Chapitre 4

Intégration, espace $\mathbf{L}^1$

Dans tout ce chapitre, $(\Omega, \mathcal{F}, \mu)$ est un espace mesuré. Nous allons définir et étudier l'intégrale sur cet espace, de fonctions mesurables de signe *non constant*. Ayant déjà vu l'intégrale de fonctions mesurables positives, on pourrait définir de manière analogue l'intégrale des fonctions mesurables *négatives* (i.e. $\forall \omega \in \Omega$, $f(\omega) \in [-\infty, 0]$) par un simple changement de signe : $\int_\Omega f \,\mathrm{d}\mu := -\int_\Omega (-f)\,\mathrm{d}\mu$. L'étape suivante serait de décomposer une fonction f de signe non constant en $f = f^+ - f^-$, différence de deux fonctions positives et de poser $\int_\Omega f\,\mathrm{d}\mu := \int_\Omega f^+\,\mathrm{d}\mu - \int_\Omega f^-\,\mathrm{d}\mu$. On voit tout de suite poindre le danger d'une écriture « $+\infty - \infty$ », dépourvue de sens. Ceci nous conduit à renoncer à définir une intégrale pour *toutes* les fonctions mesurables et à introduire l'ensemble $\mathcal{L}^1(\mu)$ des fonctions μ-intégrables pour lesquelles l'intégrale $\int_\Omega f\,\mathrm{d}\mu$ sera toujours un *réel* ou un complexe, mais jamais infini.

4.1 Fonctions μ-intégrables

Définition 4.1 (Intégrabilités). *Soit f une application définie sur Ω et à valeurs dans $\mathbb{K} = \mathbb{R}$, $\overline{\mathbb{R}}$ ou $\mathbb{C}$. On dit qu'elle est μ-intégrable si*

a) elle est mesurable $\mathcal{F}$-Bor$(\mathbb{K})$ *;*

b) l'intégrale $\int_\Omega |f|\,\mathrm{d}\mu$ est finie, *$|f|$ désignant la valeur absolue (cas $\mathbb{R}$ et $\overline{\mathbb{R}}$) ou le module (cas $\mathbb{C}$).*

Soit $A \in \mathcal{F}$, on dit que f est μ-intégrable sur A si la fonction $f\mathbf{1}_A$ est μ-intégrable au sens ci-dessus.

Remarques 4.2.

1. La mesurabilité ne concerne que les tribus ; par contre l'intégrabilité est toujours relative à une mesure particulière.
2. Une fonction mesurable positive a toujours une intégrale (et ce par rapport à n'importe quelle mesure). Elle n'est pas forcément intégrable par rapport à une mesure donnée.

3. Il importe de ne pas omettre la condition a), car $|f|$ peut très bien être mesurable et avoir une intégrale finie sans que f soit mesurable. Pour s'en convaincre, prenons pour μ une mesure de probabilité et pour f la fonction $\mathbf{1}_B - \mathbf{1}_{B^c}$ où $B \notin \mathcal{F}$. Alors f n'est pas mesurable puisque $f^{-1}(\{1\}) = B \notin \mathcal{F}$. Par contre $|f|$ est la fonction constante sur Ω valant 1, elle est mesurable et $\int_\Omega |f| \,\mathrm{d}\mu = 1 < +\infty$.
4. On peut vérifier que la μ-intégrabilité de f sur $A \in \mathcal{F}$ équivaut à la μ_A-intégrabilité de la restriction de f à A, où μ_A est la mesure restriction de μ à la tribu trace de $\mathcal{F}$ sur A (cf. remarque 3.7).

Si f est à valeurs dans $\overline{\mathbb{R}}$ ou $\mathbb{R}$, rappelons que sa partie positive f^+ et sa partie négative f^- sont les fonctions positives

$$f^+ = \max(f; 0), \qquad f^- = -\min(f; 0) = \max(-f; 0).$$

Si $f(\omega) > 0$, $f^+(\omega) = f(\omega) = |f(\omega)|$ et $f^-(\omega) = 0$. Si $f(\omega) < 0$, $f^+(\omega) = 0$ et $f^-(\omega) = -f(\omega) = |f(\omega)|$. Enfin si $f(\omega) = 0$, $f^+(\omega) = f^-(\omega) = 0$. On voit ainsi que

$$|f| = f^+ + f^-, \qquad f = f^+ - f^-, \tag{4.1}$$

en notant au passage que cette dernière égalité ne génère *jamais* l'expression non définie « $+\infty - \infty$ », car pour chaque ω, au moins l'un des deux termes $f^+(\omega)$ et $f^-(\omega)$ vaut zéro. Il résulte immédiatement de (4.1) que f est μ-intégrable si et seulement si f^+ et f^- sont mesurables et ont des intégrales finies.

Concernant les fonctions f à valeurs complexes, l'équivalence entre la mesurabilité de f et celles de $\operatorname{Re} f$ et $\operatorname{Im} f$ ainsi que les inégalités

$$\max\bigl(|\operatorname{Re} f|; |\operatorname{Im} f|\bigr) \leq |f| \leq |\operatorname{Re} f| + |\operatorname{Im} f|, \tag{4.2}$$

montrent que f est μ-intégrable si et seulement si les fonctions réelles $\operatorname{Re} f$ et $\operatorname{Im} f$ le sont.

Définition 4.3 (Intégrales). *Soit f une application définie sur Ω, à valeurs dans $\mathbb{K} = \mathbb{R}$, $\overline{\mathbb{R}}$ ou $\mathbb{C}$ et μ-intégrable. Son intégrale par rapport à μ sur Ω est alors définie par :*

– *si* $\mathbb{K} = \mathbb{R}$ *ou* $\overline{\mathbb{R}}$, $\displaystyle\int_\Omega f \,\mathrm{d}\mu := \int_\Omega f^+ \,\mathrm{d}\mu - \int_\Omega f^- \,\mathrm{d}\mu\,;$

– *si* $\mathbb{K} = \mathbb{C}$, $\displaystyle\int_\Omega f \,\mathrm{d}\mu := \int_\Omega \operatorname{Re} f \,\mathrm{d}\mu + i \int_\Omega \operatorname{Im} f \,\mathrm{d}\mu.$

Si $A \in \mathcal{F}$, l'intégrale de f sur A est définie par

$$\int_A f \,\mathrm{d}\mu := \int_\Omega f \mathbf{1}_A \,\mathrm{d}\mu.$$

Cette intégrale est aussi notée $\int_A f(\omega) \,\mathrm{d}\mu(\omega)$ ou $\int_A f(\omega)\mu(\mathrm{d}\omega)$, y compris lorsque $A = \Omega$.

Lorsque l'on considère l'intégrale comme une application définie sur un ensemble de fonctions et à valeurs dans $\mathbb{K}$, il est parfois commode d'utiliser la notation $\mu(f) := \int_\Omega f \,\mathrm{d}\mu$. De ce point de vue, $\int_A f \,\mathrm{d}\mu = \mu(f\mathbf{1}_A)$ et $\mu(A) = \mu(\mathbf{1}_A)$.

Soit f à valeurs dans $\mathbb{R}_+$ ou $\overline{\mathbb{R}}_+$ et μ-intégrable. Il est clair que la définition 4.3 de $\int_\Omega f \,\mathrm{d}\mu$ est compatible avec celle de l'intégrale d'une fonction mesurable positive au sens du chapitre 3.

Si f est à valeurs complexes et intégrable, il est clair par la définition 4.3 que

$$\operatorname{Re}\Big(\int_\Omega f \,\mathrm{d}\mu\Big) = \int_\Omega \operatorname{Re} f \,\mathrm{d}\mu, \quad \operatorname{Im}\Big(\int_\Omega f \,\mathrm{d}\mu\Big) = \int_\Omega \operatorname{Im} f \,\mathrm{d}\mu$$

Dans le langage des probabilités, la définition 4.3 est celle de l'espérance mathématique d'une variable aléatoire réelle ou complexe. Rappelons qu'une telle variable aléatoire est simplement une application mesurable $\mathcal{F}$-$\mathrm{Bor}(\mathbb{K})$, avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

Définition 4.4 (Espérance). *Soit $(\Omega, \mathcal{F}, \mathbf{P})$ un espace probabilisé et X une variable aléatoire définie sur Ω et à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$. Si X est $\mathbf{P}$-intégrable, on définit son espérance mathématique (sous $\mathbf{P}$) par*

$$\mathbf{E}X := \int_\Omega X \,\mathrm{d}\mathbf{P} = \int_\Omega X(\omega) \,\mathrm{d}\mathbf{P}(\omega).$$

Avant d'établir les propriétés générales de l'intégrale, examinons deux exemples importants dont la combinaison va nous donner le calcul d'intégrales par rapport à une mesure discrète quelconque.

Proposition 4.5. *Soit ω_0 fixé dans Ω et δ_{ω_0} la mesure de Dirac en ce point. Une fonction mesurable $f : \Omega \to \mathbb{K}$ est toujours δ_{ω_0}-intégrable si $\mathbb{K} = \mathbb{R}$, ou $\mathbb{C}$. Si $\mathbb{K} = \overline{\mathbb{R}}$, la δ_{ω_0}-intégrabilité de f équivaut à la finitude de $f(\omega_0)$. Quand f est δ_{ω_0}-intégrable, on a*

$$\int_\Omega f \,\mathrm{d}\delta_{\omega_0} = f(\omega_0).$$

Démonstration. Par mesurabilité de f, $|f| \in \mathcal{M}_+$ et par la proposition 3.23,

$$\int_\Omega |f| \,\mathrm{d}\delta_{\omega_0} = |f|(\omega_0) = |f(\omega_0)|.$$

La δ_{ω_0}-intégrabilité de f se réduit donc dans ce cas à la finitude de $f(\omega_0)$, laquelle est automatique si $\mathbb{K} = \mathbb{R}$, ou $\mathbb{C}$. Quand f est intégrable et $\mathbb{K} = \mathbb{R}$, la définition 4.3 et la proposition 3.23 appliquée aux fonctions mesurables positives f^+ et f^- nous donnent :

$$\int_\Omega f \,\mathrm{d}\delta_{\omega_0} = \int_\Omega f^+ \,\mathrm{d}\delta_{\omega_0} - \int_\Omega f^- \,\mathrm{d}\delta_{\omega_0} = f^+(\omega_0) - f^-(\omega_0) = f(\omega_0).$$

Le cas complexe se ramène au cas réel par séparation des parties réelles et imaginaires. $\square$

Proposition 4.6. *Soit J une partie de $\mathbb{N}$, $(\mu_j)_{j\in J}$ une suite (finie ou infinie) de mesures sur $(\Omega, \mathcal{F})$ et $(a_j)_{j\in J}$ une suite de réels* strictement *positifs. Notons μ la mesure $\sum_{j\in J} a_j\mu_j$ sur $(\Omega, \mathcal{F})$. Une application $f : \Omega \to \mathbb{K}$ avec $\mathbb{K} = \mathbb{R}$, $\overline{\mathbb{R}}$ ou $\mathbb{C}$, est μ-intégrable si et seulement si elle est* mesurable $\mathcal{F}$-Bor$(\mathbb{K})$ *et*

$$\sum_{j\in J} a_j \int_\Omega |f| \,\mathrm{d}\mu_j < +\infty. \tag{4.3}$$

Quand f est μ intégrable,

$$\int_\Omega f \,\mathrm{d}\mu = \sum_{j\in J} a_j \int_\Omega f \,\mathrm{d}\mu_j. \tag{4.4}$$

Démonstration. La caractérisation de la μ-intégrabilité n'est que la reproduction de la définition 4.1, dans laquelle on a implanté la formule de calcul de $\int_\Omega |f| \,\mathrm{d}\mu$ fournie par la proposition 3.24. Notons que dans la proposition 3.24, on n'interdit pas aux coefficients a_k d'être nuls, donc la formule de calcul des intégrales de fonctions mesurables *positives* vaut aussi pour les mesures $\sum_{j\in J} a_j\mu_j$ avec $J \subset \mathbb{N}$.

Supposons maintenant f μ-intégrable. La condition (4.3) et la stricte positivité de chaque a_j, $j \in J$ impliquent la finitude de $\int_\Omega |f| \,\mathrm{d}\mu_j$ et donc la μ_j-intégrabilité[1] de f.

Lorsque $\mathbb{K} = \mathbb{R}$ ou $\overline{\mathbb{R}}$, en appliquant la proposition 3.24 aux fonctions mesurables positives f^+ et f^-, on obtient

$$\int_\Omega f^+ \,\mathrm{d}\mu = \sum_{j\in J} a_j \int_\Omega f^+ \,\mathrm{d}\mu_j, \quad \int_\Omega f^- \,\mathrm{d}\mu = \sum_{j\in J} a_j \int_\Omega f^- \,\mathrm{d}\mu_j. \tag{4.5}$$

Comme $\int_\Omega f^+ \,\mathrm{d}\mu$ et $\int_\Omega f^- \,\mathrm{d}\mu$ sont majorées par $\int_\Omega |f| \,\mathrm{d}\mu$, la μ-intégrabilité de f entraîne la convergence vers un réel positif *fini* des séries[2] figurant dans (4.5). On peut donc former leur différence qui est ainsi une série convergente dans $\mathbb{R}$ et l'application de la définition 4.3 nous donne

$$\int_\Omega f \,\mathrm{d}\mu = \int_\Omega f^+ \,\mathrm{d}\mu - \int_\Omega f^- \,\mathrm{d}\mu = \sum_{j\in J} a_j \left\{ \int_\Omega f^+ \,\mathrm{d}\mu_j - \int_\Omega f^- \,\mathrm{d}\mu_j \right\} = \sum_{j\in J} a_j \int_\Omega f \,\mathrm{d}\mu_j.$$

□

Corollaire 4.7 (Représentation d'une série comme intégrale). *Soit $\sum_{k=0}^{+\infty} u_k$ une série à termes réels ou complexes,* absolument *convergente. Elle peut s'écrire comme intégrale par rapport à la mesure de comptage $\nu := \sum_{k\in\mathbb{N}} \delta_k$. En effet, en notant $f : \mathbb{N} \to \mathbb{C}$, $k \mapsto f(k) := u_k$, on a :*

$$\sum_{k=0}^{+\infty} u_k = \int_{\mathbb{N}} f \,\mathrm{d}\nu.$$

[1] C'est précisément ici que l'hypothèse $a_j > 0$ est importante. En effet si $a_j = 0$ et $\int_\Omega |f| \,\mathrm{d}\mu_j = +\infty$, on a $a_j \int_\Omega |f| \,\mathrm{d}\mu_j = 0$, ce qui n'empêche aucunement la réalisation de (4.3). Par contre $\int_\Omega f \,\mathrm{d}\mu_j$ *n'est pas défini* et (4.4) devient illégitime (ne pas confondre « non défini » et « infini »).

[2] Si J est fini, ce sont des somme finies de réels et il n'y a aucune justification à donner.

Démonstration. Munissons $\mathbb{N}$ de la tribu $\mathcal{P}(\mathbb{N})$, pour laquelle toute application $f : \mathbb{N} \to \mathbb{C}$ est mesurable. L'intégrabilité de f se réduit alors à la condition $\int_{\mathbb{N}} |f| \,\mathrm{d}\nu < +\infty$. En appliquant la proposition 3.24 à la mesure ν et à la fonction mesurable positive $|f|$, cette condition s'écrit

$$\int_{\mathbb{N}} |f| \,\mathrm{d}\nu = \sum_{k\in\mathbb{N}} \int_{\mathbb{N}} |f| \,\mathrm{d}\delta_k = \sum_{k\in\mathbb{N}} |f|(k) = \sum_{k\in\mathbb{N}} |f(k)| = \sum_{k\in\mathbb{N}} |u_k| < +\infty.$$

Ainsi la ν-intégrabilité de f résulte de la convergence absolue de la série $\sum_{k=0}^{+\infty} u_k$. Par application de (4.4), on a

$$\int_{\mathbb{N}} f \,\mathrm{d}\nu = \sum_{k\in\mathbb{N}} \int_{\mathbb{N}} f \,\mathrm{d}\delta_k = \sum_{k\in\mathbb{N}} f(k) = \sum_{k\in\mathbb{N}} u_k.$$

□

Nous passons maintenant à l'étude des propriétés générales de l'intégrale.

Définition 4.8 (Espace $\mathcal{L}^1(\mu)$). *Pour $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, on note $\mathcal{L}^1_{\mathbb{K}}(\mu)$ l'ensemble des fonctions μ-intégrables $f : \Omega \to \mathbb{K}$.*

Proposition 4.9 (Linéarité de l'intégrale). *Pour $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, l'ensemble $\mathcal{L}^1_{\mathbb{K}}(\mu)$ est un espace vectoriel sur $\mathbb{K}$. L'intégration $f \mapsto \int_\Omega f \,\mathrm{d}\mu$ considérée comme application de $\mathcal{L}^1_{\mathbb{K}}(\mu)$ dans $\mathbb{K}$ est une forme linéaire :*

$$\forall a, b \in \mathbb{K},\ \forall f, g \in \mathcal{L}^1_{\mathbb{K}}(\mu), \qquad \int_\Omega (af + bg) \,\mathrm{d}\mu = a \int_\Omega f \,\mathrm{d}\mu + b \int_\Omega g \,\mathrm{d}\mu. \tag{4.6}$$

On a exclu délibérément le cas $\mathbb{K} = \overline{\mathbb{R}}$ dans la définition de $\mathcal{L}^1_{\mathbb{K}}(\mu)$, car si f et g à valeurs dans $\mathbb{K} = \overline{\mathbb{R}}$ sont μ-intégrables, on peut très bien avoir pour certains ω, $f(\omega) = +\infty$ et $g(\omega) = -\infty$ de sorte que $(f+g)(\omega)$ n'est pas défini et $\mathcal{L}^1_{\overline{\mathbb{R}}}(\mu)$ ne peut pas être un espace vectoriel.

Preuve de la proposition 4.9. Soient f et g dans $\mathcal{L}^1_{\mathbb{K}}(\mu)$ et $h := f + g$. Alors h est mesurable et compte-tenu de la croissance et de l'additivité de l'intégrale dans $\mathcal{M}_+$, l'inégalité $|h| \le |f| + |g|$ nous donne :

$$\int_\Omega |h| \,\mathrm{d}\mu \le \int_\Omega (|f| + |g|) \,\mathrm{d}\mu = \int_\Omega |f| \,\mathrm{d}\mu + \int_\Omega |g| \,\mathrm{d}\mu < +\infty.$$

Ainsi h est μ-intégrable et $\mathcal{L}^1_{\mathbb{K}}(\mu)$ est stable par addition. Pour vérifier l'additivité de l'intégrale dans $\mathcal{L}^1_{\mathbb{K}}(\mu)$, regardons d'abord le cas $\mathbb{K} = \mathbb{R}$. On peut alors écrire

$$h^+ - h^- = f^+ - f^- + g^+ - g^-,$$

d'où l'égalité dans $\mathcal{M}_+$:

$$h^+ + f^- + g^- = h^- + f^+ + g^+.$$

Par additivité de l'intégrale dans $\mathcal{M}_+$, on obtient

$$\int_\Omega h^+ \,\mathrm{d}\mu + \int_\Omega f^- \,\mathrm{d}\mu + \int_\Omega g^- \,\mathrm{d}\mu = \int_\Omega h^- \,\mathrm{d}\mu + \int_\Omega f^+ \,\mathrm{d}\mu + \int_\Omega g^+ \,\mathrm{d}\mu$$

et la finitude de ces 6 intégrales nous permet d'écrire :

$$\int_\Omega h^+ \,\mathrm{d}\mu - \int_\Omega h^- \,\mathrm{d}\mu = \int_\Omega f^+ \,\mathrm{d}\mu - \int_\Omega f^- \,\mathrm{d}\mu + \int_\Omega g^+ \,\mathrm{d}\mu - \int_\Omega g^- \,\mathrm{d}\mu,$$

ce qui donne bien $\int_\Omega (f+g)\,\mathrm{d}\mu = \int_\Omega f\,\mathrm{d}\mu + \int_\Omega g\,\mathrm{d}\mu$.

Lorsque $\mathbb{K} = \mathbb{C}$, la définition de l'intégrale : $\int_\Omega h\,\mathrm{d}\mu = \int_\Omega \operatorname{Re} h\,\mathrm{d}\mu + i\int_\Omega \operatorname{Im} h\,\mathrm{d}\mu$ et les égalités $\operatorname{Re}(f+g) = \operatorname{Re} f + \operatorname{Re} g$, $\operatorname{Im}(f+g) = \operatorname{Im} f + \operatorname{Im} g$ nous ramènent au cas précédent.

Soient $f \in \mathcal{L}^1_{\mathbb{K}}(\mu)$ et $a \in \mathbb{K}$. L'intégrabilité de af résulte de la relation $|af| = |a||f|$ et de l'homogénéité de l'intégrale dans $\mathcal{M}_+$:

$$\forall c \in \mathbb{R}_+,\ \forall \varphi \in \mathcal{M}_+, \qquad \int_\Omega c\varphi\,\mathrm{d}\mu = c\int_\Omega \varphi\,\mathrm{d}\mu. \tag{4.7}$$

Donc $af \in \mathcal{L}^1_{\mathbb{K}}(\mu)$ et ceci achève de prouver que $\mathcal{L}^1_{\mathbb{K}}(\mu)$ est bien un espace vectoriel.

Pour vérifier que $\int_\Omega af\,\mathrm{d}\mu = a\int_\Omega f\,\mathrm{d}\mu$, commençons par le cas $\mathbb{K} = \mathbb{R}$. Si $a > 0$, $(af)^+ = af^+$ et $(af)^- = af^-$ d'où grâce à (4.7) :

$$\int_\Omega af\,\mathrm{d}\mu = \int_\Omega af^+\,\mathrm{d}\mu - \int_\Omega af^-\,\mathrm{d}\mu = a\int_\Omega f^+\,\mathrm{d}\mu - a\int_\Omega f^-\,\mathrm{d}\mu = a\int_\Omega f\,\mathrm{d}\mu.$$

Si $a < 0$, $(af)^+ = |a|f^-$ et $(af)^- = |a|f^+$, d'où

$$\begin{aligned}\int_\Omega af\,\mathrm{d}\mu &= \int_\Omega |a|f^-\,\mathrm{d}\mu - \int_\Omega |a|f^+\,\mathrm{d}\mu \\ &= |a|\int_\Omega f^-\,\mathrm{d}\mu - |a|\int_\Omega f^+\,\mathrm{d}\mu = -a\int_\Omega f^-\,\mathrm{d}\mu + a\int_\Omega f^+\,\mathrm{d}\mu = a\int_\Omega f\,\mathrm{d}\mu.\end{aligned}$$

Le cas $a = 0$ est immédiat.

Enfin, prenons $\mathbb{K} = \mathbb{C}$ et notons $\alpha = \operatorname{Re} a$, $\beta = \operatorname{Im} a$. Alors

$$af = (\alpha + i\beta)(\operatorname{Re} f + i\operatorname{Im} f) = (\alpha\operatorname{Re} f - \beta\operatorname{Im} f) + i(\alpha\operatorname{Im} f + \beta\operatorname{Re} f),$$

de sorte qu'en utilisant la définition de l'intégrale dans le cas $\mathbb{K} = \mathbb{C}$ et (4.6) déjà établie dans le cas réel, on obtient

$$\begin{aligned}\int_\Omega af\,\mathrm{d}\mu &= \int_\Omega (\alpha\operatorname{Re} f - \beta\operatorname{Im} f)\,\mathrm{d}\mu + i\int_\Omega (\alpha\operatorname{Im} f + \beta\operatorname{Re} f)\,\mathrm{d}\mu \\ &= (\alpha + i\beta)\int_\Omega \operatorname{Re} f\,\mathrm{d}\mu + (\alpha i - \beta)\int_\Omega \operatorname{Im} f\,\mathrm{d}\mu \\ &= (\alpha + i\beta)\Big(\int_\Omega \operatorname{Re} f\,\mathrm{d}\mu + i\int_\Omega \operatorname{Im} f\,\mathrm{d}\mu\Big) = a\int_\Omega f\,\mathrm{d}\mu.\end{aligned}$$

Ceci achève la vérification de (4.6) dans le cas complexe. □

Avant de poursuivre, il est utile de reformuler (4.6) dans le cadre probabiliste.

Corollaire 4.10 (Linéarité de l'espérance). *L'espérance des variables aléatoires à valeurs réelles ou complexes, définies sur* $(\Omega, \mathcal{F}, \mathbf{P})$ *et* $\mathbf{P}$*-intégrables est* linéaire.

Il est intéressant de noter qu'avec les définitions de l'espérance données en DEUG, il était *impossible* de prouver cette linéarité (à moins de se restreindre aux variables aléatoires discrètes). En effet dans le cas d'une variable aléatoire à densité g, on définissait l'espérance comme une intégrale $\int_{-\infty}^{+\infty} xg(x)\,\mathrm{d}x$. L'ennui c'est que si X et Y ont des densités, rien ne dit que $X+Y$ soit aussi à densité.

Proposition 4.11. *Avec* $\mathbb{K} = \mathbb{R}$ *ou* $\mathbb{C}$*, on a*

a) Si $f \in \mathcal{L}^1_{\mathbb{K}}(\mu)$*, alors* $|f| \in \mathcal{L}^1_{\mathbb{K}}(\mu)$ *et*

$$\Big|\int_\Omega f\,\mathrm{d}\mu\Big| \leq \int_\Omega |f|\,\mathrm{d}\mu. \tag{4.8}$$

b) Si $f, g : \Omega \to \mathbb{K}$ *sont deux fonctions mesurables telles que* $|f| \leq |g|$ *et si* g *est* μ*-intégrable, alors* f *l'est aussi.*

c) Si $f, g \in \mathcal{L}^1_{\mathbb{R}}(\mu)$ *sont telles que* $f \leq g$*, alors* $\int_\Omega f\,\mathrm{d}\mu \leq \int_\Omega g\,\mathrm{d}\mu$*.*

Démonstration. Pour prouver a), on remarque d'abord que l'intégrabilité de $|f|$ découle immédiatement de la définition de celle de f : si f est mesurable, $|f|$ l'est aussi et les deux fonctions ont même valeur absolue (module). Pour établir (4.8) dans le cas $\mathbb{K} = \mathbb{R}$, il suffit de noter que f^+ et f^- étant positives, leurs intégrales le sont aussi, d'où :

$$\begin{aligned}\Big|\int_\Omega f\,\mathrm{d}\mu\Big| = \Big|\int_\Omega f^+\,\mathrm{d}\mu - \int_\Omega f^-\,\mathrm{d}\mu\Big| &\leq \Big|\int_\Omega f^+\,\mathrm{d}\mu\Big| + \Big|\int_\Omega f^-\,\mathrm{d}\mu\Big| \\ &= \int_\Omega f^+\,\mathrm{d}\mu + \int_\Omega f^-\,\mathrm{d}\mu = \int_\Omega (f^+ + f^-)\,\mathrm{d}\mu = \int_\Omega |f|\,\mathrm{d}\mu.\end{aligned}$$

Dans le cas $\mathbb{K} = \mathbb{C}$, puisque f est dans $\mathcal{L}^1_{\mathbb{C}}(\mu)$, $w := \int_\Omega f\,\mathrm{d}\mu$ est un nombre complexe. Il existe alors $c \in \mathbb{C}$ tel que $|c| = 1$ et cw soit un réel positif ou nul (si $w \neq 0$, prendre $c = \overline{w}/|w|$, si $w = 0$, $c = 1$ convient). On a ainsi $|w| = |cw| = cw = \int_\Omega cf\,\mathrm{d}\mu \in \mathbb{R}$ d'après (4.6), d'où

$$\Big|\int_\Omega f\,\mathrm{d}\mu\Big| = \int_\Omega cf\,\mathrm{d}\mu = \mathrm{Re}\Big(\int_\Omega cf\,\mathrm{d}\mu\Big) = \int_\Omega \mathrm{Re}(cf)\,\mathrm{d}\mu \leq \int_\Omega |cf|\,\mathrm{d}\mu = \int_\Omega |f|\,\mathrm{d}\mu.$$

Le b) est quasi immédiat en se souvenant que la mesurabilité de f et g implique celle de $|f|$ et $|g|$ et que ces deux dernières fonctions mesurables positives ont toujours une intégrale élément de $\overline{\mathbb{R}}_+$. L'intégrabilité de f résulte alors de celle de g et de la croissance de l'intégrale dans $\mathcal{M}_+$:

$$\int_\Omega |f|\,\mathrm{d}\mu \leq \int_\Omega |g|\,\mathrm{d}\mu < +\infty.$$

Pour vérifier c), on observe que pour des fonctions à valeurs *réelles* (donc finies en chaque point ω) $f \leq g$ équivaut à $f^+ + g^- \leq g^+ + f^-$. La croissance de l'intégrale dans $\mathcal{M}_+$ nous donne alors

$$\int_\Omega (f^+ + g^-)\,\mathrm{d}\mu \leq \int_\Omega (g^+ + f^-)\,\mathrm{d}\mu$$

Par additivité de l'intégrale et finitude de $\int_\Omega g^-\,\mathrm{d}\mu$ et $\int_\Omega f^-\,\mathrm{d}\mu$, on en déduit

$$\int_\Omega f^+\,\mathrm{d}\mu - \int_\Omega f^-\,\mathrm{d}\mu \leq \int_\Omega g^+\,\mathrm{d}\mu - \int_\Omega g^-\,\mathrm{d}\mu,$$

d'où $\int_\Omega f\,\mathrm{d}\mu \leq \int_\Omega g\,\mathrm{d}\mu$. □

Proposition 4.12 (Intégration par rapport à une mesure à densité). *Soient* $(\Omega, \mathcal{F}, \mu)$ *un espace mesuré et* $f : \Omega \to \overline{\mathbb{R}}_+$*, mesurable* $\mathcal{F}$-$\mathrm{Bor}(\overline{\mathbb{R}}_+)$*. On note* ν *la mesure de densité* f *par rapport à* μ *(cf. théorème 3.17). Soit* $g : \Omega \to \mathbb{K} = \mathbb{R}$ *ou* $\overline{\mathbb{R}}$*, mesurable* $\mathcal{F}$-$\mathrm{Bor}(\mathbb{K})$*. L'application* g *est* ν*-intégrable si et seulement si le produit* gf *est* μ*-intégrable et dans ce cas,*

$$\int_\Omega g\,\mathrm{d}\nu = \int_\Omega gf\,\mathrm{d}\mu. \tag{4.9}$$

On a le même énoncé pour $\mathbb{K} = \mathbb{C}$*, à condition de prendre* f *à valeurs dans* $\mathbb{R}_+$ *au lieu de* $\overline{\mathbb{R}}_+$*.*

Preuve. La condition d'intégrabilité résulte immédiatement du théorème 3.17 appliqué à la fonction mesurable positive $|g|$ en notant que $|g|f = |gf|$. Remarquons au passage que les mesurabilités de gf et de f n'impliquent pas celle de g. L'hypothèse de mesurabilité de g n'est donc pas superflue.

La formule (4.9) se vérifie d'abord pour $\mathbb{K} = \mathbb{R}$, $\overline{\mathbb{R}}$ en appliquant la formule analogue pour les fonctions mesurables positives à g^+ et g^- et en remarquant que grâce à la positivité de f, $(gf)^+ = g^+f$ et $(gf)^- = g^-f$. Le cas complexe se ramène au cas réel en séparant partie réelle et imaginaire de g et en notant que f étant à valeurs *réelles*[3], $\mathrm{Re}(gf) = \mathrm{Re}(g)f$ et $\mathrm{Im}(gf) = \mathrm{Im}(g)f$. □

Théorème 4.13 (de transfert). *Soient* $(\Omega_1, \mathcal{F}_1)$ *et* $(\Omega_2, \mathcal{F}_2)$ *deux espaces mesurables et* $\varphi : \Omega_1 \to \Omega_2$ *une application mesurable* $\mathcal{F}_1$-$\mathcal{F}_2$*. Soit* μ *une mesure sur* $(\Omega_1, \mathcal{F}_1)$ *et* $\nu := \mu \circ \varphi^{-1}$ *sa mesure image sur* $(\Omega_2, \mathcal{F}_2)$*. Soit* h *une application* $h : \Omega_2 \to \mathbb{K}$*,* $\mathbb{K} = \mathbb{R}$*,* $\overline{\mathbb{R}}$ *ou* $\mathbb{C}$*, mesurable* $\mathcal{F}$-$\mathrm{Bor}(\mathbb{K})$*. Alors* h *est* ν*-intégrable si et seulement si* $h \circ \varphi$ *est* μ*-intégrable et dans ce cas,*

$$\int_{\Omega_1} (h \circ \varphi)\,\mathrm{d}\mu = \int_{\Omega_2} h\,\mathrm{d}\nu. \tag{4.10}$$

Démonstration. L'équivalence entre la μ-intégrabilité de $h \circ \varphi$ et la ν-intégrabilité de h est une conséquence directe du théorème de transfert dans $\mathcal{M}_+$ et de la relation $|h \circ \varphi| = |h| \circ \varphi$. Dans le cas $\mathbb{K} = \mathbb{R}$ ou $\overline{\mathbb{R}}$, (4.10) se vérifie en notant que $(h \circ \varphi)^+ = h^+ \circ \varphi$, et

[3]Si on autorisait $f(\omega)$ à valoir $+\infty$, on aurait un problème de définition avec le produit $g(\omega)f(\omega)$, puisque nous n'avons pas donné de sens à $z \times (+\infty)$ pour $z \in \mathbb{C}$ quelconque.

$(h \circ \varphi)^- = h^- \circ \varphi$ et en appliquant le théorème de transfert dans $\mathcal{M}^+$ à h^+ et h^-. Le cas complexe se ramène au cas réel par séparation des parties réelles et imaginaires en notant que $\operatorname{Re}(h \circ \varphi) = (\operatorname{Re} h) \circ \varphi$ et $\operatorname{Im}(h \circ \varphi) = (\operatorname{Im} h) \circ \varphi$. □

Corollaire 4.14 (Calculs d'espérances et de moments). *Soit X une variable aléatoire réelle définie sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$ et P_X sa loi. Alors X est $\mathbf{P}$-intégrable si et seulement si $\int_{\mathbb{R}} |x| \,\mathrm{d}P_X(x) < +\infty$ et dans ce cas son espérance est donnée par*

$$\mathbf{E}X := \int_\Omega X \,\mathrm{d}\mathbf{P} = \int_{\mathbb{R}} x \,\mathrm{d}P_X(x). \tag{4.11}$$

Plus généralement, si h est une application borélienne de $\mathbb{R}$ dans $\mathbb{R}$, la variable aléatoire $Y = h(X)$ est $\mathbf{P}$-intégrable si et seulement si $\int_{\mathbb{R}} |h(x)| \,\mathrm{d}P_X(x) < +\infty$ et dans ce cas

$$\mathbf{E}h(X) := \int_\Omega h(X) \,\mathrm{d}\mathbf{P} = \int_{\mathbb{R}} h(x) \,\mathrm{d}P_X(x) = \int_{\mathbb{R}} y \,\mathrm{d}P_Y(y). \tag{4.12}$$

Preuve. Pour la c.n.s. d'intégrabilité de $Y = h(X)$ et le calcul de son espérance par (4.12), il suffit d'appliquer le théorème de transfert avec $\Omega_1 = \Omega$, $\mathcal{F}_1 = \mathcal{F}$, $\Omega_2 = \mathbb{R}$, $\mathcal{F}_2 = \operatorname{Bor}(\mathbb{R})$ et $\varphi = X$. Le cas où h est l'identité sur $\mathbb{R}$ ($h(x) = x$) donne la condition d'existence et le calcul de $\mathbf{E}X$ par (4.11)... et la dernière égalité dans (4.12) en prenant cette fois $\varphi = Y$. □

Remarque 4.15. Sur le modèle du corollaire 4.14, on peut décliner plusieurs variantes dont la preuve s'obtient par une adaptation immédiate de celle du corollaire. En voici deux.

a) La variable aléatoire complexe Z sur $(\Omega, \mathcal{F}, \mathbf{P})$ est $\mathbf{P}$-intégrable si et seulement si $\int_{\mathbb{C}} |z| \,\mathrm{d}P_Z(z) < +\infty$ et dans ce cas, $\mathbf{E}Z = \int_{\mathbb{C}} z \,\mathrm{d}P_Z(z)$.

b) Soit $X : \Omega \to \mathbb{R}^d$ un vecteur aléatoire et $h : \mathbb{R}^d \to \mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, borélienne. La variable aléatoire réelle ou complexe $Y = h(X)$ est $\mathbf{P}$-intégrable si et seulement si $\int_{\mathbb{R}^d} |h(x)| \,\mathrm{d}P_X(x) < +\infty$ et dans ce cas

$$\mathbf{E}h(X) := \int_\Omega h(X) \,\mathrm{d}\mathbf{P} = \int_{\mathbb{R}^d} h(x) \,\mathrm{d}P_X(x) = \int_{\mathbb{K}} y \,\mathrm{d}P_Y(y).$$

Corollaire 4.16 (Espérance d'une variables aléatoire discrète). *Soit $(\Omega, \mathcal{F}, \mathbf{P})$ un espace probabilisé et X une variable aléatoire discrète sur $(\Omega, \mathcal{F})$, à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Alors X est $\mathbf{P}$-intégrable si et seulement si*

$$\sum_{x \in X(\Omega)} |x| \mathbf{P}(X = x) < +\infty \tag{4.13}$$

et dans ce cas l'espérance de X sous $\mathbf{P}$ peut se calculer par

$$\mathbf{E}X = \sum_{x \in X(\Omega)} x \mathbf{P}(X = x). \tag{4.14}$$

Si h est une application quelconque $\mathbb{K} \to \mathbb{C}$, $Y := h(X)$ est encore une variable aléatoire discrète. Elle est $\mathbf{P}$*-intégrable si et seulement si*

$$\sum_{x \in X(\Omega)} |h(x)| \mathbf{P}(X = x) < +\infty \tag{4.15}$$

et dans ce cas l'espérance de Y sous $\mathbf{P}$ peut se calculer par

$$\mathbf{E}h(X) = \sum_{x \in X(\Omega)} h(x) \mathbf{P}(X = x). \tag{4.16}$$

Preuve. L'ensemble $X(\Omega) = \{x_0, x_1, \dots, x_n, \dots\}$ est ici une partie au plus dénombrable de $\mathbb{R}$ ou $\mathbb{C}$. Selon la proposition 2.27, la loi de X est donnée par

$$P_X = \sum_{x_j \in X(\Omega)} \mathbf{P}(X = x_j) \delta_{x_j}.$$

Il suffit alors d'appliquer le corollaire 4.14, en utilisant les propositions 4.6 et 4.5 pour le calcul des intégrales relatives à P_X. Deux précisions s'imposent ici. Nous n'avons pas besoin de l'hypothèse h borélienne, car X est mesurable $\mathcal{F}$-$\mathcal{P}(\mathbb{K})$, donc pour n'importe quelle tribu $\mathcal{G}$ sur $\mathbb{C}$, h étant automatiquement mesurable $\mathcal{P}(\mathbb{K})$-$\mathcal{G}$, $h \circ X$ est mesurable $\mathcal{F}$-$\mathcal{G}$. Ceci est vrai en particulier pour $\mathcal{G} = \mathrm{Bor}(\mathbb{C})$. En fait on adapte légèrement la preuve du corollaire 4.14 en prenant $\mathcal{F}_2 = \mathcal{P}(\mathbb{K})$ au lieu de $\mathcal{F}_2 = \mathrm{Bor}(\mathbb{K})$ pour la mesurabilité $\mathcal{F}$-$\mathcal{F}_2$ de X dans l'application du théorème de transfert. D'autre part dans l'application de la proposition 4.6, il n'y a pas lieu d'écarter les indices j tels que $a_j := \mathbf{P}(X = x_j) = 0$. En effet, ici $\mu_j = \delta_{x_j}$ et h étant à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$, est automatiquement δ_{x_j}-intégrable (voir note 1 p. 94). □

Remarque 4.17. En DEUG (voir [ICP, 5.1]), on avait utilisé (4.13) et (4.14) pour *définir* l'existence et la valeur de $\mathbf{E}X$. Le remplacement de cette définition initiale de $\mathbf{E}X$ par l'intégrale abstraite $\int_\Omega X \, \mathrm{d}\mathbf{P}$ permet *d'unifier la théorie de l'espérance* et de bénéficier des bonnes propriétés de l'intégrale abstraite. Par exemple si X est une variable aléatoire discrète et Y une variable aléatoire à densité, ayant chacune une espérance, la variable aléatoire $Z = X + Y$ qui n'est en général ni discrète ni à densité, a aussi une espérance et $\mathbf{E}Z = \mathbf{E}X + \mathbf{E}Y$.

Exemple 4.1 (Fonction caractéristique d'un vecteur aléatoire). Soit $X : \Omega \to \mathbb{R}^d$ un vecteur aléatoire de composantes $X_1, \dots, X_d$, sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$. On appelle *fonction caractéristique* de X l'application $\varphi_X : \mathbb{R}^d \to \mathbb{C}$, définie par

$$\forall t = (t_1, \dots, t_d) \in \mathbb{R}^d, \quad \varphi_X(t) := \mathbf{E} \exp(i \langle t, X \rangle) = \mathbf{E} \exp\bigl(i(t_1 X_1 + \dots + t_d X_d)\bigr). \tag{4.17}$$

Cette fonction caractéristique joue un grand rôle en théorie des Probabilités, notamment dans l'étude de la convergence en loi. Justifions l'existence de $\varphi_X(t)$. Pour t fixé, l'application $h := \exp(i \langle t, . \rangle) : \mathbb{R}^d \to \mathbb{C}$ est continue donc borélienne et bornée donc $h(X)$

est une variable aléatoire complexe bornée, donc $\mathbf{P}$-intégrable sur Ω. En notant P_X la loi du vecteur aléatoire X, on a en appliquant la remarque 4.15 b)

$$\varphi_X(t) = \int_{\mathbb{R}^d} \exp(i\langle t, x\rangle) \, \mathrm{d}P_X(x). \tag{4.18}$$

En particulier si X a une densité f par rapport à la mesure de Lebesgue,

$$\varphi_X(t) = \int_{\mathbb{R}^d} \exp(i\langle t, x\rangle) f(x) \, \mathrm{d}\lambda_d(x). \tag{4.19}$$

La poursuite du calcul explicite de φ_X nécessite alors de savoir calculer pratiquement une intégrale par rapport à λ_d, ce que nous étudierons au chapitre 5.

Dans le cas où X est une variable aléatoire discrète, nous disposons déjà de l'outillage nécessaire pour achever le calcul de φ_X. Par exemple si X suit la loi de Poisson de paramètre α, en appliquant le corollaire 4.16, on obtient

$$\forall t \in \mathbb{R}, \quad \varphi_X(t) = \mathbf{E} \exp(itX) = \sum_{k \in \mathbb{N}} \frac{\mathrm{e}^{-\alpha}\alpha^k}{k!} \mathrm{e}^{itk} = \mathrm{e}^{-\alpha} \sum_{k \in \mathbb{N}} \frac{(\alpha \mathrm{e}^{it})^k}{k!} = \exp\bigl(\alpha(\mathrm{e}^{it} - 1)\bigr).$$

4.2 Négligeabilité

Deux fonctions intégrables qui diffèrent seulement sur un ensemble de mesure nulle ont même intégrale. Ceci motive la recherche d'une plus grande souplesse dans la théorie de l'intégration, notamment en remplaçant dans les hypothèses des théorèmes d'interversion limite intégrale, les conditions requises *en tout point* de Ω par leurs analogues *presque partout*. Une autre conséquence importante est que l'application $f \mapsto \int_\Omega |f| \, \mathrm{d}\mu$ n'est qu'une semi-norme sur $\mathcal{L}^1_{\mathbb{K}}(\mu)$. Ceci nous amènera à quotienter $\mathcal{L}^1_{\mathbb{K}}(\mu)$ par l'espace des fonctions mesurables nulles presque partout de façon à en faire un espace vectoriel normé (qui sera de plus complet).

4.2.1 Ensembles négligeables

Une propriété (π) relative à certains éléments de Ω peut toujours être vue comme une application π de Ω dans l'ensemble booléen $\{\text{Faux}, \text{Vrai}\}$ que l'on peut munir de la tribu $\mathcal{P}$ de toutes ses parties. On dit alors que la propriété (π) est $\mathcal{F}$-mesurable si l'application π est $\mathcal{F}$-$\mathcal{P}$ mesurable. Ceci est équivalent à $A := \pi^{-1}(\text{Vrai}) \in \mathcal{F}$.

Définition 4.18. *Soit* $(\Omega, \mathcal{F}, \mu)$ *un espace mesuré.*

a) Un élément A de la tribu $\mathcal{F}$ est une partie μ-négligeable *(ou simplement négligeable s'il n'y a pas d'ambiguité sur la mesure μ) si* $\mu(A) = 0$.

b) Une propriété $\mathcal{F}$-mesurable (π) est dite vraie μ-presque partout si $\mu\bigl(\pi^{-1}(\text{Faux})\bigr) = 0$. *Si $\mu = \mathbf{P}$ est une mesure de probabilité, on parle de propriété vraie $\mathbf{P}$-presque-sûrement. Dans ce cas* $\mathbf{P}\bigl(\pi^{-1}(\text{Faux})\bigr) = 0$ *est équivalent à* $\mathbf{P}\bigl(\pi^{-1}(\text{Vrai})\bigr) = 1$.

c) *Une fonction* $f : \Omega \to \mathbb{K}$, $\mathbb{K} = \mathbb{R}$ *ou* $\mathbb{C}$ *est négligeable si elle est* $\mathcal{F}$-$\mathrm{Bor}(\mathbb{K})$ *mesurable et nulle* μ*-presque partout.*

Remarquons que dans le cas d'une mesure μ telle que $\mu(\Omega) = +\infty$, on peut avoir $\mu\big(\pi^{-1}(\mathrm{Vrai})\big) = \mu(\Omega)$ sans que (π) soit vraie μ-presque partout.

Dans la suite, on abrègera « μ-presque partout » en μ-p.p. ou en p.p. s'il n'y a pas d'ambiguité sur la mesure μ concernée.

Si A est une partie mesurable (i.e. $A \in \mathcal{F}$) d'un ensemble négligeable, A est aussi négligeable. Toute réunion dénombrable d'ensembles négligeables est négligeable.

4.2.2 Finitude presque partout

Un exemple important de propriété vraie presque partout est la finitude des fonctions intégrables.

Proposition 4.19. *Soit* $f : \Omega \to \overline{\mathbb{R}}$ *une application* μ*-intégrable. Alors* f *est* finie μ *presque partout.*

La preuve de ce résultat très utile repose sur le lemme suivant.

Lemme 4.20. *Soit* $g : \Omega \to \overline{\mathbb{R}}_+$ *mesurable positive. Pour tout réel* $t > 0$,

$$\mu(\{g \geq t\}) \leq \frac{1}{t} \int_\Omega g \, \mathrm{d}\mu. \tag{4.20}$$

Dans la théorie des Probabilités, ce lemme est connu sous le nom d'inégalité de Markov. Nous le traduisons dans ce langage pour la commodité de référence ultérieure :

Lemme 4.21 (Inégalité de Markov). *Soit* Y *une variable aléatoire positive (à valeurs dans* $\overline{\mathbb{R}}_+$*) sur l'espace probabilisé* $(\Omega, \mathcal{F}, \mathbf{P})$*. Alors*

$$\forall t > 0, \quad \mathbf{P}(Y \geq t) \leq \frac{\mathbf{E}Y}{t}.$$

Remarque 4.22. L'inégalité de Markov n'a d'intérêt que si $\mathbf{E}Y < +\infty$ et dans ce cas, seulement pour $t > \mathbf{E}Y$. Pour n'importe quelle variable aléatoire positive Y telle que $\mathbf{P}(Y = +\infty) = 0$, $\mathbf{P}(Y > t)$ tend vers 0 quand t tend vers $+\infty$ par continuité séquentielle décroissante de $\mathbf{P}$. L'inégalité de Markov nous dit que si de plus $\mathbf{E}Y$ est fini, cette convergence a lieu (au moins) avec la vitesse $O(t^{-1})$. En fait on peut raffiner et montrer que cette vitesse est au moins $o(t^{-1})$, cf. T.D.

Preuve du lemme 4.20. Comme g est mesurable, $A := \{g \geq t\} \in \mathcal{F}$ et comme g est positive et minorée par la constante t sur A :

$$\int_\Omega g \, \mathrm{d}\mu \geq \int_A g \, \mathrm{d}\mu \geq \int_A t \, \mathrm{d}\mu = t\mu(\{g \geq t\}).$$

On en déduit (4.20) en divisant par t. □

Preuve de la proposition 4.19. Définissons les ensembles A_n et A dans $\mathcal{F}$ par

$$\forall n \in \mathbb{N}^*, \quad A_n := \{|f| \geq n\}, \qquad A := \{|f| = +\infty\}.$$

On a clairement $A \subset A_n$. En appliquant le lemme 4.20 on obtient

$$0 \leq \mu(A) \leq \mu(A_n) \leq \frac{1}{n} \int_\Omega |f| \,\mathrm{d}\mu$$

et comme $\int_\Omega |f| \,\mathrm{d}\mu < +\infty$ (par intégrabilité de f), ce majorant tend vers 0 quand n tend vers $+\infty$. □

Corollaire 4.23. *Soit (f_n) une suite de fonctions mesurables positives (à valeurs dans $\mathbb{R}_+$ ou $\overline{\mathbb{R}}_+$) telle que*

$$\sum_{n=0}^{+\infty} \int_\Omega f_n \,\mathrm{d}\mu < +\infty.$$

Alors $f := \sum_{n=0}^{+\infty} f_n$ est finie μ-presque partout sur Ω.

Démonstration. Il suffit d'appliquer la proposition 4.19, en notant que grâce au théorème d'interversion intégrale série dans $\mathcal{M}_+$ (corollaire du théorème de Beppo Levi),

$$\int_\Omega f \,\mathrm{d}\mu = \sum_{n=0}^{+\infty} \int_\Omega f_n \,\mathrm{d}\mu < +\infty.$$

□

En prenant pour f_n des fonctions indicatrices dans le corollaire 4.23, on obtient un résultat connu dans la théorie des Probabilités sous le nom de *premier lemme de Borel Cantelli.*

Corollaire 4.24 (Borel-Cantelli). *Soit $(A_n)_{n\geq 1}$ une suite d'évènements sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$ et A l'ensemble des ω qui appartiennent à une infinité d'évènements de la suite. On suppose que*

$$\sum_{n=1}^{+\infty} \mathbf{P}(A_n) < +\infty.$$

Alors $\mathbf{P}(A) = 0$.

Démonstration. On pose $f_n = \mathbf{1}_{A_n}$ et on remarque que

$$A = \Big\{\omega \in \Omega;\ \sum_{n=1}^{+\infty} \mathbf{1}_{A_n}(\omega) = +\infty\Big\}.$$

Alors

$$\sum_{n=1}^{+\infty} \int_\Omega f_n \,\mathrm{d}\mathbf{P} = \sum_{n=1}^{+\infty} \mathbf{P}(A_n) < +\infty.$$

Par le corollaire 4.23, $f := \sum_{n=1}^{+\infty} f_n$ est finie $\mathbf{P}$ presque partout, donc $A = \{f = +\infty\}$ est de mesure nulle. □

On peut écrire l'évènement A avec des opérations ensemblistes dénombrables à partir des A_n. En effet ω appartient à A si et seulement si il existe une infinité d'entiers k tels que $\omega \in A_k$. Ceci s'écrit

$$\forall n \in \mathbb{N}, \exists k \geq n;\ \omega \in A_k.$$

Avec la traduction automatique des quantificateurs, ceci s'écrit encore $\omega \in \bigcap_{n\in\mathbb{N}} \bigcup_{k\geq n} A_k$. Ainsi

$$A = \bigcap_{n\in\mathbb{N}} \bigcup_{k\geq n} A_k. \tag{4.21}$$

Cette écriture fait penser à la définition de la limite supérieure d'une suite de réels : $\limsup u_n = \inf_{n\in\mathbb{N}} \sup_{k\geq n} u_k$, à condition de remplacer intersection par inf et union par sup. Or pour la relation d'ordre partiel définie par l'inclusion, l'intersection donne justement la borne inférieure d'une famille d'ensembles et la réunion la borne supérieure. Ceci nous conduit à poser

$$\limsup A_n := \bigcap_{n\in\mathbb{N}} \bigcup_{k\geq n} A_k.$$

De manière analogue, on définit

$$\liminf A_n := \bigcup_{n\in\mathbb{N}} \bigcap_{k\geq n} A_k.$$

En traduisant en sens inverse les opérations ensemblistes avec des quantificateurs, on voit que ω appartient à $\liminf A_n$ si et seulement si $\exists n \in \mathbb{N}, \forall k \geq n,\ \omega \in A_k$. Autrement dit, $\liminf A_n$ est l'ensemble des ω qui appartiennent à *tous* les A_k à partir d'un certain rang $n = n(\omega)$.

En passant au complémentaire dans (4.21), on obtient

$$A^c = \bigcup_{n\in\mathbb{N}} \bigcap_{k\geq n} A_k^c = \liminf A_n^c.$$

En revenant à la conclusion du lemme de Borel Cantelli, $\mathbf{P}(A) = 0$ équivaut à $\mathbf{P}(A^c) = 1$ et cette conclusion peut alors s'interpréter ainsi : « presque sûrement, plus aucun des évènements A_n ne se réalise au-delà d'un certain rang (aléatoire) ». Nous verrons ultérieurement des exemples d'application du lemme de Borel Cantelli.

4.2.3 Égalité presque partout

Définition 4.25. *Soit* $(\Omega, \mathcal{F}, \mu)$ *une espace mesuré et* f, g *deux fonctions* $\Omega \to \mathbb{K}$, $\mathbb{K} = \mathbb{R}$, $\overline{\mathbb{R}}$ *ou* $\mathbb{C}$, $\mathcal{F}$-Bor$(\mathbb{K})$ *mesurables. On dit qu'elles sont égales* μ*-presque partout lorsque* $\{f \neq g\}$ *est négligeable. Notations :*

$$f = g \quad \mu\text{−p.p.} \qquad \textit{ou} \qquad f \overset{\text{p.p.}}{=} g,$$

s'il n'y a pas d'ambiguïté sur la mesure μ *concernée. En particulier,* f *est dite négligeable si elle est égale* μ*-p.p. à zéro.*

Proposition 4.26. *Pour* $\mathbb{K} = \mathbb{R}$ *ou* $\mathbb{C}$*, la relation d'égalité* μ*-p.p. est une relation d'équivalence sur l'ensemble des fonctions mesurables. Elle est compatible avec l'addition et la multiplication par un scalaire.*

Démonstration. La relation est évidemment réflexive ($f \overset{\text{p.p.}}{=} f$), symétrique ($f \overset{\text{p.p.}}{=} g \Rightarrow g \overset{\text{p.p.}}{=} f$). Montrons qu'elle est transitive ($f \overset{\text{p.p.}}{=} g$ et $g \overset{\text{p.p.}}{=} h$ impliquant $f \overset{\text{p.p.}}{=} h$). Les ensembles $A := \{f \neq g\}$ et $B := \{g \neq h\}$ sont négligeables, donc aussi leur réunion. Il suffit alors de remarquer que pour tout $\omega \in (A \cup B)^c$, $f(\omega) = g(\omega) = h(\omega)$.

Si $f_1 \overset{\text{p.p.}}{=} g_1$ et $f_2 \overset{\text{p.p.}}{=} g_2$, $A := \{f_1 \neq g_1\}$ et $B := \{f_2 \neq g_2\}$ sont négligeables, donc aussi leur réunion. Pour tout $\omega \in (A \cup B)^c$, $f_1(\omega) = g_1(\omega)$ et $f_2(\omega) = g_2(\omega)$ donc $(f_1+f_2)(\omega) = (g_1+g_2)(\omega)$ d'où $f_1+f_2 \overset{\text{p.p.}}{=} g_1+g_2$. De même sur A^c, on a $cf_1(\omega) = cg_1(\omega)$ pour toute constante c, donc $cf_1 \overset{\text{p.p.}}{=} cg_1$. □

Proposition 4.27. *Soient* (f_n) *et* (g_n) *sont deux suites d'applications mesurables à valeurs dans* $\mathbb{R}$ *ou* $\overline{\mathbb{R}}$*.*

a) Si $f_1 \overset{\text{p.p.}}{=} g_1$ *et* $f_2 \overset{\text{p.p.}}{=} g_2$*, alors* $\min(f_1, f_2) \overset{\text{p.p.}}{=} \min(g_1, g_2)$ *et* $\max(f_1, f_2) \overset{\text{p.p.}}{=} \max(g_1, g_2)$*.*

b) Si pour tout n*,* $f_n \overset{\text{p.p.}}{=} g_n$*, alors*

$$\inf_{n\in\mathbb{N}} f_n \overset{\text{p.p.}}{=} \inf_{n\in\mathbb{N}} g_n \qquad \sup_{n\in\mathbb{N}} f_n \overset{\text{p.p.}}{=} \sup_{n\in\mathbb{N}} g_n \tag{4.22}$$

$$\liminf_{n\to+\infty} f_n \overset{\text{p.p.}}{=} \liminf_{n\to+\infty} g_n \qquad \limsup_{n\to+\infty} f_n \overset{\text{p.p.}}{=} \limsup_{n\to+\infty} g_n \tag{4.23}$$

c) Si de plus $\lim_{n\to+\infty} f_n \overset{\text{p.p.}}{=} f$*, avec* f *mesurable,* $\lim_{n\to+\infty} g_n \overset{\text{p.p.}}{=} f$*.*

Démonstration. Le a) est un cas particulier de (4.22). Notons pour tout $n \in \mathbb{N}$, $A_n := \{\omega \in \Omega;\ f_n(\omega) \neq g_n(\omega)\}$. Alors chaque A_n est un ensemble négligeable. La réunion dénombrable $A = \bigcup_{n\in\mathbb{N}} A_n$ l'est donc aussi et

$$\forall \omega \in N^c,\ \forall n \in \mathbb{N}, \quad f_n(\omega) = g_n(\omega).$$

Le b) et le c) en résultent immédiatement. □

Proposition 4.28. *Une fonction mesurable* $f : \Omega \to \mathbb{K}$ *(*$\mathbb{K} = \mathbb{R}$*,* $\overline{\mathbb{R}}$ *ou* $\mathbb{C}$*) est négligeable si et seulement si* $\int_\Omega |f| \,\mathrm{d}\mu = 0$*.*

Démonstration. Notons $A := \{\omega \in \Omega;\ |f(\omega)| \neq 0\}$.

Si $\int_\Omega |f| \,\mathrm{d}\mu = 0$, notons pour $n \in \mathbb{N}^*$, $A_n := \{\omega \in \Omega;\ |f(\omega)| \geq 1/n\}$. Par le lemme 4.20

$$\forall n \in \mathbb{N}^*, \quad \mu(A_n) \leq n \int_\Omega |f| \,\mathrm{d}\mu = 0.$$

Comme A est limite croissante (pour l'inclusion) de la suite (A_n), on a par continuité croissante séquentielle de μ, $\mu(A) = \lim_{n\to+\infty} \mu(A_n) = 0$. Donc f est μ-négligeable.

Réciproquement, si f est négligeable, supposons d'abord $|f|$ bornée sur tout Ω par une constante $M < +\infty$. Alors il suffit d'intégrer l'inégalité $|f| \leq M\mathbf{1}_A$ pour obtenir $\int_\Omega |f| \,\mathrm{d}\mu \leq M\mu(A) = 0$. Si $|f|$ n'est pas bornée, on prend $g_n := \min(n; |f|)$ et par le cas $|f|$ bornée, $\int_\Omega g_n \,\mathrm{d}\mu = 0$. On conclut avec le théorème de Beppo Levi puisque $g_n \uparrow |f|$. □

Corollaire 4.29.

i) Soient f et g deux fonctions mesurables $\Omega \to \overline{\mathbb{R}}_+$, égales μ-p.p. Alors $\int_\Omega f \,\mathrm{d}\mu = \int_\Omega g \,\mathrm{d}\mu$.

ii) Si $f \in \mathcal{L}^1_{\mathbb{K}}(\mu)$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$) et si g est une fonction mesurable égale à f p.p. (éventuellement g à valeurs dans $\overline{\mathbb{R}}$ quand $\mathbb{K} = \mathbb{R}$), g est intégrable et $\int_\Omega f \,\mathrm{d}\mu = \int_\Omega g \,\mathrm{d}\mu$.

Démonstration. Dans les égalités

$$f = f\mathbf{1}_{\{f=g\}} + f\mathbf{1}_{\{f\neq g\}} \qquad g = g\mathbf{1}_{\{f=g\}} + g\mathbf{1}_{\{f\neq g\}},$$

les termes $f\mathbf{1}_{\{f\neq g\}}$ et $g\mathbf{1}_{\{f\neq g\}}$ sont des fonctions négligeables mesurables positives. Par la proposition 4.28, leurs intégrales sont nulles. Par additivité de l'intégrale dans $\mathcal{M}_+$ nous avons ainsi :

$$\int_\Omega f \,\mathrm{d}\mu = \int_{\{f=g\}} f \,\mathrm{d}\mu + \int_{\{f\neq g\}} f \,\mathrm{d}\mu = \int_{\{f=g\}} g \,\mathrm{d}\mu = \int_{\{f=g\}} g \,\mathrm{d}\mu + \int_{\{f\neq g\}} g \,\mathrm{d}\mu = \int_\Omega g \,\mathrm{d}\mu.$$

Pour vérifier ii), considérons d'abord le cas $\mathbb{K} = \mathbb{R}$. Par la proposition 4.27 a), $f^+ \overset{\text{p.p.}}{=} g^+$ et $f^- \overset{\text{p.p.}}{=} g^-$. On a donc d'après le i) $\int_\Omega f^+ \,\mathrm{d}\mu = \int_\Omega g^+ \,\mathrm{d}\mu$ et $\int_\Omega f^- \,\mathrm{d}\mu = \int_\Omega g^- \,\mathrm{d}\mu$. L'intégrabilité de f entraîne la finitude de ces 4 intégrales donc l'intégrabilité de g. On a alors l'égalité $\int_\Omega f \,\mathrm{d}\mu = \int_\Omega (g^+ - g^-) \,\mathrm{d}\mu = \int_\Omega g \,\mathrm{d}\mu$. Le cas complexe se ramène au cas $\mathbb{K} = \mathbb{R}$ par séparation des parties réelles et imaginaires. □

4.3 L'espace de Lebesgue $\mathrm{L}^1(\mu)$

Pour $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, notons $\mathcal{N}_{\mathbb{K}}(\mu)$ l'ensemble des applications mesurables négligeables $\Omega \to \mathbb{K}$. Grâce à la proposition 4.26, on voit immédiatement que $\mathcal{N}_{\mathbb{K}}(\mu)$ est un espace vectoriel sur $\mathbb{K}$. C'est un sous-espace vectoriel de $\mathcal{L}^1_{\mathbb{K}}(\mu)$ par la proposition 4.28. La relation d'équivalence qu'est l'égalité μ-p.p. sur $\mathcal{L}^1_{\mathbb{K}}(\mu)$ peut ainsi s'écrire

$$\forall f, g \in \mathcal{L}^1_{\mathbb{K}}(\mu), \qquad f \overset{\text{p.p.}}{=} g \text{ si et seulement si } f - g \in \mathcal{N}_{\mathbb{K}}(\mu).$$

Définition 4.30. *Soit $(\Omega, \mathcal{F}, \mu)$ une espace mesuré et $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. On note $\mathrm{L}^1_{\mathbb{K}}(\mu)$ l'espace vectoriel quotient $\mathcal{L}^1_{\mathbb{K}}(\mu)/\mathcal{N}_{\mathbb{K}}(\mu)$ de l'espace des fonctions μ-intégrables $\Omega \to \mathbb{K}$ par son sous-espace des applications μ-négligeables.*

La notation $\mathrm{L}^1_{\mathbb{K}}(\mu)$ peut être remplacée, lorsque le contexte l'exige ou le permet, par des notations encore plus explicites comme $\mathrm{L}^1_{\mathbb{K}}(\Omega, \mu)$, voire $\mathrm{L}^1_{\mathbb{K}}(\Omega, \mathcal{F}, \mu)$ ou plus brèves comme $\mathrm{L}^1(\mu)$ ou L^1, etc.

Remarque 4.31. Un élément de $\mathrm{L}^1_{\mathbb{K}}(\mu)$ est une classe d'équivalence de fonctions μ-intégrables pour l'égalité μ-presque partout. Si f et g sont deux fonctions membres de la même classe de $\mathrm{L}^1_{\mathbb{K}}(\mu)$, alors $f \overset{\text{p.p.}}{=} g$ et clairement pour tout $A \in \mathcal{F}$, $f\mathbf{1}_A \overset{\text{p.p.}}{=} g\mathbf{1}_A$. On a donc par le corollaire 4.29, $\int_A f \,\mathrm{d}\mu = \int_A g \,\mathrm{d}\mu$. Il en résulte que f et g sont indistinguables pour tout ce qui concerne le calcul d'intégrales (par rapport à μ).

Remarque 4.32. La construction de L^1 exposée ci-dessus exclut les fonctions à valeurs dans $\overline{\mathbb{R}}$. Rappelons que l'on ne peut pas munir l'ensemble des fonctions $\Omega \to \overline{\mathbb{R}}$, μ-intégrables d'une structure de $\mathbb{R}$ espace vectoriel car on ne peut garantir que la somme de deux telles fonctions soit définie sur tout Ω.

Néanmoins, si $g : \Omega \to \overline{\mathbb{R}}$ est μ-intégrable, la fonction $f := g\mathbf{1}_{\{|g|<+\infty\}}$ est égale μ-p.p. à g et est dans $\mathcal{L}^1_{\mathbb{R}}(\mu)$: en effet par la proposition 4.19, $\{f \neq g\}$ est négligeable et par le corollaire 4.29 i), appliqué à $|f|$ et $|g|$, f est μ intégrable. On s'autorisera donc parfois à parler par abus de langage de la classe de g dans $\mathrm{L}^1_{\mathbb{R}}(\mu)$.

Définition 4.33. *Soit φ un élément de $\mathrm{L}^1_{\mathbb{K}}(\mu)$. On définit son intégrale sur $A \in \mathcal{F}$ par*

$$\int_A \varphi \,\mathrm{d}\mu := \int_A f \,\mathrm{d}\mu,$$

où $f \in \mathcal{L}^1_{\mathbb{K}}(\mu)$ est un représentant quelconque de la classe φ.

D'après la remarque 4.31, cette définition est cohérente : la valeur de $\int_A \varphi \,\mathrm{d}\mu$ ne dépend pas du choix du représentant f de φ. Dans le même esprit, on peut grâce à la proposition 4.27, définir φ^+, φ^-, $|\varphi|$.

Proposition 4.34. $\mathrm{L}^1_{\mathbb{K}}(\mu)$ *est un espace vectoriel* normé *par*

$$\|\varphi\|_1 := \int_\Omega |\varphi| \,\mathrm{d}\mu. \tag{4.24}$$

L'espace $\mathrm{L}^1_{\mathbb{K}}(\mu)$ sera toujours supposé muni de cette norme. L'application $\mathrm{L}^1_{\mathbb{K}}(\mu) \to \mathbb{K}$, $\varphi \mapsto \int_\Omega \varphi \,\mathrm{d}\mu$ est une forme linéaire continue *sur $\mathrm{L}^1_{\mathbb{K}}(\mu)$. Sa norme dans l'espace des formes linéaires continues sur $\mathrm{L}^1_{\mathbb{K}}(\mu)$ (le dual topologique de $\mathrm{L}^1_{\mathbb{K}}(\mu)$) est inférieure ou égale à 1, ce qui se traduit par :*

$$\forall \varphi \in \mathrm{L}^1_{\mathbb{K}}(\mu), \quad \left|\int_\Omega \varphi \,\mathrm{d}\mu\right| \leq \|\varphi\|_1. \tag{4.25}$$

Démonstration. Vérifions d'abord que (4.24) définit bien une norme sur $\mathrm{L}^1_{\mathbb{K}}(\mu)$. Soient φ, ψ, quelconques dans $\mathrm{L}^1_{\mathbb{K}}(\mu)$ et $c \in \mathbb{K}$. Notons f et g des représentants respectifs de φ et ψ dans $\mathcal{L}^1_{\mathbb{K}}(\mu)$, alors $f + g$ est un représentant de $\varphi + \psi$ et cf un représentant de $c\varphi$. D'où l'homogénéité :

$$\|c\varphi\|_1 = \int_\Omega |cf| \,\mathrm{d}\mu = |c| \int_\Omega |f| \,\mathrm{d}\mu = |c| \cdot \|\varphi\|_1$$

et la sous-additivité

$$\|\varphi+\psi\|_1 = \int_\Omega |f+g|\,\mathrm{d}\mu \leq \int_\Omega (|f|+|g|)\,\mathrm{d}\mu = \int_\Omega |f|\,\mathrm{d}\mu + \int_\Omega |g|\,\mathrm{d}\mu = \|\varphi\|_1 + \|\psi\|_1.$$

D'autre part le réel $\|\varphi\|_1$ est clairement positif ou nul. Il reste à vérifier qu'il ne s'annule que si φ est égal au zéro de l'espace $\mathrm{L}^1_{\mathbb{K}}(\mu)$, classe d'équivalence de la fonction nulle, autrement dit $\mathcal{N}_{\mathbb{K}}(\mu)$. C'est bien le cas d'après la proposition 4.28.

L'inégalité (4.25) découle de la proposition 4.11 a) via le choix d'un représentant f. Rappelons que si ℓ est une application linéaire d'un espace vectoriel normé E dans un e.v.n. F, elle est continue si et seulement s'il existe une constante C telle que

$$\forall v \in E, \quad \|\ell(v)\|_F \leq C\|v\|_E.$$

La « norme opérateur » de ℓ dans l'espace vectoriel des applications linéaires continues de E dans F est alors définie comme la meilleure valeur possible C_ℓ pour la constante C dans l'inégalité ci-dessus, autrement dit

$$\|\ell\| := C_\ell = \sup_{\substack{v\in E\\ v\neq 0}} \frac{\|\ell(v)\|_F}{\|v\|_E} = \sup_{\|v\|_E=1} \|\ell(v)\|_F = \sup_{\|v\|_E\leq 1} \|\ell(v)\|_F.$$

Revenons à (4.25), en prenant $E = \mathrm{L}^1_{\mathbb{K}}(\mu)$ muni de la norme $\| \ \|_1$ et $F = \mathbb{K}$ muni de la norme $| \ |$, on voit que (4.25) exprime bien la continuité de la forme linéaire $\varphi \mapsto \int_\Omega \varphi\,\mathrm{d}\mu$ et que la norme opérateur de cette forme linéaire est au plus 1. En fait elle vaut exactement 1 dès qu'il existe $A \in \mathcal{F}$ tel que $0 < \mu(A) < +\infty$. En effet $\mathbf{1}_A$ est alors μ-intégrable et sa classe φ dans $\mathrm{L}^1_{\mathbb{K}}(\mu)$ vérifie

$$\Big|\int_\Omega \varphi\,\mathrm{d}\mu\Big| = \Big|\int_\Omega \mathbf{1}_A\,\mathrm{d}\mu\Big| = \mu(A) = \int_\Omega |\mathbf{1}_A|\,\mathrm{d}\mu = \int_\Omega |\varphi|\,\mathrm{d}\mu = \|\varphi\|_1,$$

ce qui montre qu'on ne peut pas choisir de constante C plus petite que 1 dans l'inégalité $|\int_\Omega \psi\,\mathrm{d}\mu| \leq C\|\psi\|_1$ $(\forall \psi \in \mathrm{L}^1_{\mathbb{K}}(\mu))$. □

Nous verrons ultérieurement, comme conséquence du théorème de convergence dominée que l'espace vectoriel normé $\mathrm{L}^1_{\mathbb{K}}(\mu)$ est *complet*. C'est donc un espace de Banach.

4.4 Le théorème de convergence dominée

Nous abordons maintenant le théorème principal d'interversion limite intégrale connu sous le nom de théorème de convergence dominée ou théorème de Lebesgue. L'appellation « convergence dominée » est en soi une bonne description de ce théorème qui dit *grosso modo* que l'interversion $\lim \int f_n\,\mathrm{d}\mu = \int \lim f_n\,\mathrm{d}\mu$ est valide dès que la suite (f_n) vérifie une hypothèse de convergence (μ-p.p. sur Ω) et une hypothèse de *domination* par une fonction μ-intégrable au sens suivant.

Définition 4.35. *La suite* (f_n) *d'applications mesurables* $\Omega \to \mathbb{K}$ *(*$\mathbb{K} = \mathbb{R}$*,* $\overline{\mathbb{R}}$ *ou* $\mathbb{C}$*) est* dominée *sur* Ω *par la fonction mesurable* $g : \Omega \to \overline{\mathbb{R}}_+$ *si*

$$\forall n \in \mathbb{N},\ \forall \omega \in \Omega, \quad |f_n(\omega)| \leq g(\omega).$$

Elle est dite dominée μ*-p.p. si l'inégalité ci-dessus a lieu* μ*-presque partout sur* Ω*.*

Il est clair que dans la définition de la domination μ-p.p., on peut intervertir « $\forall n$ » et « μ-p.p. » puisqu'une réunion dénombrable d'ensembles négligeables est encore négligeable. D'un point de vue théorique, la meilleure fonction dominant une suite (f_n) est simplement la fonction $\sup_{n\in\mathbb{N}} |f_n|$. En pratique, il est rare que l'on sache expliciter cette fonction et étudier directement son intégrabilité, c'est ce qui motive l'introduction d'une fonction g donnant plus de souplesse.

Dans sa version finale, le théorème de Lebesgue est en fait un théorème de convergence dans L^1, d'où l'intérêt de la clause « μ-p.p. » dans ses hypothèses. L'approche en pente douce (au détriment de la concision !) présentée dans les trois lemmes suivants diffère l'utilisation de cette clause. Le lecteur pressé peut sauter directement à l'énoncé du théorème 4.39 et se contenter de la démonstration plus rapide, basée sur le lemme de Fatou, qui sera exposée après la proposition 4.40.

Lemme 4.36 (de convergence décroissante). *Soit* (f_n) *une suite de fonctions mesurables* $\Omega \to [0, +\infty[$ *vérifiant les trois hypothèses*

a) La suite (f_n) *converge simplement sur* Ω *vers* 0 *:* $\forall \omega \in \Omega$*,* $\lim_{n\to+\infty} f_n(\omega) = 0$*.*

b) La suite $(f_n)_{n\geq n_0}$ *est décroissante :* $\forall \omega \in \Omega$*,* $\forall n \geq n_0$*,* $f_n(\omega) \geq f_{n+1}(\omega)$*.*

c) f_{n_0} *est* μ*-intégrable :* $\int_\Omega f_{n_0} \,\mathrm{d}\mu < +\infty$*.*

Alors

$$\lim_{n\to+\infty} \int_\Omega f_n \,\mathrm{d}\mu = 0.$$

Démonstration. Posons pour $n \geq n_0$, $h_n := f_{n_0} - f_n$. Cette fonction est bien définie sur tout Ω (parce que les f_n sont à valeurs dans $[0, +\infty[$ et non dans $\overline{\mathbb{R}}_+$). Elle est mesurable comme différence de deux fonctions mesurables à valeurs dans $\mathbb{R}$. Elle est positive d'après b). La suite $(h_n)_{n\geq n_0}$ converge en croissant vers f_{n_0} d'après a) et b).

Par b) et c), pour $n \geq n_0$, les f_n sont intégrables puisque qu'elles sont positives et

$$\forall n \geq n_0, \quad \int_\Omega f_n \,\mathrm{d}\mu \leq \int_\Omega f_{n_0} \,\mathrm{d}\mu < +\infty.$$

Les h_n sont alors elles aussi intégrables comme différences d'éléments de $\mathcal{L}^1_{\mathbb{R}}(\mu)$. Comme $f_n = f_{n_0} - h_n$, on a par linéarité de l'intégrale dans $\mathcal{L}^1_{\mathbb{R}}(\mu)$,

$$\forall n \geq n_0, \quad \int_\Omega f_n \,\mathrm{d}\mu = \int_\Omega f_{n_0} \,\mathrm{d}\mu - \int_\Omega h_n \,\mathrm{d}\mu. \tag{4.26}$$

Quand n tend vers l'infini, le membre de droite de (4.26) converge vers 0 par le théorème de Beppo Levi appliqué à la suite $(h_n)_{n\geq n_0}$ et parce que $\int_\Omega f_{n_0} \,\mathrm{d}\mu$ a une valeur finie. □

Remarquons que l'hypothèse c) est indispensable. Sur l'espace mesuré $(\mathbb{R}, \mathrm{Bor}(\mathbb{R}), \lambda)$ la suite des fonctions $f_n = \mathbf{1}_{[n,+\infty[}$ vérifie les hypothèses a) et b), mais pas la conclusion. Par contre on peut se passer de l'hypothèse de décroissance b), comme le montre le lemme suivant.

Lemme 4.37. *Soit (f_n) une suite de fonctions mesurables $\Omega \to [0, +\infty[$, convergeant simplement vers zéro sur Ω et dominée sur Ω par une fonction μ-intégrable $g : \Omega \to \overline{\mathbb{R}}_+$. Alors $\int_\Omega f_n \,\mathrm{d}\mu$ converge vers 0.*

Démonstration. Pour tout $\omega \in \Omega$, on a $0 \leq \sup_{n\in\mathbb{N}} f_n \leq g(\omega) \leq +\infty$. Comme g est μ-intégrable, l'ensemble $A := \{g < +\infty\}$ est de complémentaire négligeable (proposition 4.19). Posons alors

$$g_n := \sup_{k\geq n}(f_k \mathbf{1}_A).$$

Comme

$$0 \leq \int_\Omega f_n \,\mathrm{d}\mu = \int_A f_n \,\mathrm{d}\mu \leq \int_A g_n \,\mathrm{d}\mu = \int_\Omega g_n \,\mathrm{d}\mu,$$

il suffit de prouver la convergence vers 0 de $\int_\Omega g_n \,\mathrm{d}\mu$. Pour ce faire, on vérifie que la suite des fonctions mesurables g_n à valeurs dans $[0, +\infty[$ satisfait aux conditions du lemme 4.36. La décroissance de g_n est claire. L'intégrabilité de g_0 résulte immédiatement de la domination de la suite (f_k) et donc *a fortiori* de la suite $(f_k \mathbf{1}_A)$ par la fonction μ-intégrable g puisque

$$\int_\Omega g_0 \,\mathrm{d}\mu = \int_\Omega \sup_{k\geq 0}(f_k \mathbf{1}_A) \,\mathrm{d}\mu \leq \int_\Omega g \,\mathrm{d}\mu < +\infty.$$

D'autre part, la convergence simple sur Ω vers zéro de $(f_k \mathbf{1}_A)$ découle immédiatement de celle de (f_k) et s'écrit :

$$\forall \omega \in \Omega, \quad \forall \varepsilon > 0,\ \exists k_0(\omega),\ \forall k \geq k_0(\omega), \quad f_k(\omega)\mathbf{1}_A(\omega) < \varepsilon.$$

On en déduit que

$$\forall \omega \in \Omega, \quad \forall \varepsilon > 0,\ \exists k_0(\omega),\ g_{k_0(\omega)}(\omega) = \sup_{k\geq k_0(\omega)} f_k(\omega)\mathbf{1}_A(\omega) \leq \varepsilon,$$

puis en raison de la décroissance de (g_n), $g_n(\omega) \leq \varepsilon$ pour tout $n \geq k_0(\omega)$, ce qui prouve la convergence simple sur Ω vers zéro de (g_n). Ainsi (g_n) vérifiant toutes les hypothèses du lemme 4.36, $\int_\Omega g_n \,\mathrm{d}\mu$ converge vers zéro et il en va de même pour $\int_\Omega f_n \,\mathrm{d}\mu$. □

Lemme 4.38. *Soit (f_n) une suite de fonctions mesurables $\Omega \to \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$). On suppose que la suite (f_n) converge simplement sur Ω vers une fonction f à valeurs dans $\mathbb{K}$ et est dominée sur Ω par une fonction μ-intégrable g. Alors f et les f_n sont μ-intégrables et*

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

Démonstration. La domination sur Ω de la suite (f_n) par g s'écrit

$$\forall \omega \in \Omega, \quad |f_n(\omega)| \leq g(\omega),$$

ce qui nous donne déjà la μ-intégrabilité des f_n (cf. proposition 4.11 b). En passant à la limite dans cette inégalité, on en déduit $|f(\omega)| \leq g(\omega)$, d'où la μ-intégrabilité de f.

La suite des fonctions mesurables $h_n := |f_n - f|$ est alors dominée par la fonction μ-intégrable $2g$. Ainsi la suite (h_n) satisfait alors aux conditions du lemme 4.37 donc $\int_\Omega h_n \,\mathrm{d}\mu$ converge vers zéro. □

Théorème 4.39 (de convergence dominée). *Avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, soit (f_n) une suite de fonctions mesurables $\Omega \to \mathbb{K}$ vérifiant les deux hypothèses suivantes.*

*i) La suite (f_n) converge μ-presque partout sur Ω vers une fonction $f : \Omega \to \mathbb{K}$, mesurable $\mathcal{F}$-*Bor$(\mathbb{K})$.

*ii) La suite (f_n) est dominée μ-p.p. par une fonction μ-*intégrable $g : \Omega \to \overline{\mathbb{R}}_+$.

Sous ces conditions,

a) f et les f_n sont μ-intégrables ;

b) (f_n) converge vers f au sens L^1 : $\displaystyle\lim_{n\to+\infty} \int_\Omega |f_n - f| \,\mathrm{d}\mu = 0$;

c) on a l'interversion limite-intégrale : $\displaystyle\lim_{n\to+\infty} \int_\Omega f_n \,\mathrm{d}\mu = \int_\Omega f \,\mathrm{d}\mu.$

Démonstration. Soit A l'ensemble des $\omega \in \Omega$ où $f_n(\omega)$ converge vers $f(\omega)$. D'après l'hypothèse i), A est membre de la tribu $\mathcal{F}$ et $\mu(A^c) = 0$. Pour tout entier n, notons $B_n := \{\omega \in \Omega;\ |f_n(\omega)| \leq g(\omega)\}$. Par l'hypothèse ii), chaque B_n est membre de $\mathcal{F}$ et $\mu(B_n^c) = 0$. Notons B l'intersection de tous les B_n. Comme une réunion dénombrable d'ensembles négligeables est encore négligeable, $\mu(B^c) = 0$. Enfin, notant $E = A \cap B$, on a $\mu(E^c) = 0$ et

$$\forall \omega \in E, \qquad \lim_{n\to+\infty} f_n(\omega) = f(\omega), \tag{4.27}$$

$$\forall \omega \in E,\ \forall n \in \mathbb{N}, \qquad |f_n(\omega)| \leq g(\omega). \tag{4.28}$$

Il est clair que si l'on remplace f_n par $h_n := f_n \mathbf{1}_E$ et f par $h := f\mathbf{1}_E$, les conditions (4.27) et (4.28) deviennent valides *sur tout* Ω. La suite (h_n) vérifie donc les hypothèses du lemme 4.38 et par conséquent h et les h_n sont μ-intégrables et $\int_\Omega |h_n - h| \,\mathrm{d}\mu$ converge vers zéro.

D'autre part, les f_n et f étant mesurables et à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$, la fonction $f_n - f$ est bien *définie* en tout point de Ω et mesurable. Comme $h_n = f_n$ μ-p.p. et $h = f$ μ-p.p., f et les f_n sont μ-intégrables (cf. corollaire 4.29 ii). De plus $|h_n - h| = |f_n - f|$ μ-presque partout et $\int_\Omega |f_n - f| \,\mathrm{d}\mu = \int_\Omega |h_n - h| \,\mathrm{d}\mu$ (cf. corollaire 4.29 i). Par conséquent, $\int_\Omega |f_n - f| \,\mathrm{d}\mu$ converge vers zéro. L'interversion limite intégrale du c) est alors une conséquence immédiate de l'inégalité

$$0 \leq \left| \int_\Omega f_n \,\mathrm{d}\mu - \int_\Omega f \,\mathrm{d}\mu \right| = \left| \int_\Omega (f_n - f) \,\mathrm{d}\mu \right| \leq \int_\Omega |f_n - f| \,\mathrm{d}\mu.$$

□

En pratique, l'hypothèse de convergence i) est un peu plus simple à vérifier qu'il n'y paraît. En effet on a simplement besoin de montrer que f_n converge μ-presque partout, sans se préoccuper des valeurs à attribuer à f sur l'ensemble de divergence, ni de la mesurabilité de f. Cette simplification est une conséquence de la proposition suivante.

Proposition 4.40. *Soit* (f_n) *une suite de fonctions mesurables* $\Omega \to \mathbb{K}$ *qui converge* μ*-p.p., ce qui signifie que* $A := \{\omega \in \Omega;\ f_n(\omega)$ *converge dans* $\mathbb{K}\}$ *est un élément de la tribu* $\mathcal{F}$ *et* $\mu(A^c) = 0$. *Alors*

a) Il existe une fonction f *définie sur* tout Ω *et à valeurs dans* $\mathbb{K}$, mesurable $\mathcal{F}$-$\mathrm{Bor}(\mathbb{K})$, *telle que* (f_n) *converge* μ*-p.p. vers* f. *On peut prendre*

$$f := \lim_{n\to+\infty} \big(f_n \mathbf{1}_A\big). \tag{4.29}$$

b) (f_n) *converge* μ*-p.p. vers* h *mesurable si et seulement si* $h = f$ μ*-p.p.*

Démonstration. Vérifions que la suite de fonctions $(f_n\mathbf{1}_A)$ converge en tout point de Ω. En effet si $\omega \in A$, $f_n(\omega)$ converge vers un élément de $\mathbb{K}$ par définition de A. Si $\omega \in A^c$, $(f_n\mathbf{1}_A)(\omega) = f_n(\omega)\mathbf{1}_A(\omega) = 0$ pour tout n donc $(f_n\mathbf{1}_A)(\omega)$ converge vers 0 quand n tend vers l'infini. La formule (4.29) définit donc bien une fonction f sur *tout* Ω. Pour tout n, $f_n\mathbf{1}_A$ est mesurable comme produit de deux fonctions mesurables et donc f l'est aussi comme limite d'une suite de fonctions mesurables convergeant simplement sur Ω.

D'autre part, (f_n) converge μ-p.p. vers f puisque si $\omega \in A$, $f_n(\omega) \to f(\omega)$ et si $\omega \notin A$, $(f_n(\omega))$ ne converge pas dans $\mathbb{K}$, donc ne peut converger vers $0 = f(\omega)$. Ceci montre que A^c est exactement l'ensemble de tous les points ω tels que $f_n(\omega)$ ne converge pas vers $f(\omega)$. Comme $\mu(A^c) = 0$, on a bien vérifié que (f_n) converge μ-p.p. vers f.

Soit maintenant h mesurable, égale à f μ-presque partout. Notons B l'ensemble des $\omega \in \Omega$ tels que $(f_n(\omega))$ ne converge pas vers $h(\omega)$. Alors B est dans $\mathcal{F}$ (parce que $B = \{\limsup |f_n - h| > 0\}$). Pour $\omega \in A \cap \{h = f\}$, $(f_n(\omega))$ converge vers $h(\omega)$. On a donc l'inclusion

$$B \subset \big(A \cap \{h = f\}\big)^c = A^c \cup \{h \neq f\}$$

qui montre que $\mu(B) = 0$. Ainsi f_n converge vers h μ-presque partout.

Réciproquement, supposons que (f_n) converge vers h μ-presque partout, peut-on en déduire l'égalité μ-p.p. de f et h ? Notons

$$A_1 := \{\omega \in \Omega; f_n(\omega) \to f(\omega)\}, \quad A_2 := \{\omega \in \Omega; f_n(\omega) \to h(\omega)\}.$$

Alors $A_1 \cap A_2 \subset \{f = h\}$ par unicité de la limite de $(f_n(\omega))$ dans $\mathbb{K}$. D'où $\{f \neq h\} \subset A_1^c \cup A_2^c$ d'où $\mu(\{f \neq h\}) = 0$. □

Voici la deuxième preuve du théorème de convergence dominée, annoncée page 109.

Preuve du théorème de Lebesgue par le lemme de Fatou. Les notations et les hypothèses étant celles du théorème 4.39, posons $\Omega' := \{\omega \in \Omega;\ \forall n \geq 1,\ |f_n(\omega)| \leq g(\omega)\}$ et

$$g_n := \big(2g - |f - f_n|\big)\mathbf{1}_{\Omega'}.$$

Nous allons prouver le b) de la conclusion du théorème 4.39, le a) et le c) se vérifiant comme ci-dessus. Notons que l'hypothèse de domination μ-p.p. se traduit par $\mu(\Omega \setminus \Omega') = 0$ et que vu la définition de Ω', elle entraîne aussi la positivité des g_n *sur tout* Ω. Les g_n étant mesurables positives, on peut appliquer le lemme de Fatou à la suite (g_n) :

$$\int_\Omega \liminf_{n\to+\infty} g_n \,\mathrm{d}\mu \leq \liminf_{n\to+\infty} \int_\Omega g_n \,\mathrm{d}\mu. \tag{4.30}$$

L'hypothèse de convergence μ-p.p. de f_n vers f nous donne

$$\liminf_{n\to+\infty} g_n = 2g, \quad \mu\text{-p.p.} \tag{4.31}$$

Notons que les deux membres de cette égalité sont des fonctions définies sur tout Ω et mesurables positives. De (4.31) on déduit

$$\int_\Omega \liminf_{n\to+\infty} g_n \,\mathrm{d}\mu = \int_\Omega 2g \,\mathrm{d}\mu. \tag{4.32}$$

Regardons maintenant le deuxième membre de (4.30). Par l'hypothèse de domination, $|f - f_n|$ est μ-intégrable, on peut donc écrire $\int_\Omega g_n \,\mathrm{d}\mu$ comme différence de deux intégrales[4] d'où :

$$\begin{aligned} \liminf_{n\to+\infty} \int_\Omega g_n \,\mathrm{d}\mu &= \liminf_{n\to+\infty} \left\{ \int_\Omega 2g \,\mathrm{d}\mu - \int_\Omega |f - f_n| \,\mathrm{d}\mu \right\} &(4.33)\\ &= \int_\Omega 2g \,\mathrm{d}\mu - \limsup_{n\to+\infty} \int_\Omega |f - f_n| \,\mathrm{d}\mu. &(4.34)\end{aligned}$$

Pour passer de (4.33) à (4.34), on utilise le fait que si (u_n) est une suite de réels et c une constante (finie), $\liminf_{n\to+\infty}(c - u_n) = c - \limsup_{n\to+\infty} u_n$, dont la vérification est laissée en exercice[5]. En combinant (4.30), (4.32) et (4.34), on obtient

$$\int_\Omega 2g \,\mathrm{d}\mu \leq \int_\Omega 2g \,\mathrm{d}\mu - \limsup_{n\to+\infty} \int_\Omega |f - f_n| \,\mathrm{d}\mu.$$

Comme $\int_\Omega 2g \,\mathrm{d}\mu$ est *finie*, cette inégalité équivaut à

$$\limsup_{n\to+\infty} \int_\Omega |f - f_n| \,\mathrm{d}\mu \leq 0.$$

On en déduit immédiatement la nullité de $\limsup \int_\Omega |f - f_n| \,\mathrm{d}\mu$ puis l'existence et la nullité de $\lim \int_\Omega |f - f_n| \,\mathrm{d}\mu$. Ceci établit le b) de la conclusion du théorème 4.39. □

[4] En toute rigueur, il faudrait d'abord écrire $\int_\Omega g_n \,\mathrm{d}\mu = \int_{\Omega'} g_n \,\mathrm{d}\mu = \dots$, détails laissés au lecteur consciencieux.

[5] Noter au passage qu'ici la positivité de $\liminf(c - u_n)$ et celle de $\limsup u_n$ impliquent la *finitude* de $\limsup u_n$.

Des trois conclusions du théorème de convergence dominée, l'interversion limite intégrale c) est incontestablement la plus populaire, car c'est elle que l'on cherche à obtenir dans la plupart des applications. Néanmoins, il importe de comprendre que la convergence L^1 de f_n vers f est un résultat *nettement plus fort* que la convergence de $\int_\Omega f_n \,\mathrm{d}\mu$ vers $\int_\Omega f \,\mathrm{d}\mu$. Le résultat suivant permet de s'en convaincre.

Proposition 4.41. *Soit* $(\Omega, \mathcal{F}, \mu)$ *un espace mesuré,* (f_n) *une suite dans* $\mathrm{L}^1_{\mathbb{K}}(\mu)$ *et* $f \in \mathrm{L}^1_{\mathbb{K}}(\mu)$*. Alors* (f_n) *converge au sens* L^1 *vers* f *si et seulement si* $\int_A f_n \,\mathrm{d}\mu$ *converge vers* $\int_A f \,\mathrm{d}\mu$*,* uniformément *en* A *sur* $\mathcal{F}$*, c'est-à-dire*

$$\sup_{A\in\mathcal{F}} \left| \int_A f_n \,\mathrm{d}\mu - \int_A f \,\mathrm{d}\mu \right| \xrightarrow[n\to+\infty]{} 0.$$

La proposition 4.41 est une conséquence immédiate du lemme suivant

Lemme 4.42. *La formule*

$$N(\varphi) := \sup_{A\in\mathcal{F}} \left| \int_A \varphi \,\mathrm{d}\mu \right|, \quad \varphi \in \mathrm{L}^1_{\mathbb{K}}(\mu)$$

définit une norme sur $\mathrm{L}^1_{\mathbb{K}}(\mu)$*, équivalente à la norme* $\| \ \|_1$.

Démonstration. Vérifions d'abord que N est une norme. L'homogénéité est évidente. La sous-additivité de N résulte de l'additivité de l'intégrale, de la sous-additivité de la valeur absolue (ou du module) et de la sous-additivité du sup. Soit maintenant $\varphi \in \mathrm{L}^1_{\mathbb{K}}(\mu)$ telle que $N(\varphi) = 0$ et f un représentant de la classe de φ (f est donc une vraie fonction, $f \in \mathcal{L}^1_{\mathbb{K}}(\mu)$). Il s'agit de montrer que f est négligeable. Regardons d'abord le cas $\mathbb{K} = \mathbb{R}$. En prenant successivement $A := \{f > 0\}$ et $A' := \{f < 0\}$, on voit que $\int_A f \,\mathrm{d}\mu = 0$ et $\int_{A'} (-f) \,\mathrm{d}\mu = 0$. Or $\int_A f \,\mathrm{d}\mu = \int_\Omega f^+ \,\mathrm{d}\mu$ et $\int_{A'} (-f) \,\mathrm{d}\mu = \int_\Omega f^- \,\mathrm{d}\mu$. La proposition 4.28 appliquée aux fonctions mesurables positives f^+ et f^- montre alors qu'elles sont négligeables, donc que leur différence f l'est aussi. Le cas $\mathbb{K} = \mathbb{C}$ se ramène au cas réel en observant que les inégalités $|\operatorname{Re} z| \leq |z|$ et $|\operatorname{Im} z| \leq |z|$ impliquent $N(\operatorname{Re} f) \leq N(f)$ et $N(\operatorname{Im} f) \leq N(f)$. Ainsi N est bien une norme.

Pour montrer l'équivalence avec la norme $\| \ \|_1$, on remarque d'abord que

$$\forall \varphi \in \mathrm{L}^1_{\mathbb{K}}(\mu),\ \forall A \in \mathcal{F}, \quad \left| \int_A \varphi \,\mathrm{d}\mu \right| \leq \int_A |\varphi| \,\mathrm{d}\mu \leq \int_\Omega |\varphi| \,\mathrm{d}\mu = \|\varphi\|_1,$$

d'où en passant au sup sur $A \in \mathcal{F}$,

$$\forall \varphi \in \mathrm{L}^1_{\mathbb{K}}(\mu), \quad N(\varphi) \leq \|\varphi\|_1.$$

Dans l'autre sens, pour $\mathbb{K} = \mathbb{R}$, en prenant f représentant de la classe de φ,

$$\|f\|_1 \leq \|f^+\|_1 + \|f^-\|_1 = \int_\Omega f^+ \,\mathrm{d}\mu + \int_\Omega f^- \,\mathrm{d}\mu = \int_{\{f>0\}} f \,\mathrm{d}\mu + \int_{\{f<0\}} (-f) \,\mathrm{d}\mu \leq 2N(f),$$

d'où en revenant aux classes :

$$\forall \varphi \in \mathrm{L}^1_{\mathbb{R}}(\mu), \quad \|\varphi\|_1 \leq 2N(\varphi).$$

Dans le cas complexe, on a en utilisant le cas réel

$$\|f\|_1 \leq \|\operatorname{Re} f\|_1 + \|\operatorname{Im} f\|_1 \leq 2N(\operatorname{Re} f) + 2N(\operatorname{Im} f) \leq 4N(f).$$

En revenant aux classes on obtient :

$$\forall \varphi \in \mathrm{L}^1_{\mathbb{C}}(\mu), \quad \|\varphi\|_1 \leq 4N(\varphi),$$

ce qui achève la preuve de l'équivalence des normes. □

4.5 Comparaison avec l'intégrale de Riemann

Une des premières applications du théorème de convergence dominée est le recyclage de la plupart des intégrales de Riemann courantes en intégrales par rapport à la mesure de Lebesgue. Commençons par un bref rappel sur l'intégrabilité au sens de Riemann.

4.5.1 Intégrabilité au sens de Riemann

Soit $[a, b]$ un intervalle borné de $\mathbb{R}$. On appelle subdivision de $[a, b]$ toute suite finie du type $\Delta = \{x_0 = a < x_1 < \cdots < x_n = b\}$. Pour une fonction *bornée* $f : [a, b] \to \mathbb{R}$, $(-\infty < a < b < +\infty)$, on définit ses sommes de Darboux inférieure $S_\Delta(f)$ et supérieure $S^\Delta(f)$ par

$$S_\Delta(f) := \sum_{k=1}^{n} (x_k - x_{k-1}) \inf_{[x_{k-1}, x_k]} f, \qquad S^\Delta(f) := \sum_{k=1}^{n} (x_k - x_{k-1}) \sup_{[x_{k-1}, x_k]} f.$$

Pour une illustration, voir les figures 4.1 et 4.2.

On dit que la subdivision Δ' est un raffinement de Δ si l'ensemble des valeurs de la suite finie Δ est inclus dans celui des valeurs de la suite Δ', ce que nous noterons avec un léger abus $\Delta \subset \Delta'$. Il est facile de vérifier que

$$\Delta \subset \Delta' \quad \Rightarrow \quad S_\Delta(f) \leq S_{\Delta'}(f) \text{ et } S^\Delta(f) \geq S^{\Delta'}(f).$$

Les figures 4.3 et 4.4 illustrent l'effet de l'adjonction à la subdivision Δ des figures 4.1 et 4.2 de deux nouveaux points.

Les intégrales de Riemann *inférieure* $I_*(f)$ et *supérieure* $I^*(f)$ sont définies par

$$I_*(f) := \sup_\Delta S_\Delta(f), \quad I^*(f) := \inf_\Delta S^\Delta(f),$$

le supremum et l'infimum étant pris sur toutes les subdivisions Δ de $[a, b]$.

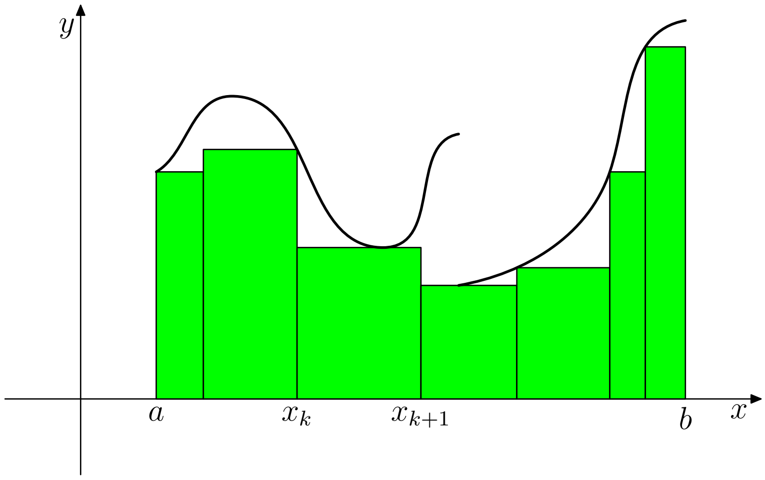

FIG. 4.1 – $S_\Delta(f)$

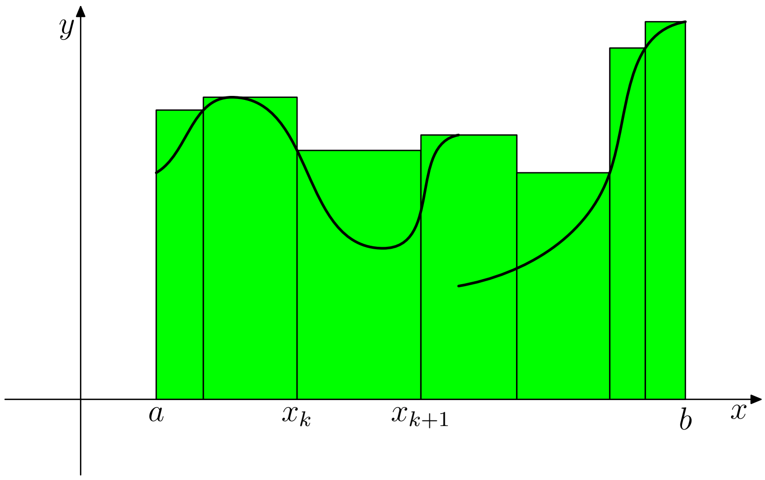

FIG. 4.2 – $S^\Delta(f)$

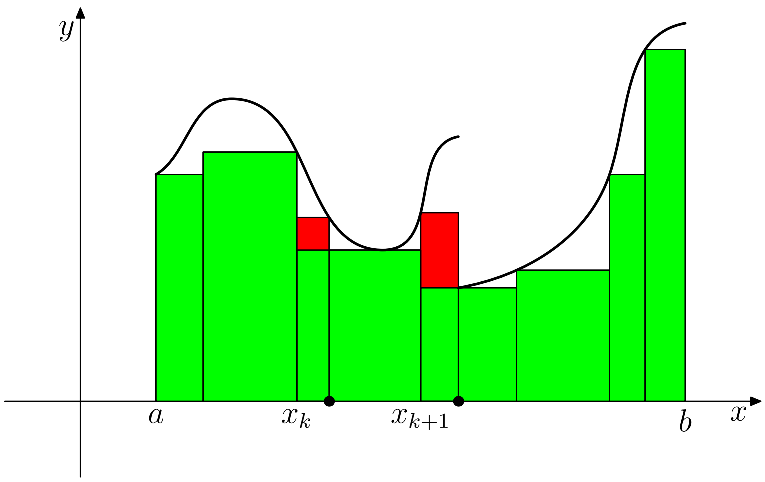

FIG. 4.3 – Si $\Delta \subset \Delta'$, $S_\Delta(f) \leq S_{\Delta'}(f)$

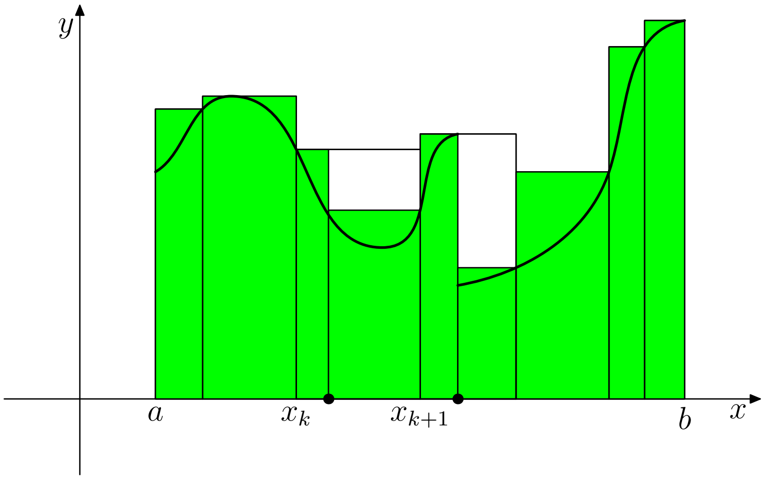

FIG. 4.4 – Si $\Delta \subset \Delta'$, $S^{\Delta}(f) \geq S^{\Delta'}(f)$

Pour Δ_1 et Δ_2 subdivisions de $[a, b]$ on a clairement

$$S_{\Delta_1}(f) \leq S_{\Delta_1 \cup \Delta_2}(f) \leq S^{\Delta_1 \cup \Delta_2}(f) \leq S^{\Delta_2}(f),$$

d'où $S_{\Delta_1}(f) \leq S^{\Delta_2}(f)$. En prenant successivement le sup sur tous les Δ_1, puis l'inf sur tous les Δ_2, on en déduit

$$I_*(f) \leq I^*(f),$$

inégalité vérifiée par toute fonction bornée $f : [a, b] \to \mathbb{R}$.

Définition 4.43. *On dit que f bornée $[a, b] \to \mathbb{R}$ est* Riemann intégrable *si avec les notations ci-dessus, $I_*(f) = I^*(f)$. Dans ce cas on définit son intégrale au sens de Riemann notée $\int_a^b f(x)\,\mathrm{d}x$ par*

$$\int_a^b f(x)\,\mathrm{d}x := I_*(f) = I^*(f).$$

Nous prendrons bien garde dans cette section de distinguer les notations $\int_a^b f(x)\,\mathrm{d}x$ intégrale de Riemann (éventuellement généralisée) et $\int_{[a,b]} f\,\mathrm{d}\lambda$, intégrale de Lebesgue (i.e. au sens de la définition 4.3) par rapport à la mesure de Lebesgue λ. Par la suite, une fois établie leur coïncidence sous des conditions assez générales, on pourra confondre les deux notations, comme il est d'usage courant dans la littérature mathématique.

Nous terminons ces « rappels » sur la Riemann intégrabilité par le cas important des fonctions continues.

Proposition 4.44. *Si $f : [a, b] \to \mathbb{R}$ est continue, elle est Riemann intégrable. De plus si F est une primitive de f sur $[a, b]$,*

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a). \tag{4.35}$$

Démonstration. Sur le *compact* $[a, b]$, la fonction f est bornée et *uniformément* continue :

$$\forall \varepsilon > 0, \quad \exists \delta > 0, \quad \forall x, x' \in [a, b], \quad |x - x'| < \delta \Rightarrow |f(x) - f(x')| < \varepsilon.$$

Pour chaque $\varepsilon > 0$, on peut trouver une subdivision $\Delta = \{x_0 = a < x_1 < \cdots < x_n = b\}$ (dépendant de δ) telle que pour $k = 1, \ldots, n$, $x_k - x_{k-1} < \delta$. Comme les bornes inférieure m_k et supérieure M_k de f sur le compact $[x_{k-1}, x_k]$ sont atteintes, on a pour tout k $0 \leq M_k - m_k < \varepsilon$. On a alors

$$0 \leq S^\Delta(f) - S_\Delta(f) = \sum_{k=1}^n (x_k - x_{k-1})(M_k - m_k) \leq \varepsilon \sum_{k=1}^n (x_k - x_{k-1}) = \varepsilon(b - a).$$

En raison de l'encadrement $S_\Delta(f) \leq I_*(f) \leq I^*(f) \leq S^\Delta(f)$, nous avons ainsi établi que

$$\forall \varepsilon > 0, \quad 0 \leq I^*(f) - I_*(f) \leq \varepsilon(b - a).$$

Comme $I^*(f) - I_*(f)$ ne dépend pas de ε, on en déduit que $I^*(f) - I_*(f) = 0$, d'où la Riemann intégrabilité de f sur $[a, b]$.

Rappelons que F est une primitive de f sur $[a, b]$ si elle est dérivable en tout point de $[a, b]$ (à droite en a et à gauche en b) et pour tout $x \in [a, b]$, $F'(x) = f(x)$. Soit $\Delta = \{x_0 = a < x_1 < \cdots < x_n = b\}$ une subdivision *quelconque* de $[a, b]$. Par le théorème des accroissements finis, il existe dans chaque $]x_{k-1}, x_k[$ un c_k tel que $F(x_k) - F(x_{k-1}) = (x_k - x_{k-1})F'(c_k) = (x_k - x_{k-1})f(c_k)$. En écrivant

$$F(b) - F(a) = \sum_{k=1}^n \big(F(x_k) - F(x_{k-1})\big) = \sum_{k=1}^n (x_k - x_{k-1}) f(c_k)$$

et en encadrant $f(c_k)$ entre les bornes inférieure et supérieure de f sur $[x_{k-1}, x_k]$, on en déduit

$$S_\Delta(f) \leq F(b) - F(a) \leq S^\Delta(f).$$

Cet encadrement est valide pour toute subdivision Δ et $F(b) - F(a)$ ne dépend pas de Δ. Par conséquent

$$I_*(f) \leq F(b) - F(a) \leq I^*(f)$$

et comme nous savons déjà que f est Riemann intégrable on en déduit $F(b) - F(a) = I_*(f) = I^*(f)$, ce qui établit (4.35). □

La preuve de (4.35) s'étend immédiatement au cas où F est dérivable sur $]a, b[$, continue à droite en a et à gauche en b et $F' = f$ sur $]a, b[$. D'autre part on peut utiliser l'intégrale de Riemann pour montrer que toute fonction continue sur $[a, b]$ admet des primitives sur $[a, b]$.

4.5.2 Comparaison

Proposition 4.45. *Soit $[a, b]$ un intervalle borné de $\mathbb{R}$. Si f borélienne bornée $[a, b] \to \mathbb{R}$ est Riemann intégrable, elle est aussi intégrable sur $[a, b]$ par rapport à la mesure de Lebesgue et*

$$\int_a^b f(x)\,\mathrm{d}x = \int_{[a,b]} f\,\mathrm{d}\lambda. \tag{4.36}$$

Démonstration. La λ intégrabilité de f résulte de sa bornitude et du fait que la mesure de Lebesgue sur $[a, b]$ est *finie*, cf. Prop. 4.11 b). Ainsi les deux intégrales dans (4.36) sont bien définies. Pour montrer leur égalité, commençons par associer à chaque subdivision Δ les fonctions *en escaliers*[6]

$$f_\Delta(x) := \sum_{k=1}^n \mathbf{1}_{[x_{k-1},x_k[}(x) \inf_{[x_{k-1},x_k]} f, \qquad f^\Delta(x) := \sum_{k=1}^n \mathbf{1}_{[x_{k-1},x_k[}(x) \sup_{[x_{k-1},x_k]} f.$$

On choisit alors une suite de subdivisions (Δ_n), croissante pour l'inclusion et telle que $S_{\Delta_n}(f) \uparrow I_*(f)$ et $S^{\Delta_n}(f) \downarrow I^*(f)$. Pour vérifier l'existence d'une telle suite (Δ_n), remarquons que la définition de $I_*(f)$ comme borne supérieure et celle de $I^*(f)$ comme borne inférieure nous fournissent deux suites (D_n) et (D'_n) pas forcément monotones de subdivisions telles que $\lim_{n\to+\infty} S_{D_n}(f) = I_*(f)$ et $\lim_{n\to+\infty} S^{D'_n}(f) = I^*(f)$. Il suffit alors de prendre $\Delta_n := \cup_{k\leq n}(D_k \cup D'_k)$. La croissance pour l'inclusion de la suite (Δ_n) implique la croissance de la suite (f_{Δ_n}) et la décroissance de (f^{Δ_n}). On a alors

$$\lim \uparrow f_{\Delta_n} =: g \leq f \leq h := \lim \downarrow f^{\Delta_n}.$$

Les fonctions g et h sont boréliennes comme limites de fonctions boréliennes. Remarquons maintenant que les fonctions en escaliers f_{Δ_n} et f^{Δ_n} étant étagées, on a

$$\int_{[a,b]} f_{\Delta_n}\,\mathrm{d}\lambda = \sum_{k=1}^n \lambda([x_{k-1}, x_k[) \inf_{[x_{k-1},x_k]} f = \sum_{k=1}^n (x_k - x_{k-1}) \inf_{[x_{k-1},x_k]} f = S_{\Delta_n}(f)$$

et de manière analogue

$$\int_{[a,b]} f^{\Delta_n}\,\mathrm{d}\lambda = S^{\Delta_n}(f).$$

Par construction de (Δ_n), $S_{\Delta_n}(f)$ et $S^{\Delta_n}(f)$ convergent respectivement vers $I_*(f)$ et $I^*(f)$, donc vers la même limite $\int_a^b f(x)\,\mathrm{d}x$ puisque f est supposée Riemann intégrable.

D'autre part, les suites (f_{Δ_n}) et (f^{Δ_n}) sont *dominées* par la fonction constante $M := \sup_{[a,b]} |f|$ qui est λ-intégrable puisque $\lambda([a, b]) = b - a < +\infty$. Par le théorème de convergence dominée, on a donc

$$\lim_{n\to+\infty} \int_{[a,b]} f_{\Delta_n}\,\mathrm{d}\lambda = \int_{[a,b]} g\,\mathrm{d}\lambda \quad \text{et} \quad \lim_{n\to+\infty} \int_{[a,b]} f^{\Delta_n}\,\mathrm{d}\lambda = \int_{[a,b]} h\,\mathrm{d}\lambda.$$

[6]Une fonction en escaliers est une combinaison linéaire d'indicatrices d'intervalles disjoints. C'est donc aussi une fonction étagée, la réciproque étant bien entendu fausse.

On voit ainsi que

$$\int_{[a,b]} g \,\mathrm{d}\lambda = I_*(f) = \int_a^b f(x)\,\mathrm{d}x = I^*(f) = \int_{[a,b]} h \,\mathrm{d}\lambda.$$

Enfin f étant λ-intégrable sur $[a,b]$, en intégrant les inégalités $g \le f \le h$, on obtient

$$\int_{[a,b]} g \,\mathrm{d}\lambda \le \int_{[a,b]} f \,\mathrm{d}\lambda \le \int_{[a,b]} h \,\mathrm{d}\lambda.$$

On en déduit que $\int_{[a,b]} f \,\mathrm{d}\lambda = \int_a^b f(x)\,\mathrm{d}x$. □

Remarque 4.46. On ne peut pas s'affranchir de l'hypothèse de Riemann intégrabilité de f dans la proposition 4.45. En effet la fonction indicatrice des rationnels de $[a,b]$ est bornée et borélienne donc λ-intégrable sur $[a,b]$. Par contre elle n'est pas Riemann intégrable car $I_*(f) = 0$ et $I^*(f) = b-a$ (exercice).

Proposition 4.47. *Soit f une fonction $]a,b[\to \mathbb{R}$ ($-\infty \le a < b \le +\infty$). On suppose que*

i) f est borélienne ;

ii) sur tout $[\alpha,\beta] \subset]a,b[$, f est bornée et Riemann intégrable ;

iii) l'intégrale de Riemann généralisée $\int_a^b f(x)\,\mathrm{d}x$ est absolument *convergente.*

Alors f est Lebesgue intégrable sur $]a,b[$ (donc aussi sur $[a,b]$) et

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

Démonstration. L'hypothèse iii) et les inégalités $0 \le f^+ \le |f|$ et $0 \le f^- \le |f|$ impliquent

$$0 \le \int_a^b f^+(x)\,\mathrm{d}x < +\infty, \quad 0 \le \int_a^b f^-(x)\,\mathrm{d}x < +\infty. \tag{4.37}$$

Soient (a_n) une suite décroissante de réels de limite a et (b_n) une suite croissante de limite b, telles que pour tout $n \in \mathbb{N}$, $a < a_n < b_n < b$. Par i), ii) et la proposition 4.45, on a pour tout n, $\int_{[a_n,b_n]} f^+ \,\mathrm{d}\lambda = \int_{a_n}^{b_n} f^+(x)\,\mathrm{d}x$ et $\int_{[a_n,b_n]} f^- \,\mathrm{d}\lambda = \int_{a_n}^{b_n} f^-(x)\,\mathrm{d}x$. De plus les suites de fonctions boréliennes positives $f^+\mathbf{1}_{[a_n,b_n]}$ et $f^-\mathbf{1}_{[a_n,b_n]}$ convergent sur tout $]a,b[$ en croissant vers f^+ et f^- respectivement. Ainsi en combinant le théorème de Beppo Levi et la convergence des intégrales de Riemann généralisées sur $]a,b[$ de f^+ et f^-, on obtient

$$\int_{]a,b[} f^+ \,\mathrm{d}\lambda = \int_a^b f^+(x)\,\mathrm{d}x, \quad \int_{]a,b[} f^- \,\mathrm{d}\lambda = \int_a^b f^-(x)\,\mathrm{d}x. \tag{4.38}$$

De (4.37) et (4.38) on déduit

$$\int_{]a,b[} |f| \,\mathrm{d}\lambda = \int_a^b f^+(x)\,\mathrm{d}x + \int_a^b f^-(x)\,\mathrm{d}x < +\infty,$$

ce qui établit la λ-intégrabilité de f. La définition de l'intégrale de f par rapport à λ et (4.38) nous donnent enfin

$$\int_{]a,b[} f \,\mathrm{d}\lambda = \int_{]a,b[} f^+ \,\mathrm{d}\lambda - \int_{]a,b[} f^- \,\mathrm{d}\lambda = \int_a^b f^+(x)\,\mathrm{d}x - \int_a^b f^-(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x.$$

□

L'hypothèse de convergence *absolue* de l'intégrale généralisée $\int_a^b f(x)\,\mathrm{d}x$ est indispensable, comme le montre le résultat suivant.

Proposition 4.48. *Soit f borélienne sur $]a,b[$ et Riemann intégrable sur tout $[\alpha,\beta] \subset]a,b[$. On suppose de plus que l'intégrale de Riemann généralisée $\int_a^b f(x)\,\mathrm{d}x$* n'est pas *absolument convergente. Alors*

$$\int_{]a,b[} |f| \,\mathrm{d}\lambda = +\infty$$

et donc f n'est pas *Lebesgue intégrable sur $]a,b[$.*

Démonstration. La croissance de l'intégrale sur $\mathcal{M}_+$ et la proposition 4.45 nous donnent la minoration

$$\int_{]a,b[} |f| \,\mathrm{d}\lambda \geq \int_{[\alpha,\beta]} |f| \,\mathrm{d}\lambda = \int_\alpha^\beta |f(x)|\,\mathrm{d}x.$$

Ce minorant tend vers $+\infty$ quand α tend vers a et β vers b en raison de la non convergence absolue de l'intégrale généralisée $\int_a^b f(x)\,\mathrm{d}x$. □

Pour conclure cette sous-section, il n'est pas inutile d'énoncer et de mémoriser le cas particulier suivant de la proposition 4.47 qui devrait couvrir la plupart des intégrales rencontrées en pratique.

Proposition 4.49. *Soient $a, b \in \overline{\mathbb{R}}$, $a < b$. Si $f :]a,b[\to \mathbb{R}$ est* continue *sur $]a,b[$ et d'intégrale de Riemann généralisée $\int_a^b f(x)\,\mathrm{d}x$* absolument *convergente, alors f est Lebesgue-intégrable sur $]a,b[$ et*

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

Preuve. Si $[\alpha,\beta] \subset]a,b[$, $a < \alpha \leq \beta < b$, donc α et β sont finis et l'intervalle $[\alpha,\beta]$ est un *compact* sur lequel f est continue donc bornée et Riemann intégrable. D'autre part f étant continue $]a,b[\to \mathbb{R}$ est clairement borélienne. Ainsi toutes les hypothèses de la proposition 4.47 sont vérifiées, d'où la conclusion. □

4.5.3 Changement de variable dans une intégrale $\int f \,\mathrm{d}\lambda$

La proposition 4.45 permet l'extension suivante aux intégrales par rapport à λ de la formule de changement de variable dans une intégrale de Riemann.

Proposition 4.50. *Soit I un intervalle ouvert (pas nécessairement borné) de $\mathbb{R}$, $\varphi : I \to \mathbb{R}$, strictement monotone et de classe C^1. On pose $J := \varphi(I)$. Alors pour toute application $f : J \to \mathbb{R}$ mesurable positive ou λ-intégrable sur J, on a*

$$\forall A \in \mathrm{Bor}(I), \qquad \int_A f\big(\varphi(x)\big)|\varphi'(x)| \,\mathrm{d}\lambda(x) = \int_{\varphi(A)} f(y) \,\mathrm{d}\lambda(y). \tag{4.39}$$

Si de plus φ' ne s'annule en aucun point de I, on a aussi

$$\forall A \in \mathrm{Bor}(I), \qquad \int_A f\big(\varphi(x)\big) \,\mathrm{d}\lambda(x) = \int_{\varphi(A)} f(y)|(\varphi^{-1})'(y)| \,\mathrm{d}\lambda(y). \tag{4.40}$$

Remarquons que les hypothèses de régularité portent sur le changement de variable φ et non sur f. On ne suppose pas f Riemann intégrable. La portée de cette formule de changement de variable dépasse donc largement celle d'une simple traduction du langage de l'intégrale de Riemann dans celui de l'intégrale de Lebesgue.

Démonstration. La première intégrale dans (4.39) peut s'écrire $\int_A f \circ \varphi \,\mathrm{d}\mu$ où μ est la mesure de densité $|\varphi'|$ par rapport à λ. On prouve (4.39) en appliquant le théorème de transfert entre les espaces mesurés :

$$(I, \mathrm{Bor}(I), \mu) \xrightarrow{\quad\varphi\quad} (J, \mathrm{Bor}(J), \nu := \mu \circ \varphi^{-1}).$$

Commençons par identifier la mesure image ν. Comme φ est continue strictement monotone, J est un intervalle ouvert. Sa tribu borélienne est engendrée par la classe des intervalles bornés $[c, d] \subset J$. De plus φ est une bijection de I sur J et l'inverse ensembliste se confond avec la bijection inverse.
Cas 1 : φ décroissante. Alors $|\varphi'| = -\varphi'$ et

$$\nu([c, d]) = \mu \circ \varphi^{-1}([c, d]) = \mu([\varphi^{-1}(d), \varphi^{-1}(c)]) = \int_{[\varphi^{-1}(d), \varphi^{-1}(c)]} (-\varphi') \,\mathrm{d}\lambda.$$

La fonction φ' est continue donc bornée et Riemann intégrable sur l'intervalle fermé borné $[\varphi^{-1}(d), \varphi^{-1}(c)]$. Par la proposition 4.45, on a

$$\int_{[\varphi^{-1}(d), \varphi^{-1}(c)]} (-\varphi') \,\mathrm{d}\lambda = -\int_{\varphi^{-1}(d)}^{\varphi^{-1}(c)} \varphi'(y) \,\mathrm{d}y = -\Big[\varphi(y)\Big]_{\varphi^{-1}(d)}^{\varphi^{-1}(c)} = d - c = \lambda([c, d]).$$

Ainsi les deux mesures ν et λ σ-finies sur J (qui n'est pas forcément borné) coïncident sur la π-classe des intervalles $[c, d]$ qui engendre $\mathrm{Bor}(J)$. Ceci montre que ν est la mesure de Lebesgue sur J.
Cas 2 : φ croissante. La vérification de l'égalité $\nu = \lambda$ étant analogue (et légèrement plus simple !) sera laissée au lecteur.

À ce stade, le théorème de transfert nous donne (4.39) pour $A = I$. Pour conclure dans le cas général, il suffit d'appliquer le théorème de transfert avec $(f \circ \varphi)\mathbf{1}_A$ en remarquant que :

$$\forall x \in I, \quad f\big(\varphi(x)\big)\mathbf{1}_A(x) = f\big(\varphi(x)\big)\mathbf{1}_{\varphi(A)}\big(\varphi(x)\big),$$

la transformation d'indicatrice se justifiant par l'équivalence de $x \in A$ et $\varphi(x) \in \varphi(A)$. Ainsi (4.39) est établie.

Pour montrer la formule jumelle (4.40), on utilise la même méthode en appliquant le théorème de transfert avec $\mu = \lambda$ et $\nu = \lambda \circ \varphi^{-1}$. Nous nous contentons de décrire l'identification de la mesure ν dans le cas φ décroissante, laissant au lecteur le soin de compléter la preuve. L'hypothèse supplémentaire que φ' ne s'annule en aucun point de I nous garantit que φ^{-1} a une dérivée *continue* en tout point de J :

$$\forall y \in J, \quad (\varphi^{-1})'(y) = \frac{1}{\varphi'\big(\varphi^{-1}(y)\big)}.$$

On peut donc écrire la variation de φ^{-1} entre deux points de J comme l'intégrale de Riemann de $(\varphi^{-1})'$ entre ces deux points. Pour $[c, d] \subset J$, $-\infty < c \leq d < +\infty$, on a ainsi

$$\begin{aligned} \nu([c,d]) = \lambda\big([\varphi^{-1}(d), \varphi^{-1}(c)]\big) = \varphi^{-1}(c) - \varphi^{-1}(d) &= \int_c^d -(\varphi^{-1})'(y)\,\mathrm{d}y \\ &= \int_c^d |(\varphi^{-1})'(y)|\,\mathrm{d}y \\ &= \int_{[c,d]} |(\varphi^{-1})'(y)|\,\mathrm{d}\lambda(y). \end{aligned}$$

Ceci montre que ν est la mesure de densité $|(\varphi^{-1})'|$ par rapport à λ. Notons que nous avons utilisé deux fois la continuité de $(\varphi^{-1})'$ sur $[c, d]$. D'abord pour écrire sa primitive φ^{-1} à l'aide d'une intégrale de Riemann, ensuite pour convertir cette intégrale de Riemann en intégrale de Lebesgue. □

4.5.4 Fonction de répartition et densité

Soit X une variable aléatoire réelle sur $(\Omega, \mathcal{F}, \mathbf{P})$, P_X sa loi et F sa fonction de répartition : $F(x) = \mathbf{P}(X \leq x) = P_X(]-\infty, x])$. Supposons de plus que P_X ait une densité f par rapport à la mesure de Lebesgue. On a alors :

$$\forall x \in \mathbb{R}, \quad F(x) = P_X(]-\infty, x]) = \int_{]-\infty,x]} f\,\mathrm{d}\lambda. \tag{4.41}$$

Il est naturel de se demander quelle relation y a-t-il entre F et f ? La réponse est que F est dérivable λ-presque partout et que sa dérivée (définie sauf sur un ensemble de mesure nulle) est égale λ-p.p. à f. La preuve de ce résultat sortirait du programme de la

Licence. Le lecteur curieux pourra se reporter à l'ouvrage de Kolmogorov, Fomine *Éléments de la Théorie des Fonctions et de l'Analyse Fonctionnelle*, chap. VI, ğ 3. Nous nous contenterons du résultat moins général suivant, qui devrait cependant couvrir la plupart des cas rencontrés en pratique.

Proposition 4.51. *Soit X une variable aléatoire réelle et F sa fonction de répartition. On suppose que F est continue sur $\mathbb{R}$ et de classe C^1 par morceaux. Alors la loi de X a une densité f par rapport à la mesure de Lebesgue et $f = F'$ λ-presque partout.*

Démonstration. Dire que F est continue et C^1 par morceaux sur $\mathbb{R}$ signifie qu'il existe une suite finie $-\infty < a_1 < a_2 < \cdots < a_n < +\infty$ telle que F est dérivable partout sur $\mathbb{R}$, sauf peut-être aux points a_i, que sa dérivée F' est continue sur $\mathbb{R} \setminus \{a_1, \dots, a_n\}$ et que F est aussi continue en chaque point a_i (les « raccords » sont continus). En raison de la croissance de F, F' est positive partout où elle est définie. Notons $a_0 := -\infty$, $a_{n+1} := +\infty$ et $I_i :=]a_i, a_{i+1}[$, $0 \le i \le n$. Comme F a pour limites 0 en $-\infty$ et 1 en $+\infty$, il est commode de la prolonger par continuité à $\overline{\mathbb{R}}$ en posant $F(a_0) := 0$ et $F(a_{n+1}) := 1$.

Si $[c, d]$ est inclus dans I_i, F' est continue sur $[c, d]$ et $\int_c^d F'(t)\,\mathrm{d}t = F(d) - F(c)$. Par continuité de F aux extrémités de I_i, on en déduit que l'intégrale généralisée $\int_{a_i}^{a_{i+1}} F'(t)\,\mathrm{d}t$ converge absolument ($F' \ge 0$ sur I_i) et que pour tout $x \in]a_i, a_{i+1}]$, $F(x) = F(a_i) + \int_{a_i}^x F'(t)\,\mathrm{d}t$. Par la proposition 4.49, on a $F(x) = F(a_i) + \int_{]a_i,x[} F'(t)\,\mathrm{d}\lambda(t)$.

Posons $f := F'$ sur $\mathbb{R} \setminus \{a_1, \dots, a_n\}$ et $f(a_i) := 0$ pour $1 \le i \le n$. Alors f est mesurable positive (exercice) et pour tout $x \in [a_i, a_{i+1}]$, $F(x) = F(a_i) + \int_{]a_i,x[} f\,\mathrm{d}\lambda$. On en déduit en particulier que

$$\int_{\mathbb{R}} |f|\,\mathrm{d}\lambda = \int_{\mathbb{R}\setminus\{a_1,\dots,a_n\}} f\,\mathrm{d}\lambda = \sum_{i=0}^{n} \int_{]a_i,a_{i+1}[} f\,\mathrm{d}\lambda = \sum_{i=0}^{n} \big(F(a_{i+1}) - F(a_i)\big) = 1 < +\infty.$$

Donc la fonction mesurable positive f est λ intégrable sur $\mathbb{R}$. Notons μ la mesure de densité f par rapport à λ. En recollant les morceaux, on voit aussi que

$$\forall x \in \mathbb{R}, \quad F(x) = \int_{]-\infty,x]} f\,\mathrm{d}\lambda = \mu(]-\infty, x]), \tag{4.42}$$

c'est clair si x n'est pas l'un des a_i (noter que l'on peut modifier l'ensemble d'intégration en rajoutant un nombre fini de points, donc un ensemble de λ-mesure nulle, sans changer l'intégrale). Si x est l'un des a_i ($1 \le i \le n$), il suffit de remarquer que les deux membres de (4.42) représentent chacun une fonction de répartition, continue à droite sur $\mathbb{R}$ et de faire tendre x vers a_i par valeurs supérieures. L'égalité (4.42) montre donc que P_X, loi de X et μ ont même fonction de répartition donc que $P_X = \mu$. Ainsi P_X a la densité f par rapport à λ. □

Exemple 4.2. Soient $X_1, \dots, X_n$, n variables aléatoires définies sur le même $(\Omega, \mathcal{F}, \mathbf{P})$, indépendantes et de même loi, de fonction de répartition F, continue sur $\mathbb{R}$ et C^1 par morceaux. On pose

$$M_n := \max_{1 \le i \le n} X_i, \quad H(x) := \mathbf{P}(M_n \le x).$$

Il est facile d'exprimer la fonction de répartition H de M_n à l'aide F. En effet, pour tout $x \in \mathbb{R}$,

$$
\begin{aligned}
H(x) = \mathbf{P}\Big(\max_{1\le i\le n} X_i \le x\Big) &= \mathbf{P}\Big(\forall i=1,\dots,n,\ X_i \le x\Big) \\
&= \mathbf{P}\Big(\bigcap_{i=1}^{n}\{X_i \le x\}\Big) \\
&= \prod_{i=1}^{n} \mathbf{P}(X_i \le x) \qquad (4.43) \\
&= \big(F(x)\big)^n, \qquad (4.44)
\end{aligned}
$$

où (4.43) se justifie (en anticipant légèrement sur le chapitre 5) par le fait que l'indépendance de la suite (X_i) implique l'indépendance mutuelle des évènements $\{X_i \le x\}$ et où (4.44) vient de ce que les X_i ont même loi de f.d.r. F, donc $\mathbf{P}(X_i \le x) = F(x)$.

Par la proposition 4.51, la loi de X_i a une densité f égale λ-presque partout à F'. Les théorèmes classiques sur la continuité et la dérivabilité des fonctions composées montrent que $H = F^n$ est elle aussi continue sur $\mathbb{R}$ et C^1 par morceaux. Il en résulte que la loi de M_n a une densité h par rapport à λ et que

$$h = H' = nF^{n-1}F', \quad \lambda - \text{p.p.}$$

Par exemple si les X_i suivent la loi exponentielle de paramètre $a > 0$,

$$F(x) = \big(1 - \mathrm{e}^{-ax}\big)\mathbf{1}_{[0,+\infty[}(x),$$

F est donc continue sur $\mathbb{R}$ et dérivable partout sauf au point $x = 0$ et M_n a pour densité :

$$h(x) = na\big(1 - \mathrm{e}^{-ax}\big)^{n-1}\mathrm{e}^{-ax}\mathbf{1}_{[0,+\infty[}(x).$$

Remarque 4.52. Attention à ne pas tranformer abusivement la proposition 4.51 en une règle pratique du style « si la fonction de répartition F est dérivable presque partout, alors il y a une densité, égale à F' p.p. ». Cette pseudo règle est doublement fausse. D'abord l'hypothèse de continuité de F *en tout point* de $\mathbb{R}$ est indispensable. Si X suit une loi discrète avec $X(\Omega)$ fini, sa f.d.r. F est dérivable et de dérivée nulle en tout point d'un ensemble de complémentaire fini. Elle est donc bien C^1 par morceaux, mais comme F a des discontinuités (en nombre fini), P_X ne peut avoir de densité par rapport à λ, car s'il y avait une densité f, on aurait :

$$\forall x \in \mathbb{R}, \quad \mathbf{P}(X = x) = \int_{\{x\}} f(t)\,\mathrm{d}\lambda(t) = 0, \quad \text{car } \lambda(\{x\}) = 0.$$

D'autre part, il existe des fonctions de répartition F, continues sur tout $\mathbb{R}$, dérivables presque partout sur $\mathbb{R}$ et de dérivée nulle presque partout sur $\mathbb{R}$. Pour un exemple, voir le problème sur « l'escalier de Cantor » dans les Annales d'IFP 2001–02. Si la loi correspondante avait une densité égale[7] presque partout à F', on aurait $P_X(\mathbb{R}) = \int_{\mathbb{R}} F'\,\mathrm{d}\lambda = 0$, ce qui est impossible puisque $P_X(\mathbb{R}) = 1$.

[7]En fait cette loi n'a *aucune* densité par rapport à λ, il s'agit d'une loi dite *singulière*.

4.6 Applications du théorème de convergence dominée

Le théorème de convergence dominée est un outil efficace pour étudier les propriétés de régularité de fonctions définies par une intégrale de la forme

$$F(x) := \int_\Omega f(x,\omega)\,\mathrm{d}\mu(\omega),$$

où f est définie sur un produit cartésien $E\times\Omega$ et à valeurs dans $\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$. Pour pouvoir ramener les questions de régularité de F à des problèmes de convergences de suites, on se limitera au cas où E est un espace *métrique*. En pratique E sera le plus souvent une partie de $\mathbb{R}$, $\mathbb{R}^d$ ou $\mathbb{C}$. Nous utilisons les notations classiques pour les applications partielles :

$$\begin{aligned} f(x,\,.\,) &: \Omega\to\mathbb{K}, \qquad \omega\mapsto f(x,\,.\,)(\omega) := f(x,\omega),\\ f(\,.\,,\omega) &: E\to\mathbb{K}, \qquad x\mapsto f(\,.\,,\omega)(x) := f(x,\omega). \end{aligned}$$

Les théorèmes que nous avons en vue établissent une régularité de F à partir d'un triplet d'hypothèses :

(Mi) Mesurabilité ou intégrabilité des applications partielles $f(x,\,.\,):\Omega\to\mathbb{K}$.

(Ré) Régularité (pour μ-presque tout ω) des applications partielles $f(\,.\,,\omega):E\to\mathbb{K}$.

(Do) Domination (localement en x) de f (ou de $\frac{\partial^k}{\partial x^k}f$) par une fonction μ-intégrable $g:\Omega\to\overline{\mathbb{R}}_+$.

Théorème 4.53 (de continuité sous le signe $\int$). *Soit (E,d) un espace métrique, $\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$ et f une application*

$$f: E\times\Omega\to\mathbb{K},\quad (x,\omega)\mapsto f(x,\omega).$$

Soit $x_0\in E$. On suppose qu'il existe un voisinage V_0 de x_0 tel que

a) Pour tout $x\in V_0$, l'application $f(x,\,.\,):\Omega\to\mathbb{K}$ est mesurable.

b) Pour μ-presque tout $\omega\in\Omega$, l'application $f(\,.\,,\omega):E\to\mathbb{K}$, est continue au point x_0.

c) Il existe une fonction μ-intégrable g (dépendant en général de V_0) telle que pour tout $x\in V_0$ et pour μ-presque tout $\omega\in\Omega$, $|f(x,\omega)|\le g(\omega)$.

Alors la fonction $F(x) := \int_\Omega f(x,\omega)\,\mathrm{d}\mu(\omega)$ est bien définie sur V_0 et est continue *au point x_0.*

Démonstration. La μ-intégrabilité de $f(x,\,.\,)$ pour tout $x\in V_0$ résulte de sa mesurabilité a) et de sa domination c) par g. La fonction F est donc bien définie au moins sur V_0. Pour la continuité, il s'agit de montrer que $\lim_{x\to x_0}F(x)=F(x_0)$. Comme E est un espace métrique, il suffit pour cela de prouver que pour *toute suite (y_n) convergeant vers* x_0 dans E, $F(y_n)$ converge vers $F(x_0)$ dans $\mathbb{K}$. Soit donc (y_n) une telle suite. Comme V_0 est un voisinage de x_0, tous les y_n sont dans V_0 à partir d'un certain rang et on ne perd pas de généralité en supposant qu'ils y sont tous. Posons alors $h_n := f(y_n, \,.\,)$. Grâce à l'hypothèse b), la suite de fonctions mesurables (h_n) converge μ-p.p. vers $h := f(x_0, \,.\,)$. Elle est dominée μ-p.p. par la fonction intégrable g grâce à c). Le théorème de convergence dominée appliqué à (h_n) donne la conclusion. □

Remarque 4.54. Ce théorème est de nature essentiellement *locale*, on montre la régularité en un point x_0 fixé, en utilisant un voisinage de ce point. Il n'est pas difficile d'en faire un théorème *global* en remplaçant V_0 par E, en supposant dans b) que $f(\,.\,,\omega)$ est continue en tout point x de E et dans c) que l'on a la même fonction dominante g pour tous les x de E. C'est d'ailleurs sous cette forme que le théorème est souvent énoncé. Si nous lui avons préféré une version plus laborieuse, c'est parce qu'elle colle mieux à l'utilisation pratique de ce théorème où il arrive souvent qu'on ne puisse choisir la même fonction dominante g pour tous les points x de E. Voir l'exemple de la transformation de Laplace ci-après.

Exemple 4.3 (Continuité des fonctions caractéristiques). Soit $X : \Omega \to \mathbb{R}^d$ un vecteur aléatoire sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$ et $\varphi_X : \mathbb{R}^d \to \mathbb{C}$, sa *fonction caractéristique* (cf. exemple 4.1, p. 100). Rappelons que

$$\forall t = (t_1, \dots, t_d) \in \mathbb{R}^d, \quad \varphi_X(t) := \mathbf{E}\exp(i\langle t, X\rangle) = \int_\Omega \exp(i\langle t, X\rangle)\,\mathrm{d}\mathbf{P}.$$

Alors φ_X est continue sur $\mathbb{R}^d$. C'est une application immédiate du théorème 4.53 (version globale) en prenant $E = \mathbb{R}^d$, $\mathbb{K} = \mathbb{C}$, $f : \mathbb{R}^d \times \Omega \to \mathbb{C}$, $(t, \omega) \mapsto \exp(i\langle t, X(\omega)\rangle)$, $V_0 = E = \mathbb{R}^d$, $\mu = \mathbf{P}$ et pour g la variable aléatoire constante 1 qui est bien $\mathbf{P}$-intégrable sur Ω.

Exemple 4.4 (Continuité d'une transformée de Laplace). Soit $f : \mathbb{R}_+ \to \mathbb{C}$ une application borélienne. On pose

$$a := \inf\Big\{x \in \mathbb{R};\ \int_{\mathbb{R}_+} \mathrm{e}^{-xu}|f(u)|\,\mathrm{d}\lambda(u) < +\infty\Big\}. \tag{4.45}$$

On suppose $a < +\infty$. Alors

$$F : x \mapsto \int_{\mathbb{R}_+} \mathrm{e}^{-xu} f(u)\,\mathrm{d}\lambda(u)$$

est définie et continue sur $]a, +\infty[$.

Preuve. On commence par noter que l'application $H : \mathbb{R} \to \overline{\mathbb{R}}_+$, $x \mapsto \int_{\mathbb{R}_+} \mathrm{e}^{-xu}|f(u)|\,\mathrm{d}\lambda(u)$ est décroissante. En effet par décroissance pour $u \geq 0$ fixé de $x \mapsto \mathrm{e}^{-xu}$, on a

$$\text{si } x_1 \leq x_2, \quad \forall u \in \mathbb{R}_+, \quad \mathrm{e}^{-x_2 u}|f(u)| \leq \mathrm{e}^{-x_1 u}|f(u)|. \tag{4.46}$$

Par croissance de l'intégrale sur $\mathcal{M}_+$, l'intégration de cette inégalité sur $\mathbb{R}_+$ (relativement à $\mathrm{d}\lambda(u)$) nous donne $H(x_2) \leq H(x_1)$.

Il en résulte que $A := \{x \in \mathbb{R};\ H(x) < +\infty\}$ est un *intervalle* non borné à droite. En effet si $x \in A$, tout $y \geq x$ vérifie $H(y) \leq H(x) < +\infty$, donc y appartient lui aussi à A. Par conséquent $A =]a, +\infty[$ ou $[a, +\infty[$ avec a défini par (4.45).

En tout cas nous disposons au moins de l'inclusion $]a, +\infty[\subset A$. On en déduit que pour $x \in]a, +\infty[$, la fonction borélienne $u \mapsto \mathrm{e}^{-xu} f(u)$ est λ-intégrable sur $\mathbb{R}_+$ et qu'ainsi $F(x)$ est bien défini[8] (à valeur dans $\mathbb{C}$).

Fixons maintenant $x_0 \in]a, +\infty[$. On se donne une marge de sécurité à gauche de x_0 en prenant $a < m_0 < x_0$, par exemple $m_0 := \frac{a+x_0}{2}$ si $a > -\infty$ ou $m_0 := x_0 - 1$ si $a = -\infty$. Alors $m_0 \in A$ et donc $H(m_0) < +\infty$. Par l'inégalité 4.46, on a

$$\forall x \in [m_0, +\infty[, \quad \forall u \in \mathbb{R}_+, \quad \mathrm{e}^{-xu}|f(u)| \leq \mathrm{e}^{-m_0 u}|f(u)|.$$

Comme $H(m_0) < +\infty$, la fonction $g : u \mapsto \mathrm{e}^{-m_0 u}|f(u)|$ est λ-intégrable sur $\mathbb{R}_+$. L'intervalle $V_0 := [m_0, +\infty[$ est un *voisinage* de x_0 et la condition de domination c) du théorème 4.53 est bien vérifiée. Peut-être est-il temps de préciser ici que nous prenons $\Omega = \mathbb{R}_+$, $\omega = u$, $\mu = \lambda$, $E =]a, +\infty[$, $\mathbb{K} = \mathbb{C}$ et que l'application f du théorème 4.53 n'a rien à voir avec le f de cet exemple. Il faudrait la remplacer par $\varphi : (x, u) \mapsto \mathrm{e}^{-xu} f(u)$ et noter que $F(x) = \int_\Omega \varphi(x, \omega)\,\mathrm{d}\lambda(\omega)$. La vérification des conditions a) et b) du théorème 4.53 étant immédiate, nous pouvons conclure que F est continue au point x_0.

Comme la seule hypothèse faite sur x_0 pour établir la continuité de F en ce point était $x_0 \in]a, +\infty[$, nous pouvons dans un deuxième temps en conclure que F est continue *en tout point* de $]a, +\infty[$. □

Théorème 4.55 (de dérivabilité sous le signe $\int$). *Soient V un ouvert de $\mathbb{R}$ et f une application $V \times \Omega \to \mathbb{K}$. On suppose que*

a) Pour tout $x \in V$, l'application $f(x, \,.\,)$ est dans $\mathcal{L}^1_{\mathbb{K}}(\Omega)$.

b) Pour μ-presque tout $\omega \in \Omega$, l'application $f(\,.\,, \omega) : V \to \mathbb{K}$, est dérivable sur V.

c) Il existe une fonction μ-intégrable g telle que pour tout $x \in V$ et pour μ-presque tout $\omega \in \Omega$, $|\frac{\partial}{\partial x} f(x, \omega)| \leq g(\omega)$.

Alors la fonction $F(x) := \int_\Omega f(x, \omega)\,\mathrm{d}\mu(\omega)$ est dérivable sur V et

$$\forall x \in V, \quad F'(x) = \int_\Omega \frac{\partial f}{\partial x}(x, \omega)\,\mathrm{d}\mu(\omega).$$

Nous avons cette fois énoncé le théorème sous une forme *globale*, mais la remarque 4.54 reste valable : en pratique, on sera souvent amené à l'utiliser comme un théorème local avec f définie sur $E \times \Omega$ en prenant pour V un « petit » intervalle ouvert contenant x_0 lorsqu'on ne parvient pas à trouver une fonction dominante g qui fasse l'affaire pour tous les x du « grand » ouvert E.

Démonstration. L'existence de $F(x)$ pour tout $x \in V$ découle de a). Par le même argument que dans la preuve du théorème 4.53, il suffit de montrer que pour une suite (y_n)

[8] Que F soit ou non définie au point $x = a$ (quand $a > -\infty$) dépend du choix de f. À titre d'exercice, examiner les cas $f(u) = u \sin u$ et $f(u) = (1 + u)^{-2}$.

convergente vers x_0 dans V (avec $y_n \neq x_0$ pour tout n),

$$\lim_{n\to+\infty} \frac{F(y_n)-F(x_0)}{y_n-x_0} = \int_\Omega \frac{\partial f}{\partial x}(x_0,\omega)\,\mathrm{d}\mu(\omega).$$

Posons

$$h_n(\omega) := \frac{1}{y_n-x_0}\big(f(y_n,\omega)-f(x_0,\omega)\big).$$

Par l'hypothèse b), il existe $D_1 \in \mathfrak{F}$, tel que $\mu(D_1^c)=0$ et pour tout $\omega\in D_1$, $f(\,.\,,\omega)$ est dérivable sur tout V. Par le théorème des accroissements finis (appliqué séparément à $\operatorname{Re} f(\,.\,,\omega)$ et $\operatorname{Im} f(\,.\,,\omega)$ quand f est à valeurs complexes), on a la majoration

$$\forall\omega\in D_1,\quad |f(y_n,\omega)-f(x_0,\omega)| \leq \sup_{x\in V}\Big|\frac{\partial f}{\partial x}(x,\omega)\Big||y_n-x_0|.$$

Par l'hypothèse c), il existe $D_2\in\mathfrak{F}$, tel que $\mu(D_2^c)=0$ et pour tout $\omega\in D_2$ et tout $x\in V$, $|\frac{\partial}{\partial x}f(x,\omega)|\leq g(\omega)$. L'ensemble $D_1\cap D_2$ est de complémentaire μ-négligeable et

$$\forall\omega\in D_1\cap D_2,\quad \lim_{n\to+\infty} h_n(\omega) = \frac{\partial f}{\partial x}(x_0,\omega) \quad \text{et} \quad \forall n\in\mathbb{N},\ |h_n(\omega)|\leq g(\omega).$$

Le théorème de convergence dominée appliqué à la suite (h_n) nous dit maintenant que sa limite μ-p.p. $\frac{\partial f}{\partial x}(x_0,\,.\,)$ est μ-intégrable et que

$$\lim_{n\to+\infty}\frac{F(y_n)-F(x_0)}{y_n-x_0} = \lim_{n\to+\infty}\int_\Omega h_n(\omega)\,\mathrm{d}\mu(\omega) = \int_\Omega \frac{\partial f}{\partial x}(x_0,\omega)\,\mathrm{d}\mu(\omega).$$

□

Pour des exemples d'application du théorème 4.55, on pourra consulter les Annales d'IFP 2001-02.

Théorème 4.56 (d'holomorphie sous le signe $\int$). *Soit V un ouvert de $\mathbb{C}$ et f : $V\times\Omega\to\mathbb{C}$. On suppose que*

a) Pour tout $z\in V$, l'application $f(z,\,.\,):\Omega\to\mathbb{C}$ est mesurable.

b) Pour μ-presque tout $\omega\in\Omega$, l'application $f(\,.\,,\omega):V\to\mathbb{C}$, est holomorphe sur V.

c) Pour tout $z_0\in V$, il existe un disque fermé *$\Delta\subset V$ de centre z_0 et une fonction μ-intégrable g (dépendant de Δ) telle que pour tout $z\in\Delta$ et pour μ-presque tout $\omega\in\Omega$, $|f(z,\omega)|\leq g(\omega)$.*

Alors la fonction $F(z):=\int_\Omega f(z,\omega)\,\mathrm{d}\mu(\omega)$ est holomorphe sur V et

$$\forall z\in V,\quad F'(z) = \int_\Omega \frac{\partial f}{\partial z}(z,\omega)\,\mathrm{d}\mu(\omega).$$

Démonstration. L'existence de F en tout point de V résulte de a) et de c). Il convient de remarquer la différence d'hypothèse avec le théorème 4.55 : on n'essaie pas de contrôler *a priori* la dérivée $\frac{\partial f}{\partial z}$, on se contente de la domination de f. On va montrer que F est dérivable (en la variable complexe z) au point z_0. Pour cela, notons R le rayon du disque fermé Δ de l'hypothèse c) et $\Delta(z_0, r)$ le disque fermé concentrique de rayon $r = R/2$. La preuve sera ramenée à une simple reproduction de celle du théorème 4.55 une fois que l'on aura justifié les deux affirmations suivantes énoncées pour alléger avec une fonction holomorphe $h : V \to \mathbb{C}$.

$$\forall v \in \Delta(z_0, r), \quad |h'(v)| \leq \frac{1}{r} \sup_{z \in \Delta(z_0, R)} |h(z)|. \tag{4.47}$$

$$\forall z_1 \in \Delta(z_0, r), \quad |h(z_1) - h(z_0)| \leq |z_1 - z_0| \sup_{v \in \Delta(z_0, r)} |h'(v)|. \tag{4.48}$$

Pour vérifier (4.47), on utilise la formule de Cauchy en notant que le disque fermé de centre v et de rayon r est contenu dans $\Delta(z_0, R)$ et que comme ce dernier disque fermé est inclus dans l'ouvert V, il existe une marge de sécurité $\varepsilon > 0$ tel que le disque ouvert de centre z_0 et de rayon $R+\varepsilon$ soit lui même inclus dans V. En notant γ le cercle (orienté positivement) de centre v et de rayon r, on a en paramétrant γ par $z = v + r\mathrm{e}^{it}$, $t \in [0, 2\pi]$,

$$h'(v) = \frac{1}{2\pi i} \int_\gamma \frac{h(z)}{(z-v)^2} \,\mathrm{d}z = \frac{1}{2\pi r} \int_0^{2\pi} h(v + r\mathrm{e}^{it})\mathrm{e}^{-it} \,\mathrm{d}t,$$

ce qui donne immédiatement (4.47).

Pour établir (4.48), paramétrons le segment radial $[z_0, z_1]$ par $z = z_0 + t(z_1 - z_0)$, $t \in [0, 1]$. On a alors

$$h(z_1) - h(z_0) = \int_{[z_0, z_1]} h'(z) \,\mathrm{d}z = \int_0^1 h'\big(z_0 + t(z_1 - z_0)\big)(z_1 - z_0) \,\mathrm{d}t,$$

d'où l'on déduit facilement (4.48). □