**Bibliographie :**
P. Malliavin : Intégration et probabilité.
W.Rudin : Analyse réelle et complexe.
J. Faraut : Calcul intégral (espace 34)
Barbe-Ledoux : Probabilités (Espace 34)

# 1 Introduction

Pourquoi une nouvelle intégrale? En L1 et L2, nous avons vu la théorie de l'intégrale de Riemann, qui permet de définir l'intégrale d'une fonction continue ou monotone par morceaux définie sur un intervalle. Cette intégrale s'étend sans problème aux fonctions définies sur $\mathbb{R}^p$. Mais elle ne permet pas de répondre à un certain nombres de questions. C'est le but de l'intégrale de Lebesgue (développée en autre par Johannes Stieltjes, Henri Lebesgue, Emile Borel et Andrei Kolmogorov dans le cadre de la théorie des probabilités).

- Certaines fonctions comme $\mathbf{1}_{\mathbb{Q}\cap[0,1]}$ sont non intégrables pour la théorie de Riemann, mais vont l'être dans la théorie de Lebesgue (avec une intégrale nulle).
- La théorie de Lebesgue permet de mesurer des ensembles très irréguliers, comme le tryadique de Cantor, ou le tapis de Sierpinski.
- Elle permet d'intégrer des fonctions définies sur des ensembles autres que $\mathbb{R}$ ou $\mathbb{R}^n$, comme les sphères.
- Elle permet d'inclure le calcul des probabilités et l'axiomatique de Kolmogorov dans le cadre de la théorie générale de l'intégration.
- Elle permet de mettre dans une même théorie la sommation des séries et le calcul intégral, et donc en particulier dans le cadre de la théorie des probabilités, de n'avoir qu'un seul formalisme pour les variables aléatoires discrètes et les variables aléatoires continues.
- Elle permet d'obtenir des théorèmes de convergence beaucoup plus puissants et robustes que dans la théorie de Riemann.

**Plan du cours**

1. Espaces mesurables, fonctions mesurables, mesures.
2. Intégrale des fonctions mesurables. Théorèmes fondamentaux
3. Produit d'espaces mesurables, théorèmes de Fubini et applications.
4. Espaces $L^p$, dualité, eléments d'analyse.
5. Transformation de Fourier.

**Rappel de théorie des ensembles.**

- Les opérations fondamentales de la théorie des ensembles sont : union, intersection, complémentaire, différence symétrique, produit d'ensembles.

- Indicatrice d'un ensemble : $\mathbf{1}_A(x)$ est la fonction qui vaut 1 sur $A$ et 0 sur $A^c$. Cette notation permet de faire des calculs rapides sur les ensembles. Par exemple,

$$\mathbf{1}_{A\cap B} = \mathbf{1}_A \mathbf{1}_B,\ \mathbf{1}_{A^c} = 1 - \mathbf{1}_A,\ \mathbf{1}_{A\Delta B} = |\mathbf{1}_A - \mathbf{1}_B|\,.$$

- Union d'une famille quelconque d'ensembles

$$\bigcup_{i\in I} A_i = \{x \ \mid\ \forall i \in I, x \in A_i\}.$$

On définit de même l'intersection.

**Ensembles dénombrables.** Un ensemble dénombrable est un ensemble en bijection avec $\mathbb{N}$. Un ensemble est dit au plus dénombrable s'il est fini ou dénombrable.

- Exemples : $\mathbb{R}$ n'est pas dénombrable, non plus qu'aucun intervalle non vide de $\mathbb{R}$. Par contre, $\mathbb{Q}$, les nombres dyadiques, les décimaux, etc., sont dénombrables.
- Une union finie ou dénombrable d'ensembles dénombrables est dénombrable.
- Un produit fini d'ensembles dénombrables est dénombrable. C'est faux pour un produit dénombrable, même d'ensembles finis.
- S'il existe une injection de $E$ dans $\mathbb{N}$, alors $E$ est au plus dénombrable.
- S'il existe une surjection de $\mathbb{N}$ sur $E$, alors $E$ est au plus dénombrable.

# 2 Espaces et fonctions mesurables, mesures.

## 2.1 Ensembles mesurables, tribus

Dans tout ce qui suit, on considère un ensemble $\Omega$. On va chercher à "mesurer" des sous-ensembles de $\Omega$. La structure de la famille d'ensembles qu'on va mesurer est donnée par la notion de $\sigma$-algèbre, ou tribu.

**Définition.** 1. **Algèbres de Boole.** *C'est un sous ensemble de $\mathcal{P}(\Omega)$ qui est stable par complémentaire, intersection, et qui contient l'ensemble vide.*

*Il est alors stable par toutes les opérations finies de la théorie des ensembles.*

Exemples d'algèbre de Boole :

1. $\mathcal{P}(\Omega)$.
2. Les réunions finies d'intervalles (sur $\mathbb{R}$).
3. La plus petite algèbre de Boole qui contient l'ensemble $A$ $(\Omega, A, A^c, \emptyset)$.
4. De façon générale, la plus petite algèbre de Boole qui contient une famille donnée $\mathcal{E}$ d'ensembles. On la note $\mathcal{A}(\mathcal{E})$. Remarquer que si $\mathcal{E}$ est finie, alors $\mathcal{A}(\mathcal{E})$ est aussi finie. Le démontrer à titre d'exercice.

**Définition.** 2. *Une* $\sigma$**-algèbre** $\mathcal{A}$ *(ou* **tribu***) est une algèbre stable par union dénombrable : si* $(A_n)_{n\in\mathbb{N}} \in \mathcal{A}$*, alors* $\bigcup_n A_n \in \mathcal{A}$*.*

**Remarque 1.** *Si* $\mathcal{A}$ *est une algèbre de Boole, pour qu'elle soit une tribu, il est équivalent de demander (au choix) que*

1. $\mathcal{A}$ *soit stable par union croissante dénombrable.*
2. $\mathcal{A}$ *soit stable par intersection dénombrable décroissante.*

*Conséquence : une algèbre de Boole finie est une* $\sigma$*-algèbre.*

**Définition.** 3. *Un espace mesurable* $(\Omega, \mathcal{A})$ *est un ensemble* $\Omega$ *muni d'une tribu* $\mathcal{A}$ *de parties de* $\Omega$*.*

Une intersection quelconque d'algèbres est une algèbre. Ce n'est pas le cas d'une réunion. Il en va de même d'une intersection quelconque de $\sigma$-algèbres.

On peut ainsi définir la plus petite algèbre de Boole qui contient $\mathcal{E}$ : $\mathcal{A}(\mathcal{E})$ : c'est l'intersection de toutes les algèbres de Boole qui contiennent $\mathcal{E}$, et la plus petite $\sigma$-algèbre (tribu) qui contienne $\mathcal{E}$. On la note $\sigma(\mathcal{E})$

**Remarque 2.** *Une partition de* $\Omega$ *est une famille de parties* $S_i$ *telle que* $S_i \cap S_j = \emptyset$ *si* $i \neq j$ *et telle que* $\cup_i S_i = \Omega$*.*

*Pour une partition finie de* $\Omega$*, soit* $\mathcal{E} = (S_1, \dots, S_n)$*, on peut entièrement décrire* $\mathcal{A}(\mathcal{E})$*. Elle contient* $2^n$ *éléments. C'est la famille de toutes les réunions finies d'éléments de* $\mathcal{E}$*. Cette tribu est alors finie.*

*Pour une partition dénombrable* $\mathcal{E}$ *de* $\Omega$*,* $\sigma(\mathcal{E})$ *est formée de toutes les réunions de sous-familles de* $\mathcal{E}$*. Elle n'est alors pas dénombrable, car l'ensemble des sous-ensemble d'un ensemble dénombrable n'est pas dénombrable.*

**Tribu induite** Si on a une tribu $\mathcal{A}$ et que $C \in \mathcal{A}$, l'ensemble des $B \in \mathcal{A}$ tels que $B \subset C$ est une tribu sur $C$. On la note $\mathcal{A}_C$

**Définition.** 4. *La tribu des* **Boréliens** *de* $\mathbb{R}$ *est la plus petite tribu qui contienne les intervalles.*

**Proposition 1.** *La tribu des boréliens est*

1. $\sigma([a,b],\ a<b)$.
2. $\sigma(]a,b[,\ a<b)$.
3. $\sigma(]-\infty,a[,\ a\in\mathbb{R})$.
4. $\sigma([a,b],\ a<b,\ a,b\in\mathbb{Q})$.
5. $\sigma(]a,b[, a<b,\ a,b\in\mathbb{Q})$.
6. $\sigma(]-\infty,a[,\ a\in\mathbb{Q})$.
7. *etc.*

**Remarque 3.** *Les réunions finies d'intervalles (de toutes les formes, bornés ou non) et une algèbre de Boole. Ce n'est pas une* $\sigma$*-algèbre. La famille des réunions dénombrables d'intervalles n'est pas une* $\sigma$*-algèbre.*

D'autres définitions de la tribu borélienne de $\mathbb{R}$.

**Proposition 2.** *La tribu borélienne est la plus petite tribu qui contienne les ouverts de* $\mathbb{R}$.

*C'est aussi la plus petite tribu qui contienne les fermés.*

*Démonstration.* — Ceci repose sur le fait élémentaire suivant : tout ouvert est une réunion dénombrable d'intervalles ouverts

On écrit, pour un ouvert $O$, $O = \cup_{I \in J} I$, où $J$ est l'ensemble des intervalles $]a, b[$, $a < b$, $a, b \in \mathbb{Q}$, $]a, b[ \subset O$.

On peut aussi démontrer qu'un ouvert est une réunion dénombrable d'intervalles ouverts disjoints, mais c'est plus difficile. ∎

Les boréliens (de même que toutes les familles d'ensembles mesurables en général) sont impossible à décrire par leurs propriétés (comme les ouverts ou les fermés). Il est d'ailleurs assez difficile d'exhiber un ensemble non borélien.

**Remarque 4.** *Pourquoi se limite-t-on à des unions dénombrables? Si l'on veut que les points soient dans une tribu, et si on acceptait les unions non dénombrables, alors toutes les parties de $\mathbb{R}$ seraient mesurables : ceci aboutirait à une contradiction, car on verra plus loin qu'il est illusoire de vouloir mesurer toutes les parties de $\mathbb{R}$ : certaines d'entre elles sont si monstrueuses qu'aucune définition raisonnable de leur longueur ne peut avoir de sens.*

Quand on dispose d'une tribu sur un ensemble $\Omega$, on en dispose aussi d'une sur toutes les parties de $\Omega$. En pratique, on se limitera aux sous-ensembles mesurables (c'est à dire appartenant à la tribu) de $\Omega$.

**Définition.** 5 (**Tribu induite**). *Soit $(\Omega, \mathcal{A})$ un ensemble mesurable, et $\Omega_0 \subset \Omega$ un élément de $\mathcal{A}$. Alors,*

$$\mathcal{A}_{\Omega_0} = \{A \in \mathcal{A},\ B \subset \Omega_0\}$$

*est une tribu sur $\Omega_0$, qu'o appelle la tribu induite de $\Omega$ sur $\Omega_0$.*

Pour démontrer qu'une propriété est vraie pour tous les boréliens, on démontrera en général qu'elle est vraie pour des boréliens simples (comme les ouverts), et on l'étendra à tous les boréliens à l'aide de la proposition suivante, qui est fondamentale.

**Théorème 1** (**Théorème des classes monotones, version ensembliste).** *Une classe $\mathcal{M}$ de parties est une classe monotone si*

1. $\Omega \in \mathcal{M}$.
2. *Si* $A, B \in \mathcal{M}$, $A \subset B$ *alors* $B \setminus A \in \mathcal{M}$.
3. *Si* $A_n$ *est une suite croissante d'élements de* $\mathcal{M}$, *alors* $\cup_n A_n \in \mathcal{M}$.

*Si $\mathcal{E}$ et une famille de parties, on note $\mathcal{M}(\mathcal{E})$ la plus petite classe monotone qui la contienne.*

*Alors, si $\mathcal{E}$ est stable par intersection finie, alors $\mathcal{M}(\mathcal{E}) = \sigma(\mathcal{E})$.*

Nous verrons plus bas en quoi ce théorème est un outil très puissant. La condition fondamentale est bien sûr la stabilité de $\mathcal{E}$ par intersection (par exemple la classe des intervalles, ou des intervalles fermés bornés, etc..).

*Démonstration.* — Pour démontrer, on commence par remarquer qu'une tribu est une classe monotone, et donc que $\mathcal{M}(\mathcal{E}) \subset \sigma(\mathcal{E})$.

On remarque aussi qu'une classe monotone stable par intersection est une tribu (car elle est déjà stable par complémentaire).

On va donc montrer que $\mathcal{M}(\mathcal{E})$ est stable par intersection.

On définit

$$\mathcal{M}_1 = \{B \in \mathcal{M}(\mathcal{E}),\ B \cap E \in \mathcal{M}(\mathcal{E}), \forall E \in \mathcal{E}\}$$

et

$$\mathcal{M}_2 = \{B \in \mathcal{M}(\mathcal{E}),\ B \cap E \in \mathcal{M}(\mathcal{E}), \forall E \in \mathcal{M}(\mathcal{E})\}$$

.

Alors, $\mathcal{M}_1$ est une classe monotone, et $\mathcal{M}_2$ aussi.

Puisque $\mathcal{M}_1$ est une classe monotone est que $\mathcal{E}$ est stable par intersection, $\mathcal{M}_1$ contient $\mathcal{E}$, donc $\mathcal{M}(\mathcal{E})$.

Puisque $\mathcal{M}_1$ contient $\mathcal{M}(\mathcal{E})$, alors $\mathcal{M}_2$ contient $\mathcal{E}$ et donc $\mathcal{M}(\mathcal{E})$. C'est exactement ce qu'on voulait démontrer.

∎

## 2.2 Mesures, fonctions additives d'ensembles.

Avant de parler de mesures, nous commençons par un lemme fondamental et facile, sur lequel repose toute la théorie de la mesure.

**Lemme 1 (Lemme fondamental).** *Soit $a_{n,p}$ une double suite de réels positifs, qui est croissante en $n$ et en $p$.*

*Alors, $\lim_n \lim_p a_{n,p} = \lim_p \lim_n a_{n,p}$.*

*De plus, si $n_k$ et $p_k$ sont des sous-suites qui convergent vers l'infini, alors on a encore*

$$\lim_k a_{n_k,p_k} = \lim_n \lim_p a_{n,p}.$$

Sa démonstration est élémentaire et laissée à titre d'exercice. Attirons l'attention sur le fait qu'ici les limites sont à considérer au sens de $[0,\infty]$.

Remarquons tout de même qu'on peut toujours se ramener au cas des suites bornées, quitte à choisir une fonction strictement $f$ croissante de $\mathbb{R}_+$ à valeurs dans un intervalle borné (par exemple arctang, et à changer $a_{n,p}$ en $b_{n,p} = f(a_{n,p})$.

**Définition.** 6. *Sur une algèbre de Boole $\mathcal{A}$, une* **fonction additive d'ensembles** *est une application $\mathcal{A} \to [0,\infty]$ telle que*

1. $\mu(\emptyset) = 0$
2. $\mu(A \cup B) = \mu(A) + \mu(B)$ *si* $A \cap B = \emptyset$.

Exemples :

1. Sur les parties de $\mathbb{R}$ ou de $\mathbb{N}$, le cardinal de $A$.
2. $\mu(A) = 0$ si $A$ est fini, et $\mu(A) = \infty$ si $A$ est de cardinal infini.

**Remarque 5.** *Pour une fonction additive d'ensembles, si $A \subset B$, alors $\mu(A) \leq \mu(B)$. De même, pour tout couple $(A,B)$ d'éléménts de $\mathcal{A}$, non nécésairement disjoints, on a*

$$\mu(A \cup B) \leq \mu(A) + \mu(B).$$

Cette définition est insuffisante. On veut un peu plus : une propriété de continuité. On appelera cela une mesure.

**Définition.** 7 (**Mesures**). *Soit $(\Omega, \mathcal{A})$ un espace mesurable ( c'est à dire un ensemble muni d'une $\sigma$-algèbre). Une application $\mu : \mathcal{A} \to [0, \infty]$ est une mesure si*

1. *C'est une fonction additive d'ensembles.*
2. *Si $A_n \in \mathcal{A}$, et $A_n$ est une suite croissante d'ensembles, avec $A = \cup_n A_n$, alors*
$$\mu(A) = \lim_n \mu(A_n).$$

*Un espace mesuré est la donnée d'un triplet $(\Omega, \mathcal{A}, \mu)$, où $\Omega$ est un ensemble, $\mathcal{A}$ est une tribu de parties de $\Omega$, et $\mu$ une mesure sur $\mathcal{A}$.*

Insistons sur le fait que la limite dans la définition précédente peut être finie ou infinie. Mais si $\Omega$ est de mesure finie ($\mu(\Omega) < \infty$), alors pour tout $A \in \mathcal{A}$, $\mu(A) < \infty$ et la limite est toujours finie.

**Remarque 6.** *Il est équivalent de demander, en plus d'être une fonction additive d'ensembles, de vérifier*

*Si $A_n$ est une suite disjointe d'élements de $\mathcal{A}$, et $A = \cup_n A_n$, alors*

$$\mu(A) = \sum_n \mu(A_n).$$

*Ou bien encore*

*$\mu$ est une application de $\mathcal{A}$ dans $[0, \infty]$ telle que $\mu(\emptyset) = 0$ et, pour toute partition $(A_n)$ de $A$ (composée éventuellement d'éléments vides),*

$$\mu(A) = \sum_n \mu(A_n).$$

*De même, lorsque $\mu(\Omega)$ est fini, il est équivalent de demander que, si $A_n$ est une suite décroissante et $A = \cap_n A_n$, alors $\mu(A) = \lim_n \mu(A_n)$. Il suffit pour cela de changer $A_n$ en $A_n^c$.*

Une fois de plus, insistons sur le fait que la somme peut être finie ou infinie. Remarquons aussi que cette série étant à termes positifs, la somme ne dépend pas de l'ordre de sommation. Cela est cohérent avec le fait que $A = \cup_n A_n$ est indépendant de l'ordre dans lequel nous avons énuméré les ensembles $A_n$.

Si $\mu$ est une mesure, alors pour toute suite décroissante $A_n$ telle que $\mu(A_0) < \infty$, $\mu(A_n)$ converge vers $\mu(\cap_n A_n)$.

Pour le voir, on applique ce qui précède à $A_0 \setminus A_n$.

Mais ceci peut être faux si $\mu(\Omega) = \infty$ (voir l'exemple plus bas de la mesure de décompte sur $\mathbb{N}$).

**Définition.** 8 (**Masse**). *$\mu(\Omega)$ s'appelle la masse de la mesure. Une mesure de masse 1 est une* **probabilité**.

L'exemple le plus simple de mesure est la masse de Dirac. Soit $x \in \Omega$. On définit
$$\delta_x(A) = \mathbf{1}_A(x).$$

C'est une mesure, comme on le vérifie immédiatement.

Donnons un autre exemple élémentaire. Sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$, pose $\mu(A) = \sharp A$ (c'est une notation pour désigner le cardinal (ou encore le nombre de points) de $A$). Alors $\mu$ est une mesure. On le vérifie aussi immédiatement. Cette mesure s'appelle la mesure de décompte sur $\mathbb{N}$. Sur cet exemple, si nous prenons $A_n = [n, \infty)$, alors nous obtenons une suite décroissante d'ensembles dont l'intersection est vide, mais $\mu(A_n) = \infty$ ne converge pas vers $\mu(\cap_n A_n) = \mu(\emptyset) = 0$.

Il est souvent plus facile de travailler avec des mesures finies. On s'y ramène souvent à condition que la mesure ait une propriété élémentaire d'éétre $\sigma$-finie.

**Définition.** 9 (**Mesures $\sigma$-finies**). *On dit que la mesure $\mu$ est $\sigma$-finie s'il existe une suite croissante $B_n$ d'éléments de $\mathcal{A}$ telle que $\cup_n B_n = \Omega$ et $\mu(B_n) < \infty$.*

La plupart des mesures intéressantes (et toutes celles rencontrées dans ce cours) seront $\sigma$-finies. Mais par exemple, la mesure, définie sur toutes les parties de $\mathbb{R}$ par $\mu(A) = 0$ si $A$ est au plus dénombrable, et $\mu(A) = \infty$ sinon, est une mesure, qui n'est pas $\sigma$-finie.

Une des conséquences du théorème des classes monotones est la

**Proposition 3.** *Soit $(\Omega, \mathcal{A})$ un ensemble mesurable et $\mathcal{E}$ une famille de parties stable par intersection finie, telle que $\sigma(\mathcal{E}) = \mathcal{A}$. Alors, si $\mu_1$ et $\mu_2$ sont deux mesures finies qui coïncident sur $\mathcal{E}$ et qui ont même masse, elles sont égales.*

*Démonstration.* — On vérifie immédiatement que l'ensemble des éléments de $A \in \mathcal{A}$ qui vérifie $\mu_1(A) = \mu_2(A)$ est une classe monotone.

∎

**Remarque 7.** *Insistons sur le fait qu'il faut que les mesures soient finies. Sinon, par exemple, si $\mu_1$ est la longueur d'un intervalle (dont on verra plus tard qu'elle se prolonge en une mesure) et si $\mu_2 = 2\mu_1$, $\mu_1$ et $\mu_2$ coincident sur les intervalles $[a, \infty[$, mais ne coïncident pas partout.*

Parmi les opérations qu'on peut faire sur les mesures, les principales sont

**Proposition 4.**

1. *Si $\mu$ est une mesure et $\lambda \in [0, \infty)$, alors $\lambda\mu$ est une mesure.*
2. *Si $\mu_1$ et $\mu_2$ sont deux mesures, sur $(\Omega, \mathcal{A})$ alors il en est de même de $\mu_1 + \mu_2$.*
3. *Si $\mu_n$ est une suite croissante de mesures, et si l'on définit $\mu(A) = \lim_n \mu_n(A)$, alors $\mu$ est une mesure.*
4. *(Restriction) Si $A \in \mathcal{A}$, alors $\mu_A(B) = \mu(B \cap A)$ définit une mesure sur la tribu induite $\mathcal{A}_A$.*
5. *(Extension) Si $A \in \mathcal{A}$ et que $\mu_A$ est une mesure sur $(A, \mathcal{A}_A)$ où $\mathcal{A}_A$ est la tribu induite, alors la formule $\mu(B) = \mu_A(A \cap B)$ définit une mesure sur $(\Omega, \mathcal{A})$. La restriction de l'extension est la mesure de départ, l'inverse n'est pas vrai. L'extension de la restriction de $\mu$ à $A$ est la mesure $\mu_A(B) = \mu(A \cap B)$.*

   *Ainsi, on peut identifier toute mesure définie sur un sous-ensemble mesurable $A$ à une mesure sur l'espace tout entier.*

Attention, dans le point 3, on autorise bien évidement la limite à avoir lieu dans $\bar{\mathbb{R}}_+ = [0, \infty]$.

Les deux premiers points sont triviaux, le dernier vient de du lemme fondamental 1.

*Démonstration.* — (Du point 3 de la proposition).

Il est évident que $\mu = \lim_n \mu_n$ est une fonction additive d'ensembles. Il n'y a qu'à montrer que c'est une mesure. Pour cela, on considère une suite croissante d'ensembles $A_p$, avec $A = \cup_p A_p$ et $a_{n,p} = \mu_n(A_p)$.

Alors, d'après le lemme fondamental (Lemme 1)

$$\mu(A) = \lim_n(\mu_n(A) = \lim_n \lim_p(\mu_n(A_p) = \lim_p \lim_n \mu_n(A_p) = \lim_p \mu(A_p).$$

La première et la dernière égalités proviennent de la définition de $\mu$ et la seconde du fait que $\mu_n$ est une mesure. ∎

Comme conséquence, on voit ainsi que si $\mu_n$ est une suite de mesures, alors $\sum_n \mu_n$ est encore une mesure.

Ainsi, la mesure de décompte sur $\mathbb{N}$ n'est rien d'autre que $\sum_n \delta_n$. C'est une mesure $\sigma$-finie.

**Remarque 8.** *Si $\mu$ est une mesure $\sigma$-finie sur l'espace mesurable $(\Omega, \mathcal{A})$, alors il existe une* **partition** *$(B_n)_{n\in\mathbb{N}}$ de $\Omega$ telle que $\mu(B_n) < \infty$, $\forall n$. Dans ce cas, si $\mu_n(B) = \mu(A \cap B_n)$, $\mu_n$ est une mesure finie sur $(\Omega, \mathcal{A})$, comme on l'a déjà vu. Alors (et puisque $(B_n)$ est une partition de $\Omega$), $\mu = \sum_n \mu_n$. Ainsi, toute mesure $\sigma$-finie est une somme (infinie) de mesures finies.*

Il n'est pas très facile de construire des mesures non triviales dans le cas non dénombrable. Mais le théorème des classes monotones nous assure de l'unicité d'une mesure sous des conditions élémentaires.

Pour construire une mesure, on part en général d'une fonction additive d'ensembles sur une algèbre $\mathcal{A}_0$, et on la prolonge à la tribu $\sigma(\mathcal{A}_0)$. On se sert pour cela du théorème suivant.

**Théorème 2 (Carathéodory).** *Soit $\mu$ une fonction additive d'ensembles finie sur une algèbre $\mathcal{A}_0$. Pour que $\mu$ se prolonge en une mesure sur $\mathcal{A} = \sigma(\mathcal{A}_0)$, il faut et il suffit que, pour toute suite décroissante $A_n$ d'éléments de $\mathcal{A}_0$ telle que $\cap_n A_n = \emptyset$, alors $\mu(A_n) \to 0$.*

*Dans ce cas, le prolongement de $\mu$ de $\mathcal{A}_0$ à $\mathcal{A}$ est unique.*

Nous n'allons pas démontrer ce théorème. La condition est évidement nécessaire. L'unicité provient du théorème précédent. Quand à l'existence, ce n'est pas très difficile mais assez long : le principe est le suivant (nous laissons à titre d'exercice la vérification des détails au lecteur) :

On commence par étendre la fonction additive d'ensembles aux unions dénombrables et aux intersections dénombrables d'éléments de $\mathcal{A}_0$, en utilisant des limites. Par exemple, si $A_n \in \mathcal{A}_0$, est $A_n$ croissante, pour $A = \cup_n A_n$, on pose $\mu(A) = \lim_n \mu(A_n)$. Ceci est possible justement à cause de la propriété de $\mu$ sur $\mathcal{A}_0$, qui montre que la limite ne dépend pas de la suite choisie.

On définit ensuite $\mu^*(A) = \inf\{\mu(B),\ A \subset B\}$, où $B$ varie dans la classe des réunions dénombrables d'éléments de $\mathcal{A}_0$. et $\mu_*(A) = \sup\{\mu(B),\ B \subset A\}$ où $B$ varie parmi les intersection dénombrables d'éléments de $\mathcal{A}_0$.

Alors, l'ensemble des éléments tels que $\mu_*(A) = \mu^*(A)$ est une $\sigma$-algèbre qui contient $\mathcal{A}_0$, et sur cette $\sigma$-algèbre $\mu_* = \mu^*$ définit bien une mesure. Puisque cette $\sigma$-algèbre contient $\mathcal{A}_0$, elle contient $\mathcal{A}$ et on a donc étendu la mesure à $\mathcal{A}$.

**Remarque 9.** *De façon générale, si $\mathcal{A}_0$ est une algèbre de Boole qui engendre une $\sigma$-algèbre $\mathcal{A}$ ($\mathcal{A} = \sigma(\mathcal{A}_0)$), et si $\mu$ est une mesure $\sigma$-finie sur $\mathcal{A}$, alors pour tout $A \in \mathcal{A}$ tel que $\mu(A) < \infty$, et pour tout $\epsilon > 0$, il existe un $B \in \mathcal{A}_0$ tel que $\mu(A\Delta B) \leq \epsilon$.*

*On le voit aisément en appliquant le théorème des classes monotones, ou bien plus simplement en remarquant que de tels ensembles $A$ forment une $\sigma$-algèbre qui contient $\mathcal{A}_0$.*

*Mais le raisonnement du théorème de Carathéodory est plus subtil, car il demande à pouvoir encadrer $A$ entre deux ensembles $A_1 \subset A \subset A_2$. On ne peut pas le faire avec des éléments de $\mathcal{A}_0$, mais on peut le faire avec respectivement des intersections dénombrables et des unions dénombrables d'éléments de $\mathcal{A}_0$.*

Tout ceci nous amène à la construction de la mesure de Lebesgue sur un intervalle.

**Théorème 3.** *Soit $I = [0,1]$, et $\mathcal{B}_I$ la classe des boréliens de $I$. Alors, il existe une unique mesure $\mu$ sur $\mathcal{B}_I$ telle que, pour tout intervalle $[a,b] \subset I$, $\mu([a,b]) = b - a$.*

Ce théorème est tout sauf trivial. On a bien sûr le même énoncé sur n'importe quel autre intervalle $[A, B]$.

*Démonstration.* — On commence par appeler $\mathcal{A}_0$ la famille formée des réunions finies de sous-intervalles de $I$, quelle que soit leur forme.

Une telle réunion peut être écrite de façon unique comme une réunion finie d'intervalles disjoints. C'est une algèbre de Boole, et pour un élément $J$ de cette famille $\mathcal{A}_0$, on définit sa mesure le Lebesgue $\lambda(J)$ comme la somme des longueurs des intervalles disjoints qui le composent.

C'est évidement une fonction additive d'ensembles (relation de Chasles). Il faut encore démontrer qu'on peut appliquer le théorème de Carathéodory.

Pour cela, on utilise la propriété suivante :

Pour tout $J \in \mathcal{A}_0$ et tout $\epsilon > 0$, il existe un élément compact $K \in \mathcal{A}_0$ tel que

1. $K \subset J$.
2. $\lambda(J \setminus K) \leq \epsilon$.

Les éléments compacts que nous considérons sont ici les réunions finies d'intervalles fermés. l'existence d'un tel $K$ est alors évidente (ou presque) si l'on raisonne séparément sur chacun intervalle disjoint qui compose $J$.

La propriété fondamentale des compacts que nous utilisons est que si $(K_n)$ est une suite décroissante de compacts et que $\cap_n K_n = \emptyset$, alors il existe un $n_0$ pour lequel $K_{n_0} = \emptyset$.

Ensuite, on considère une suite $J_n$ d'éléments de $\mathcal{A}_0$ qui décroit vers $\emptyset$. On veut démontrer que $\lambda(J_n) \to 0$. On fixe $\epsilon > 0$, et pour tout $J_n$, on considère $K_n$ compact, inclus dans $J_n$, tel que $\lambda(J_n \setminus K_n) \leq \epsilon/2^n$.

Alors, si $\bar{K}_n = K_1 \cap \ldots \cap K_n$, $\bar{K}_n$ est une suite décroissante de compacts, on a

$$J_n \setminus \bar{K}_n \subset \cup_{p=1}^n (J_p \setminus K_p),$$

et donc

$$\lambda(J_n \setminus \bar{K}_n) \leq \sum_1^n \lambda(J_p \setminus K_p) \leq \sum_{p=1}^n \frac{\epsilon}{2^p} \leq \epsilon.$$

Par ailleurs, $\cap_n \bar{K}_n \subset \cap_n J_n = \emptyset$. Puisque les $\bar{K}_n$ sont compacts, alors il existe un $n_0$ tel que $\bar{K}_{n_0} = \emptyset$, et donc $\mu(\bar{K}_{n_0}) = 0$.

Pour $n \geq n_0$

$$\lambda(J_n) \leq \lambda(J_{n_0}) = \lambda(J_{n_0} \setminus K_{n_0}) \leq \epsilon,$$

et on a fini.

∎

On peut alors définir la mesure de Lebesgue sur $\mathbb{R}$ tout entier de la façon suivante

**Définition.** 10. *Soit $I_n = [-n, n]$ et $\lambda_n$ la mesure de Lebesgue sur l'intervalle $[-n, n]$. On pose pour un borélien $A \subset \mathbb{R}$*

$$\mu_n(A) = \lambda_n(A \cap I_n).$$

*Alors $\mu_n$ est une suite croissante de mesures. La limite de cette suite est par définition la mesure de Lebesgue sur $\mathbb{R}$.*

**Fonctions de répartition et mesures bornées sur $\mathbb{R}$.**

Si $\mu$ est une mesure bornée (disons de masse 1) sur $\mathbb{R}$, alors on considère la fonction $F_\mu(x) = \mu(]-\infty, x])$. On l'appelle la fonction de répartition de $\mu$. Le théorème précédent nous montre qu'elle caractérise entièrement la mesure $\mu$.

Elle a les propriétés suivantes

1. $F_\mu$ est croissante, continue à droite.
2. $\lim_{x\to-\infty} F_\mu(x) = 0$, $\lim_{x\to+\infty} F_\mu(x) = \mu(\mathbb{R}) = 1$.

Une fonction vérifiant ces deux propriétés sera appelée une FR (fonction de répartition).

Remarquons qu'une FR a des limites à gauche en tout point. Si l'on note $F(x_-)$ la limite à gauche de $F$ en $x$, il n'y a au plus qu'un nombre fini de points pour lesquels $F(x) - F(x_-) > \frac{1}{n}$ (Il y en a au plus $nF(\infty)$.)

L'ensemble des $x$ pour lesquels $F(x)-F(x_-) > 0$ est donc au plus dénombrable (c'est la réunion des ensembles précédents).

Un ensemble dénombrable a toujours un complémentaire dense. Donc, l'ensemble des points de continuité d'une FR est toujours dense.

**Théorème 4.** *Soit $F$ une FR. Alors, il existe une (et donc une seule) mesure bornée sur $\mathbb{R}$ telle que $F = F_\mu$.*

La mesure (de Stieltjes) associée à $F$ sera noté $\mu_F$ (et plus tard parfois $dF$ lorsqu'on parlera de l'intégration de fonctions).

*Démonstration.* — La démonstration se fait exactement comme dans le cas de la mesure de Lebesgue. La difficulté est qu'il faut faire attention aux sauts de $F$ (qui vont correspondre aux masses de Dirac dans la mesure $\mu$).

Pour cela, on prend comme algèbre les réunions finies d'intervalles dont les extrémités sont dans les points de continuité de $F$. Pour un tel intervalle $(a, b)$, quels que soient ses bords (ouverts ou fermés), on pose $\mu(a, b) = F(b) - F(a)$. On applique alors la même méthode que pour la mesure de Lebesgue.

Remarquer que si $\mu(]-\infty, b]) = F(b)$, on doit avoir pour tout intervalle, dont les extrémités ne sont pas nécéssairement dans $\mathcal{D}$, $\mu(]a, b]) = F(b) - F(a)$ et de même

$$\mu([a, b]) = F(b) - F(a_-),\ \ \mu(]a, b[) = F(b_-) - F(a), \ldots$$

Cela permet de définir $\mu$ sur l'algèbre $\mathcal{A}_0$ des réunions finies d'intervalles. On considère alors l'ensemble (dense) $\mathcal{D}$ des points de continuité de la fonction $F$, et on considère l'algèbre de Boole $\mathcal{A}_1$ des réunions finies d'intervalles à extrémité dans $\mathcal{D}$. On a $\sigma(\mathcal{A}_1) = \mathcal{B}(\mathbb{R})$. Puisque les extrémités d'intervalles d'un élément de $\mathcal{A}_1$ sont des points de continuité de $F$, on peut établir le même résultat que pour la mesure de Lebesgue : pour tout élément de $A$ de $\mathcal{A}_1$ et tout $\epsilon > 0$, il existe un élément compact $K$ de $\mathcal{A}_1$ inclus dans $A$ tel que $\mu(A \setminus K) \leq \epsilon$. On peut alors faire la même démonstration que pour la mesure de Lebesgue sur $[0, 1]$.

Nous laissons les détails au lecteur.

∎

Le théorème précédent permet de décrire toutes les mesures bornées sur $\mathbb{R}$. Si la mesure n'est pas bornée, on peut faire la même chose à condition que la mesure $\mu$ donne une masse finie à tous les intervalles bornés (ou de façon équivalente à tous les intervalles compacts). Ces mesures forment une sous-classe des mesures $\sigma$-finies, qu'on appelle les mesures de **Radon**.

Pour une telle mesure, on ne pose pas $F(x) = \mu(]-\infty, x])$, qui peut être infini. On choisit plutôt

$F(x) = \mu(]0, x])$ si $x \geq 0$ et $F(x) = -\mu(]x, 0])$ si $x < 0$.

De cette façon, on aura toujours une fonction $F$ croissante continue à droite, avec $F(0) = 0$, et qui vérifie

$$\mu(]a, b]) = F(b) - F(a),$$

pour tout intervalle $]a, b]$.

On a alors

**Théorème 5.** *Soit $F$ une fonction croissante continue à droite nulle en 0. Il existe une unique mesure de Radon $\mu$ sur $\mathbb{R}$ telle que*
*$F(x) = \mu(]0, x])$ si $x \geq 0$ et $F(x) = -\mu(]x, 0])$ si $x < 0$.*

La démonstration se fait exactement comme plus haut pour les mesures finies. Le choix du point 0 pour avoir $F(0) = 0$ est arbitraire et n'a rien de particulier.

Revenons aux mesures finies. Les points de discontinuité de $F_\mu$ sont les points $x$ tels que $\mu(\{x\}) \neq 0$, puisque

$$F_\mu(x) - F_\mu(x_-) = \mu(\{x\}).$$

Ces points sont en nombre dénombrable. Enumérons les en une suite $(x_n)$ et posons $\alpha_n = \mu(\{x_n\})$.

La mesure $\nu = \sum_n \alpha_n \delta_{x_n}$ satisfait

$$\nu(A) \leq \mu(A), \ \forall A \in \mathcal{B}(\mathbb{R}).$$

Une telle mesure (somme de masses de Dirac) est dite **purement atomique**.

La mesure $\mu_1 = \mu - \nu$ est une mesure (positive) qui vérifie pour tout $x$ $\mu_1(\{x\}) = 0$. Une telle mesure est dite **sans atome**. Il est est équivalent de dire que sa fonction de répartition est continue.

(Un atome est une partie de mesure positive qu'on ne peut pas découper en ensembles mesurables de mesure strictement plus petite.)

Nous avons ainsi décomposé la mesure $\mu$ en $\mu = \mu_1 + \nu$, où $\mu_1$ est sans atome et $\mu$ est purement atomique.

Parmi toutes les mesures de Radon sur $\mathbb{R}$, le rôle particulier joué par la mesure de Lebesgue est expliqué par le théorème suivant.

**Théorème 6.** *Pour un borélien $A \in \mathcal{B}(\mathbb{R})$, et pour tout $x \in \mathbb{R}$, posons*

$$A + x = \{y \mid y - x \in A\}.$$

*Alors $A+x$ est un borélien et $\lambda(A+x) = \lambda(A)$. (En d'autres termes, la mesure de Lebesgue est invariante par translation.)*

*De plus, si une mesure de Radon $\mu$ est invariante par translation, alors il existe une constante c telle que $\mu = c\lambda$.*

*Démonstration.* — Pour montrer que $A + x$ est un borélien, nous fixons $x$ et nous remarquons que

$$\{A \mid A + x\} \in \mathcal{B}(\mathbb{R})$$

est une $\sigma$-algèbre, qui contient les intervalles et donc toute la tribu borélienne.

Montrons que la mesure de Lebesgue est invariante par translation. Pour $A \in \mathcal{B}(\mathbb{R})$ et $M \in (0, \infty)$, posons $A^M = A \cap [-M, M]$. Posons, $x$ étant fixé

$$\mathcal{M} = \{A \mid \forall M \in (0, \infty),\ \lambda(A^M + x) = \lambda(A^M)\}.$$

On vérifie immédiatement que $\mathcal{M}$ est une classe monotone. Elle contient les intervalles $[a, b]$, donc toute la tribu borélienne. On voit donc que, pour tout $x \in \mathbb{R}$, pour tout $M \in (0, \infty)$ et pour tout borélien $A$,

$$\lambda(A^M + x) = \lambda(A^M).$$

Il suffit de maintenant de faire converger $M$ vers l'infini pour obtenir le résultat.

Démontrons l'unicité. Considérons une mesure de Radon $\mu$ invariante par translation. Si $\mu(]0, 1]) = 0$, alors par translation $\mu(]p, p + 1]) = 0$ pour tout $p \in \mathbb{Z}$ et donc $\mu$ est nulle. Il n'y a rien à démontrer. Sinon, quitte à multiplier $\mu$ par une constante, on peut toujours supposer que $\mu(]0, 1]) = 1$. Démontrons qu'alors $\mu = \lambda$.

Remarquons tout d'abord que, pour $p \in \mathbb{N}$, $p \geq 1$, l'intervalle $]0, 1]$ est la réunion disjointe de $p$ translatés de l'intervalle $]0, \frac{1}{p}]$. Ces intervalles ont tous même mesure, et donc

$$p\mu(]0, \frac{1}{p}]) = \mu(]0, 1]) = 1,$$

d'où

$$\mu(]0, \frac{1}{p}]) = \frac{1}{p}.$$

Par translation, on trouve que

$$\mu(]\frac{k}{p}, \frac{k+1}{p}]) = \frac{1}{p},$$

pour tout $k \in \mathbb{Z}$, et par suite

$$\mu(]\frac{k}{p},\frac{r}{p}]) = \frac{r-k}{p},$$

pour tout $k \leq r \in \mathbb{Z}$, $p \in \mathbb{N}$, $p > 0$.

Les mesures $\mu$ et $\lambda$ coïncident donc sur les intervalles $]a,b]$ à extrémités rationnelles, qui forment une classe stable par intersection qui engendre la tribu borélienne. Elles coïncident donc partout. ∎

## 2.3 Fonctions mesurables

**Définition.** 11. *Soient $(\Omega_1, \mathcal{A}_1)$ et $(\Omega_2, \mathcal{A}_2)$ deux espaces mesurables. Une application $f : \Omega_1 \to \Omega_2$ est mesurable si*

$$\forall B \in \mathcal{A}_2,\ f^{-1}(B) \in \mathcal{A}_1.$$

Lorsque $\Omega_2 = \mathbb{R}$ (c'est à dire que $f$ est à valeurs réelles), on supposera toujours que $\mathcal{A}_2 = \mathcal{B}(\mathbb{R})$.

Rappelons que $f^{-1}(B) = \{x \in \Omega_1 \mid f(x) \in B\}$. L'application

$$f^{-1} : \mathcal{P}(\Omega_2) \to \mathcal{P}(\Omega_1)$$

vérifie

1. $f^{-1}(\emptyset) = \emptyset$;
2. $f^{-1}(B^c) = f^{-1}(B)^c$;
3. $f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2)$, et même, pour une famille quelconque $B_i$
$$f^{-1}(\cap_i B_i) = \cap_i f^{-1}(B_i),\ f^{-1}(\cup_i B_i) = \cup_i f^{-1}(B_i).$$

Par abus de notation (mais attention aux confusions), on note souvent $f^{-1}(A) = \{f \in A\}$.

Ainsi, on voit que si $\mathcal{A}_2$ est une tribu sur $\Omega_2$,

$$f^{-1}(\mathcal{A}_2) = \{A \in \Omega_1 \mid \exists B \in \mathcal{A}_2, A = f^{-1}(B)\}$$

est toujours une tribu sur $\Omega_1$.

C'est la plus petite tribu $\mathcal{A}$ sur $\Omega_1$ telle que $f$ soit mesurable de $(\Omega_1, \mathcal{A}) \to (\Omega_2, \mathcal{A}_2)$. On la note $\sigma(f)$ : la plus petite tribu qui rende $f$ mesurable (parfois aussi notée $f^{-1}(\mathcal{A}_2)$.

**Exemple fondamental de fonctions mesurable.**

Si $(\Omega, \mathcal{A})$ est un ensemble mesurable, alors $\mathbf{1}_A$ est toujours une fonction mesurable réelle. On a aussi $\sigma(\mathbf{1}_A) = \sigma(A)$.

**Proposition 5.** *La composition de deux fonctions mesurables est mesurable*

C'est immédiat et laissé au lecteur.

Pour vérifier qu'une fonction est mesurable, on n'a en général pas besoin de tester $f^{-1}(B)$ pour tous les $B \in \mathcal{A}_2$.

En effet on a

**Proposition 6.** *Soit $f : (\Omega_1, \mathcal{A}_1) \to (\Omega_2, \mathcal{A}_2)$ et supposons que $\mathcal{A}_2 = \sigma(\mathcal{E})$, où $\mathcal{E}$ est une famille quelconque d'ensembles. Pour que $f$ soit mesurable, il suffit que, pour tout $E \in \mathcal{E}$, $f^{-1}(E) \in \mathcal{A}_1$.*

*Démonstration.* — On vérifie immédiatement que

$$\mathcal{A} = \{B \subset \Omega_2 \mid f^{-1}(B) \in \mathcal{A}_1\}$$

est une tribu. Puisqu'elle contient $\mathcal{E}$, elle contient $\mathcal{A}_2$. ∎

Ainsi, pour qu'une fonction $f : \Omega_1 \to \mathbb{R}$ soit mesurable, il suffit que, pour tout $x \in \mathbb{Q}$ $\{y \mid f(y) < x\} \in \mathcal{A}_1$. (On applique ce qui précède avec $\mathcal{E} = \{(-\infty, x[,\ x \in \mathbb{Q}\}$.

De même, toute fonction continue de $\mathbb{R}$ dans $\mathbb{R}$ est mesurable, puisque la tribu borélienne est engendrée par les ouverts, et que l'image réciproque d'un ouvert par une fonction continue est un ouvert.

Il en va de meme de toute fonction monotone (remplacer dans la phrase précédente ouvert par intervalle).

**Proposition 7.** *Si $f_1$ et $f_2$ sont des fonctions mesurables réelles, il en va de même de*

1. $f_1 + f_2$;
2. $f_1 f_2$,
3. $\max(f_1, f_2)$.

*Démonstration.* — Montrons le à titre d'exemple pour $f_1 + f_2$. Les autres points sont laissés à titre d'exercice. Pour $x \in \mathbb{R}$, on écrit

$$\{f_1 + f_2 < x\} = \cup_{r \in \mathbb{Q}} \left(\{f_1 < r\} \cap \{f_2 < x - r\}\right).$$

Dans le second membre, on obtient une union dénombrable d'ensembles mesurables, qui est donc mesurable. ∎

Le dernier point nous permet de décomposer toute fonction mesurable réelle en différence de deux fonctions mesurables positives. Si l'on pose $f_+(x) = \max(f(x), 0)$ et $f_-(x) = \max(-f(x), 0)$, $f_+$ et $f_-$ sont mesurables, et on a

$$f = f_+ - f_-, \ \ |f| = f_+ + f_-.$$

Cette décomposition nous sera souvent très utile.

Il nous arrivera aussi souvent de considérer des fonctions prenant les valeurs $\pm\infty$. Dans ce cas, nous mettrons sur $\bar{\mathbb{R}}$ la tribu de $\sigma(\bar{\mathbb{R}})$, qui est la tribu engendrée par tous les intervalles (y compris ceux qui contiennent l'infini). Mais il sera en général beaucoup plus simple de changer $\bar{\mathbb{R}}$ en un intervalle fermé borné à l'aide d'une fonction monotone bornée $h$ (par exemple arctg ou th), et de dire que $f$ est mesurable à valeurs dans $\bar{\mathbb{R}}$ si et seulement si $h(f)$ est mesurable à valeurs dans $\mathbb{R}$.

La propriété de mesurabilité est très stable. Il est difficile d'exhiber une fonction non mesurable. Toutes les fonctions qu'on rencontre "naturellement" le seront.

On a en particulier

**Proposition 8.** *Soit $f_n$ une suite de fonctions mesurables $f_n : \Omega \to \mathbb{R}$. On suppose que, pour tout $x \in \Omega$, $f_n(x)$ converge vers $f(x)$. Alors, la fonction $x \to f(x)$ est mesurable.*

Ainsi, la famille des fonctions mesurables réelles est stable par limite simple (à l'opposé par exemple de la famille des fonctions continues).
*Démonstration.* — On écrit, pour $x \in \mathbb{R}$

$$\{f < x\} = \cup_{r\in\mathbb{Q}, r<x} \cup_{n_0\in\mathbb{N}} \cap_{n\geq n_0} \{f_n < r\}.$$

Cette identité ne fait que traduire en terme de théorie des ensembles la phrase

``$f(y) < x$ si et seulement si il existe $r \in \mathbb{Q}$, $r < x$, et $n_0 \in \mathbb{N}$, tel que pour tout $n \geq n_0$, $f_n(y) < r$.''

Le deuxième membre de l'identité précédente est mesurable, car construit à l'aide d'un nombre dénombrable d'opérations sur des ensembles mesurables.

∎

Ce qu'on vient d'écrire peut aussi s'étendre aux limites dans $\bar{\mathbb{R}}$.

D'ailleurs, on pourra démontrer de la même façon que, si $f_n$ est une suite de fonctions mesurables réelles, l'ensemble des $x$ pour lesquels $f_n(x)$ converge est un ensemble mesurable.

Rappelons au passage les notions de limite sup et limite inf (dans $\mathbb{R}$ ou $\bar{\mathbb{R}}$).

**Définition.** 12 (Limites sup et limites inf). *Soit $x_n$une suite de réels. On pose $\bar{x}_n = \sup_{p\geq n} x_p$, et $\underline{x}_n = \inf_{p\geq n} x_p$. Le sup et l'inf ici peuvent prendre les valeurs $+\infty$ (pour le sup) et $-\infty$ (pour l'inf). $\bar{x}_n$ est une suite décroissante. Elle converge (dans $\bar{\mathbb{R}}$) vers une limite qu'on appelle la limite sup de la suite $x_n$ et dénotée $\limsup_n x_n$. De même, $\underline{x}_n$ est une suite croissante qui converge vers une limite notés $\liminf_n x_n$. On a toujours*

$$\liminf_n x_n = \limsup_n x_n,$$

*et la suite converge si et seulement si $\liminf_n x_n = \limsup_n x_n$.*

On peut appliquer cela aux suites de fonctions à valeurs réelles. D'après tout ce qui précède, on voit que, si $f_n$ est une suite de fonctions mesurables, alors $\liminf_n f_n$ et $\limsup_n f_n$ sont des fonctions mesurables (à valeurs éventuellement dans $\bar{\mathbb{R}}$). Ainsi, on voit aisément que, pour une suite de fonction bornées, l'ensemble des $x$ où la suite converge est mesurable, puisqu'il s'écrit $\{\limsup_n f_n - \liminf_n f_n = 0\}$.

**Définition.** 13 (Fonctions étagées). *On dit qu'une fonction $\Omega \to \mathbb{R}$ est étagée si elle s'écrit*

$$f = \sum_{i=1}^{n} \lambda_i \mathbf{1}_{A_i}$$

*où $A_i \in \mathcal{A}$ et $\lambda_i \in \mathbb{R}$.*

**Remarque 10.** *Une fonction étagée n'est rien d'autre qu'une fonction mesurable réelle ne prenant qu'un nombre fini de valeurs . (A démontrer à titre d'exercice)*

*Remarquons également qu'une fonction étagée admet une représentation $f = \sum_i \mu_i \mathbf{1}_{B_i}$ où les $B_i$ sont disjoints et les $\mu_i$ sont distincts non nuls. cette représentation est unique (les $\mu_i$ sont les valeurs non nulles prises par $f$ et $B_i = f^{-1}(\mu_i)$). Nous appelerons cette réprésentation la* **représentation canonique**.

La proposition suivante nous montre l'intérêt des fonctions étagées.

**Proposition 9.** *1. Toute fonction mesurable réelle positive est limite croissante de fonctions étagées.*

*2. Toute fonction mesurable bornée est limite uniforme de fonctions étagées.*

(Rappelons qu'une fonction est dite bornée s'il existe un $M \in \mathbb{R}_+$ tel que, pour tout $x$, $f(x) \in [-M, M]$.)
*Démonstration.* — Nous ne montrons que le premier point. Le second est similaire et laissé au lecteur.

Pour tout $n \in \mathbb{N},\ n > 0$, on pose

$$f_n(x) = \sum_{k=0}^{2^{2n}} \frac{k}{2^n} \mathbf{1}_{A_{k,n}} + 2^n \mathbf{1}_{A_n},$$

où

$$A_{k,n} = \{f \in ]\frac{k}{2^n}, \frac{k+1}{2^n}],$$

at $A_n = \{f > 2^n\}$. La fonction $f_n(x)$ n'est rien d'autre que l'écriture en base 2 de $f(x)$, tronquée à la $k$-ième décimale, si $f(x) \leq 2^n$, et sinon, vaut $2^n$ si $f(x) > 2^n$.

On voit immédiatement que la suite $f_n$ est croissante, que si $f(x) < \infty$, alors

$$0 \leq f(x) - f_n(x) \leq \frac{1}{2^n}$$

dès que $n$ est assez grand pour avoir $f(x) \leq 2^n$, et que si $f(x)$ est infini, $f_n(x) = 2^n$. Dans tous les cas, $f_n(x)$ converge en croissant vers $f(x)$.

∎

Enfin, nous énonçons sans démonstration la version fonctionnelle du théorème des classes monotones.

Si $\mathcal{A}$ est une tribu, on note $\mathcal{B}(\mathcal{A})$ l'espace vectoriel de toutes les fonctions réelles $\mathcal{A}$-mesurables et bornées.

On note aussi, pour toute famillle de fonction $\mathcal{E}$, $\sigma(\mathcal{E})$ la plus petite tribu qui rend mesurable toutes les fonctions de cette famille.

Ainsi, par exemple, $\mathcal{B}(\mathbb{R}) = \sigma(x^3)$, ou bien $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{C})$, où $\mathcal{C}$ désigne l'espace de toutes les fonctions continues, etc..

**Définition.** 14. *Soit $\Omega$ un ensemble, et $\mathcal{H}$ un sous-espace vectoriel de l'espace vectoriel de toutes les fonctions bornées $\Omega \to \mathbb{R}$.On dira que $\mathcal{H}$ est un espace de classe monotone si*

1. *$\mathcal{H}$ contient les fonctions constantes*

2. *Si $f_n$ est une suite d'élements de $\mathcal{H}$ qui vérifie*

$$\forall x, \forall n,\ f_n(x) \in [0, 1],\ f_n(x) \leq f_{n+1}(x),$$

*alors la limite $f = \lim_n f_n$ est encore un élément de $\mathcal{H}$.*

<u>**Théorème**</u> **7** (Des classes monotones, version fonctionnelle)**.** *Si une classe monote $\mathcal{H}$ contient une famille de fonctions $\mathcal{E}$* **stable par multiplication***, elle contient $\mathcal{B}(\sigma(\mathcal{E}))$.*

Nous n'allons pas démontrer ce théorème, dont la démonstration est sensiblement plus compliquée que celle du théorème ensembliste correspondant. Nous renvoyons pour cela au livre de Barbe-Ledoux.

Enfin, terminons ce chapitre par une remarque sur les limsup et les liminf de suites d'ensembles, qu'il ne faut pas mélanger avec les limsup et les liminf de suites de nombres réels.

**Définition.** 15 (Limsup et Liminf d'ensembles). *Soit $(A_n)$ une suite de sous ensembles de $\Omega$. On note*

$$\limsup_n A_n = \cap_n \cup_{p\geq n} A_p, \quad \liminf_n A_n = \cup_n \cap_{p\geq n} A_p.$$

La limite sup des $A_n$ est l'ensemble des points $x$ qi appartiennent à une infinité de $A_n$. La limite inf est ceux qui appartiennent à tous les $A_n$ à partir d'un certain rang.

**Proposition 10.** *On a les propriétés élémentaires suivantes*

1. *Si $A_n \subset B_n$, alors $\limsup_n A_n \subset \limsup_n B_n$, et $\liminf_n A_n \subset \liminf_n B_n$.*
2. $\liminf_n A_n \subset \limsup_n A_n$, $(\limsup_n A_n)^c = \liminf_n A_n^c$,
   $(\liminf_n A_n)^c = \limsup_n A_n^c$.
3. *Si $A_n$ est croissante, alors $\limsup_n A_n = \liminf_n A_n = \cup_n A_n$. Si $A_n$ est décroissante, alors $\limsup_n A_n = \liminf_n A_n = \cap_n A_n$.*
4. $\limsup_n \mathbf{1}_{A_n} = \mathbf{1}_{\limsup_n A_n}$; $\liminf_n \mathbf{1}_{A_n} = \mathbf{1}_{\liminf_n A_n}$.
5. *Si $\mu$ est une mesure bornée, alors*

   $$\mu(\liminf_n A_n) \leq \liminf_n \mu(A_n) \leq \limsup_n \mu(A_n) \leq \mu(\limsup_n A_n).$$

   *En particulier, si $\limsup_n(A_n) = \liminf_n(A_n)$, alors $\mu(A_n)$ converge et on a $\lim_n \mu(A_n) = \mu(\limsup_n A_n) = \mu(\liminf_n(A_n)$.*

Ces propriétés sont presques évidentes une fois que l'on a remarqué que $\cup_{p\geq n} A_p$ est une suite décroissante d'ensembles, et que $\cap_{p\geq n} A_p$ est une suite croissante. Les détails sont laissés à titre d'exercice.

On pourra observer comme conséquence que si $\limsup_n A_n = \liminf_n A_n$ et $\limsup_n B_n = \liminf_n B_n$, alors

$$\limsup_n A_n \cap B_n = \liminf_n A_n \cap B_n, \quad \limsup_n A_n \cup B_n = \liminf_n A_n \cup B_n,$$

et par conséquent que si $A_n$ est croissante et $B_n$ décroissante, alors $\mu(A_n \cap B_n)$ converge.

# 3 Intégration de fonctions mesurables

Dans ce chapitre, nous allons considérer un espace mesuré $(\Omega, \mathcal{A}, \mu)$ et définir pour certaines fonctions mesurables $f$ l'intégrale $\int f d\mu$.

On commence par l'intégrale de fonctions simples

## 3.1 Intégration de fonctions étagées.

Appelons $\mathcal{E}$ l'ensemble des fonctions étagées à valeurs réelles

$$f(x) = \sum_{i=1}^{n} \lambda_i \mathbf{1}_{A_i}(x).$$

Rappelons que ce sont les fonctions mesurables ne prenant qu'un nombre fini de valeurs.

Parmi celles-ci, nous allons distinguer les fonctions positives ($\forall x,\ f(x) \geq 0$). On appelera $\mathcal{E}_+$ l'ensemble des fonctions étagées positives. Remarquons que

1. $\mathcal{E}$ est un espace vectoriel;
2. $\mathcal{E}_+$ est stable par $(f, g) \to \max(f, g)$ et $(f, g) \to \min(f, g)$.

D'autre part, dans la réprésentation $f = \sum_{i=1}^{k} \lambda_i \mathbf{1}_{A_i}$ d'un élément de $\mathcal{E}$, on peut toujours supposer que pour tout $i$, $\lambda_i \neq 0$.

**Définition.** 16. *Pour un élément $f \in \mathcal{E}_+$, si $f = \sum_i \lambda_i \mathbf{1}_{A_i}$, on pose*

$$\int f d\mu = \sum_i \lambda_i \mu(A_i).$$

Il faut faire attention ici à ce que, si $\mu(A_i) = \infty$ mais que $\lambda_i = 0$, on pose par convention $\lambda_i \mu(A_i) = 0$.

Ceci permet de rester compatible avec $0\mathbf{1}_\Omega = \mathbf{1}_\emptyset$.

La première chose à remarquer est que cette définition ne dépend pas de la façon dont on a écrit $f$. Pour le voir, le plus simple est de remarquer si $f = \sum_{i=1}^{n} \lambda_i \mathbf{1}_{A_i}$ a comme représentation canonique $f = \sum_{i=1}^{p} \mu_i \mathbf{1}_{B_i}$, alors les deux représentations de $f$ donnent la même valeur à $I(f)$, ce qu'on montre par exemple par récurrence sur $n$. Cela repose au bout du compte sur la formule

$$\mu(A \cup B) = \mu(A) + \mu(B) \ \text{ si } A \cap B = \emptyset.$$

Il faut faire attention ici à ce qu'on peut avoir $I(f) = \infty$, si par exemple pour un indice $i$, on a $\mu(A_i) = \infty$ et $\lambda_i > 0$.

**Proposition 11.** *L'application $f \to I(f) = \int f d\mu$ vérifie les propriétés suivantes*

1. *Si $\lambda \geq 0$, $I(\lambda f) = \lambda I(f)$;*
2. *$I(f + g) = I(f) + I(g)$.*
3. *Si $f \geq g$, $I(f) \geq I(g)$.*
4. *Si $f_n$ est une suite croissante d'éléments de $\mathcal{E}_+$ dont la limite est un élément $f$ de $\mathcal{E}_+$, alors $I(f_n)$ converge (en croissant) vers $I(f)$.*

*Démonstration.* — Les deux premiers points sont évidents. Pour le troisième, on remarque que si $f \geq g$, alors $h = f - g \in \mathcal{E}_+$ et qu'alors

$$I(f) = I(g) + I(h) \geq I(g).$$

Le quatrième point est plus délicat. Choisissons $f = \sum_{i=1}^k \lambda_i \mathbf{1}_{A_i}$, sous sa représentation canonique, avec $\lambda_i > 0$ pour tout $i$ (si tous les $\lambda_i$ sont nuls, alors $f = 0$, et il en va de même de chaque $f_n$, et il n'y a rien à dire).

De même, on peut supposer que $I(f) > 0$, sinon, il n'y a rien à démontrer. Soit alors $\alpha_n = I(f_n)$. La suite $\alpha_n$ est croissante, majorée par $I(f)$. Elle converge donc vers une limite $\alpha \leq I(f)$. Notre problème est de montrer que $\alpha = I(f)$.

Nous avons à distinguer deux cas : soit $I(f) = \infty$, soit $I(f) < \infty$.

Traitons pour commencer le second point. Montrons que, pour tout $\epsilon > 0$ assez petit, $\alpha \geq I(f) - \epsilon$. Commençons par choisir une fonction $g = \sum_{i=1}^k \mu_i \mathbf{1}_{A_i}$, telle que $\mu_i < \lambda_i$ pour tout $i$ et $I(g) \geq I(f) - \epsilon$.

La suite $f_n$ converge en croissant vers $f$, donc, sur $A_i$, $f_n(x) \geq \mu_i$ pour $n$ assez grand (disons $n \geq n_i$).

Comme il n'y a qu'un nombre fini d'indices $i$, on sait qu'il existe $n_0$ (par exemple $\max(n_i)$ ) tel que, si $n \geq n_0$, alors $f_n(x) \geq g(x)$ partout.

Donc, pour $n \geq n_0$, $\alpha_n = I(f_n) \geq I(g) \geq I(f) - \epsilon$, et donc $\alpha \geq I(f) - \epsilon$.

Le cas où $I(f)$ est infini se traite de même, en montrant que, pour tout $M > 0$, il existe un $n_0$ tel que, pour $n \geq n_0$, on a $\alpha_n \geq M$. ∎

## 3.2 L'intégrale des fonctions positives.

Nous pouvons alors donner la définition de l'intégrale d'une fonction mesurable positive

**Définition.** 17. *Soit $f$ une fonction mesurable positive. On pose*

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

*La même définition s'applique aux fonctions mesurables à valeurs dans* $[0, +\infty]$. *On note parfois, s'il y a risque d'ambiguité*

$$\int f d\mu = \int f(\omega) d\mu(\omega) = \int f(\omega) \mu(d\omega).$$

Nous savons déjà que toute fonction mesurable positive est limite croissante de fonctions étagées. Le résultat suivant (fondamental) permet de montrer comment calculer $\int f d\mu$.

**Proposition 12.** *Soit $(f_n)$ une suite de fonctions de $\mathcal{E}_+$ qui converge en croissant vers $f$. Alors, $\int f d\mu = \lim_n \int f_n d\mu$.*

*Démonstration.* — Une fois de plus, il faut distinguer le cas où $\int f d\mu < \infty$ et où $\int f d\mu = \infty$. Nous ne traitons que le premier, le second est laissé à titre d'exercice.

Soit donc $I = \int f d\mu$ et $a_n = \int f_n d\mu$, où $f_n$ est une suite croissante étagée qui converge vers $f$. Tout d'abord, $f_n$ est majorée par $f$, et donc $a_n \leq \int f d\mu$

par définition. Ensuite, $a_n$ est croissante, majorée par $\int f d\mu$, donc converge vers une limite $\alpha \leq \int f d\mu$. Tout notre problème est de montrer que $\alpha = \int f d\mu$.

On va donc montrer que, $\forall \epsilon > 0$, $\alpha \geq \int f d\mu - \epsilon$.

Tout d'abord, par définition de $\int f d\mu$, il existe un élément $g \in \mathcal{E}_+$, $g \leq f$, tel que $I(g) = \int g d\mu \geq \int f d\mu - \epsilon$.

La suite $h_n = \min(f_n, g)$ est une suite d'élément de $\mathcal{E}_+$ qui converge en croissant vers $g$. D'après la proposition précédente, si on appelle $b_n = \int h_n d\mu$, alors $b_n$ converge vers $I(g)$. Mais $b_n \leq a_n$ puisque $h_n \leq f_n$. Et donc,

$$\alpha = \lim_n a_n \geq \lim_n b_n = I(g) \geq \int f d\mu - \epsilon.$$

C'est bien ce qu'on voulait démontrer. ∎

Une fois ce résultat établi, alors les résultats suivants sont élémentaires

**Proposition 13.** *L'application $f \to \int f d\mu$ satisfait*

1. *Pour tout $\lambda \geq 0$, $\int \lambda f d\mu = \lambda \int f d\mu$*
2. $\int (f+g) d\mu = \int f d\mu + \int g d\mu$.
3. *Si $f \leq g$, alors $\int f d\mu \leq \int g d\mu$.*

*Démonstration.* — Montrons par exemple le second point. On choisit une suite $(f_n)$ et une suite $(g_n)$ de fonctions de $\mathcal{E}_+$, la première qui converge en croissant vers $f$, la seconde vers $g$.

On sait que

$$\int (f_n + g_n) d\mu = \int f_n d\mu + \int g_n d\mu,$$

puis on passe à la limite. ∎

Donnons tout d'abord quelques exemples.

1. Si $\mu = \delta_x$, $\int f d\mu = f(x)$.
2. Sur $\mathbb{N}$, si $\mu$ est la mesure de décompte, alors

   $$\int f d\mu = \sum_n f(n).$$

   Ainsi, l'intégrale par rapport à la mesure de décompte n'est rien d'autre que la somme de la série $\sum_n f(n)$.

Nous verrons plus bas que l'intégrale pour la mesure de Lebesgue d'une fonction continue ou monotone n'est rien d'autre que son intégrale de Riemann.

Ainsi, l'intégrale de Lebesgue généralise à la fois l'intégrale ordinaire et la sommation des séries, et permet d'énoncer des théorèmes généraux qui s'appliqueront dans ces deux cas.

Les deux résultats suivants sont fondamentaux

**Théorème 8.** *(Théorème de convergence monotone) Soit $(f_n)$ une suite de fonctions mesurables positives qui converge en croissant vers $f$. Alors*

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

Insistons sur le fait que ce résultat s'applique même aux fonctions qui ne sont pas finies. Remarquez que par comparaison au résultat précédent, cette fois ci la suite $f_n$ n'est plus nécessairement étagée.
*Démonstration.* — (Du théorème de convergence monotone). Soit $(f_n)$ une suite croissante de fonctions mesurables positives, et $(f_{n,p})$ une suite croissante (en $p$) de fonctions étagées qui converge vers $f_n$.

Posons alors

$$g_{n,p} = \max(f_{1,p}, f_{2,p}, \dots, f_{n,p}).$$

$(g_{n,p})$ est à la fois croissante en $n$ et en $p$, est majorée par $f_n$ et converge aussi vers $f_n$ lorsque $p \to \infty$.

Le lemme fondamental (Lemme 1) nous dit que

$$\lim_n \lim_p g_{n,p} = \lim_n g_{n,n}.$$

et donc la suite $g_{n,n}$ converge vers $f$.

Alors, d'après ce que l'on sait sur les suites croissantes de fonctions étagées,

$$\lim_n \int g_{n,n} d\mu = \int f d\mu.$$

Par ailleurs,

$$\lim_n \lim_p \int g_{n,p} d\mu = \lim_n \int f_n d\mu.$$

On réapplique le lemme fondamental (Lemme 1) avec la suite $a_{n,p} = \int g_{n,p} d\mu$, et on obtient le résultat.

∎

Une application immédiate du théorème de convergence monotone est le suivant

**Proposition 14.** *Si $(f_n)$ est une suite de fonctions mesurables positives,*

$$\int \sum_n f_n d\mu = \sum_n \int f_n d\mu.$$

(Il suffit d'appliquer le théorème de convergence monotone à la suite croissante $F_n = \sum_{p=1}^n f_p$.)

L'autre résultat fondamental est le suivant

**Théorème 9.** *(Lemme de Fatou) Soit $(f_n)$ une suite de fonctions mesurables positives. Alors*

$$\int (\liminf_n f_n) d\mu \leq \liminf_n (\int f_n d\mu).$$

*Démonstration.* — (Du lemme de Fatou) On pose $g_n = \inf_{p \geq n} f_p$. La suite $g_n$ converge en croissant (par définition) vers $\liminf_n f_n$. Et donc, par le théorème de convergence monotone,

$$\int (\liminf_n f_n) d\mu = \lim_n \int g_n d\mu.$$

Mais

$$\int g_n d\mu \leq \int f_n d\mu.$$

Or, si $a_n \leq b_n$, $\liminf a_n \leq \liminf b_n$ (Exercice immédiat d'après la définition de la lim inf.). D'où l'inégalité du lemme de Fatou. ∎

Pour se rappeler du sens de l'inégalité dans le lemme de Fatou, le mieux est de souvenir d'un contre exemple : sur $[0,1]$, avec la mesure de Lebesgue, on pose $f_n = n\mathbf{1}_{]0,1/n[}$. Alors $\int f_n dx = 1$, et $\liminf f_n = 0$. On pourra chercher à titre d'exercice un contre exemple continu.

Pour montrer la puissance de ces théorèmes, et l'intérêt de ne pas se limiter à des fonctions finies, commençons par introduire une notion.

**Définition.** 18. *On dit qu'une propriété $\mathcal{P}(x)$ a lieu presque partout (ou $\mu$-presque partout s'il y a ambiguïté), si l'ensemble des $x$ pour lesquels $\mathcal{P}(x)$ est fausse est incluse dans un ensemble mesurable de mesure nulle.*

Par exemple

**Proposition 15.** *Si $f$ est une fonction mesurable positive et que $\int f d\mu < \infty$, alors $f(x)$ est finie presque partout.*

*De même, si $f \geq 0$ est telle que $\int f d\mu = 0$, alors $f$ est nulle presque partout.*

*Démonstration.* — Commençons par le premier cas. Appelons $A = \{x \mid f(x) = \infty\}$. Pour tout $n \in \mathbb{N}$, $f \geq n\mathbf{1}_A$, et donc

$$n\mu(A) \leq \int f d\mu.$$

Donc, $\forall n > 0$, $\mu(A) \leq \frac{1}{n}\int f d\mu$, et en passant à la limite, on trouve $\mu(A) = 0$.

Pour le second, on pose de même $A_n = \{x \mid f(x) \geq 1/n\}$, de telle façon que $f \geq \frac{1}{n}\mathbf{1}_{A_n}$. D'où

$$\mu(A_n) \leq n\int f d\mu = 0.$$

Mais $A := \{f > 0\} = \cap_n A_n$ et par conséquent $\mu(A) = 0$. ∎

Voici alors une application immédiate du théorème de convergence monotone, qui est à la base de la loi des grands nombres en probabilité.

**<u>Théorème</u> 10.** *Soit $(f_n)$ une suite de fonctions mesurables positives. Si la série $\sum_n \int f_n d\mu$ converge, alors la série $\sum_n f_n(x)$ converge presque partout. En particulier, $f_n(x)$ converge vers $0$ presque partout.*

*Démonstration.* —

Soit

$$F(x) = \sum_n f_n(x).$$

On ne sait pas a priori si $F(x)$ est fini ou non. Mais le théorème de convergence monotone nous dit que de toutes façons

$$\int F(x)d\mu = \sum_n \int f_n d\mu.$$

Si la série au second membre converge, alors $\int F(x)d\mu < \infty$, et donc $F(x) < \infty$ presque partout.

Pour un $x$ pour lequel la série converge, alors le terme général $f_n(x)$ converge vers 0. ∎

## 3.3 Intégrale de fonctions non positives

Contrairement aux fonctions positives, nous ne pourrons intégrer des fonctions non positives que si elles ne sont pas "trop grosses".

**Définition.** 19. *On dit qu'une fonction mesurable réelle est intégrable si*

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

*On note souvent $f \in \mathcal{L}^1(\mu)$.*

Notons tout de suite qu'une fonction intégrable est finie presque partout.

Rappelons la décomposition $f = f_+ - f_-$ avec $|f| = f_+ + f_-$.

Puisque $f_+$ et $f_-$ sont majorées par $|f|$, alors si $f$ est intégrable, $f_+$ et $f_-$ le sont.

On a alors

**Définition.** 20. *Si $f$ est intégrable, on définit*

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

Dans la définition, on aurait pu prendre n'importe quelle autre décomposition de $f$ en différence de deux fonctions positives intégrables.

**Proposition 16.** *Si $f = f_1 - f_2$, où $f_1$ et $f_2$ sont deux fonctions positives intégrables, alors alors $f$ est intégrable et*

$$\int f d\mu = \int f_1 d\mu - \int f_2 d\mu.$$

*Démonstration.* — C'est immédiat. $f$ est intégrable puisque

$$|f| \leq f_1 + f_2 \in \mathcal{L}^1(\mu)/$$

On écrit $f_+ + f_2 = f_- + f_1$, d'où

$$\int f_+ d\mu + \int f_2 d\mu = \int f_- d\mu + \int f_1 d\mu,$$

et finalement, puisque tout le monde est fini

$$\int f_+ d\mu - \int f_- d\mu = \int f_1 d\mu - \int f_2 d\mu.$$

∎

De ceci nous déduisons immédiatement la

**Proposition 17.** *L'ensemble des fonctions intégrables est un espace vectoriel. Sur cet espace, l'application $f \mapsto \int f d\mu$ est linéaire.*

*Observons en plus qu'elle satisfait toujours*

$$f \leq g \implies \int f d\mu \leq \int g d\mu,$$

*et que*

$$\left| \int f d\mu \right| \leq \int |f|\, d\mu.$$

Le troisième théorème fondamental du calcul intégral est le

**Théorème 11.** *(Théorème de convergence dominée). Soit* $(f_n)$ *une suite de fonctions intégrables. On suppose que pour tout* $x$ $f_n(x)$ *converge vers une fonction* $f(x)$*, et qu'il existe une fonction* **intégrable** *positive* $g$ *telle que, pour tout* $n$

$$|f_n(x)| \leq g(x).$$

*Alors,* $f$ *est intégrable et*

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

*Démonstration.* — Puisque $|f| \leq g$, il est évident que $f$ est intégrable. Quitte à changer $f_n$ en $f_n - f$, on peut toujours supposer que $f = 0$. Ensuite, puisque

$$\int |f_n| \, d\mu \geq \left| \int f_n d\mu \right|,$$

quitte à remplacer $f_n$ par $|f_n|$, on peut toujours supposer que $f_n \geq 0$.

Alors, on pose $h_n = g - f_n$; $h_n$ est positive, et on peut appliquer le Lemme de Fatou à la suite $(h_n)$.

Mais $\liminf h_n = g$, et

$$\liminf_n \int h_n d\mu = \int g d\mu - \limsup_n \int f_n d\mu.$$

Le lemme de Fatou nous dit donc que

$$\int g d\mu \leq \int g d\mu - \limsup_n \int f_n d\mu,$$

d'où

$$\limsup_n \int f_n d\mu = 0.$$

Mais $\int f_n d\mu$ est une suite positive. Sa limite sup ne peut être nulle sans que sa limite inf ne le soit, et dans ce cas la limite existe et est nulle. C'est ce qu'on voulait démontrer.

∎

Dans les situations pratiques, pour appliquer le théorème, la difficulté est de trouver la fonction $g$. La plus petite qui puisse faire l'affaire est bien sûr

$$g = \sup_n |f_n| \, .$$

Mais elle n'est pas toujours très facile à calculer.

Nous allons dans ce qui suit tirer de ces résultats de nombreuses conséquences. mais nous sommes maintenant en mesure d'identifier l'intégrale de Lebesgue avec l'intégrale de Riemann.

**Théorème 12.** *Soit* $f$ *une fonction continue sur l'intervalle* $[0,1]$*, et* $\lambda$ *la mesure de Lebesgue sur* $[0,1]$*. Alors*

$$\int f d\lambda = \int_0^1 f(t) dt.$$

*Démonstration.* — Rappelons brièvement la construction de l'intégrale de Riemann : Une subdivision $\sigma$ de $[0,1]$ est une suite finie $0 = t_0 \leq t_1 \leq \ldots \leq t_n = 1$. Son pas $\tau(\sigma)$ est défini par

$$\tau(\sigma) = \sup_{i=0}^{n-1} |t_{i+1} - t_i| \, .$$

Pour toute subdivision $\sigma$, et tout $i$, on choisit $x_i \in [t_i, t_{i+1}]$, et on forme la somme

$$S_\sigma(f) = \sum_{i=0}^{n-1} f(x_i)(t_{i+1} - t_i).$$

Alors, si $f$ est continue (ou monotone, par exemple), pour toute suite $\sigma_n$ de subdivision dont le pas tend vers 0, la suite $S_{\sigma_n}(f)$ converge vers une quantité notée $\int_0^1 f(t)dt$. Les fonctions Riemann-intégrables sont exactement celes pour lesquelles toutes ces sommes convergent vers la même limite.

**Remarque 11.** *La grosse différence entre la théorie de l'intégrale de Riemann et celle de Lebesgue est que dans la première on approche les fonctions par des fonctions en escalier (combinaison linéaires d'indicatrices d'intervalles) et que dans celle de Lebesgue on approche les fonctions par des fonctions étagées (combinaisons linéaires d'indicatrices de boréliens), qui est une classe beaucoup plus vaste.*

Nous nous concentrons ici sur le cas des fonctions continues, dont on va démontrer qu'elles sont Riemann-intégrables. En fait, nous allons montrer que, pour une suite $\sigma_n$ de subdivisions dont le pas tend vers 0, alors $S_{\sigma_n}(f)$ converge vers $\int f d\lambda$.

Pour cela, on considère la suite de fonctions étagées

$$f_n = \sum_{i=0}^{n-1} f(x_i)\mathbf{1}_{[t_i, t_{i+1}]}.$$

par définition de la mesure de Lebesgue,

$$\int f_n d\lambda = S_{\sigma_n}(f).$$

Il nous reste à montrer que $f_n$ converge vers $f$ et qu'on peut appliquer le théorème de convergence dominée.

Mais le théorème de Heine nous dit que

$$\forall \epsilon > 0, \; \exists \eta, \; |x - y| \leq \eta \implies |f(x) - f(y)| \leq \epsilon.$$

Dès lors, puisque $f_n$ vaut $f(x_i)$ sur l'intervalle $[t_i, t_{i+1}]$ et que $x_i \in [t_i, t_{i+1}]$, si le pas de la subdivision est inférieur à $\eta$, pout tout $x \in [0,1]$,

$$|f_n(x) - f(x)| \leq \epsilon.$$

On voit donc que $f_n$ converge vers $f$. De plus, puisque $f$, continue sur $[0,1]$ est bornée, disons par $M$, alors il en va de même de $f_n$

$$|f_n| \leq M.$$

la fonction constante $M$ est intégrable (sur $[0,1]$, avec la mesure de Lebesgue), et donc nous pouvons appliquer le théorème de convergence dominée

$$\int f_n d\lambda \to \int f d\lambda.$$

C'est bien ce que nous voulions démontrer. ∎

**Remarque 12.** *De façon générale, dans la théorie de Riemann, on cherche à intégrer des fonctions bornées $f : [a,b] \to \mathbb{R}$. Pour cela, pour une subdivision donnée, on considère*

$$S_{*,\sigma}(f) = \sum_{i=0}^{n-1} f_{*i}(t_{i+1} - t_i)$$

*et*

$$S^*_\sigma(f) = \sum_{i=0}^{n-1} f^*_i(t_{i+1} - t_i)$$

*,où $f_{*i}$ et $f^*_i$ désignent respectivement des sup et inf de $f$ sur l'intervalle $[t_i, t_{i+1}]$. Alors, on dit que $f$ est Riemann intégrable si et seulement si, lorsque le pas de la subdivision converge vers 0, $S^*_\sigma(f)$ et $S_{*\sigma}(f)$ convergent vers la même limite.*

*Il n'est pas vrai en général qu'une fonction Riemann intégrable soit mesurable. Il nous faut introduire une tribu plus large que la tribu borélienne, la tribu de Baire. Pour la décrire, on dit qu'un ensemble est négligeable s'il est inclus dans un borélien de mesure de Lebesgue nulle. La tribu de Baire est*

$$\{B \mid \exists A \in \mathcal{B}([a,b]),\ A\Delta B \text{ est négligeable}\}.$$

*On vérifie immédiatement que c'est une tribu. La mesure de Lebesgue s'étend immédiatement à cette tribu en disant que $\lambda(A) = \lambda(B)$ si $A\Delta B$ est négligeable.*

*Alors, une fonction bornée est Riemann intégrable si et seulement si elle est mesurable pour la tribu de Baire, et que l'ensemble de ses points de discontinuité est de mesure de Lebesgue nulle.*

**Théorème 13.** *(Dérivation sous l'intégrale) Soit $I$ un intervalle ouvert de $\mathbb{R}$. On considère une application $f : I \times \Omega \to \mathbb{R}$, $f(x,\omega)$ telle que, pour tout $x$, la fonction $\omega \mapsto f(x,\omega)$ soit mesurable et intégrable, et telle que pour tout $\omega$, la fonction $x \mapsto f(x,\omega)$ soit dérivable sur $I$. On note $\partial_x f(x,\omega)$ sa dérivée.*

*Supposons qu'il existe une fonction positive intégrable $g(\omega)$ telle que, pour tout $x \in I$, $|\partial_x f(x,\omega)| \le g(\omega)$. Alors $I(x) = \int f(x,\omega)d\mu(\omega)$ est dérivable sur l'intervalle $I$ et sa dérivée est*

$$I'(x) = \int \partial_x f(x,\omega) d\mu(\omega).$$

*Démonstration.* — Nous laissons au lecteur le soin de montrer que $\partial_x f$ est mesurable. Nous choisissons $x \in I$, une suite $x_n$ qui converge vers $x$, et considérons

$$\frac{I(x_n) - I(x)}{x_n - x} = \int \frac{f(x_n,\omega) - f(x,\omega)}{x_n - x} d\mu(\omega).$$

Nous aurons montré le résultat si nous pouvons appliquer le théorème de convergence dominée à la suite

$$f_n(\omega) = \frac{f(x_n,\omega) - f(x,\omega)}{x_n - x},$$

puisque par définition de la dérivée, $f_n(x, \omega)$ converge vers $\partial_x f(x, \omega)$.

Mais le théorème des accroissements finis, appliqué à la fonction $x \mapsto f(x, \omega)$ ($\omega$ ici étant fixé) nous dit qu'il existe un point $y_n(\omega)$ de l'intervalle $[x_n, x]$ (ou bien $[x, x_n]$ selon les cas) tels que $f_n(\omega) = \partial_x f(y_n(\omega), \omega)$. L'hypothèse nous dit alors que $|f_n(\omega)| \leq g(\omega)$, et nous pouvons conclure.

∎

On peut ainsi appliquer ceci aux séries de fonctions, lorsque l'espace $\Omega$ est $\mathbb{N}$, muni de la mesure de décompte.

**Proposition 18.** *Soit $(f_n(x))$ une suite de fonctions dérivables définies sur un intervalle ouvert $I$. On suppose que, pour tout $x$ la série $S(x) = \sum_n f_n(x)$ converge et qu'il existe une série de réels $(\alpha_n)$ telle que pour tout entier $n$, et tout $x \in I$,*

$$|f_n'(x)| \leq \alpha_n$$

*avec $\sum_n \alpha_n < \infty$.*

*Alors, $S(x)$ est dérivable sur $I$ et*

$$S'(x) = \sum_n f_n'(x).$$

## 3.4 Mesures absolument continues

Une fois que nous avons construit une mesure sur un espace, alors nous pouvons en construire beaucoup d'autres.

**Définition.** 21. *Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré, et $\phi$ une fonction mesurable positive.*

*Alors, la formule*

$$\nu(A) = \int \mathbf{1}_A \phi d\mu = \int_A \phi d\mu$$

*définit une nouvelle mesure sur $(\Omega, \mathcal{A})$. Cette mesure est notée $\phi d\mu$ et est dite* **absolument continue** *par rapport à $\mu$.*

*Remarquer que, si $\mu(A) = 0$, alors $\nu$ est identiquement nulle.*

*La fonction $\phi$ s'appelle la* **densité** *de $\nu$ par rapport à $\mu$. Elle est souvent notée*

$$\phi = \frac{d\mu}{d\nu}.$$

*Deux mesures sont* **équivalentes** *si elles sont absolument continues l'une par rapport à l'autre.*

Montrons quelques exemples.

1. Si $\phi = \mathbf{1}_A$, $\phi d\mu$ n'est rien d'autre que la mesure $\mu$ restreinte à $A$.

2. Sur $\mathbb{R}$ muni de la mesure de Lebesgue, la mesure

$$\gamma(dx) = \frac{\exp(-\frac{x^2}{2})}{\sqrt{2\pi}} dx$$

s'appelle la mesure gaussienne. C'est une mesure de probabilité et elle joue un rôle central en calcul des probabilités. Elle est équivalente à la mesure de Lebesgue.

3. La mesure $\mathbf{1}_{x>0}\frac{1}{x}dx$ est une mesure qui n'est pas une mesure de Radon.

Pour calculer l'intégrale par rapport à $\phi d\mu$, nous avons la

**Proposition 19.** *Soit $\nu = \phi d\mu$ une mesure absolument continue par rapport à $\mu$. Alors, pour toute fonction mesurable positive $f$, on a*

$$\int f d\nu = \int (f\phi) d\mu.$$

*De plus, pour une fonction $f$ non nécéssairement positive, $f \in \mathcal{L}^1(\nu)$ si et seulement si $f\phi \in \mathcal{L}^1(\mu)$ et dans ce cas, on a encore*

$$\int f d\nu = \int f\phi d\mu.$$

*Démonstration.* — Commençons par le premier point, le second s'en déduit immédiatement.

La formule est vraie pour $f = \mathbf{1}_A$ par définition de la mesure $\nu$. Elle est donc vraie pour toute fonction étagée positive par linéarité.

Soit alors $f$ une fonction mesurable positive quelconque, et $f_n$ une suite de fonctions mesurables positives étagées qui converge en croissant vers $f$. D'après ce que l'on vient de voir,

$$\int f_n d\nu = \int f_n \phi d\mu,$$

et il ne nous reste plus qu'à passer à la limite dans les deux membres en utilisant le théorème de convergence monotone. ∎

Pour énoncer le principal résultat de cette section, introduisons une nouvelle notion.

**Définition.** 22. *On dit que deux mesures positives $\mu$ et $\nu$ sont* **étrangères** *s'il existe un ensemble mesurable $A$ tel que $\mu(A) = 0$ et $\nu(A^c) = 0$.*

*En d'autres termes, $\mu$ est portée par $A^c$ tandis que $\nu$ est portée par $A$.*

**Théorème 14** (Radon-Nikodym)**.** *Soit $(\Omega, \mathcal{A})$ un ensemble mesurable et $\mu$ et $\nu$ deux mesures positives. On suppose $\mu$ $\sigma$-finie. Alors, il existe une mesure $\mu_1$ étrangère à $\mu$ et une fonction mesurable positive $\phi$ telle que*

$$\nu = \phi\mu + \mu_1,$$

*et cette décomposition est unique.*

Nous n'allons pas démontrer ce théorème, dont la démonstration est donnée en appendice.

Ce théorème est souvent utilisé sous la forme suivante.

**Corollaire 1.** *Pour que $\nu$ soit absolument continue par rapport à $\mu$, il faut et il suffit que*

$$\mu(A) = 0 \implies \nu(A) = 0.$$

*Démonstration.* — (Du corollaire) La condition est bien évidement nécéssaire. Pour la réciproque, il suffit en appliquant le théorème précédent de remarquer que sous ces hypothèses, $\mu_1 = 0$. ∎

Sur $\mathbb{R}$, nous avons vu que toute mesure bornée se représente par une fonction de répartition $F_\mu$, c'est à dire une fonction croissante, continue à droite, avec $F_\mu(-\infty) = 0$ et $F_\mu(+\infty) < \infty$.

**Remarque 13.** *Alors, si $F_\mu$ est dérivable et de dérivée continue $f$, alors $\mu$ est une mesure absolument continue par rapport à la mesure de Lebesgue $\lambda$ de densité $f$. On note encore*

$$dF_\mu(x) = F'_\mu(x)dx.$$

*Il en va de même pour les mesures de Radon dont la fonction associée $F$ est dérivable de dérivée continue.*

*Pour le voir, il suffit de constater que les mesures $\mu$ et $f\lambda$ coïncident sur les intervalles $]a, b]$. Or, on a*

$$\mu(]a, b]) = F_\mu(b) - F_\mu(a)$$

*par définition de $F_\mu$, alors que, pour la mesure $\nu = f\lambda$, on a*

$$\nu(]a, b]) = \int \mathbf{1}_{]a,b]} f(x)dx = F_\mu(b) - F_\mu(a).$$

*Bien sûr, il n'est pas nécessaire pour que ce soit vrai que la fonction $F_\mu$ aie une dérivée continue. L'important est qu'elle soit l'intégrale de sa dérivée. Ce qui sera le cas par exemple dès qu'elle aura une dérivée bornée avec un nombre fini de points de discontinuités.*

Remarquons enfin qu'une mesure $\sigma$ finie est toujours équivalente à une mesure de probabilité.

Pour le voir, considérons une mesure $\mu$ $\sigma$-finie et choisissons une suite $(B_n)$ croissante, telle que $\mu(B_n) < \infty$, et telle que $\cup_n B_n = \Omega$.

On peut bien sûr supposer que tous les $\mu(B_n)$ sont non nuls à partir d'un certain rang, sinon $\mu(\Omega) = 0$ et il n'y a rien à dire. Quitte à supprimer les premiers termes de la suite, on peut donc supposer que tous les $\mu(B_n)$ sont non nuls.

Considérons alors la mesure

$$\sum_n (\mathbf{1}_{B_n} \frac{1}{2^{n+1}\mu(B_n)})\mu.$$

On voit que $\nu$ est une mesure de probabilité, qu'elle est absolument continue par rapport à $\mu$, et de densité strictement positive partout. Elle est donc équivalente à $\mu$.

## 3.5 Mesures images

**Définition.** 23. *Soit $(\Omega_1, \mathcal{A}_1, \mu_1)$ un espace mesuré, et $f$ une application mesurable de $(\Omega_1, \mathcal{A}_1)$ dans $(\Omega_2, \mathcal{A}_2)$. Alors, la formule $\mu_2(B) = \mu_1(f^{-1}(B))$ définit une mesure sur $(\Omega_2, \mathcal{A}_2)$. On l'appelle la mesure image de $\mu_1$ par $f$, et on la note $f(\mu_1)$ ou parfois aussi $f^*\mu_1$.*

Il est immédiat de vérifier l'assertion que $\mu_2$ est bien une mesure.

En voici quelques exemples.

1. L'image de la mesure de Lebesgue sur $\mathbb{R}$ par l'application $y \mapsto y + x$ (la translation par $x$) est la mesure de Lebesgue (c'est en fait ce que traduit l'invariance par translation de la mesure de Lebesgue).

2. Si $\mu_1(\Omega_1) = K$ est fini, et que $f$ est constante et vaut $x_2$, alors l'image de $\mu_1$ par $f$ est $K\delta_{x_2}$.

3. Sur $\mathbb{R}$, $f(x) = -x$, l'image de la mesure de Lebesgue par $f$ est encore la mesure de Lebesgue.

   Pour le voir, on constate que, si on appelle $\nu$ la mesure image
   $$\nu([a,b]) = \lambda([-b,-a]) = b - a,$$
   et donc $\nu$ et $\lambda$ coïncident sur les intervalles.

4. Sur $\mathbb{R}$, si $f(x) = cx$, $c \neq 0$, alors l'image de la mesure de Lebesgue $\lambda$ par $f$ est $\frac{1}{|c|}\lambda$.

   Pour le voir, on se ramène au cas où $c > 0$ (quitte à composer par $x \mapsto -x$), et ensuite, on constate que, si $\nu = f(\lambda)$,
   $$\nu([a,b]) = \lambda([\frac{a}{c}, \frac{b}{c}]) = \frac{1}{c}\lambda[a,b].$$

5. Sur $\mathbb{R}$, l'image de la mesure de Lebesgue $\lambda$ par $f(x) = 0$ est $\infty\delta_0$. (C'est en quelque sorte le cas limite de l'exemple précédent lorsque $c \to 0$). Ceci montre que l'image d'une mesure $\sigma$-finie, et même d'une mesure de Radon, n'est pas nécésairement $\sigma$-finie.

6. Sur $\mathbb{R}$ si $f$ est croissante strictement ou strictement décroissante, dérivable et de dérivée continue, alors l'image de la mesure de Lebesgue $\lambda$ par $f$ est la mesure à densité $|g'(x)|\,\lambda(dx)$, où $g$ est la fonction réciproque de $f$.

   Pour le voir, on se ramène au cas où $f$ est croissante (quitte à composer par $x \mapsto -x$), et dans ce cas, en appelant $g$ la fonction réciproque de $f$ et $\nu$ la mesure image, on écrit
   $$\nu([a,b]) = \lambda([g(a), g(b)]) = \int_{[a,b]} g'(t)\lambda(dt).$$

7. De même, la mesure image de la mesure $h(x)dx$ par une fonction monotone $f$ comme plus haut est $h(g(x))\,|g'(x)|\,dx$, où $g$ est la fonction réciproque de $f$.

8. Lorsqu'une fonction $f : \mathbb{R} \to \mathbb{R}$ n'est pas monotone mais quand même régulière, on la décompose en intervalles où elle est monotone. Sur chaque intervalle, on peut calculer comme plus haut la mesure image de la mesure de Lebesgue restreinte à cet intervalle, puis on fait la somme des mesures ainsi obtenues.

   Ainsi, la mesure image de la mesure de Lebesgue sur $\mathbb{R}$ par la fonction $f(x) = x^2$ est $\frac{1}{\sqrt{x}}\mathbf{1}_{x\geq 0}dx$.

Pour calculer les mesures images en général, nous disposons du théorème

**Théorème 15** (Théorème du tranfert)**.** *Soit $f$ une fonction mesurable de $(\Omega_1, \mathcal{A}_1, \mu_1)$ dans $(\Omega_2, \mathcal{A}_2)$, et $\mu_2 = f(\mu_1)$. Alors, pour toute fonction mesurable positive $g : \Omega_2 \to \mathbb{R}$,*
$$\int_{\Omega_2} g d\mu_2 = \int_{\Omega_1} g(f) d\mu_1.$$

*De plus, une fonction $g$ est intégrable par rapport à $\mu_2$ si et seulement si $g(f)$ est intégrable par rapport à $\mu_1$ et dans ce cas on a encore*

$$\int_{\Omega_2} g d\mu_2 = \int_{\Omega_1} g(f) d\mu_1.$$

*Démonstration.* — Le cas des fonctions non positives découle immédiatement du cas des fonctions positives, en décomposant $g$ en $g_+ - g_-$. On n'a donc à traiter que le cas des fonctions positives.

Le résultat est immédiat pour $g = \mathbf{1}_B$, car alors $g(f) = \mathbf{1}_{f^{-1}(B)}$ et le résultat n'est rien d'autre que la définition de $\mu_2$.

Par linéarité, il reste donc vrai pour les fonctions $g$ étagées.

Soit maintenant $g$ une fonction mesurable positive; $g$ est limite croissante d'une suite $(g_n)$ de fonctions étagées, et alors $g(f)$ est limite croissante des fonctions étagées $g_n(f)$. Le résultat découle alors du théorème de convergence monotone. ∎

**Remarque 14.** *On retrouve ainsi, pour une fonction monotone croissante $\mathcal{C}^1$ $g : \mathbb{R} \to \mathbb{R}$*

$$\int f(g) g'(x) dx = \int f(x) dx,$$

*en appliquant le théorème du transfert à la mesure $g'(x)dx$ dont la mesure image par $g$ est $dx$ (puisque $g'(x)dx$ est la mesure image de $dx$ par la fonction $g^{-1}$).*

## 3.6 Mesures signées ou complexes bornées

Jusqu'ici, nous ne nous sommes inéteresés qu'à des mesures à valeurs des $[0, \infty]$. Dans de nombreuses situations, il est commode d'envisager des mesures à valeurs réelles, ou meme complexes.

**Définition.** 24. *Une mesure signée bornée sur l'espace mesurable $(\Omega, \mathcal{A})$ est une application $\mu$ de $\mathcal{A}$ dans un compact $[-M, M]$ de $\mathbb{R}$ telle que*

1. *$\mu(\emptyset) = 0$;*
2. *Si $A_n$ est une suite d'éléments de $\mathcal{A}$ deux à deux disjoints, alors la série $\sum_n \mu(A_n)$ converge et*

$$\sum_n \mu(A_n) = \mu(\cup_n A_n).$$

Par exemple, la différence $\mu_1 - \mu_2$ de deux mesures bornées est une mesure signée.

Un autre exemple de mesure bornée est donnée par

$$\nu(A) = \int \mathbf{1}_A f d\mu,$$

où $\mu$ est une mesure positive et $f \in L^1(\mu)$. On dit alors, comme dans le cas où $f \geq 0$, que $\nu$ est absolument continue par rapport à $\mu$ et on note

$$d\nu = f d\mu.$$

**Remarque 15.** *Il revient au même de dire que* $\mu(\emptyset) = 0$ *et que*

$$\mu(A) = \sum_n \mu(A_n)$$

*pour toute partition dénombrable* $(A_n)$ *de* $A$*, où la série* $\sum_n \mu(A_n)$ *est convergente. Nous verrons plus bas qu'en fait la série* $\sum_n |\mu(A_n)|$ *est en fait toujours convergente.*

On peut de même considérer des mesures à valeurs complexe.

**Définition.** 25. *Une mesure complexe bornée sur l'espace mesurable* $(\Omega, \mathcal{A})$ *est une application* $\mu$ *de* $\mathcal{A}$ *dans un compact* $K$ *de* $\mathbb{C}$ *(ou ce qui revient au même dans une boule* $B(0, R)$ *de* $\mathbb{C}$ *centrée en* $0$*) telle que*

1. $\mu(\emptyset) = 0$*;*
2. *Si* $A_n$ *est une suite d'éléments de* $\mathcal{A}$ *deux à deux disjoints, alors la série* $\sum_n \mu(A_n)$ *converge et*

$$\sum_n \mu(A_n) = \mu(\cup_n A_n).$$

**Remarque 16.** *On voit immédiatement que les parties réelles et imaginaires d'une mesure bornée sont des mesures signées, bornées, et donc qu'une mesure complexe boprnée n'est rien d'autre que* $\mu_1 + i\mu_2$*, où* $\mu_1$ *et* $\mu_2$ *sont des mesures signées bornées.*

Le théorème fondamental sur les mesures bornées est le suivant, qui permet de ramener immédiatement l'étude des mesures signées bornées à celle des mesures positives bornées. Nous en donn ons la démonstration en appendice.

**Théorème 16** (Jordan-Hahn)**.** *Soit* $\mu$ *une mesure signée bornée sur* $(\Omega, \mathcal{A})$*. Alors il existe deux mesures positives* $\mu_+$ *et* $\mu_-$ *qui sont étrangères (portées par des ensembles disjoints), telle que* $\mu = \mu_+ - \mu_-$*.*

*Dans ce cas, la mesure* $|\mu| = \mu_+ + \mu_-$ *est une mesure bornée (appelée mesure de variation totale de* $\mu$*), et le nombre* $\|\mu\| = \mu_+(\Omega) + \mu_-(\Omega)$ *est appelé la norme de* $\mu$*.*

Dans le cas particulier où

$$\nu(A) = \int \mathbf{1}_A f d\mu,$$

alors

$$\nu_+(A) = \int \mathbf{1}_A f_+ d\mu, \ \nu_-(A) = \int \mathbf{1}_A f_- d\mu$$

et

$$\|\nu\| = \int |f| \, d\mu = \|f\|_{L^1(\mu)} \, .$$

Le théorème de Jordan-Hahn nous montre que le cas particulier ci-dessus est en fait le cas général, puisque si $\mu$ est une mesure bornée et si $\mu_1 = |\mu|$, alors les mesures $\mu_+$ et $\mu_-$ sont absolument continues par rapport à $\mu_1$, et si on dénote par $A$ un ensemble tel que $\mu_+(A) = \mu_-(A^c) = 0$, on a

$$d\mu = (\mathbf{1}_{A^c} - \mathbf{1}_A) d\mu_1.$$

On peut donc toujours se ramener au cas où $d\mu = fd\nu$, pour une mesure positive $\nu$, et dans ce cas l' intégrale

$$\int hd\mu = \int (hf)d\mu$$

est définie pour toute les fonctions $h$ telles que $|h| \in L^1(|\mu|)$.

On peut aussi considérer l'image d'une mesure bornée par une application mesurable, le produit tensoriel de mesures bornées, etc.

On peut aussi s'intéresser aux mesures bornées à valeurs complexes, et elles seront décomposées en $\mu_1 + i\mu_2$, où $\mu_1$ et $\mu_2$ sont des mesures réelles bornées. Elles s'écrivent toujours elles aussi sous la forme

$$\mu(A) = \int \mathbf{1}_A f d\nu,$$

où $\nu$ est une mesure positive (qu'on peut même choisir bornée) et $f$ une fonction à valeurs complexe dans $L^1(\nu)$.

L'avantage de prendre des mesures bornées est qu'elles forment un espace vectoriel. Muni de la norme $\|\mu\|$, cet espace vectoriel est un espace de Banach, comme le montre la proposition suivante.

**Proposition 20.** *Soit $\mu_n$ une suite de Cauchy de mesures bornées. Alors, elle converge vers une mesure bornée.*

*Démonstration.* — Il suffit de considérer une mesure positive $\nu$ par rapport à laquelle toutes les mesures bornées sont absolument continues, par exemple

$$\nu = \sum_n \frac{|\mu_n|}{2^n \|\mu\|_n}.$$

Alors, les densités $f_n$ de $\mu_n$ par rapport à $\nu$ forment une suite de Cauchy dans $L^1(\nu)$. Nous verrons plus bas (théorème de Riesz-Fisher, 23 ) que l'espace $L^1(\nu)$ est complet, pour toute mesure positive $\sigma$-finie. Cette suite converge dans $L^1(\nu)$ vers $f$, et $\mu_n$ converge vers $fd\nu$, comme on le vérifie immédiatement.

∎

# 4 Espaces produits

## 4.1 Tribu produit

Dans cette section, nous considérons deux espaces mesurables $(\Omega_1, \mathcal{A}_1)$ et $(\Omega_2, \mathcal{A}_2)$ et nous considérons le produit $\Omega_1 \times \Omega_2$, que nous allons munir d'une tribu

**Définition.** 26. *La tribu $\mathcal{A}_1 \otimes \mathcal{A}_2$ est la plus petite tribu sur le produit $\Omega_1 \times \Omega_2$ qui contienne les ensembles $A \times B$, où $A \in \mathcal{A}_1$ et $B \in \mathcal{A}_2$.*

Cette tribu peut être caractérisée de beaucoup d'autres façons différentes, par exemple

**Proposition 21.** *1. Si $\mathcal{A}_1 = \sigma(\mathcal{E}_1)$ et $\mathcal{A}_2 = \sigma(\mathcal{E}_2)$, et si l'on suppose que $\Omega_i \in \mathcal{E}_i$, $i = 1, 2$, alors*

$$\mathcal{A}_2 \otimes \mathcal{A}_2 = \sigma(A \times B,\ A \in \mathcal{E}_1,\ B \in \mathcal{E}_2).$$

*2. Si $\pi_1$ et $\pi_2$ désignent les projections de $\Omega_1 \times \Omega_2$ sur $\Omega_1$ et $\Omega_2$ respectivement, alors*

$$\mathcal{A}_1 \otimes \mathcal{A}_2 = \sigma(\pi_1, \pi_2).$$

*Démonstration.* — Commençons par le premier point. Appelons $\mathcal{A} = \sigma(A \times B,\ A \in \mathcal{E}_1,\ B \in \mathcal{E}_2)$. Alors clairement, $\mathcal{A} \subset \mathcal{A}_1 \otimes \mathcal{A}_2$.

Par ailleurs

$$\{B \in \mathcal{A}_1,\ B \times \Omega_2 \in \mathcal{A}\}$$

est une tribu, comme on le vérifie immédiatement. Puisqu'elle contient $\mathcal{E}_1$, elle contient $\mathcal{A}_1$. On voit donc que, pour tout $A_1 \in \mathcal{A}_1$, $A_1 \times \Omega_2 \in \mathcal{A}$.

De même, pour tout $B_2 \in \mathcal{A}_2$, $\Omega_1 \times B_2 \in \mathcal{A}$.

On a donc aussi, pour tout $A_1 \in \mathcal{A}_1$ et pour tout $A_2 \in \mathcal{A}_2$,

$$A_1 \times A_2 = (A_1 \times \Omega_2) \cap (\Omega_1 \times A_2) \in \mathcal{A},$$

et par conséquent $\mathcal{A}_1 \otimes \mathcal{A}_2 \subset \mathcal{A}$.

Pour le deuxième point, nous remarquons que $\pi_1^{-1}(A_1) = A_1 \times \Omega_2$, et $\pi_2^{-1}(A_2) = \Omega_1 \times A_2$. Ceci nous montre d'une part que $\pi_1$ et $\pi_2$ sont mesurables par rapport à $\mathcal{A}_1 \times \mathcal{A}_2$, et que la plus petite tribu qui rend mesurable à la fois $\pi_1$ et $\pi_2$ rend mesurable tous les ensembles

$$A_1 \times A_2 = \pi_1^{-1}(A_1) \cap \pi_2^{-1}(A_2),$$

où $A_i \in \mathcal{A}_i$. par conséquent

$$\mathcal{A}_1 \otimes \mathcal{A}_2 \subset \sigma(\pi_1, \pi_2).$$

∎

De la même manière, il est aisé de caractériser les fonctions mesurables à valeurs dans le produit $\Omega_1 \times \Omega_2$.

**Proposition 22.** *Une fonction $f : (\Omega, \mathcal{A}) \to (\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$ est mesurable si et seulement si les applications composées $\pi_1(f)$ et $\pi_2(f)$, à valeurs dans $\Omega_1$ et $\Omega_2$ respectivement, sont mesurables.*

Nous laissons la démonstration à titre d'exercice. Ainsi, un couple de fonction mesurable n'est rien d'autre qu'une fonction mesurable à valeurs dans le produit.

Remarquons que nous pouvons faire des produits successifs, et que le résultat est indépendant de l'ordre dans lequel nous faisons ces produits

**Proposition 23.** *Si* $(\Omega_1, \mathcal{A}_1)$*,* $(\Omega_2, \mathcal{A}_2)$*,* $(\Omega_3, \mathcal{A}_3)$ *sont trois espaces mesurables, alors, sur l'espace* $\Omega_1 \times \Omega_2 \times \Omega_3$*, on a*

$$(\mathcal{A}_1 \otimes \mathcal{A}_2) \otimes \mathcal{A}_3 = \mathcal{A}_1 \otimes (\mathcal{A}_2 \otimes \mathcal{A}_3).$$

*Cette tribu est encore*

$$\sigma(A_1 \times A_2 \times A_3,\ A_i \in \mathcal{A}_i,\ i = 1, 2, 3\}$$

*ou bien encore* $\sigma(\pi_1, \pi_2, \pi_3)$ *où* $\pi_i$ *désigne la projection sur* $\Omega_i$*.*

*Démonstration.* — Démonstration laissée au lecteur à titre d'exercice. ∎

La tribu produit peut encore se voir de bien des façons différentes. Ainsi, sur $\mathbb{R}^2$, la tribu $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^2)$ est encore la plus petite tribu qui contient toutes les boules, ou toutes les boules ouvertes, ou tous les ouverts, ou encore toutes les boules ouvertes à centre à coordonnées rationnelles et à rayon rationnel, etc. Montrons par exemple que c'est la plus petite tribu qui contienne les boules ouvertes.

Si $B(x, r)$ est la boule ouverte centrée en $x$ et de rayon $r$, nous écrivons

$$B(x, r) = \cup_I ]a, b[\times]a', b'[,$$

où $I$ est l'ensemble de tous les quadruplets de rationnels tels que $]a, b[\times]a', b'[\subset B(x, r)$.

Réciproquement, pour tout quadruplet de réels,

$$]a, b[\times]a', b'[= \cup_J B(x, r),$$

où $J$ est l'ensemble des points $x = (x_1, x_2) \in \mathbb{Q}^2$ et des points $r \in Q$, $r > 0$ tels que $B(x, r) \subset ]a, b[\times]a', b'[$.

Remarquons encore que les ensembles mesurables dans le produit ont des coupes mesurables.

**Proposition 24.** *Si* $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ *et* $\omega_1 \in \Omega_1$*, alors*

$$A_{\omega_1} = \{\omega_2,\ (\omega_1, \omega_2) \in A\} \in \mathcal{A}_2.$$

*On a bien sûr la propriété symétrique en échangeant les rôles de* $x_1$ *et* $x_2$*.*

*De même, si* $f(\omega_1, \omega_2)$ *est une fonction mesurable par rapport à* $\mathcal{A}_1 \otimes \mathcal{A}_2$ *à valeurs réelles, pour tout* $\omega_1 \in \Omega_1$*, la fonction* $f_{\omega_1}(\omega_2) : \Omega_2 \to \mathbb{R}$ *définie par*

$$f_{\omega_1}(\omega_2) = f(\omega_1, \omega_2)$$

*est mesurable par rapport à* $\mathcal{A}_2$*.*

*De plus, si* $f$ *est mesurable positive, et que* $\mu_2$ *est une mesure* $\sigma$*-finie sur* $(\Omega_2, \mathcal{A}_2)$*, la fonction*

$$\omega_1 \mapsto \int f_{\omega_1}(\omega_2) d\mu_2(\omega_2)$$

*est mesurable par rapport à* $\mathcal{A}_1$*.*

*Démonstration.* — Démontrons d'abord la première assertion.

Nous commençons par remarquer que la propriété est vraie si $A = A_1 \times A_2$, car alors $A_{\omega_1}$ est soit vide, soit égal à $A_2$, selon que $\omega_1$ est ou non dans $A_1$;

Ensuite, on considère

$$\mathcal{A} = \{A \ , \ A_{\omega_1} \in \mathcal{A}_2\}.$$

$\mathcal{A}$ est une tribu, qui contient tous les ensembles de la forme $A_1 \times A_2$, avec $A_i \in \mathcal{A}_i$, et qui contient donc toute la tribu $\mathcal{A}_1 \times \mathcal{A}_2$.

Pour la seconde assertion, nous remarquons que

$$f_{\omega_1}^{-1}(A) = (f^{-1}(A))_{\omega_1}.$$

Enfin, pour le dernier point, commençons par traiter le cas où $\mu_2$ est une mesure de probabilité. Nous commençons par le cas où $f = \mathbf{1}_{A_1 \times A_2}$, auquel cas la fonction $\int_{\Omega_2} f_{\omega_1} d\mu_2$ n'est rien d'autre que $\mathbf{1}_{A_1}(\omega_1)\mu_2(A_2)$.

Ensuite, nous traitons le cas où $f = \mathbf{1}_A$, avec $A \in \mathcal{A}_1 \times \mathcal{A}_2$. Dans ce cas, le résultat est obtenu à partir du précédent en remarquant que l'ensembles des $A$ pour lesquels $\int \mathbf{1}_{A\omega_1} d\mu_2$ est mesurable est une classe monotone. C'est ici que sert le fait que $\mu_2$ est une probabilité, pour montrer que si $A \subset B$ sont des éléments de cette classe, alors $B \setminus A$ est encore dans la classe, par la formule

$$\int \mathbf{1}_{B\setminus A\omega_1} d\mu_2 = \int \mathbf{1}_{B\omega_1} d\mu_2 - \int \mathbf{1}_{A\omega_1} d\mu_2.$$

(les autres points à vérifier sont évidents.)

Cette classe monotone contient les ensembles de la forme $A_1 \times A_2$, $A_i \in \mathcal{A}_i$, qui forment une classe stable par intersection, et donc toute la tribu $\mathcal{A}_1 \otimes \mathcal{A}_2$.

Ensuite, pour $f$ étagée, le résultat se déduit du précédent par linéarité.

Enfin, pour une fonction mesurable positive quelconque, le résultat s'obtient par passage à la limite croissante à partir du cas étagé.

Pour passer au cas où $\mu_2$ est seulement $\sigma$-finie, on écrit $d\mu_2 = h(\omega_2)d\nu_2$, où $\nu_2$ est une probabilité, et on applique ce qui précède à $\nu_2$ et à la fonction $g(\omega_1, \omega_2) = f(\omega_1, \omega_2)h(\omega_2)$.

∎

**Remarque 17.**

*Nous avons vu qu'un ensemble mesurable dans le produit a des coupes mesurables. Mais il n'est pas suffisant qu'un ensemble aie des coupes mesurables pour être mesurable dans le produit. Pour voir un contre exemple, considérons un ensemble $B$ non mesurable dans $\mathbb{R}$, et l'application diagonale $f$ de $\mathbb{R}$ dans $\mathbb{R}^2$ définie par $f(x) = (x, x)$, qui est bien évidement mesurable. Alors, l'ensemble*

$$B_1 = f(B) = \{(x, x) \ , \ x \in B\}$$

*n'est pas mesurable dans le produit, puisque $B = f^{-1}(B_1)$ et n'est pas lui même mesurable par hypothèse. Mais l'ensemble $B_1$ a toutes ses coupes soit vides soit réduites à un point, donc toutes ses coupes sont mesurables.*

*De même, nous avons vu que les projections $\pi_1$ et $\pi_2$ sur les facteurs sont mesurables, mais en général pour un ensemble mesurable $A$ dans le produit, ni $\pi_1(A)$ ni $\pi_2(A)$ ne sont mesurables.*

## 4.2 Mesure produit

Sur un produit d'espaces mesurés $\sigma$-finis $(\Omega_1, \mathcal{A}_1, \mu_1)$ et $(\Omega_2, \mathcal{A}_2, \mu_2)$ on peut construire une mesure grâce à la proposition suivante

**Proposition 25.** *Il existe une unique mesure $\mu$ sur $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$ telle que, pour tous les couples $(A_1, A_2) \in \mathcal{A}_1 \times \mathcal{A}_2$, on ait*

$$\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2).$$

*Cette mesure est appelée mesure produit et notée $\mu_1 \otimes \mu_2$.*
*Pour tout $A \in \mathcal{A}_1 \times \mathcal{A}_2$, on a*

$$\mu_1 \otimes \mu_2(A) = \int_{\Omega_1} \mu_2(A_{\omega_1}) d\mu_1(\omega_1) = \int_{\Omega_2} \mu_1(A_{\omega_2}) d\mu_2(\omega_2).$$

*Démonstration.* — Définissons

$$\nu(A) = \int_{\Omega_1} \mu_2(A_{\omega_1}) d\mu_1(\omega_1),$$

et montrons que cette quantité définit une mesure sur $\mathcal{A}_1 \otimes \mathcal{A}_2$.

Puisque $\emptyset_{\omega_1} = \emptyset$, on a $\nu(\emptyset) = 0$.

Si $A_1 \cap A_2 = \emptyset$, alors $A_{1\omega_1} \cap A_{2\omega_1} = \emptyset$ pour tout $\omega_1$, et donc

$$\mu_2((A_1 \cup A_2)_{\omega_1}) = \mu_2(A_{1\omega_1}) + \mu_2(A_{2\omega_1}),$$

et par suite

$$\nu(A_1 \cup A_2) = \nu(A_1) + \nu(A_2).$$

Soit maintenant une suite croissante $A_n$, et $A = \cup_n A_n$. Pour tout $\omega_1$, $A_{n\omega_1}$ est une suite croissante dont l'union vaut $A_{\omega_1}$, et donc, puisque $\mu_2$ est une mesure $\mu_2(A_{n\omega_1})$ converge en croissant vers $\mu_2(A_{\omega_1})$. Le théorème de convergence monotone nous montre alors que $\nu(A_n) = \int \mu_2(A_{n\omega_1}) d\mu_1$ converge en croissant vers $\nu(A) = \int \mu_2(A_{\omega_1}) d\mu_1$.

Pour $A = A_1 \times A_2$, $\mu_2(A_{\omega_1}) = \mathbf{1}_{A_1}(\omega_1)\mu_2(A_2)$, et donc

$$\nu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2).$$

Ceci montre l'existence de la mesure $\mu$. L'unicité est assurée par le fait que la classe des ensembles $A_1 \times A_2$ est stable par intersection et engendre la tribu, et que la mesure ainsi construite est $\sigma$-finie.

∎

Nous pouvons de même construire une mesure produit sur le produit d'un nombre fini d'espaces mesurés.

Nous avons alors le théorème fondamental

**Théorème 17.** *(De Fubini-Tonelli) Soit $f$ une fonction mesurable positive sur $\Omega_1 \times \Omega_2$, et $\mu = \mu_1 \otimes \mu_2$.*
*Alors,*

$$\int_{\Omega_1 \times \Omega_2} f d\mu = \int_{\Omega_1} (\int_{\Omega_2} f_{\omega_1} d\mu_2) d\mu_1 = \int_{\Omega_2} (\int_{\Omega_1} f_{\omega_2} d\mu_1) d\mu_2$$

*Le même résultat reste vrai pour une fonction non positive,* **à condition que $|f|$ soit intégrable pour la mesure $\mu$.**

*Cette condition se vérifie bien sûr en appliquant le théorème aux fonctions positives.*

*En d'autres termes, pour intégrer sur le produit par rapport à la mesure produit, on intègre d'abord en $\omega_1$, puis en $\omega_2$, ou l'inverse.*

*Démonstration.* — Commençons par le cas positif. Le théorème est vrai pour $f = \mathbf{1}_A$, où $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$, par définition de la mesure produit.

Par linéarité, il reste vrai pour les fonctions étagées.

Si $f$ est mesurable positive quelconque, il existe une suite croissante $(f_n)$ de fonctions étagées qui converge vers $f$.

Alors, $f_{n\omega_1}$ converge en croissant vers $f_{\omega_1}$ et, pour tout $\omega_1$ $h_n(\omega_1) = \int_{\Omega_2} f_{n\omega_1} d\mu_2$ converge en croissant vers $h(\omega_1) = \int_{\Omega_2} f_{\omega_1} d\mu_2$.

Par conséquent, $\int h_n(\omega_1) d\mu_1$ converge en croissant vers $\int h(\omega_1) d\mu_1$.

Par ailleurs, $\int f_n d\mu_1 \otimes \mu_2$ converge en croissant vers $\int f d\mu_1 \otimes \mu_2$.

On obtient donc le résultat à la limite.

Le résultat pour les fonctions non positives s'obtient en décomposant $f$ en $f_+ - f_-$.

∎

Ce résultat est absolument fondamental. Il a de nombreuses conséquences. Par exemple, on retrouve ainsi le théorème d'intervertion de sommes de séries et d'intégrales.

**Proposition 26.** *Soit $(f_n)$ une suite de fonctions mesurables positives sur $(\Omega, \mathcal{A}, \mu)$. Alors*

$$\sum_n \int f_n d\mu = \int \sum_n f_n d\mu.$$

*le résultat reste vrai pour des fonctions non positives à condition que*

$$\sum_n \int |f_n| \, d\mu < \infty.$$

*Démonstration.* — Il suffit d'appliquer le théorème de Fubini-Tonelli à l'espace produit $\Omega \times \mathbb{N}$, avec le produit de la mesure $\mu$ et de la mesure de décompte sur $\mathbb{N}$.

∎

**Remarque 18.** *Il faut faire attention à ce que la condition d'intégrabilité dans le cas non positif est indispensable.*

*Par exemple, regardons la fonction définie sur $[0,1]^2$ par*

$$F(x,y) = \frac{x^2 - y^2}{(x^2+y^2)^2}.$$

*On a*

$$\int_0^1 [\int_0^1 F(x,y) dy] dx = \frac{\pi}{4},$$

*tandis que, par symétrie*

$$\int_0^1 [\int_0^1 F(x,y) dx] dy = -\frac{\pi}{4}.$$

*Bien sûr, ici*

$$\int_{[0,1]^2} |F(x,y)|\, dxdy = +\infty.$$

Une autre conséquence du théorème de Fubini-Tonelli est le suivant

**Théorème 18.** *Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré, $U$ un ouvert de $\mathbb{C}^2$, et $F(z,\omega)$ une fonction qui soit analytique en $z \in U$, à $\omega$ fixé, et mesurable en $\omega$, à $z$ fixé.*

*Alors, si $\sup_{z\in U} |F(z,\omega)| = h(\omega)$ est intégrable, la fonction*

$$z \mapsto \int_\Omega F(z,\omega)\mu(d\omega)$$

*est analytique dans $U$.*

*Démonstration.* —

La preuve repose sur la formule de Cauchy. Rappelons qu'une fonction $F(z)$ est analytique dans $U$ si elle est développable en série entière au voisinage de chacun de ses points. C'est aussi équivalent au fait de dire que, pour chaque $z_0 \in U$, si la boule fermée de centre $z_0$ et de rayon $r$ est encore inclus dans $U$, alors, pour tout $z$ dans la boule ouverte,

$$F(z) = \frac{1}{2i\pi}\int_0^{2\pi} \frac{F(z_0 + re^{i\theta})}{z - z_0 - re^{i\theta}} d\theta.$$

Supposons alors que $F(z,\omega)$ satisfasse la condition de domination du théorème. On écrit alors, pour une boule de centre $z_0$ et de rayon $r$ incluse dans $U$

$$\int F(z,\omega)\mu(d\omega) = \frac{1}{2i\pi}\left(\int_\Omega \int_0^{2\pi} \frac{F(z_0 + re^{i\theta},\omega)}{z - z_0 - re^{i\theta}} d\theta\right) d\mu(\omega).$$

Il nous reste à voir que nous pouvons appliquer le théorème de Funini-Tonelli à la fonction $G(\omega,\theta) = \frac{F(z_0+re^{i\theta},\omega)}{z-z_0-re^{i\theta}}$ sur $\Omega \times [0,2\pi]$.

Mais si $\rho$ est la distance de $z$ au cercle $z_0 + re^{i\theta}$, on a

$$|G(\omega,\theta)| \le \frac{h(\omega)}{\rho},$$

et l'intégrale de cette dernière fonction vaut $2\pi \int_\Omega h(\omega)d\mu < \infty$. On peut donc appliquer le théorème d'inversion des signes somme, et on retrouve la formule de Cauchy pour $H(z) = \int F(z,\omega)d\mu(\omega)$, ce qu'on voulait démontrer.

Remarquons que dans ce théorème, et contrairement au théorème de dérivation sous le signe somme, c'est la fonction qu'on majore sur $U$ et non sa dérivée (bien qu'une fonction analytique ne soit rien d'autre qu'une fonction complexe dérivable en $z$!)

∎

## 4.3 La mesure de Lebesgue de $\mathbb{R}^n$ et la formule du changement de variables

Dans ce qui suit, nous commençons par étudier quelques propriétés de la mesure de Lebesgue sur $\mathbb{R}^n$.

Rappelons qu'une mesure de Radon est une mesure qui met une masse finie à tous les ensembles bornés.

**Proposition 27.** *La mesure de Lebesgue $\lambda$ est invariante par translation. Si une mesure de Radon $\mu$ sur $\mathbb{R}^n$ est invariante par translation, alors il existe une constante $c$ telle que $\mu = c\lambda$.*

*Démonstration.* — Il n'y a presque rien à changer par rapport à la démonstration que nous avons faite de ce résultat pour $n = 1$. ∎

Nous étudions maintenant l'action des transformations linéaires.

**Théorème 19.** *Soit $L : \mathbb{R}^n \to \mathbb{R}^n$ une application linéaire inversible. Alors, l'image de la mesure de Lebesgue $\lambda$ par $L$ est $\frac{1}{|\det(L)|}\lambda$.*

Nous commençons par un lemme en plusieurs etapes.

**Lemme 2.**

1. *Soit $L : \mathbb{R}^n \to \mathbb{R}^n$ une transformation linéaire inversible. Alors, il existe une constante $c(L)$ telle que l'image $L(\lambda)$ de la mesure de Lebesgue par $L$ soit $c(L)\lambda$.*
2. *Si $L$ est une transformation orthogonale, alors $c(L) = 1$.*
3. *Si $L = L_1L_2$, $c(L) = c(L_1)c(L_2)$.*
4. *Si $L(x_1, \ldots, x_n) = (x_1, \ldots, \lambda_k x_k, \ldots, x_n)$, alors $c(L) = \frac{1}{|\lambda_k|}$.*
5. *Si $L(x_1, \ldots, x_n) = (\lambda_1 x_1, \ldots, \lambda_n x_n)$ alors $c(L) = \frac{1}{|\det(L)|}$.*
6. *Toute application linéaire inversible $L$ se met sous la forme $L = P\Delta Q$, où $P$ et $Q$ sont des transformations orthogonales et $\Delta$ est diagonale.*

*Démonstration.* — Montrons d'abord en quoi le lemme implique le théorème. Nous voulons montrer que $c(L) = |\det(L)|^{-1}$. D'après le point 6, $L = P\Delta Q$, où $P$ et $Q$ sont orthogonales et $\Delta$ diagonale.

D'après le point 3, $c(L) = c(P)c(\Delta)c(Q)$, et d'après le point 2, $c(P) = c(Q) = 1$. D'où $c(L) = c(\Delta)$. Mais d'après le point 5, $c(\Delta) = |\det(\Delta)| = |\det(L)|$, la dernière égalité venant de ce que $|\det(Q)| = |\det(P)| = 1$, puisque $P$ et $Q$ sont orthogonales. ∎

Il nous reste à démontrer les différents points du lemme.

*Démonstration.* —

• Du point 1 : appelons $L^{-1}(B)$ l'image réciproque du borélien $B$ par $A$. Alors, pour un $x \in \mathbb{R}^n$

$$\lambda(L^{-1}(B + x)) = \lambda(L^{-1}(B) + L^{-1}(x)) = \lambda(L^{-1}(B)).$$

La mesure image de $\lambda$ par $L$ est donc invariante par translation. C'est aussi une mesure de Radon car $L$ étant inversible, l'image réciproque par $L$ d'un compact est compacte. Elle est donc proportionnelle à la mesure de Lebesgue.

• Du point 2 : si $B$ est la boule de centre 0 et de rayon 1, alors $L^{-1}(B) = B$. On a donc

$$c(L)\lambda(B) = \lambda(L^{-1}(B)) = \lambda(B).$$

Comme $\lambda(B)$ est non nulle, on obtient le résultat.

• Du point 3 : c'est immédiat d'après la définition de $c(L)$.

• Du point 4 : on se ramène tout d'abord au cas où $\lambda_k > 0$, en changeant s'il le faut $x_k$ en $-x_k$ (ce qui correspond à une transformation orthogonale de $\mathbb{R}^n$), puis on se ramène à la situation similaire en dimension 1, auquel cas ceci a déjà été vu (on peut aussi considérer directement l'image d'un pavé $[a_1, b_1] \times \ldots \times [a_n, b_n]$ sous l'action de $L$).

• Du point 5: on remarque qu'une telle transformation est la composition des transformations précédentes (qui agissent successivement sur chacune des coordonnées), et on obtient ainsi le résultat.

• Du point 6 : nous identifions $L$ et sa matrice dans la base canonique, et désignons par $M^t$ la transposée de la matrice $M$.

Soit $A = L^t L$. C'est une matrice symétrique définie positive. Elle se diagonalise dans une base orthonormée, et donc $L^t L = P^t D P$, où $P$ est une matrice orthogonale, et $D$ est une matrice diagonale à coefficients strictement positifs (rappelons que $L$ est inversible).

Appelons $\Delta$ la racine de $D$, c'est à dire la matrice diagonale à coefficients positifs telle que $\Delta^2 = D$. La matrice $Q = LP\Delta^{-1}$ vérifie $Q^t Q = Id$, donc est orthogonale et par conséquent $L = Q\Delta P^t$ qui est bien une décomposition de la forme cherchée.

∎

**Remarque 19.** *Nous pouvons ainsi calculer le volume des parallépipèdes dans $\mathbb{R}^n$. Si on se donne $n$ vecteurs $(V_1, \ldots, V_n)$ et qu'on considère l'ensemble*

$$P(V_1, \ldots, V_n) = \{\sum_{i=1}^{n} \lambda_i V_i, 0 \leq \lambda_i \leq 1\},$$

*c'est à dire la parallépipède qui s'appuie sur les vecteurs $(V_1, \ldots, V_n)$, alors*

$$\lambda_n(P(V_1, \ldots, V_n)) = |D(V_1, \ldots, V_n)|,$$

*où $D(V_1, \ldots, V_n)$ désigne le déterminant de la matrice de vecteurs colonnes $V_1, \ldots, V_n$.*

*En effet, lorsque les $V_i$ sont linéairement indépendants, nous pouvons considérer l'application linéaire $L$ définie par $L(e_i) = V_i$, où les vecteurs $e_i$ sont les vecteurs de la base canonique de $\mathbb{R}^n$. Cette application linéaire est inversible et sa matrice est (par définition) la matrice des vecteurs $(V_i)$. L'image du cube unité $C = [0,1]^n$ par $L$ est exactement $P(V_1, \ldots, V_n)$, d'où l'on déduit que*

$$\lambda_n(L^{-1}(P(V_1, \ldots, V_n)) = \lambda_n C) = 1.$$

*Par ailleurs, le résultat précédent nous dit aussi que*

$$\lambda_n^{-1}(P(V_1, \ldots, V_n)) = \frac{1}{|D(V_1, \ldots, V_n)|}\lambda_n(P(V_1, \ldots, V_n)),$$

*d'où la formule.*

*Lorsque les vecteurs ne sont pas linéairement indépendants, le déterminant est nul. Mais par ailleurs, $P(V_1, \ldots, V_n)$ est inclus dans un sous-espace vectoriel strict de $\mathbb{R}^n$. Or, les sous-espaces vectoriels stricts de $\mathbb{R}^n$ sont de mesure de Lebesgue nulle. En effet, c'est clairement le cas de $E_0 = \{0, x_2, \ldots, x_n\}$. Tout sous-espace vectoriel strict est inclus dans l'image $E_0$ par une application linéaire inversible, et ces dernières préservent les ensembles de mesure de Lebesgue nulle. La formule reste donc encore vraie lorsque les vecteurs $(V_1, \ldots, V_n)$ ne sont pas linéairement indépendants.*

Nous pouvons maintenant énoncer le théorème du changement de variables

**Théorème 20.** *Soit $U$ et $V$ deux ouverts de $\mathbb{R}^n$ et $\Phi$ un difféomorphisme entre $U$ et $V$. On désigne par $\lambda$ la mesure de Lebesgue de $\mathbb{R}^n$. Soit $D\Phi$ la matrice*

*$(\partial\Phi_i/\partial x_j)$, et $J(\Phi)$ le déterminant de cette matrice (le jacobien de $\Phi$). L'image de la mesure $\mathbf{1}_U(x)\lambda(dx)$ par $\Phi$ est*

$$\mathbf{1}_V(x)\left|J(\Phi)\circ\Phi^{-1}\right|^{-1}(x)\lambda(dx).$$

*Une autre façon d'écrire la formule est, pour toute fonction $f$ positive*

$$\int_U f(\Phi(x))\,|J(\Phi)(x)|\,\lambda(dx) = \int_V f(y)\lambda(dy).$$

*Ou, en d'autres termes, la forme "usuelle" du changement de variables, à savoir*

$$\lambda(d\Phi(x)) = |J(\Phi(x))|\,\lambda(dx).$$

Un difféomorphisme est une application de classe $\mathcal{C}^1$ qui est une bijection entre $U$ et $V$ et telle que l'inverse soit aussi de classe $\mathcal{C}^1$. Dans ce cas, le jacobien ne s'annule pas sur $U$.

Nous n'allons pas donner de démonstration de ce théorème. Remarquons cependant que ce théorème se réduit au précédent lorsque l'application $\Phi$ est linéaire. L'idée de la preuve est alors de découper un pavé $[a_1,b_1]\times\ldots\times[a_n,b_n]$ inclus dans $U$ en union de petits pavés où la fonction $\Phi$ est assez proche de l'application tangente $(\Phi(x)\simeq\Phi(x_0)+D\Phi(x_0)(x-x_0)$ , qui est affine (c'est à dire linéaire à une translation près). Il faut alors contrôler l'erreur que l'on fait en remplaçant $\Phi$ par son approximation sur chaque petit pavé, et rassembler les morceaux. Ce schéma cache en fait pas mal de difficultés techniques.

## 4.4 La mesure de Lebesgue en coordonnées polaires.

Remarquons tout d'abord que le théorème de Fubini nous autorise à calculer par récurrence le volume de la boule unité de $\mathbb{R}^n$. En effet, appelons $V_n$ le volume de la boule unité de $\mathbb{R}^n$ pour la mesure de Lebesgue. Par dilatation, nous voyons que le volume de la boule de rayon $\rho$ est $\rho^n V_n$. On a

$$V_n = 2V_{n-1}\int_0^1 (1-t^2)^{\frac{n-1}{2}}\,dt.$$

Pour le voir, appelons $B_t$ l'intersection de la boule unité avec l'hyperplan $x_1 = t$, pour $t\in(-1,1)$. Par le théorème de Pythagore, $B_t$ est une boule de l'espace euclidien de dimension $n-1$, de rayon $\sqrt{1-t^2}$, et donc de volume $V(t) = \sqrt{1-t^2}^{\,n-1}V_{n-1}$.

Le théorème de Fubini nous dit alors que

$$V_n = \int_{-1}^1 V(t)dt,$$

d'où la formule.

Un résultat plus intéressant est sans doute le suivant. On appelle dans ce qui suit $S_{n-1}$ la sphère unité de $\mathbb{R}^n$, $B_n$ la boule unité de $\mathbb{R}^n$.

**Théorème 21.** *Il existe une unique mesure de probabilité sur $S_{n-1}$ qui soit invariante par rotation. On l'appelle la mesure uniforme de la sphère.*

*Démonstration.* — L'existence est assez facile. Pour un borélien $A \subset S_{n-1}$, on considère le cône
$$\hat{A} = \{x \in B_n \setminus 0 \;\mid\; x/|x| \in A\}.$$
La mesure $\nu(A) = \lambda_n(\hat{A})$ définit une mesure sur $S_{n-1}$, dont la masse totale est le volume de la boule $V_n$. (Cette mesure n'est rien d'autre que l'image de la mesure de Lebesgue de la boule privée de 0 par l'application $x \mapsto \frac{x}{|x|}$.) Cette mesure est invariante par rotation, ce qui n'est que le reflet de l'invariance par rotation de la mesure de Lebesgue. La mesure $\nu_1(A) = \frac{\nu(A)}{V_n}$ répond bien à la question.

L'unicité est plus délicate, car on ne peut pas faire comme dans $\mathbb{R}^n$ et découper de bons boréliens de la sphère en bouts aussi petit que l'on veut et qui se déduisent l'un de l'autre par rotation (sauf dans le cas $n = 2$). Nous allons pour démontrer l'unicité utiliser une technique qui consiste à construire une mesure auxiliaire invariante par rotation sur l'espace $SO(n)$ des matrices orthogonales de déterminant 1. Il s'agira d'une mesure $\mu_n$ sur $SO(n)$ qui vérifiera, pour toute fonction positive $f$ et tout élément $R \in SO(n)$
$$\int_{SO(n)} f(M)d\mu_n(M) = \int_{SO(n)} f(RM)d\mu_n(M).$$
(On dira alors qu'elle est invariante à gauche, car elle est invariante en changeant $M$ en $RM$. Une mesure invariante à droite sera invariante en changeant $M$ en $MR$.)

Pour $n = 2$, $SO(2)$ s'identifie à $S_1$ et on prend comme mesure la mesure $\mu_1$ que nous venons de construire. Ensuite, nous allons procéder par récurrence.

Pour cela, remarquons qu'on peut identifier $SO(n)$ à $S_{n-1} \times S0(n-1)$ de la façon suivante.

Si $M$ est une matrice de $SO(n)$, on considère son premier vecteur colonne $X_1$ qui est un élément de $S_{n-1}$ (l'image par $M$ du premier vecteur de base). On considère alors l'unique rotation de $SO(n)$ qui envoie le premier élément de la base unité $e_1$ sur $X_1$ et qui laisse fixe l'orthogonal du plan engendré par $e_1$ et $X_1$ (Si $X_1 = e_1$, on prend l'identité). Appelons $\theta(R)$ cette rotation. Alors, $\theta(R)^{-1}M$ est un élément de $SO(n)$ qui laisse fixe $e_1$, et donc s'identifie à un élément de $SO(n-1)$.

Nous pouvons ainsi identifier $SO(n)$ à $S_{n-1} \times SO(n-1)$. Dans cette identification, si $R = (X, R_1)$ et $R' = (X', R'_1)$, alors $RR' = (RX', R_1R'_1)$;

Supposons que l'on ait construit sur $SO(n-1)$ une mesure de probabilité $\mu_{n-1}$ invariante par rotation et considérons la mesure de probabilité $\nu_n$ sur $S_{n-1}$ invariante par rotation que nous venons de construire. Nous posons alors
$$\mu_n = \nu_n \otimes \mu_{n-1}.$$
On vérifie immédiatement qu'elle est invariante par rotation sur $SO(n)$. En effet, si $R' = (X', R'_1)$
$$\int f(R'X, R'_1R_1)d\nu_n(X)d\mu_{n-1}(R_1) = \int f(X, R_1)d\nu_n(X)d\mu_{n-1}(R_1),$$
formule qui provient de l'invariance de $\nu_n$ par rotation dans $SO(n)$ de l'invariance par rotation de $\mu_{n-1}$ dans $SO(n-1)$.

Remarquons aussi que si nous prenons l'image de cette mesure $\mu$ par l'application $M \mapsto M^t$, nous obtenons une mesure de probabilité invariante à droite sur $SO(n)$, c'est à dire qui vérifie

$$\int f(MR)d\mu(M) = \int f(M)d\mu(M),$$

pour toute fonction positive $f$ et tout élément $R$ de $SO(n)$.

Soit alors $\nu$ une mesure de probabilité invariante par rotation sur $S_{n-1}$, et soit $\mu$ une mesure invariante par rotation à droite dans $SO(n)$. Soit $f$ une fonction positive sur $S_{n-1}$. Nous allons montrer que

$$\int_{S_{n-1}} f(X)d\nu(X) = \int_{SO(n)} f(Re_1)d\mu(R),$$

où $e_1$ est un vecteur fixe quelconque de $S_{n-1}$ (le résultat ne dépend pas du choix de ce vecteur).

Cette formule montre l'unicité puisque le calcul fait avec deux mesures différentes donnera le même résultat.

En effet, considérons sur le produit $SO(n) \times S_{n-1}$ la mesure $d\rho = d\mu \otimes d\nu$ et intégrons la fonction $g(R, X) = f(RX)$.

Par le théorème de Fubini, nous obtenons

$$\int g(R, X)d\rho = \int_{SO(n)} (\int_{S_{n-1}} f(RX)d\nu(X))d\mu(R) = \int_{S_{n-1}} (\int_{S_{n-1}} (f(RX)d\mu(R))d\nu(X).$$

Dans la première expression, en appelant $A = \int f(X)d\nu(X)$, nous obtenons

$$\int g(R, X)d\rho == \int_{SO(n)} Ad\mu(R) = A,$$

à cause de l'invariance par rotation de la mesure $\nu$.

Dans la seconde, fixons d'abord $X$ et choissons une rotation fixe $R_1$ qui envoie $e_1$ sur $X$. En appelant $B = \int_{SO(n)} f(Re_1)d\mu(R)$, nous avons

$$\int_{SO(n)} f(RX)d\mu(R) = \int f(RR_1e_1)d\mu(R) = B,$$

à cause de l'invariance par rotation à droite de la mesure $\mu$. Et donc

$$\int g(R, X)d\rho = B,$$

ce que nous voulions démontrer.

Remarquons au passage que la même démonstration montre que toute mesure de probabilité invariante à gauche sur $SO(n)$ est égale à toute mesure invariante à droite (appliquer le théorème de Fubini à la fonction $g(R, R') = f(RR')$), et donc que ces mesures sont uniques et égales.

∎

Il n'est pas très facile de décrire explicitement la mesure uniforme sur la sphère. Mais néanmoins, elle permet d'écrire la mesure de Lebesgue en coordonnées polaires.

Soit $x \in \mathbb{R}^n \setminus \{0\}$. En écrivant $|x|$ pour la norme euclidienne de $x$, on peut représenter $x$ par $(|x|, \theta = \frac{x}{|x|}) \in (0,\infty) \times S_{n-1}$. Cette application est une bijection bimesurable entre $R^n \setminus \{0\}$ et $(0,\infty) \times S_{n-1}$. Ce sont les coordonnées polaires dans $\mathbb{R}^n$. On peut dès lors se poser la question de l'image de la mesure de Lebesgue par cette application, c'est à dire de l'écriture de la mesure de Lebesgue en coordonnées polaires.

On a

**Proposition 28.** *L'image de la mesure de Lebesgue par cette application est*

$$V_n \rho^{n-1} \lambda(d\rho)\mu(d\theta),$$

*où $\lambda$ est la mesure de Lebesgue sur $\mathbb{R}_+$, $\mu$ la mesure de probabilités sur la sphère invariante par rotation, et $V_n$ le volume de la boule unité de $\mathbb{R}^n$.*

*Démonstration.* — Il suffit d'identifier cette mesure (appelons la $\rho$) sur les ensembles de la forme $A\times]0,r]$, qui forme une classe stable par intersection qui engendre la tribu.

Mais nous avons déjà vu que la mesure de $A\times]0,1]$ vaut $V_n\mu(A)$ (c'est ainsi que nous avons construit la mesure uniforme sur la sphère), et par dilatation, nous voyons que $\rho(A\times]0,r]) = r^n V_n \mu(A)$, ce qui correspond à la formule proposée.

∎

Pour décrire la mesure de Lebesgue sur la sphère de dimension unité $n-1$ (c'est à dire la sphère unité de $\mathbb{R}^n$, et calculer une intégrale sur cette sphère, il faut la paramétrer. Il y a plusieurs façons de le faire. le plus simple est sans doute la projection cylindrique.

On commence par choisir un axe, (c'est à dire un diamètre) et on enlève les deux pôles. C'est anodin puisqu'un point est de mesure nulle. Appelons $S^*$ cette sphère privée de ses deux pôles. Choissons pour fixer les idées l'axe $x_1 = \ldots = x_{n-1} = 0$. Un point $X \in S^*$ se repère par sa hauteur $x_n \in (-1,1)$ et un point $\theta \in S_{n-2}$, de telle façon que $X = (\sqrt{1-x_n^2}\theta, x_n)$.

Ceci réalise une bijection entre $S^*$ et $(-1,1) \times S_{n-2}$.

Alors, la mesure image de la mesure uniforme sur la sphère de dimension $n-1$ par cette application est

$$c_n(1-x_n^2)^{\frac{n-3}{2}} dx_n \otimes \mu_{n-2}(d\theta),$$

où $c_n$ est une constante de normalisation et $\mu_{n-2}$ est la mesure uniforme sur la sphère de dimension $n-2$.

Ceci prend une forme particulièrement simple lorsque $n = 3$. La sphère de dimension 1 (le cercle) se paramètre par $\theta \in [0, 2\pi)$, et la mesure sur cette sphère n'est rien d'autre que $\frac{d\theta}{2\pi}$. On obtient donc ainsi une image de la sphère privée des pôles sur $(-1,1) \times [0, 2\pi)$ et l'image de la mesure uniforme sur la sphère est ainsi la mesure de Lebesgue sur le rectangle (normalisée pour en faire une probabilité). Ainsi, cette projection cylindrique préserve les surfaces (ceci ne marche que pour la sphère de $\mathbb{R}^3$).

Pour les sphères de dimension plus grande, on peut ainsi travailler par récurrence sur la dimension.

# 5 Espaces $L^p$

Dans ce qui suit, nous supposerons toujours que $(\Omega, \mathcal{A}, \mu)$ est un espace mesuré $\sigma$-fini.

Rappelons que deux fonctions mesurables à valeurs réelles $f$ et $g$ sont égales $\mu$-presque partout si

$$\mu(\{\omega \mid f(\omega) \neq g(\omega)\}) = 0.$$

C'est une relation d'équivalence entre fonctions, et nous identifierons dans la suite une fonction $f$ à sa classe.

Si $f$ et $g$ sont intégrables et dans la même classe, alors

$$\int f d\mu = \int g d\mu,$$

comme nous l'avons déjà vu, et par conséquent $\int f d\mu$ ne dépend pas du choix de $f$ dans sa classe.

**Définition.** 27. *Soit $p \in [0, \infty)$. On dit que $f \in L^p(\Omega, \mathcal{A}, \mu)$ (ou plus simplement $f \in L^p$ s'il n'y a pas d'ambiguïté possible) si $f$ est une fonction mesurable à valeurs réelles et si*

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

*Dans ce cas, on note*

$$\|f\|_p = (\int |f|^p \, d\mu)^{1/p}.$$

Nous avons déjà rencontré l'espace $L^1$ des fonction intégrables.

Introduisons aussi l'espace $L^\infty$.

**Définition.** 28. *On dit que $f$ est dans $L^\infty$ (ou aussi que $f$ est essentiellement bornée) s'il existe une constante $a \geq 0$ telle que*

$$\mu(\{|f| > a\}) = 0.$$

*La borne inférieure de ces constantes est notée $\|f\|_\infty$. Notez aussi que cette borne inférieure est atteinte, c'est à dire que si $f \in L^\infty$, alors $|f| \leq \|f\|_\infty$ $\mu$-presque partout.*

Bien que nous ayons utilisé une notation de norme, nous n'avons pas encore démontré que c'est est une. C'est l'objet du paragraphe suivant.

Remarquons cependant que pour tout $p \in [1, \infty]$, et pour tout scalaire $\lambda$, on a

$$\|\lambda f\|_p = |\lambda| \, \|f\|_p \, .$$

Remarquons aussi que si $\|f\|_p = 0$, alors $|f|^p = 0$ presque partout, et donc $f = 0$ presque partout.

## 5.1 Inégalités de Hölder et de Minkowski.

Les deux inégalités suivantes sont fondamentales

**Théorème 22.** *1. (Inégalité de Hölder) Soient* $(p,q) \in [1,\infty]$*, avec*

$$\frac{1}{p} + \frac{1}{q} = 1.$$

*Alors*

$$\|fg\|_1 \leq \|f\|_p \|g\|_q .$$

*2. (Inégalité de Minkowski) Pour tout* $p \in [1,\infty]$*,*

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p .$$

*Démonstration.* — Commençons par le premier point. Traitons d'abord le cas le plus simple où $p = 1$ et $q = \infty$. En posant $c = \|g\|_\infty$, nous savons que $|fg| \leq c|f|$ $\mu$-presque partout, et donc en intégrant par rapport à $\mu$ nous obtenons

$$\int |fg|\, d\mu \leq c \int |f|\, d\mu,$$

ce qui est l'inégalité à démontrer.

Remarquons que l'inégalité de Minkowski dans les cas $p = 1$ et $p = \infty$ est tout aussi élémentaire.

Le cas général est un peu plus compliqué. Tout d'abord, remarquons qu'on peut supposer que $\|f\|_p$ et $\|g\|_q$ sont non nuls et finis, sinon il n'y a rien à prouver.

Commençons par remarquer que la fonction logarithme étant concave, on a pour tout couple de réels positifs $(x,y)$ et tout $\theta \in (0,1)$

$$\theta \log x + (1-\theta)\log y \leq \log(\theta x + (1-\theta)y).$$

Prenons les exponentielles des deux membres, et appliquons l'inégalité avec $x = a^p$, $y = b^q$ et $\theta = \frac{1}{p}$, auquel cas $1 - \theta = \frac{1}{q}$.

Nous obtenons, pour tout couple de réels positifs $a$ et $b$

$$ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q. \tag{1}$$

Cette inégalité reste bien évidement vraie si l'un des réels $a$ ou $b$ est nul.

Posons alors $f_1(x) = \frac{|f(x)|}{\|f\|_p}$ et $g_1(x) = \frac{|g(x)|}{\|g\|_q}$.

Appliquons notre inégalité avec $a = f_1(x)$ et $b = g_1(x)$, puis intégrons par rapport à $\mu$. On a

$$\int f_1^p d\mu = \int g_1^q d\mu = 1,$$

et il reste

$$\int \frac{|fg|}{\|f\|_p \|g\|_q} d\mu \leq \frac{1}{p} + \frac{1}{q} = 1,$$

et c'est bien l'inégalité que nous cherchions à démontrer.

**Remarque 20.** *Dans le cas d'égalité, la fonction* log *étant strictement concave, il n'y a égalité dans l'inégalité 1 que si* $a^p = b^q$*. Par suite, il n'y a égalité dans l'inégalité de Hölder que si* $|f|^p$ *est égale à* $|g|^q$ $\mu$*-presque partout.*

Passons maintenant à l'inégalité de Minkowski. Une fois de plus, les cas $p = 1$ et $p = \infty$ sont élémentaires. Nous ne traitons que les autres cas.

Tout d'abord, remarquons que si $f$ et $g$ sont dans $L^p$, il en va de même de $f + g$, par l'inégalité

$$|f + g|^p \leq 2^{p-1}(|f|^p + |g|^p),$$

dont nous laissons la démonstration au lecteur à titre d'exercice. (Utiliser la convexité de la fonction $x^p$.)

Ensuite, pour $p \in (1, \infty)$, nous majorons $|f + g|^p$ par $(|f|+|g|)^p$, ce qui nous permet de nous ramener au cas où $f$ et $g$ sont positives.

Enfin, nous écrivons, pour des fonctions $f$ et $g$ positives,

$$(f + g)^p = (f + g)^{p-1}f + (f + g)^{p-1}g.$$

Nous intégrons les deux membres de cette inégalité par rapport à la mesure $\mu$. Nous appliquons l'inégalité de Hölder aux deux morceaux du membre de gauche. Or, su $q$ est l'exposant conjugué de $p$, alors $q(p - 1) = p$, et par suite nous obtenons

$$\|f + g\|^p \leq \|f + g\|_p^{p/q} (\|f\|_p + \|g\|_p).$$

Il ne nous reste plus qu'à diviser les deux membres par $\|f + g\|_p^{p/q}$ pour obtenir le résultat.

∎

## 5.2 L'espace de Banach $L^p(\Omega\mathcal{A}, \mu)$

Rappelons que nous identifions deux fonctions égales presque partout.

Commençons par quelques rappels de topologie. Un espace métrique est un ensemble $E$ muni d'une distance $d : E \times E \to R_+$ qui vérifie

1. $d(x, y) = 0 \iff x = y$;
2. $d(x, y) \leq d(x, z) + d(z, y)$.

Une suite $x_n$ converge vers $x$ dans cet ensemble si $d(x_n, x)$ converge vers 0 dans $\mathbb{R}$.

Une suite de Cauchy est une suite $x_n$ telle que, si

$$a_n = \sup_{p,q \geq n} d(x_p, x_q) \to 0 \ (n \to \infty).$$

Il revient au même de demander la même propriété pour la suite

$$b_n = \sup_{p \geq n} d(x_p, x_n),$$

puisque $b_n \leq a_n \leq 2b_n$.

Un espace métrique est complet si toute suite de Cauchy y est convergente. Notons que, si une suite de Cauchy admet une sous-suite qui converge vers $x$, alors elle converge elle même vers $x$.

Lorsque $E$ est un espace vectoriel, une norme sur $E$ est une application de $E \to \mathbb{R}_+$, qu'on note $\|x\|$, telle que

1. $\|\lambda x\| = |\lambda| \, \|x\|$;

2. $\|x+y\| \leq \|x\| + \|y\|$;
3. $\|x\| = 0 \iff x = 0$.

Si $E$ est un espace vectoriel munit d'une norme, alors on en fait un espace métrique en posant $d(x,y) = \|x-y\|$.

Si cet espace est complet, on dit que c'est un **espace de Banach**.

Nous pouvons alors énoncer le

**Théorème 23** (Riesz-Fisher)**.** *L'espace $L^p$, muni de la norme $\|f\|_p$ est un espace de Banach*

*Démonstration.* — Nous avons déjà vu qu'une fonction telle que $\|f\|_p = 0$ est nulle presque partout. Par ailleurs, l'inégalité de Minkowski montre que $L^p$ est un espace vectoriel et que $\|\cdot\|_p$ est une norme.

Montrons que l'espace est complet.

Soit alors $f_n$ une suite de Cauchy, et $b_n = \sup_{q\geq n} \|f_q - f_n\|_p$.

Extrayons tout d'abord une sous-suite $n_k$ telle que $b_{n_k} \leq 2^{-k}$. Quitte à remplacer $f_n$ par $f_{n_k}$, on peut supposer que $b_n \leq 2^{-n}$, puisqu'il suffit de montrer qu'il existe une sous-suite convergente.

Montrons qu'alors, si $\mu(A) < \infty$, $f_n \mathbf{1}_A$ converge presque sûrement.

Utilisons pour cela l'inégalité de Hölder, en appelant $q$ l'exposant conjugué de $p$, nous pouvons voir que

$$c_n := \int \mathbf{1}_A \left| f_n - f_{n+1} \right| d\mu \leq \mu(A)^{1/q} b_n,$$

et donc que la série $\sum_n c_n$ est convergente. par conséquent, la série

$$\sum_n \mathbf{1}_A \left| f_{n+1} - f_n \right|$$

converge presque partout, ce qui montre que la suite $f_n$ converge presque partout sur $A$, vers une limite $f$. En choisissant une suite croissante $A_k$ telle que $\mu(A_k) < \infty$ et telle que $\cup_k A_k = \Omega$, nous voyons que $f_n$ converge presque partout vers $f$ sur $\Omega$.

Il nous reste à montrer que $f \in L^p$ et que $f_n$ converge vers $f$ dans $L^p$.

Pour cela, commençons par remarquer que la suite $\|f_n\|_p$ est bornée (majorée par $b_1 + \|f_1\|_p$ si l'on veut), et donc que, en appelant par $A$ un majorant de cette suite, on a, grâce au lemme de Fatou,

$$\int |f|^p \, d\mu \leq \liminf_n \int |f_n|^p \, d\mu \leq A^p.$$

Par ailleurs, nous écrivons toujours grâce au lemme de Fatou

$$\int |f - f_n|^p \, d\mu \leq \liminf_q \int |f_n - f_q|^p \, d\mu \leq b_n^p \leq 2^{-np}.$$

Ceci suffit à conclure.

∎

**Remarque 21.** *La démonstration du théorème précédent nous done un peu plus. Si une suite $(f_n)$ converge dans $L^p$, alors il existe une sous-suite qui converge presque partout (vers la même limite).*

On voit aisément qu'une fonction bornée à support dans un ensemble de mesure finie (c'est à dire nulle en dehors de cet ensemble) est dans tous les espaces $L^p$.

On peut voir même un peu mieux. Si $f$ est dans $L^p$, et si $A_n$ est une suite croissante d'ensembles mesurables tels que $\mu(A_n) < \infty$ et telle que $\cup_n A_n = \Omega$, alors la suite $f_n = f\mathbf{1}_{\{|f|\leq n\}\cap A_n}$ converge vers $f$ dans $L^p$. Ainsi on voit que $L^1 \cap L^\infty$ est dense dans $L^p$.

Pour une fonction $f$ donnée, l'ensemble des $p$ pour lesquels $p \in \mathbb{L}^p$ est un intervalle. Si $f \in L^{p_1} \cap L^{p_2}$, alors $f \in L^p$, $\forall p \in [p_1, p_2]$. En effet, on a

**Proposition 29.** *Si* $1 \leq p_1 \leq p_2 \leq \infty$ *et* $\frac{1}{p} = \frac{\theta}{p_1} + \frac{1-\theta}{p_2}$, *avec* $\theta \in [0,1]$, *alors*

$$\|f\|_p \leq \|f\|_{p_1}^{\theta} \|f\|_{p_2}^{1-\theta}.$$

*Démonstration.* — Quitte à approcher $f$ par une suite de fonctions qui sont dans tous les $L^p$ et à passer à la limite, on peut se ramener au cas où $f \in L^p$.

Dans ce cas, on écrit

$$|f|^p = |f|^{\theta p} |f|^{(1-\theta)p},$$

puis on applique l'inégalité de Hölder avec les exposants $q = \frac{p_1}{p\theta}$ et $r = \frac{p_2}{p(1-\theta)}$, dont on vérifie immédiatement que $\frac{1}{q} + \frac{1}{r} = 1$. On vérifie qu'on obtient bien l'inégalité annoncée. ∎

Lorsque $\mu$ est une mesure finie, alors on a un peu mieux

**Proposition 30.** *Supposons que* $\mu$ *soit finie. Alors, si* $p < q$, $L^p \subset L^q$. *Si* $\mu$ *est une probabilité, alors, si* $p \leq q$

$$\|f\|_p \leq \|f\|_q.$$

*Démonstration.* — Quitte à diviser $\mu$ par une constante, on peut supposer que $\mu$ est une probabilité. L'inclusion découle alors de l'inégalité des normes.

Pour démontrer celle ci, on applique l'inégalité de la proposition 29 avec les fonctions $|f|^q$ et 1, et les exposants conjugés $\frac{p}{q}$ et $\frac{p}{p-q}$.

∎

Dans $\mathbb{N}$, avec la mesure de décompte, les inégalités vont dans l'autre sens

**Proposition 31.** *Dans* $\mathbb{N}$ *avec la mesure* $\sum_n \delta_n$, *on a pour* $q > p$

$$\|f\|_q \leq \|f\|_p.$$

*Démonstration.* — Nous commençons par remarquer que, pour tout $x \geq 0$ et tout $p > 1$,

$$(1 + x^p)^{1/p} \leq 1 + x.$$

Pour le voir, on remarque que la fonction $g(x) = (1+x^p)^{1/p} - 1 - x$ est négative au voisinage de $x = 0$, négative à l'infini, et est convexe. Elle est donc toujours négative.

On en déduit, par homogénéïté, que pour tout $a$ et $b$ positifs, on a

$$(a^p + b^p)^{1/p} \leq a + b.$$

Puis par récurrence que pour tout $n$-uplet $(a_1, \ldots, a_n)$de réels positifs, on a

$$(a_1^p + \ldots + a_n^p)^{\frac{1}{p}} \leq a_1 + \ldots a_n.$$

En passant à la limite, on voit que pour une suite $a = (a_n)$, et pour $p > 1$,

$$\|a\|_p \leq \|a\|_1 .$$

On obtient le cas général en appliquant ce qui précède à la suite $|a|_n^p$ en changeant $p$ en $q/p$.

∎

## 5.3 Sous espaces denses de $L^p(\mathbb{R})$ ou $L^p(\mathbb{R}^n)$.

Pour établir une propriété pour toutes les fonctions de $L^p$, il est souvent commode de se ramener à des fonctions plus simples. Lorsqu'on est sur $\mathbb{R}$ ou sur $\mathbb{R}^n$, avec la tribu des boréliens et une mesure de Radon, on peut ainsi se ramener à des fonctions continues à support compact. Nous n'allons le démontrer que sur $\mathbb{R}$, bien que le même raisonnement s'applique sans difficulté à $\mathbb{R}^n$.

Rappelons qu'une fonction est à support compact s'il existe $M > 0$ telle que $f$ est nulle en dehors de $[-M, M]$.

Nous avons alors le

**Théorème 24.** *Soit $\mu$ une mesure de Radon sur $\mathbb{R}$. L'espace des fonctions continues à support compact est dense dans $L^p(\mu)$, pour tout $p \in [1, \infty)$.*

Attention à ce que la valeur $p = \infty$ est exclue ici. En effet, si une suite de fonctions continues converge vers une limite en norme $\infty$, alors la limite est continue.

*Démonstration.* — Le problème est d'approcher dans $L^p$ une fonction $f$ de $L^p$ par une fonction continue. Tout d'abord, on peut décomposer $f$ en $f_+ - f_-$, et on se ramène ainsi au cas où $f \geq 0$. Ensuite, on peut approcher $f$ par $f\mathbf{1}_{[-M,M]}$ et on se ramène au cas où $f$ est à support compact. On peut de mme remplacer $f$ par $f\mathbf{1}_{\{f \leq n\}}$ et on se ramène ainsi au cas ou $f$ est en plus bornée.

On considère alors une suite étagée qui converge en croissant vers $f$, et on voit imédiatement en utilisant le théorème de convergence dominée que cette suite converge vers $f$ dans $L^p$.

On se ramène ainsi au cas où $f$ est étagée à support compact, puis par linéarité au cas où $f$ est l'indicatrice d'un borélien borné.

Soit $B$ un tel borélien. remarquons que

$$|\mathbf{1}_B - \mathbf{1}_A|_p = \mu(A\Delta B)^{1/p}.$$

Or, nous savons que nous pouvons approcher en mesure tout borélien par une réunion finie d'intervalles (ici des intervalles bornés), et on se ramène ainsi à l'indicatrice d'un intervalle. On laisse alors finir le travail au lecteur.

∎

Une fois que l'on sait appprocher une fonction de $L^p$ par une fonction continue, on peut faire mieux.

**Corollaire 2.** *Pour $p \in [1, \infty)$ et pour toute mesure de Radon sur $\mathbb{R}$ ou $\mathbb{R}^n$, les fonctions de classe $\mathcal{C}^\infty$ (c'est à dire indéfiniment dérivable) et à support compact sont denses dans $L^p$.*

*Démonstration.* — Remarquons qu'il existe au moins une fonction $\mathcal{C}^\infty$ à support compact, positive et non identiquement nulle.

En effet, la fonction $h(x)$ qui vaut 0 sur $\mathbb{R}_-$ et $e^{-1/x}$ sur $\mathbb{R}_+$ est $\mathcal{C}^\infty$. (Nous laissons la vérification de ce point au lecteur). Alors, la fonction $\phi(x) = h(x)h(1-x)$ est $\mathcal{C}^\infty$, à support dans $[0,1]$ et strictement positive sur $(0,1)$.

On peut donc fabriquer une fonction positive, $\mathcal{C}^\infty$ à support compact et d'intégrale 1. Appelons encore $\phi$ une telle fonction.

Si $f$ est une fonction continue à support compact, alors la fonction

$$f_t(x) = \int \phi(y)f(x+ty)\lambda(dy) = \int \phi(\frac{z-x}{t})f(z)\lambda(dz)$$

est une fonction $\mathcal{C}^\infty$ à support compact, majorée par $\|f\|_\infty \|\phi\|_\infty$, et qui converge vers $f$ lorsque $t$ converge vers 0.

Donc $f_t \to f$ dans $L^p$ lorsque $t \to 0$, et c'est ce que nous voulions démontrer. ∎

## 5.4 L'espace de Hilbert $L^2$

L'espace $L^2$ joue un rôle particulier, car la norme y est définie à partir d'un produit scalaire. En effet, si l'on pose

$$\langle f, g\rangle = \int fg d\mu,$$

alors, c'est une application bilinéaire symétrique et

$$\|f\|^2 = \langle f, f\rangle.$$

Les espaces de Banach munis d'une telle norme sont appelés espaces de Hilbert. Ils jouent le rôle des espaces euclidiens en dimension infinie. On a toujours l'inégalité de Schwartz

$$|\langle f, g\rangle| \leq \|f\| \|g\|,$$

qui est générale dans les espaces de Hilbert mais qui n'est ici que le cas particulier de l'inégalité de Hölder avec $p = q = 2$.

Dans la plupart des cas, cet espace est séparable, c'est à dire qu'il existe un ensemble dénombrable dense. C'est le cas pour $L^2(\mu)$ dès que la tribu est engendrée par une famille dénombrable de parties, comme dans $\mathbb{R}$ ou $\mathbb{R}^n$, et que la mesure est $\sigma$-finie.

Pour le voir, commençons par traiter le cas où $\mu$ est finie. Si $\mathcal{A}_0$ est l'algèbre de Boole engendrée par cette famille dénombrable de parties, elle est aussi dénombrable et les combinaisons linéaires à coefficients rationnels d'indicatrices d'éléments de $\mathcal{A}_0$ forment une famille dense dans $L^2(\mu)$ par le même raisonnement que nous avons utilisé dans le chapitre précédent pour les fonctions continues.

Lorsque la mesure $\mu$ est seulement $\sigma$-finie, on commence par approcher une fonction $f \in L^2$ par $f\mathbf{1}_A$, où $A$ est de mesure finie, et on fait ensuite comme dans le cas de mesure finie.

**Remarque 22.** *De façon générale, lorsque $\mu$ est $\sigma$-finie et que $A = \sigma(\mathcal{A}_0)$, où $\mathcal{A}_0$ est une algèbre de Boole, l'espace des combinaisons linéaires d'indicatrices d'éléments de $\mathcal{A}_0$ est dense dans tous les espaces $L^p$ $(1 \leq p < \infty)$. Cela découle du fait que les fonctions étagées sont denses dans $L^p$ et du fait qu'on puisse approcher en mesure des éléments de $\mathcal{A}$ par des éléments de $\mathcal{A}_0$ (voir la remarque qui suit le théorème de Carathéodory (théorème 2).*

L'avantage des espaces de Hilbert séparables est qu'ils admettent des bases (dites bases hilbertiennes). Une base est une suite $(e_n)$ d'éléments de l'espace tels que, pour tout $n, m$

$$\langle e_n, e_m \rangle = \delta_{nm},$$

(0 si $n \neq m$ et 1 si $n = m$) et telle que les combinaisons linéaires des $(e_n)$ forment un sous-espace vectoriel dense.

A partir de toute famille d'éléments dont les combinaisons linéaires forment un sous-espace dense, on peut extraire une base hilbertienne par combinaison linéaire.

Lorsqu'on a une base, on peut décomposer une élément de l'espace dans cette base, en posant

$$x = \sum_n x_n e_n,$$

au sens où la suite $\sum_{k=1}^n x_k e_k$ converge dans l'espace, ce qui, dans ce cas, revient à dire que la série $\sum_n x_n^2$ est convergente (car $\|\sum_n^p x_q e_q\|^2 = \sum_n^p x_q^2$).

Pour un $x$ donné, on peut calculer les coefficients $x_n$ par la formule

$$x_k = \langle x, e_k \rangle,$$

exactement comme dans un espace euclidien. Le produit scalaire de $x = \sum_n x_n e_n$ et $y = \sum_n y_n e_n$ est $\sum_n x_n y_n$.

Par exemple, sur $[0, 1]$ avec la mesure de Lebesgue, une base souvent utilisée est celle des polynômes de Legendre, qui sont définis comme des polynômes $P_n$ de degré $n$, à coefficients dominants positifs, et tels que

$$\int_0^1 P_n(x) P_m(x) dx = \delta_{mn}.$$

On vérifie que cela définit de facon unique la suite $P_n$, qu'elle forme une base de $L^2([0, 1])$. On peut ainsi décomposer toute fonction $f \in L^2$ en

$$f(x) = \sum_n f_n P_n(x),$$

série convergente dans $L^2$, avec

$$f_n = \int_0^1 f(x) P_n(x) dx.$$

Toute mesure bornée sur $[0, 1]$ admet ainsi une base de polynômes orthogonaux qui permettent de décomposer les bonnes fonctions dans ces bases. C'est aussi le cas de certaines mesures bornées qui ne sont pas à support compact mais pour lesquelles la famille des polynômes est dense dans $L^2(\mu)$.

Les familles de polynômes orthogonaux les plus utilisées dans la pratique sont les polynômes de Legendre, de Chebychev, de Gegenbauer (toutes ces familles rentrent dans la famille plus large des polynômes de Jacobi), ainsi que les polynômes de Hermite, de Laguerre, de Meixner, de Hahn, etc.

L'avantage d'un espace de Hilbert est qu'il est sont propre dual.

**Proposition 32.** *Sur $L^2(\mu)$, toute forme linéaire continue $L$ se représente de façon unique sous la forme*

$$L(f) = \langle g, f \rangle,$$

*où $g$ est un élément de $L^2(\mu)$.*

*Démonstration.* — C'est un résultat général d'analyse. Nous ne faisons qu'esquisser la preuve. Dans une base donnée $(e_n)$, si $L$ est donnée, la seule fonction $g$ possible est $\sum_n g_n e_n$, avec $g_n = L(e_n)$. Il suffit de montrer que

$$\sum_n g_n^2 < \infty$$

.

Mais on voit que

$$\sum_{p=1}^{n} g_p^2 \leq \sup_{f \in B_n} L(f)^2 \leq C,$$

où

$$B_n = \{f = \sum_k \lambda_k e_k, \ \ |f|^2 \leq 1\}.$$

∎

## 5.5 Dualité dans les espaces $L^p$

Rappelons que dans un espace de Banach $B$, une application linéaire $L : B \to \mathbb{R}$ est continue si et seulement si

$$\sup_{\|x\| \leq 1} |L(x)| = C < \infty,$$

et que la plus petite constante $C$ qui vérifie cette inégalité s'appelle la norme de $L$.

L'ensemble des formes linéaires, muni de cette norme est un nouvel espace de Banach, qu'on appelle l'espace dual.

Nous avons vu au paragraphe précédent que l'espace $L^2$ était son propre dual (c'est une situation générale pour les espaces de Hilbert). Dans le cas général, nous avons

**Théorème 25** (Riesz)**.** *Soit $(, \mathcal{A}, \mu)$ un espace mesuré $\sigma$-fini. Si $p \in [1, \infty)$ et $\frac{1}{p} + \frac{1}{q} = 1$, alors le dual de $L^p$ est $L^q$, à travers l'identification, pour $g \in L^q$,*

$$\phi(g)(f) = \int fg d\mu.$$

Attention à ce que ce théorème est faux pour $p = \infty$. Le dual de $L^\infty$ est un monstre.

Il faut faire attention à ce que dit précisément ce théorème. Il identifie tout élément du dual à un élément de la forme $\phi(g)$, pour $g \in L^q$; mais il dit aussi que la norme calculée dans le dual est la norme dans $L^q$, c'est à dire que, pour $g \in L^q$,

$$\|g\|_q = \sup_{\|f\|_p \leq 1} \left| \int fg d\mu \right|.$$

*Démonstration.* — Notons tout d'abord que l'inégalité de Hölder nous dit que si $g \in L^q$, alors $\phi(g)$ est une forme linéaire continue sur $L^p$ de norme majorée par $\|g\|_q$. Par ailleurs, pour $p > 1$, pour une fonction $g \in L^q$ de norme 1, si on note $h = \epsilon(g)\,|g|^{q-1}$, où $\epsilon(g)$ est le signe de $g$, on a

$$\int |h|^p \, d\mu = \int |g|^p \, d\mu = 1,$$

et

$$\|g\|_q = \int g h d\mu = \phi(g)(h).$$

On voit donc que

$$\|\phi(g)\| \geq \|g\|_q \,,$$

d'où l'égalité.

Pour $p = 1$ et $g \in L^\infty$, le même raisonnement s'applique avec $h = \epsilon(g)$.

Il reste à montrer que toute forme linéaire continue sur $L^p$ est de cette forme. La démonstration de ce point en toute généralité repose sur le théorème de Jordan-Hahn (théorème 16). Nous n'allons traiter que le cas où $\mu$ est une mesure finie et où $p \in [1, 2]$, qui ne requiet que des outils élémentaires. La démonstration du cas général ($p \geq 2$) est renvoyée à l'appendice.

Nous savons déjà que le résultat est vrai lorsque $p = 2$, puisque $L^2$ est un espace de Hilbert.

On considère alors une forme linéaire continue $L$. Sa restriction à $L^2$ est une forme linéaire continue sur $L^2$, et il existe donc une fonction $h \in L^2$ telle que, pour toute fonction $f \in L^2$,

$$L(f) = \int h f d\mu.$$

Nous allons montrer que cette fonction $h$ est en fait dans $L^q$ et que la formule précédente s'étend aux fonctions $f$ de $L^p$.

Pour cela, on considère la fonction

$$h_n = |h|^{q-1} \epsilon(h) \mathbf{1}_{|h| \leq n},$$

où $\epsilon(h)$ est le signe de $h$. Cette fonction est bornée et donc dans $L^2$. Or

$$L(h_n) = \int |h|^q \, \mathbf{1}_{|h| \leq n} d\mu,$$

et par ailleurs

$$L(h_n) \leq C \, \|h_n\|_p \,,$$

où $C$ est la la norme de $L$.

On obtient ainsi

$$\int |h|^q \, \mathbf{1}_{|h| \leq n} d\mu \leq C (\int |h|^q \, \mathbf{1}_{|h| \leq n} d\mu)^{1/p},$$

ce qui se récrit

$$\int |h|^q \, \mathbf{1}_{|h| \leq n} d\mu \leq C^q.$$

Il reste à faire converger $n$ vers l'infini pour obtenir

$$\int |h|^q \, d\mu \leq C^q.$$

Une fois que l'on sait que $h \in L^q$, pour $f \in L^p$, on considère la suite$f_n = f\mathbf{1}_{|f|\leq n}$ qui est dans $L^2$, et on écrit

$$L(f_n) = \int f_n h d\mu.$$

Mais $f_n$ converge vers $f$ dans $L^p$, et les deux membres étant des formes linéaires continues sur $L^p$, on peut passer à la limite et obtenir

$$L(f) = \int f h d\mu.$$

Il n'est pas très difficile d'étendre cette démonstration au cas où $\mu$ est une mesure $\sigma$-finie. Par contre, la démonstration du cas $2 < p < \infty$ est beaucoup plus délicate, et nous ne la ferons pas.

∎

## 5.6 Les espaces $L^p$ complexes

Jusqu'ici, nous n'avons intégré que des fonctions à valeurs réelles. C'était essentiellement pour ne pas compliquer les choses. mais il n'y aurait eu aucune différence à intégrer des fonctions à valeurs dans $\mathbb{C}$ (nous l'avons d'ailleurs déjà fait lorsque nous avons parlé de l'intégrale de fonctions analytiques).

Une fonction mesurable à valeurs dans $\mathbb{C}$ s'écrit

$$f(\omega) = f_1(\omega) + i f_2(\omega),$$

où $f_1$ et $f_2$ sont des fonctions mesurables réelles.

On pose $|f| = \sqrt{f_1^2 + f_2^2} \leq |f_1| + |f_2|$, et on voit que $f_1$ et $f_2$ sont intégrables si et seulement si $|f|$ l'est. On dit alors que $f$ est intégrable ou que $f \in L^1$, et on pose alors

$$\int f d\mu = \int f_1 d\mu + i \int f_2 d\mu.$$

La remarque suivante est fondamentale

**Proposition 33.** *Si $f \in L^1$, alors*

$$\left| \int f d\mu \right| \leq \int |f| \, d\mu.$$

*Démonstration.* —

On choisit un nombre complexe de module 1 (soit $a + ib$) tel que

$$\left| \int f d\mu \right| = a \int f_1 d\mu + b \int f_2 d\mu.$$

(par exemple $a + ib = \exp i\theta$, où $\int f d\mu = \left|\int f d\mu\right| e^{i\theta}$.)

Puis on écrit

$$\left| a \int f_1 d\mu + b \int f_2 d\mu \right| = \left| \int (af_1 + bf_2) d\mu \right| \leq \int |af_1 + bf_2| \, d\mu \leq \int |f| \, d\mu.$$

∎

Une fois cette proposition établie, on définit de même les espaces $L^p$ complexes comme l'espace vectoriel des fonctions mesurables complexes telles que $|f|^p$ soit intégrable.

Toutes les propriétés établies précédement (inégalités de Hölder, de Minkowski, complétion, dualité) restent valables sans modifications.

Faisons une mention spéciale pour l'espace $L^2$ complexe, qui est un espace de Hilbert complexe, défini à l'aide du produit scalaire (Hermitien)

$$\langle f, g \rangle = \int \bar{f} g d\mu.$$

# 6 La convolution et la transformation de Fourier de mesures bornées

## 6.1 La convolution de mesures bornées sur $\mathbb{R}^n$

**Définition.** 29. *Soient $\mu$ et $\nu$ deux mesures bornées sur $\mathbb{R}^n$. Alors la mesure $\mu * \nu$ est la mesure image de la mesure $\mu \otimes \nu$ par l'application $(x, y) \mapsto x + y$.*

Par définition, on aura donc pour toute fonction mesurable bornée

$$\int_{\mathbb{R}^n} g(z) d\mu * \nu(z) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} g(x+y) d\mu(x) d\nu(y).$$

Notons par exemple que $\delta_x * \delta_y = \delta_{x+y}$.
La proposition suivante est immédiate

**Proposition 34.** *1. $\mu * \nu = \nu * \mu$;*

*2. $(\mu * \nu) * \rho = \mu * (\nu * \rho)$;*

*3. La convolution est bilinéaire.*

La convolution d'une mesure bornée et d'une mesure absolument continue par rapport à la mesure de Lebesgue est absolument continue (en quelque sorte, la convolution est régularisante).
En effet

**Proposition 35.** *Soit $f \in L^1(\mathbb{R}^n, \lambda_n)$ et $\nu$ une mesure bornée.*
*Alors*

1. *Pour presque tout $x \in \mathbb{R}^n$,*
$$\int |f(x-y)| \, d|\nu|(y) < \infty;$$

2. *Si $h(x) = \int f(x-y) d\nu(y)$, et si $d\mu = f d\lambda_n$, alors $\mu * \nu$ est une mesure de densité $h(x)$ par rapport à la mesure de Lebesgue $\lambda_n$.*

*Démonstration.* — Commençons par le premier point, pour lequel on peut supposer que $\nu$ est une mesure positive bornée. Alors, par le théorème de Fubini-Tonelli,

$$\int_x \int_y |f(x-y)| \, d\nu(y) \lambda_n(dx) = \int_y \int_x |f(x-y)| \, \lambda_n(dx) \nu(dy).$$

Mais pour tout $y$,

$$\int |f(x-y)| \, \lambda_n(dx) = \int |f| \, d\lambda_n = C,$$

par l'invariance par translation de la mesure de Lebesgue, et donc

$$\int_x \int_y |f(x-y)| \, d\nu(y) \lambda_n(dx) = C \nu(\mathbb{R}^n).$$

La fonction
$$g(x) = \int_y |f(x-y)|\, \nu(dy)$$
est donc dans $L^1(\mathbb{R}^n, \lambda_n)$ et donc finie presque partout.

Pour le second point, on se ramène par bilinéarité aux cas où $\mu$ et $\nu$ sont positives, avec $d\mu = fd\lambda_n$, et on écrit, pour toute fonction positive bornée $k$, avec $\rho = \mu * \nu$
$$\int k(z)d\rho = \int \int h(x+y)f(x)\lambda_n(dx)\nu(dy)$$
$$= \int \int k(z)f(z-y)\lambda_n(dy)\nu(dx) = \int k(z)h(z)\lambda_n(dz).$$

∎

## 6.2 La transformée de Fourier de mesures complexes bornées

Rappelons qu'une mesure complexe s'écrit $\mu_1 + i\mu_2$, où $\mu_1$ et $\mu_2$ sont des mesures réelles bornées, elle même différence de deux mesures positives bornées. Pour une telle mesure définie sur $\mathbb{R}^n$, on définit la transformée de Fourier par

**Définition.** 30. *Pour $t \in \mathbb{R}^n$,*
$$\hat{\mu}(t) = \int e^{it\cdot x}\mu(dx),$$
*où $t \cdot x$ désigne le produit scalaire de $t \in \mathbb{R}^n$ et de $x \in \mathbb{R}^n$.*

Cette transformée de Fourier caractérise la mesure, comme on le verra plus bas. Nous commençons par les propriétés élémentaires suivantes

**Proposition 36.** *1. Pour tout $t \in \mathbb{R}^n$, l'appplication $\mu \mapsto \hat{\mu}(t)$ est une forme linéaire (complexe) sur l'espace vectoriel des mesures (complexes) bornées.*

*2. $\hat{\mu}(0) = \mu(\mathbb{R}^n)$.*

*3. $\widehat{\mu * \nu} = \hat{\mu} * \hat{\nu}$.*

*4. Si $A$ est une application linéaire $\mathbb{R}^n \mapsto \mathbb{R}^n$ et si $A^t$ désigne sa transposée, et si $A(\mu)$ désigne l'image de $\mu$ par $A$, alors*
$$\widehat{A(\mu)}(t) = \hat{\mu}(A^t(t)).$$

*5. Si $a \in \mathbb{R}^n$ et $T_a(\mu)$ désigne l'image de $\mu$ par la translation $x \mapsto x + a$, alors*
$$\widehat{T_a(\mu)}(t) = e^{ia\cdot t}\hat{\mu}(t).$$

*6. Si $\mu$ est réelle, alors $\hat{\mu}(-t) = \bar{\hat{\mu}}(t)$.*

*7. $t \mapsto \hat{\mu}(t)$ est continue, et même uniformément continue.*

*Démonstration.* — Les deux premiers points sont évidents.

Pour le troisième, on écrit

$$\int_{R^n} g(x) d(\mu * \nu)(x) = \int_{\mathbb{R}^n \times \mathbb{R}^n} g(x+y) d\mu(x) d\nu(y).$$

Puis on applique ceci à

$$e^{it\cdot(x+y)} = e^{it\cdot x} e^{it\cdot y}$$

et le théorème de Fubini.

Pour le quatrième point, on applique le théorème du transfert, avec $\nu = A(\mu)$ :

$$\int g(x) d\nu(x) = \int g(Ax) d\mu(x),$$

qu'on applique à $e^{it\cdot x}$, et on observe que

$$e^{it\cdot Ax} = e^{iA^t t\cdot x}.$$

On traite de même la translation.

Le point suivant est évident. Pour le dernier, on se ramène au cas des mesures bornées positives. Dans ce cas, une application immédiate du théorème de convergence dominée montre que la fonction $\hat{\mu}(t)$ est continue. Pour montrer la continuité uniforme, on écrit, toujours pour une mesure posiitive bornée,

$$|\hat{\mu}(t) - \hat{\mu}(s)| \leq \int \left| e^{it\cdot x} - e^{is\cdot x} \right| d\mu = \int \left| e^{i(t-s)\cdot x} - 1 \right| d\mu(x).$$

Maintenant,

$$\int \left| e^{iu\cdot x} - 1 \right| d\mu(x)$$

converge vers 0 lorsque $u \to 0$ par le théorème de convergence dominée.

∎

La transformée de Fourier d'une mesure bornée n'est pas n'importe quelle fonction.

**Définition.** 31. *On dit qu'une fonction* $\phi : \mathbb{R}^n \mapsto \mathbb{C}$ *est* **complètement positive** *si, pour tout* $(t_1, \ldots, t_k) \in (\mathbb{R}^n)^k$ *et tout* $(c_1, \ldots, c_k) \in \mathbb{C}^k$,

$$\sum_{i,j=1}^{k} c_i \bar{c}_j \phi(t_i - t_j) \geq 0.$$

Remarquons qu'une fonction complètement positive est bornée et que $|\phi|(t) \leq \phi(0)$. (Exercice : prendre $k = 2$, $t_1 = 0$, $t_2 = t$, puis jouer sur les différentes valeurs possibles de $c_1$ et $c_2$.)

**Proposition 37.** *Soit* $\mu$ *une mesure positive bornée. Alors* $\hat{\mu}$ *est une fonction complètement positive.*

*Démonstration.* — La démonstration est presque immédiate. On a

$$\sum c_j \bar{c}_l \mu(t_j - t_l) = \int (\sum_{j,l} c_i \bar{c}_l e^{i(t_j - t_l)}) d\mu = \int \left| \sum_j c_j e^{it_j} \right|^2 d\mu \geq 0.$$

∎

En fait, le théorème suivant, que nous énonçons sans démonstration, permet de voir l'importance du résultat précédent

**Théorème 26.** *( Böchner) Si $\phi$ est une fonction complètement positive continue en $0$, alors c'est la transformée de Fourier d'une mesure positive bornée.*

Remarquons qu'une des conséquences de ce résultat est qu'une fonction complètement positive continue en 0 est continue partout et même uniformément continue.

Enfin, nous voyons que la transformée de Fourier caractérise la mesure.

**Théorème 27.** *Si deux mesures bornées ont même transformées de Fourier, elles sont égales.*

*Démonstration.* — Nous ne démontrons ce résultat que dans $\mathbb{R}$, la démonstration étant identique dans $\mathbb{R}^n$. Tout d'abord, remarquons que la transformée de Fourier de $\mu+\bar{\mu}$ est $\hat{\mu}(t)+\bar{\hat{\mu}}(-t)$, tandis que celle de $i(\mu-\bar{\mu})$ est $i(\hat{\mu}(t)-\bar{\hat{\mu}}(-t))$, ce qui nous permet de nous ramener au cas des mesures réelles.

Soit alors deux mesures réelles bornées $\mu_1$ et $\mu_2$ qui ont même transformée de Fourier $\phi$. Elles ont alors même masse $(\phi(0))$, et quitte à les normaliser, on peut supposer que cette masse est 1.

On considère alors l'espace vectoriel $\mathcal{H}$ des fonctions mesurables bornées $f$ qui sont telles que

$$\int f d\mu_1 = \int f d\mu_2.$$

$\mathcal{H}$ est un espace vectoriel de classe monotone, qui contient les constantes. Il contient donc les fonctions de la forme $x \mapsto \cos(ax)$ (en prenant les parties réelles de $\phi$) et les fonctions de la forme $x \mapsto \sin bx$. Il contient l'ensemble $\mathcal{C}$ des combinaisons lineaires de sinus et de cosinus. Cet ensemble est stable par multiplication (formule de multiplication des cosinus et des sinus).

Pour voir que $\mathcal{H}$ constient toutes les fonctions mesurables bornées (en appliquant le théorème fonctionnel des classes monotones) il reste à voir que la tribu engendrée par $\mathcal{C}$ est la tribu borélienne. Mais on a

$$\lim_{a\to 0} \frac{\sin ax}{a} = x,$$

et la tribu engendrée par $\mathcal{C}$ contient la fonction $x \mapsto x$, et donc toute la tribu borélienne.

∎

Un des théorèmes fondamentaux sur la transformée de Fourier est le suivant

**Théorème 28.** *(Riemann-Lebesgue) Si $\mu$ est une mesure bornée à densité par rapport à la mesure de Lebesgue, alors $\hat{\mu}(t)$ converge vers $0$ lorsque $\mu \to \infty$.*

Notons tout de suite que ce théorème est faux si $\mu$ n'est pas à densité, puisque par exemple la transformée de Fourier de la masse de Dirac au point $a$ vaut $e^{it\cdot a}$, qui ne converge pas vers 0 à l'infini.

*Démonstration.* — Nous n'écrivons la démonstration que dans $\mathbb{R}$, le cas de $\mathbb{R}^n$ se traitant de la même manière.

Nous nous ramenons immédiatement au cas où $\mu$ est une mesure réelle bornée. Cela revient à démontrer que si $f \in L^1(\mathbb{R})$, alors

$$\hat{f}(t) = \int e^{it\cdot x} f(x)\lambda(dx) \in \mathcal{C}_0.$$

Commencons par le cas où $f = \mathbf{1}_{[a,b]}$. Dans ce cas

$$\hat{f}(t) = \frac{e^{itb} - e^{ita}}{it} \in \mathcal{C}_0.$$

La transformée de Fourier étant linéaire, c'est encore vraie lorsque $f$ est une combinaison linéaire d'indicatrices d'intervalles fermés bornés (une fonction en escalier).

Soit alors $f \in L^1(\mathbb{R})$. Il existe une suite $(f_n)$ de fonctions en escalier qui converge vers $f$ dans $L^1(\mathbb{R})$. (Approcher une fonction étagée par une combinaison linéaire d'indicatrices d'intervalles, comme nous l'avons fait plus haut).

Mais

$$\left|\hat{f}_n(t) - \hat{f}(t)\right| \leq \int |f_n(x) - f(x)|\,\lambda(dx) = \|f_n - f\|_1\,,$$

et donc $\hat{f}_n$ converge uniformément vers $\hat{f}$, et la limite uniforme d'une suite de fonctions de $\mathcal{C}_0$ est dans $\mathcal{C}_0$.

∎

## 6.3 Formule d'inversion de Fourier

Dans le paragraphe précédent, nous avons vu que la transformée de Fourier caractérise la mesure. mais nous n'avons pas donné de méthode permettant de reconstruire $\mu$ à partir de $\hat{\mu}$. C'est l'objet de ce paragraphe.

Nous commençons par une lemme, qui est important par lui-même.

**Proposition 38.** *Soit*

$$\gamma(dx) = \exp(-\frac{x^2}{2})\frac{\lambda(dx)}{\sqrt{2\pi}}$$

*la mesure gaussienne standard sur* $\mathbb{R}$.

*Alors, la transformée de Fourier de* $\gamma$ *est*

$$\hat{\gamma}(t) = \exp(-\frac{t^2}{2}).$$

*De plus, pour tout* $z \in \mathbb{C}$, *on a*

$$\int |e^{zx}|\,\gamma(dx) < \infty$$

*et*

$$\int e^{zx}\gamma(dx) = e^{\frac{z^2}{2}}.$$

*Démonstration.* — Il est plus simple de calculer

$$L(a) = \int \exp(ax)\gamma(dx) = \exp(\frac{a^2}{2}),$$

pour $a \in \mathbb{R}$ en écrivant

$$L(a) = \exp(\frac{a^2}{2}) \int \exp(-\frac{(x+a)^2}{2}) \frac{\lambda(dx)}{\sqrt{2\pi}} = \exp(\frac{a^2}{2}).$$

Il faut ensuite justifier le passage de $a \in \mathbb{R}$ à $a \in \mathbb{C}$. Pour ce faire, on remarque que, pour $z \in \mathbb{C}$,

$$z \mapsto \int e^{zx}\gamma(dx) = \sum_n \frac{z^n}{n!} \int x^n \gamma(dx),$$

l'intervertion des signes sommes étant justifiée par le fait que

$$|\exp(zx)| \leq \exp(|zx|)$$

et que

$$\int \exp(|zx|)\gamma(dx) < \infty.$$

En appelant $a_n = \int x^n \gamma(dx)$, on a donc

$$\sum_n a_n \frac{z^n}{n!} = \exp(\frac{z^2}{2})$$

pour $z \in \mathbb{R}$, et cette formule reste donc valable pour $z \in \mathbb{C}$. (Deux séries entières qui coïncident sur $\mathbb{R}$ coïncident partout).

∎

**Théorème 29.** *(Formule d'inversion de Fourier) Soit $\mu$ une mesure bornée telle que $\hat{\mu} \in L^1(\mathbb{R})$. Alors, $\mu$ a une densité $\mathcal{C}_0$ $f$ qui est égale à*

$$f(x) = \frac{1}{2\pi} \int \hat{\mu}(t) e^{-itx} \lambda(dt).$$

*On a bien sûr une formule analogue dans $\mathbb{R}^n$.*

*Démonstration.* — Soit $f$ la fonction donnée dans le théorème. Il nous suffit de montrer que les mesures $\nu_1(dx) = f(x)\exp(-\frac{x^2}{2})\lambda(dx)$ et $\nu_2(dx) = \exp(-\frac{x^2}{2})\mu(dx)$ ont même transformées de Fourier.

Or

$$\hat{\nu}_1(s) = \int f(x) e^{-\frac{x^2}{2}} e^{isx} \lambda(dx) = \frac{1}{2\pi} \int \int e^{isx} e^{-itx} e^{-\frac{x^2}{2}} \hat{\mu}(t) \lambda(dt) \lambda(dx).$$

Remplaçant $\hat{\mu}$ par sa valeur, nous obtenons

$$\hat{\nu}_1(s) = \int \int \int e^{i(s-t)x)} e^{-\frac{x^2}{2}} e^{ity} \mu(dy) \lambda(ds) \lambda(dx).$$

On vérifie immédiatement qu'on peut appliquer le théorème de Fubini, et on intègre d'abord par rapport à $x$. On reconnaît la transformée de Fourier de la mesure gaussienne, et on obtient

$$\hat{\nu}_1(s) == \frac{1}{\sqrt{2\pi}} \int e^{-\frac{(s-t)^2}{2}} e^{ity} \lambda(dt)\mu(dy) = \frac{e^{-\frac{s^2}{2}}}{\sqrt{2\pi}} \int e^{-\frac{t^2}{2}+t(s+iy)} \lambda(dt)\mu(dy),$$

qu'on intègre à nouveau par rapport à $t$ pour obtenir

$$\int e^{isy-\frac{y^2}{2}} \mu(dy) = \hat{\nu}_2(s).$$

∎

## 6.4 La formule de Plancherel

La transformée de Fourier est a priori définie sur les mesures bornées, et par conséquent sur l'espace $L^1(\mathbb{R}^n)$ (en identifiant une fonction $f$ avec la mesure $f(x)\lambda(dx)$). Nous allons voir qu'en fait on peut la définir pour des fonctions de $L^2(\mathbb{R}^n)$, et que sur cet espace elle possède des propriétés remarquables.

Dans cette section, pour une fonction $f \in L^1(\mathbb{R}^n)$, nous poserons

$$\mathcal{F}(f)(t) = \int_{\mathbb{R}} e^{it\cdot x} f(x)\lambda(dx),$$

et

$$\mathcal{F}^{-1}(f)(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}} e^{-it\cdot x} f(x)dx.$$

Nous avons vu plus haut que si $f$ et $\mathcal{F}(f)$ sont dans $L^1(\mathbb{R})$, alors $f = \mathcal{F}^{-1}(\mathcal{F}(f))$, ce qui justifie la notation de $\mathcal{F}^{-1}$.

Mais nous ne savons pas a priori pour quelles fonctions $f$ de $L^1$ $\mathcal{F}(f)$ est dans $L^1$ (Nous savons qu'une condition nécéssaire est que $f$ soit dans $\mathcal{C}_0$, mais ce n'est pas suffisant.)

Nous allons tout d'abord décrire un espace vectoriel de "bonnes" fonctions, qui soit stable par $\mathcal{F}$ (et donc par $\mathcal{F}^{-1}$) et sur lequel $\mathcal{F}^{-1}$ soit bien l'inverse de $\mathcal{F}$. Il en existe de nombreux, mais nous allons décrire le plus couramment utilisé : l'espace $\mathcal{S}$ de Schwarz. Dans ce qui suit, pour un multindice $\alpha = (\alpha_1, \ldots, \alpha_n)$ qui est un élément de $\mathbb{N}^n$ nous noterons

$$|\alpha| = \alpha_1 + \ldots + \alpha_n$$

,

$$\partial^\alpha = \partial_{x_1}^{\alpha_1} \partial_{x_2}^{\alpha_2} \ldots \partial_{x_n}^{\alpha_n},$$

pour $x = (x_1, \ldots, x_n)$,

$$x^\alpha = x_1^{\alpha_1} \ldots x_n^{\alpha_n}.$$

On dira que $\alpha \leq \beta$ si pour tout $i$, $\alpha_i \leq \beta_i$ et on définira l'espace $\mathcal{C}^\alpha$ comme l'ensemble des fonctions $f$ telles que $\partial^\alpha f$ est continue. On vérifie que cela ne dépend pas de l'ordre dans lequel on effectue les dérivations partielles, et que si $f \in \mathcal{C}^\alpha$, et $\beta \leq \alpha$, alors $f \in \mathcal{C}^\beta$ (ce qui permet en fait de définir les espaces $\mathcal{C}^\alpha$ de façon inductive en augementant chaque indice de $\alpha$ un à un).

Nous commençons par un lemme

**Proposition 39.** *Soit $\alpha$ un multiindice.*

1. *Si, pour tout $\beta \leq \alpha$, $\left|x^\beta f\right| \in L^1(\mathbb{R}^n)$ alors $\mathcal{F}(f) \in \mathcal{C}^\alpha$ et*
$$\partial^\alpha \mathcal{F}(f) = i^{|\alpha|}\mathcal{F}(x^\alpha f).$$
2. *Si $f \in \mathcal{C}^\alpha$ et si,n pour tout $\beta \leq \alpha$, $\partial^\beta(f) \in L^1$, alors*
$$\mathcal{F}(\partial^\alpha f)(t) = (-i)^{|\alpha|}t^\alpha \mathcal{F}(f)(t).$$

*Démonstration.* — La démonstration de chacun ce des points se fait par récurrence sur chacun des indices $(\alpha_1, \ldots, \alpha_n)$. Pour des raisons de simplicité de notations, nous ne traitons que le cas de la dimension 1, le cas général s'en déduisant par récurrence et le théorème de Fubini. Nous nous contentons alors de traiter le cas $\alpha = 1$, le raisonnement par récurrence qui permet le passge au cas général étant évident.

Pour le premier point, nous écrivons
$$\mathcal{F}(f) = \int_{\mathbb{R}} e^{itx} f(x)dx.$$
Le résultat n'est rien d'autre que le théorème de dérivation sous le signe intégral en la variable $t$, puisque
$$\left|\partial_t e^{itx} f(x)\right| = \left|ix(e^{itx} f(x)\right| \leq |xf(x)|\,,$$
cette dernière quantité étant intégrable par hypothèse.

Pour le second point, nous nous ramenons comme plus haut au cas de la dimension 1.

Nous écrivons
$$\mathcal{F}(\partial_x f) = \lim_{M\to\infty}\int_{-M}^{M} e^{itx}\partial_x f(x)dx = I_M,$$
puisque $\partial_x f$ est intégrable. On intègre alors par parties et on obtient
$$\mathcal{F}(\partial_x f) = [\frac{e^{itx}f(x)}{ix}]_{-M}^{M} - it\int_{-M}^{M} e^{it} f(x)dx.$$
Puisque $f$ est dans $L^1$ le résultat sera démontré dès que nous aurons prouvé que $f(x)$ converge vers 0 à l'infini.

Or,
$$f(x) = f(0) + \int_0^x \partial_x f(t)dt,$$
et puisque $\partial_x f$ est intégrable, $f(x)$ converge vers une limite lorsque $x \to \infty$. Puisque $f$ elle même est intégrable, cette limite ne peut être que nulle. On traîte de même le cas où $x \to -\infty$.

∎

**Définition.** 32. *L'espace $\mathcal{S}$ est l'espace vectoriel des fonctions sur $\mathbb{R}^n$ qui sont $\mathcal{C}^\infty$ et telles que, pour tous les multiindices $\alpha$ et $\beta$, la fonction $x^\beta\partial^\alpha f$ est bornée.*

On vérifie immédiatement que $\mathcal{S}$ est un espace vectoriel. De plus, si $f$ est dans $\mathcal{S}$, alors pour tout couple de multiindices $\alpha$ et $\beta$, et pour tout $p \in \mathbb{N}$, la fonction

$$(1+|x|^2)^p x^\beta \partial^\alpha f$$

est bornée, et donc

$$\left|x^\beta \partial^\alpha f\right| \leq \frac{M_n}{(1+|x|^2)^n},$$

et donc est intégrable.

Par ailleurs, nous avons déjà vu que les fonctions de classe $\mathcal{C}^\infty$ et à support compact sont denses dans tous les espaces $L^p$. Il en est donc de même de $\mathcal{S}$.

**Proposition 40.** *Si $f \in \mathcal{S}$, alors $\mathcal{F}(f)$ est aussi dans $\mathcal{S}$, et*

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

*Il en va de même de $\mathcal{F}(x^\beta \partial^\alpha f)$, pour tout couple de multiindices $\alpha$ et $\beta$.*

*Démonstration.* — Nous avons vu que

$$(-it)^\beta \partial^\alpha(\mathcal{F}(f)) = \mathcal{F}((ix)^\alpha \partial^\beta f).$$

On en déduit même une borne

$$\sup_t \left|t^\beta \partial^\alpha(\mathcal{F}(f))\right| \leq \left\|x^\alpha \partial^\beta f\right\|_1.$$

De là découle les estimées sur la transformée de Fourier d'une fonction de $\mathcal{S}$, et le fait qu'on puisse lui appliquer la formule d'inversion.

∎

On peut alors énoncer le

**<u>Théorème</u> 30.** *(Formule de Plancherel) Si $f \in \mathcal{S}(\mathbb{R}^n)$, alors*

$$\|f\|_2 = (2\pi)^{-\frac{n}{2}} \|\mathcal{F}(f)\|_2.$$

*Démonstration.* — Notons $g = \mathcal{F}(f)$. On écrit

$$\|f\|_2^2 = \int f\overline{\mathcal{F}^{-1}(g)}dx,$$

qu'on développe en

$$\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n\times\mathbb{R}^n} f(x)\bar{g}(t)e^{it\cdot x}dxdt.$$

Tout est fait pour qu'on puisse appliquer le théorème de Fubini, puisque $f$ et $g$ sont intégrables. Alors, on a

$$\|f\|_2^2 = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \bar{g}(t)\mathcal{F}(f)(t)dt = \frac{1}{(2\pi)^n}\int \bar{g}(t)g(t)dt = \frac{1}{(2\pi)^n}\|g\|_2^2.$$

∎

Nous avons alors

**<u>Théorème</u> 31.** *Les applications $\mathcal{F} : \mathcal{S} \mapsto \mathcal{S}$ et $\mathcal{F}^{-1} : \mathcal{S} \mapsto \mathcal{S}$ se prolongent de manière unique en des applications de $L^2$ dans $L^2$, qui sont inverses l'une de l'autre et qui vérifient*

$$\|f\|_2 = (2\pi)^{-\frac{n}{2}} \|\mathcal{F}(f)\|_2 \,,$$

*et*

$$\|f\|_2 = (2\pi)^{\frac{n}{2}} \left\|\mathcal{F}^{-1}(f)\right\|_2 .$$

*Démonstration.* — L'espace $\mathcal{S}$ est dense dans $L^2$. La formule de Plancherel nous montre que si $f_n \in \mathcal{S}$ converge dans $L^2$ vers $f$, alors la suite $\mathcal{F}(f_n)$ est une suite de Cauchy dans $L^2$, et donc y converge. La limite ne dépend pas de la suite choisie : si on a deux suites qui convergent vers $f$, on en construit une troisième en intercalant les deux premières, et la limite obtenue par cette troisième suite est égale aux limites obtenues par les deux premières). On construit ainsi une application qui préserve la norme dans $L^2$ (à la constante $(2\pi)^{n/2}$ près). Et le même raisonnement s'applique à $\mathcal{F}^{-1}$.

∎

La construction de $\mathcal{F}(f)$ pour $f \in L^2$ n'est pas directe. mais nous pouvons au moins constater qu'elle coïncide avec la définition directe de la transformée de Fourier sur $L^1$.

Nous avons

**Proposition 41.** *Si $f \in L^2 \cap L^1$, alors les deux définitions coïncident.*

*Démonstration.* — Il suffit de montrer que si $f \in L^2 \cap L^1$, il existe une suite $(f_n)$ de fonctions de $\mathcal{S}$ qui converge vers $f$ à la fois dans $L^2$ et dans $L^1$, car dans ce cas leur transformée de Fourier convergera d'une part partout vers $\mathcal{F}(f)$ à cause de la convergence $L^1$, et d'autre part dans $L^2$ vers la transformée de Fourier de $f$ (au sens de $L^2$).

Pour cela, on constate d'abord que la suite $f_N = f\mathbf{1}_{\{|x| \leq N\} \cap \{|f| \leq N\}}$ converge vers $f$ à la fois dans $L^2$ et dans $L^1$. On se ramène donc au cas d'une fonction bornée et à support compact. Dans ce cas, il existe une suite de fonctions $\mathcal{C}^\infty$ à support compact qui converge partout vers $f$, en restant uniformément bornées et à support dans un même compact. Cette suite est dans $\mathcal{S}$ et converge vers $f$ à la fois dans $L^2$ et dans $L^1$.

∎

# 7 Appendice

Dans cet appendice, nous donnons la démonstration de quelques résultats du cours laissés en suspens

## 7.1 Le théorème fonctionnel des classes monotones

La démonstration que nous donnons de ce théorème est similaire à celle qui se trouve dans le livre de P. Barbe et M. Ledoux [1].

Rappelons l'énoncé : Sur un ensemble $\Omega$, nous appelons classe monotone $\mathcal{M}$ (de fonctions) un sous-espace vectoriel de l'ensemble de toutes les fonctions bornées $\Omega \mapsto \mathbb{R}$ qui contienne les fonctions constantes et qui soit stable par limite monotone bornée, c'est à dire que si $(f_n)$ est une suite d'éléments de $\mathcal{M}$ qui soit telle que

$$\forall n,\ 0 \leq f_n \leq 1$$

et telle que

$$\forall \omega \in \Omega,\ \forall n,\ f_n(\omega) \leq f_{n+1}(\omega),$$

alors $\lim_n f_n \in \mathcal{M}$.

Attention à ne pas confondre cette notion de classe monotone de fonctions avec celle de classe monotone d'ensembles.

Appelons $\mathcal{M}(\mathcal{C})$ la plus petite classe monotone qui contienne une famille $\mathcal{C}$ de fonctions bornées.

Par ailleurs, appelons $\mathcal{B}(\mathcal{A})$ la classe de toutes les fonctions mesurables bornées par rapport à une tribu $\mathcal{A}$, et $\mathcal{A}(\mathcal{C})$ la plus petite tribu qui rend mesurable toutes les fonctions d'une famille $\mathcal{C}$ de fonctions bornées. Alors, le théorème s'énonce comme suit

**Théorème 32.** *Soit $\mathcal{C}$ une famille de fonctions bornées, stable par multiplication. Alors*

$$\mathcal{M}(\mathcal{C}) = \mathcal{B}(\mathcal{A}(\mathcal{C})).$$

*Démonstration.* — Il est évident que $\mathcal{B}(\mathcal{A}(\mathcal{C}))$ contient $\mathcal{M}(\mathcal{C})$ puisque la classe de toutes les fonctions mesurables bornées par rapport à une tribu est une classe monotone. Tout le problème est de montrer l'inclusion inverse.

Observons tout d'abord qu'on peut se ramener au cas où $\mathcal{C}$ est un espace vectoriel de fonctions stable par multiplication et contenant les constantes. En effet, si $\mathcal{C}_1$ est le plus petit espace vectoriel de fonctions contenant les constantes et les éléments de $\mathcal{C}$ (c'est à dire l'ensemble des combinaisons linéaires finies $f = a + \sum_i \lambda_i f_i$), alors $\mathcal{C}_1$ est stable par multiplication, et

$$\mathcal{C} \subset \mathcal{C}_1 \subset \mathcal{M}(\mathcal{C}),$$

d'où

$$\mathcal{M}(\mathcal{C}) = \mathcal{M}(\mathcal{C}_1),$$

et bien évidement $\mathcal{A}(\mathcal{C}_1) = \mathcal{A}(\mathcal{C})$.

- Etape 1 : $\mathcal{M}(\mathcal{C})$ est stable par multiplication.

  Pour cela, nous raisonnons en deux étapes comme dans le théorème ensembliste des classes monotones .

Tout d'abord, si $f \in \mathcal{C}$ et $g \in \mathcal{M}(\mathcal{C})$, alors $fg \in \mathcal{M}(\mathcal{C})$. En effet, quitte à ajouter à $f$ une constante, et à utiliser le fait que $\mathcal{M}(\mathcal{C})$ est un espace vectoriel, on se ramène au cas où $f$ est positive. On peut aussi supposer que $0 \leq f \leq 1$ quitte à multiplier $f$ par un scalaire.

Il nous reste à voir que

$$\mathcal{M}_1 = \{g,\ fg \in \mathcal{M}(\mathcal{C})\}$$

est une classe monotone, ce qui est immédiat : c'est un espace vectoriel de fonctions car $\mathcal{M}(\mathcal{C})$ en est un et que le produit par $f$ est linéaire, qui contient les constantes car $\mathcal{M}(\mathcal{C})$ contient $\mathcal{C}$, et stable par convergence monotone bornée car si $(0 \leq g_n \leq 1)$ converge en croissant vers $g$, alors $0 \leq fg_n \leq 1$ converge en croissant vers $fg$. Cette classe monotone contient $\mathcal{C}$ puisque $\mathcal{C}$ est stable par multiplication, et dont contient $\mathcal{M}(\mathcal{C})$, ce qui nous donne la propriété annoncée.

Dans un deuxième temps, nous montrons que si $f \in \mathcal{M}(\mathcal{C})$ et $g \in \mathcal{M}(\mathcal{C})$ alors $fg \in \mathcal{M}(\mathcal{C})$, en se ramenant toujours au cas où $0 \leq f \leq 1$ et en considérant

$$\mathcal{M}_2 = \{g,\ fg \in \mathcal{M}(\mathcal{C})\},$$

qui est comme plus haut une classe monotone, qui contient aussi $\mathcal{C}$ par la propriété que nous venons de montrer. On en déduit que $\mathcal{M}(\mathcal{C}) \subset \mathcal{M}_2$, ce que nous voulions démontrer.

Remarquons que ce raisonnement est en tout point identique à celui que nous avions fait pour le cas ensembliste.

- Etape 2 : nous allons montrer que si $\mathcal{M}$ est une classe monotone stable par multiplication, alors c'est l'ensemble des fonctions mesurables bornées par rapport à une tribu, qui n'est rien d'autre que $\mathcal{A}(\mathcal{M})$. En d'autres termes, si une classe monotone $\mathcal{M}$ est stable par multiplication, alors

  $$\mathcal{M} = \mathcal{B}(\mathcal{A}(\mathcal{M})).$$

  C'est un plus délicat.

- Etape 2.1 Nous commencons par remarquer que, si $f \in \mathcal{M}$, alors $|f| \in \mathcal{M}$. Pour le voir, on se ramène au cas où $f$ prend ses valeurs dans l'intervalle $(-1, 1)$ (quitte à la multiplier par un scalaire puisque $f$ est bornée), puis ensuite on écrit

  $$f = \sqrt{1 - (1 - f^2)},$$

  et on développe

  $$\sqrt{1-x} = 1 - \sum_{n=1}^{\infty} a_n x^n.$$

  L'observation cruciale est que dans ce développement de Taylor, tous les coefficients $a_n$ sont positifs (faire le calcul).

  Si bien que

  $$1 - |f| = \sum_n a_n (1 - f^2)^n.$$

$\mathcal{M}$ étant stable par multiplication et un espace vectoriel, si $f \in \mathcal{M}$, il en va de même de

$$g_n = \sum_{p=1}^{n} a_p (1 - f^2)^p.$$

On a

$$0 \leq g_n \leq 1 - |f| \leq 1$$

et $1-|f|$ est bien la limite croissante d'une suite bornée d'élements positifs de $\mathcal{M}$, et donc est dans $\mathcal{M}$. Par suite, $|f| \in \mathcal{M}$.

En conséquence, $f_+ = \frac{f+|f|}{2}$ et $f_- = \frac{|f|-f}{2}$ sont encore dans $\mathcal{M}$ dès que $f$ l'est.

Finalement, si $f$ et $g$ sont dans $\mathcal{M}$, il en va de même de $\max(f, g) = (f - g)_+ + g$ et de $\min(f, g) = -\max(-f, -g)$.

- Etape 2.2 Nous voulons montrer ensuite que, si $A \in \mathcal{A}(\mathcal{M})$ alors $\mathbf{1}_A \in \mathcal{M}$. Pour cela, il nous suffit de remarquer

$$\{A, \mathbf{1}_A \in \mathcal{M}\}$$

est une tribu (car $\mathcal{M}$ est une classe monotone de fonctions stable par multiplication), et donc qu'on peut se ramener au cas où $A = \{f \geq a\}$, pour tous les $f \in \mathcal{M}$ et les $a \in \mathbb{R}$, car ainsi cette tribu contiendra tous les $\sigma(f),\ f \in \mathcal{M}$, et par suite sera égale à $\mathcal{A}(\mathcal{M})$. Quitte à remplacer $f$ par $\alpha f + \beta$, on se ramène au cas où $f \geq 0$ et $a = 1$. Mais alors,

$$\mathbf{1}_{\{f\geq 1\}} = \lim_n \min(f, 1)^n,$$

et cette dernière quantité, limite décroissante d'une suite d'éléments de $\mathcal{M}$ compris entre 0 et 1, est dans $\mathcal{M}$ (le fait de passer des suites croissantes aux suites décroissantes se fait en changeant $g_n$ en $1 - g_n$).

- Etape 2.3 Soit maintenant $f \in \mathcal{B}(\mathcal{A}(\mathcal{M}))$. Nous voulons montrer que $f \in \mathcal{M}$. On peut bien sûr supposer comme plus haut que $f$ est positive. Le point précédent nous montre que les fonctions $\mathcal{A}(\mathcal{M})$-étagées sont dans $\mathcal{M}$, puisque $\mathcal{M}$ est un espace vectoriel. Puisque toute fonction positive mesurable par rapport à une tribu est limite croissante de fonctions mesurables étagées, et que $\mathcal{M}$ est stable par limite monotone bornée, nous obtenons le résultat.

- Etape 3.Pour conclure, nous savons déjà que, si $\mathcal{C}$ est stable par multiplication, $\mathcal{M}(\mathcal{C})$ l'est aussi, et par conséquent

$$\mathcal{B}(\mathcal{A}(\mathcal{C})) \subset \mathcal{B}(\mathcal{A}(\mathcal{M}(\mathcal{C}))) = \mathcal{M}(\mathcal{C}) \subset \mathcal{B}(\mathcal{A}(\mathcal{C})),$$

d'où l'égalité que nous voulions démontrer.

∎

## 7.2 Le théorème de Carathéodory

Nous le démontrons pour les fonctions d'ensembles bornées. La démonstration que nous donnons est recopiée du livre de J. Neveu [2].

Il s'énonce alors de la façon suivante

**Théorème 33.** *Soit $\mathcal{A}_0$ une algèbre de Boole sur un ensemble $\Omega$, et $\mu$ une fonction additive d'ensembles*

$$\mu \ : \ \mathcal{A}_0 \mapsto [0,1].$$

*Pour que $\mu$ se prolonge en une mesure sur $\mathcal{A} = \sigma(\mathcal{A}_0)$, une CNS est que, pour toute suite $(A_n)$ décroissante d'éléments de $\mathcal{A}_0$ telle que $\cap_n A_n = \emptyset$, on ait $\mu(A_n) \to 0$.*

*Démonstration.* — Dans ce qui suit, puisque nous nous restreignons aux fonctions additives d'ensembles bornées, nous pourrons quitte à multiplier $\mu$ par une constante supposer que $\mu(\Omega) = 1$. Par ailleurs, nous préférons caractériser les fonctions additives d'ensemble par le fait que $\mu(\emptyset) = 0$ et que

$$\mu(A \cap B) + \mu(A \cup B) = \mu(A) + \mu(B),$$

pour tous les éléments $A$ et $B$ de $\mathcal{A}_0$.

On appelle $\bar{\mathcal{A}}$ la famille des ensembles qui s'écrivent $B = \cup_n A_n$, où $A_n$ est une suite croissante d'éléments de $\mathcal{A}_0$.

Pour $B \in \bar{\mathcal{A}}$, on pose

$$\bar{\mu}(B) = \sup_{A \in \mathcal{A}_0, A \subset B} \mu(A).$$

Remarquons que $\mathcal{A}_0 \subset \bar{\mathcal{A}}$ et que si $B \in \mathcal{A}_0$, $\bar{\mu}(B) = \mu(B)$.

Nous établissons tout d'abord le lemme

**Lemme 3.** *Si $B \in \bar{\mathcal{A}}$ et $B = \cup_n A_n$, où $A_n$ est croissante d'éléments de $\mathcal{A}_0$, alors*

$$\bar{\mu}(B) = \lim_n \mu(A_n),$$

*cette dernière limite existant puisque $\mu(A_n)$ est une suite croissante bornée de réels positifs.*

*Démonstration.* — Nous voyons tout d'abord par définition de $\bar{\mu}$ que $\bar{\mu}(B)$ que

$$\bar{\mu}(B) \geq \lim_n \mu(A_n),$$

et il nous reste à montrer l'inverse.

Soit $\alpha = \lim_n \mu(A_n)$, et choisissons $\epsilon > 0$. Par définition de $\bar{\mu}(B)$, il existe $A \in \mathcal{A}_0$, tel que $A \subset B$ et $\mu(B) \geq \bar{\mu}(B) - \epsilon$. Mais

$$\cap_n (A \cap A_n^c) \subset A^c \cap B^c = \emptyset,$$

et donc par hypothèse

$$\lim_n \mu(A \cap A_n^c) = 0,$$

ce qui se traduit encore par

$$\lim_n \mu(A \cap A_n) = \mu(A).$$

On en déduit que

$$\lim_n \mu(A_n) \geq \lim_n \mu(A \cap A_n) = \mu(A) \geq \bar{\mu}(B) - \epsilon,$$

et par conséquent

$$\lim_n \mu(A_n) \geq \bar{\mu}(B) - \epsilon.$$

En définitive, ceci étant valable pour tout $\epsilon > 0$, nous voyons que $\bar{\mu}(B) = \lim_n \mu(A_n)$.

∎

La fonction $\bar{\mu}$ satisfait les propriétés suivantes

**Lemme 4.** 1. *$\bar{\mathcal{A}}$ est stable par union et intersection;*

2. *$\bar{\mu}(A \cap B) + \bar{\mu}(A \cup B) = \bar{\mu}(A) + \bar{\mu}(B)$;*
3. *$\bar{\mu}(\emptyset) = 0$, $\bar{\mu}(\Omega) = 1$;*
4. *$\bar{\mathcal{A}}$ est stable par union monotone dénombrable et si $B_n$ est une suite croissante d'éléments de $\bar{\mathcal{A}}$, alors $\lim_n \bar{\mu}(B_n) = \bar{\mu}(\cup_n B_n)$.*

*Démonstration.* — De ces quatre points, seuls le dernier nécessite une démonstration, les autres étant soit élémentaires, soit la conséquence directe du lemme 3. Soit donc $B_n$ une suite croissante d'éléments de $\bar{\mathcal{A}}$ et posons $B = \cup_n B_n$.

Choisissons pour tout $n$ une suite $A_{p,n}$ croissante en $p$ d'éléments de $\mathcal{A}_0$ telle que $\cup_p A_{p,n} = B_n$.

La suite bi-croissante $\bar{A}_{b,n} = A_{1,n} \cup \ldots A_{p,n}$ est aussi d'union en $p$ égale à $B_n$.

La suite $(A_{n,n})$ d'léments de $\mathcal{A}_0$ est croissante, d'union $B$, et on a d'après le lemme 3

$$\lim_n \mu(A_{n,n}) = \bar{\mu}(B).$$

Mais par ailleurs

$$\lim_n \lim_p \mu(A_{p,n}) = \lim_n \mu(B_n).$$

Le résultat s'ensuit pa le lemme fondamental 1 sur les suites bi-croissantes.

∎

La seule chose qui mnque à $\bar{\mathcal{A}}$ pour avoir construit une mesure sur une $\sigma$ algèbre est d'être stable par complémentaire. Nous allons alors étendre la fonction $\bar{\mu}$ à tous les ensembles.

**Définition.** 33. *Pour toute partie $C$ de $\Omega$, nous posons*

$$\mu^*(C) = \inf_{C \subset B,\ B \in \bar{\mathcal{A}}} \bar{\mu}(B).$$

Nous avons alors le

**Lemme 5.** 1. *$\mu^*(\emptyset) = 0$, $\mu^*(\Omega) = 1$;*

2. *$\mu^*(C_1 \cap C_2) + \mu^*(C_1 \cup C_2) \leq \mu^*(C_1) + \mu^*(C_2)$.*
3. *si $C_1 \subset C_2$, $\mu^*(C_1) \leq \mu^*(C_2)$;*
4. *si $C_n$ est une suite croissante, avec $\cup_n C_n = C$, alors $\lim_n \mu^*(C_n) = \mu^*(C)$.*

*Démonstration.* — Les premiers et troisièmes points sont évidents.

Pour le second, choisissons $\epsilon > 0$ et deux éléments $B_1$ et $B_2$ de $\bar{\mathcal{A}}$ tels que $C_i \subset B_i$ et
$$\bar{\mu}(B_i) \leq \mu^*(C_i) + \epsilon.$$

Alors
$$\bar{\mu}(B_1 \cap B_2) + \bar{\mu}(B_1 \cup B_2)) = \bar{\mu}(B_1) + \bar{\mu}(B_2)) \leq \mu^*(C_1) + \mu^*(C_2) + 2\epsilon.$$

Or, $B_1 \cap B_2$ et $B_1 \cup B_2$ sont dans $\bar{\mathcal{A}}$ et $C_1 \cap C_2 \subset B_1 \cap B_2$, $C_1 \cup C_2 \subset B_1 \cup B_2$. On en déduit que
$$\mu^*(C_1 \cap C_2) + \mu^*(C_1 \cup C_2) \leq \mu^*(C_1) + \mu^*(C_2) + 2\epsilon.$$

Ceci étant vrai pour tout $\epsilon > 0$, on en déduit le résultat.

Reste à voir le dernier point. Observons tout d'abord que l'inégalité précédente nous dit qu'on a toujours
$$\mu^*(C_1 \cup C_2) \leq \mu^*(C_1) + \mu^*(C_2).$$

Choisissons $\epsilon > 0$, et, pour $n \geq 1$, choisissons $B_n \in \bar{\mathcal{A}}$, tel que $C_n \subset B_n$ et
$$\bar{\mu}(B_n) \leq \mu^*(C_n) + \frac{\epsilon}{2^n}.$$

Rendons cette suite croissante en posant
$$\bar{B}_n = B_1 \cup \ldots \cup B_n.$$

On voit par récurrence que
$$\bar{\mu}(\bar{B}_n) \leq \mu^*(C_n) + \epsilon(1 + \frac{1}{2} + \ldots + \frac{1}{2^n}).$$

En effet, on a
$$\bar{B}_{n+1} \subset \bar{B}_n \cup (B_{n+1} \setminus B_n),$$
d'où
$$\begin{aligned}\bar{\mu}(\bar{B}_{n+1}) &\leq \bar{\mu}(\bar{B}_n) + \bar{\mu}(B_{n+1}) - \bar{\mu}(B_n) \\ &\leq \bar{\mu}(\bar{B}_n) + \bar{\mu}(B_{n+1}) - \mu^*(C_n) \leq \bar{\mu}(\bar{B}_n) + \mu_*(C_{n+1}) + \frac{\epsilon}{2^n} - \mu^*(C_n),\end{aligned}$$
et la formule s'ensuit par récurrence.

On a donc
$$\bar{\mu}(\bar{B}_n) \leq \mu^*(C_n) + 2\epsilon,$$

et en posant $B = \cup_n \bar{B}_n$, qui est un élément de $\bar{\mathcal{A}}$ qui contient $C$,
$$\bar{\mu}(B) \leq \lim_n \mu^*(C_n) + 2\epsilon.$$

Ceci nous montre que
$$\mu^*(C) \leq \lim_n \mu^*(C_n) + 2\epsilon,$$
et cette inégalité étant vraie pour tout $\epsilon > 0$, on a aussi
$$\mu^*(C) \leq \lim_n \mu^*(C_n).$$

L'inégalité inverse étant évidente, nous avons démontré le dernier point.

∎

Pour conclure, nous observons que pour tout ensemble $C$, nous avons $\mu^*(C)+\mu^*(C^c)\geq 1$.

Nous avons alors le

**Lemme 6.** *Soit $\mathcal{A}=\{C \mid \mu^*(C)+\mu^*(C^c)=1\}$.*

*Alors $\mathcal{A}$ est une $\sigma$-algèbre, et sur cet ensemble, l'application $C\mapsto\mu^*(C)$ est une mesure.*

*Démonstration.* — Nous avons déjà vu que $\mathcal{A}$ contient $\Omega$ et $\emptyset$, et il est clair que $\mathcal{A}$ est stable par passage au complémentaire.

Montrons que $\mathcal{A}$ est stable par union et par instersection. Pour cela, choisissons $A_1$ et $A_1$ dans $\mathcal{A}$. Nous avons

$$\mu^*(A_1\cup A_2)+\mu^*(A_1\cap A_2)\leq\mu^*(A_1)+\mu^*(A_2),$$

et

$$\mu^*(A_1^c\cup A_2^c)+\mu^*(A_1^c\cap A_2^c)\leq\mu^*(A_1^c)+\mu^*(A_2^c).$$

Sommons ces deux inégalités termes à terme. Nous obtenons

$$\mu^*(A_1\cup A_2)+\mu^*((A_1\cup A_2)^c)+\mu^*(A_1\cap A_2)+\mu^*((A_1\cap A_2)^c)\leq$$
$$\mu^*(A_1)+\mu^*(A_2)+\mu^*(A_1^c)+\mu^*(A_2^c)=2.$$

Mais

$$\mu^*(A_1\cup A_2)+\mu^*((A_1\cup A_2)^c)\geq 1,$$

et

$$\mu^*(A_1\cap A_2)+\mu^*((A_1\cap A_2)^c)\geq 1.$$

Il ne peut donc y avoir compatibilité entre ces trois inégalités que si ce sont trois égalités, et ceci montre à la fois que $A_1\cap A_2$ et $A_1\cup A_2$ sont dans $\mathcal{A}$.

Montrons maintenant que $\mathcal{A}$ est stabe par union monotone dénombrable.

Considérons une suite croissante $(A_n)$ d'éléments de $\mathcal{A}$, et soit $A=\cup_n A_n$. On a

$$\mu^*(A_n)+\mu^*(A_n^c)=1,$$

et puisque $A^c\subset A_n^c$, on en déduit

$$\mu^*(A_n)+\mu^*(A^c)\leq 1.$$

En passant à la limite, ce qui est licite en fonctiont du point 4 du lemme 5, nous obtenons

$$\mu^*(A)+\mu^*(A^c)\leq 1.$$

Comme par ailleurs

$$\mu^*(A)+\mu^*(A^c)\geq 1,$$

nous avons le résultat.

Finalement, il nous reste à voir que $\mu^*$ est sur $\mathcal{A}$ une fonction additive d'ensembles.

Reprenons le raisonnement fait un peu plus haut. Si $A_1$ et $A_2$ sont dans $\mathcal{A}$, nous écrivons

$$\mu^*(A_1\cup A_2)+\mu^*(A_1\cap A_2)\leq\mu^*(A_1)+\mu^*(A_2),$$

et

$$\mu^*(A_1^c \cup A_2^c) + \mu^*(A_1^c \cap A_2^c) \leq \mu^*(A_1^c) + \mu^*(A_2^c).$$

En sommant terme à terme, nous trouvons 2 des deux côtés de l'inégalité, ce qui montre que dans chacune de ces inégalités, il y a égalité et donc que $\mu^*$ est une fonction additive d'ensembles.

C'est finalement à nouveau le point 4 du lemme 5 qui nous montre que $\mu^*$ est une mesure sur $\mathcal{A}$.

∎

Pour conclure, observons que $\bar{\mathcal{A}} \subset \mathcal{A}$, et donc que $\mathcal{A}$ contient l'algèbre de Boole $\mathcal{A}_0$. Sur cette algèbre $\mu$ coincide avec $\mu^*$ et on a donc bien montré que la fonction $\mu$ se prolonge en une mesure (unique grâce au théorème des classes monotones). ∎

**Remarque 23.** *La démonstration du théorème de Carathéodory nous donne en fait un peu plus. Si $\mathcal{A}_0$ est une algèbre de Boole avec $\mathcal{A} = \sigma(\mathcal{A}_0)$, si $\mu$ est une mesure sur $\mathcal{A}$ (donc le prolongement de la fonction additive d'ensembles $\mu$ sur $\mathcal{A}_0$), alors pour tout élément $A \in \mathcal{A}$, et tout $\epsilon > 0$, il existe un élemant $B$ qui est une reunion dénombrable d'éléments de $\mathcal{A}_0$ (ou ce qui revient au même une reunion monotone d'éléments de $\mathcal{A}_0$), telle que $A \subset B$ et*

$$\mu(B) \leq \mu(A) + \epsilon.$$

*En passant au complémentaire, on voit aussi qe, pour tout $\epsilon > 0$, il existe un élément $B$ qui est une intersection dénombrable d'éléments de $\mathcal{A}_0$ tel que $B \subset A$ et tel que*

$$\mu(B) \geq \mu(A) - \epsilon.$$

## 7.3 Les théorèmes de Jordan-Hahn et de Radon-Nikodym

Le théorème de Jordan-Hahn (théorème 16) est un théorème de décomposition de mesures bornées non positives. dans ce cours, on l'utilise à plusieurs endroits : dans la démonstration du théorème de Radon-Nikodym ( théorème 14), ainsi que pour établir la dualité $L^p - L^q$ (théorème de Riesz 25).

Commençons par rappeler la notion de mesure signée bornée.

**Définition.** 34. *Soit $(\Omega, \mathcal{A})$ un ensemble $\Omega$ muni d'une $\sigma$-algèbre $\mathcal{A}$ un espace mesurable.*

*On appelle mesure signée bornée sur $(\Omega, \mathcal{A})$ une application $\mu$ de $\mathcal{A}$ à valeurs dans un intervalle borné $[-M, M]$ de $\mathbb{R}$ telle que $\mu(\emptyset) = 0$, et telle que, pour tout $A \in \mathcal{A}$ et pour toute partition dénombrable $(A_n)$ de $A$, la série $\sum_n \mu(A_n)$ converge et on a*

$$\sum_n \mu(A_n) = \mu(A).$$

Rappelons le theorème de Jordan-Hahn (théorème 16-

**Théorème 34** (Jordan-Hahn)**.** *Soit $\mu$ une mesure signée bornée. Alors, il existe un élément $C \in \mathcal{A}$ tel que $A \mapsto \mu(A \cap C)$ et $A \mapsto -\mu(A \cap C^c)$ soient des mesures positives bornées.*

En d'autres termes, toute mesure signée bornée est différence de deux mesures positives bornées à support disjoints. En général, on note cette décomposition $\mu = \mu_+ - \mu_-$ et on note $|\mu|$ la mesure positive bornée $\mu_+ + \mu_-$.

Nous commençons par un lemme

**Lemme 7.** *Pour tout $A \in \mathcal{A}$, définissons*

$$\nu(A) = \sup \sum_n |\mu(A_n)| \, ,$$

*le sup étant pris sur toute les partitions $A = \cup_{n=1}^{\infty} A_n$ de $A$.*

*Alors, $A \mapsto \nu(A)$ est une mesure (positive) bornée.*

*Démonstration.* — Commençons par démontrer que c'est une mesure. Il est clair que $\mu(\emptyset) = 0$, et il reste à voir que, si $A \in \mathcal{A}$ et que $(A_n)$ est une partition de $A$, alors $\nu(A) = \sum_n \nu(A_n)$.

Montrons d'abord que $\sum_n \nu(A_n) \leq \nu(A)$. Comme il est clair que $\nu(B) \leq \nu(A)$ si $A \subset B$, on peut se ramener au cas où pour tout $n$ $\nu(A_n) < \infty$ sinon il n'y a rien à montrer. Choisissons $\epsilon > 0$, et, pour tout $n \geq 1$, une partition $(B_{n,m})$ de $A_n$ telle que

$$\nu(A_n) \leq \sum_m |\mu(B_{n,m}| + \frac{\epsilon}{2^n}.$$

$(B_{n,m})$ forme une partition dénombrable de $A$, et donc

$$\sum_{n,m} |\mu(B_{n,m})| \leq \nu(A),$$

si bien que

$$\sum_n \nu(A_n) \leq \sum_n (\sum_n |\mu(B_{n,m})| + \frac{\epsilon}{2^n} \leq \nu(A) + \sum_n \frac{\epsilon}{2^n} = \nu(A) + \epsilon.$$

Ceci étant vrai pour tout $\epsilon > 0$, on a bien l'inégalité.

Pour montrer l'inverse, commençons par le cas où $\nu(A) < \infty$. Considérons alors $\epsilon > 0$ et choisissons une partition $(B_m)$ de $A$ telle que

$$\nu(A) \leq \sum_m |\mu(B_m| + \epsilon.$$

Posons $B_{n,m} = A_n \cap B_m$. On voit que $(B_{n,m})_m$ est une partition de $A_n$ tandis que $(B_{n,m})_n$ est une partition de $B_m$. On a $\mu(B_m) = \sum_n \mu(B_{n,m})$ et donc

$$|\mu(B_m)| \leq \sum_m |\mu(B_{n,m})| \, .$$

On a donc

$$\nu(A) \leq \sum_{m,n} |\mu(B_{n,m})| + \epsilon.$$

mais

$$\sum_m |\mu(B_{n,m})| \leq \nu(A_n),$$

et par conséquent

$$\nu(A) \leq \sum_n \nu(A_n) + \epsilon,$$

et il ne reste plus qu'à faire converger $\epsilon$ vers 0.

Lorsque $\nu(A) = \infty$, on choisit $M > 0$ et une partition $(B_m)$ de $A$ telle que $\sum_m |\mu(B_m)| \geq M$, et on fait le même raisonnement.

Il nous reste à montrer que la mesure $\nu$ est bornée. Soit $M$ un majorant de $\sup_{A\in\mathcal{A}} |\mu(A)|$, et choisissons $M_1 > 2M$.

Supposons que $\nu(\Omega) = \infty$. On peut donc choisir une choisir une partition $(A_n)$ de $\Omega$ telle que

$$\sum_n |\mu(A_n)| \geq M_1.$$

Appelons $I_+$ l'ensemble des indices $n$ tels que $\mu(A_n) > 0$ et $I_-$ l'ensemble des indices $n$ tels que $\mu(A_n) < 0$. Appelons $\Omega_+ = \cup_{n\in I_+} A_n$ et $\Omega_- = \cup_{n\in I_-} A_n$. On a

$$\mu(\Omega_+) = \sum_{n\in I_+} \mu(A_n),$$

et de même

$$\mu(\Omega_-) = \sum_{n\in I_-} \mu(A_n).$$

On a

$$|\mu(\Omega_+)| \leq M, \ \ |\mu(\Omega_-)| \leq M,$$

et par ailleurs

$$\mu(\Omega_+) - \mu(\Omega_-) = \sum_n |\mu(A_n)| \geq M_1.$$

Ceci aboutit à une contradiction.

∎

Nous pouvons maintenant démontrer le théorème de Jordan-Hahn.

*Démonstration.* — Considérons la mesure $\nu$ donnée par le lemme 7, et rappelons que $|\mu(A)| \leq \nu(A)$ pour tout ensemble $A \in \mathcal{A}$. Considérons la classe

$$\mathcal{B} = \{A \in \mathcal{A}, \ \nu(A) = \mu(A)\}.$$

Alors, $\emptyset \in \mathcal{B}$ et, puisque $\mu$ et $\nu$ sont des mesures, $\mathcal{B}$ est stable par union monotone disjointe. De même, si $B \in \mathcal{B}$ et $B_1 \subset B$, alors $B_1 \in \mathcal{B}$. En effet, sinon, avec $B_2 = B \cap B_1^c$, on aurait

$$\mu(B) = \mu(B_1) + \mu(B_2) < \nu(B_1) + \nu(B_2)$$

et on aboutirait à une contradiction. Ceci montre que $\mathcal{B}$ est aussi stable par union non nécessairement disjointe, car

$$B_1 \cup B_2 = B_1 \cup (B_1 \cap B_1^c).$$

Posons

$$a = \sup_{B\in\mathcal{B}} \nu(B).$$

Nous affirmons qu'il existe un élément $D \in \mathcal{B}$ tel que $\nu(D) = a$.

En effet, il existe pour tout $n \geq 1$ un élément $B_n \in \mathcal{B}$ tel que $\nu(B_n \geq a - 1/n$. Puisque $\mathcal{B}$ est stable par union, on peut supposer la suite $B_n$ croissante, et $D = \cup_n B_n$ répond alors à la question.

Puisque $\nu$ et $\mu$ coïncident sur les sous ensembles de $D$, la restriction de $\mu$ à $D$ est une mesure positive. Il nous reste à voir que la restriction de $\mu$ à $D^c$ est une mesure négative.

Pour cela, posons, pour n'importe quel élément $A \in \mathcal{A}$,

$$\mu_-(A) = \sup_{B \subset A} -\mu(B).$$

En prenant $B = \emptyset$, nous voyons que $\mu_-(A) \geq 0$. Remarquons que $\mathcal{B} = \{A \in \mathcal{A},\ \mu_-(A) = 0\}$. En effet, nous avons vu que si $A \in \mathcal{B}$ et $B \subset A$, alors $B \in \mathcal{B}$ et donc $\mu(B) = \nu(B) \geq 0$, et par conséquent $\mu_-(B) = 0$. Par ailleurs, si $\mu_-(A) = 0$, alors toute partie $B \subset A$ est telle que $\mu(B) \geq 0$, et donc $\nu(A) = \mu(A)$ par définition de $\nu$.

Soit alors un élément $A \in \mathcal{A}$, tel que $A \subset D^c$ et $\mu(A) > 0$. Montrons que nous aboutissons à une contradiction. Tout d'abord, $\mu_-(A) > 0$, car sinon, $A \cup D \in \mathcal{B}$ et

$$\nu(A \cup D) = \nu(A) + \nu(D) > \nu(D),$$

ce qui contredit la définition de $D$.

Montrons qu'alors il existe $A_1 \subset A$ tel que $\mu(A_1) > \mu(A)$ et $\mu_-(A_1) \leq \mu_-(A)/2$.

Choisissons pour cela $B \subset A$ tel que $-\mu(B) \geq \mu_-(A)/2$, ce qui est possible par déinition de $\mu_-$. En posant $A_= A \setminus B$, on voit que, si $C \subset A_1$,

$$-\mu(C) - \mu(B) \leq \mu_-(A),$$

d'où

$$-\mu(C) \leq \mu_-(A) + \mu(B) \leq \frac{\mu_-(A)}{2}.$$

Par conséquent, $\mu_-(A_1) \leq \mu_-(A)/2$, et aussi

$$\mu(A_1) = \mu(A) - \mu(B) > \mu(A).$$

En itérant le procédé, on construit ainsi une suite décroissante $A_n$ telle que $\mu(A_n) \geq \mu(A) > 0$ et $\mu_-(A_n) \leq \mu(A)/2^n$.

En prenant enfin $\hat{A} = \cap_n A_n$, on obtient un ensemble $\hat{A} \subset D^c$, tel que $\mu(\hat{A}) > 0$ et $\mu_-(\hat{A}) = 0$. Nous avons vu que c'est impossible.

∎

**Remarque 24.** *Dans ce théorème de décomposition, l'ensemble $D$ n'est pas unique. Il est défini à un ensemble près tel que $\mu_-(A) = \mu_+(A) = 0$. Ainsi, si $\mu_1 \leq \mu_2$, on pourra toujours choisir pour $\mu_1$ et $\mu_2$ des ensembles de décomposition $D_1$ et $D_2$ (où la restriction de $\mu_i$ à $D_i$ est positive) tels que $D_1 \subset D_2$. On n'aura qu'a remplacer par exemple $D_1$ par $D_1 \cap D_2$, car dans ce cas la restriction de $\mu_1$ à $D_1 \cap D_2^c$ est nulle.*

La mesure positive $\nu$ donnée par le lemme 7 est en général notée $|\nu|$. C'est la seule mesure positive telle que $|\mu| + \mu$ et $|\mu| - \mu$ soient des mesures positives à support disjoints.

Pour démontrer le théorème de Radon-Nikodym, nous aurons besoin encore d'une autre notion.

**Définition.** 35. *Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré où $\mu$ est une mesure bornée. Soit $\mathcal{F}$ une famille quelconque de fonctions mesurables. Nous dirons qu'une fonction mesurable $f$ est un majorant de $\mathcal{F}$ pour l'inéégalité $\mu$-presque partout si $\forall g \in \mathcal{F},\ g \leq f$ ($\mu$-presque partout).*

**Proposition 42.** *Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré où $\mu$ est une mesure bornée. Soit $\mathcal{F}$ une famille quelconque de fonctions mesurables à valeurs dans $[0, \infty]$. Il existe une fonction mesurable $f$ à valeurs dans $[0, \infty]$ qui est un majorant pour l'égalité $\mu$-presque partout, et qui soit plus petite, pour l'inégalité $\mu$-presque partout, que tous les majorants.*

*On l'appelle l'essentiel supremum de la famille $\mathcal{F}$, et on le note* essup $\mathcal{F}$*. Il est unique à l'égalité $\mu$-presque partout près.*

**Remarque 25.** *Bien sûr, lorsque la famille $\mathcal{F}$ est dénombrable, alors il suffit de prendre $\sup \mathcal{F}$. Mais il faudra faire attention à faire la différence entre sup et essup pour des amilles non dénombrables. Ainsi, sur $[0,1]$ muni de la mesure de Lebesgue, si on prend $\mathcal{F} = \{I_{\{x\}},\ x \in [0,1]\}$, alors $\sup \mathcal{F} = 1$ alors que* essup $\mathcal{F} = 0$.

*Démonstration.* — On se ramène d'abord au cas où les fonctions de $\mathcal{F}$ sont à valeurs dans $[0,1]$ en changeant par exemple $f$ en $\tanh(f)$.

Pour toute sous famille dénombrable $J \subset \mathcal{F}$, posons $f_J = \sup J$. Posons $a_J = \int f_J d\mu$. Les fonctions étant toutes à valeurs dans $[0,1]$ et la mesure étant bornée, c'est une famille bornée de réels. Soit alors $A = \sup_J a_J$.

Il existe un sous-ensemble dénombrable $J_0$ de $\mathcal{F}$ tel que $a_{J_0} = A$. En effet, si $J_n$ est une famille dénombrable tele que $a_{J_n} \geq a - 1/n$, $J_0 = \cup_n J_n$ est dénombrable et pour tout $n \geq 1$, $a_{J_0} \geq a_{J_n}$.

Finalement, la fonction $f_{J_0}$ satisfait à l'énoncé. Comme c'est le sup d'une sous-famille dénombrable de $\mathcal{F}$, $f$ est majoré $\mu$-presque sûrement pour tout majorant $\mu$-presque sûr de $\mathcal{F}$. Donc, la seule chose à vérifier est qu'elle est un majorant pour l'inégalité $\mu$-presque partout. Pour cela, soit $f \in \mathcal{F}$ et $J' = J_0 \cup \{f\}$, qui est aussi dénombrable. On a

$$f_{J'} \geq f_{J_0},$$

et

$$\int f_{J'} d\mu \leq A = \int f_{J_0} d\mu.$$

ceci montre que $f_{J_0} = f_{J'}$ $\mu$-presque partout, et donc que $f \leq f_{J_0}$, $\mu$-presque partout.

∎

Finalement, rappelons les notions de mesures étrangères et absolument continues par arpport à une autre.

**Définition.** 36*. Soient $\mu$ et $\nu$ deux mesures bornées sur un ensemble mesurable $(\Omega, \mathcal{A})$. Nous dirons que $\nu$ est absolument continue par rapport à $\mu$, et on note $\nu << \mu$ s'il existe une fonction mesurable positive $f$ telle que, pour tout ensemble $A$, $\nu(A) = \int_A f d\mu$.*

*Nous dirons que $\nu$ est étrangère par rapport à $\mu$ s'il existe un ensemble mesurable $A$ tel que $\mu(A) = 0$ et $\nu(A^c) = 0$. On note alors $\mu \perp \nu$.*

**Théorème 35** (Radon-Nikodym)**.** *Soient $\mu$ et $\nu$ deux mesures bornées sur un espace mesuré $(\Omega, \mathcal{A})$. Alors, on peut décomposer $\mu$ en $\nu = \nu_1 + \nu_2$, où $\nu_1 << \mu$ et $\nu_2 \perp \mu$.*

**Remarque 26.** *On pourra étendre sans problème ce théorème au cas où $\mu$ est une mesure $\sigma$-finie et où $\nu$ est une mesure signée bornée. Remarquons qu'une conséquence importante est que si $\nu(A) = 0$ pour tous les ensembles mesurables $A$ tels que $\mu(A) = 0$, alors $\nu << \mu$. C'est sous cette forme qu'on rencontre le plus souvent le théorème.*

*Démonstration.* — Considérons la classe $\mathcal{F}$ des fonctions mesurables positives telles que, pour tout ensemble $A \in \mathcal{A}$,

$$\int_A f d\mu \leq \nu(A).$$

Montrons tout d'abord que $\mathcal{F}$ est stable par sup de deux fonctions. Etant donné $g_1$ et $g_2$ dans $\mathcal{F}$, et $g = \max(g_1, g_2)$. Soit $B = \{g_1 < g_2\}$. Nous avons

$$\int_A g d\mu = \int_{A \cap B} g_1 d\mu + \int_{A \cap B^c} g_2 d\mu \leq \nu(A \cap B) + \nu(A \cap B^c) = \nu(A).$$

Par ailleurs, $\mathcal{F}$ est stable par limite monotone : c'est une conséquence immédiate du théorème de convergence monotone.

Soit alors $g = \operatorname{essup} \mathcal{F}$, l'essentiel supremum étant calculé pour la mesure $\mu$. La fonction $g$ étant le sup d'une famille dénombrable d'éléments de $\mathcal{F}$, disons la suite $(f_n)$, c'est aussi la limite de la suite croissante $\hat{f}_n = \max(f_1, \ldots, f_n)$, et par ce qui précède $g$ est un élément de $\mathcal{F}$. Soit $\nu_1$ la mesure de densité $g$ par rapport à $\mu$, et soit $\nu_2 = \nu - \nu_1$. L'appartenance de $g$ à $\mathcal{F}$ nous dit que $\nu_2$ est une mesure positive. Pour montrer qu'on a ainsi obtenu la décomposition cherchée, il nous suffit de voir que $\nu_2$ est étrangère à $\mu$, c'est à dire d'exhiber un ensemble $D$ tel que $\mu(D) = 0$ et $\nu_2(D^c) = 0$.

Considérons alors la mesure $\mu_n = \nu_2 - \frac{1}{n}\mu$. Nous écrivons sa décomposition de Jordan-Hahn, qui exhibe un ensemble $D_n$ sur lequel $\mu_n$ est positive et tel que $-\mu_n$ est positive sur $D_n^c$. Puisque $\mu_n \leq \mu_{n+1}$, d'après la remarque 24, nous pouvons choisir la suite $D_n$ croissante. Par ailleurs, par définition de $D_n$, $g + \frac{1}{n}\mathbf{1}_{D_n}$ est un élément de $\mathcal{F}$. Mais puisque $g$ est l'essentiel supremum de $\mathcal{F}$, cela veut dire que $\frac{1}{n}\mathbf{1}_{D_n}$ est égale à 0 $\mu$-presque partout, ou encore que $\mu(D_n) = 0$.

Soit alors $D = \cup_n D_n$. Nous savons que $\mu(D) = 0$. D'autre part, puisque $\mu_n(D_n^c) \leq 0$, nous avons

$$\nu_2(D_n^c) \leq \frac{1}{n}\mu(D_n^c) \leq \frac{1}{n}\mu(\Omega).$$

Puisque $D^c \subset D_n^c$ pour tout $n$, on en déduit que $\nu_2(D^c) = 0$.

∎

**Remarque 27.** *Dans le théorème de Jordan-Hahn, nous avons construit une mesure $|\mu|$ et un ensemble $D$ telle que*

$$d\mu = (I_D - \mathbf{1}_{D^c}) d\,|\mu|\,.$$

*C'est une décomposition de Radon-Nikodym, et on voit ainsi que toute mesure bornée est à densité par rapport à une mesure positive. Par ailleurs, si on se donne une mesure positive $\nu$ et une densité intégrable $h$ non positive, la décomposition de Jordan-Hahn de la mesure signée $hd\nu$ n'est rien d'autre que la décomposition $h_+ d\nu - h_- d\nu$, et donc a posteriori cette décomposition était triviale (mais nous nous sommes servis du théorème de Jordan-Hahn pour démontrer le théorème de Radon-Nikodym).*

## 7.4 La dualité $L^p$-$L^q$

Dans cette partie, nous allons démontrer le théorème de Riesz : si $1 \leq p < \infty$ et si $\frac{1}{p} + \frac{1}{q} = 1$, alors le dual de $L^p$ est $L^q$.

Nous nous restreignons au cas des mesures bornées, le cas des mesures $\sigma$-finies s'en déduisant aisément par restriction (et les détails sont laissés au lecteur).

Quitte à multiplier la mesure par une constante, nous pouvons supposer que $\mu$ est une probabilité.

Soit alors $1 \leq p < \infty$ et $L$ une forme linéaire continue sur $L^p$, pour laquelle il existe donc une constante telle que

$$|L(f)| \leq C \|f\|_p \,,$$

pour toutes les fonctions $f$ de $L^p$. Considérons alors, pour $A \in \mathcal{A}$, $\nu(A) = L(\mathbf{1}_A)$. C'est bien évidement une fonction additive d'ensembles, et nous allons montrer que c'est une mesure bornée. Pour cela, considérons une partition $(A_n)$ d'un ensemble mesurable $A$, et observons que

$$S_n = \sum_{k=1}^{n} \nu(A_k) = \mu(B_n),$$

où $B_n = \cup_{k=1}^n A_k$.

Par suite

$$|\nu(A) - S_n| = \left|L(1_{A\setminus B_n})\right| \leq C \left\|\mathbf{1}_{A\setminus B_n}\right\|_p.$$

Mais

$$\left\|\mathbf{1}_{A\setminus B_n}\right\|_p = \mu(A \setminus B_n)^{1/p} \to 0 \ (n \to \infty),$$

et donc $\sum_1^n \nu(A_k)$ converge vers $\nu(A)$. La mesure $\nu$ est donc bien une mesure bornée. Par ailleurs, elle est clairement absolument continue par rapport à $\mu$ et par suite il existe une fonction $h$ intégrable telle que

$$L(\mathbf{1}_A) = \int h\mathbf{1}_A d\mu,$$

pour tout $A \in \mathcal{A}$.

Cette formule s'étend par linéarité aux fonctions étagées.

Par ailleurs, si $f \in L^p$, on peut décomposer $f$ en $f = f_+ - f_-$, et approcher $f_+$ et $f_-$ par des suites croissantes de fonctions étagées. Ces suites croissantes convergent en fait dans $L^p$ respectivement vers $f_+$ et $f_-$, comme on peut le vérifier aisément par convergence dominée.

Par continuité, la représentation

$$L(f) = \int hf d\mu,$$

s'étend donc à tous les éléments de $L^p$. Il nous reste à montrer que la fonction $h$ est dans $L^q$.

Mais si l'on pose $h_n = h\mathbf{1}_{-n\leq h\leq n}$, qui est une fonction bornée, donc dans tous les $L^p$, et si $\epsilon(h)$ désigne le signe de $h$, on a

$$|h_n^q| = h\epsilon(h)\,|h_n|^{q-1}\,,$$

et donc

$$\|h_n\|_q^q \leq C \left\|\epsilon(h)\,|h_n|^{q-1}\right\|_p = C\,\|h_n\|_q^{q/p}.$$

On en déduit que

$$\|h_n\|_q \leq C,$$

et nous obtenons le résultat en faisant converger $n$ vers l'infini.

# References

[1] P. Barbe and M. Ledoux. *Probabilité, De la licence à l'agrégation.* Espace 34, Belin, Montpellier, 1998.

[2] J. Neveu. *Bases mathématiques du calcul des probabilités.* Masson, Paris, 1970.