqslant \frac{\varepsilon}{3}$$

Par σ-compacité, on peut écrire $U = \bigcup_{n\in\mathbb{N}} U_n$, avec pour tout $n \in \mathbb{N}$, $U_n = U \cap E_n$ ouvert d'adhérence compacte. Ainsi, $\mathbb{1}_U(x) = \lim_{n\to\infty} \mathbb{1}_{U_n}(x)$ pour tout $x \in X$, et en raisonnant comme avant, $\|\mathbb{1}_U - \mathbb{1}_{U_n}\|_p \underset{n\to\infty}{\longrightarrow} 0$. On choisit n assez grand pour que $\|\mathbb{1}_U - \mathbb{1}_{U_n}\|_p \leqslant \varepsilon/3$, et on pose $V = U_n$. Enfin, pour $k \in \mathbb{N}^*$, on définit $\varphi_k(x) = \min\left(1, kd(x, V^\complement)\right)$. Alors φ_k est lipschitzienne, $supp(\varphi_k) \subset V$ est compact, et $\varphi_k(x) \underset{k\to\infty}{\longrightarrow} \mathbb{1}_V(x)$ pour tout $x \in X$. En particulier, $\|\mathbb{1}_V - \varphi_k\|_p \underset{k\to\infty}{\longrightarrow} 0$, donc on peut choisir $k \in \mathbb{N}^*$ assez grand pour que $\|\mathbb{1}_V - \varphi_k\|_p \leqslant \varepsilon/3$.
Conclusion : $\|\mathbb{1}_A - \varphi_k\|_p \leqslant \|\mathbb{1}_A - \mathbb{1}_U\|_p + \|\mathbb{1}_U - \mathbb{1}_V\|_p + \|\mathbb{1}_V - \varphi_k\|_p \leqslant \varepsilon$. □

Proposition 6.15. *Les fonctions en escaliers sont denses dans $L^p(\mathbb{R}^d)$ si $1 \leqslant p < \infty$.*

Démonstration. On modifie la dernière étape de la démonstration ci-dessus. Si $V \subset \mathbb{R}^d$ est un ouvert borné, il existe un pavage $P \subset V$ tel que $\mu(V \setminus P) \leqslant (\varepsilon/3)^p$, avec ε arbitraire. Ainsi, $\mathbb{1}_P$ est en escalier, et $\|\mathbb{1}_V - \mathbb{1}_P\|_p \leqslant \varepsilon/3$. □

En particulier, on montre que le complété de l'ensemble des fonctions en escaliers muni de la norme $\|\cdot\|_1$ est $L^1(\mathbb{R}^d)$, ce qui, si besoin est encore, justifie la construction de l'intégrale de Lebesgue.

Application : continuité des translations

Si $f : \mathbb{R}^d \to \mathbb{K}$, on note pour tout $y \in \mathbb{R}^d$:

$$(\tau_y f)(x) = f(x - y), \qquad x \in \mathbb{R}^d$$

$\tau_y f$ est le translaté de f par $y \in \mathbb{R}^d$.

5. Ici, on utilise le fait que $p < \infty$!

Proposition 6.16. *La translation est continue dans $L^p(\mathbb{R}^d)$ si $1 \leqslant p < \infty$. Si $f \in \mathcal{L}^p(\mathbb{R}^d)$, $\|\tau_y f - f\|_p \underset{y\to 0}{\longrightarrow} 0$.*

Ici encore, ce résultat n'est pas valable dans $L^\infty(\mathbb{R}^d)$! Un exemple simple est la fonction $H \colon \mathbb{R} \to \mathbb{R}$, dite de Heaviside, définie par

$$H(x) = \begin{cases} 1 & \text{si } x \geqslant 0 \\ 0 & \text{si } x < 0 \end{cases}$$

En effet, pour tout $y \neq 0$, on a $\|\tau_y H - H\|_\infty = 1$.

Démonstration. On commence par montrer la continuité en 0 de $y \mapsto \tau_y g$ pour une fonction g continue à support compact. Soit $M = \sup |g|$ et $R > 0$ assez grand pour que $supp(g) \subset \left\{x \in \mathbb{R}^d \mid \|x\| \leqslant R\right\}$. Ainsi, si $y \in \mathbb{R}^d$ vérifie $\|y\| \leqslant 1$, alors $|\tau_y g - g| \leqslant 2M\mathbb{1}_{\{x \mid \|x\| \leqslant R+1\}} \in \mathcal{L}^p(\mathbb{R}^d)$. Par le théorème de convergence dominée 2.25,

$$\int_{\mathbb{R}^d} |\tau_y g - g|^p \,\mathrm{d}\lambda = \int_{\mathbb{R}^d} |g(x-y) - g(x)|^p \,\mathrm{d}x \underset{y\to 0}{\longrightarrow} 0$$

donc $\|\tau_y g - g\|_p \underset{y\to 0}{\longrightarrow} 0$.

Soit $f \in \mathcal{L}^p(\mathbb{R}^d)$, et $\varepsilon > 0$. Par le théorème de densité qui précède, il existe une fonction g continue à support compact telle que $\|f - g\|_p \leqslant \varepsilon/3$. Par l'argument ci-dessus, il existe $\delta > 0$ tel que si $\|y\| \leqslant \delta$, alors on a $\|\tau_y g - g\|_p \leqslant \varepsilon/3$. Ainsi,

$$\begin{aligned} \|f - \tau_y f\|_p &\leqslant \|f - g\|_p + \|g - \tau_y g\|_p + \|\tau_y g - \tau_y f\|_p \\ &= 2\,\|f - g\|_p + \|g - \tau_y g\|_p \leqslant \varepsilon \end{aligned}$$

car $\|\tau_y g - \tau_y f\|_p = \|g - f\|_p$ du fait de l'invariance de λ par translation.

Finalement, si $y_0 \in \mathbb{R}^d$, on a :

$$\|\tau_{y_0+y} f - \tau_{y_0} f\|_p = \|\tau_y f - f\|_p \underset{y\to 0}{\longrightarrow} 0$$

□

6.4 LE PRODUIT DE CONVOLUTION

6.4.1 Définition et propriétés

Soient $f, g \colon \mathbb{R}^d \to \mathbb{K}$ deux fonctions mesurables. Le produit de convolution de f et g est défini formellement par :

$$(f * g)(x) = \int_{\mathbb{R}^d} f(x-y) g(y) \mathrm{d}y$$

Ce produit est défini pour tout $x \in \mathbb{R}^d$ si f et g sont positives, et sinon au moins pour les $x \in \mathbb{R}^d$ tels que $\int_{\mathbb{R}^d} |f(x-y)|\,|g(y)|\,\mathrm{d}y < \infty$. Dans tous les cas, on a :

$$(f * g)(x) = \int_{\mathbb{R}^d} f(x-y) g(y) \mathrm{d}y = \int_{\mathbb{R}^d} f(z) g(x-z) \mathrm{d}z = (g * f)(x)$$

ce qui signifie que le produit de convolution, quand il est défini, est commutatif.

Théorème 6.17. *Soient $p, q, r \in [1, +\infty]$ tels que $\dfrac{1}{p} + \dfrac{1}{q} = 1 + \dfrac{1}{r}$. Si $f \in \mathcal{L}^p(\mathbb{R}^d)$ et $g \in \mathcal{L}^q(\mathbb{R}^d)$, alors le produit de convolution*

$$h(x) = \int_{\mathbb{R}^d} f(x-y)g(y)\mathrm{d}y$$

est défini pour presque tout $x \in \mathbb{R}^d$, $h \in \mathcal{L}^r(\mathbb{R}^d)$, et $\|h\|_r \leqslant \|f\|_p \|g\|_q$ (inégalité de Young). En outre,
- *si $r = \infty$, h est uniformément continue.*
- *si $r = \infty$ et $1 < p, q < \infty$, alors $h(x) \underset{|x|\to\infty}{\longrightarrow} 0$.*

Démonstration. **Etape 1 :** On commence par considérer le cas où $r = \infty$, i.e. p et q sont des exposants conjugués. On peut supposer que $1 \leqslant p < \infty$. Par l'inégalité de Hölder, $h(x)$ est défini pour tout $x \in \mathbb{R}^d$, et $|h(x)| \leqslant \|f\|_p \|g\|_q$. En outre, pour tout $z \in \mathbb{R}^d$, on a :

$$h(x) - h(x-z) = \int_{\mathbb{R}^d} \left(f(x-y) - f(x-y-z)\right) g(y)\mathrm{d}y$$

donc,

$$|h(x) - h(x-z)| \leqslant \|f - \tau_z f\|_p \|g\|_q , \qquad \forall x \in \mathbb{R}^d$$

d'où

$$\sup_{x\in\mathbb{R}^d} |h(x) - h(x-z)| \leqslant \|f - \tau_z f\|_p \|g\|_q \underset{z\to 0}{\longrightarrow} 0$$

Ainsi, h est uniformément continue. En particulier, $h \in \mathcal{L}^\infty(\mathbb{R}^d)$ et $\|h\|_\infty \leqslant \|f\|_p \|g\|_q$. Si $1 < p, q < \infty$, alors il existe des suites $(f_n)_{n\in\mathbb{N}}$, $(g_n)_{n\in\mathbb{N}}$ de fonctions continues à support compact telles que :

$$\|f - f_n\|_p \underset{n\to\infty}{\longrightarrow} 0, \quad \|g - g_n\|_q \underset{n\to\infty}{\longrightarrow} 0.$$

Pour tout $n \in \mathbb{N}$, notons $h_n = f_n * g_n$. h_n est une fonction continue à support compact. On a alors

$$h - h_n = f * g - f_n * g_n = (f - f_n) * g + f_n * (g - g_n)$$

Donc

$$\|h - h_n\|_\infty \leqslant \|f - f_n\|_p \|g\|_q + \underbrace{\|f_n\|_p}_{\text{borné}} \|g - g_n\|_q \underset{n\to\infty}{\longrightarrow} 0$$

Ainsi, $h(x) \underset{|x|\to\infty}{\longrightarrow} 0$.

Etape 2 : Traitons maintenant le cas où $r = q < \infty$ et $p = 1$. On suppose d'abord que $f, g \geqslant 0$. Alors, $h \colon \mathbb{R}^d \to [0, +\infty]$ est bien définie. Il faut vérifier que h est mesurable. Soient alors $(f_n)_{n\in\mathbb{N}}$ et $(g_n)_{n\in\mathbb{N}}$ deux suites croissantes de fonctions étagées positives telles que $f_n \underset{n\to\infty}{\longrightarrow} f$ et $g_n \underset{n\to\infty}{\longrightarrow} g$ point par point. Notons $h_n = f_n * g_n$. Alors, par le théorème 2.16 de la convergence monotone (page 16),

$$h_n(x) = \int_{\mathbb{R}^d} f_n(x-y)g_n(y)\mathrm{d}y \underset{n\to\infty}{\longrightarrow} \int_{\mathbb{R}^d} f(x-y)g(y)\mathrm{d}y = h(x)$$

Or $f_n, g_n \in \mathcal{L}^1(\mathbb{R}^d) \cap \mathcal{L}^\infty(\mathbb{R}^d)$ [6]. Ainsi, comme $h_n = f_n * g_n$ est continue pour tout $n \in \mathbb{N}$, h est mesurable. Il reste à montrer que $h(x) < \infty$ pour presque tout $x \in \mathbb{R}^d$ et que $h \in \mathcal{L}^q(\mathbb{R}^d)$. Si $q > 1$, on a :

$$h(x) = \int_{\mathbb{R}^d} f(x-y)^{1-\frac{1}{q}} f(x-y)^{\frac{1}{q}} g(y)\mathrm{d}y$$

6. On remarque que $\mathcal{L}^\infty(\mathbb{R}^d)$ contient toutes les fonctions étagées.

En appliquant l'inégalité de Hölder avec les exposants conjugués $\dfrac{q}{q-1}$ et q, on trouve

$$h(x) \leqslant \left(\int_{\mathbb{R}^d} f \mathrm{d}\lambda\right)^{1-\frac{1}{q}} \left(\int_{\mathbb{R}^d} f(x-y)g(y)^q \mathrm{d}y\right)^{\frac{1}{q}}$$

En mettant à la puissance q, on obtient

$$h(x)^q \leqslant \|f\|_1^{q-1} \left(\int_{\mathbb{R}^d} f(x-y)g(y)^q \mathrm{d}y\right)$$

Et on remarque que cette relation est également vraie si $q = 1$. En intégrant selon x et en utilisant le théorème 4.6 de Fubini-Tonelli (page 59), on trouve

$$\|h\|_q^q \leqslant \|f\|_1^{q-1} \|f\|_1 \|g\|_q^q$$

D'où $\|h\|_q \leqslant \|f\|_1 \|g\|_q$.

Finalement, si f et g sont réelles ou complexes, on vérifie que h est définie presque partout et mesurable en décomposant $f = f_+ - f_-$ (si f est réelle) ou $f = \Re(f) + i\Im(f)$ (si f complexe), et on fait de même pour g. De plus, comme

$$|h(x)| \leqslant \int_{\mathbb{R}^d} |f(x-y)|\,|g(y)|\,\mathrm{d}y$$

on conclut que $\|h\|_q \leqslant \|f\|_1 \|g\|_q$.

Etape 3 : Reste à traiter le cas où $\max(p,q) < r < \infty$. Il suffit de montrer l'inégalité pour $f, g \geqslant 0$. On a :

$$h(x) = \int_{\mathbb{R}^d} \underbrace{\left(f(x-y)^{p/r} g(y)^{q/r}\right)}_{A} \underbrace{\left(f(x-y)^{1-p/r}\right)}_{B} \underbrace{\left(g(y)^{1-q/r}\right)}_{C} \mathrm{d}y$$

On applique l'inégalité de Hölder généralisée avec les exposants conjugués r, $\alpha = \dfrac{rp}{r-p}$, $\beta = \dfrac{rq}{r-q}$ (i.e. $\dfrac{1}{r} + \dfrac{1}{\alpha} + \dfrac{1}{\beta} = 1$) appliqués respectivement à A, B et C.

$$h(x) \leqslant \left(\int_{\mathbb{R}^d} f(x-y)^p g(y)^q \mathrm{d}y\right)^{1/r} \left(\int_{\mathbb{R}^d} f^p \mathrm{d}\lambda\right)^{1/\alpha} \left(\int_{\mathbb{R}^d} g^q \mathrm{d}\lambda\right)^{1/\beta}$$

d'où

$$|h(x)|^r \leqslant \left(\int_{\mathbb{R}^d} f(x-y)^p g(y)^q \mathrm{d}y\right) \|f\|_p^{r-p} \|g\|_q^{r-q}$$

En intégrant selon x et en utilisant le théorème 4.6 de Fubini-Tonelli (page 59), on trouve :

$$\|h\|_r^r \leqslant \|f\|_p^p \|g\|_q^q \|f\|_p^{r-p} \|g\|_q^{r-q} = \|f\|_p^r \|g\|_q^r$$

□

6.4.2 Régularisation par convolution

Soit $\chi \colon \mathbb{R}^d \to \mathbb{R}_+$ telle que $\int_{\mathbb{R}^d} \chi \mathrm{d}\lambda = 1$. Pour tout $\varepsilon > 0$, on définit

$$\chi_\varepsilon(x) = \frac{1}{\varepsilon^d} \chi\left(\frac{x}{\varepsilon}\right)$$

En posant $\phi\colon \mathbb{R}^d \to \mathbb{R}^d, x \mapsto \varepsilon x$ le $\mathcal{C}^1$ difféomorphisme, on a $J_\phi(x) = \varepsilon^d$ et on en déduit donc par un simple changement de variable que

$$\int_{\mathbb{R}^d} \chi_\varepsilon \mathrm{d}\lambda = \int_{\mathbb{R}^d} \frac{1}{\varepsilon^d} \chi \varepsilon^d \mathrm{d}\lambda = \int_{\mathbb{R}^d} \chi \mathrm{d}\lambda = 1, \qquad \forall \varepsilon > 0$$

Proposition 6.18. *Soit $f \in \mathcal{L}^p(\mathbb{R}^d)$ avec $p \in [1, +\infty[$. On a :*

$$\|f * \chi_e - f\|_p \underset{\varepsilon \to 0}{\longrightarrow} 0$$

Démonstration. On sait déjà que $f * \chi_\varepsilon \in \mathcal{L}^p(\mathbb{R}^d)$ et que $\|f * \chi_\varepsilon\|_p \leqslant \|f\|_p$. Pour presque tout $x \in \mathbb{R}^d$, on a :

$$\begin{aligned} (f * \chi_\varepsilon)(x) &= \int_{\mathbb{R}^d} (f(x-y) - f(x)) \chi_\varepsilon(y) \mathrm{d}y \\ &= \frac{1}{\varepsilon^d} \int_{\mathbb{R}^d} (f(x-y) - f(x)) \chi \left(\frac{y}{\varepsilon}\right) \mathrm{d}y \\ &= \int_{\mathbb{R}^d} (f(x-\varepsilon z) - f(x)) \chi(z) \mathrm{d}z \end{aligned}$$

En procédant comme dans la démonstration du théorème précédent, on a

$$|f * \chi_\varepsilon - f|(x) \leqslant \left(\int_{\mathbb{R}^d} |f(x-\varepsilon z) - f(x)|^p \chi(z) \mathrm{d}z \right)^{\frac{1}{p}} \left(\overbrace{\int_{\mathbb{R}^d} \chi(z) \mathrm{d}z}^{=1} \right)^{1-\frac{1}{p}}$$

En intégrant selon x sur $\mathbb{R}^d$ et en appliquant le théorème de Fubini-Tonelli, on en déduit que

$$\int_{\mathbb{R}^d} |f * \chi_\varepsilon - f|^p \mathrm{d}\lambda \leqslant \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |f(x-\varepsilon z) - f(x)|^p \mathrm{d}x \right) \chi(z) \mathrm{d}z$$

c'est-à-dire

$$\|f * \chi_\varepsilon - f\|_p^p \leqslant \int_{\mathbb{R}^d} \chi(z) \|\tau_{\varepsilon z} f - f\|_p^p \mathrm{d}z$$

Or, $\|\tau_{\varepsilon z} f - f\|_p^p$ converge ponctuellement vers 0 et est dominé par une fonction intégrable pour tout ε. Ainsi, par le théorème de convergence dominée, le membre de droite tend vers 0 quand ε tend vers 0.□

Lemme 6.19. *Il existe $\chi\colon \mathbb{R}^d \to \mathbb{R}_+$ une application de classe $\mathcal{C}^\infty$ telle que*

$$\int_{\mathbb{R}^d} \chi \mathrm{d}\lambda = 1$$

et $\chi(x) = 0$ si $|x| \geqslant 1$.

Démonstration. Soit $\varphi\colon \mathbb{R}^d \to \mathbb{R}_+$ l'application définie par

$$\varphi(x) = \begin{cases} \exp\left(-\dfrac{1}{1 - \|x\|^2} \right) & \text{si } \|x\| < 1 \\ 0 & \text{sinon} \end{cases}$$

On sait qu'une telle fonction est intégrable et $\mathcal{C}^\infty$, et on pose

$$\chi(x) = \frac{\varphi(x)}{\int_{\mathbb{R}^d} \varphi(y)\mathrm{d}y}$$ □

Exercice. Si χ est comme ci-dessus, et f est localement intégrable sur $\mathbb{R}^d$ (i.e. sur tous compacts), montrer que la fonction $g = \chi * f$ est $\mathcal{C}^\infty$

On en déduit alors un résultat encore plus fort que le théorème de densité :

Corollaire 6.20. *L'espace des fonctions $\mathcal{C}^\infty$ à support compact est dense dans $L^p(\mathbb{R}^d)$ si $1 \leqslant p < \infty$.*

Démonstration. Soit $f \in \mathcal{L}^p(\mathbb{R}^d)$ et $\delta > 0$. Il existe une fonction g continue à support compact telle que $\|g - f\|_p \leqslant \frac{\delta}{2}$. Soit χ comme ci-sessus, et choisissons $\varepsilon > 0$ assez petit pour que $\|g - g * \chi\|_p \leqslant \frac{\delta}{2}$. Alors $h = g * \chi_\varepsilon$ est $\mathcal{C}^\infty$ à support compact et $\|f - h\|_p \leqslant \|f - g\|_p + \|g - h\|_p \leqslant \delta$. □

Chapitre 7

Transformation de Fourier

Dans ce chapitre, pour simplifier, nous introduirons la notion de transformée de Fourier sur $\mathbb{R}$ plutôt que sur $\mathbb{R}^d$. Ce court chapitre est une application intéressante du cours d'intégration qui précède, dans le sens où de nombreux résultats sont utilisés (théorème de convergence dominée, théorèmes de continuité et dérivabilité pour les intégrales à paramètres, produit de convolution, densité des fonctions en escalier dans L^1...).

7.1 Définition et propriétés générales

Définition 7.1. Soit $f \in \mathcal{L}^1(\mathbb{R}^d)$. On appelle **transformée de Fourier** de f la fonction $\hat{f}\colon \mathbb{R} \to \mathbb{C}$ définie par

$$\hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{-ikx} \mathrm{d}x, \quad k \in \mathbb{R}$$

La correspondance $\mathcal{F}\colon f \mapsto \hat{f}$ s'appelle **la transformation de Fourier**.

Cette intégrale à un sens puisque $\left|f(x)e^{-ikx}\right| = |f(x)|$. Etudions maintenant les propriétés générales de $\hat{f}$.

Proposition 7.2. *Si $f \in \mathcal{L}^1(\mathbb{R})$, alors $\hat{f}$ est continue, $\left\|\hat{f}\right\|_\infty \leqslant \frac{1}{\sqrt{2\pi}} \|f\|_1$ et $\hat{f}(k) \underset{|k|\to\infty}{\longrightarrow} 0$ (Lemme de Riemann-Lebesgue)*

Démonstration. La fonction $x \mapsto f(x)e^{-ikx}$ est mesurable et, puisque $\left|f(x)e^{-ikx}\right| = |f(x)|$, est majorée par une fonction intégrable. De plus, la fonction $k \mapsto f(x)e^{-ikx}$ est continue pour tout $k \in \mathbb{R}$. Ainsi, la continuité de $\hat{f}$ suit du théorème 2.32 de continuité des intégrales dépendant d'un paramètre (page 30). En outre,

$$\forall k \in \mathbb{R}, \quad \left|\hat{f}(k)\right| \leqslant \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} |f(x)| \,\mathrm{d}x = \frac{1}{\sqrt{2\pi}} \|f\|_1$$

Pour montrer que $\hat{f}(k) \underset{|k|\to\infty}{\longrightarrow} 0$, on commence par le cas où $f = \mathbb{1}_I$ avec $I = [a, b]$ un intervalle de $\mathbb{R}$. On a alors, pour $k \in \mathbb{R} \setminus \{0\}$:

$$\hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-ikx} \mathrm{d}x = \frac{1}{\sqrt{2\pi}} \frac{e^{-ika} - e^{-ikb}}{ik}$$

donc

$$\left|\hat{f}(k)\right| \leqslant \frac{2}{\sqrt{2\pi}\,|k|} \underset{|k|\to\infty}{\longrightarrow} 0$$

Par linéarité, le résultat est vrai pour toute fonction f en escalier. Soient maintenant $f \in \mathcal{L}^1(\mathbb{R})$ et $\varepsilon > 0$. Par compacité des fonctions en escalier dans $\mathcal{L}^1(\mathbb{R})$, il existe une fonction g en escalier telle que $\|f-g\|_1 \leqslant \varepsilon$. De plus, il existe $k_0 > 0$ tel que $\forall\, |k| \geqslant k_0$, $|\hat{g}(k)| \leqslant \varepsilon$. Ainsi, si $|k| \geqslant k_0$,

$$\left|\hat{f}(k)\right| \leqslant \left|\hat{f}(k) - \hat{g}(k)\right| + |\hat{g}(k)| \leqslant \frac{1}{\sqrt{2\pi}} \|f-g\|_1 + \varepsilon \leqslant 2\varepsilon$$

Comme ε était arbitraire, on en déduit que $\hat{f}(k) \underset{|k|\to\infty}{\longrightarrow} 0$. □

On va maintenant étudier les conséquences sur $\hat{f}$ d'hypothèses plus fortes sur f. Par abus de notation, on notera dans la suite pour toute fonction $f\colon x \mapsto f(x)$:

$$xf\colon x \mapsto xf(x)$$

Proposition 7.3. *Soit $f \in \mathcal{L}^1(\mathbb{R})$.*

a) Si $xf \in \mathcal{L}^1(\mathbb{R})$, alors $\hat{f} \in \mathcal{C}^1(\mathbb{R})$ et

$$\hat{f}'(k) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x)(-ix)e^{-ikx}\mathrm{d}x = \widehat{-ixf}(k)$$

b) Si $f \in \mathcal{C}^1(\mathbb{R})$ et $f' \in \mathcal{L}^1(\mathbb{R})$, alors $k\hat{f} \in \mathcal{L}^\infty(\mathbb{R})$ et

$$k\hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left(-if'(x)\right) e^{-ikx}\mathrm{d}x = \widehat{-if'}(k)$$

Démonstration. a) La fonction $x \mapsto f(x)e^{-ikx}$ est mesurable, la fonction $g_x(k)\colon k \mapsto f(x)e^{-ikx}$ est dérivable à x fixé et $g_x'(k) = (-ix)f(x)e^{-ikx}$. De plus, pour tout $k \in \mathbb{R}$, $\left|(-ix)f(x)e^{-ikx}\right| = |xf(x)|$ est une fonction intégrable. Ainsi, la fonction $\hat{f}$ est dérivable par le théorème 2.33 de dérivabilité des intégrales dépendant d'un paramètre (page 30), et on a la formule voulue. La continuité de la dérivée se traite comme dans la démonstration précédente.

b) Les hypothèses impliquent que $f(x) \underset{|x|\to\infty}{\longrightarrow} 0$ (la dérivée étant intégrable admet des limites finies en $\pm\infty$, et comme f est aussi intégrable, ces limites sont nulles). Par une intégration par partie, on obtient donc

$$\begin{aligned} \widehat{f'}(k) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f'(x)e^{-ikx}\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}} \left[f(x)e^{-ikx}\right]_{x=-\infty}^{x=+\infty} + \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x)ike^{-ikx}\mathrm{d}x \\ &= ik\hat{f}(k) \end{aligned}$$

En multipliant par $-i$, on a le résultat voulu.

On va rencontrer plusieurs propositions du même genre : plus la fonction f est décroissante, plus sa transformée de Fourier est régulière. Réciproquement, plus f est régulière, plus sa transformée de Fourier est décroissante.

Exemples. – On peut calculer la transformée de Fourier de $f = \mathbb{1}_{[-a,a]}$ où $a > 0$. D'après le calcul effectué dans la démonstration précédente, on a

$$\hat{f}(k) = \begin{cases} \sqrt{\frac{2}{\pi}}\frac{\sin(ka)}{k} & \text{si } k \neq 0 \\ \sqrt{\frac{2}{\pi}}a & \text{si } k = 0 \end{cases}$$

Ici, f décroît très rapidement à l'infini (puisqu'à support compact), mais est peu régulière (deux points de discontinuité). Du coup, $\hat{f}$ est très régulière $\mathcal{C}^\infty(\mathbb{R})$ mais décroît relativement mal (comme $k \mapsto \frac{1}{k}$).

– On peut calculer la transformée de Fourier de $g(x) = e^{-a|x|}$ où $a > 0$. On a donc

$$\begin{aligned} \hat{g}(k) &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-a|x|}e^{-ikx}\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}}\int_0^\infty e^{x(-a-ik)}\mathrm{d}x + \frac{1}{\sqrt{2\pi}}\int_{-\infty}^0 e^{x(a-ik)}\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}}\left(\frac{1}{a+ik} + \frac{1}{a-ik}\right) \\ &= \sqrt{\frac{2}{\pi}}\frac{a}{a^2+k^2} \end{aligned}$$

Ici, g décroît très rapidement à l'infini et est un peu plus régulière que f (elle est continue mais non dérivable à l'origine). Du coup, $\hat{g}$ est très régulière $\mathcal{C}^\infty(\mathbb{R})$ et décroît relativement lentement (comme $k \mapsto \frac{1}{k^2}$), mais en tout cas, plus rapidement que $\hat{f}$.

– Finalement, on peut calculer la transformée de Fourier de la Gaussienne $h(x) = e^{-ax^2/2}$ où $a > 0$. En utilisant le a) de la proposition 7.3, on a $\hat{h}$ qui est $\mathcal{C}^1$ et par une intégration par parties, on trouve

$$\begin{aligned} \hat{h}'(k) &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}(-ix)e^{-ax^2/2}e^{-ikx}\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}}\left[\frac{i}{a}e^{-ax^2/2}e^{-ikx}\right]_{x=-\infty}^{x=+\infty} - \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\frac{i}{a}e^{-ax^2/2}(-ik)e^{-ikx}\mathrm{d}x \\ &= -\frac{1}{a}k\hat{h}(k) \end{aligned}$$

Ainsi, $\hat{h}(k) = \hat{h}(0)e^{-\frac{k^2}{2a}}$. Or, $\hat{h}(0) = \frac{1}{\sqrt{a}}$ donc

$$\hat{h}(k) = \frac{1}{\sqrt{a}}e^{-\frac{k^2}{2a}}$$

Ici, h est une fonction à décroissance très rapide à l'infini, et $\mathcal{C}^\infty$. Du coup, $\hat{h}$ est $\mathcal{C}^\infty$ et décroît très rapidement à l'infini.

Citons maintenant quelques propriétés immédiates de la transformation de Fourier.

Proposition 7.4. *Soit $f \in \mathcal{L}^1(\mathbb{R})$.*

1) Si f est paire (resp. impaire), alors $\hat{f}$ est paire (resp. impaire).

2) Si f est réelle, alors $\hat{f}(-k) = \overline{\hat{f}(k)}$. Du coup, si f est réelle et paire (resp. réelle et impaire), alors $\hat{f}$ est réelle et paire (resp. imaginaire pure et impaire).

3) Si $g(x) = f(x-y)$ où $y \in \mathbb{R}$, alors $\hat{g}(k) = \hat{f}(k)e^{-iky}$.

4) Si $g(x) = f(\alpha x)$ où $\alpha > 0$, alors $\hat{g}(k) = \frac{1}{\alpha}\hat{f}\left(\frac{k}{\alpha}\right)$.

Démonstration. 1) Si f est paire, on a

$$\hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(-x)e^{-ikx}\mathrm{d}x = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x)e^{-ik(-x)}\mathrm{d}x = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x)e^{-i(-k)x}\mathrm{d}x = \hat{f}(-k)$$

De même on montre le résultat pour f impaire.

2) Evident par définition. Le reste se déduit du 1).
3) On écrit $\hat{g}(k) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g(x)e^{-ikx}\mathrm{d}x$ et on fait le changement de variable $z = x - y$. On peut alors sortir e^{-iky} de l'intégrale et conclure.
4) On écrit $\hat{g}(k) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g(x)e^{-ikx}\mathrm{d}x$ et on fait le changement de variable $z = \alpha x$. La conclusion en découle immédiatement. □

Viens alors une proposition énonçant que la transformée de Fourier vérifie l'identité remarquable que l'on appelle la *formule de convolution*.

Proposition 7.5. *Si $f, g \in \mathcal{L}^1(\mathbb{R})$, et si $h\colon \mathbb{R} \to \mathbb{C}$ est la fonction définie presque partout par*

$$h(x) = \frac{1}{\sqrt{2\pi}}(f * g)(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x-y)g(y)\mathrm{d}y$$

alors $h \in \mathcal{L}^1(\mathbb{R})$ et $\hat{h}(k) = \hat{f}(k)\hat{g}(k)$.

Démonstration. Comme $h = \frac{1}{\sqrt{2\pi}} f * g$, on a vu que h est définie presque partout et est dans $\mathcal{L}^1$. Reste à calculer $\hat{h}$:

$$\hat{h}(k) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left(\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x-y)g(y)\mathrm{d}y \right) e^{-ikx}\mathrm{d}x$$

Si on appelle $F(x,y) = f(x-y)g(y)e^{-ikx}$, alors $(x,y) \mapsto |F(x,y)|$ est positive et mesurable, et son intégrale est finie, donc F est intégrable et par le théorème de Fubini-Lebesgue, on en déduit que

$$\hat{h}(k) = \frac{1}{2\pi} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x-y)e^{-ik(x-y)}\mathrm{d}y \right) g(y)e^{-iky}\mathrm{d}x = \hat{f}(k)\hat{g}(k)$$

□

Voyons maintenant un résultat fondamental qui fait toute l'utilité de la transformée de Fourier : la formule d'inversion.

Proposition 7.6 (formule d'inversion). *Soit $f \in \mathcal{L}^1(\mathbb{R})$ telle que $\hat{f} \in \mathcal{L}^1(\mathbb{R})$, alors on a*

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(k)e^{ikx}\mathrm{d}k$$

pour presque tout $x \in \mathbb{R}$. On notera $f = \mathcal{F}^{-1}(\hat{f})$.

Remarque. Si f est une application continue, alors on a l'égalité pour tout $x \in \mathbb{R}$!

On en déduit un petit corollaire qui n'est pas si évident à démontrer autrement et qui établit qu'on ne peut pas trouver de fonction orthogonale à toutes les fonctions $x \mapsto e^{-ikx}$ pour $k \in \mathbb{R}$ pour le produit scalaire dans $\mathcal{L}^1$.

Corollaire 7.7. *Si $f \in \mathcal{L}^1(\mathbb{R})$ vérifie $\displaystyle\int_{\mathbb{R}} f(x)e^{-ikx}\mathrm{d}x = 0$ pour tout $k \in \mathbb{R}$, alors $f = 0$ presque partout.*

Démonstration (Démonstration de la proposition 7.6). Notons pour tout $x \in \mathbb{R}$,

$$F(x) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \hat{f}(k)e^{ikx}\mathrm{d}k, \quad F_\varepsilon(x) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \hat{f}(k)e^{-\varepsilon|k|}e^{ikx}\mathrm{d}k, \quad \varepsilon > 0$$

On sait que F et F_ε sont continues, bornées et nulles à l'infini par la proposition 7.2. Par le théorème de convergence dominée (comme $\left|\hat{f}(k)e^{-\varepsilon|k|}e^{ikx}\right| \leqslant \left|\hat{f}(k)\right|$ pour tout $k \in \mathbb{R}$ et $\varepsilon > 0$), on déduit que $F_\varepsilon(x) \underset{x\to 0}{\longrightarrow} F(x)$ pour tout $x \in \mathbb{R}$. Si l'on pose maintenant $G(k,y) = f(y)e^{-iky}e^{-\varepsilon|k|}e^{ikx}$, alors G est intégrable car $|G(k,y)| = \left|f(y)e^{-\varepsilon|k|}\right|$ qui est intégrable, et par le théorème de Fubini-Lebesgue, on a :

$$\begin{aligned}
F_\varepsilon(x) &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\left(\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(y)e^{-iky}\mathrm{d}y\right)e^{-\varepsilon|k|}e^{ikx}\mathrm{d}k \\
&= \frac{1}{2\pi}\int_{\mathbb{R}}\underbrace{\left(\int_{\mathbb{R}} e^{-\varepsilon|k|}e^{ik(x-y)}\mathrm{d}k\right)}_{=\frac{2\varepsilon}{\varepsilon^2+(x-y)^2}} f(y)\mathrm{d}y \\
&= \frac{\varepsilon}{\pi}\int_{\mathbb{R}} f(y)\frac{1}{\varepsilon^2+(x-y)^2}\mathrm{d}y \\
&= (\chi_\varepsilon * f)(x)
\end{aligned}$$

où $\chi_\varepsilon(x) = \frac{1}{\varepsilon}\chi\left(\frac{x}{\varepsilon}\right)$ et $\chi(z) = \frac{1}{\pi}\frac{1}{1+z^2}$. Comme $f \in \mathcal{L}^1(\mathbb{R})$, on sait que $\chi_\varepsilon * f \underset{\varepsilon\to 0}{\longrightarrow} f$ dans $\mathcal{L}^1(\mathbb{R})$. Ainsi, $F_\varepsilon \in \mathcal{L}^1(\mathbb{R})$ et $F_\varepsilon \underset{\varepsilon\to 0}{\longrightarrow} f$ dans $\mathcal{L}^1(\mathbb{R})$. En particulier, il existe une suite $(\varepsilon_n)_{n\in\mathbb{N}}$ qui tend vers 0 telle que $F_{\varepsilon_n}(x) \underset{n\to\infty}{\longrightarrow} f(x)$ pour presque tout $x \in \mathbb{R}$. Ainsi, $f(x) = F(x)$ pour presque tout $x \in \mathbb{R}$. □

7.2 Etablissement d'un cadre fonctionnel

Jusqu'à présent, on a vu que les propriétés de régularité et de décroissance à l'infinie de la transformée de Fourier dépendait inversement de la régularité et de décroissance de la fonction initiale. Bref, on ne sait pas bien sur quoi les espaces sont envoyés par $\mathcal{F}$. On cherche donc un espace dans lequel la transformation de Fourier est stable et bijective. On va mentionner deux exemples de tels espaces fonctionnels : l'espace de Schwartz et l'espace $L^2(\mathbb{R})$.

7.2.1 L'espace de Schwartz

On cherche un sous-espace de $\mathcal{L}^1(\mathbb{R})$ dont l'image par la transformée de Fourier reste dans l'espace des fonctions intégrables pour pouvoir appliquer la transformée de Fourier inverse et qui soit stable par dérivation.

Définition 7.8. On dit qu'une fonction $f : \mathbb{R} \to \mathbb{C}$ de classe $\mathcal{C}^\infty$ est **à décroissance rapide** si pour tous $k, l \in \mathbb{N}$, on a :

$$\sup_{x\in\mathbb{R}}\left|x^k f^{(l)}(x)\right| < \infty$$

On note $\mathcal{S}(\mathbb{R})$ l'ensemble des fonctions $\mathcal{C}^\infty$ à décroissance rapide sur $\mathbb{R}$, et on l'appelle l'**espace de Schwartz**.

On remarque que $\mathcal{S}(\mathbb{R}) \subset \mathcal{L}^p(\mathbb{R})$ pour tout $p \in [1, +\infty[$ et que $\mathcal{S}(\mathbb{R})$ est dense dans tous les espaces $\mathcal{L}^p(\mathbb{R})$ puis qu'il contient toutes les fonctions $\mathcal{C}^\infty$ à support compact. On a de plus que la fonction de Gauss appartient à cet espace.

Proposition 7.9. *Si $f \in \mathcal{S}(\mathbb{R})$, alors $\hat{f} \in \mathcal{S}(\mathbb{R})$ et la transformation de Fourier $\mathcal{F} \colon \mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R})$ est bijective, d'inverse $\mathcal{F}^{-1}$.*

Démonstration. 1) Si $f \in \mathcal{S}(\mathbb{R})$, on montre que $\hat{f} \in \mathcal{S}(\mathbb{R})$ en itérant les arguments de la proposition 7.3.
2) En outre, on a la formule d'inversion, qui est définie aussi sur tout $\mathcal{S}(\mathbb{R})$ avec les mêmes propriétés (les démonstrations sont les mêmes), donc $\mathcal{F}$ est bijective.

7.2.2 La transformée de Fourier dans $L^2(\mathbb{R})$

Un autre espace fonctionnel répondant au cahier des charges est l'espace $L^2(\mathbb{R})$, et $\mathcal{F}$ est même une isométrie dans cet espace. Pour le moment, nous n'avons défini la transformée de fourier que pour une fonction $f \in L^1(\mathbb{R})$.

Proposition 7.10. *Soit $f \in \mathcal{L}^1(\mathbb{R}) \cap \mathcal{L}^2(\mathbb{R})$. Alors, $\hat{f} \in \mathcal{L}^2(\mathbb{R}) \cap \mathcal{L}^\infty(\mathbb{R})$ et on a*

$$\int_{\mathbb{R}} |f(x)|^2 \, \mathrm{d}x = \int_{\mathbb{R}} \left|\hat{f}(k)\right|^2 \mathrm{d}k$$

Autrement dit, $\|f\|_2 = \left\|\hat{f}\right\|_2$.

Démonstration. On applique la proposition 7.4 avec $g(x) = \overline{f(-x)}$, de sorte que $g \in \mathcal{L}^1(\mathbb{R}) \cap \mathcal{L}^2(\mathbb{R})$ et que $\hat{g}(k) = \overline{\hat{f}(k)}$ pour tout $k \in \mathbb{R}$. La fonction

$$h(x) = \frac{1}{\sqrt{2\pi}}(f * g)(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x-y)\overline{f(-y)}\mathrm{d}y = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x+y)\overline{f(y)}\mathrm{d}y \tag{7.1}$$

vérifie, par la proposition 7.5, $h \in \mathcal{L}^1(\mathbb{R})$ et $\hat{h}(k) = \hat{f}(k)\hat{g}(k) = \left|\hat{f}(k)\right|^2$. Par ailleurs, on sait par l'inégalité de Young (théorème 6.17 page 90) que h est continue et que $h(x) \underset{|x|\to\infty}{\longrightarrow} 0$. Définissons alors

$$H_\varepsilon(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{h}(k) e^{-\varepsilon|k|} e^{ikx} \mathrm{d}x, \qquad \varepsilon > 0$$

Par un calcul analogue à celui effectué dans la démonstration de la proposition 7.6, on en déduit que

$$H_\varepsilon(x) = \frac{1}{\pi} \int_{\mathbb{R}} h(x - \varepsilon z) \frac{1}{1+z^2} \mathrm{d}z$$

En particulier, en posant[1] $x = 0$, on a :

$$H_\varepsilon(0) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{h}(k) e^{-\varepsilon|k|} \mathrm{d}k = \frac{1}{\pi} \int_{\mathbb{R}} h(-\varepsilon z) \frac{1}{1+z^2} \mathrm{d}z \tag{7.2}$$

Pour tout $n \in \mathbb{N}$, posons $h_n \colon z \mapsto h(-\frac{1}{n}z)\frac{1}{1+z^2}$. Alors, $h_n(z) \underset{n\to\infty}{\longrightarrow} h(0)\frac{1}{1+z^2}$ et est majorée en module par $h(0)\left|\frac{1}{1+z^2}\right|$ (car $\forall z, |h(z)| \leqslant h(0)$ au vu de 7.1[2]) qui est intégrable, on peut appliquer le théorème 2.25 de

1. Ceci est tout à fait licite puisque la fonction est continue.
2. C'est une propriété générale des fonctions "de type positif", i.e. dont la transformée de Fourier est positive.

la convergence dominée, et

$$\frac{1}{\pi}\int_{\mathbb{R}} h(-\varepsilon z)\frac{1}{1+z^2}\mathrm{d}z \underset{\varepsilon\to 0}{\longrightarrow} h(0) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} |f(y)|^2\,\mathrm{d}y$$

alors le membre de gauche de l'équation 7.2 a aussi une limite quand $\varepsilon \to 0$. Ainsi, par le théorème de convergence monotone, on déduit que

$$\lim_{\varepsilon\to 0}\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\left|\hat{f}(k)\right|^2 e^{-\varepsilon|k|}\mathrm{d}k = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\left|\hat{f}(k)\right|^2\mathrm{d}k = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}|f(y)|^2\,\mathrm{d}y < \infty$$

□

On veut maintenant définir la transformée de Fourier $L^2(\mathbb{R})$. Soit $f \in L^2(\mathbb{R})$. Pour tout $n \in \mathbb{N}$, on note $f_n = f\mathbb{1}_{[-n,n]}$. Alors, $f_n \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ pour tout $n \in \mathbb{N}$, et $\|f_n - f\|_2 \underset{n\to\infty}{\longrightarrow} 0$. En effet,

- $f \in L^2(\mathbb{R})$ et $\mathbb{1}_{[-n,n]} \in L^2(\mathbb{R})$, $n \in \mathbb{N}$, alors par l'inégalité de Hölder (proposition 6.4, page 80), $f_n \in L^1(\mathbb{R})$.
- $\|f_n - f\|_2^2 = \int_{|x|\geqslant n}|f(x)|^2\,\mathrm{d}x$ pour tout $n \in \mathbb{N}$, quantité qui tend vers 0 quand n tend vers l'infini d'après le théorème de convergence dominée.

Comme $\|f_n - f_m\|_2 = \left\|\widehat{f_n} - \widehat{f_m}\right\|_2$ par la proposition 7.10, la suite $(\widehat{f_n})_{n\in\mathbb{N}}$ est de Cauchy dans $L^2(\mathbb{R})$. Par le théorème de Riesz-Fisher (6.8, page 82), cette suite a une limite, que l'on notera $\hat{f}$. On définit la transformation de Fourier sur $L^2(\mathbb{R})$ par $\hat{f} = \mathcal{F}f$. On vérifie alors que :

1) $\mathcal{F}$ est isométrique : $\left\|\hat{f}\right\|_2 = \|f\|_2$ (ce qui découle immédiatement de la proposition 7.10).
2) Si $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, alors la définition ci-dessus coïncide avec la définition donnée par la formule intégrale.

Définition 7.11. Soit $f \in L^2(\mathbb{R})$. On appelle **transformée de Fourier** de f la limite dans $L^2(\mathbb{R})$ de la suite $(\widehat{f_n})_{n\in\mathbb{N}}$ définie par :

$$\widehat{f_n}(k) = \frac{1}{\sqrt{2\pi}}\int_{-n}^{n} f(x)e^{-ikx}\mathrm{d}x, \quad n \in \mathbb{N}$$

On note cette limite $\hat{f} = \mathcal{F}f$.

Dans cet espace et avec cette définition de la transformée de Fourier, on a trouvé ce que l'on cherchait, à savoir un espace fonctionnel dans lequel la transformation de Fourier est stable et bijective.

Théorème 7.12. *La transformation de Fourier* $\mathcal{F}\colon \begin{array}{ccc} L^2(\mathbb{R}) & \longrightarrow & L^2(\mathbb{R}) \\ f & \longmapsto & \hat{f} \end{array}$ *est un isomorphisme d'espaces de Hilbert :*

- *$\mathcal{F}$ est linéaire et bijective.*
- *Pour tout $f, g \in L^2(\mathbb{R})$, on a $(f|g) = (\hat{f}|\hat{g})$*

Si $f \in L^2(\mathbb{R})$, alors f est la limite dans $L^2(\mathbb{R})$ de la suite $(\widetilde{f_n})_{n\in\mathbb{N}}$ définie par :

$$\widetilde{f_n}(x) = \frac{1}{\sqrt{2\pi}}\int_{-n}^{n}\hat{f}(k)e^{ikx}\mathrm{d}k$$

Démonstration. Il est clair que $\mathcal{F}$ est linéaire, et, par construction, isométrique. En particulier, $\mathcal{F}$ est donc injective. Ainsi, c'est une bijection de $L^2(\mathbb{R})$ sur un sous-espace V de $L^2(\mathbb{R})$. Comme $\mathcal{F}$ est une isométrie, alors V est fermé. On sait par ailleurs que $\mathcal{F}$ est bijective sur $\mathcal{S}(\mathbb{R})$ qui est dense dans $L^2(\mathbb{R})$. Comme $V \supset \mathcal{S}(\mathbb{R})$, V est dense dans $L^2(\mathbb{R})$ donc $V = L^2(\mathbb{R})$.
Finalement, la conservation des normes implique la conservation du produit scalaire[3]. $\square$

7.3 Formules de la transformation de Fourier sur $\mathbb{R}^d$

Tout ce qui précède (excepté les formules avec des dérivations à adapter en mettant des dérivées partielles) reste valable si $\mathbb{R}$ est remplacé par $\mathbb{R}^d$, $d \geqslant 1$. En notant $x \cdot y$ le produit scalaire entre $x, y \in \mathbb{R}^d$, on a, pour tout f :

$$\begin{cases} \hat{f}(k) = \frac{1}{(2\pi)^{d/2}} \displaystyle\int_{\mathbb{R}^d} f(x) e^{-ik\cdot x} \mathrm{d}x, & k \in \mathbb{R}^d \\ f(x) = \frac{1}{(2\pi)^{d/2}} \displaystyle\int_{\mathbb{R}^d} \hat{f}(k) e^{ik\cdot x} \mathrm{d}k, & x \in \mathbb{R}^d \end{cases}$$

7.4 Applications

Dans cette dernière partie, nous présenterons deux applications de la transformation de Fourier : l'une pour résoudre facilement une équation différentielle, et l'autre qui justifie l'introduction de la transformée de Fourier : la résolution de l'équation de la chaleur à une dimension.

7.4.1 A la rescousse des équations différentielles

Exercice. Soit $f \in \mathcal{C}^0(\mathbb{R}) \cap \mathcal{L}^1(\mathbb{R})$. Trouver toutes les solutions de l'équation différentielle :

$$-h''(x) + h(x) = f(x), \qquad x \in \mathbb{R}$$

telle que $h \in \mathcal{C}^2(\mathbb{R}) \cap \mathcal{L}^1(\mathbb{R})$.

Démonstration (Solution). On applique la transformation de Fourier à l'équation[4] différentielle. En notant $\hat{f} = \mathcal{F}f$ et $\hat{h} = \mathcal{F}h$, on trouve :

$$k^2\hat{h}(k) + \hat{h}(k) = \hat{f}(k), \qquad \text{i.e.} \qquad (1+k^2)\hat{h}(k) = \hat{f}(k)$$

Ainsi,

$$\hat{h}(k) = \frac{1}{1+k^2}\hat{f}(k), \qquad k \in \mathbb{R}$$

Notons alors

$$g(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \frac{1}{1+k^2} e^{ikx} \mathrm{d}k = \sqrt{\frac{\pi}{2}} e^{-|x|}$$

On a donc $h = \frac{1}{\sqrt{2\pi}} g * f$, d'où

$$h(x) = \frac{1}{2} \int_{\mathbb{R}} e^{-|x-y|} f(y) \mathrm{d}y$$

3. Rappel : $(u|v) = \frac{1}{4}\left(\|u+v\|^2 - \|u-v\|^2 + i\|u+iv\|^2 - i\|u-iv\|^2\right)$ (identité de polarisation)
4. La transformation de Fourier peut-être vue comme un changement de base qui diagonalise les opérateurs différentiels !

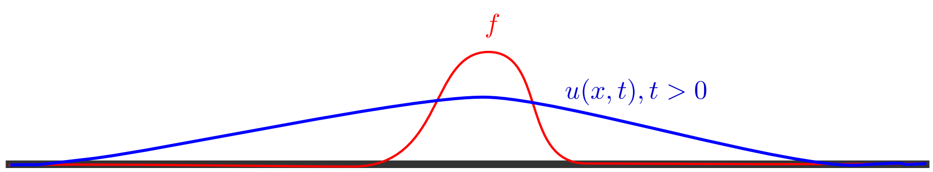
En résolvant l'équation différentielle de manière "classique", on trouve que les solutions générales ont deux termes supplémentaires de la forme Ce^{x} et De^{-x} où $C, D \in \mathbb{R}$, mais comme ces termes explosent à l'infini si C et D ne sont pas nuls, on les aurait écartés par la suite. Ainsi, la transformation de Fourier a permis immédiatement d'écarter ces termes (qui ne sont pas dans $\mathcal{L}^1(\mathbb{R})$). $\square$

7.4.2 L'équation de la chaleur à une dimension

L'analyse en fréquence apparaît en 1802 dans le traité de J. Fourier "Théorie analytique de la chaleur". Mathématiquement, on dispose d'une fonction $f\colon \mathbb{R} \to \mathbb{R}$ continue et bornée. On cherche une fonction $u\colon \mathbb{R} \times [0, +\infty[\to \mathbb{R}$, continue telle que u est de classe $\mathcal{C}^2$ sur $\mathbb{R} \times]0, +\infty[$ et telle que

$$\begin{cases} \frac{\partial u}{\partial t}(x,t) = \frac{\partial^2 u}{\partial x^2}(x,t), & x \in \mathbb{R}, t > 0 \\ u(x,0) = f(x), & x \in \mathbb{R} \end{cases}$$

Autrement dit, on dispose d'un barreau idéal de longueur infinie, chauffé initialement selon la distribution de chaleur f. On sait que la chaleur se diffuse, et on cherche de quelle façon la fonction de diffusion u se comporte.



On applique la transformation de Fourier sur la variable x avec

$$\begin{cases} u(x,t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{u}(k,t) e^{ikx} \mathrm{d}k \\ f(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(k) e^{ikx} \mathrm{d}k \end{cases}$$

On suppose ici formellement que toutes les hypothèses des théorèmes que nous sommes amenés à utiliser sont vérifiées :

$$\begin{cases} \dfrac{\partial u}{\partial t}(x,t) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{\mathbb{R}} \underbrace{\dfrac{\partial \hat{u}}{\partial t}(k,t)}_{=} e^{ikx} \mathrm{d}k \\ \dfrac{\partial^2 u}{\partial x^2}(x,t) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{\mathbb{R}} \overbrace{\left(-k^2 \hat{u}(k,t)\right)} e^{ikx} \mathrm{d}k \end{cases}$$

On trouve donc l'équation suivante :

$$\frac{\partial \hat{u}}{\partial t}(k,t) = -k^2 \hat{u}(k,t)$$

Ainsi, $\hat{u}(k,t) = \hat{u}(k,0) e^{-k^2 t} = \hat{f}(k) e^{-k^2 t}$. Notons alors

$$G_t(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-k^2 t} e^{ikx} \mathrm{d}k = \frac{1}{\sqrt{4\pi t}} e^{-\frac{|x|^2}{4t}}$$

Ainsi, $u(x,t) = (G_t * f)(x)$, c'est-à-dire

$$u(x,t) = \frac{1}{\sqrt{4\pi t}} \int_{\mathbb{R}} e^{-\frac{|x-y|^2}{4t}} f(y) \mathrm{d}y$$

Si f est continue est bornée, alors u est $\mathcal{C}^2$ et bornée pour tout $t \geqslant 0$. De plus, u vérifie l'équation de la chaleur, et à x fixé, on a $u(x,t) \underset{t \to 0}{\longrightarrow} f(x)$.

CHAPITRE A

MESURE DE HAUSDORFF

Dans cette annexe, nous étudions brièvement les mesures de Hausdorff, qui généralisent la mesure de Lebesgue. Même si elles n'ont pas la même importance pratique que la mesure de Lebesgue, il est utile, dans de nombreux domaines des mathématiques, d'être un tant soit peu familier avec celles-ci.

La théorie des mesures de Hausdorff, née une quinzaine d'années après celle de la mesure de Lebesgue, répondait à plusieurs motivations principales, notamment :

1) **la mesure d'objets de dimension inférieures**. En effet, si on se place en dimension 3, on a vu que la mesure de Lebesgue λ_3 permet d'attribuer à toutes les parties mesurables de $\mathbb{R}^3$ un "volume". Cependant, si l'on a besoin de définir "l'aire" d'une surface, ou la "longueur" d'une courbe tracée dans $\mathbb{R}^3$, la mesure de Lebesgue de tels objets est bien sûr nulle... D'où l'introduction d'une mesure en dimension 3 permettant de "classer" les ensembles négligeables du point de vue de la mesure de Lebesgue. Du coup, appliquée à un ensemble de "volume" non nul, on s'attend à ce que cette nouvelle mesure soit infinie.
2) **la notion de dimension**. Intuitivement, on a l'habitude de penser qu'une coube "régulière" est de dimension 1, car on peut localement la déformer continûment en un morceau de droite. En généralisant, on pensera donc qu'un ensemble est de dimension k si on peut l'exprimer localement au moyen de k fonctions indépendantes. En particulier l'image par une application régulière d'un ensemble de dimension k devrait être de dimension au plus k. Ceci est effectivement vrai quand on travaille avec des fonctions régulières. Cependant, il existe des courbes dites "de Peano" qui sont des courbes continues surjectives de $[0, 1]$ dans $[0, 1]^2$. Immédiatement, on voit que l'image d'une telle courbe est de dimension 2... Cet exemple montre bien qu'il est impossible de définir une notion de dimension qui soit basée sur les dimensions d'espaces de départ et d'arrivée.

A.1 COMPLÉMENTS SUR LES MESURES EXTÉRIEURES

Soit (X, d) un espace métrique séparable. Ainsi, pour tout $\delta > 0$ et partie E de X, on peut recouvrir E par une suite de boules de diamètres $\leqslant \delta$, et plus généralement, de parties de diamètres $\leqslant \delta$.

Définition A.1 (mesure métrique extérieure). Une mesure extérieure μ^* est appelée **mesure métrique extérieure** si pour tous ensembles $E, F \subset X$ tels que $d(E, F) > 0$, on a $\mu^*(E \cup F) = \mu^*(E) + \mu^*(F)$.

On rappelle que la distance entre deux parties E et F est donnée par la formule :

$$d(E, F) = \inf_{\substack{x \in E \\ y \in F}} f(x, y)$$

Définition A.2 (δ-recouvrement). Soit $\delta > 0$. Une famille $\{X_i\}_{i \in \mathbb{N}}$ de parties de X est appelée un **δ-recouvrement** si

$$X \subset \bigcup_{i \in \mathbb{N}} X_i, \qquad diam(X_i) < \delta \quad \forall i \in \mathbb{N}$$

On rappelle que le diamètre d'une partie E de X est donné par la formule :

$$diam(E) = \sup_{\substack{x \in E \\ y \in E}} d(x, y)$$

On va vouloir montrer que la mesure de Hausdorff est une mesure métrique extérieure. Pour cela, on aura besoin du lemme suivant :

Lemme A.3 (de Carathéodory). *Soient μ^* une mesure métrique extérieure, et $\{A_i\}_{i \in \mathbb{N}}$ une famille croissante de parties de X. On pose $A = \bigcup_{i \in \mathbb{N}} A_i$. Si $d(A_i, A \setminus A_{i+1}) > 0$ pour tout $i \in \mathbb{N}$, alors on a*

$$\mu^*(A) = \lim_{i \to \infty} \mu^*(A_i)$$

Démonstration. Posons $B_0 = A_0$ et, pour tout $j \geqslant 1$, $B_j = A_j \setminus A_{j-1}$. Comme la famille $\{A_i\}_{i \in \mathbb{N}}$ était croissante, $\{B_j\}_{j \in \mathbb{N}}$ est une famille de parties de X deux à deux disjointes et telle que

$$\bigcup_{j \in \mathbb{N}} B_j = \bigcup_{i \in \mathbb{N}} A_i = A$$

Supposons qu'il existe j_0 et j_1 deux entiers de même parité tels que $j_0 < j_1$ et $d(B_{j_0}, B_{j_1}) = 0$. Alors,

$$\forall \varepsilon > 0, \exists (x, y) \in B_{j_0} \times B_{j_1} : d(x, y) < \varepsilon$$

Prenons $\varepsilon = d(A_{j_0}, A \setminus A_{j_0+1}) > 0$ par hypothèse. Il existe donc $(x_0, y_0) \in B_{j_0} \times B_{j_1}$ tel que $d(x_0, y_0) < \varepsilon$. Or, $x_0 \in B_{j_0} \subset A_{j_0}$ et $y_0 \in B_{j_1} = A \setminus A_{j_1-1}$.
Comme $j_0 < j_1 - 1$ (car de même parité et $j_0 < j_1$) et que la famille $\{A_i\}_{i \in \mathbb{N}}$ est croissante, alors $y_0 \in A \setminus A_{j_0+1}$, et donc $d(x_0, y_0) \geqslant d(A_{j_0}, A \setminus A_{j_0+1}) = \varepsilon$. On obtient une absurdité.
Ainsi, par hypothèse,

$$\mu^* \left(\bigcup_k B_{2k} \right) = \sum_k \mu^*(B_{2k}), \quad \mu^* \left(\bigcup_k B_{2k+1} \right) = \sum_k \mu^*(B_{2k+1})$$

On veut montrer que $\lim_{i \to \infty} \mu^*(A_i) = \mu^*(A)$. Pour tout $i \in \mathbb{N}$, on a

$$A_i \subset A \subset A_i \cup \bigcup_{k \geqslant i+1} B_k$$

Donc, comme μ^* est une mesure extérieure

$$\mu^*(A_i) \leqslant \mu^*(A) \leqslant \mu^*(A_i) + \sum_{k \geqslant i+1} \mu^*(B_k) \tag{A.1}$$

Il suffit maintenant de montrer que $\lim_{i \to \infty} \sum_{k \geqslant i} \mu^*(B_k) = 0$. Pour cela, il suffit de montrer que la série de terme général $\mu^*(B_k)$ converge. En effet, pour tout $N \in \mathbb{N}$,

$$\begin{cases} \sum_{k=0}^{N} \mu^*(B_{2k}) = \mu^*(\bigcup_{k=0}^{N} B_{2k}) \leqslant \mu^*(A) < \infty \\ \sum_{k=0}^{N} \mu^*(B_{2k+1}) = \mu^*(\bigcup_{k=0}^{N} B_{2k+1}) \leqslant \mu^*(A) < \infty \end{cases}$$

d'où $\sum_{k\geqslant 0}\mu^*(B_k)\leqslant 2\mu^*(A)$.
Finalement, en passant à la limite dans (A.1), on a

$$\mu^*(A)=\lim_{i\to\infty}\mu^*(A_i)$$ □

Ce lemme va nous permettre de montrer qu'une mesure métrique extérieure définit une mesure borélienne.

Proposition A.4. *Tout borélien est μ^*-régulier, i.e. $\mathcal{B}(X)\subset\mathcal{M}(\mu^*)$.*

Démonstration. Soit E un ensemble fermé de X. Montrons que E est μ^*-régulier. Pour cela, soit A une partie de X. On veut montrer que $\mu(A)\geqslant\mu(A\cap E)+\mu(A\cap E^{\complement})$, où $E^{\complement}=X\setminus E$ est le complémentaire de E, l'autre inégalité provenant immédiatement de la sous-additivité de μ^*. On peut supposer que $\mu(A)<\infty$, sans quoi il n'y a rien à montrer.
Pour tout $j\geqslant 1$, posons

$$E_j=\left\{x\in X\mid d(x,E)>\frac{1}{j}\right\}$$

Comme $d(A\cap E,A\cap E_j)>0$, on a

$$\mu^*\left((A\cap E_j)\cup(A\cap E)\right)=\mu^*(A\cap E_j)+\mu^*(A\cap E)$$

d'où

$$\mu^*(A)\geqslant\mu^*(A\cap E_j)+\mu^*(A\cap E)$$

par croissance de la mesure extérieure.
En passant à la limite, comme $\bigcup_{j\in\mathbb{N}}E_j=E^{\complement}$ (E est fermé), et comme $\{A\cap E_j\}_{i\in\mathbb{N}}$ vérifie les hypothèses du lemme précédent, on obtient l'inégalité voulue.
Finalement, tout fermé est μ^*-régulier, et comme les fermés engendrent la tribu borélienne, on a la conclusion voulue. □

A.2 Définition de la mesure de Hausdorff

Maintenant que l'on a énoncé des définitions et résultats généraux pour les mesures extérieures métriques, on va définir une mesure extérieure qui, en passant à la limite sur les recouvrements de l'ensemble, donnera la définition de la mesure de Hausdorff. Pour toute partie E de X, pour tout $s\in[0,+\infty[$ et pour tout $\delta>0$, on note

$$\mathcal{H}^s_\delta(E)=\inf\left\{\sum_{i\in\mathbb{N}}diam(E_i)^s\right\}$$

où l'infimum est pris sur tous les δ-recouvrements dénombrables de E.
On a clairement

$$\mathcal{H}^s_\delta(\varnothing)=0 \tag{A.2}$$

De plus, si A et B sont deux parties de X telles que $A\subset B$, alors un δ-recouvrement de B est un δ-recouvrement A, et comme l'infimum est sur tous les recouvrements, on a

$$\mathcal{H}^s_\delta(A)\leqslant\mathcal{H}^s_\delta(B) \tag{A.3}$$

Soit $\{A_n\}_{n\in\mathbb{N}}$ une famille de parties de X, et on note $A = \bigcup_{n\in\mathbb{N}} A_n$. Soit $\varepsilon > 0$. Pour tout $n \in \mathbb{N}$, il existe un δ-recouvrement de A_n que l'on note $\{A_n^i\}_{i\in\mathbb{N}}$ tel que

$$\left|\mathcal{H}^s_\delta(A_n) - \sum_{i\in\mathbb{N}} diam(A_n^i)^s\right| \leqslant \frac{\varepsilon}{2^n}$$

On a alors $\left\{A_n^i\right\}_{\substack{n\in\mathbb{N}\\ i\in\mathbb{N}}}$ qui est un δ-recouvrement dénombrable de A, donc

$$\mathcal{H}^s_\delta(A) \leqslant \sum_{n\in\mathbb{N}}\sum_{i\in\mathbb{N}} diam(A_n^i)^s \leqslant \sum_{n\in\mathbb{N}}\left(\mathcal{H}^s_\delta(A_n) + \frac{\varepsilon}{2^n}\right) \leqslant 2\varepsilon + \sum_{n\in\mathbb{N}} \mathcal{H}^s_\delta(A_n)$$

Comme ε était arbitraire, on a

$$\mathcal{H}^s_\delta(A) \leqslant \sum_{n\in\mathbb{N}} \mathcal{H}^s_\delta(A_n) \tag{A.4}$$

Finalement, $\mathcal{H}^s_\delta$ est une mesure extérieure.

Définition A.5. Pour toute partie E de X, pour tout $s \in [0, +\infty[$, on note

$$\mathcal{H}^s(E) = \lim_{\delta\to 0} \mathcal{H}^s_\delta(E)$$

$\mathcal{H}^s$ est une mesure métrique extérieure, donc elle définit une mesure borélienne. On l'appelle **mesure de Hausdorff** (de dimension s).

Démonstration. L'application $\delta > 0 \mapsto \mathcal{H}^s_\delta(A)$, pour tout A partie de X, est décroissante (puisque l'infimum est pris sur des recouvrements de diamètre de plus en plus petits) et admet donc une limite en 0. De ce fait, on peut faire tendre δ vers 0 dans (A.2), (A.3) et (A.4), et on obtient

$$\mathcal{H}^s(\varnothing) = 0, \quad \mathcal{H}^s(A) \leqslant \mathcal{H}^s(B), \quad \mathcal{H}^s\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leqslant \sum_{n\in\mathbb{N}} \mathcal{H}^s(A_n)$$

pour toutes A, B parties de X et $\{A_n\}_{n\in\mathbb{N}}$ famille de parties de X.
Reste à montrer que cette mesure extérieure est métrique. Soient A et B deux parties de X telles que $d(A, B) > 0$. On sait déjà, $\mathcal{H}^s$ étant une mesure, que

$$\mathcal{H}^s(A \cup B) \leqslant \mathcal{H}^s(A) + \mathcal{H}^s(B)$$

Pour l'inégalité dans l'autre sens, fixons $\delta \in \left]0, \frac{d(A,B)}{3}\right[$ et $\varepsilon > 0$. Soit $\{C_i\}_{i\in\mathbb{N}}$ un δ-recouvrement de $A \cup B$ tel que

$$\left|\mathcal{H}^s_\delta(A\cup B) - \sum_{i\in\mathbb{N}} diam(C_i)^s\right| \leqslant \varepsilon$$

On pose

$$I = \{i \in \mathbb{N} \mid C_i \cap A \neq \varnothing\} \quad \text{et} \quad J = \{i \in \mathbb{N} \mid C_i \cap B \neq \varnothing\}$$

Comme $\delta < \frac{d(A,B)}{3}$, I et J sont disjoints, et de plus $\{C_i\}_{i\in I}$ (resp. $\{C_i\}_{i\in J}$) est un δ-recouvrement de A (resp. de B). On en déduit que

$$\begin{aligned} \mathcal{H}^s_\delta(A\cup B) + \varepsilon &\geqslant \sum_{i\in I} diam(C_i)^s + \sum_{i\in J} diam(C_i)^s \\ &\geqslant \mathcal{H}^s_\delta(A) + \mathcal{H}^s_\delta(B) \end{aligned}$$

En faisant tendre δ vers 0, on obtient

$$\mathcal{H}^s(A \cup B) + \varepsilon \geqslant \mathcal{H}^s(A) + \mathcal{H}^s(B)$$

Finalement, en faisant tendre ε vers 0, on obtient l'inégalité voulue. □

A.3 Propriétés et dimension de Hausdorff

Proposition A.6. *1) La mesure de Hausdorff est invariante par translation.*

2) La mesure de Hausdorff vérifie la propriété de dilatation suivante :

$$\mathcal{H}^s(\lambda A) = \lambda^s \mathcal{H}^s(A), \qquad \forall s \in [0, +\infty], \forall A \subset \mathbb{R}^n, \forall \lambda > 0$$

Démonstration. 1) En effet, si $\{A_i\}_{i\in\mathbb{N}}$ est un δ-recouvrement de $A \subset \mathbb{R}^n$, alors pour tout $b \in \mathbb{R}^n$, $\{b + A_i\}_{i\in\mathbb{N}}$ est un δ-recouvrement de $b + A$, où $b + P = \{b + p \mid p \in P\} \subset \mathbb{R}^n$ où $P \subset \mathbb{R}^n$. De plus, pour tout $i \in \mathbb{N}$, $diam(A_i) = diam(b + A_i)$ et il y a une correspondance bijective entre les recouvrements de A et ceux de $b + A$ donnée par $\{A_i\}_{i\in\mathbb{N}} \mapsto \{b + A_i\}_{i\in\mathbb{N}}$, d'où l'invariance par translation de la mesure de Hausdorff en passant à la limite quand δ tend vers 0.

2) Si $s \in [0, +\infty]$, $\delta > 0$, $A \subset \mathbb{R}^n$ et $\lambda \in \mathbb{R}$, on a pour tout δ-recouvrement $\{A_i\}_{i\in\mathbb{N}}$ de A que $\{\lambda A_i\}_{i\in\mathbb{N}}$ est un $\lambda\delta$-recouvrement de λA, donc

$$\begin{aligned} \mathcal{H}^s_{\lambda\delta}(\lambda A) &\leqslant \sum_{i\in\mathbb{N}} diam(\lambda A_i)^s \\ &= \lambda^s \sum_{i\in\mathbb{N}} diam(A_i)^s \end{aligned}$$

En passant à l'infimum sur tous les recouvrements, $\mathcal{H}^s_{\lambda\delta}(\lambda A) \leqslant \lambda^s \mathcal{H}^s_\delta(A)$. En faisant tendre δ vers 0, on a ce que l'on voulait. Réciproquement, on peut faire le même raisonnement dans l'autre sens et trouver que $\mathcal{H}^s_\delta(A) \leqslant \lambda^{-s} \mathcal{H}^s_{\lambda\delta}(\lambda A)$, d'où la conclusion désirée.

On peut alors déduire ce que représente la mesure de Hausdorff dans les cas où $s = 0$ et $s = 1$.

Proposition A.7. *i) La mesure $\mathcal{H}^0$ est la mesure de comptage.*

ii) Sur $\mathbb{R}$, la mesure $\mathcal{H}^1$ est la mesure de Lebesgue.

Démonstration. i) Soit A une partie finie de X et notons $\delta_0 = \min\{|x - y| \mid x, y \in A, x \neq y\} > 0$. On veut montrer que pour tout $\delta \in]0, \delta_0[$, $\mathcal{H}^0_\delta(A) = \text{card}(A)$. En effet, comme $A \subset \bigcup_{x\in A}\{x\}$, alors $\mathcal{H}^0_\delta(A) \leqslant \sum_{x\in A} 1 = \text{card}(A)$. Inversement, si $\{A_i\}_{i\in\mathbb{N}}$ est un δ-recouvrement de A, alors comme $\delta < \delta_0$, on a $\text{card}(I) \geqslant \text{card}(A)$ et en prenant l'infimum sur de tels recouvrements, $\mathcal{H}^0_\delta(A) \geqslant \text{card}(A)$. On en déduit l'égalité voulue. Ceci est encore vrai pour un ensemble A infini en utilisant la croissance de $\mathcal{H}^0$.

ii) On commence par découper

$$0 < \frac{1}{n} < \frac{2}{n} < \cdots < \frac{n-1}{n} < 1, \qquad n \in \mathbb{N}^*$$

et on a donc
$$\mathcal{H}^1_{\frac{1}{n}}([0,1]) \geqslant \sum_{i=1}^{n} \left(diam \left[\frac{i-1}{n}, \frac{i}{n} \right] \right)^1 = 1$$
En prenant la limite quand $n \to \infty$, on obtient $\mathcal{H}^1([0,1]) \leqslant 1$.
Soient maintenant $\delta > 0$ et $\varepsilon > 0$. Considérons $\{A_i\}_{i\in\mathbb{N}}$ un δ-recouvrement de $[0,1]$ et tel que
$$\left| H^1_\delta([0,1]) - \sum_{i\in\mathbb{N}} diam(A_i) \right| \leqslant \varepsilon$$
Sans perte de généralité, on peut supposer que les A_i sont des intervalles ouverts. Il existe une famille $\{a_k\}_{0\leqslant k\leqslant p}$, avec $0 = a_0 < a_1 < \cdots < a_p = 1$, telle que $[a_k, a_{k+1}]$ n'appartient qu'à un seul A_i pour tout $k = 0, \dots, p-1$. Par conséquent,
$$\mathcal{H}^1_\delta([0,1]) + \varepsilon \geqslant \sum_{k=0}^{p-1}$$
□

Comme $\mathcal{H}^1([0,1]) = 1$ et $\mathcal{H}^1$ est invariante par translation, on a le résultat recherché d'après la proposition 3.14 (page 49).

Introduisons maintenant la notion de dimension de Hausdorff.

Proposition A.8. *À E fixé, l'application $s \mapsto \mathcal{H}^s(E)$ est décroissante. De plus, si $s, t \in [0, +\infty]$ avec $s < t$, alors*
$$\mathcal{H}^s_\delta(E) \geqslant \delta^{s-t}\mathcal{H}^t_\delta(E)$$
En particulier, il existe un unique $s \in [0, +\infty]$ tel que $\mathcal{H}^t(E) = 0$ si $t > s$ et $\mathcal{H}^t(E) = +\infty$ si $0 \leqslant t < s$.

Définition A.9. Un tel s est appelé **la dimension de Hausdorff de** E.

Démonstration. 1) L'application $s \mapsto r^s$ est décroissante pour tout $r \in]0,1[$. On fixe E une partie de X. Soient $\delta \in]0,1[$, et $\{E_n\}_{n\in\mathbb{N}}$ un δ-recouvrement de E. En particulier, $diam(E_n) \in]0,1[$ pour tout $n \in \mathbb{N}$, donc on en déduit que $s \mapsto \mathcal{H}^s_\delta(E)$ est décroissante. En faisant tendre δ vers 0, $s \mapsto \mathcal{H}^s(E)$ est décroissante à E fixé.

2) Soit E une partie de X, et $\{E_n\}$ un δ-recouvrement de E. On a
$$\begin{aligned}
\mathcal{H}^t_\delta(E) &\leqslant \sum_{n\in\mathbb{N}} diam(E_n)^t \\
&\leqslant \delta^{t-s} \sum_{n\in\mathbb{N}} \left(\frac{diam(E_n)^t}{\delta^{t-s}} \right) \\
&\leqslant \delta^{t-s} \sum_{n\in\mathbb{N}} diam(E_n)^s \underbrace{\left(\frac{diam(E_n)}{\delta} \right)^{t-s}}_{\leqslant 1} \\
&\leqslant \delta^{t-s} \sum_{n\in\mathbb{N}} diam(E_n)^s
\end{aligned}$$

Finalement, en passant à l'infimum sur tous les δ-recouvrements, on déduit que

$$\forall s < t, \qquad \mathcal{H}^s_\delta(E) \geqslant \delta^{s-t}\mathcal{H}^t_\delta(E) \tag{A.5}$$

Par (A.5) :
– Si $\exists s < \infty$ tel que $\mathcal{H}^s(E) < \infty$, alors $\forall t > s$, on a $\mathcal{H}^t(E) = 0$.
– Si $\exists t > 0$ tel que $\mathcal{H}^t(E) > 0$, alors $\forall s < t$, on a $\mathcal{H}^s(E) = +\infty$.
Ceci assure l'existence et l'unicité de $s \in [0, +\infty]$ tel que $/H^t(E) = 0$ si $t > s$ et $\mathcal{H}^t(E) = +\infty$ si $0 \leqslant t < s$.

A.4 Exemples de mesures et de dimension de Hausdorff

On considère à présent $X = \mathbb{R}^n$, $n \geqslant 1$, muni de la distance d défini par :

$$d(x, y) = \max_{i=1,\ldots,n} |x_i - y_i|, \qquad \forall x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{R}^n$$

Proposition A.10. *Dans ce cadre, la dimension de Hausdorff de* $[0, 1[^n$ *est* n.

Remarque. C'est vrai aussi pour la distance euclidienne.

Démonstration. Notons $C_n = [0, 1[^n$ l'hypercube unité de $\mathbb{R}^n$. Pour tout $k \geqslant 1$, C_n s'écrit comme union disjointe de 2^{kn} pavés qui sont des translatés de $\left[0, 2^{-k}\right[^n$. Notons $C_{n,k} = \left[0, 2^{-k}\right[^n$.
On a donc que $\mathcal{H}^s(C_{n,k}) = 2^{-ks}\mathcal{H}^s(C_n)$, d'où par additivité-dénombrable et comme $C_n, C_{n,k}$ sont des parties $\mathcal{H}^s$-régulières :

$$\mathcal{H}^s(C_n) = 2^{nk}2^{-ks}\mathcal{H}^s(C_n)$$

Finalement, si $\mathcal{H}^s(C_n) \notin \{0, +\infty\}$, alors $s = n$ $(*)$.

Si $s > n$, alors $\mathcal{H}^s(C_n)(1 - 2^{nk}2^{-ks}) = \mathcal{H}^s(C_n)(1 - 2^{\overbrace{(n-s)}^{<0}k}) = 0$, et, en passant à la limite lorsque k tend vers l'infini, on trouve que $\mathcal{H}^s(C_n) = 0$.
Si $s < n$, pour tout recouvrement de C_n par des ensembles B_k de demi-diamètre $r_k \leqslant \frac{1}{2}$, on peut inclure B_k dans une boule euclidienne de rayon $2r_k \leqslant 1$, et $\lambda_n(B_k) \leqslant \omega_n(2r_k)^n$, où λ_n est la mesure de Lebesgue dans $\mathbb{R}^n$, et ω_n est le volume de la boule unité dans $\mathbb{R}^n$. D'où

$$1 = \lambda_n(C_n) \leqslant \sum_{k\in\mathbb{N}} \lambda_n(B_k) \leqslant \sum_{k\in\mathbb{N}} \omega_n(2r_k)^n \leqslant \omega_n \sum_{k\in\mathbb{N}} diam(B_k)^s$$

et en passant à l'infimum on voit que $\mathcal{H}^s(C_n) \geqslant 1/\omega_n > 0$, donc $\mathcal{H}^s(C_n) = +\infty$ par $(*)$.
Finalement, la dimension de Hausdorff de C_n est n. □

Application aux ensembles de Cantor

Jusqu'à présent, nous n'avons traité le cas où s n'est pas entier qu'implicitement. Mais il serait intéressant de trouver un sous-ensemble de $\mathbb{R}^n$ de dimension de Hausdorff non entière, et c'est le cas de l'ensemble de Cantor classique (i.e. le triadique). Bien entendu, d'autres ensembles de Cantor donneront des dimensions différentes !

Théorème A.11. *La dimension de Hausdorff de l'ensemble triadique de Cantor K est $s = \dfrac{\ln 2}{\ln 3}$ et $\mathcal{H}^{\frac{\ln 2}{\ln 3}}(K) = 1$*

Démonstration. Soient $k \in \mathbb{N}$ et $\varepsilon = 3^{-k}$. On peut recouvrir l'ensemble de Cantor K par 2^k segments (ceux qui interviennent dans la construction du Cantor lors de la k-ième itération) $\{I_{i,k}\}_{1 \leqslant i \leqslant 2^k}$ de longueur ε. D'où,

$$\mathcal{H}^s(K) \leqslant \sum_{i=1}^{2^k} diam(I_{i,k})^s = \sum_{i=1}^{2^k} \varepsilon^s = \sum_{i=1}^{2^k} \exp\left(-k\ln(3) \times \frac{\ln 2}{\ln 3}\right) = \sum_{i=1}^{2^k} 2^{-k} = 1$$

Cet argument montre déjà que :
- $\mathcal{H}^s(K) = 0$ si $s > \frac{\ln 2}{\ln 3}$.
- $\mathcal{H}^s(K) \leqslant 1$ si $s = \frac{\ln 2}{\ln 3}$

Soit maintenant $s \leqslant \frac{\ln 2}{\ln 3}$. On écrit $K = \bigcap_{n\in\mathbb{N}} K_n$ (on rappelle que K_n est constitué de 2^n segments de longueur 3^{-n}) et on remarque que si $\delta = 3^{-N}$, alors $\mathcal{H}^s_\delta(K) = \mathcal{H}^s_\delta(K_N) = 2^N 3^{-sN}$. La première égalité est évidemment le point qu'il faut vérifier, et qui est laissé au lecteur. Ainsi, si $s \leqslant \frac{\ln 2}{\ln 3}$,

$$\mathcal{H}^s(K) = \lim_{N\to\infty} \left(\frac{2}{3^s}\right)^N = \begin{cases} +\infty & \text{si } s < \dfrac{\ln 2}{\ln 3} \\ 1 & \text{si } s = \dfrac{\ln 2}{\ln 3} \end{cases}$$

□