

Mesure et intégration

Licence de Mathématiques
Université Lyon 1

Dragoș Iftimie

Table des matières

Introduction	3
1 Limites inférieures et supérieures. Dénombrément	3
1.1 Limites inférieures et supérieures	3
1.2 Dénombrément	4
2 Ensembles et fonctions mesurables. Tribus	5
2.1 Tribus. Espaces mesurables	5
2.2 Fonctions mesurables	5
2.3 Tribu de Borel	6
2.4 Propriétés des applications mesurables à valeurs dans \mathbb{R}	7
2.5 Boréliens de $\overline{\mathbb{R}}$, fonctions à valeurs dans \mathbb{R}	7
2.6 Fonctions étagées	8
3 Mesures positives	8
4 Intégration des fonctions positives	10
4.1 Intégration des fonctions étagées positives	10
4.2 Intégration des fonctions mesurables positives	11
5 Théorèmes de convergence	12
5.1 Théorème de convergence monotone et applications. Lemme de Fatou	12
5.2 Intégration des fonctions mesurables complexes	13
5.3 Théorème de convergence dominée de Lebesgue	14
5.4 Presque partout.	14
5.5 Tribu complétée. Mesure complétée. Tribu et mesure de Lebesgue	15
5.6 Intégrales à paramètre	17
5.7 Comparaison entre l'intégrale de Riemann et l'intégrale de Lebesgue	18
6 Intégration sur un espace produit	19
6.1 Tribu produit	19
6.2 Produits tensoriels de mesures	19
6.3 Théorèmes de Fubini	20

Mise à jour le 2 décembre 2024.

7 Mesure de Lebesgue	22
7.1 Existence	22
7.2 Unicité	23
8 Changement de variables	24
8.1 Cas linéaire	24
8.2 Cas général	24
8.3 Applications	25
8.3.1 Coordonnées polaires dans le plan \mathbb{R}^2	25
8.3.2 Coordonnées sphériques dans l'espace \mathbb{R}^3	25
8.3.3 Cas de \mathbb{R}^d	26

Introduction

Ce texte s'inspire fortement des 3 supports de cours de licence suivants :

- T. Gallay, Théorie de la mesure et de l'intégration, Grenoble.
- N. Lerner, Intégration, Rennes.
- P. Mironeanu, Mesure et intégration, Lyon.

Nous disposons de la théorie de l'intégrale de Riemann qui permet d'intégrer beaucoup de fonctions dont notamment les fonctions continues. Cette théorie est-elle suffisante pour les besoins de l'intégration ? Voici quelques problèmes posés par l'intégrale de Riemann.

- L'intégrale de Riemann est essentiellement réservée aux fonctions continues, en tout cas il faut au moins que les fonctions soient bornées. Un théorème du à Lebesgue dit d'ailleurs qu'une fonction bornée est intégrable Riemann si et seulement si l'ensemble de ses points de discontinuité est négligeable. En utilisant les intégrales généralisées, on peut aussi traiter des fonctions continues avec quelques singularités ou sur des intervalles non-bornés mais cela reste très limité et pas facile à manipuler.
- Une fonction relativement simple comme la fonction caractéristique de \mathbb{Q} n'est pas intégrable Riemann.
- Si f est continue et dérivable, dans quelle mesure l'égalité $\int_a^b f'(x) dx = f(b) - f(a)$ est-elle vraie ?
- Très important, dans quelle mesure la convergence simple de f_n vers f implique la convergence de leurs intégrales ? Le critère pour l'intégrale de Riemann nécessite la convergence uniforme de la suite f_n , condition bien trop restrictive.
- Nous avons besoin d'intégrer par rapport à d'autres mesures, par exemple en théorie des probabilités. C'est-à-dire qu'on pourrait convenir que la mesure d'un intervalle $[a, b]$ n'est pas forcément $b - a$. Que devient l'intégrale dans ce cas ?

Pour répondre à ces questions, Lebesgue a construit une autre théorie de l'intégrale que l'on connaît aujourd'hui sous le nom d'intégrale de Lebesgue. Son idée de départ a été très simple ; voici un aperçu (très) simplifié pour donner une petite idée de son approche. Pour mesurer l'aire du sous-graphe d'une fonction (qui est l'intégrale de la fonction), Riemann prenait l'intersection du sous-graphe avec une droite verticale et mesurait la longueur du segment obtenu (puis "sommait" toutes ces longueurs). Lebesgue fait la même chose mais en prenant des droites horizontales au lieu de droites verticales. Étonnamment, cette approche donne une notion d'intégrale bien plus robuste et permet d'intégrer bien plus de fonctions. Malheureusement, contrairement à l'intersection du sous-graphe avec une droite verticale qui est simplement un segment, l'intersection du sous-graphe avec une droite horizontale peut être un ensemble très compliqué. Ce n'est pas clair qu'on puisse le mesurer. On ne peut pas mesurer tous les ensembles. Citons à cet effet le paradoxe de Banach-Tarski qui date de 1923 : ces auteurs ont découpé la boule unité de \mathbb{R}^3 en un nombre fini de morceaux qui peuvent être réarrangés pour faire deux boules unité complètes ! Cela veut bien dire que l'on ne peut pas mesurer de manière raisonnable ces morceaux, sinon on aurait d'une part que la mesure de la boule unité est la somme des mesures de ces morceaux, et cette même somme fera aussi le double de la mesure de la boule unité ! Une bonne partie de ce cours sera d'ailleurs dédié à notion de mesurabilité.

1 Limites inférieures et supérieures. Dénombrement

1.1 Limites inférieures et supérieures

Rappelons d'abord que $\overline{\mathbb{R}}$, la droite réelle achevée, est définie par $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$. On peut additionner et multiplier les éléments de $\overline{\mathbb{R}}$, à l'exception des indéterminations $(+\infty) + (-\infty)$ et $0 \times (\pm\infty)$. Rappelons maintenant les définitions de sup et inf.

Définition 1.1 (inf, sup). Soit $A \subset \mathbb{R}$ non-vide. On définit $\sup A \in \mathbb{R} \cup \{+\infty\}$ comme le plus petit majorant de A et $\inf A \in \mathbb{R} \cup \{-\infty\}$ comme le plus grand minorant de A .

Soit $(x_n)_n$ une suite de \mathbb{R} et posons $y_n = \sup_{k \geq n} x_k$. La suite $y_n \in \mathbb{R} \cup \{+\infty\}$ est décroissante, elle admet donc une limite qui est égale à son inf. Cela justifie la définition suivante.

Définition 1.2 (liminf, limsup). Soit $(x_n)_n$ une suite de \mathbb{R} .

— On définit la limite supérieure de x_n comme

$$\limsup_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} \sup_{k \geq n} x_k = \inf_{n \geq 0} \sup_{k \geq n} x_k.$$

— On définit la limite inférieure de x_n comme

$$\liminf_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} \inf_{k \geq n} x_k = \sup_{n \geq 0} \inf_{k \geq n} x_k.$$

Rappelons enfin que la valeur d'adhérence d'une suite se définit comme la limite d'une sous-suite de la suite en question. Nous travaillerons ici avec des valeurs d'adhérence dans $\overline{\mathbb{R}}$.

La proposition suivante regroupe plusieurs propriétés des limites inférieures et supérieures.

Proposition 1.3. Soit $(x_n)_n$ une suite de \mathbb{R} . Nous avons les affirmations suivantes.

- a) $\limsup_{n \rightarrow \infty} x_n$ est la plus grande valeur d'adhérence de x_n et $\liminf_{n \rightarrow \infty} x_n$ est la plus petite valeur d'adhérence de x_n .
- b) Nous avons que $x_n \rightarrow l \in \overline{\mathbb{R}}$ quand $n \rightarrow \infty$ si et seulement si $\liminf_{n \rightarrow \infty} x_n = \limsup_{n \rightarrow \infty} x_n = l$.
- c) Si $(y_n)_n$ est une autre suite de \mathbb{R} alors $\limsup_{n \rightarrow \infty} (x_n + y_n) \leq \limsup_{n \rightarrow \infty} x_n + \limsup_{n \rightarrow \infty} y_n$ dès lors que cette dernière somme est bien définie.

1.2 Dénombrément

Il existe deux définitions d'un ensemble dénombrable suivant que les ensembles finis sont considérés dénombrables ou pas. Nous adopterons ici la convention qu'ils ne le sont pas.

Définition 1.4 (dénombrabilité). — Un ensemble X est dit dénombrable s'il existe une bijection de X dans \mathbb{N} .

— Un ensemble X est dit au plus dénombrable, abrégé en a.p.d., s'il est fini ou dénombrable.

Nous avons les deux propriétés suivantes.

Proposition 1.5. a) Une union a.p.d. d'ensembles a.p.d. est a.p.d..

b) Un produit fini d'ensembles a.p.d. est a.p.d..

Les exemples classiques sont \mathbb{Q} qui est dénombrable et \mathbb{R} qui ne l'est pas. Cela vient de l'écriture décimale et de la non dénombrabilité de $\mathcal{P}(\mathbb{N})$. Nous avons en effet le théorème de Cantor suivant.

Théorème 1.6 (Cantor). Si E est un ensemble non-vide, il n'existe pas de bijection entre E et $\mathcal{P}(E)$ (on a désigné par $\mathcal{P}(E)$ l'ensemble des parties de E).

2 Ensembles et fonctions mesurables. Tribus

2.1 Tribus. Espaces mesurables

Définition 2.1 (tribu, espace mesurable). Soit X un ensemble. Une famille \mathcal{M} de parties de X est une tribu sur X si :

- a) $A \in \mathcal{M}$ implique $A^c \in \mathcal{M}$;
- b) si $A_n \in \mathcal{M}$ pour tout $n \in \mathbb{N}$, alors $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{M}$;
- c) $\emptyset, X \in \mathcal{M}$.

Le couple (X, \mathcal{M}) est dit espace mesurable et les éléments de \mathcal{M} sont dits ensembles mesurables.

En anglais, une tribu est une “ σ -algebra”, d'où parfois la terminologie σ -algèbre en français.

Il est facile de voir que les propriétés a) et b) impliquent c). Une tribu est donc une famille non vide de sous-ensembles stable par union dénombrable et par passage au complémentaire. Le complémentaire d'une union étant l'intersection des complémentaires, on en déduit qu'une tribu est aussi stable par intersection dénombrable.

Exemples. Soit X un ensemble.

- $\mathcal{M} = \{\emptyset, X\}$ est une tribu.
- $\mathcal{M} = \mathcal{P}(X)$ est une tribu. Cas particulier $X = \mathbb{N}$: la tribu qui sera considérée sur \mathbb{N} est $\mathcal{P}(\mathbb{N})$.
- Si A_1, \dots, A_n forment une partition de X (disjoints deux à deux et d'union X), alors toutes les unions possibles de A_j forment une tribu sur X .

Pour définir la tribu engendré par une famille de sous-ensembles, nous avons besoin d'un résultat préliminaire.

Lemme 2.2. Soit X un ensemble et $(\mathcal{M}_i)_{i \in I}$ une famille arbitraire de tribus sur X . Alors $\bigcap_{i \in I} \mathcal{M}_i$ est une tribu sur X .

Cela justifie la définition suivante.

Définition 2.3 (tribu engendrée). Soit X un ensemble et \mathcal{F} une famille de parties de X . On définit $\mathcal{M}(\mathcal{F})$, la tribu engendrée par \mathcal{F} , comme l'intersection de toutes les tribus de X qui contiennent \mathcal{F} .

Nous avons la propriété suivante :

Proposition 2.4. La tribu engendrée par \mathcal{F} est la plus petite tribu qui contient \mathcal{F} .

Fin du cours 1 (03/09/2024).

Définissons aussi la tribu trace.

Définition 2.5. Soit (X, \mathcal{M}) un espace mesurable et $A \subset X$ un sous-ensemble arbitraire. La tribu trace sur A est la tribu $\mathcal{M}_A = A \cap \mathcal{M} = \{A \cap B ; B \in \mathcal{M}\}$.

2.2 Fonctions mesurables

Une fonction est dite mesurable si l'image réciproque d'un ensemble mesurable est mesurable :

Définition 2.6 (fonction mesurable). Soient (X_1, \mathcal{M}_1) et (X_2, \mathcal{M}_2) deux espaces mesurables et $f : X_1 \rightarrow X_2$ une application. La fonction f est dite mesurable (par rapport à ces deux tribus) si pour tout $A_2 \in \mathcal{M}_2$ nous avons que $f^{-1}(A_2) \in \mathcal{M}_1$.

La notion de mesurabilité d'une application $f : X_1 \rightarrow X_2$ dépend des tribus considérées sur X_1 et X_2 . On doit donc spécifier à chaque fois de quelles tribus il s'agit. Par abus de notation, lorsqu'une seule tribu a été définie sur X_1 et X_2 , ou que les tribus considérées sont sous-entendues, on peut omettre de préciser les tribus et parler simplement d'application mesurable de X_1 dans X_2 .

Voici quelques propriétés des applications mesurables.

- Proposition 2.7.**
- a) Soient (X_1, \mathcal{M}_1) , (X_2, \mathcal{M}_2) et (X_3, \mathcal{M}_3) trois espaces mesurables et $f : X_1 \rightarrow X_2$, $g : X_2 \rightarrow X_3$ deux applications mesurables (pour les tribus considérées). Alors $g \circ f : X_1 \rightarrow X_3$ est mesurable (pour les tribus considérées).
 - b) Soit (X, \mathcal{M}) un espace mesurable, Y un ensemble et $f : X \rightarrow Y$ une application. Soit $\mathcal{N} = \{B \subset Y ; f^{-1}(B) \in \mathcal{M}\}$. Alors \mathcal{N} est une tribu sur Y qui rend f mesurable, et c'est la plus grande tribu avec cette propriété.
 - c) Soient (X_1, \mathcal{M}_1) et (X_2, \mathcal{M}_2) deux espaces mesurables et $f : X_1 \rightarrow X_2$ une application. On suppose que \mathcal{M}_2 est engendrée par une famille $\mathcal{F} : \mathcal{M}_2 = \mathcal{M}(\mathcal{F})$. Alors f est mesurable si et seulement si pour tout $A_2 \in \mathcal{F}$ nous avons que $f^{-1}(A_2) \in \mathcal{M}_1$.

On retiendra que la composition de deux fonctions mesurables est mesurable et qu'il suffit de vérifier la mesurabilité d'une fonction sur une famille qui engendre la tribu de l'espace d'arrivée de la fonction.

2.3 Tribu de Borel

Un exemple fondamental de tribu est la tribu de Borel.

Définition 2.8 (tribu de Borel, borélien). Soit (X, d) un espace métrique. La tribu de Borel sur X , notée par $\mathcal{B}(X)$, est la tribu engendrée par les ouverts de X (par rapport à la distance d). Les éléments de la tribu de Borel sont dits ensembles boréliens, ou tout simplement boréliens.

Attention, les ouverts ne forment pas une tribu car le complémentaire d'un ouvert n'est généralement pas ouvert. Les boréliens forment une famille très grande. Voici quelques exemples de boréliens :

- les fermés ;
- toute union dénombrable d'ensembles fermés ;
- \mathbb{Q} est borélien dans \mathbb{R} ;
- les irrationnels forment un ensemble borélien dans \mathbb{R} .

Les fonctions continues sont mesurables pour la tribu de Borel.

Proposition 2.9. Soient (X_1, d_1) , (X_2, d_2) deux espaces métriques et $f : X_1 \rightarrow X_2$ une fonction continue. Alors f est mesurable lorsqu'on munit X_1 et X_2 de leurs tribus de Borel.

Définition 2.10 (pavé). Un pavé de \mathbb{R}^d est un produit d'intervalles fermés et bornés.

Proposition 2.11. Il existe une famille dénombrable de pavés de \mathbb{R}^d telle que tout ouvert de \mathbb{R}^d est union dénombrable de ces pavés. En particulier, cette famille dénombrable de pavés engendre les boréliens de \mathbb{R}^d . La même chose est vraie pour une certaine famille dénombrable de boules ouvertes, et aussi pour une certaine famille dénombrable de produits de boules ouvertes.

Dans le cas de \mathbb{R} on peut restreindre encore plus la famille qui engendre les boréliens.

Proposition 2.12. La tribu $\mathcal{B}(\mathbb{R})$ est engendrée par :

- a) les intervalles $]a, +\infty[$ quand a parcourt \mathbb{R} ;
- b) les intervalles $[a, +\infty[$ quand a parcourt \mathbb{R} ;
- c) les intervalles $]-\infty, a[$ quand a parcourt \mathbb{R} ;
- d) les intervalles $]-\infty, a]$ quand a parcourt \mathbb{R} .

2.4 Propriétés des applications mesurables à valeurs dans \mathbb{R}

Une application immédiate de la caractérisation des boréliens sur \mathbb{R} vue au-dessus et de la proposition 2.7 nous permet de montrer par exemple l'énoncé suivant :

Corollaire 2.13. *Soit (X, \mathcal{M}) un espace mesurable et $f : X \rightarrow \mathbb{R}$. On munit \mathbb{R} de la tribu de Borel. Alors f est mesurable si et seulement si $f^{-1}([a, +\infty]) \in \mathcal{M}$ pour tout $a \in \mathbb{R}$.*

Une autre application nous permet de montrer que toute application monotone est mesurable.

Proposition 2.14. *Soit $X \subset \mathbb{R}$ et $f : X \rightarrow \mathbb{R}$ monotone. Alors f est mesurable lorsqu'on munit X de la tribu trace de $\mathcal{B}(\mathbb{R})$ et \mathbb{R} de la tribu de Borel.*

Nous munirons par défaut \mathbb{R} , ou \mathbb{R}^d , par la tribu de Borel. Si aucune mention n'est faite, il faut supposer que la tribu considérée sur \mathbb{R} , ou \mathbb{R}^d , est la tribu de Borel.

La fonction indicatrice d'un ensemble est mesurable si et seulement si l'ensemble est mesurable.

Proposition 2.15. *Soit (X, \mathcal{M}) un espace mesurable, $A \subset X$. La fonction indicatrice de A , $\chi_A : X \rightarrow \mathbb{R}$, est mesurable si et seulement si A est mesurable.*

Proposition 2.16. *Soit (X, \mathcal{M}) et (Y, \mathcal{N}) deux espaces mesurables, $u_1, \dots, u_d : X \rightarrow \mathbb{R}$ des fonctions mesurables et $\phi : \mathbb{R}^d \rightarrow Y$ mesurable. Alors l'application $x \mapsto \phi(u_1(x), \dots, u_d(x))$ est mesurable de X dans Y .*

Voici un corollaire immédiat.

Corollaire 2.17. *Soit (X, \mathcal{M}) un espace mesurable.*

- a) *Une application $f : X \rightarrow \mathbb{C}$ est mesurable si et seulement $\text{Re}(f)$ et $\text{Im}(f)$ sont mesurables de X dans \mathbb{R} . De plus, si f est mesurable alors $|f|$ est mesurable aussi.*
- b) *Si $f, g : X \rightarrow \mathbb{C}$ sont mesurables, alors $f + g$ et fg sont mesurables aussi.*

Proposition 2.18. *Soit (X, \mathcal{M}) un espace mesurable et $f : X \rightarrow \mathbb{C}$ mesurable. Il existe une application $\alpha : X \rightarrow \mathbb{C}$ mesurable telle que $|\alpha| = 1$ et $f = \alpha|f|$ partout.*

Fin du cours 2 (10/09/2024).

2.5 Boréliens de $\overline{\mathbb{R}}$, fonctions à valeurs dans $\overline{\mathbb{R}}$

Sur $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ nous avons une structure d'espace métrique donnée par la distance $d(x, y) = |\arctan x - \arctan y|$ où on convient que $\arctan(\pm\infty) = \pm\frac{\pi}{2}$. On peut donc parler de boréliens et de mesurabilité sur $\overline{\mathbb{R}}$. On vérifie aisément que les boules ouvertes de $\overline{\mathbb{R}}$ sont les intervalles (où $\pm\infty$ peuvent être des extrémités d'intervalles et appartenir à l'intervalle ou pas).

Sur $\overline{\mathbb{R}}$ on peut faire des limites, prendre des sup, des inf, des limsup et des liminf. Ce qui a été fait dans la partie 1.1 reste valable dans $\overline{\mathbb{R}}$ à ceci près que lorsqu'il y a une somme il faut imposer la condition que la somme soit bien définie (c'est-à-dire qu'on ne se retrouve pas à sommer $-\infty$ et $+\infty$).

La proposition 2.12 est vraie dans $\overline{\mathbb{R}}$ aussi : la tribu $\mathcal{B}(\overline{\mathbb{R}})$ est engendrée par les intervalles $]a, +\infty]$ quand a parcourt \mathbb{R} .

Nous avons le résultat suivant :

Proposition 2.19. *Soit (X, \mathcal{M}) un espace mesurable et $f_n : X \rightarrow \overline{\mathbb{R}}$ une suite de fonctions mesurables. Alors*

- a) $\sup_n f_n$ et $\inf_n f_n$ sont des fonctions mesurables de X dans $\overline{\mathbb{R}}$;
- b) $\limsup_{n \rightarrow \infty} f_n$ et $\liminf_{n \rightarrow \infty} f_n$ sont des fonctions mesurables de X dans $\overline{\mathbb{R}}$;
- c) *Si $f : X \rightarrow \overline{\mathbb{R}}$ est limite simple de f_n , alors f est mesurable.*

Dans $\overline{\mathbb{R}}$ nous utiliserons la convention $0 \cdot \infty = 0$. L'addition $(-\infty) + (+\infty)$ reste interdite.

2.6 Fonctions étagées

Une fonction étagée est une fonction positive mesurable qui ne prend qu'un nombre fini de valeurs.

Définition 2.20 (fonction étagée). Soit (X, \mathcal{M}) un espace mesurable. Une fonction f est dite étagée sur X si :

- f est à valeurs dans \mathbb{R}_+ et mesurable de X dans \mathbb{R}_+ ;
- $f(X)$ est un ensemble fini.

Si f est étagée, en posant $f(X) = \{\alpha_1, \dots, \alpha_n\}$ (valeurs distinctes) et $A_j = f^{-1}(\alpha_j)$ nous avons que

$$f = \alpha_1 \chi_{A_1} + \cdots + \alpha_n \chi_{A_n}$$

où χ_{A_j} désigne la fonction indicatrice de A_j . Cette écriture est unique si les α_j sont distincts 2 à 2 et si les A_j forment une partition de X . On l'appelle écriture canonique de la fonction étagée f .

Théorème 2.21. Soit (X, \mathcal{M}) un espace mesurable et $f : X \rightarrow \overline{\mathbb{R}}_+$ mesurable. Alors il existe une suite de fonctions étagées positives f_n telles que

- la suite f_n est croissante : $0 \leq f_n \leq f_{n+1} \leq f$;
- la suite f_n tend vers f simplement.

Si on a de plus que f est bornée, alors on peut supposer que la suite f_n tend vers f uniformément sur X .

3 Mesures positives

Définition 3.1 (mesure positive). Soit (X, \mathcal{M}) un espace mesurable. Une mesure positive sur (X, \mathcal{M}) est une application $\mu : \mathcal{M} \rightarrow \overline{\mathbb{R}}_+$ telle que

- $\mu(\emptyset) = 0$;
- Si $(A_n)_n$ est une suite d'ensembles mesurables 2 à 2 disjoints, alors

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

Le triplet (X, \mathcal{M}, μ) est dit espace mesuré.

La deuxième propriété de la définition est dite σ -additivité. Les mesures peuvent être de plusieurs types.

Définition 3.2. Soit (X, \mathcal{M}, μ) un espace mesuré. On dit que :

- la mesure μ est finie si $\mu(X) < \infty$;
- la mesure μ est une mesure de probabilités si $\mu(X) = 1$;
- la mesure μ est σ -finie s'il existe une suite d'ensembles mesurables $(A_n)_n$ tels que $X = \bigcup_{n \in \mathbb{N}} A_n$ et $\mu(A_n) < \infty$ pour tout $n \in \mathbb{N}$;
- la mesure μ est borélienne si la tribu \mathcal{M} est la tribu de Borel ;
- la mesure μ est de Radon si elle est borélienne et finie sur les compacts (la mesure de tout compact est finie).

Exemples.

- a) Si X est un ensemble fini et $\mathcal{M} = \mathcal{P}(X)$, $\mu(A) = \text{card}(A)$ pour tout $A \subset X$ est une mesure finie.
- b) Si X est un ensemble fini et $\mathcal{M} = \mathcal{P}(X)$, $\mu(A) = \frac{\text{card}(A)}{\text{card}(X)}$ pour tout $A \subset X$ est une mesure de probabilités.

c) Si X est un ensemble arbitraire et $\mathcal{M} = \mathcal{P}(X)$, la mesure de comptage définie par

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

est une mesure.

d) Soit (X, \mathcal{M}) un espace mesurable et $a \in X$. L'application δ_a définie par

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

est une mesure appelée mesure de Dirac en a (ou masse de Dirac).

- e) Sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ il existe une unique mesure positive μ telle que $\mu([a, b]) = \mu(]a, b[) = b - a$ pour tout $a < b$ finis. Cette mesure est dite mesure de Borel. L'existence de la mesure de Borel n'est pas du tout évidente. Sa construction est difficile et fait l'objet d'un chapitre ultérieur.
- f) Sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ il existe une unique mesure positive μ telle que la mesure de tout pavé est le produit des longueurs de ses côtés : $\mu(\prod_j [a_j, b_j]) = \prod_j (b_j - a_j)$. C'est la mesure de Borel sur \mathbb{R}^d .
- g) Soit (X, \mathcal{M}, μ) un espace mesuré et $f : X \rightarrow Y$ une application. Soit $\mathcal{N} = \{B \subset Y ; f^{-1}(B) \in \mathcal{M}\}$ la tribu image sur Y . L'application $\nu(B) = \mu(f^{-1}(B))$ pour tout $B \in \mathcal{N}$ est une mesure sur \mathcal{N} dite mesure image de μ .

Avant de montrer quelques propriétés des mesures, nous avons besoin du résultat suivant sur les séries à deux indices.

Lemme 3.3. *On considère des nombres $a_{m,n} \in \overline{\mathbb{R}}_+$, $m, n \in \mathbb{N}$. Nous avons que*

$$\sum_{m \in \mathbb{N}} \left(\sum_{n \in \mathbb{N}} a_{m,n} \right) = \sum_{n \in \mathbb{N}} \left(\sum_{m \in \mathbb{N}} a_{m,n} \right).$$

La valeur commune de ces deux sommes est notée plus simplement par

$$\sum_{m,n \in \mathbb{N}} a_{m,n} = \sum_{m \in \mathbb{N}} \left(\sum_{n \in \mathbb{N}} a_{m,n} \right) = \sum_{n \in \mathbb{N}} \left(\sum_{m \in \mathbb{N}} a_{m,n} \right).$$

Voici quelques opérations sur les mesures.

- Proposition 3.4 (opérations sur les mesures).**
- a) Si μ est mesure et $\alpha \in \overline{\mathbb{R}}_+$ alors $\alpha\mu$ est une mesure. Ici $(\alpha\mu)(A) = \alpha\mu(A)$ avec la convention $0 \cdot \infty = 0$.
 - b) Si μ_1 et μ_2 sont deux mesures sur un même espace, alors $\mu_1 + \mu_2$ est une mesure. Ici on a posé $(\mu_1 + \mu_2)(A) = \mu_1(A) + \mu_2(A)$.
 - c) Si $(\mu_n)_n$ est une suite de mesures sur un même espace, alors $\sum_{n \in \mathbb{N}} \mu_n$ est une mesure. Ici on a posé $(\sum_{n \in \mathbb{N}} \mu_n)(A) = \sum_{n \in \mathbb{N}} \mu_n(A)$.

Ces opérations nous permettent d'avoir d'autres exemples de mesures.

Exemples.

- a) Probabilité de Bernoulli de paramètre $p \in [0, 1]$ donnée par : $\mu = p\delta_0 + (1 - p)\delta_1$.
- b) Probabilité de Poisson de paramètre $\lambda > 0$: $\mu = e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \delta_n$.

Nous pouvons maintenant énoncer quelques propriétés des mesures.

Proposition 3.5. Soit (X, \mathcal{M}, μ) un espace mesuré. Nous avons les propriétés suivantes :

- a) Si $A \subset B$ et A, B mesurables, alors $\mu(A) \leq \mu(B)$.
- b) Si A et B sont deux ensembles mesurables, alors $\mu(A \cup B) \leq \mu(A) + \mu(B)$.
- c) Si $(A_n)_n$ est une suite croissante d'ensembles mesurables, $A_n \subset A_{n+1}$, et $A = \bigcup_{n \in \mathbb{N}} A_n$ alors $\mu(A) = \lim_{n \rightarrow \infty} \mu(A_n)$ dans $\overline{\mathbb{R}}_+$.
- d) Si $(A_n)_n$ est une suite décroissante d'ensembles mesurables, $A_n \supset A_{n+1}$, $A = \bigcap_{n \in \mathbb{N}} A_n$ et en plus A_0 est de mesure finie, alors $\mu(A) = \lim_{n \rightarrow \infty} \mu(A_n)$.
- e) Si $(A_n)_n$ est une suite d'ensembles mesurables, alors $\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leq \sum_{n \in \mathbb{N}} \mu(A_n)$.
- f) La propriété du c) plus $\mu(\emptyset) = 0$ plus $\mu(A \cup B) \leq \mu(A) + \mu(B)$ pour tous A, B mesurables disjoints, forment un ensemble de propriétés équivalent à la définition de la mesure.

Fin du cours 3 (17/09/2024).

4 Intégration des fonctions positives

Nous allons construire l'intégrale d'une fonction positive en plusieurs étapes. D'abord pour les fonctions étagées.

4.1 Intégration des fonctions étagées positives

Définition 4.1 (intégrale d'une fonction étagée). Soit (X, \mathcal{M}, μ) un espace mesuré et f étagée positive. Si l'écriture canonique de f est donnée par $f = \alpha_1\chi_{A_1} + \dots + \alpha_n\chi_{A_n}$ (où $f(X) = \{\alpha_1, \dots, \alpha_n\}$ et $A_j = f^{-1}(\alpha_j)$), on définit l'intégrale de f par

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

Si $A \subset X$ est mesurable, on définit l'intégrale de f sur A par

$$\int_A f \, d\mu = \int_X f \chi_A \, d\mu.$$

Rappelons que nous utilisons la convention $0 \cdot \infty = 0$ ce qui fait que $\int_X 0 \, d\mu = 0$.

Exemple. Si A est mesurable nous avons que $\int_X \chi_A \, d\mu = \mu(A)$.

Proposition 4.2. Soit (X, \mathcal{M}, μ) un espace mesuré, f, g deux fonctions étagées positives et $\lambda \in \mathbb{R}_+$. Nous avons :

- a) $\int_X \lambda f \, d\mu = \lambda \int_X f \, d\mu$;
- b) $\int_X (f + g) \, d\mu = \int_X f \, d\mu + \int_X g \, d\mu$;
- c) $\int_X f \, d\mu = \sup_{\substack{h \text{ étagée} \\ 0 \leq h \leq f}} \int_X h \, d\mu$;
- d) L'application $\mathcal{M} \ni A \mapsto \lambda(A) = \int_A f \, d\mu \in \overline{\mathbb{R}}_+$ est une mesure sur (X, \mathcal{M}) . La mesure λ est dite mesure de densité f par rapport à μ .

Remarque. La partie b) au dessus implique la propriété de monotonie suivante de l'intégrale : si f, g étagées positives vérifient $f \leq g$ alors $\int_X f \, d\mu \leq \int_X g \, d\mu$.

4.2 Intégration des fonctions mesurables positives

Définition 4.3 (intégrale d'une fonction positive). Soit (X, \mathcal{M}, μ) un espace mesuré et $f : X \rightarrow \overline{\mathbb{R}}_+$ mesurable. On définit l'intégrale de f par

$$\int_X f \, d\mu = \sup_{\substack{h \text{ étagée} \\ 0 \leq h \leq f}} \int_X h \, d\mu.$$

Si $A \subset X$ est mesurable, on définit l'intégrale de f sur A par

$$\int_A f \, d\mu = \int_X f \chi_A \, d\mu.$$

Remarque. Lorsque f est étagée positive, on peut calculer son intégrale soit en la regardant comme fonction étagée avec la définition 4.1, soit comme fonction mesurable avec la définition 4.3 au-dessus. La question c) de la proposition 4.2 nous assure que le résultat est le même, il n'y a donc pas de risque de confusion.

Proposition 4.4. Soit (X, \mathcal{M}, μ) un espace mesuré, $f, g : X \rightarrow \overline{\mathbb{R}}_+$ deux fonctions mesurables. Nous avons :

- a) Si $f \leq g$ alors $\int_X f \, d\mu \leq \int_X g \, d\mu$.
- b) Si $A \subset B$ sont mesurables, alors $\int_A f \, d\mu \leq \int_B f \, d\mu$.
- c) Pour tout $\lambda \geq 0$, $\int_X \lambda f \, d\mu = \lambda \int_X f \, d\mu$.
- d) Si A est mesurable et $\mu(A) = 0$ alors $\int_A f \, d\mu = 0$ (même si $f = +\infty$).

Exemples.

- a) Soit $X = \{x_1, x_2, \dots, x_n\}$, $\mathcal{M} = \mathcal{P}(X)$ et $\mu = \text{card}$. Si $f : X \rightarrow \overline{\mathbb{R}}_+$ nous avons

$$\int_X f \, d\mu = f(x_1) + f(x_2) + \cdots + f(x_n).$$

- b) Soit $X = \{x_1, x_2, \dots, x_n\}$, $\mathcal{M} = \mathcal{P}(X)$ et $\mu = \text{card} / \text{card}(X)$. Si $f : X \rightarrow \overline{\mathbb{R}}_+$ nous avons

$$\int_X f \, d\mu = \frac{f(x_1) + f(x_2) + \cdots + f(x_n)}{n}.$$

- c) Soit $X = \{x_1, x_2, \dots, x_n\}$, $\mathcal{M} = \mathcal{P}(X)$, $g : X \rightarrow \overline{\mathbb{R}}_+$ et μ la mesure de densité g par rapport à la mesure de comptage (card). Si $f : X \rightarrow \overline{\mathbb{R}}_+$ nous avons

$$\int_X f \, d\mu = g(x_1)f(x_1) + g(x_2)f(x_2) + \cdots + g(x_n)f(x_n).$$

- d) Soit $X = (x_i)_{i \in I}$ un ensemble arbitraire, $\mathcal{M} = \mathcal{P}(X)$ et $\mu = \text{card}$. Si $f : X \rightarrow \overline{\mathbb{R}}_+$ nous avons

$$\int_X f \, d\mu = \sup_{J \text{ fini} \subset I} \sum_{j \in J} f(x_j).$$

- e) Soit \mathbb{N} muni de $\mathcal{M} = \mathcal{P}(\mathbb{N})$ et de la mesure de comptage $\mu = \text{card}$. Si $f : X \rightarrow \overline{\mathbb{R}}_+$ nous avons

$$\int_{\mathbb{N}} f \, d\mu = \sum_{n \in \mathbb{N}} f(n).$$

- f) Pour la mesure de Borel sur \mathbb{R} , on note l'intégrale avec dx comme pour l'intégrale de Riemann : $\int_{\mathbb{R}} f(x) \, dx$. Nous verrons plus tard que cela est justifié car, au moins pour les fonctions continues, l'intégrale de Riemann et l'intégrale de Lebesgue par rapport à la mesure de Borel coïncident.

5 Théorèmes de convergence

Nous allons montrer dans ce gros chapitre plusieurs théorème de convergence des intégrales et leurs conséquences sur la construction de l'intégrale. Nous allons aussi montrer des critères de convergence ou de dérivabilité pour les intégrales à paramètre.

5.1 Théorème de convergence monotone et applications. Lemme de Fatou

Le premier résultat de convergence des intégrales est le théorème de convergence monotone connu aussi sous le nom de théorème de Beppo Levi. Il s'agit d'un résultat fondamental.

Théorème 5.1 (convergence monotone, Beppo Levi). *Soit (X, \mathcal{M}, μ) un espace mesuré et $f_n : X \rightarrow \overline{\mathbb{R}}_+$ une suite de fonctions mesurables positives. On suppose que la suite est monotone croissante, $f_n \leq f_{n+1}$, et qu'elle tend simplement vers une certaine fonction $f : X \rightarrow \overline{\mathbb{R}}_+$. Alors f est mesurable et*

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

Voici quelques applications de ce théorème.

Corollaire 5.2. *Soit (X, \mathcal{M}, μ) un espace mesuré et $f, g : X \rightarrow \overline{\mathbb{R}}_+$ mesurables. Nous avons que $f + g$ est mesurable et*

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

Corollaire 5.3. *Soit (X, \mathcal{M}, μ) un espace mesuré et $f_n : X \rightarrow \overline{\mathbb{R}}_+$ une suite de fonctions mesurables positives. On pose $f(x) = \sum_{n \in \mathbb{N}} f_n(x)$. Alors f est mesurable et*

$$\int_X \left(\sum_{n \in \mathbb{N}} f_n \right) \, d\mu = \sum_{n \in \mathbb{N}} \int_X f_n \, d\mu.$$

Nous avons aussi un résultat de “convergence” des intégrales sans aucune hypothèse sur les intégrandes ! C'est le lemme de Fatou suivant.

Proposition 5.4 (Lemme de Fatou). *Soit (X, \mathcal{M}, μ) un espace mesuré et $f_n : X \rightarrow \overline{\mathbb{R}}_+$ une suite de fonctions mesurables positives. Nous avons l'inégalité suivante :*

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

Exemple. Le théorème de convergence monotone nous permet de donner la formule de l'intégrale pour la mesure image. Plus précisément, soit (X, \mathcal{M}, μ) un espace mesuré et $f : X \rightarrow Y$ une application. Soit $\mathcal{N} = \{B \subset Y ; f^{-1}(B) \in \mathcal{M}\}$ la tribu image sur Y et ν la mesure image sur \mathcal{N} (rappel : $\nu(B) = \mu(f^{-1}(B))$). Si $g : Y \rightarrow \overline{\mathbb{R}}_+$ est mesurable alors $g \circ f$ est mesurable aussi et

$$\int_Y g \, d\nu = \int_X g \circ f \, d\mu.$$

Fin du cours 4 (24/09/2024).

On peut maintenant parler de mesure à densité dans le cadre général des fonctions positives.

Proposition 5.5 (mesures à densité). Soit (X, \mathcal{M}, μ) un espace mesuré et $f : X \rightarrow \overline{\mathbb{R}}_+$ une fonction mesurable positive. L'application $\mathcal{M} \ni A \mapsto \lambda(A) = \int_A f \, d\mu \in \overline{\mathbb{R}}_+$ est une mesure sur (X, \mathcal{M}) . La mesure λ est dite mesure de densité f par rapport à μ . Si $g : X \rightarrow \overline{\mathbb{R}}_+$ est une fonction mesurable positive, nous avons que

$$\int_X g \, d\lambda = \int_X gf \, d\mu$$

ce qui justifie la notation $d\lambda = f d\mu$.

5.2 Intégration des fonctions mesurables complexes

Pour intégrer des fonctions complexes, on se ramène à des fonctions réelles en prenant la partie réelle et la partie imaginaire, puis à des fonctions positives en prenant la partie positive et la partie négative. Rappelons que pour $a \in \mathbb{R}$ on pose

$$a_+ = \max(a, 0), \quad a_- = \max(-a, 0).$$

Nous avons

$$a = a_+ - a_- \quad \text{et} \quad |a| = a_+ + a_-.$$

Pour une fonction f à valeurs réelles, on définit ainsi ses parties positives et négatives par

$$f_+(x) = (f(x))_+ \quad \text{et} \quad f_-(x) = (f(x))_-.$$

Pour une fonction f à valeurs complexes, on prend d'abord la partie réelle et la partie imaginaire, puis la partie positive et la partie négative :

$$f = \operatorname{Re}(f) + i \operatorname{Im}(f) = (\operatorname{Re}(f))_+ - (\operatorname{Re}(f))_- + i(\operatorname{Im}(f))_+ - i(\operatorname{Im}(f))_-.$$

On définit maintenant l'intégrale d'une fonction définie sur \mathbb{C} .

Définition 5.6 (intégrale d'une fonction complexe, espace $\mathcal{L}^1(X, \mu)$). Soit (X, \mathcal{M}, μ) un espace mesuré et $f : X \rightarrow \mathbb{C}$ une fonction mesurable. On dit que f est intégrable si $\int_X |f| \, d\mu < \infty$. L'espace des fonctions intégrables est noté par $\mathcal{L}^1(X, \mu)$. On définit l'intégrale de f par la formule

$$\int_X f \, d\mu = \int_X (\operatorname{Re}(f))_+ \, d\mu - \int_X (\operatorname{Re}(f))_- \, d\mu + i \int_X (\operatorname{Im}(f))_+ \, d\mu - i \int_X (\operatorname{Im}(f))_- \, d\mu.$$

Remarque. La définition de l'intégrale au-dessus est justifiée. En effet, nous avons que $(\operatorname{Re}(f))_+ \leq |\operatorname{Re}(f)| \leq |f|$ d'où $\int_X (\operatorname{Re}(f))_+ \, d\mu \leq \int_X |f| \, d\mu < \infty$. De même pour les trois autres termes.

Proposition 5.7. Soit (X, \mathcal{M}, μ) un espace mesuré. Nous avons que $\mathcal{L}^1(X, \mu)$ est un espace vectoriel et l'application $\mathcal{L}^1(X, \mu) \ni f \mapsto \int_X f \, d\mu$ est linéaire. De plus, pour tout $f \in \mathcal{L}^1(X, \mu)$ nous avons l'inégalité

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

5.3 Théorème de convergence dominée de Lebesgue

Le théorème de cette partie est peut-être le théorème le plus utile de ce cours. Il permet de passer à la limite dans des intégrales lorsque les intégrandes convergent simplement.

Théorème 5.8 (convergence dominée de Lebesgue). Soit (X, \mathcal{M}, μ) un espace mesuré et $f_n : X \rightarrow \mathbb{C}$ une suite d'applications mesurables. On suppose que

- a) (convergence simple) $f_n \rightarrow f$ simplement pour une certaine fonction $f : X \rightarrow \mathbb{C}$;
- b) (domination) il existe $g : X \rightarrow \overline{\mathbb{R}}_+$ telle que $\int_X g \, d\mu < \infty$ et $|f_n| \leq g$ pour tout n .

Alors f_n et f sont intégrables et

$$\lim_{n \rightarrow \infty} \int_X |f_n - f| \, d\mu = 0.$$

En particulier, nous avons que

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

L'hypothèse de convergence simple est une hypothèse “de bon sens”, c'est-à-dire une hypothèse sans laquelle la convergence dans la conclusion ne peut raisonnablement avoir lieu. À première vue, ce n'est pas clair pourquoi l'hypothèse de domination devrait apparaître. En fait, ce théorème admet aussi une réciproque à une sous suite près. Toute suite qui vérifie la conclusion admet une sous-suite qui vérifie les hypothèses a) et b) (cela sera vu au 2e semestre dans le cours d'éléments d'analyse fonctionnelle). Le théorème de convergence dominée de Lebesgue est donc optimal.

5.4 Presque partout.

Définition 5.9 (presque partout). Soit (X, \mathcal{M}, μ) un espace mesuré et $f : X \rightarrow \mathbb{C}$ ou $f : X \rightarrow \overline{\mathbb{R}}$ mesurable.

- On dit que f est nulle μ presque partout, ou $f = 0$ μ p.p. (ou juste $f = 0$ p.p. s'il n'y a pas de risque de confusion quant au choix de la mesure μ) si $\mu(\{x ; f(x) \neq 0\}) = 0$.
- Plus généralement, on dit qu'une propriété a lieu μ presque partout, ou μ p.p., si elle a lieu partout sauf sur un ensemble de mesure μ nulle.

Cette définition est justifiée par les propriétés suivantes de l'intégrale.

Proposition 5.10. Soit (X, \mathcal{M}, μ) un espace mesuré.

- a) Si $f : X \rightarrow \overline{\mathbb{R}}_+$ est mesurable, nous avons l'équivalence

$$\int_X f \, d\mu = 0 \iff f = 0 \text{ } \mu \text{ p.p.}$$

- b) Si $g \in \mathcal{L}^1(X, \mu)$, nous avons l'équivalence

$$\int_X |g| \, d\mu = 0 \iff g = 0 \text{ } \mu \text{ p.p.}$$

- c) Soient $f, g : X \rightarrow \overline{\mathbb{R}}_+$ mesurables ou $f, g \in \mathcal{L}^1(X, \mu)$ réelles. Si $f \leq g$ μ p.p. alors $\int_X f \, d\mu \leq \int_X g \, d\mu$.

- d) Soient $f, g : X \rightarrow \overline{\mathbb{R}}_+$ mesurables ou $f, g \in \mathcal{L}^1(X, \mu)$. Si $f = g$ μ p.p. alors $\int_X f \, d\mu = \int_X g \, d\mu$.

- e) Si $f : X \rightarrow \overline{\mathbb{R}}$ est telle que $\int_X |f| \, d\mu < \infty$ alors f est finie μ p.p..

On pourrait définir la notion d'intégrabilité des fonctions à valeurs dans $\overline{\mathbb{R}}$ en séparant la partie positive et la partie négative et en raisonnant comme dans la définition 5.6. Le résultat au-dessus montre qu'une telle définition ne serait pas très utile en pratique. En effet, une fonction f intégrable à valeurs dans $\overline{\mathbb{R}}$ est nécessairement finie p.p. On pourrait alors définir une nouvelle fonction \tilde{f} qui serait nulle là où $f = \pm\infty$ et égale à f ailleurs. On a alors que $f = \tilde{f}$ p.p. et \tilde{f} est finie partout. Le passage de f à \tilde{f} ne modifie ni la valeur de l'intégrale ni les propriétés de mesurabilité. Du point de vue de l'intégrale, les deux fonctions sont indifférenciables et on peut parfaitement travailler avec \tilde{f} au lieu de f pour les besoins de l'intégration. De même, si on a une suite f_n de fonctions intégrables à valeurs dans $\overline{\mathbb{R}}$, on posant A_n l'ensemble où $f_n = \pm\infty$ on peut rendre les f_n nulles sur $\cup_{n \in \mathbb{N}} A_n$ (qui est un ensemble de mesure nulle) et obtenir ainsi une nouvelle suite de fonctions intégrables finies, égales p.p. à la suite de départ. Attention cependant au cas particulier des fonctions positives, qui peuvent être intégrées sans que la fonction soit finie p.p. On retiendra de cette discussion que si on veut intégrer une fonction, soit elle est positive soit on peut la supposer finie (en la modifiant éventuellement sur un ensemble de mesure nulle).

Une dernière remarque sur les fonctions à valeurs dans $\overline{\mathbb{R}}$. Pour f à valeurs dans $\overline{\mathbb{R}}$, en décomposant $f = f_+ - f_-$ on pourrait imaginer une situation où par exemple f_- est d'intégrale finie et f_+ non. Dans ce cas, on peut définir l'intégrale de f par $\int_X f \, d\mu = \int_X f_+ \, d\mu - \int_X f_- \, d\mu$ qui a parfaitement un sens et est égale à $+\infty$. Dans un tel cas, f n'est pas intégrable au sens de la théorie que nous avons introduite et les théorèmes de ce cours ne s'y appliquent pas. En général, on évite de travailler avec de telles fonctions.

Les théorèmes de convergence des intégrales ont des versions "p.p.". Voici la version "p.p." du théorème de convergence dominée de Lebesgue.

Théorème 5.11 (convergence dominée de Lebesgue, version p.p.). Soit (X, \mathcal{M}, μ) un espace mesuré et $f_n : X \rightarrow \mathbb{C}$ une suite d'applications mesurables. On suppose que

- a) (convergence simple) $f_n \rightarrow f$ μ p.p. au sens suivant : il existe $f : X \rightarrow \mathbb{C}$ et A ensemble mesurable de mesure nulle tel que $f_n \rightarrow f$ simplement sur A^c et $f = 0$ sur A ;
- b) (domination) il existe $g : X \rightarrow \overline{\mathbb{R}}_+$ telle que $\int_X g \, d\mu < \infty$ et $|f_n| \leq g$ μ p.p. pour tout n .

Alors f_n et f sont intégrables et

$$\lim_{n \rightarrow \infty} \int_X |f_n - f| \, d\mu = 0.$$

En particulier, nous avons que

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

La formulation "p.p." du théorème de convergence monotone de Beppo Levi est similaire : les hypothèses $0 \leq f_n \leq f_{n+1}$ μ p.p. et $f_n \rightarrow f$ μ p.p. entraînent la même conclusion. Les détails sont laissés en exercice.

5.5 Tribu complétée. Mesure complétée. Tribu et mesure de Lebesgue

Nous montrons dans ce paragraphe qu'une mesure peut être "complétée" en ajoutant tous les sous-ensembles des ensembles de mesure nulle. C'est une notion qui est pratique, mais non essentielle dans la théorie de l'intégration.

Commençons par définir les ensembles négligeables.

Définition 5.12 (ensemble négligeable). Soit (X, \mathcal{M}, μ) un espace mesuré. Un sous-ensemble A de X est dit négligeable s'il est inclus dans un ensemble de mesure nulle de la tribu : il existe $B \in \mathcal{M}$ tel que $\mu(B) = 0$ et $A \subset B$.

On définit maintenant la tribu complétée.

Définition 5.13 (tribu complétée). Soit (X, \mathcal{M}, μ) un espace mesuré.

- La tribu complétée par rapport à la mesure μ , notée par $\overline{\mathcal{M}}$ est la tribu engendrée par \mathcal{M} et par tous les ensembles négligeables.
- La tribu est dite complète par rapport à la mesure μ si sa complétée est elle-même : $\mathcal{M} = \overline{\mathcal{M}}$. Ou, de manière équivalente, si tous les ensembles négligeables appartiennent à la tribu.

De toute évidence, les notions de tribu complétée et tribu complète ne dépendent pas uniquement de la tribu mais aussi de la mesure considérée. De ce fait, la notation $\overline{\mathcal{M}}$ est abusive car laisse entendre que cela dépend uniquement de \mathcal{M} . En pratique, le choix de la mesure μ est évident et le plus souvent n'a pas besoin d'être spécifié.

On peut facilement caractériser la tribu complétée.

Proposition 5.14. Soit (X, \mathcal{M}, μ) un espace mesuré. La tribu complétée $\overline{\mathcal{M}}$ est formé de toutes les unions possibles d'ensembles de E et d'ensembles négligeables :

$$\overline{\mathcal{M}} = \{E \cup A ; E \in \mathcal{M} \text{ et } A \text{ négligeable}\}.$$

La mesure μ admet une unique extension à $\overline{\mathcal{M}}$.

Proposition 5.15. Soit (X, \mathcal{M}, μ) un espace mesuré. Il existe une unique mesure $\overline{\mu}$ sur $\overline{\mathcal{M}}$ telle que $\overline{\mu}|_{\mathcal{M}} = \mu$.

Définition 5.16 (mesure complétée). Soit (X, \mathcal{M}, μ) un espace mesuré. La mesure $\overline{\mu}$ de la proposition précédente est dite la mesure complétée de μ .

Définition 5.17 (tribu et mesure de Lebesgue). — La tribu de Lebesgue sur \mathbb{R}^n , notée par

\mathcal{L}_n , est la tribu de Borel complétée par rapport à la mesure de Borel.

— La mesure de Lebesgue sur \mathbb{R}^n , notée par λ_n , est la mesure de Borel complétée.

Fin du cours 5 (01/10/2024).

En pratique, la mesure de Borel est quasiment la même chose que la mesure de Lebesgue. Elles diffèrent uniquement par des ensembles négligeables, les intégrales associées sont donc les mêmes. Pour différencier la mesurabilité par rapport à la tribu de Lebesgue de celle par rapport à la tribu de Borel, on utilise la terminologie de fonction Lebesgue mesurable pour les fonctions mesurables par rapport à la tribu de Lebesgue et la terminologie de fonction borélienne pour les fonctions mesurables par rapport à la tribu de Borel.

La proposition suivante dit qu'une fonction est mesurable par rapport à la tribu complétée $\overline{\mathcal{M}}$ si et seulement si elle est égale p.p. à une fonction mesurable pour la tribu \mathcal{M} .

Proposition 5.18. Soit (X, \mathcal{M}, μ) un espace mesuré et $f : X \rightarrow \overline{\mathbb{R}}$ (ou $f : X \rightarrow \mathbb{C}$). La fonction f est mesurable par rapport à la tribu complétée $\overline{\mathcal{M}}$ si et seulement si il existe $g : X \rightarrow \overline{\mathbb{R}}$ (ou $g : X \rightarrow \mathbb{C}$) mesurable par rapport à la tribu \mathcal{M} telle que $f = g$ $\overline{\mu}$ p.p.

Proposition 5.19. Si μ est une mesure complète (c'est-à-dire égale à sa mesure complétée), alors toute limite p.p. de fonctions mesurables est mesurable.

Ce résultat nous permet de simplifier la condition a) de la version p.p. du théorème de convergence dominée de Lebesgue, théorème 5.11. On peut simplement se placer dans le cadre de la tribu et de la mesure complétée et supposer dans la condition a) que les f_n sont $\overline{\mathcal{M}}$ mesurables et convergent $\overline{\mu}$ p.p. La limite p.p. sera automatiquement mesurable pour la tribu complétée, peu importe la manière dont la fonction limite est définie sur l'ensemble où la convergence simple n'a pas lieu. On peut également reformuler cette même condition a) dans le cadre de la tribu de départ (non complétée) en supposant que les f_n convergent μ p.p. vers une fonction \mathcal{M} mesurable (c'est-à-dire on rajoute comme hypothèse que la limite f est mesurable).

5.6 Intégrales à paramètre

Nous nous intéressons maintenant aux intégrales à paramètre, c'est-à-dire les intégrales où l'intégrande dépend d'un paramètre. Comment dépend l'intégrale de ce paramètre, est-elle continue, dérivable, etc. ?

Le cadre est le suivant. On se donne une fonction $f : X \times \Lambda \rightarrow \mathbb{C}$ où (X, \mathcal{M}, μ) un espace mesuré et Λ un espace métrique. On écrit $f(x, \lambda)$ où x est la variable d'intégration et λ le paramètre. On définit la fonction

$$F(\lambda) = \int_X f(x, \lambda) \, d\mu_x \equiv \int_X f(x, \lambda) \, dx$$

où $d\mu_x$ veut dire qu'on intègre par rapport à la mesure μ dans la variable x . Pour abréger et clarifier l'écriture, on écrira dans cette partie dx au lieu de $d\mu_x$.

Le théorème suivant porte sur la continuité des intégrales à paramètre.

Théorème 5.20 (continuité des intégrales à paramètre). *Soit (X, \mathcal{M}, μ) un espace mesuré, Λ un espace métrique et $f : X \times \Lambda \rightarrow \mathbb{C}$. On suppose que*

- a) pour tout $\lambda \in \Lambda$, l'application $X \ni x \mapsto f(x, \lambda) \in \mathbb{C}$ est mesurable ;
- b) Pour presque tout $x \in X$, l'application $\Lambda \ni \lambda \mapsto f(x, \lambda) \in \mathbb{C}$ est continue ;
- c) il existe $g \in \mathscr{L}^1(X, \mu)$ positive telle que

$$\forall \lambda \in \Lambda \quad |f(x, \lambda)| \leq g(x) \text{ pour presque tout } x (\mu \text{ p.p.}).$$

Alors l'application

$$\lambda \mapsto F(\lambda) = \int_X f(x, \lambda) \, dx$$

est continue de Λ dans \mathbb{C} .

Remarques.

- Si on cherche à avoir la continuité de F en un seul point λ_0 , alors il suffit de supposer dans l'hypothèse b) la continuité en λ_0 seulement.
- Étant donné que la continuité est une notion locale (i.e. dépend uniquement des valeurs de la fonction dans un voisinage), la condition du c) peut être supposée sur des boules de rayon aussi petit qu'on veut. Si Λ est un ouvert de \mathbb{R}^n alors la condition du c) peut être supposée sur les compacts de Λ seulement.

On s'intéresse maintenant à la dérivabilité des intégrales à paramètre. On doit donc supposer que Λ est un intervalle de \mathbb{R} .

Théorème 5.21 (dérivabilité des intégrales à paramètre). *Soit (X, \mathcal{M}, μ) un espace mesuré, Λ un intervalle de \mathbb{R} et $f : X \times \Lambda \rightarrow \mathbb{C}$. On suppose que*

- a) pour tout $\lambda \in \Lambda$, l'application $X \ni x \mapsto f(x, \lambda) \in \mathbb{C}$ est intégrable (appartient à $\mathscr{L}^1(X, \mu)$) ;
- b) Pour presque tout $x \in X$, l'application $\Lambda \ni \lambda \mapsto f(x, \lambda) \in \mathbb{C}$ est dérivable ;
- c) il existe $g \in \mathscr{L}^1(X, \mu)$ positive telle que pour presque tout x (μ p.p.) nous avons la majoration

$$\left| \frac{\partial f}{\partial \lambda}(x, \lambda) \right| \leq g(x) \quad \forall \lambda \in \Lambda.$$

Alors l'application

$$\lambda \mapsto F(\lambda) = \int_X f(x, \lambda) \, dx$$

est dérivable sur Λ et on peut dériver sous l'intégrale :

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

L'intégrande au-dessus $\frac{\partial f}{\partial \lambda}(x, \lambda)$ est définie μ p.p. seulement. Sur l'ensemble de mesure nulle où cette intégrande n'est pas définie, on la pose nulle.

Remarques.

- Dans le cas où λ est une extrémité de Λ , le résultat reste vrai en sous-entendant qu'on prend la dérivée à gauche ou à droite.
- Contrairement à la continuité, si on veut avoir la dérivabilité en un seul point λ_0 , il ne suffit pas de supposer dans l'hypothèse b) la dérivabilité en λ_0 seulement.
- Par contre, la dérivabilité étant aussi une notion locale, la condition du c) peut être supposée sur des boules de rayon aussi petit qu'on veut ou sur les compacts de Λ seulement.

5.7 Comparaison entre l'intégrale de Riemann et l'intégrale de Lebesgue

Nous allons maintenant comparer l'intégrale de Lebesgue à l'intégrale de Riemann. D'abord sur un intervalle compact.

Proposition 5.22. *Soit $f : [a, b] \rightarrow \mathbb{R}$ intégrable Riemann. Alors f est Lebesgue intégrable (Lebesgue mesurable et intégrable) et*

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

intégrale Riemann intégrale Lebesgue

Nous continuerons donc à noter par $\int_a^b f(x) dx$ à la fois l'intégrale de Lebesgue et l'intégrale de Riemann (si celle-ci existe).

Remarques.

- a) La réciproque de cette proposition est fausse. En effet, une fonction Lebesgue intégrable peut très bien ne pas être bornée tandis qu'une fonction intégrable Riemann est bornée par définition. On peut aussi construire des exemples de fonctions bornées Lebesgue intégrables mais non Riemann intégrables. La fonction indicatrice de \mathbb{Q} n'est pas intégrable Riemann sur $[a, b]$ mais elle intégrable Lebesgue car nulle presque partout.
- b) On peut montrer qu'une fonction bornée est Riemann intégrable sur un intervalle fermé borné si et seulement si elle est continue en dehors d'un ensemble négligeable (pour la mesure de Lebesgue).

Dans le cas des intégrales généralisées, la situation est un peu différente.

Proposition 5.23. *Soit I un intervalle non-compact et $f : I \rightarrow \mathbb{R}$ continue.*

- a) *Si $f \geq 0$, alors f est intégrable Lebesgue si et seulement si l'intégrale de Riemann généralisée de f est convergente. De plus, en cas d'intégrabilité nous avons l'égalité des intégrales :*

$$\int_I f(x) dx = \int_I f d\lambda_1.$$

intégrale généralisée intégrale Lebesgue

- b) *En général, nous avons que f est intégrable Lebesgue si et seulement si l'intégrale de Riemann généralisée de f est absolument convergente (c'est-à-dire si l'intégrale de $|f|$ est convergente). Si f est intégrable Lebesgue, alors on a l'égalité des intégrales*

$$\int_I f(x) dx = \int_I f d\lambda_1.$$

intégrale généralisée intégrale Lebesgue

Remarque. Il peut arriver qu'une intégrale généralisée existe sans que la fonction soit Lebesgue intégrable. Par exemple, la fonction $\frac{\sin x}{x}$ admet une intégrale généralisée convergente sur $[0, \infty[$ mais elle n'est pas Lebesgue intégrable car l'intégrale généralisée n'est pas absolument convergente. On appelle ça une intégrale semi-convergente. Les intégrales semi-convergentes ne rentrent donc pas dans le cadre de la théorie de Lebesgue.

6 Intégration sur un espace produit

6.1 Tribu produit

Définition 6.1 (rectangle, tribu produit). Soient (X_1, \mathcal{M}_1) et (X_2, \mathcal{M}_2) deux espaces mesurables.

- Un rectangle est un ensemble de la forme $A_1 \times A_2$, avec $A_1 \in \mathcal{M}_1$ et $A_2 \in \mathcal{M}_2$;
- La tribu produit $\mathcal{M}_1 \otimes \mathcal{M}_2$ sur $X_1 \times X_2$ est la tribu engendré par tous les rectangles.

Dans la suite, la tribu considéré par défaut sur le produit cartésien $X_1 \times X_2$ sera la tribu produit.

Proposition 6.2. La tribu produit $\mathcal{M}_1 \otimes \mathcal{M}_2$ est la plus petite tribu qui rend les projections π_1 et π_2 mesurables. On rappelle que $\pi_1 : X_1 \times X_2 \rightarrow X_1$, $\pi_1(x_1, x_2) = x_1$ et $\pi_2 : X_1 \times X_2 \rightarrow X_2$, $\pi_2(x_1, x_2) = x_2$.

Remarque. Si $f_1 : X_1 \rightarrow \mathbb{C}$ et $f_2 : X_2 \rightarrow \mathbb{C}$ sont mesurables, alors le produit tensoriel $f_1 \otimes f_2$ est mesurable. Ici, $f_1 \otimes f_2 : X_1 \times X_2 \rightarrow \mathbb{C}$ est défini par $f_1 \otimes f_2 = f_1(x_1) \times f_2(x_2)$. La mesurabilité du produit tensoriel résulte de la mesurabilité des projections et de l'écriture $f_1 \otimes f_2 = (f_1 \circ \pi_1)(f_2 \circ \pi_2)$.

Proposition 6.3. Soient (X_1, \mathcal{M}_1) , (X_2, \mathcal{M}_2) , (Y, \mathcal{N}) trois espaces mesurables et $f : X_1 \times X_2 \rightarrow Y$ mesurable.

- a) Pour $x_1 \in X_1$ on définit la fonction partielle $f(x_1, \cdot) : X_2 \rightarrow Y$, $x_2 \mapsto f(x_1, x_2)$. On définit de même pour $x_2 \in X_2$ la fonction partielle $f(\cdot, x_2) : X_1 \rightarrow Y$. Nous avons que pour tout $x_1 \in X_1$, la fonction partielle $f(x_1, \cdot)$ est mesurable et pour tout $x_2 \in X_2$ la fonction partielle $f(\cdot, x_2)$ est mesurable.
- b) Si $A \subset X_1 \times X_2$, on définit les coupes de A de la manière suivante. Pour $x_1 \in X_1$ on pose $A(x_1, \cdot) = \{x_2 \in X_2 ; (x_1, x_2) \in A\}$ et pour $x_2 \in X_2$ on pose $A(\cdot, x_2) = \{x_1 \in X_1 ; (x_1, x_2) \in A\}$. Si A est mesurable dans $X_1 \times X_2$ (i.e. $A \in \mathcal{M}_1 \otimes \mathcal{M}_2$), alors toutes ses coupes sont mesurables : pour tout $x_1 \in X_1$ on a $A(x_1, \cdot) \in \mathcal{M}_2$ et pour tout $x_2 \in X_2$ on a $A(\cdot, x_2) \in \mathcal{M}_1$.

Dans le cas de la tribu de Borel, nous avons que la tribu de Borel de l'espace produit est le produit des tribus de Borel.

Proposition 6.4. Nous avons que $\mathcal{B}(\mathbb{R}^{d_1+d_2}) = \mathcal{B}(\mathbb{R}^{d_1}) \otimes \mathcal{B}(\mathbb{R}^{d_2})$.

6.2 Produits tensoriels de mesures

La construction du produit de mesures est considérablement plus compliquée que celle de la tribu produit. Commençons par définir les classes monotones.

Définition 6.5 (classe monotone). Une famille \mathcal{D} d'ensembles est dite classe monotone si elle est stable par limite monotone :

- Si (A_n) est une suite croissante de \mathcal{D} , $A_n \subset A_{n+1}$ pour tout n , alors $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{D}$.
- Si (B_n) est une suite décroissante de \mathcal{D} , $B_n \supset B_{n+1}$ pour tout n , alors $\bigcap_{n \in \mathbb{N}} B_n \in \mathcal{D}$.

Nous avons la proposition suivante sur les unions finies de rectangles.

Proposition 6.6. Soient (X_1, \mathcal{M}_1) et (X_2, \mathcal{M}_2) deux espaces mesurables.

- a) Toute union finie de rectangles peut s'écrire comme union finie de rectangles disjoints deux à deux.
- b) L'ensemble des unions finies de rectangles est stable par intersection, réunion et différence.

Fin du cours 6 (15/10/2024).

Lemme 6.7. Soient (X_1, \mathcal{M}_1) et (X_2, \mathcal{M}_2) deux espaces mesurables. La tribu produit est la plus petite classe monotone qui contient les unions finies de rectangles.

Définition 6.8. Un espace mesuré (X, \mathcal{M}, μ) est dit σ -fini s'il existe une suite d'ensembles X_n de mesure finie et d'union X : $\mu(X_n) < \infty$ pour tout $n \in \mathbb{N}$ et $\cup_{n \in \mathbb{N}} X_n = X$.

Proposition 6.9. Soient $(X_1, \mathcal{M}_1, \mu_1)$ et $(X_2, \mathcal{M}_2, \mu_2)$ deux espaces mesurés σ -finis. Soit $A \in \mathcal{M}_1 \otimes \mathcal{M}_2$. Nous avons que les applications

$$x_1 \mapsto \int_{X_2} \chi_A(x_1, x_2) \, d\mu_2(x_2) \quad \text{et} \quad x_2 \mapsto \int_{X_1} \chi_A(x_1, x_2) \, d\mu_1(x_1)$$

sont bien définies, mesurables et de même intégrale :

$$\int_{X_1} \left(\int_{X_2} \chi_A(x_1, x_2) \, d\mu_2(x_2) \right) \, d\mu_1(x_1) = \int_{X_2} \left(\int_{X_1} \chi_A(x_1, x_2) \, d\mu_1(x_1) \right) \, d\mu_2(x_2).$$

Cette proposition motive la définition suivante de la mesure produit.

Définition 6.10 (mesure produit). Soient $(X_1, \mathcal{M}_1, \mu_1)$ et $(X_2, \mathcal{M}_2, \mu_2)$ deux espaces mesurés σ -finis. Pour $A \in \mathcal{M}_1 \otimes \mathcal{M}_2$ on pose

$$(\mu_1 \otimes \mu_2)(A) = \int_{X_1} \left(\int_{X_2} \chi_A(x_1, x_2) \, d\mu_2(x_2) \right) \, d\mu_1(x_1) = \int_{X_2} \left(\int_{X_1} \chi_A(x_1, x_2) \, d\mu_1(x_1) \right) \, d\mu_2(x_2).$$

Proposition 6.11. Soient $(X_1, \mathcal{M}_1, \mu_1)$ et $(X_2, \mathcal{M}_2, \mu_2)$ deux espaces mesurés σ -finis. Nous avons que $\mu_1 \otimes \mu_2$ est une mesure σ -finie sur la tribu produit $\mathcal{M}_1 \otimes \mathcal{M}_2$, qu'on appelle mesure produit. De plus

$$(\mu_1 \otimes \mu_2)(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$$

pour tout $A_1 \in \mathcal{M}_1$, $A_2 \in \mathcal{M}_2$.

6.3 Théorèmes de Fubini

Les théorèmes de Fubini nous disent dans quelles conditions on peut permutez deux intégrales.

Le premier théorème nous dit que c'est toujours possible pour les fonctions positives.

Théorème 6.12 (Fubini-Tonelli). Soient $(X_1, \mathcal{M}_1, \mu_1)$ et $(X_2, \mathcal{M}_2, \mu_2)$ deux espaces mesurés σ -finis et $f : X_1 \times X_2 \rightarrow \overline{\mathbb{R}}_+$ mesurable (pour la tribu produit). Nous avons que les applications

$$x_1 \mapsto \int_{X_2} f(x_1, x_2) \, d\mu_2(x_2) \quad \text{et} \quad x_2 \mapsto \int_{X_1} f(x_1, x_2) \, d\mu_1(x_1)$$

sont bien définies, mesurables et de même intégrale. De plus

$$\begin{aligned} \int_{X_1} \left(\int_{X_2} f(x_1, x_2) \, d\mu_2(x_2) \right) \, d\mu_1(x_1) &= \int_{X_2} \left(\int_{X_1} f(x_1, x_2) \, d\mu_1(x_1) \right) \, d\mu_2(x_2) \\ &= \iint_{X_1 \times X_2} f(x_1, x_2) \, d(\mu_1 \otimes \mu_2). \end{aligned}$$

On note le plus souvent

$$\iint_{X_1 \times X_2} f(x_1, x_2) \, d(\mu_1 \otimes \mu_2) = \iint_{X_1 \times X_2} f(x_1, x_2) \, d\mu_1(x_1) \, d\mu_2(x_2) = \iint_{X_1 \times X_2} f(x_1, x_2) \, d\mu_1 \, d\mu_2.$$

Dans le cas des fonctions mesurables sans signe, le même résultat est vrai à condition d'ajouter l'hypothèse que $f \in \mathcal{L}^1(X_1 \times X_2, \mu_1 \otimes \mu_2)$.

Théorème 6.13 (Fubini). Soient $(X_1, \mathcal{M}_1, \mu_1)$ et $(X_2, \mathcal{M}_2, \mu_2)$ deux espaces mesurés σ -finis et $f \in \mathscr{L}^1(X_1 \times X_2, \mu_1 \otimes \mu_2)$. Nous avons que les applications

$$x_1 \mapsto \int_{X_2} f(x_1, x_2) \, d\mu_2(x_2) \quad \text{et} \quad x_2 \mapsto \int_{X_1} f(x_1, x_2) \, d\mu_1(x_1)$$

sont bien définies presque partout, définissent des fonctions de \mathscr{L}^1 (après leur avoir attribué des valeurs nulles sur l'ensemble de mesure nulle où elle ne sont pas définies) et ont la même intégrale. De plus,

$$\begin{aligned} \int_{X_1} \left(\int_{X_2} f(x_1, x_2) \, d\mu_2(x_2) \right) \, d\mu_1(x_1) &= \int_{X_2} \left(\int_{X_1} f(x_1, x_2) \, d\mu_1(x_1) \right) \, d\mu_2(x_2) \\ &= \int_{X_1 \times X_2} f(x_1, x_2) \, d(\mu_1 \otimes \mu_2). \end{aligned}$$

Remarque. Si $f_1 \in \mathscr{L}^1(X_1, \mu_1)$ et $f_2 \in \mathscr{L}^1(X_2, \mu_2)$ ou si $f_1 : X_1 \rightarrow \overline{\mathbb{R}}_+$ et $f_2 : X_2 \rightarrow \overline{\mathbb{R}}_+$ sont mesurables, alors $f_1 \otimes f_2$ est soit intégrable sur $X_1 \times X_2$ soit positive, et dans les deux cas

$$\int_{X_1 \times X_2} f_1 \otimes f_2 \, d(\mu_1 \otimes \mu_2) = \int_{X_1} f_1 \, d\mu_1 \int_{X_2} f_2 \, d\mu_2.$$

Remarque. En pratique, lorsqu'on veut permuter deux intégrales on procède de la manière suivante :

- soit l'intégrande est positive et à ce moment-là on peut permuter les intégrales sans se poser des questions ;
- soit l'intégrande $f(x_1, x_2)$ n'a pas de signe, et dans ce cas il faut d'abord vérifier que l'intégrale avec valeur absolue sur l'intégrande est finie, peut importe le sens des intégrations :

$$\begin{aligned} \text{soit } \int_{X_1} \left(\int_{X_2} |f(x_1, x_2)| \, dx_2 \right) \, dx_1 &< \infty \\ \text{soit } \int_{X_2} \left(\int_{X_1} |f(x_1, x_2)| \, dx_1 \right) \, dx_2 &< \infty. \end{aligned}$$

Il n'est pas utile de calculer les valeurs des intégrales avec la valeur absolue, il faut seulement les majorer pour montrer qu'elles sont finies. En effet, si l'une des intégrales au-dessus est finie, le théorème de Fubini-Tonelli implique $f \in \mathscr{L}^1(X_1 \times X_2)$ et ensuite on peut appliquer le théorème de Fubini.

Remarque. Nous avons vu dans la proposition 6.4 que le produit des tribus de Borel est la tribu de Borel sur l'espace produit. Cela est aussi vrai pour les mesures de Borel. En effet, si on désigne par μ_d la mesure de Borel dans \mathbb{R}^d et si $P = \prod_{j=1}^{d_1+d_2} I_j$ est un pavé de $\mathbb{R}^{d_1+d_2}$, alors P est aussi un rectangle $P = P_1 \times P_2$ où $P_1 = \prod_{j=1}^{d_1} I_j$ et $P_2 = \prod_{j=d_1+1}^{d_1+d_2} I_j$. Par conséquent

$$\begin{aligned} (\mu_{d_1} \otimes \mu_{d_2})(P) &= (\mu_{d_1} \otimes \mu_{d_2})(P_1 \times P_2) = \mu_1(P_1)\mu_2(P_2) = \prod_{j=1}^{d_1} \text{long}(I_j) \times \prod_{j=d_1+1}^{d_1+d_2} \text{long}(I_j) \\ &= \prod_{j=1}^{d_1+d_2} \text{long}(I_j) = \mu_{d_1+d_2}(P). \end{aligned}$$

Par unicité de la mesure de Borel, il s'ensuit que $\mu_{d_1} \otimes \mu_{d_2} = \mu_{d_1+d_2}$.

Attention, la situation est un peu différente pour la tribu de Lebesgue \mathcal{L}_d . Nous avons que $\mathcal{L}_{d_1} \otimes \mathcal{L}_{d_2} \neq \mathcal{L}_{d_1+d_2}$. Plus précisément, si A_1 est un ensemble de \mathbb{R}^{d_1} qui n'est pas mesurable Lebesgue (de tels ensembles existent), alors pour tout $b \in \mathbb{R}^{d_2}$ nous avons que $A_1 \times \{b\} \in \mathcal{L}_{d_1+d_2} \setminus \mathcal{L}_{d_1} \otimes \mathcal{L}_{d_2}$. En effet, $A_1 \times \{b\}$ est négligeable car inclus dans $\mathbb{R}^{d_1} \times \{b\}$ et $\lambda_{d_1+d_2}(\mathbb{R}^{d_1} \times \{b\}) = \lambda_{d_1}(\mathbb{R}^{d_1})\lambda_{d_2}(\{b\}) = \infty \cdot 0 = 0$, donc $A_1 \times \{b\} \in \mathcal{L}_{d_1+d_2}$. Par contre $A_1 \times \{b\} \notin \mathcal{L}_{d_1} \otimes \mathcal{L}_{d_2}$ car $A_1 \notin \mathcal{L}_{d_1}$.

Fin du cours 7 (05/11/2024).

7 Mesure de Lebesgue

7.1 Existence

Voici le théorème d'existence de la mesure de Lebesgue.

Théorème 7.1 (existence de la mesure de Lebesgue). *Il existe une tribu \mathcal{L}_d sur \mathbb{R}^d stable par translation et contenant les boréliens et une mesure (positive) λ_d sur \mathcal{L}_d telle que*

- λ_d est complète ;
- λ_d est régulière au sens que :

$$\forall E \in \mathcal{L}_d, \quad \lambda_d(E) = \inf_{U \text{ ouvert } \supset E} \lambda_d(U) \quad (\text{régularité extérieure})$$

$$\forall E \in \mathcal{L}_d, \quad \lambda_d(E) = \sup_{K \text{ compact } \subset E} \lambda_d(K) \quad (\text{régularité intérieure})$$

- $\lambda_d\left(\prod_{j=1}^d I_j\right) = \prod_{j=1}^d (b_j - a_j)$ quels que soient I_j intervalles bornés de \mathbb{R} d'extrémités a_j et b_j ;
- λ_d est invariante par translation : $\lambda_d(E + a) = \lambda_d(E)$ pour tout $E \in \mathcal{L}_d$ et $a \in \mathbb{R}^d$.

Remarque. Nous verrons dans le chapitre suivant que λ_d est aussi homogène et invariante par rotation.

Par manque de temps, nous admettrons ce théorème. Mentionnons cependant qu'il y a deux manières classiques de montrer cette existence.

La première fait appel à la notion de mesure extérieure et consiste à poser, pour tout $A \subset \mathbb{R}^d$,

$$\lambda^*(A) = \inf \left\{ \sum_{j \in \mathbb{N}} \text{mes}(P_j) ; A \subset \bigcup_{j \in \mathbb{N}} P_j, P_j \text{ pavé de } \mathbb{R}^d \right\}.$$

Au-dessus, $\text{mes}(P_i)$ désigne la mesure attendue d'un pavé : $\text{mes}\left(\prod_{j=1}^d I_j\right) = \prod_{j=1}^d (b_j - a_j)$ quels que soient I_j intervalles bornés de \mathbb{R} d'extrémités a_j et b_j . Puis on montre que la restriction de λ^* aux boréliens a toutes les propriétés voulues et donne la mesure de Lebesgue.

La deuxième possibilité de construction de la mesure de Lebesgue est de montrer et d'utiliser le théorème de représentation de Riesz suivant qui a son propre intérêt.

Théorème 7.2 (représentation de Riesz). *Soit (X, d) un espace métrique, localement compact (tout point admet un voisinage compact) et séparable. Soit J une forme linéaire et positive sur $C_c^0(X)$. Alors il existe une unique mesure borélienne μ , finie sur les compacts, telle que*

$$J(f) = \int_X f \, d\mu \quad \forall f \in C_c^0(X).$$

Nous avons de plus que :

- μ est régulière ;

— Pour tout ouvert U

$$\mu(U) = \sup\{J(f) ; f \in C_c^0(X; [0, 1]), \text{ supp } f \subset U\};$$

— Pour tout compact K

$$\mu(K) = \inf\{J(f) ; f \in C_c^0(X; [0, 1]), f|_K = 1\}.$$

L'existence de la mesure de Lebesgue suit du théorème de Riesz appliqué à $X = \mathbb{R}^d$ et à la forme J donnée par l'intégrale de Riemann.

7.2 Unicité

Le théorème d'unicité de la mesure de Lebesgue peut s'énoncer de deux manières.

Théorème 7.3 (unicité de la mesure de Lebesgue). a) *Toute mesure mesurable positive μ sur \mathcal{B}_d , invariante par translation et finie sur les compacts est un multiple de la mesure de Lebesgue. En particulier, si $\mu([0, 1]^d) = 1$, alors $\mu = \lambda_d|_{\mathcal{B}_d}$.*

b) *Si μ est une mesure positive sur \mathcal{B}_d telle que $\mu(\prod_{j=1}^d [a_j, b_j]) = \prod_{j=1}^d (b_j - a_j)$ pour tout les a_j et b_j , alors $\mu = \lambda_d|_{\mathcal{B}_d}$.*

Si nous avons manqué de temps pour montrer l'existence de la mesure de Lebesgue, nous pouvons en revanche montrer l'unicité de la mesure de Lebesgue.

Remarquons d'abord que b) implique a). En effet, soit μ mesure positive sur \mathcal{B}_d , invariante par translation et telle que $\mu([0, 1]^d) = 1$. En utilisant l'invariance par translation et en mettant plusieurs cubes identiques l'un à côté de l'autre on montre l'égalité $\mu(\prod_{j=1}^d [a_j, b_j]) = \prod_{j=1}^d (b_j - a_j)$ lorsque les sommets sont de coordonnées entières, puis de coordonnées rationnelles, puis, par densité de \mathbb{Q} dans \mathbb{R} , de coordonnées quelconques.

Nous allons maintenant montrer la partie b) de ce théorème d'unicité. En se rappelant que les pavés engendrent la tribu de Borel et que $\mu = \lambda_d$ sur les pavés, on peut se poser la question si l'égalité de deux mesures sur un ensemble qui engendre la tribu implique l'égalité des mesures.

- La réponse est négative comme on peut le voir sur l'exemple de la mesure de Dirac δ_0 qui s'annule sur tous les intervalles de \mathbb{R}^* sans être la mesure nulle. Or les intervalles de \mathbb{R}^* engendrent la tribu de Borel $\mathcal{B}(\mathbb{R})$.
- La réponse est positive si on ajoute une condition supplémentaire : il faut que la famille qui engendre la tribu soit stable par intersection (finie) et qu'elle contienne une suite croissante d'ensembles de mesure finie qui tend vers l'espace entier.

Plus précisément, nous avons le résultat suivant d'unicité des mesures.

Théorème 7.4 (unicité des mesures). Soit (X, \mathcal{M}) un espace mesurable, une famille d'ensembles $\mathcal{F} \subset \mathcal{M}$ et μ, ν deux mesures positives sur \mathcal{M} . On suppose que :

- a) la famille \mathcal{F} engendre la tribu \mathcal{M} : $\mathcal{M}(\mathcal{F}) = \mathcal{M}$;
- b) la famille \mathcal{F} est stable par intersection (finie);
- c) les mesures μ et ν sont égales sur \mathcal{F} ;
- d) soit $\mu(X) = \nu(X) < \infty$ (cas des mesures finies) soit il existe $X_n \in \mathcal{F}$, $X_n \nearrow X$ (suite croissante d'ensembles d'union X) tels que $\mu(X_n) = \nu(X_n) < \infty$ pour tout $n \in \mathbb{N}$ (cas des mesures σ -finies).

Alors $\mu = \nu$.

Les pavés vérifient les hypothèses de la proposition précédente (cas σ -fini en prenant $X_n = [-n, n]^d$), donc l'unicité de la mesure de Lebesgue s'ensuit.

8 Changement de variables

Le but de cette dernière partie est de montrer une formule de changement de variables dans \mathbb{R}^d .

8.1 Cas linéaire

Nous commençons par le cas d'un changement de variables linéaires. Voici un résultat préliminaire.

Proposition 8.1. *Soit $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$ linéaire et bijective. Alors :*

- a) *E est borélien si et seulement si $T(E)$ est borélien.*
- b) *$\lambda_d(T(E)) = \lambda_d(E)\lambda_d(T([0, 1]^d))$ pour tout borélien E .*

Le résultat suivant nous donne la valeur de $\lambda_d(T([0, 1]^d))$. On identifie une application linéaire T à sa matrice dans la base canonique.

Proposition 8.2. *Si $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$ est linéaire et bijective, alors $\lambda_d(T([0, 1]^d)) = |\det T|$.*

Nous pouvons maintenant énoncer le théorème de changement de variables dans le cas linéaire.

Théorème 8.3 (changement de variables linéaire). *Soit $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$ linéaire et bijective et $f \in \mathcal{L}^1(\mathbb{R}^d)$. Alors $f \circ T \in \mathcal{L}^1(\mathbb{R}^d)$ et*

$$\int_{\mathbb{R}^d} f(y) dy = \int_{\mathbb{R}^d} f(Tx) |\det T| dx.$$

La relation au dessus reste vraie si f est positive et mesurable.

Par analogie au changement de variables sur \mathbb{R} , on pourrait penser que dans la formule au-dessus il devrait y avoir $\det T$ au lieu de $|\det T|$. La formule au-dessus est bien juste avec $|\det T|$ et elle est valable dans \mathbb{R} aussi. En effet, les intervalles dans \mathbb{R} ont une orientation : \int_a^b n'est pas égale à \int_b^a alors que dans l'intégrale de Lebesgue on écrit simplement $\int_{[a,b]}$ et il n'y a pas d'orientation possible.

8.2 Cas général

Rappelons la définition d'un C^1 difféomorphisme.

Définition 8.4 (C^1 difféomorphisme). *Soient U et V deux ouverts de \mathbb{R}^d et $\psi : U \rightarrow V$.*

- *On dit que ψ est C^1 si toutes ses dérivées partielles existent et sont continues.*
- *On dit que ψ est un C^1 difféomorphisme si ψ est bijective et si ψ et ψ^{-1} sont C^1 .*
- *Si ψ est C^1 on définit M_ψ la matrice jacobienne de ψ comme étant la matrice de ses dérivées partielles :*

$$M_\psi = (\partial_j \psi_i)_{1 \leq i,j \leq d}.$$

Le jacobien J_ψ est le déterminant de la matrice jacobienne :

$$J_\psi = \det M_\psi.$$

Voici le théorème général de changement de variables.

Théorème 8.5 (changement de variables dans \mathbb{R}^d). *Soit $\psi : U \rightarrow V$ un C^1 difféomorphisme et $f \in \mathcal{L}^1(V)$. Alors $f \circ \psi | J_\psi | \in \mathcal{L}^1(U)$ et*

$$\int_V f(y) dy = \int_U f \circ \psi(x) |J_\psi(x)| dx.$$

La relation au dessus reste vraie si f est positive et mesurable.

L'idée de la preuve est simple. Il suffit de montrer l'affirmation pour les fonctions indicatrices (puis, par linéarité, on l'a pour les fonctions étagées, puis, par Beppo Levi pour les fonctions positives et enfin dans le cas général). Par le théorème d'unicité des mesures, il suffit de faire pour les indicatrices d'un pavé. Puis on décompose le pavé en des tout petits cubes, on envoie chaque cube par f , puis on somme. Sur chaque petit cube, par la définition de la différentiabilité (qui n'est rien d'autre qu'une formule de Taylor à l'ordre 1) on peut approcher ψ par une fonction linéaire plus un petit reste. Le reste va être négligeable, et pour la partie linéaire on applique la formule du changement de variables linéaire ce qui nous fait apparaître une discrétisation du Jacobien. Lorsque la taille des petits cubes va tendre vers 0, cette discrétisation du Jacobien va tendre vers le Jacobien. Le temps manquant pour les détails de cet argument, nous admettrons la preuve de ce théorème.

8.3 Applications

8.3.1 Coordonnées polaires dans le plan \mathbb{R}^2

On pose $U = \{(r, \theta) ; r > 0, -\pi < \theta < \pi\}$ et $V = \mathbb{R}^2 \setminus \mathbb{R}_-$. Alors l'application

$$\psi(r, \theta) = (r \cos \theta, r \sin \theta)$$

est un C^1 difféomorphisme de U dans V , de jacobien $J_\psi = r$. Remarquons de plus que V est \mathbb{R}^2 privé d'un ensemble de mesure de Lebesgue λ_2 nulle. Le théorème de changement de variables 8.5 donne alors le résultat suivant.

Corollaire 8.6 (changement de variables polaires). *Soit f une fonction mesurable définie sur \mathbb{R}^2 qui est soit positive soit dans $\mathcal{L}^1(\mathbb{R}^2)$. Nous avons que*

$$\int_{\mathbb{R}^2} f(x) dx = \int_U f(r \cos \theta, r \sin \theta) r dr d\theta = \int_0^\infty \int_{-\pi}^\pi f(r \cos \theta, r \sin \theta) r dr d\theta.$$

Remarques.

a) Si f est radiale, par abus de notation $f = f(r)$, alors

$$\int_{\mathbb{R}^2} f(x) dx = \int_0^\infty f(r) 2\pi r dr.$$

b) Nous avons que, dans \mathbb{R}^2 ,

- la fonction $|x|^\alpha$ est intégrable au voisinage de 0, $|x|^\alpha \in \mathcal{L}^1(|x| < 1)$, si et seulement si $\alpha > -2$.
- la fonction $|x|^\alpha$ est intégrable à l'infini, $|x|^\alpha \in \mathcal{L}^1(|x| > 1)$, si et seulement si $\alpha < -2$.

8.3.2 Coordonnées sphériques dans l'espace \mathbb{R}^3

On pose $U = \{(r, \theta, \varphi) ; r > 0, 0 < \theta < \pi, -\pi < \varphi < \pi\}$ et $V = \mathbb{R}^3 \setminus \{(x_1, 0, x_3) ; x_1 \leq 0, x_3 \in \mathbb{R}\}$. Alors l'application

$$\psi(r, \theta, \varphi) = (r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta)$$

est un C^1 difféomorphisme de U dans V , de jacobien $J_\psi = r^2 \sin \theta$. Remarquons de plus que V est \mathbb{R}^3 privé d'un ensemble de mesure de Lebesgue λ_3 nulle. Le théorème de changement de variables 8.5 donne alors le résultat suivant.

Corollaire 8.7 (changement de variables sphériques). *Soit f une fonction mesurable définie sur \mathbb{R}^3 qui est soit positive soit dans $\mathcal{L}^1(\mathbb{R}^3)$. Nous avons que*

$$\begin{aligned} \int_{\mathbb{R}^3} f(x) dx &= \int_U f(r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta) r^2 \sin \theta dr d\theta d\varphi \\ &= \int_0^\infty \int_0^\pi \int_{-\pi}^\pi f(r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta) r^2 \sin \theta dr d\theta d\varphi. \end{aligned}$$

Remarques.

- a) Si f est radiale, par abus de notation $f = f(r)$, alors

$$\int_{\mathbb{R}^3} f(x) \, dx = \int_0^\infty f(r) 4\pi r^2 \, dr.$$

- b) Nous avons que, dans \mathbb{R}^3 ,

- la fonction $|x|^\alpha$ est intégrable au voisinage de 0, $|x|^\alpha \in \mathcal{L}^1(|x| < 1)$, si et seulement si $\alpha > -3$.
- la fonction $|x|^\alpha$ est intégrable à l'infini, $|x|^\alpha \in \mathcal{L}^1(|x| > 1)$, si et seulement si $\alpha < -3$.

8.3.3 Cas de \mathbb{R}^d

On peut trouver un changement de variables, dit coordonnées sphériques généralisées, qui convient dans le cas de \mathbb{R}^d . Il y a cependant une façon plus élégante de procéder. Commençons par définir la mesure de surface sur la sphère unité S^{d-1} de \mathbb{R}^d .

Définition 8.8. Si $A \subset S^{d-1}$ est mesurable, on définit la mesure de surface σ de A par la formule

$$\sigma(A) = \lambda_d(\gamma(A))$$

où $\gamma(A)$ est le secteur de la boule unité engendré par A :

$$\gamma(A) = \{r\omega ; \omega \in A, r \in [0, 1]\}.$$

Remarque. Il est très facile de vérifier que la mesure de surface σ est une mesure définie sur les boréliens de la sphère unité S^{d-1} .

En notant $r = |x|$ (norme euclidienne de \mathbb{R}^d) et $\omega = \frac{x}{|x|}$, toute fonction f de la variable x peut être écrite dans les variables $(r, \omega) \in \mathbb{R}_+^* \times S^{d-1}$ par la formule $f(x) = g(r, \omega) = g(|x|, \frac{x}{|x|})$ ou encore $g(r, \omega) = f(r\omega)$. La formule de changement de variables s'écrit alors dans \mathbb{R}^d sous la forme suivante.

Proposition 8.9. Soit f une fonction mesurable définie sur \mathbb{R}^d qui est soit positive soit dans $\mathcal{L}^1(\mathbb{R}^d)$. Nous avons la formule suivante

$$\int_{\mathbb{R}^d} f(x) \, dx = \int_0^\infty \int_{S^{d-1}} f(r\omega) r^{d-1} \, dr \, d\sigma(\omega).$$

Remarques. Comme dans les cas $d = 2$ et $d = 3$ on a les deux remarques suivantes pour les fonctions radiales.

- a) Si f est radiale, par abus de notation $f = f(r)$, alors

$$\int_{\mathbb{R}^d} f(x) \, dx = \sigma(S^{d-1}) \int_0^\infty f(r) r^{d-1} \, dr.$$

Ici, $\sigma(S^{d-1})$ désigne l'aire de la sphère unité de \mathbb{R}^d .

- b) Nous avons que, dans \mathbb{R}^d ,

- la fonction $|x|^\alpha$ est intégrable au voisinage de 0, $|x|^\alpha \in \mathcal{L}^1(|x| < 1)$, si et seulement si $\alpha > -d$.
- la fonction $|x|^\alpha$ est intégrable à l'infini, $|x|^\alpha \in \mathcal{L}^1(|x| > 1)$, si et seulement si $\alpha < -d$.