Chapitre 5

Intégration sur des espaces produits

5.1 Produit de deux mesures

Étant donnés deux espaces mesurés $(\Omega_1, \mathcal{F}_1, \mu_1)$ et $(\Omega_2, \mathcal{F}_1, \mu_2)$, le but de cette section est de construire une mesure μ sur $\Omega_1 \times \Omega_2$ telle que $\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$ pour tous $A_1 \in \mathcal{F}_1$ et $A_2 \in \mathcal{F}_2$. La première chose à faire est de munir $\Omega_1 \times \Omega_2$ d'une tribu adéquate.

5.1.1 Produit de deux tribus

Définition 5.1 (Tribu produit). *Soient $(\Omega_1, \mathcal{F}_1)$ et $(\Omega_2, \mathcal{F}_2)$ deux espaces mesurables. On appelle tribu produit de $\mathcal{F}_1$ par $\mathcal{F}_2$ et on note $\mathcal{F}_1 \otimes \mathcal{F}_2$, la tribu sur $\Omega_1 \times \Omega_2$ engendrée par la famille $\mathcal{R}$ des « rectangles mesurables » $A_1 \times A_2$, où $A_1 \in \mathcal{F}_1$ et $A_2 \in \mathcal{F}_2$:*

$$\mathcal{F}_1 \otimes \mathcal{F}_2 := \sigma\bigl(\{A_1 \times A_2;\ A_1 \in \mathcal{F}_1,\ A_2 \in \mathcal{F}_2\}\bigr). \tag{5.1}$$

Remarques 5.2. Il n'est sans doute pas superflu de noter d'emblée que :

a) La famille des rectangles mesurables[1] n'est pas elle même une tribu en général. Elle n'est stable ni par réunion finie ni par complémentaire.

b) Le produit de deux tribus n'est pas commutatif. La tribu $\mathcal{F}_2 \otimes \mathcal{F}_1$ n'est pas en général égale à $\mathcal{F}_1 \otimes \mathcal{F}_2$, tout simplement parce que le produit cartésien n'est pas commutatif : $\Omega_1 \times \Omega_2 \neq \Omega_2 \times \Omega_1$. Par conséquent le produit de deux mesures $\mu_1 \otimes \mu_2$, que nous allons définir sur $\mathcal{F}_1 \otimes \mathcal{F}_2$ ne sera pas non plus commutatif.

Proposition 5.3. *Les deux projections canoniques*

$$\pi_i : \Omega_1 \times \Omega_2, \quad (\omega_1, \omega_2) \mapsto \omega_i \quad (i = 1, 2)$$

sont mesurables $(\mathcal{F}_1 \otimes \mathcal{F}_2)$ - $\mathcal{F}_i$. La tribu produit $\mathcal{F}_1 \otimes \mathcal{F}_2$ est la plus petite tribu rendant mesurables les projections canoniques π_i.

[1] On l'aura compris, le mot rectangle n'est pas à prendre ici au pied de la lettre.

Preuve. En effet pour tout $A_1 \in \mathcal{F}_1$, $\pi_1^{-1}(A_1) = A_1 \times \Omega_2 \in \mathcal{F}_1 \otimes \mathcal{F}_2$, ce qui établit la mesurabilité de π_1. La vérification pour π_2 est analogue. Soit maintenant $\mathcal{F}$ une tribu sur $\Omega_1 \times \Omega_2$ telle que les projections canoniques π_i soient $\mathcal{F}$ - $\mathcal{F}_i$ mesurables ($i = 1, 2$). Nécessairement $\mathcal{F}$ contient les rectangles $\pi_1^{-1}(A_1) = A_1 \times \Omega_2$ pour tout $A_1 \in \mathcal{F}_1$ et $\pi_2^{-1}(A_2) = \Omega_1 \times A_2$ pour tout $A_2 \in \mathcal{F}_2$. La tribu $\mathcal{F}$ étant stable par intersection doit alors contenir tous les

$$(A_1 \times \Omega_2) \cap (\Omega_1 \times A_2) = A_1 \times A_2.$$

Elle contient donc tous les rectangles mesurables, donc aussi la tribu qu'ils engendrent, c'est à dire $\mathcal{F}_1 \otimes \mathcal{F}_2$. On voit ainsi que la tribu produit $\mathcal{F}_1 \otimes \mathcal{F}_2$ est la plus petite tribu rendant mesurables les projections canoniques π_i. □

On introduit maintenant la notion de *section* d'une partie d'un produit cartésien. La définition et le lemme qui suit sont purement ensemblistes et ne font pas intervenir les propriétés des tribus.

Définition 5.4. *Soit $E \subset \Omega_1 \times \Omega_2$ et $\omega_1 \in \Omega_1$ fixé. On appelle section de E en ω_1, le sous-ensemble E_{ω_1} de Ω_2 défini par*

$$E_{\omega_1} := \{\omega_2 \in \Omega_2;\ (\omega_1, \omega_2) \in E\}.$$

De même pour tout $\omega_2 \in \Omega_2$, on définit la section de E en ω_2 par

$$E_{\omega_2} := \{\omega_1 \in \Omega_1;\ (\omega_1, \omega_2) \in E\}.$$

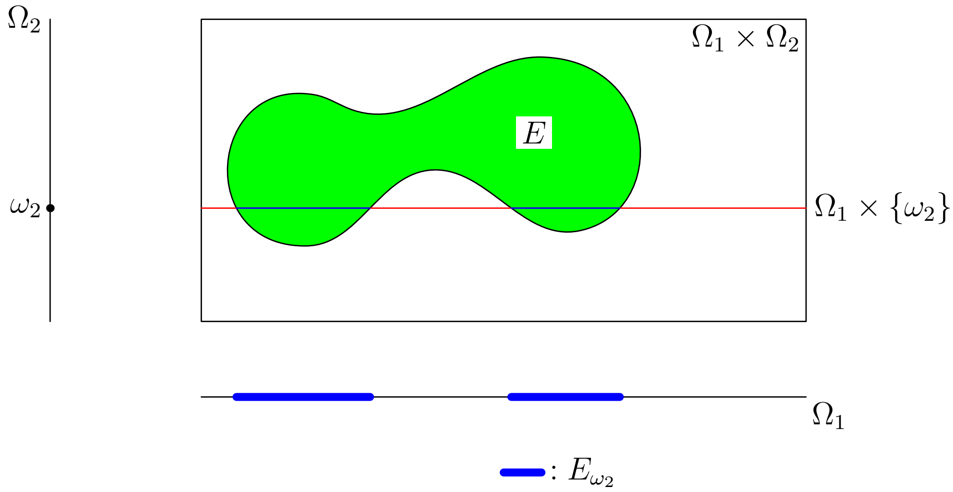

FIG. 5.1 – Section E_{ω_2} de E en ω_2

Lemme 5.5. *La section commute avec le complémentaire, la différence propre, l'union et l'intersection. Plus précisément, si $(E_i)_{i \in I}$ est une famille quelconque (pas forcément dénombrable) de parties de $\Omega_1 \times \Omega_2$, on a pour tout $\omega_1 \in \Omega_1$,* *a)* $\big((\Omega_1 \times \Omega_2) \setminus E\big)_{\omega_1} = \Omega_2 \setminus E_{\omega_1}$ *et si* $E \subset F \subset \Omega_1 \times \Omega_2$, $(F \setminus E)_{\omega_1} = F_{\omega_1} \setminus E_{\omega_1}$ *;*

b) $\Big(\bigcup\limits_{i\in I} E_i\Big)_{\omega_1} = \bigcup\limits_{i\in I} (E_i)_{\omega_1}$ *;*

c) $\Big(\bigcap\limits_{i\in I} E_i\Big)_{\omega_1} = \bigcap\limits_{i\in I} (E_i)_{\omega_1}$

et les propriétés analogues pour les sections en $\omega_2 \in \Omega_2$.

Démonstration. La vérification de a) s'écrit

$$\begin{aligned} \Omega_2 \setminus E_{\omega_1} &= \Omega_2 \setminus \{\omega_2 \in \Omega_2;\ (\omega_1, \omega_2) \in E\} \\ &= \{\omega_2 \in \Omega_2;\ (\omega_1, \omega_2) \notin E\} \\ &= \{\omega_2 \in \Omega_2;\ (\omega_1, \omega_2) \in \Omega_1 \times \Omega_2 \setminus E\} \\ &= (\Omega_1 \times \Omega_2 \setminus E)_{\omega_1}. \end{aligned}$$

Celle de b) est tout aussi facile :

$$\begin{aligned} \Big(\bigcup_{i\in I} E_i\Big)_{\omega_1} &= \{\omega_2 \in \Omega_2;\ (\omega_1, \omega_2) \in \bigcup_{i\in I} E_i\} \\ &= \{\omega_2 \in \Omega_2;\ \exists i \in I,\ (\omega_1, \omega_2) \in E_i\} \\ &= \bigcup_{i\in I} \{\omega_2 \in \Omega_2; (\omega_1, \omega_2) \in E_i\} \\ &= \bigcup_{i\in I} (E_i)_{\omega_1}. \end{aligned}$$

Enfin, c) s'obtient soit directement, soit par combinaison de a) et b). □

Proposition 5.6. *Toutes les sections d'un ensemble* $E \in \mathcal{F}_1 \otimes \mathcal{F}_2$ *sont mesurables, au sens suivant :*

$$\forall \omega_1 \in \Omega_1,\ E_{\omega_1} \in \mathcal{F}_2, \qquad \forall \omega_2 \in \Omega_2,\ E_{\omega_2} \in \mathcal{F}_1.$$

Démonstration. Il suffit bien sûr, de faire la preuve pour les sections en ω_1. Fixons donc $\omega_1 \in \Omega_1$ et définissons

$$\mathcal{G}_1 := \{E \subset \Omega_1 \times \Omega_2;\ E_{\omega_1} \in \mathcal{F}_2\}.$$

Si on montre que $\mathcal{G}_1$ est une tribu sur $\Omega_1 \times \Omega_2$ et qu'elle contient la famille des rectangles mesurables $\mathcal{R} = \{A_1 \times A_2\,;\ A_1 \in \mathcal{F}_1,\ A_2 \in \mathcal{F}_2\}$, alors elle contiendra la tribu $\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma(\mathcal{R})$ et on aura établi que pour tout $E \in \mathcal{F}_1 \otimes \mathcal{F}_2$, $E \in \mathcal{G}_1$, donc $E_{\omega_1} \in \mathcal{F}_2$. Ceci étant vrai pour tout ω_1, la proposition sera prouvée.

Pour voir que $\mathcal{G}_1$ contient tous les rectangles mesurables $A_1 \times A_2$, il suffit d'écrire

$$\begin{aligned} (A_1 \times A_2)_{\omega_1} &= \{\omega_2 \in \Omega_2;\ (\omega_1, \omega_2) \in A_1 \times A_2\} \\ &= \{\omega_2 \in \Omega_2;\ \omega_1 \in A_1 \text{ et } \omega_2 \in A_2\} \\ &= \begin{cases} \emptyset & \text{si } \omega_1 \notin A_1, \\ A_2 & \text{si } \omega_1 \in A_1. \end{cases} \end{aligned} \tag{5.2}$$

Dans les deux cas, on trouve un élément de $\mathcal{F}_2$.

Vérifions que $\mathcal{G}_1$ est une tribu. La section en ω_1 de l'ensemble vide (considéré comme sous-ensemble de $\Omega_1 \times \Omega_2$) est bien évidemment l'ensemble vide (considéré comme sous-ensemble de Ω_2) donc $\emptyset \in \mathcal{G}_1$. La stabilité de $\mathcal{G}_1$ par complémentaire résulte de celle de $\mathcal{F}_2$, via le lemme 5.5 a). La stabilité de $\mathcal{G}_1$ par réunion dénombrable résulte de même de celle de $\mathcal{F}_2$, via le lemme 5.5 b). □

Voyons maintenant la question de la mesurabilité des applications partielles.

Lemme 5.7. *Soit* $(\Omega, \mathcal{F})$ *un espace mesurable quelconque et* $f : \Omega_1 \times \Omega_2 \to \Omega$ *une application* $\mathcal{F}_1 \otimes \mathcal{F}_2$ - $\mathcal{F}$ *mesurable. Alors pour tout* $\omega_1 \in \Omega_1$*, l'application partielle*

$$f(\omega_1, \,.\,) : \Omega_2 \to \Omega, \quad \omega_2 \mapsto f(\omega_1, \omega_2),$$

est $\mathcal{F}_2$ - $\mathcal{F}$ *mesurable. De même* $f(\,.\,, \omega_2)$ *est* $\mathcal{F}_1$ - $\mathcal{F}$ *mesurable pour tout* $\omega_2 \in \Omega_2$.

Démonstration. Par symétrie, il suffit de traiter le cas de $f(\omega_1, \,.\,)$. Fixons donc ω_1 quelconque dans Ω_1 et prenons un élément quelconque B de la tribu $\mathcal{F}$.

$$\begin{aligned} f(\omega_1, \,.\,)^{-1}(B) &= \{\omega_2 \in \Omega_2;\ f(\omega_1, \,.\,)(\omega_2) \in B\} \\ &= \{\omega_2 \in \Omega_2;\ f(\omega_1, \omega_2) \in B\} \\ &= \{\omega_2 \in \Omega_2;\ (\omega_1, \omega_2) \in f^{-1}(B)\} = \big(f^{-1}(B)\big)_{\omega_1}. \end{aligned}$$

Par la $\mathcal{F}_1 \otimes \mathcal{F}_2$ - $\mathcal{F}$ mesurabilité de f, $f^{-1}(B)$ est dans la tribu $\mathcal{F}_1 \otimes \mathcal{F}_2$. La proposition 5.6 nous donne alors l'appartenance de sa section en ω_1 à la tribu $\mathcal{F}_2$. La $\mathcal{F}_2$ - $\mathcal{F}$ mesurabilité de $f(\omega_1, \,.\,)$ est ainsi établie. □

5.1.2 Construction de la mesure produit

Théorème 5.8 (mesure produit). *Pour* $i = 1, 2$*, on suppose que* μ_i *est une mesure* σ*-finie sur* $(\Omega_i, \mathcal{F}_i)$*. Alors il existe une unique mesure* μ *sur* $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$ *telle que*

$$\forall A_1 \in \mathcal{F}_1,\ \forall A_2 \in \mathcal{F}_2, \quad \mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2). \tag{5.3}$$

On la note $\mu = \mu_1 \otimes \mu_2$ *et on l'appelle* mesure produit *de* μ_1 *par* μ_2*. Elle vérifie*

$$\forall E \in \mathcal{F}_1 \otimes \mathcal{F}_2, \quad \mu_1 \otimes \mu_2(E) = \int_{\Omega_1} \mu_2(E_{\omega_1})\, \mathrm{d}\mu_1(\omega_1) = \int_{\Omega_2} \mu_1(E_{\omega_2})\, \mathrm{d}\mu_2(\omega_2). \tag{5.4}$$

Preuve de l'unicité de μ. Réglons d'abord cette question qui est la plus facile, en supposant que μ et ν sont deux mesures sur $\mathcal{F}_1 \otimes \mathcal{F}_2$ vérifiant (5.3). Elles coïncident donc sur la classe $\mathcal{R}$ des rectangles mesurables :

$$\forall A_1 \in \mathcal{F}_1,\ \forall A_2 \in \mathcal{F}_2, \quad \mu(A_1 \times A_2) = \nu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2). \tag{5.5}$$

On remarque alors que $\mathcal{R}$ est un π-système (*i.e.* stable par intersections finies). Ceci résulte de l'égalité

$$(A_1 \times A_2) \cap (B_1 \times B_2) = (A_1 \cap B_1) \times (A_2 \cap B_2),$$

vraie pour tous $A_1, B_1 \subset \Omega_1$ et $A_2, B_2 \subset \Omega_2$. Comme μ_i ($i = 1, 2$) est σ-finie, il existe une suite $(\Omega_{i,n})_{n\geq 1}$ dans $\mathcal{F}_i$, croissante pour l'inclusion et telle que

$$\forall n \geq 1,\ \mu_i(\Omega_{i,n}) < +\infty \quad \text{et} \quad \Omega_{i,n} \uparrow \Omega_i, \quad (n \to +\infty), \quad i = 1, 2.$$

Il est facile d'en déduire que $\Omega_{1,n} \times \Omega_{2,n} \uparrow \Omega_1 \times \Omega_2$ et que

$$\forall n \geq 1,\ \mu(\Omega_{1,n} \times \Omega_{2,n}) = \nu(\Omega_{1,n} \times \Omega_{2,n}) = \mu_1(\Omega_{1,n})\mu_2(\Omega_{2,n}) < +\infty.$$

On peut alors conclure par le théorème d'unicité des mesures (th. 1.34) que μ et ν coïncident sur $\sigma(\mathcal{R}) = \mathcal{F}_1 \otimes \mathcal{F}_2$. □

Preuve de l'existence de μ.

On va définir μ par la formule

$$\forall E \in \mathcal{F}_1 \otimes \mathcal{F}_2, \quad \mu(E) := \int_{\Omega_1} \mu_2(E_{\omega_1})\, \mathrm{d}\mu_1(\omega_1). \tag{5.6}$$

Pour motiver cette formule, regardons ce qu'elle donne dans le cas particulier où $E = A_1 \times A_2$ est un rectangle mesurable. Alors grâce à (5.2), on peut l'écrire

$$\begin{aligned}
\mu(A_1 \times A_2) &= \int_{A_1} \mu_2\big((A_1 \times A_2)_{\omega_1}\big)\, \mathrm{d}\mu_1(\omega_1) + \int_{\Omega_1 \setminus A_1} \mu_2\big((A_1 \times A_2)_{\omega_1}\big)\, \mathrm{d}\mu_1(\omega_1) \\
&= \int_{A_1} \mu_2(A_2)\, \mathrm{d}\mu_1(\omega_1) + \int_{\Omega_1 \setminus A_1} \mu_2(\emptyset)\, \mathrm{d}\mu_1(\omega_1) \\
&= \mu_2(A_2) \int_{A_1} \mathrm{d}\mu_1(\omega_1) + 0 = \mu_1(A_1)\mu_2(A_2).
\end{aligned}$$

Pour montrer *l'existence* de μ ayant la propriété (5.3), il suffit donc de prouver que (5.6) définit bien une fonction d'ensembles μ sur $\mathcal{F}_1 \otimes \mathcal{F}_2$ et que μ est une mesure sur cette tribu. Pour réaliser ce programme, supposons dans un premier temps que μ_2 est une mesure *finie*.

Définition de μ.

On sait par la proposition 5.6 que $\mu_2(E_{\omega_1})$ a bien un sens pour tout $E \in \mathcal{F}_1 \otimes \mathcal{F}_2$. Pour que l'intégrale définissant $\mu(E)$ dans (5.6) soit elle aussi définie, il reste à vérifier que l'application

$$h_E : \Omega_1 \to \mathbb{R}_+, \quad \omega_1 \mapsto \mu_2(E_{\omega_1})$$

est bien mesurable $\mathcal{F}_1$ - $\mathrm{Bor}(\mathbb{R}_+)$.

Mesurabilité de h_E. On utilise la même approche indirecte que dans la preuve de la proposition 5.6, en introduisant la famille d'ensembles

$$\mathcal{E} := \{E \in \mathcal{F}_1 \otimes \mathcal{F}_2;\ h_E \text{ est } \mathcal{F}_1 \text{ - } \mathrm{Bor}(\mathbb{R}_+) \text{ mesurable}\}.$$

On va montrer que $\mathcal{E}$ est une tribu qui contient la classe $\mathcal{R}$ des rectangles mesurables $A_1 \times A_2$, donc aussi $\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma(\mathcal{R})$, d'où $\mathcal{E} = \mathcal{F}_1 \otimes \mathcal{F}_2$. Ceci entraînera la mesurabilité de h_E pour tout $E \in \mathcal{F}_1 \otimes \mathcal{F}_2$.

Vérifions d'abord que tous les rectangles mesurables sont dans $\mathcal{E}$. Grâce à (5.2), on vérifie immédiatement que

$$\forall\ A_1 \times A_2 \in \mathcal{R},\ \forall \omega_1 \in \Omega_1, \quad h_{A_1\times A_2}(\omega_1) = \mu_2(A_2)\mathbf{1}_{A_1}(\omega_1).$$

Ainsi $h_{A_1\times A_2}$ est le produit de la constante positive $\mu_2(A_2)$ par la fonction $\mathcal{F}_1$ - $\mathrm{Bor}(\mathbb{R}_+)$ mesurable $\mathbf{1}_{A_1}$ (sa mesurabilité équivaut à l'appartenance de A_1 à $\mathcal{F}_1$). $h_{A_1\times A_2}$ est donc elle aussi $\mathcal{F}_1$ - $\mathrm{Bor}(\mathbb{R}_+)$ mesurable, d'où l'appartenance de $A_1 \times A_2$ à $\mathcal{E}$.

Essayons de montrer que $\mathcal{E}$ est une tribu[2]. L'ensemble vide appartient à $\mathcal{E}$. En effet, $h_\emptyset$ est la fonction nulle donc mesurable comme fonction constante[3]. Pour vérifier la stabilité par complémentaire de $\mathcal{E}$, notons $E^c := \Omega_1 \times \Omega_2 \setminus E$. On peut écrire grâce au lemme 5.5 a) et à la *finitude* de μ_2,

$$h_{E^c}(\omega_1) = \mu_2\bigl((E^c)_{\omega_1}\bigr) = \mu_2\bigl(\Omega_2 \setminus E_{\omega_1}\bigr) = \mu_2(\Omega_2) - \mu_2(E_{\omega_1}) = \mu_2(\Omega_2) - h_E(\omega_1) \geq 0.$$

Ainsi pour $E \in \mathcal{E}$, h_{E^c} est la différence de la fonction constante (finie) $\mu_2(\Omega_2)$ et de la fonction mesurable positive (à valeurs finies) h_E. Elle est donc $\mathcal{F}_1$ - $\mathrm{Bor}(\mathbb{R})$ mesurable et comme elle est positive, elle est aussi $\mathcal{F}_1$ - $\mathrm{Bor}(\mathbb{R}_+)$ mesurable, ce qui établit l'appartenance de E^c à $\mathcal{E}$. Pour la stabilité de $\mathcal{E}$ par union dénombrable, soit $(E_k)_{k\geq 1}$ une suite dans $\mathcal{E}$ et $E := \underset{k\in\mathbb{N}^*}{\cup} E_k$. En notant $F_n := \underset{1\leq k\leq n}{\cup} E_k$, on a $F_n \uparrow E$ et $h_{F_n} \uparrow h_E$ par continuité séquentielle croissante de μ_2. Il suffit donc d'établir la mesurabilité de h_{F_n} pour tout $n \geq 1$, autrement dit la stabilité de $\mathcal{E}$ par union finie. On s'est ainsi ramené à vérifier que si A et B sont dans $\mathcal{E}$, $A \cup B$ y est aussi. C'est facile dans le cas particulier où A et B sont disjoints puisque le lemme 5.5 b) et l'additivité de μ_2 nous permettent d'écrire $h_{A\cup B} = h_A + h_B$, de sorte que $h_{A\cup B}$ est mesurable comme somme de deux fonctions mesurables et $A \cup B \in \mathcal{E}$. Malheureusement dans le cas où A et B ne sont pas disjoints, on ne voit pas bien comment montrer directement la mesurabilité de $h_{A\cup B}$. Notons au passage que le cas A et B disjoints nous donne la stabilité de $\mathcal{E}$ pour la différence propre : si $C, D \in \mathcal{E}$ et $C \subset D$, en posant $A = C$ et $B = D \setminus C$, l'égalité $h_{A\cup B} = h_A + h_B$ nous donne $h_{D\setminus C} = h_D - h_C$ et assure la mesurabilité de $h_{D\setminus C}$, donc l'appartenance de $D \setminus C$ à $\mathcal{E}$. On voit aussi que l'appartenance de $A \cup B$ (dans le cas général) à $\mathcal{E}$ équivaut à celle de $A \cap B$, mais cela ne nous avance guère.

Dans cette tentative infructueuse pour montrer que $\mathcal{E}$ est une tribu, on a quand même établi entre autres que $\mathcal{E}$ contient la π-classe $\mathcal{R}$ des rectangles mesurables, qu'elle est stable par réunion croissante et par différence propre, donc que $\mathcal{E}$ est une λ-classe. Par le théorème de Dynkin (cf. th. 1.15), $\mathcal{E}$ contient $\sigma(\mathcal{R}) = \mathcal{F}_1 \otimes \mathcal{F}_2$. Comme par construction, $\mathcal{E} \subset \mathcal{F}_1 \otimes \mathcal{F}_2$, on a l'égalité $\mathcal{E} = \mathcal{F}_1 \otimes \mathcal{F}_2$, ce qui achève la preuve de la mesurabilité de h_E pour tout $E \in \mathcal{F}_1 \otimes \mathcal{F}_2$.

On a donc montré que (5.6) définissait bien une fonction d'ensembles sur $\mathcal{F}_1 \otimes \mathcal{F}_2$ lorsque μ_2 est une mesure finie. Lorsque μ_2 n'est pas finie, sa σ-finitude nous assure de

[2] L'approche exposée ci-dessous n'est délibérément pas la plus courte, elle a pour but de localiser la difficulté et de motiver l'invocation du théorème de Dynkin pour la solution de ce problème.

[3] Une fonction constante $\Omega_1 \to \mathbb{R}_+$ est déjà mesurable lorsque l'on munit Ω_1 de la tribu triviale $\{\emptyset, \Omega_1\}$, donc *a fortiori* pour toute tribu sur Ω_1.

l'existence dans $\mathcal{F}_2$ d'une suite $\Omega_{2,n} \uparrow \Omega_2$ telle que pour tout $n \geq 1$, $\mu_2(\Omega_{2,n}) < +\infty$. Le cas μ_2 finie montre alors que $\mathcal{E}$ contient tous les $E \in \mathcal{F}_1 \otimes \mathcal{F}_2$ qui sont inclus dans au moins un $\Omega_1 \times \Omega_{2,n}$. Soit alors F quelconque dans $\mathcal{F}_1 \otimes \mathcal{F}_2$. On a $F_n := F \cap (\Omega_1 \times \Omega_{2,n}) \uparrow F$ et chaque F_n est dans $\mathcal{E}$. Comme la propriété de stabilité par union croissante de $\mathcal{E}$ n'utilisait pas la finitude de μ_2, on en déduit que $F = \cup_{n\geq 1} F_n$ est lui aussi dans $\mathcal{E}$. On a ainsi établi l'inclusion $\mathcal{F}_1 \otimes \mathcal{F}_2 \subset \mathcal{E}$ et donc l'égalité $\mathcal{F}_1 \otimes \mathcal{F}_2 = \mathcal{E}$ dans le cas où μ_2 est σ-finie.

La fonction d'ensembles μ définie par (5.6) est une mesure. Il est clair que

$$\mu(\emptyset) = \int_{\Omega_1} h_\emptyset \,\mathrm{d}\mu_1 = \int_{\Omega_1} 0 \,\mathrm{d}\mu_1 = 0.$$

Soient $(E_n)_{n\geq 1}$ une suite d'éléments deux à deux disjoints de $\mathcal{F}_1 \otimes \mathcal{F}_2$ et $E := \cup_{n\geq 1} E_n$. Alors pour tout $\omega_1 \in \Omega_1$ les sections E_{n,ω_1} sont des éléments deux à deux disjoints de $\mathcal{F}_2$ (proposition 5.6 et lemme 5.5 c)). Par le lemme 5.5 b), E_{ω_1} s'écrit comme la réunion disjointe $E_{\omega_1} = \cup_{n\geq 1} E_{n,\omega_1}$ et par σ additivité de μ_2,

$$h_E = \sum_{n=1}^{+\infty} h_{E_n}.$$

En utilisant le corollaire du théorème de Beppo Levi pour les séries de fonctions mesurables positives, on en déduit

$$\mu(E) = \int_{\Omega_1} h_E \,\mathrm{d}\mu_1 = \int_{\Omega_1} \sum_{n=1}^{+\infty} h_{E_n} \,\mathrm{d}\mu_1 = \sum_{n=1}^{+\infty} \int_{\Omega_1} h_{E_n} \,\mathrm{d}\mu_1 = \sum_{n=1}^{+\infty} \mu(E_n),$$

ce qui établit la σ-additivité de μ et achève la preuve de l'existence de μ.

Dans la preuve de *l'existence* de μ, on a utilisé la σ-finitude de μ_2, pas celle de μ_1. Cependant nous avons eu besoin de la σ-finitude des deux mesures pour prouver l'unicité de μ. Un autre avantage de supposer μ_1 σ-finie est que l'on peut échanger les rôles de μ_1 et μ_2 dans la construction de μ et obtenir ainsi la formule (5.4). □

5.1.3 Application aux calculs de volumes et d'espérances

Voici une première application importante de la mesure produit.

Proposition 5.9. *Notons λ_d la mesure de Lebesgue sur $\mathbb{R}^d$, $(\Omega, \mathcal{F}, \nu)$ un espace mesuré, la mesure ν étant σ-finie et $f : \Omega \to \mathbb{R}$ une application $\mathcal{F}$ - $\mathrm{Bor}(\mathbb{R})$ mesurable positive. On définit son hypographe par*

$$G := \{(\omega, y) \in \Omega \times \mathbb{R};\ 0 \leq y \leq f(\omega)\}.$$

Alors G appartient à la tribu $\mathcal{F} \otimes \mathrm{Bor}(\mathbb{R})$ et

$$\nu \otimes \lambda_1(G) = \int_\Omega f \,\mathrm{d}\nu. \tag{5.7}$$

En particulier quand $\Omega = \mathbb{R}$ et $\nu = \lambda_1$, (5.7) donne (grâce à l'égalité $\lambda_1 \otimes \lambda_1 = \lambda_2$) l'interprétation de $\int_{\mathbb{R}} f \,\mathrm{d}\lambda_1$ comme l'aire de la région délimitée par le graphe de f et l'axe des abscisses. De même, quand $\Omega = \mathbb{R}^d$ et $\nu = \lambda_d$, identifions $\mathbb{R}^d \times \mathbb{R}$ avec $\mathbb{R}^{d+1}$, notons x un élément générique de $\mathbb{R}^d$ et y la dernière coordonnée dans $\mathbb{R}^{d+1}$. La formule (5.7) permet alors, grâce à l'égalité $\lambda_{d+1} = \lambda_d \otimes \lambda_1$, d'interpréter $\int_{\mathbb{R}^d} f(x) \,\mathrm{d}\lambda_d(x)$ comme le volume de la région délimitée par l'hyperplan de $\mathbb{R}^{d+1}$ d'équation $y = 0$ et l'hypersurface d'équation $y = f(x)$.

L'égalité $\lambda_{d+1} = \lambda_d \otimes \lambda_1$ est admise provisoirement. Elle sera établie à la sous-section 5.3.1, remarque 5.17.

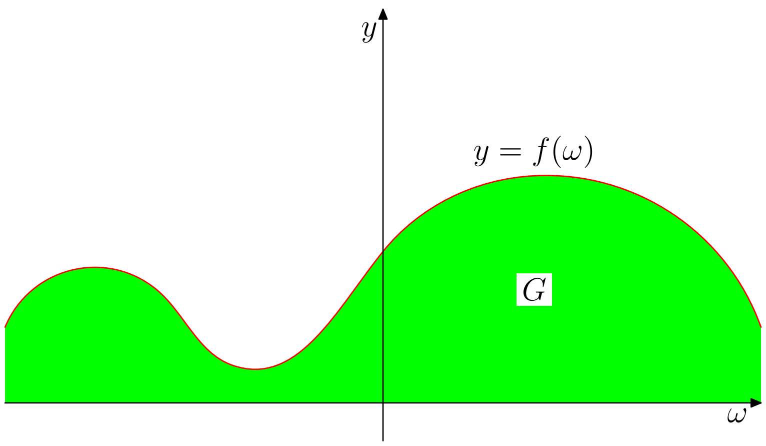

Fig. 5.2 – Hypographe de f

Démonstration. Vérifions d'abord l'appartenance de G à la tribu produit. Pour cela, on introduit l'application

$$\varphi : \Omega \times \mathbb{R} \to \mathbb{R}^2, \quad (\omega, y) \mapsto \big(f(\omega), y\big)$$

et on note que $G = \varphi^{-1}(H)$ où $H := \{(x, y) \in \mathbb{R}^2; \, 0 \leq y \leq x\}$ est fermé (donc borélien) de $\mathbb{R}^2$. Il suffit alors de montrer que φ est mesurable $\mathcal{F} \otimes \mathrm{Bor}(\mathbb{R})$ - $\mathrm{Bor}(\mathbb{R}^2)$ pour établir l'appartenance de G à $\mathcal{F} \otimes \mathrm{Bor}(\mathbb{R})$. Pour la mesurabilité de φ, on vérifie facilement que pour tout pavé de $\mathbb{R}^2$ de la forme $]a, b] \times]c, d]$,

$$\varphi^{-1}(]a, b] \times]c, d]) = f^{-1}(]a, b]) \times]c, d].$$

Comme f est $\mathcal{F}$ - $\mathrm{Bor}(\mathbb{R})$ mesurable, $f^{-1}(]a, b]) \in \mathcal{F}$ et on voit ainsi que $\varphi^{-1}(]a, b] \times]c, d])$ est un rectangle mesurable de la tribu $\mathcal{F} \otimes \mathrm{Bor}(\mathbb{R})$. La classe des $]a, b] \times]c, d]$ engendrant $\mathrm{Bor}(\mathbb{R}^2)$, on a prouvé la mesurabilité requise pour φ.

On applique alors la formule (5.6) de définition de la mesure produit en notant que la section G_ω est le segment $[0, f(\omega)]$ donc $\lambda_1(G_\omega) = f(\omega)$ (par hypothèse, $f(\omega) \geq 0$) :

$$\nu \otimes \lambda_1(G) = \int_\Omega \lambda_1(G_\omega) \,\mathrm{d}\nu(\omega) = \int_\Omega f(\omega) \,\mathrm{d}\nu(\omega).$$

□

En adaptant l'argument utilisé dans la preuve ci-dessus, il est facile de voir que le graphe de f, *i.e.* l'ensemble des couples (ω, y) tels que $y = f(\omega)$ est de $\nu \otimes \lambda_1$ mesure nulle (exercice).

Voici maintenant une application probabiliste analogue à la proposition 5.9.

Proposition 5.10. *Si X est une variable aléatoire* positive *sur l'espace probabilisé* $(\Omega, \mathcal{F}, \mathbf{P})$,

$$\mathbf{E}X = \int_{\mathbb{R}_+} \mathbf{P}(X \geq x)\, \mathrm{d}\lambda_1(x) = \int_{\mathbb{R}_+} \mathbf{P}(X > x)\, \mathrm{d}\lambda_1(x). \tag{5.8}$$

Cette formule est utile pour calculer directement l'espérance d'une variable aléatoire positive à partir de sa fonction de répartition $F(x) = \mathbf{P}(X \leq x)$ puisque $\mathbf{P}(X > x) = 1 - F(x)$. La deuxième égalité dans (5.8) n'est pas surprenante puisque les fonctions intégrées ne diffèrent qu'aux points de discontinuité de la fonction monotone F dont l'ensemble est au plus dénombrable, donc de λ_1 mesure nulle.

Démonstration. Pour la première égalité dans (5.8), posons

$$G := \{(\omega, x) \in \Omega \times \mathbb{R};\ 0 \leq x \leq X(\omega)\}.$$

Par la proposition 5.9, on a $\mathbf{E}X = \int_\Omega X\, \mathrm{d}\mathbf{P} = \mathbf{P} \otimes \lambda_1(G)$ et en utilisant [4] (5.4),

$$\begin{aligned} \mathbf{P} \otimes \lambda_1(G) = \int_{\mathbb{R}_+} \mathbf{P}(G_x)\, \mathrm{d}\lambda_1(x) &= \int_{\mathbb{R}_+} \mathbf{P}(\{\omega \in \Omega;\ 0 \leq x \leq X(\omega)\})\, \mathrm{d}\lambda_1(x) \\ &= \int_{\mathbb{R}_+} \mathbf{P}(X \geq x)\, \mathrm{d}\lambda_1(x). \end{aligned}$$

À titre d'exercice, on obtiendra la deuxième égalité de (5.8) de la même façon en remplaçant G par $\{(\omega, x) \in \Omega \times \mathbb{R}_+;\ 0 \leq x < f(\omega)\}$. □

5.2 Intégrales doubles

On étudie maintenant l'intégration par rapport à une mesure produit. Il s'agit de voir sous quelles conditions une intégrale $\int_{\Omega_1 \times \Omega_2} f\, \mathrm{d}(\mu_1 \otimes \mu_2)$ est égale à une intégrale double (ou itérée) de la forme $\int_{\Omega_1} \bigl\{ \int_{\Omega_2} f(\omega_1, \,.\,)\, \mathrm{d}\mu_2 \bigr\}\, \mathrm{d}\mu_1(\omega_1)$.

5.2.1 Cas des fonctions mesurables positives

L'intégrale par rapport à une mesure produit d'une fonction mesurable positive sur $\Omega_1 \times \Omega_2$ est toujours égale (dans $\overline{\mathbb{R}}_+$) à chacune des deux intégrales itérées $\int_{\Omega_1} \int_{\Omega_2}$ et $\int_{\Omega_2} \int_{\Omega_1}$. Ce sympathique résultat généralise aux intégrales la propriété d'échange des sommations pour les séries doubles à termes positifs. L'énoncé précis est le suivant.

[4] Remarquer que pour $x < 0$, la section G_x se réduit à l'ensemble vide d'où $\mathbf{P}(G_x) = 0$, d'où l'intégrale sur $\mathbb{R}_+$ au lieu de $\mathbb{R}$.

Théorème 5.11 (Fubini - Tonelli). *Pour $i = 1, 2$, on suppose que μ_i est une mesure σ-finie sur $(\Omega_i, \mathcal{F}_i)$. Soit $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}_+$ une application $\mathcal{F}_1 \otimes \mathcal{F}_2$ - $\mathrm{Bor}(\overline{\mathbb{R}}_+)$ mesurable positive. Les applications F_i*

$$F_1 : \omega_1 \mapsto \int_{\Omega_2} f(\omega_1, \omega_2) \,\mathrm{d}\mu_2(\omega_2) \quad \text{et} \quad F_2 : \omega_2 \mapsto \int_{\Omega_1} f(\omega_1, \omega_2) \,\mathrm{d}\mu_1(\omega_1)$$

sont $\mathcal{F}_i$ - $\mathrm{Bor}(\overline{\mathbb{R}}_+)$ mesurables et vérifient

$$\begin{aligned} \int_{\Omega_1 \times \Omega_2} f \,\mathrm{d}(\mu_1 \otimes \mu_2) &= \int_{\Omega_1} \left\{ \int_{\Omega_2} f(\omega_1, \omega_2) \,\mathrm{d}\mu_2(\omega_2) \right\} \mathrm{d}\mu_1(\omega_1) && (5.9) \\ &= \int_{\Omega_2} \left\{ \int_{\Omega_1} f(\omega_1, \omega_2) \,\mathrm{d}\mu_1(\omega_1) \right\} \mathrm{d}\mu_2(\omega_2). && (5.10) \end{aligned}$$

Démonstration. Avant d'examiner la mesurabilité des F_i, notons que grâce au lemme 5.7 appliqué avec $\Omega = \overline{\mathbb{R}}_+$, les applications partielles $f(\omega_1, \,.\,)$ et $f(\,.\,, \omega_2)$ sont mesurables positives et les intégrales définissant les F_i ont un sens comme éléments de $\overline{\mathbb{R}}_+$.

Considérons maintenant le cas particulier où f est l'indicatrice d'un ensemble E appartenant à la tribu $\mathcal{F}_1 \otimes \mathcal{F}_2$. On a alors d'après (5.4),

$$\int_{\Omega_1 \times \Omega_2} f \,\mathrm{d}(\mu_1 \otimes \mu_2) = \mu_1 \otimes \mu_2(E) = \int_{\Omega_1} \mu_2(E_{\omega_1}) \,\mathrm{d}\mu_1(\omega_1) = \int_{\Omega_2} \mu_1(E_{\omega_2}) \,\mathrm{d}\mu_2(\omega_2). \quad (5.11)$$

D'autre part, on vérifie facilement que

$$\mathbf{1}_E(\omega_1, \omega_2) = \mathbf{1}_{E_{\omega_1}}(\omega_2) = \mathbf{1}_{E_{\omega_2}}(\omega_1),$$

d'où

$$F_1(\omega_1) = \int_{\Omega_2} \mathbf{1}_E(\omega_1, \omega_2) \,\mathrm{d}\mu_2(\omega_2) = \int_{\Omega_2} \mathbf{1}_{E_{\omega_1}}(\omega_2) \,\mathrm{d}\mu_2(\omega_2) = \mu_2(E_{\omega_1}).$$

et symétriquement $F_2(\omega_2) = \mu_1(E_{\omega_2})$. La mesurabilité de $F_1 : \omega_1 \mapsto \mu_2(E_{\omega_1})$ a été établie dans la preuve du théorème 5.8. Celle de F_2 s'en déduit par symétrie. En reportant les expressions trouvées pour F_1 et F_2 dans (5.11), on a bien la double égalité (5.9)–(5.10) et le théorème est prouvé dans le cas particulier où f est l'indicatrice d'un ensemble $E \in \mathcal{F}_1 \otimes \mathcal{F}_2$.

La mesurabilité des F_i et la vérification de la double égalité (5.9)–(5.10) dans le cas général s'en déduisent par la méthode d'extension standard en considérant successivement les fonctions étagées puis les limites croissantes de fonctions étagées... la rédaction détaillée est laissée en exercice. □

5.2.2 Cas général

Théorème 5.12 (Fubini). *Pour $i = 1, 2$, on suppose que μ_i est une mesure σ-finie sur $(\Omega_i, \mathcal{F}_i)$. Soit $f : \Omega_1 \times \Omega_2 \to \mathbb{K}$, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ une application $\mathcal{F}_1 \otimes \mathcal{F}_2$ - $\mathrm{Bor}(\mathbb{K})$ mesurable.*

a) Pour que f soit $\mu_1 \otimes \mu_2$ intégrable, il faut et il suffit que l'une des deux intégrales

$$I = \int_{\Omega_1} \left\{ \int_{\Omega_2} |f(\omega_1, \omega_2)| \, \mathrm{d}\mu_2(\omega_2) \right\} \mathrm{d}\mu_1(\omega_1), \quad J = \int_{\Omega_2} \left\{ \int_{\Omega_1} |f(\omega_1, \omega_2)| \, \mathrm{d}\mu_1(\omega_1) \right\} \mathrm{d}\mu_2(\omega_2)$$

soit finie.

b) Si f est $\mu_1 \otimes \mu_2$ intégrable, la fonction $f(\omega_1, \, . \,)$ est μ_2 intégrable sur Ω_2 pour μ_1 presque tout $\omega_1 \in \Omega_1$. La fonction $\omega_1 \mapsto \int_{\Omega_2} f(\omega_1, \omega_2) \, \mathrm{d}\mu_2(\omega_2)$ est définie μ_1 presque partout sur Ω_1 et μ_1 intégrable sur Ω_1. L'énoncé analogue obtenu en permutant les rôles de μ_1 et μ_2 est aussi valable.

c) Si f est $\mu_1 \otimes \mu_2$ intégrable, on a

$$\begin{aligned} \int_{\Omega_1 \times \Omega_2} f \, \mathrm{d}(\mu_1 \otimes \mu_2) &= \int_{\Omega_1} \left\{ \int_{\Omega_2} f(\omega_1, \omega_2) \, \mathrm{d}\mu_2(\omega_2) \right\} \mathrm{d}\mu_1(\omega_1) \qquad (5.12) \\ &= \int_{\Omega_2} \left\{ \int_{\Omega_1} f(\omega_1, \omega_2) \, \mathrm{d}\mu_1(\omega_1) \right\} \mathrm{d}\mu_2(\omega_2). \qquad (5.13) \end{aligned}$$

Démonstration. Pour tout $\omega_1 \in \Omega_1$, $f(\omega_1, \, . \,)$ est $\mathcal{F}_2$ - $\mathrm{Bor}(\mathbb{K})$ mesurable par le lemme 5.7. On peut alors définir la fonction

$$G_1 : \Omega_1 \to \overline{\mathbb{R}}_+, \quad G_1(\omega_1) := \int_{\Omega_2} |f(\omega_1, \, . \,)| \, \mathrm{d}\mu_2.$$

Par le théorème de Fubini-Tonelli, G_1 est une fonction mesurable positive sur Ω_1. Il en est de même pour G_2 obtenue en échangeant les rôles de Ω_1 et Ω_2. On a alors $I = \int_{\Omega_1} G_1 \, \mathrm{d}\mu_1$ et $J = \int_{\Omega_2} G_2 \, \mathrm{d}\mu_2$.

La condition nécessaire et suffisante d'intégrabilité de f par rapport à $\mu_1 \otimes \mu_2$ s'écrit comme d'habitude $\int_{\Omega_1 \times \Omega_2} |f| \, \mathrm{d}(\mu_1 \otimes \mu_2) < +\infty$. Le théorème de Fubini-Tonelli appliqué à la fonction mesurable positive $|f|$ nous dit que cette intégrale est *toujours* égale à I et à J, qu'elle soit finie ou non. La c.n.s. d'intégrabilité a) en résulte.

Pour vérifier le point b), on suppose désormais que $I = \int_{\Omega_1} G_1 \, \mathrm{d}\mu_1 < +\infty$. Alors G_1 est finie μ_1 presque partout sur Ω_1 :

$$\Omega_1' := \{\omega_1 \in \Omega_1; \ G_1(\omega_1) < +\infty\} \in \mathcal{F}_1 \quad \text{et} \quad \mu_1(\Omega_1 \setminus \Omega_1') = 0.$$

Par définition de G_1 et Ω_1', on voit alors que pour tout $\omega_1 \in \Omega_1'$, l'application partielle mesurable $f(\omega_1, \, . \,)$ est μ_2 intégrable sur Ω_2. On peut donc définir

$$F_1 : \Omega_1 \to \mathbb{K}, \quad F_1(\omega_1) := \begin{cases} \int_{\Omega_2} f(\omega_1, \, . \,) \, \mathrm{d}\mu_2 & \text{si } \omega_1 \in \Omega_1', \\ 0 & \text{si } \omega_1 \notin \Omega_1'. \end{cases}$$

Si f est réelle, en séparant f^+ et f^- et en leur appliquant le théorème 5.11, on voit que $\int_{\Omega_2} f^+(\omega_1, \, . \,) \, \mathrm{d}\mu_2$ et $\int_{\Omega_2} f^-(\omega_1, \, . \,) \, \mathrm{d}\mu_2$ sont à valeurs positives *finies* pour chaque $\omega_1 \in \Omega_1'$ et mesurables comme fonctions $\Omega_1 \to \overline{\mathbb{R}}_+$. En écrivant

$$F_1(\omega_1) = \mathbf{1}_{\Omega_1'}(\omega_1) \int_{\Omega_2} f^+(\omega_1, \, . \,) \, \mathrm{d}\mu_2 - \mathbf{1}_{\Omega_1'}(\omega_1) \int_{\Omega_2} f^-(\omega_1, \, . \,) \, \mathrm{d}\mu_2, \qquad (5.14)$$

on voit que F_1 est $\mathcal{F}_1$ - Bor($\mathbb{R}$) mesurable. En effet, chacun des deux termes de cette différence est mesurable $\mathcal{F}_1$ - Bor($\overline{\mathbb{R}}_+$) comme produit de deux fonctions ayant cette même mesurabilité. Pour $\int_{\Omega_2} f^{\pm}(\omega_1, \,.\,) \,\mathrm{d}\mu_2$, cette mesurabilité est garantie par le théorème de Fubini-Tonelli. Grâce à l'indicatrice et à la convention arithmétique « $0 \times (+\infty) = 0$ », la fonction $\omega_1 \mapsto \mathbf{1}_{\Omega_1'}(\omega_1) \int_{\Omega_2} f^+(\omega_1, \,.\,) \,\mathrm{d}\mu_2$ ne prend que des valeurs *finies*. On en déduit facilement qu'elle est *aussi* mesurable $\mathcal{F}_1$ - Bor($\mathbb{R}$) (utiliser le corollaire 2.10). On a clairement le même résultat avec f^- à la place de f^+. Par conséquent F_1 est mesurable $\mathcal{F}_1$ - Bor($\mathbb{R}$) comme différence de deux fonctions ayant cette même mesurabilité[5].

Lorsque f est à valeurs complexes, on obtient la mesurabilité de F_1 en séparant partie réelle et partie imaginaire de f. Enfin, la μ_1 intégrabilité de F_1 résulte de l'inégalité $|F_1(\omega_1)| \leq \int_{\Omega_2} |f(\omega_1, \,.\,)| \,\mathrm{d}\mu_2$ et de l'hypothèse $I = \int_{\Omega_1} G_1 \,\mathrm{d}\mu_1 < +\infty$.

Pour vérifier c), si $\mathbb{K} = \mathbb{R}$ et f est $\mu_1 \otimes \mu_2$ intégrable, on obtient les égalités (5.12) et (5.13) en écrivant $f = f^+ - f^-$ et en appliquant les égalités analogues (5.9) et (5.10) du théorème de Fubini-Tonelli aux fonctions mesurables positives f^+ et f^-. Il n'y a aucun problème pour les différences lorsqu'on recolle les morceaux puisque toutes les intégrales concernées sont finies. Le cas complexe se déduit du cas réel par séparation des parties réelle et imaginaire. $\square$

5.2.3 Fonctions $f_1 \otimes f_2$

Un cas particulier important est celui où f peut s'écrire comme le produit $f_1 \otimes f_2$ de deux fonctions $f_i : \Omega_i \to \mathbb{K}$, $i = 1, 2$ d'une variable :

$$f_1 \otimes f_2 : \Omega_1 \times \Omega_2 \to \mathbb{K}, \quad (\omega_1, \omega_2) \mapsto f_1(\omega_1) f_2(\omega_2).$$

Étudions la mesurabilité[6] de $f_1 \otimes f_2$. Supposons d'abord que chaque f_i soit $\mathcal{F}_i$ - Bor($\mathbb{K}$) mesurable ($\mathbb{K} = \overline{\mathbb{R}}_+$, $\mathbb{R}$ ou $\mathbb{C}$). Écrivons alors $f_1 \otimes f_2 = p \circ \varphi$ où $p : (x, y) \mapsto xy$ est la multiplication $\mathbb{K}^2 \to \mathbb{K}$ et $\varphi : \Omega_1 \times \Omega_2 \to \mathbb{K}^2$, $(\omega_1, \omega_2) \mapsto \big(f_1(\omega_1), f_2(\omega_2)\big)$. Pour $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, p est continue, donc borélienne et il suffit de montrer la mesurabilité de φ. Si $\mathbb{K} = \overline{\mathbb{R}}_+$, p est discontinue aux points $(0, +\infty)$ et $(+\infty, 0)$ et on établit sa mesurabilité directement (exercice). Pour montrer la mesurabilité de φ, on utilise le fait que la tribu borélienne de $\mathbb{K}^2$ est engendrée par une classe $\mathcal{C}$ de boréliens de la forme $B_1 \times B_2$, où les B_i sont eux mêmes membres d'une classe particulière de boréliens de $\mathbb{K}$. Pour $\mathbb{K} = \mathbb{R}$ ou $\overline{\mathbb{R}}_+$, on peut prendre par exemple, les B_i de la forme $[a, b]$ avec $a, b \in \mathbb{K}$. Pour $\mathbb{K} = \mathbb{C}$, on peut en passant par l'identification de $\mathbb{C}$ et $\mathbb{R}^2$, prendre les $B = \{z \in \mathbb{C};\ \mathrm{Re}(z) \in [a, b],\ \mathrm{Im}(z) \in [c, d]\}$, avec a, b, c, d réels. Il suffit alors de noter que pour $B_1 \times B_2 \in \mathcal{C}$,

$$\begin{aligned}
\varphi^{-1}(B_1 \times B_2) &= \Big\{(\omega_1, \omega_2) \in \Omega_1 \times \Omega_2;\ \big(f_1(\omega_1), f_2(\omega_2)\big) \in B_1 \times B_2\Big\} \\
&= \Big\{(\omega_1, \omega_2) \in \Omega_1 \times \Omega_2;\ f_1(\omega_1) \in B_1 \text{ et } f_2(\omega_2) \in B_2\Big\} \\
&= \Big\{(\omega_1, \omega_2) \in \Omega_1 \times \Omega_2;\ \omega_1 \in f_1^{-1}(B_1) \text{ et } \omega_2 \in f_2^{-1}(B_2)\Big\} \\
&= f_1^{-1}(B_1) \times f_2^{-1}(B_2)
\end{aligned}$$

[5] Dans (5.14) il serait catastrophique de *mettre en facteur* $\mathbf{1}_{\Omega_1'}$. Voyez vous pourquoi ?

[6] Passage à sauter en première lecture. Allez directement au corollaire 5.13.

Chaque f_i étant $\mathcal{F}_i$ - $\mathrm{Bor}(\mathbb{K})$ mesurable, $f_i^{-1}(B_i) \in \mathcal{F}_i$ d'où $f_1^{-1}(B_1) \times f_2^{-1}(B_2) \in \mathcal{F}_1 \otimes \mathcal{F}_2$. Ceci étant vrai pour tout $B_1 \times B_2 \in \mathcal{C}$, on obtient $\varphi^{-1}(\mathcal{C}) \subset \mathcal{F}_1 \otimes \mathcal{F}_2$. On en déduit $\sigma\big(\varphi^{-1}(\mathcal{C})\big) \subset \mathcal{F}_1 \otimes \mathcal{F}_2$. D'autre part $\sigma\big(\varphi^{-1}(\mathcal{C})\big) = \varphi^{-1}\big(\sigma(\mathcal{C})\big) = \varphi^{-1}\big(\mathrm{Bor}(\mathbb{K}^2)\big)$. On a ainsi établi l'inclusion $\varphi^{-1}\big(\mathrm{Bor}(\mathbb{K}^2)\big) \subset \mathcal{F}_1 \otimes \mathcal{F}_2$ qui signifie la mesurabilité de φ.

Réciproquement, la mesurabilité de $f = f_1 \otimes f_2$ implique-t-elle celle des f_i ? La réponse est oui, sauf dans des cas pathologiques. En effet d'après le lemme 5.7, on sait que $f(\omega_1, .)$ et $f(.,\omega_2)$ sont mesurables. Traitons d'abord le cas $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Si donc on peut choisir ω_i tel que $f_i(\omega_i) \neq 0$ $(i = 1,2)$, l'égalité évidente $f(\omega_1, .) = f_1(\omega_1)f_2$ nous donne $f_2 = c_1 f(\omega_1, .)$, la constante $c = 1/f_1(\omega_1)$ étant un nombre réel ou complexe[7]. De même $f_1 = c_2 f(.,\omega_2)$. Si $\mathbb{K} = \overline{\mathbb{R}}_+$, s'il existe ω_i tel que $0 < f_i(\omega_i) < +\infty$, on est ramenés à la situation précédente. Par contre si disons f_1 ne prend que les valeurs 0 ou $+\infty$, on peut avoir f mesurable sans que f_2 le soit. Pour le voir, prendre $f_1 = (+\infty)\mathbf{1}_A$, où $A \in \mathcal{F}_1$ et $f_2 = 1 + \mathbf{1}_B$ où B est une partie de Ω_2 n'appartenant pas à $\mathcal{F}_2$. Alors $f = (+\infty)\mathbf{1}_{A\times\Omega_2}$ est mesurable puisque $A \times \Omega_2 \in \mathcal{F}_1 \otimes \mathcal{F}_2$. Par contre f_2 n'est pas mesurable car $B \notin \mathcal{F}_2$.

Corollaire 5.13 (Variables séparées). *Dans le cas où f est de la forme $f_1 \otimes f_2$, si chaque f_i est μ_i intégrable $(i = 1,2)$, alors f est $\mu_1 \otimes \mu_2$ intégrable et*

$$\int_{\Omega_1\times\Omega_2} f_1 \otimes f_2 \,\mathrm{d}(\mu_1 \otimes \mu_2) = \left\{\int_{\Omega_1} f_1 \,\mathrm{d}\mu_1\right\}\left\{\int_{\Omega_2} f_2 \,\mathrm{d}\mu_2\right\}. \tag{5.15}$$

Preuve. Vérifions d'abord la $\mu_1 \otimes \mu_2$ intégrabilité de $f_1 \otimes f_2$. Les f_i étant intégrables sont en particulier mesurables, donc $f_1 \otimes f_2$ est aussi mesurable, comme nous venons de le voir. Pour établir la finitude de $\int_{\Omega_1\times\Omega_2} |f_1 \otimes f_2| \,\mathrm{d}(\mu_1 \otimes \mu_2)$, on utilise le théorème de Fubini-Tonelli :

$$\begin{aligned}
\int_{\Omega_1\times\Omega_2} |f_1 \otimes f_2| \,\mathrm{d}(\mu_1 \otimes \mu_2) &= \int_{\Omega_1} \left\{\int_{\Omega_2} |f_1(\omega_1)||f_2(\omega_2)| \,\mathrm{d}\mu_2(\omega_2)\right\} \mathrm{d}\mu_1(\omega_1) \\
&= \int_{\Omega_1} |f_1(\omega_1)| \left\{\int_{\Omega_2} |f_2| \,\mathrm{d}\mu_2\right\} \mathrm{d}\mu_1(\omega_1) \qquad (5.16)\\
&= \left\{\int_{\Omega_2} |f_2| \,\mathrm{d}\mu_2\right\} \int_{\Omega_1} |f_1| \,\mathrm{d}\mu_1 < +\infty. \qquad (5.17)
\end{aligned}$$

Justifications. Pour (5.16), on utilise le fait que $|f_1(\omega_1)|$ est une constante (dépendant de ω_1) pour l'intégration sur Ω_2. Remarquons d'ailleurs que cette constante est finie pour μ_1 presque tout ω_1 puisque f_1 est μ_1 intégrable. Pour (5.17), on a simplement sorti la constante finie $\int_{\Omega_2} |f_2| \,\mathrm{d}\mu_2$ de l'intégrale sur Ω_1. Enfin la finitude du produit des deux intégrales résulte des intégrabilités respectives de f_1 et f_2.

Une fois acquise l'intégrabilité de $f_1 \otimes f_2$, le théorème de Fubini légitime le même calcul que ci-dessus sans les valeurs absolues et donne (5.15). □

[7] Notons que la valeur $f_1(\omega_1) = +\infty$ est exclue puisque $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

5.3 Intégration sur un produit fini d'espaces

5.3.1 Produits finis de tribus et de mesures

Nous abordons brièvement l'extension de l'étude précédente à un produit de n espaces mesurés $(\Omega_i, \mathcal{F}_i, \mu_i)$, où les mesures μ_i sont σ-finies. Notons $\mathcal{R}_n$ la classe des pavés mesurables

$$\mathcal{R}_n := \{A_1 \times A_2 \times \cdots \times A_n;\ A_i \in \mathcal{F}_i,\ 1 \leq i \leq n\}.$$

On définit la tribu produit $\mathcal{E}_n$ sur $\Omega_1 \times \cdots \times \Omega_n$ par

$$\mathcal{E}_n = \bigotimes_{i=1}^{n} \mathcal{F}_i := \sigma(\mathcal{R}_n). \tag{5.18}$$

La construction de la mesure produit $\mu_1 \otimes \cdots \otimes \mu_n$ se fait par récurrence en se ramenant au cas de deux espaces facteurs. On utilise pour cela le léger abus qui consiste à identifier les espaces[8]

$$(\Omega_1 \times \cdots \times \Omega_{n-1}) \times \Omega_n \quad \text{et} \quad \Omega_1 \times \cdots \times \Omega_n. \tag{5.19}$$

On pourrait se passer de cette identification en montrant que l'application

$$\big((\omega_1, \ldots, \omega_{n-1}), \omega_n\big) \mapsto (\omega_1, \ldots, \omega_n)$$

définit une bijection bimesurable entre ces deux espaces munis respectivement des tribus produits $\mathcal{E}_{n-1} \otimes \mathcal{F}_n$ et $\mathcal{E}_n$ ...

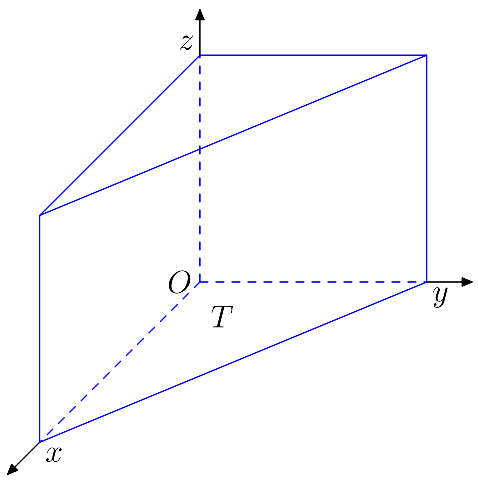

FIG. 5.3 – Prisme $C = T \times [0, 1]$

[8]Formellement, $\mathbb{R}^2 \times \mathbb{R}$ n'est pas exactement le même ensemble que $\mathbb{R}^3$. Un élément de $\mathbb{R}^2 \times \mathbb{R}$ est un *couple* $\big((x, y), z\big)$ dont la première composante est elle-même un couple de réels et la deuxième un réel, tandis qu'un élément de $\mathbb{R}^3$ est un *triplet* (x, y, z) de réels.

En réalité, *le vrai problème n'est pas tant l'identification de ces deux espaces que celle de leurs tribus*. La tribu $\mathcal{E}_{n-1}\otimes\mathcal{F}_n$ est engendrée par la classe $\mathcal{R}'$ des rectangles mesurables de la forme $E\times A_n$ avec $E\in\mathcal{E}_{n-1}$ et $A_n\in\mathcal{F}_n$. Clairement $\mathcal{R}'$ contient $\mathcal{R}_n$ (via l'identification ci-dessus) mais est plus riche que $\mathcal{R}_n$. Par exemple avec $n=3$ et $(\Omega_i,\mathcal{F}_i)=\big(\mathbb{R},\mathrm{Bor}(\mathbb{R})\big)$, $(i=1,2,3)$, le prisme de la figure 5.3

$$C:=\{(x,y,z)\in\mathbb{R}^3;\ 0\le x+y\le 1,\ x,y\ge 0,\ 0\le z\le 1\}$$

est dans $\mathcal{R}'$ comme produit cartésien du triangle $T=\{(x,y)\in\mathbb{R}_+^2;\, 0\le x+y\le 1\}$ de $\mathbb{R}^2$ par le segment $[0,1]$. Il n'est pas dans $\mathcal{R}_3$ car T n'est pas un produit cartésien. La proposition suivante permet de résoudre ce problème.

Proposition 5.14. *Avec l'identification des espaces (5.19), on a l'égalité de tribus*

$$\mathcal{E}_n=\mathcal{E}_{n-1}\otimes\mathcal{F}_n,\qquad n\ge 2.$$

Plus généralement, notons pour $0\le i<j\le n$*,*

$$\Omega_{i,j}:=\Omega_{i+1}\times\dots\times\Omega_j,\qquad \mathcal{E}_{i,j}:=\mathcal{F}_{i+1}\otimes\dots\otimes\mathcal{F}_j=\sigma(\mathcal{R}_{i,j}),$$

où $\mathcal{R}_{i,j}$ *est la classe des pavés mesurables* $A_{i+1}\times\dots\times A_j$ *avec* $A_k\in\mathcal{F}_k$*,* $i<k\le j$*. Alors en identifiant les espaces* $\Omega_{0,n}$ *et* $\Omega_{0,j}\times\Omega_{j,n}$*, on a*

$$\mathcal{E}_n=\mathcal{E}_{0,n}=\mathcal{E}_{0,j}\otimes\mathcal{E}_{j,n}.$$

Démonstration. L'égalité $\mathcal{E}_n=\mathcal{E}_{n-1}\otimes\mathcal{F}_n$ n'est que le cas $j=n-1$ de $\mathcal{E}_n=\mathcal{E}_{0,j}\otimes\mathcal{E}_{j,n}$. Pour établir cette dernière égalité, on montre l'inclusion dans les deux sens.

L'inclusion $\mathcal{E}_n\subset\mathcal{E}_{0,j}\otimes\mathcal{E}_{j,n}$ est immédiate puisque si $R=A_1\times\dots\times A_n$ est dans $\mathcal{R}_n$, on peut l'identifier avec $(A_1\times\dots\times A_j)\times(A_{j+1}\times\dots\times A_n)$. Il est donc membre de la classe $\mathcal{R}'$ des rectangles mesurables qui engendre $\mathcal{E}_{0,j}\otimes\mathcal{E}_{j,n}$. Ainsi $\mathcal{R}_n\subset\mathcal{R}'$ d'où $\mathcal{E}_n=\sigma(\mathcal{R}_n)\subset\sigma(\mathcal{R}')=\mathcal{E}_{0,j}\otimes\mathcal{E}_{j,n}$.

Pour obtenir l'inclusion dans l'autre sens, fixons d'abord un pavé quelconque $P=A_{j+1}\times\dots\times A_n$ dans $\mathcal{R}_{j,n}$ et notons

$$\mathcal{G}_P:=\{E\in\mathcal{E}_{0,j};\ E\times P\in\mathcal{E}_n\}.$$

Si $E\in\mathcal{R}_{0,j}$, $E\times P\in\mathcal{R}_n$ (via l'identification $\Omega_{0,n}=\Omega_{0,j}\times\Omega_{j,n}$). Donc $\mathcal{G}_P$ contient $\mathcal{R}_{0,j}$. On vérifie que $\mathcal{G}_P$ est une sous-tribu de $\mathcal{E}_{0,j}$ grâce au lemme élémentaire suivant qui est laissé en exercice.

Lemme 5.15. *Soient* Ω *et* Ω' *deux ensembles.*

i) Pour tout $F\subset\Omega'$*,* $\emptyset\times F=\emptyset$*.*

ii) Pour tout $E\subset\Omega$ *et tout* $F\subset\Omega'$*,* $\big(\Omega\setminus E\big)\times F=\big(\Omega\times F\big)\setminus(E\times F)$*.*

iii) Pour tout ensemble d'indices I*, toute famille* $(E_i)_{i\in I}$ *de parties de* Ω *et tout* $F\subset\Omega'$*,*

$$\Big(\bigcup_{i\in I}E_i\Big)\times F=\bigcup_{i\in I}(E_i\times F).$$

On a donc $\mathcal{E}_{0,j} \supset \mathcal{G}_P \supset \sigma(\mathcal{R}_{0,j}) = \mathcal{E}_{0,j}$. Ainsi $\mathcal{G}_P = \mathcal{E}_{0,j}$ et donc pour tout $E \in \mathcal{E}_{0,j}$, $E \times P$ est dans la tribu $\mathcal{E}_n$. Comme P a été pris quelconque dans $\mathcal{R}_{j,n}$, on a montré que

$$\forall E \in \mathcal{E}_{0,j},\ \forall P \in \mathcal{R}_{j,n}, \quad E \times P \in \mathcal{E}_n.$$

Fixons maintenant E quelconque dans $\mathcal{E}_{0,j}$ et définissons

$$\mathcal{H}_E := \{F \in \mathcal{E}_{j,n};\ E \times F \in \mathcal{E}_n\}.$$

D'après ce qui précède, $\mathcal{H}_E$ contient $\mathcal{R}_{j,n}$. On vérifie que c'est une sous-tribu de $\mathcal{E}_{j,n}$ en utilisant le lemme 5.15 (en permutant Ω et Ω'). Donc $\mathcal{E}_{j,n} \supset \mathcal{H}_E \supset \sigma(\mathcal{R}_{j,n}) = \mathcal{E}_{j,n}$ d'où $\mathcal{H}_E = \mathcal{E}_{j,n}$. Ainsi $E \times F \in \mathcal{E}_n$ pour tout $F \in \mathcal{E}_{j,n}$. Comme E est quelconque dans $\mathcal{E}_{0,j}$, on a finalement montré que

$$\forall E \in \mathcal{E}_{0,j},\ \forall F \in \mathcal{E}_{j,n}, \quad E \times F \in \mathcal{E}_n,$$

ce qui s'écrit aussi $\mathcal{R}' \subset \mathcal{E}_n$. On en déduit $\sigma(\mathcal{R}') \subset \mathcal{E}_n$, ce qui achève la preuve puisque $\sigma(\mathcal{R}') = \mathcal{E}_{0,j} \otimes \mathcal{E}_{j,n}$. □

Un exemple important de tribu produit à n facteurs est la tribu borélienne de $\mathbb{R}^n$. Cependant l'égalité

$$\mathrm{Bor}(\mathbb{R}^n) = \underbrace{\mathrm{Bor}(\mathbb{R}) \otimes \cdots \otimes \mathrm{Bor}(\mathbb{R})}_{n \text{ facteurs}} =: \big(\mathrm{Bor}(\mathbb{R})\big)^{\otimes n}, \tag{5.20}$$

n'est pas une simple conséquence de la proposition 5.14. En effet la tribu borélienne de $\mathbb{R}^n$ n'a pas été définie comme tribu *produit* des tribus boréliennes de $\mathbb{R}$, autrement dit la tribu engendrée par les $B_1 \times \cdots \times B_n$ où $B_i \in \mathrm{Bor}(\mathbb{R})$, mais comme la tribu engendrée par l'ensemble des ouverts de $\mathbb{R}^n$. La justification de (5.20) demande le petit effort supplémentaire suivant.

Proposition 5.16. *Soient p, $q \in \mathbb{N}^*$. En identifiant $\mathbb{R}^p \times \mathbb{R}^q$ et $\mathbb{R}^{p+q}$, on a*

$$\mathrm{Bor}(\mathbb{R}^p) \otimes \mathrm{Bor}(\mathbb{R}^q) = \mathrm{Bor}(\mathbb{R}^{p+q}).$$

Démonstration. Notons $\mathcal{C}$ la classe des parallélépipèdes $Q = \prod_{i=1}^{p+q}]a_i, b_i]$. L'identification $\mathbb{R}^{p+q} = \mathbb{R}^p \times \mathbb{R}^q$ nous permet d'écrire

$$Q = \prod_{i=1}^{p+q}]a_i, b_i] = \left(\prod_{i=1}^{p}]a_i, b_i]\right) \times \left(\prod_{i=p+1}^{p+q}]a_i, b_i]\right),$$

faisant ainsi apparaître Q comme un élément de la classe $\mathcal{R}$ des rectangles mesurables de la tribu produit $\mathrm{Bor}(\mathbb{R}^p) \otimes \mathrm{Bor}(\mathbb{R}^q)$. On a donc $\mathcal{C} \subset \mathcal{R}$, d'où $\sigma(\mathcal{C}) \subset \sigma(\mathcal{R}) = \mathrm{Bor}(\mathbb{R}^p) \otimes \mathrm{Bor}(\mathbb{R}^q)$. D'autre part on sait que la tribu borélienne de $\mathbb{R}^{p+q}$ est engendrée par $\mathcal{C}$. On vient ainsi d'établir l'inclusion $\mathrm{Bor}(\mathbb{R}^{p+q}) \subset \mathrm{Bor}(\mathbb{R}^p) \otimes \mathrm{Bor}(\mathbb{R}^q)$.

Pour obtenir l'inclusion dans l'autre sens, il suffit de montrer que $\mathcal{R} \subset \mathrm{Bor}(\mathbb{R}^{p+q})$, autrement dit que pour tous $A \in \mathrm{Bor}(\mathbb{R}^p)$ et $B \in \mathrm{Bor}(\mathbb{R}^q)$, $A \times B$ est dans $\mathrm{Bor}(\mathbb{R}^{p+q})$. Considérons les projections canoniques

$$\begin{aligned} \pi_1 : \mathbb{R}^{p+q} \to \mathbb{R}^p, &\qquad (x_1, \dots, x_{p+q}) \longmapsto (x_1, \dots, x_p) \\ \pi_2 : \mathbb{R}^{p+q} \to \mathbb{R}^q, &\qquad (x_1, \dots, x_{p+q}) \longmapsto (x_{p+1}, \dots, x_{p+q}). \end{aligned}$$

Ce sont des applications continues, donc boréliennes. On en déduit l'appartenance à $\mathrm{Bor}(\mathbb{R}^{p+q})$ des ensembles $\pi_1^{-1}(A) = A \times \mathbb{R}^q$ et $\pi_2^{-1}(B) = \mathbb{R}^p \times B$. On conclut en notant que $A \times B = (A \times \mathbb{R}^q) \cap (\mathbb{R}^p \times B)$. $\square$

Revenons à la construction de la mesure produit sur la tribu $\mathcal{E}_n$. On peut maintenant définir

$$\nu_n := \bigotimes_{i=1}^{n} \mu_i$$

par récurrence en posant

$$\nu_2 := \mu_1 \otimes \mu_2, \quad \nu_n := \nu_{n-1} \otimes \mu_n, \quad (n > 2).$$

Bien sûr, ν_n est l'unique mesure dont la restriction à $\mathcal{R}_n$ vérifie

$$\forall\ A_1 \times \dots \times A_n \in \mathcal{R}_n, \quad \nu_n(A_1 \times \dots \times A_n) = \mu_1(A_1)\mu_2(A_2) \dots \mu_n(A_n). \tag{5.21}$$

Remarque 5.17. En notant λ_d la mesure de Lebesgue sur $\mathbb{R}^d$, on a

$$\lambda_d = \underbrace{\lambda_1 \otimes \dots \otimes \lambda_1}_{d \text{ facteurs}} =: \lambda_1^{\otimes d}.$$

En effet par les propositions 5.14 et 5.16, on a $\mathrm{Bor}(\mathbb{R}^d) = \mathrm{Bor}(\mathbb{R})^{\otimes d}$ et grâce à (5.21) les mesures λ_d et $\lambda_1^{\otimes d}$ coïncident sur la π-classe des parallélépipèdes $]a_1, b_1] \times \dots \times]a_d, b_d]$. Par le théorème d'unicité des mesures elles coïncident sur toute la tribu engendrée, donc sur la tribu borélienne de $\mathbb{R}^d$.

5.3.2 Intégrales multiples

Les théorèmes de Tonelli et de Fubini se généralisent comme suit.

Théorème 5.18. *On suppose que pour $i = 1, \dots, n$, μ_i est une mesure σ-finie sur l'espace mesurable $(\Omega_i, \mathcal{F}_i)$ et que*

$$f : \Omega_{0,n} := \Omega_1 \times \dots \times \Omega_n \longrightarrow \overline{\mathbb{R}}_+$$

est mesurable $\bigotimes_{i=1}^n \mathcal{F}_i$ - $\mathrm{Bor}(\overline{\mathbb{R}}_+)$. Alors pour toute permutation τ des indices $1, 2, \dots, n$, on a

$$\begin{aligned} \int_{\Omega_{0,n}} f \,\mathrm{d}(\mu_1 \otimes \dots \otimes \mu_n) &= \int_{\Omega_1} \int_{\Omega_2} \dots \int_{\Omega_n} f(\omega_1, \dots, \omega_n) \,\mathrm{d}\mu_n(\omega_n) \dots \mathrm{d}\mu_2(\omega_2) \,\mathrm{d}\mu_1(\omega_1) \\ &= \int_{\Omega_{\tau(1)}} \dots \int_{\Omega_{\tau(n)}} f(\omega_1, \dots, \omega_n) \,\mathrm{d}\mu_{\tau(n)}(\omega_{\tau(n)}) \dots \mathrm{d}\mu_{\tau(1)}(\omega_{\tau(1)}). \end{aligned}$$

Toutes les intégrales itérées d'ordre k $(1 \leq k \leq n-1)$ contenues dans ces formules définissent des fonctions mesurables positives relativement aux tribus produits concernées.

Théorème 5.19. *Avec les notations du théorème précédent, si $f : \Omega_{0,n} \to \mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ est $\mathcal{E}_n$ - $\mathrm{Bor}(\mathbb{K})$ mesurable, elle est $\mu_1 \otimes \cdots \otimes \mu_n$ intégrable si et seulement si l'une des $n!$ intégrales itérées d'ordre n*

$$\int_{\Omega_{\tau(1)}} \cdots \int_{\Omega_{\tau(n)}} |f(\omega_1, \ldots, \omega_n)| \,\mathrm{d}\mu_{\tau(n)}(\omega_{\tau(n)}) \ldots \,\mathrm{d}\mu_{\tau(1)}(\omega_{\tau(1)})$$

est finie. Dans ce cas les $n!$ égalités d'intégrales du théorème précédent restent vraies et toutes les intégrales itérées d'ordre $k < n$ apparaissant dans ces formules définissent presque partout des fonctions intégrables (relativement aux mesures concernées).

Corollaire 5.20. *Si pour $i = 1, \ldots n$, $f_i : \Omega_i \to \mathbb{K}$ est μ_i intégrable, la fonction*

$$f := f_1 \otimes \cdots \otimes f_n : \Omega_{0,n} \longrightarrow \mathbb{K} = \mathbb{R}, \quad (\omega_1, \ldots, \omega_n) \longmapsto f_1(\omega_1) \ldots f_n(\omega_n)$$

est $\mu_1 \otimes \cdots \otimes \mu_n$ intégrable et

$$\int_{\Omega_{0,n}} f \,\mathrm{d}(\mu_1 \otimes \cdots \otimes \mu_n) = \prod_{i=1}^{n} \int_{\Omega_i} f_i \,\mathrm{d}\mu_i.$$

5.3.3 Application aux lois marginales d'un vecteur aléatoire

Pour terminer cette section, nous examinons une application aux vecteurs aléatoires. Rappelons que si $(\Omega, \mathcal{F}, \mathbf{P})$ est un espace probabilisé, un vecteur aléatoire sur cet espace est une application $X : \Omega \to \mathbb{R}^d$, mesurable $\mathcal{F}$ - $\mathrm{Bor}(\mathbb{R}^d)$. Sa loi P_X est la mesure image $\mathbf{P} \circ X^{-1}$ qui est une probabilité sur $\mathrm{Bor}(\mathbb{R}^d)$. Pour $i = 1, \ldots, d$, soit π_i la i-ème projection canonique de $\mathbb{R}^d$ sur $\mathbb{R}$. Notons $X_i = \pi_i \circ X$. Alors X_i est mesurable $\mathcal{F}$ - $\mathrm{Bor}(\mathbb{R})$, c'est donc une variable aléatoire réelle. Pour tout $\omega \in \Omega$, on a

$$X(\omega) = \big(X_1(\omega), X_2(\omega), \ldots, X_d(\omega)\big).$$

La loi P_{X_i} de la variable aléatoire X_i s'appelle i-ème *loi marginale* du vecteur aléatoire X. Rappelons encore que P_X a une densité f par rapport à λ_d, mesure de Lebesgue sur $\mathbb{R}^d$, si

$$\forall B \in \mathrm{Bor}(\mathbb{R}^d), \quad \mathbf{P}(X \in B) = P_X(B) = \int_B f(x_1, \ldots, x_d) \,\mathrm{d}\lambda_d(x_1, \ldots, x_d). \tag{5.22}$$

Dans ce cas, f est une fonction mesurable positive et $\int_{\mathbb{R}^d} f \,\mathrm{d}\lambda_d = 1$.

Proposition 5.21. *Soit $X = (X_1, \ldots, X_d)$ un vecteur aléatoire à valeurs dans $\mathbb{R}^d$, dont la loi a une densité f par rapport à λ_d. Alors sa i-ème loi marginale P_{X_i} admet une densité f_{X_i} par rapport à λ_1, donnée par*

$$f_{X_i}(t) = \int_{\mathbb{R}^{d-1}} f(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_d) \,\mathrm{d}\lambda_{d-1}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d).$$

Démonstration. Pour tout borélien A de $\mathbb{R}$,

$$P_{X_i}(A) = \mathbf{P}(X_i \in A) = \mathbf{P}(\pi_i \circ X \in A) = \mathbf{P}(X \in \mathbb{R}^{i-1} \times A \times \mathbb{R}^{d-i}).$$

En utilisant (5.22), la remarque 5.17 et le théorème 5.18, on en déduit

$$\begin{aligned} P_{X_i}(A) &= \int_{\mathbb{R}^{i-1}\times A\times\mathbb{R}^{d-i}} f(x_1,\dots,x_d)\,\mathrm{d}\lambda_1^{\otimes d}(x_1,\dots,x_d) \\ &= \int_A \int_{\mathbb{R}^{i-1}\times\mathbb{R}^{d-i}} f(x_1,\dots,x_d)\,\mathrm{d}\lambda_1^{\otimes(d-1)}(x_1,\dots,x_{i-1},x_{i+1},\dots,x_d)\,\mathrm{d}\lambda_1(x_i) \\ &= \int_A f_{X_i}(x_i)\,\mathrm{d}\lambda_1(x_i) \end{aligned}$$

On en déduit par (5.22) appliqué à P_{X_i}, que cette loi a pour densité f_{X_i}. □

5.4 Changement de variable dans $\mathbb{R}^d$

5.4.1 Introduction

Cette section ne relève qu'indirectement de l'intégration sur un espace produit. Les résultats qu'elle contient constituent, avec le théorème de Fubini, l'un des deux outils pour le calcul pratique d'intégrales $\int_{\mathbb{R}^d} f\,\mathrm{d}\lambda$, où λ est la mesure de Lebesgue sur $\mathbb{R}^d$. Il s'agit de donner une forme explicite au théorème de transfert pour effectuer pratiquement un changement de variable dans ce type d'intégrales. Rappelons le cadre général du théorème de transfert. On considère le diagramme suivant

$$(\Omega_1, \mathcal{F}_1, \mu) \xrightarrow{\varphi} (\Omega_2, \mathcal{F}_2, \nu := \mu\circ\varphi^{-1}) \xrightarrow{h} (\mathbb{K}, \mathrm{Bor}(\mathbb{K})),$$

où φ est mesurable $\mathcal{F}_1$ - $\mathcal{F}_2$ et h est mesurable $\mathcal{F}_2$ - $\mathrm{Bor}(\mathbb{K})$. La formule de transfert s'écrit alors

$$\int_{\Omega_1} (h\circ\varphi)\,\mathrm{d}\mu = \int_{\Omega_2} h\,\mathrm{d}\nu. \tag{5.23}$$

Cette formule est valable pour toute h mesurable positive (cas $\mathbb{K} = \overline{\mathbb{R}}_+$) et si $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, pour toute $h \in \mathcal{L}^1_{\mathbb{K}}(\nu)$ ou ce qui est équivalent, telle que $h\circ\varphi \in \mathcal{L}^1_{\mathbb{K}}(\mu)$.

Dans tout ce qui suit, on prendra $\Omega_1 = \mathbb{R}^d$ ou V ouvert de $\mathbb{R}^d$ et $\Omega_2 = \varphi(\Omega_1)$. La mesure μ sera la mesure de Lebesgue λ sur $\mathbb{R}^d$ ou sa restriction à V. Nous supposerons de plus φ *bijective*. Notons provisoirement ψ son inverse ponctuel : pour $y \in \Omega_2$, $\psi(y)$ est l'unique $x \in \Omega_1$ tel que $\varphi(x) = y$. Alors l'inverse ensembliste $\varphi^{-1}(B)$ coïncide avec l'image directe de B par l'application ψ :

$$\begin{aligned} \varphi^{-1}(B) &= \{x \in \Omega_1;\ \varphi(x) \in B\} \\ &= \{x \in \Omega_1;\ \exists y \in B, y = \varphi(x)\} \\ &= \{x \in \Omega_1;\ \exists y \in B, x = \psi(y)\} \\ &= \{\psi(y);\ y \in B\} \\ &= \psi(B). \end{aligned}$$

Cette remarque étant faite, nous revenons à la notation traditionnelle φ^{-1} pour l'inverse ponctuel ψ de φ. Posons $f := h \circ \varphi$ d'où $h = f \circ \varphi^{-1}$. La formule (5.23) peut alors s'écrire sous l'une des deux formes équivalentes :

$$\int_{\Omega_1} f(x)\,\mathrm{d}\lambda(x) = \int_{\varphi(\Omega_1)} f\big(\varphi^{-1}(y)\big)\,\mathrm{d}(\lambda \circ \varphi^{-1})(y) \tag{5.24}$$

$$\int_{\Omega_1} h\big(\varphi(x)\big)\,\mathrm{d}\lambda(x) = \int_{\varphi(\Omega_1)} h(y)\,\mathrm{d}(\lambda \circ \varphi^{-1})(y). \tag{5.25}$$

Dans (5.24), φ^{-1} intervient comme inverse ponctuel et comme inverse ensembliste, on pourrait écrire le second membre $\int_{\varphi(\Omega_1)} f\big(\psi(y)\big)\,\mathrm{d}(\lambda\circ\varphi^{-1})(y)$. Pour donner à ces formules un caractère effectif, il reste à identifier la mesure $\lambda \circ \varphi^{-1}$. Les outils à notre disposition dans le cadre du programme de la Licence nous permettent de traiter complètement le cas où φ est une bijection linéaire et nous conduiront à admettre le résultat dans le cas plus général où φ est un C^1 difféomorphisme[9].

5.4.2 Changement de variable linéaire

Soit donc φ une bijection linéaire $\mathbb{R}^d \to \mathbb{R}^d$. Elle est évidemment borélienne puisque toutes les applications linéaires entre espaces de dimension finie sont continues. Son inverse ponctuel φ^{-1} est aussi linéaire. Pour tout vecteur $u \in \mathbb{R}^d$, nous notons T_u la translation de vecteur u. C'est une bijection d'inverse $T_u^{-1} = T_{-u}$. Nous posons $\nu := \lambda \circ \varphi^{-1}$. Nous commençons par vérifier que ν est invariante par translations. Pour cela en appliquant la remarque ci-dessus sur la coïncidence entre inverse ensembliste et image directe par inverse ponctuel pour une bijection, on peut écrire pour tout borélien B de $\mathbb{R}^d$

$$(\nu \circ T_u^{-1})(B) = \nu\big(T_{-u}(B)\big) = \lambda\big((\varphi^{-1} \circ T_{-u})(B)\big).$$

Grâce à la linéarité de φ^{-1}, on a pour tout $y \in \mathbb{R}^d$,

$$(\varphi^{-1} \circ T_{-u})(y) = \varphi^{-1}(y-u) = \varphi^{-1}(y) - \varphi^{-1}(u).$$

Posant $v := \varphi^{-1}(u)$, on voit ainsi que $\varphi^{-1} \circ T_{-u} = T_{-v} \circ \varphi^{-1} = T_v^{-1} \circ \varphi^{-1}$. On en déduit

$$\lambda\big((\varphi^{-1} \circ T_{-u})(B)\big) = \lambda\big(T_v^{-1}(\varphi^{-1}(B))\big) = (\lambda \circ T_v^{-1})(\varphi^{-1}(B)).$$

Comme λ est invariante par translations, $\lambda \circ T_v^{-1} = \lambda$ d'où

$$(\lambda \circ T_v^{-1})(\varphi^{-1}(B)) = \lambda(\varphi^{-1}(B)) = \nu(B).$$

On a donc vérifié que $(\nu \circ T_u^{-1})(B) = \nu(B)$ pour tout borélien B et toute translation T_u. La mesure ν est invariante par translations.

Notons $C := [0,1]^d$ le cube unité. On a $0 < \nu(C) < +\infty$. En effet, $\varphi^{-1}(C)$ est borné comme image d'un borné par l'application linéaire continue φ^{-1}. Sa mesure de

[9] Le lecteur curieux et courageux pourra consulter W. Rudin *Analyse réelle et complexe* ou P. Billingsley *Probability and measure*…

Lebesgue est donc finie d'où $\nu(C) < +\infty$. D'autre part $\varphi^{-1}(C)$ contient $\varphi^{-1}(]0,1[^d)$ qui est *ouvert* comme image réciproque de l'ouvert $]0,1[^d$ par l'application continue φ. Cet ouvert contient au moins un cube de la forme $]x_0-\varepsilon, x_0+\varepsilon[^d$ de mesure de Lebesgue $(2\varepsilon)^d > 0$ donc $\nu(C) > 0$.

On vérifie facilement (exercice déjà fait pour $d = 1$) qu'une mesure μ invariante par translation et telle que $0 < \mu(C) < +\infty$ est proportionnelle à la mesure de Lebesgue : $\mu = K\lambda$ pour une constante $0 < K < +\infty$. Nous venons donc de prouver le résultat suivant.

Lemme 5.22. *Si $\varphi : \mathbb{R}^d \to \mathbb{R}^d$ est une bijection linéaire, il existe $K(\varphi) \in]0,+\infty[$ constante telle que*

$$\lambda \circ \varphi^{-1} = K(\varphi)\lambda. \tag{5.26}$$

Pour déterminer explicitement la constante $K(\varphi)$, il est utile d'établir les propriétés suivantes.

Lemme 5.23. *La constante $K(\varphi)$ définie par (5.26) vérifie*

a) $K(\varphi) = \lambda(\varphi^{-1}(C))$ *où* $C = [0,1]^d$.

b) Si φ et ψ sont deux bijections linéaires $\mathbb{R}^d \to \mathbb{R}^d$, $K(\varphi \circ \psi) = K(\varphi)K(\psi)$.

c) $K(\varphi) = 1/\lambda(\varphi(C))$.

Démonstration. Pour a), on note que $\lambda(C) = 1$ et on calcule $\nu(C)$ par (5.26) :

$$\lambda(\varphi^{-1}(C)) = K(\varphi)\lambda(C) = K(\varphi).$$

Le b) s'obtient en appliquant successivement a) avec $\varphi \circ \psi$, (5.26) avec ψ et a) avec φ :

$$K(\varphi\circ\psi) = \big(\lambda\circ(\psi^{-1}\circ\varphi^{-1})\big)(C) = (\lambda\circ\psi^{-1})\big(\varphi^{-1}(C)\big) = K(\psi)\lambda\big(\varphi^{-1}(C)\big) = K(\psi)K(\varphi).$$

En choisissant $\psi = \varphi^{-1}$ dans b), on obtient c). □

Arrivés à ce stade, nous avons réduit le problème d'identification de la mesure $\lambda\circ\varphi^{-1}$ au calcul du volume de $\varphi(C)$ (ou de $\varphi^{-1}(C)$). Essayons de nous en faire une idée plus précise en examinant le cas de la dimension $d = 2$. Notons (e_1, e_2) la base canonique de $\mathbb{R}^2$. Le carré unité C est

$$C = \{x = x_1e_1 + x_2e_2 \in \mathbb{R}^2;\ 0 \le x_1 \le 1,\ 0 \le x_2 \le 1\}.$$

Son image par l'application linéaire φ est donc

$$\varphi(C) = \{y = x_1\varphi(e_1) + x_2\varphi(e_2) \in \mathbb{R}^2;\ 0 \le x_1 \le 1,\ 0 \le x_2 \le 1\}.$$

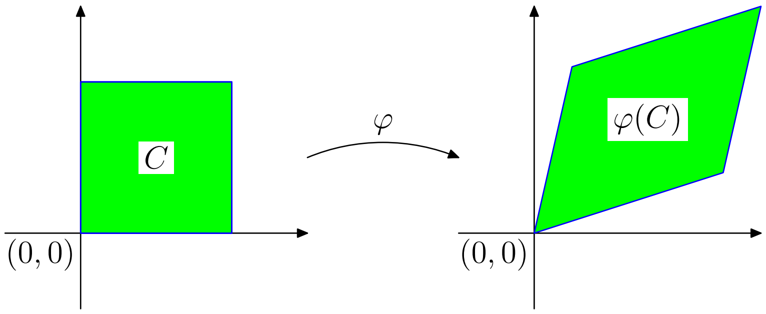

FIG. 5.4 – Image du carré unité

Il s'agit du *parallélogramme* de sommets $\bigl(0, \varphi(e_1), \varphi(e_1)+\varphi(e_2), \varphi(e_2)\bigr)$. Notons $\varphi(e_1) = ae_1 + be_2$ et $\varphi(e_2) = ce_1 + de_2$. On vérifie facilement (exercice) que l'aire de ce parallélogramme est

$$\lambda\bigl(\varphi(C)\bigr) = |ad - bc|.$$

On remarque que c'est la valeur absolue du déterminant de la matrice $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ de φ dans la base (e_1, e_2). Cette propriété est générale.

Lemme 5.24. *Si φ est une bijection linéaire $\mathbb{R}^d \to \mathbb{R}^d$, la constante $K(\varphi)$ définie par (5.26) vaut*

$$K(\varphi) = \frac{1}{|\det \varphi|} = |\det(\varphi^{-1})|. \tag{5.27}$$

Démonstration. Comme $\det(\varphi \circ \psi) = \det(\varphi)\det(\psi)$ et $K(\varphi \circ \psi) = K(\varphi)K(\psi)$, le lemme sera prouvé par décomposition de φ en un produit de bijections linéaires simples pour lesquelles on vérifie (5.27). Ces applications sont des trois types suivants décrits en donnant l'image de la base canonique $(e_1, \dots, e_d)$ de $\mathbb{R}^d$:

(I) $\bigl(\varphi(e_1), \dots, \varphi(e_d)\bigr)$ est une permutation de $(e_1, \dots, e_d)$;

(II) $\varphi(e_1) = ae_1$ $(a \neq 0)$ et $\varphi(e_i) = e_i$, $(\forall i \geq 2)$;

(III) $\varphi(e_1) = e_1 + e_2$ et $\varphi(e_i) = e_i$, $(\forall i \geq 2)$.

La décomposition de toute bijection linéaire $\mathbb{R}^d \to \mathbb{R}^d$ en un produit d'applications de ce type est un résultat purement algébrique que nous admettrons.

Si φ est du type (I), son déterminant vaut $+1$ ou -1 et comme la mesure de Lebesgue est invariante par permutation des coordonnées, $\lambda\bigl(\varphi(C)\bigr) = \lambda(C) = 1$. Donc $K(\varphi) = 1$ et (5.27) est vérifiée.

Si φ est du type (II), son déterminant vaut a. Si $a > 0$, $\varphi(C) = [0, a] \times [0, 1]^{d-1}$ et $\lambda\bigl(\varphi(C)\bigr) = (a - 0) \times 1^{d-1} = a = |a|$. Si $a < 0$, $\varphi(C) = [a, 0] \times [0, 1]^{d-1}$ et $\lambda\bigl(\varphi(C)\bigr) = (0 - a) \times 1^{d-1} = -a = |a|$. Dans les deux cas $\lambda\bigl(\varphi(C)\bigr) = |a| = |\det \varphi|$, et (5.27) est vérifiée via le lemme 5.23 c).

Si φ est du type (III), son déterminant vaut $+1$ (évident en développant suivant la première colonne). D'autre part, $\varphi(C) = A \times [0, 1]^{d-2}$, où A est le parallélogramme de $\mathbb{R}^2$ de sommets $(0, e_1 + e_2, e_1 + 2e_2, e_2)$. Comme $\lambda_d = \lambda_2 \otimes \lambda_{d-2}$, on en déduit que

$\lambda_d\big(\varphi(C)\big) = \lambda_2(A)\lambda_{d-2}([0,1]^{d-2}) = \lambda_2(A)$. L'aire de A étant clairement la même que celle du carré unité de $\mathbb{R}^2$, on a $\lambda_d\big(\varphi(C)\big) = 1 = \det\varphi$ et (5.27) est vérifiée via le lemme 5.23 c). $\square$

En revenant aux formules (5.24) et (5.25), nous pouvons maintenant énoncer :

Proposition 5.25. *Si φ est une bijection linéaire $\mathbb{R}^d \to \mathbb{R}^d$,*

$$\int_{\mathbb{R}^d} f(x)\,\mathrm{d}\lambda(x) = \int_{\mathbb{R}^d} f\big(\varphi^{-1}(y)\big)|\det(\varphi^{-1})|\,\mathrm{d}\lambda(y) \tag{5.28}$$

$$\int_{\mathbb{R}^d} h\big(\varphi(x)\big)\,\mathrm{d}\lambda(x) = \int_{\mathbb{R}^d} h(y)|\det(\varphi^{-1})|\,\mathrm{d}\lambda(y). \tag{5.29}$$

Ces égalités sont valables pour toutes fonctions mesurables positives f, h. L'égalité (5.28) est vraie pour toute f à valeurs réelles ou complexes, λ-intégrable sur $\mathbb{R}^d$ (ce qui équivaut à la λ-intégrabilité de $f\circ\varphi^{-1}$). De même (5.29) vaut pour toute $h \in \mathcal{L}^1_{\mathbb{K}}(\mathbb{R}^d,\lambda)$ ($\Leftrightarrow h\circ\varphi \in \mathcal{L}^1_{\mathbb{K}}(\mathbb{R}^d,\lambda)$).

Remarque 5.26. La portée des formules (5.28) et (5.29) est plus grande qu'il n'y paraît. Elles permettent de faire un changement de variable linéaire bijectif dans une intégrale sur *n'importe quel borélien* B de $\mathbb{R}^d$. Il suffit pour cela de remplacer f par $f\mathbf{1}_B$ et de remarquer que $\mathbf{1}_B\big(\varphi^{-1}(y)\big) = \mathbf{1}_{\varphi(B)}(y)$. On obtient ainsi

$$\int_B f(x)\,\mathrm{d}\lambda(x) = \int_{\varphi(B)} f\big(\varphi^{-1}(y)\big)|\det(\varphi^{-1})|\,\mathrm{d}\lambda(y). \tag{5.30}$$

Corollaire 5.27. *La mesure de Lebesgue de $\mathbb{R}^d$ est invariante par les isométries euclidiennes de $\mathbb{R}^d$.*

Démonstration. Soit φ une isométrie euclidienne de $\mathbb{R}^d$. Quitte à la composer avec une translation (qui laisse λ_d invariante), on se ramène au cas où $\varphi(0) = 0$. Alors φ est une application linéaire $\mathbb{R}^d \to \mathbb{R}^d$ qui conserve le produit scalaire. Il est bien connu que pour une telle application $|\det(\varphi)| = 1$, d'où la conclusion grâce à (5.26) et (5.27). $\square$

5.4.3 Le théorème de changement de variable C^1

Que peut-on dire lorsqu'on a un changement de variable non linéaire ? Rappelons que dans le cas $d = 1$, nous avons déjà établi au chapitre 4 que si I est un intervalle ouvert (pas nécessairement borné) de $\mathbb{R}$, si $\varphi : I \to \mathbb{R}$ est strictement monotone, de classe C^1 et telle que φ' ne s'annule en aucun point point de I, on a [10]

$$\int_I f(x)\,\mathrm{d}\lambda(x) = \int_{\varphi(I)} f\big(\varphi^{-1}(y)\big)|(\varphi^{-1})'(y)|\,\mathrm{d}\lambda(y).$$

Cette formule se généralise à la dimension d en remplaçant intervalle ouvert par ouvert et $(\varphi^{-1})'(y)$ par le *déterminant jacobien* de φ^{-1}, *i.e.* le déterminant de l'application linéaire tangente à φ^{-1} au point y. Nous admettrons le résultat dont l'énoncé précis est le suivant.

[10] Appliquer (4.40) avec f à la place de $f\circ\varphi$.

Théorème 5.28. *Soit V un ouvert de $\mathbb{R}^d$ et φ un C^1-difféomorphisme de $V \to \varphi(V) \subset \mathbb{R}^d$. Alors les formules de changement de variable équivalentes*

$$\begin{aligned}\int_V f(x)\,\mathrm{d}\lambda(x) &= \int_{\varphi(V)} f\big(\varphi^{-1}(y)\big)|\operatorname{Jac}(\varphi^{-1})(y)|\,\mathrm{d}\lambda(y),\\ \int_V h\big(\varphi(x)\big)\,\mathrm{d}\lambda(x) &= \int_{\varphi(V)} h(y)|\operatorname{Jac}(\varphi^{-1})(y)|\,\mathrm{d}\lambda(y),\end{aligned}$$

sont vérifiées pour toutes f, h mesurables positives ou telles que $f \in \mathcal{L}^1_{\mathbb{K}}(\lambda)$, $h \circ \varphi \in \mathcal{L}^1_{\mathbb{K}}(\lambda)$.

En notant $\psi_1, \dots, \psi_d$ les applications coordonnées de φ^{-1}, on a

$$y = (y_1, \dots, y_d) = \varphi(x) \quad\Leftrightarrow\quad x = (x_1, \dots, x_d) = \big(\psi_1(y), \dots, \psi_d(y)\big)$$

et

$$\operatorname{Jac}(\varphi^{-1})(y) := \det\left[\frac{\partial\psi_i(y)}{\partial y_j}\right]_{1\le i,j\le d}.$$

Voici une notation mnémotechnique basée sur l'analogie formelle avec le changement de variable en dimension 1, que l'on peut utiliser *au brouillon* :

$$\text{« } \mathrm{d}x_1 \dots \mathrm{d}x_d = \frac{\mathrm{D}(x_1, \dots, x_d)}{\mathrm{D}(y_1, \dots, y_d)}\,\mathrm{d}y_1 \dots \mathrm{d}y_d \text{ »} \quad \text{pour} \quad \mathrm{d}\lambda(x) = |\operatorname{Jac}(\varphi^{-1})(y)|\,\mathrm{d}\lambda(y).$$

Avant d'aborder la mise en pratique du théorème 5.28, il n'est sans doute pas superflu de faire le point sur la notion de C^1-difféomorphisme.

5.4.4 Rappels sur les C^1-difféomorphismes

Définition 5.29. *Soient E et F deux ouverts de $\mathbb{R}^d$. L'application $\varphi : E \to F$ est appelée* C^1-difféomorphisme *de E* sur *F si elle vérifie les deux conditions :*

i) φ est une bijection de E sur F.

ii) φ et φ^{-1} sont C^1, c'est-à-dire possèdent des dérivées partielles continues, en tout point de E pour φ et en tout point de F pour φ^{-1}.

Théorème 5.30 (d'inversion locale). *Soit $\varphi : E \to F$ une application C^1 sur E. Soit $a \in E$ tel que $\operatorname{Jac}(\varphi)(a) \neq 0$. Notons $b := \varphi(a) \in F$. Alors il existe un ouvert A de E contenant a et un ouvert B de F contenant b tels que la restriction φ_A de φ à A soit une bijection de A sur B, que φ_A soit C^1 sur A et que φ_A^{-1} soit C^1 sur B (autrement dit que φ_A soit un C^1-difféomorphisme de A sur B).*

Examinons l'exemple suivant. On prend $E = F = \mathbb{R}^2$ et on définit φ par $\varphi(x, y) = (x^2 - y^2, 2xy)$. On voit immédiatement que

$$\operatorname{Jac}(\varphi)(x_0, y_0) = \begin{vmatrix} 2x_0 & -2y_0 \\ 2y_0 & 2x_0 \end{vmatrix} = 4(x_0^2 + y_0^2).$$

Si $(x_0, y_0) \neq (0,0)$, on peut donc inverser localement φ au voisinage de (x_0, y_0). Par ailleurs, on vérifie directement que

$$\varphi(x,y) = \varphi(x',y') \quad \Leftrightarrow \quad (x,y) = (x',y') \text{ ou } (x,y) = (-x',-y').$$

Donc $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ n'est pas injective et il est impossible de l'inverser globalement. En particulier il est impossible de l'inverser sur un domaine contenant les deux points distincts (x_0, y_0) et $(-x_0, -y_0)$. En fait on peut inverser φ sur tout demi-plan dont la frontière passe par le point $(0,0)$. Ce problème est celui de la détermination de la racine carrée complexe puisque si $z = x + iy$, $\varphi(x,y) = \big(\mathrm{Re}(z^2), \mathrm{Im}(z^2)\big)$.

Théorème 5.31 (d'inversion globale). *Si* $\varphi : E \to F$ *est* C^1 *sur* E *et si* pour tout $a \in E$, $\mathrm{Jac}(\varphi)(a) \neq 0$, *alors l'image par* φ *de tout ouvert de* E *est un ouvert de* $\mathbb{R}^d$ *(on dit que* φ *est une application ouverte).*

Si de plus φ *est* injective, *alors* φ *est un* C^1*-difféomorphisme de* E *sur* $\varphi(E)$.

À titre d'exercice, on pourra confronter ce théorème à l'exemple précédent dans les deux cas $E := \mathbb{R}^2 \setminus \{(0,0)\}$ et $E := \{(x,y) \in \mathbb{R}^2;\ y > 0\}$.

Conséquence pratique. Pour vérifier qu'un changement de variable φ est un C^1-difféomorphisme, il *suffit* de vérifier que φ possède des dérivées partielles continues en tout point de l'ouvert de départ, que $\mathrm{Jac}(\varphi)(a) \neq 0$ en tout point a de cet ouvert et que φ est une bijection (en précisant soigneusement les ensembles de départ et d'arrivée !).

5.4.5 Méthode pratique

Pour alléger les écritures, nous nous plaçons dans le cas $d = 2$, mais les conseils donnés ci-dessous sont valables avec d quelconque, modulo le changement de notations. Soit

$$\varphi : (x,y) \longmapsto \varphi(x,y) = (s,t) \tag{5.31}$$

le changement de variable à effectuer dans l'intégrale

$$I := \int_V f(x,y)\,\mathrm{d}\lambda_2(x,y),$$

où V est un ouvert de $\mathbb{R}^2$.

1. On regarde si on peut inverser la relation $(s,t) = \varphi(x,y)$. Cela amène en général à introduire des conditions supplémentaires dans (5.31), en précisant les ensembles de départ E et d'arrivée F, de façon à considérer φ comme une bijection de E sur F. Si $V \subset E$, on passe au point suivant.

   Sinon, on découpe V en une réunion finie (éventuellement infinie dénombrable) d'ouverts V_i *disjoints* (plus éventuellement des ensembles de mesure nulle) tels que la restriction de φ à chaque V_i soit injective. On a alors

$$I = \sum_i \int_{V_i} f(x,y)\,\mathrm{d}\lambda_2(x,y).$$

2. On détermine ensuite le nouvel ensemble d'intégration $\varphi(V)$ (resp. $\varphi(V_i)$) en écrivant :
$$(s,t) \in \varphi(V) \quad \Leftrightarrow \quad (x,y) = \varphi^{-1}(s,t) \in V,$$
ce qui en pratique revient à remplacer dans les inéquations définissant V (resp. V_i) les anciennes variables x et y par leurs expressions en fonction des nouvelles variables s et t .
3. À partir de la relation $(x,y) = \varphi^{-1}(s,t)$, on calcule les dérivées partielles $\frac{\partial x}{\partial s}$, $\frac{\partial x}{\partial t}$, $\frac{\partial y}{\partial s}$ et $\frac{\partial y}{\partial t}$ de φ^{-1} et on vérifie qu'elles sont *continues* sur $\varphi(V)$ (resp. sur $\varphi(V_i)$), donc que φ^{-1} est C^1. On calcule ensuite le jacobien de φ^{-1} et on vérifie qu'il ne s'annule en aucun point de $\varphi(V)$ (resp. $\varphi(V_i)$) [11].

Une fois tout ce travail effectué, on peut appliquer le théorème 5.28 pour conclure :
$$\int_V f(x,y)\,\mathrm{d}\lambda_2(x,y) = \int_{\varphi(V)} f\big(\varphi^{-1}(s,t)\big)\lvert\operatorname{Jac}(\varphi^{-1})(s,t)\rvert\,\mathrm{d}\lambda_2(s,t),$$
(resp. $\int_{V_i} = \int_{\varphi(V_i)} \dots$).

5.4.6 Coordonnées polaires, coordonnées sphériques

Parmi les applications les plus connues du théorème 5.28 figurent le changement de variable par passage des coordonnées cartésiennes aux coordonnées polaires ($d = 2$) ou aux coordonnées sphériques ($d = 3$). Voyons cela plus en détail.

Coordonnées polaires

Le passage des coordonnées cartésiennes aux coordonnées polaires est donné par les formules
$$x = r\cos\theta, \quad y = r\sin\theta. \tag{5.32}$$
Il est commode ici de travailler avec l'application $\psi = \varphi^{-1} : (r,\theta) \mapsto (x,y)$. Cette application réalise une bijection de $]0,+\infty[\times[\alpha,\alpha+2\pi[$ sur $\mathbb{R}^2 \setminus \{(0,0)\}$, où α est un réel quelconque fixé (en pratique on prend le plus souvent $\alpha = -\pi$ ou $\alpha = 0$). Notons ici l'exclusion de tous les couples $(0,\theta)$ de l'ensemble de départ et de leur image commune $(0,0)$ dans l'ensemble d'arrivée : leur conservation eût empêché la bijectivité de ψ (du moins avec un ensemble de départ produit cartésien). Nous réduisons l'ensemble de définition de ψ à $W_\alpha :=]0,+\infty[\times]\alpha,\alpha+2\pi[$ de façon à avoir un *ouvert*. Ceci nous amène à prendre comme ensemble d'arrivée $V_\alpha := \mathbb{R}^2 \setminus D_\alpha$, où D_α est la demi-droite fermée d'origine $(0,0)$ et de vecteur directeur $(\cos\alpha, \sin\alpha)$. Comme $\lambda_2(D_\alpha) = 0$, on pourra

[11] Une variante consiste à calculer les dérivées partielles $\frac{\partial s}{\partial x}$, $\frac{\partial t}{\partial x}\dots$ de φ, vérifier leur continuité sur V, calculer $\operatorname{Jac}(\varphi)(x,y)$ et vérifier qu'il ne s'annule en aucun point de V. On établit ainsi que φ est un C^1 difféomorphisme de V sur $\varphi(V)$ et il ne reste plus qu'à calculer $\operatorname{Jac}(\varphi^{-1})(s,t)$ par la relation $\operatorname{Jac}(\varphi^{-1})(s,t) = 1/\operatorname{Jac}(\varphi)(\varphi^{-1}(s,t))$. Le choix entre ces deux méthodes sera guidé par la commodité du calcul des dérivées partielles. En tout état de cause, la nécessité d'inverser proprement φ reste incontournable.

toujours écrire $\int_{\mathbb{R}^2} f \,\mathrm{d}\lambda_2 = \int_{V_\alpha} f \,\mathrm{d}\lambda_2$ pour f λ_2 intégrable. Finalement, ψ est définie par :

$$\psi : W_\alpha =]0,+\infty[\times]\alpha,\alpha+2\pi[\longrightarrow V_\alpha = \mathbb{R}^2 \setminus D_\alpha, \quad (r,\theta) \longmapsto (x,y) = (r\cos\theta, r\sin\theta).$$

Le calcul des dérivées partielles de ψ est immédiat :

$$\frac{\partial x}{\partial r} = \cos\theta, \quad \frac{\partial y}{\partial r} = \sin\theta, \quad \frac{\partial x}{\partial \theta} = -r\sin\theta, \quad \frac{\partial y}{\partial \theta} = r\cos\theta.$$

Ces dérivées partielles sont continues en tout point de W_α, donc ψ est C^1 sur cet ouvert. D'autre part le jacobien s'écrit

$$\mathrm{Jac}(\psi)(r,\theta) = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r(\cos^2\theta + \sin^2\theta) = r.$$

On voit ainsi que $\mathrm{Jac}(\psi)(r,\theta)$ ne s'annule en aucun point de W_α. Comme ψ est une bijection de W_α sur V_α, le théorème d'inversion globale nous permet d'affirmer que ψ est un C^1-difféomorphisme de W_α sur V_α. Son inverse φ est donc un C^1-difféomorphisme de V_α sur W_α. Nous pouvons maintenant appliquer le théorème de changement de variable 5.28 pour obtenir en rappelant que $\int_{\mathbb{R}^2} f \,\mathrm{d}\lambda_2 = \int_{V_\alpha} f \,\mathrm{d}\lambda_2$,

$$\int_{\mathbb{R}^2} f(x,y) \,\mathrm{d}\lambda_2(x,y) = \int_{]0,+\infty[\times]\alpha,\alpha+2\pi[} f(r\cos\theta, r\sin\theta) r \,\mathrm{d}\lambda_2(r,\theta), \tag{5.33}$$

formule valide pour toute fonction f mesurable positive ou λ_2 intégrable sur $\mathbb{R}^2$.

Exemple 5.1 (Aire d'un disque). Le passage en coordonnées polaires permet de calculer facilement la mesure de Lebesgue d'un disque. En raison de l'invariance par translation, on ne perd pas de généralité en le supposant centré à l'origine. Soit donc $\Delta = \Delta(0,R) := \{(x,y) \in \mathbb{R}^2;\ x^2+y^2 \leq R^2\}$. En appliquant la formule (5.33) à la fonction $f = \mathbf{1}_\Delta$, on obtient :

$$\begin{aligned} \lambda_2(\Delta) = \int_{\mathbb{R}^2} \mathbf{1}_\Delta \,\mathrm{d}\lambda_2 &= \int_{\mathbb{R}^2} \mathbf{1}_{\{x^2+y^2\leq R^2\}} \,\mathrm{d}\lambda_2(x,y) \\ &= \int_{]0,+\infty[\times]0,2\pi[} \mathbf{1}_{\{r\leq R\}} r \,\mathrm{d}\lambda_2(r,\theta) \\ &= \left\{\int_0^R r \,\mathrm{d}r\right\}\left\{\int_{]0,2\pi[} \mathrm{d}\lambda_1(\theta)\right\} \\ &= (2\pi)\frac{R^2}{2}. \end{aligned}$$

Sans surprise, on trouve ainsi $\lambda_2\big(\Delta(0,R)\big) = \pi R^2$.

Exemple 5.2 (Densité gaussienne). Pour tous $\sigma > 0$ et $m \in \mathbb{R}$, les fonctions $x \mapsto (2\pi\sigma)^{-1/2}\exp\big(-(x-m)^2/(2\sigma^2)\big)$ sont des densités de probabilité par rapport à λ_1. La vérification de cette affirmation se ramène après changement de variable au calcul de l'intégrale de Riemann généralisée $I = \int_{-\infty}^{+\infty} \exp(-x^2)\,\mathrm{d}x$. En utilisant le corollaire du théorème de Fubini Tonelli pour les fonctions à variables séparées, on vérifie facilement que :

$$I^2 = \int_{\mathbb{R}^2} \exp\bigl(-(x^2+y^2)\bigr)\,\mathrm{d}\lambda_2(x,y).$$

On calcule ensuite cette intégrale double par passage en coordonnées polaires...et on trouve $I^2 = \pi$, d'où $I = \sqrt{\pi}$. La rédaction détaillée est laissée en exercice (pour un corrigé voir les Annales d'IFP, session de septembre 2002).

Coordonnées sphériques

Le passage des coordonnées cartésiennes de $\mathbb{R}^3$ aux coordonnées sphériques peut être défini par les formules :

$$\begin{cases} x &= r\cos s\cos t, \\ y &= r\sin s\cos t, \\ z &= r\sin t. \end{cases} \tag{5.34}$$

La variable r est la distance à l'origine, s peut s'interpréter comme la longitude et t comme la latitude.

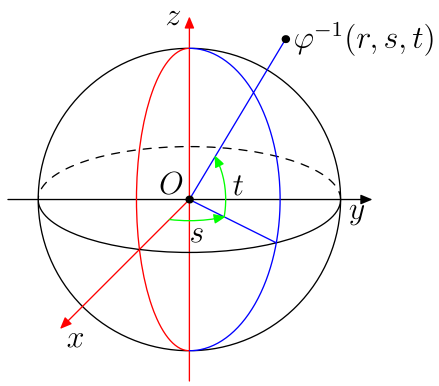

Fig. 5.5 – Coordonnées sphériques

Là aussi il est plus commode de travailler avec φ^{-1}. Pour avoir un C^1 difféomorphisme d'un ouvert de $\mathbb{R}^3$ sur son image, on prive $\mathbb{R}^3$ du demi plan méridien $\{x \geq 0,\ y = 0\}$. Les formules (5.34) définissent une bijection $\varphi^{-1} : W \to V$, où :

$$W :=]0,+\infty[\times]0,2\pi[\times\Big]\frac{-\pi}{2},\frac{\pi}{2}\Big[, \qquad V := \mathbb{R}^3 \setminus \{(x,y,z)\in\mathbb{R}^3;\ x\geq 0,\ y=0\}.$$

Les dérivées partielles de φ^{-1} apparaissent dans le calcul suivant de son jacobien et sont visiblement des fonctions continues de (r, s, t).

$$J := \operatorname{Jac}(\varphi^{-1})(r,s,t) = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial y}{\partial r} & \frac{\partial z}{\partial r} \\ \frac{\partial x}{\partial s} & \frac{\partial y}{\partial s} & \frac{\partial z}{\partial s} \\ \frac{\partial x}{\partial t} & \frac{\partial y}{\partial t} & \frac{\partial z}{\partial t} \end{vmatrix} = \begin{vmatrix} \cos s\cos t & \sin s\cos t & \sin t \\ -r\sin s\cos t & r\cos s\cos t & 0 \\ -r\cos s\sin t & -r\sin s\sin t & r\cos t \end{vmatrix}$$

En développant suivant la troisième colonne, on trouve

$$\begin{aligned} J &= \sin t \begin{vmatrix} -r\sin s\cos t & r\cos s\cos t \\ -r\cos s\sin t & -r\sin s\sin t \end{vmatrix} + r\cos t \begin{vmatrix} \cos s\cos t & \sin s\cos t \\ -r\sin s\cos t & r\cos s\cos t \end{vmatrix} \\ &= r^2\sin^2 t\cos t \begin{vmatrix} -\sin s & \cos s \\ -\cos s & -\sin s \end{vmatrix} + r^2\cos^3 t \begin{vmatrix} \cos s & \sin s \\ -\sin s & \cos s \end{vmatrix} \\ &= r^2(\sin^2 t + \cos^2 t)\cos t = r^2\cos t. \end{aligned}$$

On remarque que J ne s'annule en aucun point de W. Par le théorème d'inversion globale, φ^{-1} (et donc aussi φ) est un C^1 difféomorphisme. Ainsi en notant que $\lambda_3(\mathbb{R}^3 \setminus V) = 0$ (le demi-plan méridien exclu est inclus dans un plan qui est un sous-espace affine de $\mathbb{R}^3$ de mesure λ_3 nulle), on peut écrire la formule de changement de variable

$$\int_{\mathbb{R}^3} f \,\mathrm{d}\lambda_3 = \int_{]0,+\infty[\times]0,2\pi[\times\left]\frac{-\pi}{2},\frac{\pi}{2}\right[} f(r\cos s\cos t, r\sin s\cos t, r\sin t) r^2\cos t \,\mathrm{d}\lambda_3(r,s,t).$$

Cette formule est valide pour toute fonction f mesurable positive ou λ_3 intégrable.

Exemple 5.3 (Volume d'une boule euclidienne). En prenant dans la formule ci-dessus $f = \mathbf{1}_B$, où $B = B(0,R) := \{(x,y,z) \in \mathbb{R}^3;\ x^2+y^2+z^2 \leq R^2\}$ est la boule euclidienne de centre 0 et de rayon R, on obtient :

$$\begin{aligned} \lambda_3(B) &= \int_{]0,R[\times]0,2\pi[\times]-\pi/2,\pi/2[} r^2\cos t \,\mathrm{d}\lambda_3(r,s,t) \\ &= \left\{\int_{]0,R[} r^2 \,\mathrm{d}\lambda_1(r)\right\}\left\{\int_{]0,2\pi[} \mathrm{d}\lambda_1(s)\right\}\left\{\int_{]-\pi/2,\pi/2[} \cos t \,\mathrm{d}\lambda_1(t)\right\} \\ &= 2\pi\left\{\int_0^R r^2\,\mathrm{d}r\right\}\left\{\int_{-\pi/2}^{\pi/2} \cos t\,\mathrm{d}t\right\} \\ &= 2\pi \times \frac{R^3}{3} \times 2. \end{aligned}$$

On vérifie ainsi que le volume d'une boule euclidienne de rayon R est

$$\lambda_3\big(B(0,R)\big) = \frac{4}{3}\pi R^3.$$

5.4.7 Calculs de lois par changement de variable dans $\mathbb{R}^d$

Soit $X = (X_1, \dots, X_d)$ un vecteur aléatoire dans $\mathbb{R}^d$ ayant une densité f par rapport à la mesure de Lebesgue λ. On cherche la loi P_Y de $Y = \varphi(X)$ où l'on suppose que φ est un C^1-difféomorphisme d'un ouvert V sur $\varphi(V) \subset \mathbb{R}^d$ et que $P_X(\mathbb{R}^d \setminus V) = 0$.

Pour ce faire, on calcule de deux façons $\mathbf{E}h(Y)$ où $h : \mathbb{R}^d \to \mathbb{R}_+$ est une fonction mesurable positive quelconque.

Première façon : par transfert $\Omega \xrightarrow{Y} \mathbb{R}^d$.

$$\mathbf{E}h(Y) = \int_\Omega h(Y)\,\mathrm{d}\mathbf{P} = \int_{\mathbb{R}^d} h(y)\,\mathrm{d}P_Y(y). \tag{5.35}$$

Deuxième façon : par transfert $\Omega \xrightarrow{X} \mathbb{R}^d$.

$$\mathbf{E}h(Y) = \int_\Omega h(\varphi(X))\,\mathrm{d}\mathbf{P} = \int_{\mathbb{R}^d} h(\varphi(x)\,\mathrm{d}P_X(x) = \int_{\mathbb{R}^d} h(\varphi(x)f(x)\,\mathrm{d}\lambda(x). \tag{5.36}$$

Comme nous avons supposé que $P_X(\mathbb{R}^d \setminus V) = 0$, la densité f de X est nulle λ-p.p. sur le complémentaire de l'ouvert V d'où

$$\mathbf{E}h(Y) = \int_V h(\varphi(x)f(x)\,\mathrm{d}\lambda(x). \tag{5.37}$$

Les hypothèses du théorème 5.28 étant satisfaites, la formule de changement de variable nous donne

$$\mathbf{E}h(Y) = \int_{\varphi(V)} h(y) f\big(\varphi^{-1}(y)\big)|\operatorname{Jac}(\varphi^{-1})(y)|\,\mathrm{d}\lambda(y). \tag{5.38}$$

En notant μ la mesure sur $(\mathbb{R}^d, \operatorname{Bor}(\mathbb{R}^d))$ de densité par rapport à λ

$$g(y) := f\big(\varphi^{-1}(y)\big)|\operatorname{Jac}(\varphi^{-1})(y)|\mathbf{1}_{\varphi(V)}(y),$$

(5.38) s'écrit $\mathbf{E}h(Y) = \int_{\mathbb{R}^d} h\,\mathrm{d}\mu$. En comparant avec (5.35), nous voyons que pour toute fonction h mesurable positive, $\int_{\mathbb{R}^d} h\,\mathrm{d}P_Y = \int_{\mathbb{R}^d} h\,\mathrm{d}\mu$. Il en résulte immédiatement que $\mu = P_Y$ (prendre $h = \mathbf{1}_B$, B borélien quelconque). En conclusion la loi de Y est la mesure à densité g par rapport à λ.

Exemple 5.4. Soit (X_1, X_2) un vecteur aléatoire de densité

$$f(s,t) = \frac{1}{2\pi}\exp\Big(-\frac{s^2+t^2}{2}\Big).$$

On définit la variable aléatoire réelle

$$Y_1 := \begin{cases} X_1/X_2 & \text{si } X_2 \neq 0, \\ 0 & \text{si } X_2 = 0. \end{cases}$$

Quelle est la loi de Y_1 ?

Solution. Il est facile de vérifier que la fonction continue (donc mesurable) positive f est bien une densité, soit en remarquant que $f(s,t) = (g \otimes g)(s,t)$, où g est la densité gaussienne standard [12], soit en calculant $\int_{\mathbb{R}^2} f \, \mathrm{d}\lambda_2$ par passage en coordonnées polaires. Y_1 est bien une variable aléatoire comme fonction mesurable d'un vecteur aléatoire.

Pour pouvoir exploiter le théorème de changement de variable, on complète Y_1 en un vecteur aléatoire (Y_1, Y_2) de même dimension que (X_1, X_2) par adjonction de la variable aléatoire $Y_2 := X_2$. On va chercher la loi de (Y_1, Y_2) par la méthode du changement de variable. La loi de Y_1, première loi marginale, s'obtiendra alors par intégration partielle de la densité du vecteur. On a ainsi $(Y_1, Y_2) = \varphi(X_1, X_2)$, où φ est définie sur $\mathbb{R}^2$ par

$$(u,v) = \varphi(s,t) = \begin{cases} (s/t, t) & \text{si } t \neq 0, \\ (0,t) & \text{si } t = 0. \end{cases}$$

On vérifie facilement que φ restreinte à l'ouvert $W := \{(s,t) \in \mathbb{R}^2;\ t \neq 0\}$ est une bijection de cet ouvert sur lui même. Son inverse s'écrit

$$(s,t) = \varphi^{-1}(u,v) = (uv, v), \quad (u,v) \in \varphi(W) = W.$$

Le déterminant jacobien de φ^{-1} est donc

$$\operatorname{Jac}(\varphi^{-1})(u,v) = \begin{vmatrix} v & 0 \\ u & 1 \end{vmatrix} = v.$$

Les 4 dérivées partielles de φ^{-1} sont des fonctions continues de (u,v) et le jacobien ne s'annule en aucun point de W, on a donc bien un C^1-difféomorphisme de W sur W. Remarquons aussi que $\mathbf{P}\big((X_1, X_2) \notin W\big) = \mathbf{P}(X_2 = 0) = 0$ parce que X_2 est une variable aléatoire à densité. On aura donc pour toute fonction g mesurable positive,

$$\int_{\mathbb{R}^2} g \, \mathrm{d}P_{(X_1,X_2)} = \int_W g \, \mathrm{d}P_{(X_1,X_2)}. \tag{5.39}$$

Soit h borélienne $\mathbb{R}^2 \to \mathbb{R}_+$, calculons de deux façons $\mathbf{E}h(Y_1, Y_2)$. D'abord par transfert $\Omega \xrightarrow{(Y_1,Y_2)} \mathbb{R}^2$,

$$\mathbf{E}h(Y_1, Y_2) = \int_{\mathbb{R}^2} h(u,v) \, \mathrm{d}P_{(Y_1,Y_2)}(u,v). \tag{5.40}$$

D'autre part comme $\mathbf{E}h(Y_1, Y_2) = \mathbf{E}h\big(\varphi(X_1, X_2)\big)$, le transfert $\Omega \xrightarrow{(X_1,X_2)} \mathbb{R}^2$, (5.39) et l'expression de la densité f de (X_1, X_2) nous donnent

$$\begin{aligned} \mathbf{E}h(Y_1, Y_2) &= \int_{\mathbb{R}^2} h\big(\varphi(s,t)\big) \, \mathrm{d}P_{(X_1,X_2)}(s,t) \\ &= \int_W h\big(\varphi(s,t)\big) \, \mathrm{d}P_{(X_1,X_2)}(s,t) \\ &= \int_W h\big(\varphi(s,t)\big) \frac{1}{2\pi} \exp\Big(-\frac{s^2+t^2}{2}\Big) \, \mathrm{d}\lambda_2(s,t). \end{aligned} \tag{5.41}$$

[12] En anticipant légèrement sur la section suivante, on voit ainsi que X_1 et X_2 sont deux variables aléatoires gaussiennes indépendantes.

Le changement de variable φ est un C^1 difféomorphisme de l'ouvert W sur lui même. Appliqué à l'intégrale (5.41), il conduit à

$$\mathbf{E}h(Y_1, Y_2) = \int_W h(u,v) \frac{1}{2\pi} \exp\Bigl(-\frac{1}{2}(u^2+1)v^2\Bigr) |v| \,\mathrm{d}\lambda_2(u,v). \tag{5.42}$$

Les égalités (5.40) et (5.42) étant valables pour toute h mesurable positive, leur comparaison montre que la loi $P_{(Y_1,Y_2)}$ du vecteur aléatoire (Y_1, Y_2) est la mesure à densité par rapport à λ_2 :

$$(u,v) \longmapsto \frac{1}{2\pi} \exp\Bigl(-\frac{1}{2}(u^2+1)v^2\Bigr) |v| \mathbf{1}_W(u,v).$$

La densité f_{Y_1} de la loi marginale P_{Y_1} s'en déduit par application de la proposition 5.21 :

$$f_{Y_1}(u) = \int_{\mathbb{R}} \frac{1}{2\pi} \exp\Bigl(-\frac{1}{2}(u^2+1)v^2\Bigr) |v| \mathbf{1}_W(u,v) \,\mathrm{d}\lambda_1(v) \tag{5.43}$$

En se souvenant que $W = \{(u,v) \in \mathbb{R}^2;\ v \neq 0\}$, on voit que $\mathbf{1}_W(u,v) = \mathbf{1}_{\mathbb{R}^*}(v)$ (quel que soit $u \in \mathbb{R}$). On calcule alors l'intégrale (5.43) ainsi (les justifications suivent) :

$$\begin{aligned}
f_{Y_1}(u) &= \int_{]-\infty,0[} \frac{1}{2\pi} \exp\Bigl(-\frac{1}{2}(u^2+1)v^2\Bigr)(-v) \,\mathrm{d}\lambda_1(v) \\
&\quad + \int_{]0,+\infty[} \frac{1}{2\pi} \exp\Bigl(-\frac{1}{2}(u^2+1)v^2\Bigr) v \,\mathrm{d}\lambda_1(v) && (5.44) \\
&= -\int_{-\infty}^{0} \frac{1}{2\pi} \exp\Bigl(-\frac{1}{2}(u^2+1)v^2\Bigr) v \,\mathrm{d}v + \int_0^{+\infty} \frac{1}{2\pi} \exp\Bigl(-\frac{1}{2}(u^2+1)v^2\Bigr) v \,\mathrm{d}v \\
&= \frac{1}{\pi} \int_0^{+\infty} \exp\Bigl(-\frac{1}{2}(u^2+1)v^2\Bigr) v \,\mathrm{d}v && (5.45) \\
&= \frac{1}{\pi} \int_0^{+\infty} \exp(-z) \frac{1}{1+u^2} \,\mathrm{d}z && (5.46) \\
&= \frac{1}{\pi(1+u^2)} && (5.47)
\end{aligned}$$

Justifications :

1. Découpage (5.44) : $\{0\}$ est de λ_1 mesure nulle.
2. De (5.44) à (5.45) : la fonction à intégrer est continue et a une intégrale de Riemann généralisée absolument convergente sur chacun des intervalles $]-\infty, 0[$ et $]0, +\infty[$. Ceci justifie la conversion des intégrales de Lebesgue en intégrales de Riemann. La réduction à une seule intégrale de Riemann s'obtient par le changement de variable $v \mapsto -v$.
3. De (5.45) à (5.46) : changement de variable $z = \frac{1+u^2}{2} v^2$. Bien noter qu'ici $\frac{1+u^2}{2}$ est une constante.

Conclusion : l'égalité (5.47) montre que $Y_1 = X_1/X_2$ suit la loi de Cauchy. □

5.5 Indépendance

L'indépendance d'événements et de variables aléatoires discrètes a été étudiée [13] en DEUG. La notion de tribu produit et de mesure produit permet de généraliser et de systématiser cette étude. Dans toute cette section, on travaille avec un même espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$ sur lequel seront définis les variables et vecteurs aléatoires considérés. Nous réservons l'appelation « évènement » aux ensembles membres de la tribu $\mathcal{F}$.

5.5.1 Indépendance d'évènements

Définition 5.32. *Deux évènements A et B (i.e. $A \in \mathcal{F}$ et $B \in \mathcal{F}$) sont indépendants si*

$$\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B).$$

Remarques 5.33.

a) Si A est un événement tel que $P(A) = 0$ ou $P(A) = 1$, alors il est indépendant de tout événement, *y compris de lui même* (c'est le cas en particulier pour Ω et $\emptyset$).

b) Deux événements *incompatibles* A et B avec $P(A) > 0$ et $P(B) > 0$ ne sont *jamais indépendants*. En effet $A \cap B = \emptyset$ implique $P(A \cap B) = 0$ or $P(A)P(B) \neq 0$.

c) L'indépendance de deux événements A et B n'est pas une propriété intrinsèque aux événements, elle est toujours relative à l'espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$.

d) Si $\mathbf{P}(B) \neq 0$, l'indépendance de A et B équivaut à $\mathbf{P}(A \mid B) = \mathbf{P}(A)$. Ceci exprime bien l'idée intuitive d'indépendance : la connaissance de la réalisation de B ne modifie pas notre degré d'incertitude sur celle de A.

Définition 5.34. *Soit I un ensemble quelconque d'indices.*

a) Les évènements d'une famille $(A_i)_{i\in I}$ sont mutuellement indépendants si

$$\forall J \text{ fini } \subset I, \qquad \mathbf{P}\left(\bigcap_{j\in J} A_j\right) = \prod_{j\in J} \mathbf{P}(A_j).$$

b) Les classes d'évènements $(\mathcal{C}_i)_{i\in I}$ (i.e. $\forall i \in I$, $\mathcal{C}_i \subset \mathcal{F}$) sont mutuellement indépendantes si pour tout choix d'un A_i dans chaque $\mathcal{C}_i$, les $(A_i)_{i\in I}$ sont mutuellement indépendants au sens du a).

L'indépendance mutuelle est plus forte que l'indépendance deux à deux. Les évènements A, B, C sont indépendants deux à deux s'ils vérifient

$$\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B), \quad \mathbf{P}(B \cap C) = \mathbf{P}(B)\mathbf{P}(C), \quad \mathbf{P}(C \cap A) = \mathbf{P}(C)\mathbf{P}(A).$$

Pour être mutuellement indépendants, ils devraient vérifier *en plus* la relation

$$\mathbf{P}(A \cap B \cap C) = \mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(C),$$

laquelle ne peut se déduire des trois précédentes. Dans toute la suite, sauf mention explicite du contraire, nous allègerons « mutuellement indépendant(e)s » en « indépendant(e)s ».

[13] Il est conseillé de (re)lire les sections 2.2, 4.3 et 5.4 de [ICP].

Proposition 5.35. *Si $\mathcal{C}_1,\dots,\mathcal{C}_n$ sont des π-classes indépendantes d'évènements, alors les tribus engendrées $\sigma(\mathcal{C}_1),\dots,\sigma(\mathcal{C}_n)$ sont indépendantes.*

Démonstration. Nous nous contenterons de faire la preuve pour $n = 2$. Fixons $A_1 \in \mathcal{C}_1$ et définissons

$$\mathcal{E}_{A_1} := \{A_2 \in \sigma(\mathcal{C}_2);\ \mathbf{P}(A_1 \cap A_2) = \mathbf{P}(A_1)\mathbf{P}(A_2)\}.$$

Montrons que $\mathcal{E}_{A_1}$ est une λ-classe. Comme Ω est indépendant de tout événement, $\Omega \in \mathcal{E}_{A_1}$. Pour vérifier la stabilité par union dénombrable croissante, soit $(A_{2,n})_{n\geq 1}$ une suite croissante pour l'inclusion dans $\mathcal{E}_{A_1}$. Sa limite A_2, c'est-à-dire l'union des $A_{2,n}$, est dans $\sigma(\mathcal{C}_2)$ et il s'agit de prouver l'indépendance de A_1 et A_2. La convergence ensembliste $A_{2,n} \uparrow A_2$ implique $A_1 \cap A_{2,n} \uparrow A_1 \cap A_2$ et toutes deux donnent par continuité croissante de $\mathbf{P}$

$$\mathbf{P}(A_{2,n}) \uparrow \mathbf{P}(A_2) \quad \text{et} \quad \mathbf{P}(A_1 \cap A_{2,n}) \uparrow \mathbf{P}(A_1 \cap A_2).$$

On en déduit l'indépendance de A_1 et A_2 en passant à la limite dans les égalités

$$\mathbf{P}(A_1 \cap A_{2,n}) = \mathbf{P}(A_1)\mathbf{P}(A_{2,n})$$

qui traduisent l'appartenance des $A_{2,n}$ à $\mathcal{E}_{A_1}$. Pour vérifier la stabilité par différence propre, soient A_2 et B_2 dans $\mathcal{E}_{A_1}$ tels que $A_2 \subset B_2$. On peut alors écrire

$$\begin{aligned}
\mathbf{P}\big(A_1 \cap (B_2 \setminus A_2)\big) &= \mathbf{P}\big((A_1 \cap B_2) \setminus (A_1 \cap A_2)\big) \\
&= \mathbf{P}(A_1 \cap B_2) - \mathbf{P}(A_1 \cap A_2) \\
&= \mathbf{P}(A_1)\mathbf{P}(B_2) - \mathbf{P}(A_1)\mathbf{P}(A_2) \\
&= \mathbf{P}(A_1)\big(\mathbf{P}(B_2) - \mathbf{P}(A_2)\big) \\
&= \mathbf{P}(A_1)\mathbf{P}(B_2 \setminus A_2).
\end{aligned}$$

Donc $B_2 \setminus A_2$ est dans $\mathcal{E}_{A_1}$. Ainsi $\mathcal{E}_{A_1}$ est une λ-classe qui contient $\mathcal{C}_2$. Par le théorème de Dynkin, elle contient aussi la tribu $\sigma(\mathcal{C}_2)$, donc $\mathcal{E}_{A_1} = \sigma(\mathcal{C}_2)$. Ceci étant vrai pour tout $A_1 \in \mathcal{C}_1$, nous avons établi que

$$\forall A_1 \in \mathcal{C}_1, \forall A_2 \in \sigma(\mathcal{C}_2), \quad \mathbf{P}(A_1 \cap A_2) = \mathbf{P}(A_1)\mathbf{P}(A_2).$$

Fixons $A_2 \in \sigma(\mathcal{C}_2)$ et définissons

$$\mathcal{G}_{A_2} := \{A_1 \in \sigma(\mathcal{C}_1);\ \mathbf{P}(A_1 \cap A_2) = \mathbf{P}(A_1)\mathbf{P}(A_2)\}.$$

En procédant comme ci-dessus, on montre que $\mathcal{G}_{A_2}$ est un λ-système qui contient $\mathcal{C}_1$, on montre que $\mathcal{G}_{A_2} = \sigma(\mathcal{C}_1)$. Ceci étant vrai pour tout $A_2 \in \sigma(\mathcal{C}_2)$, on en déduit que $\mathbf{P}(A_1 \cap A_2) = \mathbf{P}(A_1)\mathbf{P}(A_2)$ pour tous $A_1 \in \sigma(\mathcal{C}_1)$ et $A_2 \in \sigma(\mathcal{C}_2)$. L'indépendance des tribus $\sigma(\mathcal{C}_1)$ et $\sigma(\mathcal{C}_2)$ est établie. □

Pour une application élémentaire de la proposition 5.35, examinons le cas où les $\mathcal{C}_i$ sont réduites à un seul évènement A_i (ce sont alors trivialement des π classes). Les tribus engendrées sont alors les $\sigma(\mathcal{C}_i) = \{A_i, A_i^c, \Omega, \emptyset\}$. On en déduit le corollaire suivant (que l'on pourra aussi démontrer directement à titre d'exercice ou trouver dans [ICP, Prop. 2.12]).

Corollaire 5.36. *Si $A_1, \dots, A_n$ est une suite finie d'évènements indépendants, alors toute suite $B_1, \dots, B_n$ telle que pour chaque i, $B_i = A_i$ ou $B_i = A_i^c$ est encore une suite d'évènements indépendants. La même propriété reste valable pour les suites infinies d'évènements indépendants.*

Voici maintenant une application de l'indépendance d'une suite d'évènements connue sous le nom de *deuxième lemme de Borel Cantelli.* Rappelons que si $(A_n)_{n\in\mathbb{N}}$ est une suite quelconque d'évènements,

$$\limsup_{n\to\infty} A_n = \bigcap_{n\in\mathbb{N}} \bigcup_{k\geq n} A_k = \{\omega \in \Omega;\ \omega \text{ réalise un infinité de } A_k\}$$

Lemme 5.37 (Borel Cantelli II). *Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'évènements* indépendants *telle que*

$$\sum_{n=0}^{+\infty} \mathbf{P}(A_n) = +\infty. \tag{5.48}$$

Alors

$$\mathbf{P}\big(\limsup_{n\to\infty} A_n\big) = 1,$$

autrement dit, presque sûrement une infinité d'évènements A_k se réalisent.

Démonstration. Posons

$$B_{n,m} := \bigcup_{n\leq k\leq m} A_k, \qquad B_n := \bigcup_{k\geq n} A_k.$$

Par le corollaire 5.36, les A_k^c sont indépendants, d'où

$$\mathbf{P}(B_{n,m}) = 1 - \mathbf{P}(B_{n,m}^c) = 1 - \mathbf{P}\Big(\bigcap_{k=n}^{m} A_k^c\Big) = 1 - \prod_{k=n}^{m}\big(1 - \mathbf{P}(A_k)\big).$$

On utilise alors l'inégalité de convexité[14] $\mathrm{e}^{-x} \geq 1 - x$ avec $x = \mathbf{P}(A_k)$ pour obtenir la minoration

$$1 \geq \mathbf{P}(B_{n,m}) \geq 1 - \prod_{k=n}^{m} \exp\big(-\mathbf{P}(A_k)\big) = 1 - \exp\Big(-\sum_{k=n}^{m} \mathbf{P}(A_k)\Big). \tag{5.49}$$

En laissant n fixe et faisant tendre m vers l'infini dans (5.49), on en déduit grâce à l'hypothèse (5.48) que $\mathbf{P}(B_{n,m})$ tend vers 1. D'autre part B_n est limite croissante pour l'inclusion des $B_{n,m}$, donc par continuité croissante séquentielle de $\mathbf{P}$,

$$\mathbf{P}(B_n) = \lim_{m\to+\infty} \mathbf{P}(B_{n,m}) = 1.$$

[14] La représentation graphique de la fonction convexe $x \mapsto \mathrm{e}^{-x}$ est toujours au-dessus de sa tangente à l'origine d'où $\mathrm{e}^{-x} \geq 1 - x$ pour tout $x \in \mathbb{R}$.

cette égalité étant vraie pour tout $n \in \mathbb{N}$, on en déduit que

$$\mathbf{P}\Big(\bigcap_{n\in\mathbb{N}} B_n\Big) = 1,$$

puisque

$$\mathbf{P}\Big(\bigcup_{n\in\mathbb{N}} B_n^c\Big) \leq \sum_{k\in\mathbb{N}} \mathbf{P}(B_n^c) = 0.$$

Comme l'intersection de tous les B_n est l'évènement $\limsup A_n$, le lemme est démontré. □

5.5.2 Indépendance de variables aléatoires

Nous définissons maintenant l'indépendance d'une famille quelconque $(X_i)_{i\in I}$ de variables ou vecteurs aléatoires. Cette indépendance est celle de la famille de sous-tribus de $\mathcal{F}$ engendrées par les X_i. Rapplons que si X est une application $(\Omega, \mathcal{F}) \to (E, \mathcal{B})$, la tribu engendrée par X est

$$\sigma(X) := X^{-1}(\mathcal{B}) = \{X^{-1}(B);\ B \in \mathcal{B}\}.$$

Si de plus, X est $\mathcal{F}$ - $\mathcal{B}$ mesurable, alors $\sigma(X)$ est une sous-tribu de $\mathcal{F}$. On peut dire alors que X est une variable aléatoire à valeurs dans l'espace mesurable $(E, \mathcal{B})$. En pratique, E peut être $\overline{\mathbb{R}}_+$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{N}$, $\mathbb{R}^d$, …et la tribu $\mathcal{B}$ est la tribu borélienne correspondante ou la tribu $\mathcal{P}(E)$ lorsque E est dénombrable.

Définition 5.38. *Soit I une famille quelconque d'indices et pour tout $i \in I$, X_i une variable aléatoire à valeurs dans l'espace mesurable $(E_i, \mathcal{B}_i)$. On dit que $(X_i)_{i\in I}$ est une famille de variables aléatoires indépendantes si la famille de tribus $(\sigma(X_i))_{i\in I}$ est indépendante, autrement dit si*

$$\forall J \textit{ fini } \subset I, \quad \forall j \in J, \forall B_j \in \mathcal{B}_j, \quad \mathbf{P}(\forall j \in J,\ X_j \in B_j) = \prod_{j\in J} \mathbf{P}(X_j \in B_j).$$

Notons que cette définition est assez souple pour englober l'indépendance d'une collection complètement hétéroclite de « variables aléatoires », certaines pouvant être des variables aléatoires réelles, d'autres des vecteurs aléatoires (de dimensions diverses), d'autres des variables aléatoires discrètes, …

Proposition 5.39. *Un vecteur aléatoire $(X_1, \dots, X_d)$ à valeurs dans $\mathbb{R}^d$ est à composantes indépendantes si et seulement si sa loi est le produit de ses lois marginales, i.e.*

$$P_{(X_1,\dots,X_d)} = P_{X_1} \otimes \cdots \otimes P_{X_d}.$$

Démonstration. Supposons d'abord que les variables aléatoires réelles $X_1, \dots, X_d$ soient indépendantes et soit B un pavé mesurable pour la tribu produit $\mathrm{Bor}(\mathbb{R}^1)^{\otimes d}$, B s'écrit donc $B = B_1 \times \cdots \times B_d$, avec pour chaque $i = 1, \dots d$, $B_i \in \operatorname{Bor}(\mathbb{R}^1)$. On a alors

$$\begin{aligned}
P_{(X_1,\dots,X_d)}(B) &= \mathbf{P}\big((X_1,\dots,X_d) \in B_1 \times \cdots \times B_d\big) && (5.50)\\
&= \mathbf{P}\big(\forall i \in \{1,\dots,d\}, X_i \in B_i\big) && (5.51)\\
&= \mathbf{P}\left(\bigcap_{i=1}^{d} X_i^{-1}(B_i)\right) && (5.52)\\
&= \mathbf{P}\big(X_1^{-1}(B_1)\big)\dots\mathbf{P}\big(X_d^{-1}(B_d)\big) && (5.53)\\
&= P_{X_1}(B_1)\dots P_{X_d}(B_d) && (5.54)\\
&= \big(P_{X_1} \otimes \cdots \otimes P_{X_d}\big)(B). && (5.55)
\end{aligned}$$

Justifications : (5.50) vient de la définition de la loi du vecteur aléatoire ; (5.51) et (5.52) utilisent la définition d'un produit cartésien et celle d'une intersection. Le passage de (5.52) à (5.53) repose sur *l'hypothèse d'indépendance* des X_i. La définition de la loi de X_i nous fait passer de (5.53) à (5.54). De là on arrive à (5.55) par la définition de la mesure produit sur la classe des pavés mesurables.

Nous venons ainsi de vérifier que les deux mesures $P_{(X_1,\dots,X_d)}$ et $P_{X_1} \otimes \cdots \otimes P_{X_d}$ coïncident sur la classe des pavés mesurables de la tribu produit $\operatorname{Bor}(\mathbb{R}^1)^{\otimes d}$. Par unicité de la mesure produit, ces deux mesures coïncident sur toute la tribu $\operatorname{Bor}(\mathbb{R}^1)^{\otimes d}$. Rappelons que cette tribu n'est autre que $\operatorname{Bor}(\mathbb{R}^d)$, d'après (5.20) et la proposition 5.16.

Réciproquement, supposons l'égalité des deux mesures $P_{(X_1,\dots,X_d)}$ et $P_{X_1} \otimes \cdots \otimes P_{X_d}$ sur la tribu borélienne de $\mathbb{R}^d$. Elles coïncident en particulier sur la classe des pavés mesurables de $\operatorname{Bor}(\mathbb{R}^1)^{\otimes d}$. On a donc pour tout $B = B_1 \times \cdots \times B_d$ avec $B_i \in \operatorname{Bor}(\mathbb{R})$,

$$P_{(X_1,\dots,X_d)}(B) = \big(P_{X_1} \otimes \cdots \otimes P_{X_d}\big)(B). \tag{5.56}$$

La justification détaillée ci-dessus des égalités (5.50) à (5.55) montre que seul le passage de (5.52) à (5.53) utilisait l'indépendance des X_i. Nous pouvons donc recycler toutes les autres égalités de la manière suivante. On a égalité des seconds membres de (5.52), (5.51) et (5.50). Les égalités (5.50) et (5.56) nous conduisent au second membre de (5.55) d'où l'on peut remonter jusqu'à celui de (5.53). Ainsi

$$\forall i \in \{1,\dots,d\},\ \forall B_i \in \operatorname{Bor}(\mathbb{R}), \quad \mathbf{P}\left(\bigcap_{i=1}^{d} X_i^{-1}(B_i)\right) = \mathbf{P}\big(X_1^{-1}(B_1)\big)\dots\mathbf{P}\big(X_d^{-1}(B_d)\big).$$

Soit J une partie quelconque de $\{1,\dots,d\}$ (automatiquement finie) et choisissons $B_i = \mathbb{R}$ lorsque $i \notin J$ dans l'égalité ci-dessus. Pour ces indices, $X_i^{-1}(B_i) = X_i^{-1}(\mathbb{R}) = \Omega$ et $\mathbf{P}\big(X_i^{-1}(B_i)\big) = 1$. On peut donc effacer sans inconvénient les évènements correspondants et leurs probabilités dans l'égalité ci-dessus. On obtient alors, avec des boréliens B_j, $(j \in J)$ quelconques :

$$\mathbf{P}\left(\bigcap_{j\in J} X_j^{-1}(B_j)\right) = \prod_{j\in J} \mathbf{P}\big(X_j^{-1}(B_j)\big),$$

ce qui prouve l'indépendance des variables aléatoires réelles $X_1,\dots,X_d$. $\square$

Corollaire 5.40. *Soit $X_1, \dots, X_n$ une suite finie de variables aléatoires réelles indépendantes. Alors les vecteurs aléatoires obtenus en découpant dans cette suite des blocs consécutifs disjoints sont indépendants.*

Corollaire 5.41. *Soit $X_1, \dots, X_n$ une suite finie de variables aléatoires réelles indépendantes. Pour tout découpage $n_0 = 0$, $1 \leq n_1 < n_2 < \cdots < n_k = n$ en blocs $Y_j = (X_{1+n_{j-1}}, \dots, X_{n_j})$ $(1 \leq j \leq k)$ et toutes fonctions mesurables $h_j = \mathbb{R}^{n_j - n_{j-1}} \to (E_j, \mathcal{B}_j)$, les variables aléatoires $h_1(Y_1), \dots, h_k(Y_k)$ sont indépendantes.*

Preuve des corollaires 5.40 et 5.41. Le corollaire 5.40 est à l'évidence un cas particulier du corollaire 5.41 en prenant pour h_j l'identité sur $\mathbb{R}^{n_j - n_{j-1}}$. Notons à ce propos que la collection de « variables aléatoires » $Z_j := h_j(Y_j)$ peut être hétéroclite : vecteurs aléatoires dans $\mathbb{R}^{d_j}$, variables aléatoires réelles, complexes ou discrètes. L'indépendance des Y_j, $(1 \leq j \leq k)$ résulte facilement de la proposition 5.39 grâce à l'associativité du produit des tribus et des mesures. Pour le corollaire 5.41, il s'agit de vérifier l'indépendance des tribus $\sigma(Z_j)$ $(1 \leq j \leq k)$. Or ces tribus s'écrivent $\sigma(Z_j) = Y_j^{-1}\bigl(h_j^{-1}(\mathcal{B}_j)\bigr)$. La mesurabilité de h_j se traduit par l'inclusion de tribus $h_j^{-1}(\mathcal{B}_j) \subset \operatorname{Bor}(\mathbb{R}^{n_j - n_{j-1}})$. On a ainsi

$$\sigma(Z_j) = Y_j^{-1}\bigl(h_j^{-1}(\mathcal{B}_j)\bigr) \subset Y_j^{-1}\bigl(\operatorname{Bor}(\mathbb{R}^{n_j - n_{j-1}})\bigr) = \sigma(Y_j).$$

L'indépendance des tribus $\sigma(Y_j)$ résulte du corollaire 5.40 et passe évidemment à leurs sous-tribus $\sigma(Z_j)$. □

Proposition 5.42.

a) Les variables aléatoires réelles $X_1, \dots, X_d$ sont indépendantes si et seulement si

$$\forall (x_1, \dots, x_d) \in \mathbb{R}^d, \quad \mathbf{P}(X_1 \leq x_1, \dots, X_d \leq x_d) = \mathbf{P}(X_1 \leq x_1) \dots \mathbf{P}(X_d \leq x_d). \tag{5.57}$$

b) Soit $(X_1, \dots, X_d)$ un vecteur aléatoire dans $\mathbb{R}^d$, de densité f. Ses composantes X_i sont indépendantes si et seulement s'il existe $f_1, \dots, f_d$, $\mathbb{R} \to \overline{\mathbb{R}}_+$, mesurables telles que

$$f = f_1 \otimes \cdots \otimes f_d.$$

Dans ce cas les densités marginales f_{X_i} sont de la forme $c_i f_i$ où les constantes $c_i > 0$ sont liées par $c_1 \dots c_d = 1$.

c) Soient $(X_1, \dots, X_d)$ un vecteur aléatoire discret à valeurs dans $\mathbb{N}^d$. Ses composantes X_i sont indépendantes si et seulement si

$$\forall (k_1, \dots, k_d) \in \mathbb{N}^d, \quad \mathbf{P}(X_1 = k_1, \dots, X_d = k_d) = \mathbf{P}(X_1 = k_1) \dots \mathbf{P}(X_d = k_d). \tag{5.58}$$

Preuve du a). Il suffit de remarquer que (5.57) exprime l'égalité des deux mesures finies $P_{(X_1, \dots, X_d)}$ et $P_{X_1} \otimes \cdots \otimes P_{X_d}$ sur la π-classe $\mathcal{C}$ des pavés mesurables de la forme

$$B =]-\infty, x_1] \times \cdots \times]-\infty, x_d], \quad (x_1, \dots, x_d) \in \mathbb{R}^d.$$

Comme $\sigma(\mathcal{C}) = \mathrm{Bor}(\mathbb{R}^d)$, le théorème d'unicité des mesures implique l'égalité de ces mesures sur toute la tribu borélienne de $\mathbb{R}^d$. L'indépendance des X_i en découle par la proposition 5.39. Réciproquement cette indépendance implique immédiatement (5.57).

Notons que le membre de gauche dans (5.57) définit la fonction de répartition $F_{X_1,\dots,X_d}$ du vecteur $(X_1, \dots, X_d)$ et que le membre de droite est le produit (tensoriel) des fonctions de répartition des X_i. On a ainsi une caractérisation de l'indépendance :

$$X_1, \dots, X_d \text{ indépendantes} \quad \Leftrightarrow \quad F_{X_1,\dots,X_d} = F_{X_1} \otimes \dots \otimes F_{X_d}. \tag{5.59}$$

exprimée en termes de fonctions de répartition. □

Preuve du b).

Supposons d'abord que la densité f du vecteur aléatoire $(X_1, \dots, X_d)$ soit de la forme $f = f_1 \otimes \dots \otimes f_d$ où les f_i sont des fonctions d'une variables réelle, mesurables positives. Par la proposition 5.21, la loi de chaque composante X_i a aussi une densité f_{X_i}, qui s'obtient en intégrant f par rapport à tous les x_j pour $j \neq i$:

$$\begin{aligned} f_{X_i}(t) &= \int_{\mathbb{R}^{d-1}} f_1(x_1) \dots f_{i-1}(x_{i-1}) f_i(t) f_{i+1}(x_{i+1}) \dots f_d(x_d) \mathrm{d}\lambda_{d-1}(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d) \\ &= f_i(t) \int_{\mathbb{R}^{d-1}} f_1(x_1) \dots f_{i-1}(x_{i-1}) f_{i+1}(x_{i+1}) \dots f_d(x_d) \mathrm{d}\lambda_{d-1}(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d) \\ &= c_i f_i(t). \end{aligned}$$

Comme f_{X_i} est une densité de probabilité, il est clair que la constante positive c_i ne peut pas être nulle. Calculons $P_{(X_1,\dots,X_d)}(B)$ pour $B = B_1 \times \dots \times B_d$, pavé mesurable quelconque de $\mathrm{Bor}(\mathbb{R})^{\otimes d}$. En utilisant le corollaire 5.20, on obtient :

$$\begin{aligned} P_{(X_1,\dots,X_d)}(B) = \int_B f \, \mathrm{d}\lambda_d &= \int_{B_1 \times \dots \times B_d} (f_1 \otimes \dots \otimes f_d) \, \mathrm{d}(\lambda_1^{\otimes d}) \\ &= \prod_{i=1}^d \int_{B_i} f_i \, \mathrm{d}\lambda_1 \\ &= \prod_{i=1}^d \frac{1}{c_i} \int_{B_i} f_{X_i} \, \mathrm{d}\lambda_1 \\ &= \prod_{i=1}^d \frac{1}{c_i} P_{X_i}(B_i). \end{aligned}$$

Ainsi

$$P_{(X_1,\dots,X_d)}(B_1 \times \dots \times B_d) = a P_{X_1}(B_1) \dots P_{X_d}(B_d), \tag{5.60}$$

où la constante a vaut $(c_1 \dots c_d)^{-1}$. En particulier en choisissant dans (5.60) tous les B_i égaux à $\mathbb{R}$, on obtient $1 = a \times 1$ d'où $a = 1$ et $c_1 \dots c_d = 1$. En reportant cette valeur de a dans (5.60), on voit alors que (5.60) exprime l'indépendance des X_i.

Réciproquement, supposons les X_i indépendantes. Par la proposition 5.21, chaque X_i a une densité f_{X_i}. Évaluons $P_{(X_1,\dots,X_d)}(B)$ pour $B = B_1 \times \dots \times B_d$ pavé mesurable quelconque de $\mathrm{Bor}(\mathbb{R})^{\otimes d}$. En utilisant l'indépendance des X_i, la définition des f_{X_i} et le corollaire 5.20, on obtient :

$$P_{(X_1,\dots,X_d)}(B) = \prod_{i=1}^{d} P_{X_i}(B_i) = \prod_{i=1}^{d} \int_{B_i} f_{X_i}\,\mathrm{d}\lambda_1.$$

En appliquant le corollaire 5.20, on en déduit

$$P_{(X_1,\dots,X_d)}(B_1 \times \cdots \times B_d) = \int_{B_1\times\cdots\times B_d} (f_{X_1} \otimes \cdots \otimes f_{X_d})\,\mathrm{d}\lambda_d. \tag{5.61}$$

En particulier en prenant dans (5.61) tous les B_i égaux à $\mathbb{R}$, on voit que

$$\int_{\mathbb{R}^d} (f_{X_1} \otimes \cdots \otimes f_{X_d})\,\mathrm{d}\lambda_d = 1,$$

ce qui montre que la fonction mesurable positive $g := f_{X_1} \otimes \cdots \otimes f_{X_d}$ est une densité de probabilité. En notant provisoirement μ la mesure de densité g par rapport à λ_d, l'égalité (5.61) s'interprète alors comme l'égalité de μ et de $P_{(X_1,\dots,X_d)}$ sur la classe $\mathcal{R}_d$ des pavés mesurables de $\mathrm{Bor}(\mathbb{R})^{\otimes d}$, donc sur toute la tribu $\mathrm{Bor}(\mathbb{R})^{\otimes d}$ par unicité de la mesure produit. Ainsi f et g sont deux densités par rapport à λ_d de la même mesure. Elles sont donc égales λ_d presque partout[15]. □

Preuve du c). La tribu concernée par la mesurabilité du vecteur aléatoire $(X_1, \dots, X_d)$ est ici $\mathcal{P}(\mathbb{N}^d)$. On voit que cette tribu coïncide avec $\mathcal{P}(\mathbb{N})^{\otimes d}$ car tout élément de $\mathcal{P}(\mathbb{N}^d)$ est une partie de $\mathbb{N}^d$, donc est au plus dénombrable et réunion au plus dénombrable de singletons $\{\mathbf{k}\} = \{(k_1, \dots, k_d)\}$. La condition (5.58) traduit l'égalité des mesures $P_{(X_1,\dots,X_d)}$ et $P_{X_1} \otimes \cdots \otimes P_{X_d}$ sur la classe des singletons de $\mathbb{N}^d$. L'indépendance des X_i se traduit elle, par l'égalité de ces mesures sur la classe des pavés $A = A_1 \times \cdots \times A_d$ où chaque A_i peut être une partie quelconque de $\mathbb{N}$. Elle implique donc (5.58) en prenant pour chaque A_i un singleton $\{k_i\}$ de $\mathbb{N}$. Pour la réciproque, on remarque que A et les A_i sont au plus dénombrables et on utilise les propriétés des séries multiples à termes

[15]Exercice : si f et g sont deux densités de la même mesure μ par rapport à une mesure ν, alors $f = g$ ν-presque partout. Indication : soit $A := \{f < g\}$, montrer que $\nu(A) = 0$ en raisonnant par l'absurde en remarquant que A est l'union croissante des $A_n := \{f + 1/n < g\}$.

positifs :

$$\begin{aligned}
P_{(X_1,\dots,X_d)}(A) &= \sum_{(k_1,\dots,k_d)\in A} P_{(X_1,\dots,X_d)}\big(\{(k_1,\dots,k_d)\}\big)\\
&= \sum_{(k_1,\dots,k_d)\in A} \mathbf{P}(X_1=k_1,\dots,X_d=k_d)\\
&= \sum_{(k_1,\dots,k_d)\in A} \mathbf{P}(X_1=k_1)\dots\mathbf{P}(X_d=k_d)\\
&= \left(\sum_{k_1\in A_1}\mathbf{P}(X_1=k_1)\right)\times\dots\times\left(\sum_{k_d\in A_d}\mathbf{P}(X_d=k_d)\right)\\
&= \mathbf{P}(X_1\in A_1)\dots\mathbf{P}(X_d\in A_d)\\
&= (P_{X_1}\otimes\dots\otimes P_{X_d})(A).
\end{aligned}$$

□

Proposition 5.43. *Soit $(X_i)_{i\in I}$ une famille quelconque de variables aléatoires réelles indépendantes. Alors pour toute partie finie J de I et toute famille $\{h_j,\ j\in J\}$ de fonctions boréliennes $\mathbb{R}\to\mathbb{C}$ telles que les $h_j(X_j)$ soient $\mathbf{P}$-intégrables, la variable aléatoire produit $\prod_{j\in J} h_j(X_j)$ est $\mathbf{P}$-intégrable et*

$$\mathbf{E}\left(\prod_{j\in J} h_j(X_j)\right) = \prod_{j\in J}\mathbf{E}h_j(X_j).$$

Démonstration. En réindexant l'ensemble fini $J=\{i_1,\dots,i_n\}$, on se ramène au cas $J=\{1,\dots,n\}$. La fonction $\omega\mapsto\prod_{j\in J} h_j\big(X_j(\omega)\big)$ est mesurable comme produit des fonctions mesurables $h_j\circ X_j$. Par le théorème de transfert et la proposition 5.39, on a

$$\begin{aligned}
\mathbf{E}\left|\prod_{j=1}^n h_j(X_j)\right| &= \int_{\mathbb{R}^n} |h_1(x_1)\dots h_n(x_n)|\,\mathrm{d}P_{(X_1,\dots,X_n)}(x_1,\dots,x_n)\\
&= \int_{\mathbb{R}^n} |h_1|\otimes\dots\otimes|h_n|\,\mathrm{d}(P_{X_1}\otimes\dots\otimes P_{X_n}).
\end{aligned}$$

Par le corollaire 5.20, cette dernière intégrale est finie dès que chacune des intégrales $\int_{\mathbb{R}}|h_j|\,\mathrm{d}P_{X_j}$ est finie, autrement dit (par transfert) dès que $\mathbf{E}|h_j(X_j)|<+\infty$. Ayant ainsi réglé la question de l'intégrabilité de la variable aléatoire $\prod_{j\in J}h_j(X_j)$, on peut écrire en utilisant successivement le transfert $\Omega\to\mathbb{R}^n$, la proposition 5.39, le corollaire 5.20 et les transferts $\mathbb{K} \xrightarrow{X_j} \Omega$:

$$\begin{aligned}
\mathbf{E}\left(\prod_{j=1}^n h_j(X_j)\right) &= \int_{\mathbb{R}^n} h_1(x_1)\dots h_n(x_n)\,\mathrm{d}P_{(X_1,\dots,X_n)}(x_1,\dots,x_n)\\
&= \int_{\mathbb{R}^n} (h_1\otimes\dots\otimes h_n)\,\mathrm{d}(P_{X_1}\otimes\dots\otimes P_{X_n})\\
&= \prod_{j=1}^n \int_{\mathbb{R}} h_j\,\mathrm{d}P_{X_j}\\
&= \prod_{j=1}^n \mathbf{E}h_j(X_j).
\end{aligned}$$

□

Corollaire 5.44. *Si les variables aléatoires réelles (ou complexes) $X_1,\dots,X_n$ sont indépendantes et intégrables, leur produit est aussi intégrable et*

$$\mathbf{E}(X_1\dots X_n) = (\mathbf{E}X_1)\dots(\mathbf{E}X_n).$$

Il est facile de voir que la réciproque du corollaire 5.44 est fausse en construisant un couple (X_1,X_2) de variables aléatoires réelles non indépendantes telles que $\mathbf{E}(X_1X_2) = \mathbf{E}X_1\mathbf{E}X_2$. Prenons par exemple X_1 de loi uniforme sur $[-1,+1]$ et $X_2 := X_1^2$. On a alors

$$\mathbf{E}X_1X_2 = \mathbf{E}X_1^3 = \int_{[-1,+1]} x^3\,\mathrm{d}\lambda(x) = 0.$$

D'autre part $\mathbf{E}X_1 = 0$ donc $\mathbf{E}X_1\mathbf{E}X_2 = 0$. Il est clair intuitivement que X_1 et X_2 ne sont pas indépendantes puisque X_2 est une fonction déterministe de X_1. Pour vérifier cette non-indépendance par le calcul, on peut remarquer que d'une part

$$\begin{aligned}
\mathbf{P}\big(X_1\in[0,1/2] \text{ et } X_2\in[0,1/4]\big) &= \mathbf{P}\big(X_1\in[0,1/2] \text{ et } X_1\in[-1/2,1/2]\big)\\
&= \mathbf{P}\big(X_1\in[0,1/2]\big) = \frac{1}{4}
\end{aligned}$$

et d'autre part

$$\mathbf{P}\big(X_1\in[0,1/2]\big)\mathbf{P}\big(X_2\in[0,1/4]\big) = \frac{1}{4}\mathbf{P}\big(X_1\in[-1/2,1/2]\big) = \frac{1}{8}.$$

5.5.3 Covariance

Nous venons de voir que si X et Y sont deux variables aléatoires indépendantes et intégrables, $\mathbf{E}XY = \mathbf{E}X\mathbf{E}Y$ et que la réciproque était fausse. Ceci est l'une des motivations pour l'étude des propriétés de la quantité $\mathbf{E}(XY) - \mathbf{E}X\mathbf{E}Y$.

Définition 5.45 (Covariance). *Soient X et Y deux variables aléatoires de carré intégrables. La quantité*

$$\operatorname{Cov}(X,Y) := \mathbf{E}\big[(X-\mathbf{E}X)(Y-\mathbf{E}Y)\big]$$

est bien définie et est appelée covariance de X et Y.

Pour légitimer cette définition, il nous faut vérifier que l'hypothèse d'intégrabilité de X^2 et Y^2 entraîne l'existence de $\mathbf{E}X$ et $\mathbf{E}Y$ et l'intégrabilité de $(X-\mathbf{E}X)(Y-\mathbf{E}Y)$. Pour le premier point, il suffit d'écrire

$$\mathbf{E}|X| = \int_{\{|X|\leq 1\}} |X|\,\mathrm{d}\mathbf{P} + \int_{\{|X|>1\}} |X|\,\mathrm{d}\mathbf{P} \leq 1 + \int_{\{|X|>1\}} X^2\,\mathrm{d}\mathbf{P} \leq 1 + \mathbf{E}X^2 < +\infty.$$

Pour le deuxième point, on utilise d'abord l'inégalité $|xy| \leq x^2+y^2$ vraie pour tous réels x, y. Posons $x = X(\omega)-\mathbf{E}X$ et $y = Y(\omega)-\mathbf{E}Y$ et intégrons l'inégalité obtenue sur Ω :

$$\mathbf{E}\big|(X-\mathbf{E}X)(Y-\mathbf{E}Y)\big| \leq \mathbf{E}\big[(X-\mathbf{E}X)^2\big] + \mathbf{E}\big[(Y-\mathbf{E}Y)^2\big]. \tag{5.62}$$

On utilise ensuite l'inégalité $(a+b)^2 \leq 2a^2+2b^2$ en posant $a = X(\omega)$ et $b = \mathbf{E}X$. On en déduit après intégration sur Ω :

$$\mathbf{E}\big[(X-\mathbf{E}X)^2\big] \leq 2\mathbf{E}(X^2) + 2(\mathbf{E}X)^2 < +\infty.$$

En reportant cette majoration et son analogue pour Y dans (5.62), on obtient l'intégrabilité de $(X-\mathbf{E}X)(Y-\mathbf{E}Y)$.

Proposition 5.46 (Formule de Koenig). *Si X et Y sont deux variables aléatoires de carré intégrables,*

$$\operatorname{Cov}(X,Y) = \mathbf{E}(XY) - \mathbf{E}X\mathbf{E}Y. \tag{5.63}$$

Démonstration. Comme nous venons de le voir, l'intégrabilité de X^2 et de Y^2 implique celles de X, de Y et de XY. Le second membre de (5.63) est donc bien défini. Introduisons pour alléger, les *constantes* $c = \mathbf{E}X$ et $c' = \mathbf{E}Y$. En utilisant la linéarité de l'espérance et le fait que l'espérance d'une constante est égale à cette constante, on obtient :

$$\begin{aligned}
\operatorname{Cov}(X,Y) = \mathbf{E}\big[(X-c)(Y-c')\big] &= \mathbf{E}(XY - cY - c'X + cc') \\
&= \mathbf{E}(XY) - c\mathbf{E}Y - c'\mathbf{E}X + cc' \\
&= \mathbf{E}(XY) - cc',
\end{aligned}$$

ce qui établit (5.63). □

Définition 5.47. *On dit que deux variables aléatoires réelles de carré intégrable X et Y sont non-corrélées si*

$$\mathbf{E}(XY) = \mathbf{E}X\mathbf{E}Y.$$

Cette condition équivaut (pour des variables de carré intégrable) à $\operatorname{Cov}(X,Y) = 0$.

Nous savons déjà que deux variables aléatoires indépendantes et de carré intégrable sont non-corrélées et que la réciproque est fausse. Notons aussi que pour des variables aléatoires X et Y indépendantes et intégrables, on a $\mathbf{E}(XY) = \mathbf{E}X\mathbf{E}Y$, même si ces variables ne sont pas de *carré* intégrable.

Définition 5.48 (Variance). *Soit X une variable aléatoire de carré intégrable. On appelle variance de X la quantité*

$$\operatorname{Var} X := \mathbf{E}\big[(X - \mathbf{E}X)^2\big] = \operatorname{Cov}(X, X).$$

L'existence a déjà été justifiée lors de la définition de la covariance. La formule de Koenig appliquée à ce cas particulier s'écrit

$$\operatorname{Var} X = \mathbf{E}(X^2) - (\mathbf{E}X)^2. \tag{5.64}$$

Proposition 5.49 (Variance d'une somme).

a) Si les variables aléatoires réelles $X_1, \dots, X_n$ sont de carré intégrable :

$$\operatorname{Var}\Big(\sum_{i=1}^n X_i\Big) = \sum_{i,j=1}^n \operatorname{Cov}(X_i, X_j) \tag{5.65}$$

$$= \sum_{i=1}^n \operatorname{Var} X_i + \sum_{\substack{i,j=1\\ i\neq j}}^n \operatorname{Cov}(X_i, X_j). \tag{5.66}$$

b) Si de plus elles sont deux à deux non corrélées ($\operatorname{Cov}(X_i, X_j) = 0$ pour $i \neq j$),

$$\operatorname{Var}\Big(\sum_{i=1}^n X_i\Big) = \sum_{i=1}^n \operatorname{Var} X_i.$$

Démonstration. Il suffit de remarquer que l'application $(X, Y) \mapsto \operatorname{Cov}(X, Y)$ est une forme bilinéaire sur l'espace vectoriel des variables aléatoires de carré intégrable. □

5.6 Convolution

Dans cette section, nous désignons par $(E, \mathcal{B})$ l'espace mesurable $E = \mathbb{R}^d$ ou $\mathbb{Z}^d$ ($d \geq 1$) muni de la tribu $\mathcal{B} = \operatorname{Bor}(\mathbb{R}^d)$ ou $\mathcal{B} = \mathcal{P}(\mathbb{Z}^d)$ respectivement. Dans les deux cas la tribu $\mathcal{B}$ est invariante par toute translation T_u de vecteur u dans E.

Définition 5.50. *Avec les notations ci-dessus, soient μ_1, μ_2 deux mesures finies sur $(E, \mathcal{B})$. Le produit de convolution de μ_1 et μ_2 est la mesure*

$$\mu_1 * \mu_2 := (\mu_1 \otimes \mu_2) \circ s^{-1},$$

où $s : E^2 \to E, \quad (x, y) \mapsto x + y$ est l'addition sur $E \times E$.

Notons que $\mu_1 * \mu_2$ est une mesure finie sur $(E, \mathcal{B})$. En effet $s^{-1}(E) = E^2$ et

$$\mu_1 * \mu_2(E) = \mu_1 \otimes \mu_2(E^2) = \mu_1(E)\mu_2(E) < +\infty.$$

En particulier quand μ_1 et μ_2 sont des probabilités sur $(E, \mathcal{B})$, $\mu_1 * \mu_2$ est aussi une probabilité sur $(E, \mathcal{B})$.

Proposition 5.51. *Pour tout $B \in \mathcal{B}$,*

$$\mu_1 * \mu_2(B) = \int_E \mu_2(B - x)\,\mathrm{d}\mu_1(x) = \int_E \mu_1(B - y)\,\mathrm{d}\mu_2(y), \tag{5.67}$$

où $B - x := \{z - x;\ z \in B\} = T_{-x}(B)$.

Démonstration. Notons $A := s^{-1}(B) = \{(x, y) \in E^2;\ x + y \in B\}$. Cherchons la section de A en x :

$$\begin{aligned} A_x = \{y \in E;\ (x, y) \in A\} &= \{y \in E;\ x + y \in B\} \\ &= \{y \in E;\ \exists z \in B, y = z - x\} \\ &= \{z - x;\ z \in B\} \\ &= B - x. \end{aligned}$$

Il suffit alors d'appliquer la formule (5.4) de calcul de la mesure produit :

$$\mu_1 * \mu_2(B) = \mu_1 \otimes \mu_2\big(s^{-1}(B)\big) = \mu_1 \otimes \mu_2(A) = \int_E \mu_2(A_x)\,\mathrm{d}\mu_1(x) = \int_E \mu_2(B - x)\,\mathrm{d}\mu_1(x).$$

La deuxième égalité dans (5.67) s'obtient en échangeant les rôles de x et y. $\square$

Remarque 5.52. Formellement, rien n'interdit de généraliser la définition 5.50 à des mesures σ-finies quelconques. Le calcul ci-dessus reste valable. Mais si on prend $E = \mathbb{R}$ et $\mu_1 = \mu_2 = \lambda$ mesure de Lebesgue, on voit que $\mu_1 * \mu_2(B) = +\infty$ pour tout borélien B tel que $\lambda(B) > 0$ puisque $\lambda(B - x) = \lambda(B)$. Ce type de pathologie nous incite à nous restreindre au cas des mesures finies.

Proposition 5.53. *Si X et Y sont deux variables aléatoires* indépendantes *à valeurs dans $(E, \mathcal{B})$, la loi de $X + Y$ est le produit de convolution des lois de X et Y :*

$$X \text{ et } Y \text{ indépendantes} \quad \Rightarrow \quad P_{X+Y} = P_X * P_Y. \tag{5.68}$$

Démonstration. Puisque $X+Y = s(X, Y)$, $P_{X+Y} = P_{(X,Y)} \circ s^{-1}$. Cette égalité est vérifiée pour la loi de n'importe quelle somme, que les termes soient indépendants ou non. La propriété supplémentaire apportée par l'indépendance est que $P_{(X,Y)} = P_X \otimes P_Y$. La conclusion en découle puisque par la définition 5.50, $(P_X \otimes P_Y) \circ s^{-1} = P_X * P_Y$. $\square$

Voyons maintenant comment intégrer par rapport à $\mu_1 * \mu_2$.

Proposition 5.54. *La formule*

$$\int_E h \, \mathrm{d}(\mu_1 * \mu_2) = \int_{E^2} h(x+y) \, \mathrm{d}(\mu_1 \otimes \mu_2)(x,y)$$

est vérifiée par

- *toute fonction h mesurable positive $E \to \overline{\mathbb{R}}_+$;*
- *toute $h \in \mathcal{L}^1_{\mathbb{K}}(\mu_1 * \mu_2)$, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, ce qui équivaut à $h \circ s \in \mathcal{L}^1_{\mathbb{K}}(\mu_1 \otimes \mu_2)$.*

Démonstration. Puisque $\mu_1 * \mu_2$ est la mesure image de $\mu_1 \otimes \mu_2$ par s, c'est simplement le théorème de transfert appliqué à $E \times E \xrightarrow{s} E$. □

Même si le produit de convolution de la mesure de Lebesgue avec elle même est sans intérêt, on peut regarder ce que donne la convolution de deux mesures finies à densité par rapport à λ.

Proposition 5.55. *Si μ_1 et μ_2 sont les mesures finies de densités respectives f et g par rapport à la mesure de Lebesgue λ sur $\mathbb{R}^d$, $\mu_1 * \mu_2$ est la mesure à densité h définie λ-p.p. par*

$$h(y) = \int_{\mathbb{R}^d} f(x) g(y-x) \, \mathrm{d}\lambda(x) = \int_{\mathbb{R}^d} f(y-x) g(x) \, \mathrm{d}\lambda(x). \tag{5.69}$$

Démonstration. Soit B un borélien quelconque de $\mathbb{R}^d$. L'application de (5.67) nous donne

$$\mu_1 * \mu_2(B) = \int_{\mathbb{R}^d} \mu_2(B-x) f(x) \, \mathrm{d}\lambda(x) = \int_{\mathbb{R}^d} \left\{ \int_{\mathbb{R}^d} \mathbf{1}_{B-x}(u) g(u) \, \mathrm{d}\lambda(u) \right\} f(x) \, \mathrm{d}\lambda(x).$$

Notons que $\mathbf{1}_{B-x}(u) = \mathbf{1}_B(x+u)$. Le changement de variable $y = x+u$ (à x fixé) dans l'intégrale interne s'écrit compte-tenu de l'invariance de λ et de $\mathbb{R}^d$ par translation :

$$\int_{\mathbb{R}^d} \mathbf{1}_{B-x}(u) g(u) \, \mathrm{d}\lambda(u) = \int_{\mathbb{R}^d} \mathbf{1}_B(y) g(y-x) \, \mathrm{d}\lambda(y).$$

En reportant cette expression dans l'intégrale itérée ci-dessus et en permutant l'ordre des intégrations grâce au théorème de Fubini Tonelli, il vient

$$\mu_1 * \mu_2(B) = \int_{\mathbb{R}^d} \mathbf{1}_B(y) \left\{ \int_{\mathbb{R}^d} f(x) g(y-x) \, \mathrm{d}\lambda(x) \right\} \mathrm{d}\lambda(y) = \int_B h(y) \, \mathrm{d}\lambda(y).$$

Nous venons ainsi d'établir que pour tout $B \in \mathrm{Bor}(\mathbb{R}^d)$, $\mu_1 * \mu_2(B) = \int_B h \, \mathrm{d}\lambda$, autrement dit que $\mu_1 * \mu_2$ est la mesure de densité h par rapport à λ. La deuxième expression de h contenue dans la formule (5.69) s'obtient de manière symétrique en permutant les rôles de μ_1 et μ_2 dans le calcul ci-dessus. Notons enfin qu'en prenant $B = \mathbb{R}^d$ dans le calcul ci-dessus, on voit que h est λ-intégrable sur $\mathbb{R}^d$, en particulier, h est finie λ-presque partout. Par ailleurs, toute fonction égale λ-p.p. à h est aussi une densité de $\mu_1 * \mu_2$. □

La proposition 5.55 incite à définir le produit de convolution de deux densités f et g (par rapport à λ de mesures finies) par $f * g = h$ où h est donnée par (5.69). Le pas suivant consiste à remarquer que si f et g sont dans $\mathrm{L}^1_{\mathbb{K}}(\lambda)$, $|f|$ et $|g|$ sont des densités par rapport à λ de mesures finies. Ceci nous conduit à la définition du produit de convolution de deux éléments de $\mathrm{L}^1_{\mathbb{K}}(\lambda)$.

Définition 5.56. *Si f et g sont dans $\mathrm{L}^1_{\mathbb{K}}(\lambda)$, leur produit de convolution est la classe de fonctions dont un représentant est défini λ-p.p. par*

$$(f * g)(y) := \int_{\mathbb{R}^d} f(x)g(y-x)\,\mathrm{d}\lambda(x) = \int_{\mathbb{R}^d} f(y-x)g(x)\,\mathrm{d}\lambda(x).$$

Proposition 5.57. *Si f et g sont dans $\mathrm{L}^1_{\mathbb{K}}(\lambda)$, alors $f * g$ est aussi dans $\mathrm{L}^1_{\mathbb{K}}(\lambda)$ et*

$$\|f * g\|_1 \leq \|f\|_1 \|g\|_1.$$

Démonstration. Il suffit d'appliquer le théorème de Fubini Tonelli et l'invariance par translation de la mesure de Lebesgue :

$$\begin{aligned}
\int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} f(x)g(y-x)\,\mathrm{d}\lambda(x) \right| \mathrm{d}\lambda(y) &\leq \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |f(x)||g(y-x)|\,\mathrm{d}\lambda(x)\,\mathrm{d}\lambda(y) \\
&= \int_{\mathbb{R}^d} |f(x)| \left\{ \int_{\mathbb{R}^d} |g(y-x)|\,\mathrm{d}\lambda(y) \right\} \mathrm{d}\lambda(x) \\
&= \int_{\mathbb{R}^d} |f(x)| \left\{ \int_{\mathbb{R}^d} |g(y)|\,\mathrm{d}\lambda(y) \right\} \mathrm{d}\lambda(x) \\
&= \left\{ \int_{\mathbb{R}^d} |g(y)|\,\mathrm{d}\lambda(y) \right\} \int_{\mathbb{R}^d} |f(x)|\,\mathrm{d}\lambda(x) \\
&= \|f\|_1 \|g\|_1.
\end{aligned}$$

□