Calcul intégral et applications

L3 – Mathématiques

Loïc Chaumont

Université d'Angers

Contents

Chapitre 1

Espaces mesurables, fonctions mesurables

La théorie de l'intégration de Lebesgue repose sur des concepts importants de théorie des ensembles et quelques notions sur la cardinalité et en particulier sur la dénombrabilité sont nécessaires avant de toute chose.

1.1. Ensembles et applications

Un ensemble est une collection éventuellement vide d'éléments. Si E est un ensemble et si x est l'un de ses éléments, alors on note $x \in E$ et ceci se dit : « x appartient à E ». Les éléments d'un ensemble sont aussi appelés les *points* de cet ensemble. Si x n'est pas un élément de l'ensemble E, alors on dit que « x n'appartient pas à E » et l'on note $x \notin E$. Un ensemble décrit à partir de ses éléments se note avec des accolades. En particulier si E contient un nombre fini n d'éléments appelés $x_1, \ldots, x_n$, alors on note $E = \{x_1, \ldots, x_n\}$ et l'ensemble E est dit *fini*. L'ensemble vide se note $\emptyset$.

Si les éléments d'un ensemble A sont des éléments d'un ensemble B, alors on dit que « A est inclus dans B » ou que « A est un sous-ensemble (ou une partie) de B » et ceci se note $A \subset B$. Dans le cas contraire, c'est à dire lorsqu'il existe au moins un élément de A qui n'appartient pas à B, on dit que « A n'est pas inclus dans B » et ceci se note $A \not\subset B$. L'ensemble vide est inclus dans tous les ensembles. L'ensemble des éléments qui appartiennent à la fois à A et à B est appelé l'*intersection* de A et B et se note $A \cap B$. On peut alors écrire $A \cap B = \{x : x \in A \text{ et } x \in B\}$. L'ensemble des éléments qui appartiennent à A ou à B est appelé l'*union* de A et B et se note $A \cup B$ (attention ce *ou* n'est pas exclusif). On peut alors écrire $A \cup B = \{x : x \in A \text{ ou } x \in B\}$.

Plus généralement si $(A_i)_{i \in I}$ est une collection d'ensembles, où I est un ensemble quelconque d'indices, alors on définit l'intersection $\cap_{i \in I} A_i$ comme l'ensemble des éléments appartenant à tous les ensembles A_i, $i \in I$ et l'union $\cup_{i \in I} A_i$ comme l'ensemble des éléments qui appartiennent à au moins l'un des ensembles A_i, $i \in I$. Autrement dit $\cap_{i \in I} A_i = \{x : x \in A_i, \text{ pour tout } i \in I\}$ et $\cup_{i \in I} A_i = \{x : \text{il existe } i \in I \text{ tel que } x \in A_i\}$.

Une *partition* d'un ensemble E est une collection $(A_i)_{i \in I}$ de sous-ensembles de E tels que $A_i \cap A_j$ pour tous $i, j \in I$ vérifiant $i \neq j$ et tels que $E = \cup_{i \in I} A_i$.

Soit E un ensemble et A un sous-ensemble de E. Le *complémentaire dans* E de l'ensemble A est l'ensemble des éléments de E qui n'appartiennent pas à A. Celui-ci se note $E \setminus A$, autrement dit, $E \setminus A = \{x \in E : x \notin A\}$.

On note $\mathcal{P}(E)$ l'ensemble de tous les sous-ensembles de E.

Le produit $A \times B$ de deux ensembles est l'ensemble des couples (x, y) tels que $x \in A$ et $y \in B$. Autrement dit, il s'agit de l'ensemble $A \times B = \{(x, y) : x \in A \text{ et } y \in B\}$.

Proposition 1.1.1. *Soit E un ensemble et $(A_i)_{i \in I}$ une collection de sous-ensembles de E, alors*

1. $E \setminus (\cap_{i \in I} A_i) = \cup_{i \in I} E \setminus A_i$.
2. $E \setminus (\cup_{i \in I} A_i) = \cap_{i \in I} E \setminus A_i$.

Proof. Nous ne montrons que la première assertion. La seconde se démontre de la même façon. Pour montrer que ces deux ensembles sont égaux, on montre qu'un élément x appartient à l'ensemble $E \setminus (\cap_{i \in I} A_i)$ si et seulement si il appartient à l'ensemble $\cup_{i \in I} E \setminus A_i$. Pour cela, on procède par équivalence de la manière suivante :

$$\begin{aligned} x \in E \setminus (\cap_{i \in I} A_i) &\Leftrightarrow x \notin \cap_{i \in I} A_i \\ &\Leftrightarrow \exists\, j \in I, x \in E \setminus A_j \\ &\Leftrightarrow x \in \cup_{i \in I} E \setminus A_i\,. \end{aligned}$$

□

Une application d'un ensemble E dans un ensemble F est une relation qui à tout élément de l'ensemble E associe *un et un seul* élément de l'ensemble F. Une application sera souvent notée f, g, h,... Les ensembles E et F sont appelés les *ensembles de départ et d'arrivée* de l'application. La notation complète d'une application f de E dans F est alors $f : E \to F$. L'élément de F auquel on associe un élément $x \in E$ est noté $f(x)$. Il s'appelle l'*image* de x. Pour un élément $y \in F$, si il existe un élément $x \in E$ tel que $y = f(x)$, alors x est appelé un *antécédent* de y (celui-ci n'est pas nécessairement unique).

Une application $f : E \to F$ est dite :

1. *Injective* si tout élément $y \in F$ a *au plus* un antécédent par f dans E. Cela est équivalent à dire que si $x, x' \in E$ vérifient $f(x) = f(x')$, alors $x = x'$.
2. *Surjective* si tout élément $y \in F$ a *au moins* un antécédent par f dans E.
3. *Bijective* si elle est à la fois injective et surjective.

Exemples : a) L'application $f : \mathbb{R} \to \mathbb{R}$ telle que $f(x) = x^2$ pour tout $x \in \mathbb{R}$ n'est ni injective, ni surjective.
b) L'application $g : [0, 1] \to [0, 2]$ telle que $g(x) = 1 - x$ est injective mais n'est pas

surjective.
c) L'application $h : [-1, 1] \to [0, 1]$ telle que $h(x) = |x|$ est surjective mais n'est pas injective.
d) L'application $k : \mathbb{R} \to]0, +\infty[$ telle que $k(x) = e^x$ est une bijection.

Si $f : E \to F$ est une bijection, alors tout élément $y \in F$ admet un et un seul antécédent x par f dans E, $y = f(x)$. Cela permet de définir une autre application notée $f^{-1} : F \to E$ et appelée la *réciproque* de f par $f^{-1}(y) = x$, où x est l'antécédent de y par f dans E.

Definition 1.1.1. *Soit $f : E \to F$ une application.*
*L'*image *par f d'un sous-ensemble $A \subset E$ est le sous-ensemble de F défini par $f(A) = \{f(x) : x \in A\}$.*
*L'*image réciproque *par f d'un sous-ensemble $B \subset F$ est le sous-ensemble de E défini par $f^{-1}(B) = \{x \in E : f(x) \in B\}$.*

Attention ! il ne faut pas confondre la notation $f^{-1}(y)$, pour $y \in F$ qui n'a de sens que si f est une bijection avec la notation $f^{-1}(B)$ qui est valide pour toute application $f : E \to F$ et pour tout sous-ensemble $B \subset F$.
D'autre part, il faut distinguer la notation $f(x)$ qui est l'image de x par f (c'est un élément de F) de la notation $f(\{x\})$ qui est l'image de l'ensemble $\{x\}$ par f (c'est un sous-ensemble de F). On pourra d'ailleurs vérifier que $f(\{x\}) = \{f(x)\}$.

Remarquons que $f(\emptyset) = \emptyset$ et que $f^{-1}(\emptyset) = \emptyset$.

La proposition suivante consiste à montrer que l'image réciproque par une application préserve l'union et l'intersection.

Proposition 1.1.2. *Soit $f : E \to F$ une application et soit $(B_i)_{i\in I}$ une collection de sous-ensembles de F, alors*

1. $f^{-1}(\cup_{i\in I} B_i) = \cup_{i\in I} f^{-1}(B_i)$.
2. $f^{-1}(\cap_{i\in I} B_i) = \cap_{i\in I} f^{-1}(B_i)$.

Proof. Nous montrons seulement la première partie, l'autre est laissée en exercice. Supposons d'abord que $f^{-1}(\cup_{i\in I} B_i) = \emptyset$. S'il existe $x \in \cup_{i\in I} f^{-1}(B_i)$, alors il existe $j \in I$ tel que $x \in f^{-1}(B_j)$. Par conséquent, il existe $y \in B_j$ tel que $f(x) = y$. Mais puisque $y \in \cup_{i\in I} B_i$, par définition, $x \in f^{-1}(\cup_{i\in I} B_i)$, ce qui est une contradiction, donc $\cup_{i\in I} f^{-1}(B_i) = \emptyset$ et l'égalité 1. est vérifiée dans ce cas.

Supposons que $f^{-1}(\cup_{i\in I} B_i)$ ne soit pas vide. Soit $x \in f^{-1}(\cup_{i\in I} B_i)$. Alors il existe $y \in \cup_{i\in I} B_i$ tel que $f(x) = y$. De plus il existe $j \in I$ tel que $y \in B_j$ et par définition, $x \in f^{-1}(B_j)$, ce qui entraîne que $x \in \cup_{i\in I} f^{-1}(B_i)$. Nous avons donc montré que $f^{-1}(\cup_{i\in I} B_i) \subset \cup_{i\in I} f^{-1}(B_i)$.

Montrons l'autre inclusion. D'après ce qui vient d'être montré, si $\cup_{i\in I} f^{-1}(B_i) = \emptyset$, alors $f^{-1}(\cup_{i\in I} B_i) = \emptyset$. Supposons que $\cup_{i\in I} f^{-1}(B_i) \neq \emptyset$. Soit $x \in \cup_{i\in I} f^{-1}(B_i)$. Alors il existe $j \in I$ tel que $x \in f^{-1}(B_j)$. Par conséquent, il existe $y \in B_j$, tel que $f(x) = y$. Mais $y \in \cup_{i\in I} B_i$, donc $x \in f^{-1}(\cup_{i\in I} B_i)$ et nous avons montré l'autre inclusion. □

L'image par une application préserve l'union, mais pas l'intersection.

Proposition 1.1.3. *Soit $f : E \to F$ une application et soit $(A_i)_{i\in I}$ une collection de sous-ensembles de E, alors*

1. $f(\cup_{i\in I} A_i) = \cup_{i\in I} f(A_i)$.
2. $f(\cap_{i\in I} A_i) \subset \cap_{i\in I} f(A_i)$.

Proof. Soit $y \in f(\cup_{i\in I} A_i)$, alors il existe $x \in \cup_{i\in I} A_i$ tel que $y = f(x)$. De plus, puisque $x \in \cup_{i\in I} A_i$, il existe $j \in I$ tel que $x \in A_j$. Par conséquent, $y = f(x) \in f(A_j)$ et donc $y \in \cup_{i\in I} f(A_i)$.

Soit $y \in \cup_{i\in I} f(A_i)$, alors il existe $j \in I$ tel que $y \in f(A_j)$. Il existe alors $x \in A_j$ tel que $y = f(x)$. Mais puisque $x \in A_j$, en particulier, $x \in \cup_{i\in I} A_i$ et donc $y = f(x) \inf(\cup_{i\in I} A_i)$. La première assertion est démontrée.

Soit $y \in f(\cap_{i\in I} A_i)$, alors il existe $x \in \cap_{i\in I} A_i$ tel que $y = f(x)$. Comme $x \in \cap_{i\in I} A_i$, on a aussi que $x \in A_i$, pour tout $i \in I$ et donc $y = f(x) \in f(A_i)$, pour tout $i \in I$. On en déduit que $y \in \cap_{i\in I} f(A_i)$. □

Remarquons que dans la seconde assertion de la proposition 1.1.3, l'inclusion réciproque $\cap_{i\in I} f(A_i) \subset f(\cap_{i\in I} A_i)$ n'est pas vraie en général. En effet, soit $f : [-2, 2] \to [0, 4]$ l'application donnée par $f(x) = x^2$. Soient $A_1 = [-2, -1]$ et $A_2 = [1, 2]$. Alors $f(A_1) = f(A_2) = [1, 4]$ et en particulier $f(A_1) \cap f(A_2) \neq \emptyset$. Pourtant $A_1 \cap A_2 = \emptyset$ et par conséquent, $f(A_1 \cap A_2) = \emptyset$.

La preuve des propositions suivantes est laissée en exercice.

Proposition 1.1.4. *Soit $f : E \to F$ une application et soient A_1 et A_2 deux sous-ensembles de E. Montrer que si $A_1 \subset A_2$, alors $f(A_1) \subset f(A_2)$. Soient B_1 et B_2 deux sous-ensembles de F. Montrer que si $B_1 \subset B_2$, alors $f^{-1}(B_1) \subset f^{-1}(B_2)$.*

Proposition 1.1.5. *Soit $f : E \to F$ une application, soit A un sous-ensemble de E et soit B un sous-ensemble de F. Montrer que*

1. $f^{-1}(F \setminus B) = E \setminus f^{-1}(B)$,
2. $A \subset f^{-1}(f(A))$,

3. $f(f^{-1}(B) \subset B$.

1.2. Cardinalité

Definition 1.2.1. *On dit que deux ensembles A et B sont équipotents (ou ont le même cardinal) et on note $Card(A) = Card(B)$, si les ensembles A et B sont en bijection.*

Exemples :
1) $\mathrm{Card}(\{1, 2, 3, 4\}) = \mathrm{Card}(\{0, \pi, 3, 8\})$.
2) $\mathrm{Card}\,(\mathbb{N}) = \mathrm{Card}\,(2\mathbb{N})$.

La proposition suivante fournit un troisième exemple.

Proposition 1.2.1. *Pour tout ensemble E, $Card(\mathscr{P}(E)) = Card(\{0,1\}^E)$.*

Proof. On note F^E l'ensemble des applications de E dans F. Soit A un ensemble quelconque. La fonction indicatrice de A est définie par :

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

L'application $A \subset E \mapsto \mathbb{1}_A \in \{0,1\}^E$ est une bijection de $\mathscr{P}(E)$ dans $\{0,1\}^E$ (à vérifier). $\square$

On convient que $\mathrm{Card}\,(\emptyset) = 0$. Un ensemble A est dit fini si il est vide ou bien si il existe un entier fini $n \geq 1$ tel que $\mathrm{Card}\,(A) = \mathrm{Card}\,(\{1, \ldots, n\})$.

Si A est un ensemble fini, $\mathrm{Card}\,(A) = \mathrm{Card}\,(B)$ si et seulement si A et B ont même nombre d'éléments. On définira donc le cardinal d'un ensemble fini comme le nombre d'éléments qu'il contient.

On peut vérifier aussi que la notion d'équipotence possède les propriétés d'une relation d'équivalence (réflexive, symétrique et transitive).

Definition 1.2.2. *On dit que A a un cardinal inférieur ou égal à B si il existe une injection de A dans B. Ceci se note $Card(A) \leq Card(B)$.*

On dit que A a un cardinal supérieur ou égal à B si il existe une surjection de A dans B. Ceci se note $Card(A) \geq Card(B)$.

Si $Card(A) \leq Card(B)$ et $Card(A) \neq Card(B)$, alors on note $Card(A) < Card(B)$ et l'on dit que le cardinal de A est strictement inférieur à celui de B.

Si $Card(A) \geq Card(B)$ et $Card(A) \neq Card(B)$, alors on note $Card(A) > Card(B)$ et l'on dit que le cardinal de A est strictement supérieur à celui de B.

Remarquons que si $A \subset B$, alors Card$(A) \leq$ Card(B) et Card$(B) \geq$ Card(A).

Il ne faut pas prêter aux relations $\leq$ et $\geq$ les mêmes propriétés que celles de la relation d'ordre dans $\mathbb{R}$. On a toutefois les propriétés suivantes. Celles-ci nécessitent une démonstration que nous ne ferons pas ici.

Propriétés :

1) Si Card$(A) \leq$ Card(B), alors Card$(B) \geq$ Card(A).

2) Si Card$(A) \geq$ Card(B), alors Card$(B) \leq$ Card(A).

(La deuxième propriété est en fait équivalente à l'axiome du choix : pour tout ensemble E non vide, il existe une fonction $f : \mathscr{P}(E) \to E$ telle que pour tout $A \subset E$ non vide, $f(A) \in A$.)

Nous ne démontrerons pas non plus le théorème suivant.

Theorem 1.2.1 (Cantor-Bernstein)**.** *Si A et B sont deux ensembles vérifiant Card$(A) \leq$ Card(B) et Card$(B) \leq$ Card(A), alors Card$(A) =$ Card(B).*

Remarquons que d'après la propriété 2) ci-dessus, on a aussi le résultat suivant : Si A et B sont deux ensembles vérifiant Card$(A) \geq$ Card(B) et Card$(B) \geq$ Card(A), alors Card$(A) =$ Card(B).

Le résultat fondamental suivant met en évidence le fait qu'étant donné un ensemble non vide quelconque E, il existe toujours un ensemble dont cardinal est strictement supérieur à Card(E).

Theorem 1.2.2 (Cantor)**.** *Si E un ensemble non vide, alors Card$(E) <$ Card$(\mathscr{P}(E))$.*

Preuve D'une part, l'injection $x \in E \mapsto \{x\} \in \mathscr{P}(E)$ montre que Card$(E) \leq$ Card$(\mathscr{P}(E))$. D'autre part, supposons qu'il existe une surjection $\varphi : E \to \mathscr{P}(E)$ et définissons $A = \{x \in E,\, x \notin \varphi(x)\}$. Puisque φ est surjective, il existe $a \in E$ tel que $\varphi(a) = A$. Alors, ou bien $a \in A$ et $a \notin \varphi(a) = A$, ce qui est une contradiction. Ou bien $a \notin \varphi(a)$ et $a \in \varphi(a) = A$, ce qui est encore une contradiction. Il n'existe donc pas de surjection de E dans $\mathscr{P}(E)$. $\square$

Definition 1.2.3. *Un ensemble est dit infini (ou de cardinal infini) si il n'est pas fini.*

Exemples :
1) $\mathbb{N}$ et $\mathbb{R}$ sont infinis.
2) L'ensemble des fonctions continues de $[0, 1]$ dans $\mathbb{R}$ est infini.

Proposition 1.2.2. *Un ensemble E est infini si et seulement si il existe une injection de $\mathbb{N}$ dans E (c'est à dire Card$(\mathbb{N}) \leq$ Card(E)).*

En conséquence, si Card$(E) <$ Card$(\mathbb{N})$, alors E est un ensemble fini.

Preuve Celle-ci est laissée en exercice. □

Definition 1.2.4. *Un ensemble E est dit dénombrable si il est en bijection avec $\mathbb{N}$, c'est à dire $Card(E) = Card(\mathbb{N})$. Il est dit au plus dénombrable si il existe une injection de E dans $\mathbb{N}$, c'est à dire $Card(E) \leq Card(\mathbb{N})$.*

Exemples :
1) $5\mathbb{N}$ est dénombrable. Plus généralement toute partie infinie de $\mathbb{N}$ est dénombrable.
2) $\mathbb{Z}$ est dénombrable.
3) $\mathbb{N}^2 = \mathbb{N} \times \mathbb{N}$ est dénombrable et par récurrence $\mathbb{N}^k$, pour tout $k \geq 1$ est dénombrable.
4) Plus généralement si $E_1, E_2, \ldots, E_n$ sont des ensembles dénombrables (avec n fini), alors leur produit cartésien $E_1 \times E_2 \times \cdots \times E_n$ est dénombrable.
5) $\mathbb{Q}$ est dénombrable.
6) $\mathscr{P}(\mathbb{N})$ et $\{0,1\}^{\mathbb{N}}$ ne sont pas dénombrables car comme nous l'avons déjà vu, $\mathrm{Card}\,(\mathbb{N}) < \mathrm{Card}\,(\mathscr{P}(\mathbb{N}))$ et $\mathrm{Card}\,(\{0,1\}^{\mathbb{N}}) = \mathrm{Card}\,(\mathscr{P}(\mathbb{N}))$.

Remarque : Un ensemble E est dénombrable si et seulement si il peut être écrit sous forme d'une suite, c'est à dire $E = \{x_n : n \in \mathbb{N}\}$ dans laquelle les termes x_n, $n \in \mathbb{N}$ sont deux à deux distincts.

Proposition 1.2.3. *Une union dénombrable d'ensembles dénombrables est dénombrable.*

Preuve : Soient $(E_n)_{n\geq 1}$ une suite d'ensembles dénombrables et $E = \cup_{n\geq 1} E_n$. Puisque $E_1 \subset E$, on a $\mathrm{Card}\,(\mathbb{N}) = \mathrm{Card}\,(E_1) \leq \mathrm{Card}\,(E)$. D'autre part, pour tout $n \geq 1$, soit $\varphi_n : \mathbb{N} \to E_n$ surjective. Alors l'application $\varphi : \mathbb{N}^2 \to E$ définie par $\varphi(n,m) = \varphi_n(m)$ est surjective. En effet, soit $x \in E$, alors il existe $n \geq 1$ tel que $x \in E_n$. Puisque $\varphi_n : \mathbb{N} \to E_n$ est surjective, il existe $m \in \mathbb{N}$ tel que $\varphi_n(m) = x$. Ainsi $\varphi_n(n) = \varphi(n,m) = x$ et (n,m) est un antécédent de X. On en déduit que $\mathrm{Card}\,(\mathbb{N}^2) = \mathrm{Card}\,(\mathbb{N}) \geq \mathrm{Card}\,(E)$. □

Proposition 1.2.4. *Un produit cartésien dénombrable d'ensembles qui contiennent chacun au moins deux éléments n'est pas dénombrable.*

Preuve : Il suffit de considérer $\{0,1\}^{\mathbb{N}}$ qui n'est pas dénombrable d'après l'exemple considéré plus haut. □

Theorem 1.2.3. $\mathbb{R}$ *n'est pas dénombrable.*

Preuve : Nous allons utiliser l'argument diagonal de Cantor pour démontrer ceci, mais il existe d'autres démonstrations (voir le résultat suivant).

Tout d'abord, puisque $[0,1] \subset \mathbb{R}$, il suffit de démontrer que $[0,1]$ n'est pas dénombrable. Supposons que ce soit le cas. Alors on peut écrire $[0,1]$ sous forme d'une suite, c'est à dire $[0,1] = \{x_n,\, n \in \mathbb{N}\}$. Pour tout $n \in \mathbb{N}$, on note $0, x_n^1 x_n^2 x_n^3 \ldots$ une écriture décimale de x_n (lorsque x_n est décimal, la suite $(x_n^k)_k$ vaut ou bien toujours 0, ou bien toujours 9 à partir d'un certain rang). On définit alors $x = 0, x^1 x^2 x^3 \ldots$ par son écriture décimale, de la manière suivante : pour tout $k \in \mathbb{N}$, $x^k = 2$, si $x_k^k = 1$ et $x^k = 1$ sinon. Par construction, $x \neq x_n$, pour tout $n \in \mathbb{N}$, donc $x \notin [0,1]$ ce qui est une contradiction. $\square$

Le résultat suivant est plus fort que le précédent car il donne la cardinalité de $\mathbb{R}$.

Theorem 1.2.4. *Les ensembles $\mathbb{R}$ et $\mathscr{P}(\mathbb{N})$ sont équipotents.*

Preuve : D'après la proposition 1.2.1, $\mathscr{P}(\mathbb{N})$ a même cardinal que $\{0,1\}^{\mathbb{N}}$. Il suffit donc de montrer que $\mathbb{R}$ et $\{0,1\}^{\mathbb{N}}$ sont équipotents.

Pour tout élément $a = (a_n)_{n \geq 0} \in \{0,1\}^{\mathbb{N}}$ on associe le réel $\varphi(a) = \sum_{n=0}^{\infty} \frac{a_n}{3^n}$. Alors l'application $\varphi : \{0,1\}^{\mathbb{N}} \to \mathbb{R}$ est injective. En effet, la suite $a = (a_n)_{n \geq 0}$ est unique dans le développement en base 3 de $\varphi(a)$ qui est : $a_0, a_1 a_2 \ldots$. On a donc $\mathrm{card}(\mathscr{P}(\mathbb{N})) \leq \mathrm{card}(\mathbb{R})$.

D'autre part, nous avons déjà vu que $\mathbb{Q}$ était dénombrable. Par conséquent, $\mathrm{card}(\mathscr{P}(\mathbb{N})) = \mathrm{card}(\mathscr{P}(\mathbb{Q}))$. Soit $\psi : \mathbb{R} \to \mathscr{P}(\mathbb{Q})$ l'application définie par $\psi(x) = \{q \in \mathbb{Q} : q < x\}$. Alors ψ est une injection de $\mathbb{R}$ dans $\mathscr{P}(\mathbb{Q})$ et par conséquent, $\mathrm{card}(\mathbb{R}) \leq \mathrm{card}(\mathscr{P}(\mathbb{Q})) = \mathrm{card}(\mathscr{P}(\mathbb{N}))$. Le résultat se déduit alors du théorème de Cantor-Bernstein. $\square$

1.3. Espaces mesurables

Désormais lorsqu'aucune confusion ne sera possible, le complémentaire du sous-ensemble A d'un ensemble E sera noté A^c, c'est-à-dire $A^c = \{x \in E : x \notin A\}$. Un sous-ensemble de $\mathscr{P}(E)$ sera appelé une *famille de parties de E*.

Definition 1.3.1. *On appelle tribu (ou σ-algèbre) sur E une famille de parties $\mathscr{A} \subset \mathscr{P}(E)$ telle que :*

1. $\emptyset \in \mathscr{A}$,
2. *si $A \in \mathscr{A}$, alors $A^c \in \mathscr{A}$ (stabilité par passage au complémentaire),*
3. *si $(A_n)_{n \geq 1}$ est une suite d'éléments de $\mathscr{A}$, alors $\cup_{n \geq 1} A_n \in \mathscr{A}$ (stabilité par union dénombrable).*

On rappelle que pour toute suite $(B_n)_{n \geq 1}$ est une suite d'éléments de $\mathscr{P}(E)$,

$$\cap_{n \geq 1} B_n = \left(\cup_{n \geq 1} B_n^c\right)^c .$$

On remarquera alors qu'étant donné le point 2., dans la définition ci-dessus, les points 1. et 3. peuvent être remplacés respectivement par :

1′. $E \in \mathscr{A}$,

3′. si $(A_n)_{n\geq 1}$ est une suite d'éléments de $\mathscr{A}$, alors $\cap_{n\geq 1} A_n \in \mathscr{A}$ (stabilité par intersection dénombrable).

Exemples :
1. $\mathscr{P}(E)$ est une tribu sur E. On l'appelle la tribu fine.
2. $\{\emptyset, E\}$ est une tribu sur E. On l'appelle la tribu grossière.
3. Pour tout $A \subset E$, $\{\emptyset, E, A, A^c\}$ est une tribu sur E.
4. $\{A \in \mathscr{P}(E)$: ou bien A est au plus dénombrable ou bien A^c est au plus dénombrable$\}$ est une tribu sur E.

L'ensemble E muni d'une tribu $\mathscr{A}$ est noté $(E, \mathscr{A})$ et est appelé un *espace mesurable*. Les éléments de la tribu $\mathscr{A}$ sont appelés les *ensembles mesurables*.

Proposition 1.3.1. *Une intersection quelconque de tribus sur E est une tribu sur E.*

Preuve. Soit I un ensemble quelconque et soit $\{\mathscr{A}_i,\, i \in I\}$ un ensemble de tribus sur E. Posons $\mathscr{A} = \cap_{i\in I} \mathscr{A}_i$. Pour tout $i \in I$, $\emptyset \in \mathscr{A}_i$, donc $\emptyset \in \mathscr{A}$. Soit $A \in \mathscr{A}$, alors pour tout $i \in I$, $A \in \mathscr{A}_i$ et puisque les $\mathscr{A}_i$ sont des tribus, pour tout $i \in I$, $A^c \in \mathscr{A}_i$ et par conséquent, $A^c \in \mathscr{A}$. Soit enfin $(A_n)_{n\geq 1}$ une suite d'éléments de $\mathscr{A}$. Alors, pour tout $i \in I$ et pour tout $n \geq 1$, $A_n \in \mathscr{A}_i$ et puisque les $\mathscr{A}_i$ sont des tribus, pour tout $i \in I$, $\cup_{n\geq 1} A_n \in \mathscr{A}_i$, donc $\cup_{n\geq 1} A_n \in \mathscr{A}$. □

Remarque : L'union de tribus n'est pas nécessairement une tribu. En fait on peut montrer que si $\mathscr{A}$ et $\mathscr{B}$ sont deux tribus sur E, alors $\mathscr{A} \cup \mathscr{B}$ est une tribu si et seulement si $\mathscr{A} \subset \mathscr{B}$ ou $\mathscr{A} \subset \mathscr{B}$. Ceci est laissé en exercice.

Proposition 1.3.2. *Soit $\mathscr{C} \subset \mathscr{P}(E)$ une famille de parties de E. L'intersection de toutes les tribus sur E qui contiennent $\mathscr{C}$ est elle-même une tribu sur E. On l'appelle la tribu engendrée par $\mathscr{C}$ et on la note $\sigma(\mathscr{C})$,*

$$\sigma(\mathscr{C}) = \bigcap_{\substack{\mathscr{A} \text{ tribu} \\ \mathscr{C} \subset \mathscr{A}}} \mathscr{A}\,.$$

Preuve. C'est une conséquence immédiate de la proposition qui précède. □

Remarquons que $\sigma(\mathscr{C})$ est la plus petite tribu sur E contenant la famille de parties $\mathscr{C}$.

Exemples :
1. Soit $A \subset E$, alors $\sigma(\{A\}) = \{\emptyset, E, A, A^c\}$.
2. La tribu engendrée par l'ensemble des singletons sur E est égale à la tribu engendrée par les ensembles dénombrables, c'est-à-dire,

$\sigma(\{\{x\} : x \in E\}) = \sigma(\{A \in \mathscr{P}(E) : A$ est au plus dénombrable.$\} =$
$\{A \in \mathscr{P}(E) :$ ou bien A est au plus dénombrable ou bien A^c est au plus dénombrable$\}$.

Rappelons qu'une topologie sur l'ensemble E est une famille de parties de $\mathscr{O} \subset E$ qui vérifie :

(*i*) $\emptyset \in \mathscr{O}$ et $E \in \mathscr{O}$,

(*ii*) si I est un ensemble d'indices quelconque et $\{A_i,\, i \in I\}$ un ensemble d'éléments de $\mathscr{O}$, alors $\cup_{i\in I} A_i \in \mathscr{O}$ (stabilité par union quelconque).

(*iii*) pour tout $n \geq 1$ et pour tous éléments $A_1, \ldots, A_n$ de $\mathscr{O}$, $\cap_{k=1}^n A_k \in \mathscr{O}$ (stabilité par intersection finie).

L'ensemble E muni d'une topologie $\mathscr{O}$ est appelé un espace topologique. On le notera $(E, \mathscr{O})$ lorsqu'aucune confusion n'est possible avec un espace mesurable. Les éléments d'une topologie sont appelés les *ouverts* de E. Les complémentaires des ouverts de E sont appelés les *fermés* de E.

De même que pour les tribus, une intersection quelconque de topologies sur E est une topologie sur E. Par conséquent, si $\mathscr{C} \subset \mathscr{P}(E)$ est une famille de parties de E, il existe une plus petite topologie sur E qui contient $\mathscr{C}$. Il s'agit de l'intersection de toutes les topologies qui contiennent $\mathscr{C}$.

Nous ne considérerons dans ce cours que le cas où $E = \mathbb{R}^d$ muni de la topologie engendrée par les pavés ouverts. Par conséquent la notation $\mathscr{O}$ sur $\mathbb{R}^d$ désgnera toujours cette topologie.

Definition 1.3.2. *Soit $(E, \mathscr{O})$ un espace topologique. On appelle tribu borélienne et on note $\mathscr{B}(E)$ la tribu engendrée par les ouverts de E, c'est-à-dire,*

$$\mathscr{B}(E) = \sigma(\mathscr{O})\,.$$

Les éléments de $\mathscr{B}(E)$ sont appelés les *boréliens* de E. Remarquons que la tribu borélienne est aussi engendrée par l'ensemble des ensembles fermés.

Dans ce cours, l'ensemble des réels $\mathbb{R}$ sera toujours muni de la topologie engendrée par les intervalles ouverts. Elle correspond aussi à la topologie est engendrée par la distance euclidienne sur $\mathbb{R}$. La tribu borélienne $\mathscr{B}(\mathbb{R})$ est donc définie relativement à cette topologie. En particulier tout intervalle de $\mathbb{R}$ est un borélien. Les boréliens de $\mathbb{R}$ peuvent toutefois être des ensembles très compliqués et il est impossible de les décrire tous. Remarquons que « très peu » de sous-ensembles de $\mathscr{P}(\mathbb{R})$ sont des boréliens puisque l'on peut montrer que $\mathrm{card}(\mathscr{B}(\mathbb{R})) = \mathrm{card}(\mathbb{R})$. Cependant la démonstration est assez difficile et nous ne la ferons pas ici.

Nous avons vu que la topologie $\mathscr{O}$ engendre la tribu borélienne. C'est même la définition de cette tribu. Nous allons voir maintenant que des familles de parties plus petites et plus élémentaires que $\mathscr{O}$ engendrent cette tribu.

Proposition 1.3.3.

$$\begin{aligned}\mathscr{B}(\mathbb{R}) &= \sigma(\{]\alpha,\beta[:\alpha<\beta,\,\alpha,\beta\in\mathbb{Q}\}) \\ &= \sigma(\{]-\infty,a[:a\in\mathbb{Q}\})=\sigma(\{]a,+\infty[:a\in\mathbb{Q}\}) \\ &= \sigma(\{]-\infty,a]:a\in\mathbb{Q}\})=\sigma(\{[a,+\infty[:a\in\mathbb{Q}\}).\end{aligned}$$

Preuve. Si $\alpha,\beta\in\mathbb{Q}$ et $\alpha<\beta$, alors $]\alpha,\beta[\in\mathscr{O}$, donc $\sigma(\{]\alpha,\beta[:\alpha<\beta,\,\alpha,\beta\in\mathbb{Q}\})\subset\mathscr{B}(\mathbb{R})$. Réciproquement, rappelons que tout ouvert $A\in\mathscr{O}$ de $\mathbb{R}$ peut s'écrire $A=\cup_{p,q\in\mathbb{Q},\,]p,q[\subset A}]p,q[$. On a alors, $A\in\sigma(\{]\alpha,\beta[:\alpha<\beta,\,\alpha,\beta\in\mathbb{Q}\})$ et donc $\mathscr{B}(\mathbb{R})\subset\sigma(\{]\alpha,\beta[:\alpha<\beta,\,\alpha,\beta\in\mathbb{Q}\})$.

Si $a\in\mathbb{Q}$, alors $]-\infty,a[\in\mathscr{O}$ et donc $\sigma(\{]-\infty,a[:a\in\mathbb{Q}\})\subset\mathscr{B}(\mathbb{R})$. D'autre part, si $\alpha,\beta\in\mathbb{Q}$, alors

$$]\alpha,\beta[=]-\infty,\beta[\setminus]-\infty,\alpha]=]-\infty,\beta[\setminus\cap_{n\geq1}]-\infty,\alpha+1/n[.$$

Par conséquent, $\mathscr{B}(\mathbb{R})=\sigma(\{]\alpha,\beta[:\alpha<\beta,\,\alpha,\beta\in\mathbb{Q}\})\subset\sigma(\{]-\infty,a[:a\in\mathbb{Q}\})$.

Les trois autres égalités se démontrent de la même façon. □

Soit maintenant $d\geq1$. Rappelons que la topologie sur $\mathbb{R}^d$ est engendrée par l'ensemble des pavés ouverts. Plus précisément, si $\Lambda_d=\{]a_1,b_1[\times]a_2,b_2[\times\cdots\times]a_d,b_d[: a_i<b_i,\,a_i,b_i\in\mathbb{Q}\,,i=1,\ldots,d\}$, alors tout A ouvert non vide de $\mathbb{R}^d$ est de la forme :

$$A=\bigcup_{\substack{P\,\in\,\Lambda_d\\ P\,\subset\,A}}P\,.$$

Cette topologie correspond aussi à la topologie engendrée par la distance euclidienne sur $\mathbb{R}^d$. On montre alors de la même manière qu'à la proposition précédente le résultat suivant.

Proposition 1.3.4. *La tribu borélienne sur $\mathbb{R}^d$ est engendrée par l'ensemble des pavés ouverts dont les extrémités sont rationnelles, c'est à dire,*

$$\mathscr{B}(\mathbb{R}^d)=\sigma(\{]a_1,b_1[\times]a_2,b_2[\times\cdots\times]a_d,b_d[:a_i,b_i\in\mathbb{Q},\,i=1,\ldots,d\})\,.$$

Nous aurons souvent besoin de considérer l'ensemble des réels auquel on ajoute $-\infty$ et $+\infty$, c'est-à-dire,

$$\overline{\mathbb{R}}=\mathbb{R}\cup\{-\infty,+\infty\}\,.$$

On prolonge ensuite l'ordre total $\leq$ sur $\mathbb{R}$ en posant $-\infty\leq x\leq+\infty$, pour tout $x\in\overline{\mathbb{R}}$. On munit alors $\overline{\mathbb{R}}$ de la topologie engendrée par $\mathscr{O}\cup\{[-\infty,a[:a\in\mathbb{R}\}\cup\{]a,+\infty]:a\in\mathbb{R}\}$. On peut alors montrer le résultat suivant.

Proposition 1.3.5.

$$\begin{aligned}\mathscr{B}(\overline{\mathbb{R}}) &= \sigma(\{[-\infty, a[: a \in \mathbb{Q}\}) = \sigma(\{]a, +\infty] : a \in \mathbb{Q}\}) \\ &= \sigma(\mathscr{B}(\mathbb{R}), \{-\infty\}, \{+\infty\})\,.\end{aligned}$$

On considérera aussi parfois la tribu borélienne sur $\mathbb{R}_+$, c'est à dire la tribu engendrée par les ouverts de $\mathbb{R}_+$ que l'on peut définir comme $\mathscr{B}(\mathbb{R}_+) = \{A \in \mathscr{B}(\mathbb{R}) : A \subset \mathbb{R}_+\}$.

De même, en posant $\overline{\mathbb{R}}_+ = \mathbb{R}_+ \cup \{+\infty\}$, on peut définir la tribu borélienne sur $\mathbb{R}_+$ par $\mathscr{B}(\overline{\mathbb{R}}_+) = \sigma(\mathscr{B}(\mathbb{R}_+), \{+\infty\})$.

Enfin on étendra la multiplication sur $\overline{\mathbb{R}}_+$ en posant $(+\infty) \times 0 = 0 \times (+\infty) = 0$.

1.4. Applications mesurables

Soient E et F deux ensembles quelconques et $f : E \to F$ une application. Pour tout sous-ensemble $C \subset F$, on rappelle que

$$f^{-1}(C) = \{x \in E : f(x) \in C\}$$

Pour toute famille de parties $\mathscr{C} \subset \mathscr{P}(F)$, nous noterons alors,

$$f^{-1}(\mathscr{C}) = \{f^{-1}(C),\, C \in \mathscr{C}\}\,.$$

(On fera attention à ne pas confondre cette notation avec celle de la réciproque d'une fonction bijective qui est notée parfois de même.)

On vérifiera que pour tout application $f : E \to F$ et pour toute suite finie ou dénombrable (B_n) de parties de F,

$$\cap f^{-1}(B_n) = f^{-1}(\cap B_n) \quad \text{et} \quad \cup f^{-1}(B_n) = f^{-1}(\cup B_n)\,.$$

Proposition 1.4.1. *Si $\mathscr{B}$ est une tribu sur F, alors la famille de parties $f^{-1}(\mathscr{B})$ est une tribu sur E. On l'appelle l'image réciproque de la tribu $\mathscr{B}$ par f.*

Preuve. 1. $f^{-1}(\emptyset) = \emptyset \in f^{-1}(\mathscr{B})$.
2. Soit $A \in f^{-1}(\mathscr{B})$. Alors il existe $B \in \mathscr{B}$, tel que $A = f^{-1}(B)$. On peut alors vérifier que $A^c = f^{-1}(B^c)$, donc $f^{-1}(B^c) \in f^{-1}(\mathscr{B})$, puisque $B^c \in \mathscr{B}$.
3. Soit (A_n) une suite d'éléments de $f^{-1}(\mathscr{B})$, alors il existe une suite (B_n) de $\mathscr{B}$ telle que $A_n = f^{-1}(B_n)$, pour tout n. De plus $\cup A_n = \cup f^{-1}(B_n) = f^{-1}(\cup B_n) \in f^{-1}(\mathscr{B})$.
□

Remarque : Si $\mathscr{A}$ est une tribu sur E, alors la famille de parties $\{f(A) : A \in \mathscr{A}\}$ n'est pas une tribu en général (l'image d'une tribu n'est pas une tribu). En revanche, la famille de parties

$$\{B \subset F : f^{-1}(B) \in \mathscr{A}\}$$

est une tribu sur F. On l'appelle la tribu image de $\mathscr{A}$ par l'application f. (Ceci est à vérifier en exercice.)

On rappelle que pour une famille de parties $\mathscr{C}$ de F, $\sigma(\mathscr{C})$ est la tribu sur F engendrée par $\mathscr{C}$.

Proposition 1.4.2. *Soit F un ensemble quelconque et soit $\mathscr{C}$ une famille de parties de F, alors pour tout ensemble E et pour toute application $f : E \to F$,*

$$\sigma(f^{-1}(\mathscr{C})) = f^{-1}(\sigma(\mathscr{C}))\,.$$

Preuve. On remarque que $f^{-1}(\mathscr{C}) \subset f^{-1}(\sigma(\mathscr{C}))$. Comme $f^{-1}(\sigma(\mathscr{C}))$ est une tribu contenant la famille $f^{-1}(\mathscr{C})$, elle contient la tribu engendrée par cette famille, c'est-à-dire $\sigma(f^{-1}(\mathscr{C})) \subset f^{-1}(\sigma(\mathscr{C}))$.

Pour montrer l'autre inclusion, il suffit de montrer que pour tout $B \in \sigma(\mathscr{C})$, $f^{-1}(B) \in \sigma(f^{-1}(\mathscr{C}))$. Soit $\mathscr{B} = \{B \subset F,\, f^{-1}(B) \in \sigma(f^{-1}(\mathscr{C}))\}$. Alors $\mathscr{B}$ est la tribu est la tribu image par f de la tribu $\sigma(f^{-1}(\mathscr{C}))$. C'est une tribu qui contient $\mathscr{C}$. Elle contient donc $\sigma(\mathscr{C})$, ce qui montre l'autre inclusion. □

Definition 1.4.1. *Soient $(E, \mathscr{A})$ et $(F, \mathscr{B})$ deux espaces mesurables. Une application $f : E \to F$ est dite $(\mathscr{A}, \mathscr{B})$-mesurable si $f^{-1}(\mathscr{B}) \subset \mathscr{A}$. (Lorsqu'il n'y aura pas d'ambiguïté, nous dirons simplement de f qu'elle est mesurable.)*

La notation $f : (E, \mathscr{A}) \to (F, \mathscr{B})$ signifie que f est une application $(\mathscr{A}, \mathscr{B})$-mesurable.

Lorsque les ensembles E et F sont des espaces topologiques munis de leur tribus boréliennes respectives, l'application f est appelée une fonction borélienne.

Dans la suite, l'espace d'arrivée F sera souvent $\mathbb{R}$, $\mathbb{R}_+$, $\overline{\mathbb{R}}$, $\overline{\mathbb{R}}_+$, $\mathbb{R}^d$, ou $\mathbb{C}$. Dans ce cas et lorsque rien ne sera précisé, nous considérerons que cet espace est muni de sa tribu borélienne.

Si la tribu dont est muni l'espace d'arrivée est engendrée par une famille $\mathscr{C}$ de parties de F, alors pour montrer la mesurabilité d'une application $f : E \to F$, il suffit de se restreindre à $\mathscr{C}$ comme le montre la proposition suivante.

Proposition 1.4.3. *Soient $(E, \mathscr{A})$ et $(F, \mathscr{B})$ deux espaces mesurables. Soit $\mathscr{C} \subset \mathscr{B}$ telle que $\mathscr{B} = \sigma(\mathscr{C})$. Alors f est $(\mathscr{A}, \mathscr{B})$-mesurable si et seulement si $f^{-1}(\mathscr{C}) \subset \mathscr{A}$.*

Preuve. Il est clair que si $f^{-1}(\mathscr{B}) \subset \mathscr{A}$, alors $f^{-1}(\mathscr{C}) \subset f^{-1}(\mathscr{B}) \subset \mathscr{A}$. Réciproquement si $f^{-1}(\mathscr{C}) \subset \mathscr{A}$, alors $\sigma(f^{-1}(\mathscr{C})) \subset \mathscr{A}$, puisque $\mathscr{A}$ est une tribu contenant $f^{-1}(\mathscr{C})$. On déduit alors de la proposition 1.4.2 que $f^{-1}(\sigma(\mathscr{C})) \subset \mathscr{A}$. □

Corollary 1.4.1. *Toute fonction monotone $f : \mathbb{R} \to \mathbb{R}$ est borélienne.*

Preuve. Il suffit de remarquer que si I est un intervalle, alors $f^{-1}(I)$ est aussi un intervalle. Comme $\mathscr{B}(\mathbb{R})$ est engendrée par les intervalles, f est mesurable. $\square$

Notation importante : Dans le cas où $F = \mathbb{R}$, on utilisera les notations suivantes :

$$\{f < a\} = f^{-1}(]-\infty, a[) \qquad \{f \leq a\} = f^{-1}(]-\infty, a])$$
$$\{f > a\} = f^{-1}(]a, +\infty[) \qquad \{f \geq a\} = f^{-1}([a, +\infty[)$$
$$\{f = a\} = f^{-1}(\{a\})\,.$$

De même, si $F = \overline{\mathbb{R}}$, alors on notera $\{f < a\} = f^{-1}([-\infty, a[)$,...

Proposition 1.4.4. *Soient E et F deux espaces topologiques munis de leurs tribus boréliennes respectives $\mathscr{B}(E)$ et $\mathscr{B}(F)$. Alors toute fonction continue $f : E \to F$ est borélienne.*

Preuve. Notons $\mathscr{O}$ la topologie de F. Par définition, la fonction f est continue si et seulement si pour tout ouvert $O \subset \mathscr{O}$ de F, $f^{-1}(O)$ est un ouvert de E, donc un borélien de E. Par conséquent, $f^{-1}(\mathscr{O}) \subset \mathscr{B}(E)$ et f est borélienne d'après la proposition 1.4.3. $\square$

Proposition 1.4.5. *Soient $(E, \mathscr{A})$, $(F, \mathscr{B})$ et $(G, \mathscr{C})$ trois espaces mesurables. Si $f : E \to F$ et $g : F \to G$ sont mesurables, alors $g \circ f$ est mesurable.*

Preuve. Remarquons que $(g \circ f)^{-1}(\mathscr{C}) = f^{-1}(g^{-1}(\mathscr{C}))$. D'autre part puisque $g^{-1}(\mathscr{C}) \subset \mathscr{B}$ et $f^{-1}(\mathscr{B}) \subset \mathscr{A}$, on a $f^{-1}(g^{-1}(\mathscr{C})) \subset \mathscr{A}$. $\square$

Application : On obtient comme conséquence des propositions 1.4.4 et 1.4.5 que
1. Si $f : E \to \mathbb{R}$ est mesurable, alors $|f|$ est mesurable.
2. Si $f : E \to \mathbb{R}^*$ est mesurable, alors $1/f$ est mesurable.

Proposition 1.4.6. *Soit $f = (f_1, f_2) : E \to \mathbb{R}^2$. Alors f est mesurable si et seulement si f_1 et f_2 le sont.*

Preuve. Soit $\pi_i : \mathbb{R}^2 \to \mathbb{R}$, $i = 1, 2$ les projections canoniques définies par $\pi_i(x_1, x_2) = x_i$. Celles-ci sont continues, donc mesurables. De plus $f_i = \pi_i \circ f$. Par conséquent, si f est mesurable, alors les f_i le sont.

Réciproquement, supposons que f_1 et f_2 sont mesurables. Alors pour tout pavé $I_1 \times I_2$ de $\mathbb{R}^2$, $f^{-1}(I_1 \times I_2) = f_1^{-1}(I_1) \cap f_2^{-1}(I_2) \in \mathscr{A}$. Puisque $\mathscr{B}(\mathbb{R}^2)$ est engendrée par les pavés, le résultat vient de la proposition 1.4.3. $\square$

Proposition 1.4.7. *Si f et g sont des fonctions mesurables de $(E, \mathscr{A})$ dans $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, alors,*

1. *$\alpha f + g$ est mesurable, pour tout $\alpha \in \mathbb{R}$.*

2. *fg est mesurable.*

Les fonctions mesurables $(E,\mathscr{A})$ dans $(\mathbb{R},\mathscr{B}(\mathbb{R}))$ forment donc une $\mathbb{R}$-algèbre.

Preuve. 1. On peut écrire $\alpha f + g = \psi \circ \varphi$, où $\varphi(x) = (f(x), g(x)) \in \mathbb{R}^2$ pour tout $x \in E$ et $\psi(x,y) = \alpha s + t$, pour tout $(s,t) \in \mathbb{R}^2$. La fonction ψ est continue de $\mathbb{R}^2$ dans $\mathbb{R}$, donc elle est $(\mathscr{B}(\mathbb{R}^2),\mathscr{B}(\mathbb{R}))$-mesurable. Comme f et g sont mesurables, la proposition 1.4.5 implique que la fonction φ est mesurable. Par conséquent, la fonction $\alpha f + g$ est mesurable comme composée de fonctions mesurables. Pour montrer le point 2., on raisonne de même en prenant $\psi : (x,y) \to xy$, qui est aussi continue. $\square$

On rappelle que pour un sous-ensemble $A \subset E$, la fonction indicatrice de A est définie par,

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

Definition 1.4.2. *Soit $(E,\mathscr{A})$ un espace mesurable. On appelle fonction étagée toute application $f : E \to \mathbb{R}$ telle que $f = \sum_{i=1}^n \alpha_i \mathbb{1}_{A_i}$, où $n \geq 1$ est un entier quelconque et où, pour tout $i = 1,\ldots,n$, $\alpha_i \in \mathbb{R}$ et $A_i \in \mathscr{A}$.*

La classe des fonctions étagées qui sera très importante par la suite. La proposition suivante est une application des résultats précédents.

Proposition 1.4.8. *Les fonctions étagées sont mesurables.*

Proof. Soit $(E,\mathscr{A})$ un espace mesurable et $A \subset E$. On montre d'abord que l'application $\mathbb{1}_A : (E,\mathscr{A}) \to (\mathbb{R},\mathscr{B}(\mathbb{R}))$ est mesurable si et seulement si $A \in \mathscr{A}$. En effet, pour tout $B \in \mathscr{B}(\mathbb{R})$, $(\mathbb{1}_A)^{-1}(B) = \emptyset$ si $0 \notin B$ et $1 \notin B$, $(\mathbb{1}_A)^{-1}(B) = E$ si $0 \in B$ et $1 \in B$, $(\mathbb{1}_A)^{-1}(B) = A$ si $0 \notin B$ et $1 \in B$ et $(\mathbb{1}_A)^{-1}(B) = A^c$ si $0 \in B$ et $1 \notin B$. On conclut alors grâce à la proposition 1.4.7. $\square$

On a alors la définition équivalente suivante d'une fonction étagée.

Definition 1.4.3. *Soit $(E,\mathscr{A})$ un espace mesurable. On appelle fonction étagée toute application mesurable $f : E \to \mathbb{R}$ qui prend un nombre fini de valeurs.*

On pourra vérifier la proposition suivante à titre d'exercice.

Proposition 1.4.9. *Soit $(E,\mathscr{A})$ une espace mesurable. Une application $f : E \to \mathbb{C}$ est mesurable si et seulement si Re(f) et Im(f) le sont. (On rappelle que par défaut, $\mathbb{C}$ est muni de sa tribu borélienne $\mathscr{B}(\mathbb{C})$.)*

Rappel : Pour une suite $(a_n)_{n\geq 0}$ à valeurs dans $\overline{\mathbb{R}}$, remarquons que $(\inf_{k\geq n} a_k)_{n\geq 0}$ est une suite croissante et que $(\sup_{k\geq n} a_k)_{n\geq 0}$ est une suite décroissante. On note alors,

$$\liminf_n a_n = \lim_{n\to+\infty} \inf_{k\geq n} a_k = \sup_{n\geq 0} \inf_{k\geq n} a_k \,,$$
$$\limsup_n a_n = \lim_{n\to+\infty} \sup_{k\geq n} a_k = \inf_{n\geq 0} \sup_{k\geq n} a_k \,.$$

La $\liminf$ et la $\limsup$ sont parfois aussi notées $\underline{\lim}$ et $\overline{\lim}$, respectivement.

Proposition 1.4.10. *Soit (f_n) une suite de fonctions mesurables de $(E,\mathscr{A})$ dans $(\overline{\mathbb{R}},\mathscr{B}(\overline{\mathbb{R}}))$.*

1. *Les fonctions $x\mapsto \inf_n f_n(x)$ et $x\mapsto \sup_n f_n(x)$ sont mesurables de $(E,\mathscr{A})$ dans $(\overline{\mathbb{R}},\mathscr{B}(\overline{\mathbb{R}}))$.*

2. *Les fonctions $x\mapsto \liminf_n f_n(x)$ et $x\mapsto \limsup_n f_n(x)$ sont mesurables de $(E,\mathscr{A})$ dans $(\overline{\mathbb{R}},\mathscr{B}(\overline{\mathbb{R}}))$.*

3. *Si la suite (f_n) converge simplement vers f sur E, alors f est mesurable $(E,\mathscr{A})$ dans $(\overline{\mathbb{R}},\mathscr{B}(\overline{\mathbb{R}}))$.*

Preuve. 1. Pour tout $a\in\mathbb{R}$, on a $\{\inf_n f_n < a\} = \cup_n\{f_n < a\} \in \mathscr{A}$ car les fonctions f_n sont mesurables, donc $\inf_n f_n$ est mesurable. D'autre part, pour les mêmes raisons, $\{\sup_n f_n \leq a\} = \cap_n\{f_n\leq a\}\in\mathscr{A}$, donc $\sup_n f_n$ est mesurable.
2. Il suffit de remarquer que $\liminf_n f_n = \sup_n(\inf_{k\geq n} f_k)$. Alors d'après 1. les fonctions $g_n = \inf_{k\geq n} f_k$ sont mesurables, donc $\sup_n g_n$ est mesurable. On traite le cas de $\limsup_n f_n$ de la même manière.
3. Il suffit de remarquer que si la suite (f_n) converge simplement vers f, alors $f = \liminf_n f_n = \limsup_n f_n$ et le résultat se déduit de 2. □

Application : Soit $f:\mathbb{R}\to\mathbb{R}$ une fonction dérivable. Alors la dérivée f' de f est une fonction borélienne. En effet, f, en tant que fonction continue est une fonction borélienne. Définissons alors la suite de fonction f_n pour tout $n\geq 1$ et pour tout $x\in\mathbb{R}$ par $f_n(x) = [f(x+1/n) - f(x)]/n$. Alors, f_n est continue sur $\mathbb{R}$. C'est donc une fonction borélienne. De plus pour tout $x\in\mathbb{R}$, $f'(x) = \lim_{n\to+\infty} f_n(x)$. On conclut grâce au point 3. de la proposition 1.4.10.

Nous allons voir maintenant que toute fonction $f:E\to\mathbb{R}$ peut être approchée par une suite de fonctions étagées au sens de la convergence simple. Ce résultat est fondamental pour la construction de l'intégrale de Lebesgue.

On renvoie à la définition 1.4.3 d'une fonction étagée. Remarquons que les fonctions étagées sont les fonctions mesurables de E dans $\mathbb{R}$ qui ne prennent qu'un nombre fini de valeurs. Ainsi une fonction étagée $f : E \to \mathbb{R}$ qui prend les valeurs $\alpha_1, \dots, \alpha_n$ peut être écrite de la manière suivante :

$$f = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{\{f=\alpha_i\}}$$

L'écriture de la définition 1.4.3 est alors unique si l'on pose $A_i = \{f = \alpha_i\}$. Ceci est équivalent à dire que $(A_i)_{1\le i\le n}$ est une partition de E. Nous ferons toujours cette hypothèse.

Notation : Nous noterons $\mathscr{E}$ l'ensemble des fonctions étagées de E dans $\mathbb{R}$ et $\mathscr{E}_+$ le sous-ensemble des fonctions étagées de E dans $\mathbb{R}_+$.

Proposition 1.4.11. *Soit $f : (E, \mathscr{A}) \to (\overline{\mathbb{R}}_+, \mathscr{B}(\overline{\mathbb{R}}_+))$ une fonction mesurable et positive. Il existe une suite croissante (f_n) de fonctions de $\mathscr{E}_+$ telle que pour tout $x \in E$, $f(x) = \lim_n f_n(x)$. De plus si f est bornée, alors la convergence de (f_n) vers f est uniforme, c'est à dire* $\lim_{n\to+\infty} \|f_n - f\|_\infty = \lim_{n\to+\infty} \sup_{x\in E} |f_n(x) - f(x)| = 0$.

Preuve. Une telle suite (f_n) peut être construite explicitement de la manière suivante,

$$f_n = \sum_{k=0}^{n2^n-1} \frac{k}{2^n} \mathbb{1}_{\left\{\frac{k}{2^n} \le f < \frac{k+1}{2^n}\right\}} + n \mathbb{1}_{\{f\ge n\}} .$$

On vérifie que pour tout n, f_n est une fonction de $\mathscr{E}_+$. Montrons alors la convergence simple. Si $x \in E$ est tel que $f(x) = +\infty$, alors pour tout n, $f_n(x) = n$ et donc $f(x) = \lim_n f_n(x)$. Si $f(x) < \infty$, alors pour tout $n > f(x)$, il existe $k \le n2^n - 1$ tel que $k \le 2^n f(x) < k + 1$, donc $|f(x) - f_n(x)| = |f(x) - \frac{k}{2^n}| \le \frac{1}{2^n} \to 0$.

Il reste à vérifier que (f_n) est une suite croissante de fonctions. Si $f(x) \ge n + 1$, alors $f_n(x) = n < n + 1 = f_{n+1}(x)$. Si $f(x) < n$, on note k l'entier tel que $k \le 2^n f(x) < k + 1$. On a alors $f_n(x) = \frac{k}{2^n}$. Comme $f(x) \ge \frac{2k}{2^{n+1}}$, on en déduit que $f_{n+1}(x) \ge \frac{2k}{2^{n+1}} = f_n(x)$. Si $n \le f(x) < n + 1$, alors $f_n(x) = n$ et il existe $k \ge n2^{n+1}$ tel que $\frac{k}{2^{n+1}} \le f(x) \le \frac{k+1}{2^{n+1}}$, donc $f_{n+1}(x) \ge n = f_n(x)$.

Si f est bornée et $N \ge \|f\|_\infty$, alors pour tout $n \ge N$, on a $\|f - f_n\|_\infty \le \frac{1}{2^n}$, ce qui montre que (f_n) converge uniformément vers f. □

Pour toute fonction $f : E \to \mathbb{R}$, on définit les fonctions f^+ et f^- par

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

Ainsi f^+ et f^- sont des fonctions positives qui vérifient

$$f = f^+ - f^- \quad \text{et} \quad |f| = f^+ + f^- .$$

Corollary 1.4.2. *Soit $f : (E, \mathscr{A}) \to (\overline{\mathbb{R}}, \mathscr{B}(\overline{\mathbb{R}}))$ une fonction mesurable. Il existe une suite (f_n) de fonctions de $\mathscr{E}$ telle que pour tout $x \in E$, $f(x) = \lim_n f_n(x)$. De plus si f est bornée, alors la convergence de (f_n) vers f est uniforme.*

Preuve. Il suffit d'appliquer la proposition précédente aux fonctions f^+ et f^-. □

Chapitre 2

Mesures, intégrale de Lebesgue

2.1. Mesures

2.1.1. Définition et propriétés générales :

On rappelle les notations $\mathbb{R}_+ = [0, +\infty[$ et $\overline{\mathbb{R}}_+ = \mathbb{R}_+ \cup \{+\infty\} = [0, +\infty]$. Les éléments d'une suite $(A_n)_{n\geq 0}$ de sous-ensembles d'un ensemble E sont dits deux à deux disjoints si pour tous entiers i et j, tels que $i \neq j$, $A_i \cap A_j = \emptyset$.

Definition 2.1.1. *Soit $(E, \mathscr{A})$ un espace mesurable. On appelle mesure sur $(E, \mathscr{A})$ une application $\mu : \mathscr{A} \to \overline{\mathbb{R}}_+$ telle que :*

1. $\mu(\emptyset) = 0$.
2. *Si $(A_n)_{n\geq 0}$ est une suite d'éléments de $\mathscr{A}$ deux à deux disjoints, alors*

$$\mu\left(\cup_{n\geq 0} A_n\right) = \sum_{n\geq 0} \mu(A_n)\,.$$

(Cette propriété est appelée la σ-additivité.) On dit que $(E, \mathscr{A}, \mu)$ est un espace mesuré et pour tout $A \in \mathscr{A}$, on appelle $\mu(A)$ la mesure de A.

Exemples :
1) La mesure telle que $\mu(A) = 0$ pour tout $A \in \mathscr{A}$ est appelée la mesure nulle.
2) L'application définie par $\mu(A) = +\infty$, pour tout $A \in \mathscr{A}$ tel que $A \neq \emptyset$ et $\mu(\emptyset) = 0$ est appelée la mesure grossière.
3) La mesure définie pour tout $A \in \mathscr{A}$ par $\mu(A) = \mathrm{Card}(A)$, si A est fini et $\mu(A) = +\infty$ sinon est appelée la mesure de comptage.
4) Pour tout $x \in E$, on définit la mesure de Dirac en x par

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

5) Si μ est une mesure sur $(E, \mathscr{A})$, alors pour tout $B \in \mathscr{A}$, l'application $\nu : \mathscr{A} \to \overline{\mathbb{R}}_+$ définie par $\nu(B) = \mu(A \cap B)$ est aussi une mesure sur $(E, \mathscr{A})$. On l'appelle la mesure trace de μ sur B.

Proposition 2.1.1. *Soit $(E, \mathscr{A}, \mu)$ un espace mesuré et $A, B \in \mathscr{A}$, alors*

1. $\mu(A) = \mu(A \setminus B) + \mu(A \cap B)$,

2. $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$,
3. $\mu(A \cup B) \leq \mu(A) + \mu(B)$,
4. *si* $A \subset B$, *alors* $\mu(A) \leq \mu(B)$.

Preuve 1. Il suffit de voir que $A \setminus B$ et $A \cap B$ sont disjoints et que leur union est égale à A.
2. Les ensembles $A \setminus B$, $B \setminus A$ et $A \cap B$ sont deux à deux disjoints et leur union est égale à $A \cup B$, donc $\mu(A \setminus B) + \mu(B \setminus A) + \mu(A \cap B) = \mu(A \cup B)$ et en ajoutant $\mu(A \cap B)$ à chaque membre, on obtient

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

D'après 1., la somme des deux premiers termes du membre de gauche vaut $\mu(A)$ et la somme des deux derniers termes du membre de gauche vaut $\mu(B)$, ce qui permet de conclure.
3. Si $\mu(A)+\mu(B) = +\infty$, l'assertion est évidente. Sinon, d'après 2., $\mu(A \cup B)+\mu(A \cap B) < +\infty$, donc $\mu(A \cap B) < +\infty$. Par conséquent, on peut retrancher $\mu(A \cap B)$ à l'égalité 2., ce qui donne

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

4. D'après 2., si $A \subset B$, alors

$$\mu(B) = \mu(B \setminus A) + \mu(A \cap B) = \mu(B \setminus A) + \mu(A) \geq \mu(A)\,,$$

ce qui donne l'égalité à démontrer. □

Definition 2.1.2. *On dit qu'une mesure* μ *sur* $(E, \mathscr{A})$ *est finie si* $\mu(E) < +\infty$. *On appelle alors* $\mu(E)$ *la masse de* E.

Si μ *est une mesure de masse* 1, *on dit que* μ *est une probabilité et* $(E, \mathscr{A}, \mu)$ *est alors appelé un espace probabilisé.*

On dit que la mesure μ *est* σ*-finie si il existe une suite* $(E_n)_{n \geq 0}$ *d'éléments de* $\mathscr{A}$ *telle que* $E = \cup_{n \geq 0} E_n$ *et* $\mu(E_n) < +\infty$, *pour tout* $n \geq 0$.

On dit qu'une suite $(A_n)_{n \geq 0}$ de sous-ensembles de E est croissante (pour l'inclusion) si pour tout $n \geq 0$, $A_n \subset A_{n+1}$ et que cette suite est décroissante (pour l'inclusion) si pour tout $n \geq 0$, $A_{n+1} \subset A_n$.

Proposition 2.1.2. *Soit* $(E, \mathscr{A}, \mu)$ *un espace mesuré et* $(A_n)_{n \geq 0}$ *une suite d'éléments de* $\mathscr{A}$.

1. *Si la suite* $(A_n)_{n \geq 0}$ *est croissante, alors*

$$\lim_{n \to +\infty} \mu(A_n) = \mu(\cup_{n \geq 0} A_n)\,.$$

2. *Si la suite* $(A_n)_{n\geq 0}$ *est décroissante et si il existe* $n_0 \geq 0$*, tel que* $\mu(A_{n_0}) < +\infty$*, alors*
$$\lim_{n\to+\infty} \mu(A_n) = \mu(\cap_{n\geq 0} A_n)\,.$$

Preuve 1. Si il existe n_0 tel que $\mu(A_{n_0}) = +\infty$, alors $\mu(\cup_{n\geq 0} A_n) = +\infty$ et $\mu(A_{n_0}) = +\infty$, pour tout $n \geq n_0$, d'où le résultat. Supposons que ce ne soit pas le cas. Alors on introduit la suite $(B_n)_{n\geq 0}$ définie par $B_0 = A_0$ et pour tout $n \geq 1$, $B_n = A_n \backslash A_{n-1}$. Les B_n sont alors deux à deux disjoints et $\cup_{n\geq 0} A_n = \cup_{n\geq 0} B_n$. Par conséquent,

$$\begin{aligned}\mu(\cup_{n\geq 0} A_n) = \mu(\cup_{n\geq 0} B_n) = \sum_{n=0}^{+\infty} \mu(B_n) &= \lim_{n\to+\infty} \left(\sum_{k=1}^{n} \mu(A_k) - \mu(A_{k-1})\right) + \mu(A_0) \\ &= \lim_{n\to+\infty} \mu(A_n)\,.\end{aligned}$$

2. Quitte à considérer la suite $(A_{n_0+n})_{n\geq 0}$, on peut supposer que $\mu(A_0) < +\infty$. Comme $(A_0 \setminus A_n)_{n\geq 0}$ est une suite croissante de $\mathscr{A}$, on peut lui appliquer le résultat du 1. On obtient alors

$$\lim_{n\to+\infty} \mu(A_0 \setminus A_n) = \mu(\cup_{n\geq 0}(A_0 \setminus A_n)) = \mu(A_0 \setminus \cap_{n\geq 0} A_n)\,.$$

Comme $\mu(A_0) < +\infty$, on en déduit que $\lim_{n\to+\infty} \mu(A_0) - \mu(A_n) = \mu(A_0) - \lim_{n\to+\infty} \mu(A_n)$, ce qui donne le résultat. □

Considérons la suite $A_n = [n, +\infty[$, $n \in \mathbb{N}$, dans l'espace mesuré $(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$. Alors (A_n) est une suite décroissante telle que $\lambda(A_n) = +\infty$ pour tout $n \geq 1$. De plus $\cap_{n\in\mathbb{N}} A_n = \emptyset$, donc $\lambda(\cap_{n\in\mathbb{N}} A_n) = 0$. Cet exemple montre que la condition : il existe $n_0 \geq 0$, tel que $\mu(A_{n_0}) < +\infty$ du point 2. de la proposition 2.1.2 est nécessaire.

Corollary 2.1.3. *Si* $(A_n)_{n\geq 0}$ *une suite quelconque d'éléments de* $\mathscr{A}$*, alors*

$$\mu(\cup_{n\geq 0} A_n) \leq \sum_{n\geq 0} \mu(A_n)\,.$$

Preuve : Posons $B_n = \cup_{k=0}^{n} A_k$. Alors (B_n) est une suite croissante d'éléments de $\mathscr{A}$ et d'après la proposition précédente, $\lim_{n\to+\infty} \mu(B_n) = \mu(\cup_{n\geq 0} B_n) = \mu(\cup_{n\geq 0} A_n)$. D'autre part, d'après la propriété 3. de la proposition 2.1.1 (généralisée à un nombre fini quelconque d'ensembles), $\mu(B_n) \leq \sum_{k=0}^{n} \mu(A_k)$ et donc $\lim_{n\to+\infty} \mu(B_n) \leq \sum_{n\geq 0} \mu(A_n)$. □

Proposition 2.1.3. *Soit* $(E, \mathscr{A}, \mu)$ *un espace mesuré,* $(F, \mathscr{B})$ *un espace mesurable et* $f : (E, \mathscr{A}) \to (F, \mathscr{B})$ *une application mesurable. Alors l'application* $\mu_f : \mathscr{B} \to \overline{\mathbb{R}}_+$ *définie par*
$$\mu_f(B) = \mu(f^{-1}(B))\,, \quad \textit{pour tout } B \in \mathscr{B}$$
définit une mesure sur $(F, \mathscr{B})$*. On appelle* μ_f *la mesure image par* f *de* μ*.*

Preuve : Notons tout d'abord que, f étant mesurable, $f^{-1}(B) \in \mathscr{A}$, pour tout $B \in \mathscr{B}$. L'application μ_f est donc bien définie de $\mathscr{B}$ dans $\overline{\mathbb{R}}_+$. De plus on a $\mu_f(\emptyset) = \mu(f^{-1}(\emptyset)) = \mu(\emptyset) = 0$ et si $(B_n)_{n\geq 0}$ est une suite d'ensembles deux à deux disjoints de $\mathscr{B}$, alors les $f^{-1}(B_n)$ sont des ensembles deux à deux disjoints de $\mathscr{A}$ et $\mu(f^{-1}(\cup_{n\geq 0} B_n)) = \mu(\cup_{n\geq 0} f^{-1}(B_n)) = \sum_{n\geq 0} \mu(f^{-1}(B_n)) = \sum_{n\geq 0} \mu_f(B_n)$. □

2.1.2. Mesure de Lebesgue :

La mesure de Lebesgue est une mesure définie sur la tribu borélienne $\mathscr{B}(\mathbb{R}^d)$ dont la valeur pour tout pavé correspond à son volume. Ceci la définit d'ailleurs de manière unique comme le montre le théorème suivant :

Theorem 2.1.1. *Il existe une unique mesure sur* $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ *telle que la mesure de tout pavé* $\prod_{i=1}^d]a_i, b_i[$ *soit égale au produit* $\prod_{i=1}^d (b_i - a_i)$. *Cette mesure est appelée la mesure de Lebesque sur* $\mathbb{R}^d$. *On la note* λ_d *ou* λ *lorsqu'il n'y a pas d'ambiguïté.*

Nous ne démontrerons pas ce résultat. De même pour le suivant qui fournit une caractérisation équivalente de la mesure de Lebesgue. Pour un borélien de $\mathscr{B}(\mathbb{R}^d)$ et pour tout $x \in \mathbb{R}^d$, on définit la translation de B par $B + x = \{x + y : y \in B\}$. Notons que $B + x$ est l'image réciproque de B par la fonction $f(y) = y - x$ qui est continue sur $\mathbb{R}^d$, donc borélienne et par conséquent $B + x \in \mathscr{B}(\mathbb{R}^d)$. Le résultat suivant montre qu'à une constante multiplicative près, la mesure de Lebesgue est l'unique mesure sur $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ qui est invariante par translation.

Proposition 2.1.4. *Soit* μ *une mesure sur* $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ *qui vérifie les deux propriétés suivantes :*

1. *pour tout* $B \in \mathscr{B}(\mathbb{R}^d)$ *et tout* $x \in \mathbb{R}^d$, $\mu(B) = \mu(B + x)$.
2. $\mu([0,1]^d) = 1$.

Alors μ *est la mesure de Lebesgue.*

Exercice : Montrer que pour tout $d \geq 1$ et tout $x \in \mathbb{R}^d$, $\lambda_d(\{x\}) = 0$. En déduire que tout ensemble dénombrable est de mesure de Lebesgue nulle.

2.1.3. Ensembles négligeables et propriété μ-p.p. :

Definition 2.1.3. *Soit* $(E, \mathscr{A}, \mu)$ *un espace mesuré. On dira qu'un sous-ensemble* N *de* E *est négligeable si il existe* $A \in \mathscr{A}$ *tel que* $N \subset A$ *et* $\mu(A) = 0$.

On dit que μ *est une mesure complète si tout ensemble négligeable appartient à la tribu* $\mathscr{A}$.

On dit qu'une propriété P *relative aux points de* E *est vraie* μ*-presque partout sur* E *si l'ensemble*

$$\{x \in E : x \text{ ne vérifie pas la propriété } P.\}$$

est négligeable. Ceci se note : « P est vraie μ-p.p. »

Exemples :
1) Les fonctions f et g définies sur $\mathbb{R}$ par $f(x) = 1$, pour tout $x \in \mathbb{R}$ et $g(x) = 1$, pour tout $x \in \mathbb{R}^*$ et $g(0) = 0$ sont égales λ-p.p. En effet $\{x \in \mathbb{R} : f(x) \neq g(x)\} = \{0\}$ et $\lambda(\{0\}) = 0$.
2) Si N est un ensemble négligeable de l'espace mesuré $(E, \mathscr{A}, \mu)$, alors $\mathbb{1}_N = 0$, μ-p.p.

Definition 2.1.4. *Soit $(E, \mathscr{A}, \mu)$ un espace mesuré et soit $\mathscr{N}$ l'ensemble des parties négligeables de E. La tribu engendrée par $\mathscr{A} \cup \mathscr{N}$ est appelée la tribu complétée par rapport à μ. On la note $\mathscr{A}_\mu$, soit $\mathscr{A}_\mu = \sigma(\mathscr{A} \cup \mathscr{N})$.*

On dira que deux sous-ensembles A et B de E sont égaux μ-p.p. si $\mathbb{1}_A = \mathbb{1}_B$, μ-p.p. Ceci sera noté $A = B$, μ-p.p.

Proposition 2.1.5. *Soit $(E, \mathscr{A}, \mu)$ un espace mesuré. La tribu complétée par rapport à μ est égale à chacune des trois familles de parties suivantes :*

1. *La famille des parties A de E pour lesquelles il existe $B \in \mathscr{A}$ et $C \in \mathscr{A}$ vérifiant*
$$B \subset A \subset C \quad \text{et} \quad \mu(C \setminus B) = 0\,.$$
2. *La famille des parties A de E pour lesquelles il existe $B \in \mathscr{A}$ et un ensemble négligeable N vérifiant*
$$A = B \cup N\,.$$
3. *La famille des parties A de E pour lesquelles il existe $B \in \mathscr{A}$ vérifiant*
$$A = B\,, \quad \mu\text{-p.p.}$$

Exercice : On définit la différence symétrique de deux sous-ensembles A et B de E par $A\Delta B = (A \setminus B) \cup (B \setminus A)$. Vérifier que $A = B$, μ-p.p. si et seulement si $A\Delta B$ est négligeable.

2.2. Intégrale de Lebesgue

2.2.1. Intégrale des fonctions étagées positives :

Rappelons que nous avons noté $\mathscr{E}_+$ l'ensemble des fonctions étagées à valeurs positives sur l'espace mesurable $(E, \mathscr{A})$, voir Définition 1.4.3. Rappelons aussi que f peut être représentée de manière unique sous la forme : $f = \sum_{i=1}^n \alpha_i \mathbb{1}_{A_i}$, où $\alpha_1, \ldots, \alpha_n$ est l'ensemble des valeurs distinctes que peut prendre f et où $A_i = \{f = \alpha_i\}$.

Definition 2.2.1. *Soit $f \in \mathscr{E}_+$. On appelle intégrale de f par rapport à une mesure μ sur $(E, \mathscr{A})$, et l'on note $\int_E f(x)\, d\mu(x)$ la quantité définie par*

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

L'intégrale pourra aussi être notée $\int_E f(x)\, \mu(dx)$ ou parfois plus simplement $\int_E f\, d\mu$.

Remarquons que $\int_E f(x)\, d\mu(x) \in \overline{\mathbb{R}}_+$. D'autre part on pourra vérifier que la valeur de $\int_E f\, d\mu$ ne dépend pas de l'écriture choisie pour f. Plus précisément, si f admet l'écriture $f = \sum_{i=1}^{m} \alpha_i \mathbb{1}_{B_i}$, alors on peut vérifier que $\sum_{i=1}^{n} \alpha_i \mu(A_i) = \sum_{i=1}^{m} \alpha_i \mu(B_i)$. Enfin notons qu'on peut écrire l'intégrale d'une fonction étagée sous la forme canonique

$$\int_E f(x)\, d\mu(x) = \sum_{\alpha \in f(E)} \alpha \mu(\{f = \alpha\})\,.$$

Dans ce cas il faut tenir compte de la convention $0 \times (+\infty) = (+\infty) \times 0 = 0$. En effet il se peut que $\mu(\{f = 0\}) = +\infty$.

Exemples :
1) Si f est nulle, alors $\int_E f\, d\mu = 0$.
2) Si $\mu = \delta_a$, alors

$$\int_E f\, d\mu = \sum_{\alpha \in f(E)} \alpha \mu(\{f = \alpha\}) = f(a)\,.$$

3) Si μ est la mesure de Lebesgue sur $\mathbb{R}$, alors

$$\int_{\mathbb{R}} \mathbb{1}_{\mathbb{Q}}\, d\lambda = \lambda(\mathbb{Q}) = 0\,.$$

4) Soit l'espace mesuré $(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$ et la fonction étagée, positive $f(x) = 2\mathbb{1}_{[-1,3]}(x) + 4\mathbb{1}_{[2,6]}(x)$, alors $\int_{\mathbb{R}} f(x) d\lambda(x) = 2\lambda([-1,3]) + 4\lambda([2,6]) = 2 \times 4 + 4 \times 4 = 24$. Notons que la décomposition canonique de f est $f = 2\mathbb{1}_{[-1,2[}(x) + 6\mathbb{1}_{[2,3]}(x) + 4\mathbb{1}_{]3,6]}(x)$. On vérifie bien que le calcul de l'intégrale à partir de cette autre forme donne le même résultat : $\int_{\mathbb{R}} f(x) d\lambda(x) = 2\lambda([-1,2[) + 6\lambda([2,3]) + 4\lambda(]3,6]) = 2\times 3 + 6\times 1 + 4\times 3 = 24$.

Proposition 2.2.1. *L'application $f \mapsto \int_E f\, d\mu$ de $\mathscr{E}_+$ dans $\overline{\mathbb{R}}_+$ satisfait aux propriétés suivantes :*

1. *Pour toutes fonctions $f, g \in \mathscr{E}_+$, $f + g \in \mathscr{E}_+$ et $\int_E f + g\, d\mu = \int_E f\, d\mu + \int_E g\, d\mu$.*
2. *Si $\alpha \in \mathbb{R}_+$ et $f \in \mathscr{E}_+$, alors $\alpha f \in \mathscr{E}_+$ et $\int_E \alpha f\, d\mu = \alpha \int_E f\, d\mu$.*
3. *Si $f, g \in \mathscr{E}_+$ sont telles que $f \leq g$, alors $\int_E f\, d\mu \leq \int_E g\, d\mu$.*

Preuve : Choisissons les représentations canoniques $f = \sum_{\alpha\in f(E)} \alpha\mu(\{f = \alpha\}) = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{A_i}$ et $g = \sum_{\beta\in f(E)} \beta\mu(\{g = \beta\}) = \sum_{i=1}^{m} \beta_i \mathbb{1}_{B_i}$, où $A_i = \{f = \alpha_i\}$ et $B_i = \{g = \beta_i\}$. Alors $(A_i)_{1\leq i\leq n}$ et $(B_i)_{1\leq i\leq m}$ sont des partitions de E, de plus $f+g$ vaut $\alpha_i + \beta_j$ sur l'ensemble $A_i \cap B_j$, donc $f + g = \sum_{i=1}^{n}\sum_{j=1}^{m}(\alpha_i + \beta_j)\mathbb{1}_{A_i\cap B_j}$. On peut alors écrire

$$\begin{aligned}\int_E f + g\,d\mu &= \sum_{i=1}^{n}\sum_{j=1}^{m}(\alpha_i + \beta_j)\mu(A_i \cap B_j) = \sum_{i=1}^{n}\alpha_i\sum_{j=1}^{m}\mu(A_i \cap B_j) + \sum_{j=1}^{m}\beta_j\sum_{i=1}^{n}\mu(A_i \cap B_j)\\ &= \sum_{i=1}^{n}\alpha_i\mu(\cup_{j=1}^{m}A_i \cap B_j) + \sum_{j=1}^{m}\beta_j\mu(\cup_{i=1}^{n}A_i \cap B_j)\\ &= \sum_{i=1}^{n}\alpha_i\mu(A_i) + \sum_{j=1}^{m}\alpha_i\mu(B_j) = \int_E f\,d\mu + \int_E g\,d\mu\,.\end{aligned}$$

2. Si $f = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{A_i}$, alors $\alpha f = \sum_{i=1}^{n} \alpha\alpha_i \mathbb{1}_{A_i}$, donc $\alpha f \in \mathscr{E}_+$ et

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

3. Comme $f \leq g$, la fonction $g - f$ est étagée et positive. On a alors par le point 1.,

$$\int_E g\,d\mu = \int_E f + (g-f)\,d\mu = \int_E f\,d\mu + \int_E (g-f)\,d\mu \geq \int_E f\,d\mu\,.$$

□

Notation : Si f est une fonction étagée positive et $A \in \mathscr{A}$, alors on vérifie que $f\mathbb{1}_A$ est aussi une fonction étagée positive et on note

$$\int_A f\,d\mu = \int_E f\mathbb{1}_A\,d\mu\,.$$

2.2.2. Intégrale des fonctions mesurables positives :

Nous noterons $\mathscr{M}_+$ l'ensemble des fonctions mesurables de $(E,\mathscr{A})$ dans $(\overline{\mathbb{R}}_+, \mathscr{B}(\overline{\mathbb{R}}_+))$.

Definition 2.2.2. *On définit l'intégrale de toute fonction $f \in \mathscr{M}_+$ par rapport à la mesure μ par*

$$\int_E f(x)\,d\mu(x) = \sup\left\{\int_E g(x)\,d\mu(x) : g \in \mathscr{E}_+,\ g \leq f\right\}\,.$$

Cette intégrale sera notée indifféremment $\int_E f(x)\,d\mu(x)$, $\int_E f(x)\,\mu(dx)$ ou $\int_E f\,d\mu$.

Si $\int_E f(x)\,d\mu(x) < +\infty$, on dira que f est μ-intégrable (ou intégrable).

Remarques : Si μ est une mesure finie, alors toute fonction de $\mathscr{M}_+$ qui est bornée est intégrable.

Il est clair d'après le point 3. de la proposition 2.2.1, que si $f \in \mathscr{E}_+$, alors l'intégrale de f selon la définition 2.2.2 correspond à l'intégrale de f selon la définition 2.2.1.

Pour toute fonction $f \in \mathscr{M}_+$ et tout $A \in \mathscr{A}$, on vérifie que $f \mathbb{I}_A \in \mathscr{M}_+$ et on note

$$\int_A f \, d\mu = \int_E f \mathbb{I}_A \, d\mu \,.$$

Exemple : Prenons $(E, \mathscr{A}) = (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ et choisissons pour mesure $\mu = \lambda$, la mesure de Lebesgue. Nous verrons au chapitre suivant que si une fonction positive f est une fonction à la fois borélienne et localement Riemann intégrable dont l'intégrale impropre est convergente, alors

$$\int_{\mathbb{R}} f \, d\lambda = \int_{-\infty}^{+\infty} f(x) \, dx \,.$$

Par exemple $f(x) = e^{-|x|}$ est continue, donc $f \in \mathscr{M}_+$. De plus l'intégrale de Riemann impropre de f sur $\mathbb{R}$ converge. On a donc,

$$\int_{\mathbb{R}} f \, d\lambda = \int_{-\infty}^{+\infty} e^{-|x|} \, dx = 2 \int_0^{+\infty} e^{-x} \, dx = 2 \,.$$

Proposition 2.2.2. *Pour toutes fonctions $f, g \in \mathscr{M}_+$, si $f \leq g$, alors $\int_E f \, d\mu \leq \int_E g \, d\mu$. En particulier $\int_E f \, d\mu \geq 0$ pour toute fonction $f \in \mathscr{M}_+$.*

Preuve : Si $h \in \mathscr{E}_+$ est telle que $h \leq f$, alors $h \leq g$, donc

$$\left\{ \int_E h(x) \, d\mu(x) : h \in \mathscr{E}_+, \, h \leq f \right\} \subset \left\{ \int_E h(x) \, d\mu(x) : h \in \mathscr{E}_+, \, h \leq g \right\} ,$$

ce qui entraîne que

$$\sup \left\{ \int_E h(x) \, d\mu(x) : h \in \mathscr{E}_+, \, h \leq f \right\} \leq \sup \left\{ \int_E h(x) \, d\mu(x) : h \in \mathscr{E}_+, \, h \leq g \right\} .$$

La dernière assertion est une conséquence de cette inégalité et du fait que l'intégrale de la fonction $x \mapsto 0$ sur E est égale à 0. $\square$

Theorem 2.2.1 (Théorème de Beppo Levi ou de convergence monotone)**.** *Si (f_n) est une suite croissante de fonctions de $\mathscr{M}_+$, alors $\lim_{n \to +\infty} f_n \in \mathscr{M}_+$. Posons $f = \lim_{n \to +\infty} f_n$, alors*

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

Preuve : Nous avons déjà vu à la Proposition 1.4.10 que $\lim_{n\to+\infty} f_n \in \mathscr{M}_+$. Montrons d'abord l'inégalité $\geq$. Comme pour tout entier n, $f_n \leq f_{n+1} \leq f$, par croissance de l'intégrale on a aussi

$$\int_E f_n \, d\mu \leq \int_E f_{n+1} \, d\mu \leq \int_E f \, d\mu \,,$$

ce qui prouve en passant à la limite que $\lim_{n\to+\infty} \int_E f_n \, d\mu \leq \int_E f \, d\mu$.

Montrons maintenant l'autre inégalité. Par définition de l'intégrale de f, il suffit de montrer que pour toute fonction étagée positive, g telle que $g \leq f$, on a $\int_E g \, d\mu \leq \lim_{n\to+\infty} \int_E f_n \, d\mu$. Soit alors $a \in [0,1[$ et $E_n = \{ag \leq f_n\}$. Comme $g \leq f$, on a l'égalité $E = \cup_n E_n$. En effet d'une part, sur $\{f = 0\}$, $f_n = g = 0$ pour tout n et donc $\{f = 0\} \subset E_n$. Par conséquent $\{f = 0\} \subset \cup_n E_n$. D'autre part sur $\{f > 0\}$, $ag < f$ car g ne prend que des valeurs finies. Donc pour tout $x \in \{f > 0\}$, il existe un rang $n(x)$ tel que pour tout $n \geq n(x)$, $ag(x) \leq f_n(x)$. Autrement dit, $x \in E_n$ et par conséquent, $x \in \cup_n E_n$. On conclut que $E = \{f = 0\} \cup \{f > 0\} \subset \cup_n E_n$ et donc $E = \cup_n E_n$.

En notant g comme $g = \sum_{i=1}^k \alpha_i \mathbb{1}_{A_i}$ et en remarquant que $\mathbb{1}_{A_i} \mathbb{1}_{E_n} = \mathbb{1}_{A_i \cap E_n}$, on obtient

$$\int_E ag \mathbb{1}_{E_n} \, d\mu = a \int_E \sum_{i=1}^k \alpha_i \mathbb{1}_{A_i \cap E_n} \, d\mu = \sum_{i=1}^k a\alpha_i \mu(A_i \cap E_n) \,.$$

Puisque la suite (E_n) croît vers E, on a $\lim_{n\to+\infty} \mu(A_i \cap E_n) = \mu(A_i)$, pour tout $i = 1, \ldots, n$, ce qui entraîne, les sommes étant finies, que

$$\lim_{n\to+\infty} \int_E ag \, \mathbb{1}_{E_n} \, d\mu = \sum_{i=1}^k a\alpha_i \mu(A_i) = a \int_E g \, d\mu \,. \tag{2.2.1}$$

D'autre part, $E_n = \{ag \leq f_n\}$, donc $ag\mathbb{1}_{E_n} \leq f_n$, d'où

$$\int_E ag \mathbb{1}_{E_n} \, d\mu \leq \int_E f_n \, d\mu \leq \lim_{n\to+\infty} \int_E f_n \, d\mu \,.$$

En passant à la limite dans la dernière inégalité et en tenant compte de (2.2.1), on obtient finalement,

$$a \int_E g \, d\mu \leq \lim_{n\to+\infty} \int_E f_n \, d\mu \,.$$

Puisque le choix de $a \in [0,1[$ est arbitraire, il peut être aussi proche de 1 qu'on le souhaite et l'on obtient ainsi l'inégalité $\leq$. $\square$

On remarque que dans ce théorème, la valeur de l'intégrale $\int_E f \, d\mu$ ne dépend pas du choix de la suite croissante (f_n).

Proposition 2.2.3. *Si f et g sont deux fonctions de $\mathscr{M}_+$, alors*

1. $\int_E f + g \, d\mu = \int_E f \, d\mu + \int_E g \, d\mu$,

2. *pour tout* $\alpha \geq 0$, $\int_E \alpha f \, d\mu = \alpha \int_E f \, d\mu$.

Preuve : 1. D'après la proposition 1.4.11 il existe des suites croissantes (f_n) et (g_n) de $\mathscr{E}_+$ telles que $f = \lim_{n\to+\infty} f_n$ et $g = \lim_{n\to+\infty} g_n$. Alors $f_n + g_n$ est une suite croissante de $\mathscr{M}_+$ telle que $f + g = \lim_{n\to+\infty} f_n + g_n$. De plus $\int_E f_n + g_n \, d\mu = \int_E f_n \, d\mu + \int_E g_n \, d\mu$ et d'après le théorème de Beppo Levi,

$$\begin{array}{ccc} \displaystyle\int_E f_n + g_n \, d\mu & \underset{n\to+\infty}{\longrightarrow} & \displaystyle\int_E f + g \, d\mu \\ \| & & \\ \displaystyle\int_E f_n \, d\mu + \int_E g_n \, d\mu & \underset{n\to+\infty}{\longrightarrow} & \displaystyle\int_E f \, d\mu + \int_E g \, d\mu \,. \end{array}$$

2. La suite (αf_n) est une suite croissante de $\mathscr{M}_+$ qui converge vers αf. De plus $\int_E \alpha f_n \, d\mu = \alpha \int_E f_n \, d\mu$ et d'après le théorème de Beppo Levi,

$$\begin{array}{ccc} \displaystyle\int_E \alpha f_n \, d\mu & \underset{n\to+\infty}{\longrightarrow} & \displaystyle\int_E \alpha f \, d\mu \\ \| & & \\ \displaystyle\alpha \int_E f_n \, d\mu & \underset{n\to+\infty}{\longrightarrow} & \displaystyle\alpha \int_E f \, d\mu \,. \end{array}$$

□

Proposition 2.2.4. *Si* $(f_n)_{n\geq 0}$ *est une suite de fonction de* $\mathscr{M}_+$, *alors*

$$\int_E \sum_{n\geq 0} f_n \, d\mu = \sum_{n\geq 0} \int_E f_n \, d\mu \,.$$

Preuve : La suite $(\sum_{k=0}^n f_k)_{n\geq 0}$ est une suite croissante de $\mathscr{M}_+$ et d'après le théorème de Beppo Levi,

$$\begin{array}{ccc} \displaystyle\int_E \sum_{k=0}^{n} f_k \, d\mu & \underset{n\to+\infty}{\longrightarrow} & \displaystyle\int_E \sum_{k\geq 0} f_k \, d\mu \\ \| & & \\ \displaystyle\sum_{k=0}^{n} \int_E f_n \, d\mu & \underset{n\to+\infty}{\longrightarrow} & \displaystyle\sum_{k\geq 0} \int_E f_k \, d\mu \,, \end{array}$$

où l'on a utilisé l'additivité des intégrales des fonctions de $\mathscr{M}_+$. □

Inégalité de Markov et ses conséquences.

Lemme 2.2.1. *Soient* $f \in \mathscr{M}_+$ *et* $a > 0$ *un réel, alors*

$$\mu(\{f \geq a\}) \leq \frac{1}{a} \int_E f \, d\mu \,.$$

Preuve : Puisque $f \geq a\mathbb{1}_{\{f\geq a\}}$, d'après la propriété de croissance de l'intégrale, on obtient

$$\int_E f\,d\mu \geq \int_E a\mathbb{1}_{\{f\geq a\}}\,d\mu = a\mu(\{f\geq a\})\,,$$

d'où le résultat. □

Corollary 2.2.4. *Si $f \in \mathscr{M}_+$ est intégrable, alors f est finie μ-p.p., autrement dit, $\mu(\{f=+\infty\})=0$.*

Preuve : Pour tout $n \geq 1$,

$$\mu(\{f=+\infty\}) \leq \mu(\{f\geq n\}) \leq \frac{1}{n}\int_E f\,d\mu \underset{n\to+\infty}{\longrightarrow} 0\,,$$

d'après l'inégalité de Markov. On conclut que $\mu(\{f=+\infty\})=0$. □

Corollary 2.2.5. *Soient $f, g \in \mathscr{M}_+$, alors*

1. *$\int_E f\,d\mu = 0$ si et seulement si $f=0$, μ-p.p.*
2. *Si $f=g$, μ-p.p., alors $\int_E f\,d\mu = \int_E g\,d\mu$.*

Preuve : 1. Supposons que $f \in \mathscr{E}_+$ et $f=0$, μ-p.p. On peut alors écrire $f = \sum_{i=1}^n \alpha_i \mathbb{1}_{A_i}$, où $\mu(A_i)=0$ dès que $\alpha_i > 0$. Par conséquent, $\int_E f\,d\mu = \sum_{i=1}^n \alpha_i\mu(A_i) = 0$. Si $f \in \mathscr{M}_+$ et $f=0$, μ-p.p., alors pour toute fonction $g \in \mathscr{E}_+$ telle que $g \leq f$, $g=0$, μ-p.p., ce qui implique que $\int_E g\,d\mu = 0$. On a donc bien démontré que $\int_E f\,d\mu = 0$ (par définition de l'intégrale).

Réciproquement, supposons que $\int_E f\,d\mu = 0$, alors pour tout $n \geq 1$, $\mu(\{f \geq \frac{1}{n}\}) \leq n\int_E f\,d\mu = 0$, d'après l'inégalité de Markov. D'autre part, $(\{f \geq \frac{1}{n}\})_{n\geq 1}$ est une suite croissante d'ensembles de $\mathscr{A}$. Par conséquent,

$$\mu(\{f\neq 0\}) = \mu\left(\cup_{n\geq 1}\left\{f \geq \frac{1}{n}\right\}\right) = \lim_{n\to+\infty}\mu\left(\left\{f\geq\frac{1}{n}\right\}\right) = 0\,.$$

2. Si $f=g$, μ-p.p., alors $f\mathbb{1}_{\{f\neq g\}}$ et $g\mathbb{1}_{\{f\neq g\}}$ sont nulles μ-p.p., ce qui implique que $\int_{\{f\neq g\}} f\,d\mu = \int_{\{f\neq g\}} g\,d\mu = 0$. On en déduit alors que

$$\int_E f\,d\mu = \int_{\{f=g\}} f\,d\mu = \int_{\{f=g\}} g\,d\mu = \int_E g\,d\mu\,.$$

□

<u>Remarque</u> : Soit $f \in \mathscr{M}_+$ et soit $A \in \mathscr{A}$ tel que $\mu(A)=0$, alors $\int_A f\,d\mu = 0$. Ceci vient du fait que $f\mathbb{1}_A = 0$, μ-p.p.

Proposition 2.2.5 (Lemme de Fatou). *Si* $(f_n)_{n\geq 0}$ *est une suite de fonctions de* $\mathscr{M}_+$*, alors*

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

Preuve : Soit $g_n = \inf_{k\geq n} f_k$ La suite $(g_n)_{n\geq 0}$ est une suite croissante de fonctions de $\mathscr{M}_+$, et d'après le théorème de Beppo Levi, $\int_E \lim_{n\to+\infty} g_n \, d\mu = \lim_{n\to+\infty} \int_E g_n \, d\mu$, soit

$$\int_E \liminf_{n\to+\infty} f_n \, d\mu = \lim_{n\to+\infty} \int_E \inf_{k\geq n} f_k \, d\mu \,. \tag{2.2.2}$$

D'autre part, pour tout $k_0 \geq n$, on a $\inf_{k\geq n} f_k \leq f_{k_0}$, ce qui implique que $\int_E \inf_{k\geq n} f_k \leq \int_E f_{k_0}$. Par passage à l'infimum pour $k_0 \geq n$, on obtient que $\int_E \inf_{k\geq n} f_k \, d\mu \leq \inf_{k\geq n} \int_E f_k \, d\mu$, ce qui montre le résultat en passant à la limite sur n et en utilisant (2.2.2). $\square$

Pour la lim sup, le résultat requiert une hypothèse comme le montre la proposition suivante.

Proposition 2.2.6. *Si* $(f_n)_{n\geq 0}$ *est une suite de fonctions de* $\mathscr{M}_+$ *telle que pour tout* $n \geq 0$*,* $f_n \leq g$*, où* $g \in \mathscr{M}_+$ *est une fonction intégrable, alors*

$$\limsup_{n\to+\infty} \int_E f_n \, d\mu \leq \int_E \limsup_{n\to+\infty} f_n \, d\mu \,.$$

Preuve : Notons tout d'abord que si $f, g \in \mathscr{M}_+$ sont telles que $f \leq g$ et si f est intégrable, alors $\int_E (g-f)\, d\mu = \int_E g \, d\mu - \int_E f \, d\mu$. En effet, $g = f + (g-f)$ et f et $g-f$ sont des fonctions mesurables positives. Le résultat vient alors de l'additivité de l'intégrale.

Appliquons le lemme de Fatou à la suite $g - f_n$ de fonctions de $\mathscr{M}_+$,

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

Remarquons que pour tout $n \geq 0$, f_n est une fonction intégrable ainsi que $\limsup_{n\to+\infty} f_n$. On conclut en utilisant la remarque du début de la preuve. Le membre de gauche devient

$$\int_E \liminf_{n\to+\infty} (g-f_n) = \int_E (g - \limsup_{n\to+\infty} f_n) \, d\mu = \int_E g \, d\mu - \int_E \limsup_{n\to+\infty} f_n \, d\mu \,.$$

et le membre de droite devient

$$\liminf_{n\to+\infty} \int_E (g-f_n) \, d\mu = \liminf_{n\to+\infty} \int_E g \, d\mu - \int_E f_n \, d\mu = \int_E g \, d\mu - \limsup_{n\to+\infty} \int_E f_n \, d\mu \,.$$

$\square$

Contre-exemple : Prenons l'espace $E = [0, 1]$ et munissons le de sa tribu borélienne $\mathscr{B}([0, 1])$. Soit la suite $f_n(x) = n\mathbb{1}_{]0,1/n]}(x)$, $n \geq 1$ de fonctions de $\mathscr{M}_+$. Alors d'une part on vérifie que pour tout $x \in [0, 1]$, $\limsup_{n\to+\infty} f_n(x) = \lim_{n\to+\infty} f_n(x) = 0$, donc $\int_E \limsup_{n\to+\infty} f_n \, d\lambda = 0$ et d'autre part les intégrales au sens de Lebesgue et au sens de Riemann correspondent (voir l'exemple du début de cette sous-section), donc $\int_E f_n \, d\lambda = \int_0^1 n\mathbb{1}_{]0,1/n]}(x) \, dx = 1$ et $\limsup_{n\to+\infty} \int_E f_n \, d\lambda = 1$. L'inégalité $\limsup_{n\to+\infty} \int_E f_n \, d\lambda \leq \int_E \limsup_{n\to+\infty} f_n \, d\lambda$ n'est donc pas vérifiée.

Mesures à densité.

Proposition 2.2.7. *Soit $g \in \mathscr{M}_+$. Posons, pour tout $A \in \mathscr{A}$,*

$$\nu(A) = \int_A g \, d\mu \, .$$

Alors ν définit une mesure sur $(E, \mathscr{A})$. On dit que ν est la mesure de densité g par rapport à la mesure μ. De plus pour toute fonction f de $\mathscr{M}_+$,

$$\int_E f \, d\nu = \int_E f g \, d\mu \, . \tag{2.2.3}$$

Preuve : Il est clair que $\nu(\emptyset) = 0$. D'autre part, si $(A_n)_{n\geq 0}$ est une suite d'éléments de $\mathscr{A}$ deux à deux disjoints, alors on pourra vérifier que $\mathbb{1}_{\cup_{n\geq 0} A_n} = \sum_{n\geq 0} \mathbb{1}_{A_n}$. Par conséquent,

$$\nu(\cup_{n\geq 0} A_n) = \int_E g \mathbb{1}_{\cup_{n\geq 0} A_n} \, d\mu = \int_E g \sum_{n\geq 0} \mathbb{1}_{A_n} \, d\mu = \sum_{n\geq 0} \int_E g \mathbb{1}_{A_n} \, d\mu = \sum_{n\geq 0} \nu(A_n) \, ,$$

d'après le théorème de Beppo Levi.

L'égalité (2.2.3) est vraie pour toute fonction indicatrice $f = \mathbb{1}_A$. Par additivité, ceci reste vrai pour toute fonction $f \in \mathscr{E}_+$. Enfin, si $f \in \mathscr{M}_+$, alors il existe une suite croissante (f_n) de fonctions de $\mathscr{E}_+$ qui converge simplement vers f. La suite $(f_n g)$ est donc une suite croissante de fonctions de $\mathscr{M}_+$ qui converge vers fg et d'après le théorème de Beppo Levi, on a

$$\begin{array}{ccc} \displaystyle\int_E f_n \, d\nu & \underset{n\to+\infty}{\longrightarrow} & \displaystyle\int_E f \, d\nu \\ \| & & \\ \displaystyle\int_E f_n g \, d\mu & \underset{n\to+\infty}{\longrightarrow} & \displaystyle\int_E f g \, d\mu \, , \end{array}$$

ce qui montre l'identité (2.2.3) dans le cas général. □

Exemple : Soit $(\Omega, \mathscr{F}, \mathbb{P})$ un espace probabilisé (espace mesuré tel que $P(\Omega) = 1$). Soit $X : \Omega \to \mathbb{R}_+$ une fonction mesurable, $\mathbb{R}_+$ étant muni de la tribu $\mathscr{B}(\mathbb{R}_+)$. La fonction X est appelée une variable aléatoire en probabilités.

La loi de X est la mesure image sur $(\mathbb{R}_+, \mathscr{B}(\mathbb{R}_+))$ de P par X, c'est à dire la mesure ν_X sur $(\mathbb{R}_+, \mathscr{B}(\mathbb{R}_+))$ définie par $\nu_X(A) = P(X^{-1}(A))$, pour tout $A \in \mathscr{B}(\mathbb{R}_+)$, voir la proposition 2.1.3. On dit que la variable aléatoire X suit la loi ν_X.

Une variable aléatoire X suit la loi normale $\mathscr{N}(m, \sigma)$ si sa loi est la mesure qui a pour densité la fonction $x \mapsto \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$ par rapport à la mesure de Lebesgue λ. Ceci se note

$$\nu_X(dx) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} \lambda(dx) .$$

2.2.3. Intégrale des fonctions mesurables de signe quelconque :

Comme dans les sections précédentes, $(E, \mathscr{A}, \mu)$ désignera un espace mesurable quelconque. Lorsqu'aucune confusion n'est possible, nous noterons $\mathscr{M}$ l'ensemble des fonctions mesurables de $(E, \mathscr{A})$ dans $(\overline{\mathbb{R}}, \mathscr{B}(\overline{\mathbb{R}}))$.

Definition 2.2.3. *Une fonction $f \in \mathscr{M}$ est dite μ-intégrable (ou intégrable) si $\int_E |f| \, d\mu < +\infty$. On notera l'ensemble des fonctions intégrables par*

$$\mathscr{L}^1(E, \mathscr{A}, \mu) = \left\{ f \in \mathscr{M} : \int_E |f| \, d\mu < +\infty \right\} .$$

L'ensemble $\mathscr{L}^1(E, \mathscr{A}, \mu)$ est parfois noté plus simplement par $\mathscr{L}^1_\mu(E)$ ou bien $\mathscr{L}^1(E)$. Nous utiliserons cette dernière notation lorsqu'aucune confusion ne sera possible.

Pour une fonction $f : E \to \overline{\mathbb{R}}$, on note

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

Notons alors que f^+ et f^- sont des fonctions positives qui vérifient

$$f = f^+ - f^- \quad \text{et} \quad |f| = f^+ + f^- .$$

En particulier, f est mesurable si et seulement si f^+ et f^- le sont.

Proposition 2.2.8. *Soit $f \in \mathscr{M}$, alors*

$$f \in \mathscr{L}^1(E) \quad \textit{si et seulement si} \quad f^+ \in \mathscr{L}^1(E) \quad \textit{et} \quad f^- \in \mathscr{L}^1(E) .$$

Preuve : Puisque $f = f^+ - f^-$, on peut écrire

$$\int_E f^\pm \, d\mu \le \int_E |f| \, d\mu = \int_E f^+ \, d\mu + \int_E f^- \, d\mu ,$$

ce qui entraîne que $\int_E |f| \, d\mu < +\infty$ si et seulement si $\int_E f^+ \, d\mu < +\infty$ et $\int_E f^- \, d\mu < +\infty$. $\square$

Definition 2.2.4. *(Intégrale de Lebesgue)*
L'intégrale de Lebesgue d'une fonction $f \in \mathscr{L}^1(E)$ est définie par

$$\int_E f\, d\mu = \int_E f^+\, d\mu - \int_E f^-\, d\mu\,.$$

Notation : Soit $f \in \mathscr{L}^1(E)$ et $A \in \mathscr{A}$. Remarquons que $f\mathbb{1}_A \in \mathscr{L}^1(E)$ car $|f\mathbb{1}_A| \leq |f|$. Nous noterons alors

$$\int_A f\, d\mu = \int_E f\mathbb{1}_A\, d\mu\,.$$

Proposition 2.2.9. *Soient $f, g : (E, \mathscr{A}) \to (\overline{\mathbb{R}}, \mathscr{B}(\overline{\mathbb{R}}))$, des fonctions mesurables telles que $f = g$, μ-p.p. Alors*

$$f \in \mathscr{L}^1(E) \quad \textit{si et seulement si} \quad g \in \mathscr{L}^1(E)$$

et dans ce cas $\int_E f\, d\mu = \int_E g\, d\mu$.

Preuve : Puisque $\{f^+ \neq g^+\} \subset \{f \neq g\}$ et $\{f^- \neq g^-\} \subset \{f \neq g\}$, l'égalité $f = g$, μ-p.p. entraîne que $f^+ = g^+$, μ-p.p. et $f^- = g^-$, μ-p.p. Par conséquent f^+ et f^- sont intégrables si et seulement si g^+ et g^- le sont et dans ce cas, $\int_E f\, d\mu = \int_E g\, d\mu$.
$\square$

Proposition 2.2.10. *Pour toute fonction $f \in \mathscr{L}^1(E)$,*

$$\mu(\{f = +\infty\}) = \mu(\{f = -\infty\}) = \mu(\{|f| = +\infty\}) = 0\,.$$

Preuve : On remarque que $\mu(\{|f| = +\infty\}) = \mu(\{f = +\infty\} \cup \{f = -\infty\}) = \mu(\{f = +\infty\}) + \mu(\{f = -\infty\})$ et que $\mu(\{f = +\infty\}) = \mu(\{f^+ = +\infty\})$ et $\mu(\{f = -\infty\}) = \mu(\{f^- = +\infty\})$. Il suffit alors d'appliquer le corollaire 2.2.4 aux fonctions positives f^+ et f^-. $\square$

Les propriétés de linéarité et de positivité de l'intégrale de Lebesgue sont héritées des propritétés de l'intégrale des fonctions mesurables et positives.

Proposition 2.2.11. *Soient $f, g \in \mathscr{L}^1(E)$, alors*

1. $f + g \in \mathscr{L}^1(E)$ *et* $\int_E (f+g)\, d\mu = \int_E f\, d\mu + \int_E g\, d\mu\,.$
2. *Pour tout $\alpha \in \mathbb{R}$, $\alpha f \in \mathscr{L}^1(E)$ et $\int_E \alpha f\, d\mu = \alpha \int_E f\, d\mu$.*

Preuve : Puisque $|f + g| \leq |f| + |g|$, on a $\int_E |f+g|\, d\mu = \int_E |f|\, d\mu + \int_E |g|\, d\mu < +\infty$ et donc $f + g \in \mathscr{L}^1(E)$. D'autre part, notons que

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

On en déduit que

$$(f+g)^+ + f^- + g^- = (f+g)^- + f^+ + g^+ .$$

Puisque les fonctions des membres de gauche et de droite sont mesurables positives, on a

$$\int_E (f+g)^+ \, d\mu + \int_E f^- \, d\mu + \int_E g^- \, d\mu = \int_E (f+g)^- \, d\mu + \int_E f^+ \, d\mu + \int_E g^+ \, d\mu .$$

Ces quantités étant finies, on en déduit que

$$\begin{aligned}
\int_E (f+g)^+ \, d\mu - \int_E (f+g)^- \, d\mu &= \int_E f^+ \, d\mu - \int_E f^- \, d\mu + \int_E g^+ \, d\mu - \int_E g^- \, d\mu \\
&= \int_E f \, d\mu + \int_E g \, d\mu .
\end{aligned}$$

Montrons maintenant la partie 2. Tout d'abord, on a $\int_E |\alpha f| \, d\mu = |\alpha| \int_E |f| \, d\mu < \infty$, donc $\alpha f \in \mathscr{L}^1(E)$. D'autre part, si $\alpha \geq 0$, alors $(\alpha f)^+ = \alpha f^+$ et $(\alpha f)^- = \alpha f^-$, d'où le résultat. Si $\alpha < 0$, alors $(\alpha f)^+ = -\alpha f^-$ et $(\alpha f)^- = -\alpha f^+$, d'où le résultat. □

Proposition 2.2.12. *Soient $f, g \in \mathscr{L}^1(E)$. Si $f \leq g$, μ-p.p., alors*

$$\int_E f \, d\mu \leq \int_E g \, d\mu .$$

Preuve : Supposons d'abord $f(x) \leq g(x)$, pour tout $x \in E$. On a $g - f \in \mathscr{M}_+$, donc $\int_E (g-f) \, d\mu \geq 0$. On conclut par la linéarité de l'intégrale. Si $f \leq g$, μ-p.p., alors on applique ce que l'on vient de montrer aux fonctions $f_1 = f \mathbb{1}_{\{f \leq g\}}$ et $g_1 = g \mathbb{1}_{\{f \leq g\}}$ qui vérifient $f_1(x) \leq g_1(x)$, pour tout $x \in E$. On conclut en remarquant que $f = f_1$ et $g = g_1$, μ-p.p. et donc $\int_E f \, d\mu = \int_E f_1 \, d\mu$ et $\int_E g \, d\mu = \int_E g_1 \, d\mu$. □

L'intégrale de Lebesgue satisfait à l'inégalité triangulaire.

Proposition 2.2.13. *Soit $f \in \mathscr{L}^1(E)$, alors*

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

Preuve : Il suffit d'écrire

$$\left| \int_E f \, d\mu \right| = \left| \int_E f^+ \, d\mu - \int_E f^- \, d\mu \right| \leq \int_E f^+ \, d\mu + \int_E f^- \, d\mu = \int_E |f| \, d\mu .$$

□

On rappelle la définition vue à la proposition 2.1.3 de la mesure image de μ sur l'espace mesurable $(F, \mathscr{B})$ par l'application mesurable $\varphi : (E, \mathscr{A}) \to (F, \mathscr{B})$. Cette mesure, que l'on note ici μ_φ est définie par

$$\mu_\varphi(B) = \mu(\varphi^{-1}(B)), \quad B \in \mathscr{B}.$$

Le résultat de la proposition suivante est souvent appelé *formule de transfert.*

Proposition 2.2.14. *Soit $f : (F, \mathscr{B}) \to (\overline{\mathbb{R}}, \mathscr{B}(\overline{\mathbb{R}}))$ une fonction mesurable. La fonction f est μ_φ-intégrable sur F si et seulement si la fonction $f \circ \varphi$ est μ-intégrable sur E et dans ce cas, on a,*

$$\int_E f \circ \varphi \, d\mu = \int_F f \, d\mu_\varphi .$$

Preuve : Supposons d'abord que $f = \mathbb{1}_B$, avec $B \in \mathscr{B}$. Alors on vérifie que $f \circ \varphi = \mathbb{1}_{\varphi^{-1}(B)}$ et l'égalité $\int_E f \circ \varphi \, d\mu = \int_F f \, d\mu_\varphi$ vient de la définition de l'intégrale d'une fonction indicatrice. Si f est une fonction étagée, le résultat se déduit de la linéarité de l'intégrale. Supposons maintenant que f soit une fonction mesurable et positive quelconque. Alors il existe une suite croissante (f_n) de fonctions étagées qui convergent simplement vers f et $(f_n \circ \varphi)$ est une suite croissante de fonctions étagées qui convergent simplement vers $f \circ \varphi$. De plus pour tout n, $\int_E f_n \circ \varphi \, d\mu = \int_F f_n \, d\mu_\varphi$. D'après le théorème de convergence monotone 2.2.1, on a alors

$$\int_E f \circ \varphi = \lim_{n \to +\infty} \int_E f_n \circ \varphi \, d\mu = \lim_{n \to +\infty} \int_F f_n \, d\mu_\varphi = \int_F f \, d\mu_\varphi .$$

Enfin soit f mesurable quelconque, alors $\int_E f^+ \circ \varphi \, d\mu = \int_F f^+ \, d\mu_\varphi$ et $\int_E f^- \circ \varphi \, d\mu = \int_F f^- \, d\mu_\varphi$. De plus $f^+ \circ \varphi = (f \circ \varphi)^+$ et $f^- \circ \varphi = (f \circ \varphi)^-$, de sorte que

$$\int_E (f \circ \varphi)^+ \, d\mu < +\infty \quad \text{et} \quad \int_E (f \circ \varphi)^- \, d\mu < +\infty$$

si et seulement si

$$\int_F f^+ \, d\mu_\varphi < +\infty \quad \text{et} \quad \int_F f^- \, d\mu_\varphi < +\infty ,$$

d'où

$$f \circ \varphi \in \mathscr{L}^1_\mu(E) \quad \text{si et seulement si} \quad f \in \mathscr{L}^1_{\mu_\varphi}(F) .$$

Enfin dans ce cas, on a

$$\int_E f \circ \varphi \, d\mu = \int_E (f \circ \varphi)^+ \, d\mu - \int_E (f \circ \varphi)^- \, d\mu = \int_E f^+ \circ \varphi \, d\mu - \int_E f^- \circ \varphi \, d\mu = \int_F f \, d\mu_\varphi .$$

□

Nous verrons en TD des applications de ce résultat au calcul de la loi de variables aléatoires.

2.2.4. Théorème de convergence dominée :

Theorem 2.2.2. *Soient $(f_n)_{n\geq 0}$ et f des fonctions mesurables de $(E,\mathscr{A})$ dans $(\overline{\mathbb{R}},\mathscr{B}(\overline{\mathbb{R}}))$ telles que*

(*i*) *la suite $(f_n)_{n\geq 0}$ converge vers f, μ-p.p.*

(*ii*) *il existe $g \in \mathscr{L}^1_\mu(E)$ telle que pour tout $n \geq 0$, $|f_n| \leq g$, μ-p.p.*

Alors les fonctions f_n, $n \geq 0$ et f sont intégrables et

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

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

Preuve : Tout d'abord, il est évident que les fonctions f_n, $n \geq 0$ et f sont intégrables. Définissons l'ensemble,

$$A = \{x \in E : \lim_{n\to+\infty} f_n(x) = f(x) \text{ et pour tout } n \geq 0,\ |f_n(x)| \leq g(x)\}\,.$$

Alors $A \in \mathscr{A}$, voir TD. De plus, d'après les hypothèses (i) et (ii),

$$\mu(A^c) = \mu\left(\{f_n \not\to f\} \cup (\cup_{n\geq 0}\{|f_n| > g\})\right) \leq \mu(\{f_n \not\to f\}) + \sum_{n\geq 0} \mu(\{|f_n| > g\}) = 0\,.$$

On remarque que si $x \in A$, alors $|f(x) - f_n(x)| \leq |f(x)| + |f_n(x)| \leq 2g(x)$, ce qui implique que $\mathbb{1}_A(2g - |f - f_n|) \in \mathscr{M}_+$. De plus $\lim_{n\to+\infty} \mathbb{1}_A(2g - |f - f_n|) = 2g\mathbb{1}_A$. Alors, par le lemme de Fatou,

$$\begin{aligned}\int_A 2g\, d\mu &= \int_A \liminf_n (2g - |f - f_n|)\, d\mu \\ &\leq \liminf_n \int_A (2g - |f - f_n|)\, d\mu = \int_A 2g\, d\mu - \limsup_n \int_A |f - f_n|\, d\mu\,.\end{aligned}$$

On en déduit que

$$\limsup_n \int_A |f - f_n|\, d\mu \leq 0$$

et donc $\limsup_n \int_A |f - f_n|\, d\mu = \liminf_n \int_A |f - f_n|\, d\mu = 0$, ce qui montre finalement que $\lim_n \int_A |f - f_n|\, d\mu = \lim_n \int_E |f - f_n|\, d\mu = 0$. $\square$

Corollary 2.2.6. *Soient $(f_n)_{n\geq 0}$ des fonctions mesurables de $(E,\mathscr{A})$ dans $(\overline{\mathbb{R}},\mathscr{B}(\overline{\mathbb{R}}))$ vérifiant $\sum_n \int_E |f_n|\, d\mu < +\infty$. Alors la fonction $\sum_n f_n$ est définie μ-p.p. et est μ-intégrable. De plus,*

$$\sum_n \int_E f_n\, d\mu = \int_E \sum_n f_n\, d\mu\,.$$

Preuve : Remarquons d'abord que puisque $\sum_n \int_E |f_n|\, d\mu < +\infty$, toutes les fonctions f_n sont μ-intégrables. De plus d'après le théorème de convergence monotone, on a

$$\int_E \sum_n |f_n|\, d\mu = \sum_n \int_E |f_n|\, d\mu < +\infty\,.$$

On en déduit que la fonction $\sum_n |f_n|$ est finie μ-p.p., donc $\sum_n f_n(x)$ converge absolument pour μ-presque tout $x \in E$. On en déduit que la fonction $\sum_n f_n$ est bien définie μ-presque partout. On pose par exemple $\sum_n f_n(x) = 0$ pour tout $x \in A = \{\sum_n |f_n| = +\infty\} \in \mathscr{A}$. La fonction $\sum_n f_n$ est alors mesurable comme limite simple de la suite de fonctions mesurables $(\sum_{k=0}^n f_k \mathbb{1}_A)_{n\geq 0}$. Par ailleurs, $\int_E |\sum_n f_n|\, d\mu \leq \int_E \sum_n |f_n|\, d\mu < +\infty$, donc $\sum_n f_n$ est intégrable.

On peut appliquer le théorème de convergence dominée à la suite de fonctions (g_n), où $g_n = \sum_{k=0}^n f_k$. En effet, on a bien

$\lim_{n\to+\infty} g_n = \sum_n f_n$, μ-p.p., d'une part et $|g_n| \leq \sum_n |f_n|$, μ-p.p., d'autre part.

On en déduit que

$$\begin{array}{ccc}
\displaystyle\lim_{n\to+\infty} \int_E \sum_{k=0}^n f_k\, d\mu & = & \displaystyle\int_E \sum_{n\geq 0} f_n\, d\mu \\
\| & & \| \\
\displaystyle\lim_{n\to+\infty} \sum_{k=0}^n \int_E f_k\, d\mu & = & \displaystyle\sum_{n\geq 0} \int_E f_n\, d\mu\,,
\end{array}$$

d'où le résultat. □

2.2.5. Intégrales dépendant d'un paramètre :

Étant donnée une fonction $f : E \times \mathbb{R} \to \overline{\mathbb{R}}$, on s'intéresse ici aux propriétés de continuité et de dérivabilité de la fonction,

$$y \mapsto \int_E f(x,y)\, d\mu(x)\,,$$

lorsque celle-ci est bien définie.

La théorème suivante est souvent appelé le *théorème de continuité sous le signe intégrale.*

Theorem 2.2.3. *Soit $y_0 \in \mathbb{R}$ et $f : E \times \mathbb{R} \to \overline{\mathbb{R}}$. On suppose qu'il existe une fonction mesurable $g \in \mathscr{L}^1_\mu(E)$ telle que*

(i) pour tout $y \in \mathbb{R}$, $x \mapsto f(x,y)$ est mesurable.

(ii) pour μ-presque tout $x \in E$, $y \mapsto f(x,y)$ est continue en y_0,

(iii) pour μ-presque tout $x \in E$ et pour tout $y \in \mathbb{R}$, $|f(x,y)| \leq g(x)$.

Alors la fonction $F : y \mapsto \int_E f(x,y)\, d\mu(x)$ est bien définie sur $\mathbb{R}$ et est continue en y_0.

Preuve : D'après les points (i) et (iii), la fonction $x \mapsto f(x,y)$ est bien intégrable pour tout $y \in \mathbb{R}$. La fonction F est donc bien définie. Pour montrer qu'elle est continue en y_0 il suffit de montrer que pour toute suite (y_n) qui tend vers y_0, $\lim_{n\to+\infty} \int_E f(x,y_n)\, d\mu(x) = \int_E f(x,y_0)\, d\mu(x)$. Mais ceci est une conséquence directe du théorème de convergence dominée. En effet, en posant $f_n(x) = f(x,y_n)$ et $f(x) = f(x,y_0)$, le point (ii) nous donne que pour μ-presque tout $x \in E$, $\lim_{n\to+\infty} f_n(x) = f(x)$ et le point (iii) nous donne que pour μ-presque tout $x \in E$ et pour tout $n \geq 1$, $|f_n(x)| \leq g(x)$, ce qui correspond bien aux hypothèses du théorème de convergence dominée. Par conséquent, nous obtenons que $\lim_{n\to+\infty} \int_E f(x,y_n)\, d\mu(x) = \int_E f(x,y_0)\, d\mu(x)$. □

Soit I un intervalle non vide de $\mathbb{R}$ et $f : E \times I \to \mathbb{R}$. On cherche maintenant à faire le lien entre la dérivabilité de $y \mapsto f(x,y)$ pour presque tout x et celle de la fonction $y \mapsto \int_E f(x,y)\, d\mu(x)$ lorsqu'elle existe.

Le théorème suivant est souvent appelé le *théorème de dérivabilité sous le signe intégrale.*

Theorem 2.2.4. *Soit I un intervalle de $\mathbb{R}$ et soit $f : E \times I \to \mathbb{R}$. On suppose qu'il existe une fonction $g \in \mathscr{L}^1_\mu(E)$ telle que*

(i) pour tout $y \in I$, $x \mapsto f(x,y)$ est intégrable.

(ii) pour μ-presque tout $x \in E$, $y \mapsto f(x,y)$ est dérivable sur I,

(iii) pour μ-presque tout $x \in E$ et pour tout $y \in \mathbb{R}$, $|\frac{\partial}{\partial_y} f(x,y)| \leq g(x)$.

Alors la fonction $F : y \mapsto \int_E f(x,y)\, d\mu(x)$ est dérivable sur I et pour tout $y \in I$,

$$F'(y) = \int_E \frac{\partial}{\partial y} f(x,y)\, d\mu(x)\,.$$

Preuve : Soit $y_0 \in I$ et $(y_n)_{n\geq 1}$ une suite de I qui converge vers y_0 et qui est telle que $y_n \neq y_0$, pour tout $n \geq 1$. On considère la suite de fonctions φ_n définie par

$$\varphi_n(x) = \frac{f(x,y_n) - f(x,y_0)}{y_n - y_0}\,.$$

Remarquons alors que par hypothèse, les fonctions φ_n sont intégrables et $\lim_{n\to+\infty} \varphi_n(x) = \frac{\partial}{\partial_y} f(x, y_0)$, μ-p.p. Par ailleurs, pour μ-presque tout $x \in E$, le théorème des accroisssements finis donne,

$$|\varphi_n(x)| \leq \sup_{y\in[y_n,y_0]} \left| \frac{\partial}{\partial_y} f(x, y) \right| \leq |g(x)| \, .$$

On peut alors appliquer le théorème de convergence dominée qui donne

$$\lim_{n\to+\infty} \frac{F(y_n) - F(y_0)}{y_n - y_0} = \lim_{n\to+\infty} \int_E \varphi_n(x) \, d\mu(x) = \int_E \frac{\partial}{\partial_y} f(x, y) \, d\mu(x) \, .$$

□

Chapitre 3

Tribu produit, mesure produit et intégrales multiples

3.1. Tribu produit

Soient $(E, \mathscr{A}, \mu)$ et $(F, \mathscr{B}, \nu)$ deux espaces mesurés. Notons que l'ensemble $\{A\times B :\ A \in \mathscr{A},\ B \in \mathscr{B}\}$ des pavés à côtés mesurables sur $E \times F$ n'est pas une tribu en général.

Definition 3.1.1. *On appelle tribu produit de $\mathscr{A}$ et $\mathscr{B}$ la tribu notée $\mathscr{A} \otimes \mathscr{B}$ sur l'espace $E \times F$ et définie par*

$$\mathscr{A} \otimes \mathscr{B} = \sigma(\{A \times B :\ A \in \mathscr{A},\ B \in \mathscr{B}\}) .$$

On étend cette définition au produit de n tribus $\mathscr{A}_1, \dots, \mathscr{A}_n$ sur les espaces respectifs $E_1, \dots, E_n$ par

$$\mathscr{A}_1 \otimes \cdots \otimes \mathscr{A}_n = \sigma(\{A_1 \times \cdots \times A_n :\ A_1 \in \mathscr{A}_1,\ \dots, A_n \in \mathscr{A}_n\}) .$$

La preuve de la proposition suivante est laissée en exercice.

Proposition 3.1.1. *Le produit de tribus est associatif. Plus précisément, si $\mathscr{A}$, $\mathscr{B}$ et $\mathscr{C}$ sont des tribus sur les espaces respectifs E, F et G, alors*

$$\mathscr{A} \otimes (\mathscr{B} \otimes \mathscr{C}) = (\mathscr{A} \otimes \mathscr{B}) \otimes \mathscr{C} = \mathscr{A} \otimes \mathscr{B} \otimes \mathscr{C} .$$

Notons π_1 et π_2 les projections canoniques de $E \times F$ sur E et F, respectivement. Ce sont les applications définies par

$$\pi_1 : (x, y) \in E \times F \mapsto x \quad \text{et} \quad \pi_2 : (x, y) \in E \times F \mapsto y .$$

Proposition 3.1.2. *Les projections π_1 et π_2 sont mesurables de $(E \times F, \mathscr{A} \otimes \mathscr{B})$ dans $(E, \mathscr{A})$ et de $(E \times F, \mathscr{A} \otimes \mathscr{B})$ dans $(F, \mathscr{B})$, respectivement.*

De plus, la tribu $\mathscr{A} \otimes \mathscr{B}$ est la plus petite tribu rendant les projections mesurables.

Preuve : Soient $A \in \mathscr{A}$ et $B \in \mathscr{B}$. Alors on vérifie que $\pi_1^{-1}(A) = A \times F \in \mathscr{A} \otimes \mathscr{B}$ et $\pi_2^{-1}(A) = E \times B \in \mathscr{A} \otimes \mathscr{B}$, ce qui montre le résultat.

Montrons la seconde assertion. Soit $\mathscr{F}$ une tribu sur $E \times F$. Supposons que π_1 et π_2 soient respectivement mesurables de $(E \times F, \mathscr{F})$ dans $(E, \mathscr{A})$ et de $(E \times F, \mathscr{F})$ dans $(F, \mathscr{B})$. Alors, pour tout élément $A \times B \in \mathscr{A} \otimes \mathscr{B}$, $A \times B = \pi_1^{-1}(A) \cap \pi_2^{-1}(B) \in \mathscr{F}$, donc $A \times B \in \mathscr{F}$, ce qui implique que $\mathscr{A} \otimes \mathscr{B} \subset \mathscr{F}$. $\square$

Proposition 3.1.3. *La tribu borélienne vérifie*

$$\mathscr{B}(\mathbb{R}\times\mathbb{R}) = \mathscr{B}(\mathbb{R})\otimes\mathscr{B}(\mathbb{R})\,.$$

Preuve : Les projections $\pi_1 : (x,y)\in\mathbb{R}\times\mathbb{R}\mapsto x$ et $\pi_2 : (x,y)\in\mathbb{R}\times\mathbb{R}\mapsto y$ sont continues, donc $\mathscr{B}(\mathbb{R}\times\mathbb{R})$-mesurables. La proposition 3.1.2 implique alors que $\mathscr{B}(\mathbb{R})\otimes\mathscr{B}(\mathbb{R})\subset\mathscr{B}(\mathbb{R}\times\mathbb{R})$. D'autre part, on a $\mathscr{B}(\mathbb{R}\times\mathbb{R}) = \sigma(\{I\times J : I \text{ et } J \text{ intervalles de } \mathbb{R}\})\subset\sigma(\{A\times B : A\in\mathscr{B}(\mathbb{R}),\, B\in\mathscr{B}(\mathbb{R})\}) = \mathscr{B}(\mathbb{R})\otimes\mathscr{B}(\mathbb{R})$. □

Cette identité se généralise facilement en

$$\mathscr{B}(\mathbb{R}^d) = \mathscr{B}(\mathbb{R})\otimes\cdots\otimes\mathscr{B}(\mathbb{R})\,.$$

Le corollaire suivant est laissé en exercice.

Corollary 3.1.7. *Soient $(E,\mathscr{A})$, $(F,\mathscr{B})$ et $(G,\mathscr{C})$ des espaces mesurables et $f = (f_1,f_2) : E\to F\times G$ une application. Alors f est mesurable de $(E,\mathscr{A})$ dans $(F\times G,\mathscr{B}\otimes\mathscr{C})$ si et seulement si f_1 et f_2 sont mesurables de $(E,\mathscr{A})$ dans $(F,\mathscr{B})$ et de $(E,\mathscr{A})$ dans $(G,\mathscr{C})$, respectivement.*

Definition 3.1.2. *Soit $C\subset E\times F$ un ensemble quelconque. On appelle sections de C les ensembles suivants :*

$$\begin{aligned} C_x &= \{y\in F : (x,y)\in C\} \quad \textit{pour tout } x\in E,\\ C^y &= \{x\in E : (x,y)\in C\} \quad \textit{pour tout } y\in F\,. \end{aligned}$$

En particulier, si C est de la forme $A\times B$, alors

$$C_x = \begin{cases} B & \text{si } x\in A\\ \emptyset & \text{sinon}\end{cases} \quad \text{et} \quad C^y = \begin{cases} A & \text{si } y\in B\\ \emptyset & \text{sinon.}\end{cases}$$

Proposition 3.1.4. *Si $C\in\mathscr{A}\otimes\mathscr{B}$, alors pour tout $x\in E$, $C_x\in\mathscr{B}$ et pour tout $y\in F$, $C^y\in\mathscr{A}$.*

Preuve : Soit $x\in E$. Alors on vérifie que $\mathscr{F}_x = \{C\in\mathscr{A}\otimes\mathscr{B} : C_x\in\mathscr{B}\}$ est une tribu sur $E\times F$ (exercice). Par ailleurs, d'après la remarque qui précède cette proposition, $\mathscr{A}\otimes\mathscr{B}\subset\mathscr{F}_x$. On en déduit que $\mathscr{F}_x = \mathscr{A}\otimes\mathscr{B}$, d'où le résultat.
Le cas de C^y est similaire. □

Corollary 3.1.8. *Soit $f : (E\times F,\mathscr{A}\otimes\mathscr{B})\to(G,\mathscr{C})$ une application mesurable. On définit les applications*

$$\begin{aligned} &f_x : y\in F\to f(x,y)\,, \quad \textit{pour tout } x\in E,\\ &f^y : x\in E\to f(x,y)\,, \quad \textit{pour tout } y\in F. \end{aligned}$$

Alors, pour tous $x\in E$ et $y\in F$, $f_x : (F,\mathscr{B})\to(G,\mathscr{C})$ et $f^y : (E,\mathscr{A})\to(G,\mathscr{C})$ sont mesurables.

Preuve : Soit $C \in \mathscr{C}$. On remarque que $f_x^{-1}(C) = \{y \in F : f(x,y) \in C\} = \{y \in F : (x,y) \in f^{-1}(C)\} = f^{-1}(C)_x \in \mathscr{B}$, car $f^{-1}(C) \in \mathscr{A} \otimes \mathscr{B}$ et d'après la proposition 3.1.4. De même on a$(f^y)^{-1}(C) = f^{-1}(C)_y \in \mathscr{A}$. □

3.2. Mesure produit

Les preuves du lemme et du théorème suivants font appel au théorème des classes monotones que nous n'avons pas vu dans ce cours. Nous ne les démontrerons pas.

Lemme 3.2.1. *On suppose que μ et ν sont σ-finies. Soit $C \in \mathscr{A} \otimes \mathscr{B}$. Alors les fonctions*

$$\begin{aligned} &x \in E \to \nu(C_x)\,, \\ &y \in F \to \mu(C^y) \end{aligned}$$

sont mesurables.

Le résultat suivant est appelé le théorème d'unicité des mesures.

Theorem 3.2.1. *Soient μ et ν deux mesures sur un espace mesurable $(E, \mathscr{A})$. Supposons que μ et ν coïncident sur une classe $\mathscr{C} \subset \mathscr{A}$ stable par intersection finie, contenant E et qui engendre $\mathscr{A}$, c'est à dire $\sigma(\mathscr{C}) = \mathscr{A}$. S'il existe une suite croissante de parties $E_n \in \mathscr{C}$ telles que $\mu(E_n) < \infty$ et $E = \cup_n E_n$, alors $\mu = \nu$.*

Theorem 3.2.2. *Soient μ et ν deux mesures σ-finies sur $(E, \mathscr{A})$ et $(F, \mathscr{B})$ respectivement. Il existe une unique mesure sur l'espace produit $(E \times F, \mathscr{A} \otimes \mathscr{B})$ que nous noterons $\mu \otimes \nu$, telle que pour tout $A \times B \in \mathscr{A} \otimes \mathscr{B}$,*

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

De plus $\mu \otimes \nu$ est σ-finie et pour tout $C \in \mathscr{A} \otimes \mathscr{B}$,

$$\mu \otimes \nu(C) = \int_E \nu(C_x)\,d\mu(x) = \int_F \mu(C^y)\,d\nu(y)\,.$$

Preuve : On pose $\mu \otimes \nu(C) = \int_E \nu(C_x)\,d\mu(x)$. Le lemme 3.2.1 montre que $\mu \otimes \nu$ est bien définie. Par ailleurs, c'est bien une mesure sur l'espace produit $(E \times F, \mathscr{A} \otimes \mathscr{B})$. En effet

_ $\mu \otimes \nu(\emptyset) = \int_E \nu(\emptyset_x)\,d\mu(x) = \int_E \nu(\emptyset)\,d\mu(x) = 0$,

_ soient $C_n \in \mathscr{A} \otimes \mathscr{B}$, $n \geq 1$, deux à deux disjoints. Alors on vérifie que $(\cup_{n\geq 1} C_n)_x = \cup_{n\geq 1}(C_n)_x$. On a donc $\mu \otimes \nu(\cup_{n\geq 1} C_n) = \int_E \sum_n \nu((C_n)_x)\,d\mu(x) = \sum_n \mu \otimes \nu(C_n)$ par le théorème de convergence monotone de Beppo Levi.

Par ailleurs, si $A \times B \in \mathscr{A} \otimes \mathscr{B}$, alors on vérifie que $\nu((A\times B)_x) = \nu(B)\mathbb{1}_A(x)$, d'où

$$\mu \otimes \nu(A \times B) = \int_E \nu(B)\mathbb{1}_A(x)\, d\mu(x) = \mu(A)\nu(B)\,.$$

Montrons maintenant que $\mu \otimes \nu$ est l'unique mesure sur $(E \times F, \mathscr{A} \otimes \mathscr{B})$ telle que $\mu\otimes\nu(A\times B) = \mu(A)\nu(B)$, $A \times B \in \mathscr{A} \otimes \mathscr{B}$. Si m est une autre mesure vérifiant ces propriétés, alors $\mu\otimes\nu$ et m coïncident sur $\{A \times B :\ A \in \mathscr{A},\ B \in \mathscr{B}\}$, qui est stable par intersection finie et contient $E \times F$. Soient (E_n) et (F_n) les suites de $\mathscr{A}$ et $\mathscr{B}$ exprimant que μ et ν sont σ-finies. Quitte à poser $E'_n = \cup_{k=1}^n E_k$ et $F'_n = \cup_{k=1}^n F_k$ pour tout $n \in \mathbb{N}$, celles-ci peuvent être supposées croissantes et l'on a

$$\mu \otimes \nu(E_n \times F_n) = m(E_n \times F_n) = \mu(E_n)\nu(F_n) < \infty\,.$$

Les hypothèses du théorème 3.2.1 d'unicité des mesures sont donc vérifiées et $\mu\otimes\nu = m$, d'où l'unicité. De plus l'égalité ci-dessus montre que $\mu \otimes \nu$ est σ-finie.

On vérifie comme ci-dessus que $C \in \mathscr{A} \otimes \mathscr{B} \mapsto \int_F \mu(C^y)\, d\nu(y)$ définit bien une mesure sur $(E \times F, \mathscr{A} \otimes \mathscr{B})$ et que pour tout $A \times B \in \mathscr{A} \otimes \mathscr{B}$, on a

$$\int_F \mu((A \times B)^y)\, d\nu(y) = \mu(A)\nu(B)\,.$$

L'unicité nous donne alors que $\mu \otimes \nu(C) = \int_F \mu(C^y)\, d\nu(y)$ pour tout $C \in \mathscr{A} \otimes \mathscr{B}$, ce qui achève la démonstration. $\square$

Rappelons que λ désigne la mesure de Lebesgue sur $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$.

Proposition 3.2.1. *La mesure* $\underbrace{\lambda \otimes \cdots \otimes \lambda}_{d}$ *coïncide avec la mesure de Lebesgue* λ_d *sur* $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$.

Preuve : On raisonne par récurrence sur d. Notons $\underbrace{\lambda \otimes \cdots \otimes \lambda}_{d} = \lambda^d$. Le résultat est vrai pour $d = 1$. Supposons que $\lambda^{d-1} = \lambda_{d-1}$. Alors pour tout pavé $P = \prod_{k=1}^d I_k$, où les I_k sont des intervalles de $\mathbb{R}$, on a

$$\lambda^d(P) = \lambda_{d-1}\left(\prod_{k=1}^{d-1} I_k\right)\lambda(I_d) = \prod_{k=1}^{d} \lambda(I_k) = \lambda_d(P)\,.$$

Puisque λ^d et λ_d coïncident sur les pavés de $\mathbb{R}^d$, le théorème 3.2.1 d'unicité de la mesure de Lebesgue sur $\mathbb{R}^d$ permet de conclure. $\square$

3.3. Théorèmes de Fubini

Le théorème suivant dit de Fubini-Tonelli définit les intégrales doubles lorsque les mesures μ et ν sont σ-finies et lorsque les fonctions mesurables définies sur l'espace produit sont positives.

Theorem 3.3.1 (Théorème de Fubini-Tonelli)**.** *Soient μ et ν deux mesures σ-finies sur $(E, \mathscr{A})$ et $(F, \mathscr{B})$ respectivement. Soit $f : (E \times F, \mathscr{A} \otimes \mathscr{B}) \to (\overline{\mathbb{R}}_+, \mathscr{B}(\overline{\mathbb{R}}_+))$ une application mesurable.*

1. *Les fonctions $x \mapsto \int_F f(x,y)\,d\nu(y)$ et $y \mapsto \int_E f(x,y)\,d\mu(x)$ sont mesurables.*
2. *On a l'égalité*

$$\int_{E\times F} f(x,y)\,d\mu\otimes\nu(x,y) = \int_E \left(\int_F f(x,y)\,d\nu(y)\right) d\mu(x) = \int_F \left(\int_E f(x,y)\,d\mu(x)\right) d\nu(y)\,.$$

Preuve : On commence par supposer que f est de la forme $\mathbb{1}_C$, où $C \in \mathscr{A} \otimes \mathscr{B}$. Puisque $\mathbb{1}_C(x,y) = \mathbb{1}_{C_x}(y) = \mathbb{1}_{C^y}(x)$, le lemme 3.2.1 implique alors le point 1. et le théorème 3.2.2 donne le point 2. On en déduit par linéarité que le résultat est vrai pour les fonctions étagées et positives.

Supposons que $f : (E \times F, \mathscr{A} \otimes \mathscr{B}) \to (\overline{\mathbb{R}}_+, \mathscr{B}(\overline{\mathbb{R}}_+))$ soit quelconque. On sait qu'il existe une suite croissante de fonctions étagées et positives (f_n) qui converge simplement vers f. Pour tout n, $((f_n)_x)$ est une suite croissante de fonctions mesurables, positives et le théorème de convergence monotone implique que

$$\int_F f(x,y)\,d\nu(y) = \lim_n \int_F f_n(x,y)\,d\nu(y)\,.$$

Ainsi la fonction $x \mapsto \int_F f(x,y)\,d(y)$ est mesurable comme limite simple de fonctions mesurables. On montre de même que $y \mapsto \int_E f(x,y)\,d\mu(x)$ est mesurable en appliquant le théorème de convergence monotone à la suite $(f_n)^y$, $n \geq 1$, pour $y \in F$.

Les fonctions $x \mapsto \int_F f_n(x,y)\,d(y)$ forment une suite croissante de fonctions mesurables et positives. On utilise à nouveau le théorème de convergence monotone, pour obtenir

$$\begin{array}{ccc} \displaystyle\int_E \left(\int_F f_n(x,y)\,d\nu(y)\right) d\mu(x) & \longrightarrow_{n\to+\infty} & \displaystyle\int_E \left(\int_F f(x,y)\,d\nu(y)\right) d\mu(x) \\ \| & & \\ \displaystyle\int_{E\times F} f_n(x,y)\,d\mu\otimes\nu(x,y) & \longrightarrow_{n\to+\infty} & \displaystyle\int_{E\times F} f(x,y)\,d\mu\otimes\nu(x,y)\,, \end{array}$$

d'où la première égalité. On montre la seconde de la même manière. □

Remarquons que l'hypothèse de σ-finitude est cruciale. En effet, soit m la mesure de comptage sur $(\mathbb{R}, \mathscr{P}(\mathbb{R}))$. Considérons la diagonale $\Delta = \{(x,y) \in \mathbb{R}^2 : x = y\}$. Alors $\Delta \in \mathscr{B}(\mathbb{R}) \otimes \mathscr{B}(\mathbb{R}) \subset \mathscr{P}(\mathbb{R}) \otimes \mathscr{B}(\mathbb{R})$, donc $\mathbb{1}_\Delta$ est $\mathscr{P}(\mathbb{R}) \otimes \mathscr{B}(\mathbb{R})$-mesurable et

$$\begin{aligned}
\int_{\mathbb{R}} \left(\int_{\mathbb{R}} \mathbb{I}_{\Delta}(x,y)\, dm(x) \right) d\lambda(y) &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \mathbb{I}_{\{y\}}(x)\, dm(x) \right) d\lambda(y) \\
&= \int_{\mathbb{R}} m(\{y\})\, d\lambda(y) = +\infty \\
\int_{\mathbb{R}} \left(\int_{\mathbb{R}} \mathbb{I}_{\Delta}(x,y)\, d\lambda(x) \right) dm(y) &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \mathbb{I}_{\{y\}}(x)\, d\lambda(x) \right) dm(y) \\
&= \int_{\mathbb{R}} \lambda(\{y\})\, dm(y) = 0\,.
\end{aligned}$$

Des exemples sont fournis par les fonctions intégrables au sens de Riemann sur des domaines simples, voir section **??** et le théorème **??** dans cette section.

Theorem 3.3.2 (Théorème de Fubini)**.** *Soient μ et ν deux mesures σ-finies sur $(E, \mathscr{A})$ et $(F, \mathscr{B})$ respectivement. Soit $f : (E \times F, \mathscr{A} \otimes \mathscr{B}) \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ une application mesurable telle que $f \in \mathscr{L}^1_{\mu\otimes\nu}(E \times F)$.*

1. *$f_x : y \mapsto f(x,y) \in \mathscr{L}^1_\nu(F)$, μ-p.p. x et $f^y : x \mapsto f(x,y) \in \mathscr{L}^1_\mu(E)$, ν-p.p. y,*
1. *$x \mapsto \int_F f(x,y)\, d\nu(y) \in \mathscr{L}^1_\mu(E)$ et $y \mapsto \int_E f(x,y)\, d\mu(x) \in \mathscr{L}^1_\nu(F)$.*
2. *On a l'égalité*

$$\int_{E\times F} f(x,y)\, d\mu\otimes\nu(x,y) = \int_E \left(\int_F f(x,y)\, d\nu(y) \right) d\mu(x) = \int_F \left(\int_E f(x,y)\, d\mu(x) \right) d\nu(y)\,.$$

Preuve : Puisque f est $\mathscr{A} \otimes \mathscr{B}$-mesurable, le théorème de Fubini-Tonelli permet d'écrire

$$\int_{E\times F} |f(x,y)|\, d\mu\otimes\nu(x,y) = \int_E \left(\int_F |f(x,y)|\, d\nu(y) \right) d\mu(x) = \int_F \left(\int_E |f(x,y)|\, d\mu(x) \right) d\nu(y) < +\infty\,.$$

On en déduit que pour ν-presque tout $y \in F$, $\int_E |f(x,y)|\, d\mu(x) < +\infty$, ce qui montre que $f^y \in \mathscr{L}^1_\mu(E)$. De même, on obtient que $f_x \in \mathscr{L}^1_\nu(F)$.

Par l'inégalité triangulaire, on en déduit alors que $x \mapsto \int_F f(x,y)\, d\nu(y) \in \mathscr{L}^1_\mu(E)$ et $y \mapsto \int_E f(x,y)\, d\mu(x) \in \mathscr{L}^1_\nu(F)$.

Enfin lorsque f est à valeurs dans $\mathbb{R}$, on sait par Fubini-Tonelli que f^+ et f^- vérifient

$$\int_{E\times F} f^\pm(x,y)\, d\mu\otimes\nu(x,y) = \int_E \left(\int_F f^\pm(x,y)\, d\nu(y) \right) d\mu(x) = \int_F \left(\int_E f^\pm(x,y)\, d\mu(x) \right) d\nu(y)\,.$$

On en déduit alors l'égalité pour f. $\square$

Remarquons que l'hypothèse d'intégrabilité de f pour la mesure produit $\mu \otimes \nu$ est cruciale. Par exemple,

$$\int_0^1 \left(\int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} \, dy \right) dx = \frac{\pi}{4} \quad \text{mais} \quad \int_0^1 \left(\int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} \, dx \right) dy = -\frac{\pi}{4} \, . \tag{3.3.1}$$

En effet, d'une part pour tout $x \in]0, 1]$ fixé, la fonction $y \mapsto \frac{x^2 - y^2}{(x^2+y^2)^2}$ est intégrable sur $[0, 1]$ et vaut

$$\int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} \, dy = \left[\frac{y}{x^2 + y^2} \right]_0^1 = \frac{1}{x^2 + 1} \, .$$

(Prolongeons la fonction $x \mapsto \int_0^1 \frac{x^2-y^2}{(x^2+y^2)^2} \, dy$ par exemple par la valeur 1 pour $x = 0$.) D'autre part, la fonction $x \mapsto \frac{1}{x^2+1}$ est intégrable sur $[0, 1]$ et vaut

$$\int_0^1 \frac{1}{x^2 + 1} \, dx = [\arctan x]_0^1 = \frac{\pi}{4} \, .$$

Le calcul de la seconde intégrale en (3.3.1) se fait de même. Ainsi les intégrales en (3.3.1) correspondent chacune à deux intégrales de Lebesgue simples calculées séparément.

Le théorème de Fubini montre que la fonction $(x, y) \mapsto \frac{x^2-y^2}{(x^2+y^2)^2}$ n'est pas intégrable sur $[0, 1] \times [0, 1]$ puisque les deux intégrales en (3.3.1) ont une valeur différente.

Il faut donc retenir qu'il ne suffit pas que les intégrales $\int_E \left(\int_F f(x, y) \, d\nu(y) \right) d\mu(x)$ et $\int_F \left(\int_E f(x, y) \, d\mu(x) \right) d\nu(y)$ existent et soient finies pour que l'on puisse changer l'ordre d'intégration sans changer la valeur de l'intégrale.

3.4. Changement de variables

Dans cette section nous étudions le changement de variables pour l'intégrale des fonctions boréliennes $f : (\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d), \lambda_d) \to (\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$. Commençons par étudier le cas d'un changement de variables affine. Nous noterons $GL_d(\mathbb{R})$ l'ensemble des matrices carrées $d \times d$, inversibles et à coefficients dans $\mathbb{R}$.

Lemme 3.4.1. *Soient $A \in GL_d(\mathbb{R})$, $b \in \mathbb{R}^d$ et $\varphi : x \in \mathbb{R}^d \mapsto Ax + b$. Alors*

$$(\lambda_d)_\varphi = \frac{1}{|det(A)|} \lambda_d \, ,$$

où $(\lambda_d)_\varphi$ désigne la mesure image de λ_d par φ.

Preuve : Tout d'abord, on a par définition $(\lambda_d)_\varphi(B) = \lambda_b(A^{-1}B - A^{-1}b) = \lambda_d(A^{-1}B)$, pour tout $B \in \mathscr{B}(\mathbb{R}^d)$, puisque λ_d est invariante par translation. On peut donc supposer, sans perde de généralité, que $b = 0$. Remarquons alors que la mesure $(\lambda_d)_\varphi$

est également invariante par translation. En effet, pour tout $B \in \mathscr{B}(\mathbb{R}^d)$ et tout $a \in \mathbb{R}^d$,

$$(\lambda_d)_\varphi(B+a) = \lambda_d(A^{-1}B + A^{-1}a) = \lambda_d(A^{-1}B) = (\lambda_d)_\varphi(B)\,.$$

D'après la proposition 2.1.4 sur l'unicité de la mesure de Lebesque, les mesures λ_d et $(\lambda_d)_\varphi$ diffèrent d'une constante. Montrons que cette constante est égale à $|\det(A)|^{-1}$.

Si A est une matrice orthogonale et B est la sphère unité de $\mathbb{R}^d$, alors on a $A^{-1}B = B$ et donc $(\lambda_d)_\varphi(B) = \lambda_d(B)$. Dans ce cas on vérifie bien que $\lambda_d = |\det(A)|^{-1}(\lambda_d)_\varphi$, puisque $|\det(A)|^{-1} = 1$.

Supposons maintenant que A soit symétrique définie positive. On sait alors qu'il existe une matrice orthogonale P et une matrice diagonale D telles que $A = {}^tPDP$. Considérons alors le borélien $B = {}^tPD^{-1}[0,1]^d$. On remarque que $\lambda_d(A^{-1}B) = \lambda_d({}^tP[0,1]^d) = \lambda_d([0,1]^d) = 1$, d'après ce qui vient d'être montré. Par ailleurs, $\lambda_d(B) = \lambda_d(D^{-1}[0,1]^d) = \lambda_d(\prod_{i=1}^d [0, 1/\alpha_i]) = (\alpha_1 \dots \alpha_d)^{-1}$, où les α_i sont les valeurs propres de A. On vérifie donc bien encore dans ce cas que $\lambda_d = |\det(A)|^{-1}(\lambda_d)_\varphi$.

Enfin supposons que A est une matrice quelconque de $GL_d(\mathbb{R})$. D'après la décomposition polaire de A, il existe une matrice orthogonale Q et une matrice symétrique définie positive S, telles que $A = QS$. On déduit alors de ce qui a été montré plus haut que pour tout $B \in \mathscr{B}(\mathbb{R}^d)$, $\lambda_d(S^{-1}Q^{-1}B) = |\det(S)|^{-1}\lambda_d(Q^{-1}B) = |\det(S)|^{-1}\lambda_d(B)$. On obtient bien que $\lambda_d = |\det(S)|^{-1}(\lambda_d)_\varphi = |\det(A)|^{-1}(\lambda_d)_\varphi$, puisque $\det(A) = \det(Q)\det(S) = \det(S)$. □

Le théorème de changement de variable dans le cas d'une transformation affine se déduit directement de ce lemme.

Corollary 3.4.9. *Soient $A \in GL_d(\mathbb{R})$ et $b \in \mathbb{R}^d$. Pour toute fonction $f \in \mathscr{L}^1_{\lambda_d}(\mathbb{R}^d)$,*

$$\int_{\mathbb{R}^d} f(y)\, d\lambda_d(y) = \int_{\mathbb{R}^d} f(Ax+b)|det(A)|\, d\lambda_d(x)\,.$$

Preuve : Il suffit d'appliquer la formule de transfert établie à la proposition 2.2.14 en exprimant $(\lambda_d)_\varphi$ à l'aide du lemme précédent. □

Nous renvoyons au Chapitre 0 dans lequel la notion de difféomorphisme de classe $\mathcal{C}^1$ a été rappelée. Le théorème général de changement de variables est le suivant.

Theorem 3.4.1. *Soient U et V des ouverts de $\mathbb{R}^d$ et soit $\varphi : U \to V$ un difféomorphisme de classe $\mathcal{C}^1$. Pour toute fonction borélienne $f : U \to \mathbb{R}_+$,*

$$\int_V f(y)\, d\lambda_d(y) = \int_U f \circ \varphi(x)|det(J_\varphi(x))|\, d\lambda_d(x)\,. \tag{3.4.1}$$

Pour toute fonction $f : V \to \mathbb{R}$, $f \in \mathscr{L}^1(V)$ si et seulement si $f\circ\varphi|det(J_\varphi)| \in \mathscr{L}^1(U)$ et dans ce cas l'égalité (3.4.1) *est vérifiée.*

Nous ne montrons pas ce théorème ici. La preuve assez technique consiste à approcher localement le difféomorphisme φ par son application linéaire tangente φ' et à utiliser le cas affine vu au corollaire précédent.

On rappelle enfin le théorème d'inversion globale qui permet de vérifier qu'une application est un difféomorphisme de classe $\mathcal{C}^1$.

Theorem 3.4.2. *Soit U un ouvert de $\mathbb{R}^d$. La fonction $\varphi : U \to \mathbb{R}^d$ est un difféomorphisme de classe $\mathcal{C}^1$ dans son image $\varphi(U)$ si et seulement si*

(i) *φ est injective sur U,*

(ii) *φ est de classe $\mathcal{C}^1$ sur U,*

(iii) *pour tout $x \in U$, $J_\varphi(x) \in GL_d(\mathbb{R})$.*