2. Théorème de Fubini

Contenu du chapitre

2.1. Questions d'unicité

On a beaucoup vanté dans le premier chapitre les mérites de l'intégration des fonctions à valeurs dans $[0, +\infty]$, mais il y a un type de raisonnement qui ne marche pas dans ce cadre : c'est celui qui dit que si $\mathrm{A}_1 \subset \mathrm{A}_2$ et si la mesure $\mu(\mathrm{A}_2)$ est finie, on peut calculer la mesure de la différence ensembliste $\mathrm{A}_2 \setminus \mathrm{A}_1$ par *différence des mesures*

$$\mu(\mathrm{A}_2 \setminus \mathrm{A}_1) = \mu(\mathrm{A}_2) - \mu(\mathrm{A}_1).$$

Pour des démonstrations qui utilisent ce type de ressort, il est préférable de se placer dans un cadre *d'espace vectoriel* de fonctions (ce qui sous-entend naturellement : des fonctions *partout finies*).

L'un des objectifs de cette section est de donner des critères qui permettent de montrer que deux mesures μ et ν sont égales, mais à partir d'une information limitée, portant sur l'égalité $\mu(\mathrm{C}) = \nu(\mathrm{C})$ pour tout C d'une « petite » classe $\mathcal{C}$ de parties. Le mieux qu'on puisse espérer est d'en déduire l'égalité sur la tribu engendrée $\sigma(\mathcal{C})$. C'est le sens des lemmes qui vont suivre.

On choisira de travailler sur des espaces de fonctions plutôt que sur des classes d'ensembles. Voici un premier exemple où on utilise cette correspondance. Une *algèbre de fonctions réelles sur* Ω est un $\mathbb{R}$-espace vectoriel E de fonctions qui est de plus stable par produit des fonctions : si $f, g \in \mathrm{E}$, alors la fonction fg est dans E.

Un exemple important pour ce chapitre est l'algèbre de fonctions $\mathcal{L}_\infty(\Omega, \mathcal{F})$, où $\mathcal{F}$ est une tribu de parties de Ω : c'est l'algèbre formée de toutes les fonctions réelles bornées et $\mathcal{F}$-mesurables, c'est-à-dire les fonctions réelles f bornées sur Ω telles que $\{f > c\} \in \mathcal{F}$ pour tout réel c.

Lemme. *Si* E *est une algèbre de fonctions sur* Ω*, contenant les fonctions constantes, la classe d'ensembles*

$$\mathcal{E} = \{\mathrm{A} \subset \Omega : \mathbf{1}_\mathrm{A} \in \mathrm{E}\}$$

est une algèbre de parties de Ω.

Démonstration. Puisque E contient la fonction constante $\mathbf{1} = \mathbf{1}_\Omega$ égale à 1 en tout point de Ω, on a d'abord $\Omega \in \mathcal{E}$; ensuite, si $\mathrm{A} \in \mathcal{E}$, on écrit que $\mathbf{1}_{\mathrm{A}^c} = \mathbf{1}_\Omega - \mathbf{1}_\mathrm{A} \in \mathrm{E}$, donc $\mathrm{A}^c \in \mathcal{E}$. Si $\mathrm{A}, \mathrm{B} \in \mathcal{E}$, l'égalité $\mathbf{1}_{\mathrm{A}\cap\mathrm{B}} = \mathbf{1}_\mathrm{A}\, \mathbf{1}_\mathrm{B}$ montre que $\mathrm{A} \cap \mathrm{B} \in \mathcal{E}$, puisque E est stable par produit. On a montré que $\mathcal{E}$ est une algèbre de parties de Ω.

Pour aller plus loin et obtenir une tribu, il va falloir s'intéresser à la convergence de suites de fonctions.

2.1.a. *Un théorème de classe monotone*

Par *suite bornée de fonctions* sur un ensemble Ω on entendra une suite (f_n) uniformément bornée : il existe un réel M tel que $|f_n(\omega)| \le \mathrm{M}$ pour tout entier n et pour tout $\omega \in \Omega$.

Disons que E, espace vectoriel de fonctions réelles bornées sur Ω, est *stable par convergence simple bornée* (en abrégé : stable par CSB) si pour toute suite bornée (f_n) d'éléments de E qui converge simplement sur Ω vers une fonction f, on a $f \in \mathrm{E}$.

Il est clair que toute intersection d'espaces de fonctions sur Ω, stables par CSB, est encore stable par CSB, ce qui donne la possibilité de définir l'espace stable par CSB engendré par une classe quelconque Φ de fonctions réelles bornées sur Ω, comme étant l'intersection de tous les espaces stables par CSB qui contiennent Φ. Cette intersection est évidemment le plus petit espace stable par CSB qui contienne la classe Φ.

L'algèbre $\mathcal{L}_\infty(\Omega, \mathcal{F})$, où $\mathcal{F}$ est une tribu, est l'exemple typique d'algèbre stable par CSB. On retrouve cet exemple un peu partout, comme l'indique le lemme qui suit. De plus, le plus grand espace de fonctions bornées sur Ω stable par CSB est $\mathcal{L}_\infty(\Omega, \mathcal{P}(\Omega))$, l'espace de *toutes* les fonctions bornées sur Ω.

Lemme 2.1.1. *Si* F *est une algèbre de fonctions sur* Ω*, stable par CSB et contenant les fonctions constantes, la classe d'ensembles*

$$\mathcal{F} = \{\mathrm{A} \subset \Omega : \mathbf{1}_\mathrm{A} \in \mathrm{F}\}$$

est une tribu de parties de Ω*, et* F *contient l'algèbre* $\mathcal{L}_\infty(\Omega, \mathcal{F})$*.*

Démonstration. On sait déjà par le lemme précédent que $\mathcal{F}$ est une algèbre de parties de Ω ; si une suite $(\mathrm{A}_n) \subset \mathcal{F}$ admet A pour réunion croissante, la fonction $\mathbf{1}_\mathrm{A}$ est limite simple bornée de la suite $(\mathbf{1}_{\mathrm{A}_n}) \subset \mathrm{F}$, donc $\mathbf{1}_\mathrm{A} \in \mathrm{F}$ par la propriété d'ESCSB (abrégé pour espace stable par CSB), et il en résulte que $\mathrm{A} \in \mathcal{F}$. On a montré que $\mathcal{F}$ est une tribu.

Puisque F est un espace vectoriel, il contient toutes les fonctions $\mathcal{F}$-étagées, par linéarité ; on sait d'après la remarque 1.1.6 que toute fonction mesurable réelle f est limite d'une suite de fonctions $\mathcal{F}$-étagées (f_n) telle que $|f_n| \le |f|$; si $f \in \mathcal{L}_\infty(\Omega, \mathcal{F})$, la suite (f_n) est bornée et il résulte alors de la stabilité de F par CSB que $f \in \mathrm{F}$. On voit donc que $\mathcal{L}_\infty(\Omega, \mathcal{F})$ est contenu dans F.

On aura besoin du lemme simple qui suit.

Lemme. *Si* F *est un espace vectoriel stable par CSB, si* φ *est une fonction réelle bornée sur* Ω*, l'ensemble*

$$\mathrm{F}_\varphi = \{f \text{ bornée } : \varphi f \in \mathrm{F}\}$$

est un espace vectoriel stable par CSB.

Démonstration. Il est clair que F_φ est un espace vectoriel. Si une suite bornée $(f_n) \subset \mathrm{F}_\varphi$ tend simplement vers une fonction f, la suite (φf_n) est bornée, elle est dans F par définition de F_φ, et elle tend simplement vers φf, donc $\varphi f \in \mathrm{F}$ puisque F est stable par CSB, donc $f \in \mathrm{F}_\varphi$ et on a montré que F_φ est stable par CSB.

Théorème 2.1.2 (un théorème de classe monotone). *On suppose que* E *est un espace vectoriel de fonctions réelles bornées sur* Ω*, stable par CSB, qui contient la fonction* $\mathbf{1}$ *et les fonctions* $\mathbf{1}_{\mathrm{C}}$ *pour tout* C *d'une classe* $\mathcal{C}$ *de parties de* Ω *stable par intersection finie ; il en résulte que* E *contient l'algèbre* $\mathcal{L}_\infty(\Omega, \sigma(\mathcal{C}))$*, en particulier l'espace* E *contient toutes les fonctions* $\mathbf{1}_{\mathrm{B}}$*, pour* $\mathrm{B} \in \sigma(\mathcal{C})$*.*

Démonstration. Désignons par $\mathcal{F}(\mathbf{1}, \mathcal{C})$ la famille de tous les ESCSB qui contiennent $\mathbf{1}$ et toutes les $\mathbf{1}_{\mathrm{C}}$, $\mathrm{C} \in \mathcal{C}$; désignons par F le plus petit élément de la famille $\mathcal{F}(\mathbf{1}, \mathcal{C})$ (c'est l'intersection de tous les espaces de la famille) ; on a évidemment $\mathrm{F} \subset \mathrm{E}$, puisque E est un élément de la famille $\mathcal{F}(\mathbf{1}, \mathcal{C})$. Faisons la remarque suivante :

si une fonction $\varphi \in \mathrm{F}$ *est telle que* $\mathbf{1}_{\mathrm{C}}\, \varphi \in \mathrm{F}$ *pour tout* $\mathrm{C} \in \mathcal{C}$*, alors*

$$\varphi\, \mathrm{F} := \{\varphi f : f \in \mathrm{F}\} \subset \mathrm{F}.$$

On a dit en effet que la classe F_φ des fonctions f bornées telles que $\varphi f \in \mathrm{F}$ est un ESCSB ; de plus F_φ contient $\mathbf{1}$ puisque $\varphi \in \mathrm{F}$, et F_φ contient toutes les $\mathbf{1}_{\mathrm{C}}$, $\mathrm{C} \in \mathcal{C}$ par hypothèse. On voit que F_φ est un membre de la famille $\mathcal{F}(\mathbf{1}, \mathcal{C})$, donc F_φ est plus grand que le plus petit élément F de cette famille, ce qui se traduit par l'inclusion $\varphi\, \mathrm{F} \subset \mathrm{F}$ annoncée.

On va appliquer cette remarque deux fois de suite. Considérons d'abord un élément $\mathrm{C}_0 \in \mathcal{C}$, fixé mais quelconque. La fonction $\varphi = \mathbf{1}_{\mathrm{C}_0}$ est dans F (par définition de F), et pour tout $\mathrm{C} \in \mathcal{C}$ on a

$$\mathbf{1}_{\mathrm{C}}\, \mathbf{1}_{\mathrm{C}_0} = \mathbf{1}_{\mathrm{C} \cap \mathrm{C}_0} \in \mathrm{F},$$

puisque $\mathrm{C}_1 = \mathrm{C} \cap \mathrm{C}_0 \in \mathcal{C}$ (stabilité de $\mathcal{C}$ par intersection) et puisque F contient toutes les $\mathbf{1}_{\mathrm{C}_1}$, $\mathrm{C}_1 \in \mathcal{C}$. L'application de la remarque donne $\mathbf{1}_{\mathrm{C}_0}\mathrm{F} \subset \mathrm{F}$, pour tout $\mathrm{C}_0 \in \mathcal{C}$.

Prenons maintenant une fonction $f_0 \in \mathrm{F}$, fixée mais quelconque. La fonction $\varphi = f_0$ est dans F (par hypothèse), et pour tout $\mathrm{C} \in \mathcal{C}$ on a

$$\mathbf{1}_{\mathrm{C}}\, f_0 \in \mathrm{F}$$

d'après la première application de la remarque. Une deuxième application de la remarque donne $f_0\, \mathrm{F} \subset \mathrm{F}$, pour toute $f_0 \in \mathrm{F}$, ce qui signifie que F est stable par produit. On a donc maintenant une algèbre F, stable par CSB et contenant les constantes. D'après le lemme 2.1.1, la classe d'ensembles

$$\mathcal{F} = \{\mathrm{B} : \mathbf{1}_{\mathrm{B}} \in \mathrm{F}\}$$

est une tribu et $\mathcal{L}_\infty(\Omega, \mathcal{F}) \subset \mathrm{F}$; puisque $\mathcal{F}$ contient la classe $\mathcal{C}$ par construction de F, on voit que $\sigma(\mathcal{C}) \subset \mathcal{F}$, et puisque $\mathrm{F} \subset \mathrm{E}$, on obtient le résultat annoncé,

$$\mathcal{L}_\infty(\Omega, \sigma(\mathcal{C})) \subset \mathcal{L}_\infty(\Omega, \mathcal{F}) \subset \mathrm{F} \subset \mathrm{E}.$$

Remarque. L'espace F ci-dessus contient l'espace $\mathcal{L}_\infty(\Omega, \mathcal{A})$, où $\mathcal{A} = \sigma(\mathcal{C})$; mais cet espace $\mathcal{L}_\infty(\Omega, \mathcal{A})$ est stable par CSB, contient $\mathbf{1}$ et les $\mathbf{1}_{\mathrm{C}}$, donc $\mathrm{F} = \mathcal{L}_\infty(\Omega, \mathcal{A})$ puisque F est le plus petit élément de $\mathcal{F}(\mathbf{1}, \mathcal{C})$.

Corollaire 2.1.3. *Si deux mesures finies μ et ν sur $(\Omega, \mathcal{A})$ coïncident sur une classe $\mathcal{C}$ de parties de Ω stable par intersection et engendrant la tribu $\mathcal{A}$, et si $\mu(\Omega) = \nu(\Omega)$, on peut conclure que $\mu = \nu$.*

Démonstration. Désignons par E l'ensemble des fonctions mesurables bornées f telles que

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

Par la linéarité de l'intégrale il est évident que E est un espace vectoriel. Puisque les mesures sont finies, on note que les fonctions constantes sont intégrables ici, et peuvent servir de majorant dans le théorème de convergence dominée. Par une double application du théorème de convergence dominée, si $(f_n) \subset \mathrm{E}$ est une suite bornée qui tend vers f simplement, on voit que

$$\int f\,d\mu = \lim_n \int f_n\,d\mu = \lim_n \int f_n\,d\nu = \int f\,d\nu,$$

donc $f \in \mathrm{E}$ et on voit que E est un ESCSB. Cet espace contient toutes les $\mathbf{1}_{\mathrm{C}}$, $\mathrm{C} \in \mathcal{C}$ puisque $\int \mathbf{1}_{\mathrm{C}}\,d\mu = \mu(\mathrm{C}) = \nu(\mathrm{C}) = \int \mathbf{1}_{\mathrm{C}}\,d\nu$, et E contient $\mathbf{1}$, puisque $\mu(\Omega) = \nu(\Omega)$. D'après le théorème précédent, E contient toutes les fonctions $\mathbf{1}_{\mathrm{B}}$ pour $\mathrm{B} \in \mathcal{A} = \sigma(\mathcal{C})$, ce qui signifie que $\mu = \nu$.

Corollaire 2.1.4. *Si deux mesures μ et ν sur $(\Omega, \mathcal{A})$ coïncident sur une classe $\mathcal{C}$ de parties de Ω stable par intersection, engendrant la tribu $\mathcal{A}$, et s'il existe une suite croissante $(\mathrm{C}_n) \subset \mathcal{C}$ d'ensembles de* **mesure finie** *dont la réunion est Ω, alors $\mu = \nu$.*

Démonstration. Pour chaque n les formules

$$\forall \mathrm{A} \in \mathcal{A}, \quad \mu_n(\mathrm{A}) = \mu(\mathrm{C}_n \cap \mathrm{A}), \quad \nu_n(\mathrm{A}) = \nu(\mathrm{C}_n \cap \mathrm{A})$$

définissent deux mesures finies $\mu_n = \mathbf{1}_{\mathrm{C}_n}\,\mu$ et $\nu_n = \mathbf{1}_{\mathrm{C}_n}\,\nu$ sur $(\Omega, \mathcal{A})$, qui sont égales sur $\mathcal{C}$ et sur Ω ; il résulte du corollaire précédent que $\mu_n = \nu_n$, et on en déduit pour tout $\mathrm{A} \in \mathcal{A}$

$$\mu(\mathrm{A}) = \lim_n \mu(\mathrm{C}_n \cap \mathrm{A}) = \lim_n \nu(\mathrm{C}_n \cap \mathrm{A}) = \nu(\mathrm{A}).$$

2.1.b. *Application aux mesures sur $\mathbb{R}^d$; propriétés de la mesure de Lebesgue*

Corollaire. *Deux mesures μ_1 et μ_2 sur la tribu borélienne de $\mathbb{R}^d$ qui donnent la même mesure finie à tous les pavés compacts de la forme $\prod_{i=1}^d [a_i, b_i]$ sont égales, $\mu_1 = \mu_2$.*

C'est une conséquence immédiate du corollaire précédent : la classe $\mathcal{C}$, formée de l'ensemble vide et de tous les pavés compacts, est stable par intersection finie, ses éléments sont de mesure finie par hypothèse ; la classe $\mathcal{C}$ engendre la tribu borélienne et de plus la suite croissante des pavés

$$\mathrm{C}_n = [-n, n]^d \in \mathcal{C}$$

recouvre $\mathbb{R}^d$. On a tout ce qu'il faut pour appliquer le corollaire 2.1.4.

Exemple. La mesure de Lebesgue est invariante par translation et par symétrie.

Si λ désigne la mesure de Lebesgue, si $v = (v_1, \ldots, v_d)$ est un vecteur fixé et si on pose

$$\forall \mathrm{A} \in \mathcal{B}_{\mathbb{R}^d}, \quad \mu(\mathrm{A}) = \lambda(\mathrm{A} + v),$$

alors λ et μ donnent la même mesure finie à tous les pavés compacts, donc $\mu = \lambda$, ce qui signifie que la mesure de Lebesgue est invariante par translation. On règle de même l'image par $x \to -x$, en posant maintenant $\mu(\mathrm{A}) = \lambda(-\mathrm{A})$.

Envisageons l'effet des changements de variables linéaires dans l'intégrale de Lebesgue. Soit T une application de $\mathbb{R}^n$ dans $\mathbb{R}^n$ telle que l'image inverse de tout compact soit compacte ; on peut définir l'image ν de la mesure de Lebesgue λ par l'application T en posant pour tout borélien A de $\mathbb{R}^n$

$$\nu(\mathrm{A}) = \lambda(\mathrm{T}^{-1}(\mathrm{A})).$$

Il est évident que ν est une mesure au sens général, de plus cette mesure est finie sur les compacts. L'intégrale par rapport à ν est donnée par la formule

$$\int f(y)\, d\nu(y) = \int f(\mathrm{T}(x))\, dx,$$

pour toute fonction borélienne positive f. Cette formule est vraie par définition pour $f = \mathbf{1}_\mathrm{A}$, A borélien, donc pour les fonctions étagées par linéarité, puis pour les mesurables positives par convergence monotone.

Si $n = 1$, si $\mathrm{T}x = ax + b$ pour tout x réel, avec $a \neq 0$, on voit que tout intervalle de longueur ℓ est transformé par T en un intervalle de longueur $|a|\,\ell$; les mesures $\nu = \mathrm{T}(\lambda)$ et $|a|^{-1}\,\lambda$ coïncident sur les intervalles, donc elles sont égales, et

$$|a| \int_\mathbb{R} f(ax + b)\, dx = \int_\mathbb{R} f(y)\, dy$$

pour toute fonction f mesurable positive. Le même exemple en dimension d donne

$$|a|^d \int_{\mathbb{R}^d} f(ax + b)\, dx = \int_{\mathbb{R}^d} f(y)\, dy.$$

Le cas où le scalaire a serait remplacé par une matrice générale A demande d'avoir le théorème de Fubini pour la justification, et le remplacement de $|a|^d$ par la valeur absolue du déterminant de la matrice A (voir la section 2.3).

Exercice. On donne une fonction mesurable $f \geq 0$ sur $\mathbb{R}$, intégrable sur tout intervalle borné, et on suppose que la fonction F, définie sur $\mathbb{R}$ par la donnée de F(0) et par la formule

$$\forall x \in \mathbb{R}, \quad \mathrm{F}(x) = \mathrm{F}(0) + \int_0^x f(t)\, dt,$$

tend vers $-\infty$ quand $x \to -\infty$ et tend vers $+\infty$ quand $x \to +\infty$ (quand f est continue, f est bien la dérivée de F). Montrer que pour toute fonction borélienne $g \geq 0$ sur $\mathbb{R}$, on a la formule de changement de variable

$$\int_\mathbb{R} g(y)\, dy = \int_\mathbb{R} g(\mathrm{F}(x)) f(x)\, dx.$$

Solution. On pose pour tout borélien $\mathrm{A} \subset \mathbb{R}$

$$\nu(\mathrm{A}) = \int_{\mathbb{R}} \mathbf{1}_{\mathrm{A}}(\mathrm{F}(x))\, f(x)\, dx,$$

et on vérifie qu'on définit ainsi une mesure (par la proposition 1.1.4 sur l'intégration des séries de fonctions positives). Si φ est étagée positive, on obtient par « linéarité positive »

$$\int_{\mathbb{R}} \varphi(y)\, d\nu(y) = \int_{\mathbb{R}} \varphi(\mathrm{F}(x))\, f(x)\, dx,$$

et par une double application du théorème de convergence monotone, en écrivant que g est limite croissante d'une suite de fonctions étagées positives (φ_n),

$$\int_{\mathbb{R}} g(y)\, d\nu(y) = \int_{\mathbb{R}} g(\mathrm{F}(x))\, f(x)\, dx.$$

Puisque F est croissante continue et tend vers $\pm\infty$ aux bouts de la droite, il en résulte que si $[a, b]$ est un intervalle borné, on a $\mathrm{F}^{-1}([a, b]) = [\alpha, \beta]$, où $\mathrm{F}(\alpha) = a$ et $\mathrm{F}(\beta) = b$. Par conséquent,

$$\nu([a, b]) = \int_{\alpha}^{\beta} f(x)\, dx = \mathrm{F}(\beta) - \mathrm{F}(\alpha) = b - a\,;$$

ainsi, cette mesure ν coïncide avec la mesure de Lebesgue.

On va anticiper sur le théorème de Fubini, en montrant l'unicité de la mesure produit tensoriel, sous l'hypothèse que les mesures sont σ-finies.

Proposition 2.1.5. *Si μ_i est σ-finie sur $(\Omega_i, \mathcal{A}_i)$, pour $i = 1, 2$, il existe* **au plus** *une mesure ν sur $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathrm{A}_2)$ telle que*

$$\forall \mathrm{A}_1 \in \mathcal{A}_1,\ \forall \mathrm{A}_2 \in \mathcal{A}_2, \quad \nu(\mathrm{A}_1 \times \mathrm{A}_2) = \mu_1(\mathrm{A}_1)\, \mu_2(\mathrm{A}_2).$$

Démonstration. Considérons la classe $\mathcal{C}$ de tous les ensembles C de la forme $\mathrm{C} = \mathrm{A}_1 \times \mathrm{A}_2$, avec $\mathrm{A}_j \in \mathcal{A}_j$, $j = 1, 2$; on montre facilement qu'elle est stable par intersection finie,

$$(\mathrm{A}_1 \times \mathrm{A}_2) \cap (\mathrm{B}_1 \times \mathrm{B}_2) = (\mathrm{A}_1 \cap \mathrm{B}_1) \times (\mathrm{A}_2 \cap \mathrm{B}_2).$$

De plus la classe $\mathcal{C}$ engendre la tribu $\mathcal{A}_1 \otimes \mathcal{A}_2$. Si μ_1 et μ_2 sont σ-finies, on peut trouver deux suites croissantes d'ensembles $(\mathrm{A}_{1,n})$ et $(\mathrm{A}_{2,n})$ telles que $\mu_i(\mathrm{A}_{i,n}) < +\infty$ et que $\Omega_i = \bigcup_n \mathrm{A}_{i,n}$, $i = 1, 2$. Posons $\mathrm{C}_n = \mathrm{A}_{1,n} \times \mathrm{A}_{2,n}$; cette suite croissante recouvre le produit $\Omega_1 \times \Omega_2$. Si ν_1 et ν_2 sont deux solutions du problème, elle coïncident sur la classe $\mathcal{C}$, les ensembles C_n sont de mesure finie pour ν_1 et ν_2, donc $\nu_1 = \nu_2$ par le corollaire 2.1.4.

2.2. Théorème de Fubini

Sur chacun des deux espaces mesurables $(\Omega_i, \mathcal{A}_i)$, $i = 1, 2$, est donnée une mesure σ-finie μ_i. Si f est une fonction $\mathcal{A}_1 \otimes \mathcal{A}_2$-mesurable positive, on veut pouvoir considérer la quantité

$$\mathrm{I}(f) = \int_{\Omega_2} \Bigl(\int_{\Omega_1} f(\omega_1, \omega_2)\, d\mu_1(\omega_1) \Bigr)\, d\mu_2(\omega_2).$$

L'écriture de cette formule demande l'obtention de deux informations préalables :

1– l'application $\omega_1 \in \Omega_1 \to f(\omega_1, \omega_2)$ est $\mathcal{A}_1$-mesurable pour tout $\omega_2 \in \Omega_2$;
2– l'application $\omega_2 \in \Omega_2 \to \int_{\Omega_1} f(\omega_1, \omega_2)\, d\mu_1(\omega_1)$ est $\mathcal{A}_2$-mesurable.

On va donc commencer par établir le lemme suivant.

Lemme. *Si μ_1 est une mesure sur $(\Omega_1, \mathcal{A}_1)$ et si $\mathrm{B} \in \mathcal{A}_1$ est de mesure finie pour cette mesure, toute fonction f qui est $\mathcal{A}_1 \otimes \mathcal{A}_2$-étagée réelle vérifie*

1– l'application $\omega_1 \in \Omega_1 \to f(\omega_1, \omega_2)$ est $\mathcal{A}_1$-mesurable pour tout $\omega_2 \in \Omega_2$;
2– l'application $\omega_2 \in \Omega_2 \to \int_{\mathrm{B}} f(\omega_1, \omega_2)\, d\mu_1(\omega_1)$ est $\mathcal{A}_2$-mesurable.

Démonstration. On considère la classe E des fonctions réelles bornées f sur $\Omega_1 \times \Omega_2$ qui vérifient les propriétés 1 et 2 de l'énoncé. Il est clair que E est un espace vectoriel de fonctions (par les propriétés des fonctions mesurables pour le point 1, et la linéarité de l'intégrale pour le point 2). Si une suite $(f_n) \subset \mathrm{E}$ bornée par M tend simplement vers f sur $\Omega_1 \times \Omega_2$, alors f vérifie la première propriété comme limite de fonctions mesurables, et la deuxième par le théorème de convergence dominée appliqué à μ_1 avec le majorant intégrable $\mathrm{M}\,\mathbf{1}_{\mathrm{B}}$, qui montre que l'application $\omega_2 \in \Omega_2 \to \int_{\mathrm{B}} f(\omega_1, \omega_2)\, d\mu_1(\omega_1)$ est limite des applications mesurables $\omega_2 \in \Omega_2 \to \int_{\mathrm{B}} f_n(\omega_1, \omega_2)\, d\mu_1(\omega_1)$. L'espace E est donc un ESCSB. Il est clair que E contient la fonction $\mathbf{1}$.

Désignons par $\mathcal{C}$ la classe des pavés $\mathrm{A}_1 \times \mathrm{A}_2$, $\mathrm{A}_i \in \mathcal{A}_i$. On a vu que $\mathcal{C}$ est stable par intersection finie. On vérifie assez facilement que $\mathbf{1}_{\mathrm{C}} \in \mathrm{E}$ pour tout $\mathrm{C} \in \mathcal{C}$; en effet, si $f = \mathbf{1}_{\mathrm{A}_1 \times \mathrm{A}_2}$, la fonction $f(., \omega_2)$ est égale à $\mathbf{1}_{\mathrm{A}_2}(\omega_2)\, \mathbf{1}_{\mathrm{A}_1}$, qui est $\mathcal{A}_1$-mesurable pour tout ω_2, et la fonction de la condition 2 est $\mu_1(\mathrm{A}_1 \cap \mathrm{B}) \mathbf{1}_{\mathrm{A}_2}$, qui est $\mathcal{A}_2$-mesurable. Puisque E est un ESCSB qui contient $\mathbf{1}$ et toutes les fonctions $\mathbf{1}_{\mathrm{C}}$, $\mathrm{C} \in \mathcal{C}$, on sait par le théorème 2.1.2 que E contient $\mathcal{L}_\infty(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$, en particulier E contient les fonctions $\mathcal{A}_1 \otimes \mathcal{A}_2$-étagées.

2.2.a. *Construction de la mesure produit tensoriel*

Après cet intermède classe monotone, reprenons la construction. Si μ_1 est σ-finie, on peut trouver une suite (B_n) croissante de sous-ensembles de Ω_1 telle que $\mu_1(\mathrm{B}_n) < +\infty$ pour tout n et $\Omega_1 = \bigcup_n \mathrm{B}_n$. Si f est une fonction $\mathcal{A}_1 \otimes \mathcal{A}_2$-étagée positive finie, l'application $\omega_2 \in \Omega_2 \to \int_{\mathrm{B}_n} f(\omega_1, \omega_2)\, d\mu_1(\omega_1)$ est $\mathcal{A}_2$-mesurable par le lemme précédent. En passant à la limite monotone, on obtient que l'application

$$\omega_2 \in \Omega_2 \to \int_{\Omega_1} f(\omega_1, \omega_2)\, d\mu_1(\omega_1)$$

est $\mathcal{A}_2$-mesurable (à valeurs dans $[0, +\infty]$). Si f est une fonction $\mathcal{A}_1 \otimes \mathcal{A}_2$-mesurable positive, elle est limite croissante d'une suite (f_n) de fonctions étagées positives par la proposition 1.1.2 ; on voit que les propriétés 1 et 2 passent à la limite,

1– l'application $\omega_1 \in \Omega_1 \to f(\omega_1, \omega_2)$ est $\mathcal{A}_1$-mesurable pour tout $\omega_2 \in \Omega_2$;
2– l'application $\omega_2 \in \Omega_2 \to \int_{\Omega_1} f(\omega_1, \omega_2)\, d\mu_1(\omega_1)$ est $\mathcal{A}_2$-mesurable.

On peut donc considérer la quantité

$$\mathrm{I}(f) = \int_{\Omega_2} \Bigl(\int_{\Omega_1} f(\omega_1, \omega_2)\, d\mu_1(\omega_1) \Bigr)\, d\mu_2(\omega_2),$$

qui est un élément de $[0, +\infty]$.

Il est clair que $\mathrm{I}(f_1 + f_2) = \mathrm{I}(f_1) + \mathrm{I}(f_2)$, par les propriétés d'additivité des deux intégrales successives. Soit (f_n) une suite croissante de fonctions mesurables positives, et posons $f = \lim_n f_n$; on aura dans un premier temps, par les propriétés de l'intégrale par rapport à μ_1

$$\int_{\Omega_1} f(\omega_1, \omega_2)\, d\mu_1(\omega_1, \omega_2) = \lim_n \nearrow \int_{\Omega_1} f_n(\omega_1, \omega_2)\, d\mu_1(\omega_1)$$

puis par une deuxième application de la convergence monotone, avec μ_2 cette fois

$$\int_{\Omega_2} d\mu_2(\omega_2) \Bigl(\int_{\Omega_1} f(\omega_1, \omega_2)\, d\mu_1(\omega_1) \Bigr) = \lim_n \int_{\Omega_2} d\mu_2(\omega_2) \Bigl(\int_{\Omega_1} f_n(\omega_1, \omega_2)\, d\mu_1(\omega_1) \Bigr),$$

c'est-à-dire que

$$\mathrm{I}\Bigl(\lim_n f_n \Bigr) = \lim_n \mathrm{I}(f_n).$$

Soit (u_k) une suite de fonctions mesurables positives, et posons $u = \sum_{k=0}^{+\infty} u_k$; on aura en utilisant l'additivité de I et la continuité de I pour les suites croissantes

$$\mathrm{I}\Bigl(\sum_{k=0}^{+\infty} u_k \Bigr) = \sum_{k=0}^{+\infty} \mathrm{I}(u_k).$$

Il en résulte immédiatement que la fonction ν sur $\mathcal{A}_1 \otimes \mathcal{A}_2$ définie par $\nu(\mathrm{A}) = \mathrm{I}(\mathbf{1}_\mathrm{A})$ pour tout $\mathrm{A} \in \mathcal{A}_1 \otimes \mathcal{A}_2$ est une mesure sur $\mathcal{A}_1 \otimes \mathcal{A}_2$. Supposons en effet que (X_n) soit une suite de parties disjointes appartenant à $\mathcal{A}_1 \otimes \mathcal{A}_2$; alors la fonction indicatrice de $\mathrm{X} = \bigcup_{n=0}^{+\infty} \mathrm{X}_n$ est égale à $\sum_{n=0}^{+\infty} \mathbf{1}_{\mathrm{X}_n}$, donc

$$\nu(\mathrm{X}) = \mathrm{I}(\mathbf{1}_\mathrm{X}) = \mathrm{I}\Bigl(\sum_{n=0}^{+\infty} \mathbf{1}_{\mathrm{X}_n} \Bigr) = \sum_{n=0}^{+\infty} \mathrm{I}(\mathbf{1}_{\mathrm{X}_n}) = \sum_{n=0}^{+\infty} \nu(\mathrm{X}_n).$$

Si $\mathrm{B} = \mathrm{A}_1 \times \mathrm{A}_2$, la fonction $f = \mathbf{1}_\mathrm{B}$ est égale à $\mathbf{1}_{\mathrm{A}_1}(\omega_1)\, \mathbf{1}_{\mathrm{A}_2}(\omega_2)$, donc

$$\int_{\Omega_1} f(\omega_1, \omega_2)\, d\mu_1(\omega_1) = \mu_1(\mathrm{A}_1)\, \mathbf{1}_{\mathrm{A}_2}(\omega_2)$$

et ensuite

$$\nu(\mathrm{A}_1 \times \mathrm{A}_2) = \mathrm{I}(f) = \int \bigl(\mu_1(\mathrm{A}_1)\, \mathbf{1}_{\mathrm{A}_2}(\omega_2) \bigr)\, d\mu_2(\omega_2) = \mu_1(\mathrm{A}_1)\, \mu_2(\mathrm{A}_2).$$

Théorème. *Supposons μ_i σ-finie sur $(\Omega_i, \mathcal{A}_i)$, $i = 1, 2$. Il existe une unique mesure $\mu_1 \otimes \mu_2$ sur $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$ telle que*

$$(\otimes) \qquad \forall \mathrm{A}_1 \in \mathcal{A}_1,\ \forall \mathrm{A}_2 \in \mathcal{A}_2, \quad (\mu_1 \otimes \mu_2)(\mathrm{A}_1 \times \mathrm{A}_2) = \mu_1(\mathrm{A}_1)\mu_2(\mathrm{A}_2).$$

Démonstration. On a déjà montré l'unicité à la proposition 2.1.5, et on vient de construire une mesure ν qui vérifie l'égalité ci-dessus.

Remarque. Si N est μ_1-négligeable, l'ensemble $\mathrm{N} \times \Omega_2$ est $(\mu_1 \otimes \mu_2)$-négligeable.

En effet,

$$(\mu_1 \otimes \mu_2)(\mathrm{N} \times \Omega_2) = \mu_1(\mathrm{N})\, \mu_2(\Omega_2) = 0 \,.\, \mu_2(\Omega_2) = 0,$$

même si $\mu_2(\Omega_2) = +\infty$.

Revenons à l'application $f \to \mathrm{I}(f)$ définie avant le théorème. Par linéarité on obtient pour toute fonction $\mathcal{A}_1 \otimes \mathcal{A}_2$-étagée $f = \sum_{i=1}^n b_i \mathbf{1}_{\mathrm{B}_i}$ (où $\mathrm{B}_i \in \mathcal{A}_1 \otimes \mathcal{A}_2$, $i = 1, \ldots, n$)

$$\int f\, d\nu = \sum_{i=1}^n b_i\, \nu(\mathrm{B}_i) = \sum_{i=1}^n b_i\, \mathrm{I}(\mathbf{1}_{\mathrm{B}_i}) = \mathrm{I}(f) = \int_{\Omega_2} \Bigl(\int_{\Omega_1} f(\omega_1, \omega_2)\, d\mu_1(\omega_1) \Bigr)\, d\mu_2(\omega_2).$$

Puisque l'intégrale par rapport à ν et la fonctionnelle I sont continues pour les suites croissantes, on en déduit le même résultat pour toute fonction f qui est $\mathcal{A}_1 \otimes \mathcal{A}_2$-mesurable positive (en effet, une telle f est limite croissante d'une suite de fonctions $\mathcal{A}_1 \otimes \mathcal{A}_2$-étagées, par la proposition 1.1.2). La formule de définition de I nous donne donc le théorème de Fubini positif qui suit.

2.2.b. *Intégration par rapport à la mesure produit*

Théorème 2.2.1 (Fubini positif). *Supposons μ_i σ-finie sur $(\Omega_i, \mathcal{A}_i)$, $i = 1, 2$. Pour toute fonction mesurable f de $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$ à valeurs dans $[0, +\infty]$,*

$$\begin{aligned} \int_{\Omega_1 \times \Omega_2} f\, d(\mu_1 \otimes \mu_2) &= \int_{\Omega_2} \Bigl(\int_{\Omega_1} f(\omega_1, \omega_2)\, d\mu_1(\omega_1) \Bigr)\, d\mu_2(\omega_2) \\ &= \int_{\Omega_1} \Bigl(\int_{\Omega_2} f(\omega_1, \omega_2)\, d\mu_2(\omega_2) \Bigr)\, d\mu_1(\omega_1). \end{aligned}$$

Démonstration. On a construit une mesure ν qui vérifie la première égalité ci-dessus ; si on refait le travail en échangeant les rôles des deux espaces dans la construction, on construira une autre mesure ν_1 qui vérifiera par construction

$$\int_{\Omega_1 \times \Omega_2} f\, d\nu_1 = \int_{\Omega_1} \Bigl(\int_{\Omega_2} f(\omega_1, \omega_2)\, d\mu_2(\omega_2) \Bigr)\, d\mu_1(\omega_1)\,;$$

mais ν_1 vérifie aussi la propriété $(\otimes)$; par l'unicité démontrée à la proposition 2.1.5, on a $\nu_1 = \nu = \mu_1 \otimes \mu_2$, d'où la deuxième égalité ci-dessus.

Remarque. Si f_i est une fonction mesurable positive sur $(\Omega_i, \mathcal{A}_i)$, $i = 1, 2$, on a

$$\int f_1(\omega_1) f_2(\omega_2)\, d(\mu_1 \otimes \mu_2)(\omega_1, \omega_2) = \Big(\int f_1(\omega_1)\, d\mu_1(\omega_1)\Big)\Big(\int f_2(\omega_2)\, d\mu_2(\omega_2)\Big).$$

Exercice. Si f est mesurable ≥ 0 montrer que

$$\int f\, d\mu = \int_0^{+\infty} \mu\big(\{f > t\}\big)\, dt.$$

Théorème 2.2.2 de Fubini. *Supposons μ_i σ-finie sur $(\Omega_i, \mathcal{A}_i)$, $i = 1, 2$. Pour toute fonction f mesurable réelle ou complexe sur $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$, et intégrable par rapport à $\mu_1 \otimes \mu_2$, on peut dire que la fonction $\mathcal{A}_1$-mesurable*

$$\omega_1 \to f(\omega_1, \omega_2)$$

est μ_1-intégrable pour presque tout ω_2 ; la fonction presque-partout définie

$$\omega_2 \to \int_{\Omega_1} f(\omega_1, \omega_2)\, d\mu_1(\omega_1)$$

est μ_2-intégrable et

$$\int_{\Omega_1 \times \Omega_2} f\, d(\mu_1 \otimes \mu_2) = \int_{\Omega_2} \Big(\int_{\Omega_1} f(\omega_1, \omega_2)\, d\mu_1(\omega_1)\Big)\, d\mu_2(\omega_2).$$

Démonstration. On a par hypothèse

$$\mathrm{I} := \int_{\Omega_1 \times \Omega_2} |f(\omega_1, \omega_2)|\, d(\mu_1 \otimes \mu_2)(\omega_1, \omega_2) < +\infty.$$

Par le théorème de Fubini positif,

$$\mathrm{I} = \int_{\Omega_2} \Big(\int_{\Omega_1} |f(\omega_1, \omega_2)|\, d\mu_1(\omega_1)\Big)\, d\mu_2(\omega_2) < +\infty$$

ce qui implique que la parenthèse est finie μ_2-presque partout. Considérons l'ensemble mesurable μ_2-négligeable défini par

$$\mathrm{N} = \{\omega_2 \in \Omega_2 : \int_{\Omega_1} |f(\omega_1, \omega_2)|\, d\mu_1(\omega_1) = +\infty\}.$$

Modifions la fonction f en posant $g(\omega_1, \omega_2) = f(\omega_1, \omega_2)$ si $\omega_2 \notin \mathrm{N}$, et $g(\omega_1, \omega_2) = 0$ si $\omega_2 \in \mathrm{N}$. Pour tout $\omega_2 \notin \mathrm{N}$, on a évidemment

$$\int_{\Omega_1} f(\omega_1, \omega_2)\, d\mu_1(\omega_1) = \int_{\Omega_1} g(\omega_1, \omega_2)\, d\mu_1(\omega_1).$$

De plus f et g sont égales $\mu_1 \otimes \mu_2$-presque partout (car $\Omega_1 \times \mathrm{N}$ est négligeable), donc

$$\int_{\Omega_1\times\Omega_2} f(\omega_1,\omega_2)\, d(\mu_1 \otimes \mu_2)(\omega_1,\omega_2) = \int_{\Omega_1\times\Omega_2} g(\omega_1,\omega_2)\, d(\mu_1 \otimes \mu_2)(\omega_1,\omega_2).$$

Posons $g = g_+ - g_-$; on sait que $\mathrm{G}_+(\omega_2) = \int g_+(\omega_1,\omega_2)\, d\mu(\omega_1)$ est μ_2-intégrable, finie partout, ainsi que $\mathrm{G}_-(\omega_2) = \int g_-(\omega_1,\omega_2)\, d\mu(\omega_1)$. La différence $\mathrm{G}_+ - \mathrm{G}_-$ est presque partout égale à la fonction presque partout définie $\mathrm{F}(\omega_2) = \int f(\omega_1,\omega_2)\, d\mu(\omega_1)$, qui est intégrable par différence. Finalement

$$\int f\, d(\mu_1 \otimes \mu_2) = \int g\, d(\mu_1 \otimes \mu_2) = \int g_+\, d(\mu_1 \otimes \mu_2) - \int g_-\, d(\mu_1 \otimes \mu_2) =$$

$$\int \mathrm{G}_+(\omega_2)\, d\mu(\omega_2) - \int \mathrm{G}_-(\omega_2)\, d\mu(\omega_2)$$

par le théorème de Fubini positif, et cette dernière expression est égale à

$$\int \mathrm{F}(\omega_2)\, d\mu(\omega_2),$$

comme attendu.

2.3. Changements de variables

2.3.a. *Changements de variables linéaires*

On veut effectuer le changement de variable $y = \mathrm{A}x$ dans une intégrale sur $\mathbb{R}^d$, où A est une matrice inversible. Ceci est réglé par la proposition suivante.

Proposition 2.3.1. *Soit* A *une application linéaire bijective de* $\mathbb{R}^d$ *; pour toute fonction borélienne positive* g *sur* $\mathbb{R}^d$*, on a*

$$\int_{\mathbb{R}^d} g(y)\, dy = \int_{\mathbb{R}^d} g(\mathrm{A}x)\, |\det \mathrm{A}|\, dx.$$

La stratégie de la preuve consiste à décomposer le changement linéaire général en une succession de trois changements linéaires plus faciles à analyser. Le premier est très simple, et concerne les changements dont la matrice P est une matrice permutation, qui échange simplement les coordonnées de $\mathbb{R}^d$. Le déterminant d'une matrice permutation vaut ± 1, selon la signature de la permutation. Par exemple, pour $d = 2$, envisageons le changement de variable

$$y_1 = x_2, \quad y_2 = x_1$$

correspondant à l'application linéaire P définie par $\mathrm{P}(x_1, x_2) = (x_2, x_1)$. Il est clair que l'image ν par P de la mesure de Lebesgue $\lambda_2 = \lambda \otimes \lambda$ de $\mathbb{R}^2$ est une mesure qui vérifie la propriété

$$\forall \mathrm{A}, \mathrm{B} \in \mathcal{B}_{\mathbb{R}}, \quad \nu(\mathrm{A} \times \mathrm{B}) = \lambda_2(\mathrm{B} \times \mathrm{A}) = \lambda(\mathrm{A})\, \lambda(\mathrm{B})$$

donc $\nu = \lambda_2$ par l'unicité du produit tensoriel.

Comme toute matrice A peut s'écrire comme produit PLU d'une matrice de permutation P et de matrices triangulaires L, U inférieure, supérieure, il suffit de montrer la formule dans chacun des cas, et d'obtenir le cas général en composant les changements élémentaires. Or le traitement des changements triangulaires résulte immédiatement de Fubini. Écrivons le cas de dimension 2, en supposant que le changement triangulaire soit donné par

$$y_1 = ax_1 + bx_2, \quad y_2 = cx_2.$$

Le déterminant de la matrice triangulaire vaut $ac \neq 0$. On a, en utilisant simplement les changements de variables linéaires sur $\mathbb{R}$ et Fubini positif

$$\int_{\mathbb{R}^2} g(\mathrm{A}x)\, |\det \mathrm{A}|\, dx = \int_{\mathbb{R}^2} g(ax_1 + bx_2, cx_2)\, |a|\, |c|\, dx_1\, dx_2$$

$$= \int_{\mathbb{R}} \Bigl(\int_{\mathbb{R}} g(ax_1 + bx_2, cx_2)\, |a|\, dx_1 \Bigr) |c|\, dx_2 = \int_{\mathbb{R}} \Bigl(\int_{\mathbb{R}} g(y_1, cx_2)\, dy_1 \Bigr) |c|\, dx_2$$
$$= \int_{\mathbb{R}} \Bigl(\int_{\mathbb{R}} g(y_1, cx_2)\, |c|\, dx_2 \Bigr) dy_1 = \int_{\mathbb{R}} \Bigl(\int_{\mathbb{R}} g(y_1, y_2)\, dy_2 \Bigr) dy_1 = \int_{\mathbb{R}^2} g(y)\, dy.$$

Corollaire 2.3.2. *Soient $v_1, \ldots, v_d$ des vecteurs de $\mathbb{R}^d$; le parallélotope* P *engendré par ces vecteurs, c'est-à-dire*

$$\mathrm{P} = \Bigl\{ \sum_{j=1}^{d} c_j\, v_j : 0 \leq c_j \leq 1 \Bigr\},$$

a pour volume $|\det(v_1, \ldots, v_d)|$.

Si les vecteurs sont liés, le volume est clairement nul (le parallélotope est dégénéré, contenu dans un hyperplan, qui est de mesure nulle pour Lebesgue). Dans le cas contraire, considérons la transformation linéaire inversible A de $\mathbb{R}^d$ qui envoie chaque vecteur $\mathbf{e}_j$ de la base canonique sur le vecteur v_j, pour $j = 1, \ldots, d$. Il est clair que P est l'image du cube $[0,1]^d$ par cette transformation. Le résultat découle donc de la formule de changement de variable linéaire.

Le prochain corollaire énonce une propriété intuitivement évidente de la mesure de Lebesgue, mais dont la démonstration nous aura demandé pas mal de travail.

Corollaire. *La mesure de Lebesgue sur $\mathbb{R}^d$ est invariante par rotation.*

C'est évident par la formule du changement de variable linéaire, puisque le déterminant des rotations est égal à ± 1.

2.3.b. *Le théorème de changement de variable*

Théorème 2.3.3. *Soit φ une bijection d'un ouvert* U *de $\mathbb{R}^d$ sur un autre ouvert* V *de $\mathbb{R}^d$, de classe C^1 ainsi que son inverse ; désignons par $\mathrm{J}_x\varphi$ la matrice jacobienne de φ au point $x \in \mathrm{U}$. Pour toute fonction mesurable positive g sur* V*, on a*

$$\int_{\mathrm{V}} g(y)\, dy = \int_{\mathrm{U}} g(\varphi(x))\, |\det(\mathrm{J}_x\varphi)|\, dx.$$

L'une des stratégies de preuve consiste à utiliser le théorème des fonctions implicites pour décomposer, au voisinage de chaque point, le changement général en produit de changements «triangulaires», tels que

$$y_1 = f_1(x_1, x_2), \quad y_2 = f_2(x_2),$$

qui se traitent à nouveau par Fubini et changement de variable dans $\mathbb{R}$. L'autre stratégie est d'attaquer à l'arme blanche avec un peu de calcul différentiel.

On va à peu près rédiger le cas $d = 2$, en supposant φ de classe C^2 pour simplifier un tout petit peu. Les hypothèses impliquent que la différentielle $d_x\varphi$ est inversible en tout point $x \in \mathrm{U}$, et que φ est un difféomorphisme, donc les images des ouverts contenus dans U sont des ouverts contenus dans V. On va comparer la mesure de Lebesgue λ_{V} sur l'ouvert V avec la mesure ν définie sur $(\mathrm{V}, \mathcal{B}_{\mathrm{V}})$ par

$$\forall \mathrm{A} \in \mathcal{B}_{\mathrm{V}}, \quad \nu(\mathrm{A}) = \int_{\mathrm{U}} \mathbf{1}_{\mathrm{A}}(\varphi(x))\, |\det(\mathrm{J}_x\varphi)|\, dx = \int_{\varphi^{-1}(\mathrm{A})} |\det(\mathrm{J}_x\varphi)|\, dx.$$

Il est clair que ν est une mesure sur V (voir l'additif 1 plus loin). La deuxième forme montre immédiatement que $\nu(\mathrm{K}) < +\infty$ pour tout compact $\mathrm{K} \subset \mathrm{V}$, car $\varphi^{-1}(\mathrm{K})$ est un compact contenu dans U, sur lequel la fonction continue $|\det(\mathrm{J}_x\varphi)|$ est bornée.

Soit U_1 un ouvert non vide contenu dans U, soit $\mathrm{V}_1 = \varphi(\mathrm{U}_1)$ l'ouvert image et soit τ un nombre réel tel que $\tau < \lambda_{\mathrm{V}}(\mathrm{V}_1)$; l'ouvert U_1 est réunion de carrés compacts, de côtés parallèles aux axes, de la forme $[k2^{-n}, (k+1)2^{-n}] \times [\ell 2^{-n}, (\ell+1)2^{-n}]$, où $n \geq 0$ et $k, \ell \in \mathbb{Z}$. L'image V_1 de U_1 est, de son côté, réunion de la famille dénombrable des images de ces carrés. On peut donc trouver un nombre fini de ces carrés, $(\mathrm{K}_i)_{i=1}^m$, dont la réunion

$$\mathrm{K} = \bigcup_{i=1}^m \mathrm{K}_i \subset \mathrm{U}_1$$

a une image $\varphi(\mathrm{K}) \subset \mathrm{V}_1$ assez grande pour que $\tau < \lambda_{\mathrm{V}}(\varphi(\mathrm{K}))$. Sur le compact K, les dérivées secondes des composantes de φ sont bornées ; il en résulte (accroissements finis, convexité des carrés) que $x \to d_x\varphi$ est lipschitzienne d'une certaine constante M sur chacun des carrés K_i qui forment K. De plus, l'application continue $x \to \|(d_x\varphi)^{-1}\|$ est également bornée sur K, disons par le même M.

Soit $\varepsilon > 0$ donné, et choisissons $h > 0$ tel que $2\mathrm{M}^2\, h < \varepsilon$; on redécoupe les carrés K_i qui forment K en un nombre fini de carrés C_j plus petits, dont la réunion est toujours égale à K, dont les intérieurs sont disjoints et dont le demi-côté est $\leq h$. Soit C un de ces petits carrés C_j, et soit c son centre ; désignons par φ_c l'application affine tangente $\varphi_c : x \to \varphi(c) + (d_c\varphi)\,(x - c)$; pour tout point $x_0 = c + v$ de C, on a par la majoration des accroissements finis, en notant que $\|x - c\| \leq \sqrt{2}\, h$ pour tout $x \in \mathrm{C}$

$$\begin{aligned} \|\varphi(x_0) - \varphi_c(x_0)\| &= \|\varphi(c+v) - \varphi(c) - (d_c\varphi)\, v\| \\ &\leq \sup_{x \in \mathrm{C}} \|d_x\varphi - d_c\varphi\|\, \|v\| \leq \mathrm{M}\sqrt{2}\, h\, \|v\| \leq 2\mathrm{M}\, h^2. \end{aligned}$$

L'idée de la preuve est que l'image du carré C par φ est très proche du parallélogramme $\mathrm{P} = \varphi_c(\mathrm{C})$, image de C par l'application affine tangente φ_c, parallélogramme dont la mesure se calcule par la formule *linéaire* 2.3.2. Cette proximité est conséquence de l'inégalité précédente : puisque $y = \varphi(x_0)$ représente un point quelconque dans $\varphi(\mathrm{C})$, et puisque $\varphi_c(x_0)$ est dans P, on a

$$(1) \qquad \forall y \in \varphi(\mathrm{C}), \quad d(y, \mathrm{P}) \le 2\mathrm{M}\, h^2.$$

On va en fait montrer que l'image $\varphi(\mathrm{C})$ est contenue dans l'image par φ_c d'un carré C_ε un peu plus grand, de même centre c mais de demi-côté $(1+\varepsilon)\, h$. Pour réaliser cet objectif il nous faut minorer la distance entre un point quelconque $y_0 = \varphi_c(x_0)$ de $\mathrm{P} = \varphi_c(\mathrm{C})$ et un point quelconque $y_1 = \varphi_c(x_1)$ extérieur à $\varphi_c(\mathrm{C}_\varepsilon)$. Ceci provient de l'estimation de la norme de $(d_x\varphi)^{-1}$ sur K, et du simple fait que, comme x_0 est dans C et x_1 hors de C_ε, leur distance est au moins εh. On a ainsi

$$\|y_0 - y_1\| = \|\varphi_c(x_0) - \varphi_c(x_1)\| = \|(d_c\varphi)\,(x_0 - x_1)\| \ge \|(d_c\varphi)^{-1}\|^{-1}\|x_0 - x_1\| \ge \mathrm{M}^{-1}\varepsilon h,$$

ce qui signifie que $d(y_1, \mathrm{P}) \ge \mathrm{M}^{-1}\varepsilon h$, pour tout point y_1 qui n'est pas dans $\varphi_c(\mathrm{C}_\varepsilon)$. Or $\mathrm{M}^{-1}\varepsilon h > 2\mathrm{M}h^2$ par notre choix de h, et il résulte donc de l'inégalité (1) et de la précédente que l'image $\varphi(\mathrm{C})$ ne peut pas sortir de $\varphi_c(\mathrm{C}_\varepsilon)$, ce qui donne l'estimation

$$\lambda_\mathrm{V}(\varphi(\mathrm{C})) \le \lambda_\mathrm{V}(\varphi_c(\mathrm{C}_\varepsilon)) = |\det(d_c\varphi)|\, \lambda_\mathrm{U}(\mathrm{C}_\varepsilon) = (1+\varepsilon)^2 |\det(d_c\varphi)|\, \lambda_\mathrm{U}(\mathrm{C}).$$

On a approché le volume de $\varphi(\mathrm{U}_1)$ par une réunion $\varphi(\mathrm{K})$ d'images de petits carrés C_j tels que le carré C précédent, chacun de centre c_j, pour $j = 1, \ldots, \mathrm{N}$. On en déduit, en introduisant dans la notation les matrices jacobiennes, $\det(d_{c_j}\varphi) = \det(\mathrm{J}_{c_j}\varphi)$

$$\lambda_\mathrm{V}(\varphi(\mathrm{K})) \le \sum_{j=1}^{\mathrm{N}} \lambda_\mathrm{V}(\varphi(\mathrm{C}_j)) \le (1+\varepsilon)^2 \sum_{j=1}^{\mathrm{N}} |\det(\mathrm{J}_{c_j}\varphi)|\, \lambda_\mathrm{U}(\mathrm{C}_j).$$

Quand la taille h des carrés (C_j) qui couvrent K tend vers 0, on obtient par la convergence des sommes de Riemann, puis en faisant tendre ε vers 0

$$\tau < \lambda_\mathrm{V}(\varphi(\mathrm{K})) \le \int_\mathrm{K} |\det(\mathrm{J}_x\varphi)|\, dx \le \int_{\mathrm{V}_1} \mathbf{1}_{\mathrm{V}_1}(\varphi(x))\, |\det(\mathrm{J}_x\varphi)|\, dx = \nu(\mathrm{V}_1),$$

et comme τ est quelconque $< \lambda_\mathrm{V}(\mathrm{V}_1)$, on déduit que $\lambda_\mathrm{V}(\mathrm{V}_1) \le \nu(\mathrm{V}_1)$. Si A est un borélien de V et m un réel tel que $\nu(\mathrm{A}) < m$, on peut d'après la théorie générale (voir l'additif 2 plus loin) trouver un ouvert $\mathrm{V}_1 \subset \mathrm{V}$ tel que $\mathrm{A} \subset \mathrm{V}_1$ et $\nu(\mathrm{V}_1) < m$. On a alors

$$\lambda_\mathrm{V}(\mathrm{A}) \le \lambda_\mathrm{V}(\mathrm{V}_1) \le \nu(\mathrm{V}_1) < m,$$

ce qui implique $\lambda_\mathrm{V}(\mathrm{A}) \le \nu(\mathrm{A})$ pour tous les boréliens $\mathrm{A} \in \mathcal{B}_\mathrm{V}$. On en déduit comme d'habitude, pour toute fonction mesurable positive g sur V,

$$(*) \qquad \int_\mathrm{V} g(y)\, dy \le \int_\mathrm{U} g(\varphi(x))\, |\det(\mathrm{J}_x\varphi)|\, dx.$$

En utilisant le changement inverse on peut transformer cette inégalité en égalité : considérons la fonction mesurable positive $h(x) = g(\varphi(x))\, |\det(\mathrm{J}_x\varphi)|$ sur U et appliquons à la transformation inverse $\psi = \varphi^{-1}$ le résultat déjà démontré ; on obtient, compte tenu du fait que $(\mathrm{J}_x\varphi) \circ (\mathrm{J}_y\psi) = \mathrm{Id}$ lorsque $x = \psi(y)$,

$$\begin{aligned} \int_\mathrm{U} h(x)\, dx &\le \int_\mathrm{V} h(\psi(y))\, |\det(\mathrm{J}_y\psi)|\, dy \\ &= \int_\mathrm{V} g(y)\, |\det(\mathrm{J}_y\psi)|^{-1}\, |\det(\mathrm{J}_y\psi)|\, dy = \int_\mathrm{V} g(y)\, dy. \end{aligned}$$

Remarque. Si on n'était pas dans une auberge de théorie de l'intégration, on aurait mis beaucoup moins de morceaux de théorie de la mesure dans notre soupe. On aurait restreint la démonstration de l'inégalité principale $(*)$ au cas d'une fonction g continue à support compact dans V, et on aurait fait beaucoup plus appel à des arguments venant de la théorie de l'intégrale de Riemann.

Additif 1. Si $\mathrm{A} = \bigcup_{n=0}^{+\infty} \mathrm{A}_n$ est la réunion d'une suite (A_n) de boréliens deux à deux disjoints contenus dans V, on a pour tout $x \in \mathrm{U}$

$$\mathbf{1}_{\mathrm{A}}(\varphi(x)) = \sum_{n=0}^{+\infty} \mathbf{1}_{\mathrm{A}_n}(\varphi(x)),$$

donc par la proposition 1.1.4 sur l'intégration des séries de fonctions positives,

$$\nu(\mathrm{A}) = \int_{\mathrm{U}} \mathbf{1}_{\mathrm{A}}(\varphi(x))\,|\det(\mathrm{J}_x\varphi)|\,dx = \sum_{n=0}^{+\infty} \int_{\mathrm{U}} \mathbf{1}_{\mathrm{A}_n}(\varphi(x))\,|\det(\mathrm{J}_x\varphi)|\,dx = \sum_{n=0}^{+\infty} \nu(\mathrm{A}_n),$$

et ν est bien une mesure sur $(\mathrm{V}, \mathcal{B}_{\mathrm{V}})$.

Additif 2. L'ouvert V est la réunion de la suite croissante d'ouverts

$$\mathrm{V}^{(n)} = \{y \in \mathrm{V} : d(y, \mathrm{V}^c) > 2^{-n}, \quad \|y\| < 2^n\}.$$

Ces ouverts ont une adhérence compacte contenue dans V, donc $\nu(\mathrm{V}^{(n)}) < +\infty$. D'après le théorème d'approximation 1.2.3, valable pour toute mesure sur la tribu borélienne $\mathcal{B}_{\mathrm{X}}$ d'un espace métrique X, qu'on applique ici à la mesure ν sur V et à l'ouvert de mesure finie $\mathrm{V}^{(n)}$, il existe pour tout borélien $\mathrm{A} \subset \mathrm{V}^{(n)}$ et tout $\varepsilon > 0$ un ouvert W tel que

$$\mathrm{A} \subset \mathrm{W} \subset \mathrm{V}^{(n)} \quad \text{et} \quad \nu(\mathrm{W}) - \nu(\mathrm{A}) < \varepsilon.$$

Si A est un borélien de V, on peut donc trouver pour tout $n \geq 1$ un ouvert W_n tel que $\mathrm{A} \cap \mathrm{V}^{(n)} \subset \mathrm{W}_n$ et $\nu(\mathrm{W}_n) - \nu(\mathrm{A} \cap \mathrm{V}^{(n)}) < 2^{-n}\varepsilon$. L'ouvert $\mathrm{W} = \bigcup_{n=1}^{+\infty} \mathrm{W}_n \subset \mathrm{V}$ contient $\mathrm{A} = \bigcup_{n=1}^{+\infty} \mathrm{A} \cap \mathrm{V}^{(n)}$ et

$$\nu(\mathrm{W}) - \nu(\mathrm{A}) \leq \sum_{n=1}^{+\infty} \big(\nu(\mathrm{W}_n) - \nu(\mathrm{A} \cap \mathrm{V}^{(n)})\big) < \varepsilon.$$