Cours Intégration MA62

Frédéric Hérau
Université de Reims

mai 2006

Table des matières

Introduction

Ce cours est la version écrite d'un cours donné partiellement ou en totalité entre 2003 et 2006 à l'Université de Reims. Il correspond à 24H de présence devant les étudiants.

A l'issue de ce cours, les étudiants doivent avoir acquis quelques automatismes d'utilisation de l'Intégrale de Lebesgue, en particulier l'utilisation des Théorèmes fondamentaux suivant :

Les Théorèmes de convergence dominée et de convergence monotone,

les Théorèmes de Fubini-Tonelli et de Fubini.

Cet objectif ne doit pas être abandonné sous prétexte de difficultés passagères à comprendre le cours. En fait la théorie est difficile mais l'usage de l'intégrale de Lebesgue (1906) est relativement facile. Pour l'intégrale de Riemann par contre, les théorèmes d'application nécessitent souvent des hypothèses fortes, par exemple de convergence uniforme.

Intégrale de Lebesgue		Intégrale de Riemann
Théorie difficile, usage facile	$\longleftrightarrow$	Théorie facile, usage délicat

L'utilisation du cours est relativement simple. Les Théorèmes, Propositions, Définitions, et Propriétés vus au fil des chapitres sont **à connaître**. Les Remarques, Lemmes (résultats intermédiaires) ne sont pas à retenir par coeur. Par manque de temps, la construction de l'intégrale ne sera que partiellement abordée en cours.

Toute remarque sur d'éventuelles coquilles ou erreurs est bienvenue. Je tiens à remercier G. Henriet pour m'avoir fourni une première version dactylographiée de la seconde partie. Voici ci-dessous une liste quelques documents et livres dont je me suis servi pour l'élaboration de ce cours. Les documents en ligne sont accessibles par moteur de recherche (auteur + intégration sous google par exemple) :

[1] Bourdarias, C. *Intégration et Applications,* Université de Savoie, dispo. en ligne, 2000.

[2] Faraut, J. *Calcul Intégral,* Belin, 2000.

[3] Gapaillard, J. *Intégration pour la Licence, cours et exercices corrigés,* Dunod, 2002.

[4] Lerner, N. *Cours d'intégration,* Université de Rennes 1, dispo. en ligne, 2002

[5] Revuz, D. *Mesure et Intégration,* Hermann, collection Méthodes, 1997.

[6] Rudin, W. *Real and complex analysis,* McGraw-Hill, New-York, 1987.

Chapitre 1

Préliminaires et Rappels

Dans tout ce chapitre X désigne un ensemble quelconque.

1.1 La droite achevée $\overline{\mathbb{R}}$

Une suite réelle n'a pas toujours de valeur d'adhérence. Ceci est du au fait que $\mathbb{R}, |.|$ n'est pas compact. En ajoutant deux éléments à $\mathbb{R}$ on peut en faire un espace compact.

Définition 1.1.1 *On appelle droite achevée l'espace topologique noté $\overline{\mathbb{R}}$ obtenu en adjoignant à $\mathbb{R}$ deux éléments , notés $+\infty$ et $-\infty$. La topologie de $\overline{\mathbb{R}}$ est alors engendrée par les ouverts de $\mathbb{R}$ et les ensembles*

$$]a,+\infty[\cup\{+\infty\}\,, \qquad \{-\infty\}\cup]-\infty,a[$$

(que l'on notera respectivement $]a,+\infty[$ et $[-\infty,a[$). On définit de manière similaire $\overline{\mathbb{R}}_+$ en adjoignant à $\mathbb{R}^+$ l'élément $+\infty$.

La relation d'ordre sur $\mathbb{R}$ s'étend à $\overline{\mathbb{R}}$ en considérant $-\infty$ comme plus petit élément et $+\infty$ comme plus grand élément. Cet ordre est compatible avec la topologie puisque les ouverts sont comme pour $\mathbb{R}$ des réunions d'intervalles. L'espace $\overline{\mathbb{R}}$ jouit des propriétés suivantes faciles à vérifier

i) $\overline{\mathbb{R}}$ est compacte. En particulier toute partie non vide possède une borne sup et une borne inf.

ii) $\overline{\mathbb{R}}$ est homéomorphe à $[0,1]$. Il suffit pour cela de par exemple prolonger la fonction tan à l'intervalle $[-\pi/2,\pi/2]$ et la fonction arctan à $[-\infty,+\infty]$. (on rappelle que $[0,1]$ et $[-\pi/2,\pi/2]$ sont homéomorphes : une simple application affine envoie continûment l'un sur l'autre). On peut également remarquer que cet homéomorphisme peut être pris croissant.

iii) Toute suite monotone de $\overline{\mathbb{R}}$ est convergente.

Nous venons de voir que l'ordre de $\mathbb{R}$ était compatible avec celui de $\overline{\mathbb{R}}$. Les opérations usuelles le sont aussi sous les conditions naturelles suivantes : soit $x\in\mathbb{R}$

$$x+(\pm\infty)=(\pm\infty), \quad x\times(\pm\infty)=(\pm\infty) \text{ si } x>0, \quad (\mp)\infty \text{ sinon .}$$

On laisse au lecteur les autres opérations naturelles, dont sont bien sur exclues $(+\infty)+(-\infty)$ et $0\times(\pm\infty)$. En fait la bonne manière de considérer ces calculs est de considérer le prolongement par continuité de l'addition et de la multiplication a $\overline{\mathbb{R}} \times \overline{\mathbb{R}}$ (privé des cas exclus ci dessus). Mentionnons dès maintenant que par convention, nous poserons systématiquement dans la suite $0 \times \infty = 0$.

Une conséquence du fait que $\overline{\mathbb{R}}$ soit compact est en particulier que l'ensemble des valeurs d'adhérence d'une suite n'est pas vide (on rappelle que l'ensemble des valeur d'adhérence Adh(u) d'une suite (u_n) est l'ensemble des limites des sous-suites de (u_n) qui convergent).

On introduit maintenant les limites inférieures et supérieures des suites réelles.

Définition 1.1.2 *Soit $(u_n)_{n\in\mathbb{N}}$ une suite de $\overline{\mathbb{R}}$. On définit les suites $\left(\inf_{k\geq n} u_k\right)_{n\in\mathbb{N}}$ et $\left(\sup_{k\geq n} u_k\right)_{n\in\mathbb{N}}$. Ces suites sont clairement monotones donc convergentes dans $\overline{\mathbb{R}}$ (la première croît et la seconde décroît). On note alors*

$$\begin{aligned} \liminf u_n &= \lim\left(\inf_{k\geq n} u_k\right) &&= \sup\left(\inf_{k\geq n} u_k\right) \\ \limsup u_n &= \lim\left(\sup_{k\geq n} u_k\right) &&= \inf\left(\inf_{k\geq n} u_k\right) \end{aligned}$$

On a alors la

Proposition 1.1.3 *Soit $(u_n)_{n\in\mathbb{N}}$ une suite de $\overline{\mathbb{R}}$. Alors $\liminf u_n$ (resp $\limsup u_n$) est la plus petite (resp. la plus grande) valeur d'adhérence de la suite. On a $\liminf u_n \leq \limsup u_n$ et l'égalité a lieu si et seulement si la suite converge vers cette valeur*

Preuve. Avant de commencer rappelons que la limite si elle existe peut-être $\pm\infty$ puisque nous travaillons dans $\overline{\mathbb{R}}$. Considérons donc $x \in \overline{\mathbb{R}}$ une valeur d'adhérence de la suite. Alors il existe une suite extraite $(u_{n_k})_{k\in\mathbb{N}}$ de (u_n) qui converge vers x, et on a pour k tendant vers l'infini

$$x \longleftarrow u_{n_k} \leq \sup_{l\geq n_k} u_l \longrightarrow \limsup u_n.$$

La limite à droite provient simplement du fait que la suite de terme général $\sup_{l\geq n_k} u_l$ est une suite extraite de la suite $\left(\sup_{l\geq n} u_l\right)$ qui, par définition, converge vers $\limsup u_n$. On a donc montré que $y \leq \limsup u_n$ et on aurait de même $y \geq \liminf u_n$.

Pour montrer que $l = \limsup u_n$ est une valeur d'adhérence, on va se restreindre au cas où $-\infty < l < \infty$ i.e. $l \in \mathbb{R}$. La preuve dans le cas $l = \pm\infty$ est tout à fait semblable. On note (v_n) la suite $\sup_{n\geq k} u_k$ qui converge donc en décroissant vers l. Alors pour tout $\varepsilon > 0$, et $n \geq 0$ il existe $n_\varepsilon \geq n$ tel que $l \leq v_{n_\varepsilon} \leq l + \varepsilon$. Mais par ailleurs, v_{n_ε} étant un sup, il existe $n'_\varepsilon \geq n_\varepsilon$ tel que

$$v_{n'_\varepsilon} - \varepsilon \leq u_{n'_\varepsilon} \leq v_{n'_\varepsilon}$$

d'où l'on déduit

$$l - \varepsilon \leq v_{n'_\varepsilon} - \varepsilon \leq u_{n'_\varepsilon} \leq v_{n'_\varepsilon} \leq l + \varepsilon.$$

En résumé on a montré que pour tout $\varepsilon > 0$ et tout $N \in \mathbb{N}$ il existe $n (= n'_\varepsilon$ ici) tel que $|u_n - l| \leq \varepsilon$. l est donc une valeur d'adhérence de (u_n).

Pour finir, si $\liminf u_n$ et $\limsup u_n$ sont toutes les deux égales à $l \in \overline{\mathbb{R}}$, alors pour n tendant vers l'infini

$$l \longleftarrow \inf_{k\geq n} u_k \leq u_n \leq \sup_{k\geq n} u_k \longrightarrow l,$$

ce qui prouve le résultat. $\square$

Un autre résultat qui sera utile dans le cours concerne les sommes doubles dans $\overline{\mathbb{R}}$. On note $\overline{\mathbb{R}}_+ = [0, +\infty]$.

Lemme 1.1.4 *Soit* $(u_{j,k})$ *une suite double de* $\overline{\mathbb{R}}_+$. *Alors*

$$\sum_j \left(\sum_k u_{j,k} \right) = \sum_k \left(\sum_j u_{j,k} \right).$$

Cette somme est simplement notée $\sum_{j,k} u_{j,k}$.

Preuve. On sait déjà que pour j fixé la série $\sum_k u_{j,k}$ converge dans $\overline{\mathbb{R}}$ puisque son terme général est positif. Pour la même raison la série $\sum_j (\sum_k u_{j,k})$ converge également, ce qui donne un sens aux séries de l'énoncé. Maintenant pour $J, K \in \mathbb{N}$ on a

$$\sum_{j \le J} \sum_{k \le K} u_{j,k} = \sum_{k \le K} \sum_{j \le J} u_{j,k} \le \sum_{k \le K} \sup_J \sum_{j \le J} u_{j,k} = \sum_{k \le K} \sum_j u_{j,k} \le \sup_K \sum_{k \le K} \sum_j u_{j,k} = \sum_k \sum_j u_{j,k}.$$

Notons alors $S = \sum_k \sum_j u_{j,k}$ on vient de montrer que pour tout j, k,

$$\sum_{j \le J} \sum_{k \le K} u_{j,k} \le S,$$

d'où il vient en prenant le sup sur K à l'intérieur de la somme finie,

$$\sum_{j \le J} \sum_k u_{j,k} \le S,$$

puis en prenant le sup sur J,

$$\sum_j \sum_k u_{j,k} \le S.$$

Par symétrie on obtient l'inégalité dans l'autre sens, le résultat est alors prouvé. □

1.2 Rappels sur les fonctions réelles

Soit X et Y deux ensembles, on désigne par $\mathcal{F}(X, Y)$ l'ensemble des fonctions de X dans Y. Lorsque $Y = \mathbb{R}$ ou $\overline{\mathbb{R}}$ on définit naturellement pour f et $g \in \mathcal{F}(X, Y)$ les fonctions $\inf(f, g)$ et $\sup(f, g)$. Par exemple les fonctions parties positives et négatives d'une fonction sont les deux fonctions positives définies par

$$f_+ : x \longmapsto \sup(f, 0) \quad \text{et} \quad f_- : x \longmapsto \sup(-f, 0),$$

pour lesquelles on a $f = f_+ - f_-$. Pour une famille de fonctions $(f_n)_{n \in \mathbb{N}}$ dans $\mathcal{F}(X, \overline{\mathbb{R}})$ on a également $\sup_n f_n$ et $\inf_n f_n \in \mathcal{F}(X, \overline{\mathbb{R}})$. Il en va de même pour les fonctions $\limsup f_n$ et $\liminf f_n$ définies respectivement sur X par

$$x \longmapsto \limsup f_n(x) \text{ et } x \longmapsto \liminf f_n(x).$$

Parmi les fonctions réelles certaines jouent un rôle important en théorie de l'intégration, il s'agit des fonctions indicatrices d'ensemble.

Définition 1.2.1 *Soit X un ensemble et $A \subset X$. On appelle fonction indicatrice de A et on note $\mathbb{1}_A$ la fonction définie par* $\mathbb{1}_A(x) = \begin{cases} 1 \text{ si } x \in A \\ 0 \text{ sinon.} \end{cases}$

L'application $\Phi : \begin{cases} \mathcal{P}(X) & \longrightarrow \mathcal{F}(X, \mathbb{R}) \\ A & \longmapsto \mathbb{1}_A \end{cases}$ est injective et a les propriétés suivantes :

i) $A \subset B \Longrightarrow \mathbb{1}_A \leq 1_B$;

ii) $\mathbb{1}_{A \cap B} = \mathbb{1}_A \times \mathbb{1}_B$;

iii) $\mathbb{1}_{A \cup B} = \sup(\mathbb{1}_A, \mathbb{1}_B)$ $\quad$ ($= \mathbb{1}_A + \mathbb{1}_B$ si A et B sont disjoints) ;

iv) $\mathbb{1}_{A^c} = 1 - \mathbb{1}_A$, ou on a noté A^c le complémentaire dans X de A ;

v) Si $(A_i)_{i \in I}$ est une partition de X, alors $\sum_i \mathbb{1}_{A_i} = 1$.

Etant donnés deux ensemble $A \subset X$, $B \subset Y$, et une fonction $f \in \mathcal{F}(X, Y)$, on rappelle les définitions d'image et d'image réciproque d'ensemble :

$$f(A) = \{f(x); x \in A\}, \qquad f^{-1}(B) = \{x \in X \text{ tels que } f(x) \in B\}. \tag{1.1}$$

On prendra garde au fait que l'emploi des notations f et f^{-1} ci-dessus constitue bien sur un abus d'écriture : la fonction f n'est *pas* définie sur les ensembles et la fonction f^{-1} ... n'est a priori pas définie du tout ! On a alors les propriétés suivantes dont la vérification est laissée au lecteur :

$\forall B \subset Y,\ f(f^{-1}(B)) \subset B$, avec égalité quand f est surjective ;

$\forall A \subset X,\ f^{-1}(f(A)) \supset A$, avec égalité quand f est injective ;

$\forall B, B' \subset Y,\ f^{-1}(B \cup B') = f^{-1}(B) \cup f^{-1}(B')$;

$\forall B, B' \subset Y,\ f^{-1}(B \cap B') = f^{-1}(B) \cap f^{-1}(B')$;

$\forall B \subset Y, f^{-1}(B^c)) = (f^{-1}(B))^c$.

1.3 Dénombrabilité

La dénombrabilité est également au coeur des processus limites utilisés dans la théorie de l'intégration de Lebesgue. On rappelle qu'un ensemble D est dit fini ou dénombrable s'il est équipotent à une partie de $\mathbb{N}$, c'est à dire s'il existe une injection φ de D dans $\mathbb{N}$. S'il est fini il peut être mis en bijection avec un ensemble du type $\{1, ..., n\}$ où n est unique et est appelé nombre d'éléments de D. S'il est infini alors il est équipotent à $\mathbb{N}$: en effet, on peut grace à l'injection φ ci-dessus considérer D comme une partie de $\mathbb{N}$, et poser

$$d_1 = \min D, \quad d_2 = \min(D/\{d_1\}), ... \quad d_k = \min(D/\{d_1, ...d_{k-1}\}).$$

Cette définition a bien un sens puisque $\mathbb{N}$ est totalement ordonné et que toute partie non vide de $\mathbb{N}$ admet un plus petit élément. Il est alors clair que l'application $k \longmapsto d_k$ constitue une bijection de $\mathbb{N}$ sur D.

On rappelle ici quelques exemples fondamentaux d'ensembles dénombrables :

i) $\mathbb{N}^*$, $2\mathbb{N}$, $\mathbb{Z}$. Pour les 2 premiers par exemple on peut utiliser les application $k \longmapsto k+1$ et resp. $k \longmapsto 2k$, qui sont des bijections de $\mathbb{N}$ sur $\mathbb{N}^*$, resp. $2\mathbb{N}$.

ii) $\mathbb{N} \times \mathbb{N}$, $\mathbb{Q}$. Pour le premier il suffit de l'application

$$(p, q) \in \mathbb{N} \times \mathbb{N} \longmapsto 2^p(2q+1) - 1 \in \mathbb{N}$$

est une bijection. Le fait que $\mathbb{Q}$ s'injecte dans $\mathbb{N} \times \mathbb{N}$ implique alors que $\mathbb{Q}$ est dénombrable. On obtient facilemnet que $\mathbb{Q}^n$ l'est également.

On peut donner une application qui sera utile dans le chapitre suivant au fait que $\mathbb{Q}$ est dénombrable.

Lemme 1.3.1 *Soit d un entier. Il existe une famille dénombrable de pavés compacts de $\mathbb{R}^d$ (produit d'intervalle compacts) telle que tout ouvert soit réunion (nécessairement dénombrable) d'une sous-famille de ces pavés. On peut choisir ces pavés comme produit d'intervalles compacts d'extrémités rationnelles.*

Preuve. Soit donc U un ouvert. Pour tout $x = (x_1, ...x_d) \in U$, il existe un cube ouvert contenant x et contenu dans U puisque U est ouvert,

$$\left\{ y \in \mathbb{R}^d ; |y_j - x_j| \leq \rho \right\} \subset U,$$

pour un ρ non nul (c'est dû par exemple à l'équivalence des normes sur $\mathbb{R}^d$). Comme $\mathbb{Q}$ est dense dans $\mathbb{R}$, il existe p_j et $q_j \in \mathbb{Q}$ tels que

$$x_j - \rho < p_j < x_j < q_j < x_j + \rho.$$

Donc le compact $\prod_{j=1...d} [p_j, q_j]$ contient x et est contenu dans U. L'ouvert U est donc réunion d'une sous famille de la famille des pavés d'extrémités rationnelles. Or cette famille est dénombrable, puisque équipotente à $\mathbb{Q}^{2d}$ qui est dénombrable. Cela complète la preuve. $\square$

Chapitre 2

Tribus et applications mesurables

Le but de ce chapitre est de définir les objets que nous seront amenés à mesurer : les ensembles mesurables et les fonctions mesurables. De nouveau X est a priori un ensemble quelconque.

2.1 Tribus, espaces mesurables

Définition 2.1.1 *Soit X un ensemble et $\mathcal{M}$ une famille de parties de X. On dit que $\mathcal{M}$ est une tribu sur X si*

$$\begin{array}{ll} (a) & A \in \mathcal{M} \Longrightarrow A^c \in \mathcal{M}, \\ (b) & (A_n)_{n\in\mathbb{N}} \subset \mathcal{M} \Longrightarrow \cup_{n\in\mathbb{N}} A_n \in \mathcal{M}, \\ (c) & X \in \mathcal{M}. \end{array}$$

On insiste sur le fait que $\mathcal{M}$ n'est pas une partie de X, c'est une famille de parties. On remarque que la propriété (c) est une conséquence immédiate de (a) et (b), pourvu que $\mathcal{M}$ soit non vide. La propriété (b) est appelée *stabilité par réunion dénombrable.* Finalement, on peut remarquer que (b) implique qu'une tribu et stable par intersection dénombrable.

$$(b') \quad (A_n)_{n\in\mathbb{N}} \subset \mathcal{M} \Longrightarrow \cap_{n\in\mathbb{N}} A_n \in \mathcal{M}.$$

Dans la définition d'une tribu on peut remplacer (b) par (b'). Pour finir les axiomes (a) et (c) impliquent que

$$(c') \quad \emptyset \in \mathcal{M}.$$

On peut de même échanger (c) et (c') dans la définition.

Pour finir on peut remarquer que dans les axiomes d'une topologie – autre ensemble de parties fondamental en analyse dont la definition est rappelée en (2.1)) – la dénombrabilité ne jouait pas ce rôle central. Ce rôle va se confirmer plus loin.

Exemples 2.1.2 On peut donner déjà quelques exemples simples de tribus.

1. La tribu triviale : $\mathcal{M} = \{\emptyset, X\}$;
2. La tribu grossière : $\mathcal{M} = \mathcal{P}(X)$;
3. La tribu associée a une partition finie : Si $(A_i)_{i\in 1,\ldots,n}$ est une partition finie, alors

$$\mathcal{M} = \{\cup_{k\in J} A_k \quad ; \quad J \subset \{1, ..., n\}\}$$

est une tribu. En effet notons $B(J) = \cup_{k \in J} A_k$, on trouve que $B(J)^c = B(J^c)$, c'est à dire la stabilité par prise du complémentaire. Par ailleurs la stabilité par union dénombrable (en fait finie) est immédiate, et implique que $X \in \mathcal{M}$.

Définition 2.1.3 *Si X est une ensemble et $\mathcal{M}$ est une tribu sur X, le couple $(X, \mathcal{M})$ s'appelle espace mesurable. Les éléments de $\mathcal{M}$ sont appelés parties mesurables de X.*

La notion de tribu engendrée permet de définir des tribus sans les décrire explicitement. On commence par remarquer que si $(\mathcal{M})_{i \in I}$ est une famille de tribu, alors

$$\mathcal{M} = \cap_{i \in I} \mathcal{M}_i$$

est aussi une tribu : Si (A_n) est une famille dénombrable de parties de $\mathcal{M}$, alors pour tout $i \in I$, $(A_n) \subset \mathcal{M}_i$, donc par union dénombrable on obtient $\cup_{n \in \mathbb{N}} A_n \in \mathcal{M}_i$. Donc $\cup_{n \in \mathbb{N}} A_n \in \mathcal{M}$. Le passage au complémentaire s'écrit de la même manière, et X appartient à chaque $\mathcal{M}_i$ donc à $\mathcal{M}$.

Un ensemble $\mathcal{C}$ de parties de X étant donné on applique maintenant cette propriété à l'ensemble des tribus contenant $\mathcal{C}$, en remarquant qu'il en existe au moins une : la tribu grossière $\mathcal{P}(X)$. Cela conduit à la définition suivante :

Définition 2.1.4 *Soit X un ensemble et $\mathcal{C} \subset \mathcal{P}(X)$. On appelle tribu engendrée par $\mathcal{C}$ et on note $\sigma(\mathcal{C})$ la plus petite tribu contenant $\mathcal{C}$:*

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

En général, étant donnés un ensemble de parties $\mathcal{C}$ et une tribu $\mathcal{M}$, pour montrer que $\mathcal{M} = \sigma(\mathcal{C})$, il suffit de montrer que toute tribu contenant $\mathcal{C}$ contient également $\mathcal{M}$. On remarquera également que $\mathcal{M}$ étant une tribu quelconque contenant $\mathcal{C}$, on a

$$\sigma(\mathcal{C}) \subset \mathcal{M},$$

puisque $\sigma(\mathcal{C})$ est la plus petite.

On retourne maintenant à quelques exemples de tribus engendrées. Il est ainsi immédiat de montrer que la tribu associée à une partition est en fait la tribu engendrée par cette partition (exemple 2.1.2). On prendra cependant garde au fait qu'en général, la tribu engendrée par un ensemble de parties $\mathcal{C}$ ne coincide pas avec l'ensemble des unions de parties de $\mathcal{C}$, ni avec l'ensemble de ses intersections.

Un des exemples fondamentaux de tribus provient de la construction précédente appliquée à une topologie (i.e. l'ensemble des ouverts d'un espace topologique).

Définition 2.1.5 *Soit X un espace topologique. On appelle tribu borélienne $\mathcal{B}(X)$ la tribu engendrée par les ouverts de X. Ses éléments sont appelés boréliens.*

Il n'est peut-être pas inutile de rappeler ce qu'est un ouvert : Une famille de parties $\mathcal{O}$ de X est une topologie, dont les éléments sont appelés ouverts si

$$\begin{aligned} &(a) \quad U_1, U_2 \in \mathcal{O} \Longrightarrow U_1 \cap U_2 \in \mathcal{O}, \\ &(b) \quad (U_i)_{i \in I} \subset \mathcal{O} \Longrightarrow \cup_{i \in I} U_i \in \mathcal{O}, \\ &(c) \quad \emptyset, X \in \mathcal{O}. \end{aligned} \tag{2.1}$$

On remarquera en particulier que l'union se fait sur un ensemble infini quelconque, pas forcément dénombrable.

Revenons maintenant aux boréliens. D'après la définition, un borélien peut avoir une structure compliquée. Par exemple une réunion dénombrable de fermés est un borélien. Ainsi $\mathbb{Q}$, est un borélien de $\mathbb{R}$ puisque réunion dénombrable de singletons. $\mathbb{R}/\mathbb{Q}$ est également un borélien par prise du complémentaire de $\mathbb{Q}$.

La tribu des boréliens de $\mathbb{R}^d$, pour $d \geq 1$, satisfait la propriété suivante.

Proposition 2.1.6 *On a*

i) La tribu $\mathcal{B}(\mathbb{R})$ est engendrée par l'une quelconque des familles d'intervalles :

$$\{]-\infty, a[;\ a \in \mathbb{R}\}, \quad \{]-\infty, a];\ a \in \mathbb{R}\}, \quad \{]a, +\infty[;\ a \in \mathbb{R}\}, \quad \{[a, +\infty[;\ a \in \mathbb{R}\}$$
$$\{[a, b];\ a, b \in \mathbb{R}\}, \quad \{]a, b[;\ a, b \in \mathbb{R}\}, \quad \{]a, b];\ a, b \in \mathbb{R}\}, \quad \{[a, b[;\ a, b \in \mathbb{R}\}.$$

ii) La tribu $\mathcal{B}(\mathbb{R}^d)$ est engendrée par les pavés compacts (produits d'intervalles fermés).

Preuve. On va d'abord prouver la troisième assertion. Soit $\mathcal{C}$ l'ensemble des pavés compacts. On a $\mathcal{C} \subset \mathcal{B}(\mathbb{R}^d)$ donc aussi

$$\sigma(\mathcal{C}) \subset \mathcal{B}(\mathbb{R}).$$

Pour obtenir l'inclusion réciproque, on considère $\sigma(\mathcal{C})$ la tribu engendrée par les pavés compact. Alors d'après le Lemme 1.3.1, tout ouvert U est réunion dénombrable de pavés compacts (que l'on peut prendre d'extrémités rationelles) donc $U \in \sigma(\mathcal{C})$ puisque $\sigma(\mathcal{C})$ est une tribu. La topologie $\mathcal{O}$ toute entière (l'ensemble des ouvert) est donc inclue dans $\sigma(\mathcal{C})$ et on obtient

$$\mathcal{O} \subset \sigma(\mathcal{C}) \Longrightarrow \sigma(\mathcal{O}) \subset \sigma(\mathcal{C}).$$

Puisque $\sigma(\mathcal{C})$ est par définition l'ensemble des boréliens, on a prouve l'inclusion réciproque

$$\mathcal{B}(\mathbb{R}^d) \subset \sigma(\mathcal{C})$$

et l'égalité en découle.

Pour la deuxième assertion, le cas des intervalles fermés du type $[a, b]$ est un cas particulier pour $d = 1$ de ce qui vient d'être prouvé. Par ailleurs pour tout $a \leq b \in \mathbb{R}$, on a

$$[a, b] = \cap_{n \in \mathbb{N}}]a - 1/n, b + 1/n[$$

On obtient que les pavés fermés sont dans la tribu des pavés ouverts, et que donc la tribu $\mathcal{B}(\mathbb{R}^d)$ est aussi engendrée par les pavés ouverts, et par la même procédure par les pavés mixtes. De même on remarque que

$$[a, b] = [a, +\infty[\cap]-\infty, b] = [a, +\infty[\cap]b, +\infty[^c = [a, +\infty[\cap \left(\cup_{n \in \mathbb{N}} [b + 1/n, +\infty[\right)^c,$$

donc la tribu $\mathcal{B}(\mathbb{R}^d)$ est aussi engendrée par les intervalles du type $[a, \infty[$. Les autres cas se traitent de manière identique. □

Remarque 2.1.7 La même preuve donne également que la tribu $\mathcal{B}(\overline{\mathbb{R}})$ est engendrée par l'une quelconque des familles d'intervalles suivante :

$$\{[-\infty, a[;\ a \in \mathbb{R}\}, \quad \{[-\infty, a];\ a \in \mathbb{R}\}, \quad \{]a, +\infty];\ a \in \mathbb{R}\}, \quad \{]a, +\infty[;\ a \in \mathbb{R}\}.$$

On introduit également les tribus associées à une fonction :

Proposition 2.1.8 (Tribu image réciproque) *Soit X un ensemble, $(Y, \mathcal{N})$ un espace mesurable et $f : X \longmapsto Y$ une application. Alors l'ensemble*

$$f^{-1}(\mathcal{N}) \stackrel{\text{def}}{=} \left\{f^{-1}(B);\ B \in \mathcal{N}\right\}$$

est une tribu, appelée tribu image réciproque de $\mathcal{N}$ par f. (On l'appelle également tribu engendrée par f, et elle est parfois notée $\sigma(f)$.)

Preuve. La preuve est une conséquence directe des propriétés sur les images réciproques sur les fonctions rappelées en fin de Section 2.2 : Si $B \in \mathcal{N}$, alors

$$\left(f^{-1}(B)\right)^c = f^{-1}(B^c) \in f^{-1}(\mathcal{N}),$$

de plus

$$f^{-1}(Y) = X \in f^{-1}(\mathcal{N}),$$

et enfin pour (B_n) une suite de parties mesurables de Y on a

$$\cup_{n\in\mathbb{N}} f^{-1}(B_n) = f^{-1}\left(\cup_{n\in\mathbb{N}} B_n\right) \in f^{-1}(\mathcal{N}),$$

ce qui implique le résultat. $\square$

Pour finir cette partie sur les tribus on introduit

Définition 2.1.9 (Tribu produit) Soit $(X_i, \mathcal{M}_i)_{1\le i\le n}$ une famille finie d'espace mesurables. On appelle tribu produit des $\mathcal{M}_i$ et on note $\bigotimes_{1\le i\le n} \mathcal{M}_i$ la tribu engendrée par la famille des produits $\prod_{1\le i\le n} A_i$ lorsque $A_i \in \mathcal{M}_i$. On appelle alors espace produit l'ensemble $\prod_{1\le i\le n} X_i$ muni de la tribu $\bigotimes_{1\le i\le n} \mathcal{M}_i$. On le notera parfois $\prod_{1\le i\le n}(X_i, \mathcal{M}_i)$.

Exemple 2.1.10 On pourra par exemple vérifier en exercice que

$$\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^2).$$

2.2 Applications mesurables

Définition 2.2.1 *Soient $(X, \mathcal{M})$ et $(Y, \mathcal{N})$ deux espaces mesurables. Une application $f : X \longrightarrow Y$ est dite mesurable si pour tout $B \in \mathcal{N}$, $f^{-1}(B) \in \mathcal{M}$. (Ceci s'écrit aussi $f^{-1}(\mathcal{N}) \subset \mathcal{M}$).*

On remarquera que si $\mathcal{M} = \mathcal{P}(X)$ ensemble des parties de X, alors toute fonction est mesurable. C'est de toute façon la plus grosse tribu possible sur X. Réciproquement on peut se demander quelle est la plus petite tribu sur X rendant f mesurable. c'est évidement $f^{-1}(\mathcal{N})$, la tribu engendrée par f.

Exemple 2.2.2 On pourra verifier que si A est un ensemble mesurable, alors la fonction caractéristique $\mathbb{1}_A$ est mesurable (l'espace d'arrivée est ici $\mathbb{R}$ muni de la tribu borélienne).

En effet, pour $B \in \mathcal{B}(\mathbb{R})$, l'ensemble $\mathbb{1}_A^{-1}(B)$ est soit le vide soit A, et dans tout les cas dans $\mathcal{M}$.

Remarque 2.2.3 Lorsque les espaces X et Y sont des espaces topologiques, ils sont munis naturellement de leur tribu borélienne. Les fonctions mesurables de X dans Y sont alors dites *boréliennes*. Le cas des fonctions à valeur réelles sera en particulier très important.

Les fonctions mesurables se composent bien :

Proposition 2.2.4 *Soient $(X, \mathcal{M})$, $(Y, \mathcal{N})$ et $(Z,)$ trois espaces mesurables et*

$$X \xrightarrow{f} Y \xrightarrow{g} Z$$

des applications mesurables. Alors $g \circ f$ est mesurable.

Preuve. Soit $C \in \mathcal{T}$. Alors $(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C)) \in \mathcal{M}$. □

La définition de mesurabilité n'est a priori pas très maniable. Dans la pratique pour montrer qu'une application est mesurable, on utilise souvent le critère suivant

Proposition 2.2.5 *Soient $(X, \mathcal{M})$ et $(Y, \mathcal{N})$ deux espaces mesurables. On suppose que $\mathcal{N} = \sigma(\mathcal{C})$ pour un ensemble $\mathcal{C}$ donné. Pour que f soit mesurable il (faut et il) suffit que $f^{-1}(\mathcal{C}) \subset \mathcal{M}$.*

Preuve. Pour montrer que f est mesurable, il faut montrer que pour tout $B \in \mathcal{N}$, on $f^{-1}(B) \in \mathcal{M}$. Considérons donc cet ensemble

$$\mathcal{T} = \left\{B \in \mathcal{N}; f^{-1}(B) \in \mathcal{M}\right\} \subset \mathcal{N}$$

et montrons qu'il vaut $\mathcal{N}$ tout entier. Pour ça il suffit de montrer que c'est une tribu qui contient $\mathcal{C}$ ce qui donnera

$$\mathcal{N} = \sigma(\mathcal{C}) \subset \sigma(\mathcal{T}) = \mathcal{T},$$

d'où le résultat. Il est clair que cet ensemble contient $\mathcal{C}$ puisque $f^{-1}(\mathcal{C}) \subset \mathcal{M}$. Montrons que c'est une tribu : Pour $B \in \mathcal{T}$, on a $f^{-1}(B^c) = (f^{-1}(B))^c \in \mathcal{M}$ puisque $\mathcal{M}$ est une tribu. De plus, $f^{-1}(Y) = X \in \mathcal{M}$. Enfin si (B_n) est une suite d'éléments de $\mathcal{T}$, on a $f^{-1}(\cup_{n\in\mathbb{N}} B_n) = \cup_{n\in\mathbb{N}}(f^{-1}(B_n)) \in \mathcal{M}$. Donc $\mathcal{T}$ est une tribu. □

La première application de cette proposition est fondamentale, ne serait-ce que pour la légitimité de la théorie de la mesure que nous sommes en train de construire.

Proposition 2.2.6 *Soient $(X_1, \mathcal{O}_1)$ et $(X_2, \mathcal{O}_2)$ deux espaces topologiques munis de leurs tribus boréliennes respectives. Alors toute fonction continue de X_1 dans X_2 est mesurable.*

Preuve. La continuité d'une fonction signifie que $f^{-1}(\mathcal{O}_2) \subset \mathcal{O}_1$ (suivant les conventions d'écritures habituelles). Notons $\mathcal{B}_1$ et $\mathcal{B}_2$ les tribus boréliennes associées à X_1 et X_2. On obtient donc que a fortiori que $f^{-1}(\mathcal{O}_2) \subset \mathcal{B}_1$. Or $\sigma(\mathcal{O}_2) = \mathcal{B}_2$ donc la Proposition 2.2.5 implique le résultat.

La deuxième application est un critère relativement simple pour montrer la mesurabilité d'une fonction à valeur dans $\mathbb{R}$ ou $\overline{\mathbb{R}}$.

Sauf mention explicite du contraire on supposera lorsque l'espace d'arrivée est $\mathbb{R}$ (resp. $\overline{\mathbb{R}}$) qu'il est toujours muni de sa tribu des boréliens. Une application sera alors simplement dite mesurable sans référence aux tribus.

Proposition 2.2.7 *Soit $(X, \mathcal{M})$ un espace mesurable et $f : X \longrightarrow \mathbb{R}$ (resp $\overline{\mathbb{R}}$) une application. Les énoncés suivants sont équivalent :*

$$\begin{gathered} i)\ f \text{ est mesurable}, \quad ii)\ \forall b \in \mathbb{R},\ \{f < b\} \in X, \quad iii)\ \forall b \in \mathbb{R},\ \{f \leq b\} \in X, \\ iv)\ \forall b \in \mathbb{R},\ \{f > b\} \in X, \quad v)\ \forall b \in \mathbb{R},\ \{f \geq b\} \in X, \end{gathered} \tag{2.2}$$

Preuve. La preuve est immédiate : il suffit de considérer les ensembles générateurs de la tribu des boréliens vus dans le i) de la Proposition 2.1.6 et d'utiliser la Proposition 2.2.5. On fera attention au fait que lorsque l'espace d'arrivée est $\mathbb{R}$, on a $\{f < b\} = f^{-1}(]-\infty, b[)$ alors que si l'espace d'arrivée est $\overline{\mathbb{R}}$, c'est $f^{-1}([-\infty, b[)$.

Exemple 2.2.8 Considérons une fonction monotone $f : X \longrightarrow \mathbb{R}$, où X est une partie mesurable de $\mathbb{R}$, que l'on muni de la tribu $\mathcal{B}(\mathbb{R}) \cap X$ (voir exercices). Pour fixer les idées supposons f par exemple croissante. Pour $b \in \mathbb{R}$, posons

$$A = f^{-1}(] - \infty, b])$$

On vérifie que si $a \in A$, alors pour tout $x \leq a$, on a $f(x) \leq f(a) \leq b$ donc $x \in A$. On vient de montrer que A est un intervalle (de X) du type $] - \infty, \sup A[\cap X$ ou $] - \infty, \sup A] \cap X$. C'est donc un élément de $\mathcal{B}(\mathbb{R}) \cap X$. Par suite f est mesurable.

En troisième application de la Proposition 2.2.5, on étend le Lemme de composition des fonctions et on étudie la stabilité de l'espace des fonctions mesurables vis a vis des opérations usuelles.

Proposition 2.2.9 *Soient $(X, \mathcal{M})$ et $(Y, \mathcal{N})$ deux espaces mesurables et $\mathbb{R}^d$ muni de la tribu des Boréliens (resp $\overline{\mathbb{R}}_+^d$). Soient alors $u_1, ..., u_d$ des applications mesurables de X dans $\mathbb{R}$ (resp $\overline{\mathbb{R}}_+$) et Φ une application mesurable de $\mathbb{R}^d$ (resp $\overline{\mathbb{R}}_+^d$) dans Y. Alors l'application*

$$\begin{array}{ll} X & \longrightarrow Y \\ x & \longmapsto \Phi(u_1(x), ..., u_d(x)) \end{array}$$

est mesurable.

En particulier si $f,\ g : X \longrightarrow \mathbb{R}$ (resp $\overline{\mathbb{R}}_+$) sont mesurables, alors alors $f + g$, fg, $\sup(f, g)$, $\inf(f, g)$ le sont aussi. (on fait la convention $0 \times \infty = 0$)

De plus, $f : X \longrightarrow \mathbb{C}$ est mesurable si et seulement si $\mathrm{Re}\,(f)$ et $\mathrm{Im}\,(f)$ le sont et dans ce cas $|f|$ l'est aussi. Si $f,\ g : X \longrightarrow \mathbb{C}$ sont mesurables, alors $f + g$ et fg le sont aussi.

Preuve. Pour montrer la première partie de la proposition, il suffit de montrer la mesurabilité de l'application $u : x \longrightarrow (u_1(x), ...u_d(x))$ de X dans $\mathbb{R}$. Vérifions que l'image réciproque d'un pavé compact est dans $\mathcal{M}$. On a :

$$u^{-1}\left(\prod_{1 \leq j \leq d} [a_j, b_j]\right) = \bigcap_{1 \leq j \leq d} u_j^{-1}([a_j, b_j]).$$

Or les u_j sont mesurables, donc cet ensemble est dans $\mathcal{M}$.

Pour la deuxième partie, la preuve est immédiate : par exemple l'addition est continue donc mesurable de $\mathbb{R}^2$ dans $\mathbb{R}$, ce qui donne la mesurabilité de $f + g$. Pour la multiplication, on a le même résultat. Pour le cas de la droite achevée positive, et avec la convention $0 \times \infty = 0$, l'addition est continue donc mesurable et la multiplication est mesurable (attention elle n'est pas continue sur $\overline{\mathbb{R}}^2$), ce qui donne le résultat. On fait de même pour les applications à valeur dans $\mathbb{C}$. $\square$

2.3 Structure des fonctions mesurables

Proposition 2.3.1 *Soit* $(X, \mathcal{M})$ *un espace mesurable et* $(f_n)_{n\in\mathbb{N}}$ *une suite de fonctions mesurables de* X *dans* $\overline{\mathbb{R}}$. *Alors les fonctions* $\sup f_n$, $\inf f_n$, $\limsup f_n$ *et* $\liminf f_n$ *sont mesurables. En particulier la limite simple d'une suite de fonctions mesurables est mesurable.*

Preuve. Posons $g = \sup f_n$. On utilise de nouveau la Proposition fondamentale 2.2.5 : Soit $b \in \mathbb{R}$, on a

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

par union dénombrable dans $\mathcal{M}$. En effet pour x dans $g^{-1}(]b, +\infty])$, on a

$$\sup_{n\in\mathbb{N}} f_n(x) = g(x) > b \Longleftrightarrow \exists n_0 \in \mathbb{N}, \text{ tel que } f_{n_0}(x) > b.$$

D'après la remarque 2.1.7, la tribu des boréliens $\mathcal{B}(\overline{\mathbb{R}})$ est engendrée par la famille des intervalles du type $]b, +\infty]$, donc $\sup f_n$ est mesurable. Par ailleurs les identités

$$\inf f_n = -\sup(-f_n), \quad \limsup f_n = \inf_n(\sup_{k\geq n} f_k), \quad \liminf f_n = \sup_n(\inf_{k\geq n} f_k)$$

assurent la mesurabilité des fonctions correspondantes. Si pour tout $x \in X$, $f_n(x)$ converge, alors d'après la proposition 1.1.3, $\lim f_n(x) = \limsup f_n(x)$ donc la limite simple de f_n est mesurable. $\square$

On a déjà vu qu'une fonction caractéristique était mesurable. La famille suivante est importante

Définition 2.3.2 *Soit* $(X, \mathcal{M})$ *un espace mesurable. Une fonction* s *est dite étagée sur* X *si* $s : X \longrightarrow \mathbb{R}_+$ *est mesurable et ne prend qu'un nombre fini de valeurs.*

Si on note $\alpha_1, ..., \alpha_m$ *les valeurs distinctes prises par* s *et si on note pour chaque* k $A_k = s^{-1}(\{\alpha_k\}) = \{x \in X;\ s(x) = \alpha_k\}$, *alors l'ensemble des* A_k *forme une partition de* X *et on appelle écriture canonique l'écriture unique suivante* $s = \sum_{1\leq k\leq m} \alpha_k \mathbb{1}_{A_k}$.

Une fonction étagée s est donc une combinaison linéaire finie de fonctions indicatrices mesurables. Il est facile de vérifier que l'ensemble des fonctions étagées forme une algèbre. La Proposition suivante sera très utile pour construire l'intégrale d'une fonction dans le prochain chapitre :

Théoreme 2.3.3 *Soit* $(X, \mathcal{M})$ *un espace mesurable et* $f : X \longrightarrow \overline{\mathbb{R}}_+$ *une application mesurable. Il existe une suite de fonctions étagées* $(s_n)_{n\geq 1}$ *telle que* *i)* $0 \leq s_1 \leq s_2 \leq .. \leq s_n \leq s_{n+1} \leq ... \leq f$, *(suite croissante de fonctions)*

ii) $\forall x \in X, \lim_{n\to\infty} s_n(x) = f(x)$, *(convergence simple),*

ce que l'on notera $\lim \uparrow s_n = f$. *Si en plus* f *est bornée, alors la convergence de* s_n *vers* f *est uniforme, i.e.*

$$\lim_{n\to\infty} \sup_{x\in X} |s_n(x) - f(x)| \to 0.$$

Bien sur la réciproque du point ii) est vraie par simple application de la Proposition 2.3.1 : Si f est limite simple d'une suite de fonction étagées, alors elle est mesurable. Pour prouver le théorème, on va utiliser la méthode des approximations dyadiques tronquées. Il faut noter que c'est la première mise en oeuvre du principe de découpage suivant l'espace d'arrivée.

Preuve. Pour tout $n \in \mathbb{N}$, on définit la fonction suivante :

$$s_n(x) = \begin{cases} \frac{k}{2^n} & \text{si } f(x) \in \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right[, \text{ pour } 0 \leq k \leq n2^n - 1 \quad \text{(découpage)} \\ n & \text{si } f(x) \geq n \quad \text{(troncature)} \end{cases}$$

ce qui correspond à un découpage à l'arrivée de l'intervalle [0,n] en $n2^n$ intervalles de même longueur. Pour tout $n \geq 1$, la fonction s_n est étagée, et on a précisément

$$s_n = \sum_{0\leq k\leq n2^n-1} \frac{k}{2^n} \mathbb{1}_{\left\{\frac{k}{2^n} \leq f < \frac{k+1}{2^n}\right\}} + n\mathbb{1}_{\{f\geq n\}}$$

ce qui implique qu'elle est mesurable puisque f l'est.

Montrons d'abord que s_n est une suite croissante de fonctions (à ne pas confondre avec une suite de fonctions croissantes...). Soit $x \in X$ et $n \in \mathbb{N}$. Si $f(x) \geq n = \frac{n2^{n+1}}{2^{n+1}}$ alors $s_{n+1}(x)$ vaut soit $n + 1 \geq n$ soit $\frac{k2^{n+1}}{2^{n+1}}$ pour un $k \geq n$, ce qui implique dans les deux cas $s_{n+1}(x) \geq s_n(x)$. Si maintenant $f(x) < n$, alors il existe un entier k unique tel que

$$\frac{k}{2^n} \leq f(x) < \frac{k}{2^n} < n,$$

ce qui implique

$$\frac{2k}{2^{n+1}} \leq f(x) < \frac{2k+1}{2^{n+1}} < n \quad \text{ou} \quad \frac{2k+1}{2^{n+1}} \leq f(x) < \frac{2k+2}{2^{n+1}} < n.$$

Dans le premier cas, $s_{n+1} = \frac{2k}{2^{n+1}} = \frac{k}{2^n} = s_n(x)$, et dans le deuxième $s_{n+1} = \frac{2k+1}{2^{n+1}} > \frac{k}{2^n} = s_n(x)$. Dans tout les cas on a montré que $s_{n+1}(x) \geq s_n(x)$.

Etablissons ensuite la convergence simple de s_n vers f : Soit x fixé, alors — si $f(x) = +\infty$, on a $s_n(x) = n \to \infty = f(x)$ lorsque n tend vers l'infini, — si $f(x)$ est fini, alors pour $n > f(x)$ on a

$$f(x) - 1/2^n \leq s_n(x) < f(x), \tag{2.3}$$

donc $s_n(x)$ tend vers $f(x)$ par le théorème des gendarmes.

Supposons maintenant que f est bornée de borne M. Alors pour tout $n \geq M$, l'inégalité 2.3 est satisfaite pour tout les x, ce qui implique que

$$\sup_{x\in X} |s_n(x) - f(x)| \leq \frac{1}{2^n} \longrightarrow_{n\to\infty} 0$$

cela conclue la preuve du Théorème. $\square$

Chapitre 3

Mesures et intégration

Dans le Chapitre précédent nous avons répertorié les ensembles susceptibles d'être mesuré, ainsi que les fonctions mesurables. Il est temps maintenant de leur associer effectivement une quantité à l'aide des mesures.

3.1 Mesures

Définition 3.1.1 *Soit $(X, \mathcal{M})$ un espace mesurable. Une mesure positive sur X est une application $\mu : \mathcal{M} \longmapsto \overline{\mathbb{R}}_+$ telle que*

(*a*) $\mu(\emptyset) = 0$,

(*b*) *si $(A_n)_{n\in\mathbb{N}}$ est une suite d'éléments de $\mathcal{M}$ deux a deux disjoints, alors*

$$\mu(\cup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} \mu(A_n) \quad (\textit{additivité dénombrable}).$$

Le triplet $(X, \mathcal{M}, \mu)$, où simplement X s'il n'y a pas d'ambiguté sur la tribu et la mesure est alors appelé espace mesurable. De plus lorsque X est de mesure finie égale à 1, on dit que μ est une mesure de probabilité.

On peut donner quelques exemples simples de telles mesures :

Exemples 3.1.2

1. Si X est une ensemble fini, muni de la tribu $\mathcal{P}(X)$, on peut considérer la mesure μ_0 définie par
$$\mu_0(A) = \mathrm{Card}(A).$$
2. Si X est une ensemble fini non vide, muni de la tribu $\mathcal{P}(X)$, on peut considérer la mesure μ_1 définie par
$$\mu_1(A) = \mathrm{Card}(A)/\mathrm{Card}(X).$$
On vérifiera alors que μ_1 est une mesure de probabilité .
3. Une extension de la première mesure est la *mesure de comptage* : Si X est une ensemble quelconque, muni de la tribu $\mathcal{P}(X)$, on peut considérer la mesure μ définie par
$$\mu_2(A) = \begin{cases} \mathrm{Card}(A) \text{ si } A \text{ est fini ,} \\ \infty \text{ si } A \text{ est infini.} \end{cases}$$

4. Si X est une ensemble non vide, muni de la tribu $\mathcal{P}(X)$, et si a est un élément de X fixé, on peut considérer la *mesure de Dirac* δ_a, définie par
$$\delta_a(A) = \begin{cases} 1 \text{ si } a \in A, \\ 0 \text{ si } a \notin A. \end{cases}$$
On pourra remarquer que formellement au moins, on a la relation $\mu_2 = \sum_{a \in X} \delta_a$.
5. L'un des but du cours est de construire une mesure sur la tribu $\mathcal{B}(\mathbb{R})$ des boréliens de $\mathbb{R}$, qui sera appelée *mesure de Borel*, avec au moins la propriété naturelle suivante : Pour $a \leq b$, la mesure de l'intervalle $[a, b]$ est donnée par
$$\mu([a, b]) = b - a.$$
Il est facile d'étendre μ à une réunion finie d'intervalle, par contre l'extension à l'ensemble des boréliens de $\mathbb{R}$ est difficile et sera abordée au Chapitre 5.
6. On se donne une fonction ν continue et positive sur $\mathbb{R}$. On cherche alors à définir la *mesure de densité* μ_ν sur la tribu des boréliens qui, sur les intervalles, vaut
$$\mu_\nu(]a, b[) = \int_a^b \nu(t)dt.$$
L'intégrale étant celle de Riemann. Il s'agit en fait de l'extension de l'exemple précédent dans lequel on avait $\nu = 1$. De nouveau il est simple de définir μ_ν sur les réunions d'intervalles, et le prolongement à tout $\mathcal{B}(\mathbb{R})$ sera vu plus tard.
7. Nous construirons également la *mesure de Borel sur* $\mathbb{R}^d$. Il s'agit d'une mesure définie sur les Boréliens et qui sur les pavés $\prod_{1 \leq j \leq d} [a_j, b_j]$ vaut
$$\mu\left(\prod_{1 \leq j \leq d} [a_j, b_j]\right) = \prod_{1 \leq j \leq d} (b_j - a_j)$$
lorsque $a_j \leq b_j$ (0 sinon). C'est la version d-dimensionnelle de l'exemple 5.
8. *Probabilité de Cauchy sur* $\mathbb{R}$ de paramètre $\alpha > 0$. Il s'agit de la mesure positive de densité
$$\nu(t) = \frac{1}{\pi} \frac{\alpha}{\alpha^2 + t^2}.$$
C'est une mesure de Probabilité puisque
$$\mu(\mathbb{R}) = \int_{-\infty}^{+\infty} \nu(t) = \left[\frac{1}{\pi} \arctan(t/\alpha)\right]_{-\infty}^{+\infty} = 1.$$
On peut rappeler la définition de la fonction de répartition d'une mesure de probabilité sur (un sous-ensemble de) $\mathbb{R}$:
$$F_\nu(t) = \mu(] - \infty, t])$$
C'est une fonction définie sur $\mathbb{R}$, continue à droite, croissante de limite 0 en $-\infty$ et 1 en $+\infty$.
9. *Probabilité de Gauss* de moyenne m et écart-type σ. C'est la mesure positive de densité
$$\nu(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp -\frac{(x - m)^2}{2\sigma^2}.$$

10. *Probabilité de Bernouilli* de paramètre $p \in [0,1]$. C'est une mesure de probabilité sur l'ensemble $\{0,1\}$ définie par $\mu = p\delta_0 + (1-p)\delta_1$.
11. *Probabilité binomiale* de paramètres $n \in \mathbb{N}^*$ et $p \in [0,1]$. C'est une mesure de probabilité sur l'ensemble fini $\{0,1,...,n\}$ muni de la tribu de ses parties, et définie par
$$\mu = \sum_{0 \le k \le n} C_n^k p^k (1-p)^{n-k} \delta_k.$$
On pourra vérifier que pour $A \subset \mathcal{P}(\mathbb{N}_n)$,
$$\mu(A) = \sum_{k \in A} C_n^k p^k (1-p)^{n-k}.$$
12. *Probabilité de Poisson* de paramètre $\lambda > 0$. C'est la mesure de Probabilité sur $\mathbb{N}$ muni de la tribu de ses parties, et définie par
$$\mu = e^{-\lambda} \sum_{k \in \mathbb{N}} \frac{\lambda^k}{k!} \delta_k.$$

Après la revue de ces exemples, passons aux propriétés des mesures positives.

Proposition 3.1.3 *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré , où μ est une mesure positive. Les propriétés suivantes sont satisfaites :*

i) Additivité forte :
$$\forall A, B \in \mathcal{M}, \quad \mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B);$$
ii) Monotonie :
$$\forall A, B \in \mathcal{M}, \quad A \subset B \Longrightarrow \mu(A) \le \mu(B);$$
iii) σ-sous additivité :
$$\forall (A_n)_{n \in \mathbb{N}} \subset \mathcal{M}, \quad \mu(\cup_{n \in \mathbb{N}} A_n) \le \sum_{n \in \mathbb{N}} \mu(A_n);$$
iv) Continuité croissante : Pour toute suite croissante $(A_n)_{n \in \mathbb{N}}$ d'ensembles mesurables on a
$$\mu(\cup_{n \in \mathbb{N}} A_n) = \lim_{n \to \infty} \mu(A_n) = \sup_{n \ge 0} \mu(A_n);$$
v) Continuité décroissante : Pour toute suite décroissante $(A_n)_{n \in \mathbb{N}}$ d'ensembles mesurables, telle que $\mu(A_0) < \infty$, on a
$$\mu(\cap_{n \in \mathbb{N}} A_n) = \lim_{n \to \infty} \mu(A_n) = \inf_{n \ge 0} \mu(A_n).$$

Preuve. i) On a $A \cup B = (A \cap B^c) \cup_{\text{disj}} B$, donc $\mu(A \cup B) = \mu(A \cap B^c) + \mu(B)$ d'où
$$\mu(A \cup B) + \mu(A \cap B) = \underbrace{\mu(A \cap B^c) + \mu(A \cap B)}_{\mu(A)} + \mu(B)$$
d'ou le résultat.

ii) Provient directement du fait que $A \cup_{\text{disj}} (A \cap B^c) = B$.

iii) On définit une suite (B_n) d'ensembles de la manière suivante :

$$B_0 = A_0, \quad \text{et } \forall n \geq 1, \quad B_n = A_n \setminus \cup_{k \leq n-1} A_k.$$

Les ensembles B_n sont dans $\mathcal{M}$, deux à deux disjoints et on a $\cup_{n\in\mathbb{N}} A_n = \cup_{n\in\mathbb{N}} B_n$. On en déduit

$$\mu(\cup_{n\in\mathbb{N}} A_n) = \mu(\cup_{n\in\mathbb{N}} B_n) = \sum_{n\in\mathbb{N}} \mu(B_n) \leq \sum_{n\in\mathbb{N}} \mu(A_n)$$

puisque pour chaque $n \in \mathbb{N}$, $B_n \subset A_n$.

iv) On définit une suite (B_n) d'ensembles de la manière que pour iii) ce qui donne ici puisque la suite est croissante

$$B_0 = A_0, \quad \text{et } \forall n \geq 1, \quad B_n = A_n \setminus \cup_{k \leq n-1} A_k.$$

Les ensembles B_n sont dans $\mathcal{M}$, deux à deux disjoints et on a de plus pour tout $n \in \mathbb{N}$,

$$A_n = \cup_{k\leq n} A_k = \cup_{k \leq n} B_k$$

ainsi que $\cup_{n\in\mathbb{N}} A_n = \cup_{n\in\mathbb{N}} B_n$. On peut alors écrire

$$\mu(\cup_{n\in\mathbb{N}} A_n) = \mu(\cup_{n\in\mathbb{N}} B_n) = \sum_{n\in\mathbb{N}} \mu(B_n) = \lim_{n\to\infty} \sum_{k\leq n} \mu(B_k) = \lim_{n\to\infty} \mu(A_n) = \sup_{n\in\mathbb{N}} \mu(A_n).$$

v) On va se ramener au cas du iv). On pose $A = \cap A_n$ On a

$$A_0 \setminus (\cap A_n) = A_0 \cap (\cup A_n^c) = \cup_{n\in\mathbb{N}} \underbrace{(A_0 \cap A_n^c)}_{\text{suite croissante}}.$$

D'après iv) appliqué à la suite $(A_0 \cap A_n^c)$, on obtient que

$$\lim_{n\to\infty} \mu(A_0 \cap A_n^c) = \mu(A_0 \cap A^c) = \mu(A_0 \setminus A)$$

Par ailleurs A_0 est de mesure finie (c'est ici qu'on utilise l'hypothèse), donc pour chaque k on

$$\infty > \mu(A_0) = \mu(A_n) + \mu(A_0 \cap A_n^c), \quad \text{ainsi que } \infty > \mu(A_0) = \mu(A) + \mu(A_0 \cap A^c).$$

En particulier $\mu(A_n)$, $\mu(A_0 \cap A_n^c)$ et $\mu(A_0)$ sont des réels ce qui donne

$$\mu(A_n) = \mu(A_0) - \mu(A_0 \cap A_n^c) \overset{\text{en décroissant}}{\longrightarrow} \mu(A_0) - \mu(A_0 \cap A^c) = \mu(A).$$

Cela termine la preuve. $\square$

Remarque 3.1.4

1. Si la mesure est finie on peut appliquer la continuité décroissante sans restriction, puisque pour tout A_0 premier terme d'une suite, $\mu(A_0)$ est fini.
2. Réciproquement on peut étudier la suite d'ensembles suivante : soit pour tout $n \in \mathbb{N}$, $A_n = \{n, n+1, n+2, ...\} \subset \mathbb{N}$ muni de la tribu de ses parties et de la mesure de comptage. C'est une suite décroissante et on a

$$\infty = \lim \mu(A_n) \neq \mu(\cap A_n) = \mu(\emptyset) = 0$$

   La condition portant sur A_0 n'est pas satisfaite pour cette suite. On pourra vérifier qu'une condition nécessaire et suffisante est en fait qu'il existe n_0 tel que $\mu(A_{n_0}) < \infty$.

3.2 Intégration des fonctions positives

Le but de cette section est de définir l'intégrale par rapport à μ des fonctions mesurables positives. Dans un premier temps on va définir l'intégrale des fonctions étagées, qui ont été définies en 2.3.2. On pourra se rappeler que les ensembles mesurables intervenant dans la définition peuvent être assez compliques, et la tribu $\mathcal{M}$ très grosse (elle contient par exemple les réunions dénombrables d'intersections dénombrables d'ouverts....)

Lemme 3.2.1 *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Soit s une fonction étagée d'écriture canonique $s = \sum_{1\leq k\leq m} \alpha_j \mathbb{1}_{A_k}$ (voir la définition 2.3.2). On pose alors*

$$I(s) \stackrel{\text{def}}{=} \sum \alpha_k \mu(A_k), \tag{3.1}$$

(avec la convention $0 \times \infty = 0$). On a alors pour s, t étagées et $\lambda \geq 0$,

$$I(s) = \sup_{\sigma \text{ étagée }, \sigma \leq s} I(\sigma), \tag{3.2}$$

ainsi que

$$I(s+t) = I(s) + I(t), \quad I(\lambda s) = \lambda I(s), \quad \text{et} \quad s \leq t \Rightarrow I(s) \leq I(t) \tag{3.3}$$

Preuve. Montrons d'abord la première assertion. Considérons deux fonctions étagées σ et s telles que $\sigma \leq s$. On sait que l'on peut écrire

$$\sigma = \sum_{k\leq m} \beta_k \mathbb{1}_{B_k} \quad \text{et} \quad s = \sum_{j\leq n} \alpha_j \mathbb{1}_{A_j}$$

où (A_j) et (B_k) forment une deux partitions de X. En particulier on peut écrire que pour chaque $k \leq m$ on a

$$\mu(A_k) = \sum_{j\leq n} \mu(A_k \cap B_j),$$

d'où

$$I(\sigma) = \sum_{k\leq m} \beta_k \mu(B_k) = \sum_{k\leq m,\ j\leq n} \beta_k \mu(B_k \cap A_j)$$

On constate alors que soit $B_k \cap A_j = \emptyset$ qui est de mesure nulle, soit $B_k \cap A_j \neq \emptyset$, auquel cas pour $x \in B_k \cap A_j$ on a

$$\beta_k = \sigma(x) \leq s(x) = \alpha_j$$

ce qui donne

$$I(\sigma) \leq \sum_{k\leq m,\ j\leq n} \alpha_j \mu(B_k \cap A_j) = \sum_{j\leq n} \alpha_j \mu(A_j) = I(s)$$

puisque (B_k) est une partition de X.

Le dernier point de la proposition est immédiat compte tenu de la définition. Pour la propriété de multiplication par un scalaire, soit $\lambda \geq 0$. Si $\lambda = 0$ l'assertion est triviale, sinon lorsque $\lambda > 0$ on écrit que pour s étagée on a

$$I(\lambda s) = I\left(\lambda \sum_{j\leq n} \alpha_j \mathbb{1}_{A_j}\right) = I\left(\sum_{j\leq n} \lambda \alpha_j \mathbb{1}_{A_j}\right) = \sum_{j\leq n} \lambda \alpha_j \mu(A_j) = \lambda I(s)$$

d'après la définition de I.

La preuve de la propriété concernant la somme est à peine plus difficile. On considère s et t deux fonctions étagées. Alors la somme $s+t$ est également étagée (elle ne prend qu'un nombre fini de valeurs) et mesurable (comme somme de fonctions mesurables). Précisément on peut écrire

$$s = \sum_{j \le n} \alpha_j \mathbb{1}_{A_j}, \quad t = \sum_{k \le m} \beta_k \mathbb{1}_{B_k}, \qquad s+t = \sum_{j \le n, k \le m} (\alpha_j + \beta_k) \mathbb{1}_{A_j \cap B_k}$$

où pour s et t on a choisi l'écriture canonique. On obtient ainsi que

$$I(s) + I(t) = \sum_{j \le n} \alpha_j \mu(A_j) + \sum_{k \le m} \beta_k \mu(B_k)$$

Mais puisque A_k et B_j forment deux partitions de X on peut écrire

$$I(s) + I(t) = \sum_{k \le m,\ j \le n} (\alpha_j + \beta_k) \mu(B_k \cap A_j)$$

Pour calculer $I(s+t)$ il faut identifier l'écriture canonique de $s+t$. Indexons par $l \le p$ l'ensemble des valeurs différentes γ_l prises par $\alpha_j + \beta_k$ lorsque $j \le n$ et $k \le m$. On a alors

$$s + t = \sum_{l \le p} \gamma_l \mathbb{1}_{C_l}$$

où on a posé $C_l = (s+t)^{-1}(\{\gamma_l\})$. On a alors par définition

$$I(s+t) = \sum_{l \le p} \gamma_l \mu(C_l).$$

Mais par ailleurs

$$C_l = \bigcup_{\substack{j,k, \text{ tels que} \\ \alpha_j + \beta_k = \gamma_l}} A_j \cap B_k$$

où l'union est finie et disjointe, donc

$$\mu(C_l) = \sum_{\substack{j,k, \text{ tels que} \\ \alpha_j + \beta_k = \gamma_l}} \mu(A_j \cap B_k)$$

ce qui donne finalement

$$\begin{aligned} I(s+t) &= \sum_{l \le p} \sum_{\substack{j,k, \text{ tels que} \\ \alpha_j + \beta_k = \gamma_l}} \gamma_l \mu(A_j \cap \beta_k) = \sum_{l \le p} \sum_{\substack{j,k, \text{ tels que} \\ \alpha_j + \beta_k = \gamma_l}} (\alpha_j + \beta_k) \mu(A_j \cap \beta_k) \\ &= \sum_{j \le n,\ k \le m} (\alpha_j + \beta_k) \mu(A_j \cap \beta_k) = I(s) + I(t), \end{aligned}$$

d'où le résultat. Le Lemme est prouvé. □

Remarques 3.2.2

1. La quantité $I(s)$ a été définie en utilisant l'écriture canonique de s, en particulier pour chaque k, on a $A_k = s^{-1}(\alpha_k)$ avec des réels positifs α_k tous différents. On peut en fait montrer (le faire) sans difficulté que la formule de définition (3.1) est valable même si l'écriture de s n'est pas canonique.
2. On notera aussi que fonction nulle est étagée et que $I(0) = 0$. Une alternative pour éviter d'avoir affaire appel à la convention $0 \times \infty = 0$ est de ne considérer que les fonctions étagées non nulles et les $\alpha_k > 0$. La fonction nulle n'est alors plus étagée et on *pose* alors $I(0) = 0$.
3. On voit dans cette définition le point fondamental de la théorie de Lebesgue qui est de découper l'ensemble d'arrivée et d'y regarder les valeurs prises par la fonction (pour les fonctions étagées suivant les valeurs des α_j) puis de considérer les images réciproques (ici les A_j) par la fonction. C'est évidement une procédure différente de celle employée pour construire l'intégrale de Riemann, où on découpe d'abord l'espace de départ en intervalles, sur lesquels on regarde les valeurs prises par la fonction à intégrer. Nous verrons plus loin que l'intégrale de Lebesgue, construite par la première procédure solutionne un certain nombres de problèmes laissés ouverts par l'intégrale de Riemann.

On peut maintenant définir l'intégrale des fonctions mesurables de X dans $\overline{\mathbb{R}}_+$.

Définition 3.2.3 *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Pour $f : X \longmapsto \overline{\mathbb{R}}_+$ mesurable, on pose*

$$\int_X f d\mu = \sup_{s \text{ étagée, } s \leq f} I(\sigma)$$

Remarque 3.2.4

1. On observe évidement tout de suite que si f est étagée, on a
$$\int_X f d\mu = I(f)$$
au sens de la définition du lemme 3.2.1. Lorsqu'il n'y a pas d'ambiguité sur l'espace d'intégration, on pourra noter $\int f d\mu$ au lieu de $\int_X f d\mu$.
2. Pour $A \in \mathcal{M}$, on définit également l'intégrale sur A d'une fonction $f : X \longmapsto \overline{\mathbb{R}}_+$ mesurable, définie par
$$\int_A f d\mu = \int_X f \mathbb{1}_A d\mu.$$
3. Dans la définition 3.2.3, le sup peut être pris sur une suite *croissante* de fonctions mesurables étagées positives. C'est un résultat important en pratique que l'on donne dans le lemme qui suit. On rappelle également que la limite de telles suites caractérise justement les fonctions mesurables positives par le Théorème 2.3.3. Le résultat qui vient peut également être considéré comme un premier résultat d'interversion de limite et d'intégrale.

Lemme 3.2.5 *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Considérons une suite croissante (s_n) de fonctions étagées positives telles que* $\lim \uparrow s_n = f$ *(limite simple) avec donc $f : X \longmapsto \overline{\mathbb{R}}_+$ mesurable. On a alors*

$$\lim \uparrow \int s_n d\mu = \int \lim \uparrow s_n d\mu = \int f d\mu.$$

Preuve. On commence par mentionner que la suite numérique $(\int s_n d\mu)$ converge bien dans $\overline{\mathbb{R}}_+$, étant croissante d'après le Lemme 3.2.1. D'après la Définition 3.2.3, l'inégalité

$$\lim \uparrow \int s_n d\mu \leq \int f d\mu$$

est immédiate. Prouvons-la dans l'autre sens. Pour cela il suffit de montrer que si $g \leq f$ est une fonction étagée, alors

$$\int g d\mu \leq \lim \uparrow \int s_n d\mu. \tag{3.4}$$

On considère pour cela $0 < \varepsilon \leq 1$ un réel, et on pose pour tout $n \in \mathbb{N}$

$$A_n = \{s_n \geq (1-\varepsilon)g\} \stackrel{\text{def}}{=} \{x \in X;\ s_n(x) \geq (1-\varepsilon)g(x)\}.$$

Comme (s_n) est une suite croissante de fonctions, (A_n) définit une suite croissante d'ensembles mesurables de réunion X puisque s_n converge simplement vers $f \geq g$. On a également grace à la définition des A_n que pour tout n

$$s_n \geq (1-\varepsilon) g \mathbb{1}_{A_n}$$

ce qui implique par la propriété de croissance du Lemme 3.2.1 que

$$\int s_n d\mu \geq (1-\varepsilon) \int g \mathbb{1}_{A_n} d\mu.$$

Si maintenant g a pour écriture canonique $g = \sum_{1 \leq k \leq m} \alpha_k \mathbb{1}_{B_k}$ alors $g\mathbb{1}_{A_n}$, qui est également une fonction étagée s'écrit $\sum_{1 \leq k \leq m} \alpha_k \mathbb{1}_{B_k \cap A_n}$ et on obtient que

$$\int g \mathbb{1}_{A_n} d\mu = \sum_{k \leq m} \mu(A_n \cap B_k) \longrightarrow \sum_{k \leq m} \mu(B_k) = \int g d\mu \text{ lorsque } n \to \infty,$$

où la convergence vient de la propriété de continuité croissante de la mesure μ, point iv) de la Proposition 3.1.3. On a ainsi obtenu que

$$\lim \uparrow \int s_n d\mu \geq (1-\varepsilon) \int g d\mu$$

ce qui implique (3.4) puisque ε est arbitrairement petit. Le lemme est donc prouvé. □

Exemples 3.2.6 Il est utile de reprendre les exemples du paragraphe précédent, et les intégrales correspondantes.

1. On considère $X = \{x_1, ..., x_n\}$ un ensemble fini, muni de la tribu $\mathcal{P}(X)$ et de la mesure définie par $\mu_0(A) = \mathrm{Card}(A)$ pour $A \subset X$. Toutes les fonctions sont alors étagées et mesurables, et on obtient pour une telle fonction f

$$\int f d\mu_0 = \int \sum_{j=1}^{n} f(x_j) \mathbb{1}_{\{x_j\}} d\mu_0 = \sum_{j=1}^{n} f(x_j).$$

2. Sur le même espace (de cardinal n) muni de la mesure μ_1 définie par $\mu_1(A) = \text{Card}(A)/n$. on obtient

$$\int f d\mu_1 = \int \sum_{j=1}^{n} f(x_j) \mathbb{1}_{\{x_j\}} d\mu_0 = \sum_{j=1}^{n} f(x_j)/n.$$

3. Pour la mesure de comptage μ_2, définie sur $X = \{x_i, i \in I\}$ quelconque muni de la tribu $\mathcal{P}(X)$, on obtient de même

$$\int f d\mu_2 = \sum_{i \in I} f(x_i) \stackrel{\text{def}}{=} \sup_{J \text{ fini } \subset I} \sum_{i \in J} f(x_i)$$

4. Si X est une ensemble non vide, muni de la tribu $\mathcal{P}(X)$, et si a est un élément de X fixé, on a lorsque μ est la mesure de Dirac δ_a,

$$\int f d\mu = f(a)$$

5. Pour la mesure de Borel m sur $\mathbb{R}^d$, que nous verrons au Chapitre 5, on notera l'intégrale par le même signe que l'intégrale de Riemann

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

   On verra que cette intégrale coincide avec l'intégrale de Riemann pour les fonctions continues à support compact. Elle permettra cependant aussi d'intégrer des fonctions beaucoup plus compliquées, par exemple

$$\int_{\mathbb{R}^d} \mathbb{1}_{\mathbb{Q}}(x) dx = m(\mathbb{Q}) = 0.$$

6. 7. Losque μ est une mesure de densité ν par rapport à la mesure de Borel, on a

$$\int_{\mathbb{R}^d} f d\mu = \int_{\mathbb{R}^d} f(x) \nu(x) dx$$

   ce qui justifie la notation infinitésimale $d\mu = \nu dx$.

8. 9. Les mesures de probabilité de Cauchy (sur $\mathbb{R}$) et de Laplace-Gauss (sur $\mathbb{R}^d$) entrent dans la catégorie précédente avec respectivement les densités $\nu(t)dt$ et $\nu(x)dx$ où ν est défini respectivement par

$$\nu(t) = \frac{1}{\pi} \frac{\alpha}{\alpha^2 + t^2} \qquad \text{et } \nu(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp -\frac{(x-m)^2}{2\sigma^2}.$$

10. 11. 12. Les probabilités de Bernouilli, binomiale et de Poisson sont construites à partir de mesures de Dirac sur des entiers et on obtient sans peine que

$$\int f d\mu = \begin{cases} pf(0) + (1-p)f(1) & \text{Bernouilli, paramètre } p \\ \sum_{0 \le k \le n} C_n^k p^k (1-p)^{n-k} f(k) & \text{binomiale, paramètres } (n,p), \\ e^{-\lambda} \sum_{k \in \mathbb{N}} \frac{\lambda^k}{k!} f(k) & \text{Poisson, paramètre } \lambda. \end{cases}$$

On passe maintenant en revue quelques propriétés fondamentales de l'intégrale d'une fonction positive. Compte tenu de leur importance, les Théorèmes de convergence seront vus dans un chapitre dédié.

Proposition 3.2.7 *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. On a alors pour f et $g : X \longmapsto \overline{\mathbb{R}}_+$ des fonctions mesurables et $\lambda \geq 0$,*

$$\int \lambda f d\mu = \lambda \int f d\mu, \quad \int (f+g) d\mu = \int f d\mu + \int g d\mu \quad \text{et } \ f \leq g \Rightarrow \int f d\mu \leq \int g d\mu.$$

(On rappelle que l'on utilise toujours la convention $0 \times \infty = 0$).

Preuve. Compte tenu du travail précédent la preuve est assez courte. On considère deux suites de fonctions étagées (s_n) et (t_n) convergeant en croissant vers respectivement f et g. C'est possible grace au Lemme 2.3.3. On a alors

$$\lim \uparrow (s_n + t_n) = f + g, \quad \text{ et } \lim \uparrow \lambda s_n = \lambda f.$$

On obtient ainsi grace au Lemmes 3.2.5 et 3.2.1 que

$$\int \lambda f d\mu = \lim \uparrow \int \lambda s_n = \lim \uparrow \lambda \int s_n d\mu = \lambda \int f d\mu.$$

De la même manière on peut écrire que

$$\int (f+g) d\mu = \lim \uparrow \int (s_n + t_n) = \lim \uparrow (\int s_n d\mu + \int t_n d\mu) = \int f d\mu + \int g d\mu.$$

Le dernier point est une conséquence directe du premier. Cela conclue la preuve de la Proposition. $\square$

Chapitre 4

Théorèmes de convergence et fonctions intégrables

Dans le Chapitre précédent, on a pu définir l'intégrale d'une fonction mesurable à valeur dans $\overline{\mathbb{R}}_+$. Nous allons principalement dans ce chapitre établir des Théorèmes d'interversion d'intégrale et de limite sous des hypothèses faible de convergence. C'est un des atouts majeurs de l'intégrale de Lebesgue.

4.1 Théorèmes de convergence

Le premier théorème que nous présentons concerne les suites croissantes de fonctions et est à mettre en parallèle avec le Lemme 3.2.5.

Théoreme 4.1.1 Théorème de convergence monotone dit de Beppo-Levi. *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Considérons une suite croissante (f_n) de fonctions mesurables de X dans $\overline{\mathbb{R}}_+$ telles que* $\lim \uparrow f_n \overset{\text{def}}{=} f$ *(limite simple), avec donc $f : X \longmapsto \overline{\mathbb{R}}_+$ mesurable. On a alors*

$$\lim \uparrow \int f_n d\mu = \int \lim \uparrow f_n d\mu = \int f d\mu.$$

Preuve. Un examen attentif de l'énoncé aura permit de se rendre compte qu'il est identique a celui du Lemme 3.2.5 sauf que l'on a remplacé la suite (s_n) par ici la suite (f_n). Pour la preuve de ce théorème il en est exactement de même, quitte à faire appel à la Proposition 3.2.7 au lieu du Lemme 3.2.1 quand il y est fait référence. □

On pourra remarquer que l'hypothèse de convergence est réduite a la convergence simple. Evidement sans l'hypothèse de croissance le résultat est faux (voir plus loin le contre exemple donné en 4.1.6).

On donne maintenant un corollaire concernant la convergence des séries de fonctions positives.

Corollaire 4.1.2 *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Considérons une suite (f_n) de fonctions mesurables de X dans $\overline{\mathbb{R}}_+$. On pose pour tout $x \in X$, $S(x) =$*

$\sum_{n\geq 0} f_n(x)$. Alors S est mesurable positive et on a

$$\sum_{n\in\mathbb{N}} \int f_n d\mu = \int \sum_{n\in\mathbb{N}} f_n d\mu = \int S d\mu.$$

Preuve. La mesurablité de S est une conséquence de la Proposition 2.3.1, puisque S est limite simple de la suite des fonctions sommes partielles définies pour tout $n \in \mathbb{N}$ par

$$S_n(x) = \sum_{k\leq n} f_k(x).$$

Bien sur chaque S_n est mesurable comme somme finie de fonctions mesurables en vertu de la Proposition 2.2.9. Comme $\lim \uparrow S_n = S$, on peut appliquer le Théorème de Beppo-Levi 4.1.1 qui donne

$$\int S d\mu \overset{B.L.}{=} \lim_{n\to\infty} \int S_n d\mu = \lim_{n\to\infty} \int \sum_{k\leq n} f_k d\mu \overset{\text{Prop. 3.2.7}}{=} \lim_{n\to\infty} \sum_{k\leq n} \int f_k d\mu = \sum_{k\geq 0} \int f_k d\mu.$$

Les égalités ci dessus on lieu dans $\overline{\mathbb{R}}_+$ et la Proposition 3.2.7 est utilisée pour écrire que l'intégrale d'une somme finie est la somme des intégrales. Le résultat est prouvé. $\square$

Une deuxième application du Théorème de Beppo-Levi concerne les mesures à densité.

Corollaire 4.1.3 *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Considérons une fonction ν mesurable de X dans $\overline{\mathbb{R}}_+$. Pour tout $A \in \mathcal{M}$, on définit $\lambda(A) = \int_A \nu d\mu$. Alors λ est une mesure positive définie sur $\mathcal{M}$, et si f est une fonction mesurable positive, on a*

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

On notera $d\lambda = \nu d\mu$ et on dira que $d\lambda$ est la mesure de densité ν par rapport à la mesure $d\mu$.

Remarque 4.1.4

1. On prendra garde au fait que le produit des deux fonctions mesurables $f\nu$ ci dessus est bien défini dans $\overline{\mathbb{R}}_+$ avec la convention $0.\infty = 0$.
2. Cette construction répond aux exemples 6 et 7 de la série d'exemples 3.1.2 et 3.2.6 : une mesure étant construite (ici μ), il est direct de construire toute une famille de mesures, dites à densité par rapport à la première. Dans le cas réel, deux mesures étant données, les rapports entre l'une et l'autre (par exemple l'une est-elle a densité par rapport à l'autre ?) font l'objet d'un théorème difficile, le Théorème de Radon Nikodym, vu en Master.

Preuve du Corollaire. Montrons que λ est une mesure. D'après la Proposition 3.2.7, on a trivialement $\lambda(\emptyset) = \int_\emptyset \nu d\mu = 0$. Considérons maintenant une suite d'éléments deux à deux disjoints $(A_n)_{n\in\mathbb{N}}$ de $\mathcal{M}$. On a alors d'après le Corollaire 4.1.2

$$\lambda(\cup_{n\in\mathbb{N}} A_n) = \int_{\cup_{n\in\mathbb{N}} A_n} \nu d\mu = \int_X \sum_{n\in\mathbb{N}} \nu.\mathbb{1}_{A_n} d\mu \overset{\text{Cor 4.1.2}}{=} \sum_{n\in\mathbb{N}} \int_X \nu \mathbb{1}_{A_n} d\mu = \sum_{n\in\mathbb{N}} \lambda(A_n).$$

On en déduit donc que λ est une mesure.

Montrons maintenant la deuxième partie du Corollaire. Considérons d'abord une fonction étagée $s = \sum_{j=1}^n \alpha_j \mathbb{1}_{A_j}$, où on peut supposer que les α_j sont tous strictement positifs. Alors

$$\int_X s d\lambda \stackrel{\text{def}}{=} \sum_{j\leq n} \alpha_j \lambda(A_j) = \sum_{j\leq n} \int_X \nu \mathbb{1}_{A_j} d\mu = \int_X \sum_{j\leq n} \nu \mathbb{1}_{A_j} d\mu = \int_X s\nu d\mu \tag{4.1}$$

d'après la Proposition 3.2.7 donnant l'intégrale d'une somme finie de fonctions. Dans le cas général, pour f fonction mesurable de X dans $\overline{\mathbb{R}}_+$, soit (s_k) une suite croissante de fonctions étagée convergeant en croissant vers f, donnée par le Théorème 2.3.3. On a alors également

$$\lim \uparrow s_k.\nu = f.\nu$$

et on peut alors écrire

$$\int_X f d\lambda \stackrel{B.L.}{=} \lim_k \int_X s_k d\lambda \stackrel{s_k \text{ étagée}}{=} \lim_k \int_X s_k \nu d\mu \stackrel{B.L.}{=} \int_X f.\nu d\mu.$$

□

Lemme 4.1.5 Lemme de Fatou. *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Considérons une suite (f_n) de fonctions mesurables de X dans $\overline{\mathbb{R}}_+$.On a alors*

$$\int (\liminf_n f_n) d\mu \leq \liminf_n \left(\int f_n d\mu \right). \tag{4.2}$$

Preuve. On rappelle d'abord que d'après la Proposition 2.3.1, la fonction $f = \liminf_n f_n$ est mesurable positive. Par ailleurs, rappelons que $\liminf_n f_n = \sup_{n\in\mathbb{N}} \inf_{k\geq n} f_n$ par définition et que donc

$$f = \sup_{n\in\mathbb{N}} g_n \text{ où } g_n = \inf_{k\geq n} f_n.$$

Les fonctions g_n sont mesurables de X dans $\overline{\mathbb{R}}_+$ d'après la Proposition 2.3.1, et on a

$$\lim \uparrow g_n = f.$$

Le Théorème de Beppo-Levi donne donc

$$\lim \uparrow \int g_n d\mu = \int f d\mu. \tag{4.3}$$

Par ailleurs de la Proposition 3.2.7, on a pour tout n

$$g_n \leq f_n \Longrightarrow \int_X g_n \leq \int_X f_n,$$

d'où l'on déduit

$$\liminf_n \int_X g_n \leq \liminf \int_X f_n,$$

et donc le résultat d'après (4.3), puisque $\liminf g_n d\mu = \lim \uparrow \int g_n d\mu = \int f d\mu$. □

Exemple 4.1.6 L'inégalité dans le Lemme de Fatou peut être stricte, comme le montre l'exemple suivant. Considérons $X =]0, \infty[$ et définissons pour tout $n \in \mathbb{N}$,

$$f_n(x) = ne^{-nx}, \quad \text{pour } x > 0.$$

Alors on a $\int_X f_n(x)dx = 1$ et pourtant pour tout $x > 0$, $\lim_{n\to\infty} f_n(x) = 0$ qui est donc d'intégrale nulle. L'inégalité (4.2) du Lemme s'écrit donc $0 \leq 1$ et est donc stricte.

On peut remarquer que cet exemple constitue également une illustration de la nécessité de l'hypothèse de croissance dans le Théorème de Beppo-Levi. En effet la suite (f_n) converge simplement vers 0 mais pas en croissant, ce qui explique que la conclusion du Théorème de Beppo-Levi ne soit pas vérifiée (elle se traduit ici par $0 \neq 1$, mais elle pourrait ne même pas avoir de sens dans le cas général). $\square$

Nous allons maintenant donner un sens à l'intégrale de fonctions complexes, et pas seulement positives :

Définition-Théorème 4.1.7 *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. On dira qu'une fonction f mesurable de X dans $\mathbb{C}$ est* intégrable *si*

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

L'ensemble des fonctions intégrables est noté $\mathcal{L}^1(\mu)$ et l'intégrale d'une fonction $f \in \mathcal{L}^1(\mu)$ est définie par

$$\int f d\mu = \int_X (\mathrm{Re}\, f)_+ d\mu - \int_X (\mathrm{Re}\, f)_- d\mu + i \int_X (\mathrm{Im}\, f)_+ d\mu - i \int_X (\mathrm{Im}\, f)_- d\mu. \tag{4.4}$$

L'espace $\mathcal{L}^1(\mu)$ est une espace vectoriel sur $\mathbb{C}$ et $f \mapsto \int_X f d\mu$ est une forme linéaire sur cet espace.

Preuve. La formule (4.4) a bien un sens puisque les inégalités $|(\mathrm{Re}\, f)_\pm| \leq |f|$ et $|(\mathrm{Im}\, f)_\pm| \leq |f|$ impliquent par la Proposition 3.2.7 que chaque terme de (4.4) est fini, et donc que la somme est un élément de $\mathbb{C}$.

Montrons ensuite que $\mathcal{L}^1(\mu)$ est un espace vectoriel : Pour cela soient f et $g \in \mathcal{L}^1(\mu)$ et α, $\beta \in \mathbb{C}$. on a alors $|\alpha f + \beta g| \leq |\alpha||f| + |\beta||g|$ et la Proposition 3.2.7 implique que ce dernier élément est d'intégrale finie. Donc $\alpha f + \beta g \in \mathcal{L}^1(\mu)$.

Montrons enfin que la prise de l'intégrale est une forme linéaire sur $\mathcal{L}^1(\mu)$. Soient $f = f_1 + if_2$ et $g = g_1 + ig_2$ deux fonctions de $\mathcal{L}^1(\mu)$ décomposées en partie réelle et imaginaires. On souhaite d'abord montrer que

$$\int (f+g) d\mu = \int f d\mu + \int g d\mu. \tag{4.5}$$

Pour montrer cela, on écrit d'après la définition de l'intégrale

$$\mathrm{Re} \int (f+g) d\mu = \int (f_1 + g_1)_+ d\mu - \int (f_1 + g_1)_- d\mu. \tag{4.6}$$

Mais par ailleurs

$$\mathrm{Re}\,(f+g) = (f_1+g_1)_+ - (f_1+g_1)_- = (f_1)_+ - (f_1)_- + (g_1)_+ - (g_1)_-,$$

d'ou cette égalité entre fonctions positives (éventuellement infinies) :

$$(f_1+g_1)_+ + (f_1)_- + (g_1)_- = (f_1+g_1)_- + (f_1)_+ + (g_1)_+.$$

On peut appliquer la Proposition 3.2.7 sur l'intégrale de la somme de fonctions positives, qui donne

$$\int (f_1+g_1)_+ d\mu + \int (f_1)_- d\mu + \int (g_1)_- d\mu = \int (f_1+g_1)_- d\mu + \int (f_1)_+ d\mu + \int (g_1)_+ d\mu.$$

D'après 4.6 on obtient

$$\begin{aligned}\mathrm{R}e \int (f+g) d\mu &= \int (f_1)_+ d\mu + \int (g_1)_+ d\mu - \int (f_1)_- d\mu + \int (g_1)_- d\mu \\ &\overset{\text{def}}{=} \int (\mathrm{Re}\, f) d\mu + \int (\mathrm{Re}\, g) d\mu,\end{aligned}$$

où la dernière égalité provient de la définition de l'intégrale. Exactement de la même manière, on montre que

$$\mathrm{I}m \int (f+g) d\mu = \int (\mathrm{Im}\, f) d\mu + \int (\mathrm{Im}\, g) d\mu.$$

Ces deux résultats donnent

$$\begin{aligned}\int (f+g) d\mu &= \int (\mathrm{Re}\, f) d\mu + \int (\mathrm{Re}\, g) d\mu + i \int (\mathrm{Im}\, f) d\mu + i \int (\mathrm{Im}\, g) d\mu \\ &= \int f d\mu + \int g d\mu,\end{aligned}$$

où de nouveau la dernière égalité provient de la définition de l'intégrale d'une fonction dans $\mathcal{L}^1(\mu)$. L'égalité (4.5) est montrée.

Il reste à étudier la multiplication par un scalaire. Soit $f = f_1 + if_2 \in \mathcal{L}^1(\mu)$ et $\alpha = \alpha_1 + i\alpha_2 \in \mathbb{C}$. D'après le résultat précédent sur la somme de deux (donc quatre) fonctions, on peut écrire

$$\int \alpha f d\mu = \int \alpha_1 f_1 - \int \alpha_2 f_2 d\mu + \int i\alpha_1 f_2 d\mu + \int i\alpha_2 f_1 d\mu.$$

D'après la Proposition 3.2.7, on a $\int \alpha_1 f_1 d\mu = \alpha_1 \int f_1 d\mu$ lorsque $\alpha_1 \geq 0$ et lorsque $\alpha_1 < 0$ il suffit d'écrire

$$\int \alpha_1 f_1 d\mu \overset{\text{def.}}{=} -\int (-\alpha_1 f_1) d\mu = -\int (-\alpha_1) f_1 d\mu \overset{\text{Prop.3.2.7}}{=} -(-\alpha_1) \int f_1 d\mu = \alpha_1 \int f_1 d\mu.$$

Il reste à montrer que $\int i f_1 d\mu = i \int f_1 d\mu$ mais ceci est une conséquence immédiate de la définition. On a donc

$$\int \alpha f d\mu = \alpha \int f d\mu$$

et ceci conclue la preuve du Théorème. □

On termine cette Section par un des Théorèmes les plus important de la théorie de l'intégration de Lebesgue.

Théoreme 4.1.8 . Théorème de convergence dominée dit de Lebesgue. *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Considérons une suite (f_n) de fonctions mesurables de X dans $\mathbb{C}$ telles que* $\lim_{n\to\infty} f_n = f$ *(limite simple), avec donc f mesurable. Si il existe une fonction $g : X \mapsto \overline{\mathbb{R}}_+$ intégrable telle que*

$$\forall x \in X,\ n \in \mathbb{N}, \qquad |f_n(x)| \leq g(x), \tag{4.7}$$

alors f est intégrable et on a

$$\lim_{n\to\infty} \int |f_n - f| d\mu = 0, \qquad et \qquad \lim_{n\to\infty} \int f_n d\mu = \int f d\mu.$$

Preuve. Quitte à séparer partie réelle et partie imaginaire, la mesurabilité de f est une conséquence de la Proposition 2.3.1. Par ailleurs puisque $|f| \leq g$ on obtient que f est intégrable grace au dernier point de la Proposition 3.2.7. De cette Proposition de nouveau et de la linéarité prouvée au Théorème-définition 4.1.7, on obtient que pour tout $n \in \mathbb{N}$,

$$\left| \int f_n d\mu - \int f d\mu \right| \leq \int |f_n - f| d\mu.$$

La dernière assertion est donc une conséquence de l'avant-dernière. On va maintenant appliquer le lemme de Fatou à la suite de fonctions

$$h_n = 2g - |f_n - f| \overset{n\to\infty}{\longrightarrow} 2g \quad \text{(limite simple)}.$$

On écrit

$$\begin{aligned}\int 2g d\mu = \int \liminf (2g - |f_n - f|) d\mu &\overset{Fatou}{\leq} \liminf \int (2g - |f_n - f|) d\mu \\ &= \int 2g d\mu - \limsup \int |f_n - f| d\mu.\end{aligned}$$

On en déduit donc que

$$\limsup \int |f_n - f| d\mu = 0,$$

et par positivité

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

le résultat est prouvé. □

4.2 Fonctions Intégrables et ensembles de mesure nulle

Les ensembles de mesure nulle sont transparents en théorie de l'intégration de Lebsgue. C'est ce que nous allons préciser dans cette Section. On commence par une définition, ou plutôt une terminologie.

Définition 4.2.1 *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. On dira d'une propriété qu'elle est* vraie μ-presque partout, *ou* presque partout *s'il n'y a pas d'ambiguité sur la mesure lorsqu'elle est vraie en dehors d'un ensemble de mesure nulle.*

On notera pp. en abrégé. En particulier pour f et g deux fonctions mesurables, on dira que $f = g$ pp. si

$$\mu(\{x \in X;\ f(x) \neq g(x)\}) = 0.$$

L'égalité presque partout de deux fonctions est une relation d'équivalence

Preuve du dernier point. Il s'agit juste de vérifier les trois points caractérisant une relation d'équivalence. Pour f, g et h mesurables, on a $f = f$ *pp.*, et

$$f = g \text{ pp.} \Longleftrightarrow \mu(\{f \neq g\}) = 0 \Longleftrightarrow \mu(\{g \neq f\}) = 0 \Longleftrightarrow g = f \text{ pp.}$$

Par ailleurs si $f = g$ *pp.* et $g = h$ *pp.*, alors on déduit de l'inclusion

$$\{f \neq h\} \subset \{f \neq g\} \cup \{g \neq h\}$$

que

$$\mu(\{f \neq h\}) \leq \mu(\{f \neq g\}) + \mu(\{g \neq h\}) = 0,$$

d'où le résultat. □

Exemples 4.2.2 On sait que $\mathbb{Q}$ est dénombrable donc de mesure nulle pour la mesure de Borel que nous construirons au prochain chapitre. On obtient ainsi que

$$1_{\mathbb{Q}} = 0 \text{ pp.} \quad \text{et} \quad 1_{\mathbb{R}/\mathbb{Q}} = 1 \text{ pp.}$$

Donnons un autre exemple concernant des ensembles : A et B sont deux ensembles mesurables, la propriété $A \subset B$ *pp.* signifie $\mu(A \cap B^c) = 0$.

Concernant l'intégration des fonctions positives on a le résultat suivant :

Proposition 4.2.3 *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive, et soient f et g deux fonctions mesurables de X dans $\overline{\mathbb{R}}_+$. On a alors*

i) $f = 0$ *pp.* $\Longleftrightarrow \int f d\mu = 0$;

ii) $f = g$ *pp.* $\Longrightarrow \int f d\mu = \int g d\mu$;

iii)$f \leq g$ *pp.* $\Longrightarrow \int f d\mu \leq \int g d\mu$;

iv) $\int f d\mu < \infty \Longrightarrow f < \infty$ *pp.*

Preuve. La preuve utilise le résultat suivant qui peut avoir un intérêt en soit : pour tout $\alpha > 0$, on a

$$\mu(\{f > \alpha\}) \leq \frac{1}{\alpha} \int f d\mu. \tag{4.8}$$

En effet, posons $E_\alpha = \{f > \alpha\}$, alors $f \geq \alpha \mathbb{1}_{E_\alpha}$ et donc d'après la proposition 3.2.7

$$\int f d\mu \geq \int \alpha \mathbb{1}_{E_\alpha} d\mu \overset{\text{Prop.3.2.1}}{=} \alpha \int \mathbb{1}_{E_\alpha} d\mu = \alpha \mu(E_\alpha),$$

d'où le résultat. Montrons maintenant le point i). Si $\int f d\mu = 0$, alors pour tout $n \in \mathbb{N}^*$ on a

$$\mu(E_{1/n}) \leq n \int f d\mu = 0.$$

Or

$$\{f > 0\} = \cup_{n\in\mathbb{N}^*} E_{1/n}$$

et l'union étant dénombrable on obtient $\mu(\{f > 0\}) = 0$ d'où le résultat. Réciproquement si $\mu(E) = 0$ avec $E = \{f > 0\}$ alors pour toute fonction étagée $s \leq f$, on a $s = 0$ sur E^c. Si s admet pour écriture $s = \sum_{j=1}^{n} \alpha_j \mathbb{1}_{A_j}$, alors pour tout j, $A_j \subset E$ et donc $\mu(A_j) = 0$. Cela donne $\int s d\mu = I(s) = 0$ d'après la définition 3.1 et $\int f d\mu = 0$ d'après la défintion de l'intégrale 3.2.3. On pourra remarquer que le résultat reste vrai même si f prend la valeur $+\infty$ là où elle n'est pas nulle.

Montrons maintenant le point ii). On pose cette fois $E = \{f > g\}$ et l'énoncé implique $\mu(E) = 0$. Par ailleurs on a l'inégalité immédiate suivante entre fonctions positives :

$$f \leq g + (f - g)\mathbb{1}_E$$

qui implique par la Proposition 3.2.7 que

$$\int f d\mu \leq \int g d\mu + \int (f - g)\mathbb{1}_E d\mu.$$

Or $0 \leq (f - g)\mathbb{1}_E = 0$ *pp.* donc cette fonction est d'intégrale nulle d'après i), et on on obtient le résultat.

Le point iii) est une application immédiate du point ii) et du fait que $f = g$ *pp.* est impliqué par $f \leq g$ *pp.* et $g \leq f$ *pp.*.

Pour le dernier point iv), on pose $E = \{f = \infty\}$ et on remarque que $\mu(E) > 0$ implique que pour tout $n \in \mathbb{N}^*$, on a

$$\int f d\mu \geq n\mu(\{f > n\}) \geq n\mu(E) \stackrel{n\to\infty}{\longrightarrow} +\infty$$

d'après (4.8), d'où le résultat. □

Avec cette notion, on peut légèrement affaiblir les hypothèses du Théorème de convergence dominée, en ignorant les ensembles de mesure nulle :

Théoreme 4.2.4 . Théorème de convergence dominée dit de Lebesgue (version II). *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Considérons une suite (f_n) de fonctions mesurables de X dans $\mathbb{C}$ telles que*

$$\textit{presque pour tout } x \in X, \quad \lim_{n\to\infty} f_n(x) = f(x).$$

Si il existe une fonction $g : X \mapsto \overline{\mathbb{R}}_+$ intégrable telle que

$$\textit{presque pour tout } x \in X, \quad n \in \mathbb{N}, \qquad |f_n(x)| \leq g(x), \tag{4.9}$$

alors f (qui n'est en général définie que presque partout) est intégrable et on a

$$\lim_{n\to\infty} \int |f_n - f| d\mu = 0, \qquad \textit{et} \qquad \lim_{n\to\infty} \int f_n d\mu = \int f d\mu.$$

Preuve. Il s'agit juste dans un premier temps de préciser l'énoncé ; la fonction f définie presque partout peut être définie partout de la façon suivante : Soit N défini par son complémentaire

$$N^c \stackrel{\text{def}}{=} \{x \in X,\ f_n(x) \text{ converge }\}.$$

L'ensemble N est de mesure nulle. On pose alors

$$f \stackrel{\text{def}}{=} \lim_{n\to\infty} \mathbb{1}_{N^c} f_n$$

ce qui donne la convergence simple de la série de fonctions $(\mathbb{1}_{N^c} f_n)_{n\in\mathbb{N}}$ vers f, et en particulier que f est mesurable. On a évidement $|\mathbb{1}_{N^c} f_n| \leq g$ et on peut donc appliquer le Théorème de Lebesgue déjà énoncé 4.1.8 qui donne que f est intégrable et que

$$\lim_{n\to\infty} \int |\mathbb{1}_{N^c} f_n - f| d\mu = 0, \quad \text{et} \quad \lim_{n\to\infty} \int \mathbb{1}_{N^c} f_n d\mu = \int f d\mu.$$

Evidement les fonctions $|\mathbb{1}_{N^c} f_n - f|$ et $|f_n - f|$ diffèrent sur un ensemble de mesure nulle, ce qui donne l'égalité des intégrales correspondantes d'après la Proposition 4.2.3 . Il en est de même pour les intégrales $\int \mathbb{1}_{N^c} f_n d\mu$ et $\int f_n d\mu$, grace à la linéarité de l'intégrale et la Proposition 4.2.3. On obtient ainsi le résultat en remarquent enfin que f n'a également besoin de n'être connue qu'en dehors d'un ensemble de mesure nulle (ici N), et que les intégrales la faisant intervenir ne changent pas si on lui assigne sur cet ensemble une autre valeur, toujours d'après la Proposition 4.2.3. □

Remarque 4.2.5 Les autres énoncés de ce chapitre concernant les interversion limite-intégrale peuvent eux aussi admettre leur version II, où les ensembles de mesure nulle sont négligés. On laisse le soin au lecteur d'écrire leurs énoncés.

Pour finir le chapitre nous introduisons maintenant une notion élargie d'ensemble mesurable qui trouvera une application naturelle avec la mesure de Lebesgue. On commence par une définition.

Définition 4.2.6 *On dit que l'espace mesuré $(X, \mathcal{M}, \mu)$ est complet si toute partie d'un ensemble de mesure nulle est mesurable (et donc également de mesure nulle).*

En général ce n'est pas forcément le cas. Cependant si un espace mesurable n'est pas complet, il est possible de le completer de la manière suivante. On définit la tribu complétée $\mathcal{M}'$ comme suit : une partie E de X appartient à $\mathcal{M}'$ si il existe deux ensembles A et $B \in \mathcal{M}$ tels que

$$A \subset E \subset B \quad \text{et} \quad \mu(B/A) = 0. \tag{4.10}$$

On prolonge alors la mesure μ à $\mathcal{M}'$ de la manière suivante

$$\mu'(E) \stackrel{\text{def}}{=} \mu(A). \tag{4.11}$$

On a alors

Proposition 4.2.7 *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré, alors le triplet $(X, \mathcal{M}', \mu')$ défini par* (4.10) *et* (4.11) *est un espace mesuré complet.*

Preuve. Montrons que $\mathcal{M}'$ satisfait les axiomes d'une tribu. On a directement $\emptyset \in \mathcal{M} \subset \mathcal{M}'$. Par ailleurs si $E \in \mathcal{M}'$ et A et B sont définis par (4.10) alors on a

$$B^c \subset E^c \subset A^c \quad \text{et} \quad \mu(A^c/B^c) = \mu(B/A) = 0,$$

puisque une inspection rapide donne $B/A = A^c/B^c$. L'ensemble de parties $\mathcal{M}'$ est donc stable par prise du complémentaire. Regardons la stabilité par union dénombrable, dernier point à vérifier d'après la définition d'une tribu 2.1.1. Soit (E_n) une suite d'ensembles de $\mathcal{M}'$. Pour tout $n \in \mathbb{N}$, il existe des ensembles A_n et B_n de $\mathcal{M}$ tels que

$$A_n \subset E_n \subset B_n \quad \text{et} \quad \mu(B_n/A_n) = 0.$$

On a alors

$$\cup_{n\in\mathbb{N}} A_n \subset \cup_{n\in\mathbb{N}} E_n \subset \cup_{n\in\mathbb{N}} B_n \quad \text{et} \quad \mu(\cup_{n\in\mathbb{N}} B_n / \cup_{n\in\mathbb{N}} A_n) = 0,$$

puisque

$$\cup_{n\in\mathbb{N}} B_n / \cup_{n\in\mathbb{N}} A_n = \cup_{n\in\mathbb{N}} (B_n/A_n),$$

qui est donc réunion dénombrable d'ensemble de mesure nulle et donc de mesure nulle. $\mathcal{M}'$ est donc une tribu. On laisse au lecteur le soin de vérifier que μ' est alors une mesure sur $(X, \mathcal{M}')$. □

Remarque 4.2.8 On pourra remarquer que la tribu complétée est exactement la tribu engendrée par la tribu elle-même et les sous-ensembles d'ensembles de mesure nulle.

Chapitre 5

La mesure de Lebesgue sur $\mathbb{R}$

Plusieurs constructions de l'intégrale sont possibles. Nous présentons ici une construction ensembliste qui répond aux exemples 5 et 6 des remarques 3.1.2 et 3.2.6. En particulier on souhaite construire une mesure λ sur l'ensemble des boréliens de $\mathbb{R}$ qui soit telle que

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

pour tout $a \leq b \in \mathbb{R}$. Quelques résultats sur la construction ne seront qu'énoncés, et on renvoie par exemple pour les preuves complètes au livre de J. Faraut [2].

5.1 Un théorème de prolongement

On commence par une définition

Définition 5.1.1 *Soit X un ensemble et $\mathcal{U}$ une famille de parties de X. On dit que $\mathcal{U}$ est une algèbre de Boole sur X si*

$$\begin{array}{ll} (a) & A \in \mathcal{U} \Longrightarrow A^c \in \mathcal{U}, \\ (b) & A, B \in \mathcal{U} \Longrightarrow A \cup B \in \mathcal{U}, \\ (c) & X \in \mathcal{M}. \end{array}$$

On appellera ensemble élémentaire un élément de l'algèbre et fonction élémentaire une combinaison linéaire de fonctions indicatrices d'ensembles élémentaires.

On peut tout de suite voir que la définition d'une algèbre est proche de celle d'une tribu. Seule la stabilité par union dénombrable a été remplacée par la stabilité par union finie. Un exemple fondamental pour la suite est donné par les intervalles de $\mathbb{R}$:

Exemple 5.1.2 L'ensemble des réunions finies d'intervalles de $\mathbb{R}$ est une algèbre de Boole $\mathcal{U}$, la vérification des points a, b, et c étant immédiate. De plus en appliquant la Proposition 2.1.6 et la définition d'une tribu, on en déduit que $\sigma(\mathcal{U}) = \mathcal{B}(\mathbb{R})$, tribu des boréliens de $\mathbb{R}$.

De même dans $\mathbb{R}^d$, la tribu engendrée par les réunions finies de pavés (c'est une algèbre de Boole) est la tribu des boréliens de $\mathbb{R}^d$.

On énonce maintenant un Théorème de prolongement.

Théoreme 5.1.3 *Soit X un ensemble et $\mathcal{U}$ une algèbre de Boole sur cet ensemble. Soit $\mu : \mathcal{U} \longmapsto \overline{\mathbb{R}}_+$ une application vérifiant la propriété suivante : si $(A_n)_{n\in\mathbb{N}}$ est une suite d'éléments de $\mathcal{U}$ deux à deux disjoints, telle que $\cup_{n\in\mathbb{N}} A_n \in \mathcal{U}$ alors*

$$\mu(\cup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} \mu(A_n) \qquad (\textit{additivité dénombrable}).$$

Si en plus μ est σ-finie (i.e. il existe une suite d'ensembles X_n de réunion X tels que pour tout $n \in \mathbb{N}$, $\mu(X_n) < \infty$), alors μ se prolonge en une mesure unique (encore notée μ) sur la tribu $\mathcal{M}$ engendrée par $\mathcal{U}$.

Idée de la preuve. Pour l'unicité, considérons deux mesures μ_1 et μ_2 prolongeant μ sur l'espace mesurable $(\mathbb{X}, \mathcal{M})$. Alors on vérifie que l'ensemble

$$\{A \subset X;\ \mu_1(A) = \mu_2(A)\}$$

est une tribu qui contient l'ensemble $\mathcal{U}$ et sur laquelle μ_1 et μ_2 coincident. Comme la tribu engendrée par $\mathcal{U}$ est $\sigma(\mathcal{U}) = \mathcal{M}$, les mesures μ_1 et μ_2 coincident aussi sur $\mathcal{M}$. Elles sont donc égales.

Donnons une idée de la preuve de l'existence du prolongement. C'est une démonstration assez longue, qui est basée sur une définition de la mesure extérieure : Pour $A \subset X$, on définit

$$\mu^*(A) = \inf_{\substack{(A_n)_{n\in\mathbb{N}} \subset \mathcal{U} \\ A \subset \cup A_n}} \sum_{n\in\mathbb{N}} \mu(A_n).$$

Dans cette définition les A_n forment donc un recouvrement de A par des ensembles élémentaires. On vérifie que lorsque A est un ensemble élémentaire on a

$$\mu^*(A) = \mu(A),$$

puis on sélectionne parmi toutes les parties de X celles qui sont proches des ensembles élémentaires au sens suivant : On dira que $A \subset X$ est *intégrable* s'il existe une suite (A_n) d'ensembles élémentaires tels que

$$\mu^*(A_n \Delta A) \longrightarrow 0 \tag{5.1}$$

où pour chaque $n \in \mathbb{N}$, $A_n \Delta A \stackrel{\text{def}}{=} (A_n/A) \cup (A/A_n)$ est appelée différence symétrique de A_n et A. Bien sur les ensmbles élémentaires sont intégrables. Il est aisé de montrer que si A et B sont des ensembles intégrables, alors $A \cup B$, $A \cap B$, A/B le sont aussi. On a également la propriété de fermeture suivante : Si $A \subset X$ est tel qu'il existe une suite (A_n) d'ensembles intégrables tels que

$$\mu^*(A_n \Delta A) \longrightarrow 0,$$

alors A est intégrable. En fait on a le résultat suivant :

> *L'ensemble $\mathcal{M}'$ des ensembles intégrables est une tribu, et la fonction μ' restriction de μ^* à $\mathcal{M}'$ est une mesure sur $\mathcal{M}'$.*

La fin est alors directe. Puisque $\mathcal{M}'$ et une tribu qui contient les ensembles élémentaires, alors elle contient aussi $\mathcal{M} = \sigma(\mathcal{U})$. La mesure μ est alors définie comme la restriction à $\mathcal{M}$ de μ'. $\square$

Remarque 5.1.4

1. D'après la définition d'ensemble intégrable, et donc a fortiori pour un ensemble $A \in \mathcal{M}$, on a
$$\mu(A) = \mu^*(A) = \inf_{\substack{(A_n)_{n\in\mathbb{N}} \subset \mathcal{U} \\ A \subset \cup A_n}} \sum_{n\in\mathbb{N}} \mu(A_n).$$
En particulier pour tout $\varepsilon > 0$, il existe d'après la définition de la borne inférieure une suite (A_n) d'ensembles intégrables tels que $A \subset \cup_{n\in\mathbb{N}} A_n$ et $\sum_{n\in\mathbb{N}} \mu(A_n) \leq \mu(A) + \varepsilon$. En fait directement de (5.1) on sait qu'il existe un ensemble élémentaire A_ε proche de A au sens où $\mu(A_\varepsilon \Delta A) \leq \varepsilon$.
2. Considérons maintenant une fonction intégrable f pour la mesure μ. On peut alors l'approcher par une fonction élémentaire. Soit en effet $\varepsilon > 0$. D'après le Théorème de convergence dominée et le Théorème 2.3.3, on sait qu'il existe une fonction étagée s telle que $\int |f - s| d\mu \leq \varepsilon/2$. Par ailleurs une fonction étagée est combinaison linéaire finie de fonctions indicatrices d'ensembles mesurables, donc d'après la remarque précédente, il existe une fonction élémentaire g telle que $\int |s - g| d\mu \leq \varepsilon/2$. On a ainsi obtenu $\int |f - g| d\mu \leq \varepsilon$ (g est proche de f *en moyenne*).

5.2 La mesure de Lebesgue sur $\mathbb{R}$

Dans cette section nous allons appliquer le Théorème de prolongement 5.1.3 aux objets définis dans l'exemple 5.1.2, c'est à dire à

— $\mathcal{U}$ l'algèbre de Boole engendrée par les intervalles (i.e. les réunions finies d'intervalles) ;

— la fonction $\lambda : \begin{cases} \mathcal{U} & \longrightarrow & \overline{\mathbb{R}}^+ \\ \cup_{j=1}^n (a_j, b_j) & \longmapsto & \sum_{j=1}^n (b_j - a_j) \end{cases}$,

où pour les intervalles (a_j, b_j) les parenthèses remplacent soient "[" soit "]". Le résultat est alors le suivant :

Théoreme 5.2.1 *Il existe une unique mesure λ sur $\mathcal{B}(\mathbb{R})$ telle que $\lambda([a, b]) = b - a$ pour tout $a \leq b \in \mathbb{R}$.*

Idée de la preuve. Nous sommes dans la situation présentée dans le Théorème de prolongement 5.1.3. On remarque d'abord que la tribu engendrée par les intervalles est justement la tribu des Boréliens, d'après la Proposition 2.1.6. Pour pouvoir prolonger λ, vérifions qu'elle satisfait les hypothèses du Théorème. On remarque d'abord qu'elle est σ-finie. En effet on peut écrire
$$\mathbb{R} = \cup_{n\in\mathbb{N}} [-n, n], \quad \text{avec} \quad \lambda([-n, n]) = 2n < \infty.$$
Donnons quelques éléments de la preuve de l'additivité dénombrable.

Considérons d'abord $I_1, I_2, ..., I_N$ des intervalles disjoints de reunion I intervalle. Alors de la définition et en réordonnant les intervalles on a

$$\sum_{n=1}^{N} \lambda(I_n) = \lambda(I). \tag{5.2}$$

Avec un peu de travail il est possible de montrer que cette relation s'étend aux suites d'intervalles $(I_n)_{n\in\mathbb{N}}$ disjoints de reunion I également :

$$\sum_{n=1}^{\infty} \lambda(I_n) = \lambda(I). \tag{5.3}$$

Ce résultat étant établi, passons aux ensembles élémentaires. Soit donc (A_n) une suite d'éléments élémentaires de réunion A un ensemble élémentaire. Puisque chaque A_n est réunion finie d'intervalles, on peut supposer que les A_n sont chacun des intervalles en utilisant (5.2). De même A est un ensemble élémentaire, donc de la forme $A = \cup_{k=1}^{K} I_k$, où les intervalles I_k sont disjoints. On a alors d'après (5.3) pour $k \leq K$,

$$\lambda(I_k) = \sum_{n=1}^{\infty} \lambda(I_k \cap A_n),$$

mais également pour tout $n \in \mathbb{N}$,

$$\lambda(A_n) = \sum_{k=1}^{K} \lambda(I_k \cap A_n)$$

d'où $\lambda(A) = \sum_{n=1}^{\infty} \lambda(A_n)$ par regroupement. □

Remarque 5.2.2 Dans le cas particulier de $\mathbb{R}$ muni de la tribu de ses Boréliens et de la mesure λ la remarque 5.1.4 implique les résultats suivant : Si A est un Borélien de $\mathbb{R}$ et $\varepsilon > 0$, alors il existe une suite (I_n) d'intervalles tels que $A \subset \cup_{n\in\mathbb{N}} I_n$ et $\sum_{n\in\mathbb{N}} \lambda(I_n) \leq \lambda(E) + \varepsilon$. De plus si f est intégrable et $\varepsilon > 0$, il existe une fonction en escalier telle que

$$\int |f - g| d\lambda \leq \varepsilon.$$

(Rappelons qu'une fonction en escalier est de la forme $g = \sum_{j=1}^{N} \alpha_j \mathbb{1}_{(x_{j-1}, x_j)}$, où les x_j sont ordonnés).

Parmi toutes les mesures définies sur la tribu des Boréliens (il y en a beaucoup, par exemple les mesures de Dirac, de comptage ...) la mesure de Lebesgue est la seule a preserver la propriété d'invariance par translation de la tribu :

Proposition 5.2.3 *La tribu borélienne $\mathcal{B}(\mathbb{R})$ est invariante par translation. De plus si μ est une mesure sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ invariante par translation et pour laquelle la mesure d'un intervalle borné est finie, alors elle est proportionnelle à la mesure de Lebesgue.*

Preuve. On sait que l'ensemble des intervalles est invariant par translation (i.e. si I est un intervalle, alors pour $a \in \mathbb{R}$, l'ensemble $a + I \stackrel{\text{def}}{=} \{a + x;\ x \in I\}$ est également un intervalle. On en déduit que si $\mathcal{I}$ est l'ensemble des intervalles,

$$\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{I}) = \sigma(a + \mathcal{I}) = a + \sigma(\mathcal{I}) = a + \mathcal{B}(\mathbb{R}).$$

L'invariance par translation de la mesure de Lebesgue est claire d'après sa construction et sa valeur sur les intervalles.

Considérons maintenant une mesure μ satisfaisant les hypothèses de la Proposition et posons $c = \mu([0,1])$. Des propriétés d'invariance par translation et d'additivité de la mesure on obtient pour tout $n \in \mathbb{N}^*$, $\mu([0,1/n]) = c/n$ ainsi que le fait que la mesure d'un singleton est nulle. On en déduit alors que pour tout a, b rationnels $\mu([a,b]) = c(b-a)$, et la Proposition vient du fait que les intervalles de ce type engendrent la tribu des boréliens, d'après la (preuve de la) Proposition 2.1.6. □

Remarque 5.2.4 On appelle souvent mesure de Lebesgue la mesure λ complétée sur la tribu complétée de celle des boréliens, i.e. celle a laquelle on a rajouté à tous les ensembles des parties d'ensembles de mesure nulle (cf. définition 4.2.7).

Remarque 5.2.5 On peut également montrer le résultat suivant que nous ne faisons qu'énoncer : si A est mesurable, alors

i) $\lambda(A) = \inf\{\lambda(U);\ U \text{ ouvert },\ A \subset U\}$;

ii) $\lambda(A) = \sup\{\lambda(K);\ K \text{ compact },\ K \subset A\}$.

5.3 Intégrale de Riemann et intégrale de Lebesgue

On rappelle d'abord une construction de l'intégrale au sens de Riemann sur un intervalle $[a,b]$. Elle est basée sur l'approximation des fonctions par des fonctions en escalier. Soit u une telle fonction définie sur $[a,b]$. Alors il existe une subdivision

$$a = x_0 < x_1 < ... < x_n = b$$

et des réels $\alpha_0, ..., \alpha_{n-1}$ telle que pour tout $x \in]x_j, x_{j+1}[$ on ait $u(x) = \alpha_j$. On définit alors l'intégrale de Riemann de u par

$$\int_a^b u(x)dx = \sum_{j=1}^{n-1} \alpha_j (x_{j+1} - x_j).$$

On vérifie que cette définition est indépendante de la subdivision choisie, que l'intégrale ainsi définie est linéaire sur l'espace des fonctions en escalier. Pour une fonction f définie sur $[a,b]$ à valeurs réelles et bornée, on pose

$$\begin{aligned} I_R(f) &= \sup\left\{\int_a^b u(x)dx\ ;\ u \text{ en escalier },u \leq f\right\} \\ J_R(f) &= \inf\left\{\int_a^b v(x)dx\ ;\ v \text{ en escalier },v \geq f\right\} \end{aligned} \tag{5.4}$$

La fonction f est dite intégrable au sens de Riemann si $I_R(f) = J_R(f)$. Dans ce cas on appelle intégrale de Riemann de f le nombre

$$\int_a^b f(x)dx \stackrel{\text{def}}{=} I_R(f) = J_R(f).$$

Pour qu'une fonction soit intégrable au sens de Riemann, il faut et il suffit que pour tout $\varepsilon > 0$, il existe deux fonctions en escalier u et v telles que

$$u \leq f \leq v \quad \text{et} \quad \int (v(x) - u(x))dx \leq \varepsilon.$$

On passe maintenant à l'intégrale de Lebesgue sur l'intervalle $[a, b]$. On considère pour cela la tribu de Borel sur $[a, b]$ complétée, muni de la mesure de Lebesgue. Pour une fonction f définie sur $[a, b]$ à valeurs réelles et bornées, on pose

$$\begin{aligned} I_L(f) &= \sup \left\{ \int_a^b u(x)dx \ ; \ u \text{ étagée }, u \leq f \right\} \\ J_L(f) &= \inf \left\{ \int_a^b v(x)dx \ ; \ v \text{ étagée }, v \geq f \right\}. \end{aligned} \tag{5.5}$$

Lorsque f est mesurable positive, on reconnait la définition de l'intégrale $\int f d\lambda = I_L(f)$ donnée en 3.2.3. En fait avec un peu de travail on peut montrer que si f est mesurable bornée alors on a $\int f d\lambda = I_L(f) = J_L(f)$. On peut même montrer la réciproque , à savoir que si $I_L(f) = J_L(f)$ pour une fonction réelle bornée, alors elle est mesurable, et intégrable d'intégrale $I_L(f)$. On retrouve donc une situation tout a fait identique à celle proposée ci dessus pour l'intégrale de Riemann :

Pour qu'une fonction soit (mesurable et) intégrable au sens de Lebesgue, il faut et il suffit que pour tout $\varepsilon > 0$, il existe deux fonctions étagées u et v telles que

$$u \leq f \leq v \quad \text{et} \quad \int_{[a,b]} (v - u)d\lambda \leq \varepsilon.$$

On peut alors comparer les deux intégrales. La remarque principale est la suivante :

Les fonction en escalier sont étagées.

Alors il est clair que lorsque u est en escalier, on a

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

On obtient donc que pour toute fonction f réelle bornée on a

$$I_R(f) \leq I_L(f) \leq J_L(f) \leq J_R(f).$$

On en déduit donc que si f est intégrable au sens de Riemann, alors elle est mesurable et intégrable au sens de Lebesgue et que les deux intégrales coincident.

Pour finir on peut mentionner qu'il existe des fonctions qui sont intégrables au sens de Lebesgue et pas au sens de Riemann. C'est le cas de la fonction indicatrice $\mathbb{1}_{\mathbb{Q}}$. En effet pour cette fonction sur l'intervalle $[0, 1]$ on a

$$0 = I_R(f) < I_L(f) = J_L(f) = J_R(f) = 1,$$

par simple inspection et grace au fait que $\mathbb{1}_{\mathbb{Q}} = 1$ *pp.*

Chapitre 6

Espaces de fonctions intégrables

6.1 Les espace L^p

Dans toute la suite, $\mathbb{K}$ désigne le corps $\mathbb{R}$ ou $\mathbb{C}$. On commence par une définition.

Définition 6.1.1 *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Soit $1 \leq p \leq \infty$ et $f : X \to \mathbb{K}$ une application mesurable. On dira que $f \in \mathcal{L}^p(\mu)$ si $\int |f|^p d\mu < \infty$ pour $p < \infty$ et, pour $p = \infty$, si il existe $M > 0$ tel que $\mu\{x \in X;\ |f(x)| > M\} = 0$.*

Les fonctions dans ces espaces peuvent avoir une structure compliquée. Les ensembles de mesure nulle ont un rôle particulier qui sera précisé plus loin.

Exemple 6.1.2 Considérons la fonction f définie par

$$\begin{array}{rccl} f: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & \left\{ \begin{array}{ll} 1 & \text{si } x \notin \mathbb{Q} \\ x & \text{si } x \in \mathbb{Q} \end{array} \right. \end{array}.$$

On vérifie d'abord que cette fonction est mesurable sur la tribu des boréliens de $\mathbb{R}$ en appliquant la Proposition 2.2.7 : Pour $a < 1$ on a $f^{-1}(]a, +\infty]) = \mathbb{Q} \cup]a, +\infty[$ et lorsque $a \geq 1$ on obtient $f^{-1}(]a, +\infty]) = \mathbb{Q}^c \cap]a, +\infty[$. Les ensembles obtenus sont des boréliens (intersection ou union de boréliens) ce qui donne le résultat.

On constate par ailleurs que f est non bornée. Montrons cependant qu'elle appartient à $\mathcal{L}^\infty(\lambda)$, où λ est la mesure de Lebesgue. En effet on a

$$\{x \in \mathbb{R} \,;|\ f(x)| > 1\} = \mathbb{Q} \cap [-1, 1]^c,$$

donc en prenant la mesure de cet ensemble on obtient

$$\lambda \{x \in \mathbb{R} \,;|\ f(x)| > 1\} \leq \lambda(\mathbb{Q}) = 0.$$

(On rappelle que $\mathbb{Q}$ est une réunion dénombrable de singletons, qui sont de mesure nulle, donc est lui-même de mesure nulle). On en déduit que $f \in \mathcal{L}^\infty(\lambda)$ (ici $M = 1$ convient).

Proposition 6.1.3 *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Soient f et g deux fonctions mesurables et $p \in]1, +\infty[$. Considérons p' l'exposant conjugué de p*

définit par $1/p + 1/p' = 1$. Alors on a les deux inégalités suivantes :

$$i) \quad \int |fg| d\mu \le \left(\int |f|^p d\mu\right)^{1/p} \left(\int |g|^{p'} d\mu\right)^{1/p'} \qquad (\text{Hölder}),$$

$$ii) \quad \left(\int |f+g|^p d\mu\right)^{1/p} \le \left(\int |f|^p d\mu\right)^{1/p} + \left(\int |g|^p d\mu\right)^{1/p} \qquad (\text{Minkowski}).$$

Preuve. i) Montrons d'abord l'inégalité de Hölder. Il suffit de montrer l'inégalité dans le cas où f et g sont à valeur dans $\mathbb{R}_+$ puisque ce sont les modules qui interviennent. On peut supposer également que $\int |f|^p d\mu > 0$ et $\int |g|^{p'} d\mu > 0$. En effet si l'un de ces deux termes est nul, alors d'après la Proposition 4.2.3 on a $f = 0$ pp. ou $g = 0$ pp. et donc $fg = 0$ pp. ce qui implique toujours par la Proposition 4.2.3 que $\int |fg| d\mu = 0$. L'inégalité se réduit alors à $0 = 0$ et est donc vraie. On peut enfin supposer que $\int |f|^p d\mu < \infty$ et $\int |g|^p d\mu < \infty$ sinon l'inégalité est évidente puisque le terme de droite est infini. On pose alors

$$A = \left(\int |f|^p d\mu\right)^{1/p}, \quad B = \left(\int |g|^{p'} d\mu\right)^{1/p'}, \quad F = f/A, \quad \text{et} \quad G = g/B.$$

On vérifie que $\int F^p d\mu = \int G^{p'} d\mu = 1$ grace à la normalisation. Nous allons utiliser l'inégalité de convexité élémentaire suivante :

$$\forall x, y \in \mathbb{R}_+, \quad \alpha \in]0,1[, \quad x^\alpha y^{1-\alpha} \le \alpha x + (1-\alpha) y,$$

qui se prouve en prenant le logarithme de l'inégalité et en utilisant la concavité du Logarithme. Dans notre cas on prend $\alpha = 1/p$, $1 - \alpha = 1/p'$ et on applique l'inégalité à $x = F^p(t)$ et $y = G^{p'}(t)$ pour chaque $t \in X$. On obtient l'inégalité entre fonctions

$$FG = (F^p)^{1/p} (G^{p'})^{1/p'} \le \frac{1}{p} F^p + \frac{1}{p'} G^{p'},$$

ce qui donne en intégrant

$$\int FG d\mu \le \frac{1}{p} \int F^p d\mu + \frac{1}{p'} \int G^{p'} d\mu \le \frac{1}{p} + \frac{1}{p'} = 1.$$

Finalement $\int fg d\mu \le AB$ ce qui est le résultat cherché.

ii) Montrons maintenant l'inégalité de Minkowski. On commence par remarquer que l'inégalité n'a d'intérêt que si f et g sont dans $\mathcal{L}^p(\mu)$, sinon le membre de droite est infini (et le membre de gauche n'est pas forcément défini). Dans ce cas la Proposition 4.2.3 implique que les fonctions f et g sont finies presque partout et que donc la somme $f + g$ est définie et finie presque partout. On peut en particulier intégrer $|f + g|$. Par ailleurs l'inégalité $|f + g| \le ||f| + |g||$ implique qu'il suffit de montrer l'inégalité lorsque f et g sont à valeurs positives. On a alors

$$\begin{aligned}\int (f+g)^p d\mu &= \int (f+g)(f+g)^{p-1} d\mu \\ &= \int f(f+g)^{p-1} d\mu + \int g(f+g)^{p-1} d\mu,\end{aligned}$$

d'où en appliquant deux fois l'inégalité de Hölder avec $\alpha = 1/p$

$$\begin{aligned}\int (f+g)^p d\mu &\le \left(\int f^p d\mu\right)^{1/p} \left(\int (f+g)^{(p-1)p'} d\mu\right)^{1/p'} + \left(\int g^p d\mu\right)^{1/p} \left(\int (f+g)^{(p-1)p'} d\mu\right)^{1/p'} \\ &\le \left(\left(\int f^p d\mu\right)^{1/p} + \left(\int g^p d\mu\right)^{1/p}\right) \left(\int (f+g)^p d\mu\right)^{1/p'},\end{aligned} \tag{6.1}$$

où pour obtenir la dernière inégalité on a utilisé le fait que $(p-1)p' = \frac{p-1}{1-1/p} = p$. Regardons le résultat suivant la valeur du membre de gauche $C = \int (f+g)^p d\mu$. Si $C = 0$ alors l'inégalité de Minkowski est prouvée. On remarque ensuite que l'inégalité de convexité

$$\left(\frac{f+g}{2}\right)^p \le \frac{f^p + g^p}{2}$$

et le fait que f et g appartiennent à $\mathcal{L}^p(\mu)$ impliquent que $C < \infty$ en intégrant. On peut donc simplifier l'expression 6.1 et on obtient le résultat. □

Les inégalités précédentes sont fondamentales dans le cadre des espaces L^p que nous introduisons maintenant :

Définition-Théorème 6.1.4 *Pour $p \in [1, +\infty]$ on définit $L^p(\mu) = \mathcal{L}^p(\mu)/\sim$ où $\sim$ est la relation d'équivalence définie par $f \sim g$ si $f = g$ pp. (c'est l'égalité presque partout introduite en 4.2.1).*

$L^p(\mu)$ est un espace vectoriel normé pour la norme définie par

$$\begin{aligned}\|f\|_{L^p} &= \left(\int |\tilde{f}|^p d\mu\right)^{1/p} \quad \text{si } p \in [1, +\infty[, \\ \|f\|_{L^\infty} &= \inf\left\{M \in \mathbb{R}_+;\ \mu\left(|\tilde{f}| > M\right) = 0\right\},\end{aligned} \tag{6.2}$$

où dans les membres de droite $\tilde{f}$ est n'importe quel représentant dans $\mathcal{L}^p(\mu)$ de $f \in L^p(\mu)$. On identifiera souvent f et $\tilde{f}$.

Preuve. Le fait que $\sim$ soit une relation d'équivalence a déjà été prouvé en 4.2.1. Les quantités introduites en (6.2) sont clairement indépendantes du représentant choisi d'après la Proposition 4.2.3. C'est encore une illustration du fait que les ensembles de mesure nulle sont transparents en théorie de l'intégration. Montrons finalement que les espaces définis sont bien des espaces vectoriels.

Considérons d'abord α, β sont deux scalaires et f, g deux éléments de $L^p(\mu)$ de représentants $\tilde{f}$ et $\tilde{g}$. Alors $\tilde{f}$ et $\tilde{g}$ sont finies presque partout d'après la Proposition 4.2.3

point iv) si $p < \infty$, et d'après la définition dans la cas $p = \infty$. Cela permet de définir la somme $\alpha\tilde{f} + \beta\tilde{g}$ presque partout, et donc la classe de fonctions $\alpha f + \beta g$ comme l'ensemble des fonctions qui lui sont égales presque partout. Cette somme étant définie, vérifions la fin du Théorème.

Cas $1 \leq p < \infty$: On vient de voir que $\alpha\tilde{f} + \beta\tilde{g}$ était fini presque partout. On peut alors appliquer l'inégalité de Minkowski qui donne ici

$$\left(\int |\alpha\tilde{f} + \beta\tilde{g}|^p d\mu\right)^{1/p} \leq \left(\int |\alpha\tilde{f}|^p d\mu\right)^{1/p} + \left(\int |\beta\tilde{g}|^p d\mu\right)^{1/p}$$

Le membre de droite étant fini, on en déduit que $\alpha\tilde{f} + \beta\tilde{g} \in \mathcal{L}^p(\mu)$ et que donc $\alpha f + \beta g \in L^p(\mu)$. On applique alors la définition 6.2 avec $\alpha = \beta = 1$, ce qui donne l'inégalité triangulaire

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

On a également pour $\alpha \in \mathbb{C}$

$$\|\alpha f\|_{L^p} = \left(\int |\alpha\tilde{f}|^p d\mu\right)^{1/p} = |\alpha| \left(\int |\tilde{f}|^p d\mu\right)^{1/p} = |\alpha| \, \|f\|_{L^p}$$

d'après la linéarité de l'intégrale de fonctions positives prouvée en Proposition 3.2.7. Considérons enfin une (classe de) fonction $f \in L^p(\mu)$ telle que $\|f\| = 0$. Alors

$$\|f\|_{L^p} = 0 \Longleftrightarrow \int |\tilde{f}|^p d\mu = 0 \overset{\text{Prop. 4.2.3}}{\Longleftrightarrow} \tilde{f} = 0 \; pp. \Longleftrightarrow f = 0 \text{ dans } L^p(\mu).$$

On a donc montré que $\text{Ł}^p(\mu)$ est un espace vectoriel dans le cas $p < \infty$.

Cas $p = \infty$: Considérons d'abord $f \in L^\infty(\mu)$ et $\tilde{f}$ un représentant de f dans $\mathcal{L}^\infty$. Alors $\forall k \in \mathbb{N}^*$ on a

$$\mu\left(x \in \mathbb{R}, \; |\tilde{f}(x)| > 1/k + \|f\|_{L^\infty}\right) = 0,$$

puisque par définition $\|f\|_{L^\infty} = \inf\left\{M \text{ tel que } \mu(|\tilde{f}| > M) = 0\right\}$ et que donc $M = \|f\|_{L^\infty} + 1/k$ convient. Comme

$$\left\{|\tilde{f}| > \|f\|_{L^\infty}\right\} = \cup_{k \in \mathbb{N}^*} \left\{|\tilde{f}| > \|f\|_{L^\infty} + 1/k\right\},$$

on obtient par sous-additivité dénombrable que

$$\mu\left(|\tilde{f}| > \|f\|_{L^\infty}\right) \leq \sum_{k \in \mathbb{N}^*} \mu\left(|\tilde{f}| > \|f\|_{L^\infty} + 1/k\right) = 0.$$

Cela implique le résultat suivant intéressant en soit :

$$\tilde{f}(x) \leq \|f\|_{L^\infty} \;\; pp. x \in X \tag{6.3}$$

Montrons que $L^\infty(\mu)$ est un espace vectoriel : soient $f, g \in L^\infty(\mu)$, $\tilde{f}, \tilde{g}$ des représentants, alors on a l'inclusion d'ensembles

$$\left\{|\tilde{f}| \leq \|f\|_{L^\infty}\right\} \cap \{|\tilde{g}| \leq \|g\|_{L^\infty}\} \subset \left\{|\tilde{f} + \tilde{g}| \leq \|f\|_\infty + \|g\|_\infty\right\},$$

donc en passant au complémentaire :

$$\{|\tilde{f}| > \|f\|_{L^\infty}\} \cup \{|g| > \|g\|_{L^\infty}\} \supset \{|\tilde{f} + \tilde{g}| > \|f\|_{L^\infty} + \|g\|_\infty\}.$$

L'ensemble de gauche est de mesure nulle donc

$$\mu\left(|\tilde{f} + \tilde{g}| > \|f\|_{L^\infty} + \|g\|_{L^\infty}\right) = 0.$$

On en déduit donc que $f + g \in L^\infty(\mu)$ et que

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

C'est l'inégalité triangulaire. On prouve de manière immédiate que si $\alpha \in \mathbb{C}$, on a $\alpha f \in L^\infty(\mu)$ et que donc $\text{Ł}^\infty(\mu)$ est une espace vectoriel. Précisément on a

$$\begin{aligned}
\|\alpha f\|_{L^\infty} &= \inf\left\{M \in \mathbb{R}_+ \text{ tel que } \mu\left(|\alpha\tilde{f}| > M\right) = 0\right\} \\
&= \inf\left\{|\alpha| N \in \mathbb{R}_+ \text{ tel que } \mu\left(|\tilde{f}| > N\right) = 0\right\} \\
&= |\alpha|\, \|f\|_{L^\infty}.
\end{aligned}$$

Considérons enfin $f \in L^\infty(\mu)$ tel que $\|f\|_{L^\infty} = 0$. Alors $\mu(|\tilde{f}| > 0) = 0$ et donc $\tilde{f} = 0$ presque partout, ce qui implique $f = 0$ dans $\text{Ł}^\infty(\mu)$. On a bien montré que $\text{Ł}^\infty(\mu)$ est un espace vectoriel normé avec la norme définie en (6.2). $\square$

On vient de voir que l'inégalité de Minkowski correspondait à l'inégalité triangulaire dans les $L^p(\mu)$. L'inégalité de Hölder se lit ainsi :

Proposition 6.1.5 *Soient $p, p' \in [1, +\infty]$ deux exposants conjugués. Alors pour tout $f \in L^p(\mu)$, $g \in L^{p'}(\mu)$, on a $fg \in L^1(\mu)$ et $\|fg\|_1 \leq \|f\|_p \|g\|_{p'}$ (*Hölder*).*

Preuve . Lorsque $p, p' \in]1, +\infty[$ on obtient directement le résultat en appliquant l'inégalité de Hölder a des représentants de f et g.

Pour $p = 1$ (et donc $p' = +\infty$) considérons $\tilde{f}$, $\tilde{g}$ des représentants de f et g. On a alors

$$|\tilde{f}(x)\tilde{g}(x)| \leq |\tilde{f}(x)| \|g\|_{L^\infty} \quad pp.x \in X,$$

d'après la formule (6.3). Comme $\tilde{f} \in \mathcal{L}^1(\mu)$ on obtient que le produit $\tilde{f}\tilde{g}$ également d'après le dernier point de la Proposition 3.2.7. Précisément on a

$$\int |\tilde{f}\tilde{g}| d\mu \leq \left(\int |\tilde{f}| d\mu\right) \|g\|_{L^\infty} < \infty,$$

ce qui implique $fg \in L^1(\mu)$ et $\|fg\|_{L^1} \leq \|f\|_{L^1} \|g\|_{L^\infty}$. $\square$

On passe maintenant au Théorème le plus important de ce chapitre :

Théoreme 6.1.6 *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où $\mu \geq 0$, $p \in [1, +\infty]$. Alors $L^p(\mu)$ est un espace de Banach et $L^2(\mu)$ est un espace de Hilbert.*

Preuve. Pour montrer que c'est un espace de Banach il suffit de montrer qu'il est complet. Considérons (f_n) une suite de Cauchy dans $L^p(\mu)$. Cela s'écrit

$$\forall \varepsilon > 0, \quad \exists N_\varepsilon \quad \text{tel que} \quad \forall n, m \geq N_\varepsilon, \quad \|f_n - f_m\|_{L^p} \leq \varepsilon.$$

Nous allons séparer de nouveau les cas.

Cas $1 \leq p < \infty$: On va construire une suite d'indices $(n_k)_{k \in \mathbb{N}}$ tel que

$$\forall k \in \mathbb{N}, \quad \|f_{n_{k+1}} - f_{n_k}\|_{L^p} \leq 2^{-k}. \tag{6.4}$$

Procédons par réccurence. Puisque (f_n) est de Cauchy, il existe $n_1 \in \mathbb{N}$ tel que $\forall p \in \mathbb{N}$, $\|f_{n_1+p} - f_{n_1}\|_{L^p} \leq 2^{-1}$ et de même il existe $p_1 \in \mathbb{N}$ tel que $\forall p \in \mathbb{N}$, $\|f_{n_1+p_1+p} - f_{n_1+p_1}\|_{L^p} \leq 2^{-2}$. On pose alors $n_2 = n_1 + p_1$ (alors $\|f_{n_2} - f_{n_1}\|_{L^p} < 2^{-1}$))

Par récurrence sur $k \in \mathbb{N}^*$, supposons construits $n_1 \leq n_2 \leq \cdots \leq n_k$ tel que $\forall p \in \mathbb{N}$, et $\forall j \in [1, k]$ on ait

$$\|f_{n_j+p} - f_{n_j}\|_{L^p} \leq 2^{-j}. \tag{6.5}$$

On remarque en particulier que cela implique $\|f_{n_j} - f_{n_j-1}\|_{L^p} \leq 2^{-(j-1)}$ dès que $j \in [2, k]$. De la propriété de Cauchy on déduit qu'il existe $p_k \in \mathbb{N}$ tel que

$$\forall p \geq 0, \qquad \|f_{n_k+p_k+p} - f_{n_k+p_k}\| \leq 2^{-(k+1)}.$$

On pose alors $n_{k+1} = n_k + p_k$ et on obtient (6.5) pour $j = k + 1$, ce qui termine la récurrence. Puisque les n_k sont ordonnés, (6.5) implique également le résultat cherché (6.4).

On construit alors les (classes de) fonctions positives suivantes

$$g_k = \sum_{j=1}^{k} |f_{n_{j+1}} - f_{n_j}|, \qquad \text{et} \quad g = \sum_{j=1}^{\infty} |f_{n_{j+1}} - f_{n_j}|.$$

Alors $g_k \in L^p(\mu)$ comme somme finie de fonctions dans $L^p(\mu)$, puisque $L^p(\mu)$ est un espace vectoriel. On identifie pour la suite de la preuve les fonctions dans $L^p(\mu)$ et un de leur représentant. On a alors

$$\|g_k\|_{L^p} \leq \sum_{j=1}^{k} \left\|f_{n_{j+1}} - f_{n_j}\right\|_{L^p} \leq \sum_{j=1}^{k} 2^{-j} \leq 1.$$

d'après (6.4). On remarque ensuite que g est mesurable comme limite d'une suite de fonctions mesurables (la valeur ∞ est permise). On en déduit directement que g^p est également mesurable d'après la Proposition 2.2.4. Son intégrale vaut

$$\int |g|^p d\mu = \int \lim |g_k|^p d\mu = \int \liminf |g_k|^p d\mu \overset{Fatou}{\leq} \liminf \int |g_k|^p d\mu \leq 1,$$

par le lemme de Fatou 4.1.5. On a donc $g \in L^p(\mu)$ et $\|g\|_{L^p} \leq 1$. Le point iv) de la Proposition 4.2.3 implique alors que $g^p < \infty$ presque partout, ce qui signifie exactement que pour presque tout $x \in X$, la série numérique $\sum_{j=1}^{k}(f_{n_{j+1}}(x) - f_{n_j}(x))$ converge absolument dans $\mathbb{R}$. Soit A l'ensemble sur lequel cette limite existe, qui est donc tel que $\mu(A^c) = 0$. On définit alors pour tout $x \in X$ $f(x)$ par la série télescopique suivante

$$f(x) = \left(f_{n_1}(x) + \sum_{j=1}^{\infty} (f_{n_{j+1}}(x) - f_{n_j}(x)) \right) \mathbb{1}_A(x).$$

Alors pour tout $x \in X$, on a $f(x) = \lim_{j\to\infty} f_{n_j}(x)\mathbb{1}_A(x)$ donc la fonction f est mesurable. C'est un bon candidat pour être la limite de la suite (f_n) :

Montrons d'abord que $f \in L^p(\mu)$. Soit $m \in \mathbb{N}$, alors $\forall x \in A$,

$$|f(x) - f_m(x)| = \lim_{k\to\infty} |f_{n_k}(x) - f_m(x)|,$$

donc en intégrant

$$\int |f - f_m|^p d\mu = \int \lim_{k\to\infty} |f_{n_k} - f_m|^p d\mu = \int \liminf |f_{n_k} - f_m|^p d\mu \overset{Fatou}{\leq} \liminf \int |f_{n_k} - f_m|^p d\mu.$$

Or (f_j) est de Cauchy dans $L^p(\mu)$ donc $\forall \varepsilon$, $\exists N_\varepsilon$ tel que $\forall m \geq N_\varepsilon$ et $n_k \geq N_\varepsilon$ on ait $\int |f_{n_k} - f_m|^p d\mu \leq \varepsilon^p$. On en déduit que pour un tel ε et $m \geq N_\varepsilon$ on a

$$\int |f - f_m|^p \leq \varepsilon^p. \tag{6.6}$$

En particulier, on obtient $f - f_m \in L^p(\mu)$ et donc $f = f - f_m + f_m \in L^p(\mu)$ puisque $L^p(\mu)$ est un espace vectoriel.

Enfin on a également obtenu que $\forall m \geq N_\varepsilon$, $\|f - f_m\|_{L^p} \leq \varepsilon$ ce qui implique que $f_m \to f$ dans $L^p(\mu)$. On a montré que la suite (f_n) converge dans $L^p(\mu)$ et donc que $L^p(\mu)$ est un espace de Banach pour $p \in [1, +\infty[$.

Cas $p = 2$**.** Montrons que $L^2(\mu)$ est un espace de Hilbert. On sait déjà que $L^2(\mu)$ est un espace de Banach pour la norme $\|f\|_{L^2} = (\int |f|^2 d\mu)^{1/2}$. Il reste juste à verifier que cette norme est associée à un produit Hermitien. Pour cela on introduit pour $f, g \in L^2(\mu)$ l'application suivante

$$B(f, g) = \int f\overline{g} d\mu.$$

Cette application est bien définie sur $L^2(\mu)$ à valeur dans $\mathbb{C}$ grace à l'inégalité de Hölder. elle est bien linéaire par rapport à la première variable, semi-linéaire par rapport à la deuxième et vérifie $B(g, f) = \overline{B(f, g)}$. C'est donc une forme sesquilinéaire. Elle est clairement positive puisque

$$B(f, f) = \|f\|_{L^2}^2 \geq 0.$$

Enfin soit $f \in L^2(\mu)$ telle que $B(f, f) = 0$. Alors $\|f\|_{L^2}^2 = 0$ et comme $f \geq 0$ cela implique d'après le point i) de la Proposition 4.2.3 que $\tilde{f} = 0$ pp. si $\tilde{f}$ est un représentant de f. Finalement $f = 0$ dans $L^2(\mu)$. C'est donc un espace de Hilbert.

Cas $p = \infty$**.** Considérons de nouveau la suite de Cauchy (f_n) cette fois-ci dans $L^\infty(\mu)$. On identifie de nouveau les classes de fonctions et leurs représentants. On pose pour tout $n, m \in \mathbb{N}$,

$$\begin{aligned} & A_n = \{x \in X;\ |f_n(x)| > \|f_n\|_\infty\}, \\ \text{et} \quad & B_{n,m} = \{x \in X;\ \ |f_n(x) - f_m(x)| > \|f_n - f_m\|_\infty\}. \end{aligned}$$

Alors d'après la remarque faite en (6.3), on a $\mu(A_n) = 0$ et $\mu(B_{n,m}) = 0$. Cela implique que l'ensemble

$$A^c = (\bigcup_n A_n) \cup (\bigcup_{n,m} B_{n,m})$$

est de mesure nulle comme union dénombrable d'ensembles de mesure nulle. D'après les définitions précédentes, on a pour tout $x \in A$

$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\|_{L^\infty} \quad \text{et} \quad |f_n(x)| \leq \|f_n\|_{L^\infty}. \tag{6.7}$$

Nous utiliserons ces majorations plus loin. On remarque ensuite que l'inégalité triangulaire implique

$$|\|f_n\|_{L^\infty} - \|f_m\|_{L^\infty}| \leq \|f_n - f_m\|_{L^\infty}.$$

Comme (f_n) est de Cauchy dans $L^p(\mu)$, on en déduit que la suite réelle $(\|f_n\|_{L^\infty})_{n\in\mathbb{N}}$ est de Cauchy dans $\mathbb{R}$ par simple majoration. En particulier elle est bornée par un réel $M_0 \geq 0$. D'après l'inégalité (6.7), on obtient que pour tout $x \in A$ on a

$$|f_n(x)| \leq M_0.$$

Pour tout $x \in A$, on a alors d'après l'inégalité triangulaire

$$||f_n(x)| - |f_m(x)|| \leq |f_n(x) - f_m(x)| \leq \|f_n - f_m\|_{L^\infty}.$$

On en déduit que pour tout $x \in A$, $(f_n(x))$ est de cauchy dans $\mathbb{C}$ donc converge. Introduisons alors la fonction définie pour $x \in X$ par

$$f(x) = \lim_{n\to\infty}(f_n(x)\mathbb{1}_A(x)).$$

Elle est est mesurable comme limite et produit de fonctions mesurables. Par ailleurs $\forall x \in A$, $|f_n(x)| \leq M_0$ donc $|f(x)| \leq M_0$ ce qui implique $f \in L^\infty(\mu)$ et $\|f\|_{L^\infty} \leq M_0$.

Montrons enfin que $f_n \xrightarrow{L^\infty} f$. On sait que $\forall x \in A$, on a

$$|(f(x) - f_m(x))\mathbb{1}_A(x)| = \lim_{n\to\infty} |(f_n(x) - f_m(x))\mathbb{1}_A(x)| \leq \lim_{n\to\infty} \|f_n - f_m\|_{L^\infty}. \tag{6.8}$$

Or (f_n) est de Cauchy, donc pour tout $\varepsilon > 0$, $\exists N_\varepsilon$ tel que $\forall n, m \geq N_\varepsilon$, $\|f_n - f_m\|_\infty \leq \varepsilon$. L'inégalité (6.8) implique donc que $\forall m \geq N_\varepsilon$ et $\forall x \in A$,

$$|(f(x) - f_m(x))\mathbb{1}_A(x)| \leq \varepsilon,$$

donc $\|f - f_m\|_{L^\infty} \leq \varepsilon$ puisque $\mu(A) = 0$. On vient de montrer que $f_m \xrightarrow{L^\infty} f$ et donc que $L^\infty(\mu)$ est complet. La preuve est terminée. □

Au cours de la preuve, on a également montré le résultat suivant :

Proposition 6.1.7 *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré ou μ est une mesure positive. Soit $p \in [1, +\infty]$ et (f_n) une suite de fonctions convergeant vers une fonction f dans $L^p(\mu)$. Alors il existe une sous suite $(f_{n_k})_{k\in\mathbb{N}}$ telle que $f_{n_k} \to f$ presque partout.*

De nouveau on a identifié les classes de fonctions et un de leur représentants. Précisément on peut rappeler que si $\tilde{f}_n$ et $\tilde{f}$ sont des représentants respectifs de f_n et f, alors $\tilde{f}_n \to \tilde{f}$ presque partout signifie qu'il existe un ensemble A tel que $\mu(A^c) = 0$ et $\forall x \in A$, $\tilde{f}_j(x) \to \tilde{f}(x)$ dans $\mathbb{C}$. On signale également que par exemple dans $L^1(\mu)$ il existe des suites (f_n) telles que $(f_n(x))$ diverge pour tout $x \in X$. Le processus d'extraction d'une sous-suite est donc indispensable.

6.2 Intégrales dépendant d'un paramètre

On va montrer dans cette section deux Théorèmes de continuité et de derivation sous le signe intégrale. Ce sont essentiellement des applications du Théorème de convergence dominée, et les preuves sont remarquablement courtes.

Théoreme 6.2.1 *(Continuité) — Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Soit Y un espace métrique, $y_0 \in Y$ et $f : X \times Y \to \mathbb{C}$ une application vérifiant :*

a) $\forall y \in Y$, l'application $\left\{ \begin{array}{ccc} X & \longrightarrow & \mathbb{C} \\ x & \longmapsto & f(x,y) \end{array} \right. \in \mathcal{L}^1(\mu)$ *;*

b) pour μ-presque tout x, $\left\{ \begin{array}{ccc} Y & \longrightarrow & \mathbb{C} \\ y & \longmapsto & f(x,y) \end{array} \right.$ *est continue en y_0 ;*

c) il existe une fonction $0 \leq g \in \mathcal{L}^1(\mu)$ telle que pour presque tout $x \in X$, $\forall y \in Y$, $|f(x,y)| \leq g(x)$.

Alors la fonction définie sur Y par $F(y) = \int f(x,y)d\mu$ est continue en y_0.

Précisons un peu les hypothèses avant de commencer la preuve. On commence par remarquer que a) permet de donner un sens à la fonction F. b) signifie qu'il existe un ensemble N mesurable tel que $\mu(N) = 0$ (N indépendant du paramètre y) et $\forall x \in N^c$ fixé, la fonction $\left\{ \begin{array}{ccc} Y & \to & \mathbb{C} \\ y & \mapsto & f(x,y) \end{array} \right.$ est continue en y_0. L'hypothèse c) est une hypothèse de domination. Elle signifie qu'il existe un autre ensemble N' tel que $\mu(N') = 0$, et $\forall x \in N'^c$, $\forall y \in Y$ on a $|f(x,y)| \leq g(x)$.

Preuve. Soit $(y_n) \longrightarrow y_0$ dans Y. Alors $\forall n \in \mathbb{N}$,

$$F(y_n) - F(y_0) = \int \underbrace{(f(x,y_n) - f(x,y_0))}_{f_n(x)} d\mu.$$

Le point b) implique que $f_n(x) \to 0$ presque pour tout $x \in X$. Par ailleurs, le point c) implique que

$$|f_n(x)| = |f(x,y_n) - f(x,y_0)| \leq |f(x,y_n)| + |f(x,y_0)| \leq 2g(x) \quad pp.\ x \in X.$$

Le Théorème de convergence dominée version II 4.2.4 implique directement que $F(y_n) - F(y_0)$ a une limite qui est $\lim \int f_n(x)d\mu = \int \lim f_n(x)d\mu = 0$. Le Théorème est prouvé. □

Remarque 6.2.2 La continuité est une propriété locale. Pour obtenir la continuité en un point y_0, il suffit de prendre un voisinage quelconque de y_0 à la place de Y. Pour la même raison, si dans la condition b) on suppose que la fonction utilisée est continue partout au lieu de l'être juste en y_0 alors la fonction F elle même devient également continue partout. De même il peut être utile de restreindre Y.

Théoreme 6.2.3 *(différentiabilité) — Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Soit Y un ouvert de $\mathbb{R}^n$ et $f : X \times Y \to \mathbb{C}$ une application vérifiant :*

a) $\forall y \in Y$, $\left\{ \begin{array}{ccc} X & \to & \mathbb{C} \\ x & \mapsto & f(x,y) \end{array} \right. \in \mathcal{L}^1(\mu)$

b) pour μ-presque tout x, $\left\{\begin{array}{ccc} Y & \to & \mathbb{C} \\ y & \mapsto & f(x,y) \end{array}\right.$ *est différentiable sur Y ;*

c) il existe une fonction $0 \leq g \in \mathcal{L}^1(\mu)$ telle que pour presque tout $x \in X$, $\forall y \in Y$, $\|d_y f(x,y)\| \leq g(x)$.

Alors la fonction F définie sur Y par $F(y) = \int d_y f(x,y) d\mu$ est différentiable et sa différentielle est donnée par

$$dF(y) = \int d_y f(x,y) d\mu.$$

De nouveau précisons les hypothèses. On rappelle que pour $x \in X$ et $y \in Y$ fixés, $d_y f(x,y)$ est l'application linéaire de $\mathbb{C}^n$ dans $\mathbb{C}$ de composantes

$$\frac{\partial f}{\partial y_1}(x,y), \cdots, \frac{\partial f}{\partial y_n}(x,y).$$

Cela signifie que pour un vecteur $h \in \mathbb{C}^n$ on a

$$d_h f(x,y).h = \frac{\partial f}{\partial y_1} h_1 + \cdots + \frac{\partial f}{\partial y_n}.$$

De plus pour tout $y \in Y$ la fonction $x \longrightarrow d_y f(x,y)$ est mesurable. En effet elle est limite d'une suite de fonctions mesurables, puisque pour $h \in \mathbb{C}^n$ on a

$$d_y f(x,y).h = \lim_{k \to \infty} \frac{f(x, y + h/k) - f(x,y)}{1/k}.$$

L'Hypothèse c) est de nouveau une hypothèses de domination. Elle implique en particulier que $\forall y \in Y$, $\int \|d_y f(x,y)\| \, d\mu < +\infty$ et que donc $d_y f(x,y) \in \mathcal{L}^1(\mu)$ (ie. les composantes de $d_y f(x,y)$ sont dans $\mathcal{L}^1(\mu)$).

Preuve. Soit $y \in Y$ et $r > 0$ tel que $\overline{B(y,r)} \subset Y$ et soit $0 \neq h \in \overline{B(y,r)}$. Un développement de Taylor de f en y à l'ordre 1 donne alors

$$F(y+h) - F(y) = \int (f(x, y+h) - f(x,y)) d\mu = \int \left(d_y f(x,y).h + \varepsilon_{x,y}(h).\|h\|\right) d\mu.$$

Comme les fonctions $x \mapsto f(x, y+h)$, $x \mapsto f(x,y)$ et $x \mapsto d_y f(x,y).h$ sont mesurables et même dans $\mathcal{L}^1(\mu)$ on obtient que $x \mapsto \varepsilon_{x,y}(h)$ est également mesurable et dans $\mathcal{L}^1(\mu)$. Cela permet d'utiliser la linéarité de l'intégrale pour obtenir

$$F(y+h) - F(y) = \left(\int d_y f(x,y) d\mu\right).h + \left(\int \varepsilon_{x,y}(h) d\mu\right).\|h\|.$$

D'après l'inégalité des accroissements finis on a alors pour presque tout $x \in X$,

$$\begin{aligned} |\varepsilon_{x,y}(h) \times \|h\|| &= |f(x, y+h) - f(x,y) - d_y f(x,y)| \\ &\leq \sup_{\theta \in [0,1]} \|d_y f(x, y + \theta h)\|.\|h\| + \|d_y f(x,y)\|.\|h\| \\ &\leq 2g(x) \|h\|. \end{aligned} \tag{6.9}$$

C'est une hypothèse de domination. Par ailleurs f étant différentiable dans la variable y d'après l'hypothèse c) on obtient pour presque tout $x \in X$ et y fixés que

$$\varepsilon_{x,y}(h) \xrightarrow[h \to 0]{} 0.$$

Soit alors $(h_k)_{k\in\mathbb{N}}$ une suite telle que tel que $h_k \to 0$. Alors la fonction ε_k définie pour $y \in Y$ par $\varepsilon_k(x) = \varepsilon_{x,y}(h_k)$ converge presque partout vers 0 et est uniformément majorée par $2g$ d'après (6.9). Le Théorème de Lebesgue version II implique donc que

$$\lim_{k\to\infty} \int \varepsilon_{x,y}(h_k)d\mu = \int \lim_{k\to\infty} \varepsilon_{x,y}(h_k)d\mu = 0.$$

On peut donc écrire que pour tout $y \in Y$ on a

$$F(y+h) = F(y) + \int d_y f(x,y).hd\mu + E(y,h).\|h\| \quad \text{avec} \quad E(y,h) \stackrel{\text{def}}{=} \int \varepsilon_{x,y}(h)d\mu.$$

On vient de montrer que $E(y,h) \xrightarrow[h\to 0]{} 0$ ce qui implique que F est différentiable de différentielle celle proposée par l'énoncé. Le résultat est prouvé. $\square$

Chapitre 7

Produit d'espaces mesurés

7.1 Produit d'espaces mesurables.

On rappelle d'abord ce que donne la Définition 2.1.9 dans le cas du produit de 2 espaces.

Définition 7.1.1 *Soit $(X_1, \mathcal{M}_1)$, $(X_2, \mathcal{M}_2)$ deux espaces mesurables. On appelle tribu produit notée $\mathcal{M}_1 \otimes \mathcal{M}_2$ sur $X_1 \times X_2$ la tribu engendrée par les rectangles (i.e. les ensembles du type $A_1 \times A_2$ où $A_1 \in \mathcal{M}_1$ et $A_2 \in \mathcal{M}_2$).*

On peut remarquer que la tribu $\mathcal{M}_1 \otimes \mathcal{M}_2$ est la plus petite tribu rendant mesurables les application coordonnées

$$\Pi_1 : \left\{ \begin{array}{ccc} X_1 \times X_2 & \longrightarrow & X_1 \\ (x_1, x_2) & \longmapsto & x_1 \end{array} \right. \quad \text{et} \quad \Pi_2 : \left\{ \begin{array}{ccc} X_1 \times X_2 & \longrightarrow & X_1 \\ (x_1, x_2) & \longmapsto & x_2 \end{array} \right. .$$

En effet, supposons d'abord que Π_1 est mesurable. Alors pour $A_1 \in \mathcal{M}_1$, $\Pi_1^{-1}(A_1) = A_1 \times X_2 \in \mathcal{M}_1 \otimes \mathcal{M}_2$. Réciproquement supposons que $\mathcal{T}$ est une tribu sur $X_1 \times X_2$ qui rende Π_1 et Π_2 mesurables, alors pour $A_1 \in \mathcal{M}_1$, $A_2 \in \mathcal{M}_2$, on a $\Pi_1^{-1}(A_1) = A_1 \times X_2 \in \mathcal{T}$ et $\Pi_2^{-1}(A_2) = X_1 \times A_2 \in \mathcal{T}$ donc

$$(A_1 \times X_2) \cap (X_1 \times A_2) = A_1 \times A_2 \in \mathcal{T}.$$

On vient de montrer que $\mathcal{T}$ contient les rectangles donc $\mathcal{M}_1 \otimes \mathcal{M}_2 \subset \mathcal{T}$ d'après la définition d'une tribu engendrée 2.1.4.

Remarque 7.1.2 Soit $f_1 : X_1 \to \mathbb{C}$ et $f_2 : X_2 \to \mathbb{C}$ deux fonctions mesurables. On définit alors

$$f_1 \otimes f_2 : \left\{ \begin{array}{ccc} X_1 \times X_2 & \longrightarrow & \mathbb{C} \\ (x_1, x_2) & \longmapsto & f_1(x_1) f_2(x_2) \end{array} \right. .$$

La fonction $f_1 \otimes f_2$ est alors mesurable. En effet $f_1 \otimes f_2 = f_1 \circ \Pi_1 \times f_2 \circ \Pi_2$ et donc $f_1 \otimes f_2$ est produit et composée de fonctions mesurables, ce qui implique le résultat d'après les Propositions 2.2.4 et 2.2.9. $\square$

On donne maintenant un résultat sur les générateurs d'une tribu produit. La preuve est laissée en exercice.

Proposition 7.1.3 *Soient $(X_1, \mathcal{M}_1)$ et $(X_2, \mathcal{M}_2)$ deux espaces mesurables. Si $\mathcal{C}_1 \subset \mathcal{M}_1$ engendre $\mathcal{M}_1$ et contient X_1, et si $\mathcal{C}_2 \subset \mathcal{M}_2$ engendre $\mathcal{M}_2$ et contient X_2, alors $\mathcal{C}_1 \times \mathcal{C}_2 \stackrel{\text{def}}{=} \{C_1 \times C_2;\ C_1 \in \mathcal{C}_1,\ C_2 \in \mathcal{C}_2\}$ engendre $\mathcal{M}_1 \otimes \mathcal{M}_2$.*

Remarque 7.1.4 Les résultats précédents s'étendent au cas du produit de n espaces $X_1 \times \cdots \times X_n$ muni de la tribu $\mathcal{M}_1 \otimes \cdots \otimes \mathcal{M}_n$. On pourra en particulier remarquer que le produit tensoriel des tribus est associatif. Ansi on peut vérifier que le produit $(\mathcal{M}_1 \otimes \cdots \otimes \mathcal{M}_d) \otimes (\mathcal{M}_{d+1} \otimes \cdots \otimes \mathcal{M}_n)$ ne dépend pas de d et est égal à $\mathcal{M}_1 \otimes \cdots \otimes \mathcal{M}_n$. En fait on sait d'après le début de ce chapitre que $(\mathcal{M}_1 \otimes \cdots \otimes \mathcal{M}_d) \otimes (\mathcal{M}_{d+1} \otimes \cdots \otimes \mathcal{M}_n)$ est la plus petite tribu rendant mesurable l'application $(x_1, \cdots, x_n) \mapsto (x_1, \cdots, x_d)$ et $(x_1, \cdots, x_n) \mapsto (x_{d+1}, \cdots, x_n)$. C'est aussi la plus petite tribu rendant mesurables les projections définies pour tout $j \in [1, n]$ par $(x_1, \cdots, x_n) \longmapsto x_j$. C'est donc également $\mathcal{M}_1 \otimes \cdots \otimes \mathcal{M}_n$ par simple vérification.

Exemple 7.1.5 Grace à la Proposition 7.1.3 on retrouve ainsi le fait que le produit de d exemplaires de $(\mathbb{R}, B(\mathbb{R}))$ est $(\mathbb{R}^d, B(\mathbb{R}^d))$.

On retourne maintenant au cas de deux espaces.

Proposition 7.1.6 *Soient $(X_1, \mathcal{M}_1)$, $(X_2, \mathcal{M}_2)$ et $(Y, \mathcal{N})$ trois ensembles mesurables. Une application $f : Y \to X_1 \times X_2$ est mesurable si et seulement si les applications coordonnées $f_1 = \Pi_1 \circ f$ et $f_2 = \Pi_2 \circ f$ le sont.*

Preuve. Supposons d'abord f mesurable. Alors f_1 et f_2 sont également mesurables comme composées de fonctions mesurables.

Réciproquement supposons que f_1 et f_2 soient mesurables. Puisque les pavés engendrent la tribu produit, il suffit d'après la Proposition 2.2.5 de montrer que l'image réciproque d'un pavé par f est mesurable. Soient donc $A_1 \in \mathcal{M}_1$ et $A_2 \in \mathcal{M}_2$, alors

$$\begin{aligned} f^{-1}(A_1 \times A_2) &= f^{-1}((A_1 \times X_2) \cap (X_1 \times A_2)) \\ &= f^{-1}\left(\Pi_1^{-1}(A_1) \cap \Pi_2^{-1}(A_2)\right) \\ &= f^{-1}\left(\Pi_1^{-1}(A_1)\right) \cap f^{-1}\left(\Pi_2^{-1}(A_2)\right) \\ &= (\Pi_1 \circ f)^{-1}(A_1) \cap (\Pi_2 \circ f)^{-1}(A_2) \\ &= f_1^{-1}(A_1) \cap f_2^{-1}(A_2) \in \mathcal{N}, \end{aligned} \tag{7.1}$$

puisque $\mathcal{M}$ est une tribu. On a utilisé les propriétés des fonctions réciproques rappelées après (1.1). On a ainsi montré que f est mesurable. Cela conclue la preuve. $\square$

Regardons pour finir cette section les propriétés de quelques fonctions et ensembles particuliers.

Proposition 7.1.7 *Soit $(X_1, \mathcal{M}_1)$, $(X_2, \mathcal{M}_2)$ deux espaces mesurables, alors :*

a) Si $(Y, \mathcal{N})$ est mesurable et si $f : X_1 \times X_2 \to Y$ est mesurable alors pour tout $x_{1,0} \in X_1$ et $x_{2,0} \in X_2$ fixés, les fonctions suivantes sont mesurables :

$$f_{x_{1,0}} : \left\{ \begin{array}{ccc} X_2 & \longrightarrow & Y \\ x_2 & \longmapsto & f(x_{1,0}, x_2) \end{array} \right. , \qquad et \qquad f_{x_{2,0}} : \left\{ \begin{array}{ccc} X_1 & \longrightarrow & Y \\ x_1 & \longmapsto & f(x_1, x_{2,0}) \end{array} \right. .$$

b) Si $A \in \mathcal{M}_1 \otimes \mathcal{M}_2$, alors $\forall x_{1,0} \in X_1$ et $x_{2,0} \in X_2$ fixés, les ensembles suivants sont mesurables :

$$A_{x_{1,0}} = \{x_2 \in X_2;\ (x_{1,0}, x_2) \in A\} \quad \text{et} \quad A_{x_{2,0}} = \{x_1 \in X_1;\ (x_1, x_{2,0}) \in A\}. \tag{7.2}$$

Preuve. Pour a) on remarque que pour $x_{1,0} \in X_1$ fixé on a $f_{x_{1,0}} = f \circ h_{x_{1,0}}$ où

$$h_{x_{1,0}} : \begin{cases} X_2 & \rightarrow & X_1 \times X_2 \\ x_2 & \mapsto & (x_{1,0}, x_2) \end{cases}.$$

Pour cette fonction les applications coordonnées sont la fonction constante $x_2 \mapsto x_{1,0}$ et l'application Identité $\mathrm{Id}_{X_2} : x_2 \mapsto x_2$. Elles sont clairement mesurables et donc d'après la proposition 7.1.6, la fonction $h_{x_{1,0}}$ aussi. La Proposition 2.2.4 implique alors que la composée $f_{x_{1,0}}$ est également mesurable. On obtient de même que $f_{x_{2,0}}$ est mesurable en échangeant le rôle des variables. Cela prouve le point a).

Pour montrer le point b), prenons $f = \mathbb{1}_A$ et donc $f_{x_{1,0}} = f(x_{1,0}, .) = \mathbb{1}_{A_{x_{1,0}}}$. Or $f_{x_{1,0}}$ est mesurable par a) donc $A_{x_{1,0}} = f_{x_{1,0}}^{-1}(\{1\})$ aussi. On obtient de même que $A_{x_{2,0}}$ est mesurable en échangeant les rôles des variables. Cela conclue la preuve. □

7.2 Mesure produit, Théorème de Fubini.

Définition-Proposition 7.2.1 *Soit $(X_1, \mathcal{M}_1, \mu_1)$, $(X_2, \mathcal{M}_2, \mu_2)$ deux espaces mesurés où les mesures μ_j sont positives et finies (resp. σ-finies). Alors il existe une unique mesure notée $\mu_1 \otimes \mu_2$ sur $(X_1 \times X_2, \mathcal{M}_1 \otimes \mathcal{M}_2)$ telle que $\forall A_1 \in \mathcal{M}_1$, $A_2 \in \mathcal{M}_2$,*

$$\mu_1 \otimes \mu_2(A_1 \times A_2) = \mu_1(A_1) \times \mu_2(A_2), \tag{7.3}$$

et cette mesure est finie (resp. σ-finie). Le triplet $(X_1 \times X_2, \mathcal{M}_1 \otimes \mathcal{M}_2, \mu_1 \otimes \mu_2)$ est appelé espace mesuré produit.

On renvoie au Théorème 5.1.3 pour la définition d'une application $\sigma-$finie. Pour la preuve de cette Proposition, on va avoir besoin du résultat suivant, dont la preuve un peu technique est omise.

Lemme 7.2.2 *Sous les hypothèses de la Proposition 7.2.1, considérons $A \in \mathcal{M}_1 \otimes \mathcal{M}_2$. Alors l'application $\varphi_1^A : \begin{cases} X_1 & \longrightarrow & \overline{\mathbb{R}}_+ \\ x_1 & \longmapsto & \mu_2(A_{x_1}) \end{cases}$ est mesurable.*

Preuve de la Proposition. On vient de voir que $\forall A \in \mathcal{M}_1 \otimes \mathcal{M}_2$, φ_1^A est mesurable. On pose alors

$$\mu_1 \otimes \mu_2(A) \stackrel{\text{def}}{=} \int_{X_1} \varphi_1^A(x_1) d\mu_1. \tag{7.4}$$

L'application $\mu_1 \otimes \mu_2$ est bien définie sur $\mathcal{M}_1 \otimes \mathcal{M}_2$ et à valeur dans $\overline{\mathbb{R}_+}$. Montrons que c'est une mesure. On observe d'abord que

$$\mu_1 \otimes \mu_2(\emptyset) = \int \mu_2(\emptyset) d\mu_1 = 0,$$

puisque lorsque $A=\emptyset$ on a $A_{x_1}=\emptyset$ dans X_2.

Montrons ensuite l'additivité dénombrable. On considère $(A_j)_{j\in\mathbb{N}}\subset\mathcal{M}_1\otimes\mathcal{M}_2$ une suite d'ensembles mesurables disjoints et on pose $A=\cup_{j\in\mathbb{N}}A_j$. Alors on observe d'une part que $X_2\supset A_{x_1}=(\cup A_j)_{x_1}=\cup(A_{j_{x_1}})$ et d'autres part que les $A_{j_{x_1}}$ sont disjoints puisque les A_j le sont. On écrit alors

$$\begin{aligned}\mu_1\otimes\mu_2(A)=\int\mu_2(A_{x_1})d\mu_1=\int\mu_2\left(\cup_{j\in\mathbb{N}}A_{j_{x_1}}\right)d\mu_1 &\overset{\text{ensembles disjoints}}{=}\int\sum_{j\in\mathbb{N}}\mu_2(A_{j_{x_1}})d\mu_1\\ &\overset{Beppo-Levi}{=}\sum_{j\in\mathbb{N}}\int\mu_2(A_{j_{x_1}})d\mu_1=\sum_{j\in\mathbb{N}}\mu_1\otimes\mu_2(A_j).\end{aligned}\tag{7.5}$$

Ici le Théorème de Beppo-Levi 4.1.1 s'applique aux fonctions positives $x_1\longmapsto\sum_{j=1}^{n}\mu_2(A_{j_{x_1}})$ qui convergent en croissant vers la somme $\sum_{j\in\mathbb{N}}\mu_2(A_{j_{x_1}})$. C'est le résultat cherché.

Pour finir montrons la dernière propriété. Pour $A_1\in\mathcal{M}_1$ et $A_2\in\mathcal{M}_2$, on a

$$\begin{aligned}\mu_1\otimes\mu_2(A_1\times A_2)&=\int\mu_2((A_1\times A_2)_{x_1})d\mu_1=\int\mu_2(A_2)\times\mathbb{1}_{A_1}(x_1)d\mu_1\\ &=\mu_2(A_2)\int\mathbb{1}_{A_1}(x_1)d\mu_1=\mu_2(A_2)\mu_1(A_1)\end{aligned}\tag{7.6}$$

En particulier, si μ_1 et μ_2 sont finies (resp. $\sigma-$ finies), alors $\mu_1\otimes\mu_2$ aussi. Par ailleurs $\mathcal{M}_1\otimes\mathcal{M}_2$ est engendré par les rectangles sur lesquels $\mu_1\times\mu_2$ a une valeur donnée par l'énoncé donc elle est uniquement déterminée sur $\mathcal{M}_1\otimes\mathcal{M}_2$ tout entier d'après le Théorème de prolongement 5.1.3. Elle est donc unique. La preuve est finie. □

On peut maintenant aborder les Théorèmes de Fubini-Tonelli et de Fubini, qui font partie des résultats fondamentaux du cours.

Théoreme 7.2.3 (de Fubini-Tonelli) *Soit $(X_1,\mathcal{M}_1,\mu_1)$, $(X_2,\mathcal{M}_2,\mu_2)$ deux espaces mesurés où les mesures μ_j sont positives et $\sigma-$ finis. Alors pour toute fonction f mesurable sur $X_1\times X_2$ à valeur dans $\overline{\mathbb{R}}_+$, les fonctions*

$$F_1:\left\{\begin{array}{ccc}X_1&\longrightarrow&\overline{\mathbb{R}}_+\\ x_1&\longmapsto&\int_{X_2}f(x_1,x_2)d\mu_2\end{array}\right.\qquad et\qquad F_2:\left\{\begin{array}{ccc}X_2&\longrightarrow&\overline{\mathbb{R}}_+\\ x_2&\longmapsto&\int_{X_1}f(x_1,x_2)d\mu_1\end{array}\right.$$

sont mesurables positives et on a

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

(Précisons que dans le premier terme le signe intégrale double est une convention d'écriture pour rappeler que l'intégration se fait sur un espace produit)

Preuve. On regarde d'abord le cas d'une fonction indicatrice. Soit $A\in\mathcal{M}_1\otimes\mathcal{M}_2$ et considérons $f=\mathbb{1}_A$. Pour cette fonction f on a alors pour tout $x_1\in X_1$,

$$F_1(x_1)=\int_{X_2}\mathbb{1}_A(x_1,x_2)d\mu_2=\mu_2(A_{x_1})=\varphi_1^A(x_1).$$

La fonction φ_1^A est mesurable d'après le lemme admis 7.2.2. De plus,

$$\iint f d\mu_1 \otimes \mu_2 = \iint \mathbb{1}_A d\mu_1 \otimes \mu_2 = \mu_1 \otimes \mu_2(A) \overset{\text{déf. (7.4)}}{=} \int \varphi_1^A d\mu_1 = \int F_1 d\mu_1.$$

En échangeant les rôles de x_1 et x_2, on obtient par unicité de la mesure produit que

$$\mu_1 \otimes \mu_2(A) = \int_{X_2} \varphi_2^A(x_2) d\mu_2 \qquad \text{où} \quad \varphi_2^A : \left\{ \begin{array}{ccc} X_2 & \to & \overline{\mathbb{R}}_+ \\ x_2 & \mapsto & \mu_1(A_{x_2}) \end{array} \right. . \tag{7.7}$$

On rappelle que A_{x_2} a été défini en (7.2). On obtient alors de manière similaire que si $f = \mathbb{1}_A$ alors $F_2 = \varphi_2^A$ et $\int F_2(x_2) d\mu_2 = \mu_1 \otimes \mu_2(A)$.

Le cas des fonctions étagées s'en déduit alors par somme finie. Pour le cas général, on travaille d'abord sur les fonctions étagées. Soit s_k une suite de fonctions étagées positives tendant en croissant vers f. On définit

$$S_{k,1}(x_1) = \int_{X_2} s_k(x_1, x_2) d\mu_2 \qquad \text{et} \qquad S_{k,2}(x_2) = \int_{X_1} s_k(x_1, x_2) d\mu_1.$$

Alors d'après le résultat précédemment prouvé sur les fonctions étagées on a :

$$\int_{X_1} S_{k,1} d\mu_1 = \int_{X_2} S_{k,2} d\mu_2 = \iint s_k d\mu_1 \otimes \mu_2 \tag{7.8}$$

On va appliquer plusieurs fois le Théorème de Beppo-Levi. On a d'abord

$$\iint s_k d\mu_1 \otimes \mu_2 \xrightarrow[\mu_1 \otimes \mu_2]{BL} \iint f d\mu_1 \otimes \mu_2 \tag{7.9}$$

Ensuite on a $0 \leq S_{k,1} = \int_{X_2} s_k(., x_2) d\mu_2$ donc $0 \leq S_{k,1} \longrightarrow F_1$ en croissant avec k, ce qui implique

$$\int s_{k_1}(x_1) d\mu_1 \xrightarrow[\text{pour } \mu_1]{BL} \int F_1 d\mu_1. \tag{7.10}$$

On en déduit

$$\iint f d\mu_1 \otimes \mu_2 \overset{(7.9)}{=} \lim_{k \to \infty} \iint s_k d\mu_1 \otimes \mu_2 \overset{(7.8)}{=} \lim_{k \to \infty} \int S_{k,1}(x_1) d\mu_1 \overset{(7.10)}{=} \int F_1 d\mu_1.$$

C'est le résultat cherché pour F_1. Le même résultat s'obtient pour F_2 en échangeant le rôle des variables. La preuve est complète. □

Exemple 7.2.4 Un exemple basique d'interversion d'intégrale est donnée par l'égalité

$$\sum_n \sum_p a_{n,p} = \sum_p \sum_n a_{n,p} = \sum_{n,p} a_{n,p}$$

(dans $\overline{\mathbb{R}}_+$) pour $a_{n,p} \in \mathbb{R}_+$. C'est le Théorème de Fubini-Tonelli sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ muni de la mesure de comptage.

Souvent dans les applications, le Théorème de Fubini-Tonelli sert à démontrer qu'une fonction est dans $\mathcal{L}^1(\mu_1 \otimes \mu_2)$. Le Théorème de Fubini qui va suivre sert alors à calculer explicitement son intégrale.

Théoreme 7.2.5 (de Fubini) *Soit* $(X_1, \mathcal{M}_1, \mu_1)$*,* $(X_2, \mathcal{M}_2, \mu_2)$ *deux espaces mesurés, où les mesures sont positives et* $\sigma-$*finies. Soit* $f : X_1 \times X_2 \to \mathbb{C}$ *une fonction mesurable. Alors :*

a) Si $\int_{X_1} \left(\int_{X_2} |f(x_1, x_2)| d\mu_2 \right) d\mu_1 < \infty$*, alors* $f \in \mathcal{L}^1(\mu_1 \otimes \mu_2)$*.*

b) Si $f \in \mathcal{L}^1(\mu_1 \otimes \mu_2)$ *alors pour presque tout* $x_1 \in X_1$ *on a* $f(x_1, .) \in \mathcal{L}^1(\mu_2)$*, de plus*

$$F_1 : x_1 \longmapsto \int f(x_1, x_2) d\mu_2$$

est intégrable sur X_1*. On obtient de même en échangeant les rôles de* x_1 *et* x_2*. Enfin toutes les intégrales sont égales, i.e. :*

$$\begin{aligned} \int_{X_1} \left(\int_{X_2} f(x_1, x_2) d\mu_2 \right) d\mu_1 &= \int_{X_2} \left(\int_{X_1} f(x_1, x_2) d\mu_1 \right) d\mu_2 \\ &= \iint_{X_1 \times X_2} f(x_1, x_2) d\mu_1 \otimes \mu_2. \end{aligned} \tag{7.11}$$

Preuve. Le point a) est une conséquence directe du Théorème de Fubini-Tonelli appliqué à la fonction positive $|f|$.

Pour le point b), on considère dans un premier temps une fonction f à valeur réelle, pour laquelle on écrit $f = f_+ - f_-$. Appliquons le Théorème de Fubini-Tonelli à la fonction positive $f_+ \leq |f|$. On obtient

$$\int_{X_1} \left(\int_{X_2} f_+(x_1, x_2) d\mu_2 \right) d\mu_1 \overset{Fubini-Tonelli}{=} \iint f_+ d\mu_1 \otimes \mu_2 \leq \iint |f| d\mu_1 \otimes \mu_2 < \infty.$$

puisque $|f| \in \mathcal{L}^1(\mu_1 \otimes \mu_2)$. Posons alors $F_{+,1} : x_1 \mapsto \int_{X_2} f_+(x_1, x_2) d\mu_2$. D'après la Proposition 4.2.3 le fait que $\int_{X_1} F_{+,1} d\mu_1 < \infty$ implique $F_{+,1} < \infty$ μ_1-presque partout, c'est à dire $f(x_1, .) \in \mathcal{L}^1(\mu_2)$ pour presque tout x_1. En privilégiant la variable x_2 au lieu de la variable x_1 on obtient de même

$$\int_{X_2} \left(\int_{X_1} f_+(x_1, x_2) d\mu_1 \right) d\mu_2 \overset{Fubini-Tonelli}{=} \iint f_+ d\mu_1 \otimes \mu_2 < \infty.$$

ce qui prouve (7.11) avec f_+ au lieu de f. En utilisant le fait que $0 \leq f_- \leq |f|$ on obtient de même (7.11) avec f_- au lieu de f. Les intégrales étant toutes finies on peut sommer et utiliser la linéarité, ce qui donne finalement (7.11) pour f.

Dans le cas où f est complexe, on écrit $f = f_+ - f_- + ig_+ - ig_-$ et on obtient le résultat en utilisant que $0 \leq f_\pm \leq |f|$ et $0 \leq g_\pm \leq |f|$. Le Théorème est démontré. □

7.3 Changements de variables et intégration

On se place dans cette courte section sans preuve sur $\mathbb{R}^d$ muni de la tribu des boréliens et de la mesure de Lebesgue.

On rappelle que pour $\mathcal{U}$ et $\mathcal{V}$ deux ouverts de $\mathbb{R}^d$, une application $\varphi : \mathcal{U} \to \mathcal{V}$ est un $\mathcal{C}^1-$difféomorphisme si elle est bijective, $\mathcal{C}^1$ et de réciproque $\mathcal{C}^1$. Une telle application est aussi appelée changement de variables. Commençons par un rappel du cours de calcul différentiel :

Proposition 7.3.1 *Soit $\mathcal{U}$ ouvert de $\mathbb{R}^d$. Alors une application $\varphi : \mathcal{U} \longrightarrow \mathbb{R}^d$ est un $\mathcal{C}^1-$difféomorphisme de $\mathcal{U}$ sur $\varphi(\mathcal{U})$ si et seulement si :*

i) φ est injective ;

ii) φ est $\mathcal{C}^1$ ($\Longleftrightarrow$ les dérivées partielles le sont) ;

iii) $\forall x \in \mathcal{U},\ J(\varphi) = det(\varphi'(x)) \neq 0$.

Le Théorème suivant se montre par récurrence sur la dimension grace à un argument de localisation. La preuve est omise.

Théoreme 7.3.2 *Soit $\mathcal{U}$ et $\mathcal{V}$ deux ouverts de $\mathbb{R}^d$ et φ un $\mathcal{C}^1-$difféomorphisme de $\mathcal{U}$ sur $\mathcal{V}$. Alors pour toute fonction positive f sur $\mathcal{V}$ on a :*

$$\int_{\mathcal{V}} f(y)dy = \int_{\mathcal{U}} f \circ \varphi(x)|J(\varphi)(x)|dx \qquad et \qquad \int_{\mathcal{U}} f \circ \varphi(x)dx = \int_{\mathcal{V}} f(y)|J(\varphi^{-1})(y)|dy.$$

En particulier, $\lambda(\varphi(A)) = \int_A |J(\varphi)(x)|dx$ (λ est la mesure de Lebesgue).

Exemples 7.3.3

1 — Changement de variables linéaire — Si φ est définie par $\varphi(x) = \alpha x + \beta$, $\alpha \in \mathbb{R}^*$, $\beta \in \mathbb{R}^d$, alors $|J(\varphi)| = |\alpha|^d$ et on a $\int_{\mathcal{V}} f(y)dy = \int_{\mathcal{U}} f(\alpha x + \beta)|\alpha|^d dx$.

2 — Coordonnées Polaires — Soit f une fonction borélienne définie sur

$$B_R = \{(x, y);\ x^2 + y^2 < R\}.$$

Alors la fonction $f_{pol} : (r, \theta) \mapsto f(r\cos\theta, r\sin\theta)$ définie sur $]0, R[\times]0, 2\pi[$ est mesurable positive, et si f intégrable alors f_{pol} aussi. De plus on a l'égalité

$$\int_{B_R} f(x, y)dxdy = \int_{[0,2\pi]}\int_{[0,R]} f(r\cos\theta, r\sin\theta)rdrd\theta = \int_{[0,2\pi]}\int_{[0,R]} f_{pol}(r, \theta)drd\theta,$$

où on a utilisé le changement de variables suivant

$$\varphi : \left\{ \begin{array}{ccc}]0, R[\times]0, 2\pi[& \longrightarrow & B_R \setminus [0, R[\times\{0\} \\ (r, \theta) & \longmapsto & (x, y) \end{array} \right. .$$

On pourra remarquer que pour la mesure produit, le segment $[0, R[\times\{0\}$ est de mesure nulle, et peut être enlevé du domaine d'intégration de f.