



Département STPI
3ème année MIC

Calcul intégral et probabilités

Aldéric Joulin

A. Joulin
Bureau 115 - GMM
ajoulin@insa-toulouse.fr

Année universitaire 2019-2020

Table des matières

1	Introduction	7
1.1	Préambule	7
1.2	Quelques rappels sur les cardinaux	8
2	Théorie de la mesure	13
2.1	Tribus ou σ -algèbres	13
2.1.1	Définitions	13
2.1.2	Tribu engendrée	14
2.2	Mesures	16
2.2.1	Définitions et premières propositions	16
2.2.2	Prolongement de mesures	19
2.3	Mesure de Lebesgue sur \mathbb{R}^n	20
2.3.1	Définition et propriétés	20
3	Intégration selon Lebesgue	23
3.1	Fonctions mesurables	23
3.1.1	Définition, opérations	23
3.1.2	Tribu engendrée par une fonction	25
3.1.3	Suites de fonctions mesurables	25
3.1.4	Fonctions étagées	26
3.2	Construction de l'intégrale	28
3.2.1	Intégrale des fonctions étagées	28
3.2.2	Intégrale des fonctions positives	29
3.3	Intégrale des fonctions de signe quelconque	31
3.3.1	Définition et premières propriétés	31
3.3.2	Le "presque partout"	33

4	Théorèmes fondamentaux d'intégration	37
4.1	Comportement face aux limites	37
4.1.1	Théorème de convergence monotone de Beppo-Levi . . .	37
4.1.2	Lemme de Fatou	40
4.1.3	Théorème de convergence dominée de Lebesgue	40
4.2	Lien avec l'intégrale de Riemann	42
4.3	Fonctions définies par une intégrale	45
4.3.1	Théorème de continuité	45
4.3.2	Théorème de dérivabilité	45
4.4	Intégration sur les espaces produits	48
4.4.1	Introduction	48
4.4.2	Tribu produit et mesure produit	48
4.4.3	Théorème de Fubini	50
4.5	Changement de variable	52
4.5.1	Image d'une mesure par une fonction mesurable et théorème de transport	52
4.5.2	Théorème de changement de variable	54
5	Espaces de Lebesgue	57
5.1	Espaces et normes L^p	57
5.1.1	Objectif	57
5.1.2	Définition de l'espace L^p	58
5.1.3	Inégalité triangulaire et complétude de L^p	60
5.2	L'espace L^∞	65
5.2.1	Définition	65
5.2.2	Complétude	66
5.3	Espaces L^p_{loc}	67
6	Relations entre espaces fonctionnels	69
6.1	Espaces de fonctions régulières	69
6.1.1	Espaces \mathcal{C}^k	69
6.1.2	Espaces \mathcal{C}^k_c	70
6.2	Relations entre les modes de convergence	72
7	Convolution des fonctions	75
7.1	Définition	75
7.1.1	Un exemple de la physique	75
7.1.2	Définition	76

7.2	Convolution et espaces fonctionnels	77
7.2.1	Espaces L^1 et L^p	78
7.2.2	Espaces L^1 et L^∞	79
7.2.3	Espaces L^p et L^q avec p et q conjugués	79
7.2.4	L'espace L^1_{loc}	80
7.2.5	Espaces \mathcal{C}^k_c et L^1_{loc}	80
7.3	Régularisation par convolution	81
8	Probabilités	87
8.1	Analogies avec la théorie de la mesure	87
8.2	Espérance et intégrabilité	89
8.3	Inégalités classiques	91
8.4	Indépendance et loi	91
8.5	Convergences	93

Introduction

SECTION 1.1

Préambule

Dans ce cours, on présente une introduction à un outil fondamental de l'analyse : la théorie de la mesure et de l'intégration, telle qu'elle a été développée au début du 20^{ème} siècle à partir des idées de Borel et Lebesgue. A la fin du 19^{ème} siècle, plusieurs mathématiciens cherchent à généraliser l'intégrale de Riemann. Pour des fonctions définies sur $[a, b] \subset \mathbb{R}$, l'idée de départ de l'intégrale de Lebesgue est la suivante : comme dans le cas de l'intégrale de Riemann, il s'agit d'approcher l'aire sous le graphe de la fonction par celle d'une union de rectangles. Mais ces rectangles sont définis de manière différente. L'intégrale de Riemann de f sur $[a, b]$ peut être approchée par des sommes de la forme

$$\sum_{j=1}^N f(x_j) |I_j|,$$

où I_1, \dots, I_N sont des intervalles disjoints dont la réunion est $[a, b]$, les $x_j \in I_j$ sont tels que $f(x_j) \leq f$ sur I_j , et $|I_j|$ désigne la longueur de I_j . Ici, on fixe un découpage de l'axe des abscisses. Au contraire, dans le cas de la théorie de Lebesgue, on fixe un découpage en fonction des valeurs de la fonction, c'est-à-dire de l'axe des ordonnées. On ne fixe donc pas la base du rectangle au départ, mais une variation dans les valeurs atteintes. Les rectangles (briques élémentaires) correspondants ne sont plus de simples rectangles, mais pourront être une infinité de rectangles.

Quel est l'intérêt d'introduire une nouvelle notion d'intégrale alors que nous avons déjà à notre disposition l'intégrale de Riemann ? Voici quelques éléments de réponse non exhaustifs quant aux limitations de cette dernière :

- toute fonction Riemann-intégrable $f : [a, b] \rightarrow \mathbb{R}$ doit être bornée. De surcroît, il existe des fonctions très simples, comme la fonction indicatrice de l'ensemble des rationnels \mathbb{Q} , qui ne sont pas Riemann-intégrables. En d'autres termes, la classe des fonctions Riemann-intégrables est relativement restreinte.

- l'intégrale de Riemann ne peut être étendue à des fonctions définies au-delà de sous-ensembles de \mathbb{R} ou de \mathbb{R}^n . En particulier, elle diffère de la théorie des séries numériques alors qu'au fond ce sont des objets similaires.
- si $(f_n)_{n \in \mathbb{N}}$ est une suite de fonctions continues de $[a, b]$ vers \mathbb{R} convergeant uniformément sur $[a, b]$ vers une fonction $f : [a, b] \rightarrow \mathbb{R}$, alors l'inversion limite/intégrale

$$\lim_{n \rightarrow +\infty} \int_a^b f_n(x) dx = \int_a^b f(x) dx,$$

a lieu. Néanmoins, il se peut que cette inversion soit valable sans pour autant avoir la convergence uniforme, qui est un mode de convergence très fort et difficile à obtenir en pratique.

La théorie de Lebesgue va nous permettre, quant à elle :

- d'intégrer beaucoup plus de fonctions.
- de considérer d'autres espaces plus géométriques comme les sphères.
- d'unifier dans une même théorie la sommation des séries et le calcul intégral.
- d'inclure le calcul des probabilités dans le cadre de la théorie générale de l'intégration.
- enfin, d'utiliser des théorèmes de convergence beaucoup plus puissants et robustes que ceux établis dans la théorie de Riemann.

SECTION 1.2

Quelques rappels sur les cardinaux

La notion de dénombrabilité apparaît tout au long de cours. On rappelle donc ici quelques définitions et propriétés.

Pour $n \in \mathbb{N}^*$, on pose $\mathbb{N}_n = \{1, \dots, n\}$.

Définition 1.1 *Un ensemble est dit fini de cardinal n s'il peut être mis en bijection avec \mathbb{N}_n .*

Un ensemble est dit strictement dénombrable s'il peut être mis en bijection avec \mathbb{N} et dénombrable s'il est fini ou strictement dénombrable.

Deux ensembles sont dits de même cardinal s'il existe une bijection de l'un vers l'autre.

Dans un ensemble E strictement dénombrable, il y a autant d'éléments que dans \mathbb{N} . Cela implique que les éléments de cet ensemble peuvent être numérotés, ou indexés, par les entiers naturels. Si $\varphi : \mathbb{N} \rightarrow E$ est une bijection, on peut écrire

$$E = \{\varphi(k), k \in \mathbb{N}\} = \{a_k, k \in \mathbb{N}\}.$$

Bien sûr, la numérotation dépend de la bijection choisie.

Exemple 1.1 \mathbb{Z} est dénombrable.

En effet, $\varphi : \mathbb{Z} \rightarrow \mathbb{N}$, $n \mapsto \begin{cases} 2n - 1 & \text{si } n \geq 1 \\ -2n & \text{si } n \leq 0 \end{cases}$ est une bijection de \mathbb{Z} dans \mathbb{N} .

On admettra la proposition suivante :

Proposition 1.1 (Bernstein) *Un ensemble est dénombrable*

- *si (et seulement si) il existe une injection de cet ensemble dans \mathbb{N} (ou dans tout autre ensemble dénombrable),*
- *si (et seulement si) il existe une surjection de \mathbb{N} (ou tout autre ensemble dénombrable) dans cet ensemble.*

Il existe une infinité d'ensemble dénombrables, d'après le résultat suivant.

Théorème 1.2 *Si A et B sont dénombrables alors $A \times B$ est dénombrable.*

Preuve : On numérote les éléments de A et B , $A = \{a_0, a_1, \dots, a_n, \dots\}$ et $B = \{b_0, b_1, \dots, b_n, \dots\}$. Alors,

$$A \times B = \{(a_0, b_0), (a_1, b_0), (a_0, b_1), (a_2, b_0), (a_1, b_1), \dots, (a_n, b_0), (a_{n-1}, b_1), \dots\}$$

Cela revient à numéroté les éléments de $\mathbb{N} \times \mathbb{N}$ en le parcourant selon des diagonales. □

Proposition 1.3 *L'ensemble \mathbb{Q} est dénombrable.*

Preuve : l'application $\mathbb{Z} \times \mathbb{N} \rightarrow \mathbb{Q}$, $(p, q) \mapsto \frac{p}{q+1}$ est surjective. \square

Théorème 1.4 *Une réunion dénombrable d'ensembles dénombrables est un ensemble dénombrable.*

Preuve : soit $(A_k)_{k \in \mathbb{N}}$ des ensembles dénombrables. Soit pour chaque ensemble, $\varphi_k : \mathbb{N} \rightarrow A_k$ une surjection. Alors, l'application

$$\mathbb{N} \times \mathbb{N} \rightarrow \bigcup_{k \in \mathbb{N}} A_k, (k, n) \mapsto \varphi_k(n)$$

est surjective. \square

Nous nous intéressons maintenant à trouver quelques ensembles non dénombrables.

Théorème 1.5 *L'ensemble \mathbb{R} n'est pas dénombrable.*

Preuve : (due à Cantor)

Montrons que $[0, 1]$ n'est pas dénombrable. Supposons, par l'absurde, que $[0, 1]$ puisse s'écrire $[0, 1] = \{x_0, x_1, \dots, x_k, \dots\}$. Pour tout $k \in \mathbb{N}$, on écrit le développement décimal de x_k : $x_k = 0, x_k^1 x_k^2 \dots x_k^n \dots$. Construisons $y = 0, y^1 \dots y^n$ de la manière suivante : pour tout $n \in \mathbb{N}$, y^n est un chiffre appartenant à l'ensemble $\{1, 2, \dots, 8\}$, choisi tel que $y^n \neq x_n^n$. Il est alors bien clair qu'il n'existe aucun k tel que $y = x_k$. Ce qui est contradictoire puisque $y \in [0, 1]$. \square

Théorème 1.6 *Il n'existe pas de bijection de A vers $\mathcal{P}(A)$.*

Preuve : supposons, par l'absurde, que f soit une bijection de A vers $\mathcal{P}(A)$. On pose $X = \{x \in A \mid x \notin f(x)\}$. X est une partie de A , donc comme f est bijective, il existe $y \in A$ tel que $f(y) = X$. On a alors, de par la définition de y et de X ,

$$\begin{aligned} y \in X &\iff y \notin f(y) && \text{définition de } X \\ &\iff y \notin X && \text{car } f(y) = X. \end{aligned}$$

On vient donc de montrer que $y \in X \iff y \notin X$, ce qui est contradictoire. Ainsi y n'existe pas. Ce qui contredit l'hypothèse f surjective. \square

Corollaire 1.7 (*admis*) *L'ensemble $\mathcal{P}(\mathbb{N})$ n'est pas dénombrable. De plus, $\mathcal{P}(\mathbb{N})$ est en bijection avec \mathbb{R} .*

Pour terminer cette partie, notons que l'ensemble des rationnels a également la propriété suivante, que l'on utilisera souvent.

Théorème 1.8 *L'ensemble \mathbb{Q} est dense dans \mathbb{R} , i.e. pour tout $x \in \mathbb{R}$, il existe une suite de rationnels $(r_n)_{n \in \mathbb{N}} \subset \mathbb{Q}$ qui converge vers x .*

Preuve : cf. TD.

□

Théorie de la mesure

SECTION 2.1

Tribus ou σ -algèbres

2.1.1 Définitions

La première définition que nous voyons est celle d'une algèbre, qui est un ensemble de parties de X stable par passage au complémentaire et par réunion finie.

Définition 2.1 Soit X un ensemble, $\mathcal{A} \subset \mathcal{P}(X)$. On dit que \mathcal{A} est une algèbre si :

- $X \in \mathcal{A}$;
- pour tous $A, B \in \mathcal{A}$, on a $A \cup B \in \mathcal{A}$;
- $A \in \mathcal{A} \implies X \setminus A \in \mathcal{A}$.

Voici quelques conséquences immédiates regroupées dans la prochaine proposition.

Proposition 2.1 Si \mathcal{A} est une algèbre,

- $\emptyset \in \mathcal{A}$;
- si $A_1, \dots, A_k \in \mathcal{A}$, on a $\bigcup_{i=1}^k A_i \in \mathcal{A}$ (stabilité par réunion finie) ;
- si $A, B \in \mathcal{A}$, on a $A \cap B = X \setminus \left((X \setminus A) \cup (X \setminus B) \right) \in \mathcal{A}$;
- \mathcal{A} est stable par intersection finie ;
- si $A, B \in \mathcal{A}$, on a $A \setminus B = A \cap (X \setminus B) \in \mathcal{A}$ (stabilité par différence).

Une tribu est une algèbre stable par réunion dénombrable.

Définition 2.2 Soit X un ensemble et $\mathcal{A} \subset \mathcal{P}(X)$. On dit que \mathcal{A} est une tribu, ou une σ -algèbre, si :

- $X \in \mathcal{A}$,
- pour toute suite $(A_k)_{k \in \mathbb{N}}$ d'éléments de \mathcal{A} , on a $\bigcup_{k \in \mathbb{N}} A_k \in \mathcal{A}$ (stabilité par réunion dénombrable)
- $A \in \mathcal{A} \Rightarrow X \setminus A \in \mathcal{A}$ (stabilité par complémentaire).

Le préfixe σ indique une propriété de stabilité vis-à-vis des opérations dénombrables.

Comme conséquence immédiate, on en déduit la proposition suivante.

Proposition 2.2 Une tribu est stable par toutes les opérations ensemblistes classiques (intersection dénombrable, différence symétrique, etc).

Exemple 2.1 $\mathcal{P}(X)$ et $\{\emptyset, X\}$ sont deux tribus sur X , respectivement la plus grosse et la plus petite qu'on puisse définir.

La définition suivante est une simple définition de vocabulaire.

Définition 2.3 On appelle espace mesurable tout couple (X, \mathcal{A}) où $\mathcal{A} \subset \mathcal{P}(X)$ est une tribu. Les éléments de \mathcal{A} sont alors appelés parties mesurables ou ensembles mesurables.

2.1.2 Tribu engendrée

Le lemme suivant stipule qu'une intersection quelconque de tribus est une tribu.

Lemme 2.3 Soit X un ensemble et $(\mathcal{A}_i)_{i \in I}$ une famille quelconque (dénombrable ou non) de tribus. Alors, $\bigcap_{i \in I} \mathcal{A}_i$ est aussi une tribu.

Preuve :

- Pour tout $i \in I$, $X \in \mathcal{A}_i$, donc $X \in \bigcap_{i \in I} \mathcal{A}_i$.

- Pour tout $n \in \mathbb{N}$, soit $A_n \in \bigcap_{i \in I} \mathcal{A}_i$. Alors pour tout $i \in I$, pour tout $n \in \mathbb{N}$, $A_n \in \mathcal{A}_i$, donc pour tout $i \in I$, $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}_i$ car \mathcal{A}_i est une tribu. Ainsi, $\bigcup_{n \in \mathbb{N}} A_n \in \bigcap_{i \in I} \mathcal{A}_i$.
- Si $A \in \bigcap_{i \in I} \mathcal{A}_i$, alors pour tout $i \in I$, $A \in \mathcal{A}_i$, donc pour tout $i \in I$, $X \setminus A \in \mathcal{A}_i$ car \mathcal{A}_i est une tribu, i.e. $X \setminus A \in \bigcap_{i \in I} \mathcal{A}_i$, ce qui conclut la démonstration.

□

En revanche, une réunion de tribus n'est pas en général une tribu. Par exemple si l'on pose :

$$X = \{0, 1, 2\}, \mathcal{A}_1 = \{\emptyset, X, \{0\}, \{1, 2\}\}, \mathcal{A}_2 = \{\emptyset, X, \{1\}, \{0, 2\}\},$$

alors \mathcal{A}_1 et \mathcal{A}_2 sont des tribus mais pas $\mathcal{A}_1 \cup \mathcal{A}_2$.

Si $\mathcal{F} \subset \mathcal{P}(X)$, le lemme précédent permet de parler de la plus petite tribu contenant \mathcal{F} .

Définition 2.4 Soit X un ensemble et $\mathcal{F} \subset \mathcal{P}(X)$. On appelle tribu engendrée par \mathcal{F} , notée $\sigma(\mathcal{F})$, la plus petite tribu contenant \mathcal{F} . C'est l'intersection de toutes les tribus contenant \mathcal{F} .

Un des types de tribu que l'on utilise le plus est les "tribus Boréliennes".

Définition 2.5 On appelle tribu borélienne sur \mathbb{R} , notée $\mathcal{B}(\mathbb{R})$, la tribu engendrée par les intervalles ouverts $]a, b[$, $a < b \in \mathbb{R}$.

La tribu borélienne $\mathcal{B}(\mathbb{R}^p)$ sur \mathbb{R}^p est la tribu engendrée par les pavés ouverts $]a_1, b_1[\times \cdots \times]a_p, b_p[$.

Les éléments de ces tribus seront appelés ensembles boréliens ou boréliens.

À titre d'exercice, on pourra vérifier que $\mathcal{B}(\mathbb{R})$ est aussi engendrée par les intervalles fermés $[a, b]$, $a, b \in \mathbb{R}$, ou par les intervalles semi-ouverts $]a, b]$, $a, b \in \mathbb{R}$, ou par les $] - \infty, a]$, $a \in \mathbb{R}$, ou même $a \in \mathbb{Q}$.

De même pour $\mathcal{B}(\mathbb{R}^p)$.

On définit de même la tribu borélienne sur la droite achevée $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$: $\mathcal{B}(\overline{\mathbb{R}})$ est engendrée par les ensembles du type $[a, +\infty]$, $a \in \mathbb{R}$.

Comme tout ouvert de \mathbb{R} peut s'écrire comme union disjointe dénombrable d'intervalles ouverts, la tribu borélienne $\mathcal{B}(\mathbb{R})$ contient tous les ensembles ouverts, et donc tous les ensembles fermés (par complémentaire). On pourra donc construire des mesures intéressantes sur cette tribu. Mais, il faut avoir à l'esprit que les ensembles boréliens ne sont pas faciles à décrire : en font partie les unions dénombrables d'ensembles ouverts ou fermés, les unions dénombrables d'intersections dénombrables d'ensembles ouverts ou fermés, les unions dénombrables d'intersections dénombrables d'unions dénombrables... Pour autant, la tribu borélienne $\mathcal{B}(\mathbb{R})$ est strictement plus petite que $\mathcal{P}(\mathbb{R})$.

La notion de tribu borélienne est en fait plus générale. Si l'on travaille sur un espace normé (ou métrique) X , on définit la tribu borélienne sur X comme la tribu engendrée par l'ensemble des ouverts de X . Elle contient, outre les ensembles ouverts, tous les ensembles fermés de X .

Dans la suite de ce cours, si la tribu sous-jacente n'est pas précisée, c'est qu'il s'agit de la tribu borélienne.

SECTION 2.2

Mesures

2.2.1 Définitions et premières propositions

Les tribus sont le cadre naturel sur lequel sont définies les mesures, au sens de Lebesgue.

Définition 2.6 Soit (X, \mathcal{A}) un espace mesurable. On appelle mesure (ou mesure positive) sur cet espace toute application μ définie sur \mathcal{A} et à valeurs dans $[0, +\infty]$ telle que :

- $\mu(\emptyset) = 0$;
- μ est σ -additive : pour toute famille dénombrable $(A_n)_{n \in \mathbb{N}}$ d'ensembles mesurables deux à deux disjoints, $\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} \mu(A_n)$.

Le triplet (X, \mathcal{A}, μ) est alors appelé espace mesuré.

On classe les mesures en fonction de certaines de leurs propriétés.

Définition 2.7 Soit (X, \mathcal{A}, μ) un espace mesuré.

- On dit que μ est finie si $\mu(X) < +\infty$.
- Si $\mu(X) = 1$, on dit que μ est une mesure de probabilité.
- On dit que μ est σ -finie s'il existe une suite $(X_k)_{k \in \mathbb{N}}$ telle que

$$\mu(X_k) < +\infty \quad \forall k \in \mathbb{N} \quad \text{et} \quad X = \bigcup_{k \in \mathbb{N}} X_k.$$

La plupart des mesures intéressantes (et toutes celles rencontrées dans ce cours) seront au moins σ -finies.

Exemple 2.2

- Mesure de comptage sur $(X, \mathcal{P}(X))$: $\mu(A) = \begin{cases} \text{card}(A) & \text{si } A \text{ fini ;} \\ +\infty & \text{sinon.} \end{cases}$
- Mesure de Dirac sur $(X, \mathcal{P}(X))$: pour $a \in X$, la mesure de Dirac en a , notée δ_a , est définie par

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

La mesure de comptage précédente peut s'exprimer comme

$$\mu(A) = \sum_{x \in X} \delta_x(A).$$

- Si $X = \{a_1, \dots, a_n\}$ et $\mathcal{A} = \mathcal{P}(X)$, alors $\forall (\alpha_1, \dots, \alpha_n) \in (\mathbb{R}^+)^n$, $\mu = \sum \alpha_i \delta_{a_i}$ est une mesure sur $(X, \mathcal{P}(X))$.

Vous avez déjà manipulé des mesures : les probabilités. Seul le vocabulaire est différent. Ce que l'on nomme espace mesurable dans ce cours est appelé espace probabilisable dans le cadre de la théorie des probabilités. Et une mesure finie telle que $\mu(X) = 1$ est appelée probabilité. Nous verrons plus en détail ces analogies au dernier chapitre.

Nous donnons maintenant deux propositions qui découlent directement de la définition des mesures.

Proposition 2.4 Soit (X, \mathcal{A}, μ) un espace mesuré.

- Pour tous ensembles mesurables A, B tels que $A \subset B$, on a $\mu(A) \leq \mu(B)$.
- Pour toute famille dénombrable d'ensembles mesurables $(A_k)_{k \in \mathbb{N}}$,

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

- Pour tous ensembles mesurables A et B ,

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

Preuve :

- On écrit $B = A \cup (B \setminus A)$, l'union étant disjointe. On a donc $\mu(B) = \mu(A) + \mu(B \setminus A)$. Comme μ est positive, on en déduit $\mu(B) \geq \mu(A)$.
- On construit à partir de la famille $(A_k)_{k \in \mathbb{N}}$, une famille $(B_k)_{k \in \mathbb{N}}$ d'ensembles mesurables deux à deux disjoints, comme suit. On pose

$$- B_0 = A_0 ;$$

$$- \text{pour tout } k \geq 1, B_k = A_k \setminus \bigcup_{j=0}^{k-1} A_j.$$

Les $(B_k)_{k \in \mathbb{N}}$ vérifient $\bigcup_{k \in \mathbb{N}} B_k = \bigcup_{k \in \mathbb{N}} A_k$, et pour tout $k \in \mathbb{N}$, $B_k \subset A_k$.

On a alors

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

- On écrit

$$A = (A \setminus B) \cup (A \cap B), \quad B = (B \setminus A) \cup (A \cap B);$$

$$A \cup B = (A \setminus B) \cup (A \cap B) \cup (B \setminus A),$$

où les unions sont disjointes. On en déduit

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

□

Proposition 2.5 Soit (X, \mathcal{A}, μ) un espace mesuré et $(A_k)_{k \in \mathbb{N}}$ une suite d'éléments de \mathcal{A} .

- Si la suite $(A_k)_{k \in \mathbb{N}}$ est croissante pour l'inclusion (i.e. $A_k \subset A_{k+1}$), alors

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

- Si la suite $(A_k)_{k \in \mathbb{N}}$ est décroissante pour l'inclusion (i.e. $A_{k+1} \subset A_k$), et si $\mu(A_0) < +\infty$, alors

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

Preuve : cf. TD.

□

2.2.2 Prolongement de mesures

On veut souvent définir *a priori* la valeur d'une mesure sur une certaine classe d'ensembles, par exemple les segments de \mathbb{R} . En général, il n'est pas évident que l'on puisse le faire. Un tel résultat est appelé théorème de prolongement.

Théorème 2.6 (admis) Soient μ et ν deux mesures finies sur (X, \mathcal{A}) . Soit $\mathcal{C} \subset \mathcal{A}$ une algèbre telle que $\sigma(\mathcal{C}) = \mathcal{A}$. Si μ et ν sont deux mesures sur (X, \mathcal{A}) qui coïncident sur \mathcal{C} , alors elles sont égales sur \mathcal{A} .

Exemple 2.3 Sur $[0, 1]$ muni de sa tribu borélienne, si l'on a $\mu([a, b]) = \nu([a, b])$ pour tout $0 \leq a < b \leq 1$, alors $\mu \equiv \nu$ sur $\mathcal{B}([0, 1])$.

Le théorème suivant assure l'unicité d'un prolongement à la tribu engendrée, mais sous l'hypothèse restrictive de mesure σ -finie (il n'y a pas unicité dans le cas contraire).

Théorème 2.7 (Carathéodory) (admis) Soit X un ensemble, \mathcal{C} une algèbre de parties de X et μ une fonction positive et σ -additive sur \mathcal{C} telle que X peut s'écrire comme union dénombrable d'éléments $(A_k)_{k \in \mathbb{N}}$ de \mathcal{C} , avec $\mu(A_k) < +\infty$. Alors, il existe un unique prolongement de μ en une mesure (σ -additive) définie sur la tribu $\sigma(\mathcal{C})$.

SECTION 2.3

Mesure de Lebesgue sur \mathbb{R}^n

La mesure de Lebesgue est celle que l'on utilise couramment sur \mathbb{R}^n , $n \geq 1$. Nous verrons plus tard dans ce cours que l'intégrale qui lui est associée prolonge le concept d'intégrale de Riemann.

2.3.1 Définition et propriétés

Théorème-Définition 2.8 *Il existe sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ une unique mesure λ telle que pour tout intervalle $[a, b]$,*

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

Cette mesure est appelée mesure de Lebesgue sur \mathbb{R} .

Théorème-Définition 2.9 *Soit $n \geq 2$ un entier. Il existe sur $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ une unique mesure λ_n telle que pour tout pavé $P = [a_1, b_1] \times \dots \times [a_n, b_n] \subset \mathbb{R}^n$,*

$$\lambda_n(P) = \prod_{j=1}^n (b_j - a_j).$$

Cette mesure est appelée "mesure de Lebesgue n -dimensionnelle".

Preuve : Nous ne démontrons pas ces théorèmes dans ce cours. La preuve utilise un théorème de prolongement du type de celui du Théorème 2.7. \square

Proposition 2.10

- *La mesure λ_n est invariante par translation : pour tout $a \in \mathbb{R}^n$ et tout $A \in \mathcal{B}(\mathbb{R}^n)$, alors $x + A \in \mathcal{B}(\mathbb{R}^n)$ et $\lambda_n(x + A) = \lambda_n(A)$.*
- *La mesure de Lebesgue λ_n est régulière : pour tout $A \in \mathcal{L}(\mathbb{R}^n)$,*

$$\lambda_n(A) = \inf\{\lambda_n(U) \mid U \text{ ouvert, } U \supset A\} \quad \text{régularité extérieure,}$$

$$\lambda_n(A) = \sup\{\lambda_n(K) \mid K \text{ compact, } K \subset A\} \quad \text{régularité intérieure.}$$

Théorème 2.11 *La mesure de Lebesgue λ_n est, à une constante multiplicative près, la seule mesure sur $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ finie sur les compacts et invariante par translation.*

Dans la suite de ce cours, si nous ne mentionnons pas la mesure sous-jacente sur l'espace mesuré considéré, c'est qu'il s'agit de la mesure de Lebesgue.

Intégration selon Lebesgue

SECTION 3.1

Fonctions mesurables

3.1.1 Définition, opérations

Nous allons maintenant définir la classe des fonctions susceptibles d'être intégrées, les *fonctions mesurables*. Remarquez que la définition ne fait intervenir que les tribus et non les mesures.

Définition 3.1 Une fonction $f : (X, \mathcal{A}) \rightarrow (Y, \mathcal{B})$ est dite mesurable (par rapport aux tribus \mathcal{A} et \mathcal{B}) si

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

Si \mathcal{A} et \mathcal{B} sont des tribus boréliennes, alors f est dite "borélienne".

Voici un critère pratique de mesurabilité.

Proposition 3.1 Avec les conditions de la définition, si $\mathcal{B} = \sigma(\mathcal{F})$, alors :

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

Preuve : \Rightarrow : évident.

\Leftarrow : Montrons que pour tout $B \in \mathcal{B}$, $f^{-1}(B) \in \mathcal{A}$. Notons

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

On a évidemment que $\mathcal{F} \subset \mathcal{C} \subset \mathcal{B}$. On veut montrer que $\mathcal{B} \subset \mathcal{C}$, c'est-à-dire que $\sigma(\mathcal{F}) \subset \mathcal{C}$: il suffit donc de montrer que \mathcal{C} est une tribu (comme \mathcal{C} contient \mathcal{F} , elle contiendra alors la tribu engendrée par \mathcal{F} , cette dernière étant l'intersection de toutes les tribus contenant \mathcal{F}) :

- on a $f^{-1}(Y) = X \in \mathcal{A}$ donc $Y \in \mathcal{C}$.
- si $B \in \mathcal{C}$, alors $f^{-1}(B) \in \mathcal{A}$ donc $f^{-1}(Y \setminus B) = X \setminus f^{-1}(B) \in \mathcal{A}$, d'où $Y \setminus B \in \mathcal{C}$.

- soit $(A_i)_{i \in \mathbb{N}}$ une suite de \mathcal{C} , i.e. $f^{-1}(A_i) \in \mathcal{A}$. Alors

$$f^{-1}\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \bigcup_{i \in \mathbb{N}} f^{-1}(A_i) \in \mathcal{A},$$

car \mathcal{A} est une tribu.

Ainsi, \mathcal{C} est bien une tribu, d'où la conclusion. \square

En particulier, une fonction à valeurs dans \mathbb{R} (resp. \mathbb{R}^n) muni de sa tribu borélienne est mesurable si l'image réciproque de tout segment $[a, b]$ (resp. tout pavé) est mesurable.

Exemple 3.1 Soit $A \subset X$. La fonction indicatrice de A est définie par

$$\mathbb{1}_A : (X, \mathcal{A}) \rightarrow \mathbb{R}, \quad \mathbb{1}_A(x) = \begin{cases} 1 & \text{si } x \in A; \\ 0 & \text{sinon.} \end{cases}$$

L'indicatrice $\mathbb{1}_A$ est mesurable si et seulement si $A \in \mathcal{A}$.

Corollaire 3.2 Toute fonction continue de \mathbb{R}^n dans \mathbb{R}^p est mesurable.

Preuve : Pour tout ouvert O de \mathbb{R}^p , l'ensemble $f^{-1}(O)$ est un ouvert de \mathbb{R}^n car f est continue. \square

La proposition suivante est relativement facile à démontrer.

Proposition 3.3

- Soient $f : (X, \mathcal{A}) \rightarrow (Y, \mathcal{B})$ et $g : (Y, \mathcal{B}) \rightarrow (Z, \mathcal{C})$ deux fonctions mesurables. Alors leur composée $g \circ f$ est mesurable.
- Soient $f, g : (X, \mathcal{A}) \rightarrow \mathbb{R}$ deux fonctions mesurables. Alors $(f, g) : (X, \mathcal{A}) \rightarrow \mathbb{R}^2$ est mesurable.
- Soient $f, g : (X, \mathcal{A}) \rightarrow \mathbb{R}$ deux fonctions mesurables. Alors $f + g$, fg , $\sup(f, g)$, $\inf(f, g)$, et si g ne s'annule pas, f/g sont mesurables de (X, \mathcal{A}) dans \mathbb{R} .

3.1.2 Tribu engendrée par une fonction

Définition 3.2 Soit f une application de $(X, \mathcal{A}) \rightarrow (Y, \mathcal{B})$. On appelle tribu engendrée par f la plus petite tribu sur X qui rend f mesurable. On la note $\sigma(f)$.

On peut vérifier que :

- $\sigma(f) = \{f^{-1}(B) : B \in \mathcal{B}\}$.
- Si, de plus, $\mathcal{B} = \sigma(\mathcal{F})$, alors $\sigma(f) = \sigma(\{f^{-1}(B) : B \in \mathcal{F}\})$.

3.1.3 Suites de fonctions mesurables

Dans cette section, nous nous intéressons à savoir si le concept de fonction mesurable passe à la limite.

Théorème 3.4 Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions mesurables de $(X, \mathcal{A}) \rightarrow \mathbb{R}$. Si $(f_n)_{n \in \mathbb{N}}$ converge simplement vers f , alors f est mesurable.

Preuve : comme $\mathcal{B}(\mathbb{R})$ est engendrée par les intervalles de la forme $] -\infty, a[$, $a \in \mathbb{R}$, il suffit de montrer que pour $U =] -\infty, a[$, on a $f^{-1}(U) \in \mathcal{A}$. Pour $k \in \mathbb{N}^*$, on définit $U_k =] -\infty, a - \frac{1}{k}[$. Notons

$$H = \bigcup_{k \in \mathbb{N}^*, m \in \mathbb{N} \ n \geq m} \bigcap_{n \geq m} f_n^{-1}(U_k),$$

qui appartient à la tribu \mathcal{A} (pourquoi ?). On montre par double inclusion que $f^{-1}(U) = H$, ce qui permet de conclure. En effet :

- si $x \in H$, alors il existe $k_0 \in \mathbb{N}^*$ et $m_0 \in \mathbb{N}$ tels que $\forall n \geq m_0$, $f_n(x) < a - \frac{1}{k_0}$. En passant à la limite $n \rightarrow +\infty$, on obtient $f(x) \leq a - \frac{1}{k_0} < a$, i.e. $x \in f^{-1}(U)$.
- réciproquement, si $f(x) < a$, i.e. $x \in f^{-1}(U)$, alors il existe $k \in \mathbb{N}^*$ tel que $f(x) < a - \frac{1}{k}$. Comme $f_n(x) \xrightarrow[n \rightarrow +\infty]{} f(x)$, il existe $m \in \mathbb{N}$ tel que $\forall n > m$, $|f_n(x) - f(x)| < 1/(2k)$, donc

$$f_n(x) \leq |f_n(x) - f(x)| + f(x) < \frac{1}{2k} + a - \frac{1}{k} < a - \frac{1}{2k}.$$

Ainsi, pour tout $n > m$, $f_n(x) \in U_{2k}$. Autrement dit $x \in H$.

□

On rappelle à présent la définition de la limite inférieure et limite supérieure d'une suite $(u_n)_{n \in \mathbb{N}}$ de \mathbb{R} . On pose

$$\liminf_{n \rightarrow +\infty} u_n = \lim_{n \rightarrow +\infty} \inf_{m \geq n} u_m = \sup_{n \in \mathbb{N}} \inf_{m \geq n} u_m,$$

et

$$\limsup_{n \rightarrow +\infty} u_n = \lim_{n \rightarrow +\infty} \sup_{m \geq n} u_m = \inf_{n \in \mathbb{N}} \sup_{m \geq n} u_m.$$

Ci-dessus, la première limite est un sup car c'est la limite d'une suite croissante, $(\inf_{m \geq n} u_m)_{n \in \mathbb{N}}$. De même, la seconde est un inf car il s'agit de la limite d'une suite décroissante, $(\sup_{m \geq n} u_m)_{n \in \mathbb{N}}$. Par monotonie, ces limites existent dans $\overline{\mathbb{R}}$ et l'on a

- $\liminf_{n \rightarrow +\infty} u_n \leq \limsup_{n \rightarrow +\infty} u_n$.
- $\lim_{n \rightarrow +\infty} u_n = l \iff \liminf_{n \rightarrow +\infty} u_n = l = \limsup_{n \rightarrow +\infty} u_n$.

On a alors le corollaire immédiat.

Corollaire 3.5 Soit $(f_k)_{k \in \mathbb{N}}$ une suite de fonctions mesurables de (X, \mathcal{A}) dans \mathbb{R} . Alors $\limsup_{k \rightarrow +\infty} f_k$ et $\liminf_{k \rightarrow +\infty} f_k$ sont mesurables de (X, \mathcal{A}) dans $\overline{\mathbb{R}}$.

3.1.4 Fonctions étagées

Les fonctions étagées vont être utiles à la construction de l'intégrale.

Définition 3.3 Une fonction étagée est une fonction $(X, \mathcal{A}) \rightarrow \mathbb{R}$ mesurable, qui ne prend qu'un nombre fini de valeurs.

Si f est étagée alors f s'écrit $f = \sum_1^k \alpha_i \mathbb{1}_{A_i}$ avec

$$A_i = \{x \in X \mid f(x) = \alpha_i\} = f^{-1}(\{\alpha_i\}) \in \mathcal{A}.$$

L'écriture est unique avec les conditions suivantes :

$$A_i \neq \emptyset, A_i \cap A_j = \emptyset \text{ et } \alpha_i \neq \alpha_j \text{ pour } i \neq j, \bigcup_{i=1}^k A_i = X. \quad (3.1)$$

Dans la suite du cours, nous supposerons que les fonctions étagées considérées satisferont (3.1).

Attention : ne pas confondre avec les fonctions en escalier, qui sont des combinaisons linéaires de fonctions indicatrices d'*intervalles*. En particulier, une fonction en escalier positive est étagée mais la réciproque est fausse.

Théorème 3.6 *Toute fonction mesurable de $(X, \mathcal{A}) \rightarrow \overline{\mathbb{R}}$ est limite simple de fonctions étagées. De plus, si $f \geq 0$, la suite peut être choisie croissante et positive.*

Preuve : 1^{er} cas : $f \geq 0$.

Pour tout $n \in \mathbb{N}^*$, tout $k \in \{1, \dots, n2^n\}$, on pose

$$A_k^n = f^{-1} \left(\left[\frac{k-1}{2^n}, \frac{k}{2^n} \right] \right) = \left\{ x \in X : \frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n} \right\}$$

$$\text{et } B^n = f^{-1}([n, +\infty[) = \{x \in X : f(x) \geq n\}.$$

Ces ensembles sont mesurables et forment une partition de X . Posons

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

qui, par construction, est une fonction étagée positive. Alors,

- la suite (f_n) est croissante. En effet, pour $x \in X$, s'il existe $k \in \{1, \dots, n2^n\}$ tel que $x \in A_k^n$ alors soit $x \in A_{2k}^{n+1}$, soit $x \in A_{2k-1}^{n+1}$, mais dans les deux cas,

$$f_{n+1}(x) \geq \frac{(2k-1)-1}{2^{n+1}} = \frac{k-1}{2^n} = f_n(x).$$

Si par contre $x \in B^n$, i.e. $f(x) \geq n$ auquel cas $f_n(x) = n$, alors soit $f(x) \geq n+1$ et dans ce cas $x \in B^{n+1}$ et donc $f_{n+1}(x) = n+1 \geq n = f_n(x)$, soit $f(x) \in [n, n+1[$ et donc $x \in A_k^{n+1}$ pour un $k \geq k_0 (= n2^{n+1} + 1)$, d'où

$$f_{n+1}(x) = \frac{k-1}{2^{n+1}} \geq \frac{k_0-1}{2^{n+1}} = n = f_n(x).$$

Ainsi, la suite de fonctions (f_n) est croissante.

- la suite (f_n) converge simplement vers f . Soit $x \in X$ fixé.
 - si $f(x) = +\infty$, pour tout $n \in \mathbb{N}$, $x \in B^n$, donc $f_n(x) = n \rightarrow +\infty = f(x)$.
 - si $f(x) < +\infty$, dès que $n > f(x)$, il existe $k \in \{1, \dots, n2^n\}$ tel que $f(x) \in \left[\frac{k-1}{2^n}, \frac{k}{2^n} \right]$, donc

$$|f_n(x) - f(x)| = f(x) - \frac{k-1}{2^n} < \frac{k}{2^n} - \frac{k-1}{2^n} = \frac{1}{2^n},$$

quantité qui tend vers 0 lorsque n tend vers l'infini. Ainsi, (f_n) est croissante et converge simplement vers f .

2^{ème} cas : f quelconque.

Avec les notations $f^+ = \sup(f, 0)$ et $f^- = \sup(-f, 0) = -\inf(f, 0)$, on a $f = f^+ - f^-$, les fonctions f^+ et f^- étant toutes les deux positives. On applique le cas précédent à f^+ et f^- . \square

SECTION 3.2

Construction de l'intégrale

3.2.1 Intégrale des fonctions étagées

Définition 3.4 Soit $f : (X, \mathcal{A}, \mu) \rightarrow [0, +\infty[$ une fonction étagée positive de la forme $f = \sum_{i=1}^k \alpha_i \mathbb{1}_{A_i}$. On appelle intégrale de f sur $A \in \mathcal{A}$ par rapport à la mesure μ la quantité

$$\int_A f d\mu = \sum_{i=1}^k \alpha_i \mu(A_i \cap A).$$

Lemme 3.7 Soit $f = \sum_{i=1}^k \alpha_i \mathbb{1}_{A_i}$ une fonction étagée positive.

- (Croissance) Si $f \leq g$ où g est une fonction étagée positive, alors pour tout $A \in \mathcal{A}$, $\int_A f d\mu \leq \int_A g d\mu$.

- (Linéarité) Pour tous $A \in \mathcal{A}$ et $c \geq 0$, on a $\int_A (cf + g)d\mu = c \int_A f d\mu + \int_A g d\mu$.
- Pour tout $A \in \mathcal{A}$, $\int_A f d\mu = \int_X f \mathbb{1}_A d\mu$.
- Si $A \in \mathcal{A}$ avec $\mu(A) = 0$, alors $\int_A f d\mu = 0$.
- (Relation de Chasles) Pour tous $A, B \in \mathcal{A}$ tels que $A \cap B = \emptyset$, on a

$$\int_{A \cup B} f d\mu = \int_A f d\mu + \int_B f d\mu.$$

Preuve : Nous ne démontrons que la relation de Chasles, les autres points étant laissés en exercice. On a

$$\begin{aligned} \int_{A \cup B} f d\mu &= \sum_{i=1}^k \alpha_i \mu(A_i \cap (A \cup B)) \\ &= \sum_{i=1}^k \alpha_i \mu((A_i \cap A) \cup (A_i \cap B)) \\ &= \sum_{i=1}^k \alpha_i (\mu(A_i \cap A) + \mu(A_i \cap B)) \quad (\text{union disjointe}) \\ &= \sum_{i=1}^k \alpha_i \mu(A_i \cap A) + \sum_{i=1}^k \alpha_i \mu(A_i \cap B) \\ &= \int_A f d\mu + \int_B f d\mu. \end{aligned}$$

□

3.2.2 Intégrale des fonctions positives

On étend maintenant cette définition aux fonctions mesurables positives.

Définition 3.5 Soit $f : (X, \mathcal{A}, \mu) \rightarrow [0, +\infty]$ une fonction mesurable positive. On appelle intégrale de f sur $A \in \mathcal{A}$ par rapport à μ la quantité

$$\int_A f d\mu = \sup \left\{ \int_A g d\mu : g \text{ étagée, positive, } g \leq f \right\}.$$

Notons que cette quantité est bien définie dans $\overline{\mathbb{R}}$. De plus, si f est étagée, cette définition coïncide avec la précédente. En revanche, cette définition ne permet pas en général de calculer l'intégrale de f .

Proposition 3.8 Les propriétés du lemme 3.7 s'étendent aux fonctions mesurables positives.

Preuve : Le premier point est évident car comme il s'agit d'une inégalité, la définition précédente nous permet de conclure.

En revanche, les autres points sont plus délicats car il nous faut démontrer des égalités et cela sous-entend qu'il faut peu ou prou permuter un supremum avec une intégrale, ce qui loin d'être évident. La stratégie consiste alors à utiliser le théorème de convergence monotone (ou théorème de Beppo-Levi) que l'on verra au chapitre suivant. Une fois ce théorème démontré, les résultats que l'on souhaite démontrer en découlent directement.

Regardons ce qu'il en est au sujet de la linéarité. Soient $(f_n)_{n \in \mathbb{N}}$ et $(g_n)_{n \in \mathbb{N}}$ des suites croissantes de fonctions étagées positives qui convergent simplement vers f et g respectivement. Alors $(cf_n + g_n)_{n \in \mathbb{N}}$ est une suite croissante de fonctions étagées positives qui converge vers $cf + g$. Par le théorème de convergence monotone,

$$\begin{aligned} c \int_X f d\mu + \int_X g d\mu &= c \lim_{n \rightarrow +\infty} \int_X f_n d\mu + \lim_{n \rightarrow +\infty} \int_X g_n d\mu \quad (\text{Beppo-Levi}) \\ &= \lim_{n \rightarrow +\infty} c \int_X f_n d\mu + \int_X g_n d\mu \\ &= \lim_{n \rightarrow +\infty} \left(\int_X (cf_n + g_n) d\mu \right) \quad (\text{car fonctions étagées}) \\ &= \int_X (cf + g) d\mu, \quad (\text{Beppo-Levi}) \end{aligned}$$

ce qui conclut la démonstration de la linéarité. Les autres points se démontrent de la même manière. \square

Le résultat suivant découle directement de la définition.

Proposition 3.9 (Inégalité de Chebyshev)

Soit $f : (X, \mathcal{A}, \mu) \rightarrow [0, +\infty]$ une fonction mesurable positive. Pour tout $\alpha > 0$,

$$\mu(\{x \in X \mid f(x) \geq \alpha\}) \leq \frac{1}{\alpha} \int_X f d\mu.$$

Preuve : La preuve de ce résultat, ainsi que du corollaire qui suit, sera faite en TD. \square

Corollaire 3.10 Soit $f : (X, \mathcal{A}, \mu) \rightarrow [0, +\infty]$ une fonction mesurable positive, telle que

$$\int_X f d\mu < +\infty.$$

Alors $\mu(\{x \in X \mid f(x) = +\infty\}) = 0$.

SECTION 3.3

Intégrale des fonctions de signe quelconque

3.3.1 Définition et premières propriétés

Soit f une fonction mesurable de $(X, \mathcal{A}, \mu) \rightarrow \overline{\mathbb{R}}$. On décompose f en ses parties positive et négative :

$$f = f^+ - f^-, \quad \text{avec} \quad f^+ = \sup(f, 0) \quad \text{et} \quad f^- = \sup(-f, 0).$$

Notons que l'on a aussi $|f| = f^+ + f^-$.

Définition 3.6 On dit que f est intégrable sur $A \in \mathcal{A}$ par rapport à la mesure μ si $\int_A |f| d\mu < \infty$. On définit alors l'intégrale de f sur $A \in \mathcal{A}$ par rapport à μ par

$$\int_A f d\mu = \int_A f^+ d\mu - \int_A f^- d\mu.$$

Attention : contrairement au cas des fonctions positives, on ne définit pas l'intégrale de toute fonction mesurable réelle.

Si f est à valeurs complexes, on écrit $f = \mathcal{R}ef + i\mathcal{I}mf$ et l'on définit, pour f une fonction intégrable par rapport à μ ,

$$\int_X f d\mu = \int_X (\mathcal{R}ef) d\mu + i \int_X (\mathcal{I}mf) d\mu.$$

Notation : Dans le cas $X = \mathbb{R}$ ou $X = \mathbb{R}^n$ muni de la mesure de Lebesgue, on notera souvent

$$\int_{\mathbb{R}^n} f d\lambda_n = \int_{\mathbb{R}^n} f(x) dx.$$

Proposition 3.11 *Les propriétés du lemme 3.7 s'étendent aux fonctions intégrables de signe quelconque. De surcroît, on a pour toute fonction intégrable f sur $A \in \mathcal{A}$ par rapport à μ ,*

$$\left| \int_A f d\mu \right| \leq \int_A |f| d\mu.$$

Preuve : Évidente en utilisant la proposition 3.8 pour les parties positive et négative de la fonction intégrable. \square

En pratique, on cherche souvent à intégrer des fonctions mesurables bornées. Le résultat suivant assure l'intégrabilité de telles fonctions sur des ensembles de mesure finie.

Proposition 3.12 *On suppose que la mesure μ est finie. Alors toute fonction $f : (X, \mathcal{A}, \mu) \rightarrow \mathbb{R}$ mesurable et bornée sur X est intégrable.*

Preuve : Il suffit de majorer $|f|$ par la fonction constante égale à $\sup_{x \in X} |f(x)|$ sur X , auquel cas on a

$$\int_X |f| d\mu \leq \sup_{x \in X} |f(x)| \mu(X) < +\infty,$$

car la mesure est supposée finie. \square

Cas particulier : si $f : [a, b] \rightarrow \mathbb{R}$ est une fonction continue sur un intervalle fermé borné (i.e. un intervalle compact), alors par le théorème de Weierstrass f est bornée et atteint ses bornes. La fonction f est donc intégrable sur $[a, b]$ et l'on a

$$\left| \int_a^b f(x) dx \right| \leq (b - a) \max_{x \in [a, b]} |f(x)|.$$

3.3.2 Le "presque partout"

Définition 3.7 On dit qu'une propriété est vraie presque partout ou pour presque tout $x \in X$ si elle est vraie sur $A \in \mathcal{A}$ avec $\mu(X \setminus A) = 0$. Cette notion dépendant de la mesure choisie sur X , on note μ -p.p. (ou simplement p.p. s'il n'y a pas de risque de confusion).

Lemme 3.13

- Soient f et g deux fonctions mesurables positives ou intégrables telles que $f = g$ μ -p.p. Alors $\int_X f d\mu = \int_X g d\mu$.
- Soient f et g deux fonctions mesurables positives ou intégrables telles que $f \leq g$ μ -p.p. Alors $\int_X f d\mu \leq \int_X g d\mu$.
- Si f est une fonction mesurable positive telle que $\int_X f d\mu < +\infty$, alors f est finie μ -p.p.
- Si f est une fonction mesurable positive telle que $\int_X f d\mu = 0$, alors $f = 0$ μ -p.p.

Preuve : Écrire que $f = g$ μ -p.p. signifie que $f = g$ sur un ensemble $A \in \mathcal{A}$ tel que $\mu(X \setminus A) = 0$. Ainsi, on a

$$\int_X f d\mu = \int_A f d\mu + \int_{X \setminus A} f d\mu = \int_A f d\mu + 0 = \int_A g d\mu + 0 = \int_X g d\mu.$$

Le second point se démontre de la même manière.

Le troisième, quant à lui, est la version presque partout du corollaire 3.10.

Le dernier point est un peu plus délicat. On veut montrer que $f = 0$ μ -p.p., c'est-à-dire que $\mu(\{x \in X : f(x) > 0\}) = 0$, la fonction f étant supposée positive. Notons que l'ensemble borélien $\{x \in X : f(x) > 0\} = f^{-1}(]0, +\infty[)$ peut être obtenu par double inclusion comme

$$\{x \in X : f(x) > 0\} = \bigcup_{n \in \mathbb{N}^*} \left\{ x \in X : f(x) > \frac{1}{n} \right\},$$

une union de boréliens que l'on note $\bigcup_{n \in \mathbb{N}^*} A_n$. Observons que l'on a par l'inégalité de Chebyshev : pour tout $n \in \mathbb{N}^*$,

$$\mu(A_n) \leq n \int_X f d\mu = 0,$$

l'intégrale de f sur X par rapport à μ étant supposée nulle, d'où le fait que $\mu(A_n) = 0$ pour tout $n \in \mathbb{N}^*$. Enfin, la famille $(A_n)_{n \in \mathbb{N}^*}$ étant croissante en n pour l'inclusion, on a par la proposition 2.5 que

$$\mu(\{x \in X : f(x) > 0\}) = \mu\left(\bigcup_{n \in \mathbb{N}^*} A_n\right) = \lim_{n \rightarrow +\infty} \mu(A_n) = \lim_{n \rightarrow +\infty} 0 = 0,$$

ce qui achève la démonstration. \square

Exemple 3.2 *Considérons la fonction $f = \mathbb{1}_{\mathbb{Q}}$ sur \mathbb{R} muni de la mesure de Lebesgue. On a $f = 0$ p.p. En effet,*

$$\lambda(\{x \in \mathbb{R} : f(x) = 1\}) = \lambda(\mathbb{Q}) = \lambda\left(\bigcup_{x \in \mathbb{Q}} \{x\}\right) = \sum_{x \in \mathbb{Q}} \lambda(\{x\}) = \sum_{x \in \mathbb{Q}} 0 = 0.$$

Un concept important qui sera illustré dans les chapitres à venir est la convergence presque partout.

Définition 3.1 *Une suite $(f_n)_{n \in \mathbb{N}}$ de fonctions mesurables de (X, \mathcal{A}, μ) vers \mathbb{R} converge " μ -presque partout" vers une fonction mesurable $f : (X, \mathcal{A}, \mu) \rightarrow \mathbb{R}$ si elle converge simplement vers f sur un ensemble dont le complémentaire est de mesure nulle, i.e.*

$$\mu\left(\left\{x \in X : \lim_{n \rightarrow +\infty} f_n(x) \neq f(x)\right\}\right) = 0.$$

En particulier, on a le résultat trivial suivant.

Proposition 3.1 *La convergence simple entraîne la convergence p.p.*

En pratique, on commence par tenter de montrer la convergence simple. Une fois cette analyse effectuée, on regarde en détail les différents cas obtenus pour passer à la convergence p.p.

Exemple 3.3 *Considérons sur $X = \mathbb{R}$ muni de la mesure de Lebesgue la suite de fonctions $f_n(x) = \mathbb{1}_{[-1/n, 1/n]}(x)$, $n \in \mathbb{N}^*$. On remarque que cette suite converge simplement sur \mathbb{R} vers la fonction $= \mathbb{1}_{\{0\}}$. En effet, si $x = 0$ alors $f_n(0) = 1$ pour tout $n \in \mathbb{N}^*$ tandis que si $x \neq 0$, alors il existe $n \in \mathbb{N}^*$ assez grand de sorte que $x \notin [-1/n, 1/n]$ et donc $f_n(x) = 0$ pour n assez grand.*

Qu'en est-il de la convergence p.p. ? Étant donné que nous n'avons que deux cas, $\{0\}$ et \mathbb{R}^ et que le singleton $\{0\}$ est un ensemble de mesure de Lebesgue nulle, cette suite converge p.p. vers la fonction nulle.*

Exemple 3.4 *Considérons sur $X = \mathbb{R}$ muni de la mesure de Lebesgue la suite de fonctions $f_n(x) = (\cos x)^n$, $n \in \mathbb{N}$. On remarque que cette suite converge simplement vers 1 sur l'ensemble $2\pi\mathbb{Z}$, vers 0 sur $\mathbb{R} \setminus \pi\mathbb{Z}$ car dans ce cas $\cos x \in]-1, 1[$ et enfin ne converge pas simplement sur $(2\mathbb{Z} + 1)\pi$ car dans ce cas $f_n(x) = (-1)^n$ qui n'admet pas de limite.*

Cependant, elle converge p.p. vers 0 car l'ensemble $\mathbb{Z}\pi$ est dénombrable donc de mesure de Lebesgue nulle.

Théorèmes fondamentaux d'intégration

SECTION 4.1

Comportement face aux limites

Le problème est le suivant. Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions mesurables positives ou intégrables convergeant simplement vers f . A-t-on toujours

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

Autrement dit, peut-on toujours intervertir limite et intégrale ?

Le réponse à cette question est : non, en général. Pour pouvoir intervertir limite et intégrale, il faut faire des hypothèses supplémentaires sur la suite de fonctions $(f_n)_{n \in \mathbb{N}}$. Nous allons énoncer trois résultats :

- le théorème de convergence monotone (ou théorème de Beppo-Levi, déjà entrevu au chapitre précédent) : on suppose que $(f_n)_{n \in \mathbb{N}}$ est une suite croissante de fonctions mesurables positives convergeant simplement vers f ;
- le théorème de convergence dominée : outre la convergence simple, on fait une hypothèse de domination sur la suite $(f_n)_{n \in \mathbb{N}}$;
- le lemme de Fatou : ici, on ne suppose pas la convergence de la suite. On obtient une inégalité entre les $\lim \inf$.

4.1.1 Théorème de convergence monotone de Beppo-Levi

Le théorème de convergence monotone permet d'intervertir limite et intégrale dans le cas où la suite est croissante.

Théorème 4.1 (Beppo-Levi)

Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions mesurables positives de (X, \mathcal{A}, μ) vers

$[0, +\infty]$ qui est croissante, i.e. pour tout $x \in X$, tout $n \in \mathbb{N}$, $f_n(x) \leq f_{n+1}(x)$. Alors pour tout $A \in \mathcal{A}$,

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

Preuve : Contrairement aux démonstrations des théorèmes fondamentaux à venir qui se basent sur ce résultat, la démonstration du théorème de convergence monotone est quelque peu délicate. Notons f la limite des f_n qui est mesurable et positive comme limite simple de fonctions mesurables et positives (l'ensemble des valeurs de f est a priori $[0, +\infty]$). Notons que l'on peut se restreindre au cas $A = X$ puisque la suite $(f_n \mathbb{1}_A)_{n \in \mathbb{N}}$ converge en croissant vers $f \mathbb{1}_A$.

Comme la suite $(f_n)_{n \in \mathbb{N}}$ est croissante, la suite réelle $\left(\int_X f_n d\mu \right)_{n \in \mathbb{N}}$ l'est aussi, donc $\left(\int_X f_n d\mu \right)_{n \in \mathbb{N}}$ a une limite $\alpha \in [0, +\infty]$.

De plus, pour tout $n \in \mathbb{N}$, $f_n \leq f$ donc $\int_X f_n d\mu \leq \int_X f d\mu$. À la limite $n \rightarrow +\infty$, on a $\alpha \leq \int_X f d\mu$.

Soit maintenant g une fonction étagée positive et telle que $g \leq f$, avec $g = \sum_{j=1}^m b_j \mathbb{1}_{B_j}$.

Étant donné $c \in]0, 1[$ fixé, définissons pour tout $n \in \mathbb{N}$, l'ensemble

$$C_n = \{f_n \geq cg\} = \{x \in X : f_n(x) \geq cg(x)\}.$$

Ces ensembles sont mesurables car $C_n = (f_n - cg)^{-1}([0, +\infty])$. De plus, la suite $(C_n)_{n \in \mathbb{N}}$ est croissante pour l'inclusion car la suite $(f_n)_{n \in \mathbb{N}}$ est croissante.

Par ailleurs, on a $\bigcup_{n \in \mathbb{N}} C_n = X$. En effet, si $x \in X$, on a

$$f_n(x) \xrightarrow{n \rightarrow +\infty} f(x) \geq g(x),$$

donc il existe n_0 tel que $f_{n_0}(x) \geq cg(x)$. Ainsi $x \in C_{n_0}$ et $x \in \bigcup_{n \in \mathbb{N}} C_n$.

À présent, on a pour tout $n \in \mathbb{N}$,

$$\int_X f_n d\mu \geq \int_{C_n} f_n d\mu \geq \int_{C_n} cg d\mu = c \int_{C_n} g d\mu = c \sum_{j=1}^m b_j \mu(B_j \cap C_n).$$

Or $B_j = \bigcup_{n \in \mathbb{N}} (B_j \cap C_n)$, donc

$$\mu(B_j) = \mu \left(\bigcup_{n \in \mathbb{N}} B_j \cap C_n \right) = \lim_{n \rightarrow \infty} \mu(B_j \cap C_n),$$

par la proposition 2.5. Ainsi,

$$\alpha = \lim_{n \rightarrow +\infty} \int_X f_n d\mu \geq \lim_{n \rightarrow +\infty} c \sum_{j=1}^m b_j \mu(B_j \cap C_n) = c \sum_{j=1}^m b_j \mu(B_j) = c \int_X g d\mu.$$

Ceci étant vrai pour toute fonction g étagée telle que $g \leq f$, on en déduit en passant au sup que $\alpha \geq c \int_X f d\mu$. Ceci étant vrai pour tout $c \in [0, 1[$, il suffit de faire tendre c vers 1 pour obtenir l'inégalité $\alpha \geq \int_X f d\mu$, ce qui achève la démonstration du théorème. \square

Faisons quelques remarques.

- Les séries et les intégrales pourront être traitées dans le même cadre. En effet, prenons $X = \mathbb{N}$, $\mathcal{A} = \mathcal{P}(\mathbb{N})$ et μ la mesure de comptage. Soit $(u_n)_{n \in \mathbb{N}}$ une série à termes positifs ou nuls. Posons $f(n) = u_n$. Alors la somme de la série de terme général u_n s'écrit comme une intégrale par rapport à μ :

$$\sum_{n=0}^{+\infty} u_n = \int_{\mathbb{N}} f d\mu.$$

En effet, la suite de fonctions (f_m) définie par $f_m(k) = u_k$ si $k \leq m$ et $f_m(k) = 0$ si $k > m$, est croissante. On a $\int_{\mathbb{N}} f_m d\mu = \sum_{k=0}^m u_k$. On voit aisément que $f_m \nearrow f$.

- Soit maintenant (f_n) une suite de fonctions mesurables positives et posons $f = \sum_{n=0}^{+\infty} f_n$. Alors

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

En effet, posons $g_n = \sum_{m=0}^n f_m$. On a $g_n \nearrow f$ et donc d'après le théorème de Beppo-Levi, $\lim_{n \rightarrow +\infty} \int_X g_n d\mu = \int_X \lim_{n \rightarrow +\infty} g_n d\mu$, c'est-à-dire

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

4.1.2 Lemme de Fatou

Lemme 4.2 (Lemme de Fatou)

Soit $f_n : (X, \mathcal{A}, \mu) \rightarrow [0, +\infty]$ une suite de fonctions mesurables positives. On a alors

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

Preuve : Posons pour tout $n \in \mathbb{N}$, $g_n = \inf_{m \geq n} f_m$. Alors, la suite $(g_n)_{n \in \mathbb{N}}$ est une suite de fonctions mesurables positives, qui converge en croissant vers $\sup_{n \in \mathbb{N}} g_n = \liminf_{n \rightarrow +\infty} f_n$. D'après le théorème de convergence monotone, on a donc

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

Mais comme $g_n \leq f_m$ pour tout $n \in \mathbb{N}$ et tout $m \geq n$, on a $\int_X g_n d\mu \leq \int_X f_m d\mu$ pour tout $m \geq n$ et donc $\int_X g_n d\mu \leq \inf_{m \geq n} \int_X f_m d\mu$, d'où

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

On en déduit le résultat désiré. □

4.1.3 Théorème de convergence dominée de Lebesgue

Lemme 4.3 Soit $g : (X, \mathcal{A}, \mu) \rightarrow \mathbb{R}$ une fonction intégrable et $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions intégrables de (X, \mathcal{A}, μ) vers \mathbb{R} .

- Si $g \leq f_n$ pour tout $n \in \mathbb{N}$, alors $\int_X \liminf_{n \rightarrow +\infty} f_n d\mu \leq \liminf_{n \rightarrow +\infty} \int_X f_n d\mu$.
- Si $g \geq f_n$ pour tout $n \in \mathbb{N}$, alors $\limsup_{n \rightarrow +\infty} \int_X f_n d\mu \leq \int_X \limsup_{n \rightarrow +\infty} f_n d\mu$.

Preuve : Ce lemme est une conséquence directe du lemme de Fatou. Pour tout $n \in \mathbb{N}$, $f_n - g$ est positive et mesurable, donc d'après le lemme de Fatou, on a

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

Comme $\liminf_{n \rightarrow +\infty} (f_n - g) = (\liminf_{n \rightarrow +\infty} f_n) - g$, et que l'intégrale est linéaire, on en déduit le premier point.

Pour le second, on utilise la propriété évidente :

$$\liminf_{n \rightarrow +\infty} (-f_n) = -\limsup_{n \rightarrow +\infty} f_n.$$

□

Théorème 4.4 (Théorème de convergence dominée de Lebesgue)

Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions mesurables de (X, \mathcal{A}, μ) dans \mathbb{R} . On suppose que :

- $(f_n)_{n \in \mathbb{N}}$ converge vers f μ -p.p.
- il existe une fonction intégrable $g : X \rightarrow [0, +\infty[$ telle que pour tout $n \in \mathbb{N}$, pour tout $x \in X$, $|f_n(x)| \leq g(x)$.

Alors, f_n est intégrable pour tout $n \in \mathbb{N}$, f est intégrable et

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

Notons que l'on peut supposer que la condition de domination a lieu seulement presque partout.

Preuve : La condition de domination implique que pour tout $n \in \mathbb{N}$,

$$\int_X |f_n| d\mu \leq \int_X g d\mu < +\infty,$$

donc que $|f_n|$ est intégrable. Pour μ -presque tout $x \in X$, $f(x) = \lim_{n \rightarrow +\infty} f_n(x)$, donc $|f| \leq g$ μ -p.p. Ainsi,

$$\int_X |f| d\mu \leq \int_X g d\mu < +\infty,$$

et $|f|$ est intégrable. Comme on a, pour tout $n \in \mathbb{N}$, $-g \leq f_n \leq g$, d'après le lemme 4.3 on peut écrire que

$$\begin{aligned} \limsup_{n \rightarrow +\infty} \int_X f_n d\mu &\leq \int_X \limsup_{n \rightarrow +\infty} f_n d\mu \\ &= \int_X f d\mu \\ &= \int_X \liminf_{n \rightarrow +\infty} f_n d\mu \\ &\leq \liminf_{n \rightarrow +\infty} \int_X f_n d\mu. \end{aligned}$$

Ainsi $\limsup_{n \rightarrow +\infty} \int_X f_n d\mu = \liminf_{n \rightarrow +\infty} \int_X f_n d\mu$, ce qui prouve que $\lim_{n \rightarrow +\infty} \int_X f_n d\mu$ existe et vaut, d'après les inégalités ci-dessus, $\int_X f d\mu$. \square

Exemple 4.1 Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ une fonction intégrable. Alors

$$\lim_{n \rightarrow +\infty} \int_{\mathbb{R}} f(x)(\cos x)^n dx = 0.$$

En effet, la fonction $x \mapsto f(x)(\cos x)^n$, qui est évidemment mesurable sur \mathbb{R} , converge p.p. vers 0 lorsque n tend vers l'infini (cf. l'exemple 3.4) et est dominée sur \mathbb{R} en valeur absolue par $|f|$ qui est intégrable par hypothèse et indépendante de n . Le théorème de convergence dominée donne alors le résultat désiré.

SECTION 4.2

Lien avec l'intégrale de Riemann

On va voir ci-dessous que l'intégrale de Lebesgue généralise l'intégrale de Riemann lorsque X est un intervalle ou un produit d'intervalles dans le cas

multidimensionnel. En particulier, on verra en TD qu'en général pour démontrer qu'une fonction mesurable est intégrable au sens de Lebesgue, il est la plupart du temps suffisant de démontrer que sa valeur absolue est intégrable au sens de Riemann.

Nous nous limiterons ici à la dimension 1, les démonstrations se généralisant sans problème à la dimension $n \geq 2$. Pour simplifier, nous nous limiterons également aux fonctions positives.

Théorème 4.5 *Soit $[a, b]$ un segment de \mathbb{R} et $f : [a, b] \rightarrow \mathbb{R}$ une fonction positive, bornée, et intégrable au sens Riemann sur $[a, b]$. Alors f est intégrable sur $[a, b]$ par rapport à la mesure de Lebesgue, et les intégrales sont les mêmes.*

Preuve : Rappelons tout d'abord la définition de " f Riemann-intégrable". Pour toute subdivision pointée de $[a, b]$, i.e. la donnée de $\sigma = (x_j)_{0 \leq j \leq n}$, $\xi = (\xi_j)_{1 \leq j \leq n}$ tels que $a = x_0 < x_1 \dots < x_n = b$ et $\xi_j \in]x_{j-1}, x_j[$, on pose

$$I(f, \sigma, \xi) = \sum_{j=1}^n (x_j - x_{j-1}) f(\xi_j). \text{ On note } |\sigma| = \sup_j x_j - x_{j-1} \text{ le pas de la}$$

subdivision. La fonction f est intégrable au sens de Riemann sur $[a, b]$ si et seulement si pour tous σ, ξ , la quantité $I(f, \sigma, \xi)$ tend vers une unique limite I quand $|\sigma| \rightarrow 0$. Le temps de cette preuve, on notera $I = \mathcal{R}(f, a, b)$.

Revenons à la démonstration. Soit f positive, Riemann-intégrable et bornée sur $[a, b]$. On considère la suite de subdivisions emboîtées définies pour chaque n par :

$$u_0^n = a < u_1^n < \dots < u_{2^n}^n = b \quad \left(\text{de pas } \frac{b-a}{2^n} \right).$$

On pose pour tout $n \in \mathbb{N}$, $1 \leq j \leq 2^n$,

$$m_j^n = \inf_{x \in [u_{j-1}^n, u_j^n]} f(x) \text{ et } M_j^n = \sup_{x \in [u_{j-1}^n, u_j^n]} f(x).$$

Les deux sommes :

$$I_n^- = \frac{b-a}{2^n} \sum_{j=1}^{2^n} m_j^n \quad \text{et} \quad I_n^+ = \frac{b-a}{2^n} \sum_{j=1}^{2^n} M_j^n$$

convergent toutes deux vers l'intégrale de Riemann $\mathcal{R}(f, a, b)$. Posons maintenant

$$g_n = m_1^n \mathbb{1}_{[u_0^n, u_1^n]} + \sum_{j=2}^{2^n} m_j^n \mathbb{1}_{[u_{j-1}^n, u_j^n]}$$

et

$$h_n = M_1^n \mathbb{1}_{[u_0^n, u_1^n]} + \sum_{j=2}^{2^n} M_j^n \mathbb{1}_{[u_{j-1}^n, u_j^n]}.$$

On a pour tout n , $g_n \leq f \leq h_n$. De plus, (g_n) est croissante alors que (h_n) est décroissante. Si g et h sont leurs limites (simples) respectives (notons que g et h sont mesurables comme limites de fonctions mesurables), alors on a $g \leq f \leq h$.

Les suites (g_n) et (h_n) sont majorées par $\sup_{x \in [a, b]} f(x)$ qui est intégrable sur $[a, b]$ par rapport à la mesure de Lebesgue. Ainsi par le théorème de convergence dominée les suites (I_n^-) et (I_n^+) convergent respectivement vers $\int_a^b g d\lambda$ et $\int_a^b h d\lambda$, qui sont donc toutes deux égales à l'intégrale de Riemann $\mathcal{R}(f, a, b)$. Ainsi, $g = h = f$ μ -pp, et $\mathcal{R}(f, a, b) = \int_a^b f d\lambda$. \square

On rappelle l'extension de l'intégrale de Riemann à un intervalle non borné. On se limite encore aux fonctions positives. L'intégrale au sens de Riemann ainsi définie : $\int_0^{+\infty} f(x) dx = \lim_{b \rightarrow +\infty} \int_0^b f(x) dx$ est là aussi égale à l'intégrale de Lebesgue de f sur $[0, +\infty[$. En effet, définissons la suite (f_n) par $f_n = f \mathbb{1}_{[0, n]}$. Comme f est positive, $\int_0^{+\infty} f(x) dx = \lim_{n \rightarrow +\infty} \int_0^n f(x) dx = \lim_{n \rightarrow +\infty} \int_0^{+\infty} f_n(x) dx$. Pour tout n , f_n est intégrable au sens de Lebesgue d'après le théorème précédent, et comme $f_n \nearrow f$, on obtient le résultat par le théorème de convergence monotone.

On procéderait de manière analogue pour les autres extensions (fonction non définie aux bornes de l'intervalle d'intégration).

Attention : dans le cas où f n'est pas positive, ce que l'on vient de faire se généralise au cas des intégrales absolument convergentes. Les intégrales semi-convergentes ne rentrent pas dans ce cadre.

Exemple 4.2 La fonction $x \mapsto \frac{\sin x}{x}$ n'est pas intégrable sur $[0, +\infty[$, i.e.

$$\int_0^{+\infty} \left| \frac{\sin x}{x} \right| dx = +\infty, \text{ mais on peut donner un sens à } \int_0^{+\infty} \frac{\sin x}{x} dx \text{ comme}$$

$$\lim_{M \rightarrow +\infty} \int_0^M \frac{\sin x}{x} dx.$$

SECTION 4.3

Fonctions définies par une intégrale

4.3.1 Théorème de continuité

Les théorèmes ci-dessous permettent d'invertir les notions de continuité/dérivabilité et d'intégrale, et à ce titre sont très importants.

Théorème 4.6 *Soit (X, \mathcal{A}, μ) un espace mesuré, I un ouvert de \mathbb{R} contenant t_0 , et $f : I \times X \rightarrow \mathbb{R}$ une fonction telle que*

- pour tout $t \in I$, $x \mapsto f(t, x)$ est mesurable ;
- la fonction $t \mapsto f(t, x)$ est continue en t_0 pour μ -presque tout $x \in X$;
- il existe $g : X \rightarrow [0, +\infty]$ telle que $\int_X g d\mu < +\infty$ et $|f(t, x)| \leq g(x)$ pour tout $t \in I$ et μ -presque tout $x \in X$.

Alors la fonction $F : I \rightarrow \mathbb{R}$, $F(t) = \int_X f(t, x) d\mu(x)$ est bien définie sur I et continue en t_0 .

Preuve : Tout d'abord, la condition de domination assure que pour tout $t \in I$, $x \mapsto f(t, x)$ est intégrable, donc $F(t)$ est bien défini.

Pour montrer ce théorème, on utilise la caractérisation de la continuité à l'aide des suites. Soit (t_n) une suite de I tendant vers t_0 . Pour tout n , on note $f_n : X \rightarrow \mathbb{R}$, $f_n(x) = f(t_n, x)$. Alors, f_n est majorée par g presque partout et $\lim_{n \rightarrow +\infty} f_n(x) = f(t_0, x)$ pour presque tout x . Par le théorème de convergence dominée de Lebesgue, on obtient donc que $F(t_n) \rightarrow F(t_0)$. Ceci étant vrai pour toute suite (t_n) tendant vers t_0 , F est continue en t_0 . \square

4.3.2 Théorème de dérivabilité

Théorème 4.7 *Soit (X, \mathcal{A}, μ) un espace mesuré, I un intervalle ouvert de \mathbb{R} , et $f : I \times X \rightarrow \mathbb{R}$ une fonction tels que :*

- pour tout $t \in I$, la fonction $x \mapsto f(t, x)$ est mesurable ;
- il existe $t_0 \in I$ tel que $x \mapsto f(t_0, x)$ est intégrable ;

- pour presque tout x , $t \mapsto f(t, x)$ est de classe C^1 sur I ;
- il existe $g : X \rightarrow [0, +\infty]$ telle que $\int_X g d\mu < +\infty$ et pour tout $t \in I$ et pour μ -presque tout x ,

$$\left| \frac{\partial f}{\partial t}(t, x) \right| \leq g(x).$$

Alors, la fonction $F : I \rightarrow \mathbb{R}$, $F(t) = \int_X f(t, x) d\mu$ est bien définie et de classe C^1 sur I , de dérivée

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

Preuve : Soit $t \in I$, (t_n) une suite de I tendant vers t . Justifions d'abord que la dernière intégrale ci-dessus existe. La fonction $x \mapsto \frac{\partial f}{\partial t}(t, x)$ étant la limite simple de $\varphi_n : x \mapsto \frac{f(t_n, x) - f(t, x)}{t_n - t}$ qui est mesurable, elle est donc mesurable elle aussi. De plus, elle est majorée par g donc intégrable.

À présent, soient $x \in X$ et $t' \in I$. Par le théorème des accroissements finis, il existe $u \in]t, t'[$ (ou $]t', t[$) tel que

$$\left| \frac{f(t', x) - f(t, x)}{t' - t} \right| = \left| \frac{\partial f}{\partial t}(u, x) \right| \leq g(x).$$

- Avec $t' = t_0$, on obtient que

$$|f(t, x)| \leq g(x)|t - t_0| + |f(t_0, x)|,$$

ce qui assure que $F(t)$ est bien définie dans \mathbb{R} .

- Pour tout n , $\frac{F(t_n) - F(t)}{t_n - t} = \int_X \varphi_n d\mu$. Or, $|\varphi_n| \leq g$ (prendre $t' = t_n$), et (φ_n) tend simplement vers la fonction $x \mapsto \frac{\partial f}{\partial t}(t, x)$. Ainsi, par le théorème de convergence dominée,

$$\lim_{n \rightarrow +\infty} \int_X \varphi_n d\mu = \int_X \frac{\partial f}{\partial t}(t, x) d\mu(x).$$

Comme $F'(t) = \lim_{n \rightarrow +\infty} \frac{F(t_n) - F(t)}{t_n - t}$, ceci achève la démonstration (F est alors bien de classe C^1 car la continuité de F' est immédiate par l'utilisation

du théorème de continuité au niveau de la dérivée). \square

Notons que l'on peut remplacer la 4ème hypothèse par
 • pour tout intervalle compact K inclus dans I , il existe g_K intégrable sur X telle que $\forall t \in K$, $|\frac{\partial f}{\partial t}(t, x)| \leq g_K(x)$ μ -p.p. sur X .

Exemple 4.3 Posons $\Gamma(t) = \int_0^{+\infty} x^{t-1} e^{-x} dx$ (fonction Gamma d'Euler).

○ Quel est le domaine de définition de Γ , c'est-à-dire l'ensemble des valeurs de t pour lesquelles la fonction $f_t : x \mapsto x^{t-1} e^{-x}$ est intégrable ? Pour tout $t \in \mathbb{R}$, la fonction f_t est mesurable sur l'intervalle $]0, +\infty[$ car continue sur cet intervalle et une courte analyse indique que f_t est intégrable sur $]0, +\infty[$ si et seulement si $t > 0$. Ainsi, la fonction Γ est bien définie sur $]0, +\infty[$.

○ Est-elle continue sur $]0, +\infty[$? Tout d'abord pour tout $x > 0$, la fonction $t \mapsto f_t(x)$ est continue en $t > 0$. Ensuite pour tout $t \in [a, b] \subset]0, +\infty[$, on a pour tout $x > 0$,

$$|x^{t-1} e^{-x}| \leq x^{a-1} \mathbb{1}_{[0,1]}(x) + \mathbb{1}_{[1,\infty[}(x) x^{b-1} e^{-x},$$

qui est intégrable sur $]0, +\infty[$ et indépendant du paramètre t . Ainsi, par le théorème de continuité, Γ est continue sur l'intervalle $[a, b]$, donc sur $]0, +\infty[$, a et b étant arbitrairement choisis dans $]0, +\infty[$. Notons qu'il aurait été difficile de trouver une fonction intégrable dominant $|f_t|$ indépendamment de $t > 0$, d'où l'intérêt de localiser t sur l'intervalle compact $[a, b]$. Ce type de localisation est couramment utilisé en pratique.

○ La fonction Γ est-elle de classe C^1 sur $]0, +\infty[$? Pour tout $x > 0$, la fonction $t \mapsto f_t(x)$ est de classe C^1 sur $]0, +\infty[$ et

$$\frac{\partial f_t(x)}{\partial t} = \ln(x) e^{-x} x^{t-1}.$$

De plus, pour tout $t \in [a, b]$, on a

$$|\ln(x) e^{-x} x^{t-1}| \leq |\ln(x)| e^{-x} (\mathbb{1}_{[0,1]}(x) x^{a-1} + \mathbb{1}_{[1,+\infty[}(x) x^{b-1}),$$

qui est intégrable sur $]0, +\infty[$ et indépendant du paramètre t . Ainsi, par le théorème de dérivabilité, Γ est de classe C^1 sur l'intervalle $]0, +\infty[$ (car C^1 sur $[a, b]$ où $0 < a < b$ sont arbitraires) et pour tout $t > 0$,

$$\Gamma'(t) = \int_0^{+\infty} \ln(x) e^{-x} x^{t-1} dx.$$

Intégration sur les espaces produits

L'objectif de cette section est de calculer des intégrales de plusieurs variables et notamment d'échanger l'ordre des intégrales. Pour ce faire, nous avons besoin de plusieurs résultats théoriques pas très difficiles à démontrer mais malgré tout assez techniques : soit nous les admettrons, soit nous en donnerons seulement une esquisse de démonstration.

4.4.1 Introduction

Soit à calculer $I = \int \int_{[0,1] \times [0,1]} \frac{x^2 - y^2}{(x^2 + y^2)^2} dx dy$.

On commence par intégrer par rapport à x , on trouve :

$$I = \int_0^1 \left[\int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dx \right] dy = \int_0^1 \left[\frac{-x}{x^2 + y^2} \right]_{x=0}^1 dy = \int_0^1 \frac{-1}{1 + y^2} dy = -\frac{\pi}{4}.$$

En intervertissant les ordres d'intégration on aurait trouvé

$$I = \int_0^1 \left[\int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dy \right] dx = \int_0^1 \left[\frac{y}{x^2 + y^2} \right]_{y=0}^1 dx = \frac{\pi}{4}$$

Ce paradoxe nous prouve en fait que I n'a pas de sens. La suite va nous procurer un théorème simple d'emploi pour effectuer les calculs en toute quiétude.

4.4.2 Tribu produit et mesure produit

Définition 4.1 Soient (X, \mathcal{A}) et (Y, \mathcal{B}) deux espaces mesurables. On appelle tribu produit la tribu engendrée par les produits $\{A \times B : A \in \mathcal{A}, B \in \mathcal{B}\}$. Elle est notée $\mathcal{A} \otimes \mathcal{B}$.

Exemple 4.4 $\mathcal{B}(\mathbb{R}^{n+p}) = \mathcal{B}(\mathbb{R}^n) \otimes \mathcal{B}(\mathbb{R}^p)$.

Notons que l'on a besoin de considérer la tribu engendrée car la famille $\{A \times B : A \in \mathcal{A}, B \in \mathcal{B}\}$ n'est pas une tribu, n'étant pas stable par passage au complémentaire.

Définition 4.2 Soient $E \in \mathcal{A} \otimes \mathcal{B}$ et $f : (X \times Y, \mathcal{A} \otimes \mathcal{B}) \rightarrow (Z, \mathcal{C})$.

On note

$$E_x = \{y \in Y : (x, y) \in E\}, \quad \text{et} \quad E^y = \{x \in X : (x, y) \in E\},$$

appelées sections de E , et

$$f_x : (Y, \mathcal{B}) \rightarrow (Z, \mathcal{C}), \quad y \mapsto f(x, y), \quad \text{et} \quad f^y : (X, \mathcal{A}) \rightarrow (Z, \mathcal{C}), \quad x \mapsto f(x, y).$$

Lemme 4.8 (admis) Soit $f : (X \times Y, \mathcal{A} \otimes \mathcal{B}) \rightarrow (Z, \mathcal{C})$ une fonction mesurable. Soit $E \in \mathcal{A} \otimes \mathcal{B}$. Alors,

1. Pour tout $x \in X$, $E_x \in \mathcal{B}$, et pour tout $y \in Y$, $E^y \in \mathcal{A}$.
2. Pour tout $x \in X$, $f_x : (Y, \mathcal{B}) \rightarrow (Z, \mathcal{C})$ est mesurable et pour tout $y \in Y$, $f^y : (X, \mathcal{A}) \rightarrow (Z, \mathcal{C})$ est mesurable.

Proposition 4.9 (admis) Soient (X, \mathcal{A}, μ) et (Y, \mathcal{B}, ν) deux espaces mesurés de mesures σ -finies et $E \in \mathcal{A} \otimes \mathcal{B}$. Alors,

1. Pour tout $x \in X$, $x \mapsto \nu(E_x)$ est mesurable de $(X, \mathcal{A}) \rightarrow [0, +\infty]$.
2. Pour tout $y \in Y$, $y \mapsto \mu(E^y)$ est mesurable de $(Y, \mathcal{B}) \rightarrow [0, +\infty]$.

Théorème 4.10 (admis) Soient (X, \mathcal{A}, μ) et (Y, \mathcal{B}, ν) deux espaces mesurés de mesures σ -finies. Il existe une unique mesure $\mu \otimes \nu$ sur $(X \times Y, \mathcal{A} \otimes \mathcal{B})$ telle que pour tous $A \in \mathcal{A}$, $B \in \mathcal{B}$,

$$\mu \otimes \nu(A \times B) = \mu(A)\nu(B).$$

Cette mesure est appelée mesure produit. Elle est définie par :

$$\mu \otimes \nu(E) = \int_X \nu(E_x) d\mu(x) = \int_Y \mu(E^y) d\nu(y), \quad \forall E \in \mathcal{A} \otimes \mathcal{B},$$

ce qui peut encore s'écrire

$$\mu \otimes \nu(E) = \int_X \left[\int_Y \mathbb{1}_E(x, y) d\nu(y) \right] d\mu(x) = \int_Y \left[\int_X \mathbb{1}_E(x, y) d\mu(x) \right] d\nu(y).$$

Exemple 4.5 La mesure de Lebesgue λ_{n+p} définie sur $\mathcal{B}(\mathbb{R}^{n+p})$ est égale à la mesure produit $\lambda_n \otimes \lambda_p$:

$$\lambda_{n+p} = \lambda_n \otimes \lambda_p.$$

4.4.3 Théorème de Fubini

Théorème 4.11 (Théorème de Fubini-Tonelli)

Soient (X, \mathcal{A}, μ) et (Y, \mathcal{B}, ν) deux espaces mesurés de mesures σ -finies et $f : (X \times Y, \mathcal{A} \otimes \mathcal{B}) \rightarrow [0, +\infty]$ une fonction mesurable positive. Alors :

- La fonction $x \mapsto \int_Y f_x d\nu = \int_Y f(x, y) d\nu(y)$ est mesurable sur (X, \mathcal{A}) ,
- La fonction $y \mapsto \int_X f^y d\mu = \int_X f(x, y) d\mu(x)$ est mesurable sur (Y, \mathcal{B}) .
- On a l'égalité

$$\int_{X \times Y} f d(\mu \otimes \nu) = \int_X \left(\int_Y f_x d\nu \right) d\mu = \int_Y \left(\int_X f^y d\mu \right) d\nu,$$

ce qui se réécrit

$$\begin{aligned} \int_{X \times Y} f(x, y) d(\mu \otimes \nu)(x, y) &= \int_X \left(\int_Y f(x, y) d\nu(y) \right) d\mu(x) \\ &= \int_Y \left(\int_X f(x, y) d\mu(x) \right) d\nu(y). \end{aligned}$$

Preuve : Le schéma de la preuve est le schéma habituel. On montre le résultat pour

- $f = \mathbb{1}_A$, puis
- f étagée (par linéarité), puis
- f positive (par convergence monotone).

La seule difficulté réside dans l'établissement du résultat pour $\mathbb{1}_A$. C'est exactement le théorème 4.10, que nous avons admis. \square

On généralise immédiatement ce résultat aux fonctions intégrables de signe quelconque en l'utilisant pour les parties positive et négative. On a alors le théorème suivant.

Théorème 4.12 (Théorème de Fubini)

Soient (X, \mathcal{A}, μ) et (Y, \mathcal{B}, ν) deux espaces mesurés de mesures σ -finies et $f : (X \times Y, \mathcal{A} \otimes \mathcal{B}) \rightarrow \mathbb{R}$ mesurable telle que $\int_{X \times Y} |f(x, y)| d(\mu \otimes \nu)(x, y) < +\infty$ (i.e. f est $\mu \otimes \nu$ -intégrable), alors

- pour μ -presque tout $x \in X$, f_x est ν -intégrable et pour ν -presque tout $y \in Y$, f^y est μ -intégrable.
- $x \mapsto \int_Y f_x d\nu$ est définie presque partout et est μ -intégrable,
 $y \mapsto \int_X f^y d\mu$ est définie presque partout et est ν -intégrable.
- on a l'égalité

$$\begin{aligned} \int_{X \times Y} f(x, y) d(\mu \otimes \nu)(x, y) &= \int_X \left(\int_Y f(x, y) d\nu(y) \right) d\mu(x) \\ &= \int_Y \left(\int_X f(x, y) d\mu(x) \right) d\nu(y). \end{aligned}$$

En prenant pour μ et ν la mesure de comptage sur \mathbb{N} , on obtient le résultat suivant.

Corollaire 4.13 Pour toute suite double $(a_{m,n})_{m,n \in \mathbb{N}}$ de nombres réels ou complexes vérifiant :

- $a_{m,n} \geq 0, \forall m, n \in \mathbb{N}$,
- ou $\sum_{m,n} |a_{m,n}| < +\infty$,

on a alors : $\sum a_{m,n} = \sum_m (\sum_n a_{m,n}) = \sum_n (\sum_m a_{m,n})$.

Contre-exemple :

$$a_{m,n} = \begin{cases} 1 & \text{si } m = n ; \\ -1 & \text{si } m = n + 1 ; \\ 0 & \text{sinon.} \end{cases}$$

On vérifie aisément que: $\sum_m (\sum_n a_{m,n}) \neq \sum_n (\sum_m a_{m,n})$.

En pratique, après avoir vérifié les hypothèses μ, ν σ -finies et f mesurable, on prend la valeur absolue et l'on calcule $\int_{X \times Y} |f| d(\mu \otimes \nu)$ dans n'importe quel ordre (par rapport à x puis à y , ou le contraire). Si l'on trouve une valeur finie, on peut alors appliquer le théorème Fubini pour calculer $\int_{X \times Y} f d(\mu \otimes \nu)$.

SECTION 4.5

Changement de variable

Dans cette section, on cherche à transporter des mesures sur un autre ensemble par le biais d'une application.

Nous allons d'abord considérer des changements de variables dans un cadre mesurable. Bien sûr, dans un cadre abstrait, il n'est pas question de parler de difféomorphisme, de jacobien... Nous ferons le lien avec les notions habituelles de changement de variables ensuite seulement.

4.5.1 Image d'une mesure par une fonction mesurable et théorème de transport

Proposition-définition 4.14 *Soit f une application mesurable de (X, \mathcal{A}, μ) dans (Y, \mathcal{B}) . Posons*

$$\mu_f(B) = \mu(f^{-1}(B)), \quad B \in \mathcal{B}.$$

Alors μ_f définit une mesure sur (Y, \mathcal{B}) . C'est la "mesure-image" de μ par f .

Preuve : Il est immédiat que μ_f est une mesure puisque f^{-1} conserve les opérations ensemblistes. □

On donne maintenant des exemples de mesure image.

Exemple 4.6

- *Plongement :*

Soit $f : [-1, 1] \mapsto \mathbb{R}^2$, $f(x) = (x, y)$ étant l'unique point du demi-cercle $\mathcal{C} = \{(x, y); x^2 + y^2 = 1, y > 0\}$ d'abscisse x .

On munit $[-1, 1]$ de la mesure de Lebesgue λ . La mesure λ_f (mesure sur \mathbb{R}^2) est portée par \mathcal{C} . Elle n'a pas de densité sur \mathbb{R}^2 . Elle ne peut

être connue que par paramétrage de la courbe \mathcal{C} . Par exemple, pour tout borélien A de \mathbb{R}^2 , $\lambda_f(A) = \int \mathbb{1}_A(x, \sqrt{1-x^2}) dx$.

On peut encore poser $x = \cos t$ si $t \in [0, \pi]$. On a alors $\lambda_f(A) = \int \mathbb{1}_A(\cos t, \sin t) \sin t dt$.

- *Projection :*

Soit μ la mesure de Lebesgue sur le disque $\mathcal{D} = \mathcal{D}(0, 1)$ de \mathbb{R}^2 . Soit $f : \mathcal{D} \rightarrow \mathbb{R}$, $(x, y) \mapsto x$. La mesure μ_f (sur \mathbb{R}) est donnée par

$$\mu_f([a, b]) = \int_{\mathbb{R}^2} \mathbb{1}_{[a, b]}(x) \underbrace{\mathbb{1}_{\mathcal{D}}(x, y) dx dy}_{d\mu(x, y)},$$

soit, en utilisant Fubini-Tonelli,

$$\mu_f([a, b]) = \int_a^b \left(\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} dy \right) \mathbb{1}_{[-1, 1]}(x) dx = \int_a^b 2\sqrt{1-x^2} \mathbb{1}_{[-1, 1]}(x) dx.$$

Donc μ_f a pour densité $h(x) = 2\sqrt{1-x^2} \mathbb{1}_{[-1, 1]}(x)$ par rapport à la mesure de Lebesgue sur \mathbb{R} .

- *Probabilités :*

La loi d'une variable aléatoire est la mesure image de la mesure de probabilité \mathbb{P} par cette variable aléatoire, cf. le chapitre 8.

Théorème 4.15 (Théorème de transport)

Soit $f : (X, \mathcal{A}, \mu) \rightarrow \mathbb{R}$ une fonction mesurable et μ_f la mesure image de μ par f .

- Pour toute fonction $\Phi : \mathbb{R} \rightarrow [0, +\infty]$ mesurable positive, on a

$$\int_{\mathbb{R}} \Phi(t) d\mu_f(t) = \int_X \Phi \circ f(x) d\mu(x).$$

- Pour toute fonction $\Phi : \mathbb{R} \rightarrow \mathbb{R}$ mesurable, Φ est μ_f -intégrable si et seulement si $\Phi \circ f$ est μ -intégrable. Dans ce cas, on a

$$\int_{\mathbb{R}} \Phi(t) d\mu_f(t) = \int_X \Phi \circ f(x) d\mu(x).$$

Preuve : La démonstration de ce théorème suit un schéma classique. On choisit tout d'abord $\Phi = \mathbb{1}_B$. Comme $\mathbb{1}_B(f(x)) = \mathbb{1}_{f^{-1}(B)}(x)$, on a

$$\int_{\mathbb{R}} \mathbb{1}_B d\mu_f = \mu_f(B) = \mu(f^{-1}(B)) = \int_X \mathbb{1}_B(f(x)) d\mu(x).$$

On continue la démonstration avec les fonctions Φ étagées, puis positives (qui sont limite croissante de fonctions étagées) à l'aide du théorème de Beppo-Levi pour conclure.

Enfin, on traite le cas des fonctions intégrables de signe quelconque ($\Phi = \Phi^+ - \Phi^-$). \square

4.5.2 Théorème de changement de variable

Théorème 4.16 *Soit $T \in \mathcal{M}_n(\mathbb{R})$ une matrice inversible. Alors*

$$(\lambda_n)_T = |\det T|^{-1} \lambda_n.$$

Autrement dit, pour tout borélien A de \mathbb{R}^n ,

$$\lambda_n(T(A)) = |\det T| \lambda_n(A).$$

Preuve : Notons $\mu(A) = \lambda_n(T(A))$. On vérifie aisément que μ est une mesure sur $\mathcal{B}(\mathbb{R}^n)$, finie sur les compacts et invariante par translation. Ainsi, il existe $c_T \in \mathbb{R}$ tel que $\mu = c_T \lambda_n$. En effet,

- $\mu(\emptyset) = 0$.
- Si les A_n sont mesurables 2 à 2 disjoints alors les $T(A_n)$ le sont également (car T est inversible), donc

$$\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \lambda_n\left(\bigcup_{n \in \mathbb{N}} T(A_n)\right) = \sum_{n \in \mathbb{N}} \lambda_n(T(A_n)) = \sum_{n \in \mathbb{N}} \mu(A_n).$$

- Si K est un compact de \mathbb{R}^n , $T(K)$ l'est aussi donc $\mu(K) = \lambda_n(T(K)) < +\infty$.
- $\mu(A + a) = \lambda_n(T(A + a)) = \lambda_n(T(A) + T(a)) = \lambda(T(A)) = \mu(A)$.

Il reste à calculer c_T . Or $c_T = \lambda_n(T([0, 1]^n))$. L'idée est la suivante (nous ne détaillerons pas la preuve ici). On décompose T en produit de matrices élémentaires, pour lesquelles il est facile de calculer $T([0, 1]^n)$. \square

L'interprétation géométrique de ce résultat est le suivant. Prenons pour A le cube unité $C = [0, 1]^n$. L'image par T de C est le parallélépipède dont les arêtes sont données par les vecteurs colonnes de T . Le théorème conclut que le volume (non orienté) de ce parallélépipède est le déterminant (non signé) de T .

Ce résultat constitue la brique élémentaire de la démonstration du théorème de changement de variable (habituel) suivant.

Théorème 4.17 (*admis*)

Soit U, V deux ouverts de \mathbb{R}^n et $f : U \rightarrow V$ un \mathcal{C}^1 -difféomorphisme.

On note $J_f(x) = \det \left(\frac{\partial f_i}{\partial x_j}(x) \right)_{1 \leq i, j \leq n}$ le jacobien de f au point $x \in U$. Dans la suite, λ_n , dx , dy désignent la mesure de Lebesgue sur \mathbb{R}^n .

- Pour tout borélien $B \subset U$, $\lambda_n(f(B)) = \int_B |J_f(x)| dx$.
- pour toute fonction $\Phi : V \rightarrow [0, +\infty]$ mesurable positive, on a

$$\int_V \Phi(y) dy = \int_U \Phi(f(x)) |J_f(x)| dx.$$

- pour toute fonction $\Phi : V \rightarrow \mathbb{R}$ mesurable, on a
- $$\int_V |\Phi(y)| dy < +\infty \iff \int_U |\Phi \circ f(x)| |J_f(x)| dx < +\infty. \text{ Et alors,}$$

$$\int_V \Phi(y) dy = \int_U \Phi(f(x)) |J_f(x)| dx.$$

Exercice : On sait que $\int_{\mathbb{R}} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}$. Soit q la forme quadratique sur \mathbb{R}^n $q(x) = \langle Mx, x \rangle$ où M est une matrice symétrique définie positive de taille n . Montrer que

$$\int_{\mathbb{R}^n} e^{-\frac{1}{2}q(x)} dx = \frac{(2\pi)^{\frac{n}{2}}}{(\det M)^{\frac{1}{2}}}.$$

Espaces de Lebesgue

Dans la suite du cours, nous noterons \mathcal{L} l'ensemble des fonctions mesurables de l'espace mesuré (X, \mathcal{A}, μ) dans \mathbb{R} .

SECTION 5.1

Espaces et normes L^p

5.1.1 Objectif

Nous voulons introduire des normes sur l'ensemble des fonctions mesurables \mathcal{L} . Comme une norme définit une distance qui elle même définit une topologie (les ouverts), la norme définit en fait le concept de "convergence". Ainsi, si nous avons à notre disposition plusieurs normes sur le même espace, nous avons alors plusieurs notions de convergence, qui d'ailleurs peuvent ne pas être comparables. On rappelle la définition d'une norme.

Définition 5.1 Une norme N sur un espace vectoriel E vérifie les propriétés suivantes :

- Pour tout $f \in E$, $N(f) \in [0, +\infty[$, notamment $N(f) \neq +\infty$.
- Définie : $N(f) = 0 \Rightarrow f = 0$.
- Homogénéité : Pour tout $f \in E$, $\lambda \in \mathbb{R}$, $N(\lambda f) = |\lambda|N(f)$.
- Inégalité triangulaire : Pour tout $(f, g) \in E$, $N(f + g) \leq N(f) + N(g)$.

On va de plus vouloir vérifier que l'espace vectoriel E muni de cette norme est complet, c'est-à-dire que toute de Cauchy de E (pour cette norme) converge vers une limite dans E : on parle alors d'espace de Banach. On rappelle qu'une suite $(f_n)_{n \in \mathbb{N}}$ d'éléments de l'espace E muni de la norme N est dite de Cauchy si

$$\lim_{m, n \rightarrow +\infty} N(f_m - f_n) = 0.$$

Comme une suite $(f_n)_{n \in \mathbb{N}}$ d'éléments de E muni de la norme N converge vers une limite f si

$$\lim_{n \rightarrow +\infty} N(f_n - f) = 0,$$

l'inégalité triangulaire entraîne immédiatement que toute suite convergente est de Cauchy.

5.1.2 Définition de l'espace L^p

Soit $1 \leq p < +\infty$ et soit

$$N(f) = \left(\int_X |f(x)|^p d\mu(x) \right)^{1/p},$$

définie pour tout $f \in \mathcal{L}$. On a bien $N(f) \geq 0$, mais cependant pour certaines fonctions de \mathcal{L} , on a $N(f) = +\infty$. Nous allons limiter l'espace E à

$$E = \{f \in \mathcal{L} : N(f) < +\infty\}.$$

L'ensemble E est un espace vectoriel dès que nous démontrons que N est homogène et vérifie l'inégalité triangulaire. L'homogénéité de N est très simple car :

$$N(\lambda f) = \left(\int_X |\lambda f(x)|^p d\mu(x) \right)^{1/p} = |\lambda| \left(\int_X |f(x)|^p d\mu(x) \right)^{1/p} = |\lambda| N(f).$$

Nous démontrerons l'inégalité triangulaire plus tard. Nous nous intéressons maintenant à montrer que N est définie, c'est-à-dire que $N(f) = 0 \Rightarrow f = 0$. Mais cela est manifestement faux, car

$$N(f) = 0 \implies |f|^p = 0 \text{ } \mu - \text{p.p.} \implies |f| = 0 \text{ } \mu - \text{p.p.}$$

Donc f est nulle presque partout et non en tout point. Par exemple prenons comme espace de départ \mathbb{R} avec la mesure de Lebesgue et $f = \mathbf{1}_{\{0\}}$, alors $N(f) = 0$ mais $f \neq 0$. Donc N ne peut pas définir une norme sur notre espace E .

Dans ce cas, l'idée est de restreindre notre espace E pour que N soit bien une norme. Supposons que μ soit la mesure de Lebesgue sur $X \subset \mathbb{R}$ et définissons

$$E = \{f \in \mathcal{L} \text{ et continue} : N(f) < +\infty\}.$$

On vient donc de rajouter la continuité aux fonctions considérées. Dans ce cas effectivement, si $N(f) = 0$ alors $f = 0$ μ -p.p. et donc $f = 0$ en tout point car f est continue. Ainsi, quand on aura montré que N vérifie l'inégalité

triangulaire, on aura montré que N est une norme sur E . Cependant cet espace n'est pas forcément complet. Par exemple, prenons $X = [-1, 1]$ et $p = 1$ ainsi

$$E = \{f \in \mathcal{L} \text{ et continue sur } [-1, 1] : \int_{-1}^1 |f(x)| dx < +\infty\}.$$

On considère la suite de fonction (f_n) définie par

$$f_n(x) = \begin{cases} -1 & \text{si } x \leq -\frac{1}{n}, \\ nx & \text{si } -\frac{1}{n} < x < \frac{1}{n}, \\ 1 & \text{si } x \geq \frac{1}{n}. \end{cases}$$

La suite (f_n) converge simplement vers la fonction $f(x) = -\mathbb{1}_{[-1, 0[} + \mathbb{1}_{]0, 1]}$. On peut vérifier que

$$\lim_{n \rightarrow \infty} \int_{-1}^1 |f_n(x) - f(x)| dx = 0,$$

et donc que $(f_n)_{n \in \mathbb{N}^*}$ est de Cauchy mais f n'est pas dans E . Ainsi $(f_n)_{n \in \mathbb{N}^*}$ ne converge pas dans E .

Au vu des essais précédents, on voit qu'il faut assez réduire l'espace E pour permettre à la norme d'être définie mais pas trop pour garder E complet.

Définition 5.2 La relation “ $f = g$ μ -presque partout sur X ” est une relation d'équivalence sur \mathcal{L} . Cette relation, notée \mathcal{R} , est compatible avec la structure d'espace vectoriel de \mathcal{L} .

On désigne par L l'espace quotient $L = \mathcal{L}/\mathcal{R}$. Ainsi $f \in L$ désigne n'importe quel représentant de la classe de f , i.e. n'importe quelle fonction mesurable \tilde{f} telle que $f = \tilde{f}$ μ -p.p. sur X .

Exemple 5.1

- Soient $f, g, h : \mathbb{R} \rightarrow \mathbb{R}$ les fonctions définies par

$$f(x) = \mathbb{1}_{]0, +\infty[}(x), \quad g(x) = \mathbb{1}_{]0, +\infty[}(x), \quad h(x) = \mathbb{1}_{]0, +\infty[}(x) - \mathbb{1}_{\{0\}}(x).$$

Alors $f \mathcal{R} g$ et $f \mathcal{R} h$ pour la mesure de Lebesgue.

- $0\mathcal{R}1_{\mathbb{Q}}$ pour la mesure de Lebesgue.

On utilise généralement le mot de régularité pour caractériser les propriétés de continuité et de dérivabilité d'une fonction : une fonction continue par morceaux est moins régulière qu'une fonction continue, qui elle-même est moins régulière qu'une fonction dérivable.

Pour des éléments de L , la valeur en un point $x \in X$ n'a pas forcément de sens. Quand c'est possible, on identifiera toujours $f \in L$ à son représentant le plus régulier, c'est-à-dire à son représentant qui a le plus de propriétés en terme de continuité, dérivabilité, ...

Exemple 5.2 Dans L , on identifiera $1_{\mathbb{Q}}$ à la fonction nulle et $1_{\mathbb{R} \setminus \mathbb{Q}}$ à la fonction constante égale à 1.

On peut maintenant définir les espaces L^p .

Définition 5.3 On note L^p l'espace

$$L^p = \left\{ f \in L : \int_X |f(x)|^p d\mu(x) < +\infty \right\},$$

et l'on note

$$\|f\|_{L^p} = \left(\int_X |f(x)|^p d\mu(x) \right)^{1/p}.$$

Exemple 5.3 Dans ces exemples, la mesure sous-jacente est la mesure de Lebesgue.

- Si $X = [1, +\infty[$, la fonction définie par $f(x) = \frac{1}{x}$ appartient à L^2 mais pas à L^1 .
- Si $X =]0, 1[$, la fonction définie par $f(x) = \frac{1}{\sqrt{x}}$ appartient à L^1 mais pas à L^2 .

5.1.3 Inégalité triangulaire et complétude de L^p

Il nous reste à montrer que $\|\cdot\|_{L^p}$ est une norme pour l'espace L^p (donc l'inégalité triangulaire) et que L^p est complet. Commençons par rappeler une inégalité classique du cours d'optimisation.

Définition 5.4 Un ensemble X est convexe si et seulement si $\forall x, y \in X$, $\forall \theta \in [0, 1]$,

$$(1 - \theta)x + \theta y \in X.$$

dans ce cas, une fonction $\varphi : X \rightarrow \mathbb{R}$ est dite convexe si et seulement si $\forall x, y \in X$, $\forall \theta \in [0, 1]$,

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

Exemple 5.4

- Soit φ dérivable sur $]a, b[$, alors φ est convexe si et seulement si φ' est croissante.
- Soit φ , deux fois dérivable sur $]a, b[$, alors φ est convexe si et seulement si $\varphi''(x) \geq 0$ pour tout $x \in]a, b[$.

L'inégalité suivante, dite inégalité de Young, est un cas particulier des inégalités de Jensen.

Lemme 5.1 Si $a \geq 0$, $b \geq 0$ et si $p, q > 0$ sont tels que $\frac{1}{p} + \frac{1}{q} = 1$ (on dit que p et q sont conjugués), alors $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$.

Preuve : Le cas $ab = 0$ est trivial et sinon, on pose $s = p \ln a$, $t = q \ln b$, et la convexité de la fonction exponentielle entraîne que $e^{\frac{s}{p} + \frac{t}{q}} \leq \frac{e^s}{p} + \frac{e^t}{q}$. \square

Avant de démontrer l'inégalité triangulaire, on démontre une inégalité qui a son importance propre.

Théorème 5.1 (Inégalité de Hölder) Soient $1 < p, q < +\infty$ que l'on suppose conjugués. Soient $f \in L^p$ et $g \in L^q$. Alors on a l'inégalité

$$\int_X |f(x)g(x)| d\mu(x) \leq \left(\int_X |f(x)|^p d\mu(x) \right)^{\frac{1}{p}} \left(\int_X |g(x)|^q d\mu(x) \right)^{\frac{1}{q}},$$

qui s'écrit aussi

$$\|fg\|_{L^1} \leq \|f\|_{L^p} \|g\|_{L^q}.$$

Preuve : Notons $\alpha = \|f\|_{L^p}$ et $\beta = \|g\|_{L^q}$. Si $\alpha = 0$ (resp. $\beta = 0$) alors $f = 0$ μ -p.p (resp. $g = 0$ μ -p.p) et ainsi $fg = 0$ μ -p.p donc $\int_X |fg| d\mu = 0$.
 À présent, considérons le cas $\alpha > 0$ et $\beta > 0$. Notons $F = \frac{|f|}{\alpha}$ et $G = \frac{|g|}{\beta}$.

Ces quantités vérifient :

$$\int_X F^p d\mu = \int_X G^q d\mu = 1.$$

On a $0 \leq F(x)G(x) < +\infty$ μ -p.p. et par l'inégalité de Young,

$$F(x)G(x) \leq \frac{F(x)^p}{p} + \frac{G(x)^q}{q} \quad \mu - p.p.$$

En intégrant, on obtient finalement

$$\int_X F(x)G(x)d\mu \leq 1,$$

c'est-à-dire la conclusion désirée. \square

Si $p = 2$, $q = 2$ l'inégalité de Hölder est alors la célèbre inégalité de Cauchy-Schwarz car

$$\langle f, g \rangle = \int_X f(x)g(x) d\mu(x),$$

définit un produit scalaire sur L^2 .

On a égalité dans l'inégalité de Hölder si f et g sont proportionnelles μ -p.p.

Nous allons utiliser à présent l'inégalité de Hölder pour montrer l'inégalité triangulaire, ce qui démontre que $\|\cdot\|_{L^p}$ est une norme et que L^p est stable par addition : L^p est alors un espace vectoriel.

Théorème 5.2 (Inégalité de Minkowski) *Soit $1 \leq p < +\infty$ et $f, g \in L^p$. Alors on a l'inégalité*

$$\left(\int_X |f(x) + g(x)|^p d\mu(x) \right)^{\frac{1}{p}} \leq \left(\int_X |f(x)|^p d\mu(x) \right)^{\frac{1}{p}} + \left(\int_X |g(x)|^p d\mu(x) \right)^{\frac{1}{p}},$$

qui s'écrit aussi

$$\|f + g\|_{L^p} \leq \|f\|_{L^p} + \|g\|_{L^p}.$$

Preuve : Pour $p = 1$, l'inégalité est évidente. Pour $p > 1$, on utilise l'inégalité de Hölder en écrivant $(f + g)^p = f(f + g)^{p-1} + g(f + g)^{p-1}$ et on pose $q = \frac{p}{p-1}$. On a bien $p^{-1} + q^{-1} = 1$, $q(p-1) = p$ et donc

$$\begin{aligned} \int_X f (f + g)^{p-1} d\mu &\leq \left(\int_X f^p d\mu \right)^{\frac{1}{p}} \left(\int_X (f + g)^p d\mu \right)^{\frac{1}{q}}, \\ \int_X g (f + g)^{p-1} d\mu &\leq \left(\int_X g^p d\mu \right)^{\frac{1}{p}} \left(\int_X (f + g)^p d\mu \right)^{\frac{1}{q}}. \end{aligned}$$

En additionnant membre à membre les inégalités précédentes, on obtient

$$\int_X |f + g|^p d\mu \leq \left(\left(\int_X |f|^p d\mu \right)^{\frac{1}{p}} + \left(\int_X |g|^p d\mu \right)^{\frac{1}{p}} \right) \left(\int_X |f + g|^p d\mu \right)^{\frac{1}{q}}.$$

Et comme $1 - \frac{1}{q} = \frac{1}{p}$, on obtient l'inégalité désirée en divisant par

$$\left(\int_X |f + g|^p d\mu \right)^{\frac{1}{q}}.$$

□

Ainsi, le corollaire suivant est immédiat par ce qui précède.

Corollaire 5.1 *L'application $\|\cdot\|_{L^p}$ est une norme sur L^p .*

On peut donc à présent définir la convergence dans l'espace L^p .

Définition 5.5 *Une suite $(f_n)_{n \in \mathbb{N}}$ de fonctions de L^p converge dans l'espace L^p vers une fonction $f \in L^p$ si*

$$\lim_{n \rightarrow +\infty} \|f_n - f\|_{L^p} = 0.$$

Remarquons que si l'on revient sur le théorème de convergence dominée, la conclusion est que $\lim_{n \rightarrow +\infty} \int_X f_n d\mu = \int_X f d\mu$, qui se réécrit comme

$$\lim_{n \rightarrow +\infty} \left| \int_X (f_n - f) d\mu \right| = 0,$$

une quantité plus petite que la limite de la norme L^1 de $f_n - f$ (pour le voir, il suffit simplement de majorer en mettant la valeur absolue à l'intérieur de l'intégrale). Cependant, en utilisant le théorème de convergence dominée non pas pour la suite $(f_n)_{n \in \mathbb{N}}$ mais pour $(|f_n - f|)_{n \in \mathbb{N}}$ (pourquoi est-ce possible ?), on obtient alors que $\lim_{n \rightarrow +\infty} \|f_n - f\|_{L^1} = 0$, qui est en fait la vraie conclusion du théorème de convergence dominée.

Revenons au problème qui nous intéresse, c'est-à-dire de montrer la complétude de l'espace L^p lorsque $1 \leq p < +\infty$.

Théorème 5.3 (Riesz-Fischer) *Soit $(f_n)_{n \in \mathbb{N}}$ une suite de Cauchy dans L^p où $1 \leq p < \infty$. Alors il existe $f \in L^p$ telle que*

$$\lim_{n \rightarrow +\infty} \|f - f_n\|_{L^p} = 0.$$

Preuve : Ce théorème se démontre en deux étapes :

- On note d'abord que si $(u_n)_{n \in \mathbb{N}}$ est une suite de fonctions de L^p telle que $\sum_{n=0}^{+\infty} \|u_n\|_{L^p} < +\infty$, alors la série numérique de terme général $u_n(x)$ converge absolument μ -p.p. sur X vers une fonction $f \in L^p$, définie presque partout par

$$f(x) = \sum_{n=0}^{+\infty} u_n(x).$$

En effet, si $s_n(x) = \sum_{k=0}^n |u_k(x)|$, la suite de fonction $(s_n^p)_{n \in \mathbb{N}}$ est une suite croissante de fonctions positives mesurables, et le théorème de convergence monotone entraîne

$$\lim_{n \rightarrow +\infty} \int_X s_n(x)^p d\mu = \int_X \lim_{n \rightarrow +\infty} s_n(x)^p d\mu,$$

que l'on peut écrire

$$\lim_{n \rightarrow +\infty} \|s_n\|_{L^p}^p = \int_X \left(\sum_{k=0}^{+\infty} |u_k(x)| \right)^p dx.$$

Mais avec l'inégalité de Minkowski, et pour tout $n \in \mathbb{N}$, on a

$$\|s_n\|_{L^p} \leq \sum_{k=0}^n \|u_k\|_{L^p} < \sum_{k=0}^{+\infty} \|u_k\|_{L^p},$$

de sorte que la limite ci-dessus est finie. On en déduit que la fonction $(\sum_{k=0}^{+\infty} |u_k|)^p$ est finie μ -p.p. sur X : la série converge absolument μ -p.p. donc elle converge μ -p.p. Soit f sa somme. En utilisant l'inégalité de Minkowski généralisée aux sommes infinies, on obtient la convergence vers f dans L^p :

$$\|f - \sum_{k=0}^n u_k\|_{L^p} = \left\| \sum_{k=n+1}^{+\infty} u_k \right\|_{L^p} \leq \sum_{k=n+1}^{+\infty} \|u_k\|_{L^p}.$$

- Soit $(f_n)_{n \in \mathbb{N}}$ une suite de Cauchy dans L^p . Il existe une fonction φ , définie sur \mathbb{N} , croissante, telle que si $l, m \geq \varphi(n)$ alors $\|f_l - f_m\|_{L^p} < 2^{-n}$. Soit $(g_n)_{n \in \mathbb{N}}$ la sous-suite de $(f_n)_{n \in \mathbb{N}}$ définie par $g_n = f_{\varphi(n)}$. On applique le premier point à la série des $u_n = g_{n+1} - g_n$ pour montrer la convergence de g_n vers f dans L^p .

Comme la suite $(f_n)_{n \in \mathbb{N}}$ est de Cauchy, la convergence d'une sous-suite suffit à entraîner celle de la suite complète :

$$\lim_{n \rightarrow +\infty} \|f_n - f\|_{L^p} = 0,$$

ce qui termine la démonstration. \square

Au passage, on a montré que pour $1 \leq p < +\infty$, la convergence de la suite (f_n) vers f dans L^p entraîne l'existence d'une sous-suite (f_{n_k}) qui converge vers f presque partout.

SECTION 5.2

L'espace L^∞

Moralement, cet espace est celui des fonctions bornées. Cependant, la définition est un peu complexe car s'agissant d'un espace de fonctions définies seulement presque partout, il n'est pas évident de définir les bornes.

5.2.1 Définition

Définition 5.6 Soit $f \in L$. Le réel positif M est un "majorant essentiel" de f si et seulement si

$$\mu(\{x \in X; |f(x)| > M\}) = 0.$$

L'ensemble des majorants essentiels n'est pas vide ($+\infty$ est un majorant essentiel) et il est minoré (aucun nombre négatif ne peut être majorant essentiel).

On appelle "borne supérieure essentielle" de f sur X , et l'on note $\|f\|_{L^\infty}$ la borne inférieure de l'ensemble des majorants essentiels, éventuellement égale à $+\infty$.

Définition 5.7 L'espace L^∞ est l'ensemble des classes d'équivalence modulo \mathcal{R} de fonctions mesurables f telles que $\|f\|_{L^\infty} < +\infty$. L'application $\|\cdot\|_{L^\infty}$ définit une norme sur L^∞ et L^∞ est un espace vectoriel normé.

Le fait que $\|\cdot\|_{L^\infty}$ soit une norme est trivial et de l'homogénéité de la norme suit la stabilité de L^∞ pour la multiplication scalaire. La stabilité de L^∞ par rapport à l'addition est une conséquence de l'inégalité triangulaire.

On peut donc à présent parler de convergence dans l'espace L^∞ .

5.2.2 Complétude

Pour terminer cette section, nous nous proposons maintenant de montrer que l'espace L^∞ muni de sa norme est un espace complet.

Proposition 5.1 L'espace L^∞ est complet.

Preuve : Si $(f_n)_{n \in \mathbb{N}}$ est une suite de Cauchy dans L^∞ , alors

$$\forall \epsilon > 0, \exists k(\epsilon) > 0 \text{ tel que } n, m > k(\epsilon) \implies \|f_n - f_m\|_{L^\infty} < \epsilon.$$

Les ensembles suivants sont évidemment négligeables :

$$\begin{aligned} E_n &= \{x \in I; |f_n(x)| > \|f_n\|_{L^\infty}\}, \\ E_{n,m} &= \{x \in I; |f_n(x) - f_m(x)| > \|f_n - f_m\|_{L^\infty}\}. \end{aligned}$$

Soit \mathcal{N} leur réunion : c'est une réunion dénombrable de parties négligeables, donc \mathcal{N} est négligeable. Pour tout $x \in \mathcal{N}^c$, la suite $(f_n(x))_{n \in \mathbb{N}}$ est de Cauchy dans \mathbb{R} , donc convergente. Définissons alors la fonction f par :

$$f(x) = \begin{cases} \lim_{n \rightarrow +\infty} f_n(x) & \text{si } x \in \mathcal{N}^c, \\ 0 & \text{sinon.} \end{cases}$$

La fonction f est mesurable et vérifie $\|f - f_m\|_{L^\infty} = \lim_{n \rightarrow +\infty} \|f_n - f_m\|_{L^\infty}$.
Comme la suite (f_n) est de Cauchy,

$$\lim_{m \rightarrow +\infty} \|f - f_m\|_{L^\infty} = \lim_{n, m \rightarrow +\infty} \|f_n - f_m\|_{L^\infty} = 0.$$

Pour $m > k(\epsilon)$ on a en outre que $\|f - f_m\|_{L^\infty} < \epsilon$ et donc que

$$\|f\|_{L^\infty} \leq \|f_m\|_{L^\infty} + \epsilon < \infty.$$

Ainsi, on a finalement montré que $f \in L^\infty$. □

SECTION 5.3

Espaces L^p_{loc}

Remarquons que sur \mathbb{R} des fonctions aussi simples que les polynômes ne font pas partie de l'espace L^1 . Cependant, comme toutes les autres fonctions continues, elles sont intégrables sur tout compact de \mathbb{R} . D'où l'intérêt d'introduire les espaces suivants.

Définition 5.8 *L'espace L^p_{loc} , avec $1 \leq p \leq +\infty$, est l'espace des fonctions localement de puissance $p^{\text{ème}}$ intégrables (localement intégrables si $p = 1$ et localement essentiellement bornées si $p = \infty$) sur X .*

- On a que $f \in L^p_{loc}$ si et seulement si pour tout compact $K \subset X$, $\|f \mathbf{1}_K\|_{L^p} < +\infty$.
- On a $f_n \xrightarrow[n \rightarrow +\infty]{} f$ dans L^p_{loc} si et seulement si pour tout compact $K \subset X$, $\lim_{n \rightarrow \infty} \|(f_n - f) \mathbf{1}_K\|_{L^p} = 0$.

Relations entre les espaces fonctionnels classiques

Ce chapitre est destiné à faire un bilan provisoire des différents espaces fonctionnels avec lesquels nous travaillons et les relations entre eux. Dans toute la suite du cours, bien que tout ce qui suit peut s'étendre à d'autres espaces mesurés, on ne s'intéresse qu'aux fonctions à valeurs réelles définies sur $X = \mathbb{R}$ ou sur un sous-ensemble borélien $X \subset \mathbb{R}$, naturellement muni de sa tribu borélienne et de la mesure de Lebesgue.

SECTION 6.1

Espaces de fonctions régulières

Nous commençons par introduire ou rappeler certaines définitions.

6.1.1 Espaces \mathcal{C}^k

Définition 6.1 Soit $k \in \mathbb{N}$. On note \mathcal{C}^k (respectivement \mathcal{C} , respectivement \mathcal{C}^∞) l'espace vectoriel des fonctions continues k fois continûment dérivables (resp. continues si $k = 0$, resp. infiniment continûment dérivables) sur X .

On munit l'espace \mathcal{C} de la norme uniforme définie par

$$\|f\|_\infty = \sup_{x \in X} |f(x)|.$$

Sur l'espace \mathcal{C} , cette norme coïncide avec la norme de l'espace L^∞ , auquel cas on utilisera dans la suite la notation $\|f\|_{L^\infty}$ plutôt que $\|f\|_\infty$.

Proposition 6.1 Si X est compact, alors l'espace \mathcal{C} est complet pour la norme uniforme.

Preuve : Soit $(f_n)_{n \in \mathbb{N}}$ une suite de Cauchy de l'espace \mathcal{C} muni de la norme $\|\cdot\|_{L^\infty}$. Comme cet espace est inclus dans L^∞ dont on sait qu'il est complet, alors la suite de Cauchy $(f_n)_{n \in \mathbb{N}}$ converge pour la norme $\|\cdot\|_{L^\infty}$ vers un élément $f \in L^\infty$, qui en fait est la limite uniforme d'une suite de fonctions

continues sur X . Elle est donc elle-même continue sur X , d'où le fait que $f \in \mathcal{C}$. \square

La compacité de l'ensemble X ne sert pas dans la démonstration. Cependant, il nous faut la supposer sinon la norme $\|\cdot\|_{L^\infty}$ peut ne pas être définie.

Exemple 6.1

- Sur $X =]0, 1[$, les fonctions $f(x) = \frac{1}{x}$ et $g(x) = \frac{1}{x^2}$ sont telles que

$$\|f\|_{L^\infty} = \|g\|_{L^\infty} = \|f - g\|_{L^\infty} = +\infty.$$

- Si $X = \mathbb{R}$, il suffit de prendre $f(x) = x$ et $g(x) = x^2$

On utilise alors la notion de convergence uniforme sur les compacts, en tenant compte des dérivées lorsque $k > 0$:

Définition 6.2 On dit que la suite $(f_n)_{n \in \mathbb{N}}$ converge vers f dans \mathcal{C}^k si les dérivées $f_n^{(l)}$ converge vers $f^{(l)}$ uniformément sur tout compact de X , pour tout $0 \leq l \leq k$.

6.1.2 Espaces \mathcal{C}_c^k

Définition 6.3 Le support d'une fonction continue f est la fermeture de l'ensemble des points où elle n'est pas nulle :

$$\text{supp}(f) = \overline{\{x \in X : f(x) \neq 0\}}.$$

Notons que cette définition n'est pas adaptée aux fonctions définies presque partout : avec cette définition, le support de la fonction $\mathbf{1}_{\mathbb{Q}}$ serait \mathbb{R} , alors que cette fonction est nulle p.p. C'est pourquoi pour une fonction générale $f \in L(X)$, son support est le complémentaire du plus grand ouvert sur lequel f est nulle p.p. (pour la fonction $\mathbf{1}_{\mathbb{Q}}$, son support est alors l'ensemble vide).

Définition 6.4 Soit $k \in \mathbb{N}$. On note \mathcal{C}_c^k (respectivement \mathcal{C}_c , respectivement \mathcal{C}_c^∞) l'espace vectoriel des fonctions continues k fois continûment dérivables (resp. continues si $k = 0$, resp. infiniment continûment dérivables) sur X et à support compact inclus dans X .

Exemple 6.2 Soit g la fonction définie sur \mathbb{R} par

$$g(x) = \exp\left(\frac{-1}{1-x^2}\right) \mathbb{1}_{]-1,1[}(x).$$

Bien qu'il y ait manifestement un problème aux points ± 1 , cette fonction est prolongeable par continuité en ces points, tout comme toutes ses dérivées. En effet, pour tout $k \geq 1$ et pour tout $|x| < 1$, on a

$$g^{(k)}(x) = \frac{P(x)}{(1-x^2)^{2k}} \exp\left(\frac{-1}{1-x^2}\right),$$

où P est un polynôme, de sorte que $\lim_{|x| \rightarrow 1^-} g^{(k)}(x) = 0$. On en déduit que $g \in \mathcal{C}_c^\infty$.

Exemple 6.3 Considérons la suite $(f_n)_{n \in \mathbb{N}}$ de fonctions définies sur \mathbb{R} par

$$f_n(x) = \begin{cases} 0 & \text{si } x \notin [-n, n], \\ (x+n)e^{-x^2} & \text{si } -n \leq x < -n+1, \\ e^{-x^2} & \text{si } -n+1 \leq x \leq n-1, \\ (n-x)e^{-x^2} & \text{si } n-1 < x \leq n. \end{cases}$$

Pour tout $n \in \mathbb{N}$, on a $f_n \in \mathcal{C}_c$. La suite $(f_n)_{n \in \mathbb{N}}$ converge uniformément sur \mathbb{R} vers la fonction $f(x) = e^{-x^2}$, mais cette fonction n'est pas à support compact.

Ainsi, l'espace \mathcal{C}_c n'est pas un sous-espace fermé de l'espace \mathcal{C} muni de la topologie de la convergence uniforme. Or, comme tout sous-espace complet d'un espace vectoriel normé (non nécessairement complet) est fermé (admis), on en déduit que l'espace \mathcal{C}_c n'est pas complet pour la norme uniforme.

À défaut de définir une topologie pour les \mathcal{C}_c^k , on va préciser une définition de la convergence des suites qui en fait un espace complet.

Définition 6.5 Une suite $(f_n)_{n \in \mathbb{N}}$ converge vers une fonction f dans \mathcal{C}_c^k si et seulement si :

- il existe un compact $K \subset X$ tel que $\text{supp}(f_n) \subset K$ pour tout $n \in \mathbb{N}$,
- $f_n^{(l)} \xrightarrow[n \rightarrow +\infty]{} f^{(l)}$ uniformément sur X pour tout $0 \leq l \leq k$.

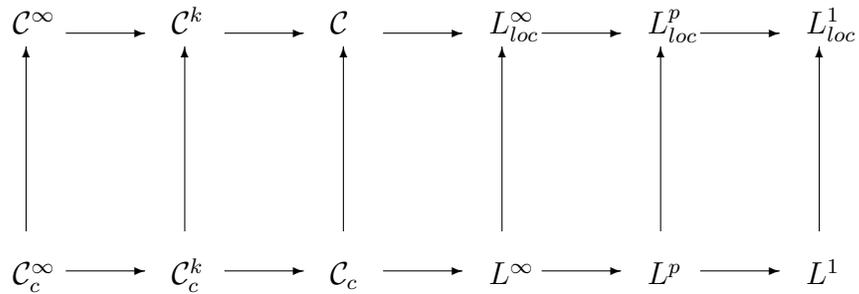
Ainsi, il aurait fallu dans l'exemple précédent que le support des fonctions f_n , en l'occurrence $[-n, n]$, soit inclus dans un compact ne dépendant pas de n pour que la fonction limite f soit elle-aussi à support compact.

SECTION 6.2

Relations entre les modes de convergence

Les inclusions entre les espaces fonctionnels sont résumées dans le tableau suivant. Une flèche y relie un espace E à un espace F si et seulement si E s'injecte continûment dans F , c'est-à-dire que :

- $E \subset F$,
- $f_n \rightarrow f$ dans E entraîne $f_n \rightarrow f$ dans F .



L'espace \mathcal{C}_c^∞ s'injecte continûment dans chacun des espaces de ce tableau, et chaque espace s'injecte continûment dans L_{loc}^1 .

Attention : il y a quelques restrictions à faire pour obtenir ce diagramme. Toutes les inclusions sont vraies sans hypothèse additionnelle sauf pour les inclusions concernant les espaces L^p . En effet, il faut supposer l'espace X borné entraînant que sa mesure de Lebesgue est finie pour que l'on ait $L^\infty \subset L^p \subset L^1$ (voir le contre-exemple 5.3 du chapitre précédent).

En revanche, les convergences p.p. et dans les espaces L^p ne sont pas directement comparables, comme on le verra en TD. Néanmoins, nous avons vu précédemment les propriétés suivantes : si $1 \leq p < +\infty$,

- la convergence de la suite $(f_n)_{n \in \mathbb{N}}$ vers f dans L^p entraîne la convergence p.p. vers cette même limite f mais seulement d'une sous-suite $(f_{n_k})_{k \in \mathbb{N}}$.
- si la suite $(f_n)_{n \in \mathbb{N}}$ converge p.p. vers f et est telle que $|f_n|^p$ est dominée par une fonction intégrable et indépendante de n , alors $f \in L^p$ et la convergence vers f a aussi lieu dans L^p (une application du théorème de convergence dominée).

Enfin, ces deux modes de convergence entraînent une convergence commune, la convergence dite en mesure, qui n'est autre que la version "théorie de la mesure" de la convergence en probabilité, comme on le verra au dernier chapitre. On admet cette propriété.

Définition 6.6 *Une suite de fonctions $(f_n)_{n \in \mathbb{N}}$ de L converge en mesure vers une fonction $f \in L$ si*

$$\forall \varepsilon > 0, \lim_{n \rightarrow +\infty} \mu(\{x \in X : |f_n(x) - f(x)| > \varepsilon\}) = 0.$$

Convolution des fonctions

SECTION 7.1

Définition

7.1.1 Un exemple de la physique

Exemple 7.1 *Le circuit RC représenté à la figure 7.1 est constitué d'une résistance et d'un condensateur. Son entrée est la tension $x(t)$ fournie par le générateur ; sa sortie la tension $v(t)$ aux bornes du condensateur.*

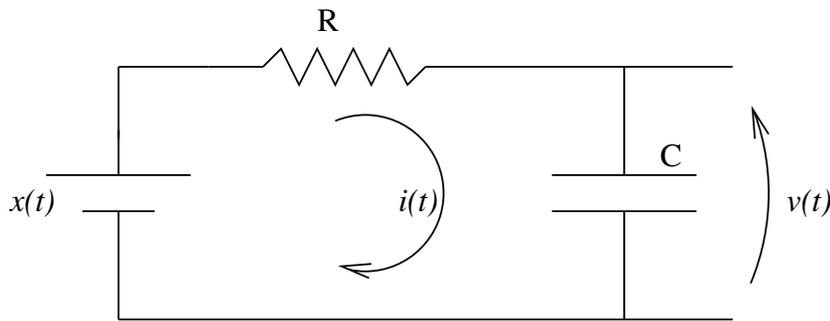


Figure 7.1: Circuit RC

Si la charge est $q(t)$, on sait que $v(t) = \frac{q(t)}{C}$, et la loi d'Ohm fournit :

$$Ri(t) + v(t) = x(t),$$

et en tenant compte de ce que $i(t) = q'(t)$,

$$RC v'(t) + v(t) = x(t).$$

C'est une équation différentielle linéaire du premier ordre à coefficients constants. Elle se résout classiquement en cherchant une solution générale de l'équation homogène, puis en recherchant une solution particulière de l'équation avec second membre par la méthode de variation de la constante. On pose $v(t) = k(t) e^{-\frac{t}{RC}}$.

On obtient alors $k'(t) = \frac{1}{RC} e^{+\frac{t}{RC}} x(t)$, puis, sous des hypothèses raisonnables sur $x(t)$:

$$k(t) = \frac{1}{RC} \int_{-\infty}^t e^{+\frac{s}{RC}} x(s) ds + k,$$

qui fournit :

$$v(t) = \frac{1}{RC} \int_{-\infty}^t e^{-\frac{t-s}{RC}} x(s) ds + k e^{-\frac{t}{RC}}, \quad (7.1)$$

la constante k étant alors déterminée par une condition supplémentaire, qui peut être par exemple que si la tension d'entrée $x(t)$ est nulle, alors $v(t) = 0$.

La sortie $v(t)$ apparaît ci-dessus comme une fonction de l'entrée $x(t)$, à travers un opérateur de type intégral. La variable de temps y apparaît à la fois comme borne de l'intégrale et sous le symbole \int . En introduisant la fonction $\mathbb{1}_{]0, \infty[}$, appelée fonction de Heaviside (ou échelon unité en théorie du signal), on observe que

$$\int_{-\infty}^t e^{-\frac{t-s}{RC}} x(s) ds = \int_{\mathbb{R}} e^{-\frac{t-s}{RC}} \mathbb{1}_{]0, \infty[}(t-s) x(s) ds,$$

de sorte que si l'on note $h(t) = \frac{1}{RC} e^{-\frac{t}{RC}} \mathbb{1}_{]0, \infty[}(t)$, on pourra écrire :

$$v(t) = \int_{\mathbb{R}} h(t-s) x(s) ds = (h \star x)(t). \quad (7.2)$$

On dit que v est égal au produit de convolution des fonctions h et x .

Les formules (7.1) et (7.2) sont formelles : on ne sait pas si les intégrales sont bien définies. On verra des conditions garantissant l'existence du produit de convolution. En particulier, les fonctions h et x pourront être de la forme $x = x \mathbb{1}_{\mathbb{R}^+}$ et $h = h \mathbb{1}_{\mathbb{R}^+}$ pour décrire l'évolution du système après sa mise sous tension.

7.1.2 Définition

Définition 7.1 Soient $f, g \in L$. Alors la fonction définie par

$$(x, y) \mapsto f(x-y) g(y),$$

est mesurable de \mathbb{R}^2 dans \mathbb{R} . Les fonctions f et g sont dites convolables si la fonction $y \mapsto f(x-y)g(y)$ est intégrable sur \mathbb{R} pour presque tout $x \in \mathbb{R}$. On définit dans ce cas le produit de convolution :

$$(f \star g)(x) = \int_{\mathbb{R}} f(x-y)g(y)dy.$$

Proposition 7.1 *Le produit de convolution est bilinéaire et symétrique au sens suivant : s'ils sont bien définis, on a*

- $f \star g = g \star f$.
- $(\alpha f + \beta g) \star h = \alpha f \star h + \beta g \star h$ pour tous $\alpha, \beta \in \mathbb{R}$.
- $\text{supp}(f \star g) \subset \overline{\text{supp}(f) + \text{supp}(g)}$.

Preuve : Évidente pour les deux premières propriétés. Pour celle associée aux supports, soit $x \notin \overline{\text{supp}(f) + \text{supp}(g)}$. Montrons que $f \star g(x) = 0$. On a

$$\begin{aligned} f \star g(x) &= \int_{\mathbb{R}} f(x-y)g(y) \mathbb{1}_{\text{supp}(f)}(x-y) \mathbb{1}_{\text{supp}(g)}(y) dy \\ &= \int_{\mathbb{R}} f(x-y)g(y) \mathbb{1}_{(\text{supp}(g)) \cap (x-\text{supp}(f))}(y) dy \\ &= 0, \end{aligned}$$

car $(\text{supp}(g)) \cap (x - \text{supp}(f)) = \emptyset$. En effet, si cet ensemble n'était pas vide, il contiendrait un t dans $\text{supp}(g)$ qui s'écrirait $t = x - u$ avec $u \in \text{supp}(f)$. Mais alors $x = t + u$ serait dans la somme des supports, ce qui est contraire à l'hypothèse de départ. \square

SECTION 7.2

Convolution et espaces fonctionnels

Pour quels espaces fonctionnels le produit de convolution est-il bien défini, et dans quel espace appartient-il ? Ce sont deux questions auxquelles on se propose de répondre. Lorsque l'on pourra y répondre de façon positive, on notera $E \star F \subset G$.

7.2.1 Espaces L^1 et L^p

Théorème 7.1 *Pour tout $1 \leq p < \infty$, on a*

$$L^1 \star L^p \subset L^p.$$

De surcroît on a l'inégalité suivante : si $f \in L^1$ et $g \in L^p$, alors

$$\|f \star g\|_{L^p} \leq \|f\|_{L^1} \|g\|_{L^p}.$$

Preuve : La fonction $y \mapsto |f(x-y)|^{\frac{1}{p}} |g(y)|$ appartient à L^p pour presque tout x de \mathbb{R} , d'après le théorème de Fubini, car

$$\int_{\mathbb{R}} |f(x-y)| |g(y)|^p dx \leq |g(y)|^p \|f\|_{L^1},$$

et par suite, $\int_{\mathbb{R}^2} |f(x-y)| |g(y)|^p dx dy \leq \|g\|_{L^p}^p \|f\|_{L^1}$.

On applique l'inégalité de Hölder à $|f(x-y)|^{\frac{1}{p}} |g(y)|$, élément de L^p et $|f(x-y)|^{\frac{1}{q}}$, qui appartient à L^q :

$$(|f| \star |g|)(x) \leq \left(\int_{\mathbb{R}} |f(x-y)| |g(y)|^p dy \right)^{\frac{1}{p}} \|f\|_{L^1}^{\frac{1}{q}},$$

d'où

$$\int_{\mathbb{R}} (|f| \star |g|)^p(x) dx \leq \|f\|_{L^1}^{\frac{1}{p}} \|g\|_{L^p}^p \|f\|_{L^1}^{\frac{1}{q}} = \|f\|_{L^1}^p \|g\|_{L^p}^p.$$

□

Théorème 7.2 *L'espace L^1 est une algèbre de convolution : pour tous $f, g \in L^1$ alors $f \star g \in L^1$ et pour tout $h \in L^1$,*

$$f \star (g \star h) = (f \star g) \star h.$$

Preuve : Démontrons l'associativité du produit de convolution. Si f, g et h sont des fonctions de L^1 , il en est de même pour $f \star g$ et $g \star h$. Pour presque tout $x \in \mathbb{R}$, la fonction

$$y \mapsto \int_{\mathbb{R}} |f(x-y-z) g(z) h(y)| dz = \int_{\mathbb{R}} |f(x-u) g(u-y) h(y)| du,$$

est intégrable et l'on a, d'après le théorème de Fubini,

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x-y-z) g(z) dz \right) h(y) dy = \int_{\mathbb{R}} f(x-u) \left(\int_{\mathbb{R}} g(u-y) h(y) dy \right) du,$$

c'est-à-dire

$$((f \star g) \star h)(x) = (f \star (g \star h))(x).$$

□

7.2.2 Espaces L^1 et L^∞

Lorsque l'espace L^p est remplacé par L^∞ , on obtient le même type de résultat, en gagnant de la continuité, ce qui est remarquable car les fonctions initialement prises dans L^1 et L^p ne sont pas forcément continues !!

Théorème 7.3 *On a*

$$L^1 \star L^\infty \subset \mathcal{C} \cap L^\infty,$$

et pour tous $f \in L^1$ et $g \in L^\infty$,

$$\|f \star g\|_{L^\infty} \leq \|f\|_{L^1} \|g\|_{L^\infty}. \quad (7.3)$$

Preuve : L'inégalité (7.3) est triviale, ce qui nous donne que $L^1 \star L^\infty \subset L^\infty$. Pour montrer que le produit de convolution $f \star g$ est continue sur \mathbb{R} , c'est plus délicat et l'on admettra cette propriété (bien que la fonction $y \mapsto f(y)g(x-y)$ soit dominée en valeur absolue par $|f| \|g\|_{L^\infty}$ qui est bien intégrable, rien ne nous dit qu'elle est continue sur \mathbb{R} : on ne peut donc pas utiliser directement le théorème de continuité sous l'intégrale). □

7.2.3 Espaces L^p et L^q avec p et q conjugués

Lorsque p et q sont conjugués, l'inégalité de Hölder entraîne immédiatement le résultat suivant.

Théorème 7.4 *Lorsque p et q sont conjugués, on a*

$$L^p \star L^q \subset L^\infty,$$

et pour tous $f \in L^p$ et $g \in L^q$,

$$\|f \star g\|_{L^\infty} \leq \|f\|_{L^p} \|g\|_{L^q}. \quad (7.4)$$

7.2.4 L'espace L^1_{loc}

Si f et g sont localement intégrables, leur produit de convolution n'est généralement pas défini. Cependant, le théorème suivant décrit une situation dont les applications sont importantes.

On notera u la fonction échelon unité de Heaviside,

$$u(x) = \mathbb{1}_{]0, +\infty[}(x).$$

Théorème 7.5 *Soient f et g deux fonctions de L^1_{loc} . Les fonctions uf et ug sont convolables et*

$$(uf) * (ug)(x) = u(x) \int_0^x f(x-y)g(y) dy.$$

Preuve : La fonction $(x, y) \mapsto |u(x-y)f(x-y)| |u(y)g(y)|$ est positive et mesurable sur \mathbb{R}^2 . D'autre part, on vérifie que

$$u(x-y)u(y) = u(x) \mathbb{1}_{]0, x[}(y) = u(x) \mathbb{1}_{]0, x[}(x-y).$$

Pour presque tout $x \in \mathbb{R}$, on a donc

$$\begin{aligned} |uf| \star |ug|(x) &= u(x) \int_0^x |f(x-y)g(y)| dy, \\ &= u(x) \int_0^{+\infty} |f(x-y) \mathbb{1}_{]0, x[}(x-y)| | \mathbb{1}_{]0, x[}(y)g(y)| dy. \end{aligned}$$

Comme f et g sont dans L^1_{loc} , les fonctions $\mathbb{1}_{]0, x[}f$ et $\mathbb{1}_{]0, x[}g$ sont dans L^1 : elles sont convolables d'après le théorème 7.1. \square

7.2.5 Espaces \mathcal{C}^k_c et L^1_{loc}

Intéressons-nous à présent à la régularité du produit de convolution.

Théorème 7.6 *Pour tout $1 \leq k \leq +\infty$, on a*

$$\mathcal{C}^k_c \star L^1_{loc} \subset \mathcal{C}^k,$$

et pour tous $f \in \mathcal{C}^k_c$ et $g \in L^1_{loc}$,

$$(f \star g)^{(l)} = f^{(l)} \star g \text{ pour } 1 \leq l \leq k.$$

Preuve : On le démontre pour $k \geq 1$, et d'abord pour $l = 1$, en utilisant le théorème de dérivabilité d'une fonction définie par une intégrale :

- on a $f \in \mathcal{C}_c^k$, donc il existe $M > 0$ tel que $f(y) = 0$ dès que $|y| > M$. Soit $K = [-a, a]$ un intervalle compact de \mathbb{R} , et $b = a + M$. On note $g_b = g \mathbb{1}_{[-b, b]}$. Alors, pour tout $x \in K$, on a

$$\int_{\mathbb{R}} |g(x-y) f(y)| dy = \int_{-M}^M |f(y) g(x-y)| dy \leq \|g_b\|_{L^1} \|f\|_{L^\infty} < +\infty.$$

Ainsi, les fonctions f et g sont donc convolables.

- la fonction $x \mapsto f(x-y) g(y)$ est dérivable pour presque tout $y \in \mathbb{R}$, de dérivée $x \mapsto f'(x-y) g(y)$.
- pour tout $x \in K$ et presque tout $y \in \mathbb{R}$,

$$|f'(x-y) g(y)| \leq \|f'\|_{L^\infty} |g_b(y)| =: g_K(y),$$

avec $g_K \in L^1$.

Ainsi, les hypothèses du théorème sont vérifiées et

$$(f \star g)'(x) = \int_{\mathbb{R}} f'(x-y) g(y) dy.$$

Ceci est vrai pour tout x réel. Enfin, un raisonnement par récurrence sur $1 \leq l < k$ nous permet d'obtenir le cas général. \square

Bien évidemment, ce résultat est vrai lorsqu'on remplace l'espace L_{loc}^1 par L^1 , vu que $L^1 \subset L_{loc}^1$.

Enfin, notons que le produit de convolution de deux fonctions à support compact est également à support compact. En effet, ceci est immédiat car si f et g sont de telles fonctions, on a vu que $\text{supp}(f \star g) \subset \overline{\text{supp}(f) + \text{supp}(g)}$.

SECTION 7.3

Régularisation par convolution

L'un des intérêts principaux du produit de convolution est qu'il permet d'approcher des fonctions de l'espace L^p , donc éventuellement non régulières,

par des suites de fonctions régulières : ce qui se cache derrière n'est autre qu'un résultat de densité d'un espace fonctionnel dans un autre. En pratique, il s'avère que ce type de résultat est très important. En effet, lorsqu'on désire établir un résultat pour des fonctions appartenant à un certain espace fonctionnel, il est souvent suffisant de l'établir seulement pour les fonctions appartenant à un sous-espace dense dans le premier, l'extension à l'espace initial se faisant alors par un passage à la limite. Ceci est intéressant lorsque les fonctions du sous-espace sont plus faciles à manipuler en pratique, comme par exemple les fonctions C^∞ à support compact.

Commençons par un premier résultat que l'on admet car difficile et fastidieux à démontrer. Il s'agit d'un premier théorème de densité.

Théorème 7.7 (*admis*) *Pour tout $1 \leq p < +\infty$, l'espace \mathcal{C}_c est dense dans L^p .*

Ainsi, pour toute fonction de L^p , il existe une suite de fonctions $(\varphi_n)_{n \in \mathbb{N}}$ de l'espace \mathcal{C}_c telle que

$$\lim_{n \rightarrow +\infty} \|\varphi_n - f\|_{L^p} = 0.$$

A priori, cette suite n'est pas associée à un produit de convolution. C'est l'utilisation de ce dernier qui va nous permettre de gagner en régularité, c'est-à-dire de passer de l'espace \mathcal{C}_c à l'espace \mathcal{C}_c^∞ . Pour ce faire, introduisons tout d'abord la notion d'unité approchée.

Définition 7.2 *On appelle unité approchée dans L^1 une suite de fonctions $(\varphi_n)_{n \in \mathbb{N}}$ de la forme $\varphi_n(x) = n \varphi(nx)$, avec φ une fonction telle que*

- $\varphi \geq 0$,
- $\text{supp}(\varphi) \subset [-1, 1]$,
- $\int_{\mathbb{R}} \varphi(x) dx = 1$.

Théorème 7.8 *Soit $(\varphi_n)_{n \in \mathbb{N}}$ une unité approchée dans L^1 et $f \in L^p$ avec $1 \leq p < +\infty$. Alors*

$$\lim_{n \rightarrow +\infty} \|\varphi_n \star f - f\|_{L^p} = 0.$$

Preuve : D'après le théorème 7.7, l'espace \mathcal{C}_c est dense dans L^p , donc pour tout $\varepsilon > 0$, on peut trouver $g \in \mathcal{C}_c$ telle que $\|f - g\|_{L^p} < \varepsilon/3$. Fixons-nous un tel $\varepsilon > 0$. Pour tout $n \in \mathbb{N}$, on a :

$$\|f \star \varphi_n - f\|_{L^p} \leq \|f \star \varphi_n - g \star \varphi_n\|_{L^p} + \|g \star \varphi_n - g\|_{L^p} + \|f - g\|_{L^p}.$$

Comme $\|f - g\|_{L^p} < \varepsilon/3$ et $\varphi_n \in L^1$ avec de surcroît $\|\varphi_n\|_{L^1} = 1$, le théorème 7.1 entraîne que

$$\|\varphi_n \star (f - g)\|_{L^p} \leq \|\varphi_n\|_{L^1} \|f - g\|_{L^p} < 1 \times \varepsilon/3.$$

Pour tout $n \in \mathbb{N}$, on a donc

$$\|f \star \varphi_n - f\|_{L^p} \leq 2\varepsilon/3 + \|g \star \varphi_n - g\|_{L^p}.$$

Il reste donc à montrer que $\|g \star \varphi_n - g\|_p < \varepsilon/3$ quitte à prendre n assez grand. La fonction g appartenant à \mathcal{C}_c est uniformément continue sur \mathbb{R} , donc il existe $\alpha > 0$ tel que si $|y| < \alpha$, alors pour tout $x \in \mathbb{R}$, on a $|g(x) - g(x - y)| < \varepsilon/3$. Ainsi, on a :

$$\begin{aligned} g(x) - g \star \varphi_n(x) &= g(x) \int_{\mathbb{R}} \varphi_n(y) dy - g \star \varphi_n(x), \\ &= \int_{\mathbb{R}} (g(x) - g(x - y)) \varphi_n(y) dy, \\ &= \int_{\{|y| < \alpha\}} (g(x) - g(x - y)) \varphi_n(y) dy \\ &\quad + \int_{\{|y| \geq \alpha\}} (g(x) - g(x - y)) \varphi_n(y) dy. \end{aligned}$$

On peut majorer chaque terme de cette somme :

- dès que $n > \frac{1}{\alpha}$, on a $\int_{\{|y| > \alpha\}} |\varphi_n(y)| dy = 0$, car $\text{supp}(\varphi_n) \subset \left[-\frac{1}{n}, \frac{1}{n}\right]$,
d'où

$$\left| \int_{\{|y| > \alpha\}} (g(x) - g(x - y)) \varphi_n(y) dy \right| \leq 2 \|g\|_{L^\infty} \int_{\{|y| > \alpha\}} |\varphi_n(y)| dy = 0.$$

- pour le premier terme, on a

$$\left| \int_{\{|y| < \alpha\}} (g(x) - g(x - y)) \varphi_n(y) dy \right| \leq \|\varphi_n\|_{L^1} \varepsilon/3 = \varepsilon/3.$$

On en déduit que, pour n assez grand, on a $|g(x) - g \star \varphi_n(x)| \leq \varepsilon/3$ et ce pour tout $x \in \mathbb{R}$: la suite $(g \star \varphi_n)_{n \in \mathbb{N}}$ converge donc uniformément vers g . Comme le produit de convolution de g et φ_n est celui de deux fonctions à support compact, il est aussi à support compact et même indépendant de n : on en déduit que la convergence de la suite $(g \star \varphi_n)_{n \in \mathbb{N}}$ a bien lieu dans \mathcal{C}_c , un espace qui s'injecte continûment dans L^p , de sorte que $\|g \star \varphi_n - g\|_{L^p}$ tend vers 0 lorsque n tend vers l'infini. Pour terminer la démonstration, il suffit de prendre n assez grand pour assurer que $\|g \star \varphi_n - g\|_{L^p} < \varepsilon/3$. \square

Un mot à propos de la terminologie "unité approchée". Si $p = 1$, la convergence de la suite $(\varphi_n \star f)_{n \in \mathbb{N}}$ vers f a lieu dans L^1 , dont on sait qu'elle est une algèbre de convolution, cf. le théorème 7.2. Cette algèbre n'a pas d'unité : cela signifie qu'il n'existe aucun élément de L^1 , que l'on pourrait par exemple noter δ tel que $\delta \star f = f \star \delta = f$ pour toute fonction f non nulle de L^1 . On comprend donc pourquoi l'on parle d'unité approchée.

En procédant à la même démonstration pour la fonction $f_K = f \mathbb{1}_K$, où K est un compact de \mathbb{R} , on obtient le résultat suivant.

Corollaire 7.1 *Soit $(\varphi_n)_{n \in \mathbb{N}}$ une unité approchée dans L^1 et $f \in L^p_{loc}$ avec $1 \leq p < +\infty$. Alors*

$$\lim_{n \rightarrow +\infty} \|\varphi_n \star f - f\|_{L^p_{loc}} = 0.$$

En particulier, si les φ_n sont dans \mathcal{C}_c^∞ , alors le résultat précédent combiné au théorème 7.6 indique que toute fonction $f \in L^1$ ou de L^1_{loc} peut être approchée dans L^1 ou dans L^1_{loc} par une suite de fonctions dont les éléments sont dans \mathcal{C}^∞ , ce qui est très utile en pratique. Avec un peu plus d'effort, on peut construire la suite $(\varphi_n)_{n \in \mathbb{N}}$ à partir d'une fonction φ particulière, de sorte que chacun de ses éléments soit à support compact : on obtient alors le résultat de densité suivant.

Théorème 7.9 *Pour $1 \leq p < \infty$, l'espace \mathcal{C}_c^∞ est dense dans les espaces L^1 et L^1_{loc} .*

Pour terminer ce chapitre, illustrons ce résultat sur quelques situations.

Exemple 7.2 *Pour chacune des fonctions suivantes, on peut trouver une suite de fonctions de \mathcal{C}_c qui l'approche dans L^1 :*

- $f(x) = \frac{1}{1+x^2}$.
- $g(x) = \mathbb{1}_{[-1,1]}(x)$.
- $h(x) = \frac{1}{\sqrt{|x|}} \mathbb{1}_{[-1,1]}(x)$.

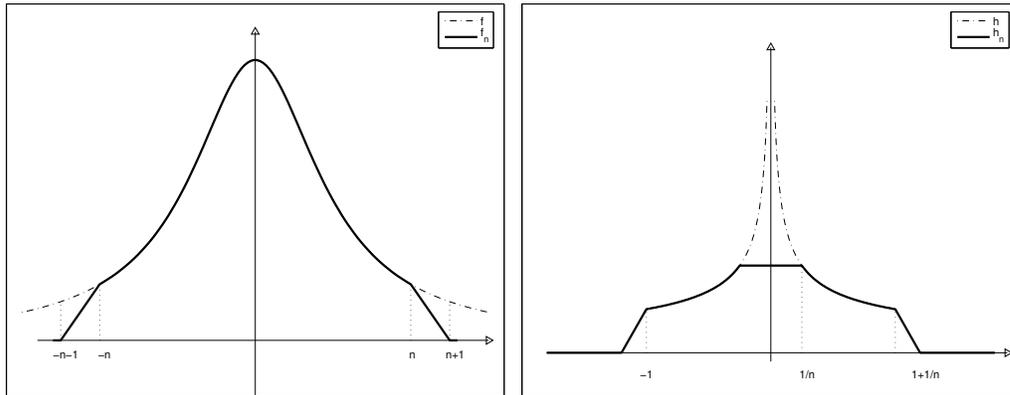


Figure 7.2: On peut approcher une fonction de L^1 par des fonctions continues à support compact.

Attention : le théorème 7.9 peut être généralisé aux espaces L^p avec $p > 1$, mais il est faux si $p = +\infty$. On peut le vérifier sur la fonction g de l'exemple 7.2 qui n'est pas continue sur \mathbb{R} , donc ne peut être la limite uniforme d'une suite de fonctions continues.

Probabilités

Dans ce dernier chapitre, nous allons revoir brièvement la théorie des probabilités en l'inscrivant dans le cadre général de la théorie de la mesure et de l'intégration. Bien évidemment, il ne s'agit pas de refaire un cours exhaustif sur ce trop vaste sujet mais plutôt de voir comment les concepts probabilistes vus en MIC2 se réécrivent dans le langage que nous utilisons depuis le début de ce cours.

SECTION 8.1

Analogies avec la théorie de la mesure

Avant de commencer à établir les résultats importants de la théorie des probabilités, reformulons tout d'abord les objets probabilistes d'intérêt à travers le langage de la théorie de la mesure. Pour ce faire, nous avons le tableau suivant :

Langage des probabilités	Théorie de la mesure
espace probabilisé : $(\Omega, \mathcal{A}, \mathbb{P})$ avec de plus $\mathbb{P}(\Omega) = 1$	espace mesuré : (X, \mathcal{A}, μ)
paramètre d'aléa : ω	inconnue : x
variable aléatoire réelle : $Z : \Omega \rightarrow \mathbb{R}$	fonction mesurable : $f : X \rightarrow \mathbb{R}$
espérance : $\mathbb{E}[Z] = \int_{\Omega} Z d\mathbb{P}$	intégrale : $\int_X f d\mu$

Une variable aléatoire réelle n'est rien d'autre qu'une fonction quantifiant, en un certain sens, une expérience aléatoire (par exemple le jeu de pile ou face est codé numériquement par une variable aléatoire Z prenant la valeur 1 si l'on obtient pile et 0 sinon). Afin de pouvoir calculer des quantités intéressantes, il nous faut introduire une structure probabiliste sur l'espace d'arrivée de la variable aléatoire réelle, en l'occurrence $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$: c'est l'objet des lois (de probabilité). La loi de la variable aléatoire réelle Z est définie sur $\mathcal{B}(\mathbb{R})$ par

$$P_Z(A) = \mathbb{P}(\{\omega \in \Omega : Z(\omega) \in A\}) = \mathbb{P}(Z^{-1}(A)) = \mathbb{E}[\mathbf{1}_A(Z)], \quad A \in \mathcal{B}(\mathbb{R}).$$

Autrement dit, c'est la mesure image (voir la Section 4.5.1) de la mesure de probabilité \mathbb{P} par la variable aléatoire Z . Par abus de notation, on la note

plus communément

$$P_Z(A) = \mathbb{P}(Z \in A).$$

En particulier, il est facile de vérifier qu'elle satisfait les axiomes d'une mesure de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Par ailleurs, sachant que toute fonction mesurable peut être approchée simplement par une suite de fonctions étagées, on en déduit que caractériser la loi d'une variable aléatoire réelle Z revient à calculer la quantité $\mathbb{E}[f(Z)]$ pour toute fonction f mesurable positive ou bornée.

Dans la théorie de la mesure, on peut montrer que toute mesure σ -finie (donc toute mesure de probabilité) sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ se décompose en une partie absolument continue et une partie étrangère pour la mesure de Lebesgue. Sans trop rentrer dans les détails, cela signifie dans notre cadre probabiliste que la loi P_Z de toute variable aléatoire réelle Z peut se décomposer de la manière suivante par rapport à la mesure de Lebesgue : il existe f_Z une fonction mesurable positive et μ une mesure satisfaisant la propriété

$$\mu(A) \neq 0 \Rightarrow m(A) = 0, \quad A \in \mathcal{B}(\mathbb{R}),$$

telles que

$$P_Z(A) = \int_A f_Z(x) dx + \mu(A).$$

Si $\mu = 0$ alors on dit que f_Z est la densité (de probabilité) de la variable aléatoire réelle Z . On dit aussi que Z est continue (attention, cela n'a rien à voir avec la notion usuelle de la continuité) et que P_Z est absolument continue par rapport à la mesure de Lebesgue. Dans le cas où $f_Z = 0$ et μ est une somme dénombrable de mesures de Dirac δ_{x_n} , alors la variable aléatoire est dite discrète. Une variable aléatoire est donc discrète si et seulement s'il existe des coefficients $(c_n)_{n \in \mathbb{N}}$ positifs ou nuls et une suite de points $(x_n)_{n \in \mathbb{N}}$ tels que

$$P_Z = \sum_{n \in \mathbb{N}} c_n \delta_{x_n} \quad \text{avec} \quad \sum_{n \in \mathbb{N}} c_n = 1.$$

Si une variable aléatoire ne prend qu'un nombre dénombrable de valeurs (x_n) , alors elle est discrète et

$$P_Z = \sum_{n \in \mathbb{N}} \mathbb{P}(Z = x_n) \delta_{x_n}.$$

Par exemple, si l'on revient au jeu du pile ou face, la variable aléatoire Z prenant la valeur 1 si l'on obtient pile et 0 sinon, toutes deux avec probabilité

1/2, suit une loi de Bernoulli de paramètre 1/2. Dans le langage de la théorie de la mesure, la loi P_Z s'écrit comme une somme de masses de Dirac pondérées

$$P_Z = \frac{1}{2} (\delta_0 + \delta_1).$$

De même, si Z correspond au nombre de résultats pile obtenus lorsqu'on lance à n reprises la même pièce de monnaie, alors Z est une variable aléatoire binomiale et sa loi est donnée par

$$P_Z = \left(\frac{1}{2}\right)^n \sum_{k=0}^n C_n^k \delta_k.$$

Bien évidemment, les autres exemples classiques comme la loi uniforme sur un ensemble fini ou la loi géométrique rentrent dans ce cadre. Dans le cas des variables aléatoires continues, nous considérons les lois uniforme sur un intervalle borné, normale (ou gaussienne), exponentielle, Gamma, etc...

SECTION 8.2

Espérance et intégrabilité

Étant donnée une variable aléatoire réelle Z et $p \in [1, \infty[$, on dit qu'elle est dans l'espace L^p si

$$\mathbb{E}[|Z|^p] = \int_{\Omega} |Z(\omega)|^p d\mathbb{P}(\omega) < +\infty.$$

En particulier si elle est dans l'espace L^1 , alors elle est dite intégrable et son espérance est bien définie (on rappelle que l'espérance n'est pas toujours définie pour une variable aléatoire donnée !!) et si de surcroît elle est dans l'espace L^2 , alors on dit qu'elle est de carré intégrable et sa variance

$$\text{Var}(Z) = \mathbb{E}[(Z - \mathbb{E}[Z])^2] = \mathbb{E}[Z^2] - \mathbb{E}[Z]^2,$$

est elle aussi bien définie (par l'inégalité de Cauchy-Schwarz à venir). En utilisant le Théorème 4.15 (voir la Section 4.5.1), la première quantité ci-dessus se réécrit comme

$$\mathbb{E}[|Z|^p] = \int_{\mathbb{R}} |x|^p dP_Z(x).$$

Plus généralement, si h est une fonction borélienne alors $h(Z)$ est une variable aléatoire réelle et sous réserve que h est à valeurs positives ou que $h(Z) \in L^1$, on a

$$\mathbb{E}[h(Z)] = \int_{\Omega} h(Z(\omega)) d\mathbb{P}(\omega) = \int_{\mathbb{R}} h(x) dP_Z(x).$$

Ainsi, on a $h(Z) \in L^1$ si et seulement si h est intégrable sur \mathbb{R} pour la mesure correspondant à la loi P_Z . Dans le cas des variables aléatoires discrètes, si $P_Z = \sum_n \mathbb{P}(Z = x_n) \delta_{x_n}$, alors

$$\mathbb{E}[h(Z)] = \sum_n h(x_n) \mathbb{P}(Z = x_n),$$

tandis que si Z est continue, on a

$$\mathbb{E}[h(Z)] = \int_{\mathbb{R}} h(x) f_Z(x) dx.$$

À titre d'exemple, une fonction indicatrice étant intégrable par rapport à toute mesure finie, la fonction de répartition F_Z de la variable aléatoire réelle Z est définie par le choix de la fonction $h(x) = \mathbb{1}_{]-\infty, t]}(x)$:

$$F_Z(t) = \mathbb{E}[\mathbb{1}_{]-\infty, t]}(Z)] = \mathbb{P}(Z \leq t) = \int_{\mathbb{R}} \mathbb{1}_{]-\infty, t]}(x) dP_Z(x), \quad t \in \mathbb{R}.$$

En étendant la notion la notion d'espérance pour des fonctions h boréliennes à valeurs dans le plan complexe \mathbb{C} , le choix de l'exponentielle complexe $h(x) = e^{itx}$, qui est bien une fonction intégrable car de module 1, donne lieu à la fonction caractéristique de Z :

$$\Phi_Z(t) = \mathbb{E}[e^{itZ}] = \int_{\mathbb{R}} e^{itx} dP_Z(x), \quad t \in \mathbb{R}.$$

De plus, lorsque Z admet une densité f_Z , alors la fonction caractéristique est la transformée de Fourier de la densité f_Z (ce concept ne sera abordé qu'en 4GMM). Pour terminer cette section, on mentionnera deux points très importants en pratique : la fonction de répartition et la fonction caractéristique sont deux objets qui caractérisent la loi d'une variable aléatoire réelle, c'est-à-dire que si deux variables aléatoires réelles ont même fonction de répartition (respectivement même fonction caractéristique), alors elles ont la même loi.

SECTION 8.3

Inégalités classiques

Les inégalités vues au chapitre 5 admettent une version probabiliste.

• **Inégalité de Hölder** : si Z_1 et Z_2 sont deux variables aléatoires réelles telles que $Z_1 \in L^p$ et $Z_2 \in L^q$, où les exposants $p, q > 1$ sont conjugués, alors

$$\mathbb{E}[|Z_1 Z_2|] \leq \mathbb{E}[|Z_1|^p]^{1/p} \mathbb{E}[|Z_2|^q]^{1/q}.$$

• **Inégalité de Cauchy-Schwarz** : c'est le cas précédent avec $p = q = 2$. En particulier en prenant $Z_1 = Z$ et $Z_2 = 1$ (une probabilité étant une mesure finie, 1 est de carré intégrable), on obtient

$$\mathbb{E}[|Z|] \leq \mathbb{E}[|Z|^2]^{1/2}.$$

• **Inégalité de Markov** : on peut aussi mentionner l'inégalité de Markov : si Z est une variable aléatoire intégrable, alors

$$\mathbb{P}(|Z| \geq x) \leq \frac{\mathbb{E}[|Z|]}{x}, \quad x > 0.$$

• **Inégalité de Chebyshev** : en choisissant dans l'inégalité précédente $Z = (Y - \mathbb{E}[Y])^2$, où Y est une variable aléatoire réelle de carré intégrable, on obtient

$$\mathbb{P}(|Y - \mathbb{E}[Y]| \geq x) \leq \frac{\text{Var}(Y)}{x^2}, \quad x > 0.$$

SECTION 8.4

Indépendance et loi

L'indépendance est une notion-clé de la théorie des probabilités. Dans ce qui suit, nous ne considérons par souci de simplicité que des couples de variables aléatoires réelles, mais ceci se généralise aisément au cas des variables aléatoires multidimensionnelles. Rappelons que deux variables aléatoires réelles Z_1 et Z_2 sont dites indépendantes si pour tous $A, B \in \mathcal{B}(\mathbb{R})$, on a l'égalité

$$\mathbb{P}(Y \in A; Z \in B) = \mathbb{P}(Y \in A) \mathbb{P}(Z \in B),$$

où le point-virgule représente l'intersection entre deux ensembles. Du point de vue des lois de Y et Z , l'égalité précédente se réécrit comme

$$P_{Y,Z}(A \times B) = P_Y(A) P_Z(B).$$

Ci-dessus, la loi $P_{Y,Z}$ est la loi jointe du couple (Y, Z) , vu comme une variable aléatoire à valeurs dans l'espace produit \mathbb{R}^2 . Ainsi, l'indépendance entre deux variables aléatoires réelles se traduit simplement par le fait que la loi jointe $P_{Y,Z}$ du couple (Y, Z) est le produit tensoriel $P_Y \otimes P_Z$ des deux lois marginales P_Y et P_Z . Dans le cas des variables continues, la densité jointe est le produit des densités marginales.

Notons que l'égalité précédente se réécrit comme

$$\mathbb{E}[\mathbf{1}_A(Y) \mathbf{1}_B(Z)] = \mathbb{E}[\mathbf{1}_{A \times B}(Y, Z)] = \mathbb{E}[\mathbf{1}_A(Y)] \mathbb{E}[\mathbf{1}_B(Z)].$$

Autrement dit, deux variables aléatoires réelles Y et Z sont indépendantes si et seulement si l'espérance du produit d'indicatrices vaut le produit des espérances de ces indicatrices. Sachant que toute fonction mesurable peut être approchée simplement par une suite de fonctions étagées, on en déduit que l'indépendance peut être reformulée de la manière suivante : deux variables aléatoires réelles Y et Z sont indépendantes si et seulement si l'on a l'égalité

$$\mathbb{E}[f(Y)g(Z)] = \mathbb{E}[f(Y)] \mathbb{E}[g(Z)],$$

pour toutes fonctions f, g boréliennes positives ou bornées. En particulier, l'indépendance est une notion plus forte que la non-corrélation, c'est-à-dire que

$$\mathbb{E}[YZ] = \mathbb{E}[Y] \mathbb{E}[Z].$$

En effet, pour obtenir cette égalité, le choix des deux fonctions f et g est imposé (la fonction identité pour les deux).

Enfin, terminons cette section par la notion de produit de convolution. On s'intéresse maintenant au calcul de la loi P_{Y+Z} de la somme $Y+Z$ lorsque Y et Z sont des variables aléatoires réelles indépendantes. Pour ce faire, il nous faut définir le produit de convolution de deux mesures. Soient μ et ν deux mesures sur un espace mesuré (X, \mathcal{A}, m) . On définit la mesure $\mu \star \nu$ sur (X, \mathcal{A}, m) de la façon suivante : pour tout $A \in \mathcal{A}$,

$$\mu \star \nu(A) = \int_X \int_X \mathbf{1}_A(x+y) d(\mu \otimes \nu)(x, y) = \int_X \int_X \mathbf{1}_A(x+y) d\mu(x) d\nu(y). \quad (8.1)$$

En d'autres termes, la mesure $\mu \star \nu$ est la mesure image de $\mu \otimes \nu$ par l'application $(x, y) \mapsto x + y$. On peut vérifier que la définition de la convolution de mesures est compatible avec la notion de convolution des fonctions

au sens où s'il existe deux fonctions positives $f, g \in L^1$ qui soient les densités respectives des mesures μ et ν par rapport à la mesure de Lebesgue, alors $f \star g$ est la densité de la mesure $\mu \star \nu$ par rapport à la mesure de Lebesgue. On déduit de la formule (8.1), par définition de l'intégrale, que pour toute fonction $\phi : X \rightarrow \mathbb{R}$ mesurable positive, on a

$$\int_X \phi(z) d(\mu \star \nu)(z) = \int_X \int_X \phi(x + y) d\mu(x) d\nu(y).$$

Ainsi, dans le cadre probabiliste des lois de variables aléatoires réelles indépendantes, on a pour toute fonction ϕ borélienne positive ou bornée,

$$\begin{aligned} \int_{\mathbb{R}} \phi dP_{Y+Z} &= \int_{\mathbb{R}^2} \phi(x + y) dP_{Y,Z}(x, y) \\ &= \int_{\mathbb{R}^2} \phi(x + y) dP_Y(x) dP_Z(y) \\ &= \int_{\mathbb{R}} \phi d(P_Y \star P_Z). \end{aligned}$$

On en déduit que les deux mesures P_{Y+Z} et $P_Y \star P_Z$ coïncident dès que les variables aléatoires sont indépendantes. En particulier, lorsque les variables sont continues, la densité f_{Y+Z} de la somme $Y + Z$ n'est rien d'autre que le produit de convolution $f_Y \star f_Z$ des densités de Y et de Z .

SECTION 8.5

Convergences

Terminons ce polycopié en énumérant les différents modes de convergence que l'on rencontre principalement. Dans la suite, on se donne une suite de variables aléatoires réelles $(Z_n)_{n \in \mathbb{N}}$ ainsi qu'une variable aléatoire réelle Z .

• **Convergence presque sûre :** la suite $(Z_n)_{n \in \mathbb{N}}$ converge presque sûrement vers Z si pour tout $\omega \in \Omega$ en dehors d'un ensemble négligeable pour la probabilité \mathbb{P} , c'est-à-dire de probabilité nulle, on a $Z_n(\omega) \rightarrow Z(\omega)$ lorsque n tend vers l'infini. Le complémentaire d'un ensemble négligeable étant de probabilité 1, ceci se réécrit comme

$$\mathbb{P}\left(\lim_{n \rightarrow +\infty} Z_n = Z\right) = 1.$$

Il s'agit de la version probabiliste de la convergence presque partout. C'est celle qui apparaît dans la loi forte des grands nombres pour des variables aléatoires indépendantes, de même loi et intégrables.

• **Convergence en probabilité** : la suite $(Z_n)_{n \in \mathbb{N}}$ converge en probabilité vers Z si pour tout $\varepsilon > 0$, on a

$$\lim_{n \rightarrow +\infty} \mathbb{P}(|Z_n - Z| > \varepsilon) = 0.$$

Il s'agit de la version probabiliste de la convergence en mesure.

• **Convergence en loi** : la suite $(Z_n)_{n \in \mathbb{N}}$ converge en loi vers Z si pour toute fonction $\phi \in \mathcal{C} \cap L^\infty$, on a

$$\lim_{n \rightarrow +\infty} \mathbb{E}[\phi(Z_n)] = \mathbb{E}[\phi(Z)].$$

Il s'agit de la convergence étroite pour les analystes, que nous n'avons pas introduite dans ce cours. Néanmoins, il est très importante au sens où c'est cette convergence qui a lieu dans le théorème central limite pour des variables aléatoires indépendantes, de même loi et de carré intégrable.

• **Convergence dans L^p** : étant donné $1 \leq p < +\infty$, la suite $(Z_n)_{n \in \mathbb{N}}$ converge dans L^p vers Z si l'on a

$$\lim_{n \rightarrow +\infty} \mathbb{E}[|Z_n - Z|^p] = 0.$$

Bien évidemment, pour que cette convergence ait un sens, il faut supposer que toutes les variables sont dans l'espace L^p .

Une fois ces modes de convergence introduits, on peut reformuler tous les théorèmes fondamentaux du chapitre 4, à savoir les théorèmes de convergence monotone de Beppo-Levi et de convergence dominée de Lebesgue, ainsi que les théorèmes de continuité et de dérivabilité. Pour ce faire, on remplace les (suites de) fonctions par des (suites de) variables aléatoires réelles, la convergence presque partout par la convergence presque sûre et enfin les intégrales par des espérances. Par exemple, la version probabiliste du théorème de convergence dominée de Lebesgue est la suivante :

si la suite $(Z_n)_{n \in \mathbb{N}}$ converge presque sûrement vers Z et s'il existe une variable aléatoire positive Y intégrable et telle que $|Z_n| \leq Y$ pour tout $n \in \mathbb{N}$, alors $Z \in L^1$ et

$$\lim_{n \rightarrow +\infty} \mathbb{E}[|Z_n - Z|] = 0 \quad \text{et en particulier} \quad \lim_{n \rightarrow +\infty} \mathbb{E}[Z_n] = \mathbb{E}[Z].$$

Par ailleurs, lorsqu'on considère la version probabiliste du théorème de dérivabilité sous l'intégrale, on est en mesure de relier la fonction caractéristique et les moments d'une variable aléatoire : s'il existe $p \geq 1$ tel que $Z \in L^p$, alors on a

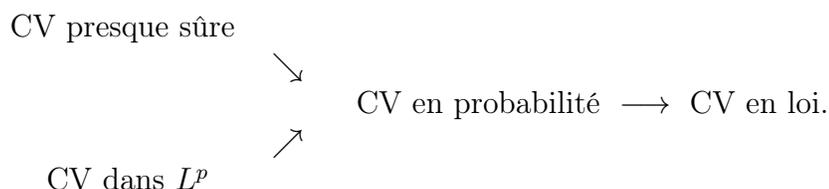
$$\Phi_Z^{(p)}(t) = i^p \mathbb{E}[Z^p e^{itZ}], \quad t \in \mathbb{R},$$

et en particulier lorsqu'on choisit $t = 0$ on obtient que

$$\mathbb{E}[Z^p] = \frac{\Phi_Z^{(p)}(0)}{i^p}.$$

Cette observation est à la base de la démonstration du théorème central limite.

À présent, on peut se demander si les différents modes de convergence énoncés ci-dessus sont comparables. Le schéma général des implications entre convergences est le suivant :



Par ailleurs, on a aussi les résultats suivants :

- La convergence en loi n'entraîne pas la convergence en probabilité, sauf lorsque la limite Z est une constante.

- La convergence en probabilité n'entraîne pas la convergence p.s.

- Les convergences p.s. et dans L^p ne sont pas comparables en général. En revanche, si la convergence p.s. a lieu et s'il existe une variable aléatoire positive et intégrable Y telle que $|Z_n|^p \leq Y$, alors la limite Z est dans l'espace L^p et la convergence a aussi lieu dans L^p : c'est une application de la version probabiliste du théorème de convergence dominée.

- Enfin, deux faits connus et utiles en pratique : la convergence en loi est équivalente à :

- la convergence simple des fonctions de répartition en tout point où la fonction de répartition de la limite est continue ;

- la convergence simple des fonctions caractéristiques.