Intégrale de Lebesgue

L3 Mathématiques

Jean-Christophe Breton

Université de Rennes 1

Septembre–Décembre 2016

version du 29 novembre 2016

Table des matières

Introduction à la théorie de la mesure

Historiquement, comme l'indique le nom, le but de cette théorie est de mesurer des ensembles. Sans s'en rendre compte, plusieurs types de « mesures » ont déjà été rencontrées :

— Le cardinal d'un ensemble discret, par exemple le cardinal de $\{1, 2, 3, 4\}$ est 4, celui de $\{1, 9, 26, 74, 106\}$ est 5, celui de $\mathbb{N}$ est $+\infty$.

— La longueur, l'aire, le volume d'une courbe, d'une figure plane, d'un solide en dimension 3.
Par exemple, la longueur de l'intervalle $[-3, 5]$ est $5 - (-3) = 8$, l'aire du disque $D(0, R)$ est πR^2, le volume du cylindre de base $D(0, R)$ et de hauteur h est $\pi R^2 h$.

— La probabilité d'un évènement : par exemple si on lance un dé équilibré la probabilité d'avoir un quatre est $1/6$, celle d'avoir une face impaire est $1/2$, celle de gagner au loto (au premier rang) est $1/\binom{49}{5} \sim 5,24 \times 10^{-7}$.

Ces mesures sont des cas particuliers d'une notion plus générale de mesure, outil de base pour une nouvelle théorie de l'intégration, dite intégrale de Lebesgue (1902). Elle généralise la notion déjà vue de l'intégrale de Riemann (cf. [JCB-Riemann]), donc ce qui est déjà connu avec Riemann n'est pas perdu mais généralisé. Cependant, cette nouvelle théorie

— s'applique à une classe de fonctions beaucoup plus grande (les fonctions *mesurables*) ;

— a des théorèmes de convergence beaucoup plus puissants : théorème de convergence monotone, théorème de convergence dominée pour avoir des résultats du type

$$\lim_{n\to+\infty} \int f_n(x)\, dx = \int \lim_{n\to+\infty} f_n(x)\, dx$$

ou

$$\sum_{n\geq 1} \int f_n(x)\, dx = \int \sum_{n\geq 1} f_n(x)\, dx;$$

— traite sans difficulté les intégrales multiples (théorèmes de Fubini-Tonelli et de Fubini) ;

— unifie les différentes façons de mesurer, par exemple le calcul d'une espérance de variables aléatoires, d'une série, d'une intégrale classique sont des cas particuliers d'intégrales au sens de Lebesgue.

Cette théorie unifiante éclaire les analogies souvent constatées en L1, L2 entre les résultats liés aux séries et aux intégrales de Riemann. De plus, cette théorie sert de cadre pour une théorie des probabilités moderne due à Kolmogorov (cf. [JCB-proba]).

Remarque : Notons qu'une mesure est toujours associée à une famille d'ensembles à mesurer. On appelera bientôt ces familles des *tribus* ou des σ-*algèbres*.

Une référence classique pour ce cours est [Rud]. D'autres références sont [ACMR], [Bouyssel], [BP] et [LCP] (en anglais), dont la partie ″théorie de la mesure″ a inspiré une partie de ces notes.

Chapitre 1

Tribus (σ-algèbres) et mesures

Dans ce chapitre, on introduit les notions clefs de théorie de la mesure : les tribus (appelées aussi σ-algèbre) en Section 1.2 et les mesures en Section 1.3. On présente les principales propriétés des mesures en Section 1.4. On présente également la notion de classe monotone, à la base de l'argument du même nom en Section 1.5. On montre comment compléter une tribu en Section 1.6 . On commence ce chapitre par rappeler les opérations ensemblistes de base en Section 1.1.

1.1 Rappels ensemblistes

Dans toute la suite, on considère un ensemble de base X dont on considère des sous-ensembles $E, F, \dots$ et des familles de sous-ensembles. On rappelle que $\mathcal{P}(X)$ désigne la famille de tous les sous-ensembles de X. Par exemple si $X = \{1, 2, 3\}$ alors

$$\mathcal{P}(X) = \big\{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, X = \{1, 2, 3\}\big\}.$$

En général si $\mathrm{card}(X) = n$ alors $\mathrm{card}(\mathcal{P}(X)) = 2^n$. En effet, cela se voit facilement par récurrence : si $n = 0$ alors $X = \emptyset$ et $\mathcal{P}(X) = \{\emptyset\}$ est de cardinal $2^0 = 1$. Si le résultat est acquis lorsque $\mathrm{card}(X) = n$ alors considérons $X' = X \cup \{x'\}$ de cardinal $n + 1$ et notons que les sous-ensembles de X' sont de deux types :

- ceux ne contenant pas x', ce sont alors exactement des sous-ensembles de X, en nombre 2^n (hypothèse de récurrence) ;
- ceux contenant x' et ils sont alors exactement de la forme $E \cup \{x'\}$ où E est un des 2^n sous-ensembles de X ;

Finalement, X' a $2^n + 2^n = 2^{n+1}$ sous-ensembles, ce qui achève la récurrence.

Opérations ensemblistes

On rappelle maintenant les principales opérations ensemblistes sur des sous-ensembles E, F d'un ensemble de base X :

- union : $E \cup F = \{x \in X : x \in E \text{ ou } x \in F\}$;

— intersection : $E \cap F = \{x \in X : x \in E \text{ et } x \in F\}$;
— différence (ensembliste) : $E \setminus F = \{x \in E : x \notin F\}$;
— différence propre : $E \setminus F$ lorsque $F \subset E$;
— différence symétrique : $E \Delta F = (E \setminus F) \cup (F \setminus E)$;
— complémentaire : $E^c = X \setminus E = \{x \in X : x \notin E\}$.

On rappelle quelques règles :
— commutativité : $E \cup F = F \cup E$, $E \cap F = F \cap E$;
— associativité : $(E \cup F) \cup G = E \cup (F \cup G)$, $(E \cap F) \cap G = E \cap (F \cap G)$;
— distributivité : $(E \cup F) \cap G = (E \cap G) \cup (F \cap G)$, $(E \cap F) \cup G = (E \cup G) \cap (F \cup G)$;
— involution : $(E^c)^c = E$;
— lois de Morgan : $(E \cap F)^c = E^c \cup F^c$, $(E \cup F)^c = E^c \cap F^c$;
— $E \setminus F = E \cap F^c$.

On dit que E et F sont **disjoints** si $E \cap F = \emptyset$.
On rappelle que pour montrer une égalité ensembliste $E = F$, le plus simple est de montrer la double inclusion $E \subset F$ et $F \subset E$.

Noter enfin qu'en mathématiques le « ou » est un ou **inclusif** alors que dans le langage usuel il s'agit d'un ou exclusif (thé ou café ? C'est l'un ou l'autre mais pas les deux alors que le « ou » mathématique autorise à prendre les deux).

Les opérations sur les ensembles peuvent faire intervenir plus de deux ensembles. Ainsi si $(E_i)_{i \in I}$ est une famille quelconque d'ensemble indéxée par I alors $\bigcup_{i \in I} E_i$ est l'ensemble des $x \in X$ qui sont dans au moins un des E_i pour $i \in I$. De même $\bigcap_{i \in I} E_i$ est l'ensemble des $x \in X$ qui sont dans tous les E_i pour $i \in I$.

Dénombrabilité

Dans la suite de ce cours, la dénombrabilité est une notion fondamentale. De nombreuses propriétés feront intervenir des familles dénombrables.

Définition 1.1.1 (Dénombrabilité) *On rappelle qu'un ensemble E est dénombrable s'il peut être mis en bijection avec (une partie de) $\mathbb{N}$, ie. il existe une injection f de E dans $\mathbb{N}$.*

Concrètement, un ensemble E est dénombrable si on peut énumérer tous ses éléments. L'ensemble $\mathbb{N}$, bien sûr, est dénombrable mais $\mathbb{Z}$, $\mathbb{Q}$ le sont aussi. Une réunion dénombrable d'ensembles dénombrables reste dénombarable. Par contre $[0, 1]$ ou $\mathbb{R}$ ne le sont pas, ni $\mathbb{N}^{\mathbb{N}}$ ou de façon plus exotique le Cantor non plus, cf. Remarque 2.3.1.
D'ordinaire, le terme dénombrable est utilisé pour les parties infinies dénombrables, mais bien sûr, les ensembles finis sont aussi dénombrables. En général, les familles dénombrables ou les propriétés qui s'expriment en termes de dénombrabilité sont notées avec le préfixe σ pour témoigner de leur caractère dénombrable (exemples : σ-algèbre, σ-additivité).

Limites d'ensembles

Définition 1.1.2 (Suite monotone d'ensembles) — *Une suite $(E_i)_{i\geq 1}$ est dite croissante si pour tout $i \geq 1$, on a $E_i \subset E_{i+1}$. On note alors $\lim_i E_i = \bigcup_{i\geq 1} E_i$.*
— *Une suite $(E_i)_{i\geq 1}$ est dite décroissante si pour tout $i \geq 1$, on a $E_i \supset E_{i+1}$. On note alors $\lim_i E_i = \bigcap_{i\geq 1} E_i$.*

De façon générale, on peut définir les limites inférieure et supérieure d'une suite d'ensembles $(E_i)_{i\geq 1}$ de X.

Définition 1.1.3 (Limites inférieure et supérieure) *Étant donnée une suite d'évènements $(E_i)_{i\geq 1}$, on définit*

$$\text{la limite supérieure : } \limsup_{i\to+\infty} E_i = \bigcap_{i\geq 1}\bigcup_{j>i} E_j$$

$$\text{et la limite inférieure : } \liminf_{i\to+\infty} E_i = \bigcup_{i\geq 1}\bigcap_{j>i} E_j.$$

Noter que

$$\liminf_{i\to+\infty} E_i \subset \limsup_{i\to+\infty} E_i. \tag{1.1}$$

En effet $\bigcap_{k>n} E_k \subset E_q$ pour tout $q > n$. On a donc $\bigcap_{k>n} E_k \subset \bigcup_{q>p} E_q$ pour tout p et tout n.
On a alors pour tout n :

$$\bigcap_{k>n} E_k \subset \bigcap_{p\geq 1}\bigcup_{q>p} E_q = \limsup_{n\to+\infty} E_n.$$

Finalement

$$\bigcup_{n\geq 1}\bigcap_{k>n} E_k \subset \limsup_{n\to+\infty} E_n,$$

c'est à dire l'inclusion (1.1).

De plus, les régles élémentaires sur les $\bigcup$, $\bigcap$ et ^{c} (lois de Morgan) donnent sans difficulté :

$$\left(\limsup_{n\to+\infty} E_n\right)^c = \liminf_{n\to+\infty} E_n^c.$$

Définition 1.1.4 (Suite convergente d'ensembles) *Une suite $(E_i)_{i\geq 1}$ est dite convergente si $\liminf_{i\to+\infty} E_i = \limsup_{i\to+\infty} E_i$.*

Lorsque $(E_i)_{i\geq 1}$ est croissante alors $\liminf_{i\to+\infty} E_i = \limsup_{i\to+\infty} E_i = \bigcup_{i\geq 1} E_i$ et lorsque $(E_i)_{i\geq 1}$ est décroissante alors $\liminf_{i\to+\infty} E_i = \limsup_{i\to+\infty} E_i = \bigcap_{i\geq 1} E_i$. Dans les deux cas, il s'agit évidemment de suites convergentes.

L'intérêt des limites inférieure et supérieure provient notamment de l'interprétation suivante qui permet de « traduire » en langage ensembliste une assertion logique :

Proposition 1.1.1 *Soit E_i, $i \geq 1$, une collection infinie d'ensembles. Alors*

— *« À partir d'un certain rang, x est dans tous les E_i » s'écrit*

$$x \in \bigcup_{i\geq 1}\bigcap_{j>i} E_j \qquad \big(= \liminf_{n\to+\infty} E_n\big).$$

— *« x est dans une infinité de E_i » s'écrit*

$$x \in \bigcap_{i\geq 1}\bigcup_{j>i} E_j \qquad \big(= \limsup_{n\to+\infty} E_n\big).$$

Démonstration :

• Pour le premier point : Soit x qui, à partir d'un certain rang, est dans tous les E_i. On traduit cela de la façon suivante : il existe un rang p tel que pour tout rang $q > p$, x est dans E_q. D'après la signification des symboles $\forall, \exists, \cap, \cup$, cela revient à écrire

$$x \in \underbrace{\bigcup_{p\geq 1}}_{\text{il existe } p\geq 1} \quad \underbrace{\bigcap_{q>p}}_{\text{pour tout } q>p} \quad \underbrace{E_q}_{x \text{ est dans } E_q} .$$

• Pour le second point, dire que x est dans une infinité de E_i est équivalent à dire que

« pour tout p, il existe $q > p$ avec x dans E_q. »

En effet, si tel est le cas, x est bien dans une infinité de E_i car, d'après cette propriété,

— avec $p = 0$, il existe $p_1 > p$ tel que x est dans E_{p_1},
— avec $p = p_1$, il existe $p_2 > p_1$ tel que x est dans E_{p_2},
— avec $p = p_2$, il existe $p_3 > p_2$ tel que x est dans E_{p_3},
— ...
— avec $p = p_n$, il existe $p_{n+1} > p_n$ tel que x est dans $E_{p_{n+1}}$,
— ...

et finalement, x est dans chaque E_{p_n}, $n \geq 1$, c'est à dire dans une infinité de E_i. Réciproquement, s'il est dans une infinité de E_i, alors pour tout p, on trouve $q > p$ tel que $x \in E_q$, sinon, ce serait qu'il existe p tel que pour $q > p$, x n'est pas dans E_q. Ou encore : x ne peut appartenir qu'aux E_i d'indice $i \leq p$, c'est à dire seulement à un nombre fini d'entre eux, ce qui est faux.

Donc, pour ce deuxième point, pour tout p, on trouve $q > p$, tel que $x \in E_q$, en langage $\forall, \exists$, cela s'écrit

$$x \in \underbrace{\bigcap_{p\geq 1}}_{\text{pour tout } p\geq 1} \quad \underbrace{\bigcup_{q>p}}_{\text{il existe } q>p} \quad \underbrace{E_q}_{x \text{ est dans } E_q} .$$

□

1.2 Algèbres et σ-algèbres

On considère dans la suite X un ensemble fixé (des exemples typiques auxquels penser sont $X = \mathbb{N}$ et $X = \mathbb{R}$).

Définition 1.2.1 (Algèbre) *Une famille $\mathcal{A}$ de sous-ensembles de X est une algèbre si*

1. $X \in \mathcal{A}$;
2. *$\mathcal{A}$ est stable par complémentaire : si $A \in \mathcal{A}$ alors $A^c \in \mathcal{A}$;*
3. *$\mathcal{A}$ est stable par réunion (finie) : si $A, B \in \mathcal{A}$ alors $A \cup B \in \mathcal{A}$.*

Remarque 1.2.1

— Nécessairement, l'ensemble vide $\emptyset \in \mathcal{A}$ puisque $\emptyset = X^c$.

— On peut remplacer $X \in \mathcal{A}$ par $\mathcal{A}$ non vide car alors si $E \in \mathcal{A}$, on a aussi $E^c \in \mathcal{A}$ et $X = E \cup E^c \in \mathcal{A}$.

— Une algèbre $\mathcal{A}$ est stable par intersection : si $A, B \in \mathcal{A}$ alors $A \cap B \in \mathcal{A}$ (un ensemble vérifiant une telle propriété est appelée **π-système**).

— Par une récurrence immédiate, une algèbre $\mathcal{A}$ est stable par intersection finie et par union finie.

— Une algèbre $\mathcal{A}$ est stable par différence (ensembliste) : si $A, B \in \mathcal{A}$ alors $A \backslash B \in \mathcal{A}$.

Par exemple $\mathcal{A} = \{\emptyset, X\}$, $\mathcal{P}(X)$, $\{A \subset X \ : \ A \text{ fini ou } X \setminus A \text{ fini}\}$ sont des algèbres de X.

Définition 1.2.2 (σ-algèbre, tribu) *Une famille $\mathcal{A}$ de sous-ensembles de X est une tribu ou une σ-algèbre si*

i) $X \in \mathcal{A}$;

ii) $\mathcal{A}$ est stable par complémentaire : si $A \in \mathcal{A}$ alors $A^c \in \mathcal{A}$;

iii) $\mathcal{A}$ est stable par réunion dénombrable : si pour tout $n \geq 1$, $A_n \in \mathcal{A}$ alors $\bigcup_{n\geq 1} A_n \in \mathcal{A}$.

Conséquences immédiates : pour une σ-algèbre $\mathcal{A}$:

— l'ensemble vide $\emptyset \in \mathcal{A}$;

— une σ-algèbre est une algèbre ;

— $\mathcal{A}$ est stable par intersection dénombrable : si pour tout $n \geq 1$, $A_n \in \mathcal{A}$, alors $\bigcap_{n\geq 1} A_n \in \mathcal{A}$;

— en particulier, $A \cap B \in \mathcal{A}$ et $A \cup B \in \mathcal{A}$ quand A et B sont dans $\mathcal{A}$;

— si $A, B \in \mathcal{A}$ alors $A \setminus B \in \mathcal{A}$.

Exemples : $\{\emptyset, X\}$ (la tribu grossière), $\mathcal{P}(X)$ (la tribu totale), famille d'observables $\mathcal{F}$ en probabilités sont des σ-algèbres.

Remarque 1.2.2 (Explication des axiomes d'une tribu) Une tribu est une famille d'ensembles sur laquelle une « mesure » va être définie. C'est donc une famille d'ensembles à mesurer (ce qu'on appelle en probabilités une famille d'observables).

On comprend bien les axiomes en les interprétant en termes d'évènements probabilistes :

— Si un évènement A est observable, alors l'évènement contraire A^c doit l'être aussi : c'est ce que dit la stabilité par complémentaire.

— Si deux évènements A et B sont observables alors l'évènement A ou B (c'est à dire $A \cup B$) doit l'être aussi : c'est ce que dit la stabilité par réunion.

— La stabilité par réunion finie ne peut suffire. L'exemple suivant illustre ce fait. On lance un dé jusqu'à l'obtention du premier as. Un tel évènement ne pas être décrit par un nombre fini d'évènements élementaires (*a priori* le numéro du lancer du premier as peut être arbitrairement grand). Si on note A l'évènement « on obtient un as » et A_i « on obtient un as au i-ème lancer », alors

$$A = A_1 \cup A_2 \cdots \cup A_n \cdots \cup \cdots = \bigcup_{n\geq 1} A_n$$

et pour que A soit observable quand les A_n le sont, il faut la stabilité par réunion dénombrable de la tribu (des observables) $\mathcal{A}$.
L'axiome de stabilité par réunion dit donc pour une collection dénombrable d'évènements $(A_n)_{n\geq 1}$: si chaque A_n est observable alors l'évènement A_1 ou A_2 ou $\cdots$ ou A_n ou $\cdots$ l'est encore.

— L'évènement certain X est observable, c'est ce qui se cache sous l'axiome X est dans la tribu.

— Les autres opérations sur les évènements observables comme « si A et B sont observables alors A et B aussi » se déduisent des axiomes de base de la définition 1.2.2.

Les axiomes de la définition d'une tribu ne sont donc rien d'autre que la formalisation mathématique des opérations (logiques) qu'on peut faire sur des ensembles « à mesurer » (tels que les évènements).

Définition 1.2.3 (Espace mesurable) *Un ensemble muni d'une tribu $(X, \mathcal{A})$ est appelé un espace mesurable.*

Les ensembles $A \in \mathcal{A}$ s'appellent aussi les (ensembles) mesurables.

Comme l'ensemble des parties $\mathcal{P}(X)$ de X est une tribu, cela semble la tribu la plus commode à considérer sur un ensemble X (puisqu'elle est toujours disponible). Malheureusement, pour beaucoup d'ensembles, cette tribu est trop grande pour définir de bons outils dessus (les mesures) sans incohérence interne. Par exemple, pour généraliser la notion de longueur sur $\mathbb{R}$, on ne peut pas considérer $\mathcal{P}(\mathbb{R})$, il va falloir introduire une nouvelle tribu : la tribu borélienne. De façon générale, pour définir, les bonnes tribus à utiliser (ni trop pauvre ni trop riche, telle que la tribu borélienne), on introduit la notion de tribu engendrée et ce sont les tribus engendrées par de bonnes familles qui nous intéresserons.

Proposition 1.2.1 *Une intersection quelconque de tribus est une tribu. (Les tribus sont des famille de Moore.)*

Démonstration : Soit $\mathcal{A}_i, i \in I$, des tribus. On montre que $\mathcal{A} = \bigcap_{i\in I} \mathcal{A}_i$ en est une aussi.

— D'abord $X \in \mathcal{A}_i$ pour tout $i \in I$ donc $X \in \bigcap_{i\in I} \mathcal{A}_i = \mathcal{A}$.
— Soit $A \in \mathcal{A}$ alors pour tout $i \in I$, $A \in \mathcal{A}_i$, stable par complémentaire donc $A^c \in \mathcal{A}_i$ pour chaque $i \in I$. Mais alors, $A^c \in \mathcal{A} = \bigcap_{i\in I} \mathcal{A}_i$. La famille $\mathcal{A}$ est donc stable par complémentaire.
— Soit pour $j \geq 1$, $A_j \in \mathcal{A} = \bigcap_{i\in I} \mathcal{A}_i$. Pour chaque $i \in I$, $A_j \in \mathcal{A}_i$ donc $\bigcup_{j\geq 1} A_j \in \mathcal{A}_i$ car $\mathcal{A}_i$ est une tribu. Finalemant, $\bigcup_{j\geq 1} A_j \in \bigcap_{i\in I} \mathcal{A}_i = \mathcal{A}$, qui est stable par réunion dénombrable.

La famille $\mathcal{A}$ satisfait tous les axiomes caractéristiques d'une tribu, c'en est donc une. □

Définition 1.2.4 (Tribu engendrée) *Soit $\mathcal{M}$ une famille de sous-ensemble de X ($\mathcal{M} \subset \mathcal{P}(X)$). On note $\sigma(\mathcal{M})$ la plus petite tribu de X (pour l'inclusion) contenant $\mathcal{M}$. On l'appelle la tribu engendrée par $\mathcal{M}$.*

Proposition 1.2.2 *La tribu engendrée par une famille $\mathcal{M}$ est donnée par*

$$\sigma(\mathcal{M}) = \bigcap_{\mathcal{A} \text{ tribu contenant } \mathcal{M}} \mathcal{A}.$$

Démonstration : Notons $\mathcal{B} = \bigcap_{\mathcal{A} \text{ tribu contenant } \mathcal{M}} \mathcal{A}$. D'après la proposition précédente $\mathcal{B}$ est une tribu et clairement elle contient $\mathcal{M}$, si bien qu'on a $\sigma(\mathcal{M}) \subset \mathcal{B}$ d'après la définition de $\sigma(\mathcal{M})$. Puis $\sigma(\mathcal{M})$ est une tribu contenant $\mathcal{M}$ donc par définition de $\mathcal{B}$, intersection de telles tribus, on a $\mathcal{B} \subset \sigma(\mathcal{M})$. Finalement, on a bien $\mathcal{B} = \sigma(\mathcal{M})$. □

Exemple fondamental : Tribu borélienne

On considère un ensemble X muni d'une topologie $\mathcal{T} \subset \mathcal{P}(X)$. On rappelle qu'une **topologie** est une famille d'ensembles $\mathcal{T}$
— contient X ;
— stable par intersection finie ;
— stable par réunion quelconque.

Cette définition est à comparer avec celle d'une tribu en Définition 1.2.2. Les ensembles de $\mathcal{T}$ sont appelés les ouverts de la topologie.

Définition 1.2.5 (Tribu borélienne) *La tribu engendrée par une topologie (c'est à dire engendrée par la famille $\mathcal{M} = \mathcal{T}$ des ouverts d'une topologie) est la tribu (ou σ-algèbre) borélienne, notée $\mathcal{B}(X)$. Les élements de la tribu borélienne $\mathcal{B}(X)$ s'appelent les (ensembles) boréliens de X.*

Il s'agit de la plus petite tribu contenant tous les ouverts de X.

Remarque 1.2.3 La tribu borélienne $\mathcal{B}(X)$ de X contient
— les ouverts U_i ;

— les intersections $\bigcap_{i\in I} U_i$ d'ouverts (I dénombrable) ;
— les réunions d'intersection $\bigcup_{j\in J}\bigcap_{i\in I} U_i$ d'ouverts I (I, J dénombrable) ;
— en généralisant le procédé : les réunions d'intersections de réunions de ... d'ouverts

$$\dots\bigcap_{i\in I}\bigcup_{j\in J}\dots\bigcap_{k\in K}\bigcup_{l\in L}\dots\bigcap_{m\in M}\bigcup_{n\in N}\dots U_n$$

avec des ensembles d'indexation $I, J, \dots, K, L, \dots, M, N, \dots$ dénombrables.

Comme le processus ne s'arrête pas, on ne peut pas décrire tous les boréliens par des réunions d'intersections (etc) d'ouverts. Par contre, dans le cas de $X = \mathbb{R}$, on peut optimiser la famille qui engendre $\mathcal{B}(\mathbb{R})$, c'est à dire choisir une famille encore plus petite que celle des ouverts qui suffit pour retrouver toute la tribu borélienne $\mathcal{B}(\mathbb{R})$ avec les opérations $\bigcup, \bigcap, {}^c$.

Boréliens réels

Les boréliens jouent un rôle essentiel dans l'intégration (réelle). Dans cette section, on considère $X = \mathbb{R} =]-\infty, +\infty[$ muni de sa topologie usuelle (topologie de l'ordre qui coïncide avec la topologie engendrée par la distance usuelle $|\cdot|$). On considère alors sur $\mathbb{R}$ la tribu borélienne $\mathcal{B}(\mathbb{R})$ engendrée par les ouverts de sa topologie usuelle. On rappelle que les ouverts de $\mathbb{R}$ sont des réunions dénombrables d'intervalles ouverts $\bigcup_{n\geq 1}]a_n, b_n[$ (réunion finie ou dénombrable).

Typiquement, les boréliens de $\mathbb{R}$ sont

— tout ouvert, tout fermé
— tout intervalle ouvert, fermé, semi-fermé, fini, infini ;
— tout singleton $\{x\}$, $x \in \mathbb{R}$;
— tout ensemble dénombrable $\{x_i : i \in I\}$, $I \subset \mathbb{N}$, $x_i \in$ *rit*.

En effet, $]a, b[\in \mathcal{B}(\mathbb{R})$ car est ouvert, $[a, b] = \bigcap_{n\geq 1}]a-1/n, b+1/n[$, $[a, b[= \bigcap_{n\geq 1}]a-1/n, b[$, $]a, b] = \bigcap_{n\geq 1}]a, b+1/n[$ sont dans $\mathcal{B}(\mathbb{R})$.
Puis $]-\infty, b] = \Big(\bigcup_{\substack{n\leq [b]\\ n\in\mathbb{Z}}}]n-1, n]\Big)\cup]n, b]$, $]-\infty, b[= \bigcup_{n\geq 1}]-\infty, b-1/n]$, $[a, +\infty] =]-\infty, a[^c$, $]a, +\infty[=]-\infty, a]^c$ sont aussi dans $\mathcal{B}(\mathbb{R})$.
Enfin, $\{x\} = \bigcap_{n\in\mathbb{N}}]x - \frac{1}{n}, x] \in \mathcal{B}(\mathbb{R})$.
Un ensemble dénombrable est une union disjointe de singletons donc est dans $\mathcal{B}(\mathbb{R})$.

Proposition 1.2.3 *Les boréliens de* $\mathbb{R}$ *sont engendrés par*

— *les ouverts ;*
— *les fermés ;*
— *les intervalles ouverts* $]a, b[$ *;*
— *les intervalles fermés* $[a, b]$ *;*
— *les intervalles semi-ouverts* $[a, b[,]a, b]$.
— *les demi-droites fermées* $]-\infty, a]$ *ou* $[b, +\infty[$

Démonstration : On utilise que tout ouvert O de $\mathbb{R}$ s'écrit (par un argument de connexité) comme une réunion dénombrable d'intervalles ouverts $]a_i, b_i[$:

$$O = \bigcup_{i \in I}]a_i, b_i[. \tag{1.2}$$

Puis on utilise des expressions du type

$$\begin{aligned}]a,b[&= \bigcup_{n\geq 1} \big[a+1/n, b-1/n\big], \quad]a,b[= \bigcup_{n\geq 1} \big[a+1/n, b\big[, \\ ]a,b[&= \bigcup_{n\geq 1} \big]a, b-1/n\big] \quad]a,b[=]-\infty, a]^c \cap]-\infty, b[\end{aligned}$$

pour montrer que les familles énoncées permettent de retrouver tous les intervalles ouverts $]a, b[$ et donc par (1.2) les ouverts de $\mathbb{R}$. Les tribus engendrées par ces familles contiennent donc $\mathcal{B}(\mathbb{R})$. Comme en plus elles sont incluses dans $\mathcal{B}(\mathbb{R})$, il y a égalité entre toutes ces tribus. □

Remarque 1.2.4 Ces familles suffisent en appliquant les opérations licites dans les tribus (réunion, intersection, complémentaire) pour retrouver tous les ensembles de la tribu borélienne. Attention toutefois, cela n'empêche pas que certains ensembles de cette tribu peuvent être plus exotiques (ensemble de Cantor, Remarque 2.3.1).

1.3 Mesure

On a déjà évoqué que le cardinal (d'un ensemble), la longueur (d'une courbe), l'aire (d'une surface plane), le volume (d'un solide) ou encore la probabilité (d'un évènement) sont différentes façons de mesurer des objets. Toutes ces notions sont des cas particuliers de la notion générale de mesure. Ces mesures particulières sont associées aux types d'ensemble qu'elles mesurent. Pour une mesure abstraite, la famille d'éléments « mesurables » sera une tribu : on définit une mesure sur une tribu. Considérons un espace mesurable $(X, \mathcal{A})$.

Définition 1.3.1 (Mesure) *Une mesure μ sur $(X, \mathcal{A})$ est une application de $\mathcal{A} \to [0, +\infty]$ telle que*

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

ii) si $(A_n)_{n\geq 1}$ est une suite dénombrable d'ensembles de $\mathcal{A}$ deux à deux disjoints alors

$$\mu\Big(\bigcup_{n\geq 1} A_n\Big) = \sum_{n\geq 1} \mu(A_n) \quad (\sigma\text{-additivité}).$$

En particulier, la mesure peut prendre la valeur $+\infty$ (ce n'est pas choquant, par exemple $\mathbb{N}$ a un cardinal infini, et $\mathbb{R}$ une longueur infinie), mais elle doit être à valeurs positives. Comme $+\infty$ est une valeur autorisée, il convient de définir de nouvelles règles de calcul sur $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$. Ainsi, aux règles réelles usuelles, on rajoute

— pour tout $a \in \mathbb{R}$: $-\infty \pm a = -\infty$, $+\infty \pm a = +\infty$;
— pour $0 < a < +\infty$: $a \times (+\infty) = +\infty$, $a \times (-\infty) = -\infty$;
— pour $-\infty < a < 0$: $a \times (-\infty) = +\infty$, $a \times (+\infty) = -\infty$;
— $+\infty + (+\infty) = +\infty$, $-\infty - (-\infty) = -\infty$;
— $0 \times (\pm\infty) = 0$.

Cette dernière convention se justifie en considérant la surface de $\mathbb{R}$ considérée comme une surface de longueur ∞ et de largeur 0 : son aire est nulle et donc $0 \times (+\infty)$ doit faire zéro. Attention, cette convention induit un problème de continuité pour le produit $(x, y) \in \bar{\mathbb{R}}^2 \mapsto xy$ puisqu'avec $x_n = 1/n$ et $y_n = n$, on a $x_n y_n = 1$, qui n'est pas $\big(\lim_{n\to+\infty} x_n\big)\big(\lim_{n\to+\infty} y_n\big) = 0 \times (+\infty) = 0$.

Définition 1.3.2 (Espace mesuré) *Le triplet $(X, \mathcal{A}, \mu)$ est appelé un espace mesuré (espace mesurable + mesure).*

Remarque 1.3.1 (En pratique)
— Si l'espace X est discret, par exemple $\mathbb{N}$, la tribu totale $\mathcal{P}(X)$ est une bonne tribu à considérer.
— Si l'espace X est topologique, la tribu borélienne $\mathcal{B}(X)$ est une bonne tribu à considérer. Par exemple pour $\mathbb{R}$: $\mathcal{B}(\mathbb{R})$.

Exemples :

• Mesure de comptage (ou de dénombrement) sur $(X, \mathcal{P}(X))$:

$$\eta(A) = \begin{cases} \text{card } A & \text{si } A \text{ est fini} \\ +\infty & \text{sinon.} \end{cases}$$

• Mesure de Dirac sur $(X, \mathcal{P}(X))$: soit $a \in X$,

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

La mesure de Dirac δ_a indique si un ensemble contient ou non le point a.
Par exemple $\delta_0([-1, 1]) = 1, \delta_0(]0, 1]) = \delta_0(\mathbb{R}^*) = 0, \delta_2(\mathbb{N}) = 1, \delta_\pi(\mathbb{Q}) = 0$.

• Exemple fondamental : la mesure de Lebesgue (cf. Chap. 2)

Théorème 1.3.1 (Mesure de Lebesgue) *Il existe une unique mesure λ sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ telle que*
— *pour tout intervalle $[a, b]$ borné, on a : $\lambda([a, b]) = \lambda(]a, b[) = b - a$.*
— *invariance par translation : pour tout $A \in \mathcal{B}(\mathbb{R})$, $\lambda(A + x) = \lambda(A)$ où $A + x = \{a + x : a \in A\}$.*

La mesure de Lebesgue λ généralise la notion de longueur d'un intervalle à tous les boréliens $A \in \mathcal{B}(\mathbb{R})$. Il aurait été vain de chercher à généraliser la longueur à une mesure qui mesure toutes les parties de $\mathbb{R}$ (c'est à dire une mesure sur tout $\mathcal{P}(\mathbb{R})$) car on montre qu'une telle généralisation sur $\mathcal{P}(\mathbb{R})$ contient une incohérence. Il faut se contenter d'une généralisation sur la tribu borélienne $\mathcal{B}(\mathbb{R})$.

• Probabilité : Traditionnellement, dans le cadre probabiliste, on note $(\Omega, \mathcal{F}, \mathbb{P})$ plutôt que $(X, \mathcal{A}, \mu)$. Dans $(\Omega, \mathcal{F}, \mathbb{P})$, Ω est un ensemble (dit espace de probabilité), $\mathcal{F}$ une tribu appelée la famille des observables, et $\mathbb{P}$ une mesure appelée probabilité, c'est à dire une mesure de poids 1 : $\mathbb{P}(\Omega) = 1$.

Définition 1.3.3 *Soit μ une mesure sur $(X, \mathcal{A})$.*
- *On appelle* **poids** *de μ la quantité $\mu(X)$.*
- *Si $\mu(X) < +\infty$, alors la mesure μ est dite* **finie**.
- *Si X se décompose en $X = \bigcup_{n\geq 1} X_n$ avec $\mu(X_n) < +\infty$ alors la mesure est dite* **σ-finie**.
- *Si $\mu(X) = 1$, alors la mesure μ est dite de* **probabilité**.
- *La mesure μ est dite* **borélienne** *si $\mathcal{A} = \mathcal{B}(X)$, ie. μ est définie sur une tribu borélienne.*
- *La mesure μ est dite* **de Borel** *si elle est borélienne et est finie sur tous les ensembles compacts.*

Par exemple :
- Sur $\mathbb{R}$, la mesure de Lebesgue λ est
  - borélienne,
  - une mesure σ-finie car
  $$\mathbb{R} = \bigcup_{n\in\mathbb{N}} [-n, n] \text{ et } \lambda([-n, n]) = 2n;$$
  - une mesure de Borel car pout tout compact K, il existe n tel que $K \subset [-n, n]$ et $\lambda(K) \leq \lambda([-n, n]) = 2n$.
- Une mesure de Dirac δ_a est une mesure de probabilité (dégénérée) car $\delta_a(X) = 1$.
- $\big([0, 1], \mathcal{B}([0, 1]), \lambda\big)$ est un espace de probabilité car $\lambda([0, 1]) = 1$.
- Un exemple de mesure finie est la mesure de dénombrement sur un ensemble fini X puisque la mesure de X est $\mathrm{card}(X)$ qui est fini.

Définition 1.3.4 *Soit μ une mesure sur $(X, \mathcal{A})$.*
- *(Atome) On appelle* **atome** *de μ tout $a \in X$ tel que $\{a\} \in \mathcal{A}$ et $\mu(\{a\}) > 0$.*
- *Si $(X, \mathcal{A})$ est un espace de Borel (ie. X espace topologique et $\mathcal{A}$ tribu de Borel), on appelle* **support** *(topologique) de μ le plus petit ensemble F fermé de X tel que $\mu(F^c) = 0$, il s'agit de l'ensemble des $x \in X$ tel que pour tout voisinage U de x on a $\mu(U) > 0$. Dans la suite, on notera $\mathcal{S}upp(\mu)$ ce support.*

1.4 Propriétés des mesures

• **Croissance :** Une mesure μ est une application croissante : si $A, B \in \mathcal{A}$ avec $A \subset B$, alors $\mu(A) \leq \mu(B)$.
En effet $A = B \cup (A \setminus B)$, comme l'union est disjointe, par additivité, on a

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

et $\mu(A \setminus B) \geq 0$.

• **Additivité :** Si $A, B \in \mathcal{A}$ avec $A \cap B = \emptyset$ alors

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

• **Différence propre :** Si $A, B \in \mathcal{A}$, $B \subset A$ et $\mu(A) < +\infty$, alors $\mu(A \setminus B) = \mu(A) - \mu(B)$.
En effet, cela vient de la décomposition $A = B \cup (A \setminus B)$, union disjointe.

• **Croissance séquentielle** : si $A_n \in \mathcal{A}$ et $A_n \subset A_{n+1}$ pour $n \geq 1$ alors

$$\mu\Big(\bigcup_{n\geq 1} A_n\Big) = \lim_{n\to+\infty} \mu(A_n). \tag{1.3}$$

En effet, on note que $\bigcup_{k=1}^n A_k = A_n$. Soit $C_n = A_n \setminus A_{n-1}$, $n \geq 1$, (avec $A_0 = \emptyset$). On montre facilement que $\bigcup_{k=1}^n C_k = \bigcup_{k=1}^n A_k$: une inclusion vient de ce que $C_k \subset A_k$ pour tout k, l'autre de ce que si $x \in \bigcup_{k=1}^n A_k$ alors en notant k le plus petit entier tel que $x \in A_k$ alors on a $x \in C_k$ car $x \notin A_{k-1}$ et donc $x \in \bigcup_{k=1}^n C_k$. De plus les C_k, $k \geq 1$, sont deux à deux disjoints : si $x \in C_k$ alors $x \notin A_{k-1}$ et donc $x \notin C_l$ pour $l < k$. Comme les C_k sont disjoints, on a

$$\begin{aligned}
\lim_{n\to+\infty} \mu(A_n) &= \lim_{n\to+\infty} \mu\Big(\bigcup_{k=1}^n A_k\Big) = \lim_{n\to+\infty} \mu\Big(\bigcup_{k=1}^n C_k\Big) = \lim_{n\to+\infty} \sum_{k=1}^n \mu(C_k) \quad \text{(additivité)}\\
&= \sum_{k=1}^{+\infty} \mu(C_k) = \mu\Big(\bigcup_{k=1}^{+\infty} C_k\Big) \quad (\sigma\text{-additivité}) \quad = \mu\Big(\bigcup_{k=1}^{+\infty} A_k\Big).
\end{aligned}$$

• **Décroissance séquentielle** : si $A_n \in \mathcal{A}$ et $A_n \supset A_{n+1}$, $n \geq 1$ et si $\mu(A_1) < +\infty$ alors

$$\mu\Big(\bigcap_{n=1}^{+\infty} A_n\Big) = \lim_{n\to+\infty} \mu(A_n). \tag{1.4}$$

Comme $\mu(A_1) < +\infty$, on peut passer aux complémentaires dans A_1 : en notant $B_i = A_1 \setminus A_i$, la suite B_i, $i \geq 1$, est croissante et d'après le cas précédent

$$\lim_{n\to+\infty} \mu(B_n) = \mu\Big(\bigcup_{n\geq 1} B_n\Big).$$

Comme $\mu(A_1) < +\infty$, on a $\mu(B_n) = \mu(A_1) - \mu(A_n)$ et $\bigcup_{n\geq 1} B_n = A_1 \setminus \bigcap_{n\geq 1} A_n$ de mesure $\mu\Big(\bigcup_{n\geq 1} B_n\Big) = \mu(A_1) - \mu\Big(\bigcap_{n\geq 1} A_n\Big)$, on a donc

$$\lim_{n\to+\infty} \big(\mu(A_1) - \mu(A_n)\big) = \mu(A_1) - \mu\Big(\bigcap_{n\geq 1} A_n\Big)$$

et il suffit de retrancher $\mu(A_1) < +\infty$ (et changer le signe) pour conclure.

Contre-exemple pour la décroissance sans hypothèse de finitude sur la mesure : sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, soit $A_n =]n, +\infty[$, on a $\displaystyle\bigcap_{n=1}^{+\infty} A_n = \emptyset$ de mesure nulle tandis que $\lim_{n\to+\infty} \lambda(A_n) = +\infty$.

• μ est **sous-additive** : si $A_n \in \mathcal{A}$ pour $n \geq 1$, alors

$$\mu\Big(\bigcup_{n\geq 1} A_n\Big) \leq \sum_{n\geq 1} \mu(A_n) \qquad \text{(sous-additivité)}.$$

En effet, il suffit de prouver $\mu(A \cup B) \leq \mu(A) + \mu(B)$, la preuve se complète ensuite par récurrence pour avoir $\mu\Big(\bigcup_{i=1}^n A_i\Big) \leq \sum_{i=1}^n \mu(A_i)$ puis par croissance séquentielle (1.3)

$$\mu\Big(\bigcup_{n\geq 1} A_n\Big) = \lim_{m\to+\infty} \mu\Big(\bigcup_{n=1}^m A_n\Big) \leq \lim_{m\to+\infty} \sum_{n=1}^m \mu(A_n) = \sum_{n\geq 1} \mu(A_n).$$

On utilise la décomposition $A \cup B = (A \setminus B) \cup (A \cap B) \cup (B \setminus A)$ en ensembles disjoints, par additivité :

$$\begin{aligned} \mu(A \cup B) &= \mu(A \setminus B) + \mu(A \cap B) + \mu(B \setminus A) \\ &= \mu(A) + \mu(B \setminus A) \\ &\leq \mu(A) + \mu(B). \end{aligned}$$

Définition 1.4.1 (Négligeable, presque partout) — *Un ensemble N de $(X, \mathcal{A}, \mu)$ est dit μ-négligeable s'il existe $A \in \mathcal{A}$ tel que $N \subset A$ et $\mu(A) = 0$.*

— *On dit qu'une propriété est vraie μ-presque partout (μ-p.p.) sur $(X, \mathcal{A}, \mu)$ si l'ensemble des $x \in X$ pour lesquelles elle n'est pas vraie est négligeable. S'il n'y a pas d'ambiguïté sur la mesure μ dont on se sert, on écrit seulement presque partout et on note p.p.*

1.5 Classe monotone

Dans cette section, on introduit un procédé d'extension des définitions de certains objets sur les tribus après les avoir définis sur des classes restreintes d'ensemble.

Définition 1.5.1 (Classe monotone ou λ-système) *Une famille $\mathcal{M}$ de parties de X est appelée classe monotone si*

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

ii) $\mathcal{M}$ *est stable par différence propre : lorsque* $A, B \in \mathcal{M}$ *et* $B \subset A$, *alors* $A \setminus B \in \mathcal{M}$;

iii) $\mathcal{M}$ *est stable par réunion dénombrable croissante* ($A_j \in \mathcal{M}$, $j \geq 1$, $A_j \subset A_{j+1} \Rightarrow \bigcup_{j\geq 1} A_j \in \mathcal{M}$).

Remarque 1.5.1 — Une classe monotone est stable par complémentaire : il suffit d'écrire $A^c = X \setminus A$ pour $X, A \in \mathcal{M}$.

— Une classe monotone est stable par intersection dénombrable décroissante : si $(B_i)_{i\geq 1}$ est une suite de $\mathcal{M}$ telle que $B_i \supset B_{i+1}$, $i \geq 1$, alors $A_i = B_i^c \in \mathcal{M}$ car $A_i = X \setminus B_i$ et $(A_i)_{i\geq 1}$ forme une suite croissante de $\mathcal{M}$ pour laquelle $\bigcup_{i\geq 1} A_i \in \mathcal{M}$ mais alors $\bigcap_{i\geq 1} B_i = \left(\bigcup_{i\geq 1} A_i\right)^c \in \mathcal{M}$.

Une classe monotone est donc stable par limite monotone d'ensembles (croissante ou décroissante), cela justifie la terminologie.

Quelques propriétés

Proposition 1.5.1 *Une intersection d'un nombre quelconque de classes monotones est encore une classe monotone.*

Démonstration : Même type de preuve que pour les tribus dans la Prop. 1.2.1. □

Définition 1.5.2 (Classe monotone engendrée) *Pour $\mathcal{E} \subset \mathcal{P}(X)$, on appelle classe monotone engendrée par $\mathcal{E}$, la plus petite classe monotone contenant $\mathcal{E}$, c'est à dire l'intersection de toutes les classes monotones contenant $\mathcal{E}$. On la note $\mathcal{M}(\mathcal{E})$.*

Proposition 1.5.2 *Une tribu est une classe monotone.*

Démonstration : Les premier et troisième points de la définition1.5.1 sont immédiat pour une tribu $\mathcal{A}$. Pour le deuxième point, il suffit de voir que $A \setminus B = A \cap B^c$ pour s'assurer que $\mathcal{A}$ est stable par différence (propre). □

Par contre, toute classe monotone n'est pas une tribu ; par exemple sur $X = \{a, b, c, d\}$,

$$\mathcal{M} = \big\{\emptyset, \{a, b\}, \{c, d\}, \{b, c\}, \{a, d\}, \{a, b, c, d\}\big\}$$

est bien une classe monotone mais pas un tribu car par exemple $\{a, b\} \in \mathcal{M}$ et $\{b, c\} \in \mathcal{M}$ alors que $\{a, b\} \cup \{b, c\} = \{a, b, c\} \notin \mathcal{M}$. Pour avoir un résultat réciproque à la Prop. 1.5.2, il faut une condition en plus :

Proposition 1.5.3 *Une classe monotone stable par intersection finie est une tribu.*

Démonstration : Une classe monotone est stable aussi par réunion finie en vertu de l'axiome ii) de la Définition 1.5.1, écrire : $\bigcup_{i=1}^n A_i = (\bigcap_{i=1}^n A_i^c)^c$. On utilise alors la reécriture d'une réunion dénombrable comme une réunion croissante : pour toute famille A_j, $j \geq 1$, $\bigcup_{j\geq 1} A_j = \bigcup_{j\geq 1} \Big(\bigcup_{k\leq j} A_k\Big)$.

Théorème 1.5.1 (des classes monotones) *Soit $\mathcal{E}$ une famille de parties de X stable par intersection finie (ie. $\mathcal{E}$ est un π-système). Alors $\mathcal{M}(\mathcal{E}) = \sigma(\mathcal{E})$.*

En pratique, on utilise le résultat sous la forme suivante :

Corollaire 1.5.1 (Classes monotones) *Soit $\mathcal{M}$ une classe monotone contenant la famille de parties $\mathcal{E}$, stable par intersection finie (ie. $\mathcal{E}$ est un π-système). Alors $\sigma(\mathcal{E}) \subset \mathcal{M}$.*

Démonstration : Par le Th. 1.5.1, on a $\sigma(\mathcal{E}) = \mathcal{M}(\mathcal{E})$. Mais comme $\mathcal{M}$ est une classe monotone contenant $\mathcal{E}$ on a aussi par définition de C : $\mathcal{M}(\mathcal{E}) \subset \mathcal{M}$. Finalement, $\sigma(\mathcal{E}) \subset \mathcal{M}$. $\square$

Démonstration : En vertu de l'exemple 2) ci-dessus, $\sigma(\mathcal{E})$ est une classe monotone qui contient $\mathcal{E}$ et donc $\mathcal{M}(\mathcal{E}) \subset \sigma(\mathcal{E})$. Pour prouver l'inclusion réciproque, on montre que $\mathcal{M}(\mathcal{E})$ est stable par intersection finie car alors, d'après l'exemple 3) ci-dessus, $\mathcal{M}(\mathcal{E})$ sera une tribu contenant $\mathcal{E}$, et on aura $\sigma(\mathcal{E}) \subset \mathcal{M}(\mathcal{E})$. Il suffit donc de prouver que si $A, B \in \mathcal{M}(\mathcal{E})$, alors $A \cap B \in \mathcal{M}(\mathcal{E})$. Pour cela, soit

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

Comme $\mathcal{E}$, π-système, est stable par intersection finie, on constate que $\mathcal{M}_1$ contient $\mathcal{E}$. Puis on vérifie facilement que $\mathcal{M}_1$ est une classe monotone car

- $X \in \mathcal{M}_1$ car $X \in \mathcal{M}(\mathcal{E})$ et pour $B \in \mathcal{E}$, $B \cap X = B \in \mathcal{E} \subset \mathcal{M}(\mathcal{E})$.
- si $A_1, A_2 \in \mathcal{M}_1$ avec $A_2 \subset A_1$ alors $A_1 \setminus A_2 \in \mathcal{M}(\mathcal{E})$ car $\mathcal{M}(\mathcal{E})$ est une classe monotone puis pour $B \in \mathcal{E}$, on a $(A_1 \setminus A_2) \cap B = (A_1 \cap B) \setminus (A_2 \cap B)$; mais $A_1 \cap B, A_2 \cap B \in \mathcal{M}(\mathcal{E})$ car $A_1, A_2 \in \mathcal{M}_1$; puis comme $\mathcal{M}(\mathcal{E})$ est stable par différence propre, on a $(A_1 \setminus A_2) \cap B \in \mathcal{M}(\mathcal{E})$; finalement $A_1 \setminus A_2 \in \mathcal{M}_1$.
- si A_j, $j \geq 1$, est dans $\mathcal{M}_1$ avec $A_j \subset A_{j+1}$ alors $\bigcup_{j\geq 1} A_j \in \mathcal{M}(\mathcal{E})$ car $\mathcal{M}(\mathcal{E})$ est stable par réunion croissante ; puis, pour $B \in \mathcal{E}$, $\Big(\bigcup_{j\geq 1} A_j\Big) \cap B = \bigcup_{j\geq 1}(A_j \cap B) \in \mathcal{M}(\mathcal{E})$ car $A_j \cap B \in \mathcal{M}(\mathcal{E})$ et la stabilité par réunion monotone s'ensuit.

Finalement, $\mathcal{M}_1$ est une classe monotone contenant $\mathcal{E}$ donc contient aussi $\mathcal{M}(\mathcal{E})$ et, par définition est contenu dans $\mathcal{M}(\mathcal{E})$, ce qui donne $\mathcal{M}_1 = \mathcal{M}(\mathcal{E})$.
Soit maintenant

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

L'ensemble $\mathcal{M}_2$ est aussi une classe monotone (faire comme précédemment avec $\mathcal{M}_1$ à la place de $\mathcal{E}$). De plus il contient $\mathcal{E}$: on doit pour cela montrer que si $A \in \mathcal{E}$, alors pour

tout $B \in \mathcal{M}(\mathcal{E})$ on a $A \cap B \in \mathcal{M}(\mathcal{E})$. Mais comme $B \in \mathcal{M}(\mathcal{E}) = \mathcal{M}_1$, et $A \in \mathcal{E}$, on a $A \cap B = B \cap A \in \mathcal{M}(\mathcal{E})$ (par définition de $\mathcal{M}_1$). On a donc $\mathcal{M}_2$ classe monotone contenant $\mathcal{E}$ et donc $\mathcal{M}(\mathcal{E}) \subset \mathcal{M}_2$. Comme par définition on a aussi $\mathcal{M}_2 \subset \mathcal{M}(\mathcal{E})$, il vient $\mathcal{M}_2 = \mathcal{M}(\mathcal{E})$ ce qui montre que $\mathcal{M}(\mathcal{E})$ est stable par intersection finie.
D'après l'argument qui commence la preuve, le théorème des classes monotones (Th. 1.5.1) est prouvé. □

Application du théorème des classes monotones

Théorème 1.5.2 (Dynkin) *Soient deux mesures finies μ_1 et μ_2 sur $(X, \mathcal{A})$ de même poids ($\mu_1(X) = \mu_2(X) < +\infty$), qui coïncident sur $\mathcal{C} \subset \mathcal{A}$, sous-famille stable par intersections finies (on parle de π-système) et qui engendre $\mathcal{A}$. Alors μ_1 et μ_2 sont égales sur $\mathcal{A}$.*

Démonstration : Soit $\mathcal{M} = \{A \in \mathcal{A} : \mu_1(A) = \mu_2(A)\}$. Alors, à nouveau, on constate facilement que $\mathcal{M}$ est une classe monotone qui contient $\mathcal{C}$ (hypothèse) :

- $X \in \mathcal{M}$ car $\mu_1(X) = \mu_2(X)$ par hypothèse.
- Si $A, B \in \mathcal{M}$ avec $B \subset A$ alors
$$\mu_1(A \setminus B) = \mu_1(A) - \mu_1(B) = \mu_2(A) - \mu_2(B) = \mu_2(A \setminus B).$$
- Si $A_j \in \mathcal{M}$ avec $A_j \subset A_{j+1}$, $i \geq 1$, alors par croissance séquentielle (1.3) :
$$\mu_1\Big(\bigcup_{j\geq 1} A_j\Big) = \lim_{j\to+\infty} \mu_1(A_j) = \lim_{j\to+\infty} \mu_2(A_j) = \mu_2\Big(\bigcup_{j\geq 1} A_j\Big).$$

Comme $\mathcal{C}$ est un π-système, le Théorème des classes monotones (Th. 1.5.1) garantit que $\sigma(\mathcal{C}) \subset \mathcal{M}$ ce qui conclut la preuve du théorème de Dynkin. □

Le résultat suivant est une version σ-finie du théorème de Dynkin (Th. 1.5.2) :

Corollaire 1.5.2 *Soit $(X, \mathcal{A})$ un espace mesuré et μ_1, μ_2 deux mesures telles que $X = \bigcup_{n\geq 1} X_n$ avec $\mu_1(X_n) = \mu_2(X_n) < +\infty$. Si μ_1 et μ_2 coïncident sur une famille $\mathcal{C} \subset \mathcal{A}$ stable par intersection (π-système) contenant les ensembles X_n, $n \geq 1$, et qui engendre $\mathcal{A}$. Alors $\mu_1 = \mu_2$ sur $\mathcal{A}$.*

Démonstration : Quitte à remplacer les X_n par $Y_n = \bigcup_{k=1}^n X_k$, $n \geq 1$, et à utiliser la croissance séquentielle de μ_1, μ_2, on peut supposer les X_n, $n \geq 1$, croissants. Le théorème de Dynkin (Th. 1.5.2) s'applique à $\mu_1^{(n)} = \mu_{1|X_n}$ et $\mu_2^{(n)} = \mu_{2|X_n}$ puisque ce sont des mesures finies de même poids. On a pour tout $A \in \sigma(\mathcal{C})$: $\mu_1^{(n)}(A) = \mu_2^{(n)}(A)$, ie. $\mu_1(A \cap X_n) = \mu_2(A \cap X_n)$. Mais par croissance séquentielle (1.3), on a

$$\lim_{n\to+\infty} \mu_1(A \cap X_n) \quad = \quad \mu_1\Big(\bigcup_{n\geq 1}(A \cap X_n)\Big) = \mu_1(A)$$

$$\lim_{n\to+\infty}\mu_2(A\cap X_n) \quad = \quad \mu_2\Big(\bigcup_{n\geq 1}(A\cap X_n)\Big) = \mu_2(A).$$

On a donc $\mu_1(A) = \mu_2(A)$ pour tout $A \in \sigma(\mathcal{C})$ en passant à la limite dans $\mu_1^{(n)}(A) = \mu_2^{(n)}(A)$.
□

Remarque 1.5.2 Avec $\mathcal{C}$ l'ensemble des intervalles de $\mathbb{R}$ (stable par intersection finie) et $X_n = [-n, n]$, on montre l'unicité de la mesure de Lebesgue sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (car on a $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{C})$).

1.6 Tribu complète

Si μ est une mesure sur une famille $\mathcal{E}$ et $E \in \mathcal{E}$ est tel que $\mu(E) = 0$ alors $\mu(F) = 0$ pour tout $F \in \mathcal{E}$ tel que $F \subset E$. Cependant si $F \subset E$ mais $F \notin \mathcal{E}$, $\mu(F)$ n'est pas définie et on ne peut pas dire $\mu(F) = 0$. Ce constat déplaisant suggère que la mesure μ n'est pas définie pour assez d'ensemble et incite à compléter la mesure en une mesure sur une famille $\bar{\mathcal{E}}$ qui vérifierait : si $\mu(E) = 0$ pour $E \in \bar{\mathcal{E}}$, alors nécessairement $F \subset E$ vérifie $F \in \bar{\mathcal{E}}$ et $\mu(F) = 0$. On dit alors que la tribu $\bar{\mathcal{E}}$ et la mesure $\bar{\mu}$ sont complètes. Lorsqu'une tribu $\mathcal{A}$ n'est pas complète (pour une mesure μ), on peut considérer la plus petite tribu complète la contenant, on l'appelle la tribu complétée de $\mathcal{A}$. Cette complétion est l'objet de cette section, pour cela on élargit la famille sur laquelle μ est définie en ajoutant tous les ensembles qui devraient avoir pour mesure 0.

On considère une mesure μ sur une σ-algèbre $\mathcal{A}$ et on étend μ sur une tribu complétée. Pour cela, on considère la famille d'ensembles suivante :

$$\mathcal{A}_\mu = \big\{E \subset X \ : \ \exists A_1, A_2 \in \mathcal{A}, \text{ tels que } A_1 \subset E \subset A_2 \text{ et } \mu(A_2 \setminus A_1) = 0\big\}. \tag{1.5}$$

Proposition 1.6.1 *En notant $\mathcal{N}$ la classe des ensembles μ-négligeables, ie. $\mathcal{N} = \{N \subset X : \exists A \in \mathcal{A} : N \subset A, \mu(A) = 0\}$, on a $\mathcal{A}_\mu = \sigma(\mathcal{A} \cup \mathcal{N})$.*

Démonstration : D'abord, on montre que $\mathcal{A}_\mu$ en (1.5) est une σ-algèbre :
- $X \in \mathcal{A}_\mu$ puisqu'avec $A_1 = A_2 = X$ on vérifie la définition.
- si $E \in \mathcal{A}_\mu$ alors $\exists A_1, A_2 \in \mathcal{A}$, tels que $A_1 \subset E \subset A_2$ et $\mu(A_2 \setminus A_1) = 0$ et on a $A_1^c, A_2^c \in \mathcal{A}$, tels que $A_2^c \subset E^c \subset A_1^c$ et $\mu(A_1^c \setminus A_2^c) = \mu(A_1^c \cap A_2) = \mu(A_2 \setminus A_1) = 0$. On a donc bien $E^c \in \mathcal{A}_\mu$.

— Soit $(E_i)_{i\geq 1}$ une suite $\mathcal{A}_\mu$. Avec des notations évidentes, on a

$$\bigcup_{i\geq 1} A_1^i \subset \bigcup_{i\geq 1} E_i \subset \bigcup_{i\geq 1} A_2^i$$

avec $\bigcup_{i\geq 1} A_1^i, \bigcup_{i\geq 1} A_2^i \in \mathcal{A}$ et

$$\Big(\bigcup_{i\geq 1} A_2^i\Big) \setminus \Big(\bigcup_{i\geq 1} A_1^i\Big) \subset \bigcup_{i\geq 1} (A_2^i \setminus A_1^i)$$

et donc

$$\mu\left(\Big(\bigcup_{i\geq 1} A_2^i\Big) \setminus \Big(\bigcup_{i\geq 1} A_1^i\Big)\right) \leq \mu\Big(\bigcup_{i\geq 1} (A_2^i \setminus A_1^i)\Big) \leq \sum_{i\geq 1} \mu(A_2^i \setminus A_1^i) = 0.$$

Comme $\mathcal{A} \subset \mathcal{A}_\mu$ et $\mathcal{N} \subset \mathcal{A}_\mu$ on a $\mathcal{A}\cup\mathcal{N} \subset \mathcal{A}_\mu$ et puisque $\mathcal{A}_\mu$ est une tribu, $\sigma(\mathcal{A}\cup\mathcal{N}) \subset \mathcal{A}_\mu$. Puis si $E \in \mathcal{A}_\mu$ alors $E = A_1 \cup (E \setminus A_1) \in \sigma(\mathcal{A}\cup\mathcal{N})$ puisque $A_1 \in \mathcal{A}$ et $E \setminus A_1 \in \mathcal{N}$. □

Théorème 1.6.1 *On définit $\bar{\mu}$ sur $\mathcal{A}_\mu$ par $\bar{\mu}(E) = \mu(A_1) = \mu(A_2)$ lorsque $A_1 \subset E \subset A_2$ avec $A_1, A_2 \in \mathcal{A}$ et $\mu(A_2 \setminus A_1) = 0$. La fonction d'ensemble $\bar{\mu}$ est bien définie et il s'agit de la seule mesure sur $\mathcal{A}_\mu$ qui prolonge μ (ie. qui coïncide avec μ sur $\mathcal{A}$).*

Démonstration : Si on a $A_1 \subset E \subset A_2$ et $B_1 \subset E \subset B_2$ alors $A_1 \cup B_1 \subset E \subset A_2 \cap B_2$ avec $A_1 \cup B_1 \in \mathcal{A}$ et $A_2 \cap B_2 \in \mathcal{A}$ et comme $(A_2 \cap B_2) \setminus (A_1 \cup B_1) \subset A_2 \setminus A_1$, on a

$$\mu\big((A_2 \cap B_2) \setminus (A_1 \cup B_1)\big) = 0.$$

Comme $A_1 \subset A_1 \cup B_1 \subset A_2 \cap B_2 \subset A_2$ et $\mu(A_1) = \mu(A_2)$, on a en fait

$$\mu(A_1 \cup B_1) = \mu(A_2 \cap B_2) = \mu(A_1) = \mu(A_2) = \mu(B_1) = \mu(B_2)$$

ce qui assure de la cohérence de la définition de $\bar{\mu}$ via A_1, A_2 ou B_1, B_2.
Il s'agit d'une mesure car $\bar{\mu}(\emptyset) = 0$ et si les $E_i \in \mathcal{A}_\mu$, $i \geq 1$, sont disjoints alors avec des notations évidentes

$$\bigcup_{i\geq 1} A_1^i \subset \bigcup_{i\geq 1} E_i \subset \bigcup_{i\geq 1} A_2^i$$

et

$$\bar{\mu}\Big(\bigcup_{i\geq 1} E_i\Big) = \mu\Big(\bigcup_{i\geq 1} A_1^i\Big) = \sum_{i\geq 1} \mu(A_1^i) = \sum_{i\geq 1} \bar{\mu}(E_i)$$

car les A_1^i, $i \geq 1$, sont disjoints, et ce qui établit la σ-additivité.
De plus, il est clair que $\bar{\mu}$ prolonge μ puisque si $E \in \mathcal{A}$ alors $A_1 = E = A_2$ et $\bar{\mu}(E) = \mu(E)$.
Puis, si ν est une autre mesure sur $\mathcal{A}_\mu$ prolongeant μ. Soit $E \in \mathcal{A}_\mu$ alors

$$\nu(E) = \nu\big(A_1 \cup (E \setminus A_1)\big) = \nu(A_1) + \nu(E \setminus A_1)$$

mais $\nu(A_1) = \mu(A_1) = \bar{\mu}(E)$ et $\nu(E \setminus A_1) \leq \nu(A_2 \setminus A_1) = \mu(A_2 \setminus A_1) = 0$. Ainsi on a bien $\nu(E) = \bar{\mu}(E)$. □

On poursuivra l'étude d'espace mesuré complété avec les fonctions $\mathcal{A}_\mu$-mesurables au Chap. 3 en montrant qu'on peut les approximer par des fonctions mesurables par rapport à $\mathcal{A}$, cf. Prop. **??**.

Chapitre 2

Mesure de Lebesgue

Dans ce chapitre, on construit la mesure de Lebesgue avec la notion de mesure extérieure qu'on introduit en Section 2.1. La mesure de Lebesgue est construite en Section 2.2 et on discute les principales propriétés de la mesure de Lebesgue Section 2.3.

2.1 Mesure extérieure

On introduit la notion de mesure extérieure. Ce nouveau type de mesure présente l'avantage de pouvoir être défini sur tous les sous-ensembles d'un ensemble donné X, a priori donc sans structure de tribu. L'inconvénient est qu'une mesure extérieure dispose de bien moins de propriétés, cf. Section 1.4. Mais dans la section suivante, les mesures extérieures seront un outil pour construire des (vraies) mesures intéressantes (en particulier la mesure de Lebesgue).

Définition 2.1.1 (Mesure extérieure) *On appelle mesure extérieure toute fonction d'ensemble μ^* positive, définie pour tous les sous-ensembles E de X telle que*

- $\mu^*(\emptyset) = 0$;
- *μ^* est monotone, ie. $A \subset B$ implique $\mu^*(A) \leq \mu^*(B)$;*
- *μ^* est σ-sous-additive, ie. si $(E_i)_{i\geq 1}$ est une famille dénombrable quelconque alors*

$$\mu^*\Big(\bigcup_{i\geq 1} E_i\Big) \leq \sum_{i\geq 1} \mu^*(E_i).$$

On associe à une mesure extérieure une notion de mesurabilité :

Définition 2.1.2 (μ^*-mesurabilité) *Un ensemble E est dit μ^*-mesurable si pour tout sous-ensemble A de X, on a*

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c) \tag{2.1}$$

ie. E est μ^*-mesurable si tout ensemble se décompose additivement relativement à μ^*. On note $\mathcal{S}_{\mu^*}$ la famille des ensembles μ^* mesurables.

Comme l'autre sens est due à la sous-additivité, il suffit de montrer $\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c)$ pour établir (2.1).

Proposition 2.1.1 *Soient $E, F \in \mathcal{S}_{\mu^*}$ et $A \subset X$.*

1. $\mu^*(A) = \mu^*(A \cap E \cap F) + \mu^*(A \cap E \cap F^c) + \mu^*(A \cap E^c \cap F) + \mu^*(A \cap E^c \cap F^c)$;
2. $\mu^*\big(A \cap (E \cup F)\big) = \mu^*(A \cap E \cap F) + \mu^*(A \cap E^c \cap F) + \mu^*(A \cap E \cap F^c)$;
3. *Si E, F sont de plus disjoints alors*
$$\mu^*\big(A \cap (E \cup F)\big) = \mu^*(A \cap E) + \mu^*(A \cap F).$$

Démonstration : 1) Comme E est μ^*-mesurable, on a
$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c) \tag{2.2}$$
Puis comme F est μ^*-mesurable, on a
$$\mu^*(A \cap E) = \mu^*(A \cap E \cap F) + \mu^*(A \cap E \cap F^c) \tag{2.3}$$
$$\mu^*(A \cap E^c) = \mu^*(A \cap E^c \cap F) + \mu^*(A \cap E^c \cap F^c) \tag{2.4}$$
ce qui donne le résultat en reportant (2.3), (2.4) dans (2.2).
2) vient de 1) en écrivant remplaçant A par $A \cap (E \cup F)$ et en utilisant
$$A \cap (E \cup F) \cap E \cap F = A \cap E \cap F, \quad A \cap (E \cup F) \cap E^c \cap F^c = \emptyset.$$
3) suit immédiatement de 1) avec $E \cap F = \emptyset$. □

Proposition 2.1.2 *Si μ^* est une mesure extérieure, la famille $\mathcal{S}_{\mu^*}$ des μ^*-mesurables est une σ-algèbre. Si $(E_i)_{i \geq 1}$ est une suite d'ensembles disjoints de $\mathcal{S}_{\mu^*}$ et $E = \bigcup_{i \geq 1} E_i$ alors $\mu^*(E) = \sum_{i \geq 1} \mu^*(E_i)$. Ainsi, la restriction $\mu = \mu^*_{|\mathcal{S}_{\mu^*}}$ de μ^* à $\mathcal{S}_{\mu^*}$ est une mesure sur $\mathcal{S}_{\mu^*}$.*

Démonstration : On commence par montrer que $\mathcal{S}_{\mu^*}$ est une σ-algèbre. Il suit immédiatement de la définition de μ^* que si $E \in \mathcal{S}_{\mu^*}$ alors $E^c \in \mathcal{S}_{\mu^*}$ si bien que $\mathcal{S}_{\mu^*}$ est stable par complémentaire.
Puis, si $E, F \in \mathcal{S}_{\mu^*}$ et $A \subset X$ alors d'après la Prop. 2.1.1, on a
$$\begin{aligned} \mu^*(A) &= \mu^*\big(A \cap (E \cup F)\big) + \mu^*(A \cap E^c \cap F^c) \\ &= \mu^*\big(A \cap (E \cup F)\big) + \mu^*\big(A \cap (E \cup F)^c\big) \end{aligned}$$
si bien que $E \cup F \in \mathcal{S}_{\mu^*}$ et donc $\mathcal{S}_{\mu^*}$ est une algèbre (non vide car $X \in \mathcal{S}_{\mu^*}$). Pour montrer que $\mathcal{S}_{\mu^*}$ est une σ-algèbre, il reste à établir la stabilité par union dénombrable.

En fait, il suffit de voir la stabibilité par union dénombrable d'ensembles disjoints car une union dénombrable $\bigcup_{n \geq 1} A_n$ s'écrit comme une union dénombrable disjointe $\bigcup_{n \geq 1} B_n$ avec $N_n \in \mathcal{S}_{\mu^*}$ quand $A_n \in \mathcal{S}_{\mu^*}$: en effet, on écrit d'abord $\bigcup_{n \geq 1} A_n = \bigcup_{n \geq 1} C_n$ avec les

$C_n = \bigcup_{k=1}^n A_k$ croissants, puis on prend $B_1 = C_1 = A_1$, $B_2 = C_2 \setminus C_1, \dots B_n = C_n \setminus C_{n-1}$, les B_n sont bien disjoints (croissance des C_n) et $\bigcup_{k=1}^n B_k = \bigcup_{k=1}^n C_k = \bigcup_{k=1}^n A_k$ donc $\bigcup_{n\geq 1} B_n = \bigcup_{n\geq 1} A_n$; puis comme $\mathcal{S}_{\mu^*}$ est une algb̀re, si les A_n sont dans $\mathcal{S}_{\mu^*}$, alors les C_n aussi (stabilité par union finie) et les $B_n = C_n \cap C_{n_1}^c$ aussi (stabilité par intersection et par complémentaire).

On montre donc que si $(E_i)_{i\geq 1}$ est une suite d'ensembles disjoints de $\mathcal{S}_{\mu^*}$ alors $E = \bigcup_{i\geq 1} E_i \in \mathcal{S}_{\mu^*}$. Par récurrence, d'après la Prop. 2.1.1, on a

$$\mu^*\Big(A \cap \bigcup_{i=1}^n E_i\Big) = \sum_{i=1}^n \mu^*(A \cup E_i).$$

En écrivant $F_n = \bigcup_{i=1}^n E_i$, on a $F_n \in \mathcal{S}_{\mu^*}$ (car c'est une algèbre) et donc pour tout A, on a

$$\begin{aligned} \mu^*(A) &= \mu^*(A \cap F_n) + \mu^*(A \cap F_n^c) \\ &= \sum_{i=1}^n \mu^*(A \cap E_i) + \mu^*(A \cap F_n^c) \\ &\geq \sum_{i=1}^n \mu^*(A \cap E_i) + \mu^*(A \cap E^c) \end{aligned}$$

puisque $E^c \subset F_n^c$ et μ^* est monotone. Comme cela est vraie pour tout $n \geq 1$

$$\begin{aligned} \mu^*(A) &\geq \sum_{i\geq 1} \mu^*(A \cap E_i) + \mu^*(A \cap E^c) && (2.5) \\ &\geq \mu^*(A \cap E) + \mu^*(A \cap E^c) && (2.6) \end{aligned}$$

puisque $A \cap E = \bigcup_{i=1}^{+\infty} (A \cap E_i)$ et par σ-sous-additive $\mu^*(A \cap E) \leq \sum_{i=1}^{+\infty} \mu^*(A \cap E_i)$. L'inégalité obtenue est complétée par sous-additivité de μ^* pour avoir une égalité et prouver que $E \in \mathcal{S}_{\mu^*}$. Ainsi, $\mathcal{S}_{\mu^*}$ est stable par union dénombrable disjointe et union dénombrable quelconque. Il s'agit donc bien d'une σ-algèbre.

Pour montrer que μ^* est une mesure, noter que comme $\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c)$, les inégalités dans (2.5) sont en fait des égalités et donc pour toute suite $(E_i)_{i\geq 1}$ d'ensembles disjoints de $\mathcal{S}_{\mu^*}$ avec $E = \bigcup_{i\geq 1} E_i$, on a

$$\mu^*(A) = \sum_{i\geq 1} \mu^*(A \cap E_i) + \mu^*(A \cap E^c).$$

Avec $A = E$, le dernier terme s'annule et on a $A \cap E_i = E_i$, soit

$$\mu^*(E) = \sum_{i\geq 1} \mu^*(E_i),$$

ie. μ^* est bien σ-additive et donc une mesure, ce qui achève de prouver le theorème. □

Proposition 2.1.3 *La tribu des mesurables $\mathcal{S}_{\mu^*}$ d'une mesure extérieure μ^* est complète pour la mesure $\mu = \mu^*_{|\mathcal{S}_{\mu^*}}$ induite.*

Démonstration : Soient $A \in \mathcal{S}_{\mu^*}$ tel que $\mu(A) = \mu^*(A) = 0$ et $B \subset A$. Par croissance de μ^*, on a $\mu^*(E \cap B) \leq \mu^*(E) \leq \mu^*(A) = 0$ et donc $\mu^*(E \cap B) = 0$. Puis par sous-additivité et croissance comme $E = (E \cap B) \cup (E \cap B^c)$, on a :

$$\mu^*(E) \leq \mu^*(E \cap B) + \mu^*(E \cap B^c) = \mu^*(E \cap B) \leq \mu^*(E) \tag{2.7}$$

on a donc égalité dans (2.7) et $B \in \mathcal{S}_{\mu^*}$. On a alors $\mu(B) = \mu^*(B) \leq \mu^*(A) = \mu(A) = 0$, soit $\mu(B) = 0$. □

2.2 Construction de la mesure de Lebesgue

On considère l'ensemble $X = \mathbb{R}$. L'objectif est de construire une mesure qui étend la notion de longueur ℓ définie correctement définie seulement sur l'ensemble des intervalles bornés semi-ouverts

$$\mathcal{P} = \big\{]a, b] : -\infty < a \leq b < +\infty \big\}$$

par $\ell(]a, b]) = b - a$. Pour appliquer la stratégie de la Section 2.1, on définit la mesure extérieure

$$\ell^*(E) = \inf \left(\sum_{n\geq 1} (b_n - a_n) : E \subset \bigcup_{n\geq 1}]a_n, b_n[\right). \tag{2.8}$$

Proposition 2.2.1 *La fonction d'ensemble ℓ^* définie en (2.8) est effectivement une mesure extérieure.*

Démonstration : Il est clair que $\ell^*(\emptyset) = 0$ car $\emptyset \subset] - \varepsilon, \varepsilon[$ et $\ell^*(\emptyset) \leq 2\varepsilon$ pour tout $\varepsilon > 0$. On a ℓ^* croissante car pour $E \subset F$ alors si l'union $\bigcup_{n\geq 1}]a_n, b_n[$ couvre F, elle couvre aussi E si bien que l'inf définissant E porte sur plus de termes et de ce fait on a $\ell^*(E) \leq \ell^*(F)$. Soit maintenant $(E_n)_{n\geq 1}$ un suite de sous-ensembles de $\mathbb{R}$. S'il existe n_0 tel que $\ell^*(E_{n_0}) = +\infty$ alors $\sum_{n\geq 1} \ell^*(E_n) = +\infty$ et on a immédiatement

$$\ell^*(E) \leq \sum_{n\geq 1} \ell^*(E_n). \tag{2.9}$$

On peut donc supposer $\ell^*(E_n) < +\infty$ pour chaque $n \geq 1$. Étant donné $\varepsilon > 0$, pour chaque $n \geq 1$, il existe des $]a_{n,k}, b_{n,k}[[\in \mathcal{P}$ tels que

$$E_n \subset \bigcup_{k\geq 1}]a_{n,k}, b_{n,k}[, \quad \text{et} \quad \sum_{k\geq 1} (b_{n,k} - a_{n,k}) \leq \ell^*(E_n) + \varepsilon/2^n.$$

Mais alors
$$\bigcup_{n\geq 1} E_n \subset \bigcup_{n\geq 1}\bigcup_{k\geq 1}]a_{n,k}, b_{n,k}[$$
et
$$\ell^*\Big(\bigcup_{n\geq 1} E_n\Big) \leq \sum_{n\geq 1}\sum_{k\geq 1}(b_{n,k}-a_{n,k}) \leq \sum_{n\geq 1}\Big(\ell^*(E_n)+\varepsilon/2^n\Big) \leq \sum_{n\geq 1}\ell^*(E_n)+\varepsilon. \tag{2.10}$$
Comme $\varepsilon > 0$ est quelconque, on déduit de (2.9) la σ-additivité (2.9). □

D'après la Prop. 2.2.1, on peut appliquer la Prop. 2.1.2, et $\lambda = \ell^*_{|\mathcal{S}(\ell^*)}$ est une mesure sur $\mathcal{S}(\ell^*)$. On a alors

Proposition 2.2.2 *Les boréliens de $\mathbb{R}$ sont ℓ^*-mesurables, ie. $\mathcal{B}(\mathbb{R}) \subset \mathcal{S}_{\ell^*}$. De plus $\lambda(]a,b[) = \lambda([a,b]) = b-a$.*

Démonstration : 1) Comme $\mathcal{B}(\mathbb{R})$ est engendrée par la famille des ensembles du type $]-\infty, x]$, $x\in\mathbb{R}$, il suffit de montrer que $]-\infty,x]\in\mathcal{S}_{\ell^*}$.

Pour cela, soit $E\subset\mathbb{R}$ tel que $E\subset\bigcup_{n\geq 1}]a_n,b_n[$. Comme $]-\infty,x]\cap]a_n,b_n[\subset]a_n\wedge x,(b_n\vee x)+\varepsilon/2^n[$ et $]-\infty,x]^c\cap]a_n,b_n[\subset]a_n\vee x, b_n\vee x[$, on a
$$]-\infty,x]\cap E\subset\bigcup_{n\geq 1}\Big]a_n\wedge x,(b_n\wedge x)+\varepsilon/2^n\Big[,\quad]-\infty,x]^c\cap E\subset\bigcup_{n\geq 1}\Big]a_n\vee x, b_n\vee x\Big[$$
et
$$\begin{aligned}\ell^*\big(]-\infty,x]\cap E\big) &\leq \sum_{n\geq 1}\Big((b_n\wedge x)-(a_n\wedge x)+\varepsilon/2^n\Big)\\ \ell^*\big(]-\infty,x]^c\cap E\big) &\leq \sum_{n\geq 1}\Big((b_n\vee x)-(a_n\vee x)\Big)\end{aligned}$$
d'où
$$\ell^*\big(]-\infty,x]\cap E\big)+\ell^*\big(]-\infty,x]^c\cap E\big)\leq\sum_{n\geq 1}\big(b_n-a_n\big)+\varepsilon.$$
Comme $\varepsilon>0$ est arbitraire, on a
$$\ell^*\big(]-\infty,x]\cap E\big)+\ell^*\big(]-\infty,x]^c\cap E\big)\leq\sum_{n\geq 1}\big(b_n-a_n\big)$$
et comme ceci est vrai pour toute union $\bigcup_{n\geq 1}]a_n,b_n[\supset E$, on a
$$\ell^*\big(]-\infty,x]\cap E\big)+\ell^*\big(]-\infty,x]^c\cap E\big)\leq \ell^*(E)$$
et avec la sous-additivité de la mesure extérieure ℓ^*, on a l'égalité, ce qui prouve que $]-\infty,x]\in\mathcal{S}_{\ell^*}$ et donc $\mathcal{B}(\mathbb{R})\subset\mathcal{S}_{\ell^*}$.

2) Comme $[a,b] \subset]a-\varepsilon, b+\varepsilon[$, on a $\ell^*([a,b]) \leq (b-a)+2\varepsilon$ et donc, en faisant tendre $\varepsilon \searrow 0$, $\ell^*([a,b]) \leq (b-a)$. Il reste à voir $\ell^*([a,b]) \geq (b-a)$. Supposons alors $[a,b] \subset \bigcup_{n\geq 1}]a_n, b_n[$. Par compacité, on peut extraire une couverture par un nombre fini d'ouverts, disons $[a,b] \subset \bigcup_{i=1}^N]a_i, b_i[$. On a alors par des manipulations algébriques élémentaires :

$$(b-a) \leq \sum_{i=1}^N (b_i - a_i) \leq \sum_{n\geq 1} (b_n - a_n).$$

Comme cela est valable pour toute union $\bigcup_{n\geq 1}]a_n, b_n[\supset [a,b]$, on en déduit $b-a \leq \ell^*([a,b])$ et finalement, on a $\ell^*([a,b]) = b-a$.

Notons pour conclure que comme $\{a\} \subset]a-\varepsilon, a+\varepsilon[$, on a $\ell^*(\{a\}) \leq 2\varepsilon$ et donc, en faisant $\varepsilon \searrow 0$, $\ell^*(\{a\}) = 0$. On a donc aussi

$$\ell^*(]a,b[) = \ell^*([a,b]) - \ell^*(\{a\}) - \ell^*(\{b\}) = \ell^*([a,b]) = b-a.$$

□

On a $\mathcal{B}(\mathbb{R}) \subset \mathcal{S}_{\ell^*}$ (Prop. 2.2.2) et $\mathcal{S}_{\ell^*}$ est complète pour $\lambda = \ell^*_{|\mathcal{S}_{\ell^*}}$ (Prop. 2.1.3), en fait on a mieux :

Proposition 2.2.3 *La tribu complétée de $\mathcal{B}(\mathbb{R})$ est $\mathcal{S}_{\ell^*}$. Dans la suite, on note $\mathcal{L}(\mathbb{R})$ et elle s'appelle la tribu de Lebesgue.*

Démonstration : Notons $\widetilde{\mathcal{B}}(\mathbb{R})$ la tribu complétée de $\mathcal{B}(\mathbb{R})$.
On a vu que si $B \subset \mathbb{R}$ est tel qu'il existe $N \in \mathcal{B}(\mathbb{R})$ avec $\lambda(N) = 0$ alors $\ell^*(B) = 0$ et $B \in \mathcal{S}_{\ell^*}$, ie. $\mathcal{B}(\mathbb{R}) \subset \mathcal{S}_{\ell^*}$.

Réciproquement si $A \in \mathcal{S}_{\ell^*}$, on montre qu'il existe $B, C \in \mathcal{B}(\mathbb{R})$ tels que $C \subset A \subset B$ et $\lambda(B \setminus C) = 0$. On écrit $A = \bigcup_{n\geq 1} A_n$ avec $A_n = A \cap]-n, n[$, comme $\ell^*(A_n) \leq \ell^*(]-n,n[) = \lambda(]-n,n[) = 2n$, pour tout $p \geq 1$, il existe une couverture $A_n = \bigcup_{k\geq 1}]a_{k,p}, b_{k,p}[$ telle que $\sum_{k\geq 1} \big(b_{k,p} - a_{k,p}\big) \leq \ell^*(A_n) + 1/p$. Prenons

$$B_n = \bigcap_{p\geq 1} \bigcup_{k\geq 1}]a_{k,p}, b_{k,p}[.$$

Alors $A_n \subset B_n$, $B_n \in \mathcal{B}(\mathbb{R})$ et

$$\ell^*(B_n) \leq \ell^*\Big(\bigcup_{k\geq 1}]a_{k,p}, b_{k,p}[\Big) \leq \sum_{k\geq 1} \big(b_{k,p} - a_{k,p}\big) \leq \ell^*(E_n) + 1/p.$$

Comme ceci est valable pour tout $p \geq 1$, on a $\lambda(B_n) = \ell^*(B_n) = \ell^*(A_n)$.
En faisant la même chose pour $]-n,n[\setminus A_n$, on trouve $C_n \in \mathcal{B}(\mathbb{R})$ tel que $\lambda(C_n) = \ell^*(C_n) = \ell^*(A)$ et $C_n \subset A_n$. En particulier comme $C_n \subset A_n \subset B_n$, on a $\lambda(B_n \setminus C_n) = \lambda(B_n) - \lambda(C_n) = 0$. On a donc $A_n \in \widetilde{\mathcal{B}}(\mathbb{R})$ et $A = \bigcup_{n\geq 1} A_n \in \widetilde{\mathcal{B}}(\mathbb{R})$. □

Remarque 2.2.1 (Borélien et Lebesguien) On a

$$\mathcal{B}(\mathbb{R}) \subsetneq \mathcal{L}(\mathbb{R}) \subsetneq \mathcal{P}(\mathbb{R})$$

avec des inclusions strictes (argument de cardinalité pour la première, axiome du choix pour la deuxième).

Théorème 2.2.1 (Existence et unicité de la mesure de Lebesgue) *Il existe une unique mesure λ sur $\mathcal{B}(\mathbb{R})$ qui étend la notion delongueur des intervalles : $\lambda([a,b]) = b - a$ et λ est σ-finie. On l'appelle la mesure de Lebesgue.*

Démonstration : D'après la Prop. 2.2.1, ℓ^* définie en (2.8) est une mesure extérieure. La Prop. 2.1.2 affirme alors que $\mathcal{S}_{\ell^*}$ est une tribu sur laquelle la restriction $\lambda = \ell^*_{\mathcal{S}_{\ell^*}}$ est une mesure avec $\lambda([a,b]) = \lambda(]a,b[) = b - a$ d'après la Prop. 2.2.2. Il est clair que λ est σ-finie car $\lambda([-n,n]) = 2n < +\infty$ et $\mathbb{R} = \bigcup_{n\geq 1}[-n,n]$. Il reste à voir l'unicité.

Supposons que m, m' sont deux mesures sur $\mathcal{B}(\mathbb{R})$ telles que

$$m([a,b]) = m'([a,b]) = b - a.$$

Le but est de montrer $m(A) = m'(A)$ pour tout $A \in \mathcal{B}(\mathbb{R})$.
Sur $[-n,n]$, m et m' ont mêmes poids $2n$ et coïncident sur les l'ensemble $\mathcal{P}_n$ des intervalles $[a,b] \subset [-n,n]$. Comme $\mathcal{P}_n$ est stable par intersection, le théorème de Dynkin (Th. 1.5.2, conséquence du théorème des classes monotones, Th. 1.5.1) s'applique et donne $m = m'$ sur $\sigma(P_n) = \mathcal{B}([-n,n])$.

Si $A \in \mathcal{B}(\mathbb{R})$ alors $A \cap [-n,n] \in \mathcal{B}([-n,n])$ et donc

$$m\big(A \cap [-n,n]\big) = m'\big(A \cap [-n,n]\big). \tag{2.11}$$

Mais par croissance séquentielle des mesures

$$\lim_{n\to+\infty} m\big(A \cap [-n,n]\big) = m\Big(A \cap \big(\bigcup_{n\geq 1}[-n,n]\big)\Big) = m(A)$$

et

$$\lim_{n\to+\infty} m'\big(A \cap [-n,n]\big) = m'\Big(A \cap \big(\bigcup_{n\geq 1}[-n,n]\big)\Big) = m'(A).$$

En passant à la limite dans (2.11), on a $m(A) = m'(A)$. □

Remarque 2.2.2 (Carathéodory) La construction de la mesure de Lebesgue λ à partir d'une mesure extérieure ℓ^* est un cas particulier d'application du théorème d'extension de Carathéodory.

Théorème 2.2.2 (Carathéodory) *Soit m une fonction d'ensemble sur une "semi-algèbre" $\mathcal{P}$ vérifiant*

— *$m(\emptyset) = 0$;*
— *$m(A \cup B) = m(A) + m(B)$ pour $A, B \in \mathcal{P}$, $A \cap B = \emptyset$;*
— *(condition de Carathéodory) si $(A_n)_{n\geq 1}$ est une suitde dcroissante de $\mathcal{P}$ ($A_n \supset A_{n+1}$) avec $\bigcap_{n\geq 1} A_n = \emptyset$ alors $m(A_n) \to 0$, $n \to +\infty$.*

Alors il existe une unique mesure $\bar{m}$ sur $\sigma(\mathcal{P})$ étendant m (ie. $\bar{m}(E) = m(E)$ pour $E \in \mathcal{R}$). De plus, si m est σ-finie sur $\mathcal{P}$ alors $\bar{m}$ est l'unique mesure étendant μ sur $\sigma(\mathcal{P})$ et elle est elle-même σ-finie sur $\sigma(\mathcal{P})$.

La construction de la mesure de Lebesgue consiste à appliquer ce théorème avec ℓ définie en (2.8) et la "semi-algèbre" $\mathcal{P} = \big\{]a, b] : -\infty < a \leq b < +\infty\big\}$.

2.3 Propriétés de la mesure de Lebesgue

Proposition 2.3.1 *La mesure de Lesbesgue n'a pas d'***atome***, ie. il n'existe pas de $a \in \mathbb{R}$ tel que $\lambda(\{a\}) > 0$.*

Démonstration : Déjà vue en Prop. 2.2.2 mais se revoit facilement : en effet, pour tout $a \in \mathbb{R}$, on a $\lambda(]a - 1/n, a]) = 1/n$ et par continuité décroissante de λ, on a $\lambda(\{a\}) = \lim_{n\to+\infty} \lambda(]a - 1/n, a]) = 0$. □

Par σ-additivité, on déduit facilement :

Proposition 2.3.2 *Tout ensemble dénombrable A est de mesure de Lebesgue $\lambda(A) = 0$.*

Et on déduit de plus que les intervalles bornés sont de même mesure de Lebesgue qu'ils soient fermés, ouverts ou semi-ouverts :

$$\lambda([a, b]) = \lambda(]a, b]) = \lambda([a, b[)\lambda(]a, b[) = b - a.$$

Cela reste vrai pour les intervalles non bornés puisque leur mesure sont toute $+\infty$.

Remarque 2.3.1 (Cantor) Il existe des ensembles mesurable non dénombrables de mesure de Lebesgue nulle. L'exemple typique est le triadique de Cantor

$$K = \bigcap_{n\geq 0} K_n \tag{2.12}$$

où $K_0 = [0, 1]$ et K_n est défini par récurrence par

$$K_{n+1} = \Big(\frac{1}{3}K_n\Big) \cup \Big(\frac{1}{3}(K_n + 2)\Big).$$

On voit facilement par récurrence que K_n est la réunion de 2^n intervalles de longueur $1/3^n$. En particulier, $K_n \in \mathcal{B}(\mathbb{R})$ et donc $K \in \mathcal{B}(\mathbb{R})$ avec $\lambda(K) = 0$ car $\lambda(K) \leq \lambda(K_n) = 2^n/3^n$ pour tout $n \geq 1$.
De plus, K n'est pas dénombrable car il contient (exactement) tous les nombres de la forme $\sum_{n\geq 1} \frac{2a_n}{3^n}$ lorsque $(a_n)_{n\geq 1}$ parcourt $\{0, 1\}^{\mathbb{N}^*}$.

Proposition 2.3.3 (Invariance par translation) *La mesure de Lebesgue est invariante par translation, ie. pour tout $A \in \mathcal{B}(\mathbb{R})$ et $\alpha \in \mathbb{R}$, $\lambda(A+\alpha) = \lambda(A)$.*

Démonstration : Cela découle de l'unicité dans le Th. 2.2.1. Notons $\tau_\alpha(x) = x + \alpha$ la translation. On définit λ_{τ_α} par $\lambda_{\tau_\alpha}(A) = \lambda(\tau_\alpha^{-1}A)$. D'après la Prop. **??**, $\tau_\alpha^{-1}A \in \mathcal{B}(\mathbb{R})$ quand $A \in \mathcal{B}(\mathbb{R})$ et donc $\lambda(\tau_\alpha^{-1}A)$ est bien défini. On montre facilement que λ_{τ_α} est une mesure. Il s'agit d'un cas particulier de mesure image comme défini en Déf. 3.2.1. La mesure λ_{τ_α} vérifie

$$\lambda_{\tau_\alpha}(]a,b]) = \lambda(\tau_\alpha^{-1}]a,b]) = \lambda(]a+\alpha, b+\alpha]) = (b+\alpha) - (a+\alpha) = b - a = \lambda(A).$$

Ainsi, λ_{τ_α} et λ coïncident sur $\mathcal{P}$. L'unicité dans le Th. 2.2.1 assure alors $\lambda_{\tau_\alpha} = \lambda$.

□

Proposition 2.3.4 (Invariance par symétrie) *La mesure de Lebesgue est invariante par symétrie : $\lambda(-A) = \lambda(A)$ pour tout $A \in \mathcal{B}(\mathbb{R})$.*

Démonstration : On définit $s(x) = -x$ et on considère λ_s définie par $\lambda_s(A) = \lambda(s^{-1}A)$. À nouveau, d'après la Prop. **??**, $s^{-1}A \in \mathcal{B}(\mathbb{R})$ ssi $A \in \mathcal{B}(\mathbb{R})$ ce qui assure que $\lambda_s(A)$ est bien défini. On montre facilement que λ_s est une mesure (cas particulier de mesure image, cf. Def 3.2.1). On a

$$\lambda_s(]a,b]) = \lambda(s^{-1}]a,b]) = \lambda([-b,-a[) = \lambda(]-b,-a]) = -a - (-b) = b - a.$$

L'unicité du Th. 2.2.1 assure alors $\lambda_s = \lambda$. □

Proposition 2.3.5 *Soit T une application affine définie par $Tx = \alpha x + \beta$, $\alpha \in \mathbb{R}^*, \beta \in \mathbb{R}$. Si $E \in \mathcal{B}(\mathbb{R})$, on a $\lambda(TE) = |\alpha|\lambda(E)$.*

Démonstration : On sait déjà que TE est borélien ssi E l'est . On observe d'abord que $\lambda(TE) = |\alpha|\lambda(E)$ pour tout $E \in \mathcal{B}(\mathbb{R})$. En effet, pour cela, on définit deux mesures sur $\mathcal{B}(\mathbb{R})$ en posant $\nu_1(E) = \lambda(TE)$ et $\nu_2(E) = |\alpha|\lambda(E)$, $E \in \mathcal{B}(\mathbb{R})$. Ces mesures ν_1 et ν_2 coïncident sur les intervalles $[a,b]$ donc aussi sur $\mathcal{B}(\mathbb{R})$ de, la même façon que pour l'unicité dans le Th. 2.2.1. □

Chapitre 3

Fonctions mesurables

Les fonctions mesurables sont les fonctions de base en théorie de la mesure. Il s'agit de fonctions qui se comportent convenablement par rapport à des tribus sur les espaces sur lesquels elle est définie, elles sont présentées en Section 3.1. La notion de mesurabilité est extrêmement flexible et se conserve par la plupart des opérations usuelles sur les fonctions, cf. Section 3.2 (même le passage à la limite ! cf. Section 3.3). On présente en Section 3.4 les fonctions étagées qui vont servir de fonctions élémentaires pour la construction de l'intégrale dans le Chapitre 4.

Rappel (Images directe et réciproque) Si $f : X \to Y$ est une application quelconque

— pour toute partie A de X, l'image directe de A par f est la partie de Y donnée par

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

— pour toute partie B de Y, l'image réciproque de B par f est la partie de X donnée par

$$f^{-1}(B) = \big\{x \in X : f(x) \in B\big\}.$$

Si $(A_i)_{i\in I}$ et $(B_j)_{j\in J}$ sont des parties de X et Y respectivement, on rappelle que le comportement de f vis à vis de $\cup, \cap$ et de ^{c}

$$f\Big(\bigcup_{i\in I} A_i\Big) = \bigcup_{i\in I} f(A_i) \quad f\Big(\bigcap_{i\in I} A_i\Big) \subset \bigcap_{i\in I} f(A_i)$$

tandis que le comportement de f^{-1} est meilleur :

$$f^{-1}\Big(\bigcup_{j\in J} B_j\Big) = \bigcup_{j\in J} f^{-1}(B_j) \quad f^{-1}\Big(\bigcap_{j\in J} B_j\Big) = \bigcap_{j\in J} f^{-1}(B_j) \quad f^{-1}(B^c) = (f^{-1}(B))^c.$$

Dans tout le chapitre, on considère un espace mesuré $(X, \mathcal{A}, \mu)$.

3.1 Mesurabilité de fonction

Définition 3.1.1 (Fonction mesurable) *Soient $(X, \mathcal{A})$, $(Y, \mathcal{B})$ deux espaces munis de tribus. Une fonction $f : X \to Y$ est dite $(\mathcal{A}, \mathcal{B})$-mesurable si et seulement si $\forall B \in \mathcal{B}$, $f^{-1}(B) \in \mathcal{A}$.*

Remarque 3.1.1
— Cette définition est à comparer avec la définition de la continuité : l'image réciproque d'un ouvert doit être ouverte.
— Quand Y est un espace topologique et que rien n'est précisé, on prendra la tribu borélienne $\mathcal{B}(Y)$ de Y.
— Dans le contexte probabiliste, les fonctions mesurables s'appellent les variables aléatoires; dans ce cas, on note traditionnellement $(X, \mathcal{A}, \mu) = (\Omega, \mathcal{F}, \mathbb{P})$ et une fonction $X : (\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ mesurable s'appelle une variable aléatoire.

Définition 3.1.2 (Indicatrice) *Une fonction indicatrice $\mathbf{1}_A$ d'un ensemble A est la fonction définie par*

$$\mathbf{1}_A(x) = \begin{cases} 1 & si \quad x \in A, \\ 0 & si \quad x \notin A. \end{cases}$$

Cette fonction ne prend donc que deux valeurs 1 ou 0 selon qu'elle est évaluée sur A ou non.

On vérifie facilement les propriétés suivantes :

Proposition 3.1.1 *On a $A = B$ ssi $\mathbf{1}_A = \mathbf{1}_B$; On a $\mathbf{1}_A \mathbf{1}_B = \mathbf{1}_{A \cap B}$, $1 - \mathbf{1}_A = \mathbf{1}_{A^c}$; si A et B sont disjoints, on a $\mathbf{1}_A + \mathbf{1}_B = \mathbf{1}_{A \cup B}$; si $A \subset B$, on a $\mathbf{1}_A \leq \mathbf{1}_B$.*

Proposition 3.1.2 *La fonction indicatrice $\mathbf{1}_A$ de A est mesurable (en tant que fonction) de $(X, \mathcal{A})$ dans $\mathbb{R}$ ssi A est mesurable (en tant qu'ensemble).*

Démonstration : On suppose $A \in \mathcal{A}$. Soit $B \in \mathcal{B}(\mathbb{R})$ alors $(\mathbf{1}_A)^{-1}(B) = \{x \in X : \mathbf{1}_A(x) \in B\}$ et
— si 0 et $1 \in B$ alors $(\mathbf{1}_A)^{-1}(B) = X \in \mathcal{A}$,
— si $1 \in B$ mais $0 \notin B$ alors $(\mathbf{1}_A)^{-1}(B) = A \in \mathcal{A}$,
— si $0 \in B$ mais $1 \notin B$ alors $(\mathbf{1}_A)^{-1}(B) = A^c \in \mathcal{A}$,
— si 0 et $1 \notin B$ alors $(\mathbf{1}_A)^{-1}(B) = \emptyset \in \mathcal{A}$,
Finalement on a pour tout $B \in \mathcal{A}$, $(\mathbf{1}_A)^{-1}(B) \in \mathcal{A}$: la fonction $\mathbf{1}_A$ est mesurable. Réciproquement si $\mathbf{1}_A$ est mesurable alors $A = (\mathbf{1}_A)^{-1}(\{1\}) \in \mathcal{A}$. □

Proposition 3.1.3 *Si $\mathcal{M}$ engendre la tribu $\mathcal{B}$ de Y, f est mesurable ssi $f^{-1}(B) \in \mathcal{A}$ pour tout $B \in \mathcal{M}$.*

Démonstration : En effet notons $\mathcal{C}$ l'ensemble des $B \in \mathcal{B}$ tels que $f^{-1}(B) \in \mathcal{A}$. Alors $\mathcal{C}$ est une tribu de Y car

— $f^{-1}(Y) = X \in \mathcal{A}$ donc $Y \in \mathcal{C}$,
— si $B \in \mathcal{C}$, $f^{-1}(B^c) = (f^{-1}(B))^c \in \mathcal{A}$ car $f^{-1}(B) \in \mathcal{A}$ et $\mathcal{A}$ est stable par complémentaire.
— Si A_n, $n \geq 1$, sont dans $\mathcal{C}$ alors $f^{-1}(A_n) \in \mathcal{A}$ d'où $f^{-1}\Big(\bigcup_{n\geq 1} A_n\Big) = \bigcup_{n\geq 1} f^{-1}(A_n) \in \mathcal{A}$. Et donc $\bigcup_{n\geq 1} A_n \in \mathcal{C}$, qui est stable par réunion.

Finalement, $\mathcal{C}$ est une tribu puis par hypothèse $\mathcal{C}$ contient $\mathcal{M}$ qui engendre $\mathcal{B}$. Donc $\mathcal{B} \subset \mathcal{C}$ et on a en particulier pour tout $B \in \mathcal{B}$, $f^{-1}(B) \in \mathcal{A}$, c'est à dire f est mesurable. □

Ainsi, quand $Y = \mathbb{R}$ ou $\mathbb{C}$ (ou un espace topologique) muni de la tribu borélienne, f est mesurable
— ssi $\forall B \in \mathcal{B}(Y)$ alors $f^{-1}(B) \in \mathcal{A}$.
— ssi $\forall O$ ouvert, $f^{-1}(O) \in \mathcal{A}$
— ssi, pour tout $a \in \mathbb{R}$, $f^{-1}(]a, +\infty[) \in \mathcal{A}$ dans le cas $Y = \mathbb{R}$.

Corollaire 3.1.1 *Une fonction continue de $(X, \mathcal{T})$ dans $(Y, \mathcal{T}')$ est mesurable pour les tribus boréliennes $\mathcal{B}(X)$ et $\mathcal{B}(Y)$ associées à X et à Y.*

On connaît donc maintenant beaucoup de fonctions mesurables (pour les tribus boréliennes) : toutes les fonctions continues.

Démonstration : En effet, comme les ouverts de Y engendrent $\mathcal{B}(Y)$, il suffit de voir que l'image réciproque $f^{-1}(O)$ d'un ouvert O de Y est dans $\mathcal{B}(X)$. Or par continuité de f, $f^{-1}(O)$ est ouvert dans X donc borélien. □

3.2 Propriétés des fonctions mesurables

La mesurabilité des fonctions est une propriété stable par toutes les opérations usuelles sur les fonctions :

Proposition 3.2.1 (Composition 1) *Si $f : (X, \mathcal{A}) \to (Y, \mathcal{B})$ et $g : (Y, \mathcal{B}) \to (Z, \mathcal{C})$ sont mesurables alors $g \circ f : (X, \mathcal{A}) \to (Z, \mathcal{C})$ est mesurable.*

Démonstration : Soit $C \in \mathcal{C}$, $(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C))$. Or par mesurabilité de g, $g^{-1}(C) \in \mathcal{B}$, puis par celle de f, $f^{-1}(g^{-1}(C)) \in \mathcal{A}$. □

En particulier, on a :

Proposition 3.2.2 (Composition 2) *Si $f : (X, \mathcal{A}) \to Y$ espace topologique est mesurable et $g : Y \to Z$, espace topologique est continue alors $g \circ f : (X, \mathcal{A}) \to Z$ est mesurable.*

Démonstration : Soit $C \in \mathcal{T}_Z$, $(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C))$. Or par continuité de g, $g^{-1}(C) \in \mathcal{T}_Y$, puis par mesurabilité de f, $f^{-1}(g^{-1}(C)) \in \mathcal{A}$. □

Par exemple

— si $f : (X, \mathcal{A}) \to \mathbb{R}$ est mesurable alors $|f|$, f^+, f^- le sont.
En effet, on peut appliquer la Prop. 3.2.2 avec les applications continues

$$\left\{\begin{array}{ccc} \mathbb{R} & \to & \mathbb{R}_+ \\ x & \mapsto & |x| \end{array}\right., \quad \left\{\begin{array}{ccc} \mathbb{R} & \to & \mathbb{R}_+ \\ x & \mapsto & \max(x,0) \end{array}\right., \quad \left\{\begin{array}{ccc} \mathbb{R} & \to & \mathbb{R}_+ \\ z & \mapsto & \min(-x,0) \end{array}\right..$$

— si $f : (X, \mathcal{A}) \to \mathbb{C}$ est mesurable alors $|f|$, $\mathrm{Im}(f)$, $\mathrm{Re}(f)$ le sont.
En effet, on peut appliquer la Prop. 3.2.2 avec les applications continues

$$\left\{\begin{array}{ccc} \mathbb{C} & \to & \mathbb{R}_+ \\ z & \mapsto & |z| \end{array}\right., \quad \left\{\begin{array}{ccc} \mathbb{C} & \to & \mathbb{R}_+ \\ z & \mapsto & Re(z) \end{array}\right., \quad \left\{\begin{array}{ccc} \mathbb{C} & \to & \mathbb{R}_+ \\ z & \mapsto & Im(z) \end{array}\right..$$

Proposition 3.2.3 (Couple) *Soit $(X, \mathcal{A})$ un espace mesurable et $f : (X, \mathcal{A}) \to \mathbb{R}$, $g : (X, \mathcal{A}) \to \mathbb{R}$ sont mesurables, alors $h = (f, g) : (X, \mathcal{A}) \to \mathbb{R}^2$ est mesurable.*

Démonstration : On munit $\mathbb{R}^2$ de la topologie produit pour laquelle les ouverts sont des produits d'ouverts $U \times V$. Soit donc $U \times V$ un ouvert produit de $\mathbb{R}^2$. D'après la Prop. 3.1.3, (f, g) est mesurable ssi

$$(f, g)^{-1}(U \times V) \in \mathcal{A}.$$

Or

$$\begin{aligned} (f,g)^{-1}\big(U \times V\big) &= \big\{x \in X : (f,g)(x) \in U \times V\big\} \\ &= \big\{x \in X : (f(x), g(x)) \in U \times V\big\} \\ &= \big\{x \in X : f(x) \in U, g(x) \in V\big\} \\ &= \big\{x \in X : f(x) \in U\big\} \cap \big\{x : g(x) \in V\big\} \\ &= f^{-1}(U) \cap g^{-1}(V) \in \mathcal{A} \end{aligned}$$

car $f^{-1}(U) \in \mathcal{A}$ et $g^{-1}(V) \in \mathcal{A}$ par mesurabilité de f, g. □

On en déduit si f et g sont mesurables, a est scalaire
— af, $f + g$, $f - g$, $f \times g$, f/g (si $g(x) \neq 0, \forall x$), $\max(f, g)$, $\min(f, g)$ sont mesurables.
— Toute combinaison linéaire de fonctions mesurables est mesurable.
— $f : (X, \mathcal{A}) \to \mathbb{C}$ est mesurable ssi $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ le sont.
— $f : (X, \mathcal{A}) \to \mathbb{R}$, est mesurable ssi $f^+ = \max(f, 0)$ et $f^- = \min(f, 0)$ le sont.

On énonce des résultats analogues sur $\mathbb{C}$ et en fait sur tout espace topologique Y qui a une base dénombrable d'ouverts.

Démonstration : Soit O un ouvert de $\mathbb{R}^2$, comme $\mathbb{R}^2$ a une base dénombrable d'ouverts, il s'écrit comme réunion dénombrables de pavés ouverts $O = \bigcup_{i=1}^n \big(]a_i, b_i[\times]c_i, d_i[\big)$. On a alors

$$h^{-1}(O) = h^{-1}\Big(\bigcup_{i=1}^n \big(]a_i, b_i[\times]c_i, d_i[\big)\Big) = \bigcup_{i=1}^n h^{-1}\big(]a_i, b_i[\times]c_i, d_i[\big).$$

Or $h^{-1}\big(]a_i,b_i[\times]c_i,d_i[\big) = \big\{x \in X : (f(x),g(x)) \in]a_i,b_i[\times]c_i,d_i[\big\} = f^{-1}(]a_i,b_i[) \cap g^{-1}(]c_i,d_i[)$.
Donc $h^{-1}(O) = \bigcup_{i=1}^n f^{-1}(]a_i,b_i[) \cap g^{-1}(]c_i,d_i[) \in \mathcal{A}$. □

Définition 3.2.1 (Mesure image) *Soit $f : (X,\mathcal{A},\mu) \to (Y,\mathcal{B})$ une fonction mesurable. On définit sur $(Y,\mathcal{B})$ la mesure image de f notée μf^{-1} ou μ_f :*

$$\mu_f(B) = \mu(f^{-1}(B)).$$

On montre facilement que μ_f est bien une mesure : $\mu_f(\emptyset) = \mu(f^{-1}(\emptyset)) = \mu(\emptyset) = 0$ et pour des $B_i \in \mathcal{B}$, $i \geq 1$, disjoints

$$\begin{aligned} \mu_f\Big(\bigcup_{i\geq 1} B_i\Big) &= \mu\Big(f^{-1}\Big(\bigcup_{i\geq 1} B_i\Big)\Big) = \mu\Big(\bigcup_{i\geq 1} f^{-1}(B_i)\Big) \\ &= \sum_{i\geq 1} \mu\big(f^{-1}(B_i)\big) = \sum_{i\geq 1} \mu_f(B_i). \end{aligned}$$

car les B_i étant disjoints, les $f^{-1}(B_i)$ le sont aussi. **Exemple :**(Loi de probabilité) Dans le contexte probabiliste, la mesure image d'une variable aléatoire s'appelle sa loi : si $(X,\mathcal{A},\mu) = (\Omega,\mathcal{F},\mathbb{P})$ et $X : (\Omega,\mathcal{F},\mathbb{P}) \to \mathbb{R}$ est une variable aléatoire alors $\mathbb{P}_X$ est la loi de la variable aléatoire X :

$$\mathbb{P}_X(B) = \mathbb{P}(X \in B), \quad B \in \mathcal{B}.$$

3.3 Limite de fonctions mesurables

Topologie métrique sur $\overline{\mathbb{R}}$

On considère les ensembles $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty,-\infty\}$ et $[0,+\infty] = [0,+\infty[\cup\{+\infty\}$. On définit une distance (et donc une topologie métrique associée à cette distance) sur ces ensembles : Soit ϕ une bijection de $\overline{\mathbb{R}}$ sur un compact (par exemple $\phi = \arctan$, bijection de $\overline{\mathbb{R}}$ sur $[-\pi/2,\pi/2]$). On pose

$$d(x,y) = |\phi(x) - \phi(y)|$$

avec $\arctan(\pm\infty) = \pm\pi/2$. Alors $(\overline{\mathbb{R}},d)$ et $([0,+\infty],d)$ sont métriques.
Les boréliens associés à ces ensembles (avec la topologie définie par la métrique indiquée) sont engendrés par $\big\{]a,+\infty], a \in \mathbb{R}\big\}$.

On adopte les règles de calculs suivantes dans $[0,+\infty]$:

Proposition 3.3.1 (Règles de calcul) *Soient $a,b \in \mathbb{R}$, on a*

$$\begin{aligned} a \times b = ab &\quad si \quad a,b \neq +\infty, \\ a \times (+\infty) = +\infty &\quad si \quad a > 0, \\ 0 \times (+\infty) = 0 &\quad si \quad a = 0. \end{aligned}$$

Remarque 3.3.1 — Ce produit n'est pas continu dans $[0, +\infty]$ en effet avec $a_n = n$ et $b_n = 1/n$ on a $a_n \to +\infty$, $b_n \to 0$ puis $a_n b_n = 1 \not\to ab = +\infty \times 0 = 0$.

— La convention $0 \times (+\infty) = 0$ est naturelle malgré tout quand on pense à $0 \times (+\infty)$ comme le calcul de la surface de $\mathbb{R}$ de longueur $+\infty$ et de largeur 0.

Liminf et limsup de fonctions

Pour une suite réelle $u = (u_n)_{n\geq 1}$, on définit ses limites supérieure et inférieure

$$\begin{aligned}\liminf_{n\to+\infty} u_n &= \sup_{n\geq 1} \inf_{k\geq n} u_k,\\ \limsup_{n\to+\infty} u_n &= \inf_{n\geq 1} \sup_{k\geq n} u_k.\end{aligned}$$

Ce sont les plus petites et plus grandes valeurs d'adhérence de la suite u, elles existent toujours. On a toujours

$$\liminf_{n\to+\infty} u_n \leq \limsup_{n\to+\infty} u_n$$

et il y a égalité ssi la suite u converge ; de plus si tel est le cas

$$\lim_{n\to+\infty} u_n = \liminf_{n\to+\infty} u_n = \limsup_{n\to+\infty} u_n.$$

En plus, en changeant le signe, les limites inférieure et supérieure s'échangent :

$$\begin{aligned}\liminf_{n\to+\infty}(-u_n) &= -\limsup_{n\to+\infty} u_n,\\ \limsup_{n\to+\infty}(-u_n) &= -\liminf_{n\to+\infty} u_n.\end{aligned}$$

Ces notions rappellent celle du Chapitre 1 pour les limites inférieure et supérieure d'ensembles.

Considérons par exemple

— la suite $(u_n)_{\geq 1}$ donnée par $u_{2n} = 1$ et $u_{2n+1} = 0$ alors sa limite supérieure est 1, sa limite inférieure est 0.

— la suite $(v_n)_{n\geq 1}$ donnée par $v_n = \cos(n)$, on montre que sa limite supérieure est 1, sa limite inférieure est -1.

— la suite $(w_n)_{n\geq 1}$ donnée par $w_{2n} = n$ et $w_{2n+1} = \frac{2n}{n+1}$, on montre que sa limite supérieure est $+\infty$, sa limite inférieure est 2.

Pour une suite de fonctions $(f_n)_{n\geq 1}$, on définit des fonctions limites inférieure et supérieure de la façon suivante :

$$\left(\liminf_{n\to+\infty} f_n\right)(x) = \liminf_{n\to+\infty} f_n(x), \qquad \left(\limsup_{n\to+\infty} f_n\right)(x) = \limsup_{n\to+\infty} f_n(x).$$

Proposition 3.3.2 *Soient $f_n : (X, \mathcal{A}) \to \overline{\mathbb{R}}$, $n \geq 1$, une suite de fonctions mesurables alors $\sup_{n\geq 1} f_n$ et $\inf_{n\geq 1} f_n$ sont mesurables.*

Démonstration : On le montre pour le sup, le raisonnement s'adapterait facilement à l'inf.

$$\begin{aligned}\big(\sup_{n\geq 1} f_n\big)^{-1}(]a,+\infty]) &= \big\{x : \big(\sup_{n\geq 1} f_n\big)(x) \in]a,+\infty]\big\} = \big\{x\in X : \sup_{n\geq 1} f_n(x) \in]a,+\infty]\big\}\\ &= \big\{x\in X : \sup_{n\geq 1} f_n(x) > a\big\} = \big\{x \in X : \exists n\geq 1, f_n(x) > a\big\}\\ &= \bigcup_{n\geq 1}\big\{x\in X : f_n(x) > a\big\} = \bigcup_{n\geq 1} f_n^{-1}\big(]a,+\infty[\big) \in \mathcal{A}\end{aligned}$$

car pour chaque $n \geq 1$, $f_n^{-1}(]a,+\infty[) \in \mathcal{A}$ (f_n mesurable) qui est stable par réunion. □

Proposition 3.3.3 *Soit* $(X,\mathcal{A})$ *un espace mesurable,* $f_n : (X,\mathcal{A}) \to \overline{\mathbb{R}}$ *ou* $[0,+\infty]$, $n\geq 1$, *des fonctions mesurables. Alors* $\limsup_n f_n$ *et* $\liminf_n f_n : (X,\mathcal{A}) \to \overline{\mathbb{R}}$ *ou* $[0,+\infty]$ *sont mesurables.*

Démonstration : Pour $\limsup_n f_n$, on note $g_k = \sup_{n\geq k} f_n(x)$. D'après le résultat pour le sup, les fonctions g_k, $k\geq 1$, sont toutes mesurables. Puis $\limsup_n f_n = \inf_{k\geq 1} g_k$ est mesurable car inf de fonctions mesurables.
Pour $\liminf_n f_n$ enfin, l'argument est analogue ou on utilise

$$\liminf_n f_n = -\limsup_n (-f_n).$$

□

On en déduit que la mesurabilité se conserve même en passant à la limite :

Théorème 3.3.1 *Soit* $(f_n)_{n\geq 1}$ *une suite de fonctions mesurables sur* $(X,\mathcal{A})$ *dans un espace métrique* (E,d). *Si cette suite de fonctions converge simplement vers* f *(c'est à dire pour tout* $x\in X$, $f(x) = \lim_{n\to+\infty} f_n(x)$*) alors* f *est une fonction mesurable à valeurs dans* E.

C'est un résultat très agréable si on le compare avec le résultat analogue pour la continuité où on a besoin de la convergence uniforme pour que la continuité se conserve à la limite. Une fonction obtenue comme limite (simple) de mesurables est donc mesurable !

Démonstration : D'après la Proposition 3.1.3, il suffit de montrer que si O est un ouvert de E alors $f^{-1}(O) \in \mathcal{A}$. Pour cela, on pose

$$O_r = \{x\in O : d(x, E\setminus O) > 1/r\}, \quad r\geq 1.$$

Comme la distance d est continue et $E\setminus O$ est fermé, l'application $g(x) = d(x, E\setminus O)$ est continue et donc $O_r = g^{-1}(]1/r,+\infty[)$ est ouvert (continuité de g). L'ensemble O_r est donc borélien de E et

$$\bigcup_{r\geq 1} O_r = \{x\in O \ : \ d(x,E\setminus O) > 0\} = O$$

(noter que $d(x, \backslash O) = 0$ ssi $x \in E \setminus 0$ ie. $x \notin 0$). On a

$$\begin{aligned} f^{-1}(O) &= \big\{x \in X : f(x) \in O\big\} = \big\{x \in X : \lim_{n\to+\infty} f_n(x) \in O\big\} \\ &= \Big\{x \in X : \lim_{n\to+\infty} f_n(x) \in \bigcup_{r\geq 1} O_r\Big\} \\ &= \big\{x \in X : \exists r \geq 1, \exists m \geq 1, \forall n \geq m \ : \ f_n(x) \in O_r\big\} \\ &= \bigcup_{r,m\in\mathbb{N}^*} \bigcap_{n\geq m} f_n^{-1}(O_r) \end{aligned}$$

est un ensemble mesurable de $\mathcal{A}$. □

3.4 Fonctions étagées (simples)

On rappelle la notion de fonction indicatrice en Déf. 3.1.2. En particulier, l'indicatrice $\mathbf{1}_A$ est mesurable ssi A l'est (Prop. 3.1.2).

Définition 3.4.1 (Fonction simple ou étagée) *Une fonction $f : (X, \mathcal{A}) \to \mathbb{R}^+$ est dite étagée positive si c'est une combinaison linéaire finie à coefficients positifs de fonctions indicatrices $\mathbf{1}_{A_i}$ pour des ensembles mesurables A_i deux à deux disjoints (pour simplifier) :*

$$f(x) = \sum_{i=1}^{n} \alpha_i \mathbf{1}_{A_i}, \quad \alpha_i \in \mathbb{R}^+, \quad A_i \in \mathcal{A}, \quad n \geq 1.$$

— Cela ressemble à une fonction en escalier mais c'est plus général car pour une fonction en escalier les ensembles A_i doivent être des intervalles de $\mathbb{R}$, alors qu'ici, il s'agit d'ensembles mesurables quelconques sur X quelconque. En fait quand $X = \mathbb{R}$, les fonctions en escalier sont des cas particuliers de fonctions étagées (avec des A_i égaux à des intervalles disjoints, plutôt qu'à des ensembles mesurables généraux).

— Une fonction étagée est mesurable car combinaison linéaire de fonctions indicatrices qui le sont clairement.

— Une fonction étagée prend un nombre fini de valeur : elle vaut α_i sur l'ensemble A_i.

— Si les α_i, $1 \leq i \leq n$, sont les valeurs possibles pour une fonction étagée f, alors avec $A_i = f^{-1}(\alpha_i)$, qui est mesurable par mesurabilité de f, on peut écrire

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

— Une combinaison linéaire de fonctions étagées est encore une fonction étagée.

Exemples :

• Avec $E(x)$ désignant la partie entière de x, soit

$$f(x) = \begin{cases} 0 & x < 0 \\ E(x) & x \in [0,5] \\ 3 & \text{sinon.} \end{cases}$$

Montrer que f est étagée en l'écrivant comme combinaison linéaire de fonctions indicatrices.

• La fonction $x \mapsto E(x)$ est-elle étagée ?

• On considère $\mathbb{N}$ muni de la tribu $\mathcal{P}(\mathbb{N})$ et les suites suivantes

— $u = (u_n)_{n\geq 1}$ avec $u_n = 0$ pour $n \geq 6$,

— $v = (v_n)_{n\geq 1}$ avec $v_n = 3$ pour $n > 4$,

— $w = (w_n)_{n\geq 1}$.

Montrer que u, v sont étagées mais pas w.

Pour l'intégrale de Riemann, on définit une fonction Riemann intégrable quand elle est (uniformément) approchable par une fonction en escalier et l'intégrale se définit comme la limite de celles en escalier, cf. [JCB-Riemann]. Les fonctions étagées vont jouer un rôle analogue en théorie de la mesure pour l'intégrale de Lebesgue, mais les choses se passent mieux : toute fonction mesurable est limite de fonctions étagées.

Proposition 3.4.1 (Approximation) *Toute fonction mesurable f à valeurs dans $\mathbb{R}^d$ est limite simple de fonctions étagées. Si de plus f est réelle positive, la limite peut être choisie croissante.*

Démonstration : Soit d'abord f fonction mesurable réelle positive. On définit pour tout n et $k \geq 1$

$$\begin{aligned} A_{n,k} &= \left\{ x \in X : \frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n} \right\}, \quad 1 \leq k \leq n2^n \\ B_n &= \{ x \in X : f(x) \geq n \}. \end{aligned}$$

Les ensembles $A_{n,k} = f^{-1}\big([(k-1)/2^n, k/2^n[\big)$ et $B_n = f^{-1}([n, +\infty[$ sont des images réciproques par f, mesurable, d'intervalles. Ils sont donc dans $\mathcal{A}$. La suite

$$f_n(x) = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \mathbf{1}_{A_{n,k}}(x) + n\mathbf{1}_{B_n}(x)$$

converge en croissant vers $f(x)$. En effet, comme

$$\left[\frac{k-1}{2^n}, \frac{k}{2^n}\right[\subset \left[\frac{2k-2}{2^{n+1}}, \frac{2k-1}{2^{n+1}}\right[\cup \left[\frac{2k-1}{2^{n+1}}, \frac{2k}{2^{n+1}}\right[$$

on a $A_{n,k} = A_{n+1,2k-1} \cup A_{n+1,2k}$. Ainsi pour $f_n(x) = \frac{k-1}{2^n}$ on a soit $f_{n+1}(x) = \frac{2k-2}{2^{n+1}} = f_n(x)$ soit $f_{n+1}(x) = \frac{2k-1}{2^{n+1}} \geq f_n(x)$ donc en tout cas $f_{n+1}(x) \geq f_n(x)$.
Puis si $f(x) \geq n$ alors soit $f(x) \geq n+1$ et alors $f_{n+1}(x) = n+1 \geq f_n(x) = n$, soit $f(x)$ est dans

$$[n, n+1[= \bigcup_{k=2^{n+1}n+1}^{2^{n+1}(n+1)} \left[\frac{k-1}{2^{n+1}}, \frac{k}{2^{n+1}}\right[.$$

Pour $k \in [2^{n+1}n+1, 2^{n+1}(n+1)]$, on a alors $f_{n+1}(x) = \frac{k-1}{2^{n+1}} \geq \frac{2^{n+1}n}{2^{n+1}} = n = f_n(x)$ et donc $f_{n+1}(x) \geq f_n(x)$.

Pour la convergence : si $f(x) = +\infty$ alors on a clairement $f_n(x) = n \to +\infty$ soit $\lim_{n\to+\infty} f_n(x) = f(x)$. Tandis que si $f(x) < +\infty$ alors pour $n \geq [f(x)] + 1$, on a $f(x) \leq n$ et donc $f_n(x) = \frac{k-1}{2^n}$ pour $f(x) \in \left[\frac{k-1}{2^n}, \frac{k}{2^n}\right[$, ie. $|f(x) - f_n(x)| \leq 2^{-n}$. On a donc $\lim_{n\to+\infty} f_n(x) = f(x)$.

Si f est réelle mais de signe quelconque, on écrit $f = f^+ - f^-$ avec $f^+(x) = \max(f(x), 0)$, $f^-(x) = \max(-f(x), 0)$ et on approxime f^+ et f^- comme précédemment.

Si $f = (f_1, \dots, f_d)$ est à valeurs dans $\mathbb{R}^d$, on applique la stratégie précédente à chaque fonction coordonnée f_i, $1 \leq i \leq d$. □

Ce résultat va permettre de définir l'intégrale d'une fonction mesurable à partir de celle des fonctions étagées car toute fonction mesurable est limite de fonctions étagées croissantes.

Chapitre 4

Intégrales des fonctions mesurables positives

Dans ce chapitre, on commence à construire l'intégrale de fonction par rapport à une mesure abstraite comme définie en Déf. 1.3.1. C'est cette nouvelle intégrale qu'on appelle l'intégrale de Lebesgue. Pour cela, on commence avec ce chapitre par les intégrales de fonctions mesurables positives. La construction de cette nouvelle intégrale suit un cheminement classique, analogue par exemple à celle de l'intégrale de Riemann où on commence par la définir pour les fonctions en escalier puis on généralise ensuite à une classe de fonctions plus grande (les fonctions dites Riemann-intégrables). Pour l'intégration par rapport à une mesure abstraite μ (intégrale de Lebesgue), on commence aussi par définir l'intégrale pour des fonctions assez simples, les fonctions étagées (ou simples) avant de généraliser aux fonctions mesurables positives en Section 4.1 et de voir les premières propriétés de cette intégrale en Section 4.2.

On donne ensuite des théorèmes de convergence pour les intégrales de fonctions postives : le théorème de convergence monotone (Th. 4.3.1) et ses conséquences en Section 4.3 puis le lemme de Fatou (Th. 4.4.1 en Section 4.4).

Dans toute la suite, on considère un espace mesuré $(X, \mathcal{A}, \mu)$.

4.1 Intégrale des fonctions positives

Définition 4.1.1 *Soit $f = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i}$ une fonction étagée positive ($\alpha_i \geq 0$) avec les A_i deux à deux disjoints, on définit l'intégrale de f par rapport à la mesure μ par*

$$\int f \, d\mu = \int_X f \, d\mu = \sum_{i=1}^n \alpha_i \mu(A_i).$$

Comme $f\mathbf{1}_E = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i \cap E}$ est encore étagée, on définit l'intégrale sur $E \in \mathcal{A}$ par

$$\int_E f d\mu = \int (f\mathbf{1}_E) \, d\mu \quad \Big(= \sum_{i=1}^n \alpha_i \mu(A_i \cap E) \Big).$$

Remarque 4.1.1 — L'intégrale de f étagée ne dépend pas de son écriture sous forme de combinaison linéaire de fonctions étagées, une autre écriture de f donne la même valeur à l'intégrale : cette définition a donc bien un sens.

— Par exemple

$$\int (2\mathbf{1}_A + 3\mathbf{1}_B + \frac{1}{2}\mathbf{1}_C)\, d\mu = 2\mu(A) + 3\mu(B) + \frac{1}{2}\mu(C).$$

— Noter en particulier que pour une fonction indicatrice :

$$\int \mathbf{1}_A \, d\mu = \mu(A), \quad \int_E \mathbf{1}_A \, d\mu = \mu(A \cap E).$$

Ce résultat est à la fois élémentaire et important : il fait le lien entre une mesure $\mu(A)$ et une intégrale $\int \mathbf{1}_A d\mu$, toute mesure d'un ensemble peut se voir comme une intégrale.

Définition 4.1.2 (Intégrale) *Soit $f : (X, \mathcal{A}, \mu) \to [0, +\infty]$ mesurable. On définit son intégrale comme le sup des intégrales de fonctions étagées majorées par f :*

$$\int f \, d\mu = \int_X f \, d\mu = \sup\Big(\int s \, d\mu : s \textit{ étagée } \leq f\Big).$$

De même pour $E \in \mathcal{A}$ de X, on définit son intégrale sur E par

$$\int_E f \, d\mu = \int (f\mathbf{1}_E) \, d\mu = \sup\Big(\int_E s \, d\mu : s \textit{ étagée } \leq f \textit{ sur } E\Big).$$

Définition 4.1.3 (Intégrabilité) *Une fonction mesurable positive est dite μ-intégrable si $\int f \, d\mu < +\infty$.*

Exemple :

— On considère $(X, \mathcal{A})$ muni de la mesure δ_a. Dans ce cas,

$$\int f \, d\delta_a = f(a).$$

En effet, si $f = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i}$ (avec $(A_i)_{1 \leq i \leq n}$ une partition de X) est étagée alors il existe un unique indice $1 \leq i_0 \leq n$ tel que $a \in A_{i_0}$ et

$$\int f \, d\delta_a = \sum_{i=1}^n \alpha_i \delta_a(A_i) = \alpha_{i_0}\delta_a(A_{i_0}) = \alpha_{i_0} = f(a)$$

car $f(a) = \alpha_{i_0}\mathbf{1}_{A_{i_0}}(a) = \alpha_{i_0}$. De façon générale,

$$\int f d\delta_a = \sup\left(\int s \, d\delta_a : s \text{ étagée } \leq f\right) = \sup\big(s(a) : s \text{ étagée } \leq f\big) = f(a).$$

— On considère $(X, \mathcal{A}, \mu) = (\mathbb{N}, \mathcal{P}(\mathbb{N}), \eta)$ où η est la mesure de dénombrement, alors si on se donne une suite $u = (u_n)_{\geq 1}$, on peut l'écrire $u = \sum_{n=0}^{+\infty} u_n \mathbf{1}_{\{n\}}$. En effet, $u(k) = \sum_{n=0}^{+\infty} u_n \mathbf{1}_{\{n\}}(k) = u_k$ car le seul terme non nul dans la somme est $\mathbf{1}_{\{k\}}(k) = 1$. On a alors

$$\int_{\mathbb{N}} u \, d\eta = \sum_{n=0}^{+\infty} \int_{\{n\}} u \, d\eta = \sum_{n=0}^{+\infty} u(n)\eta(\{n\}) = \sum_{n=0}^{+\infty} u_n.$$

Autrement dit une série se voit comme une intégrale par rapport à une mesure discrète : la mesure de dénombrement.
Une série positive converge ssi la suite associée est intégrable pour la mesure de dénombrement.

4.2 Propriétés de l'intégrale

On considère des fonctions mesurables positives f et g et des ensembles E, F mesurables :

• **(Croissance 1)** Si $f \leq g$ sur X, alors $\int f \, d\mu \leq \int g \, d\mu$.
En effet si $f \leq g$ toute fonction étagée s majorée par f l'est aussi par g si bien que le sup qui définit $\int g \, d\mu$ est pris sur un ensemble plus grand que celui définissant $\int f \, d\mu$, il est donc plus grand :

$$\int f \, d\mu = \sup\Big(\int s \, d\mu : s \text{ étagée} \leq f\Big) \leq \sup\Big(\int s \, d\mu : s \text{ étagée } \leq g\Big) = \int g \, d\mu$$

car $\{s \text{ étagée} \leq f\} \subset \{s \text{ étagée} \leq g\}$.

Proposition 4.2.1 (Inégalité de Markov) *Soit $f : (X, \mathcal{A}, \mu) \to [0, +\infty]$ mesurable et $\alpha > 0$, alors*

$$\mu\big(x \in X : f(x) \geq \alpha\big) \leq \frac{\int_X f \, d\mu}{\alpha}.$$

Démonstration : On a

$$\alpha \mathbf{1}_{\{x \in X : f(x) \geq \alpha\}}(x) \leq f(x) \tag{4.1}$$

car soit x vérifie $f(x) \geq \alpha$ et (4.1) devient $\alpha \leq f(x)$, donc est vraie ; soit x ne vérifie pas $f(x) \geq \alpha$ et (4.1) devient $0 \leq f(x)$, ce qui est encore vraie car f est positive.

En intégrant l'inégalité (4.1), il vient

$$\int_X \alpha \mathbf{1}_{\{x \in X : f(x) \geq \alpha\}} \, d\mu \leq \int_X f \, d\mu,$$

ce qui conclut puisque $\int_X \alpha \mathbf{1}_{\{x\in X : f(x)\geq\alpha\}} d\mu = \alpha\mu\big(x \in X : f(x) \geq \alpha\big)$ d'après la définition des intégrales des fonctions simples. $\square$

• **(Croissance 2)** Si $E \subset F$ alors $\int_E f \ d\mu \leq \int_F f \ d\mu$.
Comme $E \subset F$, ensembles mesurables, on a $\mathbf{1}_E \leq \mathbf{1}_F$. Puis comme f est une fonction positive, on a aussi $f\mathbf{1}_E \leq f\mathbf{1}_F$. Et donc par croissance de l'intégrale (premier point) $\int f\mathbf{1}_E \ d\mu \leq \int f\mathbf{1}_F \ d\mu$, c'est à dire pour f mesurable positive :

$$E \subset F \Longrightarrow \int_E f \ d\mu \leq \int_F f \ d\mu.$$

• **(Nullité 1)** Si $f(x) = 0$ pour $x \in E$ alors $\int_E f \ d\mu = 0$.
En effet $\int_E f \ d\mu = \int f\mathbf{1}_E \ d\mu = \sup\Big(\int s \ d\mu\Big)$ où le sup est pris sur les fonctions étagées positives s majorées par $f\mathbf{1}_E$, ie. $s(x) \leq f(x)\mathbf{1}_E(x)$.
Or sur E, f est nulle donc nécéssairement, $s(x) = 0$ sur E et pour les fonctions s à considérer l'intégrale est $\int s\mathbf{1}_E \ d\mu = 0$, dont le sup ne peut manquer d'être aussi 0.

• **(Nullité 2)** Si $\mu(E) = 0$ alors $\int_E f \ d\mu = 0$.
En effet

$$\int_E f \ d\mu = \int f\mathbf{1}_E \ d\mu = \sup\Big(\int s\mathbf{1}_E \ d\mu\Big)$$

où le sup est pris sur les fonctions étagées positives s majorées par $f\mathbf{1}_E$. Or pour une telle fonction $s = \sum_{i=1}^n a_i\mathbf{1}_{A_i}$, on a

$$\int s\mathbf{1}_E \ d\mu = \sum_{i=1}^n a_i\mu(A_i \cap E) = 0$$

car $\mu(A_i \cap E) \leq \mu(E) = 0$. Finalement, on prend le sup sur 0, ce qui donne 0.

• **(Linéarité 1)** Si $c \geq 0$ alors $\int cf \ d\mu = c\int f \ d\mu$.
C'est clair d'abord si $f = \sum_{i=1}^n \alpha_i\mathbf{1}_{A_i}$ est une fonction étagée car alors cf l'est aussi et

$$\int cf \ d\mu \ = \ \sum_{i=1}^n c\alpha_i\mu(A_i) = c\sum_{i=1}^n \alpha_i\mu(A_i) = c\int f \ d\mu.$$

Puis si f est quelconque, s est une fonction étagée majorée par f ssi cs en est une majorée par cf. Donc l'ensemble des fonctions étagées majorées par cf est exactement $\big\{cs : s \text{ étagée} \leq f\big\}$,

$$\begin{aligned}
\int cf \ d\mu &= \sup\Big(\int s' \ d\mu : s' \text{ étagée} \leq cf\Big)\\
&= \sup\Big(\int s' \ d\mu : s' = cs, s \text{ étagée} \leq f\Big)
\end{aligned}$$

$$\begin{aligned}
&= \sup\Big(\int cs\ d\mu : s \text{ étagée} \leq f\Big)\\
&= \sup\Big(c\int s\ d\mu : s \text{ étagée} \leq f\Big)\\
&= c\sup\Big(\int s\ d\mu : s \text{ étagée} \leq f\Big)\\
&= c\int f\ d\mu
\end{aligned}$$

où on a utilisé le fait (déjà vérifié) pour une fonction étagée que $\int cs\ d\mu = c\int s\ d\mu$.

• **(Linéarité 2)** $\int (f+g)d\mu = \int fd\mu + \int gd\mu$ (On le prouve d'abord pour des fonctions étagées puis on utilisera le théorème de convergence monotone (Th. 4.3.1) pour généraliser aux fonctions positives et de signe quelconque. D'abord donc si f et g sont étagées :

$$f = \sum_{i=1}^n a_i \mathbf{1}_{A_i}, \quad g = \sum_{j=1}^p b_j \mathbf{1}_{B_j}$$

avec les A_i deux à deux disjoints et les B_j aussi. On suppose (sans restriction) que $\bigcup_{1\leq i\leq n} A_i = X$ et $\bigcup_{1\leq j\leq p} B_j = X$. On a donc pour chaque $1 \leq i \leq n$ et $1 \leq j \leq p$

$$A_i = A_i \cap \bigcup_{j=1}^p B_j = \bigcup_{j=1}^p (A_i \cap B_j) \quad \text{et} \quad B_j = B_j \cap \bigcup_{i=1}^n A_i = \bigcup_{i=1}^n (A_i \cap B_j),$$

puis comme les unions sont disjointes

$$\mathbf{1}_{A_i} = \sum_{j=1}^p \mathbf{1}_{A_i\cap B_j} \quad \text{et} \quad \mathbf{1}_{B_j} = \sum_{i=1}^n \mathbf{1}_{A_i\cap B_j}.$$

On peut alors reécrire f et g comme combinaisons des mêmes indicatrices $\mathbf{1}_{A_i\cap B_j}$, $1 \leq i \leq n, 1 \leq j \leq p$:

$$f = \sum_{\substack{1\leq i\leq n\\ 1\leq j\leq p}} a_i \mathbf{1}_{A_i\cap B_j}, \quad g = \sum_{\substack{1\leq i\leq n\\ 1\leq j\leq p}} b_i \mathbf{1}_{A_i\cap B_j}.$$

D'où $f+g = \sum_{\substack{1\leq i\leq n\\ 1\leq j\leq p}} a_i \mathbf{1}_{A_i\cap B_j} + \sum_{\substack{1\leq i\leq n\\ 1\leq j\leq p}} b_j \mathbf{1}_{A_i\cap B_j} = \sum_{\substack{1\leq i\leq n\\ 1\leq j\leq p}} (a_i+b_j)\mathbf{1}_{A_i\cap B_j}$ et

$$\begin{aligned}
\int (f+g)\ d\mu &= \sum_{\substack{1\leq i\leq n\\ 1\leq j\leq p}} (a_i+b_j)\mu(A_i\cap B_j)\\
&= \sum_{\substack{1\leq i\leq n\\ 1\leq j\leq p}} a_i\mu(A_i\cap B_j) + \sum_{\substack{1\leq i\leq n\\ 1\leq j\leq p}} b_j\mu(A_i\cap B_j)\\
&= \sum_{i=1}^n a_i \sum_{j=1}^p \mu(A_i\cap B_j) + \sum_{j=1}^p b_j \sum_{i=1}^n \mu(A_i\cap B_j)
\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^n a_i\mu\Big(A_i \cap \bigcup_{j=1}^p B_j\Big) + \sum_{j=1}^p b_j\mu\Big(\bigcup_{i=1}^n A_i \cap B_j\Big) \\
&= \sum_{i=1}^n a_i\mu(A_i) + \sum_{j=1}^p b_j\mu(B_j) \\
&= \int f\ d\mu + \int g\ d\mu.
\end{aligned}$$

Si f et g sont des fonctions mesurables positives quelconques, soient $(s_n)_{n\geq 1}$ et $(t_n)_{n\geq 1}$ des suites croissantes de fonctions étagées positives qui convergent vers f et g. Alors s_n+t_n sont aussi des fonctions étagées positives et elles convergent vers $f+g$. D'après le théorème de convergence monotone (cf. Théorème 4.3.1, et Corol. 4.3.1), $\int s_n\ d\mu$, $\int t_n\ d\mu$ $\int (s_n+t_n)\ d\mu$ convergent respectivement vers $\int f\ d\mu$, $\int g\ d\mu$, $\int (f+g)\ d\mu$. Donc la linéarité vient en passant à la limite dans $\int (s_n+t_n)\ d\mu = \int s_n\ d\mu + \int t_n\ d\mu$.

• **(Relation de Chasles)** Si E et F sont mesurables disjoints, alors pour une fonction mesurable

$$\int_{E\cup F} f\ d\mu = \int_E f\ d\mu + \int_F f\ d\mu.$$

Comme E et F sont disjoints, $\mathbf{1}_{E\cup F} = \mathbf{1}_E + \mathbf{1}_F$, il vient alors

$$\begin{aligned}
\int_{E\cup F} f\ d\mu &= \int f\mathbf{1}_{E\cup F}\ d\mu = \int f(\mathbf{1}_E + \mathbf{1}_F)\ d\mu = \int \big(f\mathbf{1}_E + f\mathbf{1}_F\big)\ d\mu \\
&= \int f\mathbf{1}_E\ d\mu + \int f\mathbf{1}_F\ d\mu = \int_E f\ d\mu + \int_F f\ d\mu
\end{aligned}$$

en utilisant la linéarité 2.

On généralisera dans le chapitre suivant au cas de fonctions mesurables à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$.

Notation. Dans la suite, pour insister sur la variable d'intégration, on notera les intégrales

$$\int f\ d\mu = \int f(x)\ \mu(dx).$$

En particulier, quand $\mu = \lambda$ (la mesure de Lebesgue), on notera $\lambda(dx) = dx$

$$\int f\ d\lambda = \int f(x)\ \lambda(dx) = \int f(x)\ dx$$

retrouvant la notation habituelle des intégrales de Riemann, cf. Chap. 6.

4.3 Convergence monotone

Théorème 4.3.1 (Convergence monotone, Beppo Levi) *Soit* $(X, \mathcal{A}, \mu)$ *un espace mesuré et* $f_n : (X, \mathcal{A}, \mu) \to [0, +\infty]$, $n \geq 1$, *une suite croissante de fonctions mesurables* $(f_n \leq f_{n+1})$. *On pose* $f(x) = \lim_{n\to+\infty} f_n(x)$ *la limite dans* $[0, +\infty]$. *Alors*

$$\int f \, d\mu = \lim_{n\to+\infty} \int f_n \, d\mu. \tag{4.2}$$

Remarque 4.3.1 — Dans l'énoncé, c'est la suite $(f_n)_{n\geq1}$ qui est croissante (i.e. $f_n(x) \leq f_{n+1}(x)$ et non pas les fonctions f_n (ce n'est pas $f_n(x) \leq f_n(y)$ pour $x \leq y$).
— La limite simple f des f_n peut aussi prendre la valeur $+\infty$!
— le théorème est faux pour une suite positive décroissante : sur $(X, \mathcal{A}, \mu) = (\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+), \lambda)$, prenons $f_n(x) = 1/(x + n)$ pour $n \geq 1$. Pour tout $n \geq 1$, on a $f_n(x) \leq f_{n+1}(x)$ et $\lim_{n\to+\infty} f_n(x) = 0$. Mais $\int f d\mu = 0$ alors que

$$\int f_n d\mu = \int_0^{+\infty} \frac{dx}{x+n} = +\infty.$$

On a d'abord besoin d'un résultat préliminaire :

Lemme 4.3.1 *Soit* s *une fonction étagée alors* $\varphi : E \longmapsto \int_E s \, d\mu$ *définit une mesure.*

Démonstration : On vérifie les deux axiomes d'une mesure.
— $\varphi(\emptyset) = \int_\emptyset s d\mu = 0$ car on obtient 0 en intégrant sur un ensemble de mesure 0.
— Soient E_k, $k \geq 1$, des ensembles mesurables deux à deux disjoints de réunion E. On suppose que s s'écrit

$$s = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i}.$$

Alors

$$\begin{aligned}
\varphi(E) &= \varphi\Big(\bigcup_{k\geq1} E_k\Big) = \int_{\bigcup_{k\geq1} E_k} s \, d\mu = \sum_{i=1}^n \alpha_i \mu\Big(A_i \cap \big(\bigcup_{k\geq1} E_k\big)\Big) \\
&= \sum_{i=1}^n \alpha_i \mu\Big(\bigcup_{k\geq1} (A_i \cap E_k)\Big) = \sum_{i=1}^n \alpha_i \sum_{k\geq1} \mu(A_i \cap E_k) \\
&= \sum_{k\geq1} \sum_{i=1}^n \alpha_i \mu(A_i \cap E_k) = \sum_{k\geq1} \int_{E_k} s \, d\mu = \sum_{k\geq1} \varphi(E_k).
\end{aligned}$$

L'application φ est donc bien une mesure sur $\mathcal{A}$.

□

Passons à la preuve du théorème de convergence monotone (Th. 4.3.1).

Démonstration : On a pour tout $x \in X$ et tout $n \geq 1$, $f_n(x) \leq f_{n+1}(x)$ et en passant à la limite $f_n(x) \leq f(x)$, ce qui donne en intégrant

$$\int_X f_n \, d\mu \leq \int_X f \, d\mu. \tag{4.3}$$

D'autre part, on a aussi $\displaystyle\int_X f_n \, d\mu \leq \int_X f_{n+1} \, d\mu$ donc $\displaystyle\int_X f_n \, d\mu$ est une suite (numérique) croissante de $[0, +\infty]$. Elle a donc une limite, notée $\alpha \in [0, +\infty]$:

$$\lim_{n\to+\infty} \int_X f_n \, d\mu = \alpha.$$

En passant à la limite dans (4.3), on a déjà $\alpha \leq \displaystyle\int_X f \, d\mu$.

Soit maintenant, s une fonction étagée majorée par f et $c \in]0, 1[$, on introduit les ensembles mesurables

$$E_n = \big\{x \in X \ : \ f_n(x) \geq cs(x)\big\} = (f_n - cs)^{-1}([0, +\infty[).$$

Comme $(f_n)_{n\geq 1}$ est une suite croissante, on constate que $E_n \subset E_{n+1}$ car si $f_n(x) \geq cs(x)$ *a fortiori* $f_{n+1}(x) \geq cs(x)$. De plus, $f_n(x) \to f(x)$ ce qui donne l'existence d'un entier N tel que pour tout $n \geq N$, $f_n(x) \geq cf(x) \geq cs(x)$, ie. $x \in E_n$ et $x \in \bigcup_{n=1}^{+\infty} E_n$. Autrement écrit

$$X = \bigcup_{n=1}^{+\infty} E_n.$$

On a

$$\int_X f_n d\mu \geq \int_{E_n} f_n \, d\mu \geq \int_{E_n} cs \, d\mu = c \int_{E_n} s \, d\mu = c\varphi(E_n). \tag{4.4}$$

Puis comme d'après le Lemme 4.3.1 φ définit une mesure et $(E_n)_{n\geq 1}$ est croissante pour l'inclusion, on a par croissance séquentielle (1.3) de la mesure φ,

$$\lim_{n\to+\infty} \varphi(E_n) = \varphi\Big(\bigcup_{n\geq 1} E_n\Big) = \varphi(X) = \int_X s \, d\mu,$$

on a donc en passant à la limite dans (4.4) :

$$\alpha \geq c \int_X s \, d\mu.$$

Comme c'est vrai pour tout $c \in]0, 1[$, en faisant tendre c vers 1, on obtient $\alpha \geq \int_X s \, d\mu$ puis en prenant le sup sur les fonctions s étagées majorées par f, on déduit $\alpha \geq \int_X f \, d\mu$. Finalement, on a obtenu

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

□

Quelques conséquences de la convergence monotone

Corollaire 4.3.1 (Linéarité) *Soient f et g des fonctions mesurables positives, on a*

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

Démonstration : Si f et g sont des fonctions mesurables positives quelconques, soient $(s_n)_{n\geq 1}$ et $(t_n)_{n\geq 1}$ des suites croissantes de fonctions étagées positives qui convergent vers f et g. Alors s_n+t_n sont aussi des fonctions étagées positives et elles convergent vers $f+g$. D'après le théorème de convergence monotone (cf. Théorème 4.3.1), $\int s_n\ d\mu$, $\int t_n\ d\mu$ et $\int(s_n+t_n)\ d\mu$ convergent respectivement vers $\int f\ d\mu$, $\int g\ d\mu$, $\int(f+g)\ d\mu$. Donc la linéarité vient en passant à la limite dans $\int(s_n+t_n)\ d\mu = \int s_n\ d\mu + \int t_n\ d\mu$, due à la linéarité déjà vue pour les fonctions étagées. □

Corollaire 4.3.2 (Intégrale et série) *Soient $(f_n)_{n\geq 1}$ une suite de fonctions mesurables de X dans $[0,+\infty]$ alors*

$$\int \sum_{n=1}^{+\infty} f_n\ d\mu = \sum_{n=1}^{+\infty} \int f_n\ d\mu.$$

Démonstration : Appliquer le théorème de convergence monotone à $g_p = \sum_{n=1}^{p} f_n$ qui est mesurable, et forme une suite croissante car $g_{p+1} - g_p = f_p \geq 0$. Comme la limite g de g_p est $g = \sum_{n=1}^{+\infty} f_n$, on a

$$\begin{aligned}
\int \sum_{n=1}^{+\infty} f_n\ d\mu = \int g\ d\mu &= \lim_{p\to+\infty} \int g_p\ d\mu = \lim_{p\to+\infty} \int \sum_{n=1}^{p} f_n\ d\mu \\
&= \lim_{p\to+\infty} \sum_{n=1}^{p} \int f_n\ d\mu = \sum_{n=1}^{+\infty} \int f_n\ d\mu
\end{aligned}$$

où on a juste utilisé la linéarité de l'intégrale pour échanger la somme finie $\sum_{n=1}^{p}$ et l'intégrale $\int$. □

Corollaire 4.3.3 (Relation de Chasles dénombrable) *Soient E_i, $i \geq 1$, ensembles mesurables disjoints et f une fonction mesurable positive*

$$\int_{\bigcup_{i\geq 1} E_i} f\ d\mu = \sum_{i\geq 1} \int_{E_i} f\ d\mu.$$

Démonstration : Comme les E_i sont disjoints on a $\mathbf{1}_{\bigcup_{i\geq 1} E_i} = \sum_{i\geq 1} \mathbf{1}_{E_i}$ et

$$\int_{\bigcup_{i\geq 1} E_i} d\mu \quad = \quad \int f\mathbf{1}_{\bigcup_{i\geq 1} E_i}\ d\mu = \int \sum_{i\geq 1} (f\mathbf{1}_{E_i})\ d\mu = \sum_{i\geq 1} \int (f\mathbf{1}_{E_i})\ d\mu$$

$$= \sum_{i\geq 1}\int_{E_i} f\ d\mu$$

avec le Corollaire 4.3.2 appliqué aux fonctions mesurables $f\mathbf{1}_{E_i}$ positives. □

Le résultat suivant est l'analogue du Lemme 4.3.1 pour des fonctions mesurables positives quelconques :

Corollaire 4.3.4 *Soit $f : (X, \mathcal{A}, \mu) \to [0, +\infty]$ mesurable alors la fonction sur $\mathcal{A}$*

$$\varphi(E) = \int_E f\ d\mu$$

est une mesure (elle est finie ssi f est intégrable). Puis pour une fonction mesurable g

$$\int_X g\ d\varphi = \int_X g\ f d\mu. \tag{4.5}$$

Remarque 4.3.2 — En quelque sorte, on a $d\varphi = f d\mu$.
- — Le théorème de convergence monotone est la clef de nombreux raisonnements typiques. Pour justifier une propriété, souvent, on la montre d'abord pour les fonctions indicatrices $\mathbf{1}_A$, on la généralise par linéarité aux fonctions étagées positives puis enfin aux fonctions mesurables quelconques positives par convergence monotone en approximant par des fonctions simples. Enfin, les fonctions f de signe quelconque s'écrivent $f = f^+ - f^-$.

Illustrons cette façon de faire par la deuxième partie de cette preuve.

Démonstration : φ est une mesure car
- — $\varphi(\emptyset) = \int_\emptyset f d\mu = 0$ car on obtient 0 en intégrant sur un ensemble de mesure 0.
- — Soient E_k, $k \geq 1$, des ensembles mesurables deux à deux disjoints de réunion E. On a $\mathbf{1}_E = \sum_{k\geq 1}\mathbf{1}_{E_k}$ car les E_k sont deux à deux disjoints

$$\varphi(E) = \int_X \mathbf{1}_E f\ d\mu = \int_X \sum_{k\geq 1}\mathbf{1}_{E_k} f\ d\mu = \sum_{k\geq 1}\int_X \mathbf{1}_{E_k} f\ d\mu = \sum_{k\geq 1}\varphi(E_k).$$

L'application φ est donc bien une mesure.

Pour la seconde partie, si $g = \mathbf{1}_A$ est une indicatrice, on a

$$\int_X \mathbf{1}_A\ d\varphi = \varphi(A) = \int_X \mathbf{1}_A f\ d\mu = \int_X gf\ d\mu$$

et (4.5) est satisfaite dans ce cas. Pour une fonction étagée, ça reste vraie par linéarité en utilisant le premier cas : si $g = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i}$ alors

$$\int_X g\ d\varphi = \int_X \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i}\ d\varphi = \sum_{i=1}^n \alpha_i \int_X \mathbf{1}_{A_i}\ d\varphi = \sum_{i=1}^n \alpha_i \int_X \mathbf{1}_{A_i} f\ d\mu$$

$$= \int_X \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i} f \, d\mu = \int_X gf \, d\mu.$$

Pour g mesurable positive, on prend $(s_n)_{n\geq 1}$ suite de fonctions étagées croissantes qui converge vers g et on applique le théorème de convergence monotone

$$\lim_{n\to+\infty} \int_X s_n \, d\varphi = \int_X g \, d\varphi, \quad \lim_{n\to+\infty} \int_X s_n \, f d\mu = \int_X gf \, d\mu$$

car $s_n \nearrow g$ et $s_n f \nearrow gf$ $(f \geq 0)$. D'après le cas étagé (Lemme 4.3.1), pour s_n, on a :

$$\int_X s_n \, d\varphi = \int_X s_n f \, d\mu$$

ce qui donne en passant à la limite par convergence monotone :

$$\int_X g \, d\varphi = \int_X gf \, d\mu.$$

Si g est de signe quelconque on écrit $g = g^+ - g^-$ et on utilise le cas positif et la linéarité pour obtenir (4.5) dans ce cas réel de signe quelconque.
Pour g complexe, on écrit $g = Re(g) + iIm(g)$ et on applique le cas réel à chaque terme réel. □

4.4 Lemme de Fatou

Théorème 4.4.1 (Lemme de Fatou) *Soit $f_n : (X, \mathcal{A}, \mu) \to [0, +\infty]$, $n \geq 1$, une suite de fonctions mesurables positives, alors*

$$\int \liminf_{n\to+\infty} f_n \, d\mu \leq \liminf_{n\to+\infty} \int f_n \, d\mu.$$

Corollaire 4.4.1 *Soit $(f_n)_{n\geq 1}$ une suite de fonctions mesurables positives qui converge simplement vers f et telle que la suite des intégrales $\int f_n d\mu$ est majorée par M alors*

$$\int f d\mu \leq M.$$

Démonstration : $\liminf_n f_n = \sup_{n\geq 1} \inf_{k\geq n} f_k = \sup_{n\geq 1} g_n$ avec $g_n = \inf_{k\geq n} f_k$. Les fonctions g_n sont mesurables, positives et $g_n \leq g_{n+1}$. Par le théorème de convergence monotone (Th. 4.3.1),

$$\int \lim_{n\to+\infty} g_n \, d\mu = \lim_{n\to+\infty} \int g_n \, d\mu.$$

Or comme la suite $(g_n)_{n\geq 1}$ est croissante sa limite est égale à son sup et

$$\lim_{n\to+\infty} g_n = \sup_{n\geq 1} g_n = \sup_{n\geq 1}\inf_{k\geq n} f_k = \liminf_n f_n.$$

Puis comme $g_n \leq f_n$, on a $\int g_n \, d\mu \leq \int f_n \, d\mu$. D'où

$$\int \liminf_n f_n \, d\mu = \lim_{n\to+\infty} \int g_n \, d\mu = \liminf_n \int g_n \, d\mu \leq \liminf_{n\to+\infty} \int f_n \, d\mu$$

où on a utilisé que la limite de $\int g_n \, d\mu$ coïncide avec sa liminf car la limite existe. □

Ce résultat explique que par passage à la limite, les intégrales de fonctions positives ne peuvent que diminuer.

Applications

En anticipant les intégrales pour les fonctions de signe quelconque (Chapitre 5) , la proposition suivante suit facilement du Lemme de Fatou (Th. 4.4.1). Cette proposition permettra de prouver facilement le théorème de convergence dominée (Th. 5.4.1).

Proposition 4.4.1 *Soient, sur $(X, \mathcal{A}, \mu)$, g une fonction mesurable intégrable et $(f_n)_{n\geq 1}$ une suite de fonctions mesurables intégrables (le signe de ces fonctions n'est pas prescrit).*

1. *Si $g \leq f_n$ alors $\int \liminf_{n\to+\infty} f_n \, d\mu \leq \liminf_{n\to+\infty} \int f_n \, d\mu$.*
2. *Si $f_n \leq g$ alors $\limsup_{n\to+\infty} \int f_n \, d\mu \leq \int \limsup_{n\to+\infty} f_n \, d\mu$.*

Démonstration : Par le lemme de Fatou (Th. 4.4.1) appliqué à $f_n - g \geq 0$, on a

$$\int \liminf_{n\to+\infty} (f_n - g) \, d\mu \leq \liminf_{n\to+\infty} \int (f_n - g) \, d\mu$$

ce qui prouve 1) en rajoutant $\int g \, d\mu$ qui est finie par hypothèse.

Pour le 2), de la même façon, le lemme de Fatou (Th. 4.4.1) appliqué à $g - f_n \geq 0$ donne

$$\int \liminf_{n\to+\infty} (g - f_n) \, d\mu \leq \liminf_{n\to+\infty} \int (g - f_n) \, d\mu$$

on retranche $\int g \, d\mu < +\infty$ et on change le signe pour récupérer l'inégalité sur la $\limsup$. □

Chapitre 5

Intégrales des fonctions mesurables de signe quelconque

En Section 5.1, on généralise le chapitre précédent pour l'intégrale des fonctions mesurables positives à celle des fonctions mesurables de signe quelconque. En Section 5.2, on prouve une importante formule de transfert ou formule de changement de variable abstrait (Th. 5.2.1). On présente également en Section 5.4 le théorème de convergence dominée (Th. 5.4.1) qui montre que, sous une hypothèse de domination, on peut échanger limite et intégrale. Ce théorème permet d'étudier les fonctions définies comme des intégrales à paramètres. Sauf mention contraire, on considère dans tout ce chapitre un espace mesuré $(X, \mathcal{A}, \mu)$.

5.1 Définition et propriétés

Par défaut, dans la suite, on considère une fonction $f : (X, \mathcal{A}, \mu) \to \mathbb{R}$ mesurable (ou selon les cas, à valeurs dans $\mathbb{C}$. Pour une telle fonction f, la fonction $|f|$ est mesurable et son intégrale $\int f\, d\mu$ est bien définie d'après le Chapitre 4 (mais peut valoir $+\infty$).

Définition 5.1.1 (Intégrabilité) *Une fonction f est dite μ-intégrable si la fonction positive $|f|$ est d'intégrale finie :*

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

Dans la suite, on écrit une fonction $f : (X, \mathcal{A}, \mu) \to \mathbb{C}$ sous la forme

$$\begin{aligned} f &= u + iv \\ &= u^+ - u^- + i(v^+ - v^-) \end{aligned} \tag{5.1}$$

avec $u = Re(f)$ et $v = Im(f)$. Comme les fonctions positives u^+, u^-, v^+, v^- sont toutes majorées par $|f|$ alors leurs intégrales sont toutes finies car par exemple $\int u^+\, d\mu \leq \int |f|\, d\mu <$

$+\infty$ et on pose alors par définition

$$\begin{aligned}\int f\ d\mu &= \int u^+\ d\mu - \int u^-\ d\mu + i\left(\int v^+\ d\mu - \int v^-\ d\mu\right)\\ &= \int u\ d\mu + i\int v\ d\mu\end{aligned}$$

ou, plus simplement, si $f : X \longrightarrow \mathbb{R}$, $\int f\ d\mu = \int f^+\ d\mu - \int f^-\ d\mu$. Puis l'intégrale sur $E \in \mathcal{A}$ de f fonction intégrable est définie par

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

ce qui a bien un sens car la fonction $f\mathbf{1}_E$ est intégrable puisque $|f\mathbf{1}_E| \leq |f|$.

Remarque 5.1.1 — Notons que si une fonction mesurable est positive, on peut toujours définir son intégrale mais elle peut valoir $+\infty$. Elle est dite alors μ-intégrable si son intégrale par rapport à μ est finie.
- — Tandis que si f est de signe quelconque, on ne définit son intégrale que si elle est intégrable, c'est à dire lorsque $|f|$ est d'intégrale (forcément définie) finie.
- — En particulier dans la théorie de l'intégrale de Lebesgue, il n'y a pas de notion de simple intégrabilité, c'est à dire de fonction non intégrable dont l'intégrale converge simplement comme par exemple l'intégrale de Riemann $\int_0^{+\infty} \frac{\sin x}{x} dx$ (à voir).

Proposition 5.1.1 (Quelques propriétés) — *Si f, g sont mesurables réelles et $f \leq g$ alors $\int f d\mu \leq \int g d\mu$.*
- — *Si $\mu(E) = 0$, alors $\int_E f\ d\mu = 0$ pour toute fonction intégrable.*
- — **(Linéarité)** *Soient $\alpha, \beta \in \mathbb{C}$ et f, g intégrables de X dans $\mathbb{C}$ alors*

$$\int (\alpha f + \beta g)\ d\mu = \alpha \int f\ d\mu + \beta \int g\ d\mu.$$

- — **(Chasles)** *Si E et F sont disjoints, alors $\int_{E\cup F} f\ d\mu = \int_E f\ d\mu + \int_F f\ d\mu$.*

Démonstration : Tout ceci provient facilement des propriétés déjà vues pour les intégrales de fonctions positives et passe au cas des fonctions de signe quelconque en les écrivant sous la forme (5.1) et en appliquant à chaque terme les propriétés correspondantes des intégrales de fonctions mesurables positives.
Pour la linéarité, c'est fastidieux : écrire $f = Re(f)^+ - Re(f)^- + i(Im(f)^+ - Im(f^-))$, $\alpha = Re(\alpha)^+ - Re(\alpha)^- + i(Im(\alpha)^+ - Im(\alpha^-))$ et de même pour g et β. Développer $(\alpha f + \beta g)$ et utiliser la linéarité pour les fonctions positives avec des coefficients positifs et reformer α, β, f, g à la fin. $\square$

Proposition 5.1.2 (Intégrale et valeurs absolues ou module) *Soit $f : (X, \mathcal{A}, \mu) \to \mathbb{C}$ intégrable alors*
$$\left|\int f\ d\mu\right| \leq \int |f|\ d\mu.$$

Démonstration : Si $\int f d\mu = 0$, c'est immédiat. Sinon, soit $\alpha = \left|\int f\ d\mu\right| / (\int f d\mu)$. Alors $|\alpha| = 1$. Puis
$$\begin{aligned}\left|\int f\ d\mu\right| &= \alpha \int f\ d\mu = \int \alpha f\ d\mu = \int Re(\alpha f)\ d\mu + i\int Im(\alpha f)\ d\mu = \int Re(\alpha f)\ d\mu \\ &\leq \int |\alpha f|\ d\mu = \int |f|\ d\mu\end{aligned}$$
en utilisant la majoration $Re(z) \leq |z|$ valable pour tout $z \in \mathbb{C}$ puis $|\alpha| = 1$. □

5.2 Formule de transfert

Soient $(X, \mathcal{A})$ et $(Y, \mathcal{B})$ deux espaces mesurables et $\varphi : (X, \mathcal{A}) \to (Y, \mathcal{B})$ une fonction mesurable. Si on considère une mesure μ sur $(X, \mathcal{A})$ alors la mesure image $\mu\varphi^{-1} = \mu_\varphi$ est une mesure sur $(Y, \mathcal{B})$ (Définition 3.2.1). On a un lien entre les intégrales sur X par rapport à μ et celle sur Y par rapport à $\mu\varphi^{-1} = \mu_\varphi$:

Théorème 5.2.1 (Formule de transfert) *Soit $h : (Y, \mathcal{B}) \to \mathbb{K} = \mathbb{R}, \overline{\mathbb{R}}, \mathbb{C}$ mesurable alors h est μ_φ-intégrable ssi $h \circ \varphi$ est μ-intégrable et*
$$\int_X h \circ \varphi\ d\mu = \int_Y h\ d\mu_\varphi. \tag{5.2}$$

Il s'agit d'une formule de changement de variable abstraite, très générale puisque la seule condition pour le changement de variable $y = \varphi(x)$ est que φ soit mesurable ! Pour une formule de changement de variable plus concrète, voir le Chap. 7 et le Th. 7.4.1.

Démonstration : La stratégie est à nouveau de commencer par $h = \mathbf{1}_B$ une indicatrice, puis par linéarité de généraliser aux fonctions étagées et par convergence monotone aux fonctions mesurables positives ; enfin par linéarité aux fonctions mesurables quelconques.

- Si $h = \mathbf{1}_B$, $B \in \mathcal{B}$, alors $\varphi^{-1}(B) \in \mathcal{A}$ car φ est mesurable et
$$\int_Y \mathbf{1}_B\ d\mu_\varphi = \mu_\varphi(B) = \mu(\varphi^{-1}(B)) = \int_X \mathbf{1}_{\varphi^{-1}(B)}\ d\mu = \int_X \mathbf{1}_B \circ \varphi\ d\mu$$
car $\mathbf{1}_{\varphi^{-1}(B)} = \mathbf{1}_B \circ \varphi$ et (5.2) est vérifiée.
- Si $h = \sum_{i=1}^n \alpha_i \mathbf{1}_{B_i}$ est étagée, alors
$$h \circ \varphi = \left(\sum_{i=1}^n \alpha_i \mathbf{1}_{B_i}\right) \circ \varphi = \sum_{i=1}^n \alpha_i (\mathbf{1}_{B_i} \circ \varphi) = \sum_{i=1}^n \alpha_i \mathbf{1}_{\varphi^{-1}(B_i)}$$

et

$$\begin{aligned}
\int_X (h\circ\varphi)\ d\mu &= \sum_{i=1}^n \alpha_i \int_X \mathbf{1}_{\varphi^{-1}(B_i)}\ d\mu = \sum_{i=1}^n \alpha_i \mu(\varphi^{-1}(B_i))\\
&= \sum_{i=1}^n \alpha_i\ \mu_\varphi(B_i) = \sum_{i=1}^n \alpha_i \int_Y \mathbf{1}_{B_i}\ d\mu_\varphi\\
&= \int_Y h\ d\mu_\varphi.
\end{aligned}$$

et (5.2) est à nouveau vérifiée.

• Si h est mesurable positive alors il existe s_n, $n\geq 1$, des fonctions étagées positives qui forment une suite croissante qui converge vers h. Mais alors $s_n\circ\varphi$ sont aussi des fonctions étagées qui forment une suite croissante et qui converge vers $h\circ\varphi$. On a donc $\lim_{n\to+\infty} s_n = h$ et $\lim_{n\to+\infty} s_n\circ\varphi = h\circ\varphi$. D'après le point précédent, on a

$$\int_X s_n\circ\varphi\ d\mu = \int_Y s_n\ d\mu_\varphi.$$

Par convergence monotone (Th. 4.3.1), le membre de gauche converge vers $\int_X h\circ\varphi\ d\mu$, celui de droite vers $\int_Y h\ d\mu_\varphi$, ce qui justifie donc encore (5.2) dans ce cas.

• Si h est quelconque alors $|h\circ\varphi| = |h|\circ\varphi$ et le point 3, appliquée à $|h|$ positive, donne $\int_X |h\circ\varphi|\ d\mu = \int_Y |h|\ d\mu_\varphi$ et donc l'équivalence de la μ_φ-intégrabilité de h et de la μ-intégrabilité de $h\circ\varphi$.
On écrit alors $(h\circ\varphi)^+ = h^+\circ\varphi$ et $(h\circ\varphi)^- = h^-\circ\varphi$ puis

$$\begin{aligned}
\int_X (h\circ\varphi)\ d\mu &= \int_X (h\circ\varphi)^+\ d\mu - \int_X (h\circ\varphi)^-\ d\mu\\
&= \int_Y h^+\ d\mu_\varphi - \int_Y h^-\ d\mu_\varphi\\
&= \int_Y h\ d\mu_\varphi
\end{aligned}$$

en utilisant le point 3 pour les fonctions positives h^+ et h^-, ce qui achève la preuve de (5.2) dans le cas général. □

Exemple (Espérance probabiliste). Une application importante d'utilisation de la formule de transfert (5.2) est en probabilités pour exprimer une espérance d'une variable aléatoire X comme une intégrale par rapport à la loi de la variable aléatoire $\mathbb{P}_X$. On considère $(X,\mathcal{A},\mu) = (\Omega,\mathcal{F},\mathbb{P})$ et $X : (\Omega,\mathcal{F},\mathbb{P}) \to \mathbb{R}$ joue le rôle de f et $h(x) = x$, on a alors

$$\mathbb{E}[X] = \int X\ d\mathbb{P} = \int_{\mathbb{R}} x\ \mathbb{P}_X(dx)$$

dans ce cas la mesure image $\mathbb{P}_X$ est la loi sur $\mathbb{R}$ de la variable aléatoire.

5.3 Ensembles négligeables

Définition 5.3.1 *Un ensemble $A \in \mathcal{A}$ est dit μ-négligeable si $\mu(A) = 0$.*

Exemples : Sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, $\{x_0\}, \mathbb{N}, \mathbb{Z}, \mathbb{Q}$ sont négligeables. Sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \eta)$, où η est la mesure de comptage, seul $\emptyset$ est négligeable.

Définition 5.3.2 *Une fonction f mesurable est dite μ-négligeable s'il existe A négligeable tel que $\forall x \notin A$ alors $f(x) = 0$ (l'ensemble des points où f est non nulle est dans un négligeable).*

Exemples : Sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, $f(x) = \mathbf{1}_{\mathbb{Q}}(x)$, $g(x) = x\mathbf{1}_{\mathbb{N}}(x)$ sont négligeables.

Définition 5.3.3 (Presque partout) *Une propriété est dite vraie (μ-)presque partout, si l'ensemble des points où elle n'est pas vraie est μ-négligeable.*
En particulier, $f = g$ presque partout si $\{x \in X : f(x) \neq g(x)\}$ est dans un négligeable.
On écrit p.p. pour presque partout.

Proposition 5.3.1

- *Si $f = g$ p.p. alors f est intégrable ssi g l'est et*
$$\int f\ d\mu = \int g\ d\mu.$$
- *Si $f \leq g$ p.p. alors $\int f\ d\mu \leq \int g\ d\mu$.*
- *Si $f \geq 0$ et $\int f\ d\mu < +\infty$ alors f est finie p.p.*
- *Si $\left|\int f\ d\mu\right| = \int |f|\ d\mu$ alors il existe $\alpha \in \mathbb{C}$ de module 1 tel que $\alpha f = |f|$ ie. f est d'argument constant.*

Démonstration : • Soit $A = \{x \in X : f(x) \neq g(x)\}$ alors $\mu(A) = 0$. Si $x \notin A$, on a $f(x) = g(x)$ donc $|f(x)| = |g(x)|$. Donc si f est intégrable alors g l'est aussi puisque
$$\begin{aligned}\int |g|\ d\mu &= \int_A |g|\ d\mu + \int_{A^c} |g|\ d\mu = \int_{A^c} |g|\ d\mu \\ &= \int_{A^c} |f|\ d\mu = \int_{A^c} |f|\ d\mu + \int_A |f|\ d\mu = \int |f|\ d\mu < +\infty.\end{aligned}$$
Puis le même calcul sans les $|\cdot|$ montre que $\int f d\mu = \int g\ d\mu$.

• Si $A = \{x \in X : f(x) = +\infty\}$ n'est pas négligeable alors
$$\int f d\mu \geq \int_A f d\mu = +\infty$$
car $f = +\infty$ sur A qui est de mesure $\mu(A) > 0$. On contredit la finitude de $\int f\ d\mu$. Il faut donc avoir $\mu(A) = 0$.

• Reprendre la preuve de $\left|\int f d\mu\right| \leq \int |f|\ d\mu$, page 53. Pour avoir égalité, il faut qu'il y ait égalité dans toutes les étapes de cette preuve, ce qui nécessite $Re(\alpha f) = |\alpha f|$. Mais $Re(z) = |z|$ exige $z = |z|$. Ici on a donc $|\alpha f| = \alpha f$, soit, puisque $|\alpha| = 1$, $|f| = \alpha f$. □

Remarque 5.3.1 (Tribu complète) Il se peut que A mesurable de mesure $\mu(A)$ nulle contienne des sous-ensembles B non mesurables, ce qui cause bien des soucis. Dans ce cas, on complète la tribu comme expliqué en Section 1.6.

5.4 Convergence dominée et applications

Théorème 5.4.1 (Convergence dominée) *Soit sur $(X, \mathcal{A}, \mu)$, $f_n : (X, \mathcal{A}, \mu) \to \mathbb{C}$, $n \geq 1$, mesurables telles que $f_n \to f$ p.p. quand $n \to +\infty$. S'il existe g une fonction mesurable $(X, \mathcal{A}, \mu) \to [0, +\infty]$ telle que*

1. *$|f_n| \leq g$ p.p.*
2. *g est μ-intégrable.*

Alors

$$\lim_{n\to+\infty} \int f_n \, d\mu = \int f \, d\mu. \tag{5.3}$$

Conséquence : Si la convergence est dominée, on peut intervertir limite et intégrale. En fait, on a même mieux que (5.3) : on a

$$\lim_{n\to+\infty} \int |f_n - f| \, d\mu = 0. \tag{5.4}$$

Cela asure en particulier (5.3) puisque

$$\left| \int f_n \, d\mu - \int f \, d\mu \right| = \left| \int (f_n - f) \, d\mu \right| \leq \int |f_n - f| \, d\mu.$$

et donc $\lim_{n\to+\infty} \left| \int f_n \, d\mu - \int f \, d\mu \right| \leq \lim_{n\to+\infty} \int |f_n - f| \, d\mu$.

On propose deux preuves différentes du Th. 5.4.1. L'une à partir de la Proposition 4.4.1, l'autre pour la conclusion plus forte (5.4).

Démonstration : Comme $\lim_{n\to+\infty} f_n = f$ p.p. et $-g \leq f_n \leq g$, la Proposition 4.4.1 assure que

$$\begin{aligned}
\limsup_{n\to+\infty} \int f_n \, d\mu \quad &\leq \quad \int \limsup_{n\to+\infty} f_n \, d\mu = \int f \, d\mu = \int \liminf_{n\to+\infty} f_n \, d\mu \\
&\leq \quad \liminf_{n\to+\infty} \int f_n \, d\mu
\end{aligned}$$

ce qui prouve le résultat. □

Démonstration : (de (5.4)) On considère la fonction $2g - |f - f_n|$. Comme $|f_n| \leq g$ et (à la limite) $|f| \leq g$, il vient facilement $|f - f_n| \leq 2g$, c'est à dire $2g - |f - f_n| \geq 0$. De plus, comme $f_n \to f$ simplement, on a $\lim_{n\to+\infty} 2g - |f - f_n| = 2g$. On applique alors le lemme de Fatou (Th. 4.4.1) à cette suite de fonctions :

$$\begin{aligned}
\int \liminf_{n\to+\infty} \big(2g - |f - f_n|\big)\, d\mu &\leq \liminf_{n\to+\infty} \int 2g - |f - f_n|\, d\mu \\
\int 2g\, d\mu &\leq \liminf_{n\to+\infty} \int 2g\, d\mu - \int |f - f_n|\, d\mu \\
\int 2g\, d\mu &\leq \int 2g\, d\mu + \liminf_{n\to+\infty} \left(-\int |f - f_n|\, d\mu\right) \\
0 &\leq -\limsup_{n\to+\infty} \int |f - f_n|\, d\mu \\
\limsup_{n\to+\infty} \int |f - f_n|\, d\mu &\leq 0
\end{aligned}$$

en simplifiant par $\int g\, d\mu$ finie. Comme en plus $\liminf_{n\to+\infty} \int |f - f_n|\, d\mu \geq 0$, on a $\lim_{n\to+\infty} \int |f - f_n|\, d\mu = 0$ et a fortiori

$$\lim_{n\to+\infty} \int f_n\, d\mu = \int f\, d\mu.$$

□

Exemple : Montrer que si $f_n : [0,1] \to [0,1]$ est continue et pour tout $x \in [0,1]$ on a $f_n(x) \to 0$ alors $\int_{[0,1]} f_n d\lambda \to 0$. Essayer sans convergence dominée.

Exemple : Considérer la suite de fonctions $(f_n)_{n\geq 1}$ mesurables définies sur $\mathbb{R}$ et telle que pour tout $n \geq 1$, et $x \in \mathbb{R}$: $|f_n(x)| \leq 1/(1 + x^2)$. Supposons que pour tout $x \in \mathbb{R}$, $f_n(x)$ converge vers $f(x)$. Montrer alors que

$$\lim_{n\to+\infty} \int_{-\infty}^{+\infty} f_n(x)\, dx = \int_{-\infty}^{+\infty} f(x)\, dx.$$

Application aux intégrales à paramètres

Les résultats suivants donnent des conditions assez faibles sur f pour que la fonction $F(t) = \int_X f(t,x)\, \mu(dx)$ soit continue ou dérivable en $t \in T$.

Théorème 5.4.2 (Continuité sous l'intégrale) *Soit $(X, \mathcal{A}, \mu)$ un espace mesuré et T un espace métrique (en général $T = \mathbb{R}$ ou à $\mathbb{C}$). On considère $f : T \times X \to \mathbb{C}$ tel que pour tout $t \in T$, $x \mapsto f(t,x)$ est mesurable et*

$$\forall (t,x) \in T \times X, \quad |f(t,x)| \leq g(x)$$

avec g fonction μ-intégrable ($\int_X g(x)\, \mu(dx) < +\infty$). On pose $F(t) = \int_X f(t,x)\, \mu(dx)$. Soit $t_0 \in T$, on suppose que pour presque tout $x \in X$, $t \mapsto f(t,x)$ est continue en t_0. Alors la fonction F est continue en t_0.

Démonstration : Comme $\mathbb{R}$ et $\mathbb{C}$ sont métriques, la continuité est donnée par la continuité séquentielle (c'est à dire avec les suites : F est continue en t_0 ssi pour toute suite $(t_n)_{n\geq 1}$ qui converge vers t_0, on a $\lim_{n\to+\infty} F(t_n) = F(t_0)$).

Il faut alors voir que pour toute suite $(t_n)_{n\geq 1}$ qui converge vers t_0, on a $F(t_n) \to F(t_0)$. Mais $F(t_n) = \int_X f(t_n, x)\ \mu(dx)$. Les fonctions $x \mapsto f(t_n, x)$ sont mesurables et dominées par g, intégrable. De plus quand $n \to +\infty$, $f(t_n, x) \to f(t_0, x)$ pour presque chaque $x \in X$ par la continuité presque partout de $f(t_0, \cdot)$. La conclusion découle directement du théorème de convergence dominée (Th. 5.4.1) :

$$\lim_{n\to+\infty} F(t_n) = \lim_{n\to+\infty} \int_X f(t_n, x)\ \mu(dx) = \int_X f(t_0, x)\ \mu(dx) = F(t_0).$$

□

Exemples. La fonction définie par $F(t) = \int_0^{+\infty} e^{-t^3 x} dx$ est continue sur $\mathbb{R}_+^*$.
La fonction Gamma définie par $\Gamma(t) = \int_0^{+\infty} x^{t-1} e^{-x}\ dx$ est continue sur $\mathbb{R}_+^*$.

Théorème 5.4.3 (Dérivabilité sous l'intégrale) *Soit $(X, \mathcal{A}, \mu)$, I un intervalle ouvert de $\mathbb{R}$ et $f : I \times X \to \mathbb{C}$ telle que*

— *$\forall t \in I$, $x \mapsto f(t, x)$ est mesurable ;*
— *$\exists t_0 \in I$, $x \mapsto f(t_0, x)$ est μ-intégrable.*

On suppose qu'il existe $A \in \mathcal{A}$ tel que $\mu(A^c) = 0$ et

— *$\forall x \in A$, $t \mapsto f(t, x)$ est dérivable en tout $t \in I$;*
— *$\forall x \in A, \forall t \in I$, $\left|\frac{\partial f}{\partial t}(t, x)\right| \leq g(x)$, avec $\int_X g(x)\ \mu(dx) < +\infty$.*

Alors $F(t) = \int_X f(t, x)\ \mu(dx)$ est finie et la fonction F est dérivable de dérivée

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

Démonstration : Pour $x \in A$, $\frac{\partial f}{\partial t}(t, x) = \lim_{t_n \to t} \frac{f(t_n,x)-f(t,x)}{t_n - t}$ est mesurable car limite de fonctions mesurables.
Pour $t_1 \in I$ quelconque et t_0 donné dans l'énoncé, par le théorème des accroissements finis, on a $\theta \in]t_0, t_1[$:

$$|f(t_1, x) - f(t_0, x)| = \left|\frac{\partial f}{\partial t}(\theta, x)\right| |t_1 - t_0| \leq g(x)|t_1 - t_0|.$$

Il suit $|f(t_1, x)| \leq g(x)|t_1 - t_0| + |f(t_0, x)|$ qui est intégrable car g et $f(t_0, \cdot)$ le sont. L'intégrale $F(t_1)$ est donc finie pour tout $t_1 \in I$.

Soit maintenant $t \in I$ fixé et $t_n \to t$, on a

$$\frac{F(t_n) - F(t)}{t_n - t} = \int_A \frac{f(t_n, x) - f(t, x)}{t_n - t}\ \mu(dx).$$

Or $\frac{f(t_n,x)-f(t,x)}{t_n-t}$ est une suite de fonctions qui tend vers $\frac{\partial f}{\partial t}(t,x)$ et qui vaut par le théorème des accroissements finis $\frac{\partial f}{\partial t}(\theta_n,x)$ pour un certain $\theta_n \in]t,t_n[$. Or $\frac{\partial f}{\partial t}(\theta_n,x)$ est dominée par $g(x)$, intégrable. Le théorème de convergence dominée (Th. 5.4.1) s'applique et donne

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

□

Exemples.

— La fonction définie par $F(t) = \int_0^{+\infty} e^{-t^3x}\ dx$ est dérivable sur $\mathbb{R}_+^*$.

— La fonction Gamma définie par $\Gamma(t) = \int_0^{+\infty} x^{t-1}e^{-x}\ dx$ est dérivable sur $\mathbb{R}_+^*$.

En fait par un argument récursif, on montre que ces fonctions sont mêmes C^∞ sur $\mathbb{R}_+^*$.

Contre-exemple. Soit $f:\mathbb{R}\to\mathbb{R}$ mesurable et sa primitive

$$F(t) = \int_0^t f(x)dx = \int f(x)\mathbf{1}_{[0,t]}(x)\ dx$$

On est tenté d'appliquer le théorème de dérivation (Th. 5.4.3) à $f(t,x) = f(x)\mathbf{1}_{[0,t]}(x)$ en constatant que $\frac{\partial f}{\partial t}$ existe presque partout (sauf en $t=x$) et vaut alors 0. On aurait donc $F'(t)=0$. Le problème vient du fait que l'ensemble presque sûr où $f(\cdot,x)$ est dérivable dépend de x, ce qui n'est pas autorisé pour appliquer le Th. 5.4.3.

Chapitre 6

Lien avec l'intégrale de Riemann

L'intégrale de Riemann est une intégrale pour les fonctions définies sur $\mathbb{R}$. La mesure classique à considérer sur $\mathbb{R}$ étant la mesure de Lebesgue λ, on va donc faire le lien entre l'intégrale de Riemann et l'intégrale de Lebesgue pour la mesure de Lebesgue λ. On commence par faire le lien entre les deux types d'intégrales sur des intervalles $[a,b]$ bornés pour les fonctions en escalier en Section 6.1, pour les fonctions continues en Section 6.2, et pour les fonctions Riemann-intégrables en Section 6.3. Enfin en Section 6.4, on fait le lien avec les intégrales de Riemann impropres.

6.1 Fonctions en escalier

Soit $[a,b]$ un intervalle borné de $\mathbb{R}$ et $S = \{a = t_0 \leq t_1 \cdots \leq t_n = b\}$ une subdivision de $[a,b]$. On note $\rho(S) = \max_{0 \leq i \leq n-1} |t_i - t_{i-1}|$ le pas de la subdivision S.

On a déjà remarqué que les fonctions en escalier

$$f(x) = \sum_{i=0}^{n-1} \alpha_i \mathbf{1}_{[t_i, t_{i+1}[}(x)$$

sont des cas spéciaux de fonctions étagées (car les intervalles $[t_i, t_{i+1}[$ sont des boréliens). Leur intégrale de Riemann sur $[a,b]$ est donnée par

$$\int_a^b f(x)\ dx = \sum_{i=0}^{n-1} \alpha_i (t_{i+1} - t_i).$$

Tandis (qu'en tant que fonctions étagées), leur intégrale de Lebesgue sur $[a,b] \in \mathcal{B}(\mathbb{R})$ est donnée par

$$\int_{[a,b]} f\ d\lambda = \sum_{i=0}^{n-1} \alpha_i \lambda([t_i, t_{i+1}[) = \sum_{i=0}^{n-1} \alpha_i (t_{i+1} - t_i).$$

Les deux types d'intégrale de Riemann et de Lebesgue coïncident donc pour les fonctions en escalier :

$$\int_{[a,b]} f\ d\lambda = \int_a^b f(x)\ dx. \tag{6.1}$$

6.2 Fonctions continues de $[a, b] \to \mathbb{R}$

Soit $f : [a, b] \to \mathbb{R}$ une fonction continue. Elle est mesurable et bornée par M : $|f(x)| \leq M$. Donc f est λ-intégrable (au sens de Lebesgue) sur $[a, b]$ fini :

$$\int_{[a,b]} |f|\ d\lambda \leq \int_{[a,b]} M\ d\lambda = M\lambda([a, b]) = (b - a)M < +\infty.$$

En fait, la continuité sur $[a, b]$ donne l'uniforme continuité de f (théorème de Heine). La fonction f est donc limite uniforme de fonctions en escalier E_n (on dit dans ce cas que f est une fonction **réglée**).

On a vu en Section 6.1 que pour les fonctions en escalier E_n, les intégrales de Riemann $\int_a^b E_n(x)\ dx$ et de Lebesgue $\int_{[a,b]} E_n\ d\lambda$ coïncident. Par définition, $\int_a^b f(x)\ dx$ est limite de $\int_a^b E_n(x)\ dx$ et $\int_{[a,b]} f\ d\lambda$ est sup, donc limite croissante des intégrales $\int_{[a,b]} E_n\ d\lambda$. On a donc avec le théorème de convergence monotone (Th. 4.3.1) :

$$\int_a^b f(x)\ dx = \lim_{n\to+\infty} \int_a^b E_n(x)\ dx = \lim_{n\to+\infty} \int_{[a,b]} E_n(x)\ d\lambda = \int_{[a,b]} f\ d\lambda.$$

Les intégrales de Riemann et de Lebesgue coïncident donc pour des fonctions continues sur des intervalles bornés. Elles sont même λ-intégrables.

Remarque 6.2.1
- — Le même argument s'applique en fait aux fonctions réglées (limites uniformes de fonctions en escalier) : leurs intégrale de Riemann et de Lebesgue sur des intervalles bornés sont égales.
- — On a deux notations (a priori) différentes pour les intégrales de Riemann et de Lebesgue d'une fonction (continue) f mais finalement les deux intégrales coïncident :

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

- — Tous les calculs que l'on sait faire avec l'intégrale de Riemann restent donc valables avec l'intégrale de Lebesgue. Les calculs de primitives, intégration par parties, changements de variables usuels, etc, ne disparaissent pas.

6.3 Fonctions Riemann-intégrables

Mesurabilité

D'abord, on montre que les fonctions Riemann-intégrables ont un sens pour l'intégrale de Lebesgue : elles sont égales λ-presque partout à des fonctions mesurables (pour la tribu borélienne de $\mathbb{R}$). On commence par le critère suivant de Riemann-intégrabilité.

Théorème 6.3.1 (Lebesgue) *Une fonction f bornée sur $[a,b]$ est Riemann-intégrable ssi l'ensemble de ses points de discontinuité est de mesure de Lebesgue nulle. On dit alors que f est continue λ-presque partout.*

Notons D l'ensemble des points de discontinuité de f. En notant $Osc_x f$ l'oscillation de f en x,

$$Osc_x(f) = \inf_{\substack{I \text{ voisinnage} \\ \text{ouvert de } x}} \Big(\sup \big(|f(u) - f(v)| : u, v \in I \cap [a,b] \big) \Big),$$

on utilise le résultat suivant d'intégrale de Riemann :

Théorème 6.3.2 (Darboux) *Soit f bornée sur $[a,b]$. Elle est Riemann-intégrable ssi pour tout $\varepsilon > 0$, il existe $\alpha > 0$ et une famille d'intervalles ouverts $]c_i, d_i[$, $1 \leq i \leq N$, tels que $\sum_{i=1}^N (d_i - c_i) \leq \varepsilon$ et*

$$D_\alpha = \Big\{ x \in [a,b] : Osc_x(f) \geq \alpha \Big\} \subset \bigcup_{i=1}^N]c_i, d_i[.$$

Intuitivement, on englobe les oscillations trop grandes de f dans des intervalles dont la somme des longueurs est aussi petite qu'on veut.

Démonstration :(Darboux, Th. 6.3.2) Soit f Riemann-intégrable sur $[a,b]$. On se donne $\varepsilon > 0$ et α. D'après la Riemann-intgrabilité (cf. [JCB-Riemann]), il existe une subdivision $d : a = x_0 < x_1 < \cdots < x_n = b$ de $[a,b]$ telle que pour les sommes de Darboux associées, on ait :

$$0 \leq S_f(d) - s_f(d) = \sum_{i=0}^{n-1} (x_{i+1} - x_i)(M_i(f) - m_i(f)) \leq \frac{\varepsilon\alpha}{2}$$

ici $M_i(f) = \sup_{x \in [x_i, x_{i+1}]} f(x)$ et $m_i(f) = \inf_{x \in [x_i, x_{i+1}]} f(x)$. Notons que $(M_i(f) - m_i(f))$ est l'oscillation de f sur $[x_i, x_{i+1}]$. On pose $I_\alpha = \Big\{ i \in \{0, \dots, n-1\} : Osc_f([x_i, x_{i+1}]) \geq \alpha \Big\}$. Pour $i \notin I_\alpha$, on a $Osc_f([x_i, x_{i+1}]) < \alpha$ et pour $x \in]x_i, x_{i+1}[$, on a $Osc_x(f) \leq Osc_f(]x_i, x_{i+1}[) < \alpha$. Il vient $]x_i, x_{i+1}[\cap D_\alpha = \emptyset$ et donc

$$D_\alpha \subset \Big(\bigcup_{i \in I_\alpha}]x_i, x_{i+1}[\Big) \cup \{x_i : 0 \leq i \leq n\}.$$

Pour $i \in I_\alpha$, on a $\alpha \leq Osc_f([x_i, x_{i+1}]) = M_i(f) - m_i(f)$ et donc

$$\sum_{i \in I_\alpha} \alpha(x_{i+1} - x_i) \leq \sum_{i \in I_\alpha} (x_{i+1} - x_i)\big(M_i(f) - m_i(f)\big) \leq S_f(d) - s_f(d) \leq \frac{\varepsilon\alpha}{2}$$

et $\sum_{i \in I_\alpha}(x_{i+1} - x_i) \leq \frac{\varepsilon}{2}$.

Pour $i \in \{0, \dots, n\}$, on pose $u_i = x_i - \varepsilon/(4(n+1))$ et $v_i = x_i + \varepsilon/(4(N+1))$. On a alors $\sum_{i=0}^n (v_i - u_i) = (N+1)\frac{2\varepsilon}{4(n+1)} = \frac{\varepsilon}{2}$ et donc

$$D_\alpha \subset \Big(\bigcup_{i \in I_\alpha}]x_i, x_{i+1}[\Big) \cup \Big(\bigcup_{i=0}^n]u_i, v_i[\Big).$$

avec

$$\sum_{i \in I_\alpha}(x_{i+1} - x_i) + \sum_{i=0}^n (v_i - u_i) \leq \varepsilon$$

ce qui prouve le sens direct du Th. 6.3.2.

Pour la réciproque : soit f bornée sur $[a, b]$ vérifiant la condition du Th. 6.3.2. Soit $\alpha > 0$ et $\varepsilon' > 0$, il existe une famille finie de q intervalles $]c_i, d_i[$, $1 \leq i \leq q$, tels que

$$D_\alpha \subset \Big(\bigcup_{i \in I_\alpha}]c_i, d_i[\Big) \quad \text{et} \quad \sum_{i=1}^q (d_i - c_i) \leq \varepsilon'.$$

On peut supposer les intervalles $]c_i, d_i[$ disjoints (sinon on redécoupe les intervalles) et on peut supposer l'indexation telle que $c_1 < d_1 < c_2 < d_2 < \dots < c_q < d_q$. Si $a < c_1$, on pose $I_0 = [a, c_1]$ puis $I_1 = [d_1, c_2], \dots, I_{q-1} = [d_{q-1}, c_q]$ et $I_q = [d_q, b]$ si $d_q < b$ (si $c_1 \leq a$ alors I_0 n'existe pas ; si $d_q \geq b$ alors I_q n'existe pas).

Comme $D_\alpha \subset \bigcup_{i=1}^q]c_i, d_i[$, on a $D_\alpha \cap I_k = \emptyset$ et donc pour tout $x \in I_k$, on a $Osc_x(f) < \alpha$. La borne supérieure des $Osc_x(f)$ sur I_k est donc inférieure à α et il existe alors une subdivision δ_k de I_k ($y_0^k < y_1^k \dots < y_{n_k}^n$) telle que $Osc_f([y_j^k, y_{j+1}^k]) \leq 2\alpha$ pour tout $0 \leq j \leq n_k - 1$.

Soit alors d la subdivision de $[a, b]$ formée de a, b et des point y_j^k des subdivisions δ_k, $1 \leq k \leq q$, qu'on reordonne en $z_0 = a < z_1 < \dots < z_p = b$. Un segment $[z_s, z_{s+1}]$ de la subdivision d est du type $[y_j^k, y_{j+1}^k]$ ou $[c_i, d_i]$ et on peut regrouper les termes $S_f(d) - s_f(d)$ en les $(d_i - c_i)Osc_f([c_i, d_i])$ et les

$$\sum_{j=0}^{n_k-1} (y_{j+1}^k - y_j^k) Osc_f([y_j^k, y_{j+1}^k]) \leq 2\alpha \sum_{j=0}^{n_k-1} (y_{j+1}^k - y_j^k) \leq 2\alpha\lambda(I_k).$$

On a donc

$$S_f(d) - s_f(d) \leq \sum_{k=0}^{q-1} (2\alpha)\lambda(I_k) + \sum_{i=1}^q (d_i - c_i) Osc_f([c_i, d_i]).$$

Comme les I_k sont des intervalles disjoints et l'oscillation de f sur $[c_i, d_i]$ est majorée par celle de f sur $[a, b]$, on a

$$\begin{aligned} S_f(d) - s_f(d) &\leq (2\alpha)(b-a) + Osc_f([a,b]) \sum_{i=1}^{q} (d_i - c_i) \\ &\leq (2\alpha)(b-a) + \varepsilon' Osc_f([a,b]). \end{aligned}$$

Finalement, étant donné $\eta > 0$, les choix $\alpha \leq \eta/(4(b-a))$ et $\varepsilon' \leq \eta/(2Osc_f([a,b]))$ assurent l'existence d'une subdivision d de $[a, b]$ telle que $S_f(d) - s_f(d) \leq \eta$ ce qui signifie la Riemann-intégrabilité de f. □

Démonstration :(Lebesgue, Th. 6.3.1) Soit f Riemann-intégrable sur $[a, b]$. On note que l'ensemble D de discontinuité est l'ensemble des $x \in [a, b]$ tels que $Osc_x(f) > 0$ qu'on peut encore écrire

$$D = \bigcup_{n\geq 1} D_n, \qquad \text{avec} \qquad D_n = \big\{x \in [a,b] : Osc_x(f) \geq 1/n\big\}. \tag{6.2}$$

Comme f est Riemann-intégrable sur $[a, b]$, le Th. 6.3.2 justifie que pour $\varepsilon > 0$, chaque D_n peut être recouvert par une famille finie d'intervalles ouverts, dont la somme des longueurs est inférieure à $\varepsilon/2^n$:

$$D_n \subset \bigcup_{i=1}^{p_n}]c_{n,i}, d_{n,i}[\qquad \text{avec} \qquad \sum_{i=1}^{p_n} (d_{n,i} - c_{n,i}) \leq \varepsilon/2^n.$$

Ainsi D est couvert par une réunion dénombrable d'ouverts de longueur totale inférieure à ε :

$$D \subset \bigcup_{n\geq 1} \bigcup_{i=1}^{p_n}]c_{n,i}, d_{n,i}[\qquad \text{avec} \qquad \sum_{n\geq 1} \sum_{i=1}^{p_n} (d_{n,i} - c_{n,i}) \leq \sum_{n\geq 1} \varepsilon/2^n = \varepsilon.$$

On a donc $\lambda(D) \leq \varepsilon$ et comme $\varepsilon > 0$ est quelconque, on a $\lambda(D) = 0$.

Réciproquement, si $\lambda(D) = 0$, soit $\varepsilon > 0$ alors il existe des intervalles ouverts $]c_n, d_n[$, $n \geq 1$, tels que

$$D \subset \bigcup_{n\geq 1}]c_n, d_n[, \qquad \sum_{n\geq 1} (d_n - c_n) \leq \varepsilon.$$

Comme D_n donné en (6.2) est fermé de $[a, b]$, D_n est compact. De pus $D_n \subset D$ est couvert par $\bigcup_{n\geq 1}]c_n, d_n[$. On peut donc extraire un recouvrement fini de D_n par des intervalles de longueur inférieure à ε. D'après le Th. 6.3.2, f est Riemann-intégrable. □

Proposition 6.3.1 *Les fonctions Riemann-intégrables $f : [a, b] \to \mathbb{R}$ sont égales λ-presque partout à des fonctions mesurables pour les tribus boréliennes de $[a, b]$ et de $\mathbb{R}$.*

Démonstration : On rappelle qu'étant donnée une subdivision $S = \{a = t_0 \leq \cdots \leq t_n = b\}$ de $[a, b]$, on encadre f Riemann-intégrable par les fonctions en escalier de type Darboux :

$$E^-_{(f,S)}(x) = \sum_{i=0}^{n-1} m_i \, \mathbf{1}_{[t_i,t_{i+1}[}(x), \quad E^+_{(f,S)}(x) = \sum_{i=0}^{n-1} M_i \, \mathbf{1}_{[t_i,t_{i+1}[}(x) \tag{6.3}$$

où $m_i = \inf_{x\in[t_i,t_{i+1}[} f(x)$ et avec $M_i = \sup_{x\in[t_i,t_{i+1}[} f(x)$. On a alors

$$E^-_{(f,S)} \leq f \leq E^+_{(f,S)}$$

et aussi

$$\sup_S E^-_{(f,S)} \leq f \leq \inf_S E^+_{(f,S)}$$

où le sup est pris sur les subdivisions S de $[a, b]$. Puis comme f est Riemann-intégrable, en notant

$$A^-(f, S) = \sum_{i=0}^{n-1} m_i(t_{i+1} - t_i), \quad A^+(f, S) = \sum_{i=0}^{n-1} M_i(t_{i+1} - t_i)$$

ses sommes de Darboux inférieure et supérieure, on a

$$\lim_{\rho(S)\to 0} \int_a^b E^-_{(f,S)}(x)\, dx = \lim_{\rho(S)\to 0} A^-(f, S) = \lim_{\rho(S)\to 0} A^+(f, S) = \lim_{\rho(S)\to 0} \int_a^b E^+_{(f,S)}(x)\, dx.$$

Comme pour une subdivision S fixée, on a

$$E^-_{(f,S)} \leq \sup_S E^-_{(f,S)} \leq \inf_S E^+_{(f,S)} \leq E^+_{(f,S)}$$

en intégrant, il vient

$$\int_a^b E^-_{(f,S)}(x)\, dx \leq \int_a^b \sup_S E^-_{(f,S)}(x)\, dx \leq \int_a^b \inf_S E^+_{(f,S)}(x)\, dx \leq \int_a^b E^+_{(f,S)}(x)\, dx$$

et en passant à la limite

$$\lim_{\rho(S)\to 0} \int_a^b E^-_{(f,S)}(x)\, dx \leq \int_a^b \sup_S E^-_{(f,S)}(x)\, dx \leq \int_a^b \inf_S E^+_{(f,S)}(x)\, dx \leq \lim_{\rho(S)\to 0} \int_a^b E^+_{(f,S)}(x)\, dx.$$

Comme par définition de l'intégrale de Riemann $\int_a^b f(x)\, dx$, les deux côtés de l'inégalité précédente ont des limites égales, il y a égalité partout. On a donc $\int_a^b \sup_S E^-_{(f,S)}(x)\, dx = \int_a^b \inf_S E^+_{(f,S)}(x)\, dx$. Comme en plus $\sup_S E^-_{(f,S)} \leq f \leq \inf_S E^+_{(f,S)}$, cela exige

$$\sup_S E^-_{(f,S)}(x) = f(x) = \inf_S E^+_{(f,S)}(x)$$

en tous les points x de continuité, c'est à dire d'après le théorème précédent de Lebesgue (Th. 6.3.1), égalité λ-presque partout. On en déduit que pour λ-presque chaque $x \in [a,b]$, on a

$$f(x) = \sup_S E^-_{(f,S)}(x) = \inf_S E^+_{(f,S)}(x).$$

Comme on peut prendre le sup et l'inf précédents sur une famille dénombrable de subdivisions $S_n = \big\{a + \frac{k}{n}(b-a) : 0 \le k \le n\big\}$, $n \ge 1$, et comme $E^-_{(f,S_n)}$ et $E^+_{(f,S_n)}$ sont mesurables pour les tribus boréliennes (car en escalier), les fonctions $\sup_n E^-_{(f,S_n)}$ et $\inf_n E^+_{(f,S_n)}$ sont mesurables. Finalement, la fonction f est bien égale λ-p.p. à une fonction mesurable. □

Remarque 6.3.1 Attention si $f = g$ pp avec g mesurable, on n'a pas nécesairement f mesurable sauf si la tribu est complète.

Comparaison des intégrales

Proposition 6.3.2 *Soit $f : [a,b] \to \mathbb{R}$ une fonction Riemann-intégrable. Alors elle est λ-intégrable au sens de Lebesgue et les intégrales de Riemann et de Lebesgue sont égales :*

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

Démonstration : Pour commencer, on suppose pour simplifier f positive. D'après le Th. 6.3.1, l'ensemble des points de discontinuité D de f est de mesure $\lambda(D) = 0$. On a

$$\int_{[a,b]} f\, d\lambda = \int_{[a,b]\cap D^c} f\, d\lambda \le M(b-a) < +\infty$$

où $M = \sup_{x\in[a,b]\cap D^c} f(x) < +\infty$ car f est continue donc bornée sur $[a,b] \cap D^c$, ensemble borné. Avec $S_n = \big\{a + \frac{k}{2^n}(b-a) : 0 \le k \le 2^n\big\}$, on a $\rho(S_n) = 2^{-n} \to 0$, et les partitions S_n, $n \ge 1$, sont emboitées. On a λ-p.p.

$$E^-_{(f,S_n)} \nearrow f \quad \text{et} \quad E^+_{(f,S_n)} \searrow f$$

avec les notations de (6.3). Mais par définition de l'intégrale de Riemann de f, on a

$$\int_a^b f(x)\, dx = \lim_{\rho(S_n)\to 0} A^-(f, S_n) = \lim_{\rho(S_n)\to 0} \int_a^b E^-_{(f,S_n)}(x)\, dx.$$

Comme les partitions S_n, $n \ge 1$, sont emboitées, on a $E^-_{(f,S_n)} \le E^-_{(f,S_{n+1})}$, le théorème de convergence monotone (Th. 4.3.1) assure

$$\int_{[a,b]} f\, d\lambda = \lim_{\rho(S_n)\to 0} \int_{[a,b]} E^-_{(f,S_n)}\, d\lambda.$$

Comme pour les fonctions en escalier, les deux types d'intégrales coïncident $\int_{[a,b]} E^-_{(f,S)}\, d\lambda = \int_a^b E^-_{(f,S)}(x)\, dx$ et l'égalité persiste à la limite :

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

Si f n'est pas positive, alors $|f|$ est encore Riemann-intégrable et on écrit $f = f^+ - f^-$ avec f^+ et f^- ses parties positive et négative, puis on applique ce qui précède à f^+ et f^-. Si f est complexe, on écrit $f = Re(f) + iIm(f)$ et on applique ce qui précède à $Re(f)$ et à $Im(f)$. $\square$

Dans la suite, pour simplifier, on utilisera souvent la notation $\int_a^b f(x)\, dx$ pour l'intégrale $\int_{[a,b]} f\, d\lambda$ par rapport à la mesure λ.

6.4 Cas des intégrales de Riemann impropres

Soit $-\infty \le a < b \le +\infty$ et $f :]a,b[\to \mathbb{C}$ une fonction localement Riemann-intégrable, c'est à dire qu'on suppose que pour tout $a < c < d < b$, on a $\int_c^d |f(x)| dx < +\infty$. Par définition, la fonction f a une intégrale (de Riemann) impropre convergente sur $]a,b[$ ssi

$$\lim_{\substack{\alpha \searrow a \\ \beta \nearrow b}} \int_\alpha^\beta f(x)\, dx$$

existe, c'est à dire ssi pour toutes suites $\alpha_n \searrow a$ et $\beta_n \nearrow b$, on a

$$\int_{\alpha_n}^{\beta_n} f(t)\, dt = \int_{[\alpha_n,\beta_n]} f\, d\lambda \tag{6.4}$$

a une limite, notée $\int_a^b f(x)\, dx$.

Or si f est λ-intégrable sur $]a,b[$ (au sens de Lebesgue), par le théorème de convergence dominée (Th. 5.4.1), on a

$$\int_{[\alpha_n,\beta_n]} |f|\, d\lambda = \int \mathbf{1}_{[\alpha_n,\beta_n]} |f|\, d\lambda \longrightarrow \int \mathbf{1}_{]a,b[} |f|\, d\lambda = \int_{]a,b[} |f| d\lambda, \quad n \to +\infty,$$

car $f\mathbf{1}_{]a,b[}$ est λ-intégrable. Dans ce cas, un passage à la limite dans (6.4) montre que f a une intégrale impropre de Riemann absolument convergente.

Par contre si f n'est pas λ-intégrable (au sens de Lebesgue) sur $]a,b[$, rien ne permet de passer à la limite dans $\int_{]\alpha_n,\beta_n]} f\, d\lambda$ et on ne peut pas considérer l'intégrale de Lebesgue de f sur $]a,b[$ dans ce cas.

Pour **résumer**, il y a deux cas de figure :

— les intégrales impropres convergent absolument, dans ce cas, c'est que $f \in L^1(]a, b[)$ (ie. elles sont λ-intégrables au sens de Lebesgue) ;

— les intégrales sont seulement semi-convergentes et $\int_{]a,b[} f \, d\lambda$ n'est pas défini au sens de Lebesgue.

Exemple : L'intégrale $\int_0^{+\infty} \frac{\sin x}{x} \, dx$ est seulement semi-convergente et n'a pas de sens en tant qu'intégrale de Lebesgue alors qu'en tant qu'intégrale impropre de Riemann, elle vaut $\pi/2$.

Il faudrait définir une notion d'intégrale impropre de Lebesgue pour généraliser les intégrales semi-convergentes. On ne le fera pas dans ce cours.

Chapitre 7

Intégrale multiple

Dans ce chapitre, on considère des intégrales sur des espaces produits, définissant ainsi des intégrales multiples. Pour intégrer sur un espace produit, il est nécessaire de considérer une tribu sur l'espace produit, la plus naturelle est la tribu produit, elle est introduite en Section 7.1. Sur cette tribu, on introduit la mesure produit en Section 7.2. Ces intégrales sur des espaces produits (intégrales multiples) se ramènent à des intégrales sur les espaces facteurs (intégrales simples) grâce aux théorèmes de Fubini en Section 7.3.
On introduit finalement un autre outil de calcul d'intégrales multiples en Section 7.4 avec les changements de variables.

Pour commencer, on rappelle :

Définition 7.0.1 (Produit d'ensembles) *Le produit cartésien de deux ensembles A et B noté $A \times B$ est l'ensemble des couples (a, b) avec $a \in A$ et $b \in B$, ie. $A \times B = \big\{(a, b) : a \in A, b \in B\big\}$.*

Dans ce chapitre, on considère deux espaces mesurables $(X, \mathcal{A})$ et $(Y, \mathcal{B})$ et l'espace produit

$$X \times Y = \big\{(x, y) : x \in X, y \in Y\big\}.$$

7.1 Tribu produit

Définition 7.1.1 (Tribu produit) *La σ-algèbre produit $\mathcal{A} \otimes \mathcal{B}$ est la tribu de $X \times Y$ engendrée par les pavés, c'est à dire les produits de mesurables $A \times B$ pour $A \in \mathcal{A}$, $B \in \mathcal{B}$:*

$$\mathcal{A} \otimes \mathcal{B} = \sigma(\mathcal{P}) \qquad \textit{où} \quad \mathcal{P} = \big\{A \times B : A \in \mathcal{A}, B \in \mathcal{B}\big\}.$$

Remarque 7.1.1 — Le produit de la tribu borélienne sur $\mathbb{R}$ par elle même donne la tribu borélienne sur $\mathbb{R}^2$. En effet $\mathcal{B}(\mathbb{R}^2)$ est engendrée par les produits d'intervalles ouverts $]a, b[\times]c, d[$ qui engendrent aussi la tribu produit $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$. On a donc $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^2)$.

— Plus généralement, la tribu borélienne sur $\mathbb{R}^n$ (engendrée donc par les ouverts de $\mathbb{R}^n$) s'obtient comme le produit de celles de $\mathbb{R}$:

$$\mathcal{B}(\mathbb{R}^n) = \underbrace{\mathcal{B}(\mathbb{R}) \otimes \cdots \otimes \mathcal{B}(\mathbb{R})}_{n \text{ fois}}.$$

— Les ensembles mesurables "de base" de $X \times Y$ sont donc les produits $A \times B$ de mesurables de X et de Y. Lorsque $X = Y = \mathbb{R}$, les plus simples sont mêmes les produits d'intervalles $[a, b] \times [c, d]$. Cependant, il y a des mesurables de $\mathcal{A} \otimes \mathcal{B}$ plus généraux qui ne peuvent se voir comme des produits de mesurables, par exemple le disque unité $D(0,1) = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2\} \in \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ puisque on l'écrit comme l'image réciproque d'un intervalle (donc mesurable) par une fonction mesurable :

$$D(0,1) = F^{-1}([0,1])$$

avec $F(x, y) = x^2 + y^2$ qui est mesurable car continue. De même $K(0,1) = \{(x, y) \in \mathbb{R}^2 : |x| + |y| \leq 1\} \in \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ puisque

$$K(0,1) = G^{-1}([0,1])$$

avec $G(x, y) = |x| + |y|$ qui est mesurable car continue. Noter que $D(0,1)$ est la boule unité pour la norme euclidienne (ou ℓ_2), $K(0,1)$ est celle de la norme ℓ_1 et celle de la norme uniforme ℓ_∞ est $[-1,1] \times [-1,1]$ mesurable car produit de mesurables.

Définition 7.1.2 (Sections) *Soit $E \subset X \times Y$ et $f : X \times Y \to Z$, $x \in X$, $y \in Y$. On définit les sections de E :*

$$E_x = \{y \in Y : (x, y) \in E\} \subset Y, \quad E^y = \{x \in X : (x, y) \in E\} \subset X$$

et celles de f

$$f_x : \left\{ \begin{array}{ccc} (Y, \mathcal{B}) & \to & Z \\ y & \mapsto & f(x, y) \end{array} \right. , \quad f^y : \left\{ \begin{array}{ccc} (X, \mathcal{B}) & \to & Z \\ x & \mapsto & f(x, y). \end{array} \right.$$

Exemples : Avec $X = Y = \mathbb{R}$, on a

$$([1,3] \times [-4,-2])_{x=2} = [-4,-2], \qquad ([1,3] \times [-4,-2])_{x=5} = \emptyset,$$
$$([1,3] \times [-4,-2])^{y=-3} = [1,3], \qquad ([1,3] \times [-4,-2])^{y=0} = \emptyset.$$

Plus généralement, soit $A \in \mathcal{A}$, $B \in \mathcal{B}$ des mesurables de X et de Y, alors

$$(A \times B)_x = \left\{ \begin{array}{lll} B & \text{si} & x \in A, \\ \emptyset & \text{si} & x \notin A \end{array} \right. \qquad (A \times B)^y = \left\{ \begin{array}{lll} A & \text{si} & y \in B, \\ \emptyset & \text{si} & y \notin B. \end{array} \right.$$

Proposition 7.1.1 *Soit $E \in \mathcal{A} \otimes \mathcal{B}$ et $f : (X \times Y, \mathcal{A} \otimes \mathcal{B}) \to (Z, \mathcal{C})$ mesurable alors*

1. *$E_x \in \mathcal{B}$, $E^y \in \mathcal{A}$ pour tout $x \in X$ et $y \in Y$.*
2. *$f_x : (Y, \mathcal{B}) \to (Z, \mathcal{C})$ est mesurable et $f^y : (X, \mathcal{A}) \to (Z, \mathcal{C})$ est mesurable.*

Démonstration : 1) Soit $x \in X$, alors $\mathcal{F} = \{E \subset X \times Y : E_x \in \mathcal{B}\}$ est une tribu car $\mathcal{B}$ en est une (facile à voir) et pour $A \in \mathcal{A}$, $B \in \mathcal{B}, x \in X$, on a $(A \times B) \in \mathcal{F}$ car

$$(A \times B)_x = \begin{cases} B & \text{si } x \in A, \\ \emptyset & \text{sinon.} \end{cases} \tag{7.1}$$

La tribu $\mathcal{F}$ contient donc tous les ensembles du type $A \times B$ pour $A \in \mathcal{A}$ et $B \in \mathcal{B}$. Par définition de la tribu produit, on a $\mathcal{A} \otimes \mathcal{B} \subset \mathcal{F}$. Donc, pour tout $x \in X$ et $E \in \mathcal{A} \otimes \mathcal{B}$, on a $E_x \in \mathcal{B}$. On fait de la même façon pour E^y avec $y \in Y$.

2) On a

$$f_x^{-1}(C) = \{y \in Y : f(x, y) \in C\} = \{y \in Y : (x, y) \in f^{-1}(C)\} = f^{-1}(C)_x.$$

Or $C \in \mathcal{C}$ implique $f^{-1}(C) \in \mathcal{A} \otimes \mathcal{B}$ par mesurabilité de f et $f^{-1}(C)_x \in \mathcal{B}$ par 1). La section f_x est donc mesurable. De même pour f^y, $y \in Y$. □

7.2 Mesure produit

Rappel (mesure σ-finie) : La notion de mesure σ-finie est essentielle dans ce chapitre. On rappelle qu'une mesure μ sur un espace mesurable $(X, \mathcal{A})$ est σ-finie s'il existe une décomposition dénombrable de X en $X = \bigcup_{n \geq 1} X_n$ avec les $X_n \in \mathcal{A}$ de mesure $\mu(X_n) < +\infty$. Un exemple fondamental est la mesure de Lebesgue λ sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ puisque dans ce cas $\mathbb{R} = \bigcup_{n \geq 1} [-n, n]$ avec $\lambda([-n, n]) = 2n < +\infty$. Un autre exemple de mesure σ-finie est une mesure de probabilité $\mathbb{P}$, la mesure est même finie puisque $\mathbb{P}(\Omega) = 1$.

On considère dans toute la suite $(X, \mathcal{A}, \mu)$ et $(Y, \mathcal{B}, \nu)$ deux espaces mesurés avec μ et ν des mesures σ-finies.

Proposition 7.2.1 *Soit $E \in \mathcal{A} \otimes \mathcal{B}$ alors les applications suivantes sont mesurables :*

$$\begin{cases} (X, \mathcal{A}) & \to & [0, +\infty] \\ x & \mapsto & \nu\big(E_x\big) \end{cases}, \quad \begin{cases} (Y, \mathcal{B}) & \to & [0, +\infty] \\ y & \mapsto & \mu\big(E^y\big). \end{cases}$$

Démonstration : On fait la preuve que $x \mapsto \nu(E_x)$ est mesurable pour tout $E \in \mathcal{A} \otimes \mathcal{B}$, l'autre partie étant analogue.

• **1er cas :** on suppose ν mesure finie. On note

$$\mathcal{D} = \big\{E \subset \mathcal{A} \otimes \mathcal{B} : x \mapsto \nu(E_x) \text{ est mesurable}\big\}$$

et

$$\mathcal{P} = \big\{A \times B : A \in \mathcal{A}, B \in \mathcal{B}\big\}. \tag{7.2}$$

On a $\nu\big((A\times B)_x\big)=\nu(B)$ si $x\in A$ et $=0$ si $x\notin A$. On a donc

$$\nu\big((A\times B)_x\big)=\nu(B)\mathbf{1}_A(x)$$

qui est donc mesurable ($\nu(B)<+\infty$), ie. $\mathcal{P}\subset\mathcal{D}$. Observons de plus que $\mathcal{P}$ est stable par intersection (ie. c'est un π-système car $(A_1\times B_1)\cap(A_2\times B_2)=(A_1\cap A_2)\times(B_1\cap B_2)$).
On montre que $\mathcal{D}$ est une classe monotone de $X\times Y$:

— $X\times Y\in\mathcal{P}\subset\mathcal{D}$.

— Si $E,F\in\mathcal{D}$ avec $E\subset F$ alors

$$\nu\big((F\setminus E)_x\big)=\nu\big(F_x\setminus E_x\big)=\nu(F_x)-\nu(E_x) \tag{7.3}$$

est mesurable en tant que différence de fonctions mesurables (car $E,F\in\mathcal{D}$). On a donc $F\setminus E\in\mathcal{D}$. À noter que l'égalité (7.3) utilise ν finie.

— Si $E_n\in\mathcal{D}$, $n\geq 1$, avec $E_n\subset E_{n+1}$ alors comme $(E_n)_x\subset(E_{n+1})_x$, par croissance séquentielle (1.3), on a

$$\nu\Big(\big(\bigcup_{n\geq 1}E_n\big)_x\Big)=\nu\Big(\bigcup_{n\geq 1}(E_n)_x\Big)=\lim_{n\to+\infty}\nu\big((E_n)_x\big)$$

est mesurable en tant que limite de fonctions mesurables (car $E_n\in\mathcal{D}$).

D'après le théorème des classes monotones (Th. 1.5.1), on a $\sigma(\mathcal{P})\subset\mathcal{D}$. Mais par définition, $\sigma(\mathcal{P})=\mathcal{A}\otimes\mathcal{B}$, on a donc $\mathcal{A}\otimes\mathcal{B}\subset\mathcal{D}$, ce qui prouve le premier cas.

• **2ème cas :** La mesure ν n'est pas une mesure finie mais seulement σ-finie. On a donc $Y=\bigcup_{n\geq 1}Y_n$ avec $Y_n\subset Y_{n+1}$ et $\nu(Y_n)<+\infty$ pour tout $n\geq 1$. On pose $\nu_n(B)=\nu(B\cap Y_n)$ pour chaque $n\geq 1$. La mesure ν_n est finie et d'après le premier cas $x\mapsto\nu_n(E_x)$ est mesurable lorsque $E\in\mathcal{A}\otimes\mathcal{B}$.
Par croissance séquentielle, on a $\nu(B)=\lim_{n\to+\infty}\nu_n(B)$ en particulier lorsque $E\in\mathcal{A}\otimes\mathcal{B}$, on a $\nu(E_x)=\lim_{n\to+\infty}\nu_n(E_x)$ mesurable en tant que limite de fonctions mesurables. □

Proposition 7.2.2 *Soient $(X,\mathcal{A},\mu)$ et $(Y,\mathcal{B},\nu)$ deux espaces mesurés avec des mesures μ et ν qui sont σ-finies. Alors il existe une seule mesure (dite mesure produit) notée $\mu\otimes\nu$ sur $\mathcal{A}\otimes\mathcal{B}$ telle que*

$$(\mu\otimes\nu)(A\times B)=\mu(A)\nu(B),\quad A\in\mathcal{A},\ B\in\mathcal{B}. \tag{7.4}$$

De plus pour tout $E\in\mathcal{A}\otimes\mathcal{B}$,

$$(\mu\otimes\nu)(E)=\int_X\nu(E_x)\,d\mu=\int_Y\mu(E^y)\,d\nu.$$

Démonstration : • L'**unicité** découle du théorème des classes monotones (Th. 1.5.1). Si m et m' conviennent alors par définition, m, m' coïncident sur $\mathcal{P}$ en (7.2), stable par intersection finie. Comme μ et ν sont σ-finies, on a

$$X=\bigcup_{n\geq 1}X_n,\quad Y=\bigcup_{n\geq 1}Y_n$$

avec, pour chaque $n \geq 1$, $X_n \subset X_{n+1}$, $\mu(X_n) < +\infty$ et $Y_n \subset Y_{n+1}$, $\nu(Y_n) < +\infty$. Ainsi, par croissance,

$$X \times Y = \bigcup_{n\geq 1} (X_n \times Y_n)$$

avec $(X_n \times Y_n) \subset (X_{n+1} \times Y_{n+1})$ et $m(X_n \times Y_n) = m'(X_n \times Y_n) < +\infty$. Le théorème de Dynkin (Th. 1.5.2) assure sur chaque $X_n \times Y_n$ que $m = m'$. On a donc pour tout $E \in \mathcal{A} \otimes \mathcal{B}$:

$$m\Big(E \cap (X_n \times Y_n)\Big) = m'\Big(E \cap (X_n \times Y_n)\Big). \tag{7.5}$$

Mais par croissance séquentielle (1.3) des mesures, on a

$$m(E) = \lim_{n\to+\infty} m\Big(E \cap (X_n \times Y_n)\Big), \quad m'(E) = \lim_{n\to+\infty} m'\Big(E \cap (X_n \times Y_n)\Big).$$

En passant à la limite dans (7.5), on a alors $m(E) = m'(E)$, ce qui prouve l'unicité.

• **Existence.** D'après le résultat précédent (Prop. 7.2.1) pour $E \in \mathcal{A} \otimes \mathcal{B}$,

$$m_1(E) = \int_X \nu(E_x)\ d\mu, \quad m_2(E) = \int_Y \mu(E^y)\ d\nu$$

ont un sens car $x \mapsto \nu(E_x)$ est $\mathcal{A}$-mesurable donc sa μ-intégrale existe et $y \mapsto \mu(E^y)$ est $\mathcal{B}$-mesurable donc sa ν-intégrale existe. On montre maintenant qu'il s'agit de mesures :

— $m_1(\emptyset) = m_2(\emptyset) = 0$ car

$$m_1(\emptyset) = \int_X \nu(\emptyset_x)\ d\mu = \int_X \nu(\emptyset)\ d\mu = \int_X 0\ d\mu = 0.$$

— Soit $E = \bigcup_{n=1}^{+\infty} E_n$ avec $E_n \in \mathcal{A} \otimes \mathcal{B}$ deux à deux disjoints, on a

$$E_x = \left(\bigcup_{n=1}^{+\infty} E_n\right)_x = \bigcup_{n=1}^{+\infty} (E_n)_x$$

avec les ensembles $(E_n)_x$, $n \geq 1$, deux à deux disjoints. D'où

$$\begin{aligned} m_1(E) &= \int_X \nu(E_x)\ d\mu = \int_X \nu\Big(\bigcup_{n=1}^{+\infty} (E_n)_x\Big)\ d\mu \\ &= \int_X \sum_{n=1}^{+\infty} \nu\big((E_n)_x\big)\ d\mu = \sum_{n=1}^{+\infty} \int_X \nu\big((E_n)_x\big)\ d\mu = \sum_{n=1}^{+\infty} m_1(E_n) \end{aligned}$$

où on a utilisé le théorème de convergence monotone (Th. 4.3.1) pour échanger $\int \sum_{n=1}^{+\infty} = \sum_{n=1}^{+\infty} \int$. De la même façon, on a

$$m_2\Big(\bigcup_{n=1}^{+\infty} E_n\Big) = \sum_{n=1}^{+\infty} m_2(E_n).$$

Les applications m_1 et m_2 sont donc deux mesures. Puis comme avec (7.1) :

$$m_1(A\times B)=\int_X \nu\big((A\times B)_x\big)\ d\mu=\int_A \nu(B)\ d\mu+\int_{A^c}\nu(\emptyset)\ d\mu=\int_A \nu(B)\ d\mu=\mu(B)\mu(A)$$

et de même $m_2(A\times B)=\mu(A)\nu(B)$, ces mesures conviennent pour le théorème.

Par l'unicité vue, la mesure cherchée est $(\mu\otimes\nu)=m_1=m_2$. □

Exemple :(Mesure de Lebesgue sur $\mathbb{R}^n$) Comme $\mathcal{B}(\mathbb{R}^n)=\mathcal{B}(\mathbb{R})^{\otimes n}$, on peut munir $(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n))$ de la mesure produit $\lambda\otimes\cdots\otimes\lambda=\lambda^{\otimes n}$. Il s'agit de la mesure de Lebesgue sur $\mathbb{R}^n$ et on la note λ_n. Elle est invariante par translation et vérifie

$$\lambda_n\Big(\prod_{i=1}^n [a_i,b_i]\Big)=\prod_{i=1}^n (b_i-a_i).$$

7.3 Théorèmes de Fubini

Sous de bonnes conditions, le théorème de Fubini (Th. 7.3.2) permet de permuter les intégrations dans des intégrales multiples. Ainsi les intégrales multiples, ou par rapport à des mesures produits, se ramènent à des intégrales simples emboitées.

Théorème 7.3.1 (Fubini-Tonelli) *Soit $f:(X\times Y,\mathcal{A}\otimes\mathcal{B})\to[0,+\infty]$ une fonction mesurable* **positive**, *avec μ et ν des mesures σ-finies sur $(X,\mathcal{A})$ et $(Y,\mathcal{B})$. Alors*

1. *$x\mapsto\int_Y f_x\ d\nu$ est $\mathcal{A}$-mesurable et $y\mapsto\int_X f^y\ d\mu$ est $\mathcal{B}$-mesurable.*
2. *Puis, on a*

$$\int_{X\times Y} f\ d(\mu\otimes\nu)=\int_X\left(\int_Y f_x\ d\nu\right)d\mu=\int_Y\left(\int_X f^y\ d\mu\right)\ d\nu.$$

Démonstration : • Pour $f=\mathbf{1}_E$, il s'agit de la Prop. 7.2.2 précédente.
• Pour f fonction étagée positive, on obtient 1) et 2) par linéarité grâce au cas précédent des fonctions indicatrices.
• Pour f fonction mesurable positive quelconque, on trouve $(s_n)_{n\geq1}$ une suite croissante de fonctions étagées positives qui converge vers f (Prop. 3.4.1). Mais alors la suite des sections $(s_n)_x$ est étagée, croissante, positive et converge vers la section f_x de f. D'après le théorème de convergence monotone (Th. 4.3.1), on a :

$$\int_Y f_x\ d\nu=\lim_{n\to+\infty}\int_Y (s_n)_x\ d\nu$$

est donc une fonction mesurable car limite de fonctions mesurables. Puis en appliquant trois fois la convergence monotone (Th. 4.3.1, pour $(\mu\otimes\nu)$, pour μ et pour ν) et le résultat déjà vérifié pour les fonctions étagées s_n, on a

$$\begin{aligned}\int f\ d(\mu\otimes\nu) &= \lim_{n\to+\infty}\int s_n\ d(\mu\otimes\nu) = \lim_{n\to+\infty}\int_X\left(\int_Y (s_n)_x\ d\nu\right)\ d\mu\\ &= \int_X \lim_{n\to+\infty}\left(\int_Y (s_n)_x\ d\nu\right)\ d\mu = \int_X\left(\int_Y \lim_{n\to+\infty}(s_n)_x\ d\nu\right)\ d\mu\\ &= \int_X\left(\int_Y f_x\ d\nu\right)\ d\mu.\end{aligned}$$

Le résultat avec les intégrations dans l'autre sens se justifie de la même façon. □

Le théorème de Fubini-Tonelli (Th. 7.3.1) ne s'applique qu'à des fonctions mesurables positives (sans conditions supplémentaires). Pour des fonctions quelconques, on a le résultat suivant :

Théorème 7.3.2 (Fubini) *Soient $(X,\mathcal{A},\mu)$ et $(Y,\mathcal{B},\nu)$ des espaces mesurés σ-finis et $f:(X\times Y,\mathcal{A}\otimes\mathcal{B},\mu\otimes\nu)\to\mathbb{R}$ ou $\mathbb{C}$ $(\mu\otimes\nu)$-***intégrable***. Alors*

1. *Pour μ-presque chaque x, f_x est ν-intégrable et pour ν-presque chaque y, f^y est μ-intégrable.*
2. *La fonction $F(x)=\int_Y f_x\ d\nu$ est $\mathcal{A}$-mesurable et μ-intégrable ; la fonction $G(y)=\int_X f^y\ d\mu$ est $\mathcal{B}$-mesurable et ν-intégrable.*
3. *On a*
$$\int_{X\times Y} f\ d(\mu\otimes\nu)=\int_X F\ d\mu=\int_Y G\ d\nu.$$

On a donc en écrivant les variables d'intégration :

$$\int_{X\times Y} f(x,y)\ (\mu\otimes\nu)(dx,dy)=\int_X\int_Y f(x,y)\ \nu(dy)\mu(dx)=\int_Y\int_X f(x,y)\ \mu(dx)\nu(dy).$$

Démonstration : En appliquant le théorème de Fubini-Tonelli (Th. 7.3.1) à $|f|$, fonction mesurable positive, on a

$$\int_X\left(\int_Y |f_x|\ d\nu\right)d\mu=\int_Y\left(\int_X |f^y|\ d\mu\right)\ d\nu=\int_{X\times Y}|f|\ d(\mu\otimes\nu)<+\infty$$

car f est intégrable pour $(\mu\otimes\nu)$. On en déduit que $x\mapsto\int_Y |f_x|\ d\nu$ est finie μ-pp et $y\mapsto\int_X |f^y|\ d\nu$ est finie ν-pp, car ces fonctions sont positives et d'intégrales finies (Prop. 5.1.1). Cela justifie le point 1).

Pour le 2), d'après Th. 7.3.1, F et G sont mesurables. Puis on écrit $f=(u^+-u^-)+i(v^+-v^-)$ où $u=Re(f)$ et $v=Im(f)$. On a alors

$$I(x)=\int f_x\ d\nu=\int u_x^+\ d\nu-\int u_x^-\ d\nu+i\int v_x^+\ d\nu-i\int v_x^-\ d\nu.$$

Par le théorème de Fubini-Tonelli (Th. 7.3.1), les quatre intégrales sont des fonctions mesurables donc F aussi puis

$$|F(x)| \le \int_Y |f_x|\, d\nu < +\infty$$

pour μ-presque chaque $x \in X$. On a donc f intégrable. On fait de même pour G.

Pour 3), on écrit à nouveau $f = u^+ - u^- + i(v^+ - v^-)$, on utilise la linéarité et le théorème de Fubini-Tonelli (Th. 7.3.1) pour les intégrales de u^+, u^-, v^+, v^- puis on reforme f à la fin, les somme étant licites par intégrabilité. □

Remarque 7.3.1 — En pratique, on raisonne de la façon suivante :

1. On s'assure que f est mesurable (arguments généraux).
2. Pour montrer que f est intégrable, on calcule $\int |f| d(\mu \otimes \nu)$ en appliquant le théorème de Fubini-Tonelli (Th. 7.3.1) à la fonction positive $|f|$:
$$\int |f| d(\mu \otimes \nu) = \int_Y \left(\int_X |f(x,y)| d\mu \right) d\nu = \int_X \left(\int_Y |f(x,y)| d\nu \right) d\mu$$
en choisissant la forme la plus convenable (intégrer d'abord en x ou en y) pour faire le calcul.
3. On applique le théorème de Fubini (Th. 7.3.2).

— Si F est positive, on peut intervertir directement les intégrations (par la version Fubini-Tonelli (Th. 7.3.1) du résultat). Si f ne l'est pas, il faut vérifier l'intégrabilité en calculant l'intégrale de $|f|$ en appliquant par exemple la version Fubini-Tonelli à $|f| > 0$ pour se ramener à des intégrales simples.

— L'utilisation du théorème de Fubini (Th. 7.3.2) permet de ramener de nombreux calculs d'intégrales doubles (ou triples ou plus généralement multiples) à des calculs successifs d'intégrales simples (aussi bien pour des calculs effectifs que pour montrer des convergences d'intégrales).

— Un autre outil essentiel est le changement de variables présenté en Section 7.4.

— Application : grâce au théorème de Fubini (Th. 7.3.2), on peut justifier des interversions $\sum \int = \int \sum$: en effet une somme peut se voir comme une intégrale par rapport à une mesure discrète de comptage et il s'agit bien dès lors d'intervertir deux intégrales.

7.4 Changement de variables

On a vu un résultat de changement de variable abstrait (la formule de transfert, Th. 5.2.1) avec pour seule condition d'avoir un changement de variable $y = \varphi(x)$ mesurable.

Cependant, dans cette formule, la nouvelle mesure $\nu = \mu_f$ n'est pas explicite du tout. Ce résultat est donc essentiellement abstrait et difficile à utiliser pour des calculs explicites. On propose dans cette section, pour les intégrales sur $\mathbb{R}^n$ par rapport à la mesure de Lebesgue λ_n sur $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, des résultats plus explicites, avec des conditions plus restrictives sur le changement de variables (difféomorphisme).

Définition 7.4.1 (Difféomorphisme) *Soit $F : D \subset \mathbb{R}^n \to D' \subset \mathbb{R}^n$ où D et D' sont des ouverts. La fonction F est appelée un difféomorphisme si c'est une bijection de classe C^1 dont la bijection réciproque est aussi de classe C^1.*

Définition 7.4.2 (Jacobien) *La matrice jacobienne d'un changement de variable*

$$y = F(x) \iff (y_1, \dots, y_n) = \big(F_1(x_1, \dots, x_n), \dots, (F_n(x_1, \dots, x_n)\big) = \Big(\frac{\partial F_i}{\partial x_j}(x)\Big)_{1\le i,j\le n}$$

est

$$J_F(x) = J_F(x_1, \dots, x_n) = \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \dots & \frac{\partial F_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial F_n}{\partial x_1} & \dots & \frac{\partial F_n}{\partial x_n} \end{pmatrix}.$$

Le jacobien est le déterminant de la matrice jacobienne $\det J_F(x)$.

La matrice jacobienne est donc la matrice des dérivées partielles.

Rappel : Calculs des déterminants.

— Déterminants d'ordre 2 : $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$,

— Déterminants d'ordre 3 : Règle de Sarrus $\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1b_2c_3 + b_1c_2a_3 + c_1a_2b_3 - a_3b_2c_1 - b_3c_2a_1 - c_3a_2b_1$.

— Déterminants d'ordre quelconque : développements selon une ligne ou une colonne pour se ramener à des déterminants d'ordre inférieur.

La règle de dérivation composée (ie. $(f\circ g)'(x) = f'(g(x))g'(x)$) assure que

$$\frac{\partial (F\circ G)_i}{\partial x_j}(x) = \sum_{k=1}^n \frac{\partial F_i}{\partial x_k}(G(x))\frac{\partial G_k}{\partial x_j}(x)$$

ce qui se résume en

$$J_{F\circ G}(x) = J_F(G(x))J_G(x).$$

Si F est une bijection C^1 de U sur V de jacobienne inversible sur U alors le théorème d'inversion locale assure que F^{-1} est aussi C^1 (et réciproquement). En fait, F est un C^1-difféomorphisme de U sur $V = F(U)$ ssi son jacobien (déterminant de la matrice jacobienne) est non nul sur U ; de plus, la jacobienne de F^{-1} est $J_{F^{-1}}(y) = J_F\big(F^{-1}(y)\big)^{-1}$.

L'objet de la section est la preuve du résultat suivant :

Théorème 7.4.1 (Changement de variables) *Soient U, V des ouverts de $\mathbb{R}^n$ et $\varphi : U \to V$ un difféomorphisme de classe C^1 de jacobien $J_\varphi(x) = \left(\frac{\partial \varphi_i}{\partial x_j}\right)_{1\le i,j\le n}$ où $\varphi = (\varphi_1, \dots, \varphi_n)$. Alors*

1. *pour tout $B \in \mathcal{B}(\mathbb{R}^n)$, $B \subset U$:*
$$\lambda_n\big(\varphi(B)\big) = \int_B |\det J_\varphi|\ d\lambda_n; \tag{7.6}$$
2. *pour toute fonction $f : V \longrightarrow \mathbb{C}$ intégrable :*
$$\int_V f\ d\lambda_n = \int_U (f \circ \varphi)\ |\det J_\varphi|\ d\lambda_n \tag{7.7}$$
(ie. si une des deux intégrales a un sens alors l'autre aussi et il y a égalité).

On commence par prouver 1) pour φ application linéaire en Section 7.4.2 avant de prouver le cas général de 1) en Section 7.4.3. On prouve ensuite 2) en Section 7.4.4.

On présente enfin quelques changements de variable classiques explicites avec les changements en polaires (pour $\mathbb{R}^2$) ou en sphérique (pour $\mathbb{R}^3$) et l'analogue pour $\mathbb{R}^n$.

On commence par un rappel sur le changement de variable pour les intégrales de Riemann en Section 7.4.1.

7.4.1 Rappel : intégrale de Riemann

Soit I un intervalle de $\mathbb{R}$ et $\varphi : I \to \mathbb{R}$ strictement monotone et C^1 tel que φ' ne s'annule pas sur I. Alors on a
$$\int_I f(x)\ dx = \int_{\varphi(I)} f(\varphi^{-1}(y))|(\varphi^{-1})'(y)|\ dy.$$
Pour cela, on pose $y = \varphi(x)$ ou $x = \varphi^{-1}(y)$ et en dérivant on a la relation (formelle) entre dx et dy en dérivant
$$\frac{dx}{dy} = (\varphi^{-1})'(y) \qquad \text{c'est à dire} \qquad dx = (\varphi^{-1})'(y)dy.$$

7.4.2 Changement de variables linéaire

Lemme 7.4.1 1. *Soit $T : \mathbb{R}^n \to \mathbb{R}^n$ linéaire et inversible et B borélien alors*
$$\lambda_n\big(T(B)\big) = |\det T|\ \lambda_n(B).$$
2. *Si $T : \mathbb{R}^n \to \mathbb{R}^n$ est affine inversible, ie. $T(x) = \alpha + T_0(x)$ avec $\alpha \in \mathbb{R}^n$ et T_0 linéaire inversible, alors on a encore pour tout B borélien alors*
$$\lambda_n\big(T(B)\big) = |\det T_0|\ \lambda_n(B).$$

Démonstration : 1) On pose $\lambda'_n(B) = \lambda_n\big(T(B)\big)$. Alors

- λ'_n est une mesure car T est mesurable ;
- λ'_n est finie sur les compacts car si K est compact alors $T(K)$ aussi et donc $\lambda_n(T(K)) < +\infty$ puisque la mesure de Lebesgue est finie sur les compacts ;
- on a $\lambda'_n(B+x) = \lambda_n(T(B+x)) = \lambda_n(T(B)+T(x)) = \lambda_n(T(B)) = \lambda'_n(B)$ puisque λ_n est invariante par translation.

D'après l'unicité de la mesure de Lebesgue (Th. 2.2.1), on a $\lambda'_n = c_T\lambda_n$ avec $c_T = \lambda_n\big(T([0,1[^n)\big)$. Il reste alors à montrer que $c_T = |\det T|$. Pour cela, on utilise que toute matrice T inversible s'écrit sous la forme $T = T_1 \cdots T_k$ où les T_i, $1 \le i \le k$, sont de l'un des trois types suivants :

- matrice de permutation (permute les lignes ou les colonnes selon que l'on multiplie à gauche ou à droite) ;
- $\mathrm{Diag}(1,\dots,1,\alpha,1\dots,1)$ (multiplie une colonne ou une ligne par α) ;
- $\begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ & & & & & \\ 0 & \dots & \dots & \dots & 0 & 1 \end{pmatrix}$ (additionne les deux 1ères lignes ou colonnes).

Étant donnée une matrice, par la multiplication à gauche ou à droite par ces matrices, on peut permuter les lignes ou les colonnes de cette matrice ou en faire des combinaisons linéaires.

Comme T est inversible, il existe $t_{i,j} \neq 0$. En multipliant par une matrice de type 1 puis de type 2, on peut supposer que ce coefficient est $t_{1,1}$ et qu'il vaut 1. En multipliant par des matrices de type 3, on se ramène alors à

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & * & \dots & * \\ & & & \\ 0 & * & \dots & * \end{pmatrix}.$$

Par récurrence, en répétant ce procédé, on arrive à la matrice identité I_n. Cela signifie que T est bien un produit de matrices de type 1, 2, ou 3, ie. $T = T_1 \dots T_k$.

On a alors $T(B) = T_1(T_2(\dots T_k(B)\dots))$ et on déduit $c_T = c_{T_1}\dots c_{T_k}$. La preuve est alors achevée si on montre $c_T = |\det T|$ pour T de type 1, 2, ou 3.

- Si T est de type 1, on a

  $$c_T = \lambda_n\big(T([0,1[^n)\big) = \lambda_n([0,1[^n) = 1 = \det T$$

  car T conserve le cube $[0,1[^n$.
- Si T est de type 2, on a

  $$c_T = \lambda_n\big(T([0,1[^n)\big) = \lambda_n\big([0,1[^{n-1}\times[0,\alpha[\big) = |\alpha| = |\det T|.$$

— Si T est de type 3, par exemple comme donné précédemment, on a $T([0,1]^n) = D \times [0,1]^{n-2}$ avec $D = \{(x,y) \in \mathbb{R}^2 : 0 \le x \le 1, x \le y \le 2\}$, le triangle de sommets $(0,0)$, $(0,1)$, $(1,2)$ et donc d'aire 1, il suit

$$c_T = \lambda_n\big(T([0,1[^n))\big) = \lambda_n\big(D \times [0,1]^{n-2}\big) = \lambda_2\big(D\big)\lambda_{n-2}\big([0,1]^{n-2}\big) = 1 = \det T.$$

2) On a

$$\lambda_n\big(T(B)\big) = \lambda_n\big(\alpha + T_0(B)\big) = \lambda_n\big(T_0(B)\big) = |\det T_0|\ \lambda_n(B).$$

□

Corollaire 7.4.1 *Si $B \in \mathcal{B}(\mathbb{R}^n)$ est inclus dans un hyperplan affine alors $\lambda_n(B) = 0$.*

Démonstration : D'après le théorème de Fubini, on a $\lambda_n\big(\mathbb{R}^{n-1} \times \{0\}\big) = 0$. Ainsi si $B \subset \mathbb{R}^{n-1} \times \{0\}$, on a $\lambda_n(B) = 0$.
Ensuite, si B est inclus dans un hyperplan vectoriel H, il existe T linéaire bijective tel que $H = T\big(\mathbb{R}^{n-1} \times \{0\}\big)$ donc, d'après le Lemme 7.4.1, $\lambda_n(H) = 0$ et $\lambda_n(B) = 0$.
Enfin, si B est inclu dans un hyperplan affine, translaté d'un hyperplan vectoriel, par invariance par translation, on a $\lambda_n(H + a) = \lambda_n(H) = 0$ et donc $\lambda_n(B) = 0$. □

Dans la suite, pour calculer les volumes, on choisit la norme sup ; dans ce cas, la boule de rayon R est le cube d'arête $2R$ et on a $\lambda_n(C^\circ) = \lambda_n(\overline{C})$.

7.4.3 Preuve de 1) dans le Théorème 7.4.1

L'idée repose sur le fait que localement une application C^1 est proche de son application linéaire affine tangente :

$$\varphi(x + h) \approx \varphi(x) + D_x\varphi(h).$$

On découpe alors l'ouvert d'intérêt U en petits cubes $U \approx \bigcup_{C \subset U} C$ sur lesquels on utilise cette approximation. On a alors heuristiquement

$$\begin{aligned}
\lambda_n\big(\varphi(B)\big) &\approx \sum_{C \subset B} \lambda_n\big(\varphi(C)\big) \approx \sum_{C \subset B} \lambda_n\big(D_x\varphi(C)\big) \approx \sum_{C \subset B} \det(D_x\varphi)\ \lambda_n(C) \\
&= \sum_{C \subset B} \int_C \det(D_x\varphi)\ \lambda_n(dx) = \sum_{C \subset B} \int_C |\det J_\varphi|\ d\lambda_n = \int_B |\det J_\varphi|\ d\lambda_n.
\end{aligned}$$

Pour cela, il faut que l'approximation soit uniformément proche de son application linéaire affine tangente. Cela est assuré par

Lemme 7.4.2 (Uniformité) *Soit O un ouvert de $\mathbb{R}^n$ tel que $\overline{O}$ est compact et inclus dans U (donné par l'énoncé du Th. 7.4.1). Alors pour tout $\eta > 0$, il existe $\delta > 0$ tel que $\|x - a\| < \delta$ et $[x, a] \subset \overline{O}$ impliquent*

$$\|\varphi(x) - \varphi(a) - D\varphi(a)(x - a)\| \le \eta \|x - a\|.$$

Démonstration : Comme φ est un difféomorphisme C^1 , $x \mapsto D\varphi(x)$ est une application continue donc uniformément continue sur $\overline{O}$ compact (théorème de Heine).
Soit alors $\theta(y) = \varphi(y) - D\varphi(a)(y)$, on a $D\theta(y) = D\varphi(y) - D\varphi(a)$. L'uniforme continuité associe à $\eta > 0$ un $\delta > 0$ tel que si $x, a \in O$ avec $[x, a] \subset \overline{O}$ et $\|x - a\| < \delta$ alors pour tout $y \in [x, a]$, on a bien $\|y - a\| < \delta$ et $\|D\theta(y)\| = \|D\varphi(y) - D\varphi(a)\| < \eta$. Le théorème des accroissements finis donne alors $\|\theta(x) - \theta(a)\| \leq \eta\|x - a\|$ c'est à dire

$$\|\varphi(x) - \varphi(a) - D\varphi(a)(x - a)\| \leq \eta\|x - a\|.$$

□

Lemme 7.4.3 *Soit O ouvert tel que $\overline{O}$ est compact et inclus dans U. Pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que pour tout $C = \overline{B(a, R)} \subset \overline{O}$ et $R < \delta$, on a*

$$\lambda_n\big(\varphi(C)\big) \leq (1 + \varepsilon)|J_{\varphi(a)}|\lambda_n(C), \quad \textit{et} \quad \lambda_n\big(\varphi(C^\circ)\big) \geq (1 - \varepsilon)|J_{\varphi(a)}|\lambda_n(C). \tag{7.8}$$

Démonstration : Comme φ est C^1, $x \mapsto D\varphi(x)^{-1}$ est inversible et $\|D\varphi(x)\|^{-1} \leq M$ pour $x \in \overline{O}$ compact. On choisit η tel que

$$1 - \varepsilon < (1 - \eta M)^n < (1 + \eta M)^n < 1 + \varepsilon.$$

On applique alors le Lemme 7.4.2 avec ce $\eta > 0$. Pour $C = \overline{B(a, R)}$ et $R < \delta$, on a $C \subset \overline{O}$ et pour tout $x \in C$, $\|x - a\| < \delta$, $[x, a] \subset C$ par convexité, on a

$$\|\varphi(x) - \varphi(a) - D\varphi(a)(x - a)\| \leq \eta\|x - a\| \leq \eta R.$$

En appliquant $D\varphi(a)^{-1}$, il vient :

$$\Big\|D\varphi(a)^{-1} \circ \varphi(x) - D\varphi(a)^{-1} \circ \varphi(a) - (x - a)\Big\| \leq M\eta R \tag{7.9}$$

pour tout $x \in C$, donc

$$D\varphi(a)^{-1} \circ \varphi(C) \subset \overline{B\big(D\varphi(a)^{-1} \circ \varphi(a), (1 + M\eta)R\big)}$$

et donc

$$\lambda_n\Big(D\varphi(a)^{-1} \circ \varphi(C)\Big) \leq (1 + M\eta)^n(2R)^n < (1 + \varepsilon)\lambda_n(C). \tag{7.10}$$

Par ailleurs, par le Lemme 7.4.1, $\lambda_n\Big(D\varphi(a)^{-1} \circ \varphi(C)\Big) = |J_{D\varphi(a)}|^{-1}\lambda_n(\varphi(C))$, ce qui avec (7.10) fournit la première partie de (7.8).

Si $\|x - a\| = R$, d'après (7.9),

$$\Big\|D\varphi(a)^{-1} \circ \varphi(x) - D\varphi(a)^{-1} \circ \varphi(a)\Big\| \geq R - M\eta R.$$

Avec $E = B\big(D\varphi(a)^{-1}\circ\varphi(a),(1-M\eta)R\big)$, on a $E\cap D\varphi(a)^{-1}\big(\varphi(\partial C)\big)=\emptyset$ où on note $\partial C=\{x\in\mathbb{R}^n : \|x-a\|=R\}$. On a

$$E \;=\; \Big(E\cap D\varphi(a)^{-1}\big(\varphi(C^\circ)\big)\Big)\cup\Big(E\cap D\varphi(a)^{-1}\big(\varphi(\partial C)\big)\Big)\cup\Big(E\setminus D\varphi(a)^{-1}\big(\varphi(\overline{C})\big)\Big).$$

Les trois ensembles sont disjoints ; le premier est ouvert car C est ouvert et φ est un difféomorphisme, le deuxième est vide et le troisième est ouvert car $\overline{C}$ étant compact, $D\varphi(a)^{-1}\big(\varphi(\overline{C})\big)$ l'est aussi. On a donc écrit E comme la réunion d'ouverts disjoints. Comme la boule E est connexe, cela exige que l'un des deux ouverts soit vide mais le premier ne l'est pas car contient $D\varphi(a)^{-1}\circ\varphi(a)$. Il vient

$$E\subset D\varphi(a)^{-1}\circ\varphi(C^\circ)$$

et par le Lemme 7.4.1,

$$(2R)^n(1-\eta M)^n=\lambda_n(E)\le\lambda_n\big(D\varphi(a)^{-1}\circ\varphi(C^\circ)\big)=|J\varphi(a)|^{-1}\lambda_n\big(\varphi(C^\circ)\big).$$

On a donc

$$\lambda_n(C)(1-\varepsilon)\le|J\varphi(a)|^{-1}\lambda_n\big(\varphi(C^\circ)\big)$$

c'est à dire la deuxième partie de (7.8). □

Lemme 7.4.4 . *Étant donné $\beta>0$, tout ouvert O de $\mathbb{R}^n$ peut s'écrire comme une union dénombrable de cubes semi-ouverts C_i deux à deux disjoints et d'arêtes inférieures à 2β.*

Démonstration : Soit $k_0\ge1$ tel que $\frac{1}{2^{k_0}}<2\beta$. On considère tous les cubes de bord dyadique $\prod_{i=1}^n\left[\frac{l_i}{2^{k_0}},\frac{l_i+1}{2^{k_0}}\right[$ contenus dans O. Puis pour l'indice k_0+1, on considère les cubes $\prod_{i=1}^n\left[\frac{l_i}{2^{k_0+1}},\frac{l_i+1}{2^{k_0+1}}\right[$ contenus dans O mais pas dans les dyadiques déjà considérés. On continue ainsi le découpage de O. On recouvre ainsi bien l'ouvert O car comme $\mathbb{R}^n\setminus O$ est fermé, pour $x\in O$, $d(x,\mathbb{R}^n\setminus O)>0$. On choisit alors k_1 tel que $\frac{1}{2^{k_1}}<d(x,\mathbb{R}^n\setminus O)$; le cube contenant x construit à l'étape k_1 est contenu dans O. □

Lemme 7.4.5 *Soit O un ouvert contenu dans U tel que $\overline{O}$ est compact inclus dans U alors* (7.6) *est vraie pour O, ie.*

$$\lambda_n\big(\varphi(O)\big)=\int_O|J_\varphi|\,d\lambda_n.$$

Démonstration : L'application $x\mapsto|\det J_\varphi(x)|$ est uniformément continue sur $\overline{O}$, compact (théorème de Heine). Pour $\varepsilon>0$, il existe β tel que $\|x-a\|<\beta$ implique $\big\|\det J_\varphi(x)-\det J_\varphi(a)\big\|<\varepsilon$ et il existe alors δ donné par le Lemme 7.4.3. On applique alors le Lemme 7.4.4 pour $\min(\beta,\delta)$: $O=\bigcup_{i\ge1}C_i$ où les C_i sont semi-ouverts et deux à deux disjoints (on note a_i le centre de C_i). On a

$$\lambda_n\big(\varphi(O)\big)\;=\;\lambda_n\Big(\bigcup_{i\ge1}\varphi(C_i)\Big)=\sum_{i\ge1}\lambda_n(\varphi(C_i))\le\sum_{i\ge1}\lambda_n(\varphi(\overline{C_i}))$$

$$\begin{aligned}
&\leq \sum_{i\geq 1}(1+\varepsilon)|\det J_\varphi(a_i)|\lambda_n(\overline{C_i}) = \sum_{i\geq 1}(1+\varepsilon)|\det J_\varphi(a_i)|\lambda_n(C_i)\\
&\leq (1+\varepsilon)\sum_{i\geq 1}\int_{C_i}|\det J_\varphi(a_i)|\;d\lambda_n\\
&\leq (1+\varepsilon)\sum_{i\geq 1}\left(\int_{C_i}|\det J_\varphi(x)|\;d\lambda_n+\varepsilon\lambda_n(C_i)\right)\\
&\leq (1+\varepsilon)\left(\int_O|\det J_\varphi|\;d\lambda_n+\varepsilon\lambda_n(O)\right).
\end{aligned}$$

Comme $\overline{O}$ est compact, on a $\lambda_n(O)<+\infty$. Alors en faisant $\varepsilon\searrow 0$, on a

$$\lambda_n\big(\varphi(O)\big)\leq\int_O|\det J_\varphi|\;d\lambda_n.$$

En partant de $\lambda_n\big(\varphi(O)\big)\geq\sum_{i\geq 1}\lambda_n\big(\varphi(C_i^\circ)\big)$, on montre de la même façon

$$\lambda_n\big(\varphi(O)\big)\geq(1-\varepsilon)\left(\int_O|\det J_\varphi|\;d\lambda_n+\varepsilon\lambda_n(O)\right)$$

dont on déduit

$$\lambda_n\big(\varphi(O)\big)\geq\int_O|\det J_\varphi|\;d\lambda_n.$$

Finalement,

$$\lambda_n\big(\varphi(O)\big)=\int_O|\det J_\varphi|\;d\lambda_n.$$

□

Lemme 7.4.6 *Pour tout ouvert $O\subset U$ et compact $K\subset U$. Alors (7.6) est vraie pour O et pour K, ie. on a*

$$\lambda_n\big(\varphi(O)\big))=\int_O|\det J_\varphi|\;d\lambda_n,\qquad \textit{et}\qquad \lambda_n\big(\varphi(K)\big))=\int_K|\det J_\varphi|\;d\lambda_n.$$

Démonstration : On écrit $O=\bigcup_{k\geq 1}O_k$ avec $\overline{O_k}$ compact et $O_k\subset O_{k+1}$, par exemple

$$O_k=B(0,k)\cap\big\{x\in\mathbb{R}^n : d(x,\mathbb{R}^n\setminus O)>1/k\big\}.$$

Comme $\varphi(O_k)\subset\varphi(O_{k+1})$, on a d'après le Lemme 7.4.5 :

$$\lambda_n\big(\varphi(O)\big)\;=\;\lambda_n\Big(\bigcup_{k\geq 1}\varphi(O_k)\Big)=\lim_{k\to+\infty}\lambda_n\big(\varphi(O_k)\big)=\lim_{k\to+\infty}\int_U\mathbf{1}_{O_k}\;|\det J_\varphi|\;d\lambda_n$$

$$= \int_U \lim_{k\to+\infty} \mathbf{1}_{O_k} \, |\det J_\varphi| \, d\lambda_n = \int_O |\det J_\varphi| d\lambda_n$$

par convergence monotone (Th. 4.3.1). Si K est compact inclus dans U, un résultat de topologie (lemme de Urysohn) montre qu'on peut intercaler un ouvert O et son adhérence $\overline{O}$ qui est compact, ie.

$$K \subset O \subset \overline{O} \subset U.$$

Comme $K = O \setminus (O \setminus K)$, et $\lambda_n\big(\varphi(O)\big)$ est fini (car λ_n est finie sur $\varphi(\overline{O})$ compact), puis $\varphi(K) = \varphi(O) - \varphi(O \setminus K)$, on a alors, O et $O \setminus K$ étant ouverts

$$\lambda_n\big(\varphi(K)\big) = \lambda_n\big(\varphi(O)\big) - \lambda_n\big(\varphi(O \setminus K)\big) = \int_O |\det J_\varphi| \, d\lambda_n - \int_{O\setminus K} |\det J_\varphi| \, d\lambda_n = \int_K |\det J_\varphi| \, d\lambda_n.$$

□

Proposition 7.4.1 *Si B est un borélien inclus dans U alors (7.6) est vraie pour B, ie.*

$$\lambda_n\big(\varphi(B)\big) = \int_B |\det J_\varphi| \, d\lambda_n.$$

Démonstration : On considère φ borélien, comme φ^{-1} est continue, $\varphi(B) = (\varphi^{-1})^{-1}(B)$ est un borélien.
1er cas : B est borné. Dans ce cas, $\varphi(B)$ est borné et $\lambda_n\big(\varphi(B)\big) < +\infty$. Par régularité intérieure et extérieure, il existe

- K_p compact inclus dans B, $K_p \subset K_{p+1}$, et $\lambda_n\big(\varphi(K_p)\big) \nearrow \lambda_n\big(\varphi(B)\big)$,
- O_p ouvert contenant B, $O_p \supset O_{p+1}$ et $\lambda_n\big(\varphi(O_p)\big) \searrow \lambda_n\big(\varphi(B)\big)$.

Mais alors

$$\begin{aligned}
\lambda_n\big(\varphi(B)\big) &= \lim_{p\to+\infty} \lambda_n\big(\varphi(K_p)\big) = \lim_{p\to+\infty} \int_{K_p} |\det J_\varphi| \, d\lambda_n \\
&= \int_{\bigcup_{p\geq 1} K_p} |\det J_\varphi| \, d\lambda_n \quad \text{(convergence monotone, Th. 4.3.1)}
\end{aligned}$$

et

$$\begin{aligned}
\lambda_n\big(\varphi(B)\big) &= \lim_{p\to+\infty} \lambda_n\big(\varphi(O_p)\big) = \lim_{p\to+\infty} \int_{O_p} |\det J_\varphi| \, d\lambda_n \\
&= \int_{\bigcap_{p\geq 1} O_p} |\det J_\varphi| \, d\lambda_n \quad \text{(convergence dominée).}
\end{aligned}$$

Mais

$$\bigcup_{p\geq 1} K_p \subset B \subset \bigcap_{p\geq 1} O_p$$

assure

$$\int_{\bigcup_{n\geq 1} K_n} |\det J_\varphi| \, d\lambda_n \leq \int_B |\det J_\varphi| \, d\lambda_n \leq \int_{\bigcap_{p\geq 1} O_p} |\det J_\varphi| \, d\lambda_n$$

et finalement

$$\lambda_n\big(\varphi(B)\big) = \int_B |\det J_\varphi| \, d\lambda_n.$$

2ème cas. B n'est contenu dans aucun compact contenu dans U. On écrit alors $U = \bigcup_{k\geq 1} U_k$ avec U_k compact, $U_k \subset U_{k+1}$, et on pose $B_k = B \cap U_k$. D'après le premier cas appliqué au borélien borné B_k,

$$\lambda_n\big(\varphi(B_k)\big) = \int_{B_k} |\det J_\varphi| \, d\lambda_n.$$

Finalement, en utilisant le théorème de convergence monotone (Th. 4.3.1),

$$\lambda_n\big(\varphi(B)\big) = \lim_{k\to+\infty} \lambda_n\big(\varphi(B_k)\big) = \lim_{k\to+\infty} \int_{B_k} |\det J_\varphi| \, d\lambda_n = \int_B |\det J_\varphi| \, d\lambda_n.$$

□

Corollaire 7.4.2 *1. Soit B borélien tel que $\lambda_n(B) = 0$ alors $\lambda_n\big(\varphi(B)\big) = 0$.*

2. Si $A \in \mathcal{L}(\mathbb{R}^n)$ alors $\varphi(A) \in \mathcal{L}(\mathbb{R}^n)$.

Démonstration : 1) Comme $\lambda_n(B) = 0$, on a

$$\lambda_n\big(\varphi(B)\big) = \int_B |\det J_\varphi| \, d\lambda_n = 0.$$

2) Soit $A \in \mathcal{L}(\mathbb{R}^n)$ alors il existe $E, F \in \mathcal{B}(\mathbb{R}^n)$ tels que $E \subset A \subset F$ et $\lambda_n(F \setminus E) = 0$. On a alors $\varphi(E) \subset \varphi(A) \subset \varphi(F)$ avec $\lambda_n\big(\varphi(F) \setminus \varphi(E)\big) = \lambda_n\big(\varphi(F \setminus E)\big) = 0$ d'après le 1). On a donc $\varphi(A) \in \mathcal{L}(\mathbb{R}^n)$ et $\lambda_n\big(\varphi(A)\big) = \int_A |\det J_\varphi| \, d\lambda_n$. □

Cela achève la preuve de 1) dans le Théorème 7.4.1, il reste à voir la preuve de 2).

7.4.4 Preuve de 2) dans le Théorème 7.4.1

Il reste à voir la preuve de (7.7) dans le Théorème 7.4.1, ie. pour $f : V \to [0, +\infty[$ ou $\mathbb{C}$, on a

$$\int_V f \, d\lambda_n = \int_U (f \circ \varphi) \, |\det J_\varphi| \, d\lambda_n.$$

Le résultat (7.7) est déjà acquis pour $f = \mathbf{1}_B$ d'après la Prop. 7.4.1 et par linéarité pour f fonction simple (étagée).

Si f est mesurable positive, il existe $(f_p)_{p\geq 1}$ suite croissante de fonctions étagées avec $f_p \nearrow f$ (Prop. 3.4.1), d'après le théorème de convergence monotone (Th. 4.3.1) :

$$\begin{aligned}\int_V f\ d\lambda_n &= \int_V \lim_{p\to+\infty} f_p\ d\lambda_n = \lim_{p\to+\infty}\int_V f_p\ d\lambda_n \\ &= \lim_{p\to+\infty}\int_U (f_p\circ\varphi)\ |\det J_\varphi|\ d\lambda_n = \int_U (f\circ\varphi)\ |\det J_\varphi|\ d\lambda_n\end{aligned}$$

en utilisant (7.7) déjà prouvé pour les fonctions étagées f_p. Puis si $f : V \to \mathbb{C}$ alors d'après le cas des fonctions positives :

$$\int_V |f|\ d\lambda_n = \int_U |f\circ\varphi|\ |\det J_\varphi|\ d\lambda_n$$

ainsi f est λ_n-intégrable sur V ssi $|f\circ\varphi|\ |\det J_\varphi|$ l'est sur U. Lorsque tel est le cas, on écrit $f = Re(f)^+ - Re(f)^- + iIm(f)^+ - iIm(f)^-$ et on applique le cas déjà prouvé de (7.7) à chaque fonction positive $Re(f)^+$, $Re(f)^-$, $Im(f)^+$, $Im(f)^-$, ce qui prouve le cas général de (7.7) et achève la preuve du Th. 7.4.1.

7.4.5 Coordonnées polaires et sphériques

Un changement de variables utile dans le plan $\mathbb{R}^2$ est le changement de variables en polaire qui consiste à passer de (x, y) représentant des coordonnées cartésiennes dans un repère orthonormé à (r, θ) les coordonnées polaires correspondantes données par

$$(x, y) = \varphi(r, \theta) \iff \varphi : \left\{\begin{array}{lcl} x &=& r\cos\theta \\ y &=& r\sin\theta\end{array}\right., \quad r \in [0, +\infty[,\ \theta \in [0, 2\pi[.$$

Formellement, on remplace alors $dxdy$ par $rdrd\theta$ car le jacobien du changement de variable est r :

$$J_\varphi(r,\theta) = \begin{vmatrix}\cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta\end{vmatrix} = r\cos^2\theta + r\sin^2\theta = r.$$

Ainsi :

$$\begin{aligned}\int\int_{\mathbb{R}^2} f(x,y)\ dxdy &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\ dxdy \\ &= \int\int_{[0,+\infty[\times[0,2\pi[} f(r\cos\theta, r\sin\theta)\ rdrd\theta.\end{aligned}$$

Exemples :

- Normalisation de la loi normale $\displaystyle\int_{-\infty}^{+\infty} e^{-x^2/2}dx = \sqrt{2\pi}$.

Notons $I = \displaystyle\int_{-\infty}^{+\infty} e^{-x^2/2}\ dx$ et montrons que $I^2 = 2\pi$. On a

$$I^2 = \int_{-\infty}^{+\infty} e^{-x^2/2}\ dx \times \int_{-\infty}^{+\infty} e^{-y^2/2}\ dy$$

$$\begin{aligned}
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-x^2/2}e^{-y^2/2}\ dxdy = \int\int_{\mathbb{R}\times\mathbb{R}} e^{-(x^2+y^2)/2}\ dxdy \\
&= \int_0^{2\pi}\int_0^{+\infty} e^{-r^2/2} r\ drd\theta \\
&= \int_0^{2\pi} d\theta \int_0^{+\infty} re^{-r^2/2}\ dr = 2\pi\left[-e^{-r^2/2}\right]_0^{+\infty} = 2\pi
\end{aligned}$$

où on a utilisé le théorème de Fubini (Th. 7.3.2) à la 2ème ligne puis on a fait un changement de variables en polaires à la 3ème ligne.

• Aire d'un disque : $\Delta = \{(x,y) : x^2+y^2 \leq R^2\}$:

$$\lambda_2(B(0,R^2)) = \iint_{B(0,R)} dxdy = \iint_{[0,R]\times[0,2\pi[} rdrd\theta = \int_0^{2\pi} d\theta \int_0^R rdr = 2\pi\left[\frac{r^2}{2}\right]_0^R = \pi R^2.$$

En dimension 3, le changement de variables utile est le changement en coordonnées sphériques donné par

$$(x,y,z) = \phi(r,\theta,\varphi) \quad\Longleftrightarrow\quad \phi : \left\{\begin{array}{lcl} x &=& r\cos\theta\cos\varphi \\ y &=& r\cos\theta\sin\varphi \\ z &=& r\sin\theta \end{array}\right.$$

où $\theta \in [-\pi/2, \pi/2]$ est la latitude, $\varphi \in [0, 2\pi[$ est la longitude et $r \in [0,+\infty[$ la distance à l'origine. Le jacobien du changement de variables est

$$J_\phi(r,\theta,\varphi) = \left|\begin{array}{ccc} \cos\theta\cos\varphi & -r\sin\theta\cos\varphi & -r\cos\theta\sin\varphi \\ \cos\theta\sin\varphi & -r\sin\theta\sin\varphi & r\sin\theta\cos\varphi \\ \sin\theta & r\cos\theta & 0 \end{array}\right| = r^2\cos\theta.$$

Ainsi :

$$\begin{aligned}
&\int\int\int_{\mathbb{R}^3} f(x,y,z)\ dxdydz \\
&= \int\int\int_{[0,+\infty[\times[0,2\pi[\times[-\frac{\pi}{2},\frac{\pi}{2}]} f(r\cos\theta\cos\varphi, r\cos\theta\sin\varphi, r\sin\theta)\ r^2\cos\theta drd\theta d\varphi.
\end{aligned}$$

Ce type de changement de variables (polaires en dimension 2, sphériques en dimension 3) se généralise en dimension n avec

$$\left\{\begin{array}{lcl}
x_1 &=& r\cos\theta_1\cos\theta_2\dots\cos\theta_{n-2}\cos\theta_{n-1}, \\
x_2 &=& r\cos\theta_1\cos\theta_2\dots\cos\theta_{n-2}\sin\theta_{n-1}, \\
x_3 &=& r\cos\theta_1\cos\theta_2\dots\cos\theta_{n-3}\sin\theta_{n-2}, \\
x_4 &=& r\cos\theta_1\cos\theta_2\dots\cos\theta_{n-4}\sin\theta_{n-3}, \\
\dots &=& \dots \\
x_{n-1} &=& r\cos\theta_1\sin\theta_2, \\
x_n &=& r\sin\theta_n.
\end{array}\right.$$

Exemple : Calcul du volume d'une boule euclidienne de rayon R en dimension 3 est

$$\begin{aligned}
\lambda_3(B(0,R)) &= \iiint_{B(0,R)} dxdydz = \iiint_{[0,R]\times[-\pi/2,\pi/2]\times[0,2\pi[} r^2 \cos\theta dr d\theta d\varphi \\
&= \int_0^{2\pi} d\varphi \int_{-\pi/2}^{\pi/2} \cos\theta \; d\theta \int_0^R r^2 \; dr = 2\pi[\sin\theta]_{-\pi/2}^{\pi/2} \left[\frac{r^3}{3}\right]_0^R \\
&= \frac{4}{3}\pi R^3
\end{aligned}$$

où λ_3 désigne la mesure de Lebesgue en dimension 3.

Chapitre 8

Espaces L^p

Dans ce chapitre, on s'intéresse aux fonctions dont les puissances sont intégrables sur un espace $(X, \mathcal{A}, \mu)$. On commence en Section 8.1 par des rappels de convexité sur lesquels la plupart des résultats de ce chapitre sont fondés. On introduit les espaces $\mathcal{L}^p(X, \mathcal{A}, \mu)$ en Section 8.2, puis $L^p(X, \mathcal{A}, \mu)$ en Section 8.3. Les principales inégalités (Hölder, Cauchy-Schwarz, Minkowski) sont prouvées en Section 8.4.

8.1 Convexité

On considère I un intervalle de $\mathbb{R}$. On donne dans cette section des rappels sur les fonctions convexes (pour lesquels on renvoie à des références d'analyse).

Définition 8.1.1 (Convexité) *Une fonction $\varphi : I \to \mathbb{R}$ est convexe ssi pour tout $x, y \in I$ et $t \in [0, 1]$ alors*

$$\varphi\big(tx + (1-t)y\big) \leq t\varphi(x) + (1-t)\varphi(y). \tag{8.1}$$

La fonction φ est dite concave si $-\varphi$ est convexe.

Géométriquement, une fonction φ est convexe si entre deux points x et y, la corde entre $(x, f(x))$ et $(y, f(y))$ est sous le segment entre $(x, f(x))$ et $(y, f(y))$.

Proposition 8.1.1 *Une fonction convexe s'exprime comme le sup des fonctions affines qu'elle majore. De ce fait, une fonction convexe est mesurable.*

Proposition 8.1.2 *Une fonction φ est convexe sur I ssi $\forall x < y < z$ dans I, on a*

$$\frac{\varphi(y) - \varphi(x)}{y - x} \leq \frac{\varphi(z) - \varphi(y)}{z - y}. \tag{8.2}$$

Démonstration : le réel $y \in]x, z[$ s'écrit $\alpha x + (1-\alpha)z$ pour un certain $\alpha \in]0, 1[$ et (8.2) s'écrit alors

$$\frac{\varphi(\alpha x + (1-\alpha)z) - \varphi(x)}{(1-\alpha)(z-x)} \leq \frac{\varphi(z) - \varphi(\alpha x + (1-\alpha)z)}{\alpha(z-x)}.$$

soit

$$\alpha\big(\varphi(\alpha x+(1-\alpha)z)-\varphi(x)\big)\leq(1-\alpha)\big(\varphi(z)-\varphi(\alpha x+(1-\alpha)z)\big),$$

ce qui se simplifie en

$$\varphi(\alpha x+(1-\alpha)z)\leq\alpha\varphi(x)+(1-\alpha)\varphi(z),$$

soit la convexité de φ. □

Exemples :

— Une fonction φ de classe C^1 est convexe ssi sa dérivée φ' est croissante. La condition (8.2) signifie la même chose dans le cas non dérivable (croissance du taux d'accroissements).

— Une fonction φ de classe C^2 est convexe ssi pour tout x on a $\varphi''(x)\geq 0$.

— $f(x)=x^{2n}$ est convexe sur $\mathbb{R}$, $g(x)=x^{2n+1}$ n'est convexe que sur $\mathbb{R}^+$, $h(x)=ax+b$ affine est convexe, exp est convexe sur $\mathbb{R}$, ln est concave (c'est à dire $-\ln$ est convexe).

Le résultat suivant est propre aux mesures de probabilités.

Théorème 8.1.1 (Jensen) *Soit $(X,\mathcal{A},\mu)$ un espace de* **probabilité** *et $f:X\to I\subset\mathbb{R}$ intégrable et $\varphi:I\to\mathbb{R}$ convexe. Alors si $\varphi\circ f$ est intégrable ou positive, on a*

$$\varphi\left(\int_X f\,d\mu\right)\leq\int_X\varphi\circ f\,d\mu.$$

Démonstration : Posons $y=\int f\,d\mu$. Par convexité de f, en interprétant $\int f\,d\mu$ comme un barycentre, on montre que $y\in I$ car f est à valeurs dans I : si $I=(a,b)$, on a $a\leq f\leq b$ et donc $a\mu(X)\leq\int f\,d\mu\leq b\mu(X)$, soit $\int f\,d\mu\in(a,b)=I$. Pour cette valeur de y, par convexité de φ (cf. (8.1)), il existe $\alpha_y\in\mathbb{R}$ tel que

$$\sup_{x<y}\frac{\varphi(y)-\varphi(x)}{y-x}\leq\alpha_y\leq\inf_{z>y}\frac{\varphi(z)-\varphi(y)}{z-y}.$$

En particulier, on a $\varphi(x)\geq\varphi(y)+\alpha_y(x-y)$ pour tout $x\in I$. Ainsi, avec $x=f(u)\in I$, on a $\varphi\circ f(u)\geq\varphi(y)+\alpha_y(f(u)-y)$. Comme φ, convexe, est mesurable, $\varphi\circ f$ l'est aussi. On a ainsi

$$\begin{aligned}\int\varphi\circ f(u)\,\mu(du)&\geq\int\varphi(y)\,\mu(du)+\alpha_y\Big(\int f(u)\,\mu(du)-y\mu(X)\Big)\\&\geq\varphi(y)\mu(X)+\alpha_y\times 0=\varphi\Big(\int f\,d\mu\Big)\end{aligned}$$

car $\mu(X)=1$ (μ est une mesure de probabilité) et par définition $y=\int f\,d\mu$. □

8.2 Espace $\mathcal{L}^p(X, \mathcal{A}, \mu)$

Définition 8.2.1 ((semi)-norme L^p) *Soit $(X, \mathcal{A}, \mu)$ un espace mesuré et $f : X \to \mathbb{C}$ une fonction mesurable.*

— *Pour $p \in [1, +\infty[$, on définit*

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

— *Puis $M \in \mathbb{R}$ est un majorant essentiel de f si $\mu\big(x \in X : |f(x)| > M\big) = 0$.*

— *On définit alors*

$$\|f\|_\infty = \begin{cases} +\infty & \text{si } f \text{ n'a pas de majorant essentiel,} \\ \inf & \text{des majorants essentiels sinon.} \end{cases}$$

Par exemple, sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, $\|\mathbf{1}_\mathbb{Q}\|_\infty = 0$.

Définition 8.2.2 (Espace $\mathcal{L}^p(X, \mathcal{A}, \mu)$) — *On note $\mathcal{L}^p(X, \mathcal{A}, \mu)$ l'ensemble des fonctions mesurables de puissance p-ième μ-intégrable.*

— *On note aussi $\mathcal{L}^\infty(X, \mathcal{A}, \mu)$ l'ensemble des fonctions mesurables (essentiellement) bornées.*

Proposition 8.2.1 *Les ensembles $\mathcal{L}^p(X, \mathcal{A}, \mu)$ et $\mathcal{L}^\infty(X, \mathcal{A}, \mu)$ sont des espaces vectoriels.*

Démonstration : On montre qu'il s'agit d'un sous-espace vectoriel de l'espace des fonctions de X dans $\mathbb{C}$. Pour cela, on montre la stabilité par addition et par multiplication par un scalaire.
Pour $p = +\infty$, soit $M < +\infty$ un majorant essentiel de $|f|$ et $M' < +\infty$ un majorant essentiel de $|g|$. On a

$$\big\{x \in X : |f(x) + g(x)| > M + M'\big\} \subset \big\{x \in X : |f(x)| > M\big\} \cup \big\{x \in X : |g(x)| > M'\big\}$$

donc, par sous-additivité de μ, on a

$$\begin{aligned} & \mu\big(x \in X : |f(x) + g(x)| > M + M'\big) \\ \leq\ & \mu\big(x \in X : |f(x)| > M\big) + \mu\big(x \in X : |g(x)| > M'\big) \\ =\ & 0 + 0 = 0. \end{aligned} \tag{8.3}$$

Ainsi, $M + M' < +\infty$ est un majorant essentiel de $f + g$ et donc $f + g \in L^\infty(X, \mathcal{A}, \mu)$.

Pour $p \in [1, \infty[$: Par convexité de la fonction $x \mapsto x^p$ sur $\mathbb{R}^+$, on a pour $a, b > 0$: $\left(\frac{a+b}{2}\right)^p \leq \frac{1}{2}a^p + \frac{1}{2}b^p$. Avec $a = |f(x)|$ et $b = |g(x)|$, il vient

$$\left(\frac{|f(x) + g(x)|}{2}\right)^p \leq \left(\frac{|f(x)| + |g(x)|}{2}\right)^p \leq \frac{1}{2}|f(x)|^p + \frac{1}{2}|g(x)|^p.$$

En intégrant par rapport à μ, et en multipliant des deux côtés par 2^p, on déduit :

$$\int_X |f+g|^p \, d\mu \leq 2^{p-1} \int_X |f|^p \, d\mu + 2^{p-1} \int_X |g|^p \, d\mu$$

qui est fini lorsque f et g sont dans $\mathcal{L}^p(X, \mathcal{A}, \mu)$. C'est donc que $f+g \in \mathcal{L}^p(X, \mathcal{A}, \mu)$ comme par ailleurs $af \in \mathcal{L}^p(X, \mathcal{A}, \mu)$ quand $f \in \mathcal{L}^p(X, \mathcal{A}, \mu)$, on a bien un espace vectoriel. □

Rappel (Norme et semi-norme) : $\|\cdot\| : E \to \mathbb{R}_+$ est une norme sur un espace vectoriel E si

— $x = 0 \Longleftrightarrow \|x\| = 0$,
— pour tout $x \in E$ et $\lambda \in \mathbb{C}$, $\|\lambda x\| = |\lambda| \|x\|$,
— inégalité triangulaire : pour tout $x, y \in E$ $\|x+y\| \leq \|x\| + \|y\|$.

$\|\cdot\|$ est seulement une semi-norme si le premier point est remplacé par $x = 0 \Rightarrow \|x\| = 0$, c'est à dire un vecteur x peut être de norme $\|x\| = 0$ sans être nul.

En fait, on définit des (semi-)normes à partir de ces quantités $\|\cdot\|_p$.
On vérifie facilement que $\|\alpha f\|_p = |\alpha| \|f\|_p$.
Par contre notons que $\|f\|_p = 0$ n'entraîne pas $f = 0$ mais seulement que $f = 0$ μ-p.p. En effet, par l'inégalité de Markov, on a

$$\mu\big(x \in X : |f(x)| \geq 1/n\big) \leq n^p \int |f|^p \, d\mu = n \|f\|_p^p = 0$$

et alors comme $\{x \in X : |f(x)| > 0\} = \bigcup_{n=1}^{+\infty} \{x \in X : |f(x)| \geq 1/n\}$, par la croissance séquentielle de la mesure μ :

$$\mu\big(x : f(x) \neq 0\big) = \mu\big(x : |f(x)| > 0\big) = \lim_{n \to +\infty} \mu\big(x : |f(x)| \geq 1/n\big) = \lim_{n \to +\infty} 0 = 0.$$

Pour avoir vraiment une norme, il faudrait que $\|f\|_p = 0$ entraîne $f = 0$, c'est à dire $f(x) = 0$ pour tout x et pas seulement pour μ-presque tous les x. Pour remédier à ce problème, on va identifier les fonctions qui ont la même valeur presque partout. Ainsi, $f = 0$ p.p. est identifiée à la fonction nulle (qui vaut zéro toujours).

8.3 Espaces $L^p(X, \mathcal{A}, \mu)$

Formellement, on définit la relation d'équivalence $\sim_\mu$ par $f \sim_\mu g$ ssi il existe $A \in \mathcal{A}$ tel que $\mu(X \setminus A) = 0$ (c'est à dire A^c est μ-négligeable) et pour tout $x \in A$, $f(x) = g(x)$ (c'est à dire f égale g μ-p.p.) et on considère désormais l'espace quotient :

Définition 8.3.1 (Espace L^p) *Pour $p \in [1, \infty]$, on pose*

$$L^p(X, \mathcal{A}, \mu) = \mathcal{L}^p(X, \mathcal{A}, \mu) / \sim_\mu .$$

— $L^p(X, \mathcal{A}, \mu)$ *est l'espace vectoriel des classes d'équivalence des fonctions modulo* $f = g$ μ*-p.p. telles que* $\|f\|_p < +\infty$ *(ie. telles que* $|f|^p$ *est* μ*-intégrable).*

— $L^\infty(X, \mathcal{A}, \mu)$ *est l'espace vectoriel des classes d'équivalence des fonctions modulo* $f = g$ μ*-p.p. telles que* $\|f\|_\infty < +\infty$ *(i.e. bornées* μ*-p.p.).*

Par exemple $\mathbf{1}_\mathbb{Q}$ est identifiée à 0 dans les espaces $L^p\big(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda\big)$. Plus simplement :

Si $\mathcal{A} = \mathcal{P}(X)$ et μ est la mesure de comptage (ou de dénombrement), on note traditionnellement $\ell^p(X)$, ainsi

$$\begin{aligned} \ell^p(\mathbb{N}) &= \Big\{(u_n)_{n\geq 1} \text{ avec } \sum_{n=1}^{+\infty} |u_n|^p < +\infty\Big\} \\ \ell^\infty(\mathbb{N}) &= \big\{\text{suites bornées}\big\}. \end{aligned}$$

Sur les espaces $L^p(X, \mathcal{A}, \mu)$, les semi-normes $\|\cdot\|_p$ deviennent de vraies normes car on a vu que $\|f\|_p = 0$ entraîne $f = 0$ p.p., c'est à dire que f est la classe nulle (f est identifiée à la fonction nulle).

Il reste quand même à voir l'inégalité triangulaire pour $\|\cdot\|_p$. On l'obtiendra par l'inégalité de Minkowski (8.6) ci-dessous. Avant, on définit :

Définition 8.3.2 (Exposants conjugués) *Soient* $p, q \in [1, +\infty]$*. Ce sont des exposants conjugués ssi*

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

Par exemple, 1 et $+\infty$ sont conjugués, 2 est son propre conjugué, 3 a pour conjugué $\frac{3}{2}$.

8.4 Inégalités de convexité

Théorème 8.4.1 (Hölder, Cauchy-Schwarz, Minkowski) *Soient* $(X, \mathcal{A}, \mu)$ *un espace mesuré et* $f, g : X \to \mathbb{C}$ *des fonctions mesurables,* p, q *des exposants conjugués dans* $[1, +\infty]$*. Si* $f \in L^p(X, \mathcal{A}, \mu)$ *et* $g \in L^q(X, \mathcal{A}, \mu)$*, alors on a* $fg \in L^1(X, \mathcal{A}, \mu)$ *avec*

$$\|fg\|_1 = \int_X |fg|\ d\mu \leq \|f\|_p \|g\|_q \quad (\textit{inégalité de } \textbf{Hölder}). \tag{8.4}$$

Pour $p = q = 2$*, il s'agit de l'inégalité de* **Cauchy-Schwarz** *:*

$$\left|\int fg\ d\mu\right|^2 \leq \left(\int |f|^2\ d\mu\right)\ \left(\int |g|^2\ d\mu\right). \tag{8.5}$$

Puis, on a aussi pour $f, g \in L^p(X, \mathcal{A}, \mu)$ *:*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p \quad (\textit{inégalité de } \textbf{Minkowski}). \tag{8.6}$$

Démonstration :(Hölder) • Si p (ou q) $= +\infty$, on a $|f(x)| \leq \|f\|_\infty$ μ-p.p. d'où $|f(x)g(x)| \leq \|f\|_\infty |g(x)|$ p.p., ce qui donne en intégrant

$$\int_X |fg|\ d\mu \leq \|f\|_\infty \int_X g\ d\mu = \|f\|_\infty \|g\|_1.$$

• Soient $p, q \neq +\infty, 1$ avec pour commencer $\|f\|_p = \|g\|_q = 1$. Si $f(x), g(x) \neq 0, +\infty$ on peut écrire $|f(x)|^p = e^u$ et $|g(x)|^q = e^v$ (prendre $u = \ln(|f(x)|^p)$ et $v = \ln(|g(x)|^q)$). Par convexité de exp, on a alors :

$$\exp\left(\frac{1}{p}u + \frac{1}{q}v\right) \leq \frac{1}{p}e^u + \frac{1}{q}e^v,$$

ce qui se reécrit $|f(x)g(x)| \leq \frac{1}{p}|f(x)|^p + \frac{1}{q}|g(x)|^q$. L'inégalité est en fait encore vraie si $f(x)$ ou $g(x)$ vaut 0 ou $+\infty$. En intégrant l'inégalité précédente sur tout X, on a alors

$$\|fg\|_1 = \int_X |fg|\ d\mu \leq \frac{1}{p}\int_X |f|^p\ d\mu + \frac{1}{q}\int_X |g|^q\ d\mu = \frac{1}{p}\|f\|_p^p + \frac{1}{q}\|g\|_q^q = \frac{1}{p} + \frac{1}{q} = 1$$

car d'abord $\|f\|_p = \|g\|_q = 1$ puis ensuite parce que p, q sont conjugués. Ce qui prouve l'inégalité de Hölder dans ce cas.

• Si $p, q \neq 1, +\infty$ et si $\|f\|_p$ et $\|g\|_q \neq 0, +\infty$ alors on pose

$$\widetilde{f} = f/\|f\|_p, \quad \widetilde{g} = g/\|g\|_q.$$

Comme $\|\widetilde{f}\|_p = \|\widetilde{g}\|_q = 1$, le cas précédent donne :

$$\|\widetilde{f}\widetilde{g}\|_1 \leq 1 \Longleftrightarrow \|fg\|_1 \leq \|f\|_p \|g\|_q.$$

• Si $\|f\|_p = 0$ alors $f = 0$ μ-p.p. et donc $fg = 0$ μ-p.p. si bien que $\int_X |fg|\ d\mu = \|fg\|_1 = 0$ et l'inégalité cherchée se réécrit $0 \leq 0$, ce qui est vrai.

• Si $\|f\|_p = +\infty$ alors il suffit de considérer le cas $\|g\|_q \neq 0$ sinon on se ramène au cas précédent (avec g à la place de f). Mais dans ce cas, l'inégalité devient une majoration par $+\infty$ ce qui est nécessairement vrai. □

Démonstration :(Cauchy-Schwarz) On applique l'inégalité de Hölder (8.4) avec $p = q = 2$:

$$\left|\int fg d\mu\right|^2 \leq \left(\int |fg| d\mu\right)^2 = \|fg\|_1^2 \leq \|f\|_2^2 \|g\|_2^2 = \left(\int f^2 d\mu\right)\left(\int g^2 d\mu\right).$$

□

Démonstration :(Minkowski) Pour $p=1$, soit $f,g\in L^1(X,\mathcal{A},\mu)$, en intégrant $|f+g|\leq |f|+|g|$, on a

$$\int_X |f+g|\,d\mu \leq \int_X |f|\,d\mu + \int |g|\,d\mu \quad\Longleftrightarrow\quad \|f+g\|_1\leq \|f\|_1+\|g\|_1.$$

Pour $p=+\infty$, soit $f,g\in L^\infty(X,\mathcal{A},\mu)$. Si $M>\|f\|_\infty$ et $M'>\|g\|_\infty$ alors on a vu précédemment en (8.3) que $M+M'$ est un majorant essentiel de $f+g$ mais alors par définition de $\|f+g\|_\infty$, on a :

$$\|f+g\|_\infty \leq M+M'. \tag{8.7}$$

En faisant $M\searrow \|f\|_\infty$ et $M'\searrow \|g\|_\infty$, on déduit de (8.7) : $\|f+g\|_\infty\leq \|f\|_\infty+\|g\|_\infty$.
Maintenant on suppose $p\in]1,+\infty[$ et on commence par écrire

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

et

$$|f+g|^p\leq |f|\ |f+g|^{p-1}+|g|\ |f+g|^{p-1}.$$

Soit q conjugué de p alors d'après l'inégalité de Hölder (8.4)

$$\begin{aligned}
&\int |f+g|^p\,d\mu = \int |f|\ |f+g|^{p-1}\,d\mu + \int |g|\ |f+g|^{p-1}\,d\mu\\
&\leq \left(\int |f|^p\,d\mu\right)^{1/p}\left(\int |f+g|^{(p-1)q}\,d\mu\right)^{1/q} + \left(\int |g|^p\,d\mu\right)^{1/p}\left(\int |f+g|^{(p-1)q}d\mu\right)^{1/q}\\
&\leq \left[\left(\int |f|^p\,d\mu\right)^{1/p}+\left(\int |g|^p\,d\mu\right)^{1/p}\right]\left(\int |f+g|^{(p-1)q}\,d\mu\right)^{1/q}.
\end{aligned}$$

Comme $(p-1)q=p$ et $1/q=(p-1)/p$, on a

$$\begin{aligned}
\int |f+g|^p\,d\mu &\leq \left[\left(\int |f|^p\,d\mu\right)^{1/p}+\left(\int |g|^p\,d\mu\right)^{1/p}\right]\left(\int |f+g|^p\,d\mu\right)^{\frac{p-1}{p}}\\
\|f+g\|_p^p &\leq [\|f\|_p+\|g\|_p]\ \|f+g\|_p^{p-1}.
\end{aligned} \tag{8.8}$$

• Si $\|f+g\|_p\neq 0,+\infty$, l'inégalité de Minkowski (8.6) suit en simplifiant par $\|f+g\|_p^{p-1}$ dans (8.8).

• Si $\|f+g\|_p=0$, l'inégalité de Minkowski (8.6) est immédiate.

• Si $\|f+g\|_p=+\infty$, par convexité de la fonction $x\in\mathbb{R}_+\mapsto x^p$ pour $p\geq 1$, on a

$$\left(\frac{|f|+|g|}{2}\right)^p\leq \frac{1}{2}|f|^p+\frac{1}{2}|g|^p \quad\Longleftrightarrow\quad (|f|+|g|)^p\leq 2^{p-1}|f|^p+2^{p-1}|g|^p.$$

Comme $|f+g| \leq |f|+|g|$, on a donc $\displaystyle\int |f+g|^p\,d\mu \leq 2^{p-1}\int |f|^p\,d\mu + 2^{p-1}\int |g|^p\,d\mu$. Mais alors si $\int |f+g|^p\,d\mu$ est infini, $2^{p-1}\int |f|^p\,d\mu + 2^{p-1}\int |g|^p d\mu$ l'est aussi donc $\int |f|^p\,d\mu$ ou $\int |g|^p\,d\mu$ l'est si bien que l'inégalité de Minkowski (8.6) devient $+\infty \leq +\infty$, ce qui est encore vrai. □

Corollaire 8.4.1 *L'application $f \mapsto \|f\|_p$ est une norme sur $L^p(X,\mathcal{A},\mu)$.*

Démonstration : Il ne manquait plus que l'inégalité triangulaire pour prouver que $\|\cdot\|_p$ est une norme : c'est précisément l'inégalité de Minkowski (8.6). □

L'espace $L^p(X,\mathcal{A},\mu)$ est donc un espace vectoriel normé. On a mieux avec le résultat suivant :

Théorème 8.4.2 (Riesz-Fisher) *Muni de la norme $\|\cdot\|_p$, l'espace vectoriel $L^p(X,\mathcal{A},\mu)$ est complet, il s'agit donc d'un espace de Banach.*

Démonstration : Soit $(f_n)_{n\geq 1}$ une suite de Cauchy pour $\|\cdot\|_p$, on montre qu'elle converge dans $L^p(X,\mathcal{A},\mu)$.

Si $p=+\infty$, soit $A_{n,m} = \big\{x \in X : |f_n(x)-f_m(x)| > \|f_n-f_m\|_\infty\big\}$. Alors $\mu(A_{n,m})=0$ et donc $\mu\Big(\bigcup_{n,m\geq 1} A_{n,m}\Big) = 0$ car

$$\mu\Big(\bigcup_{n,m\geq 1} A_{n,m}\Big) \leq \sum_{n,m\geq 1} \mu(A_{n,m}) = 0.$$

Pour $x \in E = X\setminus\bigcup_{n,m\geq 1} A_{n,m}$, on a donc pour tout $n,m\geq 1$, $|f_n(x)-f_m(x)| \leq \|f_n-f_m\|_\infty$. Comme $(f_n)_{n\geq 1}$ est de Cauchy pour $\|\cdot\|_\infty$, pour tout $\varepsilon>0$, il existe $N\geq 1$ tel que si $n,m\geq N$, on a $\|f_n-f_m\|_\infty<\varepsilon$ et a fortiori $|f_n(x)-f_m(x)|<\varepsilon$ pour $x\in E$.
Ainsi $(f_n(x))_{n\geq 1}$ est une suite de Cauchy de $\mathbb{C}$ donc convergente vers un certain $f(x)\in\mathbb{C}$. On construit ainsi $f(x)$ pour tout $x\in E$ et on complète la définition de f sur tout X en posant (par exemple) $f(x)=0$ lorsque $x\notin E$.
Pour tout $\varepsilon>0$, il existe $N\geq 1$ tel que si $n,m\geq N$, pour tout $x\in E$, $|f_n(x)-f_m(x)|\leq\varepsilon$. À la limite quand $m\to+\infty$, on a $|f_n(x)-f(x)|\leq\varepsilon$.
Comme ceci est valable pour tout $x\in E$ avec $\mu(E^c)=0$, on a $\|f_n-f\|_\infty\leq\varepsilon$. Ainsi $f\in L^\infty(X,\mathcal{A},\mu)$ puisque $\|f\|_\infty \leq \|f_n\|_\infty + \|f_n-f\|_\infty < +\infty$ et $f_n\to f$ dans $L^\infty(X,\mathcal{A},\mu)$, ce qui prouve que $L^\infty(X,\mathcal{A},\mu)$ est complet pour $\|\cdot\|_\infty$.

On considère maintenant $p\in[1,+\infty[$. La condition de Cauchy s'écrit : pour tout $\varepsilon>0$, il existe $N\geq 1$ tel que si $n,m\geq N$, on a $\|f_n-f_m\|_p<\varepsilon$. On construit alors par récurrence $n_1<n_2<\cdots<n_k<\cdots$ tels que pour tout $i\geq 1$, on ait

$$\|f_{n_i}-f_{n_{i+1}}\|_p \leq \frac{1}{2^i} \tag{8.9}$$

et on pose $g_k = \sum_{i=1}^k |f_{n_i} - f_{n_{i+1}}|$ et

$$g(x) = \lim_{k\to+\infty} g_k(x) = \sum_{i=1}^{+\infty} |f_{n_i}(x) - f_{n_{i+1}}(x)| \in [0, +\infty].$$

Par l'inégalité de Minkowski (8.6), on a

$$\|g_k\|_p \leq \sum_{i=1}^k \|f_{n_i} - f_{n_{i+1}}\|_p \leq \sum_{i=1}^k \frac{1}{2^i} \leq 1.$$

Comme $g(x)^p$ est la limite croissante de $g_k(x)^p$, le théorème de convergence monotone (Th. 4.3.1) assure

$$\int_X g^p \, d\mu = \lim_{k\to+\infty} \int_X g_k^p \, d\mu = \lim_{k\to+\infty} \|g_k\|_p^p \leq 1.$$

Comme $\int_X g^p \, d\mu \leq 1$, on a $g(x)$ fini sur $E \in \mathcal{A}$ avec $\mu(E^c) = 0$ (inégalité de Markov, Prop. 4.2.1).
Mais alors pour $x \in E$, on a $\sum_{i=1}^{+\infty} |f_{n_i}(x) - f_{n_{i+1}}(x)| < +\infty$. Comme $\mathbb{C}$ est complet la convergence absolue de la série $\sum_{i=1}^{+\infty} |f_{n_i}(x) - f_{n_{i+1}}(x)|$ implique sa convergence simple, ie. $\sum_{i=1}^{+\infty} \big(f_{n_i}(x) - f_{n_{i+1}}(x)\big)$ converge. Posons alors

$$f(x) = f_{n_1}(x) + \sum_{i=1}^{+\infty} \big(f_{n_i}(x) - f_{n_{i+1}}(x)\big)$$

lorsque $x \in E$ et (par exemple) $f(x) = 0$ lorsque $x \notin E$. Comme $f_{n_1}(x) + \sum_{i=1}^k f_{n_i}(x) - f_{n_{i+1}}(x) = f_{n_k}(x)$, on a alors $f(x) = \lim_{k\to+\infty} f_{n_k}(x)$ lorsque $x \in E$, ie.

$$f_{n_k} \text{ converge vers } f \ \mu\text{-p.p.} \tag{8.10}$$

Puis par la condition de Cauchy, pour tout $\varepsilon > 0$, il existe $N \geq 1$ tel que si $n, m \geq N$, on a

$$\int_X |f_n - f_m|^p \, d\mu < \varepsilon^p.$$

Avec $m = n_k$ et k assez grand pour que $m \geq N$, le lemme de Fatou (Th. 4.4.1) donne : pour tout $\varepsilon > 0$, il existe N tel que pour $n \geq N$, on a

$$\int_X \liminf_{k\to+\infty} |f_n - f_{n_k}|^p \, d\mu \leq \liminf_{k\to+\infty} \int_X |f_n - f_{n_k}|^p \, d\mu \leq \varepsilon^p.$$

Comme $f_{n_k} \to f$ μ-p.p., on a finalement

$$\int_X |f_n - f|^p \, d\mu \leq \varepsilon^p$$

ie. $\|f_n - f\|_p \leq \varepsilon$. On a donc $f \in L^p(X, \mathcal{A}, \mu)$ puisque $\|f\|_p \leq \|f_n - f\|_p + \|f\|_p < +\infty$ et $f_n \to f$ dans $L^p(X, \mathcal{A}, \mu)$, ce qui prouve que $L^p(X, \mathcal{A}, \mu)$ est complet pour $\|\cdot\|_p$. $\square$

Au cours de la preuve du théorème de Riesz-Fisher (Th. 8.4.2), on a établi le résultat suivant (cf. (8.10)) qui donne un lien entre la convergence d'une suite de fonctions $(f_n)_{n\geq 1}$ vers f dans $L^p(X, \mathcal{A}, \mu)$ et la convergence μ-presque partout (pour une sous-suite).

Théorème 8.4.3 (Convergences L^p et p.p.) *Si $f_n \to f$ dans $L^p(X, \mathcal{A}, \mu)$, c'est à dire $\lim_{n\to+\infty} \|f_n - f\|_p = 0$ alors il existe une sous-suite $(f_{n_k})_{k\geq 1}$ de $(f_n)_{n\geq 1}$ qui converge μ-p.p. vers f.*

Comparaison des espaces $L^p(X, \mathcal{A}, \mu)$

En général, il n'y a pas de relations d'inclusion entre les espaces $L^p(X, \mathcal{A}, \mu)$ pour différents exposants p. Par exemple, considérons $(X, \mathcal{A}, \mu) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ et les espaces $L^1\big(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda\big)$, $L^2\big(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda\big)$ associés. Alors avec

$$f(x) = \frac{1}{\sqrt{x}}\mathbf{1}_{]0,1]}(x), \quad g(x) = \frac{1}{x}\mathbf{1}_{[1,+\infty[}(x),$$

on a $f \in L^1\big(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda\big)$ mais $f \notin L^2\big(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda\big)$ puisque

$$\begin{aligned} \|f\|_1 &= \int |f(x)|d\lambda = \int_0^1 \frac{dx}{\sqrt{x}} = \frac{1}{2} \\ \|f\|_2^2 &= \int |f(x)|^2 d\lambda = \int_0^1 \frac{dx}{x} = +\infty \end{aligned}$$

alors que $g \in L^2\big(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda\big)$ mais $g \notin L^1\big(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda\big)$ puisque

$$\begin{aligned} \|g\|_1 &= \int |g(x)|d\lambda = \int_1^{+\infty} \frac{dx}{x} = +\infty \\ \|g\|_2^2 &= \int |g(x)|^2 d\lambda = \int_1^{+\infty} \frac{dx}{x^2} = 1. \end{aligned}$$

On n'a donc aucune inclusion entre les espaces $L^1\big(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda\big)$ et $L^2\big(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda\big)$. L'exemple se généralise pour les espaces $L^p\big(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda\big)$ et $L^q\big(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda\big)$.

Par contre si μ est une mesure **finie** (typiquement une mesure de probabilité), les espaces $L^p(X, \mathcal{A}, \mu)$ sont ordonnés pour l'inclusion :

Théorème 8.4.4 *Soit $(X, \mathcal{A}, \mu)$ un espace mesuré fini (c'est à dire $\mu(X) < +\infty$) alors pour $p \geq q$:*

$$L^p\big(X, \mathcal{A}, \mu\big) \subset L^q\big(X, \mathcal{A}, \mu\big).$$

Démonstration : • D'abord si $p = +\infty$: si $f \in L^\infty(X, \mathcal{A}, \mu)$ alors $f \in L^q(X, \mathcal{A}, \mu)$, en effet $|f(x)| \leq \|f\|_\infty$ pour μ-presque chaque x et donc,

$$\|f\|_q = \left(\int |f|^q \, d\mu\right)^{1/q} \leq \left(\int \|f\|_\infty^q \, d\mu\right)^{1/q} \leq \left(\|f\|_\infty^q \int d\mu\right)^{1/q} = \|f\|_\infty \mu(X)^{1/q} < +\infty.$$

Donc $f \in L^q(X, \mathcal{A}, \mu)$.

• Puis si $f \in L^p(X, \mathcal{A}, \mu)$ et $q \leq p$, alors $f \in L^q(X, \mathcal{A}, \mu)$, en effet comme $\frac{q}{p} + \frac{p-q}{p} = 1$, $\alpha = p/q$ et $\beta = \frac{p}{p-q}$ sont conjugués, l'inégalité d'Hölder (8.4) pour ces exposants α, β donne alors

$$\|f\|_q^q = \left(\int |f|^q \times 1 \, d\mu\right) \leq \left(\int |f|^{q \times \frac{p}{q}} \, d\mu\right)^{q/p} \left(\int 1^{p/(p-q)} d\mu\right)^{1-q/p}.$$

D'où

$$\begin{aligned}
\|f\|_q^q &\leq \mu(X)^{1-q/p} \left(\int |f|^p \, d\mu\right)^{q/p}, \\
\|f\|_q &\leq \mu(X)^{1/q-1/p} \left(\int |f|^p \, d\mu\right)^{1/p} < +\infty.
\end{aligned}$$

□

Remarque 8.4.1 Plus précisément, le Théorème 8.4.4 montre une inclusion topologique $L^p(X, \mathcal{A}, \mu) \hookrightarrow L^q(X, \mathcal{A}, \mu)$ car l'inclusion canonique $f \in L^p(X, \mathcal{A}, \mu) \mapsto f \in L^q(X, \mathcal{A}, \mu)$ est continue.

Chapitre 9

Convolution

Dans ce chapitre, on commence en Section 9.1 par présenter l'opération de convolution entre deux fonctions. On donne des résultats de dérivabilité des convolutions en Section 9.3. L'approximation et la régularisation de fonctions sont des applications importantes de la convolution, on donne de tels résultats en Section 9.4 (Stone-Weierstrass, densité dans les L^p, Fejér).

Sauf mention contraire, on considère dans ce chapitre $(X, \mathcal{A}, \mu) = \big(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \lambda_n\big)$ et toutes les fonctions considérées sont mesurables de $\mathbb{R}^n$ dans $\mathbb{R}$.

9.1 Définition et propriétés

Définition 9.1.1 *La convolution de deux fonctions f et g réelles est la fonction $f * g$ définie sur $\mathbb{R}^n$ et donnée par*

$$(f * g)(x) = \int_{\mathbb{R}^n} f(x - y)g(y)\, dy.$$

On parle encore de la convolée $f * g$ de f et de g.

Remarque 9.1.1 Attention, il n'est pas clair pour quels $x \in \mathbb{R}^n$ la fonction $f * g$ est bien définie. Certains résultats suivent pour donner des conditions d'existence de $(f * g)(x)$.

Proposition 9.1.1 *Soient f, g des fonctions mesurables, alors lorsque c'est bien défini :*

— $(f * g)(x) = (g * f)(x)$

— $(f, g) \mapsto f * g$ *est bilinéaire.*

Démonstration : • Pour le premier point, le changement de variable $y \to z = x - y$ donne

$$(f * g)(x) = \int_{\mathbb{R}^n} f(x - y)g(y)\, dy = \int_{\mathbb{R}^n} f(z)g(x - z)|(-1)^n|\, dz = (g * f)(x).$$

• Pour le second point, par linéarité de l'intégrale, il est facile de voir que

$$(a_1 f_1 + a_2 f_2) * g\,(x) = a_1(f_1 * g)(x) + a_2(f_2 * g)(x).$$

La linéarité par rapport à g se montre de la même façon ou en utilisant celle par rapport à f combinée avec la symétrie de $*$ vue au premier point. □

Proposition 9.1.2 *Si f est nulle hors de A et g nulle hors de B, alors $(f * g)(x)$ existe et vaut 0 en dehors de $A + B = \{a + b : a \in A, b \in B\}$.*

Démonstration : L'intégrale définissant $(f * g)(x)$ est bien définie si l'intégrale

$$\int_{\mathbb{R}^n} |f(y)|\, |g(x-y)|\, dy$$

est convergente. Or si $x = y + (x - y) \notin A + B$, alors soit $y \notin A$ et $f(y) = 0$, soit $y \in A$ mais alors $x - y \notin B$ et $g(x - y) = 0$. Donc dans tous les cas pour $x \notin A + B$, on a $f(y)g(x - y) = 0$ pour tout y et en intégrant, il vient $\int_{\mathbb{R}^n} |f(y)|\, |g(x - y)|\, dy = 0$ si bien que $(f * g)(x)$ est bien définie et vaut *a fortiori* 0 puisque

$$\left| \int_{\mathbb{R}^n} f(y)\ g(x-y)\ dy \right| \leq \int_{\mathbb{R}^n} |f(y)|\, |g(x-y)|\, dy.$$

□

Proposition 9.1.3 *Soient f et g bornées et nulles respectivement hors de A compact et de B alors $(f * g)(x)$ existe pour tout x et $(f * g)(x) = 0$ pour $x \notin A + B$.*

Démonstration : Pour $x \notin A + B$, c'est le résultat précédent qui s'applique. Il ne reste qu'à montrer l'existence de l'intégrale définissant $(f * g)(x)$ lorsque $x \in A + B$. Pour cela, on a

$$\int_{\mathbb{R}^n} |f(y)|\, |g(x-y)|\, dy = \int_A |f(y)|\, |g(x-y)|\, dy < \lambda(A) \|f\|_\infty \|g\|_\infty < +\infty.$$

On en déduit que $y \mapsto f(y)g(x - y)$ est intégrable pour tout $x \in \mathbb{R}^n$. La fonction $f * g$ est donc bien définie sur tout $\mathbb{R}^n$. □

9.2 Normes des convolutions

Les résultats suivants donnent des inégalités de type Hölder (8.4) pour les convolutions, ce faisant on obtient aussi des conditions d'existence de la convolée $f * g$.

Proposition 9.2.1 *Soient $f, g \in L^1(\mathbb{R}^n)$, alors $f * g$ est définie p.p. et*

$$\|f * g\|_1 \le \|f\|_1 \|g\|_1.$$

Démonstration : En appliquant le théorème de Fubini-Tonelli (Th. 7.3.1), on a

$$\begin{aligned}
\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |f(y)|\,|g(x-y)|\,dy\,dx &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |f(y)|\,|g(x-y)| dx\,dy \\
&= \int_{\mathbb{R}^n} |f(y)| \left(\int_{\mathbb{R}^n} |g(x-y)|\,dx \right) dy \\
&= \int_{\mathbb{R}^n} |f(y)| \left(\int_{\mathbb{R}^n} |g(z)|\,dz \right) dy \\
&\quad \text{(changement de variable } x \to z = x - y) \\
&= \int_{\mathbb{R}^n} |f(y)|\,dy \int_{\mathbb{R}^n} |g(z)|\,dz \\
&= \|f\|_1 \|g\|_1. \qquad (9.1)
\end{aligned}$$

Par conséquent, $I(x) = \int_{\mathbb{R}^n} |f(y)|\,|g(x-y)|\,dy$ est d'intégrale finie bornée par $\|f\|_1\|g\|_1$, c'est donc que $I(x) = +\infty$ sur un ensemble de mesure nulle et donc $f * g$ est bien définie p.p. Puis

$$\begin{aligned}
\|f * g\|_1 = \int_{\mathbb{R}^n} |(f * g)(x)|\,dx &= \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} f(y) g(x-y)\,dy \right| dx \\
&\le \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |f(y) g(x-y)|\,dy\,dx \le \|f\|_1 \|g\|_1
\end{aligned}$$

en utilisant la borne (9.1). □

Ainsi $*$ est une loi de composition interne de $L^1(\mathbb{R}^n)$. On peut vérifier qu'elle est associative. $(L^1(\mathbb{R}^n), +, \cdot, *)$ est alors une algèbre ; comme elle est complète pour $\|\cdot\|_1$, il s'agit même d'une algèbre de Banach.

On a aussi, le résultat suivant :

Proposition 9.2.2 *Soit $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$ avec $r \ne +\infty$ alors pour $f \in L^p(\mathbb{R}^n)$, $g \in L^q(\mathbb{R}^n)$, $f * g$ existe p.p. et $f * g \in L^r(\mathbb{R}^n)$, avec $\|f * g\|_r \le \|f\|_p\,\|g\|_q$. Lorsque $r = +\infty$, on a même $(f * g)(x)$ existe pour tout $x \in \mathbb{R}^n$ avec $|(f * g)(x)| \le \|f\|_p\,\|g\|_q$.*

Si $q = 1$, on constate que la convolution par une fonction de $g \in L^1(\mathbb{R}^n)$ est un opérateur de $L^p(\mathbb{R}^n)$ dans $L^p(\mathbb{R}^n)$:

$$f \in L^p(\mathbb{R}^n) \mapsto f * g \in L^p(\mathbb{R}^n).$$

On utilise la généralisation suivante de l'inégalité de Hölder (8.4) :

Proposition 9.2.3 (Hölder généralisé) *Soit* $\frac{1}{\alpha}+\frac{1}{\beta}+\frac{1}{\gamma}=1$ *avec* $\alpha,\beta,\gamma\in[1,+\infty]$ *alors pour* $f\in L^\alpha(X,\mathcal{A},\mu)$, $g\in L^\beta(X,\mathcal{A},\mu)$, $h\in L^\gamma(X,\mathcal{A},\mu)$, *on a*

$$\|fgh\|_1\leq\|f\|_\alpha\|g\|_\beta\|h\|_\gamma. \tag{9.2}$$

Plus généralement si $p_i\in[1,+\infty]$, $1\leq i\leq N$, *avec* $\sum_{i=1}^N\frac{1}{p_i}=1$ *alors pour* $f_i\in L^{p_i}(X,\mathcal{A},\mu)$, $1\leq i\leq N$, *on a*

$$\Big\|\prod_{i=1}^N f_i\Big\|_1\leq\prod_{i=1}^N\|f_i\|_{p_i}. \tag{9.3}$$

Démonstration : En effet d'après l'inégalité de Hölder (8.4) avec $p_1=\frac{\alpha\beta}{\alpha+\beta}$ et $q_1=\gamma$ (conjugués) en (9.4), puis avec $p_2=\frac{\alpha}{p_1}$, $q_2=\frac{\beta}{p_1}$ (conjugués) en (9.5), on a

$$\begin{aligned}
\|fgh\|_1 &\leq \|fg\|_{p_1}\|h\|_{q_1}=\|f^{p_1}g^{p_1}\|_1^{1/p_1}\|h\|_\gamma && (9.4)\\
&\leq \Big(\|f^{p_1}\|_{p_2}\|g^{p_1}\|_{q_2}\Big)^{1/p_1}\|h\|_\gamma && (9.5)\\
&= \Big(\|f^{p_1p_2}\|_1^{1/p_2}\|g^{p_1q_2}\|_1^{1/q_2}\Big)^{1/p_1}\|h\|_\gamma\\
&= \|f\|_\alpha\|g\|_\beta\|h\|_\gamma.
\end{aligned}$$

L'inégalité (9.3) se prouve de la même façon par récurrence.

Démonstration :[Prop. 9.2.2] De $\frac{1}{p}+\frac{1}{q}=1+\frac{1}{r}$, on déduit $r>p$ et $r>q$. On a $\frac{1}{r}+\left(\frac{1}{p}-\frac{1}{r}\right)+\left(\frac{1}{q}-\frac{1}{r}\right)=1$. Posons

$$\alpha=r,\quad \beta=\frac{rp}{r-p},\quad \gamma=\frac{rq}{r-q}.$$

On a alors $\frac{1}{\alpha}+\frac{1}{\beta}+\frac{1}{\gamma}=1$ et par l'inégalité de Hölder généralisée (9.2), il vient :

$$\begin{aligned}
&\int_{\mathbb{R}^n}|f(y)||g(x-y)|\,dy\\
&=\int_{\mathbb{R}^n}|f(y)|^{p/r}|g(x-y)|^{q/r}|f(y)|^{1-p/r}|g(y)|^{1-q/r}\,dy\\
&\leq\left(\int_{\mathbb{R}^n}|f(y)|^p|g(x-y)|^q\,dy\right)^{1/r}\left(\int_{\mathbb{R}^n}|f(y)|^p\,dy\right)^{(r-p)/(rp)}\left(\int_{\mathbb{R}^n}|g(x-y)|^q\,dy\right)^{(r-q)/(rq)}\\
&=\left(\int_{\mathbb{R}^n}|f(y)|^p|g(x-y)|^q\,dy\right)^{1/r}\|f\|_p^{(r-p)/r}\|g\|_q^{(r-q)/r}
\end{aligned}$$

avec le changement de variable $y\to z=x-y$ dans $\int_{\mathbb{R}^n}|g(x-y)|^q\,dy$. Lorsque $r=+\infty$, la borne précédente s'écrit

$$\int_{\mathbb{R}^n}|f(y)||g(x-y)|\,dy\leq\|f\|_p\|g\|_q$$

et assure que $(f*g)(x)$ est bien défini pour tout $x\in\mathbb{R}^n$ et $|(f*g)(x)|\leq\|f\|_p\|g\|_q$. Puis pour $r\neq+\infty$, on a

$$\begin{aligned}
\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n}|f(y)||g(x-y)|\ dy\right)^r dx &\leq \|f\|_p^{r-p}\|g\|_q^{r-q}\int_{\mathbb{R}^n}|f(y)|^p|g(x-y)|^q\ dydx\\
&\leq \|f\|_p^{r-p}\|g\|_q^{r-q}\int_{\mathbb{R}^n}|f(y)|^p\left(\int_{\mathbb{R}^n}|g(x-y)|^q\ dx\right)dy\\
&= \|f\|_p^{r-p}\|g\|_q^{r-q}\ \|f\|_p^p\|g\|_q^q=\|f\|_p^r\|g\|_q^r. \qquad (9.6)
\end{aligned}$$

On en déduit que $\left(\int_{\mathbb{R}^n}|f(y)||g(x-y)|\ dy\right)^r<+\infty$ p.p. et donc $(f*g)(x)$ est bien définie pour presque tout $x\in\mathbb{R}^n$. Il vient alors avec (9.6) :

$$\begin{aligned}
\|f*g\|_r &= \left(\int_{\mathbb{R}^n}|(f*g)(x)|^r\ dx\right)^{1/r}\\
&\leq \left(\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n}|f(y)||g(x-y)|\ dy\right)^r\ dx\right)^{1/r}\\
&\leq \big(\|f\|_p^r\|g\|_q^r\big)^{1/r}=\|f\|_p\|g\|_q.
\end{aligned}$$

□

Lorsque $f\in L^p(\mathbb{R}^n)$ et $g\in L^q(\mathbb{R}^n)$, avec p,q conjugués, la convolée $f*g$ est régulière (sans aucune hypothèse de régularité de f,g !). Pour cela, on a besoin d'abord de :

Théorème 9.2.1 *Soit $f\in L^p(\mathbb{R}^n)$ pour $p<+\infty$ alors pour $h\in\mathbb{R}^n$, $\tau_hf(x)=f(x+h)$ converge vers f dans $L^p(\mathbb{R}^n)$ quand $h\to0$, ie. l'opérateur de translation par une constante est continu dans $L^p(\mathbb{R}^n)$.*

Démonstration : Comme $p<+\infty$, on peut utiliser la densité des fonctions continues à support compact dans $L^p(\mathbb{R}^n)$ (Th. **??**).

• On commence donc par considérer φ une telle fonction, de support dans $B(0,R)$. Pour $\|h\|\leq1/2$, on a

$$\begin{aligned}
\int_{\mathbb{R}^n}|\tau_h\varphi(x)-\varphi(x)|^p\ dx &= \int_{\mathbb{R}^n}|\varphi(x+h)-\varphi(x)|^p\ dx\\
&\leq \int_{\|x\|\leq R+1/2}|\varphi(x+h)-\varphi(x)|^p\ dx
\end{aligned}$$

car φ est nulle hors de $B(0,R)$ et $\varphi(\cdot+h)$ l'est hors de $B(0,R+\frac12)$ quand $\|h\|<\frac12$. Comme φ est uniformément continue sur le compact $\overline{B(0,R+1/2)}$ (théorème de Heine), pour tout $\varepsilon>0$, il existe $\alpha>0$ tel que si $\|h\|\leq\alpha$ alors pour tout $x\in B(0,R+1/2)$ on a

$$|\varphi(x+h)-\varphi(x)|\leq\varepsilon.$$

D'où

$$\int|\tau_h\varphi(x)-\varphi(x)|^p\ dx \ \leq\ \int_{B(0,R+1/2)}\varepsilon^p\ d\lambda_n=\varepsilon^p\ \lambda_n\big(B(0,R+1/2)\big)$$

$$\|\tau_h\varphi - \varphi\|_p \leq \varepsilon \left(\lambda_n(B(0, R + 1/2))\right)^{1/p}.$$

C'est à dire $\lim_{h\to 0} \|\tau_h\varphi - \varphi\|_p = 0$.

• Pour $f \in L^p(\mathbb{R}^n)$ et $\varepsilon > 0$, par densité des fonctions continues à support compact, il existe une telle fonction φ telle que $\|f - \varphi\|_p \leq \varepsilon/3$. Puis

$$\begin{aligned}
\|\tau_h f - f\|_p &\leq \|\tau_h f - \tau_h\varphi\|_p + \|\tau_h\varphi - \varphi\|_p + \|\varphi - f\|_p \\
&= 2\|f - \varphi\|_p + \|\tau_h\varphi - \varphi\|_p \\
&\leq 2\varepsilon/3 + \|\tau_h\varphi - \varphi\|_p
\end{aligned}$$

en utilisant $\|\tau_h f - \tau_h\varphi\|_p = \|f - \varphi\|_p$ dû à un changement de variables immédiat. D'après le premier cas comme φ est continue à support compact, $\lim_{h\to 0} \|\tau_h\varphi - \varphi\|_p = 0$ quand $h \to 0$, donc pour h assez petit, $\|\tau_h\varphi - \varphi\|_p \leq \varepsilon/3$ si bien que

$$\|\tau_h f - f\|_p \leq 2\varepsilon/3 + \varepsilon/3 = \varepsilon.$$

Le résultat suit car $\varepsilon > 0$ est arbitraire. □

On a ensuite pour p et q sont conjugués le résultat de convolution suivant :

Théorème 9.2.2 *Soit* $\frac{1}{p} + \frac{1}{q} = 1$ *et* $f \in L^p(\mathbb{R}^n)$*,* $g \in L^q(\mathbb{R}^n)$ *alors*

1. *la convolée* $f * g$ *est continue et même uniformément continue ;*
2. *pour* $p, q \neq +\infty$*, on a de plus* $\lim_{\|x\|\to+\infty}(f * g)(x) = 0$.

Démonstration : 1) Avec l'inégalité d'Hölder et un changement de variable immédiat $y \to x - y$ on a :

$$\begin{aligned}
\big|(f * g)(x + h) - (f * g)(x)\big| &\leq \int_{\mathbb{R}^n} |f(y)||g(x + h - y) - g(x - y)|\, dy \\
&\leq \|f\|_p \left(\int_{\mathbb{R}^n} |g(x + h - y) - g(x - y)|^q\, dy\right)^{1/q} \\
&\leq \|f\|_p \left(\int_{\mathbb{R}^n} |g(z + h) - g(z)|^q\, dz\right)^{1/q} \\
&\leq \|f\|_p\|\tau_h g - g\|_q.
\end{aligned}$$

Quitte à échanger les rôles de f et g, on peut supposer $q \neq +\infty$. Or le Th. 9.2.1 donne pour tout $\varepsilon > 0$, l'existence de $\alpha > 0$ tel que pour $\|h\| \leq \alpha$, on a $\|\tau_h g - g\|_q \leq \varepsilon$, c'est à dire

$$|(f * g)(x + h) - (f * g)(x)| \leq \varepsilon\|f\|_p.$$

On a donc pour tout $x \in \mathbb{R}^n$ et $\|h\|$ assez petit $|(f * g)(x + h) - (f * g)(x)|$ petit, ce qui donne la continuité uniforme de $f * g$.

2) Par le Th. **??**, soient φ et ϕ des fonctions continues à support compact qui approchent respectivement $f \in L^p(\mathbb{R}^n)$ et $g \in L^q(\mathbb{R}^n)$ en norme $\|\cdot\|_p$, $\|\cdot\|_q$ respectivement. On approche alors $f * g$ par $\varphi * \phi$ de la façon suivante :

$$\begin{aligned} |(f*g)(x) - (\varphi*\phi)(x)| &\leq |(f*g)(x) - (\varphi*g)(x)| + |(\varphi*g)(x) - (\varphi*\phi)(x)| \\ &\leq \big|\big((f-\varphi)*g\big)(x)\big| + \big|\big(\varphi*(g-\phi)\big)(x)\big| \\ &\leq \|f-\varphi\|_p\,\|g\|_q + \|\varphi\|_p\,\|g-\phi\|_q \end{aligned}$$

en majorant grâce au premier point. On a alors

$$\begin{aligned} |(f*g)(x)| &\leq |(f*g)(x) - (\varphi*\phi)(x)| + |(\varphi*\phi)(x)| \\ &\leq \|f-\varphi\|_p\,\|g\|_q + \|\varphi\|_p\,\|g-\phi\|_q + |(\varphi*\phi)(x)|. \end{aligned}$$

Choisissons $\varphi \in \mathcal{C}_c^0(\mathbb{R}^n)$ telle que

$$\|f-\varphi\|_p \leq \frac{\varepsilon}{2\|g\|_q},$$

puis $\phi \in \mathcal{C}_c^0(\mathbb{R}^n)$ telle que

$$\|g-\phi\|_q \leq \frac{\varepsilon}{2\|\varphi\|_p},$$

ce qui est possible par densité des fonctions continues à support compact dans $L^p(\mathbb{R}^n)$ et $L^q(\mathbb{R}^n)$ ($p, q < +\infty$, Th. **??**). Mais alors comme $\varphi*\phi$ est nulle hors d'un compact (car φ et ϕ sont elles mêmes nulles hors d'un compact, cf. Prop. 9.1.3), on en déduit que $|(f*g)(x)| < \varepsilon$ pour $\|x\|$ assez grand. On a donc bien prouvé :

$$\lim_{\|x\|\to+\infty} (f*g)(x) = 0.$$

□

9.3 Dérivation des convolutions

Proposition 9.3.1 *1. Si $f \in L^1(\mathbb{R}^n)$ et g est $C_b^1(\mathbb{R}^n)$ (fonction bornée de dérivées partielles $\partial g/\partial x_i$ bornées), alors $f*g$ est C^1 de dérivées*

$$\frac{\partial}{\partial x_i}\big(f*g\big) = f * \frac{\partial g}{\partial x_i}.$$

2. Si $f \in L^p(\mathbb{R}^n)$ et g est $C_c^1(\mathbb{R}^n)$ (C^1 à support compact), la même conclusion reste vraie.

3. Précédemment, si $g \in C_b^\infty(\mathbb{R}^n)$ (dans 1)) ou $g \in C_c^\infty(\mathbb{R}^n)$ (dans 2)) alors $f*g$ est C^∞ avec pour, tout multi-indice $\alpha \in \{1,\dots,n\}^p$, $\partial^\alpha(f*g) = f*\partial^\alpha g$.

Démonstration : 1) La fonction $f*g$ existe et est continue d'après le théorème précédent car on convole $L^1 * L^\infty$. Puis, pour tout $y \in \mathbb{R}^n$, $x \in \mathbb{R}^n \mapsto f(y)g(x-y)$ est dérivable par rapport à x_i de dérivée

$$f(y)\,\frac{\partial g}{\partial x_i}(x-y).$$

Comme $|f(y)\frac{\partial g}{\partial x_i}(x-y)| \leq M|f(y)|$ (dérivées partielles de g bornées), le théorème de dérivation sous l'intégrale (Th. 5.4.3) s'applique et donne la dérivabilité de $f*g$ par rapport à x_i avec la dérivée

$$\Big(f * \frac{\partial g}{\partial x_i}\Big)(x) = \int f(y)\frac{\partial g}{\partial x_i}(x-y)\;dy,$$

convolée à nouveau de $L^1 * L^\infty$ donc continue. Finalement, $f*g$ est de classe C^1.

2) Comme g est à support compact dans $B(0,R)$, on a $g \in L^q(\mathbb{R}^n)$ (q conjugué de p) et $f*g$ est bien définie par le Th. 9.2.2. Puis pour $\|x\| \leq K$ et pour tout $y \in \mathbb{R}^n$, $x \in K \mapsto f(y)g(x-y)$ est dérivable par rapport à x_i de dérivée $f(y)\frac{\partial g}{\partial x_i}(x-y)$ dominée par

$$|f(y)|\,\left\|\frac{\partial g}{\partial x_i}\right\|_\infty \mathbf{1}_{\{\|y\|\leq R+K\}}$$

car si $\|y\| > R+K$ alors $\|x-y\| > R$ et $\frac{\partial g}{\partial x_i}(x-y) = 0$. Par l'inégalité de Hölder (8.4), on a :

$$\begin{aligned}
\int_{\mathbb{R}^n} |f(y)|\,\left\|\frac{\partial g}{\partial x_i}\right\|_\infty \mathbf{1}_{\{\|y\|\leq R+K\}}dy &\leq \|f\|_p\left(\int\left(\left\|\frac{\partial g}{\partial x_i}\right\|_\infty \mathbf{1}_{\{\|y\|\leq R+K\}}\right)^q dy\right)^{1/q}\\
&\leq \|f\|_p\left\|\frac{\partial g}{\partial x_i}\right\|_\infty \lambda_n\big(\overline{B(0,(R+K))}\big)^{1/q} < +\infty.
\end{aligned}$$

On peut appliquer le théorème de dérivation sous le signe intégrale (Th. 5.4.3) et on en déduit que $f*g$ est dérivable par rapport à x_i de dérivée $f*\frac{\partial g}{\partial x_i}$; on en déduit aussi la continuité. Finalement, $f*g$ est de classe C^1.

3) Le résultat vient par applications successives des parties 1) ou 2). □

Définition 9.3.1 *Une fonction $g : \mathbb{R}$ ou $\mathbb{C} \to \mathbb{C}$ est* **entière** *si elle s'écrit $g(x) = \sum_{n\geq 0} c_n x^n$ pour une suite $(c_n)_{n\geq 0}$ avec un rayon de convergence infini.*

Pour les convolées avec des fonctions entières, on a la version suivante du résultat de dérivation (Prop. 9.3.1) :

Proposition 9.3.2 *Si $f \in L^1(\mathbb{C})$ (ou $L^1(\mathbb{R})$) est à support compact et g est une fonction entière. Alors $f*g$ est une fonction entière.*

Démonstration : Soit $0 < M < +\infty$ tel que $f(x) = 0$ quand $\|x\| > M$. Pour $x \in B(0,R)$, on a

$$\begin{aligned}
(f*g)(x) &= \int_{B(0,M)} f(y)\Big(\sum_{n\geq 0} c_n (x-y)^n\Big)\, dy \\
&= \sum_{n\geq 0} c_n \int_{B(0,M)} f(y)(x-y)^n\, dy \qquad (9.7)\\
&= \sum_{n\geq 0} c_n \int_{B(0,M)} f(y)\Big(\sum_{k=0}^{n} \binom{n}{k} x^k (-1)^{n-k} y^{n-k}\Big)\, dy \\
&= \sum_{k\geq 0} x^k \sum_{n\geq k} c_n \binom{n}{k} (-1)^{n-k}\Big(\int_{B(0,M)} f(y) y^{n-k}\, dy\Big) \\
&= \sum_{k\geq 0} d_k x^k \qquad (9.8)
\end{aligned}$$

avec $d_k = \sum_{n\geq k} c_n \binom{n}{k} (-1)^{n-k}\Big(\int_{B(0,M)} f(y) y^{n-k}\, dy\Big)$. L'échange (9.7) $\sum / \int$ se justifie avec le théorème de convergence dominée (Th. 5.4.1) puisque lorsque $x \in B(0,R)$ alors $\|x-y\| \leq R+M$ et

$$\Big|f(y)\sum_{n=0}^{p} c_n (x-y)^n\Big| \leq \Big(\sum_{n\geq 0} |c_n|(R+M)^n\Big)|f(y)| \in L^1(\mathbb{C})$$

en utilisant la convergence normale de la série $\sum_{n\geq 0} c_n x^n$ sur $\overline{B(0,R+M)}$. De plus, il y a convergence normale de la série (9.8) sur $\overline{B(0,R)}$ car

$$\begin{aligned}
\sum_{k\geq 0} |d_k| R^k &\leq \sum_{k\geq 0}\Big(\sum_{n\geq k} |c_n| \binom{n}{k}\Big(\int_{B(0,M)} |f(y)| M^{n-k}\, dy\Big)\Big) R^k \\
&\leq \sum_{n\geq 0} |c_n| \Big(\sum_{k=0}^{k}\sum_{n\geq k} \binom{n}{k} R^k M^{n-k}\Big)\Big(\int_{B(0,M)} |f(y)|\, dy\Big) \\
&= \Big(\sum_{n\geq 0} |c_n| (R+M)^k\Big)\|f\|_1 < +\infty
\end{aligned}$$

puisque $\sum_{n\geq 0} c_n x^n$ a un rayon de convergence infini. Finalement comme R est quelconque, le rayon de convergence de $\sum_{k\geq 0} d_k x^k$ est infini et $f*g$ est entière. □

9.4 Approximation et régularisation

Un des inconvénients majeurs de l'opération convolution est de ne pas avoir d'unité, c'est à dire de fonction $\mathbf{e}$ telle que toute fonction f : $f*\mathbf{e} = \mathbf{e}*f = f$. Pour compenser ce défaut, on introduit la notion d'approximation de l'unité :

Définition 9.4.1 (Approximation de l'unité) *Soient $e_k : \mathbb{R}^n \to \mathbb{R}^+$, $k \geq 1$, des fonctions mesurables vérifiant*

$$\int_{\mathbb{R}^n} e_k(x)\ dx = 1 \quad \text{et pour tout } \eta > 0\ :\ \lim_{k\to+\infty} \int_{|x|\geq\eta} e_k(x)\ dx = 0.$$

La suite $(e_k)_{k\geq1}$, est appelée approximation de l'unité.

Proposition 9.4.1 *Soit $e : \mathbb{R}^n \to \mathbb{R}_+$ une fonction mesurable telle que $\int_{\mathbb{R}^n} e(x)\ dx = 1$. Alors $e_k(x) = k^n e(kx)$, $k \geq 1$ définit une approximation de l'unité.*

Démonstration : En effet,

- d'abord e_k est une fonction mesurable positive ;
- puis $\int_{\mathbb{R}^n} e_k(x)\ dx = \int_{\mathbb{R}^n} e(y)\ dy = 1$ avec le changement de variable $x \to y = kx$.
- avec le même changement de variable $x \to y = kx$, par convergence dominée (Th. 5.4.1) :

$$\int_{\{\|x\|>\eta\}} e_k(x)\ dx = \int_{\{\|y\|>k\eta\}} e(y)\ dy \to 0, \quad k \to +\infty.$$

□

Exemples : • Soit

$$e(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}. \tag{9.9}$$

Il s'agit d'une fonction C^∞ entière d'intégrale 1 ($e^{-x^2/2} = \sum_{n\geq0} \frac{(-1)^n}{2^n n!} x^{2n}$ et $\int_{-\infty}^{+\infty} e^{-x^2/2}\ dx = \sqrt{2\pi}$). D'après la Prop. 9.4.1, on définit une approximation de l'unité en prenant $e_k(x) = ke(kx)$.

• Soit

$$\varphi(x) = \left\{ \begin{array}{lll} \mathbb{R} \to & \mathbb{R} & \\ x \mapsto & e^{-1/(1-x^2)} & |x| < 1 \\ & 0 & |x| \geq 1. \end{array} \right.$$

On montre que φ est C^∞ (par récurrence en utilisant que pour toute fraction rationnelle P/Q, on a $\lim_{x\to\pm1} \frac{P(x)}{Q(x)} \exp(-1/(1-x^2)))$ et à support compact $\overline{B(0,1)}$, ce qui permet de définir sur $\mathbb{R}^n$

$$e(x) = \frac{\varphi(\|x\|^2)}{\int_{\mathbb{R}^n} \varphi(\|y\|^2)\ dy} \tag{9.10}$$

où $\|x\|^2 = x_1^2 + \cdots + x_n^2$. Alors e est C^∞, à support dans $B(0,1)$ et $\int_{\mathbb{R}^n} e(x)\ dx = 1$. D'après la Prop. 9.4.1, on définit à nouveau une approximation de l'unité en prenant $e_k(x) = k^n e(kx)$.

La terminologie "approximation de l'unité" est justifiée par les deux résultats suivants :

Théorème 9.4.1 *Soit $f \in L^\infty(\mathbb{R}^n)$ et $(e_k)_{k\geq1}$ une approximation de l'unité, on suppose que* *i) soit f est continue sur $\{\|x\| \leq R\}$,*

ii) soit f est uniformément continue sur $\mathbb{R}^n$.

Alors $f * e_k$, $k \geq 1$, converge uniformément vers f.

Démonstration : Comme f est continue sur $\overline{B(0,R)}$, elle y est aussi uniformément continue (théorème de Heine) et il suffit de prouver ii). Notons que comme $\int_{\mathbb{R}^n} e_k(y)\,dy = 1$, on peut écrire $f(x) = \int_{\mathbb{R}^n} f(x)e_k(y)\,dy$. Puis par continuité uniforme de f : pour $\varepsilon > 0$, il existe $\eta > 0$ tel que pour $\|y\| \leq \eta$, on a $|f(x-y) - f(x)| \leq \varepsilon 2$, si bien que

$$\begin{aligned}
\big|f * e_k(x) - f(x)\big| &= \left|\int_{\mathbb{R}^n} f(x-y)e_k(y)\,dy - \int_{\mathbb{R}^n} f(x)e_k(y)\,dy\right| \\
&= \left|\int_{\mathbb{R}^n} \big(f(x-y) - f(x)\big)\,e_k(y)\,dy\right| \\
&\leq \int_{\mathbb{R}^n} |f(x-y) - f(x)|\,e_k(y)\,dy \\
&\leq \int_{\|y\|\leq\eta} |f(x-y) - f(x)|\,e_k(y)\,dy + \int_{\|y\|>\eta} |f(x-y) - f(x)|\,e_k(y)\,dy.
\end{aligned}$$

On a alors

$$\begin{aligned}
\sup_{x\in\mathbb{R}^n} \big|f * e_k(x) - f(x)\big| &\leq \varepsilon \int_{\|y\|\leq\eta} e_k(y)\,dy + 2\|f\|_\infty \int_{\|y\|>\eta} e_k(y)\,dy \\
&\leq \frac{\varepsilon}{2} + 2\|f\|_\infty \int_{\|y\|>\eta} e_k(y)\,dy \\
&\leq \varepsilon
\end{aligned}$$

avec $k \geq k_0$ assez grand pour que $\int_{\|y\|>\eta} e_k(y)\,dy \leq \varepsilon/(4\|f\|_\infty)$ puisque $\lim_{k\to+\infty} \int_{\|y\|>\eta} e_k(y)\,dy = 0$. On a donc pour tout $\varepsilon > 0$ et $k \geq k_0$: $\sup_{x\in\mathbb{R}^n} \big|f * e_k(x) - f(x)\big| \leq \varepsilon$, ce qui prouve le Th. 9.4.1. □

On a aussi :

Théorème 9.4.2 *Soit $f \in L^p(\mathbb{R}^n)$ pour $p < +\infty$ et $(e_k)_{k\geq1}$ une approximation de l'unité alors $f * e_k \to f$ dans $L^p(\mathbb{R}^n)$ lorsque $k \to +\infty$.*

Démonstration : Comme $e_k \in L^1(\mathbb{R}^n)$, $f \in L^p(\mathbb{R}^n)$, on a $f * e_k \in L^p(\mathbb{R}^n)$ (Prop. 9.2.2). Puis

$$\begin{aligned}
\big|f * e_k(x) - f(x)\big| &= \left|\int_{\mathbb{R}^n} \big(f(x-y) - f(x)\big)\,e_k(y)\,dy\right| \\
&\leq \int_{\mathbb{R}^n} |f(x-y) - f(x)|\,e_k(y)^{1/p} e_k(y)^{1/q}\,dy
\end{aligned}$$

$$\leq \left(\int_{\mathbb{R}^n} |f(x-y) - f(x)|^p e_k(y)\, dy\right)^{1/p} \left(\int_{\mathbb{R}^n} e_k(y)\, dy\right)^{1/q}$$

par l'inégalité de Hölder (8.4). Puis comme $\int_{\mathbb{R}^n} e_k(y)\, dy = 1$, on a

$$\begin{aligned}
\|f * e_k - f\|_p^p &= \int_{\mathbb{R}^n} |f * e_k(x) - f(x)|^p \, dx \\
&\leq \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |f(x-y) - f(x)|^p e_k(y)\, dy dx \\
&\leq \int_{\mathbb{R}^n} e_k(y) \left(\int_{\mathbb{R}^n} |f(x-y) - f(x)|^p \, dx\right) dy \\
&\quad \text{(Th. de Fubini-Tonelli, Th. 7.3.1)} \\
&= \int_{\mathbb{R}^n} e_k(y) \|\tau_{-y} f - f\|_p^p \, dy.
\end{aligned}$$

Comme $p < +\infty$, par le Th. 9.2.1, pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que pour $\|y\| < \eta$, $\|\tau_{-y} f - f\|_p^p \leq \varepsilon^p/2$, puis par l'inégalité de Minkowski (8.6) : $\|\tau_{-y} f - f\|_p \leq \|\tau_{-y} f\|_p + \|f\|_p = 2\|f\|_p$. Ainsi :

$$\begin{aligned}
\|f * e_k - f\|_p^p &\leq \int_{\|y\| \leq \eta} e_k(y) \|\tau_{-y} f - f\|_p^p \, dy + \int_{\|y\| > \eta} e_k(y) \|\tau_{-y} f - f\|_p^p \, dy \\
&\leq \frac{\varepsilon^p}{2} \int_{\|y\| \leq \eta} e_k(y)\, dy + \big(2\|f\|_p\big)^p \int_{\|y\| > \eta} e_k(y)\, dy \\
&\leq \frac{\varepsilon^p}{2} + \frac{\varepsilon^p}{2} = 2\varepsilon^p
\end{aligned}$$

pour $k \geq k_0$ assez pour que $\int_{\|y\| > \eta} e_k(y)\, dy \leq \frac{\varepsilon^p}{2(2\|f\|_p)^p}$ puisque $\lim_{k \to +\infty} \int_{|y| > \eta} e_k(y)\, dy = 0$. Finalement, pour tout $\varepsilon > 0$, il existe $k_0 \in \mathbb{N}$ tel que pour $k \geq k_0$, on ait $\|f * e_k - f\|_p \leq \varepsilon$, ce qui prouve le Th. 9.4.2. □

Applications des approximations de l'unité

Les approximations de l'unité permettent de prouver de jolis résultats d'analyse : le théorème de Stone-Weierstrass (Th. 9.4.3), le théorème de densité de $C_c^\infty(\mathbb{R}^n)$ dans $L^p(\mathbb{R}^n)$ lorsque $p < +\infty$ (Th. 9.4.4) et le théorème de Fejér sur la convergence en mode Césaro des séries de Fourier d'une fonction f périodique continue (Th. 9.4.5).

Stone-Weierstrass

Avec l'approximation de l'unité $e_k(x) = k^n e(kx)$, $k \geq 1$, où e est donnée par (9.9), on a :

Théorème 9.4.3 (Stone-Weierstrass) *Soit $f : K \to \mathbb{R}$ continue sur K compact. Alors f peut être uniformément approchée par des polynômes (ie. sur les compacts, l'ensemble des fonctions polynomiales est dense dans l'ensemble des fonctions continues pour la norme uniforme).*

Démonstration : On commence par prolonger f en une fonction $\widetilde{f}$ définie sur $\mathbb{R}$ et à support compact. On obtient ainsi une fonction $\widetilde{f}$ uniformément continue et bornée. Par la Prop. 9.3.2, les fonctions $\widetilde{f} * e_k$, $k \geq 1$, sont entières car les e_k, $k \geq 1$, le sont puis elles convergent uniformément vers $\widetilde{f}$ sur $\mathbb{R}$ donc convergent uniformément vers f sur K (Th. 9.4.1). Or toute fonction entière est limite uniforme sur un compact de ses sommes partielles qui sont des polynômes. Ainsi pour tout $\varepsilon > 0$, il existe k_0 tel que

$$\sup_{x\in K} \big|f(x) - (\widetilde{f} * e_{k_0})(x)\big| < \frac{\varepsilon}{2}$$

puis en notant

$$(\widetilde{f} * e_{k_0})(x) = \sum_{n\geq 0} c_n x^n$$

le développement en série entiére de $\widetilde{f} * e_{k_0}$, il existe n_0 et que

$$\sup_{x\in K} \Big| \sum_{n=0}^{n_0} c_n x^n - (\widetilde{f} * e_{k_0})(x)\Big| < \frac{\varepsilon}{2}.$$

Finalement, en notant $P_{n_0}(x) = \sum_{n=0}^{n_0} c_n x^n$, on a

$$\begin{aligned}
\sup_{x\in K} \big|f(x) - P_{n_0}(x)\big| &\leq \sup_{x\in K} \big|f(x) - (\widetilde{f} * e_{k_0})(x)\big| + \sup_{x\in K} \big|(\widetilde{f} * e_{k_0})(x) - P_{n_0}(x)\big| \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon
\end{aligned}$$

et f a un polynôme dans chacun de ses voisinages uniformes sur K, ie. f est limite uniforme de polynômes sur le compact K. □

Densité de C_c^∞ dans L^p

Avec l'approximation de l'unité $e_k(x) = ke(kx)$, $k \geq 1$, où e est donnée par (9.10), le résultat suivant complète le théorème de densité (Th. ??) du Chapitre 8 :

Théorème 9.4.4 *Pour $p < +\infty$, $C_c^\infty(\mathbb{R}^n) = \big\{$fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$ de classe C^∞ à support compact$\big\}$ est dense dans $L^p(\mathbb{R}^n)$.*

Démonstration : Troncature et convolution.

- Troncature : soit $f \in L^p(\mathbb{R}^n)$ et $f_m = f\mathbf{1}_{B(0,m)}$ alors pour tout $x \in \mathbb{R}^n$ $f_m(x) \to f(x)$,

$m \to +\infty$, et $|f_m(x) - f(x)|^p \leq 2^p |f(x)|^p \in L^1$. Ainsi par convergence dominée (Th. 5.4.1), on a

$$\lim_{m \to +\infty} \int |f_m - f|^p \, d\lambda_n = 0$$

et donc $f_m \to f$ dans $L^p(\mathbb{R}^n)$ quand $m \to +\infty$. Pour $\varepsilon > 0$ fixé, soit donc m_0 tel que $\|f_{m_0} - f\|_p \leq \varepsilon/2$.

• La fonction $f_{m_0} * e_k$ est C^∞ (car $e_k \in C_c^\infty$ et $f \in L^p$, Prop. 9.3.1) et à support compact (f_{m_0} et e_k à support compact, Prop. 9.1.2) donc $f_{m_0} * e_k \to f_{m_0}$ dans $L^p(\mathbb{R}^n)$ (Th. 9.4.2). On trouve donc k_0 tel que $\|f_{m_0} * e_{k_0} - f_{m_0}\|_p \leq \varepsilon/2$. Finalement,

$$\begin{aligned} \big\|f - (f_{m_0} * e_{k_0})\big\|_p &\leq \|f - f_{m_0}\|_p + \|f_{m_0} - (f_{m_0} * e_{k_0})\|_p \\ &\leq \varepsilon/2 + \varepsilon/2 = \varepsilon. \end{aligned}$$

Dans tout voisinage de $f \in L^p(\mathbb{R}^n)$ pour la norme $\|\cdot\|_p$, il y a donc une fonction $(f_{m_0} * e_{k_0})$ de classe C^∞ et à support compact, ce qui prouve le résultat de densité du Th. 9.4.4. □

Série de Fourier et Fejér

Sur le cercle $\mathcal{C} = \mathbb{R}/(2\pi\mathbb{Z})$, on adapte les notions de convolution et d'approximation de l'unité comme suit :

— convolution :

$$(f * g)(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) g(x - y) \, dy;$$

— approximation de l'unité sur $[0, 2\pi]$: ce sont des fonctions $e_k : \mathcal{C} \to \mathbb{C}$, $k \geq 1$, positives telles que $\frac{1}{2\pi} \int_{-\pi}^{\pi} e_k(t) \, dt = 1$ et

$$\lim_{k \to +\infty} \frac{1}{2\pi} \int_{\{-\eta < |y| \leq \pi\}} e_k(y) \, dy = 0.$$

Avec ce type d'approximation de l'unité, les résultats d'approximation vus sur $\mathbb{R}^n$ (Th. 9.4.1, Th. 9.4.2) restent valables sur le cercle $\mathcal{C}$. On a alors le résultat suivant sur une convergence en mode Césaro des séries de Fourier. On rappelle :

Théorème de Césaro : *Si $(u_n)_{n \geq 1}$ est une suite numérique qui converge vers $l \in \mathbb{R}$ alors $\lim_{n \to +\infty} \frac{1}{n}\big(u_1 + \dots + u_n\big) = l$.*

Théorème 9.4.5 (Fejér) *Soit f une fonction 2π-périodique, continue de $\mathbb{R}$ dans $\mathbb{C}$ et*

$$\widetilde{f}(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-ikt} \, dt, \quad s_n(x) = \sum_{k=-n}^{n} \widetilde{f}(k) e^{ikx}, \quad \sigma_N(x) = \frac{1}{N+1} \sum_{n=0}^{N} s_n(x).$$

Alors σ_N converge uniformément vers f quand $N \to +\infty$.

Les $\widetilde{f}(k)$, $k \in \mathbb{Z}$, sont les coefficients de Fourier de f ; s_n est la somme de Fourier associée à f.

Le Th. 9.4.5 exprime donc f continue 2π-périodique comme limite uniforme de polynômes trigonométriques (analogue du théorème de Stone-Weierstrass, Th. 9.4.3 sur le cercle $\mathcal{C}$, ie. dans le cas périodique).

Démonstration : On a $\widetilde{f}(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)e^{iky}dy$ et

$$\begin{aligned} s_n(x) &= \frac{1}{2\pi}\sum_{k=-n}^{n}\int_{-\pi}^{\pi} f(y)\exp\big(ik(x-y)\big)\,dy \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)\Big(\sum_{k=-n}^{n}\exp\big(ik(x-y)\big)\Big)\,dy \\ &= (f * D_n)(x) \end{aligned}$$

en notant $D_n(y) = \sum_{k=-n}^{n}\exp(iky)$. Notons que

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} D_n(y)\,dy = \sum_{k=-n}^{n}\frac{1}{2\pi}\int_{-\pi}^{\pi}\exp(iky)\,dy = 1.$$

On a $D_n(y) = \frac{\exp(i(n+1)y)-\exp(-iny)}{\exp(iy)-1}$ et donc

$$D_n(y) = \frac{\sin\big((n+1/2)y\big)}{\sin(y/2)} \quad \text{quand } y \neq 0 \qquad \text{et} \qquad D_n(0) = 2n+1.$$

Par définition de σ_N, on a $\sigma_N = f * K_N$ avec

$$\begin{aligned} K_n &:= \frac{1}{N+1}\sum_{n=0}^{N} D_n \\ K_n(y) &= \frac{1}{(N+1)\sin(y/2)}\sum_{n=0}^{N} Im\big(\exp(i(n+1/2)y)\big) \\ &= \frac{1}{(N+1)\sin(y/2)} Im\Big(\exp(iy/2)\frac{\big(\exp(i(N+1)y)-1\big)}{\exp(iy)-1}\Big) \\ &= \frac{1}{(N+1)\sin(y/2)} Im\Big(\exp(i(N+1/2)y)\frac{\sin((N+1)y/2)}{\sin(y/2)}\Big) \\ &= \frac{1}{N+1}\left(\frac{\sin\big((N+1)y/2\big)}{\sin(y/2)}\right)^2. \end{aligned}$$

On a donc $K_n(y) \geq 0$, $\int_{-\pi}^{\pi} K_n(y)\,dy = \frac{1}{N+1}\sum_{n=0}^{N}\frac{1}{2\pi}\int_{-\pi}^{\pi} D_n(y)\,dy = 1$ et pour $\eta > 0$, $\eta < |y| \leq \pi$,

$$0 \leq K_n(y) \leq \frac{1}{N+1}\frac{1}{\big(\sin(\eta/2)\big)^2} \longrightarrow 0, \quad N \to +\infty,$$

soit
$$\lim_{N\to+\infty} \frac{1}{2\pi} \int_{\eta<|y|\leq\pi} K_N(y)\ dy = 0$$
et $(K_N)_{N\geq 1}$ est donc une approximation de l'unité. De plus comme f est continue sur le compact $\mathcal{C}$, f y est uniformément continue. D'après le Th. 9.4.1 (adapté sur $\mathcal{C}$), on a $\sigma_N = f * K_N \to f$, $N \to +\infty$, uniformément, ce qui prouve le théorème de Fejér (Th. 9.4.5).
□

Chapitre 10

Absolue continuité

Dans ce chapitre, on géneralise la notion de mesure sur un espace mesurable $(X,\mathcal{A})$ en mesure à valeurs non nécessairement positives, appelée mesure signée, et mesure à valeurs complexes dite mesure complexe (!) en Section 10.1. Pour une mesure signée μ, on introduit alors deux types de décomposition en Section 10.2 : la décomposition de Hahn de l'espace $(X,\mathcal{A})$ en ensembles dit positif et négatif pour μ et la décomposition de Jordan de μ en la différence de deux vraies mesures (positives). On discute la notion d'intégrale par rapport à ces mesures en Section 10.3. L'absolue continuité et la singularité sont des relations entre deux mesures (signées) qu'on présente en Section 10.4. L'ensemble des mesures absolument continues par rapport à une mesure donnée μ est parfaitement décrit en Section 10.5 dans le cas σ-fini avec les théorèmes de Lebesgue (Th. 10.5.1) et de Radon-Nikodym (Th. 10.5.2).

10.1 Mesure signée

Définition 10.1.1 *On appelle mesure signée toute fonction d'ensembles μ sur un espace mesurable $(X,\mathcal{A})$ à valeurs dans $\mathbb{R}\cup\{-\infty\}$ ou $\mathbb{R}\cup\{+\infty\}$ telle que $\mu(\emptyset)=0$ et μ est σ-additive.*

Ainsi une mesure signée ne peut pas prendre à la fois la valeur $-\infty$ et la valeur $+\infty$: c'est lun ou l'autre.
Une mesure signée est dite finie sur une famille $\mathcal{E}$ si $|\mu(E)|<+\infty$ pour tout $E\in\mathcal{E}$ et σ-finie sur $\mathcal{E}$ si pour tout $E\in\mathcal{E}$ il existe une suite $(E_n)_{n\geq1}$ de $\mathcal{E}$ telle que $|\mu(E_n)|<+\infty$ et $E\subset\bigcup_{n\geq1}E_n$.

De nombreuses propriétés usuelles des (vraies) mesures (cf. Section 1.4) restent vraies pour les mesures signées :

Proposition 10.1.1 *Soit μ une mesure signée sur $(X,\mathcal{A})$ alors*

i) Si $A,B\in\mathcal{A}$ et $A\subset B$ alors $|\mu(B)|<+\infty$ implique $|\mu(A)|<+\infty$.

ii) Si $A,B\in\mathcal{A}$ et $A\subset B$ alors $|\mu(B)|<+\infty$ implique $\mu(B\setminus A)=\mu(B)-\mu(A)$.

iii) Si $(A_n)_{n\geq 1}$ est une suite disjointe de $\mathcal{A}$ telle que $\left|\mu\Big(\bigcup_{n\geq 1} A_n\Big)\right| < +\infty$ alors la série $\sum_{n\geq 1}\mu(A_n)$ converge absolument.

iv) Si $(A_n)_{n\geq 1}$ est une suite monotone de $\mathcal{A}$, avec en plus dans le cas décroissant $|\mu(A_{n_0})| < +\infty$ pour un certain n_0, alors

$$\mu\Big(\lim_n A_n\Big) = \lim_{n\to+\infty}\mu(A_n).$$

Démonstration : Si $A, B \in \mathcal{A}$ et $A \subset B$ avec $|\mu(B)| < +\infty$ alors $B = A \cup (B\setminus A)$, union de deux ensembles disjoints de $\mathcal{A}$ pour laquelle l'additivité de μ donne $\mu(B) = \mu(A) + \mu(B\setminus A)$. On en déduit i) puisque si $\mu(B)$ est finie alors $\mu(A)$ et $\mu(B\setminus A)$ aussi (on ne peut pas avoir de forme indéterminée $\mu(A) + \infty$ et $\mu(B\setminus B) = \infty$). Dans ce cas, on peut retrancher $\mu(A)$ et en déduire ii).
Pour iii), On note $A_n^+ = A_n$ ou $\emptyset$ et $A_n^- = \emptyset$ ou A_n selon que $\mu(A_n) \geq 0$ ou $\mu(A_n) < 0$. On a alors

$$\sum_{n\geq 1}\mu(A_n^+) = \mu\Big(\bigcup_{n\geq 1} A_n^+\Big), \quad \sum_{n\geq 1}\mu(A_n^-) = \mu\Big(\bigcup_{n\geq 1} A_n^-\Big)$$

ce qui avec i) assure que les sommes $\sum_{n\geq 1}\mu(A_n^+)$ et $\sum_{n\geq 1}\mu(A_n^-)$ sont finies. Il vient

$$\sum_{n\geq 1}|\mu(A_n)| = \sum_{n\geq 1}\big(\mu(A_n^+) - \mu(A_n^-)\big) = \sum_{n\geq 1}\mu(A_n^+) - \sum_{n\geq 1}\mu(A_n^-)$$

est finie comme souhaité.
Enfin, iv) se montre comme pour la propriété analogue pour les (vraies) mesures. □

Au delà encore des mesures à valeurs réelles, on peut considérer des mesures à valeurs complexes. Une mesure complexe est une fonction d'ensembles à valeurs complexes, σ-additive et telle que $\mu(\emptyset) = 0$. Comme la convergence d'une série complexe exige la convergence des parties réelle et imaginaire, il s'en suit que les parties réelle et imaginaire de μ sont des mesures signées (réelles). En fait, une mesure μ est complexe ssi elle s'écrit $\mu = \mu_1 + i\mu_2$ où μ_1, μ_2 sont des mesures signées réelles.

10.2 Décompositions de Hahn et de Jordan

Si μ_1, μ_2 sont deux mesures sur une tribu $\mathcal{A}$ alors leur somme $\mu_1 + \mu_2$, définie par $(\mu_1 + \mu_2)(A) = \mu_1(A) + \mu_2(A)$, $A \in \mathcal{A}$, est clairement une mesure sur $\mathcal{A}$. Cependant la différence $\mu_1 - \mu_2$ définie par $(\mu_1 - \mu_2)(A) = \mu_1(A) - \mu_2(A)$, $A \in \mathcal{A}$, ne l'est pas forcément (la positivité n'est pas garantie). Mais il s'agit d'une mesure signée. En fait, on montre dans cette section que toute mesure signée s'exprime comme la différence de deux (vraies) mesures dont l'une au moins est finie (cf. Th. 10.2.2, décomposition de Jordan).

Si μ est une mesure signée sur $(X, \mathcal{A})$, un ensemble $A \in \mathcal{A}$ est dit positif (resp. négatif, nul) si $\mu(B) \geq 0$ (resp. $\mu(B) \leq 0$, $\mu(B) = 0$) pour tout $B \subset A$. Noter que les sous-ensembles mesurables d'ensembles positifs le sont aussi. Et l'union d'ensemble A_n, $n \geq 1$, positifs l'est aussi : si $B \in \mathcal{A}$ et $B \subset \bigcup_{n\geq 1} A_n$ alors $B = \bigcup_{n\geq 1}(B \cap A_n) =$. On peut même écire $B = \bigcup_{n\geq 1} B_n$ avec des $B_n \in \mathcal{A}$ disjoints et $B_n \subset B \cap A_n$. Ainsi $\mu(B_n) \geq 0$ et $\mu(B) = \sum_{n\geq 1} \mu(B_n) \geq 0$ (σ-additvité). De même pour les ensembles négatifs.

Théorème 10.2.1 (Décomposition de Hahn) *Si μ est une mesure signée sur l'espace mesurable $(X, \mathcal{A})$ alors il existe des ensembles disjoints A, B tels que A est positif, B est négatif et $A \cup B = X$.*

La décomposition de $X = A \cup B$ en ensembles positif et négatif s'appelle la **décomposition de Hahn** de X. D'après le Th. 10.2.1, cette décomposition existe mais elle n'est clairement pas unique (un ensemble nul peut être rajouté à A ou B). Une telle décomposition permettra cependant de donner une décomposition d'une mesure signée μ qui ne dépendra pas de la décomposition de Hahn considérée (Th. 10.2.2).

Démonstration : Comme μ est supposée prendre au plus une des deux valeurs $-\infty, +\infty$, pour fixer les idées on suppose que $-\infty < \mu(A) \leq +\infty$ pour tout $A \in \mathcal{A}$. On pose

$$\alpha = \inf\big(\mu(A) : A \in \mathcal{A} \text{ négatif}\big).$$

Comme l'ensemble $\emptyset$ est négatif, il vient de suite $\alpha \leq 0$. Soit maintenant $(B_n)_{n\geq 1}$ une suite d'ensembles négatifs tels que $\alpha = \lim_{n\to+\infty} \mu(B_n)$, on pose $B = \bigcup_{n\geq 1} B_n$. On prouve le théorème par étapes :

i) *B est négatif.* En effet, l'union d'ensembles négatifs reste négative.

ii) *$\mu(B) = \alpha$ et donc $-\infty < \alpha \leq 0$.* D'après i) et la définition de α, on a $\alpha \leq \mu(B)$. Puis pour tout n, $B = (B \setminus B_n) \cup B_n$ et donc

$$\mu(B) = \mu(B \setminus B_n) + \mu(B_n) \leq \mu(B_n)$$

puisque $B \setminus B_n \subset B$ est négatif. Il suit $\mu(B) \leq \lim_{n\to+\infty} \mu(B_n) = \alpha$ ce qui prouve $\mu(B) = \alpha$ et ii) est acquis.

iii) *Soit $A = X \setminus B$. Si $F \subset A$ est négatif alors F est nul.* En effet, soit $F \subset A$ négatif et $G \in \mathcal{A}$ tel que $G \subset F$. Alors G est négatif et $E = B \cup G$ est négatif. Par définition de α et avec ii), on a $\alpha \leq \mu(E) = \mu(B) + \mu(G) = \alpha + \mu(G)$. D'où $\mu(G) \geq 0$. Comme par ailleurs G est négatif, on a $\mu(G) = 0$ et donc F est nécessairement nul.

iv) *L'ensemble $A = X \setminus B$ est positif.* Par l'absurde, si A est négatif, il existe $E_0 \subset A$, $E_0 \in \mathcal{A}$, avec $\mu(E_0) < 0$. Comme E_0 n'est pas nul, d'aprés iii), il n'est pas négatif non plus. On note k_1 alors le plus petit entier tel qu'il existe un mesurable $E_1 \subset E_0$ avec $\mu(E_1) \geq 1/k_1$. Comme $\mu(E_0)$ est finie ($-\infty < \mu(E_0) < 0$) et $E_1 \subset E_0$, la Prop. 10.1.1-i) et ii) donne $\mu(E_0 \setminus E_1) = \mu(E_0) - \mu(E_1) < 0$ comme $\mu(E_0) < 0$ et $\mu(E_1) > 0$. En appliquant le même argument à $E_0 \setminus E_1$: on trouve k_2 le plus petit entier tel qu'il existe un mesurable $E_2 \subset E_0 \setminus E_1$ avec $\mu(E_2) \geq 1/k_2$. En procédant par récurrence, on note k_n le plus petit entier tel qu'il existe un mesurable $E_n \subset E_0 \setminus \left(\bigcup_{i=1}^{n-1} E_i\right)$ avec $\mu(E_n) \geq 1/k_n$. On pose alors $F_0 = E_0 \setminus \left(\bigcup_{i=1}^{+\infty} E_i\right)$. On a $\bigcup_{n\geq 1} E_n \subset E_0$ et

$|\mu(E_0)| < +\infty$ si bien que $\sum_{n\geq 1}\mu(E_n) = \mu\Big(\bigcup_{n\geq 1} E_n\Big)$ converge et donc $\mu(E_n)\to 0$, ce qui exige $k_n \to +\infty$.
Maintenant, pour chaque $n \geq 1$, $F_0 \subset E_0 \setminus \bigcup_{i=1}^{n-1} E_i$. Ainsi, pour tout $F \in \mathcal{A}$ avec $F \subset F_0$, par définition de k_n, on a $\mu(F) < 1/(k_n - 1)$ et comme $k_n \to +\infty$ on a $\mu(F) \leq 0$. Ainsi F_0 est négatif et donc par iii), F_0 est en fait nul. Mais

$$\mu(F_0) = \mu(E_0) - \sum_{i\geq 1}\mu(E_i) < 0$$

puisque $\mu(E_0) < 0$ et $\mu(E_i) > 0$, $i \geq 1$. Cette conclusion ($\mu(F_0) < 0$) contredit F_0 ensemble nul. L'hypothèse de départ était donc absurde et finalement, A est positif. Cela achève la preuve du Th. 10.2.1.

□

Théorème 10.2.2 (Décomposition de Jordan) *Soit μ une mesure signée sur un espace mesurable $(X, \mathcal{A})$. Si $X = A \cup B$ est une décomposition de Hahn de X pour μ (A positif et B négatif) alors on définit des mesures μ_+, μ_- sur $\mathcal{A}$ par*

$$\mu_+(E) = \mu(E\cap A), \quad \mu_-(E) = -\mu(E\cap B), \quad E \in \mathcal{A}$$

dont l'une au moins est finie et $\mu = \mu_+ - \mu_-$. Les mesures μ_+, μ_- ne dépendent pas de la décomposition de Hahn utilisée et la décomposition $\mu = \mu_+ - \mu_-$ s'appelle la décomposition de Jordan de la mesure signée μ.

Démonstration : Comme $A \cap E \subset A$ (positif) et $B \cap E \subset B$ (négatif), les fonctions d'ensembles μ_+, μ_- sont positives. Il suit alors que ce sont de vraies mesures (la σ-additivité étant facilement due à celle de μ). Comme μ prend au plus une des deux valeurs $\pm\infty$, une au moins des deux mesures $\mu_\pm$ est finie. Puis pour tout $E \in \mathcal{A}$,

$$\mu(E) = \mu(E\cap A) + \mu(E\cap B) = \mu_+(E) - \mu_-(E)$$

et donc $\mu = \mu_+ - \mu_-$.

Il reste à montrer que $\mu_\pm$ ne dépend pas de la décomposition de Hahn utilisée. Pour cela, on considère deux décompositions de Hahn $X = A_1 \cup B_1 = A_2 \cup B_2$ par rapport à μ et on montre que

$$\mu(E\cap A_1) = \mu(E\cap A_2), \quad \mu(E\cap B_1) = \mu(E\cap B_2).$$

Noter que l'ensemble $E\cap(A_1\setminus A_2)$ est un sous-ensemble de A_1 positif et donc $\mu\big(E\cap(A_1\setminus A_2)\big) \geq 0$ mais est aussi un sous-ensemble de B_2 négatif et donc $\mu\big(E\cap(A_1\setminus A_2)\big) \leq 0$. Ainsi $\mu\big(E\cap(A_1\setminus A_2)\big) = 0$ pour tout $E \in \mathcal{A}$. De même, $\mu\big(E\cap(A_2\setminus A_1)\big) = 0$ pour tout $E \in \mathcal{A}$. Il vient alors

$$\mu(E\cap A_1) = \mu(E\cap A_1\cap A_2) = \mu(E\cap A_2).$$

De la même façon, on a $\mu(E \cap B_1) = \mu(E \cap B_2)$, $E \in \mathcal{A}$, ce qui complète la preuve du Th. 10.2.2. □

Il est clair qu'une mesure signée se décompose en vraies mesures de plusieurs façons : par exemple $\mu = (\mu_+ + \nu) - (\mu_- + \nu)$, où ν est une (vraie) mesure finie quelconque. Cependant, la décomposition de Jordan est caractérisée par une propriété d'unicité et de minimalité (μ_+, μ_- ont des supports disjoints). Par

$$|\mu|(A) = \mu_+(A) + \mu_-(A), \quad A \in \mathcal{A}$$

on définit une (vraie) mesure $|\mu|$ qu'on appelle la **variation totale** de μ.
Noter qu'un ensemble $E \in \mathcal{A}$ est positif ssi $\mu_-(E) = 0$ puisque si E est positif $E \cap B$ est un ensemble nul en tant qu'ensemble positif (sous-ensemble de E) et ensemble négatif (sous-ensemble de B). Réciproquement si $\mu_-(E) = 0$ alors pour $F \in \mathcal{A}$ et $F \subset E$, on a $\mu_-(F) = 0$ et $\mu(F) = \mu_+(F) \geq 0$ justifiant que E est positif.
De la même façon, E est négatif ssi $\mu_+(E) = 0$.
On a aussi $|\mu(E)| \leq |\mu|(E)$, avec égalité seulement si E est positif ou négatif. Enfin, noter que $|\mu|(E) = 0$ implique que E est un ensemble nul pour $|\mu|, \mu_+, \mu_-, \mu$.

Un exemple important de mesure signée est fourni par des intégrales indéfinies d'une fonction dont l'intégrale est bien définie :

Théorème 10.2.3 *Soient $(X, \mathcal{A}, \mu)$ un espace mesuré et f une fonction mesurable définie p.p. sur X et telle que f^+* **ou** *$f^- \in L^1(X, \mathcal{A}, \mu)$. Alors on définit une mesure signée ν sur $\mathcal{A}$ par*

$$\nu(A) = \int_A f \, d\mu, \quad A \in \mathcal{A}.$$

Démonstration : On a de suite $\nu(\emptyset) = 0$ et si $f \in L^1(X, \mathcal{A}, \mu)$ alors ν est finie ($\nu(X) = \int_X f d\mu$). On complète la preuve en montrant que ν est σ-additive. Pour cela, soit $(A_n)_{n\geq 1}$ une suite de $\mathcal{A}$ disjointe et $A = \bigcup_{n\geq 1} A_n$. Alors $f_\pm \, \mathbf{1}_A = \sum_{n\geq 1} f_\pm \, \mathbf{1}_{A_n}$ p.p. et par le théorème de convergence monotone (Th. 4.3.1), on a

$$\int_A f_\pm \, d\mu = \int f_\pm \, \mathbf{1}_A \, d\mu = \int \Big(\sum_{n\geq 1} f_\pm \, \mathbf{1}_{A_n} \Big) \, d\mu = \sum_{n\geq 1} \Big(\int f_\pm \, \mathbf{1}_{A_n} \, d\mu \Big) = \sum_{n\geq 1} \int_{A_n} f_\pm \, d\mu.$$

Comme $f_+ \in L^1(X, \mathcal{A}, \mu)$ ou $f_- \in L^1(X, \mathcal{A}, \mu)$, une des deux séries prédédentes converge vers un nombre fini et on a

$$\begin{aligned} \nu(A) &= \int_A f_+ \, d\mu - \int_A f_- \, d\mu = \sum_{n\geq 1} \Big(\int_{A_n} f_+ \, d\mu - \int_{A_n} f_- \, d\mu \Big) \\ &= \sum_{n\geq 1} \int_{A_n} f \, d\mu = \sum_{n\geq 1} \nu(A_n) \end{aligned}$$

justifiant la σ-additivité et que ν est bien une mesure. □

Une décomposition de Hahn naturelle de X par rapport à ν est $X = A \cup B$ où $A = \{x \in X : f(x) \geq 0\}$ et $B = A^c = \{x \in X : f(x) < 0\}$. Si l'ensemble $\{x \in X : f(x) = 0\}$ est non vide, une autre décomposition de Hahn est $A_1 \cup B_1$ avec $A_1 = \{x \in X : f(x) > 0\}$ et $B_1 = A_1^c = \{x \in X : f(x) \leq 0\}$. La décomposition de Jordan $\nu = \nu_+ - \nu_-$ de ν est donnée dans les deux cas par

$$\nu_+(E) = \int_E f^+ \, d\mu, \quad \nu_-(E) = \int_E f^- \, d\mu, \quad E \in \mathcal{A}$$

et la variation totale $|\nu|$ de ν par

$$|\nu|(E) = \nu_+(E) + \nu_-(E) = \int_E f^+ \, d\mu + \int_E f^- \, d\mu = \int_E (f^+ + f^-) \, d\mu = \int_E |f| \, d\mu.$$

Enfin, on montre facilement grâce à la décomposition de Jordan que l'extension d'une mesure σ-finie signée a une propriété d'unicité analogue aux (vraies) mesures.

Proposition 10.2.1 *Soient μ, ν des mesures signées sur un espace mesuré $(X, \mathcal{A})$, égales sur un π-système $\mathcal{P} \subset \mathcal{A}$ tel que $\sigma(\mathcal{P}) = \mathcal{A}$. Si μ est σ-finie sur $\mathcal{P}$ alors $\mu = \nu$ sur $\mathcal{A}$.*

Démonstration : On écrit $\mu = \mu_+ - \mu_-$ et $\nu = \nu_+ - \nu_-$. Pour $E \in \mathcal{P}$, on a

$$\mu_+(E) - \mu_-(E) = \nu_+(E) - \nu_-(E)$$

et donc $\mu_+(E) + \nu_-(E) = \nu_+(E) + \mu_-(E)$ lorsque les quatre termes sont finis. Mais si l'un des termes est infini par exemple $\mu_+(E) = +\infty$ alors $\nu_+(E) = +\infty$ (et $\mu_-(E)$, $\nu_-(E)$ sont finis) et l'égalité précédente persiste. On a donc $\mu_+ + \nu_- = \nu_+ + \mu_+$. Comme il s'agit de mesure σ-finies, égale sur $\mathcal{P}$, elle sont égales sur $\mathcal{S}(\mathcal{P}) = \mathcal{A}$ (extension déjà utilisée du théorème de Dynkin, Th. 1.5.2). On en déduit $\mu = \nu$ sur $\mathcal{A}$ en réarrangeant les termes à l'envers. □

10.3 Intégrale par rapport à une mesure signée

Si μ est une mesure signée sur $(X, \mathcal{A})$ de décomposition de Jordan $\mu = \mu_+ - \mu_-$, l'intégrale par rapport à μ s'exprime comme la différence des intégrales par rapport à μ_+ et μ_- pour les fonctions $f \in L^1(X, \mathcal{A}, \mu_+) \cap L^1(X, \mathcal{A}, \mu_-)$:

$$\begin{aligned} \int f d\mu &= \int f \, d\mu_+ - \int f \, d\mu_- \\ &= \int f^+ \, d\mu_+ - \int f^- \, d\mu_+ - \int f^+ \, d\mu_+ + \int f^- \, d\mu_-. \end{aligned} \tag{10.1}$$

Comme $|\mu| = \mu_+ + \nu_-$, on a pour toute fonction mesurable

$$\int |f| \, d|\mu| = \int |f| \, d\mu_+ + \int |f| \, d\mu_-$$

et donc $f \in L^1(X, \mathcal{A}, |\mu|)$ ssi $f \in L^1(X, \mathcal{A}, \mu_+)$ et $f \in L^1(X, \mathcal{A}, \mu_-)$, ie. $L^1(X, \mathcal{A}, |\mu|) = L^1(X, \mathcal{A}, \mu_+) \cap L^1(X, \mathcal{A}, \mu_-)$.

Lorsque f est mesurable mais $f \notin L^1(X, \mathcal{A}, |\mu|)$, on peut avoir $\int f\, d\mu = +\infty$ lorsque les deux termes négatifs dans la définition (10.1) de $\int f\, d\mu$ sont finis et un des termes positifs est $+\infty$. Autrement dit, $\int f\, d\mu = +\infty$ quand $f_- \in L^1(X, \mathcal{A}, \mu_+), f_+ \in L^1(X, \mathcal{A}, \mu_-)$ et $f_+ \notin L^1(X, \mathcal{A}, \mu_+)$ ou $f_- \in L^1(X, \mathcal{A}, \mu_-)$. De même pour avoir $\int f\, d\mu = -\infty$.

L'intégrale par rapport à une mesure signée vérifie des propriétés analogues à celle de l'intégrale par rapport à une (vraie) mesure.

Théorème 10.3.1 *i) Si μ est une mesure signée et $f \in L^1(X, \mathcal{A}, \mu)$ alors*

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

ii) (Convergence dominée) Soient μ une mesure signée et $(f_n)_{n \geq 1}$ une suite de fonctions de $L^1(X, \mathcal{A}, |\mu|)$ et $g \in L^1(X, \mathcal{A}, |\mu|)$ avec $|f_n| \leq g$ $|\mu|$-p.p. pour tout $n \geq 1$. Si f est mesurable telle que $f_n \to f$ $|\mu|$-p.p. alors $f \in L^1(X, \mathcal{A}, |\mu|)$ et

$$\int |f_n - f|\, d|\mu| \to 0, \quad \int f_n\, d\mu \to \int f\, d\mu, \quad n \to +\infty.$$

Démonstration : i) On utilise les propriétés analogues des (vraies) mesures avec la définition $\int f\, d\mu = \int f\, d\mu_+ - \int f\, d\mu_-$:

$$\left| \int f\, d\mu \right| \leq \left| \int f\, d\mu_+ \right| + \left| \int f\, d\mu_- \right| \leq \int |f|\, d\mu_+ + \int |f|\, d\mu_- = \int |f|\, d\mu.$$

ii) La première limite s'obtient par convergence dominée pour la (vraie) mesure $|\mu|$ et la deuxième suit de la première limite combinée avec i). □

Le résultat suivant est le théorème de transfert pour les mesures signées. Comme dans le cas de vraies mesures, on peut étendre le résultat aux fonctions non-intégrables lorsque les intégrales sont quand même bien définies.

Théorème 10.3.2 *Soient $(X, \mathcal{A})$ et $(Y, \mathcal{B})$ des espaces mesurés et μ une mesure signée sur $(X, \mathcal{A})$. On considère une fonction $\varphi : (X, \mathcal{A}) \to (Y, \mathcal{B})$ mesurable (définie $|\mu|$-p.p. sur X suffit). Alors*

i) On définit une mesure signée μ_φ sur $(Y, \mathcal{B})$ par

$$\mu_\varphi(F) = \mu(\varphi^{-1}(F)), \quad F \in \mathcal{B}.$$

ii) Si f est $\mathcal{B}$-mesurable définie μ_φ-p.p. et telle que $f \circ \varphi \in L^1(|\mu|)$ alors $f \in L^1(|\mu_\varphi|)$ et

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

Démonstration : Comme dans le cas μ vraie mesure, on montre que μ_φ est σ-additive et que $\mu_\varphi(\emptyset) = 0$. Puis comme μ prend au plus une des deux valeurs $-\infty, +\infty$, il en est de même pour μ_φ. Il suit que μ_φ est une mesure signée, ie. i) est prouvé.

Pour simplifier, on suppose que φ est définie sur tout X. Alors $\varphi^{-1}(\mathcal{B})$ est une tribu et on note $\widetilde{\mu}$ la restriction de μ à $\varphi^{-1}(\mathcal{B}) \subset \mathcal{A}$.
Soit $Y = A \cup B$ la décomposition de Hahn de Y pour $\widetilde{\mu}_\varphi$ (A est positif, B est négatif). On montre que $X = \varphi^{-1}(A) \cup \varphi^{-1}(B)$ est une décomposition de Hahn de X pour $\widetilde{\mu}$. En effet, $\varphi^{-1}(A)$ et $\varphi^{-1}(B)$ sont dans $\varphi^{-1}(\mathcal{B}) \subset \mathcal{A}$ et disjoints avec pour union X. Maintenant, soit $E \subset \varphi^{-1}(\mathcal{B})$ un sous-ensemble de $\varphi^{-1}(A)$ alors $E = \varphi^{-1}(G)$ pour un certain $G \in \mathcal{B}$. Comme $E = \varphi^{-1}(G) \subset \varphi^{-1}(A)$, on a $E = \varphi^{-1}(G \cap A)$ et donc $\widetilde{\mu}(E) = \widetilde{\mu}_\varphi(G \cap A) \geq 0$ puisque A est positif pour $\widetilde{\mu}_\varphi$. On en déduit que $\varphi^{-1}(A)$ est positif pour $\widetilde{\mu}$. De même, on montre que $\varphi^{-1}(B)$ est négatif pour $\widetilde{\mu}$.

Maintenant soit $\widetilde{\mu} = \mu_+ - \widetilde{\mu}_-$ la décomposition de Jordan de $\widetilde{\mu}$. On montre que $\widetilde{\mu}_\varphi = (\widetilde{\mu}_+ - \widetilde{\mu}_-)_\varphi = \widetilde{\mu}_{+,\varphi} - \widetilde{\mu}_{-,\varphi}$ est la décomposition de Jordan de $\widetilde{\mu}_\varphi$. En effet pour chaque $E \in \mathcal{B}$, on a

$$\widetilde{\mu}_{+,\varphi}(E) = \widetilde{\mu}\big(\varphi^{-1}E \cap \varphi^{-1}A\big) = \widetilde{\mu}\big(\varphi^{-1}(E \cap A)\big) = \widetilde{\mu}_\varphi(E \cap A) = (\widetilde{\mu}_\varphi)_+(E)$$

puisque $Y = A \cup B$ est la décomposition de Hahn de Y pour $\widetilde{\mu}_\varphi$. Donc $\widetilde{\mu}_{+,\varphi} = (\widetilde{\mu}_\varphi)_+$ et de même $\widetilde{\mu}_{-,\varphi} = (\widetilde{\mu}_\varphi)_-$. Il suit que $\widetilde{\mu}_\varphi = \widetilde{\mu}_{+,\varphi} - \widetilde{\mu}_{-,\varphi}$ est la décomposition de Jordan de $\widetilde{\mu}_\varphi$ et

$$|\widetilde{\mu}_\varphi| = \widetilde{\mu}_{+,\varphi} + \widetilde{\mu}_{-,\varphi} = (\widetilde{\mu}_+ + \widetilde{\mu}_-)_\varphi = |\widetilde{\mu}|_\varphi.$$

Noter que $|\widetilde{\mu}|(E) \leq |\widetilde{\mu}(E)|$ pour tout $E \in \varphi^{-1}(\mathcal{B})$ puisque

$$\begin{aligned}
|\widetilde{\mu}|(E) &= \widetilde{\mu}_+(E) + \widetilde{\mu}_-(E) = \mu\big(E \cap \varphi^{-1}A\big) - \mu\big(E \cap \varphi^{-1}B\big) \\
&= \mu\big(E \cap \varphi^{-1}A\big) - \mu\big(E \cap \varphi^{-1}B\big) \leq |\mu|\big(E \cap \varphi^{-1}A\big) + |\mu|\big(E \cap \varphi^{-1}B\big) \\
&= |\mu|(E).
\end{aligned}$$

On déduit alors du cas vraie mesure :

$$\int_Y |f|\ d|\mu_\varphi| = \int_Y |f|\ d\mu_\varphi = \int_Y |f|\ d\mu_\varphi = \int_X |f \circ \varphi|\ d\mu \leq \int_X |f \circ \varphi|\ d|\mu|.$$

Par conséquent, $f \circ \varphi \in L^1(X, \mathcal{A}, |\mu|)$ implique $f \in L^1(|\mu_\varphi|)$ et à nouveau d'après le cas vraie mesure

$$\begin{aligned}
\int_Y f\ d\mu_\varphi &= \int_Y f\ d\mu_\varphi = \int_Y f\ d\mu_{+,\varphi} - \int_Y f\ d\mu_{-,\varphi} \\
&= \int_X |f \circ \varphi|\ d\mu_+ - \int_X |f \circ \varphi|\ d\mu_- = \int_X f \circ \varphi\ d\mu
\end{aligned}$$

ce qui prouve le théorème lorsque φ est définie sur tout X. On peut adapter cette preuve au cas où φ est définie seulement $|\mu|$-p.p. sur X. Si φ est définie sur E avec $|\mu|(E^c) = 0$ alors on définit la fonction $\varphi' : X \to Y$ par $\varphi'(x) = \varphi(x)$ si $x \in E$ et $\varphi'(x) = y_0$ si $x \notin E$ où $y_0 \in Y$ est quelconque. □

10.4 Absolue continuité et singularité

Dans cette section, on fixe un espace mesuré $(X, \mathcal{A})$ et on y considère μ, ν deux mesures signées.

Définition 10.4.1 (Absolue continuité and co) *On dit que ν est absolument continue par rapport à μ, vraie mesure, et on note $\mu \ll \nu$, si $\mu(A) = 0 \implies \nu(A) = 0$.*
Si μ est signée, on définit encore $\nu \ll \mu$ par $\nu \ll |\mu|$.
On dit que μ et ν sont équivalentes, et on note $\mu \sim \nu$, si elles sont mutuellement absolument continues, ie. $\nu \ll \mu$ et $\mu \ll \nu$.
On dit que μ et ν sont singulières, et on note $\mu \perp \mu$, s'il existe $A \in \mathcal{A}$ et que $\mu(A) = 0$ et $\nu(A^c) = 0$ (ie. les supports des mesures sont disjoints, cf. Déf 1.3.4).

Les notions d'absolue continuité et de singularité par rapport à une même mesure sont orthogonales car :

Proposition 10.4.1 *Soient μ, ν deux mesures telles $\nu \ll \mu$ et $\nu \perp \mu$, alors $\nu = 0$.*

Démonstration : En effet il existe $A \in \mathcal{A}$ tel que $\mu(A) = 0$ et $\nu(A^c) = 0$. Mais alors pour $E \in \mathcal{A}$ quelconque, on a $\nu(E) = \nu(E \cap A) + \nu(E \cap A^c)$ qui est nul car $E \cap A \subset A$ donc $\mu(E \cap A) = 0$ et $\nu(E \cap A) = 0$ d'après $\nu \ll \mu$, puis $E \cap A^c \subset A^c$ avec $\nu(A^c) = 0$ donc $\nu(E \cap A^c) = 0$. □

Proposition 10.4.2 *Soient μ, ν des vraies mesures. On a $\nu \ll \mu$ ssi $\forall \varepsilon > 0$, il existe $\delta > 0$ tel que $\mu(A) < \delta$ implique $\nu(A) < \varepsilon$.*

Démonstration : La condition est suffisante car si $\mu(A) = 0$, elle donne $\nu(A) < \varepsilon$ pour tout $\varepsilon > 0$ qu'on peut faire tendre vers 0, ce qui donne $\nu(A) = 0$ et $\nu \ll \mu$.
La condition est nécessaire puisque si elle est en défaut, il existe $\varepsilon > 0$ et pour chaque $n \geq 1$, $A_n \in \mathcal{A}$ tel que $\mu(A_n) \leq 1/n^2$ et $\nu(A_n) > \varepsilon$. On considère alors $A = \limsup_n A_n = \bigcap_{m \geq 1} \bigcup_{n \geq m} A_n$ Pour tout $m \geq 1$, on a $\mu(A) \leq \sum_{n \geq m} (1/n^2)$ et donc $\mu(A) = 0$. Alors que $\nu(A) = \lim_{m \to +\infty} \nu\Big(\bigcup_{n \geq m} A_n\Big) > \varepsilon$ puisque $\bigcup_{n \geq m} A_n \supset A_m$ avec $\nu(A_m) > \varepsilon$. Le mesurable A contredit alors la définition de $\nu \ll \mu$. □

Un exemple de mesure ν absolument continue par rapport à μ est fourni par la Section 10.3 avec

$$\nu(A) = \int_A f \, d\mu, \quad A \in \mathcal{A}$$

lorsque $f^+ \in L^1(X, \mathcal{A}, \mu)$ ou $f^- \in L^1(X, \mathcal{A}, \mu)$. Le théorème de Radon-Nikodym (Th. 10.5.2 ci-dessous) montre que les mesures ν absolument continues par rapport à μ sont en fait exactement celles-ci quand μ, ν sont σ-continues.

Proposition 10.4.3 *Soient μ, ν deux mesures signées sur $(X, \mathcal{A})$. On a équivalence entre* *i)* $\nu \ll \mu$,

ii) $\nu_+ \ll \mu$ *et* $\nu_- \ll \mu$,

iii) $|\nu| \ll |\mu|$.

Démonstration : Pour voir que i) implique ii), on considère $E \in \mathcal{A}$ avec $|\mu|(E) = 0$ et $X = A \cup B$ une décomposition de Hahn de X par rapport à ν. Alors comme $|\mu|$ est une mesure $|\mu|(E) = 0$ implique $|\mu|(E \cap A) = |\mu|(E \cap B) = 0$. Comme $\nu \ll \mu$, on a $\nu(E \cap A) = \nu(E \cap B) = 0$ et donc $\nu_+(E) = \nu_-(E) = 0$. On en déduit $\nu_+ \ll \mu$ et $\nu_- \ll \mu$ et $|\nu| \ll \mu$.
ii) implique facilement iii) puisque $|\nu|(E) = \nu_+(E) + \nu_-(E) = 0$ lorsque $|\mu|(E) = 0$.
Finalement, lorsque iii) est vrai, soit $E \in \mathcal{A}$ tel que $|\mu|(E) = 0$. iii) implique $|\nu|(E) = \nu_+(E) + \nu_-(E) = 0$ et donc $|\nu(E)| \leq |\nu|(E) = 0$, c'est à dire i). □

D'après la Prop. 10.4.3, on a $\nu \ll \mu$ ssi $|\nu| \ll |\mu|$. On a aussi $\nu \sim \mu$ ssi $|\nu| \sim |\mu|$.

Exemple : Sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, la mesure de référence est la mesure de Lebesgue λ. On considére fréquemment sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ des mesures absolument continues par rapport à λ. Par exemple, lorsque ce sont des mesures de probabilités, il s'agit de lois à densité. Si sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, on a $\mu \ll \lambda$ alors nécessairement μ n'a pas d'atome mais ne pas avoir d'atome est loin d'être suffisant pour avoir l'absolue continuité.

10.5 Théorème de Radon-Nikodym

Les théorèmes de Lebesgue (Th. 10.5.1) et de Radon-Nikodym (Th. 10.5.2) assurent que, étant donnée une mesure μ σ-finie, une mesure signée σ-finie peut s'écrire comme la somme de deux mesures signées, l'une absolument continué par rapport à μ et l'autre singulière par rapport à μ.
On commence par prouver le résultat pour des mesures finies (Lemme 10.5.1) avant de considérer le cas général de mesures σ-finies et signées (Théorèmes 10.5.1 et 10.5.2). Le théorème de Radon-Nikodym trouve une place essentielle en probabilités pour définir les variables aléatoires à densité.

Lemme 10.5.1 *Soient $(X, \mathcal{A}, \mu)$ un espace mesuré fini et ν une mesure finie sur $\mathcal{A}$. Alors, il existe deux mesures finies uniquement déterminées ν_1 et ν_2 sur $\mathcal{A}$ telles que*

$$\nu = \nu_1 + \nu_2, \quad \nu_1 \ll \mu, \quad \nu_2 \perp \mu.$$

De plus, il existe une fonction f μ-integrable, p.p. unique, telle que pour tout $A \in \mathcal{A}$: $\nu_1(A) = \int_A f \, d\mu$.

Démonstration : Unicité : si $\nu_1 + \nu_2 = \nu_3 + \nu_4$ alors $\nu_1 - \nu_3 = \nu_4 - \nu_2$ est à la fois absolument continue par rapport à μ (car ν_1, ν_3 le sont) et singulière par rapport à μ (car ν_2, ν_4 le sont), ce qui exige sa nullité (Prop. **??**). On a donc $\nu_1 = \nu_3$ et $\nu_2 = \nu_4$.

De plus, si $\nu_1(A) = \int_A f\,d\mu = \int_A g\,d\mu$ alors en choisissant $A = \{f > g\}$, on a $\int_A (f-g)\,d\mu = 0$ ce qui exige $\mu(A) = 0$. De même $\mu(f < g) = 0$ et donc $f = g$ μ-p.p.

Existence de ν_1, ν_2 et f : on note $\mathcal{K}$ la famille des fonctions f mesurables positives sur X telles que

$$\int_A f\,d\mu \le \nu(A) \quad \text{pour tout } A \in \mathcal{A}.$$

La preuve consiste à trouver $f \in \mathcal{K}$ qui maximise $\int f d\mu$ de façon à "extraire" autant que possible μ de ν par $\int f\,d\mu$. On espère alors que le reste $\nu_2 = \nu - \nu_1$ sera singulier par rapport à μ.
On remarque que $\mathcal{K}$ est non vide car il contient la fonction nulle. On écrit alors $\alpha = \sup\left(\int_X f\,d\mu : f \in \mathcal{K}\right)$ et on considère $(f_n)_{n\ge1}$ suite de $\mathcal{K}$ telle que $\int_X f_n\,d\mu \nearrow \alpha$.
On pose $g_n(x) = \max(f_1(x), \dots, f_n(x)) \ge 0$. Pour $A \in \mathcal{A}$ et $n \ge 1$, A se décompose en $\bigcup_{i=1}^n A_i$ où les A_i sont mesurables disjoints et tels que $g_n(x) = f_i(x)$ pour $x \in A_i$: en effet, prendre $A_1 = \{x \in X : g_n(x) = f_1(x)\}$ puis $A_2 = \{x \in X : g_n(x) = f_2(x)\} \setminus A_1$, etc. Ainsi

$$\int_A g_n\,d\mu = \sum_{i=1}^n \int_{A_i} g_n\,d\mu = \sum_{i=1}^n \int_{A_i} f_i\,d\mu \le \sum_{i=1}^n \nu(A_i) = \nu(A)$$

ce qui assure $g_n \in \mathcal{K}$. Comme $(g_n)_{n\ge1}$ est une suite croissante, sa limite $f(x) = \lim_{n\to+\infty} g_n(x)$ est bien définie en tout $x \in X$ et par convergence monotone (Th. 4.3.1), on a

$$\int_A f\,d\mu = \lim_{n\to+\infty} \int_A g_n\,d\mu \le \nu(A).$$

Il suit que $f \in \mathcal{K}$ et $\int_X f\,d\mu = \lim_{n\to+\infty} \int_X g_n\,d\mu \ge \lim_{n\to+\infty} \int_X f_n\,d\mu = \alpha$, soit $\int_X f\,d\mu = \alpha$. On pose alors pour tout $A \in \mathcal{A}$:

$$\nu_1(A) = \int_A f\,d\mu \quad \text{et} \quad \nu_2(A) = \nu(A) - \nu_1(A).$$

Alors ν_1 est clairement une mesure avec $f \ge 0$, $f \in L^1(X, \mathcal{A}, \mu)$ et $\nu_1 \ll \mu$. Puis ν_2 est finie, σ-additive avec $\nu_2(A) \ge 0$ pour tout $A \in \mathcal{A}$ puisque $f \in \mathcal{K}$ assure $\nu_1(A) = \int_A f d\mu \le \nu(A)$. Ainsi ν_2 est une mesure finie et il reste à justifier que $\nu_2 \perp \mu$.

Pour cela, on considère la mesure (signée) $\lambda_n = \nu_2 - \frac{1}{n}\mu$, $n \ge 1$, et on pose $X = E_n \cup F_n$, décomposition de Hahn de X pour λ_n ($\lambda_{n|E_n}$ est positive, $\lambda_{n|F_n}$ est négative). Si $h_n = f + \frac{1}{n}\mathbf{1}_{E_n}$, alors pour tout $A \in \mathcal{A}$, on a

$$\begin{aligned}\int_A h_n\,d\mu &= \int_A f\,d\mu + \frac{1}{n}\mu(E_n \cap A) = \nu(A) - \nu_2(A) + \frac{1}{n}\mu(E_n \cap A)\\ &= \nu(A) - \nu_2(A \cap F_n) - \lambda_n(E_n \cap A) \le \nu(A)\end{aligned}$$

puisque ν_2 est une mesure et $\lambda_{n|E_n}$ est positive. Ainsi $h_n \in \mathcal{K}$ et

$$\alpha \ge \int_X h_n\,d\mu = \int_X f\,d\mu + \frac{1}{n}\mu(E_n) = \alpha + \frac{1}{n}\mu(E_n)$$

ce qui exige $\mu(E_n) = 0$. Si $E = \bigcup_{n\geq 1} E_n$ alors $\mu(E) = 0$. Comme $E^c \subset E_n^c = F_n$, on a $\lambda_n(E^c) \leq 0$ et donc $\nu_2(E^c) \leq \frac{1}{n}\mu(E^c)$ pour tous les $n \geq 1$.
Il suit $\nu_2(E^c) = 0 = \mu(E)$, ce qui établit $\nu_2 \perp \mu$ et achève la preuve du Lemme 10.5.1. □

On énonce maintenant le résultat sous sa forme générale :

Théorème 10.5.1 (Décomposition de Lebesgue) *Soient $(X, \mathcal{A}, \mu)$ un espace mesuré σ-fini et ν une mesure signée σ-finie sur $\mathcal{A}$. Alors, il existe deux mesures σ-finies uniquement déterminées ν_1 et ν_2 sur $\mathcal{A}$ telles que*

$$\nu = \nu_1 + \nu_2, \quad \nu_1 \ll \mu, \quad \nu_2 \perp \mu.$$

De plus, si ν est une (vraie) mesure alors ν_1 et ν_2 en sont aussi. La décomposition $\nu = \nu_1 + \nu_2$ s'appelle la décomposition de Lebesgue de ν par rapport à μ.

Démonstration : On montre d'abord l'existence de ν_1 et de ν_2 lorsque μ et ν sont toutes les deux de vraies mesures σ-finies. Pour cela on écrit $X = \bigcup_{n\geq 1} X_n$ avec $X_n \in \mathcal{A}$ disjoints et $0 \leq \mu(X_n) < +\infty$ et $0 \leq \nu(X_n) < +\infty$. Pour chaque $n \geq 1$, on pose :

$$\mu^{(n)}(A) = \mu(A \cap X_n), \quad \nu^{(n)}(A) = \nu(A \cap X_n), \quad A \in \mathcal{A}.$$

Comme $\mu^{(n)}, \nu^{(n)}$ sont finies, le Lemme 10.5.1 s'applique et donne les décompositions

$$\nu^{(n)} = \nu_1^{(n)} + \nu_2^{(n)} \quad \text{avec } \nu_1^{(n)} \ll \mu^{(n)}, \nu_2^{(n)} \perp \mu^{(n)}.$$

On définit maintenant les fonctions d'ensembles ν_1, ν_2 par

$$\nu_1(A) = \sum_{n\geq 1} \nu_1^{(n)}(A), \quad \nu_2(A) = \sum_{n\geq 1} \nu_2^{(n)}(A), \quad A \in \mathcal{A}.$$

On a $\nu = \nu_1 + \nu_2$ car

$$\nu(A) = \sum_{n\geq 1} \nu^{(n)}(A) = \sum_{n\geq 1} \big(\nu_1^{(n)}(A) + \nu_2^{(n)}(A)\big) = \sum_{n\geq 1} \nu_1^{(n)}(A) + \sum_{n\geq 1} \nu_2^{(n)}(A) = \nu_1(A) + \nu_2(A).$$

On voit maintenant que ν_1 et ν_2 sont des mesures σ-finies : pour la σ-additivité, si $A = \bigcup_{n\geq 1} A_n$ est une union disjointe alors

$$\nu_1(A) = \sum_{n\geq 1} \nu_1^{(n)}(A) = \sum_{n\geq 1}\sum_{k\geq 1} \nu_1^{(n)}(A_k) = \sum_{k\geq 1}\sum_{n\geq 1} \nu_1^{(n)}(A_k) = \sum_{k\geq 1} \nu_1(A_k)$$

en échangeant des sommations de termes positifs (théorème de Fubini-Tonelli, Th. 7.3.1, pour la mesure de comptage sur $\mathbb{N}$). Pour la σ-finitude : comme $X = \bigcup_{n\geq 1} X_n$, on a

$$\nu_1(X_n) = \sum_{m\geq 1} \nu_1^{(m)}(X_n) \leq \sum_{m\geq 1} \nu^{(m)}(X_n) = \nu(X_n) < +\infty.$$

De même pour ν_2.

Pour montrer que $\nu_1 \ll \mu$, on fixe $A \in \mathcal{A}$ avec $\mu(A) = 0$. Alors $\mu^{(n)}(A) = \mu(A \cap X_n) = 0$ et comme $\nu_1^{(n)} \ll \mu^{(n)}$, on a $\nu_1^{(n)}(A) = 0$. Il vient alors $\nu_1(A) = \sum_{n\geq 1} \nu^{(n)}(A) = 0$, soit $\nu_1 \ll \mu$.

On termine la preuve (lorsque ν est positive) en montrant que $\nu_2 \perp \mu$. Pour cela, comme pour $n \geq 1$, on a $\nu_2^{(n)} \perp \mu^{(n)}$, il existe $E_n \in \mathcal{A}$ tel que $\mu^{(n)}(E_n) = 0$ et $\nu^{(n)}(E_n^c) = 0$. Soient alors $F_n = E_n \cap X_n$ et $F = \bigcup_{n\geq 1} F_n$. Les ensembles F_n sont disjoints et

$$\mu(F) = \sum_{n\geq 1} \mu(F_n) = \sum_{n\geq 1} \mu^{(n)}(E_n) = 0,$$

tandis que $\nu^{(n)}(X_n^c) = \nu(X_n \cap X_n^c) = 0$ et donc $\nu_2^{(n)}(X_n^c) = 0$. Comme $F_n^c = E_n^c \cup X_n^c$, il vient $\nu_2^{(n)}(F_n^c) = 0$. Finalement,

$$\nu_2(F^c) = \sum_{n\geq 1} \nu_2^{(n)}(F^c) \leq \sum_{n\geq 1} \nu_2^{(n)}(F_n^c) = 0$$

puisque $F^c \subset F_n^c$. Ainsi $\mu(F) = 0 = \nu_2(F^c)$, soit $\nu_2 \perp \mu$.

Il reste à traiter le cas de ν mesure signée. Sa décomposition de Jordan (Th. 10.2.2) s'écrit $\nu = \nu_+ - \nu_-$ où au moins une des deux mesures ν_+, ν_- est finie et l'autre σ-finie. D'après la partie du théorème déjà prouvée dans le cas de (vraies) mesures σ-finies, on peut écrire

$$\begin{array}{rcll} \nu_+ & = & \nu_{+,1} + \nu_{+,2} & \text{avec } \nu_{+,1} \ll \mu,\ \nu_{+,2} \perp \mu \\ \nu_- & = & \nu_{-,1} + \nu_{-,2} & \text{avec } \nu_{-,1} \ll \mu,\ \nu_{-,2} \perp \mu. \end{array}$$

Par exemple si ν_- est finie alors $\nu_{-,1}$ et $\nu_{-,2}$ le sont aussi et $\nu = (\nu_{+,1} - \nu_{-,1}) + (\nu_{+,2} - \nu_{-,2}) = \nu_1 + \nu_2$ avec $\nu_1 = \nu_{+,1} - \nu_{-,1} \ll \mu$ et $\nu_2 = \nu_{+,2} - \nu_{-,2} \perp \mu$.

La décomposition de Lebesgue en découle donc lorsque ν est σ-finie signée.

Enfin, pour voir l'unicité, on suppose d'abord ν vraie mesure σ-finie et $\nu = \nu_1 + \nu_2 = \nu_3 + \nu_4$ où $\nu_1, \nu_3 \ll \mu$ et $\nu_2, \nu_4 \perp \mu$. Comme μ et ν sont des vraies mesures σ-finies, on peut écrire $X = \bigcup_{n\geq 1} X_n$ avec les X_n disjoints mesurables de mesures $\mu(X_n), \nu(X_n)$ finies. Pour chaque $n \geq 1$, on définit alors $\mu^{(n)}$ et $\nu_i^{(n)}$, $i = 1, 2, 3, 4$ par

$$\mu^{(n)}(A) = \mu(A \cap X_n), \quad \nu_i^{(n)}(A) = \nu_i(A \cap X_n), \quad A \in \mathcal{A}.$$

On a alors

$$\nu_1^{(n)} + \nu_2^{(n)} = \nu_3^{(n)} + \nu_4^{(n)}, \quad \nu_1^{(n)}, \nu_3^{(n)} \ll \mu^{(n)}, \quad \nu_2^{(n)}, \nu_4^{(n)} \perp \mu^{(n)}.$$

L'unicité du Lemme 10.5.1 assure alors $\nu_1^{(n)} = \nu_3^{(n)}$ et $\nu_2^{(n)} = \nu_4^{(n)}$ pour chaque $n \geq 1$. Il vient alors

$$\nu_1 = \sum_{n\geq 1} \nu_1^{(n)} = \sum_{n\geq 1} \nu_3^{(n)} = \nu_3$$

et de même $\nu_2 = \nu_4$. On a obtenu l'unicité lorsque ν est une vraie mesure σ-finie. Pour conclure, lorsque ν est signée σ-finie avec deux décomposition $\nu_1 + \nu_2 = \nu_3 + \nu_4$, l'unicité provient de la décomposition de Jordan pour chaque ν_i en réarrangeant l'équation de façon que chaque côté soit positif et en appliquant l'unicité déjà acquise pour les (vraies) mesures. □

Théorème 10.5.2 (Radon-Nikodym) *Soient $(X, \mathcal{A}, \mu)$ un espace σ-fini avec ν une mesure signée σ-finie sur $\mathcal{A}$. Si $\nu \ll \mu$ alors il existe une fonction f sur X, presque sûrement unique, à valeurs finies telle que pour tout $A \in \mathcal{A}$:*

$$\nu(A) = \int_A f \, d\mu.$$

La fonction f s'appelle la densité de nu par rapport à mu, elle est μ-intégrable ssi ν est finie. En général, au moins une des deux fonctions f^+, f^- est μ-intégrable et cela arrive lorsque ν_+ ou ν_- est finie. Si ν est une vraie mesure alors f est positive.

Démonstration : L'existence de f vient du Lemme 10.5.1 lorsque μ, ν sont finies et positives En effet, d'après l'unicité de la décomposition de Lebesgue en $\nu = \nu_1 + \nu_2 = \nu + 0$, on doit avoir $\nu = \nu_1$ et donc $\nu(A) = \nu_1(A) = \int_A f \, d\mu$ (Lemme 10.5.1). De plus, f prend des valeurs finies et est μ-intégrable car $\int_X f \, d\mu = \nu(X) < +\infty$.

On suppose maintenant que μ, ν sont des (vraies) mesures σ-finies. Comme dans les preuves précédentes, on écrit $X = \bigcup_{n \geq 1} X_n$ avec les $X_n \in \mathcal{A}$ disjoints tels que $\mu(X_n) < +\infty$ et $\nu(X_n) < +\infty$ et on pose à nouveau

$$\mu^{(n)}(A) = \mu(A \cap X_n), \quad \nu^{(n)}(A) = \nu(A \cap X_n).$$

Comme $\mu^{(n)}$ et $\nu^{(n)}$ sont des mesures finies sur $\mathcal{A}$ avec $\nu^{(n)} \ll \mu^{(n)}$ par la première partie de cette présente preuve, on a $\nu^{(n)}(A) = \int_A f_n \, d\mu^{(n)}$, $n \geq 1$, pour tout $A \in \mathcal{A}$ et pour une fonction f_n positive à valeurs finies. Alors

$$\nu(A \cap X_n) = \nu^{(n)}(A) = \int \mathbf{1}_A f_n \, d\mu^{(n)} = \int_{X_n} \mathbf{1}_A f_n \, d\mu = \int_A \mathbf{1}_{X_n} f_n \, d\mu.$$

On pose alors $f = \sum_{n \geq 1} \mathbf{1}_{X_n} f_n$ et par le théorème de convergence monotone (Th. 4.3.1), on a

$$\nu(A) = \sum_{n \geq 1} \nu(A \cap X_n) = \sum_{n \geq 1} \int_A \mathbf{1}_{X_n} f_n \, d\mu = \int_A \Big(\sum_{n \geq 1} \mathbf{1}_{X_n} f_n \Big) \, d\mu = \int_A f \, d\mu.$$

La fonction f est mesurable positive avec des valeurs finies (les f_n le sont et les X_n sont disjoints si bien que $f(x) = f_n(x)$ lorsque $x \in X_n$). On a donc prouvé l'existence de f lorsque μ, ν sont des mesures σ-finies.

Quand ν est seulement signée, on dispose de sa décomposition de Jordan $\nu = \nu_+ - \nu_-$ (Th. **??**) avec au moins une des deux mesures ν_+, ν_- finie et l'autre σ-finie. D'aprés le cas précédent de la preuve (vraies mesures σ-finies), on a $\nu_+(A) = \int_A f_+ \, d\mu$ et $\nu_-(A) = \int_A f_- \, d\mu$, $A \in \mathcal{A}$, pour des fonctions f_+, f_- positives et à valeurs finies et dont l'une au moins est μ-intégrable. La décomposition de Hahn de X pour ν, $X = E \cup F$, avec $\nu_+(F) = 0$ et $\nu_-(E) = 0$ montre qu'on peut prendre f_+ sur F et $f_- = 0$ sur E. Ainsi, clairement $\nu(A) = \int_A f \, d\mu$ pour tout $A \in \mathcal{A}$ où $f = f_+ - f_-$ (où f_+, f_- sont les parties positives et négatives de f) a toutes les propriétés requises pour le théorème, ce qui prouve l'existence de f.

Pour montrer la (presque sûre) unicité de f, on considère g une autre fonction avec les mêmes propriétés. On écrit $X = \bigcup_{n\geq 1} X_n$ avec les $X_n \in \mathcal{A}$ disjoints et de mesure $\mu(X_n), \nu(X_n)$ finies. Alors pour tout $n \geq 1$,

$$\nu^{(n)}(A) = \nu(A \cap X_n) = \int_A f\mathbf{1}_{X_n} \, d\mu = \int_A g\mathbf{1}_{X_n} \, d\mu, \quad A \in \mathcal{A}.$$

Comme $\nu^{(n)}$ est une mesure signée finie, $f\mathbf{1}_{X_n}$ et $g\mathbf{1}_{X_n}$ sont μ-intégrables et $f\mathbf{1}_{X_n} = g\mathbf{1}_{X_n}$ p.p. pour chaque $n \geq 1$ (a voir). On en déduit $f = g$ p.p. sur X. La fonction f est donc presque sûrement unique. □

La décomposition de Lebesgue (Th. 10.5.1) et le théorème de Radon-Nikodym (Th. 10.5.2) peuvent être combinés en un énoncé unique comme ci-dessous :

Théorème 10.5.3 *Soient $(X, \mathcal{A}, \mu)$ un espace mesuré σ-fini et ν une mesure signée σ-finie sur $\mathcal{A}$. Alors il existe deux mesures ν_1, ν_2 uniquement déterminées telles que*

$$\nu = \nu_1 + \nu_2, \quad \nu_1 \ll \mu, \; \nu_2 \perp \mu$$

et une fonction f, presque sûrement unique, mesurable, à valeurs finies telle que f^+ ou f^- est μ-intégrable et pour tout $A \in \mathcal{A}$: $\nu_1(A) = \int_A f \, d\mu$.
Ainsi pour un certain $A_0 \in \mathcal{A}$ avec $\mu(A_0) = 0$, on a pour $A \in \mathcal{A}$

$$\nu(A) = \int_A f \, d\mu + \nu_2(A \cap A_0) = \int_A f \, d\mu + \nu(A \cap A_0)$$

puisque $\mu(A_0) = 0$ implique $\nu_1(A \cap A_0) = 0$. La fonction f est μ-intégrable ssi ν_1 est finie. On a $\nu \ll \mu$ ssi $\nu(A_0) = 0$. Si ν est une mesure alors ν_1, ν_2 et f sont positives.

Noter pour finir que les théorèmes de Lebesgue (Th. 10.5.1) et de Radon-Nikodym (Th. 10.5.2) peuvent être en défaut sans la condition de σ-finitude. (ex à voir [LCP]).

Bibliographie

[ACMR] Guy Auliac, Christiane Cocozza-Thivent, Sophie Mercier, Michel Roussignol. *Mathématiques : Intégration et probabilités.* Edisciences, Coll. Objectif Licence, 2005.

[BL] Philippe Barbé, Michel Ledoux. *Probabilité.* EDP science, 2007.

[Bouyssel] Michel Bouyssel. *Intégrale de Lebesgue, mesure et intégration.* Cépaduès, 1996.

[JCB-Riemann] Jean-Christophe Breton. *Intégrale de Riemann.* Notes de cours de L3 Mathématiques, Université de La Rochelle, 2009.

[JCB-proba] Jean-Christophe Breton. *Probabilités.* Notes de cours de L3 Mathématiques, Université de Rennes 1, 2014.

[BP] Marc Briane, Gilles Pagès. *Théorie de l'intégration*, 5ème édition. Coll. Vuibert Supérieur, Ed. Vuibert, 2012.

[LCP] Ross Leadbetter, Stamatis Cambanis, Vladas Pipiras. *A basic course in Measure and Probability.* Cambridge university press, 2014.

[Rud] Walter Rudin. *Analyse réelle et complexe.* 3ème édition, Dunod, 1998.

[Suq] Charles Suquet. *Intégration, Fourier, Probabilités.* Note de cours de L3, Université Lille 1, 2004.