Université Pierre & Marie Curie (Paris 6)

Licence de Mathématiques L3

UE LM364 Intégration 1 & UE LM365 Intégration 2

Année 2014–15

Théorie de la Mesure et Intégration

Amaury Lambert [1]

1. Responsable des deux UE de l'année universitaire 2008–09 à l'année 2011–12, puis uniquement de l'UE LM365, en 2012–13 et 2013–14. Mél : amaury.lambert@upmc.fr

Table des matières

Première partie

LM364 – Intégration 1

Chapitre 1

Suites, ensembles et suites d'ensembles

1.1 La droite achevée

Définition 1.1 *On appelle* droite achevée *l'ensemble* $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$.

On considérera toujours la droite achevée comme l'espace métrique associé à l'une des distances du type $d(x, y) := |f(x) - f(y)|$ où $f(x) = \frac{x}{\sqrt{x^2+1}}$ si $x \in \mathbb{R}$ et $f(\pm\infty) = \pm 1$. Autrement dit, $\bar{\mathbb{R}}$ est muni de la topologie usuelle de $\mathbb{R}$, complétée avec les notions usuelles de convergence vers $+\infty$ et vers $-\infty$.
La droite achevée est munie d'un ordre total que le lecteur aura deviné : pour tous $x \leq y \in \mathbb{R}$,

$$-\infty < x \leq y < +\infty.$$

La droite achevée est également munie des opérations algébriques usuelles, avec les conventions suivantes :

$$+\infty + \infty = +\infty, \quad -\infty - \infty = -\infty, \quad a + \infty = +\infty, \quad a - \infty = -\infty,$$

pour tout $a \in \mathbb{R}$, ainsi que

$$0 \times \infty = 0,$$

et

$$a \in]0, \infty] \;\Rightarrow\; a \times \infty = +\infty, \quad a \in [-\infty, 0[\;\Rightarrow\; a \times \infty = -\infty.$$

Remarque 1.2 *Tout au long de ce cours, il faudra acquérir le réflexe de* NE JAMAIS ÉCRIRE *aucune des opérations interdites suivantes :* $(+\infty) - (+\infty)$*, ainsi que* $(-\infty) - (-\infty)$*, et encore* $(\pm\infty)/(\pm\infty)$.

Une suite numérique est une suite à valeurs dans $\mathbb{R}$ ou dans $\bar{\mathbb{R}}$.

1.2 Rappels sur les suites et séries numériques

Définition 1.3 *On dit que* $a \in \bar{\mathbb{R}}$ *est une* valeur d'adhérence *de la suite* (u_n) *s'il existe une suite extraite* $(u_{\varphi(n)})$ *qui converge vers* a.

Exemple 1.4 *Les valeurs d'adhérence de la suite* $(\cos(\pi n/2))$ *sont* $-1, 0$ *et* 1. *Celles de la suite* $((-1)^n + \frac{1}{n})$ *sont* -1 *et* $+1$.

Notation 1.5 (importante) *Lorsqu'une suite* (u_n) *est croissante (resp. décroissante), on notera souvent* $\lim_n \uparrow u_n$ *(resp.* $\lim_n \downarrow u_n$*) sa limite, pour rappeler que la suite est monotone, et surtout pour indiquer que cette limite existe* donc *toujours (dans* $\bar{\mathbb{R}}$*).*

Définition 1.6 *La borne supérieure (*$\in \bar{\mathbb{R}}$*) de l'ensemble des valeurs d'adhérence de la suite* (u_n) *est aussi une valeur d'adhérence de* (u_n). *On la note* $\overline{\lim}_{n\to\infty} u_n$ *ou* $\limsup_{n\to\infty} u_n$. *C'est donc la plus grande valeur d'adhérence de* (u_n) *et elle vérifie*

$$\overline{\lim_n}\, u_n = \lim_n \downarrow (\sup_{k\geq n} u_k) = \inf_n (\sup_{k\geq n} u_k).$$

De façon analogue, la plus petite valeur d'adhérence de (u_n) *est notée* $\underline{\lim}_{n\to\infty} u_n$ *ou* $\liminf_{n\to\infty} u_n$...

Définition 1.7 *On dit que la série de terme général* (u_n) *est* absolument convergente *si la suite des sommes partielles* $(\sum_{k=0}^n |u_k|)_n$ *converge dans* $\mathbb{R}$, *ce que l'on note également* $\sum_n |u_n| < \infty$.

Théorème 1.8 *Si la série de terme général* (u_n) *est absolument convergente, alors elle est convergente, c'est-à-dire que la suite des sommes partielles* $(\sum_{k=0}^n u_k)_n$ *converge dans* $\mathbb{R}$.

Proposition 1.9 *La* somme *de la série de terme général* $u_n \geq 0$ *(c'est-à-dire la limite de la suite des sommes partielles, qui existe toujours dans* $\bar{\mathbb{R}}_+$*) ne dépend pas de l'ordre de sommation.*

Démonstration. Soit une bijection $\varphi : \mathbb{N} \longrightarrow \mathbb{N}$. On veut montrer que la suite $S'_n := \sum_{k=0}^n u_{\varphi(k)}$ a même limite dans $\bar{\mathbb{R}}_+$ que $S_n := \sum_{k=0}^n u_k$.

Soit $n \geq 0$ et $N := \max\{\varphi(0), \ldots, \varphi(n)\}$. Alors $S'_n = u_{\varphi(0)} + \cdots + u_{\varphi(n)} \leq \sum_{j=0}^N u_j = S_N$, donc $S'_n \leq S_N \leq S_\infty$. Faisant tendre $n \to \infty$ on obtient que $S'_\infty \leq S_\infty$. L'inégalité opposée s'obtient par symétrie. □

1.3 Ensembles

1.3.1 Terminologie

Soit E un ensemble. Mettons-nous d'accord sur un peu de terminologie.

- $A \subseteq E$ sera appelé *sous-ensemble* ou *partie* de E ;
- $\mathscr{P}(E) := \{\text{parties de } E\}$;
- $\mathscr{A} \subseteq \mathscr{P}(E)$ sera appelé *famille* de parties de E ou *classe* de parties de E plutôt qu'ensemble de sous ensembles de E ou partie de $\mathscr{P}(E)$;
- dans quelques rares cas, nous serons amenés à considérer des ensembles de familles de parties, que l'on appellera alors *collections* de familles de parties de E.

1.3.2 Opérations classiques

Recensons quelques opérations classiques sur les parties d'un ensemble E. Soient A_1 et A_2 deux parties de E.

— *La réunion* de A_1 et A_2, notée $A_1 \cup A_2$: $\forall x \in E$,

$$x \in A_1 \cup A_2 \Leftrightarrow \exists i \in \{1, 2\}, x \in A_i$$

— *L'intersection* de A_1 et A_2, notée $A_1 \cap A_2$: $\forall x \in E$,

$$x \in A_1 \cap A_2 \Leftrightarrow \forall i \in \{1, 2\}, x \in A_i$$

— *Le complémentaire* de A_1, noté ${}^cA_1$: $\forall x \in E$,

$$x \in {}^cA_1 \Leftrightarrow x \notin A_1$$

— *La différence* de A_1 avec A_2, notée $A_1 \setminus A_2$ et dite *différence propre* dans le cas où $A_2 \subseteq A_1$: $\forall x \in E$,

$$x \in A_1 \setminus A_2 \Leftrightarrow x \in A_1 \text{ et } x \notin A_2$$

— *La différence symétrique* de A_1 et A_2, notée $A_1 \Delta A_2$: $\forall x \in E$,

$$x \in A_1 \Delta A_2 \Leftrightarrow x \in A_1 \cup A_2 \text{ et } x \notin A_1 \cap A_2.$$

Remarque 1.10 *Remarquer l'association de la réunion avec le quantificateur « $\exists$ », de l'intersection avec le quantificateur « $\forall$ », ainsi que l'association du passage au complémentaire avec la négation et de l'inclusion avec l'implication : $A_1 \subseteq A_2$ ssi $\forall x \in E$,*

$$x \in A_1 \Rightarrow x \in A_2.$$

Exercice 1.11 *Montrer les identités suivantes :* @

$$^c(A_1 \cup A_2) = {}^cA_1 \cap {}^cA_2$$
$$^c(A_1 \cap A_2) = {}^cA_1 \cup {}^cA_2$$
$$A_1 \setminus A_2 = A_1 \cap {}^cA_2$$
$$A_1 \Delta A_2 = (A_1 \cup A_2) \setminus (A_1 \cap A_2) = (A_1 \setminus A_2) \cup (A_2 \setminus A_1).$$

1.3.3 Suites de parties d'un ensemble

Nous allons définir ici les notions de limite, limite supérieure et limite inférieure d'une suite de parties. Soit (A_n) une suite de parties de E.

Définition 1.12 *On rappelle que la suite (A_n) est dite* croissante *(resp.* décroissante*) lorsque pour tout entier n, $A_n \subseteq A_{n+1}$ (resp. $A_{n+1} \subseteq A_n$). Dans ce cas, la limite de la suite (A_n) est définie naturellement comme la* réunion *(resp. l'*intersection*) de tous les A_n :*

$$\lim_{n\to\infty} A_n := \bigcup_n A_n \text{ (resp. } \bigcap_n A_n\text{)}.$$

Par analogie avec le cas réel, on notera cette limite $\lim \uparrow$ *(resp.* $\lim \downarrow$*) pour faire référence au fait que la suite (A_n) est croissante et que la limite est donc la réunion (resp. l'intersection) de tous ses éléments.*

Définition 1.13 *On définit les deux parties de E suivantes :*

$$\limsup_{n\to\infty} A_n (ou \ \overline{\lim_{n\to\infty}} A_n) := \lim_{n\to\infty} \downarrow \bigcup_{k\geq n} A_k = \bigcap_n \bigcup_{k\geq n} A_k,$$

où la notation $\lim \downarrow$ *fait référence au fait que la suite* $\left(\bigcup_{k\geq n} A_k\right)_n$ *est décroissante, si bien que sa limite existe toujours (et est l'intersection de tous ses éléments, ce qu'indique la dernière égalité) ;*

$$\liminf_{n\to\infty} A_n (ou \ \underline{\lim_{n\to\infty}} A_n) := \lim_{n\to\infty} \uparrow \bigcap_{k\geq n} A_k = \bigcup_n \bigcap_{k\geq n} A_k,$$

où la notation $\lim \uparrow$ *fait référence au fait que la suite* $\left(\bigcap_{k\geq n} A_k\right)_n$ *est croissante, si bien que sa limite existe toujours (et est la réunion de tous ses éléments, ce qu'indique la dernière égalité).*

Remarque 1.14 *On peut aussi caractériser la limite supérieure et la limite inférieure par les assertions suivantes : pour tout $x \in E$,*

$$\begin{aligned} x \in \limsup_{n\to\infty} A_n &\Leftrightarrow \forall n \ \exists k \geq n, \ x \in A_k \\ &\Leftrightarrow \{n : x \in A_n\} \textit{ est infini.} \\ x \in \liminf_{n\to\infty} A_n &\Leftrightarrow \exists n \ \forall k \geq n, \ x \in A_k \\ &\Leftrightarrow \{n : x \notin A_n\} \textit{ est fini.} \end{aligned}$$

Noter que $\liminf_n A_n \subseteq \limsup_n A_n$.

Définition 1.15 *On dit que la suite* (A_n) *converge si* $\liminf_n A_n = \limsup_n A_n$. *Lorsque c'est le cas on définit* $\lim_n A_n := \liminf_n A_n = \limsup_n A_n$.

Remarque 1.16 *Soit A la limite d'une suite (A_n) qui converge. Alors A est caractérisée par :*

$$\begin{cases} \forall x \in A & \exists n_0 \quad \forall n \geq n_0 \quad x \in A_n \\ \forall x \notin A & \exists n_1 \quad \forall n \geq n_1 \quad x \notin A_n. \end{cases}$$

Exercice 1.17 *Montrer les deux égalités suivantes* @

$$\begin{aligned} \limsup_n {}^c A_n &= {}^c(\liminf_n A_n) \\ \liminf_n {}^c A_n &= {}^c(\limsup_n A_n). \end{aligned}$$

1.3.4 Fonctions et fonctions indicatrices

Définition 1.18 *On appelle* indicatrice *ou* fonction indicatrice *de la partie A, et l'on note $\mathbb{1}_A$, la fonction*

$$\begin{aligned} \mathbb{1}_A : E &\longrightarrow \{0,1\} \\ x &\longmapsto \begin{cases} 0 & si \quad x \notin A \\ 1 & si \quad x \in A. \end{cases} \end{aligned}$$

Remarque 1.19 *Noter que* $\mathbb{1}_{^cA} = 1 - \mathbb{1}_A$.

Proposition 1.20 *Au sens de la convergence simple,*

$$\overline{\lim_n} \mathbb{1}_{A_n} = \mathbb{1}_{\overline{\lim}_n A_n}$$

et

$$\underline{\lim_n} \mathbb{1}_{A_n} = \mathbb{1}_{\underline{\lim}_n A_n}$$

Dém. Pour tout $x \in E$,

$$\begin{aligned} \overline{\lim_n} \mathbb{1}_{A_n}(x) = 1 &\Leftrightarrow \forall n\ \exists k \geq n,\ \mathbb{1}_{A_k}(x) = 1 \\ &\Leftrightarrow \forall n\ \exists k \geq n,\ x \in A_k \\ &\Leftrightarrow x \in \overline{\lim_n} A_n \\ &\Leftrightarrow \mathbb{1}_{\overline{\lim}_n A_n}(x) = 1. \end{aligned}$$

L'autre assertion se démontre de la même manière, ou alors en se servant de l'assertion précédente :

$$\underline{\lim_n} \mathbb{1}_{A_n} = \underline{\lim_n}(1 - \mathbb{1}_{^cA_n}) = 1 - \overline{\lim_n} \mathbb{1}_{^cA_n} = 1 - \mathbb{1}_{\overline{\lim}_n {^cA_n}} = 1 - \mathbb{1}_{^c(\underline{\lim}_n A_n)} = \mathbb{1}_{\underline{\lim}_n A_n},$$

ce qui achève la démonstration. □

Remarque 1.21 *Conséquence de cette proposition : la suite de parties* (A_n) *converge ssi la suite de fonctions* $(\mathbb{1}_{A_n})$ *converge simplement (et lorsque c'est le cas, la convergence a lieu vers* $\mathbb{1}_{\lim_n A_n}$*).*

Définition 1.22 *Soient* E, F *deux ensembles et* $f : E \longrightarrow F$.

— *pour tout* $A \subseteq E$, *on note* $f(A)$ *l'*image directe *de* A *par* f *:*

$$f(A) := \{y \in F : \exists x \in A,\ f(x) = y\}.$$

— *pour tout* $B \subseteq F$, *on note* $f^{-1}(B)$ *l'*image réciproque *de* B *par* f *:*

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

Remarque 1.23 *La notation* f^{-1} *ne fera que très rarement, sinon jamais, référence à l'application inverse ou réciproque de l'application* f *dans les cas où elle serait par hasard bijective. Néanmoins, noter la cohérence de ces notations, au sens où si* f *est bijective, alors on a bien égalité entre l'image réciproque* $f^{-1}(B)$ *de* B *par* f *et l'image directe* $f^{-1}(B)$ *de* B *par l'inverse* f^{-1} *de* f.

Exercice 1.24 *Montrer les formules de Hausdorff(cf feuille de TD). Pour tous I et J ensembles d'indices non vides, pour toute famille $(A_i)_{i\in I}$ de parties de E et pour toute famille $(B_j)_{j\in J}$ de parties de F, pour toute fonction $f : E \longrightarrow F$,* @

$$f\left(\bigcup_i A_i\right) = \bigcup_i f(A_i),$$

$$f\left(\bigcap_i A_i\right) \subseteq \bigcap_i f(A_i)$$

avec égalité si f est injective ;

$$f^{-1}\left(\bigcup_j B_j\right) = \bigcup_j f^{-1}(B_j),$$

$$f^{-1}\left(\bigcap_j B_j\right) = \bigcap_j f^{-1}(B_j),$$

et pour tout $B \subseteq F$,

$${}^c\big(f^{-1}(B)\big) = f^{-1}\left({}^cB\right).$$

Chapitre 2

Théorie des cardinaux

2.1 Cardinaux, équipotence, dénombrabilité

Définition 2.1 *Deux ensembles E et F sont dits* équipotents*, ou* avoir même cardinal*, ou encore* même puissance*, s'il existe une* bijection *de l'un sur l'autre. On note alors* $\mathrm{Card}(E) = \mathrm{Card}(F)$*.*

Définition 2.2 *On notera* $\mathrm{Card}(E) \leq \mathrm{Card}(F)$ *s'il existe une* injection *de E dans F, c'est-à-dire si E a même puissance qu'une partie de F. Si de plus E et F n'ont pas même puissance, on notera* $\mathrm{Card}(E) < \mathrm{Card}(F)$*.*

Exemple 2.3 *Quelques exemples d'équipotences :*
- *Les ensembles $\mathscr{P}(E)$ et $\{0,1\}^E$ (= ensemble des applications : $E \longrightarrow \{0,1\}$) sont équipotents car l'application $A \mapsto \mathbb{1}_A$ est une bijection de l'un sur l'autre ;*
- *les ensembles $\mathbb{N}$ et $2\mathbb{N}$ (entiers pairs) sont équipotents car l'application $n \mapsto 2n$ est une bijection de l'un sur l'autre ;*
- *les ensembles $\mathbb{N}$ et $\mathbb{N} \times \mathbb{N}$ sont équipotents car on peut bien énumérer de manière injective les couples d'entiers (par exemple en suivant les points des droites d'équation $y = -x + c$, lorsque c croît dans $\mathbb{N}$) ;*
- *par récurrence, $\mathbb{N}$ est équipotent avec tous les produits cartésiens du type $\mathbb{N}^p$ ($p \in \mathbb{N}^\star$).* @

Théorème 2.4 (théorème de Cantor–Bernstein, admis) *Si* $\mathrm{Card}(E_1) \leq \mathrm{Card}(E_2)$ *et* $\mathrm{Card}(E_2) \leq \mathrm{Card}(E_1)$*, alors* $\mathrm{Card}(E_1) = \mathrm{Card}(E_2)$*.*

Remarque 2.5 *La relation $\leq$ est une* relation d'ordre*. En effet elle est*

1. *réflexive : il existe une injection de E dans E (l'injection canonique, c'est-à-dire ici l'identité), donc* $\mathrm{Card}(E) \leq \mathrm{Card}(E)$ *;*
2. *antisymétrique, grâce au théorème de Cantor–Bernstein ;*
3. *transitive : si* $\mathrm{Card}(E_1) \leq \mathrm{Card}(E_2)$ *et* $\mathrm{Card}(E_2) \leq \mathrm{Card}(E_3)$*, alors il existe une injection $f_1 : E_1 \longrightarrow E_2$ et une injection $f_2 : E_2 \longrightarrow E_3$, donc il existe une injection $f_3 : E_1 \longrightarrow E_3$ qui n'est autre que... $f_2 \circ f_1$, par conséquent* $\mathrm{Card}(E_1) \leq \mathrm{Card}(E_3)$*.* @

Remarque 2.6 *Ces énoncés ne sont pas des évidences, car il faut bien garder à l'esprit que les cardinaux ne sont pas des nombres réels (sauf pour le cas très particulier des ensembles finis).*

La proposition suivante, dont la démonstration est très jolie, assure en particulier qu'il existe une suite infinie strictement croissante de cardinaux :

$$\mathrm{Card}(E) < \mathrm{Card}(\mathscr{P}(E)) < \mathrm{Card}(\mathscr{P}(\mathscr{P}(E))) < \cdots$$

Proposition 2.7 $\mathrm{Card}(E) < \mathrm{Card}(\mathscr{P}(E))$.

Dém. Soit $f : E \to \mathscr{P}(E)$. Montrons que f ne peut être surjective (et donc ne peut être bijective). Soit

$$\Omega := \{x \in E : x \notin f(x)\}.$$

Montrons que par l'absurde que Ω ne peut avoir d'antécédent par f. Si $\exists z \in E$ tel que $f(z) = \Omega$ alors

— soit $z \in \Omega$ alors $z \notin f(z)$, c'est-à-dire $z \notin \Omega$;

— soit $z \notin \Omega$ alors $z \in f(z)$, c'est-à-dire $z \in \Omega$,

ce qui constitue une contradiction. D'autre part il existe clairement une injection de E dans $\mathscr{P}(E)$, par exemple celle qui à x associe $\{x\}$. □

Définition 2.8 *On définit les notions d'infini et de dénombrable comme suit :*

— *E est dit* infini *s'il existe $x_0 \in E$ et une injection de E dans $E \setminus \{x_0\}$, et est dit* fini *sinon ;*

— *E est dit* dénombrable *si* $\mathrm{Card}(E) \leq \mathrm{Card}(\mathbb{N})$ *;*

— *E est dit* infini dénombrable *si* $\mathrm{Card}(E) = \mathrm{Card}(\mathbb{N})$ *;*

— *E est dit (infini)* non dénombrable *si* $\mathrm{Card}(E) > \mathrm{Card}(\mathbb{N})$ *;*

— *une partie A de E est dite* cofinie *si ${}^cA$ est fini.*

Remarque 2.9 *L'ensemble $\mathbb{N}$ est (bien !) infini car par exemple l'application*

$$\begin{aligned} f : \mathbb{N} &\longrightarrow \mathbb{N}^\star \\ n &\longmapsto n+1 \end{aligned}$$

est bien une injection.

Définition 2.10 $\mathrm{Card}(\mathbb{N})$ *est souvent noté $\aleph_0$ (« aleph zéro »).*

La proposition suivante, laissée en exercice (indication : montrer par récurrence sur $n \in \mathbb{N}^\star$ qu'il existe n éléments distincts $x_1, \ldots, x_n$ de E et une injection $i_n : E \to E \setminus \{x_1, \ldots, x_n\}$), assure que les ensembles équipotents à $\mathbb{N}$ sont les plus petits ensembles infinis au sens des cardinaux.

Proposition 2.11 *E est infini ssi* $\mathrm{Card}(E) \geq \mathrm{Card}(\mathbb{N})$. @

2.2 Cardinaux classiques et propriétés

Proposition 2.12 *Les ensembles $\mathbb{Z}$, $\mathbb{N}^p$ $(p \in \mathbb{N}^\star)$ et $\mathbb{Q}$ sont dénombrables.*

Dém. On a déjà vu que $\mathbb{N}^p$ était équipotent à $\mathbb{N}$. Pour ce qui est de $\mathbb{Z}$, la fonction

$$\begin{aligned} f : \mathbb{Z} &\longrightarrow \mathbb{N} \\ n &\longmapsto \begin{cases} -2n & \text{si} \quad n \leq 0 \\ 2n-1 & \text{si} \quad n > 0 \end{cases} \end{aligned}$$

est une bijection.

Enfin, rappelons que pour tout $x \in \mathbb{Q}^\star$, $\exists!(p,q) \in \mathbb{Z}^\star \times \mathbb{N}^\star$ tel que $x = p/q$ et $p \wedge q = 1$. Ainsi la fonction qui à 0 associe $(0,1)$ et qui est définie sur $\mathbb{Q}^\star$ par

$$\begin{aligned} f : \ \mathbb{Q}^\star &\longrightarrow \mathbb{Z} \times \mathbb{N}^\star \\ p/q &\longmapsto (p,q) \end{aligned}$$

est une injection de $\mathbb{Q}$ dans $\mathbb{Z} \times \mathbb{N}^\star$, donc $\mathrm{Card}(\mathbb{Q}) \leq \mathrm{Card}(\mathbb{Z} \times \mathbb{N}^\star)$. Or il existe une injection $g : \mathbb{Z} \to \mathbb{N}$, donc l'application qui à (x,y) associe $(g(x),y)$ est une injection de $\mathbb{Z} \times \mathbb{N}^\star$ dans $\mathbb{N}^2$, ce qui montre que $\mathrm{Card}(\mathbb{Z} \times \mathbb{N}^\star) \leq \mathrm{Card}(\mathbb{N}^2) = \mathrm{Card}(\mathbb{N})$, donc $\mathrm{Card}(\mathbb{Q}) \leq \mathrm{Card}(\mathbb{N})$. □

Proposition 2.13 *Toute réunion dénombrable d'ensembles dénombrables est dénombrable.*

Dém. Soit $E = \bigcup_{n \in \mathbb{N}} E_n$, où pour tout $n \in \mathbb{N}$, E_n est dénombrable. Alors par définition, pour tout $n \in \mathbb{N}$ il existe une injection $\varphi_n : E_n \to \mathbb{N}$. Pour tout $x \in E$ on définit alors

$$N(x) := \min\{n \geq 0 : x \in E_n\} < \infty.$$

Alors la fonction

$$\begin{aligned} \phi : E &\longrightarrow \mathbb{N}^2 \\ x &\longmapsto (N(x), \varphi_{N(x)}(x)) \end{aligned}$$

est une injection car pour tous $x, y \in E$ tels que $\phi(x) = \phi(y)$, on a $N(x) = N(y) =: n$ puis $\varphi_{N(x)}(x) = \varphi_{N(y)}(y)$, c'est-à-dire $\varphi_n(x) = \varphi_n(y)$, donc $x = y$, puisque φ_n est injective. Par conséquent, $\mathrm{Card}(E) \leq \mathrm{Card}(\mathbb{N}^2) = \mathrm{Card}(\mathbb{N})$. □

Proposition 2.14 *Tout produit cartésien fini d'ensembles dénombrables est dénombrable.*

Dém. Pour $i = 1, \ldots, n$, soit E_i dénombrable et une injection $\varphi_i : E_i \to \mathbb{N}$. Alors la fonction

$$\begin{aligned} \phi : \qquad \Pi_{i=1}^n E_i &\longrightarrow \mathbb{N}^n \\ (x_1, \ldots, x_n) &\longmapsto (\varphi_1(x_1), \ldots, \varphi_n(x_n)) \end{aligned}$$

est clairement injective donc $\mathrm{Card}(\Pi_i E_i) \leq \mathrm{Card}(\mathbb{N}^n) = \mathrm{Card}(\mathbb{N})$. □

Proposition 2.15 *Tout produit cartésien infini dénombrable d'ensembles non vides (même finis) est non-dénombrable pourvu qu'une infinité d'entre eux ne soient pas réduits à un singleton.*

Dém. Admettons pour simplifier que pour tout $i \in \mathbb{N}$, $\mathrm{Card}(E_i) \geq 2$. Alors pour tout i, il existe une injection $\varphi_i : \{0,1\} \to E_i$. Donc l'application

$$\begin{aligned} \phi : \quad \{0,1\}^{\mathbb{N}} &\longrightarrow E_0 \times E_1 \times \cdots \\ (x_0, x_1, \ldots) &\longmapsto (\varphi_0(x_0), \varphi_1(x_1), \ldots) \end{aligned}$$

est injective, donc $\mathrm{Card}(\Pi_i E_i) \geq \mathrm{Card}(\{0,1\}^{\mathbb{N}}) = \mathrm{Card}\mathscr{P}(\mathbb{N}) > \mathrm{Card}(\mathbb{N})$. □

Théorème 2.16 *Les ensembles* $\mathbb{R}$ *et* $\mathscr{P}(\mathbb{N})$ *sont équipotents.*

Définition 2.17 *On dit d'un ensemble équipotent à* $\mathbb{R}$ *qu'il a la* puissance du continu.

Dém. Première étape : montrons que toute partie de $\mathbb{R}$ contenant un intervalle ouvert a la puissance du continu. Soit $A \subseteq \mathbb{R}$ contenant un intervalle I qu'on écrira sous la forme $I =]b-a, b+a[$, alors A s'injecte bien sûr dans $\mathbb{R}$, mais $\mathbb{R}$ s'injecte aussi dans A par exemple par l'application

$$\begin{aligned} \phi : \mathbb{R} &\longrightarrow A \\ x &\longmapsto a\frac{x}{\sqrt{x^2+1}} + b \end{aligned}$$

.

Deuxième étape : montrons que $\mathrm{Card}(\mathscr{P}(\mathbb{N})) \leq \mathrm{Card}([0,1/2])$ dont on sait d'après l'étape précédente que ce cardinal vaut $\mathrm{Card}(\mathbb{R})$. Soit l'application

$$\begin{aligned} \phi : \quad \{0,1\}^{\mathbb{N}} &\longrightarrow [0,1/2] \\ x = (x_n) &\longmapsto \sum_{n \geq 0} \frac{x_n}{3^{n+1}}. \end{aligned}$$

Montrons que ϕ est bien injective. Pour tous $x \neq y$, soit $n := \min\{k \geq 0 : x_k \neq y_k\} < \infty$. Alors

$$\begin{aligned} |\phi(x) - \phi(y)| &= \left| \frac{x_n - y_n}{3^{n+1}} + \sum_{k \geq n+1} \frac{x_k - y_k}{3^{k+1}} \right| \\ &\geq \frac{|x_n - y_n|}{3^{n+1}} - \left| \sum_{k \geq n+1} \frac{y_k - x_k}{3^{k+1}} \right| \\ &\geq \frac{1}{3^{n+1}} - \sum_{k \geq n+1} \frac{1}{3^{k+1}} \\ &= \frac{1}{3^{n+1}} - \frac{1}{3^{n+2}} \frac{1}{1 - 1/3} = \frac{1}{2 \cdot 3^{n+1}} > 0, \end{aligned}$$

ce qui prouve que $\phi(x) \neq \phi(y)$.

Troisième et dernière étape : montrons que $\mathrm{Card}(\{0,1\}^{\mathbb{N}}) \geq \mathrm{Card}([0,1[)$, ce qui équivaut à $\mathrm{Card}(\mathscr{P}(\mathbb{N})) \geq \mathrm{Card}(\mathbb{R})$. Soit $\psi : [0,1[\to \{0,1\}^{\mathbb{N}}$ l'application qui à $x \in [0,1[$ associe son développement dyadique propre, c'est-à-dire la suite (x_n) de 0 et de 1 définie récursivement par $x_0 := [2x]$, et

$$x_n := \left[2^{n+1} \left(x - \sum_{k=0}^{n-1} \frac{x_k}{2^{k+1}} \right) \right].$$

Alors comme $x = \sum_{k \geq 0} \frac{x_k}{2^{k+1}}$, ψ est clairement injective (car si $\psi(x) = \psi(y)$, $x = y$). $\square$

Chapitre 3

Tribus de parties d'un ensemble

3.1 Définitions et exemples

Définition 3.1 *Une classe* $\mathscr{A}$ *de parties d'un ensemble* E *est appelée* tribu *ou* σ-algèbre *si*

(i) elle contient E *:* $E \in \mathscr{A}$ *;*

(ii) elle est stable par passage au complémentaire *: pour tout* $A \subseteq E$, $A \in \mathscr{A} \Leftrightarrow {}^cA \in \mathscr{A}$ *;*

(iii) elle est stable par réunion dénombrable *: si* (A_n) *est une famille dénombrable d'éléments de* $\mathscr{A}$*, alors* $\cup_n A_n \in \mathscr{A}$.

On dit alors que $(E, \mathscr{A})$ *est un* espace mesurable.

Remarque 3.2 *Cette définition a quelques conséquences immédiates :*
- $\varnothing \in \mathscr{A}$ *car* $\varnothing = {}^cE$ *;*
- *stabilité par intersection dénombrable car* $\cap_n A_n = {}^c(\cup_n {}^cA_n)$ *;*
- *stabilité par différence car* $A \setminus B = A \cap {}^cB$ *;*
- *stabilité par différence symétrique car* $A \Delta B = (A \setminus B) \cup (B \setminus A)$ *;*
- *stabilité par limite supérieure car* $\overline{\lim}_n A_n = \cap_n \cup_{k \geq n} A_k$ *;*
- *stabilité par limite inférieure...*

Exercice 3.3 *Il aurait été équivalent de définir une tribu comme une classe* $\mathscr{A}$ *de parties de* E *vérifiant (par exemple) les propriétés suivantes :* $\mathscr{A}$ *contient* $\varnothing$*, est stable par passage au complémentaire et est stable par intersection dénombrable.*

Exemple 3.4 *Quelques exemples de tribus :*
- $\{\varnothing, E\}$ *est une tribu (parfois appelée la* tribu grossière*) ;*
- $\mathscr{P}(E)$ *est bien sûr une tribu (parfois appelée la* tribu triviale*) ;*
- *si* $(A_n)_{n \in \mathbb{N}}$ *est une partition de* E *dénombrable (finie ou infinie), alors*

$$\mathscr{A} := \{\cup_{i \in I} A_i : I \subseteq \mathbb{N}\}$$

  est une tribu sur E *;*
- *si* $A \subseteq E$*, la plus petite (voir section suivante) tribu contenant* A *est* $\{\varnothing, E, A, {}^cA\}$ *;*

— *enfin,*
$$\mathscr{A} := \{A \subseteq E : A \text{ ou } {}^cA \text{ est dénombrable}\}$$
est une tribu, ce que nous démontrons ci-dessous.

Dém. *Nous démontrerons uniquement la stabilité par réunion dénombrable. Soient* $(A_n)_n \in \mathscr{A}$*. De deux choses l'une :*

— *soit pour tout* n*,* A_n *est dénombrable et alors* $\cup_n A_n$ *est dénombrable ;*
— *soit* $\exists n_0$ *tel que* A_{n_0} *est non dénombrable, et alors* ${}^cA_{n_0}$ *est dénombrable, donc* $\cap_n {}^cA_n \subseteq {}^cA_{n_0}$ *est dénombrable, et par conséquent* $\cup_n A_n$ *est de complémentaire dénombrable (car égal à* $\cap_n {}^cA_n$*) ;*

Dans les deux cas $\cup_n A_n \in \mathscr{A}$*.* □

3.2 Tribu engendrée

Proposition 3.5 (et définition) *a) l'intersection d'une collection non vide quelconque*[1] *de tribus de parties de* E *est elle-même une tribu ;* @

b) pour toute classe $\mathscr{C}$ *de parties de* E*, l'intersection de toutes les tribus contenant*[2] $\mathscr{C}$ *est (donc*[3]*) une tribu : elle est (appelée)* la plus petite tribu *contenant* $\mathscr{C}$*, ou tribu* engendrée *par* $\mathscr{C}$*, et notée* $\sigma(\mathscr{C})$ *:*
$$\sigma(\mathscr{C}) := \bigcap_{\mathscr{A} \text{ tribu}, \mathscr{C} \subseteq \mathscr{A}} \mathscr{A}.$$

Remarque 3.6 — *On rappelle que le terme collection désigne un ensemble de famille de parties, c'est-à-dire un ensemble d'ensembles de sous-ensembles de* E*... ;*
— *je ne devrais pas préciser, mais il faut garder à l'esprit que si* $\mathscr{A}$ *et* $\mathscr{B}$ *sont des familles de parties de* E *alors* $C \in \mathscr{A} \cap \mathscr{B}$ *ssi* $C \in \mathscr{A}$ *et* $C \in \mathscr{B}$ *(on n'intersecte pas ici les parties de* E*) ;*
— *le terme de plus petite tribu n'a de sens qu'à la lumière de la définition précédente, car il n'existe pas d'ordre total sur les tribus.*

Remarque 3.7 — *pour toute classe* $\mathscr{B}$ *de parties de* E*,* $\mathscr{B} \subseteq \sigma(\mathscr{B})$*, par définition ;*
— *si* $\mathscr{C}$ *est une classe de parties de* E *et* $\mathscr{A}$ *est une tribu de parties de* E *telle que* $\mathscr{C} \subseteq \mathscr{A}$*, alors* $\mathscr{A}$ *est élément de la collection des tribus contenant* $\mathscr{C}$*, donc contient son intersection* $\sigma(\mathscr{C})$*, autrement dit* $\sigma(\mathscr{C}) \subseteq \mathscr{A}$ *;*
— *première conséquence, si* $\mathscr{A}$ *est une tribu de parties de* E*, alors* $\sigma(\mathscr{A}) = \mathscr{A}$ *;*
— *deuxième conséquence, si* $\mathscr{C} \subseteq \mathscr{B}$ *alors* $\mathscr{B} \subseteq \sigma(\mathscr{B})$ *implique* $\mathscr{C} \subseteq \sigma(\mathscr{B})$*, et comme* $\mathscr{B}$ *est une tribu,* $\sigma(\mathscr{C}) \subseteq \sigma(\mathscr{B})$*.*

Remarque 3.8 (méthodologie) — *Si* $\mathscr{A}$ *est une tribu, pour montrer que* $\mathscr{A} = \sigma(\mathscr{C})$*, on montre que* $\mathscr{A} \subseteq \sigma(\mathscr{C})$ *et que* $\mathscr{C} \subseteq \mathscr{A}$ *;*

1. quelconque au sens de « pas forcément dénombrable »
2. au sens de l'inclusion
3. cette collection est non vide car un de ses éléments est $\mathscr{P}(E)$

— *pour montrer que* $\sigma(\mathscr{C}_1) = \sigma(\mathscr{C}_2)$, *on montre que* $\mathscr{C}_1 \subseteq \sigma(\mathscr{C}_2)$ *et que* $\mathscr{C}_2 \subseteq \sigma(\mathscr{C}_1)$.

Définition 3.9 *On note* $\mathscr{B}(\mathbb{R})$, *ou* $\mathscr{B}or(\mathbb{R})$, *et on appelle* tribu de Borel *sur* $\mathbb{R}$ *la tribu engendrée par les intervalles ouverts. La tribu de Borel sur* $\bar{\mathbb{R}}$ *est l'ensemble des parties de* $\mathbb{R}$ *prenant l'une des formes* A, $A \cup \{+\infty\}$, $A \cup \{-\infty\}$ *ou* $A \cup \{-\infty, +\infty\}$, *où* $A \in \mathscr{B}or(\mathbb{R})$.

Proposition 3.10 *Soit* S *une partie dense de la droite réelle*[4]. *Alors* $\mathscr{B}or(\mathbb{R})$ *est la tribu engendréee par les intervalles du type*

a) $[a, +\infty[, a \in S$; *b)* $]b, +\infty[, b \in S$; *c)* $]-\infty, c[, c \in S$; *d)* $]-\infty, d], d \in S$.

Il en est de même pour $\mathscr{B}or(\bar{\mathbb{R}})$ *avec les intervalles du type* $[a, +\infty]$,...

Dém. [de a)] Soit $\mathscr{I}_S$ l'ensemble des intervalles de la forme $[a, +\infty[$ pour $a \in S$. Tout d'abord, $\mathscr{B}(\mathbb{R})$ contient tous les intervalles fermés de $\mathbb{R}$ car est stable par passage au complémentaire donc on a l'inclusion $\sigma(\mathscr{I}_S) \subseteq \mathscr{B}(\mathbb{R})$. Soit maintenant $a \in [-\infty, +\infty[$. Comme S est dense, il existe une suite décroissante (a_n) d'éléments de S tels que $a_n \neq a$ @
pour tout n, et $\lim_n \downarrow a_n = a$. Comme $[a_n, +\infty[\in \mathscr{I}_S$, on a $[a_n, +\infty[\in \sigma(\mathscr{I}_S)$, donc par stabilité par réunion dénombrable de la tribu $\sigma(\mathscr{I}_S)$,

$$]a, +\infty[= \cup_n [a_n, +\infty[\in \sigma(\mathscr{I}_S).$$

On démontre avec une suite croissante que $[a, +\infty[\in \sigma(\mathscr{I}_S)$. Maintenant pour tous $a, b \in [-\infty, +\infty[$, l'intervalle $]a, b[$ s'écrit $]a, +\infty[\setminus [b, +\infty[\in \sigma(\mathscr{I}_S)$. Par conséquent $\mathscr{I} \subseteq \sigma(\mathscr{I}_S)$, où $\mathscr{I}$ est l'ensemble des intervalles ouverts de $\mathbb{R}$ et $\mathscr{B}(\mathbb{R}) = \sigma(\mathscr{I}) \subseteq \sigma(\mathscr{I}_S)$. □

3.3 Tribus image et image réciproque

Soit $f : E_1 \longrightarrow E_2$.

Proposition 3.11 *Si* $\mathscr{A}_2$ *est une tribu sur* E_2,

$$f^{-1}(\mathscr{A}_2) := \{f^{-1}(Y), Y \in \mathscr{A}_2\}$$

est une tribu sur E_1, *appelée* tribu image réciproque *(de* $\mathscr{A}_2$ *par* f*).*

Dém. Par les formules de Hausdorff :
i) $f^{-1}(E_2) = E_1 \in f^{-1}(\mathscr{A}_2)$;
ii) pour tout $Y \in \mathscr{A}_2$, $\complement(f^{-1}(Y)) = f^{-1}(\complement Y) \in f^{-1}(\mathscr{A}_2)$;
iii) pour toute suite $(Y_n) \in \mathscr{A}_2$, $\cup_n f^{-1}(Y_n) = f^{-1}(\cup_n Y_n) \in f^{-1}(\mathscr{A}_2)$ car $\cup_n Y_n \in \mathscr{A}_2$.□

Proposition 3.12 *Si* $\mathscr{A}_1$ *est une tribu sur* E_1,

$$\mathscr{B} = \{Y \subseteq E_2 : f^{-1}(Y) \in \mathscr{A}_1\}$$

est une tribu sur E_2, *appelée* tribu image *(de* $\mathscr{A}_1$ *par* f*).*

Remarque 3.13 *La tribu image n'est PAS* $f(\mathscr{A}_1)$ *qui n'est en général pas une tribu.*

4. c'est-à-dire telle que tout nombre réel est limite d'une suite à valeurs dans S ; par exemple $S = \mathbb{Q}$

Dém. Par les formules de Hausdorff également. □ @

Définition 3.14 (et proposition) *Soit* $(E, \mathscr{A})$ *un ensemble mesurable et* X *une partie de* E*. La classe* $\mathscr{C} = \{A \cap X : A \in \mathscr{A}\}$ *de parties de* X *est une tribu sur* X *appelée* tribu trace *de* $\mathscr{A}$ *sur* X*.*

Remarque 3.15 *Cette définition a surtout de l'intérêt dans le cas où* $X \notin \mathscr{A}$*.*

Dém. La classe $\mathscr{C}$ est la tribu image réciproque de $\mathscr{A}$ par l'injection canonique $i : X \to E$: en effet pour tout $A \in \mathscr{A}$, $i^{-1}(A) = A \cap X$. □

Théorème 3.16 (lemme de transport) *Soit* $f : E_1 \longrightarrow E_2$ *et* $\mathscr{C}$ *une classe de parties de* E_2*. Alors* $\sigma(f^{-1}(\mathscr{C})) = f^{-1}(\sigma(\mathscr{C}))$*.*

Dém. Montrons l'inclusion $\subseteq$. Tout d'abord $\mathscr{C} \subseteq \sigma(\mathscr{C})$, donc $f^{-1}(\mathscr{C}) \subseteq f^{-1}(\sigma(\mathscr{C}))$. Mais $f^{-1}(\sigma(\mathscr{C}))$ est une tribu donc $\sigma(f^{-1}(\mathscr{C})) \subseteq \sigma(f^{-1}(\sigma(\mathscr{C}))) = f^{-1}(\sigma(\mathscr{C}))$.

Montrons maintenant l'inclusion $\supseteq$. Soit $\mathscr{B}$ la tribu image de $\sigma(f^{-1}(\mathscr{C}))$ par f, c'est-à-dire

$$\mathscr{B} := \{Y \subseteq E_2 : f^{-1}(Y) \in \sigma(f^{-1}(\mathscr{C}))\}.$$

Alors $\mathscr{C} \subseteq \mathscr{B}$, et comme $\mathscr{B}$ est une tribu, $\sigma(\mathscr{C}) \subseteq \mathscr{B}$, puis $f^{-1}(\sigma(\mathscr{C})) \subseteq f^{-1}(\mathscr{B})$. Mais par définition de $\mathscr{B}$, $f^{-1}(\mathscr{B}) \subseteq \sigma(f^{-1}(\mathscr{C}))$ donc $f^{-1}(\sigma(\mathscr{C})) \subseteq \sigma(f^{-1}(\mathscr{C}))$. □

Chapitre 4

Fonctions mesurables

4.1 Définitions

Notation 4.1 *Soit* $f : E_1 \longrightarrow E_2$ *et* $B \subseteq E_2$. *On utilise très fréquemment la notation* $\{f \in B\}$ *à la place de* $f^{-1}(B)$, *ce qui peut se voir comme une écriture condensée de* $\{x : f(x) \in B\}$. *Par exemple, dans le cas où* $E_2 = \mathbb{R}$ *et* $B = [a, +\infty[$, *on écrira* $f^{-1}(B)$ *sous la forme* $\{f \geq a\}$.

Définition 4.2 *Une fonction* $f : (E_1, \mathscr{A}_1) \longrightarrow (E_2, \mathscr{A}_2)$ *est dite mesurable*[1] *si* $f^{-1}(\mathscr{A}_2) \subseteq \mathscr{A}_1$ *(c'est-à-dire : pour tout* $B \in \mathscr{A}_2$, $f^{-1}(B) \in \mathscr{A}_1$*).*

Notation 4.3 *On notera* $\mathscr{F}(\mathscr{A}_1, \mathscr{A}_2)$ *l'ensemble des fonctions mesurables :* $(E_1, \mathscr{A}_1) \to (E_2, \mathscr{A}_2)$.

Remarque 4.4 *Si on ne se donne que la tribu* $\mathscr{A}_1$, *alors la tribu image de* $\mathscr{A}_1$ *par* f *est la plus grande tribu sur* E_2 *qui rende* f *mesurable.*

Si on ne se donne que $\mathscr{A}_2$, *alors la tribu image réciproque de* $\mathscr{A}_2$ *par* f *est la plus petite tribu sur* E_1 *qui rende* f *mesurable. On note aussi cette tribu* $\sigma(f)$.

Remarque 4.5 *Une fonction indicatrice* $\mathbb{1}_A : (E, \mathscr{A}) \longrightarrow (\{0,1\}, \mathscr{P}(\{0,1\}))$ *est mesurable ssi* $A \in \mathscr{A}$, *ce que l'on dira aussi « A est mesurable »*[2].

4.2 Exemples et opérations stables pour la mesurabilité

La proposition suivante est une conséquence du lemme de transport.

Proposition 4.6 *Soit* $\mathscr{C}$ *une classe de parties de* F *et* $\mathscr{B} := \sigma(\mathscr{C})$. *Alors* $f : (E, \mathscr{A}) \to (F, \mathscr{B})$ *est mesurable ssi* $f^{-1}(\mathscr{C}) \subseteq \mathscr{A}$.

Dém. L'application f est mesurable ssi $f^{-1}(\mathscr{B}) \subseteq \mathscr{A}$, mais $f^{-1}(\mathscr{B}) = f^{-1}(\sigma(\mathscr{C})) = \sigma(f^{-1}(\mathscr{C}))$ et $\sigma(f^{-1}(\mathscr{C})) \subseteq \mathscr{A}$ ssi $f^{-1}(\mathscr{C}) \subseteq \mathscr{A}$. □

1. sous-entendu par rapport aux deux tribus $\mathscr{A}_1$ et $\mathscr{A}_2$
2. toujours en référence sous-entendue à la tribu $\mathscr{A}$

Application. Soit S une partie dense de $\mathbb{R}$. Alors la fonction $f : (E, \mathscr{A}) \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ est mesurable ssi $\{f \geq a\} \in \mathscr{A}$ pour tout $a \in S$ (et l'on peut bien sûr remplacer $\{f \geq a\}$ par $\{f > a\}$, $\{f \leq a\}$ ou $\{f < a\}$).

Proposition 4.7 *Soient $f_1 : (E_1, \mathscr{A}_1) \to (E_2, \mathscr{A}_2)$ et $f_2 : (E_2, \mathscr{A}_2) \to (E_3, \mathscr{A}_3)$. Si f_1 et f_2 sont mesurables, alors $f_2 \circ f_1 : (E_1, \mathscr{A}_1) \to (E_3, \mathscr{A}_3)$ est aussi mesurable.*

Dém. Pour tout élément A_3 de $\mathscr{A}_3$, on vérifie que $(f_2 \circ f_1)^{-1}(A_3) = f_1^{-1}(f_2^{-1}(A_3))$. @
Comme f_2 est mesurable, $f_2^{-1}(A_3) \in \mathscr{A}_2$. De plus, comme f_1 est mesurable $f_1^{-1}(f_2^{-1}(A_3)) \in \mathscr{A}_1$. □

Proposition 4.8 *Soit une suite (f_n) de $\mathscr{F}(\mathscr{A}, \mathscr{B}or(\bar{\mathbb{R}}))$. Alors*

a) $\sup_n f_n$ *et* $\inf_n f_n$ *sont mesurables ;*

b) $\limsup_n f_n$ *et* $\liminf_n f_n$ *sont mesurables ;*

c) Si (f_n) converge simplement *vers une fonction f (dans $\bar{\mathbb{R}}$), alors f est mesurable.*

Dém. a) Pour tout $a \in \mathbb{R}$, $\{\sup_n f_n \leq a\} = \cap_n \{f_n \leq a\} \in \mathscr{A}$ et $\{\inf_n f_n \geq a\} = \cap_n \{f_n \geq a\} \in \mathscr{A}$.

b) D'après a), pour tout $n \in \mathbb{N}$, la fonction $\sup_{k \geq n} f_k$ est mesurable, donc la fonction $\limsup_n f_n = \inf_n \sup_{k \geq n} f_k$ est mesurable. De même pour $\liminf_n f_n$.

c) Si $f_n \to f$, alors $f = \limsup_n f_n$, qui est mesurable d'après b). □

On peut raffiner le résultat sur la mesurabilité de la limite d'une suite de fonctions mesurables de la manière suivante[3].

Théorème 4.9 *Soit $C := \{x \in E :$ la suite $(f_n(x))_n$ converge dans $\bar{\mathbb{R}}\}$. Alors ($C \in \mathscr{A}$ et) si $\mathscr{C}$ désigne la tribu trace de $\mathscr{A}$ sur C alors la fonction $f := \lim_n f_n : (C, \mathscr{C}) \to (\bar{\mathbb{R}}, \mathscr{B}(\bar{\mathbb{R}}))$ est mesurable.*

Dém. On note $f^{\downarrow} := \liminf_n f_n$ et $f^{\uparrow} := \limsup_n f_n$. Alors C est mesurable car

$$C = {}^c\{f^{\uparrow} \neq f^{\downarrow}\} = {}^c(\cup_{r \in \mathbb{Q}} \{f^{\uparrow} > r\} \cap \{f^{\downarrow} < r\}).$$

Rappelons que la mesurabilité de C n'est pas nécessaire pour définir la tribu trace $\mathscr{C}$. Néanmoins, pour tout borélien B de $\bar{\mathbb{R}}$, $f^{-1}(B) = C \cap (f^{\uparrow})^{-1}(B) \in \mathscr{C}$. En effet, $f^{-1}(B) = \{x \in E : f^{\downarrow}(x) = f^{\uparrow}(x)$ et $f^{\uparrow}(x) \in B\}$. □

4.3 Fonctions étagées, en escalier, réglées

Définition 4.10 *Une fonction $f \in \mathscr{F}(\mathscr{A}, \mathscr{B}or(\mathbb{R}))$ est dite* étagée *si elle ne prend qu'un nombre fini de valeurs. Alors il existe une partition* finie *$(A_i, i \in I)$ de E, $\mathscr{A}$-mesurable[4], et des nombres réels $(\alpha_i, i \in I)$ tels que $f = \sum_{i \in I} \alpha_i \mathbb{1}_{A_i}$.*

3. Ce résultat est hors de la portée stricte du cours, mais la question à laquelle il répond est tellement naturelle...

4. au sens où $A_i \in \mathscr{A}$ pour tout $i \in I$

Notation 4.11 *On note $\mathscr{E}(\mathscr{A})$ l'ensemble des fonctions étagées :* $(E, \mathscr{A}) \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$.

Remarque 4.12 *Il existe une représentation canonique de f sous la forme $\sum_{i \in I} \alpha_i \mathbb{1}_{A_i}$ où les α_i sont deux à deux distincts et $A_i = \{f = \alpha_i\}$. On notera qu'une fonction indicatrice est bien sûr étagée car $\mathbb{1}_A = 1 \cdot \mathbb{1}_A + 0 \cdot \mathbb{1}_{^cA}$.*

Proposition 4.13 *Pour toutes fonctions étagées f, g et pour tout $\lambda \in \mathbb{R}$, $\lambda f + g$ est étagée (autrement dit $\mathscr{E}(\mathscr{A})$ est un espace vectoriel), ainsi que fg, $f \wedge g$ et $f \vee g$.*[5]

Dém. On écrit f et g sous la forme $f = \sum_{i \in I} \alpha_i \mathbb{1}_{A_i}$ et $g = \sum_{j \in J} \beta_j \mathbb{1}_{B_j}$. Alors $(A_i \cap B_j; (i,j) \in I \times J)$ est une partition finie de E et on peut écrire $\lambda f + g = \sum_{(i,j) \in I \times J} (\lambda \alpha_i + \beta_j) \mathbb{1}_{A_i \cap B_j}$, $fg = \alpha_i \beta_j \mathbb{1}_{A_i \cap B_j}$, etc. □

Théorème 4.14 (lemme fondamental d'approximation) *Pour toute $f \in \mathscr{F}(\mathscr{A}, \mathscr{B}(\bar{\mathbb{R}}))$, il existe une suite (f_n) de fonctions étagées convergeant simplement vers f*[6]. *De plus,*

a) si f est positive, *on peut choisir la suite (f_n) positive et croissante*[7] *;*

b) si f est bornée, *on peut choisir (f_n) de sorte que la convergence soit* uniforme[8].

Dém. Commençons par le cas où f est positive. On définit alors

$$f_n := \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \mathbb{1}_{\{(k-1)2^{-n} < f \leq k2^{-n}\}} + n \mathbb{1}_{\{f > n\}}.$$

Alors pour tout $x \in E$, la suite $(f_n(x))_n$ est bien (positive et) croissante et converge vers @
$f(x)$, en effet : si $f(x) = +\infty$, alors $f_n(x) = n \to \infty$; sinon il existe n_0 tel que $f(x) < n_0$, ce qui implique que pour tout $n \geq n_0$, $|f_n(x) - f(x)| \leq 2^{-n} \to 0$.

Si f est bornée et positive, alors il existe n_0 tel que pour tout $x \in E$, $f(x) < n_0$, donc pour tout $x \in E$, pour tout $n \geq n_0$, $|f_n(x) - f(x)| \leq 2^{-n} \to 0$, ce qui n'est autre qu'une convergence uniforme.

Si f est de signe quelconque, on écrit f sous la forme $f = f^+ - f^-$, où

$$f^+ := f \mathbb{1}_{\{f>0\}} \qquad \text{et} \qquad f^- := -f \mathbb{1}_{\{f<0\}}.$$

La somme $f^+ - f^-$ n'est jamais indéterminée, car pour tout $x \in E$, au moins un des deux termes $f^+(x)$ ou $f^-(x)$ est nul. On notera également que f^+ (et f^-, par un même raisonnement) est mesurable car pour tout $a \geq 0$, $\{f^+ \geq a\} = \{f \geq a\}$ et pour tout $a < 0$, $\{f^+ \geq a\} = E$. À présent, comme f^+ et f^- sont positives, il existe deux suites croissantes (u_n) et (v_n) de fonctions étagées positives convergeant resp. vers f^+ et f^-. De plus, si l'on utilise la construction de ces suites proposée plus haut, on a $u_n v_n = 0$, de sorte que l'on peut toujours définir $f_n := u_n - v_n$, qui définit une suite de fonctions étagées convergeant vers $f^+ - f^- = f$.

5. $a \wedge b$ est une notation alternative pour $\min(a, b)$, et $a \vee b$ pour $\max(a, b)$
6. autrement dit (rappel...) : $\forall x \in E$, $f_n(x) \to f(x)$ lorsque $n \to \infty$
7. autrement dit : $\forall x \in E$, $\forall n \in \mathbb{N}$, $0 \leq f_n(x) \leq f_{n+1}(x)$ – rien à voir avec des fonctions croissantes, ce qui n'aurait d'ailleurs pas de sens ici...
8. autrement dit (rappel...) : $\sup_{x \in E} |f_n(x) - f(x)| \to 0$ lorsque $n \to \infty$

Si f est de signe quelconque mais bornée, f^+ et f^- sont bornées, donc on peut choisir les suites (u_n) et (v_n) pour que les convergences vers f^+ et f^- soient toutes deux uniformes. Alors la suite $(u_n - v_n)$ converge uniformément vers f. $\square$

Définition 4.15 *Une fonction $f : [a,b] \longrightarrow \mathbb{R}$ est dite en escalier s'il existe une* subdivision finie $a = a_0 < a_1 < \cdots < a_n = b$ *de l'intervalle $[a,b]$ telle que f soit constante sur chaque intervalle $]a_i, a_{i+1}[$.*

Remarque 4.16 *Les valeurs prises exactement en chaque point $a_0, a_1, \ldots, a_n$ sont sans importance.*

Remarque 4.17 *Une fonction en escalier a toujours pour espace de départ un intervalle compact de $\mathbb{R}$, ce qui en fait un objet beaucoup moins général qu'une fonction étagée. D'ailleurs, une fonction en escalier est toujours un cas particulier de fonction étagée, au sens où elle est un élément de $\mathscr{E}(\mathscr{B}or([a,b]))$, car elle ne prend qu'un nombre fini de valeurs et elle est mesurable, en effet : les parties de $[a,b]$ sur lesquelles f est constante sont des intervalles (les singletons sont bien sûr des intervalles) ou des réunions d'intervalles, donc des boréliens, donc l'image réciproque de toute partie de $\mathbb{R}$ est toujours un borélien de $[a,b]$.*

Le contre-exemple classique de la réciproque est $\mathbb{1}_{\mathbb{Q}}$, qui est étagée mais n'est en escalier sur aucun intervalle de $\mathbb{R}$ (non réduit à un point).

Remarque 4.18 *L'intégrale de Riemann est définie par approximation à partir de l'intégrale des fonctions en escalier, tandis que celle que nous étudions dans ce cours (parfois dite de Lebesgue) est construite à partir des fonctions étagées. Dans le premier cas, on approche l'intégrale d'une fonction quelconque par celle d'une fonction en escalier, c'est-à-dire en découpant l'espace de départ (un intervalle) en petits morceaux (les subdivisions), tandis que dans le second cas, c'est l'espace d'arrivée (qui est toujours $\mathbb{R}$ ou $\bar{\mathbb{R}}$) qui est découpé. Cette différence est fondamentale car la première approche ne peut se généraliser facilement à des fonctions ayant un autre espace de départ que $\mathbb{R}$. Mais surtout les espaces de fonctions mesurables (celles qui admettront une intégrale au sens de Lebesgue) sont beaucoup plus grands que celui des fonctions Riemann-intégrables et ils sont stables sous l'action de multiples opérations comme le passage à la limite. Enfin, nous allons définir dans ce cours l'intégrale par rapport à une mesure quelconque, et pas seulement l'intégrale par rapport à la mesure de Lebesgue (celle qui a ceci de commun avec l'intégrale de Riemann qu'elle donne un sens mathématique à la notion physique de volume).*

Définition 4.19 *Une fonction $f : [a,b] \longrightarrow \mathbb{R}$ est dite* réglée *si elle est limite* uniforme *de fonctions en escalier.*

Remarque 4.20 *Toute fonction réglée est mesurable (on dira ici borélienne car les tribus de départ et d'arrivée sont des tribus de Borel) car limite de fonctions mesurables (et même étagées) que sont les fonctions en escalier.*

Théorème 4.21 (admis) *Une fonction f est réglée ssi elle admet une limite à gauche en tout point de $]a,b]$ et une limite à droite en tout point de $[a,b[$.*

Corollaire 4.22 *Toute fonction* $f : \mathbb{R} \longrightarrow \mathbb{R}$ *monotone est réglée.*

Remarque 4.23 *Toute fonction monotone est borélienne, car réglée. Mais cela peut se voir directement : toute fonction monotone est borélienne car pour tout* $a \in \mathbb{R}$, $\{f \geq a\}$ *est une demie-droite, en effet : si* $m(a) := \inf\{x : f(x) \geq a\}$, *alors dans le cas où* f *est croissante par exemple,* $\{f \geq a\}$ *coïncide soit avec* $[m(a), +\infty[$, *soit avec* $]m(a), +\infty[$.

Chapitre 5

Le cas borélien

5.1 (Rappels de) Topologie

Définition 5.1 *Une famille $\mathscr{O}(E)$ de parties d'un ensemble E est appelée* topologie*, et ses éléments des* ouverts*, si*

i) elle contient $\varnothing$ et E : $\varnothing \in \mathscr{O}(E)$ et $E \in \mathscr{O}(E)$;

ii) elle est stable par intersections finies : $\forall U, V \in \mathscr{O}(E)$, $U \cap V \in \mathscr{O}(E)$;

iii) elle est stable par réunion quelconque [1] : pour tout I ensemble d'indices et pour toute famille d'ouverts $(O_i, i \in I)$, $\cup_{i \in I} O_i$ est un ouvert.

Les complémentaires des ouverts sont appelés des fermés.

Remarque 5.2 *Les ouverts $\varnothing$ et E sont aussi des fermés ; les fermés sont stables par réunions finies et par intersections quelconques.*

Définition 5.3 *On appelle* voisinage *de $x \in E$ toute partie $\mathcal{V}$ de E telle qu'il existe un ouvert O pour lequel $x \in O \subseteq \mathcal{V}$. Tout ouvert est donc voisinage de chacun de ses points.*

Définition 5.4 *Dans un espace* métrique (E, d)*, la topologie dite relative à la distance d est constituée des réunions quelconques de parties du type*

$$B(x, r) := \{y \in E : d(x, y) < r\}$$

appelée boule ouverte *de centre x et de rayon r.*

Remarque 5.5 *Une partie O de l'espace métrique (E, d) est ouverte ssi $\forall x \in O$, $\exists r > 0$, $B(x, r) \subseteq O$ (un ouvert O d'un espace métrique est la réunion des boules ouvertes contenues dans O).*

Une partie A de l'espace métrique (E, d) est fermée ssi pour toute suite (x_n) à valeurs dans A et convergeant vers une limite x, $x \in A$.

Remarque 5.6 *La topologie de $\mathbb{R}$ relative à la distance usuelle est donc constituée des réunions quelconques d'intervalles ouverts.*

---

1. au sens où l'on ne fait pas d'hypothèse sur le cardinal de I

Définition 5.7 (et proposition) *Le plus grand ouvert contenu dans $A \subseteq E$, c'est-à-dire la réunion de tous les ouverts contenus dans A, est noté $\mathring{A}$ et appelé* intérieur *de A, ou ensemble des points* intérieurs *à A. En particulier, A est ouvert ssi $A = \mathring{A}$.*

Dans le cas métrique, un point $x \in E$ est intérieur à A ssi $\exists \varepsilon > 0$ tel que $B(x, \varepsilon) \subseteq A$.

Définition 5.8 (et proposition) *Le plus petit fermé contenant $A \subseteq E$, c'est-à-dire l'intersection de tous les fermés contenant A, est noté $\bar{A}$ et appelé* adhérence *de A, ou ensemble des points* adhérents *à A. En particulier, A est fermé ssi $A = \bar{A}$.*

Dans le cas métrique, un point $x \in E$ est adhérent *à A ssi il existe une suite (x_n) à valeurs dans A telle que $\lim_n x_n = x$.*

Remarque 5.9 *Pour tout $A \subseteq E$, l'intérieur de ${}^cA$ est le complémentaire de $\bar{A}$ et l'adhérence de ${}^cA$ est le complémentaire de $\mathring{A}$.*

Définition 5.10 *La* frontière *de A est le fermé $\partial A := \bar{A} \setminus \mathring{A}$.*

Définition 5.11 *Soient E et F deux espaces topologiques. Une fonction $f : E \longrightarrow F$ est dite* continue *si l'image réciproque par f de tout ouvert est un ouvert (ce qui est équivalent à dire que l'image réciproque par f de tout fermé est un fermé).*

Proposition 5.12 *Soient E et F deux espaces métriques. Une fonction $f : E \longrightarrow F$ est dite* continue *ssi pour toute suite (x_n) de E convergeant vers x, la suite $(f(x_n))$ est aussi convergente et $\lim_n f(x_n) = f(x)$.*

Définition 5.13 *Soit $X \subseteq E$. La* topologie trace[2] *de $\mathscr{O}(E)$ sur X est constituée des intersections des ouverts de E avec X. Dans le cas métrique, la topologie trace est la topologie relative à la restriction de la distance à $X \times X$.*

Définition 5.14 *La* topologie produit *de $E \times F$ est constituée des réunions quelconques de pavés à côtés ouverts :*

$$\mathscr{O}(E \times F) := \{\cup_{i \in I} U_i \times V_i, U_i \in \mathscr{O}(E), V_i \in \mathscr{O}(F), I \text{ ensemble d'indices quelconque}\}.$$

Proposition 5.15 *La topologie produit est aussi la plus petite topologie qui rendent les projections canoniques π_E et π_F continues :*

$$\begin{aligned} \pi_E : E \times F &\longrightarrow E \\ (x, y) &\longmapsto x \end{aligned}$$

et

$$\begin{aligned} \pi_F : E \times F &\longrightarrow F \\ (x, y) &\longmapsto y \end{aligned}$$

2. dite aussi topologie induite

Dans le cas métrique, la topologie produit est la topologie relative à toute distance classique du type

$$\begin{aligned} d((x,y),(x',y')) \quad &:= \quad d_E(x,x') + d_F(y,y') \\ &\text{OU} \quad \sqrt{d_E(x,x')^2 + d_F(y,y')^2} \\ &\text{OU} \quad d_E(x,x') \vee d_F(y,y'). \end{aligned}$$

Définition 5.16 *On dit qu'une famille dénombrable d'ouverts $(\omega_n)_{n\in\mathbb{N}}$ de E est une* base dénombrable *d'ouverts si tout ouvert de E s'écrit comme réunion d'éléments de cette famille, autrement dit : $\forall O \in \mathscr{O}(E)$, $\exists I \subseteq \mathbb{N}$: $O = \cup_{i\in I}\omega_i$; ou de manière équivalente : $\forall O \in \mathscr{O}(E)$, $\forall x \in O$, $\exists n \in \mathbb{N}$: $x \in \omega_n \subseteq O$.*

Proposition 5.17 *Un espace métrique (E, d) est à base dénombrable d'ouverts ssi il contient une suite dense*[3]*. On dit alors que E est* séparable.

Dém. Sens $\Rightarrow$: soit (x_n) une suite de E telle que pour tout n, $x_n \in \omega_n$. Alors la suite (x_n) est dense, en effet : pour tout $x \in E$, l'ouvert $B(x, 1/n)$ s'écrit comme réunion d'ouverts du type ω_i, donc $\exists i(n)$ tel que $\omega_{i(n)} \subseteq B(x, 1/n)$. Soit $y_n := x_{i(n)}$, alors $d(y_n, x) \leq 1/n$, donc $y_n \to x$.

Sens $\Leftarrow$: si (x_n) est une suite dense, alors la famille $\{B(x_n, r), n \in \mathbb{N}, r \in \mathbb{Q}_+^\star\}$ est une base dénombrable d'ouverts car elle s'injecte dans $\mathbb{N} \times \mathbb{Q}$ (qui est dénombrable) et pour tout $O \in \mathscr{O}(E)$,

$$O = \bigcup_{n,r:B(x_n,r)\subseteq O} B(x_n, r),$$

ce qui achève la démonstration. $\square$

Remarque 5.18 $\mathbb{R}^d$ *est séparable car $\mathbb{Q}^d$ est une suite dense. Les rectangles ouverts (produits d'intervalles ouverts) à extrémités rationnelles forment une base dénombrable d'ouverts de $\mathbb{R}^d$.*

Définition 5.19 (Borel-Lebesgue) *Une partie A d'un espace topologique E est dite* compacte *si de tout recouvrement ouvert de A on peut extraire un sous-recouvrement fini, autrement dit pour toute famille $(\Omega_i)_{i\in I}$ d'ouverts de E telle que $A \subseteq \cup_{i\in I}\Omega_i$, $\exists J$ fini $\subseteq I$ tel que $A \subseteq \cup_{j\in J}\Omega_j$.*

Théorème 5.20 (Bolzano-Weierstrass) *Une partie A d'un espace métrique E est compacte ssi toute suite à valeurs dans A admet au moins une valeur d'adhérence dans A*[4]*.*

Corollaire 5.21 *Tout compact est fermé. De plus, toute partie compacte d'un espace vectoriel normé*[5] *est bornée.*

3. autrement dit : il existe un ensemble dénombrable (une suite, quoi) A tel que $\bar{A} = E$ (A est alors dit dense dans E)
4. autrement dit : admet au moins une sous-suite convergente de limite $\in A$
5. un espace vectoriel normé est un espace métrique, donc topologique

Théorème 5.22 *Dans un espace vectoriel normé de dimension finie, un fermé est compact ssi il est borné.*

Proposition 5.23 *Un fermé contenu dans un compact est compact. L'image d'un compact par une fonction continue est compacte.*

Remarque 5.24 *On rappelle que la topologie de $\bar{\mathbb{R}}$ est la topologie relative à la distance d définie par $d(x,y) = |f(x) - f(y)|$ pour tous $x, y \in \bar{\mathbb{R}}$, où $f(x) = x/\sqrt{x^2+1}$ et $f(\pm\infty) = \pm 1$. En particulier, $\bar{\mathbb{R}}$ est compact et $[x, +\infty]$ est un compact de $\bar{\mathbb{R}}$.*

5.2 Tribu borélienne et fonctions boréliennes

Définition 5.25 *Si E est un espace topologique, on note $\mathscr{B}or(E)$ ou $\mathscr{B}(E)$ et on appelle* tribu de Borel *ou* tribu borélienne, *la tribu engendrée par les ouverts de E, autrement dit, $\mathscr{B}(E) := \sigma(\mathscr{O}(E))$. Les éléments de $\mathscr{B}(E)$ sont appelés* parties boréliennes *de E, ou plus simplement* boréliens *de E.*

Remarque 5.26 *La tribu de Borel est aussi la tribu engendrée par la classe $\mathscr{C}$ des fermés de E, en effet : $\mathscr{C} \subseteq \mathscr{B}(E)$ (donc $\sigma(\mathscr{C}) \subseteq \mathscr{B}(E)$) car tout fermé est le complémentaire d'un ouvert, qui appartient à $\mathscr{B}(E)$, donc appartient aussi à $\mathscr{B}(E)$; $\mathscr{O}(E) \subseteq \sigma(\mathscr{C})$ (donc $\sigma(\mathscr{O}(E)) \subseteq \sigma(\mathscr{C})$) car tout ouvert est le complémentaire d'un fermé, qui appartient à $\sigma(\mathscr{C})$, donc appartient aussi à $\sigma(\mathscr{C})$ (même raisonnement).*

Remarque 5.27 *Il existe des parties de $\mathbb{R}$ non boréliennes (voir dernière section de ce chapitre). En revanche, si E est dénombrable, muni de la topologie discrète : toute partie est ouverte (et fermée), donc borélienne : $\mathscr{B}(E) = \mathscr{P}(E)$.*

Proposition 5.28 *Si E admet une base dénombrable d'ouverts $(\omega_n)_{n\in\mathbb{N}}$, alors $\mathscr{B}or(E) = \sigma(\{\omega_n; n \in \mathbb{N}\})$.*

Dém. Par double inclusion : $\{\omega_n; n \in \mathbb{N}\} \subseteq \mathscr{O}(E) \subseteq \mathscr{B}(E)$, donc $\sigma(\{\omega_n; n \in \mathbb{N}\}) \subseteq \mathscr{B}(E)$. Dans l'autre sens, on sait que tout ouvert O s'écrit comme réunion d'éléments de $\{\omega_n; n \in \mathbb{N}\}$. Comme une telle réunion est forcément dénombrable, O est un élément de $\sigma(\{\omega_n; n \in \mathbb{N}\})$. On a donc $\mathscr{O}(E) \subseteq \sigma(\{\omega_n; n \in \mathbb{N}\})$, ce qui implique $\mathscr{B}(E) \subseteq \sigma(\{\omega_n; n \in \mathbb{N}\})$. □

Corollaire 5.29 *La tribu $\mathscr{B}or(\mathbb{R}^d)$ est la tribu engendrée par la classe des rectangles ouverts*[6]*, mais est aussi la tribu engendrée par les rectangles ouverts à extrémités à coordonnées dans $\mathbb{Q}$ ou dans toute autre partie dense de $\mathbb{R}$.*

Proposition 5.30 *La tribu trace de $\mathscr{B}or(E)$ sur une partie X de E est la tribu engendrée par la topologie trace de X.*

---

6. rectangle = produit d'intervalles ; rectangle ouvert = produit d'intervalles ouverts

Dém. Soit $i : X \to E$ l'injection canonique. La tribu trace est $i^{-1}(\mathscr{B}(E)) = i^{-1}(\sigma(\mathscr{O}(E)) = \sigma(i^{-1}(\mathscr{O}(E))$, par le lemme de transport. Mais $i^{-1}(\mathscr{O}(E)$ n'est autre que la topologie trace, c'est-à-dire $\{A \cap X, A \in \mathscr{O}(E)\}$. □

Définition 5.31 (terminologie) *Si E_1 et E_2 sont des espaces topologiques, en notant $\mathscr{A}_i := \mathscr{B}or(E_i)$, les éléments de $\mathscr{F}(\mathscr{A}_1, \mathscr{A}_2)$ sont appelés* fonctions boréliennes.

Proposition 5.32 *Soit $f : (E, \mathscr{A}) \to (F, \mathscr{B})$. Si F est topologique et $\mathscr{B} = \mathscr{B}or(F)$, alors f est mesurable ssi pour tout ouvert O de F, $f^{-1}(O) \in \mathscr{A}$.*

Dém. Lemme de transport : $f^{-1}(\sigma(\mathscr{O}(F))) = \sigma(f^{-1}(\mathscr{O}(F)))$, or f est mesurable ssi $f^{-1}(\mathscr{B}(F)) \subseteq \mathscr{A}$, donc ssi $\sigma(f^{-1}(\mathscr{O}(F))) \subseteq \mathscr{A}$, c'est-à-dire ssi $f^{-1}(\mathscr{O}(F)) \subseteq \mathscr{A}$. □

Corollaire 5.33 *Si E et F sont topologiques, alors toute fonction continue est borélienne.*

Proposition 5.34 *Soit*

$$\begin{aligned} f : (E, \mathscr{A}) &\longrightarrow (\mathbb{R}^2, \mathscr{B}(\mathbb{R}^2)) \\ x &\longmapsto (f_1(x), f_2(x)) \end{aligned}$$

Alors f est mesurable ssi $f_i \in \mathscr{F}(\mathscr{A}, \mathscr{B}(\mathbb{R}))$ pour tout $i = 1, 2$.

Remarque 5.35 *Si $\mathbb{C}$ est identifié à $\mathbb{R}^2$, une fonction complexe f est mesurable ssi $\Re(f)$ et $\Im(f)$ le sont.*

Dém. Sens $\Rightarrow$: pour tout $i = 1, 2$, la projection canonique $\pi_i : \mathbb{R}^2 \to \mathbb{R}$ est continue par définition de la topologie produit, donc borélienne, ainsi comme f est mesurable, $f_i = \pi_i \circ f$ est mesurable.

Sens $\Leftarrow$: on sait que $\mathscr{B}or(\mathbb{R}^2)$ est engendrée (par exemple) par les rectangles ouverts. Donc par le lemme de transport, f est mesurable ssi pour tous intervalles ouverts U et V, $f^{-1}(U \times V) \in \mathscr{A}$. Or $f^{-1}(U \times V) = \{f_1 \in U\} \cap \{f_2 \in V\}$. Mais par hypothèse $\{f_1 \in U\} \in \mathscr{A}$ et $\{f_2 \in V\} \in \mathscr{A}$, donc leur intersection est aussi dans $\mathscr{A}$. □

Applications. Pour toutes fonctions $f, g \in \mathscr{F}(\mathscr{A}, \mathscr{B}(\mathbb{R}))$ et pour tout $\lambda \in \mathbb{R}$, $\lambda f + g$ est mesurable (autrement dit, $\mathscr{F}(\mathscr{A}, \mathscr{B}(\mathbb{R}))$ est un espace vectoriel), ainsi que les fonctions suivantes : fg, $f \wedge g$, $f \vee g$, f^+, f^-, $|f|$, $|f|^p$,... Il suffit pour le voir d'utiliser la continuité des applications qui à (x, y) associent $\lambda x + y$, xy, $x \wedge y$, etc. ainsi que le fait que la composée de deux applications mesurables est mesurable.

5.3 L'ensemble triadique de Cantor

L'ensemble triadique de Cantor est un sous-ensemble de l'intervalle $[0, 1]$. C'est un exemple de partie de $\mathbb{R}$ qui ne contient aucun point isolé mais ne contient pas non plus d'intervalle ouvert. Il est défini comme la limite d'une suite décroissante de réunions finies d'intervalles fermés, ce qui en fait un fermé (comme intersection de fermés). Plus précisément, soit A_0 l'intervalle $[0,1]$, A_1 la réunion de l'intervalle $[0,1/3]$ et de l'intervalle $[2/3,1]$, et plus généralement A_{n+1} la partie de A_n obtenue en divisant chaque composante connexe de A_n en trois sous-intervalles de tailles égales et en lui en ôtant le sous-intervalle central. Plus rigoureusement, $A_{n+1} := \frac{1}{3}A_n \cup \frac{1}{3}(2+A_n)$.

Définition 5.36 *Le fermé $K := \lim_n \downarrow A_n$ est appelé* ensemble triadique de Cantor.

Dans la proposition suivante, on appelle (provisoirement sans précautions mathématiques) « mesure de Lebesgue » d'une partie de $\mathbb{R}$, sa longueur totale. L'objet ultérieur de ce cours sera en partie de donner une définition rigoureuse de ce concept...

Proposition 5.37 *L'ensemble triadique de Cantor peut s'écrire sous la forme*

$$K = \left\{ \sum_{n\geq 1} \frac{x_n}{3^n}, x_n \in \{0,2\} \right\}.$$

Il est compact, d'intérieur vide, équipotent à $\mathbb{R}$, de mesure de Lebesgue nulle.

Remarque 5.38 *Tout ensemble dénombrable est de mesure de Lebesgue nulle, comme réunion* dénombrable *d'ensembles de mesure nulle (les singletons le constituant). On voit ici que la réciproque est fausse : K est un exemple d'ensemble de mesure de Lebesgue nulle mais non dénombrable.*

Dém. K est fermé borné dans $\mathbb{R}$ donc compact. Par une récurrence immédiate, on voit que les composantes connexes de A_n qui sont des intervalles fermés de longueur 3^{-n} dont les extrémités sont les nombres réels de la forme $\sum_{k=1}^n \frac{x_k^{(n)}}{3^n} + \frac{\varepsilon_n}{3^n}$, où $x_k^{(n)} \in \{0,2\}$ et $\varepsilon_n \in \{0,1\}$: pour chaque intervalle, $\varepsilon_n = 0$ correspond à l'extrémité gauche, et $\varepsilon_n = 1$ correspond à l'extrémité droite. Montrons l'égalité annoncée par double inclusion : @

$\supseteq$: pour toute suite (x_k) à valeurs dans $\{0,2\}$, pour tout entier n, $\sum_{k=1}^n \frac{x_k}{3^n} \in A_n \subseteq K$, donc comme K est fermé, la limite $\sum_{k=1}^\infty \frac{x_k}{3^n} \in K$.

$\subseteq$: soit $x \in K$ et soit $x^{(n)}$ l'extrémité gauche de la composante connexe de A_n qui contient x. En particulier $|x^{(n)} - x| \leq 3^{-n}$. Cherchons une relation entre $x^{(n)}$ et $x^{(n+1)}$. Lorsqu'on passe de A_n à A_{n+1}, soit x est dans le sous-intervalle de gauche, auquel cas $x^{(n+1)} = x^{(n)}$, soit x est dans le sous-intervalle de droite, auquel cas $x^{(n+1)} = x^{(n)} + \frac{2}{3^{n+1}}$. On peut donc écrire $x^{(n+1)} = x^{(n)} + \frac{x_{n+1}}{3^{n+1}}$, où $x_{n+1} \in \{0,2\}$, et comme $x^{(0)} = 0$, cela donne $x^{(n)} = \sum_{k=1}^n \frac{x_k}{3^k}$, qui converge en croissant vers $y := \sum_{k=1}^\infty \frac{x_k}{3^k}$. Or $|x^{(n)} - x| \leq 3^{-n}$ donc la suite $(x^{(n)})$ converge vers x, ce qui implique $y = x$.

Montrons que $\overset{\circ}{K} = \varnothing$. Soit $x \in K$ et $\varepsilon > 0$. La boule $B(x,\varepsilon)$ intersecte ${}^cA_n$ pour tout n dès que $3^{-n} < \varepsilon$. Donc $B(x,\varepsilon)$ intersecte $\cup_n {}^cA_n$, qui n'est autre que le complémentaire de $\cap_n A_n = K$. Ainsi, K ne contient aucune boule ouverte centrée sur x, c'est-à-dire que x n'est pas intérieur à K.

Montrons que K a la puissance du continu. L'application

$$\begin{aligned} f : \{0,2\}^{\mathbb{N}^\star} &\longrightarrow K \\ (x_n) &\longmapsto \sum_{n\geq 1} \frac{x_n}{3^n} \end{aligned}$$

est une injection donc $\mathrm{Card}(K) \geq \mathrm{Card}(\{0,2\}^{\mathbb{N}^\star}) = \mathrm{Card}(\mathbb{R})$. D'autre part $\mathrm{Card}(\mathbb{R}) \leq$ @ $\mathrm{Card}(K)$ puisque $K \subseteq \mathbb{R}$.

Enfin K est de mesure de Lebesgue nulle car $K = \lim_n \downarrow A_n$ donc[7] $\lambda(K) = \lim_n \downarrow \lambda(A_n) = \lim_n \downarrow \left(\frac{2}{3}\right)^n = 0$. □

5.4 Une partie de $\mathbb{R}$ non borélienne

Les tribus sont des familles de parties qui sont destinées à être mesurées. Pour pouvoir mesurer des parties suffisamment compliquées comme celles qui ne peuvent être définies que par des passages à la limite (comme l'ensemble triadique de Cantor), les tribus doivent être assez fines pour être stables par des opérations relativement générales comme le passage au complémentaire, les réunions et intersections dénombrables. Néanmoins, elles ne doivent pas être si fines qu'elles contiennent des parties non mesurables, comme l'exemple qui va suivre.

On définit la relation d'équivalence $\sim$ sur $\mathbb{R}$:

$$x \sim y \Leftrightarrow x - y \in \mathbb{Q}.$$

En se servant de l'axiome du choix, on peut supposer l'existence d'une partie A de $]0,1[$ *qui contient exactement un représentant et un seul de chaque classe d'équivalence de la relation* $\sim$. En particulier, A n'est pas dénombrable, mais surtout nous allons montrer que A *ne peut admettre de mesure de Lebesgue*. Cette assertion implique (mais est plus forte que) l'assertion suivante : A n'est pas borélienne. En effet, nous verrons (plus tard) que tout borélien admet une mesure de Lebesgue.

Montrons par l'absurde que A ne peut admettre de mesure de Lebesgue : soit $\lambda(A) \in [0,+\infty]$ la mesure de A (nous verrons que λ est la notation usuelle de la mesure de Lebesgue). Soit

$$L := \bigcup_{r\in\mathbb{Q}\cap]-1,1[} (r + A),$$

où $r + A = \{r + x, x \in A\}$. Comme A admet une mesure, alors chaque partie $r + A$ en admet une aussi, qui vaut d'ailleurs $\lambda(A)$ par invariance par translation de la mesure de Lebesgue. Comme L est réunion dénombrable de parties admettant une mesure, ce doit être également son cas.

Montrons que $]0,1[\subseteq L$. Pour tout $x \in]0,1[$, désignons par $a = a(x)$ le représentant de sa classe d'équivalence contenu dans A. Alors en particulier, $x-a \in \mathbb{Q}$, et $x-a \in]-1,1[$,

7. Propriété de continuité de la mesure pour les suites décroissantes dont un élément est de mesure finie, ce que nous verrons bientôt...

donc $r := x - a \in \mathbb{Q} \cap]-1, 1[$, et comme $x \in r + A$, $x \in L$. On a aussi $L \subseteq]-1, 2[$, donc on en déduit

$$1 \leq \lambda(L) \leq 3.$$

Montrons que les parties $r + A$ ($r \in \mathbb{Q}$) sont deux à deux disjointes. Soient $r, s \in \mathbb{Q}$. Si $(r + A) \cap (s + A) \neq \varnothing$, alors il existe $a, b \in A$ tels que $z = r + a = s + b$, donc $b - a = r - s \in \mathbb{Q}$. Par conséquent $a \sim b$, mais comme $a, b \in A$ qui ne contient qu'un représentant de chaque classe d'équivalence, $a = b$, donc $r = s$.

Par σ-additivité, nous en déduisons

$$\lambda(L) = \lambda(\cup_r (r + A)) = \sum_r \lambda(r + A) = \sum_r \lambda(A).$$

Cette somme ne peut être qu'infinie (si $\lambda(A) \neq 0$) ou nulle (si $\lambda(A) = 0$), ce qui contredit l'inégalité $1 \leq \lambda(L) \leq 3$. $\square$

Chapitre 6

Mesures

6.1 Définitions et propriétés

Définition 6.1 *Une mesure*[1] *sur l'espace mesurable* $(E, \mathscr{A})$ *est une application* $\mu : \mathscr{A} \to [0, +\infty]$ *qui :*

(i) associe la valeur 0 à l'ensemble vide : $\mu(\varnothing) = 0$ *;*

(ii) est σ-additive *: pour toute suite* (A_n) *d'éléments de* $\mathscr{A}$ *deux à deux disjoints,*

$$\mu(\cup_n A_n) = \sum_n \mu(A_n).$$

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*.*

Remarque 6.2 *On a besoin de la* σ*-additivité pour pouvoir calculer la mesure de parties compliquées construites comme limites d'ensembles plus simples que l'on sait mesurer.*

Remarque 6.3 *Dans l'égalité* $\mu(\cup_n A_n) = \sum_n \mu(A_n)$*, on remarque que l'ordre de sommation (membre de droite) n'intervient pas car la série est à termes positifs, ce qui est cohérent avec le membre de gauche.*

On remarquera également que la σ*-additivité implique l'additivité finie grâce à (i) : si l'on définit* $A_i = \varnothing$ *pour tout* $i \geq n+1$*, alors* $\mu(\cup_{i=1}^n A_i) = \mu(\cup_{i=1}^\infty A_i) = \sum_{i\geq 1} \mu(A_i) = \sum_{i=1}^n \mu(A_i)$*.*

Proposition 6.4 *Une mesure* μ *sur un* $(E, \mathscr{A})$ *vérifie pour tous* $A, B \in \mathscr{A}$ *:*

(i) Additivité finie : $\mu(A) = \mu(A \setminus B) + \mu(A \cap B)$ *;*

(ii) Additivité forte : $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$ *;*

(iii) Sous-additivité : $\mu(A \cup B) \leq \mu(A) + \mu(B)$ *;*

(iv) Croissance : si $A \subseteq B$*,* $\mu(A) \leq \mu(B)$*.*

Remarque 6.5 *En (ii), prendre garde de ne pas écrire* $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$*, qui pourrait être une forme indéterminée, si* $\mu(A \cap B) = +\infty$*.*

1. Il est sous-entendu que nous ne considérons dans ce cours que des mesures *positives*

Dém. (i) $A \setminus B$ et $A \cap B$ sont disjoints et leur réunion est A.

(ii) $A \setminus B$, $A \cap B$ et $B \setminus A$ sont disjoints et leur réunion est $A \cup B$, donc

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

donc en ajoutant $\mu(A \cap B)$ à chaque membre on obtient

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

mais dans le premier membre, grâce à (i), la somme des deux premiers termes vaut $\mu(A)$ et la somme des deux derniers termes vaut $\mu(B)$.

(iii) Si $\mu(A) + \mu(B) = +\infty$, l'assertion est évidente, tandis que dans le cas contraire, grâce à (ii), $\mu(A \cup B) + \mu(A \cap B) < \infty$, donc en particulier $\mu(A \cap B) < \infty$. Par conséquent on peut retrancher $\mu(A \cap B)$ à l'égalité (ii), ce qui donne

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

(iv) D'après (ii) si $A \subseteq B$, alors

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

qui est l'inégalité souhaitée. □

Proposition 6.6 *Une application $\mu : \mathscr{A} \to [0, +\infty]$ est une mesure ssi :*

(i) $\mu(\varnothing) = 0$;

(ii) μ est finiment additive *: pour tous éléments A_i ($i \in I$) deux à deux disjoints de la tribu $\mathscr{A}$, si I est fini, alors $\mu(\cup_{i \in I} A_i) = \sum_{i \in I} \mu(A_i)$.*

(iii) μ est continue à gauche[2] *: pour toute suite* croissante *$(A_n)_{n \in \mathbb{N}}$ d'éléments de $\mathscr{A}$,*

$$\mu(\lim_n \uparrow A_n) = \lim_n \uparrow \mu(A_n).$$

Remarque 6.7 *On se rappellera qu'ici, la suite $(A_n)_{n \in \mathbb{N}}$ étant croissante, $\lim_n \uparrow A_n$ n'est autre que $\cup_n A_n$.*

Dém. Montrons d'abord le sens $\Rightarrow$ et supposons donc que μ est une mesure. On a déjà vu que (i) et (ii) sont vraies. Montrons la continuité à gauche. Soit (A_n) une suite croissante de parties mesurables et soient $B_0 := A_0$, et pour tout entier naturel non nul n, $B_n := A_n \setminus A_{n-1}$. Alors les (B_n) sont des éléments de $\mathscr{A}$ deux à deux disjoints, donc

$$\mu(\cup_n B_n) = \sum_n \mu(B_n).$$

Mais d'une part, $\cup_n B_n = \cup_n A_n = \lim_n \uparrow A_n$ et d'autre part,

$$\sum_n \mu(B_n) = \lim_n \sum_{k=0}^{n} \mu(B_k) = \lim_n \mu(\cup_{k=0}^{n} B_k) = \lim_n \mu(A_n).$$

2. il s'agit d'une expression figurée qui signifie 'continue pour les suites croissantes' et est utilisée par analogie avec les fonctions : $\mathbb{R} \to \mathbb{R}$ pour qui ces deux expressions sont synonymes

Montrons maintenant $\Leftarrow$. Soit donc μ vérifiant les trois propriétés de la proposition. Il nous suffit de montrer que μ est bien σ-additive. Soient (A_n) mesurables et deux à deux disjointes. Soit $B_n := \cup_{k=0}^n A_k$, alors (B_n) est une suite croissante donc $\mu(\cup_n B_n) = \lim_n \mu(B_n)$. Mais d'une part $\mu(\cup_n B_n) = \mu(\cup_n \cup_{k=0}^n A_k) = \mu(\cup_n A_n)$, et d'autre part, comme μ est finiment additive, $\mu(B_n) = \mu(\cup_{k=0}^n A_k) = \sum_{k=0}^n \mu(A_k)$. Ainsi

$$\mu(\cup_n A_n) = \mu(\cup_n B_n) = \lim_n \mu(B_n) = \lim_n \sum_{k=0}^n \mu(A_k) = \sum_n \mu(A_n),$$

ce qui montre la σ-additivité de μ. □

Corollaire 6.8 *Toute mesure μ est* sous σ-additive*, au sens où pour toute suite (A_n) d'éléments de $\mathscr{A}$, $\mu(\cup_n A_n) \leq \sum_n \mu(A_n)$.*

Dém. Soit $B_n := \cup_{k=0}^n A_k$. Par sous-additivité, $\mu(B_n) \leq \sum_{k=0}^n \mu(A_k)$. Mais comme la suite (B_n) croît et converge vers $\cup_n A_n$, en passant à la limite dans l'inégalité précédente, on obtient

$$\mu(\cup_n A_n) = \mu(\cup_n B_n) = \lim_n \mu(B_n) \leq \lim_n \sum_{k=0}^n \mu(A_k) = \sum_n \mu(A_n),$$

où la deuxième égalité est due à la continuité à gauche des mesures. □

Exemple 6.9 *Quelques exemples de mesures :*

- *la* mesure nulle *est définie sur $\mathscr{P}(E)$ (et donc sur toute autre tribu) par $\mu(A) := 0$ pour tout $A \subseteq E$;*
- *la* mesure grossière *sur $\mathscr{P}(E)$: $\mu(A) := +\infty$ dès que $A \neq \varnothing$ (et $\mu(\varnothing) = 0$) ;*
- *pour tout $a \in E$, la* mesure de Dirac *au point a est définie pour tout $A \in \mathscr{P}(E)$ par*

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

Cette mesure est souvent notée δ_a ;

- *la* mesure de comptage *sur $\mathscr{P}(E)$:*

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

où Card(A) désigne ici le nombre d'éléments de l'ensemble A.

- *soit un espace mesuré $(E, \mathscr{A}, \mu)$ et X une partie de E.* ***SI*** *$X \in \mathscr{A}$, alors on peut définir la* mesure trace *μ_X de μ sur X par $\mu_X(A) := \mu(A \cap X)$ pour tout $A \in \mathscr{A}$.*

Exercice 6.10 *Démontrer que la mesure de comptage est bien une mesure en prouvant qu'elle vérifie les trois propriétés de la Proposition 6.6 : elle prend la valeur 0 en $\varnothing$, elle est finiment additive et elle est continue à gauche.*

6.2 Mesure de Lebesgue

La *mesure de Lebesgue* est une mesure définie sur la tribu de Borel de $\mathbb{R}^d$. Elle donne un sens mathématique à la notion physique de volume (de surface si $d = 2$, de longueur si $d = 1$). Rappelons que cette mesure n'est pas définie sur tout $\mathscr{P}(\mathbb{R}^d)$, ce que nous avons démontré au chapitre précédent dans le cas $d = 1$ à l'aide de l'axiome du choix.

Théorème 6.11 *Il existe une unique mesure sur les boréliens de $\mathbb{R}^d$ telle que la mesure de tout rectangle $\prod_{i=1}^{d}]a_i, b_i[$ soit égale au produit $\prod_{i=1}^{d}(b_i - a_i)$. Cette mesure est appelée* mesure de Lebesgue *et est ordinairement notée λ_d, voire λ s'il n'y a pas d'ambiguïté sur la dimension.*

Remarque 6.12 *Montrer qu'il existe une unique mesure qui vérifie certaines propriétés se dit « construire une mesure ». Le théorème dont on se sert pour montrer l'unicité s'appelle* théorème de la classe monotone*, et celui dont on se sert pour l'existence s'appelle* théorème de Caratheodory*. Nous énoncerons ces théorèmes au second semestre (Intégration II) et démontrerons même le premier. Pour le moment le théorème qui précède reste admis.*

Exercice 6.13 *Montrer que si A est un borélien de $\mathbb{R}^d$ alors tous les translatés de A sont des boréliens (se servir du fait qu'une translation est une application bijective et continue).*

Proposition 6.14 *Soit μ une mesure sur $\mathscr{B}or(\mathbb{R}^d)$ qui vérifie les deux propriétés suivantes :*

(i) invariance par translation : pour tout borélien A et toute translation f, $\mu(f(A)) = \mu(A)$;

(ii) le rectangle unité est de mesure 1 : $\mu\left([0,1]^d\right) = 1$.

Alors μ est la mesure de Lebesgue.

Dém. Nous ne détaillons ici que le cas $d = 1$. Le cas général est laissé au lecteur. @

a) Nous montrons d'abord par l'absurde que μ est nulle sur les singletons. S'il existe $x \in \mathbb{R}$ tel que $\mu(\{x\}) = \varepsilon > 0$, alors par invariance par translation, $\mu(\{y\}) = \varepsilon$ pour tout $y \in \mathbb{R}$. Par conséquent, $\mu(\mathbb{Q} \cap [0,1]) = \sum_{y \in \mathbb{Q} \cap [0,1]} \varepsilon = +\infty$, ce qui constitue une contradiction puisque $\mu(\mathbb{Q} \cap [0,1]) \le \mu([0,1]) = 1$.

b) D'après ce qui précède, pour tout entier naturel $n \ge 1$,

$$1 = \mu([0,1]) = \sum_{k=1}^{n} \mu\left(\left]\frac{k-1}{n}, \frac{k}{n}\right[\right) = \sum_{k=1}^{n} \mu\left(\left]0, \frac{1}{n}\right[\right) = n\mu\left(\left]0, \frac{1}{n}\right[\right),$$

d'où $\mu(]0, 1/n[) = 1/n$. De plus, pour tous entiers $k_1 \le k_2$,

$$\mu\left(\left]\frac{k_1}{n}, \frac{k_2}{n}\right[\right) = \sum_{j=k_1+1}^{k_2} \mu\left(\left]\frac{j-1}{n}, \frac{j}{n}\right[\right) = \sum_{j=k_1+1}^{k_2} \mu\left(\left]0, \frac{1}{n}\right[\right) = \frac{k_2 - k_1}{n}.$$

c) Soient $r < r'$ deux rationnels, que l'on peut écrire sous la forme $r = p/q$ et $r' = p'/q'$, où p, p', q, q' sont des entiers. Alors d'après ce qui précède,

$$\mu\left(]r, r'[\right) = \mu\left(\left]\frac{p}{q}, \frac{p'}{q'}\right[\right) = \mu\left(\left]\frac{pq'}{qq'}, \frac{p'q}{qq'}\right[\right) = \frac{p'q - pq'}{qq'} = r' - r.$$

d) Passons maintenant à la limite sur les rationnels. Soient $a < b$ deux nombres réels. Alors il existe une suite décroissante (a_n) et une suite croissante (b_n), toutes deux constituées de nombres rationnels, dont les limites sont resp. a et b. Alors la suite d'intervalles $(]a_n, b_n[)$ est une suite croissante qui converge vers $]a, b[$, donc par continuité à gauche des mesures,

$$\mu(]a, b[) = \mu(\lim_n \uparrow]a_n, b_n[) = \lim_n \uparrow \mu(]a_n, b_n[) = \lim_n \uparrow (b_n - a_n) = b - a,$$

ce qui montre que la mesure de tout intervalle est sa longueur, et garantit ainsi que μ est la mesure de Lebesgue sur $\mathbb{R}$. $\square$

6.3 Autres définitions et autres propriétés

Définition 6.15 *Une mesure μ sur un espace mesurable $(E, \mathscr{A})$:*

- *est dite* finie*, ou* bornée*, si $\mu(E) < \infty$ (ce qui équivaut à : $\mu(A) < \infty$ pour tout $A \in \mathscr{A}$). Le nombre réel $\mu(E)$ est alors appelé* masse totale *de μ ;*
- *est appelée (mesure de)* probabilité *si sa masse totale vaut 1 ;*
- *est dite σ-*finie *s'il existe une suite (E_n) de parties mesurables de E telles que $\mu(E_n) < \infty$ et $\cup_n E_n = E$;*
- *est appelée* mesure de Borel[3] *si E est topologique, localement compact[4] et séparable[5], que $\mathscr{A}$ est la tribu borélienne de E et que μ est finie sur les compacts : $\mu(K) < \infty$ pour tout compact K de E[6].*

Remarque 6.16 *Si μ est une mesure de Borel alors elle est σ-finie car E étant localement compact et séparable, il peut s'écrire comme réunion dénombrable de compacts (on dit qu'il est σ-compact) : $E = \cup_{n\in\mathbb{N}} E_n$ où tous les E_n sont compacts donc vérifient $\mu(E_n) < \infty$.*

En revanche la réciproque est fausse. Soit μ la mesure sur $\mathbb{R}$ définie par $\mu = \sum_n \alpha_n \delta_{x_n}$ (voir Corollaire 6.21) où $\sum_n \alpha_n = \infty$ et (x_n) est une suite de réels deux à deux distincts et de limite x finie. Alors μ est une mesure σ-finie (prendre $E_n = \mathbb{R} \setminus \{x_k ; k \geq n\}$) mais elle n'est pas de Borel car tout voisinage compact de x est de mesure infinie.

Proposition 6.17 (Continuité pour les suites décroissantes de mesure finie) *Si (A_n) est une suite décroissante de $\mathscr{A}$ telle que $\mu(A_n) < \infty$ à partir d'un certain rang,*

3. ou parfois *mesure de Radon*. En fait, le terme « mesure de Radon » fait référence à une forme linéaire positive sur un espace de fonction continues à support compact. Le théorème de représentation de Riesz assure que toute mesure de Radon est en fait une intégrale par rapport à une mesure de Borel

4. autrement dit : pour tout point $x \in E$, il existe un ouvert contenant x et inclus dans un compact de E

5. rappel : admettant une suite dense

6. rappel : un compact est fermé donc borélien

alors

$$\lim_n \downarrow \mu(A_n) = \mu(\lim_n \downarrow A_n),$$

qui n'est autre que $\mu(\cap_n A_n)$.

Remarque 6.18 *Un corollaire immédiat de la proposition précédente est que* les mesures finies sont continues à droite. *La mesure de Lebesgue est un exemple de mesure non continue à droite : si* $A_n := [n, +\infty[$, *alors* (A_n) *est une suite décroissante de limite* $\varnothing$, *mais comme* $(\lambda(A_n))$ *est identiquement égale à* $+\infty$, *elle converge vers* $+\infty$, *et non pas vers* $\lambda(\varnothing) = 0$.

Dém. Par hypothèse, il existe n_0 tel que pour tout $n \geq n_0$, $\mu(A_n) < \infty$. Soit alors $B_n := A_{n_0} \setminus A_n$. La suite (B_n) est croissante et converge vers $A_{n_0} \setminus \cap_n A_n$, donc

$$\mu(A_{n_0}) - \mu(\cap_n A_n) = \mu(\lim_n \uparrow B_n) = \lim_n \uparrow \mu(B_n) = \lim_n \uparrow (\mu(A_{n_0}) - \mu(A_n)) = \mu(A_{n_0}) - \lim_n \downarrow \mu(A_n),$$

ce qui donne bien $\mu(\cap_n A_n) = \lim_n \downarrow \mu(A_n)$. □

Proposition 6.19 *Pour toute suite* (A_n) *d'éléments de la tribu* $\mathscr{A}$, *si* μ *est une mesure finie (ou s'il existe* B *de mesure finie tel que* $A_n \subseteq B$ *à partir d'un certain rang), alors*

$$\mu\left(\liminf_n A_n\right) \leq \liminf_n \mu(A_n) \leq \limsup_n \mu(A_n) \leq \mu\left(\limsup_n A_n\right).$$

La première inégalité reste valable sans les hypothèses qui précèdent.

Dém. Soit $B_n := \cap_{k \geq n} A_k$. Alors (B_n) est une suite croissante qui converge vers $\liminf_n A_n$, donc $\mu(\liminf_n A_n) = \lim_n \uparrow \mu(B_n)$. Or $B_n \subseteq A_n$ donc $\mu(B_n) \leq \mu(A_n)$ et par conséquent $\lim_n \mu(B_n) = \liminf_n \mu(B_n) \leq \liminf_n \mu(A_n)$, ce qui assure la première inégalité.

Concernant les limites supérieures, supposons que μ est finie (mais sous l'hypothèse plus faible de l'énoncé, la démonstration est la même). Alors @

$$\begin{aligned} \mu\left(\limsup_n A_n\right) &= \mu(E) - \mu\left(\liminf_n {}^c A_n\right) \geq \mu(E) - \liminf_n \mu({}^c A_n) \\ &= \mu(E) - \liminf_n (\mu(E) - \mu(A_n)) = \limsup_n \mu(A_n), \end{aligned}$$

où l'inégalité est due à la conclusion précédente. □

Terminons ce chapitre par la

Proposition 6.20 *a) Si* (μ_n) *est une suite croissante de mesures, au sens où pour tout* $A \in \mathscr{A}$, $\mu_n(A) \leq \mu_{n+1}(A)$, *alors l'égalité* $\mu(A) := \lim_n \uparrow \mu_n(A) \in [0, +\infty]$ *définit une mesure* μ *sur* $\mathscr{A}$.

b) Tout combinaison linéaire dénombrable, à coefficients positifs, de mesures, est une mesure.

Dém. Pour b), il suffit de montrer qu'une combinaison linéaire finie, à coefficients positifs, de mesures, soit $\sum_{k=0}^{n} \alpha_k \mu_k$, est toujours une mesure, car alors a) impliquera b). En effet, une combinaison linéaire à coefficients positifs dénombrable est simplement la limite croissante d'une suite de sommes partielles. La démonstration se fait (par exemple) sur le même modèle que celle qui suit. @

Démontrons a) grâce à la Proposition 6.6.

(i) comme $\mu_n(\varnothing) = 0$, $\mu(\varnothing) = \lim_n \mu_n(\varnothing) = 0$.

(ii) pour tout ensemble d'indices fini I, pour toutes parties mesurables $(A_i)_{i\in I}$ deux à deux disjointes, l'additivité finie de chaque μ_n s'écrit

$$\mu_n\left(\cup_{i\in I} A_i\right) = \sum_{i\in I} \mu_n(A_i).$$

L'additivité finie de μ s'obtient en faisant tendre $n \to \infty$ dans chaque membre (car le membre de droite est une somme finie).

(iii) soit maintenant une suite croissante (A_k) d'éléments de la tribu $\mathscr{A}$. La suite doublement indicée $(\mu_n(A_k))$ est croissante en k ET en n, ce qui garantit que l'on peut @ intervertir les limites en n et en k, d'où :

$$\mu\left(\lim_k \uparrow A_k\right) = \lim_n \uparrow \mu_n\left(\lim_k \uparrow A_k\right) = \lim_n \uparrow \lim_k \uparrow \mu_n(A_k) = \lim_k \uparrow \lim_n \uparrow \mu_n(A_k) = \lim_k \uparrow \mu(A_k),$$

où la deuxième égalité est due à la continuité à gauche de chaque mesure μ_n. □

Corollaire 6.21 *Pour toute suite (x_n) d'éléments d'un ensemble E, pour toute suite (α_n) de nombre réels positifs, $\sum_n \alpha_n \delta_{x_n}$ est une mesure sur $\mathscr{P}(E)$.*

Remarque 6.22 *Le résultat précédent montre que l'on peut définir sur n'importe quel espace des mesures qui sont un peu moins élémentaires que les exemples généraux donnés dans la première section.*

Chapitre 7

Intégrale par rapport à une mesure des fonctions mesurables positives

7.1 Intégrale des fonctions étagées positives

Notation 7.1 *Pour tout espace mesurable $(E, \mathscr{A})$, on notera $\mathscr{E}_+(\mathscr{A})$ l'ensemble des éléments de $\mathscr{E}(\mathscr{A})$ (fonctions étagées) à valeurs* positives.

Définition 7.2 *Pour toute fonction $f \in \mathscr{E}_+(\mathscr{A})$, on appelle intégrale de f par rapport à une mesure μ sur $(E, \mathscr{A})$, et l'on note $\int_E f d\mu$ l'élément de $[0, +\infty]$*

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

avec la convention habituelle $0 \times \infty = 0$.

Remarque 7.3 *La définition précédente ne dépend (heureusement) pas de la représentation de f sous la forme $f = \sum_{i \in I} \alpha_i \mathbb{1}_{A_i}$, car on a toujours l'égalité* @

$$\int_E f\, d\mu = \sum_{i \in I} \alpha_i \mu(A_i).$$

Notation 7.4 *On notera indifféremment l'intégrale de f par rapport à la mesure μ sous une des formes suivantes*

$$\int_E f\, d\mu, \qquad \int_E f(x)\, d\mu(x), \qquad \int_E f(x)\, \mu(dx),$$

voire en omettant l'indice E du signe intégral.

Proposition 7.5 *Pour tout $f \in \mathscr{E}_+(\mathscr{A})$,*

$$\int_E f\, d\mu < \infty \Leftrightarrow \mu(\{f \neq 0\}) < \infty.$$

Dém. Soit $f = \sum_{i \in I} \alpha_i \mathbb{1}_{A_i}$. Alors

$$\begin{aligned}\sum_{i \in I} \alpha_i \mu(A_i) < \infty &\iff \forall i \in I \quad (\alpha_i \neq 0 \Rightarrow \mu(A_i) < \infty) \\ &\iff \sum_{i \in I : \alpha_i \neq 0} \mu(A_i) < \infty \\ &\iff \mu\left(\bigcup_{i \in I : \alpha_i \neq 0} A_i\right) < \infty,\end{aligned}$$

ce qui achève la démonstration, car $\bigcup_{i \in I : \alpha_i \neq 0} A_i = \{f \neq 0\}$. □

Exemple 7.6 *Si f est nulle alors $\int_E f \, d\mu = 0$.*

Si $\mu = \delta_a$, alors

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

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

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

Proposition 7.7 *L'application $f \mapsto \int_E f \, d\mu$ du cône $\mathscr{E}_+(\mathscr{A})$ dans $\bar{\mathbb{R}}_+$ jouit des propriétés suivantes :*

(i) additivité : $\int (f+g) \, d\mu = \int f \, d\mu + \int g \, d\mu$;
(ii) positive homogénéité : pour tout réel positif a, $\int (af) \, d\mu = a \int f \, d\mu$;
(iii) croissance : pour tous $f, g \in \mathscr{E}_+(\mathscr{A})$, $f \leq g \Rightarrow \int f \, d\mu \leq \int g \, d\mu$.

Dém. Soient $f = \sum_{i \in I} \alpha_i \mathbb{1}_{A_i}$ et $g = \sum_{j \in J} \beta_j \mathbb{1}_{B_j}$, où les α_i, β_j sont des réels positifs ou nuls, et $(A_i)_{i \in I}$, $(B_j)_{j \in J}$ sont des partitions finies de E.

(i) Remarquons que $(A_i \cap B_j)_{(i,j) \in I \times J}$ est une partition finie de E et que

$$f + g = \sum_{(i,j) \in I \times J} (\alpha_i + \beta_j) \mathbb{1}_{A_i \cap B_j}.$$

Par conséquent,

$$\begin{aligned}\int_E (f+g) \, d\mu &= \sum_{i,j} (\alpha_i + \beta_j) \mu(A_i \cap B_j) \\ &= \sum_{i,j} \alpha_i \mu(A_i \cap B_j) + \sum_{i,j} \beta_j \mu(A_i \cap B_j) \\ &= \sum_{i \in I} \alpha_i \sum_{j \in J} \mu(A_i \cap B_j) + \sum_{j \in J} \beta_j \sum_{i \in I} \mu(A_i \cap B_j) \\ &= \sum_{i \in I} \alpha_i \mu(A_i) + \sum_{j \in J} \beta_j \mu(B_j) \\ &= \int_E f \, d\mu + \int_E g \, d\mu.\end{aligned}$$

(ii) Pour tout $a \geq 0$, $af = \sum_{i \in I} a\alpha_i \mathbb{1}_{A_i}$, d'où

$$\int_E (af)\, d\mu = \sum_{i \in I} a\alpha_i \mu(A_i) = a \sum_{i \in I} \alpha_i \mu(A_i) = a \int_E f\, d\mu.$$

(iii) En écrivant $g = f + (g - f)$, où $g - f$ est étagée positive, d'après (i), $\int g\, d\mu = \int f\, d\mu + \int (g-f)\, d\mu$, donc $\int g\, d\mu \geq \int f\, d\mu$. □

7.2 Intégrale des fonctions mesurables positives

Notation 7.8 *Pour tout espace mesurable $(E, \mathscr{A})$, on notera $\mathscr{F}_+(\mathscr{A})$ l'ensemble des élémnets de $\mathscr{F}(\mathscr{A}, \mathscr{B}or(\bar{\mathbb{R}}))$ (fonctions mesurables à valeurs dans $\bar{\mathbb{R}}$) à valeurs* positives.

Définition 7.9 *Pour tout $f \in \mathscr{F}_+(\mathscr{A})$, on appelle intégrale de f par rapport à μ, et l'on note*[1] *$\int_E f\, d\mu$ l'élément de* $[0, +\infty]$

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

Si $\int_E f\, d\mu < \infty$, on dira que f est intégrable.

Proposition 7.10 (croissance de l'intégrale) *Pour toutes $f, g \in \mathscr{F}_+(\mathscr{A})$, si $f \leq g$, alors $\int_E f\, d\mu \leq \int_E g\, d\mu$.*

Dém. Si $\varphi \in \mathscr{E}_+(\mathscr{A})$ est telle que $\varphi \leq f$ alors $\varphi \leq g$ donc

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

ce qui est l'inégalité recherchée. □

Théorème 7.11 (Théorème de Beppo Levi, ou de convergence monotone) *Si (f_n) est une suite* croissante *de $\mathscr{F}_+(\mathscr{A})$, alors nous savons que $f := \lim_n \uparrow f_n \in \mathscr{F}_+(\mathscr{A})$, mais surtout*

$$\int_E f\, d\mu = \lim_n \uparrow \int_E f_n\, d\mu.$$

Corollaire 7.12 *L'intégrale $\int_E f\, d\mu$ est la limite des intégrales $\int_E f_n\, d\mu$, où (f_n) est une suite* arbitraire *de fonctions étagées positives croissant vers f.*

Remarque 7.13 *Le corollaire précédent assure qu'on aurait pu* définir *$\int_E f\, d\mu$ comme la* limite *(et non la borne supérieure etc.) des intégrales de toute suite de fonctions étagées positives croissant vers f, mais l'inconvénient est qu'il aurait fallu auparavant montrer que cette limite ne dépend effectivement pas de la suite de fonctions choisie.*

1. la même notation est encore utilisée, car il s'agit d'un *prolongement* de l'intégrale initialement définie pour les fonctions *étagées* positives, aux fonctions *mesurables* positives

Dém. du théorème de Beppo Levi. Montrons d'abord l'inégalité $\geq$. Comme pour tout entier n, on a $f_n \leq f_{n+1} \leq f$, par croissance de l'intégrale on a également

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

ce qui prouve en passant à la limite que $\lim_n \uparrow \int_E f_n \, d\mu \leq \int_E f \, d\mu$.

Montrons maintenant l'autre inégalité. Par définition de l'intégrale de f, il suffit de montrer que pour toute fonction étagée positive φ telle que $\varphi \leq f$, on a $\int \varphi \, d\mu \leq \lim_n \uparrow \int_E f_n \, d\mu$. Soit alors $a \in [0, 1[$ et $E_n := \{a\varphi \leq f_n\}$. Comme $\varphi \leq f$, on a l'égalité $E = \cup_n E_n$, en effet :

- sur $\{f = 0\}$, $f_n = \varphi = 0$ pour tout entier n, donc $\{f = 0\} \subseteq E_n$ et par conséquent $\{f = 0\} \subseteq \cup_n E_n$;
- sur $\{f > 0\}$, $a\varphi < f$ car φ ne prend que des valeurs finies. Donc pour tout $x \in \{f > 0\}$, il existe un rang $N(x)$ tel que pour tout $n \geq N(x)$, $a\varphi(x) \leq f_n(x)$, autrement dit $x \in E_n$, et par conséquent $x \in \cup_n E_n$.

En conclusion, $E = \{f = 0\} \cup \{f > 0\} \subseteq \cup_n E_n$. Or en notant $\varphi = \sum_{i \in I} \alpha_i \mathbb{1}_{A_i}$,

$$\int_E a\varphi \mathbb{1}_{E_n} \, d\mu = a \int_E \sum_{i \in I} \alpha_i \mathbb{1}_{A_i \cap E_n} \, d\mu = \sum_{i \in I} a\alpha_i \mu(A_i \cap E_n).$$

Ainsi comme les E_n croissent vers E, par continuité à gauche de la mesure, $\lim_n \uparrow \mu(A_i \cap E_n) = \mu(A_i)$ pour tout $i \in I$, ce qui s'écrit, I étant fini,

$$\lim_n \uparrow \int_E a\varphi \mathbb{1}_{E_n} \, d\mu = \sum_{i \in I} a\alpha_i \mu(A_i) = a \int_E \varphi \, d\mu.$$

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

$$\int_E a\varphi \mathbb{1}_{E_n} \, d\mu \leq \int_E f_n \, d\mu \leq \lim_n \uparrow \int_E f_n \, d\mu.$$

En se servant des deux équations qui précédent et notamment en passant à la limite dans la dernière inégalité, on trouve

$$a \int_E \varphi \, d\mu \leq \lim_n \uparrow \int_E f_n \, d\mu.$$

L'inégalité cherchée est donc prouvée, car a est arbitrairement proche de 1. □

Proposition 7.14 (Lemme de Fatou) *Pour toute suite (f_n) de $\mathscr{F}_+(\mathscr{A})$, nous savons que $\liminf_n f_n \in \mathscr{F}_+(\mathscr{A})$, mais surtout*

$$\int_E \liminf_n f_n \, d\mu \leq \liminf_n \int_E f_n \, d\mu.$$

Remarque 7.15 *Pour $f_n = \mathbb{1}_{A_n}$ où $A_n \in \mathscr{A}$, le lemme de Fatou se traduit par une inégalité que nous connaissions déjà* @

$$\mu\left(\liminf_n A_n\right) \leq \liminf_n \mu(A_n).$$

Dém. Soit $g_n := \inf_{k\geq n} f_k$ et $g := \liminf_n f_n = \lim_n \uparrow g_n$. Comme g est la limite de la suite croissante (g_n), le théorème de Beppo Levi assure que

$$\int_E g\, d\mu = \lim_n \uparrow \int_E g_n\, d\mu.$$

D'autre part, $g_n \leq f_n$ donc par croissance de l'intégrale, $\int_E g_n\, d\mu \leq \int_E f_n\, d\mu$ et

$$\liminf_n \int_E g_n\, d\mu \leq \liminf_n \int_E f_n\, d\mu.$$

Mais d'après ce qui précède, $\liminf_n \int_E g_n\, d\mu = \lim_n \int_E g_n\, d\mu = \int_E g\, d\mu$, ce qui fournit l'inégalité souhaitée. □

Proposition 7.16 *Les propriétés de positive homogénéité et d'additivité passent (comme celle de croissance) aux intégrales de fonctions mesurables positives. En d'autres termes pour tout $a \geq 0$ et pour tout $f \in \mathscr{F}_+(\mathscr{A})$,*

$$\int_E (af)\, d\mu = a \int_E f\, d\mu \quad \text{et} \quad \int_E (f+g)\, d\mu = \int_E f\, d\mu + \int_E g\, d\mu.$$

Dém. Comme f et g sont mesurables et positives, il existe d'après le lemme fondamental d'approximation des suites croissantes (f_n) et (g_n) de fonctions étagées positives croissant vers f et g respectivement. La proposition se prouve en écrivant les propriétés de positive homogénéité et d'additivité pour ces fonctions étagées et en appliquant à chaque suite d'intégrales le théorème de Beppo Levi. □ @

Proposition 7.17 *Pour toute suite (f_n) de $\mathscr{F}_+(\mathscr{A})$, nous savons que $\sum_n f_n \in \mathscr{F}_+(\mathscr{A})$, mais surtout*

$$\int_E \left(\sum_n f_n \right) d\mu = \sum_n \int_E f_n\, d\mu.$$

Dém. Il suffit de poser $g_n := \sum_{k=0}^n f_k$, d'utiliser l'additivité de l'intégrale et d'appliquer le théorème de Beppo Levi à la suite croissante (g_n). □ @

Corollaire 7.18 *Pour tout $f \in \mathscr{F}_+(\mathscr{A})$, l'application*

$$\begin{aligned} \nu : \mathscr{A} &\longrightarrow [0, +\infty] \\ A &\longmapsto \int_E f \mathbb{1}_A\, d\mu \end{aligned}$$

est une mesure sur $(E, \mathscr{A})$ appelée mesure de densité f par rapport à μ.

Notation 7.19 *On note souvent $\int_A f\, d\mu$ à la place*[2] *de $\int_E f \mathbb{1}_A\, d\mu$.*

2. ce qui d'ailleurs est cohérent avec le cas trivial $A = E$...

Dém. Vérifions les deux propriétés caractérisant les mesures. Tout d'abord $\nu(\varnothing) = 0$ car $f\mathbb{1}_\varnothing$ est la fonction étagée nulle partout. Montrons à présent que ν est σ-additive. Soient (A_n) une suite d'éléments de $\mathscr{A}$ deux à deux disjoints. D'après la proposition qui précède le corollaire, nous pouvons échanger sommation et intégrale de sorte que

$$\nu\left(\cup_n A_n\right) = \int_E f\mathbb{1}_{\cup_n A_n}\,d\mu = \int_E f\sum_n \mathbb{1}_{A_n}\,d\mu$$
$$= \int_E \sum_n f\mathbb{1}_{A_n}\,d\mu = \sum_n \int_E f\mathbb{1}_{A_n}\,d\mu = \sum_n \nu(A_n),$$

ce qui achève la démonstration. □

Proposition 7.20 *Si μ est* finie *alors pour tout $f \in \mathscr{F}_+(\mathscr{A})$,*

$$f \textit{ bornée} \implies f \textit{ intégrable.}$$

Dém. Si f est bornée, il existe un nombre réel positif a tel que $f \leq a\mathbb{1}_E$, donc $\int_E f\,d\mu \leq a\mu(E) < \infty$, car par hypothèse μ est finie. □

Proposition 7.21 (Inégalité de Markov) *Pour tout $f \in \mathscr{F}_+(\mathscr{A})$, pour tout $a > 0$,*

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

Dém. Comme $f \geq a\mathbb{1}_{\{f\geq a\}}$, par croissance $\int_E f\,d\mu \geq a\mu(\{f \geq a\})$. □

Proposition 7.22 *Pour tout $f \in \mathscr{F}_+(\mathscr{A})$,*

$$\int_E f\,d\mu = 0 \Longleftrightarrow \mu(\{f \neq 0\}) = 0.$$

Dém. Pour le sens $\Rightarrow$, soit $A_n := \{f \geq 1/n\}$. Par l'inégalité de Markov, $\mu(A_n) \leq n\int_{A_n} f\,d\mu$, donc comme $\int_E f\,d\mu = 0$, $\mu(A_n) = 0$. Or $A := \{f \neq 0\} = \lim_n \uparrow A_n$, donc par continuité à gauche de μ, $\mu(A) = \lim_n \uparrow \mu(A_n) = 0$. Traitons maintenant le sens $\Leftarrow$. Par additivité

$$\int_E f\,d\mu = \int_A f\,d\mu + \int_{{}^cA} f\,d\mu = \int_A f\,d\mu,$$

car f est nulle sur ${}^cA$, donc si $\mu(A) = 0$ on a bien $\int_E f\,d\mu = 0$.

On pouvait aussi procéder de la manière suivante, qui est un peu moins simple, mais peut servir d'entraînement. On commence par montrer la proposition pour une fonction étagée positive φ :

$$\int_E \varphi\,d\mu = 0 \Leftrightarrow \sum_{i\in I}\alpha_i\mu(A_i) = 0 \Leftrightarrow \sum_{i:\alpha_i\neq 0}\alpha_i\mu(A_i) = 0$$
$$\Leftrightarrow \sum_{i:\alpha_i\neq 0}\mu(A_i) = 0 \Leftrightarrow \mu\left(\cup_{i:\alpha_i\neq 0}A_i\right) = 0 \Leftrightarrow \mu(\{\varphi \neq 0\}) = 0.$$

Maintenant soit $f \in \mathscr{F}_+(\mathscr{A})$ et (f_n) une suite croissante d'éléments de $\mathscr{E}_+(\mathscr{A})$ convergeant vers f. Montrons la double implication en commençant par $\Rightarrow$. Supposons donc que $\int_E f\,d\mu = 0$. Pour tout entier n, comme $0 \leq f_n \leq f$, on a également $\int_E f_n\,d\mu = 0$. D'après ce qui précède, on a donc $\mu(\{f_n \neq 0\}) = 0$. Si nous montrons que la suite $\{f_n \neq 0\}$ converge en croissant vers $\{f \neq 0\}$, alors la continuité à gauche de μ impliquera $\mu(\{f \neq 0\}) = 0$. La croissance de cette suite de parties est due à la croissance de la suite (f_n), en effet si $f_n(x) \neq 0$ alors $f_{n+1}(x) \neq 0$ car $f_{n+1}(x) \geq f_n(x)$. Montrons que la limite A de cette suite, qui n'est autre que sa réunion, est $\{f \neq 0\}$ par double inclusion. Si $x \in A$, il existe n tel que $f_n(x) \neq 0$, mais comme $f(x) \geq f_n(x)$, on a aussi $f(x) \neq 0$. Réciproquement, si $f(x) \neq 0$ alors comme $f(x)$ est la limite de la suite réelle $(f_n(x))_n$, à partir d'un certain rang $f_n(x) \neq 0$.

Montrons enfin l'inégalité opposée. Si $\mu(\{f \neq 0\}) = 0$, alors comme $\{f_n \neq 0\} \subseteq \{f \neq 0\}$ pour tout n, on a aussi $\mu(\{f_n \neq 0\}) = 0$, ce qui implique $\int_E f_n\,d\mu = 0$. Le théorème de Beppo Levi permet de conclure que $\int_E f\,d\mu = 0$. □

Remarque 7.23 *Noter que le raisonnement qui permet de montrer que la suite $\{f_n \neq 0\}$ converge vers $\{f \neq 0\}$ ne tient plus si la suite (f_n) converge vers f sans croître, en effet avec $f_n(x) = x^n$ sur l'intervalle $[0,1]$: on a $\{f_n \neq 0\} =]0,1]$ pour tout n, donc la limite de cette suite est $]0,1]$, alors que si f désigne la limite de la suite (f_n) alors $\{f \neq 0\} = \{1\}$.*

Notation 7.24 (importante) *Au lieu d'écrire $\mu(\{f \neq 0\}) = 0$, on notera souvent*

$$f = 0 \quad \mu\text{-presque partout} \qquad \text{ou} \qquad f = 0 \quad \mu\text{-p.p.}$$

De manière générale, soit $N \in \mathscr{A}$ tel que $\mu(N) = 0$, et une certaine propriété $P(x)$ qui dépend de $x \in E$. Si $\{x \in E : P(x) \text{ est fausse}\} \subseteq N$, on dira que $P(x)$ est vraie « pour μ-presque tout x », ou « $\mu(dx)$-presque partout », ou que P est vraie μ-p.p.

L'ensemble N est appelé ensemble négligeable*, ou μ-négligeable. Les ensembles dénombrables, l'ensemble triadique de Cantor, sont des ensembles λ-négligeables.*

Dans certains contextes, une partie de E sera dite négligeable même si elle n'est pas mesurable mais si elle est incluse dans une partie mesurable de mesure nulle.

Proposition 7.25 *Pour tous $f, g \in \mathscr{F}_+(\mathscr{A})$,*

$$f = g \quad \mu\text{-p.p.} \Rightarrow \int_E f\,d\mu = \int_E g\,d\mu.$$

Dém. On pose

$$h := \begin{cases} \max(f,g) - \min(f,g) & \text{sur } \{\min(f,g) < \infty\} \\ 0 & \text{sur } \{f = g = \infty\}. \end{cases}$$

Comme $\{f = g\} = \{h = 0\}$, par passage au complémentaire $\{h \neq 0\} = \{f \neq g\}$ donc $\mu(\{h \neq 0\}) = 0$ (ce qui s'écrit aussi $h = 0$ μ-p.p.), par conséquent $\int_E h\,d\mu = 0$ (par la Proposition 7.22). Mais comme $\max(f,g) = \min(f,g) + h$, par additivité on a

$$\int_E \max(f,g)\,d\mu = \int_E \min(f,g)\,d\mu + \int_E h\,d\mu = \int_E \min(f,g)\,d\mu.$$

Et comme $f \wedge g \leq f \leq f \vee g$ et $f \wedge g \leq g \leq f \vee g$, par croissance on a

$$\int_E \max(f,g)\,d\mu = \int_E \min(f,g)\,d\mu = \int_E f\,d\mu = \int_E g\,d\mu,$$

ce qui achève la démonstration. □

Proposition 7.26 *Pour tout $f \in \mathscr{F}_+(\mathscr{A})$,*

$$\int_E f\,d\mu < +\infty \Longrightarrow \mu(\{f = +\infty\}) = 0.$$

Dém. Soit $A := \{f = +\infty\}$. Par contraposée, si $\mu(A) \neq 0$, alors $\int_E f\,d\mu \geq \int_A f\,d\mu = (+\infty)\mu(A) = +\infty$.

Autre possibilité (pour l'entraînement...) : se servir de l'inégalité de Markov. Soit $A_n := \{f \geq n\}$, alors $A = \lim_n \downarrow A_n$. Or par l'inégalité de Markov, $\mu(A_1) \leq \int_E f\,d\mu < +\infty$. Or comme toute mesure, μ est continue pour les suites décroissantes (on dit aussi continue à droite) dont un des termes est de mesure finie, donc $\mu(A) = \lim_n \downarrow \mu(A_n)$. Mais par l'inégalité de Markov à nouveau, $\mu(A_n) \leq n^{-1} \int_E f\,d\mu \longrightarrow 0$ quand $n \to \infty$. □

Corollaire 7.27 (Lemme de Borel–Cantelli) *Soit (A_n) une suite d'éléments de $\mathscr{A}$. Alors*

$$\sum_{n \geq 0} \mu(A_n) < +\infty \Longrightarrow \mu\left(\limsup_n A_n\right) = 0.$$

Dém. Soit $f := \sum_n \mathbb{1}_{A_n}$. Par hypothèse, $\sum_n \mu(A_n) = \int_E f\,d\mu < \infty$. Donc par le résultat précédent, $\mu(\{f = +\infty\}) = 0$. Montrons que $\limsup_n A_n = \{f = +\infty\}$.

$$\begin{aligned}\sum_n \mathbb{1}_{A_n}(x) < \infty &\Leftrightarrow \mathbb{1}_{A_n}(x) = 0 \text{ à partir d'un certain rang } n_0 \\ &\Leftrightarrow \exists n_0,\ \forall n \geq n_0,\quad x \in {}^c\!A_n \Leftrightarrow x \in \liminf_n {}^c\!A_n \Leftrightarrow x \notin \limsup_n A_n.\end{aligned}$$

Exemple 7.28 (mesure de comptage) *L'intégration par rapport à la mesure de comptage sur $\mathbb{N}$ est tout simplement la* sommation de série*. En effet $u \in \mathscr{F}_+(\mathscr{P}(\mathbb{N}))$ est tout simplement une suite (u_n) de réels positifs et pour tout $N \in \mathbb{N}$, la suite $\varphi_N := (u_n \mathbb{1}_{n \leq N})$ est une fonction étagée positive qui converge en croissant vers u. Donc si m désigne la mesure de comptage sur $(\mathbb{N}, \mathscr{P}(\mathbb{N}))$, alors*

$$\int_{\mathbb{N}} u\,dm = \lim_{N\to\infty} \int_{\mathbb{N}} \varphi_N\,dm = \lim_{N\to\infty} \sum_{n=0}^{N} u_n\,m(\{n\}) = \lim_{N\to\infty} \sum_{n=0}^{N} u_n = \sum_n u_n.$$

Chapitre 8

Intégrale des fonctions mesurables de signe quelconque et l'espace $\mathscr{L}^1(\mu)$

8.1 Intégrale des fonctions mesurables de signe quelconque

On rappelle que $\mathscr{F}(\mathscr{A})$ désigne l'ensemble des fonctions mesurables : $(E, \mathscr{A}) \to (\bar{\mathbb{R}}, \mathscr{B}or(\bar{\mathbb{R}}))$.

Définition 8.1 *Pour tout $f \in \mathscr{F}(\mathscr{A})$, on dira que f* admet une intégrale *si $\int_E f^+ \, d\mu < \infty$ ou $\int_E f^- \, d\mu < \infty$, et l'on définit alors l'intégrale de f par rapport à μ, notée*[1] *$\int_E f \, d\mu$, l'élément de $\bar{\mathbb{R}}$ suivant*

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

Si les deux intégrales $\int_E f^+ \, d\mu$ et $\int_E f^- \, d\mu$ sont finies, autrement dit $\int_E |\, f \,| \, d\mu < \infty$, ou encore dit $|\, f \,|$ est μ-intégrable, alors on dira (un peu) plus simplement que f est μ-intégrable. *En particulier,*

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

Dém. de l'inégalité. Par définition,

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

par additivité. De même, on démontre que $-\int_E f \, d\mu \leq \int_E |\, f \,| \, d\mu$. □

Proposition 8.2 *Soient $g, h \in \mathscr{F}_+(\mathscr{A})$ tels que $\int_E g \, d\mu < \infty$ ou $\int_E h \, d\mu < \infty$. Alors* $\min(g, h) < \infty$ *μ-p.p. et $f := (g - h) \mathbb{1}_{\{\min(g,h) < \infty\}}$ qui est bien définie (et vaut d'ailleurs $g - h$ μ-p.p.) admet une intégrale*

$$\int_E f \, d\mu = \int_E g \, d\mu - \int_E h \, d\mu.$$

1. encore !

Dém. Supposons que $\int_E h\, d\mu < \infty$ (la démonstration est identique si c'est $\int_E g\, d\mu$ qui est finie), alors $h < \infty$ μ-p.p. par la Proposition 7.26 (donc $\min(g, h) < \infty$ μ-p.p.). Ensuite par définition de f, on a $f^- \leq h$, donc f^- est μ-intégrable. Enfin, comme @
$f^+ + h = f^- + g$, par additivité

$$\int_E f^+\, d\mu + \int_E h\, d\mu = \int_E f^-\, d\mu + \int_E g\, d\mu,$$

égalité à laquelle on peut retrancher les deux intégrales de f^- et de h, qui sont finies, ce qui donne

$$\int_E f^+\, d\mu - \int_E f^-\, d\mu = \int_E g\, d\mu - \int_E h\, d\mu,$$

qui est bien défini dans $]-\infty, +\infty]$, ce qui implique que f admet bien une intégrale égale à $\int_E g\, d\mu - \int_E h\, d\mu$. □

Définition 8.3 (et proposition) *L'espace $\mathscr{L}^1(\mathscr{E}, \mathscr{A}, \mu)$, noté aussi $\mathscr{L}^1(\mu)$, des fonctions μ-intégrables, est un espace vectoriel et l'application $f \mapsto \int_E f\, d\mu$ est une forme linéaire positive donc croissante.*

Remarque 8.4 *Bien noter $\mathscr{L}^1$ car la notation L^1 fera plus tard référence à un autre espace.*

Dém. Pour tous $f, g \in \mathscr{L}^1(\mu)$, pour tout $\lambda \in \mathbb{R}$, $|\,\lambda f + g\,| \leq |\,\lambda\,|.|\,f\,| + |\,g\,|$ donc $\lambda f + g$ est intégrable par additivité. De plus, en se servant des égalités du type $f = f^+ - f^-$ appliquées à f, g et $f + g$, on obtient $f + g = (f+g)^+ - (f+g)^- = f^+ - f^- + g^+ - g^-$ et par conséquent

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

d'où en passant aux intégrales,

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

Comme toutes ces quantités sont finies, on peut les retrancher à loisir ce qui donne

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

autrement dit $\int_E (f+g)\, d\mu = \int_E f\, d\mu + \int_E g\, d\mu$. De même, en utilisant les égalités $(\lambda f)^+ = \lambda f^+$, $(\lambda f)^- = \lambda f^-$ lorsque $\lambda > 0$ et $(\lambda f)^+ = -\lambda f^-$, $(\lambda f)^- = -\lambda f^+$ lorsque $\lambda < 0$, et en utilisant la positive homogénéité de l'intégrale sur $\mathscr{F}_+$, on obtient :

a) dans le cas $\lambda > 0$,

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

b) dans le cas $\lambda < 0$,

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

L'intégrale est donc bien une forme linéaire. Elle est positive car si $f \geq 0$, alors $f = f^+$ et par définition $\int_E f\,d\mu = \int_E f^+\,d\mu \geq 0$. Elle est donc croissante car si $g \geq h$, alors grâce à la Proposition 8.2, $(g-h)\mathbb{1}_{\{\min(g,h)<\infty\}}$ est positive et son intégrale, qui est positive d'après ce qui précède, vaut $\int_E g\,d\mu - \int_E h\,d\mu \geq 0$, ce qui montre que $\int_E g\,d\mu \geq \int_E h\,d\mu$. □

Remarque 8.5 *Si m est la mesure de comptage sur $\mathbb{N}$ alors $\mathscr{L}^1(m)$ est l'ensemble, souvent noté ℓ^1, des suites dont la série est* absolument convergente.

Lemme 8.6 *Si $f = g$ μ-p.p., alors f est intégrable (resp. admet une intégrale) ssi g est intégrable (resp. admet une intégrale).*

Dém. Comme $f = g$ μ-p.p., on voit facilement que $f^+ = g^+$ μ-p.p. et que $f^- = g^-$ μ-p.p. Il suffit donc de montrer que pour tous f, g *positives*, si $f = g$ μ-p.p. alors $\int f\,d\mu < \infty$ ssi $\int g\,d\mu < \infty$. En fait on a déjà montré mieux (Corollaire 7.25) : si f et g sont positives et que $f = g$ μ-p.p. alors on a l'égalité $\int f\,d\mu = \int g\,d\mu$ dans $\bar{\mathbb{R}}_+$. □ @

8.2 Les grands théorèmes de convergence

Nous allons maintenant donner une version plus générale du théorème de convergence monotone et de ses corollaires, auquel nous ajouterons un dernier théorème de convergence, dit de convergence dominée.

Théorème 8.7 *Soit $g \in \mathscr{L}^1(\mu)$ et (f_n) une suite d'éléments de $\mathscr{F}(\mathscr{A})$ telle que pour tout entier n, $f_{n+1} \geq f_n \geq g$ μ-p.p. Alors $\lim_n \uparrow f_n$ est définie μ-p.p. et toute fonction f μ-p.p. égale[2] à $\lim_n f_n$ admet une intégrale et vérifie*

$$\int_E f\,d\mu = \lim \uparrow \int_E f_n\,d\mu.$$

Corollaire 8.8 *Soit (φ_n) une suite d'éléments de $\mathscr{F}(\mathscr{A})$ telle que pour tout entier n, $\varphi_n \geq 0$ μ-p.p. Alors $\sum_n \varphi_n$ admet une intégrale et*

$$\int_E \left(\sum_n \varphi_n\right) d\mu = \sum_n \int_E \varphi_n\,d\mu.$$

Dém. du corollaire. On prend bien sûr $f_n = \sum_{k=0}^n \varphi_k$ qui est positive μ-p.p. car positive sur $\cap_k\{\varphi_k \geq 0\}$, partie mesurable dont le complémentaire est de mesure inférieure ou égale à $\sum_k \mu(\{\varphi_k < 0\}) = 0$. En particulier la fonction $\sum_n \varphi_n$ est définie μ-p.p. et vaut $\sum_n \varphi_n^+$ μ-p.p. donc admet une intégrale. Le corollaire découle du théorème qui précède et de l'additivité de l'intégrale □ @

2. comme par exemple $f = \liminf_n f_n$ si l'on veut

Dém. du théorème. On prend d'abord $g \equiv 0$ et on définit $A := \cap_{n\geq 0}\{f_{n+1} \geq f_n \geq g\}$. Alors A est mesurable, avec

$$\mu({}^cA) = \mu\left(\bigcup_{n\geq 0} {}^c\{f_{n+1} \geq f_n \geq g\}\right) \leq \sum_{n\geq 0} \mu\left({}^c\{f_{n+1} \geq f_n \geq g\}\right) = 0,$$

et $(f_n \mathbb{1}_A)$ est une suite croissante de fonctions mesurables positives. Soit h sa limite, et f telle que $h = f$ μ-p.p. Alors d'après le théorème de convergence monotone,

$$\int_E f\, d\mu = \int_E h\, d\mu = \lim_n \uparrow \int_E f_n \mathbb{1}_A\, d\mu = \lim_n \uparrow \int_E f_n\, d\mu,$$

puisque $\mu({}^cA) = 0$ (en particulier f_n admet bien une intégrale, car $f_n = f_n^+$ μ-p.p.). Maintenant si $g \not\equiv 0$, on applique bien sûr ce qu'on vient d'obtenir à $g_n := f_n - g \geq 0$ μ-p.p. Comme $f_n = g + g_n = g^+ - g^- + g_n^+ - g_n^- = (g^+ + g_n^+) - (g_n^- + g^-)$, et que $g_n^- = 0$ μ-p.p., on a $\int (g_n^- + g^-)\, d\mu < \infty$, donc f_n admet une intégrale. De plus,

$$\int_E f_n\, d\mu = \int_E g^+\, d\mu + \int_E g_n^+\, d\mu - \int_E g^-\, d\mu = \int_E g\, d\mu + \int_E g_n\, d\mu,$$

qui converge en croissant vers $\int_E g\, d\mu + \int_E (\lim_n \uparrow g_n)\, d\mu$. Or $g_n = f_n - g$, donc $\lim_n \uparrow g_n = (\lim_n f_n) - g$, et comme $g \in \mathscr{L}^1$, la suite $(\int_E f_n\, d\mu)$ converge en croissant vers $\int_E g\, d\mu + \int_E ((\lim_n f_n) - g)\, d\mu = \int_E (\lim_n f_n)\, d\mu$. □

Donnons une version plus générale du lemme de Fatou.

Proposition 8.9 *Soit $g \in \mathscr{L}^1(\mu)$ et (f_n) une suite d'éléments de $\mathscr{F}(\mathscr{A})$. Alors*
a) $f_n \geq g$ μ-p.p. pour tout $n \Longrightarrow \int (\liminf_n f_n)\, d\mu \leq \liminf_n \int f_n\, d\mu$
et
b) $f_n \leq g$ μ-p.p. pour tout $n \Longrightarrow \int (\limsup_n f_n)\, d\mu \geq \limsup_n \int f_n\, d\mu$.

Dém. Remarquons que a)$\Rightarrow$ b) en remplaçant f_n par $-f_n$ et en multipliant les deux membres de l'inégalité par -1. Montrons a) avec $g \equiv 0$. Il faut bien sûr vérifier que f_n admet une intégrale pour tout n et que $\liminf_n f_n$ également admet une intégrale.

Si $f_n \geq 0$ μ-p.p. alors $f_n^- = 0$ μ-p.p. et $\int_n f_n^-\, d\mu = 0$ donc f_n admet une intégrale. Soit alors $A := \cup_n \{f_n < 0\}$, alors A (est mesurable et) $\mu(A) \leq \sum_n \mu(\{f_n < 0\}) = 0$, donc

$$\liminf_n f_n = \mathbb{1}_A \liminf f_n + \liminf (f_n \mathbb{1}_{{}^cA}),$$

où dans le membre de droite, le premier terme vaut 0 μ-p.p. et le deuxième terme est positif, donc $\liminf_n f_n$ admet une intégrale, et d'après la première version du lemme de Fatou,

$$\int_E \left(\liminf_n f_n \mathbb{1}_{{}^cA}\right)\, d\mu \leq \liminf_n \left(\int_E f_n \mathbb{1}_{{}^cA}\, d\mu\right).$$

Or comme $\mu(A) = 0$, on a $\int_E (\mathbb{1}_A \liminf_n f_n)\, d\mu = 0$ et $\int_E f_n \mathbb{1}_A\, d\mu = 0$. En ajoutant ces deux dernières quantités (nulles !) à l'inégalité précédente, on obtient l'inégalité voulue. @

Si $g \not\equiv 0$, on applique bien sûr ce que l'on vient de faire à $g_n := f_n - g$ qui est positive μ-p.p. Comme dans la démonstration du théorème précédent, on voit que f_n admet une intégrale et que $\int_E g_n\, d\mu = \int_E f_n\, d\mu - \int_E g\, d\mu$, ainsi que $\int_E(\liminf_n g_n)\, d\mu = \int_E(\liminf_n f_n)\, d\mu - \int_E g\, d\mu$, ce qui donne le résultat.

On aurait également pu appliquer à notre deuxième version du théorème de convergence monotone, la méthode utilisée pour déduire la première version lemme de Fatou de la première version du théorème de convergence monotone... □ @

Nous pouvons maintenant énoncer un autre des résultats majeurs de ce cours.

Théorème 8.10 (Théorème de Lebesgue, ou de convergence dominée) *Si (f_n) est une suite d'éléments de $\mathscr{L}^1(\mu)$ convergeant*[3] *μ-p.p. vers une fonction f et qu'il existe $g \in \mathscr{L}^1(\mu)$ telle que pour tout $n \in \mathbb{N}$, $|\, f_n \,| \leq g$ μ-p.p., alors $f \in \mathscr{L}^1(\mu)$ et*

a) $\displaystyle\lim_{n\to\infty} \int_E |\, f - f_n \,|\, d\mu = 0$

et

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

Dém. Soit $A := \{x \in E : \lim_n f_n(x) = f(x)\} \cap \cap_n \{|\, f_n \,| \leq g\}$. Alors A est de complémentaire négligeable et pour tout $x \in A$, $|\, f_n(x) \,| \leq g(x)$, donc $|\, f(x) \,| \leq g(x)$. Par conséquent, ayant $|\, f \,| \leq g$ μ-p.p., nous avons $\int |\, f \,|\, d\mu \leq \int g\, d\mu$, donc $\int |\, f \,|\, d\mu < \infty$.

Comme $f_n \geq -g$ μ-p.p. et que $-g \in \mathscr{L}^1(\mu)$, d'après le lemme de Fatou,

$$\int_E f\, d\mu = \int_E (\liminf_n f_n)\, d\mu \leq \liminf_n \int_E f_n\, d\mu.$$

De même, comme $f_n \leq g$ μ-p.p. et que $g \in \mathscr{L}^1(\mu)$, d'après le lemme de Fatou,

$$\int_E f\, d\mu = \int_E (\limsup_n f_n)\, d\mu \geq \limsup_n \int_E f_n\, d\mu.$$

On a donc

$$\limsup_n \int_E f_n\, d\mu \leq \int_E f\, d\mu \leq \liminf_n \int_E f_n\, d\mu,$$

ce qui implique $\int_E f\, d\mu = \lim_n \int_E f_n\, d\mu$, autrement dit b). Pour obtenir a), on applique le lemme de Fatou à $|\, f - f_n \,| \leq |\, f \,| + |\, f_n \,| \leq 2g \in \mathscr{L}^1(\mu)$. Donc

$$\int_E \left(\limsup_n |\, f - f_n \,|\right) d\mu \geq \limsup_n \int_E |\, f - f_n \,|\, d\mu,$$

ce qui implique $\limsup_n \int_E |\, f - f_n \,|\, d\mu = 0$, autrement dit a). □

3. simplement

8.3 Intégrale des fonctions à valeurs complexes

Définition 8.11 *Une fonction $f : (E, \mathscr{A}) \to (\mathbb{C}, \mathscr{B}or(\mathbb{C}))$ est dite μ-intégrable si f est mesurable et $|f|$ est intégrable. Ceci entraîne que $\Re(f)$ et $\Im(f)$ sont intégrables et l'intégrale de f par rapport à μ est définie comme le nombre complexe*

$$\int_E f\, d\mu = \int_E \Re(f)\, d\mu + i \int_E \Im(f)\, d\mu.$$

Théorème 8.12 *L'ensemble $\mathscr{L}^1_{\mathbb{C}}(E, \mathscr{A}, \mu)$ des fonctions complexes μ-intégrables est un espace vectoriel sur $\mathbb{C}$. L'application $f \mapsto \int_E f\, d\mu$ est une $\mathbb{C}$-forme linéaire et pour tout $f \in \mathscr{L}^1_{\mathbb{C}}(\mu)$,*

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

Il suffit de montrer la dernière inégalité, car le reste découle facilement de la linéarité de l'intégrale réelle et de la définition.

Il existe $z \in \mathbb{C}$ tel que $z \int f\, d\mu$ est réel et $|z| = 1$, donc

$$\left| \int f\, d\mu \right| = |z| \left| \int f\, d\mu \right| = \left| z \int f\, d\mu \right|.$$

Or par linéarité de l'intégrale complexe,

$$z \int f\, d\mu = \int (zf)\, d\mu = \int_E \Re(zf)\, d\mu + i \int_E \Im(zf)\, d\mu.$$

Or on a choisi z pour que $z \int f\, d\mu$ soit réel donc $\int_E \Im(zf)\, d\mu = 0$. De plus, $|\Re(zf)| \leq |zf| = |f|$, donc comme

$$\left| \int_E \Re(zf)\, d\mu \right| \leq \int_E |\Re(zf)|\, d\mu \leq \int_E |f| d\mu,$$

on a

$$\left| \int_E f\, d\mu \right| = \left| z \int_E f\, d\mu \right| = \left| \int_E \Re(zf)\, d\mu \right| \leq \int_E |f| d\mu,$$

ce qui constitue l'inégalité souhaitée. □

Proposition 8.13 *Soit (φ_n) une suite de fonctions mesurables à valeurs dans $\bar{\mathbb{R}}$ ou $\mathbb{C}$. Si la série de terme général $\int_E |\varphi_n| d\mu$ est convergente, alors la fonction $\sum_n |\varphi_n|$ est intégrable ainsi que la fonction définie μ-p.p. $\sum_n \varphi_n$ et*

$$\int_E \left(\sum_n \varphi_n \right) d\mu = \sum_n \left(\int_E \varphi_n\, d\mu \right).$$

Dém. Soit $g := \sum_n |\varphi_n| < \infty$ μ-p.p. car $\int_E g\,d\mu < \infty$ par le corollaire du théorème de convergence monotone sur les séries. Donc il existe une partie mesurable A de E, de complémentaire μ-négligeable, telle que $g(x) < \infty$ pour tout $x \in A$. Autrement dit, pour tout $x \in A$, la série de terme général $\varphi_n(x)$ est absolument convergente, donc convergente. Si $f_n := \sum_{k \leq n} \varphi_k$, on a donc la convergence sur A de la suite (f_n) vers une fonction $f := \sum_n \varphi_n$, assortie de la domination des $|f_n|$ par g, donc par le théorème de convergence dominée,

$$\int_E f\,d\mu = \lim_{n\to\infty} \int_E f_n\,d\mu = \lim_{n\to\infty} \sum_{k=0}^{n} \int_E \varphi_k\,d\mu = \sum_n \int_E \varphi_n\,d\mu,$$

ce qui achève la démonstration. $\square$

Chapitre 9

Applications

9.1 Intégrale de Lebesgue et intégrale de Riemann

Définition 9.1 *Une fonction $f : [a,b] \to \mathbb{R}$ est dite Riemann-intégrable si pour tout $\varepsilon > 0$ il existe une fonction ϕ_ε en escalier sur $[a,b]$ et une fonction ψ_ε* positive *et en escalier sur $[a,b]$ telles que*

$$|f - \phi_\varepsilon| \leq \psi_\varepsilon \qquad \text{et} \qquad \int_a^b \psi_\varepsilon \leq \varepsilon, \tag{9.1}$$

*où l'*intégrale de Riemann *d'une fonction en escalier φ, notée $\int_a^b \varphi$, est définie*[1] *comme $\sum_i \alpha_i(a_i - a_{i-1})$ si φ vaut α_i sur $]a_{i-1}, a_i[$.*

De plus, pour tous ϕ_ε, ψ_ε vérifiant (9.1), la limite lorsque $\varepsilon \to 0$ de $\int_a^b \phi_\varepsilon$ est toujours la même, et est égale, par définition, au nombre réel noté[2] *$\int_a^b f$ et appelé*[3] *intégrale de Riemann de f.*

Remarque 9.2 *Toute fonction Riemann-intégrable est bornée (car ϕ_ε et ψ_ε le sont).*

Remarque 9.3 *Toute fonction réglée est Riemann-intégrable, car limite uniforme de fonctions en escalier : il suffit de prendre $\psi_\varepsilon = \varepsilon/(b-a)$.*

Théorème 9.4 *Pour tout f Riemann-intégrable sur $[a,b]$, il existe $g \in \mathscr{L}^1([a,b], \mathscr{B}or([a,b]), \lambda)$ tel que*

a) $f = g$ μ-p.p.

b) $\displaystyle\int_a^b f = \int_{[a,b]} g \, d\lambda$.

Remarque 9.5 *D'un point de vue pratique, si f est borélienne, alors on peut prendre $g = f$.*

1. cette définition ne dépend bien sûr pas de la subdivision choisie $a < a_1 < \cdots < b$ pour représenter φ
2. aussi...
3. encore...

Dém. Pour tout n il existe ϕ_n, ψ_n en escalier telles que $|f-\phi_n| \leq \psi_n$ et $\int_a^b \psi_n \leq 1/n$. Rappelons déjà que par définition,

$$\int_a^b f = \lim_{n\to\infty} \int_a^b \phi_n.$$

Soit alors

$$\alpha_n := \phi_n - \psi_n \qquad \text{et} \qquad \beta_n := \phi_n + \psi_n.$$

Comme $|f-\phi_n| \leq \psi_n$, on a $\phi_n - f \leq \psi_n$ et $f - \phi_n \leq \psi_n$, c'est-à-dire $\alpha_n \leq f \leq \beta_n$. D'autre part, comme $\beta_n - \alpha_n = 2\psi_n$, on a

$$\lim_n \int_a^b (\beta_n - \alpha_n) = 0,$$

et donc

$$\int_a^b f = \lim_n \int_a^b \alpha_n = \lim_n \int_a^b \beta_n.$$

Soient à présent

$$\tilde{\alpha}_n := \max(\alpha_1, \ldots, \alpha_n) \qquad \text{et} \qquad \tilde{\beta}_n := \min(\beta_1, \ldots, \beta_n).$$

On obtient alors l'encadrement

$$\alpha_n \leq \tilde{\alpha}_n \leq f \leq \tilde{\beta}_n \leq \beta_n.$$

On définit encore

$$\tilde{\alpha} := \lim_n \uparrow \alpha_n \qquad \text{et} \qquad \tilde{\beta} := \lim_n \downarrow \beta_n,$$

ce qui donne

$$\tilde{\alpha} \leq f \leq \tilde{\beta}.$$

De plus, comme une fonction en escalier est étagée, pour tout n, ϕ_n et ψ_n sont étagées donc boréliennes, ainsi que α_n, β_n, puis $\tilde{\alpha}_n$, $\tilde{\beta}_n$ par stabilité de la mesurabilité par passage à la borne supérieure ou inférieure, et enfin $\tilde{\alpha}$, $\tilde{\beta}$ sont boréliennes par stabilité de la mesurabilité par passage à la limite.

Il est clair par la définition donnée plus haut que pour les fonctions en escalier (donc étagées), intégrales de Riemann et de Lebesgue (i.e., par rapport à la mesure de Lebesgue) coïncident, donc pour tout n

$$\int_a^b \alpha_n = \int_{[a,b]} \alpha_n \, d\lambda \qquad \text{et} \qquad \int_a^b \beta_n = \int_{[a,b]} \beta_n \, d\lambda,$$

d'où

$$\int_a^b \alpha_n = \int_{[a,b]} \alpha_n \, d\lambda \leq \int_{[a,b]} \tilde{\alpha}_n \, d\lambda \leq \int_{[a,b]} \tilde{\alpha} \, d\lambda \leq \int_{[a,b]} \tilde{\beta} \, d\lambda \leq \int_{[a,b]} \tilde{\beta}_n \, d\lambda \leq \int_{[a,b]} \beta_n \, d\lambda = \int_a^b \beta_n.$$

En passant à la limite, comme $\int_a^b \alpha_n$ et $\int_a^b \beta_n$ convergent toutes deux vers $\int_a^b f$, on en déduit que

$$\int_{[a,b]} \tilde{\alpha}\, d\lambda = \int_{[a,b]} \tilde{\beta}\, d\lambda = \int_a^b f.$$

Remarquons que $\alpha_1 \leq \tilde{\alpha} \leq \tilde{\beta} \leq \beta_1$, si bien que $\tilde{\alpha}$ et $\tilde{\beta}$ sont bornées. Ainsi ces deux fonctions sont λ-intégrables sur $[a,b]$, la fonction $\gamma := \tilde{\beta} - \tilde{\alpha}$ est bien définie, et

$$\int_{[a,b]} \gamma\, d\lambda = \int_{[a,b]} \tilde{\alpha}\, d\lambda - \int_{[a,b]} \tilde{\beta}\, d\lambda = 0.$$

Comme $\gamma \geq 0$, ceci implique que $\gamma = 0$ λ-p.p. Mais $\tilde{\alpha} \leq f \leq \tilde{\beta}$, donc $\tilde{\alpha} = f$ sur $\{\gamma = 0\}$ et si $g := f\mathbb{1}_{\{\gamma=0\}}$, alors $g = \tilde{\alpha}\mathbb{1}_{\{\gamma=0\}}$ et

$$\int_{[a,b]} g\, d\lambda = \int_{[a,b]} \tilde{\alpha}\mathbb{1}_{\{\gamma=0\}}\, d\lambda = \int_{[a,b]} \tilde{\alpha}\, d\lambda = \int_a^b f,$$

ce qui donne bien l'égalité entre l'intégrale de Riemann de f et l'intégrale de Lebesgue d'une certaine fonction g égale à f λ-p.p. □

9.2 Dérivées et primitives

Proposition 9.6 *Soit $a \in \mathbb{R}$ et $f : [a, +\infty[\to \mathbb{R}$ borélienne. Si f est λ-localement intégrable, au sens où pour tout $b > a$, $f\mathbb{1}_{[a,b]} \in \mathscr{L}^1(\lambda)$ alors la fonction $F : x \mapsto \int_{[a,x]} f\, d\lambda$ est* continue*.*

Dém. Soit $x \geq a$ et une suite (x_n) convergeant vers x en croissant, et telle que $x_n \neq x$ pour tout n. Alors les fonctions $f\mathbb{1}_{[a,x_n]}$ convergent vers $f\mathbb{1}_{[a,x[}$ tout en étant dominées par $|\, f\,|\mathbb{1}_{[a,x]}$ qui est λ-intégrable par hypothèse. Donc par convergence dominée,

$$\lim_n F(x_n) = \lim_n \int f\mathbb{1}_{[a,x_n]}\, d\lambda = \int f\mathbb{1}_{[a,x[}\, d\lambda = \int f\mathbb{1}_{[a,x]}\, d\lambda = F(x).$$

Ceci prouve que F est continue à gauche. La démonstration est identique lorsque (x_n) est décroissante, ce qui prouve que F est aussi continue à droite. □ @

Théorème 9.7 *Soit $[a,b]$ un intervalle compact de $\mathbb{R}$ et $f : [a,b] \to \mathbb{R}$ une fonction dérivable de dérivée f'* bornée*. Alors f' est mesurable et l'intégrale de f' par rapport à λ est égale à sa primitive f au sens où*

$$\int_{[a,b]} f'\, d\lambda = f(b) - f(a).$$

Dém. Soit

$$g_n(x) := \begin{cases} n\left(f\left(x+\frac{1}{n}\right) - f(x)\right) & \text{si} \quad x \in [a, b-1/n] \\ 0 & \text{si} \quad x \in]b-1/n, b]. \end{cases}$$

Alors pour tout $x \in [a, b[$, $\lim_n g_n(x) = f'(x)$, ce qui montre que $f' \mathbb{1}_{[a,b[}$ est mesurable comme limite de fonctions mesurables, et donc $f' \mathbb{1}_{[a,b]}$ est mesurable. Par l'inégalité des accroissements finis, pour tous $n \in \mathbb{N}$ et $x \in [a, b]$, $|\, g_n(x) \,| \leq M := \sup_{[a,b]} |\, f' \,|$ qui est fini par hypothèse. Or $M \mathbb{1}_{[a,b]} \in \mathscr{L}^1(\lambda)$ donc par convergence dominée,

$$\lim_{n\to\infty} \int_{[a,b]} g_n \, d\lambda = \int_{[a,b[} f' \, d\lambda = \int_{[a,b]} f' \, d\lambda.$$

Montrons à présent que l'on a également

$$\lim_{n\to\infty} \int_{[a,b]} g_n \, d\lambda = f(b) - f(a).$$

Dans les égalités suivantes, nous utilisons la linéarité des intégrales de Lebesgue et de Riemann, ainsi que l'égalité entre ces intégrales due à la continuité de f :

$$\begin{aligned} \int_{[a,b]} g_n(x) \, d\lambda(x) &= n \int_{[a,b-1/n]} \left(f\left(x + \frac{1}{n}\right) - f(x) \right) d\lambda(x) \\ &= n \left(\int_a^{b-1/n} f\left(\cdot + \frac{1}{n}\right) - \int_a^{b-1/n} f \right) \\ &= n \left(\int_{a+1/n}^{b} f - \int_a^{b-1/n} f \right) \\ &= n \left(\int_{b-1/n}^{b} f - \int_a^{a+1/n} f \right) \\ &= n \int_{[b-1/n,b]} f \, d\lambda - n \int_{[a,a+1/n]} f \, d\lambda. \end{aligned}$$

Ces passages entre intégrale de Riemann et de Lebesgue servent uniquement à justifier le changement de variable (translation, troisième égalité) que nous ne connaissons pas (encore) dans le cas de l'intégrale de Lebesgue.

Il ne reste plus qu'à montrer que le second terme de la dernière différence tend vers $f(a)$ quand $n \to \infty$, et la même méthode s'appliquera pour montrer que le premier terme tend vers $f(b)$. Soit $\alpha_n := \inf_{[a,a+1/n]} f$ et $\beta_n := \sup_{[a,a+1/n]} f$. Alors

$$\alpha_n = n \int_{[a,a+1/n]} \alpha_n \, d\lambda \leq n \int_{[a,a+1/n]} f \, d\lambda \leq n \int_{[a,a+1/n]} \beta_n \, d\lambda = \beta_n,$$

mais comme f est continue, $\lim_n \alpha_n = \lim_n \beta_n = f(a)$, ce qui achève la démonstration. □

Remarque 9.8 *Le théorème précédent serait faux si l'on ne faisait pas l'hypothèse que f' est bornée, comme on peut le voir sur le contre-exemple suivant :*

$$f_n(x) := \left(\frac{3}{2}\right)^n \int_{[0,x]} \mathbb{1}_{A_n}\, d\lambda \qquad x \geq 0,$$

où A_n est le n-ième élément de la suite qui converge vers l'ensemble triadique de Cantor. En effet, la limite f de la suite (f_n), appelée fonction de Lebesgue*, ou « escalier du diable », est continue, dérivable λ-p.p. avec pour dérivée $f' = 0$ λ-p.p. (mais non bornée), donc $\int_{[0,1]} f'\, d\lambda = 0$. Pourtant $f(1) - f(0) = 1 - 0 = 1$.*

9.3 Intégrales dépendant d'un paramètre

Théorème 9.9 *Soit $f : I \times E \to \mathbb{R}$, où I est un intervalle de $\mathbb{R}$ tel que pour tout $t \in I$, $f(t, \cdot) : (E, \mathscr{A}) \to (\mathbb{R}, \mathscr{B}(\mathbb{R})$ est mesurable. Alors*

a) S'il existe $t_0 \in I$ tel que

$$\text{pour } \mu\text{-presque tout } x,\ t \mapsto f(t,x) \text{ est continue en } t_0,$$

et s'il existe $g : (E, \mathscr{A}) \to (\mathbb{R}, \mathscr{B}(\mathbb{R})$ intégrable *telle que*

$$\text{pour } \mu\text{-presque tout } x, \text{ pour tout } t \in I, \quad |\, f(t,x)\, | \leq g(x),$$

alors la fonction $h : t \mapsto \int_E f(t,x)\, d\mu(x)$ est bien définie et elle est continue *en t_0.*

b) Si pour tout $t \in I$, $f(t, \cdot)$ est intégrable[4], si

$$\text{pour } \mu\text{-presque tout } x,\ t \mapsto f(t,x) \text{ est dérivable sur tout l'intervalle } I,$$

et s'il existe $g_1 : (E, \mathscr{A}) \to (\mathbb{R}, \mathscr{B}(\mathbb{R})$ intégrable *telle que*

$$\text{pour } \mu\text{-presque tout } x, \text{ pour tout } t \in I, \quad \left| \frac{\partial f}{\partial t}(t,x) \right| \leq g_1(x),$$

alors la fonction h est bien définie et est dérivable sur tout I, de dérivée

$$h'(t) = \int_E \frac{\partial f}{\partial t}(t,x)\, d\mu(x) \qquad t \in I.$$

Remarque 9.10 *Les hypothèses* commençant *par « pour μ-presque tout x » peuvent toutes (sauf une, voir plus bas) être* affaiblies *en échangeant les quantificateurs. Pour voir la différence, à titre d'exemple, l'assertion*

$$\text{pour } \mu\text{-presque tout } x, \text{ pour tout } t \in I, \quad |\, f(t,x)\, | \leq g(x),$$

signifie qu'il existe un élément A de la tribu $\mathscr{A}$ de complémentaire négligeable, tel que pour tout $x \in A$, pour tout $t \in I$, $|\, f(t,x)\, | \leq g(x)$. D'autre part, l'assertion

$$\text{pour tout } t \in I, \text{ pour } \mu\text{-presque tout } x, \quad |\, f(t,x)\, | \leq g(x),$$

4. pour que h soit bien définie

signifie que pour tout $t \in I$ il existe un élément A_t de la tribu $\mathscr{A}$, dépendant de t, *de complémentaire négligeable, tel que pour tout $x \in A_t$, $|\, f(t,x)\,| \leq g(x)$. Si I était dénombrable les deux assertions seraient équivalentes, quitte à définir A, dans la première assertion, comme l'intersection de tous les ensembles A_t de la seconde assertion. Ici I est un intervalle, donc la seconde assertion peut inclure strictement la première. Lorsque nous n'utilisons dans la démonstration que des* suites *à valeurs dans I, cette distinction est sans conséquence. En revanche, la domination des dérivées partielles doit être énoncée telle quelle, pour avoir une inégalité des accroissements finis vraie p.p.*

Dém. a) Pour tout $t \in I$, la domination p.p. de $f(t,\cdot)$ par la fonction intégrable g garantit que $f(t,\cdot)$ est intégrable et donc que h est bien définie. Soit une suite (s_n) de I de limite t_0. Montrons que $\lim_n h(s_n) = h(t_0)$. Soit $f_n(x) := f(s_n, x)$. De la continuité pour p.t. x de la fonction $f(\cdot, x)$ en t_0, on déduit la convergence p.p. de f_n vers $f(t_0,\cdot)$. Comme la suite (f_n) est dominée p.p. par la fonction intégrable g, le théorème de convergence dominée assure que

$$\lim_n h(s_n) = \lim_n \int_E f_n \, d\mu = \int_E f(t_0, x)\, d\mu(x) = h(t_0),$$

ce qui est la limite souhaitée.

b) Soit $t \in I$ et (s_n) une suite de I convergeant vers t telle que $s_n \neq t$ pour tout n. Soit $f_n : (E, \mathscr{A}) \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ la fonction intégrable suivante

$$f_n(x) := \frac{f(s_n, x) - f(t,x)}{s_n - t} \qquad x \in E.$$

Alors pour p.t. x, la suite (f_n) converge vers $\partial f/\partial t(t,x)$, ce qui fait de cette dérivée partielle une fonction mesurable de x. L'hypothèse de domination p.p. de cette fonction mesurable par la fonction intégrable g_1 garantit que $\int_E \partial f/\partial t(t,x)\, d\mu(x)$ est bien définie pour tout $t \in I$. De plus, par l'intégrabilité de f_n

$$\int_E f_n \, d\mu = \frac{1}{s_n - t}\left(\int_E f(s_n, \cdot)\, d\mu - \int_E f(t, \cdot)\, d\mu\right) = \frac{h(s_n) - h(t)}{s_n - t}.$$

Or par l'inégalité des accroissements finis, pour μ-presque tout x,

$$|\, f_n(x)\,| \leq \sup_{s \in I} \left| \frac{\partial f}{\partial t}(s,x) \right| \leq g_1(x),$$

et comme g_1 est intégrable, on obtient

$$\lim_n \frac{h(s_n) - h(t)}{s_n - t} = \lim_n \int_E f_n \, d\mu = \int_E \frac{\partial f}{\partial t}(t,x)\, d\mu(x)$$

par le théorème de convergence dominée. □

9.4 Applications

9.4.1 Dérivation sous la somme

Soit (u_n) une suite de fonctions dérivables sur un intervalle I de $\mathbb{R}$ telle que
(i) pour tout $t \in I$, $\sum_n |\, u_n(t) \,|$ converge ;
(ii) pour tout $t \in I$, $|\, u_n'(t) \,| \leq w_n$ pour une suite (w_n) telle que $\sum_n w_n < \infty$.
Alors $S(t) := \sum_n u_n(t)$ est bien définie et est dérivable en tout $t \in I$, avec $S'(t) = \sum_n u_n'(t)$.

9.4.2 Convolution

Soit $f \in \mathscr{L}^1(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$ et φ dérivable de dérivée bornée. Alors la fonction $f \star \varphi$ définie par

$$f \star \varphi(t) := \int_{\mathbb{R}} \varphi(t-x)\, f(x)\, d\lambda(x) \qquad t \in R,$$

est bien définie et dérivable sur $\mathbb{R}$ de dérivée

$$(f \star \varphi)'(t) = \int_{\mathbb{R}} \varphi(t-x)\, f(x)\, d\lambda(x) = f \star \varphi'(t).$$

9.4.3 Transformée de Fourier

Soit $f : \mathbb{R} \to \mathbb{R}$ λ-intégrable et $F(t) := \int_{\mathbb{R}} e^{itx}\, f(x)\, d\lambda(x)$. Si $x \mapsto xf(x)$ est intégrable, alors F est dérivable sur $\mathbb{R}$ et

$$F'(t) = i \int_{\mathbb{R}} e^{itx}\, xf(x)\, d\lambda(x).$$

Chapitre 10

Inégalités et espaces $\mathscr{L}^p$

10.1 Inégalité de Jensen

Proposition 10.1 *Si μ est une* probabilité *sur $(E, \mathscr{A})$ et $f \in \mathscr{L}^1(E, \mathscr{A}, \mu)$ est à valeurs dans un intervalle I de $\mathbb{R}$, alors pour toute fonction* convexe $\varphi : I \to \mathbb{R}$, *on a $\int_E f\, d\mu \in I$, $\int_E \varphi(f)\, d\mu$ existe dans $]-\infty, +\infty]$ et*

$$\int_E \varphi(f)\, d\mu \geq \varphi\left(\int_E f\, d\mu\right).$$

Dém. Soient $a := \inf(I)$ et $b := \sup(I)$. Ayant $-\infty \leq a \leq f \leq b \leq +\infty$, comme μ est une probabilité,

$$a = \int_E a\, d\mu \leq \int_E f\, d\mu \leq \int_E b\, d\mu = b,$$

donc $m := \int_E f\, d\mu \in I$. Comme φ est convexe, il existe au moins une droite située en-dessous du graphe de φ et passant par $(m, \varphi(m))$, d'équation $y = \alpha(x - m) + \varphi(m)$. Ceci se traduit par

$$\varphi(u) \geq \alpha(u - m) + \varphi(m) \qquad u \in I,$$

et donc pour tout $x \in E$,

$$\varphi \circ f(x) \geq \alpha(f(x) - m) + \varphi(m).$$

La fonction $\varphi \circ f$ étant minorée par une fonction intégrable, elle admet une intégrale (qui ne peut être égale à $-\infty$) et

$$\int_E \varphi \circ f\, d\mu \geq \alpha \int_E (f - m)\, d\mu + \int_E \varphi(m)\, d\mu = 0 + \varphi(m),$$

par linéarité, et parce que μ est une probabilité. $\square$

10.2 Inégalités de Hölder et de Minkowski

10.2.1 Semi-normes $\mathscr{L}^p$, $p \in [1, +\infty]$

Définition 10.2 ($p \in [1, +\infty[$) *On note $\mathscr{L}^p(E, \mathscr{A}, \mu)$, ou $\mathscr{L}^p(\mu)$, l'ensemble de toutes les fonctions mesurables à valeurs dans $\bar{\mathbb{R}}$ telles que $| f |^p$ est μ-intégrable. Pour tout $f \in \mathscr{L}^p(\mu)$, on pose*

$$\| f \|_p := \left(\int_E | f |^p \, d\mu \right)^{\frac{1}{p}},$$

qu'on appelle[1] norme $\mathscr{L}^p$ *de f.*

Définition 10.3 ($p = \infty$) *On note $\mathscr{L}^\infty(E, \mathscr{A}, \mu)$, ou $\mathscr{L}^\infty(\mu)$, l'ensemble de toutes les fonctions mesurables f à valeurs dans $\bar{\mathbb{R}}$ qui sont μ-*essentiellement bornées*, c'est-à-dire telles qu'il existe $a > 0$ pour lequel $\mu(\{| f | \geq a\}) = 0$. Si $f \in \mathscr{L}^\infty(\mu)$, on pose*

$$\| f \|_\infty := \inf\{a > 0 : \mu(\{| f | \geq a\}) = 0\} \in [0, +\infty[,$$

qu'on appelle[2] supremum essentiel *de f.*

10.2.2 Inégalité de Hölder

Proposition 10.4 (Inégalité de Hölder) *Pour tous $p, q \in [1, +\infty]$ tels que $\dfrac{1}{p} + \dfrac{1}{q} = 1$ –on dit que ces nombres sont* conjugués[3]*– alors si $f \in \mathscr{L}^p$ et $g \in \mathscr{L}^q$, $fg \in \mathscr{L}^1$ et*

$$\| fg \|_1 \leq \| f \|_p \, \| g \|_q.$$

Dém. Premier cas : p ou q vaut $+\infty$. Supposons $q = +\infty$, et donc $p = 1$. Alors par définition du supremum essentiel, $g \leq \| g \|_\infty$ μ-p.p., donc $| fg | \leq | f | \, \| g \|_\infty$ μ-p.p. Par conséquent

$$\| fg \|_1 = \int_E | fg | \, d\mu \leq \| g \|_\infty \int_E | f | \, d\mu = \| g \|_\infty \, \| f \|_1.$$

Deuxième cas : p et q sont finis. L'inégalité de Hölder est évidente si $\| f \|_p = 0$ ou $\| g \|_q = 0$. En effet si par exemple $\| f \|_p = 0$ alors $| f | = 0$ μ-p.p. donc $| fg | = 0$ μ-p.p. et par conséquent $\| fg \|_1 = 0$. Nous supposons donc maintenant que $\| f \|_p \neq 0$ et $\| g \|_q \neq 0$.

Rappelons la concavité du logarithme :

$$\forall \alpha \in [0, 1], \ \forall x, y > 0, \quad \ln(\alpha x + (1 - \alpha) y) \geq \alpha \ln(x) + (1 - \alpha) \ln(y).$$

En posant $\alpha = 1/p$, et donc $1 - \alpha = 1/q$, ainsi que $x = | a |^p$ et $y = | b |^q$, on obtient

$$\ln\left(\frac{| a |^p}{p} + \frac{| b |^q}{q} \right) \geq \frac{1}{p} \ln(| a |^p) + \frac{1}{q} \ln(| b |^q) = \ln(| ab |),$$

1. mais dont nous verrons qu'il ne s'agit en fait que d'une semi-norme
2. mais on devrait dire μ-supremum essentiel
3. typiquement $p = q = 2$ ou $p = 1$ et $q = +\infty$

et donc par croissance de l'exponentielle,

$$\frac{|\,a\,|^p}{p} + \frac{|\,b\,|^q}{q} \geq |\,ab\,|.$$

Ainsi avec

$$a := \frac{|\,f(x)\,|}{\|\,f\,\|_p} \quad \text{et} \quad b := \frac{|\,g(x)\,|}{\|\,g\,\|_q},$$

on obtient

$$\frac{|\,fg(x)\,|}{\|\,f\,\|_p\ \|\,g\,\|_q} \leq \frac{1}{p}\ \frac{|\,f(x)\,|^p}{\|\,f\,\|_p^p} + \frac{1}{q}\ \frac{|\,g(x)\,|^q}{\|\,g\,\|_q^q}.$$

En intégrant chaque membre, la croissance de l'intégrale implique

$$\frac{\|\,fg\,\|_1}{\|\,f\,\|_p\ \|\,g\,\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1,$$

ce qui est l'inégalité attendue. □

10.2.3 Inégalité de Minkowski

Proposition 10.5 *Pour tout $p \in [1, +\infty]$, si $f, g \in \mathscr{L}^p$, alors $f + g \in \mathscr{L}^p$ et*

$$\|\,f + g\,\|_p \leq \|\,f\,\|_p + \|\,g\,\|_p.$$

Dém. Premier cas : $p = 1$. Par l'inégalité triangulaire $|\,f + g\,| \leq |\,f\,| + |\,g\,|$, donc $f + g \in \mathscr{L}^1$ et par croissance de l'intégrale,

$$\|\,f + g\,\|_1 \leq \int_E |\,f\,|\,d\mu + \int_E |\,g\,|\,d\mu = \|\,f\,\|_1 + \|\,g\,\|_1.$$

Deuxième cas : $p = +\infty$. Par définition du supremum essentiel, $|\,f\,| \leq \|\,f\,\|_\infty$ et $|\,g\,| \leq \|\,g\,\|_\infty$ μ-p.p. donc par l'inégalité triangulaire, $|\,f + g\,| \leq \|\,f\,\|_\infty + \|\,g\,\|_\infty$ μ-p.p. Mais par définition du supremum essentiel à nouveau, pour tout $A \geq 0$ tel que $\mu\left(\{|\,f + g\,| > A\}\right) = 0$, on a $\|\,f + g\,\|_\infty \leq A$, donc ici $\|\,f + g\,\|_\infty \leq \|\,f\,\|_\infty + \|\,g\,\|_\infty$.

Troisième cas : $p \in]1, +\infty[$. Notons $h := |\,f + g\,|$. Alors $h \in \mathscr{L}^p$ car pour tous $x, y \geq 0$,

$$(x + y)^p \leq 2^p\big(\max(x, y)\big)^p = 2^p \max\left(x^p, y^p\right) \leq 2^p\left(x^p + y^p\right),$$

ce qui implique que $\int h^p\,d\mu \leq 2^p\big(\int |\,f\,|^p\,d\mu + \int |\,g\,|^p\,d\mu\big) < \infty$. Soit q le nombre conjugué de p. Comme $h \in \mathscr{L}^p$, $h^{p-1} \in \mathscr{L}^q$ puisque $q(p - 1) = p$. Donc grâce à l'inégalité de Hölder appliquée à $h^{p-1} \in \mathscr{L}^q$ et $f \in \mathscr{L}^p$,

$$\int_E |\,f\,|\,h^{p-1}\,d\mu = \|\,fh^{p-1}\,\|_1 \leq \|\,f\,\|_p\ \|\,h^{p-1}\,\|_q.$$

De même,

$$\int_E |\,g\,|\,h^{p-1}\,d\mu \leq \|\,g\,\|_p\ \|\,h^{p-1}\,\|_q.$$

Or
$$h^p = |\, f + g\,| h^{p-1} \leq |\, f\,| h^{p-1} + |\, g\,| h^{p-1},$$
donc
$$\int_E h^p \, d\mu \leq \|\, f\,\|_p \;\|\, h^{p-1}\,\|_q + \|\, g\,\|_p \;\|\, h^{p-1}\,\|_q = (\|\, f\,\|_p + \|\, g\,\|_p) \left(\int_E h^{q(p-1)} \, d\mu \right)^{\frac{1}{q}},$$
autrement dit
$$\|\, h\,\|_p^p \leq (\|\, f\,\|_p + \|\, g\,\|_p) \,\|\, h\,\|_p^{\frac{p}{q}}.$$
On peut supposer que $\|\, h\,\|_p \neq 0$, faute de quoi l'inégalité de Minkowski devient triviale, donc en divisant chaque membre par $\|\, h\,\|_p^{p/q}$, on obtient
$$\|\, h\,\|_p \leq \|\, f\,\|_p + \|\, g\,\|_p,$$
qui est l'inégalité souhaitée. □

10.3 Espace $\mathscr{L}^p$ et espace L^p

On rappelle la notion de norme sur un espace vectoriel F.

Définition 10.6 *Une fonction $N : F \to \mathbb{R}_+$ est appelée* norme *si*
(i) pour tout $v \in F$, $N(v) = 0$ ssi $v = 0$
(ii) [homogénéité] pour tous $a \in \mathbb{R}$ et $v \in F$, $N(av) = |\, a\,|\; N(v)$
(iii) [inégalité triangulaire] pour tous $u, v \in F$, $N(u + v) \leq N(u) + N(v)$.
Si (i) est remplacé par la propriété plus faible (i') : $N(0) = 0$, alors N est appelée une semi-norme.

Remarque 10.7 *On rappelle qu'une norme N sur un e.v. F induit une topologie sur F, la topologie relative à la distance $d(u, v) := N(u - v)$.*

Proposition 10.8 *Pour tout $p \in [1, +\infty]$, $(\mathscr{L}^p, \|\cdot\|_p)$ est un espace vectoriel semi-normé.*

Dém. Soient $f, g \in \mathscr{L}^p$ et $a \in \mathbb{R}^\star$.
(i') Si $f = 0$, alors $|\, f\,|^p = 0$, dont l'intégrale est nulle, donc $\|\, f\,\|_p = 0$.
(ii) évident par linéarité de l'intégrale si $p < \infty$. Si $p = +\infty$, @

$$\begin{aligned}\|\, af\,\|_\infty &= \inf\{m > 0 : \mu(\{|\, af\,| \geq m\}) = 0\} \\ &= |\, a\,| \inf\{m' > 0 : \mu(\{|\, af\,| \geq |\, a\,| m'\}) = 0\} \\ &= |\, a\,| \inf\{m' > 0 : \mu(\{|\, f\,| \geq m'\}) = 0\} = |\, a\,| \,\|\, f\,\|_\infty.\end{aligned}$$

(iii) Inégalité de Minkowski. □

Si N est une semi-norme sur un e.v. F, il existe un procédé classique consistant à modifier F pour que N devienne une norme, et qui consiste à identifier u et $v \in F$ dès que $N(u-v)=0$. Rigoureusement, il s'agit de définir l'*espace quotient* de F par la relation d'équivalence

$$u \sim v \Longleftrightarrow N(u-v)=0,$$

c'est-à-dire l'ensemble constitué des classes d'équivalence de $\sim$. Ici, quel que soit $p \in [1,+\infty]$, pour tous $f,g \in \mathscr{L}^p$,

$$\| f-g \|_p = 0 \Leftrightarrow \mu(\{f \neq g\}) = 0 \Leftrightarrow f = g \;\; \mu\text{-p.p.}$$

La relation d'équivalence associée à chaque semi-norme $\| \cdot \|_p$ ne dépend donc pas de p et est la même pour tous les espaces $\mathscr{L}^p$:

$$f \sim g \Longleftrightarrow f = g \;\; \mu\text{-p.p.}$$

Définition 10.9 *Pour $p \in [1,+\infty]$, on note $L^p(E,\mathscr{A},\mu)$, ou $L^p(\mu)$, l'ensemble des classes d'équivalence des éléments de $\mathscr{L}^p(\mu)$ par la relation d'équivalence définie par l'égalité μ-p.p.*

Soit $\overline{f} := \{g : g = f \ \mu\text{-p.p.}\}$ la classe d'équivalence de f. Les opérations classiques s'étendent aux classes d'équivalence, avec $\overline{af} = a\overline{f}$ et $\overline{f+g} = \overline{f} + \overline{g}$.

On peut également définir $\| \cdot \|_p$ sur $L^p(\mu)$ par $\| \overline{f} \|_p = \| f \|_p$, qui ne dépend pas du représentant choisi, car $f = g$ μ-p.p. implique $\| f \|_p = \| g \|_p$.

Remarque 10.10 *On fera systématiquement l'abus de notation qui consiste à ne pas différencier fonctions et classes d'équivalences, c'est-à-dire à utiliser le même symbole pour une fonction f et pour sa classe d'équivalence $\overline{f}$.*

On a alors immédiatement ce que l'on cherchait.

Théorème 10.11 *L'ensemble $(L^p(\mu), \| \cdot \|_p)$ est un espace vectoriel normé.*

Exemple 10.12 *On note ℓ^p l'espace $\mathscr{L}^p(\mathbb{N},\mathscr{P}(\mathbb{N}),m)$, où m est la mesure de comptage. Soit $u \in \ell^p$. Si $p < \infty$, alors*

$$\| u \|_p = \left(\sum_n | u_n |^p \right)^{\frac{1}{p}},$$

tandis que si $p = +\infty$,

$$\| u \|_\infty = \sup_n | u_n |.$$

Il n'est pas besoin ici de quotienter $\mathscr{L}^p$ car $\| u \|_p = 0$ implique $u = 0$.

Proposition 10.13 *a) pour tous $p \leq q$, on a l'inclusion $\ell^p \subseteq \ell^q$.*

b) Si μ est finie, *alors pour tous $p \leq q$, on a l'inclusion $L^p(\mu) \supseteq L^q(\mu)$.*

Dém. En exercice. $\square$ @

Remarque 10.14 *Si μ n'est pas finie, il n'y a pas d'inclusion générale, comme en atteste le contre-exemple suivant avec $(E, \mathscr{A}, \mu) = (\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$: avec*

$$f(x) = \frac{1}{\sqrt{x}} \mathbb{1}_{]0,1]}(x) \quad \text{et} \quad g(x) = \frac{1}{\sqrt{x^2+1}},$$

on voit facilement que f est dans $\mathscr{L}^1$ mais pas dans $\mathscr{L}^2$, et que g est dans $\mathscr{L}^2$ mais pas dans $\mathscr{L}^1$.

Deuxième partie

LM365 – Intégration 2

Chapitre 11

Construction d'une mesure : existence et unicité

11.1 Quelques rappels et nouvelles définitions

11.1.1 Rappels

Définition 11.1 *Une classe $\mathscr{A}$ de parties d'un ensemble E est appelée* tribu *ou σ-*algèbre *si*

(i) elle contient E : $E \in \mathscr{A}$;

(ii) elle est stable par passage au complémentaire *: pour tout $A \subseteq E$, $A \in \mathscr{A} \Leftrightarrow {}^{c}A \in \mathscr{A}$;*

(iii) elle est stable par réunion dénombrable *: si (A_n) est une famille dénombrable d'éléments de $\mathscr{A}$, alors $\cup_n A_n \in \mathscr{A}$.*

Définition 11.2 *Une mesure sur l'espace mesurable $(E, \mathscr{A})$ est une application $\mu : \mathscr{A} \to [0, +\infty]$ qui :*

(i) associe la valeur 0 à l'ensemble vide : $\mu(\varnothing) = 0$;

*(ii) est σ-*additive *: pour toute suite (A_n) d'éléments de $\mathscr{A}$ deux à deux disjoints,*

$$\mu(\cup_n A_n) = \sum_n \mu(A_n).$$

11.1.2 Définitions utiles dans le cadre de l'unicité des mesures

Définition 11.3 *Une classe Λ de parties d'un ensemble E est appelée* λ-système *si*

(i) elle contient $\varnothing$: $\varnothing \in \Lambda$;

(ii) elle est stable par différence propre *: pour tous $A, B \in \Lambda$, $A \subseteq B \Rightarrow B \setminus A \in \Lambda$;*

(iii) elle est stable par réunion dénombrable croissante *: si (A_n) est une suite* croissante *d'éléments de Λ, alors $\cup_n A_n \in \Lambda$.*

Remarque 11.4 *Dans la définition précédente de λ-système, la propriété (i) est seulement là pour rappeler que Λ est non vide car c'est une conséquence de (ii).*

Remarque 11.5 *Une tribu est un λ-système, et donc en particulier $\mathscr{P}(E)$ en est un.*

Définition 11.6 *Un λ-système qui contient E est appelé* classe monotone. *Une définition de classe monotone peut donc être obtenue par la modification suivante de la définition de λ-système : en changeant (i) pour (i') : $E \in \Lambda$.*

Proposition 11.7 *a) L'intersection d'une collection quelconque non vide de λ-systèmes est un λ-système.*

b) Pour toute classe $\mathscr{C}$ de parties de E, l'intersection[1] *de tous les λ-systèmes contenant tous les éléments de $\mathscr{C}$ est donc un λ-système, noté $\Lambda(\mathscr{C})$, et appelé λ-système engendré par $\mathscr{C}$ ou plus petit λ-système contenant $\mathscr{C}$.*

Remarque 11.8 *On rappelle que l'on définit de la même manière la tribu $\sigma(\mathscr{C})$ engendrée par $\mathscr{C}$. La démonstration de la proposition précédente est laissée en exercice.* @

Proposition 11.9 *Si Λ est une classe monotone stable par intersections finies, alors Λ est une tribu.*

Dém. Vérifions une par une les trois propriétés caractéristiques des tribus.

(i) $E \in \Lambda$ puisque Λ est une classe monotone.

(ii) Comme $E \in \Lambda$, pour tout $A \in \Lambda$, le complémentaire de A est la différence propre $E \setminus A$, donc ${}^cA \in \Lambda$.

(iii) Soir (A_n) une suite d'éléments de Λ. Pour tout entier n, soit $B_n := \cup_{k=0}^n A_k$. Comme $\cup_n A_n = \cup_n B_n$ et que (B_n) est une suite croissante, la propriété (iii) des classes monotones implique qu'il suffit de montrer que $B_n \in \Lambda$ pour tout n. Autrement dit, il suffit de montrer que Λ est stable par réunions finies. Or on se souvient que Λ est stable par passage au complémentaire et, par hypothèse, stable par intersections finies, donc pour tous $A, B \in \Lambda$, $A \cup B = {}^c({}^cA \cap {}^cB) \in \Lambda$, ce qui achève la démonstration. $\square$

Définition 11.10 *Une classe $\mathscr{C}$ de parties de E est appelée* π-système *si*

(i) elle contient E : $E \in \mathscr{C}$;

(ii) elle est stable par intersections finies *: pour tous $A, B \in \mathscr{C}$, $A \cap B \in \mathscr{C}$.*

Remarque 11.11 *Avec cette définition, la proposition qui précède peut s'énoncer ainsi : si $\mathscr{A}$ est à la fois un λ-système et un π-système, alors $\mathscr{A}$ est une tribu.*

Remarque 11.12 *Dans $\mathbb{R}^d$, l'ensemble des rectangles (produits cartésiens d'intervalles), l'ensemble des rectangles ouverts (produits cartésiens d'intervalles ouverts) forment chacun un π-système.* @

---

1. non vide puisque $\mathscr{P}(E)$ est un λ-système

11.1.3 Définitions utiles dans le cadre de l'existence des mesures

Définition 11.13 *Une classe $\mathscr{B}$ de parties d'un ensemble E est appelée* algèbre *ou* algèbre de Boole *si*

(i) elle contient E : $E \in \mathscr{B}$;

(ii) elle est stable par passage au complémentaire *: pour tout $A \subseteq E$, $A \in \mathscr{B} \Leftrightarrow {}^{c}A \in \mathscr{B}$;*

(iii) elle est stable par réunions finies *: pour tous $A, B \in \mathscr{B}$, $A \cup B \in \mathscr{B}$.*

Remarque 11.14 *Une tribu est donc une algèbre de Boole stable par réunion dénombrable, d'où le nom de σ-algèbre.*

Remarque 11.15 *Dans $\mathbb{R}^d$, l'ensemble des réunions finies de rectangles forment une algèbre, ainsi que l'ensemble des réunions finies de rectangles disjoints.* @

11.2 Unicité d'une mesure

11.2.1 Théorème de la classe monotone et corollaires

Théorème 11.16 (Théorème de la classe monotone) *Pour tout π-système $\mathscr{C}$, $\Lambda(\mathscr{C}) = \sigma(\mathscr{C})$. En particulier, le plus petit λ-système $\Lambda(\mathscr{C})$ contenant $\mathscr{C}$ est donc une tribu.*

Dém. Supposons que $\Lambda(\mathscr{C})$ est une tribu et montrons qu'alors $\Lambda(\mathscr{C}) = \sigma(\mathscr{C})$. De manière générale, comme $\sigma(\mathscr{C})$ est une tribu contenant $\mathscr{C}$, c'est un λ-système contenant $\mathscr{C}$, donc on a toujours $\Lambda(\mathscr{C}) \subseteq \sigma(\mathscr{C})$ (puisque $\Lambda(\mathscr{C})$ est le plus petit λ-système contenant $\mathscr{C}$). De plus, d'après l'assertion qui précède, $\Lambda(\mathscr{C})$ est une tribu contenant $\mathscr{C}$, donc $\sigma(\mathscr{C}) \subseteq \Lambda(\mathscr{C})$ (puisque $\sigma(\mathscr{C})$ est la plus petite tribu contenant $\mathscr{C}$).

Montrons à présent que $\Lambda(\mathscr{C})$ est une tribu. D'après la proposition qui précède, il suffit de montrer que $\Lambda(\mathscr{C})$ contient E et est stable par intersections finies. Il est évident que $E \in \Lambda(\mathscr{C})$ car $E \in \mathscr{C}$ ($\mathscr{C}$ est un π-système) et $\mathscr{C} \subseteq \Lambda(\mathscr{C})$. En particulier $\Lambda(\mathscr{C})$ est une classe monotone. Montrons donc que $\Lambda(\mathscr{C})$ est stable par intersections finies, en utilisant le fait que $\mathscr{C}$ l'est.

Pour tout $C \subseteq E$ fixé, on définit

$$\Lambda_C := \{A \in \Lambda(\mathscr{C}) : A \cap C \in \Lambda(\mathscr{C})\}.$$

Nous allons montrer que Λ_C est toujours un λ-système.

(i) comme $A := \varnothing \in \Lambda(\mathscr{C})$ (qui est une classe monotone) et que $A \cap C = \varnothing \in \Lambda(\mathscr{C})$ (toujours...), on a bien que $\varnothing \in \Lambda_C$.

(ii) Pour tous $A, B \in \Lambda_C$ tels que $A \subseteq B$, on a $A, B \in \Lambda(\mathscr{C})$ avec $A \cap C$ et $B \cap C$ éléments de $\Lambda(\mathscr{C})$. Par stabilité par différence propre de $\Lambda(\mathscr{C})$, comme $A \cap C \subseteq B \cap C$, on a $(B \cap C) \setminus (A \cap C) \in \Lambda(\mathscr{C})$, mais $(B \cap C) \setminus (A \cap C) = (B \setminus A) \cap C$, donc $B \setminus A \in \Lambda_C$.

(iii) Soit (A_n) une suite croissante d'éléments de Λ_C, de sorte que pour tout entier n, $A_n \in \Lambda(\mathscr{C})$ et $A_n \cap C \in \Lambda(\mathscr{C})$. Or $(A_n \cap C)$ est croissante, donc comme $\Lambda(\mathscr{C})$ est stable par réunion dénombrable croissante, on a $\cup_n (A_n \cap C) \in \Lambda(\mathscr{C})$. Mais $\cup_n (A_n \cap C) = (\cup_n A_n) \cap C$, donc $\cup_n A_n \in \Lambda_C$.

En conclusion, Λ_C est bien un λ-système. Supposons maintenant que $C \in \mathscr{C}$. Alors Λ_C contient tous les éléments de $\mathscr{C}$, en effet pour tout $A \in \mathscr{C}$, comme $C \in \mathscr{C}$ et que $\mathscr{C}$ est stable par intersections finies, $A \cap C \in \mathscr{C}$, donc $A \cap C \in \Lambda(\mathscr{C})$, de sorte que $A \in \Lambda_C$. Par conséquent, Λ_C est un λ-système contenant $\mathscr{C}$, et comme par définition $\Lambda_C \subseteq \Lambda(\mathscr{C})$, qui est le plus petit λ-système contenant $\mathscr{C}$, $\Lambda_C = \Lambda(\mathscr{C})$. Comme C est arbitraire, on peut donc écrire

$$\forall C \in \mathscr{C}, \quad \forall A \in \Lambda(\mathscr{C}), \quad A \cap C \in \Lambda(\mathscr{C}).$$

On voudrait maintenant remplacer la première occurrence de $\mathscr{C}$ par $\Lambda(\mathscr{C})$ de manière à obtenir la stabilité par intersections finies de $\Lambda(\mathscr{C})$. Soit $A \in \Lambda(\mathscr{C})$. D'après ce qui précède, pour tout $C \in \mathscr{C}$, $A \cap C \in \Lambda(\mathscr{C})$, autrement dit $\mathscr{C} \subseteq \Lambda_A \subseteq \Lambda(\mathscr{C})$. Comme Λ_A est un λ-système, $\Lambda_A = \Lambda(\mathscr{C})$. Ceci s'écrit

$$\forall A \in \Lambda(\mathscr{C}), \quad \forall B \in \Lambda(\mathscr{C}), \quad B \cap A \in \Lambda(\mathscr{C}),$$

ce qui n'est autre que la stabilité de $\Lambda(\mathscr{C})$ par intersections finies. □

On en déduit les deux résultats d'unicité suivants :

Corollaire 11.17 *Soient μ et ν deux mesures finies sur un espace mesurable $(E, \mathscr{A})$ qui coïncident[2] sur un π-système $\mathscr{C} \subseteq \mathscr{A}$ qui engendre[3] $\mathscr{A}$, alors μ et ν coïncident sur $\mathscr{A}$.*

Corollaire 11.18 *Soient μ et ν deux mesures σ-finies sur un espace mesurable $(E, \mathscr{A})$ telles que :*
a) il existe une suite mesurable croissante (E_n) telle que $\cup_n E_n = E$;
b) pour tout entier n, $\mu(E_n) = \nu(E_n) < \infty$;
c) μ et ν coïncident sur un π-système $\mathscr{C}$ engendrant $\mathscr{A}$ et contenant chaque E_n.
Alors μ et ν coïncident sur $\mathscr{A}$.

Remarque 11.19 *Le fait que μ et ν sont σ-finies pourrait ne pas être mentionné dans l'énoncé qui précède, car c'est en fait une conséquence des conditions (a) et (b).*

Démonstration du Corollaire 11.17. Soit $\Lambda := \{A \in \mathscr{A} : \mu(A) = \nu(A)\}$. Alors Λ est un λ-système car :
(i) Λ contient l'ensemble vide, puisque $\mu(\varnothing) = 0 = \nu(\varnothing)$;
(ii) pour tous éléments A, B de Λ, si $A \subseteq B$, alors comme μ et ν sont finies,

$$\mu(B \setminus A) = \mu(B) - \mu(A) = \nu(B) - \nu(A) = \nu(B \setminus A).$$

(iii) pour toute suite croissante (A_n) d'éléments de Λ, par continuité à gauche de la mesure,

$$\mu(\cup_n A_n) = \lim_n \uparrow \mu(A_n) = \lim_n \uparrow \nu(A_n) = \nu(\cup_n A_n).$$

Comme Λ contient $\mathscr{C}$, $\Lambda(\mathscr{C}) \subseteq \Lambda \subseteq \mathscr{A}$. Mais d'après le théorème de la classe monotone, comme $\mathscr{C}$ est un π-système, $\sigma(\mathscr{C}) = \Lambda(\mathscr{C})$. Enfin par hypothèse, $\mathscr{A} = \sigma(\mathscr{C})$, donc $\mathscr{A} = \sigma(\mathscr{C}) = \Lambda(\mathscr{C}) \subseteq \Lambda \subseteq \mathscr{A}$, ce qui montre bien que $\mathscr{A} = \Lambda$. □

2. c'est-à-dire que pour tout $A \in \mathscr{C}$, $\mu(A) = \nu(A)$
3. c'est-à-dire que $\mathscr{A} = \sigma(\mathscr{C})$

Démonstration du Corollaire 11.18. On applique le corollaire 11.17 aux mesures traces $\mu_n := \mu(\cdot \cap E_n)$ et $\nu_n := \nu(\cdot \cap E_n)$ qui sont finies grâce à l'hypothèse (b). Elles coïncident bien sur $\mathscr{C}$ par l'hypothèse (c), car pour tout $C \in \mathscr{C}$, comme $C \cap E_n \in \mathscr{C}$ (car $E_n \in \mathscr{C}$ et $\mathscr{C}$ est stable par intersection),

$$\mu_n(C) = \mu(C \cap E_n) = \nu(C \cap E_n) = \nu_n(C).$$

Donc μ_n et ν_n coïncident sur $\mathscr{A}$. Maintenant l'hypothèse (a) permet de conclure en utilisant la continuité à gauche de la mesure, car pour tout $A \in \mathscr{A}$,

$$\mu(A) = \mu(\cup_n E_n \cap A) = \mu(\lim_n \uparrow (E_n \cap A)) = \lim_n \uparrow \mu(E_n \cap A) = \lim_n \uparrow \mu_n(A),$$

et de même pour ν. Or $\mu_n(A) = \nu_n(A)$, donc

$$\mu(A) = \lim_n \uparrow \mu_n(A) = \lim_n \uparrow \nu_n(A) = \nu(A),$$

ce qui montre que μ et ν coïncident sur $\mathscr{A}$. □

11.2.2 Applications

Unicité de la mesure de Lebesgue

Supposons qu'il existe deux mesures μ et ν sur $\mathscr{B}(\mathbb{R}^d)$ telles que pour tout rectangle ouvert $R = \prod_{k=1}^d I_k$, où chaque $I_k =]a_k, b_k[$ est un intervalle (ouvert, mais cela est sans importance ici) éventuellement infini de $\mathbb{R}$,

$$\mu(R) = \prod_{k=1}^d (b_k - a_k) = \nu(R),$$

avec la convention habituelle $0 \times \infty = 0$. Montrons qu'alors μ et ν coïncident sur $\mathscr{B}(\mathbb{R}^d)$. Ceci prouvera l'unicité de la mesure de Lebesgue (dont nous montrerons l'existence à la section suivante).

Soit $\mathscr{C}$ l'ensemble des rectangles ouverts de $\mathbb{R}^d$ (produits d'intervalles ouverts pouvant être infinis, donc en particulier pouvant être égaux à $\mathbb{R}$ tout entier, ce qui garantit que $\mathbb{R}^d \in \mathscr{C}$). En particulier $\mathscr{C}$ est un π-système et l'on sait que la tribu engendrée par $\mathscr{C}$ est @ $\mathscr{B}(\mathbb{R}^d)$. Soit E_n le produit des intervalles $]-n, n[$, c'est-à-dire $E_n := \prod_{k=1}^d]-n, n[$. Alors les propriétés du corollaire 11.18 sont bien vérifiées : a) $\cup_n E_n = \mathbb{R}^d$; b) pour tout entier n, $\mu(E_n) = \nu(E_n) = (2n)^d < \infty$; c) $E_n \in \mathscr{C}$, et $\sigma(\mathscr{C}) = \mathscr{B}(\mathbb{R}^d)$, où $\mathscr{C}$ est un π-système. On peut donc conclure que μ et ν coïncident sur $\mathscr{B}(\mathbb{R}^d)$.

Caractérisation d'une mesure par sa fonction de répartition

Définition 11.20 *Si μ est une mesure finie sur* $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, *on appelle* fonction de répartition *de μ la fonction $F : \mathbb{R} \to \mathbb{R}_+$ définie par $F(x) := \mu(]-\infty, x])$.*

Proposition 11.21 *La fonction de répartition F d'une mesure finie est continue à droite, croissante, et vérifie*

$$\lim_{x\to-\infty} F(x) = 0 \quad \text{et} \quad \lim_{x\to+\infty} F(x) = \mu(\mathbb{R}).$$

De plus, pour tous réels $a < b$

$$\begin{array}{llll} i) & \mu(]a,b]) = F(b) - F(a) & ii) & \mu([a,b]) = F(b) - F(a-) \\ iii) & \mu(]a,b[) = F(b-) - F(a) & iv) & \mu([a,b[) = F(b-) - F(a-). \end{array}$$

Dém. À faire en exercice. □ @

Exemple 11.22 *Si $\mu = \delta_a$, alors $F = \mathbb{1}_{[a,+\infty[}$. De manière générale, si $\mu = \sum_n \alpha_n \delta_{x_n}$, alors F est discontinue en tout point x_n tel que $\alpha_n > 0$ et continue partout ailleurs*

$$F(x) = \sum_n \alpha_n \mathbb{1}_{[x_n,+\infty[}.$$

Définition 11.23 *Si μ est une mesure de Borel*[4] *sur $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, on appelle* fonction de répartition généralisée *de μ la fonction $G : \mathbb{R} \to \mathbb{R}$ définie par*

$$G(x) = \begin{cases} -\mu(]x,0]) & si \quad x < 0 \\ \mu(]0,x]) & si \quad x > 0 \\ 0 & si \quad x = 0. \end{cases}$$

Proposition 11.24 *La fonction de répartition G d'une mesure de Borel est continue à droite, croissante, et vérifie*

$$\lim_{x\to-\infty} G(x) = -\mu(\mathbb{R}_-) \quad \text{et} \quad \lim_{x\to+\infty} G(x) = \mu(\mathbb{R}_+^\star).$$

De plus, pour tous réels $a < b$, G vérifie les quatre propriétés énoncées à la proposition précédente pour les mesures finies.

Dém. À faire en exercice. □ @

Exemple 11.25 *Si μ est la mesure de Lebesgue sur $\mathbb{R}$, alors $G(x) = x$.*

Théorème 11.26 *Si μ et ν sont deux mesures finies (resp. de Borel) sur $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, et qu'elles ont la même fonction de répartition F (resp. la même fonction de répartition généralisée G), alors elles sont égales.*

4. c'est-à-dire une mesure finie sur les compacts

Dém. Traitons d'abord le cas où μ et ν sont finies. Soit alors

$$\mathscr{C} := \{]-\infty, x] : x \in \mathbb{R}\} \cup \{\mathbb{R}\}.$$

On voit facilement que $\mathscr{C}$ est un π-système engendrant $\mathscr{B}(\mathbb{R})$ sur lequel μ et ν coïncident, car $\mu(]-\infty, x]) = F(x) = \nu(]-\infty, x])$ (l'égalité $\mu(\mathbb{R}) = \nu(\mathbb{R})$ s'obtient par passage à la limite). Le corollaire 11.17 permet de conclure que μ et ν coïncident sur $\mathscr{B}(\mathbb{R})$.

Dans le cas où μ et ν sont seulement finies sur les compacts, on définit

$$\mathscr{C} := \{]x, y] : -\infty < x \leq y < +\infty\} \cup \{\mathbb{R}\},$$

et $E_n :=]-n, n]$. Alors $E_n \in \mathscr{C}$, $\cup_n E_n = \mathbb{R}$ et $\mathscr{C}$ est un π-système engendrant $\mathscr{B}(\mathbb{R})$ sur lequel μ et ν coïncident, car $\mu(]x, y]) = G(y) - G(x) = \nu(]x, y])$ (les égalités $\mu(\mathbb{R}_-) = \nu(\mathbb{R}_-)$ et $\mu(\mathbb{R}_+^\star) = \nu(\mathbb{R}_+^\star)$ s'obtiennent pas passages à la limite, et impliquent $\mu(\mathbb{R}) = \nu(\mathbb{R})$). Comme on a $\mu(E_n) = \nu(E_n) < \infty$, le corollaire 11.18 permet de conclure que μ et ν coïncident sur $\mathscr{B}(\mathbb{R})$. □

11.3 Existence d'une mesure

11.3.1 Théorème de Caratheodory

Définition 11.27 *Soit $\mathscr{B}$ une algèbre de Boole sur un ensemble E. Une* mesure d'algèbre *sur $(E, \mathscr{B})$ est une application $m : \mathscr{B} \to [0, +\infty]$ qui :*
(i) associe la valeur 0 à l'ensemble vide : $m(\varnothing) = 0$;
(ii) est finiment additive *: pour tous $A, B \in \mathscr{B}$ tels que $A \cap B = \varnothing$, $m(A \cup B) = m(A) + m(B)$;*
(iii) satisfait la popriété suivante : il existe une suite croissante (E_n) d'éléments de $\mathscr{B}$ convergeant vers E telle que $m(E_n) < \infty$ pour chaque entier n et telle que pour tout $A \in \mathscr{B}$, $\lim_n \uparrow m(A \cap E_n) = m(A)$;
(iv) satisfait la propriété de Caratheodory *: pour toute suite décroissante (A_n) d'éléments de $\mathscr{B}$ convergeant vers $\varnothing$ et telle que $m(A_0) < \infty$, $\lim_n \downarrow m(A_n) = 0$.*

Proposition 11.28 *Une mesure d'algèbre m sur $(E, \mathscr{B})$ vérifie pour tous $A, B \in \mathscr{B}$:*
(i) Additivité finie : $m(A) = m(A \setminus B) + m(A \cap B)$;
(ii) Additivité forte : $m(A \cup B) + m(A \cap B) = m(A) + m(B)$;
(iii) Sous-additivité : $m(A \cup B) \leq m(A) + m(B)$;
(iv) Croissance : si $A \subseteq B$, $m(A) \leq m(B)$.

Théorème 11.29 (de prolongement de Caratheodory) *Soit $\mathscr{B}$ une algèbre de Boole sur un ensemble E. Si m est une mesure d'algèbre sur $(E, \mathscr{B})$, alors il existe une mesure μ sur la tribu $\sigma(\mathscr{B})$ qui coïncide[5] avec m sur $\mathscr{B}$.*

Remarque 11.30 *On dit alors que μ est un prolongement de la mesure (d'algèbre) m, qui elle est seulement définie sur l'algèbre $\mathscr{B}$, à la tribu $\sigma(\mathscr{B})$. Ce théorème de prolongement est admis.*

5. c'est-à-dire que pour tout $B \in \mathscr{B}$, $\mu(B) = m(B)$

Remarque 11.31 *En fait on aurait directement pu dire dans le théorème qu'il existe un* unique *tel prolongement. En effet, si μ et ν sont deux prolongements d'une même mesure d'algèbre $\mathscr{B}$, alors μ et ν coïncident sur $\mathscr{B}$, qui est aussi un π-système, et comme il existe une suite mesurable (E_n) convergeant vers E telle que $\mu(E_n) = \nu(E_n) < \infty$, alors μ et ν coïncident sur $\sigma(\mathscr{B})$ d'après le corollaire 11.18.*

11.3.2 Applications

Existence de la mesure de Lebesgue

Montrons l'existence d'une mesure sur $\mathscr{B}or(\mathbb{R}^d)$ telle que la mesure d'un rectangle $R = \prod_{k=1}^d I_k$, où chaque I_k est un intervalle de $\mathbb{R}$ d'extrémité gauche $a_k \geq -\infty$ et d'extrémité droite $b_k \leq +\infty$ (les extrémités pouvant être fermées ou ouvertes), vaut $\prod_{k=1}^d (b_k - a_k)$. On définit $\mathscr{B}$ l'ensemble des réunions finies de rectangles deux à deux disjoints. Alors pour tout $A \in \mathscr{B}$, A s'écrit de manière unique [6] sous la forme $A = \cup_{i=1}^j R_i$, où les (R_i) sont des rectangles deux à deux disjoints, et l'on peut définir sans ambiguïté la mesure d'algèbre m sur $\mathscr{B}$ par $m(A) = \sum_{i=1}^j m(R_i)$, où la mesure d'un rectangle a été définie précédemment. On peut alors vérifier que $\mathscr{B}$ est une algèbre et que m est une mesure d'algèbre sur $(E, \mathscr{B})$ avec E_n définie comme le produit des intervalles $]-n, n[$ (pour montrer la propriété de Caratheodory, on s'inspirera de l'application suivante). Comme $\sigma(\mathscr{B}) = \mathscr{B}or(\mathbb{R}^d)$, le théorème de Caratheodory permet bien de déduire l'existence d'une mesure, appelée mesure de Lebesgue, prolongeant la mesure m à tous les boréliens de $\mathbb{R}^d$.

Définition d'une mesure par sa fonction de répartition

Théorème 11.32 *Si $F : \mathbb{R} \to \mathbb{R}_+$ est une fonction croissante, bornée, continue à droite et telle que $\lim_{x \to -\infty} F(x) = 0$ alors il existe une (unique) mesure μ sur $\mathscr{B}or(\mathbb{R})$ qui admet F pour fonction de répartition.*

Théorème 11.33 *Si $G : \mathbb{R} \to \mathbb{R}$ est une fonction croissante, continue à droite et telle que $G(0) = 0$, alors il existe une (unique) mesure μ de Borel sur $\mathscr{B}or(\mathbb{R})$ qui admet G pour fonction de répartition généralisée.*

Remarque 11.34 *L'existence de la mesure de Lebesgue sur $\mathbb{R}$ peut également se déduire du théorème précédent en prenant $G(x) = x$.*

Dém. du théorème 11.33. Soit $\mathscr{B}$ l'ensemble de toutes les réunions finies d'intervalles disjoints, qui est une algèbre. Pour tous $-\infty \leq x \leq y \leq \infty$, on définit alors (avec $G(-\infty) = \lim_{x \to -\infty} G(x)$ et $G(\infty-) = \lim_{x \to \infty} G(x)$)

$$m(]x, y[) := G(y-) - G(x) \quad \text{et} \quad m(\{x\}) = G(x) - G(x-),$$

6. Cette écriture n'est unique que si elle est supposée minimale, c'est-à-dire utilisant un minimum de rectangles ; de plus, si les rectangles ne sont plus supposés disjoints, il n'existe même plus de décomposition minimale unique, et il faut donc montrer que la définition de m ne dépend pas de la décomposition choisie

ce qui définit m sur tout intervalle de $\mathbb{R}$, et pour tout $A = \cup_{j=1}^{i} I_j \in \mathscr{B}$, où les (I_j) sont des intervalles disjoints, $m(A) := \sum_{j=1}^{i} m(I_j)$. Montrons que les quatre propriétés du théorème de prolongement de Caratheodory sont satisfaites. Le résultat découlera alors de ce théorème car $\sigma(\mathscr{B}) = \mathscr{B}or(\mathbb{R})$.

Il est immédiat de vérifier que c'est bien le cas des propriétés (i) et (ii). Pour (iii), définissons $E_n :=]-n, n]$ et fixons $A \in \mathscr{B}$. En distinguant suivant que A contient un intervalle infini ou non, on montre alors que $\lim_n \uparrow m(A \cap E_n) = m(A)$. En effet, d'après la définition de la mesure d'algèbre m, on peut toujours se ramener au cas où A est un seul intervalle. Si par exemple $A =]x, +\infty[$, où x est fini, alors à partir d'un certain rang, $A \cap E_n =]x, n]$ et $\lim_n m(A \cap E_n) = \lim_n G(n) - G(x) = G(\infty-) - G(x) = m(A)$. Les autres cas se démontrent de la même manière.

Le plus difficile reste à faire, à savoir montrer que la propriété (iv), dite de Caratheodory, est satisfaite. Soit donc une suite décroissante (A_n) d'éléments de $\mathscr{B}$ telle que $m(A_0) < \infty$. On peut donc écrire A_n de manière unique sous la forme

$$A_n = \bigcup_{k=1}^{K_n} I_{k,n},$$

où les K_n intervalles $(I_{k,n})_k$ sont disjoints. On fixe alors ε et l'on cherche à définir une suite de compacts (A'_n) telle que $A'_n \subseteq A_n$ et $m(A_n \setminus A'_n) \leq \varepsilon 2^{-n}$. Il suffit pour cela d'exhiber des intervalles *compacts* $I'_{k,n} \subseteq I_{k,n}$ suffisamment grands pour que

$$m\left(I_{k,n} \setminus I'_{k,n}\right) \leq \frac{\varepsilon}{K_n 2^n} \quad n \geq 0, 1 \leq k \leq K_n.$$

En effet, en définissant

$$A'_n := \bigcup_{k=1}^{K_n} I'_{k,n},$$

on aura alors $A_n \subseteq A'_n$ et $A_n \setminus A'_n = \cup_{k=1}^{K_n} \left(I_{k,n} \setminus I'_{k,n}\right)$, si bien que

$$m\left(A_n \setminus A'_n\right) = \sum_{k=1}^{K_n} m\left(I_{k,n} \setminus I'_{k,n}\right) \leq \sum_{k=1}^{K_n} \frac{\varepsilon}{K_n 2^n} = \frac{\varepsilon}{2^n}.$$

Remarquons d'abord que pour tous $-\infty \leq x \leq y \leq +\infty$ tels que $m(]x, y[) < \infty$, pour toute suite décroissante (x_n) convergeant vers x et pour toute suite croissante (y_n) convergeant vers y,

$$m(]x, y[\setminus[x_n, y_n]) = m(]x, x_n[) + m(]y_n, y[) = G(x_n-) - G(x) + G(y-) - G(y_n) \longrightarrow 0,$$

lorsque $n \to \infty$, car G est continue à droite (avec des limites à gauche, car croissante) et l'on a supposé que $G(y-)$ et $G(x)$ sont finis. En conclusion, pour tout $\alpha > 0$ et tout intervalle $I =]x, y[$ de $\mathbb{R}$ (resp. $I = [x, y[$, resp. $I =]x, y]$) tel que $m(I) < \infty$, il existe un intervalle compact $I' = [x', y']$ (resp. $I' = [x, y']$, resp. $I' = [x', y]$) inclus dans I tel que $m(I' \setminus I) < \alpha$. Puisque $m(A_0) < \infty$, on peut donc bien trouver des intervalles compacts $I'_{k,n}$ vérifiant la propriété demandée.

Comme $A'_n \subseteq A_n$, on a $\cap_n A'_n \subseteq \cap_n A_n = \varnothing$. Or chaque A'_n est compact (comme réunion finie de compacts) donc il existe un entier N tel que $\cap_{n=0}^N A'_n = \varnothing$. Or @

$$\left(\bigcap_{n=0}^N A_n\right) \cap {}^c\left(\bigcup_{n=0}^N A_n \setminus A'_n\right) = \left(\bigcap_{n=0}^N A_n\right) \cap \left(\bigcap_{n=0}^N ({}^cA_n \cup A'_n)\right)$$
$$= \bigcap_{n=0}^N A_n \cap ({}^cA_n \cup A'_n) = \bigcap_{n=0}^N A_n \cap A'_n = \bigcap_{n=0}^N A'_n = \varnothing.$$

De cette intersection vide, on déduit que

$$A_N = \bigcap_{n=0}^N A_n \subseteq \bigcup_{n=0}^N (A_n \setminus A'_n)\,.$$

Mais comme une mesure d'algèbre est croissante et sous-additive, pour tout entier $k \geq N$, @

$$m(A_k) \leq m(A_N) \leq m\left(\bigcup_{n=0}^N (A_n \setminus A'_n)\right) \leq \sum_{n=0}^N m\,(A_n \setminus A'_n) \leq \sum_{n=0}^N \varepsilon 2^{-n} \leq 2\varepsilon.$$

En conclusion, pour tout $\varepsilon > 0$, il existe un entier N tel que pour tout $k \geq N$, $m(A_k) \leq 2\varepsilon$, ce qui n'est autre que la propriété de Caratheodory. □

Chapitre 12

Tribu produit, mesure produit et « intégrales multiples »

12.1 Tribu produit

12.1.1 Cas général

Soient $(E_1, \mathscr{A}_1)$ et $(E_2, \mathscr{A}_2)$ deux espaces mesurables.

Définition 12.1 *On appelle* tribu produit *sur $E_1 \times E_2$, et l'on note $\mathscr{A}_1 \otimes \mathscr{A}_2$, la plus petite tribu contenant les pavés à côtés mesurables :*

$$\mathscr{A}_1 \otimes \mathscr{A}_2 := \sigma\left(\mathscr{A}_1 \times \mathscr{A}_2\right),$$

où l'on a noté

$$\mathscr{A}_1 \times \mathscr{A}_2 := \{A_1 \times A_2 : A_1 \in \mathscr{A}_1, A_2 \in \mathscr{A}_2\}.$$

Le couple $(E_1 \times E_2, \mathscr{A}_1 \otimes \mathscr{A}_2)$ est appelé espace mesurable produit.

Remarque 12.2 *Bien entendu, la famille $\mathscr{A}_1 \times \mathscr{A}_2$ des pavés à côtés mesurables n'est en général pas une tribu.*

Proposition 12.3 *La tribu $\mathscr{A}_1 \otimes \mathscr{A}_2$ est aussi la tribu engendrée par les projections canoniques π_1 et π_2, c'est-à-dire la plus petite tribu sur $E_1 \times E_2$ qui rende π_1 et π_2 mesurables*[1].

Dém. Soit $\mathscr{B}$ la tribu engendrée par π_1 et π_2. Par définition, $\mathscr{B}$ est la plus petite tribu contenant les parties de $E_1 \times E_2$ de la forme $\pi_1^{-1}(A_1)$ et $\pi_2^{-1}(A_2)$ où $A_i \in \mathscr{A}_i$, $i = 1, 2$. Or $\pi_1^{-1}(A_1) = A_1 \times E_2$ et $\pi_2^{-1}(A_2) = E_1 \times A_2$, donc $\mathscr{B}$ est aussi la plus petite tribu qui contient les parties de $E_1 \times E_2$ de la forme $(A_1 \times E_2) \cap (E_1 \times A_2) = A_1 \times A_2$, c'est-à-dire $\sigma(\mathscr{A}_1 \times \mathscr{A}_2)$. $\square$

1. On rappelle que π_1 et π_2 sont définies par : $\pi_1(x, y) = x$ et que $\pi_2(x, y) = y$

Proposition 12.4 *Soit*

$$\begin{aligned} f : (X, \mathscr{T}) &\longrightarrow (E_1 \times E_2, \mathscr{A}_1 \otimes \mathscr{A}_2) \\ x &\longmapsto f(x) = (f_1(x), f_2(x)) \end{aligned}$$

Alors la fonction f est mesurable ssi f_1 et f_2 sont mesurables[2].

Dém. Sens direct de l'équivalence : si f est mesurable alors pour tout $i = 1, 2$, $f_i = \pi_i \circ f$ est mesurable comme composée de fonctions mesurables.

Autre sens : supposons que f_1 et f_2 sont mesurables. Alors pour tous $A_1 \in \mathscr{A}_1$ et $A_2 \in \mathscr{A}_2$,

$$f^{-1}(A_1 \times A_2) = \{x \in X : f(x) \in A_1 \times A_2\} = \{x \in X : f_1(x) \in A_1, f_2(x) \in A_2\} = f_1^{-1}(A_1) \cap f_2^{-1}(A_2).$$

Par hypothèse $f_1^{-1}(A_1) \in \mathscr{T}$ et $f_2^{-1}(A_2) \in \mathscr{T}$, donc $f^{-1}(A_1 \times A_2) \in \mathscr{T}$ par stabilité des tribus par intersection. En conclusion, $f^{-1}(\mathscr{A}_1 \times \mathscr{A}_2) \subseteq \mathscr{T}$, donc

$$f^{-1}(\mathscr{A}_1 \otimes \mathscr{A}_2) = \sigma\left(f^{-1}(\mathscr{A}_1 \times \mathscr{A}_2)\right) \subseteq \mathscr{T},$$

où la première égalité est une application du lemme de transport, et l'inclusion n'est autre que la mesurabilité de f. □

Remarque 12.5 *Ce qui précède peut bien sûr s'énoncer de manière similaire pour tout produit cartésien fini d'ensembles. Si $((E_i, \mathscr{A}_i))_{1 \leq i \leq d}$ sont d espaces mesurables, alors on définit $\mathscr{A}_1 \otimes \cdots \otimes \mathscr{A}_d$ comme la plus petite tribu sur $E_1 \times \cdots \times E_d$ contenant tous les pavés de la forme $A_1 \times \cdots \times A_d$, où $A_i \in \mathscr{A}_i$ pour tous $i = 1, \ldots, d$; c'est aussi la tribu engendrée par les projections canoniques*

$$\begin{aligned} \pi_i : \qquad \prod_{j=1}^{d} E_j &\longrightarrow E_i \\ (x_1, \ldots, x_d) &\longmapsto x_i \end{aligned}$$

pour i parcourant $\{1, \ldots, d\}$.

Proposition 12.6 (associativité de $\otimes$) *On a l'égalité suivante entre tribus*

$$(\mathscr{A}_1 \otimes \cdots \otimes \mathscr{A}_j) \otimes (\mathscr{A}_{j+1} \otimes \cdots \otimes \mathscr{A}_d) = \mathscr{A}_1 \otimes \cdots \otimes \mathscr{A}_d,$$

où l'on a bien sûr identifié $(E_1 \times \cdots \times E_j) \times (E_{j+1} \times \cdots \times E_d)$ et $E_1 \times \cdots \times E_d$.

2. comme fonctions de $(X, \mathscr{T})$ vers $(E_1, \mathscr{A}_1)$ et $(E_2, \mathscr{A}_2)$ respectivement

Dém. Montrons la proposition dans le cas où $j = 2$ et $d = 3$.

Première démonstration possible. La tribu $(\mathscr{A}_1 \otimes \mathscr{A}_2) \otimes \mathscr{A}_3$ est la plus petite tribu qui rende mesurables les applications

$$\begin{aligned} f_{1,2} : E_1 \times E_2 \times E_3 &\longrightarrow (E_1 \times E_2, \mathscr{A}_1 \otimes \mathscr{A}_2) \\ ((x_1, x_2), x_3) &\longmapsto (x_1, x_2) \end{aligned}$$

et

$$\begin{aligned} f_3 : E_1 \times E_2 \times E_3 &\longrightarrow (E_3, \mathscr{A}_3) \\ ((x_1, x_2), x_3) &\longmapsto x_3 \end{aligned}$$

Or $f_{1,2}$ est mesurable ssi ses applications coordonnées le sont. Par conséquent, ces deux applications sont mesurables ssi les trois applications $(x_1, x_2, x_3) \mapsto x_i$, pour $i = 1, 2, 3$, sont mesurables. Donc $(\mathscr{A}_1 \otimes \mathscr{A}_2) \otimes \mathscr{A}_3$ est la tribu engendrée par ces trois applications, c'est donc $\mathscr{A}_1 \otimes \mathscr{A}_2 \otimes \mathscr{A}_3$.

Deuxième démonstration possible, par double inclusion. Cette démonstration est plus compliquée, mais constitue un bon exercice. Démontrons d'abord l'inclusion $\mathscr{A}_1 \otimes \mathscr{A}_2 \otimes \mathscr{A}_3 \subseteq (\mathscr{A}_1 \otimes \mathscr{A}_2) \otimes \mathscr{A}_3$. Pour tous $A_i \in \mathscr{A}_i$, $(i = 1, 2, 3)$, $A_1 \times A_2 \in \mathscr{A}_1 \otimes \mathscr{A}_2$, donc

$$A_1 \times A_2 \times A_3 \in (\mathscr{A}_1 \otimes \mathscr{A}_2) \times \mathscr{A}_3 \subseteq (\mathscr{A}_1 \otimes \mathscr{A}_2) \otimes \mathscr{A}_3,$$

donc $(\mathscr{A}_1 \otimes \mathscr{A}_2) \otimes \mathscr{A}_3$ contient les pavés à côtés mesurables de $E_1 \times E_2 \times E_3$, donc contient tous les éléments de la tribu qu'ils engendrent, ce qui est l'inclusion annoncée.

Montrons l'inclusion inverse. Fixons $A_3 \in \mathscr{A}_3$ et définissons

$$\mathscr{T} := \{B \in \mathscr{A}_1 \otimes \mathscr{A}_2 : B \times A_3 \in \mathscr{A}_1 \otimes \mathscr{A}_2 \otimes \mathscr{A}_3\}.$$

On veut montrer que $\mathscr{T}$ est une tribu : si $\mathscr{T}$ est une tribu, alors comme $\mathscr{T}$ contient les pavés de $E_1 \times E_2$ à côtés mesurables, $\mathscr{T}$ contient tous les éléments de la tribu $\mathscr{B} := \mathscr{A}_1 \otimes \mathscr{A}_2$ qu'ils engendrent. Ceci implique que pour tout $B \in \mathscr{B}$, $B \times A_3 \in \mathscr{A}_1 \otimes \mathscr{A}_2 \otimes \mathscr{A}_3$, où A_3 est un élément arbitraire de $\mathscr{A}_3$. Autrement dit, $\mathscr{B} \times \mathscr{A}_3 \subseteq \mathscr{A}_1 \otimes \mathscr{A}_2 \otimes \mathscr{A}_3$ et donc

$$\mathscr{B} \otimes \mathscr{A}_3 = \sigma(\mathscr{B} \times \mathscr{A}_3) \subseteq \mathscr{A}_1 \otimes \mathscr{A}_2 \otimes \mathscr{A}_3,$$

qui est l'inclusion annoncée. Vérifions donc que $\mathscr{T}$ est une tribu :

i) $\varnothing \in \mathscr{T}$ car $\varnothing \times A_3 = \varnothing \in \mathscr{A}_1 \otimes \mathscr{A}_2 \otimes \mathscr{A}_3$.

ii) pour tout $B \in \mathscr{T}$, ${}^cB \in \mathscr{T}$ car

$${}^cB \times A_3 = (E_1 \times E_2 \times A_3) \cap {}^c(B \times A_3)$$

qui est bien élément de $\mathscr{T}$ car $E_1 \times E_2 \times A_3 \in \mathscr{A}_1 \otimes \mathscr{A}_2 \otimes \mathscr{A}_3$ et $B \times A_3 \in \mathscr{A}_1 \otimes \mathscr{A}_2 \otimes \mathscr{A}_3$, qui est une tribu, donc stable par passage au complémentaire et intersection.

iii) pour toute suite (B_n) d'éléments de $\mathscr{T}$,

$$(\cup_n B_n) \times A_3 = \cup_n (B_n \times A_3),$$

qui est bien élément de $\mathscr{T}$ car pour tout n, $B_n \times A_3 \in \mathscr{A}_1 \otimes \mathscr{A}_2 \otimes \mathscr{A}_3$, qui est une tribu, donc stable par réunion dénombrable. □

12.1.2 Le cas borélien

Lorsque E_1 et E_2 sont des espaces topologiques, nous disposons déjà d'une tribu sur $E_1 \times E_2$ qui est la tribu borélienne $\mathscr{B}or(E_1 \times E_2)$, ou tribu engendrée par la topologie produit, dont on rappelle que les éléments sont les réunions (quelconques) de produits d'ouverts (dits aussi pavés à côtés ouverts).

Proposition 12.7 *a) On a toujours l'inclusion*

$$\mathscr{B}or(E_1) \otimes \mathscr{B}or(E_2) \subseteq \mathscr{B}or(E_1 \times E_2).$$

b) Si E_1 et E_2 sont tous deux à base dénombrable d'ouverts (en particulier si E_1 et E_2 sont des espaces métriques séparables), alors l'inclusion précédente devient une égalité.

Dém. a) Par définition de la topologie produit, $\pi_i : E_1 \times E_2 \to E_i$ est continue pour $i = 1, 2$ ($i = 1$: si O_1 est un ouvert de E_1, $\pi_1^{-1}(O_1) = O_1 \times E_2$ est un ouvert de $E_1 \times E_2$), et par conséquent π_i est borélienne[3]. Or la plus petite tribu qui rende mesurables π_1 et π_2 est $\mathscr{B}or(E_1) \otimes \mathscr{B}or(E_2)$, d'où le résultat.

b) Pour tout $i = 1, 2$, soit $\mathscr{U}_i = (U_n^{(i)})_{n \in \mathbb{N}}$ une base dénombrable d'ouverts de E_i, c'est-à-dire que tout ouvert de E_i peut s'écrire comme réunion (forcément dénombrable, donc) d'éléments de $\mathscr{U}_i$. Par définition de la topologie produit, tout ouvert Ω de $E_1 \times E_2$ est une réunion (quelconque, cette fois) de produits d'ouverts

$$\Omega = \bigcup_{j \in J} O_j^{(1)} \times O_j^{(2)},$$

où J est un ensemble d'indices quelconque et pour tous i, j, $O_j^{(i)}$ est un ouvert de E_i. Comme $O_j^{(i)}$ est un ouvert de E_i, $O_j^{(i)}$ s'écrit comme réunion d'éléments de $\mathscr{U}_i$, c'est-à-dire qu'il existe une partie $K_j^{(i)}$ de $\mathbb{N}$ telle que

$$O_j^{(i)} = \bigcup_{h \in K_j^{(i)}} U_h^{(i)},$$

et ainsi

$$O_j^{(1)} \times O_j^{(2)} = \left(\bigcup_{h \in K_j^{(1)}} U_h^{(1)} \right) \times \left(\bigcup_{k \in K_j^{(2)}} U_k^{(2)} \right) = \bigcup_{(h,k) \in K_j^{(1)} \times K_j^{(2)}} U_h^{(1)} \times U_k^{(2)}.$$

En conclusion,

$$\Omega = \bigcup_{(h,k) \in \cup_{j \in J} K_j^{(1)} \times K_j^{(2)}} U_h^{(1)} \times U_k^{(2)},$$

3. sans ambiguïté : l'espace d'arrivée E_i est muni de sa tribu borélienne $\mathscr{B}or(E_i)$ et l'espace de départ $E_1 \times E_2$ est muni de sa tribu borélienne $\mathscr{B}or(E_1 \times E_2)$

qui est une réunion dénombrable de produits d'ouverts car $\cup_{j\in J} K_j^{(1)} \times K_j^{(2)} \subseteq \mathbb{N}^2$. Comme un produit d'ouverts est élément de $\mathscr{B}or(E_1) \otimes \mathscr{B}or(E_2)$, c'est le cas également de Ω, par stabilité des tribus par réunion dénombrable. Ainsi les ouverts de $E_1 \times E_2$ sont des éléments de la tribu $\mathscr{B}or(E_1) \otimes \mathscr{B}or(E_2)$, et par conséquent la plus petite tribu contenant les ouverts de $E_1 \times E_2$, à savoir $\mathscr{B}or(E_1 \times E_2)$, est incluse dans $\mathscr{B}or(E_1) \otimes \mathscr{B}or(E_2)$. □

Corollaire 12.8 *Comme $\mathbb{R}$ est un espace métrique séparable, $\mathscr{B}or(\mathbb{R}) \otimes \mathscr{B}or(\mathbb{R}) = \mathscr{B}or(\mathbb{R}^2)$ et plus généralement, pour tout entier $d \geq 2$,*

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

Ce corollaire permet, par exemple, de voir rapidement pourquoi, si $f, g : \mathbb{R} \to \mathbb{R}$ sont deux fonctions boréliennes, alors $f + g$ et fg sont aussi boréliennes. En effet, on sait que l'application somme S

$$\begin{aligned} S : (\mathbb{R}^2, \mathscr{B}or(\mathbb{R}^2)) &\longrightarrow (\mathbb{R}, \mathscr{B}or(\mathbb{R})) \\ (x, y) &\longmapsto x + y \end{aligned}$$

et l'application produit P

$$\begin{aligned} P : (\mathbb{R}^2, \mathscr{B}or(\mathbb{R}^2)) &\longrightarrow (\mathbb{R}, \mathscr{B}or(\mathbb{R})) \\ (x, y) &\longmapsto xy \end{aligned}$$

sont boréliennes car continues. De plus, on sait que l'application

$$\begin{aligned} C : (\mathbb{R}, \mathscr{B}or(\mathbb{R})) &\longrightarrow (\mathbb{R}^2, \mathscr{B}or(\mathbb{R}) \otimes \mathscr{B}or(\mathbb{R})) \\ x &\longmapsto (f(x), g(x)) \end{aligned}$$

est mesurable, car les deux applications coordonnées f et g sont mesurables. Ayant l'égalité entre $\mathscr{B}or(\mathbb{R}^2)$ et $\mathscr{B}or(\mathbb{R}) \otimes \mathscr{B}or(\mathbb{R})$, on a donc la mesurabilité de $f + g = S \circ C$ et de $fg = P \circ C$.

12.1.3 Sections

Définition 12.9 *Si $C \in \mathscr{A}_1 \otimes \mathscr{A}_2$, pour tous $x_1 \in E_1$ et $x_2 \in E_2$, on note*

$$C_{x_1} := \{y_2 \in E_2 : (x_1, y_2) \in C\} \quad \text{et} \quad C^{x_2} := \{y_1 \in E_1 : (y_1, x_2) \in C\},$$

que l'on appelle sections *de C.*

Proposition 12.10 *Soit $f \in \mathscr{F}(\mathscr{A}_1 \otimes \mathscr{A}_2, \mathscr{B}or(\mathbb{R}))$. Alors pour tout $x_1 \in E_1$, l'application partielle*

$$\begin{aligned} f_{x_1} : (E_2, \mathscr{A}_2) &\longrightarrow (\mathbb{R}, \mathscr{B}or(\mathbb{R})) \\ x_2 &\longmapsto f(x_1, x_2) \end{aligned}$$

est mesurable.

Remarque 12.11 *Attention, la réciproque est fausse : le fait que toutes les applications partielles soient mesurables n'implique pas forcément que f soit mesurable.*

Dém. L'application

$$\begin{aligned} g_{x_1} : (E_2, \mathscr{A}_2) &\longrightarrow (E_1 \times E_2, \mathscr{A}_1 \otimes \mathscr{A}_2) \\ x_2 &\longmapsto (x_1, x_2) \end{aligned}$$

est mesurable car chacune des applications coordonnées l'est de façon évidente. Donc $f_{x_1} = f \circ g_{x_1}$ est mesurable. □

Proposition 12.12 *Les sections d'éléments de la tribu produit sont mesurables. Autrement dit, pour tout $C \in \mathscr{A}_1 \otimes \mathscr{A}_2$ et pour tous $x_1 \in E_1$ et $x_2 \in E_2$: $C_{x_1} \in \mathscr{A}_2$ et $C^{x_2} \in \mathscr{A}_1$.*

Dém. Il suffit d'appliquer la proposition précédente à la fonction $f = \mathbb{1}_C$, qui est mesurable par hypothèse. On obtient donc que f_{x_1} est mesurable, mais $f_{x_1} = \mathbb{1}_{C_{x_1}}$, donc $C_{x_1} \in \mathscr{A}_2$. □

12.2 Mesure produit

Soient μ_1 et μ_2 deux mesures σ-**finies**, sur $(E_1, \mathscr{A}_1)$ et $(E_2, \mathscr{A}_2)$ respectivement.

Lemme 12.13 *Pour tout $C \in \mathscr{A}_1 \otimes \mathscr{A}_2$, l'application*

$$\begin{aligned} h_C : (E_1, \mathscr{A}_1) &\longrightarrow (\bar{\mathbb{R}}_+, \mathscr{B}or(\bar{\mathbb{R}}_+)) \\ x_1 &\longmapsto \mu_2(C_{x_1}) \end{aligned}$$

est mesurable.

Dém. Supposons d'abord que μ_2 est finie. Soit

$$\Lambda := \{C \in \mathscr{A}_1 \otimes \mathscr{A}_2 : h_C \text{ est mesurable}\}.$$

Montrons que Λ est un λ-système.

i) $C = \varnothing \in \Lambda$ car alors h_C est la fonction nulle, qui est toujours mesurable.

ii) Soient $C \subseteq D$ tous deux éléments de Λ. Alors h_C et h_D sont mesurables, et $h_{D\setminus C} = h_D - h_C$, car $(D \setminus C)_{x_1} = D_{x_1} \setminus C_{x_1}$, avec $C_{x_1} \subseteq D_{x_1}$. Donc $h_{D\setminus C}$ est mesurable, et $D \setminus C \in \Lambda$.

iii) Soit $(C^{(n)})_n$ une suite croissante d'éléments de Λ et C sa limite. Alors la suite $(C^{(n)}_{x_1})$ est croissante, donc par continuité à gauche de la mesure μ_2,

$$h_C(x_1) = \mu_2((\cup_n C^{(n)})_{x_1}) = \mu_2(\cup_n C^{(n)}_{x_1}) = \mu_2(\lim_n \uparrow C^{(n)}_{x_1}) = \lim_n \uparrow \mu_2(C^{(n)}_{x_1}) = \lim_n h_{C_n}(x_1),$$

qui est bien mesurable, comme limite de fonctions mesurables.

En conclusion, Λ est bien un λ-système. Montrons que Λ contient le π-système $\mathscr{A}_1 \times \mathscr{A}_2$. En effet, pour tout $C = A_1 \times A_2 \in \mathscr{A}_1 \times \mathscr{A}_2$,

$$C_{x_1} = \begin{cases} A_2 & \text{si } x_1 \in A_1 \\ \varnothing & \text{sinon,} \end{cases}$$

donc $h_C = \mu_2(A_2)\mathbb{1}_{A_1}$, qui est mesurable car étagée. Le théorème de la classe monotone assure alors que Λ contient $\sigma(\mathscr{A}_1 \times \mathscr{A}_2) = \mathscr{A}_1 \otimes \mathscr{A}_2$. La proposition est donc démontrée dans le cas où μ_2 est finie.

Si μ_2 est seulement σ-finie, alors par définition, il existe une suite croissante $(E_2^{(n)})_n$ d'éléments de $\mathscr{A}_2$ convergeant vers E_2 telle que $\mu_2(E_2^{(n)}) < \infty$ pour tout entier n. En particulier, pour tout $A_2 \in \mathscr{A}_2$, $\mu_2(A_2) = \lim_n \uparrow \mu_2(A_2 \cap E_2^{(n)})$. En appliquant ce qui précède à la mesure trace de μ_2 sur $E_2^{(n)}$, on obtient que l'application $h_n : x_1 \mapsto \mu_2(C_{x_1} \cap E_2^{(n)})$ est mesurable, et par conséquent l'application h_C est également mesurable, comme limite (croissante) de la suite de fonctions (h_n). □

Théorème 12.14 *Il existe une unique mesure m sur l'espace produit $(E_1 \times E_2, \mathscr{A}_1 \otimes \mathscr{A}_2)$ vérifiant*

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

pour tous $A_1 \in \mathscr{A}_1$ et $A_2 \in \mathscr{A}_2$. Cette mesure est σ-finie et est appelée mesure produit*. On la note $m = \mu_1 \otimes \mu_2$. De plus, pour tout $C \in \mathscr{A}_1 \otimes \mathscr{A}_2$,*

$$\int_{E_1} \mu_2(C_{x_1})\, d\mu_1(x_1) = \mu_1 \otimes \mu_2(C) = \int_{E_2} \mu_1(C^{x_2})\, d\mu_2(x_2).$$

Remarque 12.15 *On pourrait énoncer un résultat qui assure que le produit de mesures $\otimes$ est associatif, et le démontrer en utilisant la coïncidence des différents produits de mesures possibles sur les pavés à côtés mesurables. Une conséquence de cette remarque est la proposition suivante.* @

Proposition 12.16 *La mesure de Lebesgue λ_d sur $(\mathbb{R}^d, \mathscr{B}or(\mathbb{R}^d))$ est aussi la mesure produit $\lambda_1^{\otimes d}$.*

Remarque 12.17 *Le théorème 12.14 est faux lorsque μ_1 ou μ_2 n'est pas σ-finie comme on le voit en prenant par exemple la mesure de Lebesgue sur $\mathbb{R}$ pour μ_1 (qui est bien σ-finie) mais la mesure de comptage sur $\mathbb{R}$ pour μ_2 (qui n'est pas σ-finie). En prenant par exemple $(E_1, \mathscr{A}_1) = (\mathbb{R}, \mathscr{B}or(\mathbb{R}))$, $(E_2, \mathscr{A}_2) = (\mathbb{R}, \mathscr{P}(\mathbb{R}))$, et $C = \{(x, x) : x \in \mathbb{R}\}$ la première bissectrice de $\mathbb{R}^2$, alors $C_{x_1} = \{x_1\}$ et $C^{x_2} = \{x_2\}$, donc $\mu_1(C^{x_2}) = 0$, tandis que $\mu_2(C_{x_1}) = 1$. Par conséquent,*

$$\int_{E_1} \mu_2(C_{x_1})\, d\mu_1(x_1) = \mu_1(E_1) = +\infty \neq 0 = \int_{E_2} \mu_1(C_{x_2})\, d\mu_2(x_2).$$

Dém. du théorème 12.14. a) Unicité. Si m et m' vérifient la propriété du théorème, c'est qu'elles coïncident sur le π-système $\mathscr{A}_1 \times \mathscr{A}_2$ qui engendre $\mathscr{A}_1 \otimes \mathscr{A}_2$. De plus, comme μ_1 et μ_2 sont toutes deux σ-finies, alors pour $i = 1, 2$, il existe une suite croissante $(E_i^{(n)})$ d'éléments de $\mathscr{A}_i$ convergeant vers E_i et tels que $\mu_i(E_i^{(n)}) < \infty$. Alors si l'on définit $C_n = E_1^{(n)} \times E_2^{(n)}$, $m(C_n) = m'(C_n) = \mu_1(E_1^{(n)})\, \mu_2(E_2^{(n)}) < \infty$ et $\cup_n C_n = E_1 \times E_2$, ce qui permet de conclure que $m = m'$ par le Corollaire 11.18.

b) Existence. On définit $m_1 : \mathscr{A}_1 \otimes \mathscr{A}_2 \to \bar{\mathbb{R}}_+$ par

$$m_1(C) := \int_{E_1} \mu_2(C_{x_1})\, d\mu_1(x_1).$$

Montrons que m_1 est une mesure.

i) $m_1(\varnothing) = 0$ car toutes les sections de l'ensemble vide sont vides.

ii) Soit $(C^{(n)})$ une suite d'éléments de $\mathscr{A}_1 \otimes \mathscr{A}_2$ deux à deux disjoints. Alors pour tout $x_1 \in E_1$, les sections $(C^{(n)}_{x_1})$ sont deux à deux disjointes et $\left(\cup_n C^{(n)}\right)_{x_1} = \cup_n C^{(n)}_{x_1}$, si bien que, par le théorème de Beppo Levi,

$$m_1\left(\cup_n C^{(n)}\right) = \int_{E_1} \left(\sum_n \mu_2\left(C^{(n)}_{x_1}\right)\right) d\mu_1(x_1) = \sum_n \int_{E_1} \mu_2\left(C^{(n)}_{x_1}\right)\, d\mu_1(x_1) = \sum_n m_1\left(C^{(n)}\right),$$

ce qui prouve la σ-additivité de m_1. De plus, pour tous $A_1 \in \mathscr{A}_1$ et $A_2 \in \mathscr{A}_2$, et pour tout $x_1 \in E_1$, la section $(A_1 \times A_2)_{x_1}$ vaut A_2 si $x_1 \in A_1$ et est vide sinon. Ainsi,

$$m_1(A_1 \times A_2) = \int_{A_1} \mu_2(A_2)\, d\mu_1(x_1) = \mu_1(A_1)\mu_2(A_2).$$

Il existe donc bien une mesure $m = m_1$ satisfaisant $m(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$ et cette mesure vérifie

$$m(C) := \int_{E_1} \mu_2(C_{x_1})\, d\mu_1(x_1).$$

De même, on définit la mesure m_2 par

$$m_2(C) := \int_{E_2} \mu_1(C^{x_2})\, d\mu_2(x_2),$$

et l'on montre que m_2 est une mesure qui coïncide avec m sur $\mathscr{A}_1 \times \mathscr{A}_2$, donc est égale à m (cf. a)). On se reportera aussi à a) pour voir que m est σ-finie. □

12.3 Théorèmes de Fubini

12.3.1 Théorème de Fubini–Tonelli

Théorème 12.18 (de Fubini–Tonelli) *Si $f : (E_1 \times E_2, \mathscr{A}_1 \otimes \mathscr{A}_2) \to (\bar{\mathbb{R}}_+, \mathscr{B}or(\bar{\mathbb{R}}_+))$ est mesurable, alors les fonctions ϕ et ψ définies resp. sur E_1 et E_2 par*

$$\phi(x_1) := \int_{E_2} f(x_1, x_2)\, d\mu_2(x_2) \quad \text{et} \quad \psi(x_2) := \int_{E_1} f(x_1, x_2)\, d\mu_1(x_1)$$

sont toutes deux mesurables et l'on a la double égalité dans $\bar{\mathbb{R}}_+$

$$\int_{E_1} \phi\, d\mu_1 = \int_{E_1 \times E_2} f\, d(\mu_1 \otimes \mu_2) = \int_{E_2} \psi\, d\mu_2. \tag{12.1}$$

Dém. Si $f = \mathbb{1}_C$ pour $C \in \mathscr{A}_1 \otimes \mathscr{A}_2$, alors $\phi(x_1) = \mu_2(C_{x_1})$ et $\psi(x_2) = \mu_1(C_{x_2})$, et donc 12.13 assure que ϕ et ψ sont mesurables et les trois termes de l'équation (12.1) sont égaux à $\mu_1 \otimes \mu_2(C)$ par le théorème 12.14. Cette assertion s'étend aux fonctions étagées positives pas linéarité de l'intégrale, puis aux fonctions mesurables positives par le lemme fondamental d'approximation et le théorème de Beppo Levi. □

12.3.2 Théorème de Fubini–Lebesgue

Théorème 12.19 (de Fubini–Lebesgue) *Soit f comme dans le théorème de Fubini–Tonelli mais de signe quelconque. Alors si f est $\mu_1 \otimes \mu_2$-intégrable*[4]*, alors les fonctions ϕ et ψ du théorème sont resp. définies μ_1-p.p. et μ_2-p.p., sont resp. μ_1-intégrables et μ_2-intégrables, et vérifient la double égalité* (12.1).

Dém. On définit
$$\phi_+(x_1) = \int_{E_2} f^+(x_1, x_2)\, d\mu_2(x_2),$$
ainsi que de manière évidente, ϕ_-, ψ_+ et ψ_-. D'après le théorème de Fubini–Tonelli,
$$\int_{E_1} \phi_+\, d\mu_1 = \int_{E_1 \times E_2} f^+\, d(\mu_1 \otimes \mu_2) = \int_{E_2} \psi_+\, d\mu_2,$$
qui est un nombre réel fini par hypothèse (se référer au terme du milieu). Par conséquent, ϕ_+ est finie μ_1-p.p. et ψ_+ est finie μ_2-p.p., ainsi que ϕ_- et ψ_- respectivement. Donc la fonction ϕ est définie μ_1-p.p. (comme différence de deux fonctions finies p.p.) et l'intégrale de $|\,\phi\,|$ est finie car égale à la somme des intégrales de ϕ_+ et de ϕ_-, qui sont toutes deux finies. Le résultat analogue se démontre de la même manière pour ψ, et ainsi l'égalité (12.1) s'obtient en faisant la différence de deux quantités finies. □

Remarque 12.20 *Si f est positive, le théorème de Fubini–Tonelli assure que l'intégrale de f par rapport à $\mu_1 \otimes \mu_2$ peut toujours se calculer en faisant deux intégrales « simples » successives dans l'ordre que l'on souhaite. Si f est de signe quelconque, il faut, pour appliquer le théorème de Fubini-Lebesgue, d'abord vérifier l'intégrabilité de $|\,f\,|$ par rapport à $\mu_1 \otimes \mu_2$ en utilisant le théorème de Fubini–Tonelli (il suffit de vérifier que l'une des intégrales« doubles » est finie).*

Remarque 12.21 *Par convention d'écriture, on écrira invariablement :*
$$\begin{aligned}\int_{E_1 \times E_2} f\, d(\mu_1 \otimes \mu_2) &= \int_{E_1} d\mu_1(x_1) \int_{E_2} d\mu_2(x_2)\, f(x_1, x_2)\\ &= \int_{E_2} d\mu_2(x_2) \int_{E_1} d\mu_1(x_1)\, f(x_1, x_2) = \int_{E_1}\int_{E_2} d\mu_1(x_1) d\mu_2(dx_2)\, f(x_1, x_2),\end{aligned}$$
mais on évitera en général d'écrire des intégrales « multiples » comme
$$\int_E \cdots \int_E f(x_1, \ldots, x_n)\, d\mu(x_1) \cdots d\mu(x_n),$$

4. ce qui se vérifie grâce au théorème de Fubini–Tonelli...

auquel on préférera

$$\int_{E^n} f \, d\mu^{\otimes n}.$$

Remarque 12.22 *Si* $(a_{n,m})$ *est une suite doublement indicée de nombres réels positifs et si* μ *est la mesure de comptage sur* $(\mathbb{N}, \mathscr{P}(\mathbb{N}))$, *alors l'interversion suivante*

$$\sum_n \sum_m a_{n,m} = \sum_m \sum_n a_{n,m}$$

peut être vue comme une application du théorème de Beppo Levi, car $\sum_m a_{n,m} = \int_{\mathbb{N}} a_{n,m} \, d\mu(m)$, *comme du théorème de Fubini–Tonelli, car les termes de l'équation sont tous deux égaux à* $\int_{\mathbb{N}\times\mathbb{N}} a_{n,m} d\mu^{\otimes 2}(n, m)$.

Chapitre 13

Mesure image et changement de variable

13.1 Mesure image

Soient $(E_1, \mathscr{A}_1)$ et $(E_2, \mathscr{A}_2)$ deux espaces mesurables, μ une mesure sur $(E_1, \mathscr{A}_1)$ et $h \in \mathscr{F}(\mathscr{A}_1, \mathscr{A}_2)$.

Définition 13.1 (et proposition) *L'égalité*

$$\nu(A_2) := \mu\left(h^{-1}(A_2)\right) \qquad A_2 \in \mathscr{A}_2$$

définit une mesure ν sur $(E_2, \mathscr{A}_2)$ appelée mesure image *et notée $\mu \circ h^{-1}$, ou $h(\mu)$, ou encore μ_h.*

Dém. i) $\nu(\varnothing) = \mu(h^{-1}(\varnothing)) = \mu(\varnothing) = 0$.

ii) Soit (B_n) une suite d'éléments de $\mathscr{A}_2$ deux à deux disjoints, alors d'après les formules de Hausdorff, pour tous $i \neq j$,

$$h^{-1}(B_i) \cap h^{-1}(B_j) = h^{-1}(B_i \cap B_j) = h^{-1}(\varnothing) = \varnothing,$$

et ainsi

$$\nu\left(\cup_n B_n\right) = \mu\left(h^{-1}(\cup_n B_n)\right) = \mu\left(\cup_n h^{-1}(B_n)\right) = \sum_n \mu\left(h^{-1}(B_n)\right) = \sum_n \nu(B_n),$$

par σ-additivité de μ. □

Théorème 13.2 *Soit $f \in \mathscr{F}(\mathscr{A}_2, \mathscr{B}or(\mathbb{R}))$. Si f est positive μ_h-p.p. alors on a l'égalité suivante dans $\bar{\mathbb{R}}_+$*

$$\int_{E_2} f\, d\mu_h = \int_{E_1} f \circ h\, d\mu. \tag{13.1}$$

De même, f est μ_h-intégrable ssi $f \circ h$ est μ-intégrable, et si c'est le cas, on a l'égalité (13.1) *dans $\mathbb{R}$.*

Dém. Si $f = \mathbb{1}_B$, où $B \in \mathscr{A}_2$, alors

$$\int_{E_1} f \circ h \, d\mu = \int_{E_1} \mathbb{1}_B \circ h \, d\mu = \int_{E_1} \mathbb{1}_{h^{-1}(B)} \, d\mu = \mu\left(h^{-1}(B)\right) = \mu_h(B) = \int_{E_2} f \, d\mu_h.$$

La linéarité de l'intégrale, le lemme fondamental d'approximation et le théorème de Beppo Levi impliquent (13.1) dès que f est positive μ_h-p.p. L'extension aux fonctions mesurables de signe quelconque et l'équivalence des intégrabilités se déduisent classiquement @ de la décomposition $f = f^+ - f^-$. □

Application. Soit $h \in \mathscr{F}(\mathscr{A}, \mathscr{A})$ telle que $\mu_h = \mu$. Alors

$$\int_E f \, d\mu = \int_E f \circ h \, d\mu$$

pour toute fonction positive μ-p.p. ou μ-intégrable. En particulier, si μ est la mesure de Lebesgue sur $\mathbb{R}^d$ et $h = \tau_a$ est la translation de vecteur $a \in \mathbb{R}^d$, alors $\mu_h = \mu$ et donc

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

À partir de maintenant on supposera que $\mu = \lambda_d$, que l'on notera λ s'il n'y a pas d'ambiguïté.

Proposition 13.3 *Soit $A \in GL_d(\mathbb{R})$ et $b \in \mathbb{R}^d$. Soit h l'application affine définie par $h(x) = Ax + b$. Alors*

$$h(\lambda) = |\det A|^{-1} \lambda.$$

En particulier, pour tout f positive ou λ-intégrable, on a

$$\int_{\mathbb{R}^d} f \circ h \, d\lambda = \frac{1}{|\det A|} \int_{\mathbb{R}^d} f \, d\lambda.$$

Application (vue en détail en TD) : calcul du volume de la boule unité. Soit $B_n(r)$ la boule (centrée sur l'origine) de rayon r dans $\mathbb{R}^n$ muni de la norme euclidienne. Alors par la proposition précédente, si h est l'homothétie de paramètre r^{-1},

$$\mathrm{Vol}(B_n(r)) = \lambda(h^{-1}(B_n(1))) = |\det(h)|^{-1} \lambda(B_n(1)) = r^n \lambda(B_n(1)).$$

D'autre part si $c_n := \mathrm{Vol}(B_n(1))$, alors

$$c_n = \int_{-1}^{1} dx_1 \int_{x_2^2 + \cdots + x_n^2 \leq 1 - x_1^2} dx_2 \cdots dx_n = \int_{-1}^{1} dx_1 \mathrm{Vol}\left(B_{n-1}\left(\sqrt{1 - x_1^2}\right)\right) = c_{n-1} I_{n-1},$$

où l'on a défini

$$I_n := \int_{-1}^{1} \left(1 - x^2\right)^{n/2} \, dx.$$

Une intégration par parties permet de voir que $I_n = nI_{n-2}/(n+1)$, ce qui indique pourquoi le résultat dépend de la parité de n. En effet, après calculs, on obtient pour tout entier k

$$c_{2k} = \frac{\pi^k}{k!},$$

tandis que

$$c_{2k+1} = \frac{2^{k+1}\pi^k}{1 \cdot 3 \cdot 5 \cdots (2k+1)}.$$

On retrouve ainsi que $c_1 = 2$, $c_2 = \pi$ et $c_3 = 4\pi/3$.

Dém. de la proposition. Montrons qu'on peut supposer que $b = 0$. Admettons la proposition dans le cas où $b = 0$, c'est-à-dire que $\lambda \circ A^{-1} = |\det A|^{-1}\lambda$. Maintenant si $h(x) = Ax + b$, c'est-à-dire $h = \tau_b \circ A$, on a

$$\lambda \circ h^{-1} = \lambda \circ (\tau_b \circ A)^{-1} = \lambda \circ A^{-1} \circ \tau_b^{-1} = |\det A|^{-1}\lambda \circ \tau_b^{-1} = |\det A|^{-1}\lambda.$$

On peut donc supposer dorénavant que $b = 0$. Soit $\nu := A(\lambda)$. Il faut montrer que $\nu = |\det A|^{-1}\lambda$. Montrons d'abord que ν est invariante par translation. En effet, comme pour tout $c \in \mathbb{R}^d$, $A^{-1} \circ \tau_c^{-1}(x) = A^{-1}(x - c) = A^{-1}(x) - A^{-1}(c) = \tau_{-A^{-1}(c)} \circ A^{-1}(x)$,

$$\nu \circ \tau_c^{-1} = \lambda \circ A^{-1} \circ \tau_c^{-1} = \lambda \circ \tau_{-A^{-1}(c)} \circ A^{-1} = \lambda \circ A^{-1} = \nu.$$

Soit C_d le rectangle unité $[0,1]^d$. Montrons que $\nu(C_d) > 0$. Comme $\mathbb{R}^d \subseteq \cup_{x\in\mathbb{Z}^d}(x + C_d)$, par invariance par translation de ν,

$$\nu\left(\mathbb{R}^d\right) \leq \sum_{x\in\mathbb{Z}^d} \nu(x + C_d) = \sum_{x\in\mathbb{Z}^d} \nu(C_d),$$

donc si $\nu(C_d) = 0$, $\nu\left(\mathbb{R}^d\right) = 0$, ce qui n'est pas possible car $\nu\left(\mathbb{R}^d\right) = \lambda \circ A^{-1}\left(\mathbb{R}^d\right) = \lambda\left(\mathbb{R}^d\right) = +\infty$. Montrons que $\nu(C_d) < \infty$. L'application A^{-1} est linéaire, donc continue, donc l'image C'_d du compact C_d par A^{-1} est également compacte. Comme λ est finie sur les compacts, $\nu(C_d) = \lambda(C'_d) < \infty$.

Soit $c = c(A) = \lambda \circ A^{-1}\left([0,1]^d\right)$. D'après ce qui précède, $c \in]0,\infty[$ et $\nu' = c^{-1}\nu$ est une mesure invariante par translation telle que $\nu'\left([0,1]^d\right) = 1$, donc ν' est la mesure de Lebesgue sur $\mathbb{R}^d$. Il suffit donc de montrer que $c(A) = |\det A|^{-1}$. Montrons que c est un morphisme. Si φ_1 et φ_2 sont deux endomorphismes inversibles de $\mathbb{R}^d$, alors d'après ce qui précède,

$$\lambda \circ (\varphi_1 \circ \varphi_2)^{-1} = c(\varphi_1 \circ \varphi_2)\,\lambda,$$

mais également

$$\lambda \circ (\varphi_1 \circ \varphi_2)^{-1} = \lambda \circ \varphi_2^{-1} \circ \varphi_1^{-1} = c(\varphi_2)\,\lambda \circ \varphi_1^{-1} = c(\varphi_2)\,c(\varphi_1)\,\lambda,$$

ce qui implique effectivement que $c(\varphi_1 \circ \varphi_2) = c(\varphi_1)\,c(\varphi_2)$. Comme tout endomorphisme inversible Φ de $\mathbb{R}^d$ s'écrit comme produit fini d'endomorphismes du type φ_1, φ_2, φ_3, où (en écrivant e_i le i-ème vecteur de la base canonique)

$$\varphi_1(e_i) = e_{\sigma(i)} \quad \text{pour } \sigma \text{ une permutation de } \{1, \ldots, d\},$$

$$\varphi_2(e_1) = \alpha e_1 \quad \text{et} \quad \varphi_2(e_j) = e_j \ \forall j \neq 1 \quad (\text{avec } \alpha \neq 0),$$
$$\varphi_3(e_1) = e_1 + e_2, \quad \text{et} \quad \varphi_3(e_j) = e_j \ \forall j \neq 1,$$

il suffit de montrer que $c(\varphi) = |\det \varphi|^{-1}$ pour chacune de ces trois (sortes d')applications. En effet, ceci étant démontré, nous aurons pour $\Phi = \Pi_{i=1}^n \phi_i$, où les ϕ_i sont du type ci-avant (et où le produit est un produit matriciel, c'est à-dire une composition),

$$c(\Phi) = c\left(\Pi_{i=1}^n \phi_i\right) = \Pi_{i=1}^n c(\phi_i) = \left(\Pi_{i=1}^n |\det \phi_i|^{-1}\right) = |\Pi_{i=1}^n \det \phi_i|^{-1} = |\det \Phi|^{-1}.$$

Le rectangle unité C_d est invariant par φ_1 donc $c(\varphi_1) = \lambda(C_d) = 1 = |\det \varphi_1|^{-1}$. Dans le cas de φ_2,

$$\varphi_2^{-1}(C_d) = I_\alpha \times C_{d-1},$$

où $I_\alpha = [0, 1/\alpha]$ si $\alpha > 0$ et $I_\alpha = [1/\alpha, 0]$ si $\alpha < 0$. Par conséquent $c(\varphi_2) = \lambda_1(I_\alpha)\lambda_{d-1}(C_{d-1}) = |\alpha|^{-1} = |\det \varphi_2|^{-1}$. Enfin,

$$\varphi_3^{-1}(C_d) = P_2 \times C_{d-2},$$

où P_2 est un losange du plan d'aire 1, donc $c(\varphi_3) = \lambda_2(P_2)\lambda_{d-2}(C_{d-2}) = 1 = |\det \varphi_3|^{-1}$, ce qui achève la démonstration. □

13.2 Formule du changement de variable

Soient U et V deux ouverts de $\mathbb{R}^d$ et ϕ un C^1-difféomorphisme entre U et V, c'est-à-dire une bijection $\phi : U \to V$ telle que ϕ est de classe C^1 sur U et ϕ^{-1} est de classe C^1 sur V. Pour tout $u \in U$, on note $\phi'(u)$ la matrice carrée $d \times d$ des dérivées partielles de ϕ évaluées en u, autrement dit la matrice représentative de l'application linéaire tangente à ϕ en u, appelée *matrice jacobienne* de ϕ en u. On note $J_\phi(u)$ le déterminant de $\phi'(u)$, appelé *jacobien* de ϕ en u. Nous allons montrer que l'image par ϕ de la mesure de densité $|J_\phi|$ par rapport à λ (sur U) est λ (sur V), et que l'image par ϕ de λ (sur U) est la mesure de densité $|J_{\phi^{-1}}|$ par rapport à λ (sur V).

Théorème 13.4 (formule de changement de variable) *Soit f une fonction borélienne sur V. Si f est positive ou λ-intégrable, alors*

$$\int_V f\, d\lambda = \int_U f \circ \phi\, |J_\phi|\, d\lambda.$$

De manière équivalente, si f est positive ou que $f \circ \phi$ est λ-intégrable, alors

$$\int_U f \circ \phi\, d\lambda = \int_V f\, |J_{\phi^{-1}}|\, d\lambda.$$

Remarque 13.5 *En dimension 1, si $\phi :]\alpha, \beta[\longrightarrow]a, b[$ est un C^1-difféomorphisme, alors ϕ' ne peut pas s'annuler et en particulier, ϕ est de signe constant. Avec les notations de l'intégrale de Riemann, si l'on applique la formule de changement de variable apprise au lycée, on retombe bien entendu sur la formule du théorème précédent. Si $\phi' > 0$, alors*

$$\int_a^b f(x)\, dx = \int_{\phi^{-1}(a)}^{\phi^{-1}(b)} f \circ \phi(u)\, \phi'(u)\, du = \int_\alpha^\beta f \circ \phi(u)\, |\phi'(u)|\, du.$$

Si $\phi' < 0$, alors

$$\int_a^b f(x)\,dx = \int_{\phi^{-1}(a)}^{\phi^{-1}(b)} f\circ\phi(u)\,\phi'(u)\,du = -\int_\alpha^\beta f\circ\phi(u)\,\phi'(u)\,du = \int_\alpha^\beta f\circ\phi(u)\,|\,\phi'(u)\,|\,du.$$

Corollaire 13.6 *Si μ est la mesure de densité f par rapport à λ et si $\phi : \mathbb{R}^d \to \mathbb{R}^d$ est un C^1-difféomorphisme, alors la mesure image de μ par ϕ admet une densité g par rapport à λ, et g est donnée par*

$$g(x) = f\circ\phi^{-1}(x)\,|\,J_{\phi^{-1}}(x)\,| \qquad x \in \mathbb{R}^d.$$

Dém. Pour toute fonction borélienne positive h, en se servant du théorème précédent,

$$\int h\,d\mu_\phi = \int h\circ\phi\,f\,d\lambda = \int h\circ\phi\,f\circ\phi^{-1}\circ\phi\,d\lambda = \int h\,f\circ\phi^{-1}\,|\,J_{\phi^{-1}}\,|\,d\lambda,$$

ce qui prouve le corollaire (prendre tout simplement une indicatrice pour h). □

Remarque 13.7 *Pour vérifier que ϕ est un C^1-difféomorphisme, on applique ordinairement le théorème d'inversion locale : soit U un ouvert de $\mathbb{R}^d$ et $\phi : U \to \mathbb{R}^d$. Soit $V := \phi(U)$. Alors ϕ est un C^1-difféomorphisme ssi*

i) ϕ est injective ;

ii) ϕ est de classe C^1 ;

iii) pour tout $u \in U$, $J_\phi(u) \neq 0$.

Sous ces conditions, V est un ouvert et pour tout $x \in V$, $\left(\phi^{-1}\right)'(x) = \left(\phi'\circ\phi^{-1}(x)\right)^{-1}$.

Exemple 13.8 (coordonnées polaires) *Par le théorème d'inversion locale, la fonction*

$$\begin{aligned}\phi :]0,\infty[\times]0,2\pi[&\longrightarrow \mathbb{R}^2 \setminus ([0,\infty[\times\{0\})\\ (\rho,\theta) &\longmapsto (\rho\cos\theta, \rho\sin\theta)\end{aligned}$$

est un C^1-difféomorphisme, avec

$$\phi'(\rho,\theta) = \begin{pmatrix} \cos\theta & -\rho\sin\theta \\ \sin\theta & \rho\cos\theta \end{pmatrix}$$

et $J_\phi(\rho,\theta) = \rho$. Ainsi pour toute fonction borélienne f λ_2-intégrable,

$$\begin{aligned}\int_{\mathbb{R}^2} f(x,y)\,dx\,dy &= \int_{\mathbb{R}^2\setminus([0,\infty[\times\{0\})} f(x,y)\,dx\,dy\\ &= \int_{]0,\infty[\times]0,2\pi[} f\circ\phi(\rho,\theta)\,J_\phi(\rho,\theta)\,d\rho\,d\theta = \int_{[0,\infty[\times[0,2\pi]} f\circ\phi(\rho,\theta)\,\rho\,d\rho\,d\theta.\end{aligned}$$

En particulier, l'intégrale $I = \int_{\mathbb{R}} e^{-x^2}\,dx$ peut se calculer comme suit, grâce à deux applications du théorème de Fubini–Tonelli :

$$I^2 = \int_{\mathbb{R}^2} e^{-(x^2+y^2)}dx\,dy = \int_{[0,\infty[\times[0,2\pi]} e^{-\rho^2}\rho\,d\rho\,d\theta = 2\pi\int_{[0,\infty[} e^{-\rho^2}\rho\,d\rho = \pi,$$

d'où l'égalité bien connue $I = \sqrt{\pi}$.

La démonstration de la formule du changement de variable est plutôt technique. Nous renvoyons le lecteur à la démonstration par récurrence p.64 du cours de Jean Jacod, ou à la démonstration p.242 du livre de Marc Briane et Gilles Pagès. Nous donnons ci-après l'idée de cette dernière. On recouvre l'ouvert U par une réunion dénombrable d'hyper-cubes semi-ouverts (C_i) deux à deux disjoints et de mesure de Lebesgue arbitrairement petite fixée. On note u_i le centre de C_i. Comme ϕ est bijective, $V = \phi(U)$ s'écrit à son tour comme réunion disjointe des $\phi(C_i)$, donc pour toute fonction borélienne f positive,

$$\int_V f\,d\lambda = \sum_i \int_{\phi(C_i)} f\,d\lambda \approx \sum_i f(\phi(u_i))\lambda(\phi(C_i)).$$

Mais localement, ϕ peut être approchée par son application linéaire tangente $\phi'(u_i)$, aussi comme $\lambda(\phi(C_i))$ est la mesure de C_i par la mesure image de λ par ϕ^{-1}, ayant $\phi^{-1}(x) \approx Ax + b$, avec $A = (\phi^{-1})'$ (et $b = \phi^{-1}(u_i) - Au_i$), on a

$$\lambda(\phi(C_i)) \approx |\det A|^{-1}\lambda(C_i) = |\det \phi'(u_i)|\lambda(C_i).$$

Ainsi,

$$\begin{aligned}\int_V f\,d\lambda &\approx \sum_i f(\phi(u_i))\,|\,J_\phi(u_i)\,|\,\lambda(C_i)\\ &\approx \sum_i \int_{C_i} f\circ\phi(u)\,|\,J_\phi(u)\,|\,d\lambda(u)\\ &= \int_U f\circ\phi\,|\,J_\phi\,|\,d\lambda,\end{aligned}$$

ce qui achève cette esquisse de démonstration.

Chapitre 14

Les espaces L^p

Dans tout ce chapitre, on se place sur un espace mesuré $(E, \mathscr{A}, \mu)$ et pour tout $p \in \bar{\mathbb{R}}_+$, on abrégera $L^p(E, \mathscr{A}, \mu)$ en $L^p(\mu)$, voire en L^p. On désignera par (f_n) une suite de fonctions mesurables à valeurs dans $\bar{\mathbb{R}}$ muni de sa tribu borélienne.

14.1 Les espaces de Banach L^p

14.1.1 Convergence dans L^p et convergence simple

Rappelons que la topologie usuelle d'un espace vectoriel normé est la topologie relative à la distance $d(f, g) = \| f - g \|$. Ainsi on dira que la suite (f_n) converge dans L^p si

a) pour tout $n \in \mathbb{N}$, $f_n \in L^p$ et $f \in L^p$;

b) $\lim_n \| f - f_n \|_p = 0$.

On rappelle que la suite (f_n) converge simplement vers f si $\lim_n f_n(x) = f(x)$ pour μ-presque tout x.

Proposition 14.1 *Soit $p \in [1, +\infty[$.*

a) [convergence L^p-dominée] Si $f_n \to f$ μ-p.p. et qu'il existe $g \in L^p$ tel que $| f_n | \leq g$ pour tout entier n, alors $f_n \overset{L^p}{\to} f$.

b) i) Si $f_n \overset{L^p}{\to} f$, alors il existe une suite extraite de (f_n) qui converge vers f μ-p.p.

b) ii) Si $f_n \overset{L^\infty}{\to} f$, alors $f_n \to f$ uniformément en dehors d'un ensemble négligeable, donc $f_n \to f$ μ-p.p.

Remarque 14.2 *Dans le cas de l'espace ℓ^p (pour $p < \infty$), une suite (de fonctions, aussi appelées suites ici...) $(u^{(n)})$ converge vers la fonction $u \in \ell^p$ si $\sum_k | u_k^{(n)} |^p < \infty$, si $\sum_k | u_k |^p < \infty$ et si*

$$\lim_n \sum_k | u_k^{(n)} - u_k |^p = 0.$$

Ceci implique en particulier que $u_k^{(n)} \longrightarrow u_k$ lorsque $n \to \infty$. En conclusion, dans l'espace ℓ^p (vrai aussi si $p = +\infty$ par b)ii)),

$$f_n \overset{\ell^p}{\to} f \qquad \Longrightarrow \qquad f_n \to f \qquad \textit{simplement (partout).}$$

Évidemment, on n'a pas la réciproque, comme on peut le voir sur le contre-exemple $u^{(n)} = \mathbb{1}_{\{n\}}$. Alors la suite $(u^{(n)})$ converge simplement vers la fonction nulle car $u_k^{(n)} = 0$ pour tout $k > n$. Néanmoins pour tout n, la fonction $u^{(n)}$ est à distance 1 de la fonction nulle : $\| u^{(n)} - 0 \|_p = (\sum_k | u_k^{(n)} |^p)^{1/p} = 1$ pour tout p (même $p = \infty$), et donc ne converge pas vers la suite nulle dans ℓ^p. En effet, ici la plus petite fonction dominant la suite $(u^{(n)})$ est la fonction v constante à 1. Pour $p < \infty$, cette fonction n'est pas dans ℓ^p, donc on ne peut pas appliquer a). De plus, $v \in \ell^\infty$, ce qui montre aussi que a) n'est pas vrai en général pour $p = \infty$.

Dém. a) On applique le théorème de convergence dominée. En effet, $| f_n - f |^p \leq (| f_n | + | f |)^p \leq 2^p | g |^p$ μ-p.p., et par hypothèse $| g |^p$ est intégrable, donc comme $| f_n - f |^p \to 0$, μ-p.p., on a la convergence vers 0 de $\int | f_n - f |^p \, d\mu$.

b)i) On applique le lemme de Borel–Cantelli. Construisons une suite extraite $(f_{\varphi(n)})$ par récurrence. Soit $\varphi(0) = 0$ et

$$\varphi(n+1) := \min\{k > \varphi(n) : \| f - f_k \|_p \leq 2^{-(n+1)}\},$$

qui est toujours un nombre fini puisque $\| f - f_n \|_p \to 0$. On a donc $\| f - f_{\varphi(n)} \|_p \leq 2^{-n}$ et pour tout $\varepsilon > 0$,

$$\mu(| f - f_n | \geq \varepsilon) = \mu(| f - f_n |^p \geq \varepsilon^p) \leq \varepsilon^{-p} \| f - f_n \|_p^p,$$

par l'inégalité de Markov. Par conséquent, avec $A_n(\varepsilon) := \{| f - f_{\varphi(n)} | \geq \varepsilon\}$,

$$\sum_n \mu(A_n(\varepsilon)) \leq \sum_n \varepsilon^{-p} \| f - f_{\varphi(n)} \|_p^p \leq \sum_n \varepsilon^{-p} 2^{-np} < \infty,$$

et le lemme de Borel–Cantelli assure donc que $\mu(\limsup_n A_n(\varepsilon)) = 0$. Définissons alors

$$B_\varepsilon := \limsup_n A_n(\varepsilon) \quad \text{et} \quad B := \cup_{\varepsilon > 0} B_\varepsilon.$$

Remarquons que B est bien mesurable puisque $B = \lim_{k \in \mathbb{N}^\star} \uparrow B_{1/k}$ et que $\mu(B) \leq \sum_k \mu(B_{1/k}) = 0$. Concluons en remarquant que pour tout $x \in {}^c\!B$, $f_{\varphi(n)}(x) \to f(x)$. En effet, pour tout ε, $x \in {}^c\!B_\varepsilon = \liminf_n {}^c\!A_n(\varepsilon)$, donc il existe n_0 tel que pour tout $n \geq n_0$, $x \in {}^c\!A_n(\varepsilon)$, c'est-à-dire $| f - f_{\varphi(n)} |(x) < \varepsilon$.

b)ii) Dire que $f_n \to f$ dans L^∞, c'est dire qu'il existe une partie mesurable Ω de complémentaire négligeable telle que $\sup_{x \in \Omega} | f_n - f |(x) \to 0$. En effet, par définition du supremum essentiel, pour tout entier n il existe une partie mesurable Ω_n de complémentaire négligeable telle que $| f_n - f |(x) \leq \| f_n - f \|_\infty$. Si l'on définit $\Omega := \cap_n \Omega_n$, Ω est bien de complémentaire négligeable et

$$\sup_{x \in \Omega} | f_n - f |(x) \leq \| f_n - f \|_\infty \to 0.$$

La suite (f_n) converge donc uniformément vers f sur Ω, donc elle converge simplement sur Ω. □

Corollaire 14.3 *Soit $p \in [1, +\infty]$. Si l'on a la convergence de la suite (f_n) vers f dans L^p et vers g μ-p.p. alors f et g sont égales μ-p.p.*

Dém. On sait qu'il existe une suite extraite $(f_{\varphi(n)})$ qui converge μ-p.p. vers f. Or la suite (f_n) converge μ-p.p. vers g, donc la sous-suite $(f_{\varphi(n)})$ également. Ainsi $f = g$ μ-p.p. □

14.1.2 Complétude des espaces L^p

Lemme 14.4 *Soit $(H, \|\cdot\|)$ un e.v. normé. Si toute série de terme général (u_n) telle que $\sum_n \| u_n \| < \infty$ est convergente*[1]*, alors $(H, \|\cdot\|)$ est complet.*

Dém. Soit (x_n) une suite de Cauchy à valeurs dans $(H, \|\cdot\|)$. On définit alors par récurrence une injection croissante φ de $\mathbb{N}$, par $\varphi(0) = 0$, et

$$\varphi(n+1) := \min\{k > \varphi(n) : \forall j, \ell \geq k, \| x_j - x_\ell \| \leq 2^{-(n+1)}\},$$

qui est un nombre fini puisque (x_n) est de Cauchy. Soit alors $u_n := x_{\varphi(n+1)} - x_{\varphi(n)}$. Alors $x_{\varphi(n)} = x_0 + \sum_{k=0}^{n-1} u_k$ et

$$\| u_n \| = \| x_{\varphi(n+1)} - x_{\varphi(n)} \| \leq 2^{-n},$$

donc $\sum_n \| u_n \| < \infty$. Par hypothèse, la suite des sommes partielles $(\sum_{k=0}^n u_k)$ converge donc dans H, et par conséquent, la suite $(x_{\varphi(n)})$ également. En conclusion, (x_n) est une suite de Cauchy dont une suite extraite converge, c'est donc une suite convergente. □

Théorème 14.5 (de Riesz–Fisher) *Pour tout $p \in [1, +\infty]$, $L^p(\mu)$ est un espace de Banach*[2]*.*

Dém. Nous allons montrer que $L^p(\mu)$ vérifie les hypothèses du lemme précédent. Soit (u_n) une suite de L^p telle que $\sum_n \| u_n \|_p =: a < \infty$ et soit $f_n := \sum_{k=0}^n u_k$ la n-ième somme partielle. Il nous suffit donc de montrer que la suite (f_n) converge dans L^p. Soit alors

$$h_n := \sum_{k=0}^{n} | u_k |.$$

La suite (h_n) est une suite croissante, qui converge donc simplement vers la fonction positive $h := \lim_n \uparrow h_n$. Montrons d'abord que $h \in L^p$ avec $\| h \|_p \leq a$ en différenciant suivant que p est fini ou infini. Si $p = \infty$, il existe une partie négligeable N de E telle que pour tout $x \in {}^cN$, $| u_n(x) | \leq \| u_n \|_\infty$, et donc

$$h_n(x) \leq \sum_{k=0}^{n} | u_k(x) | \leq \sum_{k=0}^{n} \| u_k \|_\infty \leq a.$$

Par conséquent pour tout $x \in {}^cN$, $h(x) = \lim_n h_n(x) \leq a$, et $h \in L^\infty$ avec $\| h \|_\infty \leq a$. Si $p < \infty$, l'inégalité triangulaire implique que $\| h_n \|_p \leq \sum_{k=0}^n \| u_n \|_p \leq a$, donc par le

1. au sens où la suite des sommes partielles converge dans H muni de sa norme $\|\cdot\|$
2. c'est-à-dire un espace vectoriel normé *complet*

théorème de convergence monotone,

$$\int_E h^p \, d\mu = \lim_n \int_E h_n^p \, d\mu = \lim_n \| h_n \|_p^p \leq a^p.$$

Par conséquent, $\| h \|_p \leq a$ ici encore.

Il existe donc une partie négligeable N de E telle que pour tout $x \in {}^cN$, $h(x) = \sum_n | u_n(x) | < \infty$. Autrement dit, pour tout $x \in {}^cN$, la série de terme général $u_n(x)$ est absolument convergente (dans $\mathbb{R}$) et est donc convergente. Autrement dit la suite (f_n) converge μ-p.p., vers une certaine fonction f. D'après l'inégalité triangulaire de $\mathbb{R}$, $| f_n(x) | \leq h_n(x) \leq h(x)$, donc $| f | \leq h$ μ-p.p., et par conséquent, $\| f \|_p \leq \| h \|_p \leq a$, ce qui s'écrit

$$\| \sum_{k \geq 0} u_k \|_p \leq \sum_{k \geq 0} \| u_k \|_p.$$

Si l'on applique cette inégalité à la suite $u' = (u_{n+k+1})_k$, ayant $f - f_n = \sum_k u'_k$, on obtient

$$\| f - f_n \|_p \leq \sum_{k \geq n+1} \| u_k \|_p,$$

qui tend vers 0 lorsque $n \to \infty$, ce qui revient à dire que la suite (f_n) converge vers f dans L^p. □

14.2 L'espace L^2 et les espaces de Hilbert

14.2.1 L'espace de Hilbert $L^2(\mu)$

Définition 14.6 (et proposition) *Soit H un e.v. réel et*

$$\begin{aligned} \langle \cdot, \cdot \rangle : H \times H &\longrightarrow \mathbb{R} \\ (u, v) &\longmapsto \langle u, v \rangle \end{aligned}$$

une forme bilinéaire symétrique positive*, c'est-à-dire telle que*

i) positivité : $\langle u, u \rangle \geq 0$ pour tout $u \in H$;

ii) symétrie : $\langle u, v \rangle = \langle v, u \rangle$ pour tous $u, v \in H$;

iii) bilinéarité : pour tout $v \in H$, l'application $u \mapsto \langle u, v \rangle$ est linéaire.

Si de plus on a l'implication $[\langle u, u \rangle = 0 \;\Rightarrow\; u = 0]$, alors on parle de forme bilinéaire symétrique strictement positive*, ou* définie positive*, ou plus simplement de* produit scalaire*. Lorsque c'est le cas, on vérifie facilement que $\| u \| := \sqrt{\langle u, u \rangle}$ définit une norme sur H, que l'on appelle alors espace* préhilbertien*. Si de plus H est complet, on dit que H est un* espace de Hilbert*.* @

Remarque 14.7 *Dans le cas complexe, on parle de forme* sesquilinéaire*, et la propriété ii) doit être remplacée par ii') $\langle u, v \rangle = \overline{\langle v, u \rangle}$.*

Théorème 14.8 *L'application*

$$\begin{aligned}\langle\cdot,\cdot\rangle : L^2(\mu)\times L^2(\mu) &\longrightarrow \mathbb{R}\\ (f,g) &\longmapsto \langle f,g\rangle := \int_E fg\,d\mu\end{aligned}$$

est (bien définie grâce à l'inégalité de Hölder et est) un produit scalaire. D'après le théorème de Riesz–Fisher, l'espace L^2 muni de ce produit scalaire est donc un espace de Hilbert.

Remarque 14.9 *L'inégalité de Hölder (avec $p = q = 1/2$) est appelée ici inégalité de Cauchy-Schwarz des espaces préhilbertiens :*

$$|\langle f,g\rangle| \leq \| f \|_2 \, \| g \|_2.$$

La démonstration du théorème est laissée au lecteur. @

Remarque 14.10 *Dans $L^2_{\mathbb{C}}(\mu)$, le produit scalaire $\langle f,g\rangle$ est défini par*

$$\langle f,g\rangle = \int_E f\,\overline{g}\,d\mu.$$

14.2.2 Théorème de projection

Soit C une partie *fermée* et *convexe*[3] d'un espace de Hilbert $(H,\langle\cdot,\cdot\rangle)$.

Théorème 14.11 *Pour tout $x\in H$, il existe un unique élément z de C, appelé* projection orthogonale *de x sur C, tel que*

$$\| x - z \| = \inf_{y\in C} \| x - y \|.$$

Dém. Le théorème repose sur l'identité du parallélogramme : pour tous $u, v \in H$,

$$\| u+v \|^2 + \| u-v \|^2 = 2\| u \|^2 + 2\| v \|^2,$$

qui est une conséquence immédiate de la bilinéarité du produit scalaire. @

Soit $d := \inf_{y\in C} \| x-y \|$ la distance de x à C. Par définition de la borne inférieure, il existe une suite (y_n) de C telle que $\lim_n \| x - y_n \| = d$. Nous allons nous servir de l'identité du parallélogramme pour montrer qu'une telle suite est de Cauchy, en posant $u = \frac{1}{2}(y_n - x)$ et $v = \frac{1}{2}(y_m - x)$ pour m, n deux entiers quelconques. On obtient alors

$$\| \frac{1}{2}(y_n + y_m) - x \|^2 + \frac{1}{4}\| y_n - y_m \|^2 = \frac{1}{2}\| y_n - x \|^2 + \frac{1}{2}\| y_m - x \|^2.$$

3. c'est-à-dire que pour tous $x, y \in C$ et pour tout $\lambda\in[0,1]$, la 'combinaison convexe' $\lambda x + (1-\lambda)y$ est toujours dans C, autrement dit C contient tous les segments joignant deux de ses éléments

Or C est convexe, donc $\frac{1}{2}(y_n + y_m) \in C$ et par conséquent $\| \frac{1}{2}(y_n + y_m) - x \| \geq d$. Ceci donne

$$\| y_n - y_m \|^2 \leq 2\| y_n - x \|^2 + 2\| y_m - x \|^2 - 4d^2.$$

Soit $\varepsilon > 0$. Comme $\lim_n \| x - y_n \|^2 = d^2$, il existe un entier N tel que pour tout $n \geq N$, $\| x - y_n \|^2 < d^2 + \frac{\varepsilon^2}{4}$. Grâce à l'inégalité obtenue précédemment, pour tous $n, m \geq N$, $\| y_n - y_m \| < \varepsilon$, autrement dit la suite (y_n) est de Cauchy. Comme H est complet, cette suite converge vers un certain z qui doit appartenir à C, puisque C est fermé. Ayant $\lim_n \| x - y_n \| = d$, par continuité de la norme, $\| x - z \| = d$.

Il ne reste plus à montrer que z est unique. Soit $z' \in C$ tel que $\| x - z' \| = d$. On définit une nouvelle suite (y_n) à valeurs dans C par $y_n = z$ si n est pair et $y_n = z'$ si n est impair. Alors $\| x - y_n \| = d$, donc par le même raisonnement que précédemment (y_n) converge, ce qui implique que $\| z' - z \| = 0$, c'est-à-dire $z' = z$. □

Proposition 14.12 *Soit F un sous-espace vectoriel fermé*[4] *de H. Alors pour tout $x \in H$, la projection orthogonale $p_F(x)$ de x sur F vérifie*

$$\langle x - p_F(x), y \rangle = 0 \qquad \textit{pour tout } y \in F.$$

Dém. Soient $y \in F$ et $t \in \mathbb{R}$. On note $z = p_F(x)$ comme dans le théorème de projection et $u = x - z$. Comme $z + ty \in F$ et que z minimise la distance de x à F,

$$\| u \|^2 = \| x - z \|^2 \leq \| x - (z + ty) \|^2 = \| u - ty \|^2 = \| u \|^2 - 2t\langle u, y \rangle + t^2\| y \|^2.$$

Donc $-2t\langle u, y \rangle + t^2\| y \|^2 \geq 0$ pour tout réel t, ce qui implique que $\langle u, y \rangle = 0$. □

Corollaire 14.13 *Si F est un sous-espace vectoriel fermé de H, alors H se décompose en somme directe suivant $F \oplus F^\perp$, où*

$$F^\perp := \{u \in H : \ \forall y \in F, \ \langle u, y \rangle = 0\}.$$

Dém. Soit $x \in H$. En notant comme précédemment $u = x - p_F(x)$, on peut écrire x sous la forme d'une somme $u + p_F(x)$, où $p_F(x) \in F$, et d'après la proposition précédente $u \in F^\perp$. D'autre part $F \cap F^\perp$ est bien réduit au singleton $\{0\}$ car pour tout $x \in F \cap F^\perp$, $\langle x, x \rangle = 0$, donc $x = 0$. □

Remarque 14.14 *Un développement très intéressant de cette section d'algèbre bilinéaire, que nous avons pourtant choisi de ne pas suivre, concerne les parties totales et les bases orthonormales de H. On dit qu'une partie K de H est* totale *si le sous-espace vectoriel* fermé *engendré par K, c'est-à-dire l'adhérence du sous-espace vectoriel engendré par K*[5] *est H tout entier. On appelle* système orthonormal *une partie de H dont tous les éléments sont de norme 1 et orthogonaux deux à deux ; on appelle* base orthonormale *un système orthonormal total. On peut démontrer que K est totale ssi $K^\perp = \{0\}$. De plus, si*

4. un sous-espace vectoriel est toujours convexe car une combinaison convexe est un cas particulier de combinaison linéaire ; un sous-espace vectoriel n'est forcément fermé que si H est de dimension finie

5. qui n'est lui-même que l'ensemble des combinaisons linéaires *finies* d'éléments de K

H_u est l'espace vectoriel fermé engendré par un système orthonormal dénombrable (u_n), alors l'application qui à une suite (a_n) de ℓ^2 associe $\sum_n a_n u_n \in H_u$ est un isomorphisme d'espaces de Hilbert (préserve le produit scalaire). Enfin, si μ est une mesure σ-finie, alors l'espace de Hilbert $L^2(\mu)$ admet une base orthonormale dénombrable.

14.2.3 Lemme de Riesz–Fisher

Nous allons appliquer les résultats de la section précédente à l'espace de Hilbert $H = L^2(E, \mathscr{A}, \mu)$.

Définition 14.15 *Soit H un $\mathbb{K}$-espace vectoriel ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$). On rappelle qu'une* forme linéaire *sur H est une application linéaire $\Phi : H \longrightarrow \mathbb{K}$ et que le* dual topologique *H' de H est l'espace vectoriel des formes linéaires* continues *Φ sur H, muni de la norme usuelle des espaces vectoriels d'applications linéaires, c'est-à-dire*

$$\| \Phi \| = \sup_{x \in H : x \neq 0} \frac{| \Phi(x) |}{\| x \|} = \sup_{x \in H : \| x \| = 1} | \Phi(x) |.$$

Exemple 14.16 (fondamental) *Soit $g \in L^2(\mu)$ et Φ_g l'application définie par*

$$\begin{aligned} \Phi_g : L^2(\mu) &\longrightarrow \mathbb{R} \\ f &\longmapsto \int_E f g \, d\mu. \end{aligned}$$

L'inégalité de Hölder assure que Φ_g est bien définie. La linéarité de l'intégrale assure aussi que Φ_g est linéaire. Grâce encore à l'inégalité de Hölder, $| \Phi_g(f) | \leq \| g \|_2 \| f \|_2$, si bien que Φ_g est $\| g \|_2$-lipschitzienne (et de norme $\| \Phi_g \| \leq \| g \|_2$), donc continue. Ainsi Φ_g est un élément du dual topologique de $L^2(\mu)$. Le lemme de Riesz–Fisher établit la réciproque de cet exemple. On remarquera également que si Φ_g n'est pas nulle, c'est-à-dire si g n'est pas nulle, alors si l'on choisit $f = g / \| g \|_2$, on a $\Phi_g(f) = \| g \|_2$, et donc $\| \Phi_g \| = \| g \|_2$.

Théorème 14.17 (Lemme de Riesz–Fisher) *Soit $\Phi : L^2(\mu) \longrightarrow \mathbb{R}$ une forme linéaire continue. Alors $\exists ! \varphi \in L^2(\mu)$ tel que*

$$\text{pour tout } f \in L^2(\mu), \qquad \Phi(f) = \int_E f \varphi \, d\mu.$$

Remarque 14.18 *Le même énoncé est vrai dans $\mathbb{C}$.*

Remarque 14.19 *D'après le lemme de Riesz–Fisher, si l'on note $H = L^2(\mu)$, l'application*

$$\begin{aligned} I : H &\longrightarrow H' \\ g &\longmapsto \Phi_g \end{aligned}$$

est un isomorphisme (et même une isométrie d'après l'exemple qui précède l'énoncé du lemme). Ainsi $L^2(\mu)$ est isomorphe (isométriquement) à son dual topologique. On parle de dualité L^2–L^2.

Dém. Montrons d'abord l'unicité. Si φ_1 et φ_2 sont deux éléments de $L^2(\mu)$ tels que $\Phi(f) = \langle f, \varphi_1 \rangle = \langle f, \varphi_2 \rangle$, alors $\langle f, \varphi_1 - \varphi_2 \rangle = 0$. Si l'on choisit $f = \varphi_1 - \varphi_2 \in L^2(\mu)$, on obtient $\| \varphi_1 - \varphi_2 \|_2 = 0$, autrement dit $\varphi_1 = \varphi_2$.

Montrons à présent l'existence. Soit $F := \ker \Phi$. Comme Φ est continue, $F = \Phi^{-1}(\{0\})$ est un fermé, c'est donc un sous-espace vectoriel fermé de $L^2(\mu)$. On sait donc d'après le dernier corollaire que $L^2(\mu) = F \oplus F^\perp$.

Premier cas : $F^\perp = \{0\}$. Alors $F = L^2(\mu)$, ce qui signifie que Φ est la forme linéaire nulle et le choix de la fonction nulle pour φ convient.

Deuxième cas : $F^\perp \neq \{0\}$. Nous allons voir que $F^\perp$ est une droite vectorielle et que φ est un élément de cette droite. Soit g un élément non nul de $F^\perp$ et $\varphi = cg$, où

$$c := \frac{\Phi(g)}{\| g \|_2^2} \in \mathbb{R}.$$

Comme g est non nul et que $g \in F^\perp$, $g \notin F$ et par conséquent $\Phi(g) \neq 0$. Alors pour tout $f \in L^2(\mu)$, avec $\lambda = \Phi(f)/\Phi(g) \in \mathbb{R}$, $f - \lambda g \in F$, car $\Phi(f - \lambda g) = \Phi(f) - \lambda \Phi(g) = 0$. Par conséquent, puisque $g \in F^\perp$, $\langle f - \lambda g, g \rangle = 0$, ce qui s'écrit

$$\lambda = \frac{\langle f, g \rangle}{\| g \|_2^2},$$

donc

$$\Phi(f) = \lambda \Phi(g) = \frac{\langle f, g \rangle}{\| g \|_2^2} \Phi(g) = c \langle f, g \rangle = \langle f, \varphi \rangle,$$

ce qui achève la démonstration. □

14.3 Théorème de Radon–Nikodym

Définition 14.20 *Soient μ et ν deux mesures sur un espace mesurable $(E, \mathscr{A})$. On dit que ν est* absolument continue *par rapport à μ, et on note $\nu \ll \mu$, si pour tout $A \in \mathscr{A}$,*

$$\mu(A) = 0 \implies \nu(A) = 0.$$

Exemple 14.21 (très important) *Si $f : (E, \mathscr{A}) \to (\bar{\mathbb{R}}_+, \mathscr{B}(\bar{\mathbb{R}}_+))$ est mesurable, alors la mesure ν de densité f par rapport à μ, définie par*

$$\nu(A) = \int_E \mathbb{1}_A \, f \, d\mu \qquad A \in \mathscr{A},$$

est absolument continue par rapport à μ. On note souvent $f = \frac{d\nu}{d\mu}$ et on l'appelle dérivée de Radon–Nikodym, *en référence au théorème qui établit la réciproque de cet exemple sous la condition que μ et ν sont toutes deux σ-finies. Voici en effet un contre-exemple avec la mesure de Lebesgue λ sur $[0, 1]$ et m la mesure de comptage. On voit que $\lambda \ll m$ car si $m(A) = 0$ alors $A = \varnothing$ et par conséquent $\lambda(A) = 0$. Pourtant $\frac{d\lambda}{dm}$ n'existe pas,* *comme nous allons le montrer par l'absurde. Supposons que $f = \frac{d\lambda}{dm}$ existe et notons $D = \{f \neq 0\}$. Alors $\int_{[0,1]} f\, dm = \lambda([0,1]) = 1$, et pour tout $\varepsilon > 0$,*

$$1 \geq \int_{[0,1]} f \mathbb{1}_{f \geq \varepsilon}\, dm \geq \varepsilon m(\{f \geq \varepsilon\}),$$

ce qui implique que $\{f \geq \varepsilon\}$ est fini. En particulier $D = \cup_{n \geq 1}\{f \geq 1/n\}$ est dénombrable et donc $\lambda(D) = 0$. On a donc

$$1 = \lambda(\{f = 0\}) = \int_{[0,1]} \mathbb{1}_{\{f=0\}} f\, dm = 0,$$

ce qui est la contradiction annoncée.

Théorème 14.22 (de Radon–Nikodym) *Soient μ et ν deux mesures σ-finies sur $(E, \mathscr{A})$ telles que $\nu \ll \mu$. Alors il existe une fonction mesurable $f : (E, \mathscr{A}) \to (\bar{\mathbb{R}}_+, \mathscr{B}(\bar{\mathbb{R}}_+))$, unique à un ensemble μ-négligeable près, telle que pour tout $A \in \mathscr{A}$,*

$$\nu(A) = \int_E \mathbb{1}_A f\, d\mu.$$

Dém. La démonstration se fait en 4 étapes : 1) μ et ν finies, $\nu \leq \mu$; 2) μ et ν finies ; 3) μ et ν σ-finies ; 4) unicité.

Étape 1. On suppose ici que μ et ν sont finies et que $\nu \leq \mu$, autrement dit pour tout $A \in \mathscr{A}$, $\nu(A) \leq \mu(A)$. Remarquons que ceci implique l'absolue continuité $\nu \ll \mu$. Il est d'ailleurs équivalent de supposer (par le lemme fondamental d'approximation et le théorème de convergence monotone) que pour toute fonction mesurable positive g, $\int g\, d\nu \leq \int g\, d\mu$. Soit alors la fonction Φ définie par

$$\begin{aligned} \Phi : L^2(\mu) &\longrightarrow \mathbb{R} \\ g &\longmapsto \int_E g\, d\nu \end{aligned}$$

qui est bien définie car $L^2(\mu) \subseteq L^1(\nu)$. En effet, comme $\nu \leq \mu$, on a $L^2(\mu) \subseteq L^2(\nu)$, et $L^2(\nu) \subseteq L^1(\nu)$ puisque ν est finie. La fonction Φ est donc une forme linéaire sur $L^2(\mu)$. De plus, grâce à l'inégalité de Hölder,

$$|\,\Phi(g)\,| \leq \int_E |\,g\,|\, d\nu = \int_E |\,g\,|.1\, d\nu \leq \|\,g\,\|_{L^2(\nu)} \|\,1\,\|_{L^2(\nu)} \leq \|\,g\,\|_{L^2(\mu)} \sqrt{\nu(E)},$$

autrement dit Φ est $\sqrt{\nu(E)}$-lipschitzienne, et est donc continue. Nous pouvons donc appliquer le lemme de Riesz-Fisher à la forme linéaire continue Φ sur $L^2(\mu)$, et exhiber un élément φ de $L^2(\mu)$ tel que pour tout $g \in L^2(\mu)$, $\Phi(g) = \int_E g\varphi\, d\mu$. Notons aussi que $\varphi \in L^1(\mu)$, car $L^2(\mu) \subseteq L^1(\mu)$. De plus, puisque μ est finie, toute fonction indicatrice $g = \mathbb{1}_A$ est dans $L^2(\mu)$ et ainsi

$$\nu(A) = \int_E \mathbb{1}_A\, d\nu = \int_E \mathbb{1}_A \varphi\, d\mu.$$

Montrons que $\varphi \in [0,1]$ μ-p.p. Soit $\varepsilon > 0$.

$$0 \leq \nu(\{\varphi \leq -\varepsilon\}) = \int_{\{\varphi \leq -\varepsilon\}} \varphi \, d\mu \leq -\varepsilon \mu(\{\varphi \leq -\varepsilon\}) \leq 0,$$

ce qui implique que $\mu(\{\varphi \leq -\varepsilon\}) = \nu(\{\varphi \leq -\varepsilon\}) = 0$. De même, comme $\nu \leq \mu$,

$$\nu(\{\varphi \geq 1+\varepsilon\}) = \int_{\{\varphi \geq 1+\varepsilon\}} \varphi \, d\mu \geq (1+\varepsilon)\mu(\{\varphi \geq 1+\varepsilon\}) \geq (1+\varepsilon)\nu(\{\varphi \geq 1+\varepsilon\}),$$

ce qui implique que $\nu(\{\varphi \geq 1+\varepsilon\}) = \mu(\{\varphi \geq 1+\varepsilon\}) = 0$. Ainsi pour tout $\varepsilon > 0$, $\mu(\{\varphi \notin [-\varepsilon, 1+\varepsilon]\}) = 0$. Comme $\{\varphi \notin [0,1]\} = \cup_{n \geq 1}\{\varphi \notin [-1/n, 1+1/n]\}$, $\mu(\{\varphi \notin [0,1]\}) = 0$. Conclusion : il existe $\varphi \in L^1(\mu)$ prenant ses valeurs dans $[0,1]$ μ-p.p. tel que pour tout $A \in \mathscr{A}$, $\nu(A) = \int_A \varphi \, d\mu$.

Étape 2. On traite ici le cas de deux mesures μ et ν finies telles que $\nu \ll \mu$. On applique alors l'étape 1 au couple $(\nu, \nu+\mu)$, puisque $\mu+\nu$ est une mesure finie qui vérifie $\nu \leq \mu + \nu$. Il existe donc $\varphi \in L^1(\mu+\nu)$ prenant ses valeurs dans $[0,1]$ $(\mu+\nu)$-p.p., donc μ-p.p. et ν-p.p., tel que pour tout $A \in \mathscr{A}$, $\nu(A) = \int_A \varphi \, d\mu + \int_A \varphi \, d\nu$. Ceci implique que pour tout $A \in \mathscr{A}$, $\int_A (1-\varphi) \, d\nu = \int_A \varphi \, d\mu$, puis par approximation et convergence monotone, que pour toute fonction g mesurable positive,

$$\int_E (1-\varphi) g \, d\nu = \int_E \varphi g \, d\mu. \tag{14.1}$$

Si $N := \{\varphi = 1\}$, alors avec $g = \mathbb{1}_N$,

$$\int_E (1-\varphi)\mathbb{1}_N \, d\nu = 0 = \int_E \varphi \mathbb{1}_N \, d\mu = \mu(N).$$

Comme de plus $\nu \ll \mu$, on a $\nu(N) = 0$. Donc pour tout $A \in \mathscr{A}$, comme $1-\varphi \neq 0$ sur ${}^c\!N$, on peut définir $g = \mathbb{1}_{A \cap {}^cN}/(1-\varphi)$, et avec ce choix de g, on obtient alors

$$\int_E (1-\varphi) \frac{\mathbb{1}_{A\cap {}^cN}}{1-\varphi} \, d\nu = \nu(A \cap {}^cN) = \nu(A) = \int_E \varphi \frac{\mathbb{1}_{A\cap {}^cN}}{1-\varphi} \, d\mu = \int_A \varphi \frac{\mathbb{1}_{{}^cN}}{1-\varphi} \, d\mu.$$

On définit alors la fonction μ-p.p. positive f par

$$f := \frac{\varphi}{1-\varphi} \mathbb{1}_{{}^cN},$$

ce qui assure que pour tout $A \in \mathscr{A}$, $\nu(A) = \int_A f \, d\mu$, où il suffit de prendre $A = E$ pour voir que $f \in L^1(\mu)$.

Étape 3. Extension au cas σ-fini. Il existe donc deux partitions de E dénombrables et $\mathscr{A}$-mesurables (F_n) et (G_n) telles que $\mu(F_n) < \infty$ et $\nu(G_n) < \infty$. De ces deux partitions on peut tirer une partition de E dénombrable et $\mathscr{A}$-mesurable (E_n) telle que $\mu(E_n) < \infty$ et $\nu(E_n) < \infty$, par exemple en considérant toutes les intersections $F_j \cap G_k$ et en utilisant une bijection allant de $\mathbb{N}^2$ dans $\mathbb{N}$. Pour n fixé, soient μ_n et ν_n les mesures traces de μ et ν respectivement, sur E_n. Alors μ_n et ν_n sont des mesures finies et puisque $\nu \ll \mu$, $\nu_n \ll \mu_n$, donc il existe une fonction positive $f_n \in L^1(\mu_n)$ telle que pour tout $A \in \mathscr{A}$,

$$\nu_n(A) = \int_A f_n \, d\mu_n = \int_A f_n \mathbb{1}_{E_n} \, d\mu.$$

Soit maintenant $f := \sum_n f_n \mathbb{1}_{E_n}$. Alors pour tout $A \in \mathscr{A}$,

$$\int_A f \, d\mu = \int_A \sum_n f_n \mathbb{1}_{E_n} \, d\mu = \sum_n \int_A f_n \mathbb{1}_{E_n} \, d\mu = \sum_n \nu_n(A) = \sum_n \nu(A \cap E_n) = \nu(A).$$

Étape 4. Montrons l'unicité de la dérivée de Radon–Nikodym. Supposons qu'il existe deux fonctions mesurables positives f et g telles que pour tout $A \in \mathscr{A}$,

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

Ainsi avec $A_n = \{f > g + \varepsilon\} \cap E_n$, où ε un nombre réel positif quelconque,

$$\nu(A_n) = \int_{A_n} f \, d\mu = \int_{A_n} g \, d\mu.$$

Comme $\nu(A_n) < \infty$, $\int_{A_n} g \, d\mu < \infty$, et l'on peut donc retrancher ce terme à la dernière égalité, de manière à obtenir

$$0 = \int_{\{f > g+\varepsilon\} \cap E_n} (f - g) \, d\mu \geq \varepsilon \mu(\{f > g + \varepsilon\} \cap E_n).$$

Par conséquent, $\mu(\{f > g + \varepsilon\} \cap E_n) = 0$, et en passant à la limite $n \to \infty$, on obtient $\mu(\{f > g + \varepsilon\}) = 0$, puis en passant à la limite $\varepsilon \to 0$ (le long d'une suite dénombrable), on obtient $\mu(\{f > g\}) = 0$. Par symétrie, on obtient $\mu(\{f < g\}) = 0$, et en conclusion $\mu(\{f \neq g\}) = 0$. □

14.4 Dualité L^p–L^q

Le lemme de Riesz–Fisher montre que le dual topologique de $L^2(\mu)$ est isomorphe à $L^2(\mu)$, ce que l'on a appelé la dualité L^2–L^2. Nous allons énoncer ici une extension de cette propriété, qui est la dualité L^p–L^q, où p et q sont deux éléments conjugués de $[1, +\infty]$, c'est-à-dire tels que $\frac{1}{p} + \frac{1}{q} = 1$. Comme dans l'exemple 14.16, nous pouvons démontrer que pour tout $g \in L^q(\mu)$, l'application

$$\begin{aligned} \Phi_g : L^p(\mu) &\longrightarrow \mathbb{R} \\ f &\longmapsto \int_E fg \, d\mu \end{aligned}$$

est une forme linéaire continue, et que si $g \geq 0$ μ-p.p., alors Φ_g est *positive*, au sens où si $f \geq 0$ μ-p.p., alors $\Phi_g(f) \geq 0$.

Théorème 14.23 (dualité L^p–L^q) *Soit μ une mesure σ-finie sur un espace mesurable $(E, \mathscr{A})$ et $p \in [1, +\infty[$. Soit Φ une forme linéaire continue sur $L^p(\mu)$. Alors $\exists! \varphi \in L^q(\mu)$ tel que*

$$\text{pour tout } f \in L^p(\mu), \qquad \Phi(f) = \int_E f\varphi \, d\mu.$$

Remarque 14.24 *Le précédent théorème reste vrai lorsque μ n'est pas σ-finie pourvu que $1 < p < \infty$.*

Dém. Nous allons seulement donner une esquisse de la démonstration, et uniquement dans le cas où l'on suppose que Φ est aussi positive (et alors $\varphi \geq 0$ μ-p.p.).

1) Soit une partition dénombrable $\mathscr{A}$-mesurable (E_n) de E telle que $\mu(E_n) < \infty$ pour tout n. On définit alors l'application $\nu_n : \mathscr{A} \to \mathbb{R}_+$ par

$$\nu_n(A) := \Phi\left(\mathbb{1}_{A\cap E_n}\right) \qquad A \in \mathscr{A}.$$

Alors ν_n est une mesure (finie). D'abord, $\nu_n(\varnothing) = \Phi(0) = 0$. Ensuite pour toute suite (A_k) d'éléments de $\mathscr{A}$ deux à deux disjoints, si l'on note $f_k := \mathbb{1}_{\cup_{j=0}^k A_j \cap E_n} = \sum_{j=0}^k \mathbb{1}_{A_j\cap E_n}$, alors (f_k) converge simplement vers $f := \mathbb{1}_{A\cap E_n}$, où $A = \cup_k A_k$. De plus $f_k \leq \mathbb{1}_{E_n} \in L^p$ car $\| \mathbb{1}_{E_n} \|_p = \mu(E_n)^{1/p} < \infty$, donc par convergence L^p-dominée, la suite (f_k) tend vers f dans L^p. Par conséquent, comme Φ est continue,

$$\nu_n\left(\cup_k A_k\right) = \Phi\left(\lim_{k\to\infty} f_k\right) = \lim_{k\to\infty} \Phi(f_k) = \lim_{k\to\infty} \sum_{j=0}^{k} \Phi(\mathbb{1}_{A_k\cap E_n}) = \sum_k \nu_n(A_k).$$

2) Appliquer le théorème de Radon–Nikodym à ν_n et μ. Comme Φ est linéaire et continue, elle est $\| \Phi \|$-lipschitzienne, et donc pour tout $A \in \mathscr{A}$ tel que $\mu(A) = 0$,

$$\nu_n(A) = \Phi(\mathbb{1}_{A\cap E_n}) \leq \| \Phi \| \,.\, \| \mathbb{1}_{A\cap E_n} \|_p = \| \Phi \| \, \mu(A\cap E_n)^{1/p} = 0.$$

Ainsi $\nu_n \ll \mu$ et il existe donc une application mesurable positive $\varphi_n \in L^1(\mu)$ (car ν_n est finie) telle que pour tout $A \in \mathscr{A}$, $\nu_n(A) = \int_A \varphi_n \, d\mu$. De plus, quitte à remplacer φ_n par $\varphi_n \mathbb{1}_{E_n}$, on peut supposer que φ_n est nulle sur ${}^c E_n$.

3) Soit $\varphi := \sum_n \varphi_n$. Il s'agit ici de montrer que pour tout $f \in L^p(\mu)$,

$$\Phi(f) = \int_E f\varphi \, d\mu.$$

Pour $f \geq 0$ élément de $L^p(\mu)$, la suite $(\sum_{j=0}^n f\mathbb{1}_{E_j})$ converge simplement vers f et est dominée par $f \in L^p$, donc par convergence L^p-dominée, converge vers f dans L^p, ainsi par continuité de Φ,

$$\Phi(f) = \Phi\left(\lim_n \sum_{j=0}^n f\mathbb{1}_{E_j}\right) = \lim_n \Phi\left(\sum_{j=0}^n f\mathbb{1}_{E_j}\right) = \lim_n \sum_{j=0}^n \Phi\left(f\mathbb{1}_{E_j}\right) = \lim_n \sum_{j=0}^n \int_E f \, d\nu_j.$$

Mais par le lemme fondamental d'approximation et le théorème de convergence monotone, il est facile de voir que $\int_E f \, d\nu_j = \int_E f\varphi_j \, d\mu$, de sorte que par convergence monotone à nouveau, $\Phi(f) = \sum_n \int_E f\varphi_n \, d\mu = \int_E f\varphi \, d\mu$. L'extension aux fonctions de signe quelconque est classique.

4) Montrer que $\varphi \in L^q(\mu)$, en distinguant suivant que $p > 1$ ou $p = 1$. Dans la suite on suppose que $\varphi \geq 0$ partout, quitte à remplacer φ par $\varphi \mathbb{1}_{\varphi>0}$ (car $\varphi \geq 0$ μ-p.p.). Supposons d'abord que $p > 1$ (et ainsi $q < \infty$). Soit

$$f_{m,n} := \varphi^{q-1} \mathbb{1}_{F_n \cap \{\varphi \leq m\}},$$

où $F_n = \cup_{k=0}^n E_k$ est une suite croissante d'éléments de $\mathscr{A}$ convergeant vers E. Montrons que $f_{m,n} \in \mathcal{L}^p(\mu)$. En effet,

$$\| f_{m,n} \|_p^p = \int_{F_n} \varphi^{p(q-1)} \mathbb{1}_{\{\varphi \leq m\}} \, d\mu = \int_{F_n} \varphi^q \mathbb{1}_{\{\varphi \leq m\}} \, d\mu,$$

car $p(q-1) = q$ et donc $\| f_{m,n} \|_p^p \leq m^q \mu(F_n) < \infty$. De plus,

$$\Phi(f_{m,n}) = \int_{F_n} \varphi^q \mathbb{1}_{\{\varphi \leq m\}} \, d\mu \leq \| \Phi \| \, . \, \| f_{m,n} \|_p = \| \Phi \| \left(\int_{F_n} \varphi^q \mathbb{1}_{\{\varphi \leq m\}} \, d\mu \right)^{1/p} .$$

Ceci implique que

$$\left(\int_{F_n} \varphi^q \mathbb{1}_{\{\varphi \leq m\}} \, d\mu \right)^{1/q} \leq \| \Phi \|,$$

et ainsi par une double application du théorème de convergence monotone, que $\| \varphi \|_q \leq \| \Phi \| < \infty$, ce qui assure que $\varphi \in L^q(\mu)$.

Supposons maintenant que $p = 1$ (et donc $q = \infty$). Alors pour tout $A \in \mathscr{A}$ tel que $\mu(A) < \infty$,

$$\int_A \varphi \, d\mu = \Phi(\mathbb{1}_A) \leq \| \Phi \| \, \mu(A),$$

donc $\int_A (\varphi - \| \Phi \|) \, d\mu \leq 0$ et ce pour A quelconque de mesure finie, donc $\varphi \leq \| \Phi \|$ @ μ-p.p., ce qui assure que $\varphi \in L^\infty(\mu)$.

5) Unicité de φ (à un ensemble négligeable près). Soient φ_1 et φ_2 telles que pour tout $A \in \mathscr{A}$,

$$\Phi(\mathbb{1}_{A \cap E_n}) = \int_A \varphi_1 \mathbb{1}_{E_n} \, d\mu = \int_A \varphi_2 \mathbb{1}_{E_n} \, d\mu.$$

Comme ces trois termes sont finis, avec $A = \{\varphi_1 > \varphi_2\}$

$$\int_A (\varphi_1 - \varphi_2) \mathbb{1}_{E_n} \, d\mu = 0,$$

ce qui implique que $(\varphi_1 - \varphi_2) \mathbb{1}_{A \cap E_n} = 0$ μ-p.p. et ainsi $\mu(A) = 0$. On conclut par un argument de symétrie. $\square$

Chapitre 15

Régularité d'une mesure et théorèmes de densité

15.1 Régularité d'une mesure sur un espace métrique

Soit μ une mesure sur un espace métrique E muni de sa tribu borélienne $\mathscr{B}(E)$.

Définition 15.1 *On dit que μ est*

a) extérieurement régulière *si pour tout $A \in \mathscr{B}(E)$,*

$$\mu(A) = \inf\{\mu(O), O \textit{ ouvert}, O \supseteq A\};$$

b) intérieurement régulière *si pour tout $A \in \mathscr{B}(E)$,*

$$\mu(A) = \sup\{\mu(K), K \textit{ compact}, K \subseteq A\};$$

c) régulière*, si elle est à la fois extérieurement et intérieurement régulière.*

Proposition 15.2 *Si μ est finie, alors elle est extérieurement régulière et pour tout $A \in \mathscr{B}(E)$,*

$$\mu(A) = \sup\{\mu(F), F \textit{ fermé}, F \subseteq A\}.$$

Dém. Il suffit de montrer que pour tout $A \in \mathscr{B}(E)$,

$$\forall \varepsilon > 0, \text{ il existe un ouvert } O \supseteq A \text{ et un fermé } F \subseteq A \text{ tels que } \mu(O \setminus F) < \varepsilon. \quad (15.1)$$

Soit $\mathscr{T} := \{A \in \mathscr{B}(E)$ vérifiant (15.1)$\}$. Montrons que $\mathscr{T}$ est une tribu contenant $\mathscr{O}(E)$, et par conséquent égale à $\mathscr{B}(E)$.

Montrons que $\mathscr{T}$ contient bien les ouverts. Soit A un ouvert. Pour O, il suffit de prendre $O = A$. Pour F, définissons

$$F_n := \{x \in A : d(x, {}^cA) \geq \frac{1}{n}\}.$$

Comme la fonction $x \mapsto d(x, {}^cA)$ est continue, F_n est fermé. De plus ${}^cA$ est fermé donc pour tout $x \in A$, $d(x, {}^cA) \neq 0$. En effet, $d(x, {}^cA) = 0$ ssi x est dans l'adhérence de ${}^cA$, qui n'est autre que le complémentaire de l'intérieur de A, c'est-à-dire le complémentaire de A. Ainsi, $A = \lim_n \uparrow F_n$ et donc $\mu(A \setminus F_n) \downarrow \mu(\varnothing) = 0$ car μ est finie donc continue à droite.

Montrons que $\mathscr{T}$ est une tribu.

i) $E \in \mathscr{T}$, car il suffit alors de prendre $O = F = E$.

ii) Passage au complémentaire. Soit $A \in \mathscr{T}$ et $\varepsilon > 0$. Alors il existe $F \subseteq A \subseteq O$ tels que $\mu(O \setminus F) < \varepsilon$, où O est ouvert et F fermé. Avec $O' = {}^cF$ et $F' = {}^cO$, O' est ouvert, F' est fermé, et $F' \subseteq A \subseteq O'$. De plus, $O' \setminus F' = O' \cap {}^cF' = {}^cF \cap O = O \setminus F$, donc $\mu(O \setminus F) < \varepsilon$.

iii) Réunion dénombrable. Soit (A_n) une suite d'éléments de $\mathscr{T}$ et $\varepsilon > 0$. Alors pour chaque entier n, il existe un ouvert O_n et un fermé F_n tels que $F_n \subseteq A_n \subseteq O_n$ et

$$\mu(O_n \setminus F_n) \leq \frac{\varepsilon}{2^{n+1}}.$$

Comme $(\cup_{k\leq n} F_k)_n$ est une suite croissante de limite $\cup_n F_n$, $\mu(\cup_{k\leq n} F_k)$ converge vers $\mu(\cup_n F_n)$. Comme $\mu(\cup_n F_n) < \infty$, il existe un entier n_ε tel que

$$\mu(\cup_{k\leq n_\varepsilon} F_k) \geq \mu(\cup_n F_n) - \frac{\varepsilon}{2}.$$

Soient maintenant l'ouvert $O' = \cup_n O_n$ (réunion d'ouverts) et le fermé $F' = \cup_{k\leq n_\varepsilon} F_k$ (réunion finie de fermés). On a alors

$$F' = \cup_{k\leq n_\varepsilon} F_k \subseteq \cup_n F_n \subseteq \cup_n A_n \subseteq \cup_n O_n = O',$$

et de plus

$$\begin{aligned}
\mu(O' \setminus F') &= \mu\left(\cup_n O_n\right) - \mu\left(\cup_{k\leq n_\varepsilon} F_k\right) \\
&= \mu\left(\cup_n O_n\right) - \mu\left(\cup_n F_n\right) + \mu\left(\cup_n F_n\right) - \mu\left(\cup_{k\leq n_\varepsilon} F_k\right) \\
&= \mu\left(\cup_n O_n \setminus \cup_k F_k\right) + \mu\left(\cup_n F_n\right) - \mu\left(\cup_{k\leq n_\varepsilon} F_k\right) \\
&\leq \mu\left((\cup_n O_n) \cap (\cap_k {}^cF_k)\right) + \frac{\varepsilon}{2} \\
&= \mu\left(\cup_n \cap_k \left(O_n \cap {}^cF_k\right)\right) + \frac{\varepsilon}{2} \\
&\leq \mu\left(\cup_n (O_n \cap {}^cF_n\right) + \frac{\varepsilon}{2} \\
&\leq \sum_n \mu\left(O_n \setminus F_n\right) + \frac{\varepsilon}{2} \leq \sum_n \frac{\varepsilon}{2^{n+1}} + \frac{\varepsilon}{2} \leq \varepsilon,
\end{aligned}$$

ce qui achève la démonstration. □

Théorème 15.3 *Toute mesure de Borel sur un espace localement compact séparable est régulière.*

Dém. On utilise le fait qu'un tel espace est σ-compact, c'est-à-dire qu'il existe une suite croissante (E_n) de compacts dont la limite $\cup_n E_n$ est égale à E. En fait, nous allons utiliser le résultat plus fort que $E = \cup_n \mathring{E}_n$, où l'on rappelle que $\mathring{E}_n$ désigne l'intérieur de E_n.

Régularité intérieure. Soit $\varepsilon > 0$ et $A \in \mathscr{B}(E)$. Soit μ_n la mesure trace de μ sur E_n. Comme μ_n est finie, d'après la proposition précédente, il existe un fermé $F_n \subseteq A$ tel que

$$\mu(A \cap E_n) \leq \mu(F_n \cap E_n) + \frac{\varepsilon}{2}.$$

Soit le compact $K_n = F_n \cap E_n$ (intersection d'un fermé et d'un compact). Comme $F_n \subseteq A$, on a $K_n \subseteq A$ et l'on réécrit l'équation précédente

$$\mu(A \cap E_n) \leq \mu(K_n) + \frac{\varepsilon}{2}. \tag{15.2}$$

Observons que $\lim_n \mu(A \cap E_n) = \mu(A)$. Si $\mu(A) = \infty$, alors d'après (15.2) $\lim_n \mu(K_n) = +\infty$, autrement dit $\lim_n \mu(K_n) = \mu(A)$. Si $\mu(A) < \infty$, alors il existe un entier n tel que

$$\mu(A) \leq \mu(A \cap E_n) + \frac{\varepsilon}{2} \leq \mu(K_n) + \varepsilon,$$

ce qui nous permet de conclure que μ est régulière intérieurement.

Régularité extérieure. Soit $\varepsilon > 0$ et $A \in \mathscr{B}(E)$. Soit μ_n la mesure trace de μ sur $\mathring{E}_n$. Comme μ_n est finie, d'après la proposition précédente, elle est régulière extérieurement. Ainsi il existe un ouvert O_n tel que $O_n \supseteq A$ et

$$\mu(A \cap \mathring{E}_n) \geq \mu(O_n \cap \mathring{E}_n) - \frac{\varepsilon}{2^n}. \tag{15.3}$$

Montrons que $O := \cup_n (O_n \cap \mathring{E}_n)$ vérifie $\mu(A) \geq \mu(O) - \varepsilon$. Comme $A \cap \mathring{E}_n \subseteq O_n \cap \mathring{E}_n$, on a $A = \cup_n (A \cap \mathring{E}_n) \subseteq O$, ce qui donnera le résultat car O est ouvert (c'est une réunion d'ouverts). Soit $U_n := \cup_{k \leq n} (O_k \cap \mathring{E}_k)$. Nous allons montrer par récurrence sur $n \geq 1$ que

$$\mu(U_n) \leq \mu(A \cap \mathring{E}_n) + \sum_{1 \leq k \leq n} \frac{\varepsilon}{2^k}.$$

L'égalité est vraie pour $n = 1$ car elle se réduit à (15.3). En se servant de (15.3),

$$\begin{aligned}
\mu(U_{n+1}) &= \mu(U_n) + \mu(O_{n+1} \cap \mathring{E}_{n+1}) - \mu(U_n \cap O_{n+1} \cap \mathring{E}_{n+1}) \\
&\leq \mu(A \cap \mathring{E}_n) + \sum_{k \leq n} \frac{\varepsilon}{2^k} + \mu(A \cap \mathring{E}_{n+1}) + \frac{\varepsilon}{2^{n+1}} - \mu(U_n \cap O_{n+1} \cap \mathring{E}_{n+1}) \\
&= \mu(A \cap \mathring{E}_{n+1}) + \sum_{k \leq n+1} \frac{\varepsilon}{2^k} + a_n,
\end{aligned}$$

où

$$a_n := \mu(A \cap \mathring{E}_n) - \mu(U_n \cap O_{n+1} \cap \mathring{E}_{n+1}) \leq 0,$$

car $A \cap \mathring{E}_n \subseteq U_n \cap O_{n+1} \cap \mathring{E}_{n+1}$. En effet, d'une part $A \subseteq O_{n+1}$ et $\mathring{E}_n \subseteq \mathring{E}_{n+1}$ donc $A \cap \mathring{E}_n \subseteq O_{n+1} \cap \mathring{E}_{n+1}$; d'autre part, $A \subseteq O_n$, donc $A \cap \mathring{E}_n \subseteq O_n \cap \mathring{E}_n \subseteq U_n$. On a donc bien

$$\mu\left(U_{n+1}\right) \leq \mu(A \cap \mathring{E}_{n+1}) + \sum_{k \leq n+1} \frac{\varepsilon}{2^k}.$$

Ainsi, comme (U_n) est une suite croissante de limite O, on obtient l'inégalité souhaitée en passant à la limite en n. □

15.2 Théorèmes de densité

Proposition 15.4 *Pour tout $p \in [1, +\infty[$, l'ensemble des (représentants des) fonctions étagées intégrables est dense dans $L^p(\mu)$.*

Dém. Soit $f \in L^p$. Quitte à raisonner sur f^+ et f^-, nous pouvons supposer que f est positive. Par le lemme fondamental d'approximation, il existe une suite croissante (φ_n) de fonctions étagées positives convergeant simplement vers f. Il nous suffit alors de montrer que $\varphi_n \in L^1(\mu) \cap L^p(\mu)$ pour tout n et que la convergence a aussi lieu dans $L^p(\mu)$. Comme $f \in L^p(\mu)$ et que $0 \leq \varphi_n \leq f$, on a bien $\varphi_n \in L^p(\mu)$. Mais une fonction étagée positive φ de $L^p(\mu)$ est aussi élément de $L^1(\mu)$, en effet, pour tout $y \in \varphi(E)$, si $y \neq 0$ alors $\mu(\varphi = y) < \infty$ car $\int \varphi^p \, d\mu = \sum_{y \in \varphi(E), y \neq 0} y^p \mu(\varphi = y) < \infty$. Mais comme $\varphi(E)$ est fini, $\int \varphi \, d\mu = \sum_{y \in \varphi(E)} y \mu(\varphi = y) < \infty$. Enfin, $f - \varphi_n$ est dominée par $f \in L^p(\mu)$, donc par convergence L^p-dominée, la suite (φ_n) converge vers f dans $L^p(\mu)$. □

Théorème 15.5 *Soit μ une mesure de Borel sur $\mathbb{R}^d$. Alors pour tout $p \in [1, +\infty[$, l'espace vectoriel $C_K^\infty(\mathbb{R}^d, \mathbb{R})$ des fonctions réelles à support compact indéfiniment différentiables est dense dans $L^p(\mu)$.*

Dém. D'après la proposition précédente, il suffit de montrer que l'espace C_K^∞ est dense dans l'espace des fonctions étagées intégrables [1], c'est-à-dire les fonctions de la forme $\sum_{i=1}^n \alpha_i \mathbb{1}_{A_i}$ où pour tout i, $\alpha_i \mu(A_i) < \infty$. Par linéarité, il suffit de montrer que pour tout $A \in \mathscr{B}(\mathbb{R}^d)$ de mesure finie, la fonction indicatrice de A est limite dans $L^p(\mu)$ d'une suite d'éléments de C_K^∞.

Soit $\varepsilon > 0$. Comme μ est une mesure de Borel, elle est extérieurement régulière, donc A étant de mesure finie, il existe un ouvert $O \supseteq A$ tel que $\mu(O \setminus A) < (\varepsilon/3)^p$, autrement dit $\| \mathbb{1}_O - \mathbb{1}_A \|_p < \varepsilon/3$. Comme O est la réunion (dénombrable) des rectangles ouverts et bornés à extrémités rationnelles qu'il contient, et que $\mu(O) < \infty$, par continuité à gauche de la mesure, il existe une famille *finie* (et disjointe) de rectangles bornés et ouverts contenus dans O dont la réunion Ω est telle que $\mu(O \setminus \Omega) < (\varepsilon/3)^p$, autrement dit $\| \mathbb{1}_O - \mathbb{1}_\Omega \|_p < \varepsilon/3$. Observons ici que pour tout intervalle ouvert $]a, b[$, il est aisé de @ construire une suite croissante $(\psi_n^{(a,b)})$ de fonctions indéfiniment différentiables dominées par $\mathbb{1}_{]a,b[}$ et convergeant simplement vers $\mathbb{1}_{]a,b[}$. Ainsi pour tout rectangle ouvert $R =$

1. en effet, un résultat classique de topologie assure que si A est dense dans B et B est dense dans C, alors A est dense dans C

$\prod_{i=1}^{d}]a_i, b_i[$, les fonctions $\varphi_n^R(x_1, \ldots, x_d) := \prod_{i=1}^{d} \psi_n^{(a_i,b_i)}(x_i)$ sont dominées par $\mathbb{1}_R$ et convergent simplement vers $\mathbb{1}_R$. À présent, rappelons-nous que Ω est réunion finie et disjointe de rectangles ouverts bornés, soit $\Omega = \cup_j R_j$. En définissant $\phi_n = \sum_j \varphi_n^{R_j}$, on obtient une suite de fonctions indéfiniment différentiables dominées par $\mathbb{1}_\Omega$ et convergeant simplement vers $\mathbb{1}_\Omega$. Comme Ω est borné, ces fonctions sont à support compact, et comme $\mu(\Omega) < \infty$, par convergence L^p-dominée, la convergence a également lieu dans L^p. On peut donc trouver une fonction indéfiniment différentiable à support compact ϕ telle que $\| \mathbb{1}_\Omega - \phi \|_p < \varepsilon/3$. L'inégalité triangulaire nous permet de conclure que $\| \mathbb{1}_A - \phi \|_p < \varepsilon$.
$\square$

Chapitre 16

Produit de convolution

16.1 Convolution de mesures et de fonctions positives

Nous allons définir une nouvelle opération sur les mesures, qui seront toujours supposées ici **définies sur $\mathscr{B}(\mathbb{R}^d)$ et σ-finies**.

Définition 16.1 *Soient μ et ν deux mesures σ-finies sur $\mathscr{B}(\mathbb{R}^d)$. On appelle* produit de convolution, *et l'on note $\mu \star \nu$, l'image de $\mu \otimes \nu$ par l'application $(x, y) \mapsto x + y$.*

Remarque 16.2 *Par le théorème de Fubini–Tonelli, pour tout borélien A,*

$$\begin{aligned} \mu \star \nu(A) &= \int \mathbb{1}_A(x+y)\, d(\mu \otimes \nu)(x,y) \\ &= \int d\mu(x) \int d\nu(y)\, \mathbb{1}_A(x+y) = \int d\nu(y) \int d\mu(x)\, \mathbb{1}_A(x+y) \\ &= \int d\mu(x)\, \nu(A-x) = \int d\nu(y)\, \mu(A-y). \end{aligned}$$

En particulier, $\mu \star \nu(\mathbb{R}^d) = \mu(\mathbb{R}^d)\nu(\mathbb{R}^d)$, et donc si μ et ν sont des probabilités alors $\mu \star \nu$ en est également une.

Proposition 16.3 *le produit de convolution est commutatif, associatif et possède un élément neutre qui est δ_0.*

Dém. La démonstration de la commutativité est immédiate grâce à la commutativité @ du produit $\otimes$ de mesures. Pour l'associativité, voir la démonstration de l'associativité du produit de convolution de fonctions positives ci-dessous. Le fait que δ_0 est un élément neutre est une conséquence immédiate de la remarque précédente. □

Remarque 16.4 *Avant de convoler $\mu \star \nu$ avec une troisième mesure, il faut tout d'abord vérifier que $\mu \star \nu$ est elle-même σ-finie, car ce n'est pas toujours le cas, comme on peut le voir avec le contre-exemple $\lambda \star \lambda$. En effet $\lambda \star \lambda(A) = \int d\lambda(x)\, \lambda(A - x) = \int d\lambda(x)\, \lambda(A) = \lambda(\mathbb{R}^d)\lambda(A)$, qui vaut $+\infty$ dès que A est de mesure de Lebesgue non nulle.*

Proposition 16.5 *Si ν admet une densité g par rapport à la mesure de Lebesgue, alors $\mu \star \nu$ admet également une densité, (encore) notée $\mu \star g$, où*

$$\mu \star g(x) := \int d\mu(y)\, g(x-y) \qquad x \in \mathbb{R}^d.$$

Dém. Par le théorème de Fubini–Tonelli, $\mu \star g$ est borélienne car l'application $(x, y) \mapsto g(x-y)$ est borélienne (voir chapitre sur les tribus produits). De plus, par la formule du changement de variable puis par Fubini–Tonelli,

$$\begin{aligned} \mu \star \nu(A) &= \int d\mu(y) \int d\nu(x) \mathbb{1}_A(x+y) \\ &= \int d\mu(y) \int d\lambda(x)\, g(x)\, \mathbb{1}_A(x+y) \\ &= \int d\mu(y) \int d\lambda(u)\, g(u-y)\, \mathbb{1}_A(u) \\ &= \int d\lambda(u) \mathbb{1}_A(u) \int d\mu(y)\, g(u-y), \end{aligned}$$

qui n'est autre que $\int d\lambda(u)\, \mathbb{1}_A(u)\, \mu \star g(u)$. $\square$

Corollaire 16.6 *Si μ (resp. ν) admet une densité f (resp. g) par rapport à la mesure de Lebesgue, alors $\mu \star \nu$ admet également une densité, (encore !) notée $f \star g$, où*

$$f \star g(x) := \int d\lambda(y)\, f(y)\, g(x-y).$$

Proposition 16.7 *Le produit de convolution $f \star g$ défini pour tout couple (f, g) de fonctions boréliennes positives sur $\mathbb{R}^d$ est commutatif et associatif. De plus, $f \star g$ est borélienne et positive, $\{f \star g \neq 0\} \subseteq \{f \neq 0\} + \{g \neq 0\}$, et*

$$\int_{\mathbb{R}^d} f \star g\, d\lambda = \left(\int f\, d\lambda \right) . \left(\int g\, d\lambda \right).$$

Exemple 16.8 *La fonction $f \star \mathbb{1}_{\mathbb{R}^d}$ est la fonction constante égale à $\int f\, d\lambda$.*

Dém. La commutativité se démontre par un changement de variable affine $u = x - y$ (x étant fixé)

$$\int d\lambda(y)\, f(y)\, g(x-y) = \int d\lambda(u)\, f(x-u)\, g(u) \qquad x \in \mathbb{R}^d.$$

En ce qui concerne l'associativité, si f, g et h sont trois fonctions boréliennes positives, alors

$$\begin{aligned}(f\star g)\star h(x) &= \int dz\, h(z)\, f\star g(x-z)\\ &= \int dz\, h(z)\int dy\, f(y)\, g(x-z-y)\\ &= \int dy\, f(y)\int dz\, h(z)\, g(x-z-y)\\ &= \int dy\, f(y)\, g\star h(x-y)\end{aligned}$$

qui n'est autre que $f\star(g\star h)(x)$.

En désignant par μ la mesure de densité f par rapport à la mesure de Lebesgue, comme $f\star g=\mu\star g$, la proposition précédente donne la mesurabilité de $f\star g$.

Soit $x\in\mathbb{R}^d$ tel que $x\notin\{f\neq 0\}+\{g\neq 0\}$. Alors pour tout $y\in\mathbb{R}^d$, $g(y)=0$ ou $f(x-y)=0$ (sans quoi $y\in\{g\neq 0\}$ et $x-y\in\{f\neq 0\}$, et alors $x=y+(x-y)$ appartient à la somme). Par conséquent $g(y)f(x-y)=0$ pour tout y, si bien que $f\star g(x)=0$.

En désignant par ν la mesure de densité g par rapport à la mesure de Lebesgue, $\int f\star g\,d\lambda=\mu\star\nu(\mathbb{R}^d)=\mu(\mathbb{R}^d)\nu(\mathbb{R}^d)=\left(\int f\,d\lambda\right).\left(\int g\,d\lambda\right)$. □

16.2 Convolution de fonctions boréliennes de signe quelconque

Définition 16.9 *Soient deux fonctions boréliennes $f,g:\mathbb{R}^d\to\mathbb{R}$. Pour tout $x\in\mathbb{R}^d$ tel que $|\,f\,|\star|\,g\,|(x)<\infty$, on définit le nombre réel*

$$f\star g(x):=\int_{\mathbb{R}^d} d\lambda(y)\, f(y)\, g(x-y),$$

que l'on appelle convolée, *ou* produit de convolution *de f et g au point x.*

Proposition 16.10 *Le produit de convolution des fonctions boréliennes jouit des propriétés suivantes :*

a) dès que l'un des deux nombres $f\star g(x)$ ou $g\star f(x)$ est bien défini, ils sont égaux, et

$$|\,f\star g(x)\,|\leq|\,f\,|\star|\,g\,|(x).$$

b) Si pour tout $x\in\mathbb{R}^d$, $|\,f\,|\star|\,g\,|(x)<\infty$, alors la fonction $f\star g$ est borélienne.

c) $\{f\star g\neq 0\}\subseteq\{f\neq 0\}+\{g\neq 0\}$.

d) Si $f_1=f_2$ λ-p.p. et $g_1=g_2$ λ-p.p. alors pour tout $x\in\mathbb{R}^d$, $|\,f_1\,|\star|\,g_1\,|(x)=|\,f_2\,|\star|\,g_2\,|(x)$, et si ce nombre est fini, $f_1\star g_1(x)=f_2\star g_2(x)$.

Remarque 16.11 *Le produit de convolution des fonctions boréliennes de signe quelconque n'est pas associatif en général.*

Dém. a) L'égalité s'obtient par un changement de variable affine et l'inégalité par croissance de l'intégrale.

b) Par croissance de l'intégrale, toutes les fonctions $f^\pm \star g^\pm$ sont bien définies et par linéarité

$$f \star g = f^+ \star g^+ - f^+ \star g^- - f^- \star g^+ + f^- \star g^-,$$

et tous les termes de cette somme sont des fonctions boréliennes (voir section précédente), d'où le résultat.

c) Par l'inégalité $|\, f \star g\,| \leq |\, f\,| \star |\, g\,|$, on a $\{|\, f \star g\,| \neq 0\} \subseteq \{|\, f\,| \star |\, g\,| \neq 0\}$, d'où

$$\{f \star g \neq 0\} = \{|\, f \star g\,| \neq 0\} \subseteq \{|\, f\,| \star |\, g\,| \neq 0\} \subseteq \{|\, f\,| \neq 0\} + \{|\, g\,| \neq 0\} = \{f \neq 0\} + \{g \neq 0\}.$$

d) Pour tout $x \in \mathbb{R}^d$, si l'on définit le borélien $A(x)$ par @

$$A(x) := \{y \in \mathbb{R}^d : f_1(x-y)\, g_1(y) \neq f_2(x-y)\, g_2(y)\},$$

on a alors

$$A(x) \subseteq \{y \in \mathbb{R}^d : f_1(x-y) \neq f_2(x-y)\} \cup \{y \in \mathbb{R}^d : g_1(y) \neq g_2(y)\} = (x - \{f_1 \neq f_2\}) \cup \{g_1 \neq g_2\},$$

si bien que

$$\lambda(A(x)) \leq \lambda(\{f_1 \neq f_2\}) + \lambda(\{g_1 \neq g_2\}) = 0.$$

Par conséquent,

$$\begin{aligned} |\, f_1\,| \star |\, g_1\,|(x) &= \int |\, f_1(x-y)\,|.|\, g_1(y)\,|\, \mathbb{1}_{{}^cA(x)}(y)\, dy \\ &= \int |\, f_2(x-y)\,|.|\, g_2(y)\,|\, \mathbb{1}_{{}^cA(x)}(y)\, dy = |\, f_2\,| \star |\, g_2\,|(x). \end{aligned}$$

Si cette quantité est finie, on fait le même raisonnement avec $f_1 \star g_1$ et $f_2 \star g_2$. □

On souhaite à présent exhiber des conditions suffisantes d'existence partout ou λ-presque partout de $f \star g$.

Proposition 16.12 *1) $f \star g(x)$ existe* pour tout $x \in \mathbb{R}^d$ *si l'une des deux conditions suivantes est réalisée :*

a) f est localement intégrable[1] *et g est essentiellement bornée et à support compact ;*

b) il existe p et q conjugués[2] *tels que $f \in \mathscr{L}^p(\lambda)$ et $g \in \mathscr{L}^q(\lambda)$. Dans ce cas, $|\, f \star g\,|$ est en plus bornée, majorée par $\|\, f\,\|_p\, \|\, g\,\|_q$.*

2) $f \star g(x)$ existe pour λ-presque tout $x \in \mathbb{R}^d$ *dès que $f \in \mathscr{L}^1(\lambda)$ et $g \in \mathscr{L}^p(\lambda)$ (pour $p \in [1, +\infty]$). Dans ce cas, $f \star g \in L^p(\lambda)$ et de plus $\|\, f \star g\,\|_p \leq \|\, f\,\|_1\, \|\, g\,\|_p$.*

1. autrement dit $\int_K |\, f\,|\, d\lambda < \infty$ pour tout compact K de $\mathbb{R}^d$
2. $p, q \in [1, +\infty]$ tels que $p^{-1} + q^{-1} = 1$

Dém. 1) Il nous faut montrer que dans les deux cas $|\, f\,|\star|\, g\,|(x) < \infty$ pour tout $x \in \mathbb{R}^d$.
Cas a :

$$\begin{aligned}
|\, f\,| \star |\, g\,|(x) &= \int |\, f(x-y)\,|.|\, g(y)\,|\, \mathbb{1}_{\{|\, g\,|\leq \|\, g\,\|_\infty\}}\, \mathbb{1}_{\{|\, g\,|\neq 0\}}\, dy \\
&\leq \|\, g\,\|_\infty \int |\, f(x-y)\,|\, \mathbb{1}_{\{|\, g\,|\neq 0\}}\, dy \\
&= \|\, g\,\|_\infty \int_{x-\{|\, g\,|\neq 0\}} |\, f\,|\, d\lambda,
\end{aligned}$$

qui est fini car $x - \{|\, g\,| \neq 0\}$ est borné. Passons au cas b. Dans le cas où $p = 1$ et $q = \infty$, la même démonstration que précédemment montre que

$$\begin{aligned}
|\, f\,| \star |\, g\,|(x) &= \int |\, f(x-y)\,|.|\, g(y)\,|\, \mathbb{1}_{\{|\, g\,|\leq \|\, g\,\|_\infty\}}\, dy \\
&\leq \|\, g\,\|_\infty \int |\, f(x-y)\,|\, dy \\
&= \|\, g\,\|_\infty \int |\, f\,|\, d\lambda = \|\, f\,\|_1\, \|\, g\,\|_\infty,
\end{aligned}$$

ce qui montre que pour tout x, $|\, f\,|\star|\, g\,|(x) \leq \|\, f\,\|_1\, \|\, g\,\|_\infty < \infty$ et donc que $|\, f \star g\,|(x) \leq \|\, f\,\|_1\, \|\, g\,\|_\infty$. Pour $p \in]1, +\infty[$, par l'inégalité de Hölder,

$$|\, f\,| \star |\, g\,|(x) \leq \left(\int |\, f(x-y)\,|^p\, dy\right)^{1/p} \left(\int |\, g(y)\,|^q\, dy\right)^{1/q} = \|\, f\,\|_p\, \|\, g\,\|_q,$$

et l'on aboutit à la même conclusion que précédemment.

2) Soit $f \in \mathscr{L}^1(\lambda)$ et $g \in \mathscr{L}^p(\lambda)$. Le cas où $p = \infty$ est un cas particulier de ce qui précède, aussi nous pouvons supposer que $p < \infty$. En notant μ la mesure de probabilité de densité $|\, f\,|/\|\, f\,\|_1$ (le cas où $\|\, f\,\|_1 = 0$ étant trivial) par rapport à la mesure de Lebesgue et en lui appliquant l'inégalité de Jensen avec l'application convexe $\mathbb{R}_+ : x \mapsto x^p$,

$$\begin{aligned}
\int (|\, f\,| \star |\, g\,|)^p\, (x)\, dx &= \int dx \left(\int |\, f(y)\,|.|\, g(x-y)\,|\, dy\right)^p \\
&= \|\, f\,\|_1^p \left(\int |\, g(x-y)\,|\, d\mu(y)\right)^p \\
&\leq \|\, f\,\|_1^p \int |\, g(x-y)\,|^p\, d\mu(y) \\
&= \|\, f\,\|_1^{p-1} \int dx \int dy\, |\, f(y)\,|.|\, g(x-y)\,|^p \\
&= \|\, f\,\|_1^{p-1} \int dy\, |\, f(y)\,| \int dx\, |\, g(x-y)\,|^p \\
&= \|\, f\,\|_1^p\, \|\, g\,\|_p^p,
\end{aligned}$$

ce qui donne l'inégalité $\|\, |\, f\,|\star|\, g\,|\,\|_p \leq \|\, f\,\|_1\, \|\, g\,\|_p$. Ainsi $|\, f\,|\star|\, g\,|(x) < \infty$ pour λ-presque tout x, donc $f \star g(x)$ est défini pour λ-presque tout x et $\|\, f \star g\,\|_p \leq \|\, |\, f\,| \star |\, g\,|\,\|_p \leq \|\, f\,\|_1 \|\, g\,\|_p$. $\square$

Chapitre 17

Transformée de Fourier

17.1 Définition et premières propriétés

Définition 17.1 *a) Soit μ une mesure* finie *sur $\mathscr{B}or(\mathbb{R}^d)$. On définit la* transformée de Fourier *de μ, et l'on note $\hat{\mu}$, la fonction $\hat{\mu} : \mathbb{R}^d \to \mathbb{C}$ définie par*

$$\hat{\mu}(u) = \int_{\mathbb{R}^d} e^{-i\langle u,x\rangle}\, d\mu(x) \qquad u \in \mathbb{R}^d,$$

où l'on rappelle que $\langle \cdot, \cdot \rangle$ désigne le produit scalaire canonique de $\mathbb{R}^d$.

b) Soit $f : \mathbb{R}^d \to \mathbb{C}$ un élément de $L^1_{\mathbb{C}}(\lambda)$. On définit également la transformée de Fourier *de f, et l'on note (encore) $\hat{f}$, la fonction $\hat{f} : \mathbb{R}^d \to \mathbb{C}$ définie par*

$$\hat{f}(u) = \int_{\mathbb{R}^d} e^{-i\langle u,x\rangle} f(x)\, d\lambda(x) \qquad u \in \mathbb{R}^d.$$

Remarque 17.2 *Si f est positive, alors en désignant par μ la mesure de densité f par rapport λ, on a par définition $\hat{f} = \hat{\mu}$.*

Proposition 17.3 *a) La transformée de Fourier d'une mesure finie ou d'une fonction intégrable est toujours* continue.

b) Les applications $\mu \mapsto \hat{\mu}$ et $f \mapsto \hat{f}$ sont linéaires, et pour tout $u \in \mathbb{R}^d$,

$$|\,\hat{\mu}(u)\,| \leq \mu(\mathbb{R}^d), \qquad\qquad |\,\hat{f}(u)\,| \leq \int |\,f\,|\, d\lambda.$$

c) La transformée de Fourier est un morphisme de groupes pour le produit de convolution, au sens où

$$\widehat{\mu \star \nu} = \hat{\mu}\hat{\nu}, \qquad \widehat{\mu \star f} = \hat{\mu}\hat{f}, \qquad \widehat{f \star g} = \hat{f}\hat{g}.$$

Dém. a) En notant $g(u,x) = e^{-i\langle u,x\rangle} f(x)$, on voit que pour tout x l'application $u \mapsto g(u,x)$ est continue et dominée par $|\,f(x)\,|$ qui par hypothèse est λ-intégrable en x, donc l'application $\hat{f} : u \mapsto \int g(u,x)\, d\lambda(x)$ est continue (et la démonstration est bien sûr la même pour $\hat{\mu}$).

b) évident. @
c) Calculons la transformée de Fourier de $\mu \star \nu$:

$$\widehat{\mu \star \nu} = \int e^{-i\langle u,x\rangle}\, d(\mu \star \nu)(x) = \int\int e^{-i\langle u,x+y\rangle}\, d\mu(x)\, d\nu(y)$$
$$= \int e^{-i\langle u,x\rangle}\, d\mu(x) \int e^{-i\langle u,y\rangle}\, d\nu(y)$$

par le théorème de Fubini–Lebesgue, et cette dernière expression n'est autre que $\hat{\mu}(u)\,\hat{\nu}(u)$ (la démonstration est bien sûr la même pour $\mu \star f$ et $f \star g$). □

Remarque 17.4 *Soit $f; \mathbb{R}^d \to \mathbb{C}$ une fonction λ-intégrable. On pourra montrer en exercice les égalités suivantes.* @

— *Si g est définie par $g(x) = f(-x)$ alors $\hat{g}(u) = \hat{f}(-u)$;*
— *si g est définie par $g(x) = \bar{f}(x)$ (complexe conjugué), alors $\hat{g}(u) = \overline{\hat{f}(-u)}$;*
— *si g est définie par $g(x) = f(x/a)$ (où a est un réel non nul quelconque), alors $\hat{g}(u) = |\, a\,|^d\, \hat{f}(au)$.*

17.2 Injectivité de la transformée de Fourier

Nous allons à présent montrer que la transformée de Fourier d'une mesure finie la caractérise, en nous servant de la fonction $g : \mathbb{R}^d \to \mathbb{R}$ définie par

$$g(x) := (2\pi)^{-d/2} \exp\left(-\frac{\|\, x\,\|^2}{2}\right),$$

où $\|\cdot\|$ désigne la norme euclidienne usuelle de $\mathbb{R}^d$.

Lemme 17.5 *La fonction positive g est une densité*[1] *de probabilité sur $\mathbb{R}^d$ et*

$$\hat{g}(u) = \exp\left(-\frac{\|\, u\,\|^2}{2}\right).$$

Dém. On sait que $\int_{\mathbb{R}} e^{-x^2/2}\, dx = \sqrt{2\pi}$. Par Fubini–Tonelli, en notant x_i la i-ème composante de $x \in \mathbb{R}^d$, on a

$$\int g\, d\lambda = (2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{-\frac{1}{2}\sum_{i=1}^d x_i^2}\, d\lambda_d(x) = \left((2\pi)^{-1/2} \int_{\mathbb{R}} e^{-x^2/2}\, d\lambda_1(x)\right)^d = 1^d = 1.$$

1. sous-entendu : par rapport à la mesure de Lebesgue

Calculons à présent la transformée de Fourier de g :

$$\begin{aligned}
\hat{g}(u) &= (2\pi)^{-d/2} \int_{\mathbb{R}^d} d\lambda_d(x)\, e^{-i\langle u,x\rangle}\, e^{-\frac{1}{2}\langle x,x\rangle} \\
&= (2\pi)^{-d/2} \int_{\mathbb{R}^d} d\lambda_d(x)\, e^{-\frac{1}{2}\langle x+iu,x+iu\rangle}\, e^{-\frac{1}{2}\langle u,u\rangle} \\
&= (2\pi)^{-d/2} e^{-\frac{1}{2}\| u \|^2} \int_{\mathbb{R}^d} d\lambda_d(x)\, e^{-\frac{1}{2}\sum_{j=1}^d (x_j+iu_j)^2} \\
&= e^{-\frac{1}{2}\| u \|^2} \prod_{j=1}^d (2\pi)^{-1/2} \int_{\mathbb{R}} d\lambda_1(x)\, e^{-\frac{1}{2}(x+iu_j)^2},
\end{aligned}$$

par Fubini–Lebesgue. Il suffit donc de montrer que pour tout $u \in \mathbb{R}$,

$$(2\pi)^{-1/2} \int_{\mathbb{R}} d\lambda_1(x)\, e^{-\frac{1}{2}(x+iu)^2} = 1.$$

Soit

$$F(u) := \int_{\mathbb{R}} d\lambda_1(x)\, e^{-\frac{1}{2}(x+iu)^2} \qquad u \in \mathbb{R}.$$

Soit $R > 0$. Comme l'application $u \mapsto e^{-\frac{1}{2}(x+iu)^2}$ est dérivable sur $[-R, R]$ et que sa dérivée $-i(x+iu)\, e^{-\frac{1}{2}(x+iu)^2}$ est dominée sur $[-R, R]$ par $C_R(|\, x \,| + R)\, e^{-\frac{1}{2}x^2}$ (où C_R est une constante qui dépend seulement de R), qui est intégrable en x sur $\mathbb{R}$, F est dérivable sur $[-R, R]$ et sa dérivée vaut

$$F'(u) = -i \int_{\mathbb{R}} d\lambda_1(x)\, (x+iu)\, e^{-\frac{1}{2}(x+iu)^2}.$$

En décomposant l'intégrand suivant sa partie réelle et sa partie imaginaire, dont des primitives sont les parties réelle et imaginaire respectivement, de $i\, e^{-\frac{1}{2}(x+iu)^2}$, il n'est pas difficile de montrer que cette intégrale est nulle. Ainsi, comme R est arbitraire, F @ est dérivable en tout point de $\mathbb{R}$ et de dérivée nulle, donc est constante sur $\mathbb{R}$ égale à $F(0) = \sqrt{2\pi}$, ce qui achève la démonstration. □

Pour tout $\sigma > 0$, définissons à présent

$$g_\sigma(x) := \sigma^{-d}\, g\left(\frac{x}{\sigma}\right) = \left(\sigma\sqrt{2\pi}\right)^{-d}\, e^{-\frac{\| x \|^2}{2\sigma^2}}.$$

On sait alors que

$$\hat{g}_\sigma(u) = \sigma^{-d}\left(\sigma^d \hat{g}(\sigma u)\right) = e^{-\frac{\sigma^2 \| u \|^2}{2}}.$$

Lemme 17.6 *Soit μ une mesure finie sur $\mathscr{B}or(\mathbb{R}^d)$.*

a) On a l'égalité

$$g_\sigma \star \mu(x) = (2\pi)^{-d} \int_{\mathbb{R}^d} d\lambda(u)\, \hat{\mu}(u)\, e^{i\langle u,x\rangle - \frac{\sigma^2}{2}\| u \|^2}.$$

b) Pour toute fonction $h : \mathbb{R}^d \to \mathbb{R}$ continue bornée,

$$\int_{\mathbb{R}^d} h\,d\mu = \lim_{\sigma\downarrow 0} \int_{\mathbb{R}^d} g_\sigma \star \mu(x)\,h(x)\,d\lambda(x).$$

Dém. a) Calculons $g_\sigma \star \mu(x)$:

$$\begin{aligned} g_\sigma \star \mu(x) &= \int g_\sigma(x-y)\,d\mu(y) \\ &= \left(\sigma\sqrt{2\pi}\right)^{-d} \int e^{-\frac{\|x-y\|^2}{2\sigma^2}}\,d\mu(y) \\ &= \left(\sigma\sqrt{2\pi}\right)^{-d} \int d\mu(y)\hat{g}_{\frac{1}{\sigma}}(y-x) \\ &= \left(\sigma\sqrt{2\pi}\right)^{-d} \int d\mu(y) \int d\lambda(z)\,g_{\frac{1}{\sigma}}(z)\,e^{-i\langle y-x,z\rangle}. \end{aligned}$$

Par le théorème de Fubini-Lebesgue (μ est finie),

$$\begin{aligned} g_\sigma \star \mu(x) &= \left(\sigma\sqrt{2\pi}\right)^{-d} \int d\lambda(u)\,g_{\frac{1}{\sigma}}(u) \int d\mu(y)\,e^{-i\langle y-x,u\rangle} \\ &= \left(\sigma\sqrt{2\pi}\right)^{-d} \int d\lambda(u)\,e^{i\langle x,u\rangle}\,g_{\frac{1}{\sigma}}(u)\,\hat{\mu}(u) \\ &= \left(\sigma\sqrt{2\pi}\right)^{-d} \int d\lambda(u)\,e^{i\langle x,u\rangle} \left(\frac{\sqrt{2\pi}}{\sigma}\right)^{-d} e^{-\frac{\sigma^2\|u\|^2}{2}}\,\hat{\mu}(u) \\ &= (2\pi)^{-d} \int d\lambda(u)\,e^{i\langle u,x\rangle-\frac{\sigma^2\|u\|^2}{2}}\,\hat{\mu}(u). \end{aligned}$$

b) Soit

$$\begin{aligned} I_\sigma &:= \int_{\mathbb{R}^d} g_\sigma \star \mu(x)\,h(x)\,d\lambda(x) \\ &= \int_{\mathbb{R}^d} d\lambda(x)\,h(x) \int_{\mathbb{R}^d} d\mu(y)\,g_\sigma(x-y), \end{aligned}$$

double intégrale à laquelle nous allons appliquer le théorème de Fubini–Lebesgue. En effet, comme h est bornée (par une certaine constante C)

$$\int_{\mathbb{R}^d} d\mu(y) \int_{\mathbb{R}^d} d\lambda(x)\,|\,h(x)\,|.|\,g_\sigma(x-y)\,| \le C \int_{\mathbb{R}^d} d\mu(y) \int_{\mathbb{R}^d} d\lambda(x)\,g_\sigma(x-y) = C\,\mu(\mathbb{R}^d) < \infty,$$

en faisant le changement de variable $u = (x-y)/\sigma$, car on se souvient que g est une densité de probabilité. Nous en déduisons donc, grâce au même changement de variable,

$$\begin{aligned} I_\sigma &= \int_{\mathbb{R}^d} d\mu(y) \int_{\mathbb{R}^d} d\lambda(x)\,h(x)\,g_\sigma(x-y) \\ &= \int_{\mathbb{R}^d} d\mu(y) \int_{\mathbb{R}^d} d\lambda(u)\,\sigma^d\,h(y+\sigma u)\,g_\sigma(\sigma u) \\ &= \int_{\mathbb{R}^d} d\mu(y) \int_{\mathbb{R}^d} d\lambda(u)\,h(y+\sigma u)\,g(u). \end{aligned}$$

Comme la fonction $y \mapsto \int_{\mathbb{R}^d} d\lambda(u)\, h(y+\sigma u)\, g(u)$ est dominée par $C \int g\, d\lambda = C$, et que les fonctions constantes sont μ-intégrables, le théorème de convergence dominée nous permet de conclure

$$\lim_{\sigma\downarrow 0} I_\sigma = \int_{\mathbb{R}^d} d\mu(y) \lim_{\sigma\downarrow 0} \int_{\mathbb{R}^d} d\lambda(u)\, h(y+\sigma u)\, g(u).$$

Il suffit alors de voir que la limite dans l'intégrale est égale à $h(y)$, par une autre application du théorème de convergence dominée. En effet, l'interversion de la limite et de l'intégrale est permise car l'intégrand est dominé par Cg qui est λ-intégrable. Enfin, par continuité de h, $\lim_{\sigma\downarrow 0} h(y+\sigma x) = h(y)$ et g est une densité de probabilité, ce qui permet de voir que

$$\lim_{\sigma\downarrow 0} \int_{\mathbb{R}^d} d\lambda(u)\, h(y+\sigma u)\, g(u) = h(y).$$

$\square$

Lemme 17.7 *Soit $(E, \mathscr{A}, \mu)$ un espace mesuré quelconque. Soient $f, g : (E, \mathscr{A}) \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ deux fonctions mesurables.*

a) Si f et g sont μ-intégrables et que $\int_A f\, d\mu = \int_A g\, d\mu$ pour tout $A \in \mathscr{C}$, où $\mathscr{C}$ est un π-système engendrant[2] $\mathscr{A}$, alors $f = g$ μ-p.p.

b) Si $\mathscr{C}$ est seulement stable par intersections finies, mais qu'il existe une suite croissante (E_n) d'éléments de $\mathscr{A}$ et convergeant vers E, telle que $\int_{E_n} |f|\, d\mu < \infty$ et $\int_{E_n} |g|\, d\mu < \infty$ pour tout n, et que $\int_{A\cap E_n} f\, d\mu = \int_{A\cap E_n} g\, d\mu$ pour tout $A \in \mathscr{C}$, alors $f = g$ μ-p.p.

Dém. a) On définit les mesures μ^+, μ^-, ν^+ et ν^- comme les mesures de densités respectives f^+, f^-, g^+ et g^- par rapport à μ. Comme f et g sont μ-intégrables, ces quatre mesures sont finies et par hypothèse, pour tout $A \in \mathscr{C}$, $\mu^+(A) - \mu^-(A) = \nu^+(A) - \nu^-(A)$, ce que nous pouvons également écrire

$$\mu^+(A) + \nu^-(A) = \nu^+(A) + \mu^-(A).$$

En d'autres termes les mesures $\mu^+ + \nu^-$ et $\nu^+ + \mu^-$ coïncident sur un π-système engendrant $\mathscr{A}$, donc coïncident sur $\mathscr{A}$. L'égalité précédente est donc toujours satisfaite pour $A \in \mathscr{A}$ et comme les quatres termes sont finis, on peut écrire $\mu^+(A) - \mu^-(A) = \nu^+(A) - \nu^-(A)$, autrement dit

$$\int_A f\, d\mu = \int_A g\, d\mu \qquad A \in \mathscr{A}.$$

Il est ensuite classique de voir pourquoi ceci implique que $f = g$ μ-p.p., en prenant $A = \{f > g\}$ puis en raisonnant par symétrie sur f et g.

b) Il suffit d'appliquer la méthode précédente aux mesures traces de μ et ν sur E_n et au π-système $\mathscr{C} \cup \{E\}$. On obtient alors que pour tout n $f\mathbb{1}_{E_n} = g\mathbb{1}_{E_n}$ μ-p.p., ce qui implique l'égalité μ-p.p. entre f et g. $\square$

2. au sens où $\sigma(\mathscr{C}) = \mathscr{A}$

Théorème 17.8 (injectivité de la transformée de Fourier) *a) Si μ et ν sont deux mesures finies sur $\mathscr{B}(\mathbb{R}^d)$ telles que $\hat{\mu} = \hat{\nu}$, alors $\mu = \nu$.*

b) Si f et g sont deux fonctions complexes λ-intégrables sur $\mathbb{R}^d$ telles que $\hat{f} = \hat{g}$, alors $f = g$ λ-p.p.

Dém. a) D'après le Lemme 17.6 a), si $\hat{\mu} = \hat{\nu}$, alors pour tout $\sigma > 0$, $g_\sigma \star \mu = g_\sigma \star \nu$, ce qui implique, d'après le Lemme 17.6 b), que pour toute fonction h continue bornée, $\int h\,d\mu = \int h\,d\nu$. Ceci implique que $\mu = \nu$. En effet, si $\mathscr{C}$ est la classe des rectangles de la forme $A = \prod_{j=1}^d]-\infty, a_j[$ (où a_j peut éventuellement être égal à $+\infty$), alors il est facile de construire une suite (h_n) de fonctions continues, toutes bornées par $\mathbb{1}_{\mathbb{R}^d}$, telles que $\lim_{n\to\infty} h_n = \mathbb{1}_A$. Alors par convergence dominée, $\lim_n \int h_n\,d\mu = \mu(A)$, et comme $\int h_n\,d\mu = \int h_n\,d\nu$, on a l'égalité $\mu(A) = \nu(A)$. Comme $\mathscr{C}$ est un π-système engendrant $\mathscr{B}(\mathbb{R}^d)$, μ et ν coïncident sur $\mathscr{B}(\mathbb{R}^d)$.

b) On pourrait montrer une version du lemme 17.6 où l'on a remplacé μ par $f.\lambda$, où @ $f \in \mathscr{L}^1_{\mathbb{C}}(\lambda)$ (d'abord pour $f \geq 0$, puis par linéarité...). Donc si $\hat{f} = \hat{g}$, par les mêmes arguments que a), on peut montrer que $\int_A f\,d\lambda = \int_A g\,d\lambda$ pour tout $A \in \mathscr{C}$, ce qui prouve, grâce au lemme 17.7 a) appliqué aux parties réelles et imaginaires de f et g, que $f = g$ λ-p.p. □

Remarque 17.9 *En fait, grâce au Lemme 17.6, on a montré que pour toute fonction h continue bornée,*

$$\int h\,d\mu = \lim_{\sigma\downarrow 0} (2\pi)^{-d} \int d\lambda(x)\,h(x) \int d\lambda(u)\,\hat{\mu}(u)\,e^{i\langle u,x\rangle - \frac{\sigma^2}{2}\|u\|^2},$$

ce qui constitue une formule d'inversion *pour les mesures finies, comme le précise le théorème qui suit.*

Théorème 17.10 *a) Si μ est une mesure finie dont la transformée de Fourier $\hat{\mu}$ est λ-intégrable, alors elle admet une densité continue et bornée g par rapport à λ donnée par*

$$g(x) = (2\pi)^{-d} \int_{\mathbb{R}^d} e^{i\langle u,x\rangle}\,\hat{\mu}(u)\,d\lambda(u) \qquad x \in \mathbb{R}^d.$$

*b) Si $f \in \mathscr{L}^1_{\mathbb{C}}(\lambda)$ est telle que $\hat{f} \in \mathscr{L}^1_{\mathbb{C}}(\lambda)$, alors pour λ-**presque tout** x,*

$$f(x) = (2\pi)^{-d} \int_{\mathbb{R}^d} e^{i\langle u,x\rangle}\,\hat{f}(u)\,d\lambda(u).$$

Remarque 17.11 *On ne peut bien sûr pas espérer s'affranchir de l'égalité presque partout dans b), puisque le membre de droite est une fonction continue et bornée, ce qui n'est pas forcément le cas de f.*

Dém. a) Soit g définie comme dans l'énoncé du théorème. Comme $\hat{\mu}$ est intégrable, g est bornée. De la remarque qui précède l'énoncé, pour toute fonction h continue bornée,

$$\int h\,d\mu = \lim_{\sigma\downarrow 0}(2\pi)^{-d}\int d\lambda(x)\,h(x)\int d\lambda(u)\,\hat{\mu}(u)\,e^{i\langle u,x\rangle-\frac{\sigma^2}{2}\|\,u\,\|^2}.$$

Si h est à support compact, on peut appliquer deux fois le théorème de convergence dominée pour obtenir l'égalité $\int h\,d\mu = \int hg\,d\lambda$. En effet, tout d'abord, la fonction $x \mapsto h(x)\int d\lambda(u)\,\hat{\mu}(u)\,e^{i\langle u,x\rangle-\frac{\sigma^2}{2}\|\,u\,\|^2}$ est dominée en module par $C\,|\,h\,|$, qui est λ-intégrable (car $\hat{\mu}$ est supposée intégrable et h à support compact), donc

$$\int h\,d\mu = (2\pi)^{-d}\int d\lambda(x)\,h(x)\lim_{\sigma\downarrow 0}\int d\lambda(u)\,\hat{\mu}(u)\,e^{i\langle u,x\rangle-\frac{\sigma^2}{2}\|\,u\,\|^2}.$$

Ensuite, la fonction $u \mapsto \hat{\mu}(u)\,e^{i\langle u,x\rangle-\frac{\sigma^2}{2}\|\,u\,\|^2}$ est dominée en module par $|\,\hat{\mu}\,|$, qui est intégrable par hypothèse, donc

$$\lim_{\sigma\downarrow 0}\int \hat{\mu}(u)\,e^{i\langle u,x\rangle-\frac{\sigma^2}{2}\|\,u\,\|^2} = (2\pi)^d\ g(x).$$

Par la même méthode que précédemment, avec $\mathscr{C}$ la famille des rectangles de la forme $\prod_{j=1}^d]a_j,b_j]$, où les a_j et b_j sont finis, on obtient l'égalité $\mu(A)=\int_A g\,d\lambda$ pour tout $A\in\mathscr{C}$. Comme $\mathscr{C}$ est stable par intersections finies, et que la suite des rectangles $]-n,n]^d$ est une suite d'éléments de $\mathscr{C}$ qui converge vers $\mathbb{R}^d$, on applique le lemme 17.7 b) à la fonction nulle et à la partie imaginaire de g, pour voir que l'égalité $\mu(A)=\int g\,d\lambda = \int_A \Re(g)\,d\lambda + i\int_A \Im(g)\,d\lambda$ pour tout $A\in\mathscr{C}$, implique que $\int_A \Im(g)\,d\lambda = 0$ et donc que $\Im(g)=0$ λ-p.p. (on vérifie que $\Im(g)$ est localement intégrable car elle est bornée). Donc g est réelle λ-p.p., mais comme g est continue, g est réelle partout. Soient alors

$$\nu^+(A) := \int_A g^+\,d\lambda \quad , \quad \nu^-(A) := \int_A g^-\,d\lambda.$$

Comme g est bornée, ν^+ et ν^- sont finies sur $\mathscr{C}$, et pour tout $A\in\mathscr{C}$, $\mu(A)=\nu^+(A)-\nu^-(A)$, ce qui implique que $\mu+\nu^-$ et ν^+ coïncident sur $\mathscr{C}$, et donc sur $\mathscr{B}(\mathbb{R}^d)$.

Or pour tout $N\in\mathscr{B}(\mathbb{R}^d)$ de λ-mesure nulle, $\nu^+(N)=\int_N g^+\,d\lambda = 0$ et de la même manière $\nu^-(N)=0$. L'égalité $\mu+\nu^-=\nu^+$ implique alors que $\mu(N)=0$, c'est-à-dire que μ est absolument continue par rapport à la mesure de Lebesgue. Comme μ et λ sont σ-finies, le théorème de Radon–Nikodym assure qu'il existe une fonction positive $\varphi\in\mathscr{L}^1(\lambda)$ telle que pour tout $A\in\mathscr{B}(\mathbb{R}^d)$, $\mu(A)=\int_A\varphi\,d\lambda$. Et si $A\in\mathscr{C}$, ceci peut s'écrire $\int_A g\,d\lambda = \int_A\varphi\,d\lambda$, car g est bornée. Donc d'après le lemme 17.7 b), $\varphi = g$ λ-p.p., et en particulier g est positive λ-p.p., mais comme g est continue, g est en fait positive partout. Enfin, l'égalité $\varphi=g$ λ-p.p. implique que $g\in\mathscr{L}^1(\lambda)$, puis que

$$\mu(A)=\int_A\varphi\,d\lambda = \int_A g\,d\lambda \qquad A\in\mathscr{B}(\mathbb{R}^d).$$

b) Si $f\geq 0$, on applique a) à la mesure μ de densité f par rapport à λ (car alors $\hat{\mu}=\hat{f}$). On peut alors conclure que μ admet une densité g par rapport à λ (où g est donnée par la formule d'inversion), ce qui implique que $f=g$ λ-p.p. Le cas général se traite en décomposant f suivant $f=\Re(f)^+-\Re(f)^-+i\Im(f)^+-i\Im(f)^-$. $\square$

Remarque 17.12 *On pourra lire avec avantage les trois dernières pages du polycopié de Jean Jacod concernant la transformée de Fourier dans* L^2*.*