Cours de Mesure et Intégration

Thierry Fack

Professeur à l'Université de Lyon

Faculté de Sciences et Technologie
Institut Camille Jordan
Bât. Braconnier
43 bd. Du 11 Novembre 1918
69622 Villeurbanne Cedex

2012/13

Mesure et intégration - MAT3014L
Semestre d'automne 2012

PROGRAMME

1. Théorie de la mesure. Algèbres et tribus, questions d'engendrement, tribu borélienne. Mesures positives définies sur les algèbres et sur les tribus. Mesure extérieure. Fonctions mesurables.

2. Intégration des fonctions. Intégration des fonctions mesurables. Comparaison entre l'intégrale de Lebesgue et celle de Riemann. Fonctions absolument continues. Le théorème de convergence dominée de Lebesgue et ses conséquences.

3. Intégration sur les espaces produits. Mesure produit, théorème de Fubini. Mesure de Lebesgue sur R n . Théorème du changement de variables.

4. Probabilités. Mesures de probabilités, variables aléatoires.

Université Claude Bernard Lyon 1
Licence de Sciences et Technologies

COURS DE CALCUL INTÉGRAL

Sommaire

CHAPITRE I

PRIMITIVES ET INTÉGRALES

1. — Primitives d'une fonction

1.1. Intégrales indéfinie et définie. — Soient E un espace de Banach réel (par exemple $\mathbb{R}^n$ muni de la norme euclidienne), I un intervalle de $\mathbb{R}$ et $f : I \to E$ une fonction à valeurs dans E.

DÉFINITION 1. — *On appelle primitive de* f *sur* I *toute fonction continue* $F : I \to E$, *dérivable sur* I *sauf peut-être aux points d'un ensemble dénombrable* $D \subset I$, *et qui vérifie* $F'(x) = f(x)$ *pour* $x \in I - D$.

Si F est une primitive de f sur I, la fonction $x \to F(x) + C$ est aussi une primitive de f pour tout $C \in E$. Il résulte du théorème de la moyenne[1] que toutes les primitives de f sont de cette forme. Rappelons que le théorème de la moyenne affirme

---

[1] Voir par exemple le livre de J. DIEUDONNÉ, *Éléments d'Analyse*, tome I, chapitre VIII, théorème 8.5.1.

qu'une fonction continue $f:[a,b]\to E$, dérivable hors d'un ensemble dénombrable $D\subset[a,b]$ et telle que $\|f'(x)\|\le M$ quel que soit $x\notin D$ vérifie :

$$\|f(b)-f(a)\|\le M(b-a).$$

Soient alors F_1, F_2 deux primitives de f sur I et montrons que la différence $G=F_1-F_2$ est constante. Fixons $x_o\in I$ et $\varepsilon>0$. Comme $G'(x)=0$ sauf peut-être pour x dans un ensemble dénombrable D, on a $\|G'(x)\|\le\varepsilon$ pour $x\notin D$ et donc $\|G(x)-G(x_o)\|\le\varepsilon|x-x_o|$ pour tout $x\in I$ en vertu du théorème de la moyenne. En faisant tendre ε vers 0, on en déduit que $G=F_1-F_2$ est constante sur I.

La famille des fonctions $F(x)+C$ est appelée *l'intégrale indéfinie* de f ; on la note :

$$\int f(x)dx=F(x)+C.$$

Si a,b sont deux points quelconques de I, la différence $F(b)-F(a)$ est un vecteur qui ne dépend pas du choix de la primitive F de f sur I en vertu de ce qui précède. On l'appelle *l'intégrale définie* de f entre a et b et on le note :

$$\int_a^b f(x)dx=F(b)-F(a)\in E.$$

Ainsi :

$$\int \cos x\,dx=\sin x+C,\quad \int_0^{\frac{\pi}{2}}\cos x\,dx=\sin\tfrac{\pi}{2}-\sin 0=1.$$

On observera que, si f possède une primitive F, la fonction $\tilde{f}$ obtenue en changeant arbitrairement les valeurs de f aux points d'un ensemble dénombrable admet encore F pour primitive. Il s'ensuit que :

$$\int_a^b f(x)dx = \int_a^b \tilde{f}(x)dx,$$

de sorte que l'intégrale définie de f est inchangée lorsqu'on modifie les valeurs de f sur un ensemble dénombrable. Ainsi, pour toute fonction continue $f:[a,b] \to E$ dérivable sauf peut être aux points d'un ensemble dénombrable D, on a :

$$f(x) = f(a) + \int_a^x f'(t)dt$$

quel que soit $x \in [a,b]$ (les valeurs de $f'(t)$ aux points $t \in D$ peuvent être choisies arbitrairement).

Rappelons qu'une fonction $f:[a,b] \to E$ est dite *en escalier* sur $[a,b]$ s'il existe une subdivision

$$a = a_o < a_1 < ... < a_{n-1} < a_n = b$$

du segment $[a,b]$ telle que f soit égale à une constante c_i sur chacun des intervalles $]a_i, a_{i+1}[$. On notera $\mathcal{E}([a,b],E)$ l'ensemble des fonctions en escalier sur $[a,b]$, à valeurs dans E.

PROPOSITION 1. — *Toute fonction en escalier sur un intervalle fermé et borné* $[a,b] \subset \mathbb{R}$ *admet des primitives sur* $[a,b]$.

Soit $f \in \mathcal{E}([a,b],E)$ et définissons la fonction F en posant :

$$F(x) = c_i(x - a_i) + \sum_{k=0}^{i-1} c_k(a_{k+1} - a_k)$$

pour x dans l'intervalle $[a_i, a_{i+1}]$ ($0 \le i \le n-1$). On vérifie aisément que F est une primitive de f sur $[a,b]$, d'où la proposition 1. ■

1.2. Intégrale définie et aire. — L'intégrale définie d'une fonction positive s'interprète géométriquement comme une aire plane (cf. fig. 1). Plus précisément, on a :

Fig.1. — Interprétation de l’intégrale définie comme une aire.

PROPOSITION 2. — *L’intégrale définie entre* a *et* b *d’une fonction positive continue* $f:[a,b]\to\mathbb{R}$ *s’interprète comme l’aire plane du sous-graphe*

$$G(f)=\{(u,v)\in[a,b]\times\mathbb{R}\mid 0\le v\le f(u)\}\subset\mathbb{R}^2$$

de f *, i.e.*

$$\int_a^b f(x)dx = Aire\ de\ G(f).$$

Pour tout $x\in[a,b]$, notons $F(x)$ l’aire de la partie du plan (cf. fig. 2) définie par :

$$A(x)=\{(u,v)\in[a,x]\times\mathbb{R}\mid 0\le v\le f(u)\}\subset\mathbb{R}^2.$$

On a donc $F(a)=0$. Pour démontrer la proposition 2, il suffit d’établir que $F'(x)=f(x)$ pour tout $x\in]a,b[$, car F est continue et l’aire de $G(f)$ est par définition égale à $F(b)-F(a)=F(b)$. Fixons $x\in]a,b[$ et montrons que la dérivée à droite de F en x est égale à $f(x)$. Pour tout $\varepsilon>0$ il existe, en vertu de la continuité de f en x, un réel $\eta>0$ tel que :

$$(1)\ x\le u\le x+\eta \Rightarrow f(x)-\varepsilon\le f(u)\le f(x)+\varepsilon.$$

Pour $x<u$, l’accroissement $F(u)-F(x)$ s’interprète comme l’aire de la portion de plan :

$$\Delta = \{(t,s) \in [x,u] \times \mathbb{R} \mid 0 \leq s \leq f(t)\}.$$

Fig.2. — Encadrement de la variation de l'aire $A(x)$.

Un argument simple d'encadrement de l'aire de Δ par des aires de rectangles, basé sur l'inégalité (1), montre alors que l'on a pour $x < u \leq x + \eta$:

$$(u-x)(f(x)-\varepsilon) \leq Aire(\Delta) \leq (u-x)(f(x)+\varepsilon).$$

On a donc, lorsque $x < u \leq x + \eta$:

$$f(x)-\varepsilon \leq \frac{F(u)-F(x)}{u-x} \leq f(x)+\varepsilon,$$

ce qui prouve que F est dérivable à droite au point x et a pour dérivée $f(x)$. On démontre de même que F est dérivable à gauche au point x, de dérivée $f(x)$, d'où la proposition 2. ■

Observons que la notion d'aire plane n'a pas été ici précisément définie, de sorte la proposition 2 est plus une heuristique qu'une véritable proposition. En particulier, elle ne peut être considérée comme une véritable preuve de l'existence de primitives pour toute fonction continue (et positive) sur $[a,b]$. En revanche, elle permet de calculer de nombreuses aires planes à partir du calcul des primitives, ce qui a été abondamment utilisé au XVII^e^ pour « quarrer » les surfaces.

2. — Primitives des fonctions réglées

L'intégrale simple d'une fonction $f:[a,b]\to E$ n'a été définie que pour les fonctions f admettant des primitives. Comme il existe des fonctions dépourvues de primitives sur $[a,b]$, il est naturel de se demander quelles fonctions possèdent des primitives. Nous montrons ci-dessous que c'est le cas des fonctions réglées. Dans toute ce paragraphe, on désignera par E un espace de Banach réel.

2.1. Fonctions réglées. — Pour toute fonction $f:[a,b]\to E$, on pose :

$$\|f\|_\infty = \operatorname*{Sup}_{a\le x\le b}\|f(x)\|.$$

La fonction f est donc *bornée* sur $[a,b]$ si et seulement si $\|f\|_\infty<+\infty$. L'ensemble des fonctions bornées de $[a,b]$ dans l'espace de Banach E constitue un espace de Banach[2] pour la norme $f\to\|f\|_\infty$ de la convergence uniforme. On le note $\mathcal{B}([a,b],E)$.

Rappelons qu'une suite $(f_1,f_2,...)$ de fonctions de $[a,b]$ dans E *converge uniformément* vers la fonction $f:[a,b]\to E$ si :

---

[2] Voir par exemple le livre de J. DIEUDONNÉ, *Éléments d'Analyse,* tome I, chapitre VII, 7.1.3.

$$\|f_n - f\|_\infty = \sup_{a \leq x \leq b} \|f_n(x) - f(x)\| \xrightarrow[n\to\infty]{} 0 .$$

On sait que la limite uniforme d'une suite de fonctions bornées est bornée. Rappelons également que la limite uniforme d'une suite de fonctions continues en un point $x \in [a,b]$ est continue en ce point.

DÉFINITION 2. — *Une fonction* $f : I \to E$ *définie sur un intervalle* $I \subset \mathbb{R}$ *est dite réglée sur* I *si elle admet en tout point* x *une limite à droite* $f(x+0)$ *et une limite à gauche* $f(x-0)$.

On observera que, si $f : I \to E$ est une fonction réglée sur I, alors $x \to g(f(x))$ est réglée pour toute fonction continue $g : E \to F$ à valeurs dans un espace (de Banach) F. En particulier, la fonction $x \to \|f(x)\|$ est réglée sur I. On notera $\mathfrak{R}(I,E)$ l'ensemble des fonctions réglées sur I à valeurs dans l'espace de Banach E ; c'est un sous-espace vectoriel de toute les fonctions de I dans E.

EXEMPLES. — Toute fonction continue sur un intervalle I de $\mathbb{R}$, à valeurs dans E, est réglée sur I. Toute fonction monotone (croissante ou décroissante) $f : I \to \mathbb{R}$ admet une limite à droite et une limite à gauche, donc est réglée. Les fonctions en escalier sur $[a,b]$ sont également réglées, comme on le vérifie aisément.

PROPOSITION 3. — *La limite uniforme d'une suite de fonctions réglées est une fonction réglée.*

Supposons que f soit limite uniforme d'une suite de fonctions réglées $f_n : [a,b] \to E$. Pour montrer que f

admet une limite à droite au point x, il suffit en vertu du critère de Cauchy de montrer qu'il existe, pour tout $\varepsilon > 0$, un réel $\eta > 0$ tel que l'on ait :

$$(1) \quad x < u,v \leq x+\eta \Rightarrow \| f(u) - f(v) \| \leq \varepsilon .$$

Or on a $\| f - f_n \|_\infty \leq \frac{\varepsilon}{4}$ pour n assez grand, d'où :

$$(2) \quad \begin{aligned} \| f(u) - f(v) \| &\leq 2 \| f - f_n \|_\infty + \| f_n(u) - f_n(v) \| \\ &\leq \tfrac{\varepsilon}{2} + \| f_n(u) - f_n(v) \| . \end{aligned}$$

Puisque f_n vérifie la relation (1) pour $\frac{\varepsilon}{2}$, on déduit immédiatement de (2) l'existence de $\eta > 0$ tel que (1) soit vérifiée pour f, qui admet donc une limite à droite au point x. Le même raisonnement montre que f admet une limite à gauche en tout point, donc est réglée. ■

Il résulte de la proposition 3 qu'une limite uniforme de fonctions en escalier est réglée. Montrons que toute fonction réglée est limite uniforme de fonctions en escalier :

THÉORÈME 1. — *Une fonction* $f : [a,b] \to E$ *à valeurs dans un espace de Banach* E *est réglée si et seulement si elle est limite uniforme sur* $[a,b]$ *d'une suite de fonctions en escalier.*

Il suffit de montrer qu'une fonction réglée est limite uniforme de fonctions en escalier. Si f est réglée sur $[a,b]$ il existe, pour tout entier $n \geq 1$ et tout point $x \in \,]a,b[$, un intervalle :

$$I(x) = \,]x-\eta, x+\eta[\, \subset [a,b]$$

tel que l'on ait :

$$\begin{cases} u \in I(x) \text{ et } x < u \Rightarrow \| f(u) - f(x+0) \| \leq \frac{1}{n} \\ u \in I(x) \text{ et } u < x \Rightarrow \| f(u) - f(x-0) \| \leq \frac{1}{n} . \end{cases}$$

En d'autres termes, la restriction de f à l'intervalle $I(x)$ est proche à $\frac{1}{n}$ près (pour la norme de la convergence uniforme sur $I(x)$) de la fonction en escalier qui vaut $f(x)$ au point x, $f(x-0)$ sur l'intervalle $]x-\eta,x[$ et $f(x+0)$ sur l'intervalle $]x,x+\eta[$. De même, on peut trouver un intervalle $I(a)=[a,a+\delta[\subset[a,b]$ (resp. un intervalle $I(b)=]b-\delta,b]\subset[a,b]$) tel que la restriction de f à cet intervalle soit proche à $\frac{1}{n}$ en norme uniforme d'une fonction en escalier. Comme $[a,b]$ est compact, il peut être recouvert par un nombre fini d'intervalles $I(x)$ ($a\le x\le b$) sur lesquels f est proche à $\frac{1}{n}$ en norme uniforme d'une fonction en escalier. On en déduit alors facilement l'existence d'une fonction en escalier f_n telle que :

$$\|f-f_n\|_\infty \le \frac{1}{n},$$

ce qui démontre le théorème 1. ■

COROLLAIRE 1. — *Toute fonction réglée sur un intervalle $[a,b]$ (à valeurs dans un espace de Banach) est bornée et ses points de discontinuité forment un ensemble (au plus) dénombrable.*

Soit f une fonction réglée et notons $(f_1, f_2,...)$ une suite de fonctions en escalier convergeant uniformément vers f. Comme chaque f_n est bornée, la limite uniforme f est bornée. Par ailleurs, puisque chaque fonction f_n n'a qu'un nombre fini de points de discontinuité, les points x où toutes les f_n sont continues forment un ensemble de complémentaire dénombrable. Mais une limite uniforme de fonctions continues en un point x est continue en ce point, de sorte que le corollaire 1 est démontré. ■

2.2. Existence de primitives. — Nous sommes maintenant en mesure de démontrer le résultat principal de ce chapitre :

THÉORÈME 2. — *Soit* $f:[a,b]\to E$ *une fonction réglée sur* $[a,b]$*, à valeurs dans un espace de Banach* E*. Soit* $(f_1, f_2,...)$ *une suite de fonctions en escalier convergeant uniformément vers* f *sur* $[a,b]$*, et désignons par* F_n *la primitive de* f_n *qui s'annule au point* a*. Alors, on a :*

(i) La fonction f *possède des primitives sur l'intervalle* $[a,b]$ *;*

(ii) L'unique primitive de f *qui s'annule au point* a *est limite uniforme sur* $[a,b]$ *de la suite* $(F_1,F_2,...)$ *;*

(iii) Toute primitive F *de* f *sur* $[a,b]$ *est dérivable en tout point* x *où* f *est continue et vérifie* $F'(x)=f(x)$*.*

Pour démontrer le théorème, il suffit de prouver que la suite $(F_1,F_2,...)$ converge uniformément sur $[a,b]$ vers une fonction continue F qui admet pour dérivée $f(x)$ en tout point x où f est continue. Pour tout $x\in[a,b]$, on a en vertu du théorème de la moyenne :

$$(1)\quad \left\|F_n(x)-F_p(x)\right\|\leq(b-a)\left\|f_n-f_p\right\|_\infty ,$$

puisque $F_n(a)-F_p(a)=0$. Comme $\left\|f_n-f_p\right\|_\infty$ tend vers 0 quand $n,p\to+\infty$, la suite $(F_n(x))_{n\geq 1}$ est de Cauchy dans E en vertu de (1), donc converge vers une limite que nous notons $F(x)$. On définit ainsi une fonction $F:[a,b]\to E$ qui vérifie $F(a)=0$. En faisant tendre p vers l'infini dans (1), on obtient :

$$\left\|F_n(x)-F(x)\right\|\leq(b-a)\left\|f_n-f\right\|_\infty ,$$

ce qui prouve que la fonction F est limite uniforme sur $[a,b]$ des fonctions F_n. Comme les F_n sont continues, la fonction F est également continue sur $[a,b]$. Pour achever la démonstration du théorème, il suffit donc de montrer que $F'(x) = f(x)$ en tout point x où f est continue[3]. Soit x un point de continuité de f dont nous supposerons, pour simplifier, qu'il vérifie $a < x < b$. Fixons $\varepsilon > 0$; nous devons montrer qu'il existe $\eta > 0$ tel que :

$$(2)\ |u-x| \le \eta \Rightarrow \|F(u)-F(x)-(u-x)f(x)\| \le |u-x|\varepsilon .$$

Comme f est continue en x, il existe $\eta > 0$ tel que :

$$|u-x| \le \eta \Rightarrow \|f(u)-f(x)\| \le \tfrac{\varepsilon}{2} .$$

Soit $N \ge 1$ un entier tel que :

$$n \ge N \Rightarrow \|f_n - f\|_\infty \le \tfrac{\varepsilon}{4} .$$

Pour $n \ge N$ et $|u-x| \le \eta$, on a :

$$(3)\ \|f_n(u)-f_n(x)\| \le 2\|f_n - f\|_\infty + \|f(u)-f(x)\| \le \varepsilon .$$

Or F_n est dérivable de dérivée f_n, sauf peut-être aux points d'un ensemble dénombrable. L'inégalité (3) implique donc, en vertu du théorème de la moyenne :

$$\|F_n(u)-F_n(x)-(u-x)f_n(x)\| \le |u-x|\varepsilon ,$$

pour tout $n \ge N$ et $|u-x| \le \eta$. En faisant tendre n vers $+\infty$ dans cette dernière inégalité, on obtient (2), ce qui achève la démonstration du théorème. ■

Il résulte immédiatement du théorème 2 qu'une fonction f réglée sur un intervalle $I \subset \mathbb{R}$ (et à valeurs dans un espace de Banach) admet une primitive sur I.

En effet, tout intervalle I peut s'écrire comme réunion d'une suite croissante d'intervalles compacts $I_n = [a_n, b_n]$, $n \ge 1$. Choisissons $a \in I_1$ et notons F_n la primitive de f sur I_n qui s'annule au point a. On vérifie aisément que

---

[3] Si $x = a$ (resp. $x = b$), $F'(x)$ désigne ici la dérivée à droite (resp. à gauche) de F au point a (resp. au point b).

l'on définit sans ambiguïté une primitive F de f sur I (qui s'annule en a) en posant :

$$F(x) = F_n(x) \text{ si } x \in I_n .$$

3. — L'intégrale des fonctions réglées

3.1. Expressions de l'intégrale. — Toute fonction réglée $f : [a,b] \to E$ à valeurs dans un espace de Banach possède une primitive F sur $[a,b]$, donc une intégrale définie par :

$$\int_a^b f(x)dx = F(b) - F(a) .$$

On a vu que la valeur de cette intégrale ne change pas lorsque l'on modifie les valeurs de f aux points d'un ensemble dénombrable. Ainsi, l'intégrale d'une fonction en escalier $f : [a,b] \to E$ égale à c_i sur chaque intervalle $]a_i, a_{i+1}[$ d'une subdivision

$$a = a_o < a_1 < \ldots < a_{n-1} < a_n = b$$

de $[a,b]$ est donnée par :

$$(1) \quad \int_a^b f(x)dx = \sum_{i=0}^{n-1} c_i(a_{i+1} - a_i) .$$

C'est une conséquence immédiate de la proposition 1.

Il résulte du théorème 2 *(ii)* que l'intégrale simple de $f \in \Re([a,b], E)$ est également donnée par :

$$\int_a^b f(x)dx = \lim_{n \to +\infty} \int_a^b f_n(x)dx ,$$

où $(f_1, f_2, \ldots)$ est une suite de fonctions en escalier convergeant uniformément vers f sur $[a,b]$.

L'intégrale définie s'interprète également comme une limite de sommes de RIEMANN de la forme

$$\sum_{i=0}^{n-1} f(x_i)\Delta x_i\,,$$

associées à des partitions de $[a,b]$ en petits intervalles de longueurs Δx_i. Plus précisément, on a :

PROPOSITION 4. — *Soit $f:[a,b]\to E$ une fonction réglée. Pour tout $\varepsilon>0$, il existe $\eta>0$ tel que l'on ait, pour toute subdivision*

$$\sigma : a = x_o < x_1 < \ldots < x_{n-1} < x_n = b$$

de $[a,b]$ vérifiant $\delta(\sigma) = \underset{0\le i\le n-1}{Max}\left|a_{i+1}-a_i\right| \le \eta$ et tout choix de $\xi_i \in [x_i, x_{i+1}]$:

$$\left\|\int_a^b f(x)dx - \sum_{i=0}^{n-1} f(\xi_i)(x_{i+1}-x_i)\right\| \le \varepsilon\,.$$

La proposition est facile à vérifier pour une fonction en escalier. Posons, avec les notations utilisées :

$$S(f,\sigma,\xi) = \sum_{i=0}^{n-1} f(\xi_i)(x_{i+1}-x_i)\,.$$

Soit f une fonction réglée et choisissons une suite $(f_n)_{n\ge 1}$ de fonctions en escalier convergeant uniformément vers f sur $[a,b]$. Des relations :

$$\left\|S(f,\sigma,\xi)-S(f_n,\sigma,\xi)\right\| \le (b-a)\left\|f-f_n\right\|_\infty\,,$$

$$\left\|\int_a^b f(x)dx - \int_a^b f_n(x)dx\right\| \le (b-a)\left\|f-f_n\right\|_\infty\,,$$

on déduit que :

$$(1)\quad \left\|\int_a^b f(x)dx - S(f,\sigma,\xi)\right\| \le 2(b-a)\left\|f-f_n\right\|_\infty + \left\|\int_a^b f_n(x)dx - S(f_n,\sigma,\xi)\right\|.$$

Choisissons $n\ge 1$ tel que $2(b-a)\left\|f-f_n\right\|_\infty \le \frac{\varepsilon}{2}$. Comme la proposition 4 est vraie pour f_n, on déduit immédiatement de *(1)* qu'elle est vraie pour f. ■

On déduit en particulier de la proposition 4 que l'on a, pour toute fonction réglée $f:[a,b]\to E$:

$$\int_a^b f(x)dx = \lim_{n\to\infty} \tfrac{b-a}{n} \sum_{i=0}^{n-1} f(a + i\tfrac{b-a}{n}).$$

3.2. Propriétés de l'intégrale. — La proposition ci-dessous résume les principales propriétés de l'intégrale des fonctions réglées :

PROPOSITION 5. — *Soient* E *un espace de Banach et* $a < b$. *L'intégrale définie entre* a *et* b *est une application linéaire* $f \to \int_a^b f(x)dx$ *de* $\mathfrak{R}([a,b],E)$ *dans* E *qui possède les propriétés suivantes :*

(i) $\left\|\int_a^b f(x)dx\right\| \le \int_a^b \|f(x)\|dx \le (b-a)\|f\|_\infty$ *pour toute* $f \in \mathfrak{R}([a,b],E)$ *;*

(ii) Si $E = \mathbb{R}$ *et si les fonctions réglées* f,g *vérifient* $f(x) \le g(x)$ *sauf peut-être pour un ensemble dénombrable de valeurs de* x, *alors on a :*

$$\int_a^b f(x)dx \le \int_a^b g(x)dx.$$

En particulier, si $m \le f(x) \le M$ *sauf peut-être pour un ensemble dénombrable de valeurs de* x, *on a :*

$$m(b-a) \le \int_a^b f(x)dx \le M(b-a);$$

(iii) Si $a < c < b$, *on a (relation de* CHASLES *pour les intégrales) :*

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$$

quelles que soient $f,g \in \mathfrak{R}([a,b],E)$ *;*

(iv) Pour toute fonction $f \in \mathfrak{R}([a,b],E)$, *la fonction* F *définie par :*

$$F(x) = \int_a^x f(t)dt.$$

est dérivable en tout point de continuité x *de* f *, de dérivée* $F'(x) = f(x)$.

La linéarité de l'intégrale résulte immédiatement du fait que, si F, G sont respectivement des primitives de f, g sur $[a,b]$, alors $\lambda F + \mu G$ est une primitive de $\lambda f + \mu g$ sur $[a,b]$ quels que soient $\lambda, \mu \in \mathbb{R}$. Pour les fonctions en escalier, la relation *(i)* se déduit immédiatement de l'expression de l'intégrale définie rappelée ci-dessus (section 3.1, formule *(1)*). Par passage à la limite, en utilisant une suite $(f_1, f_2, ...)$ de fonctions en escalier qui converge uniformément vers f sur $[a,b]$, on en déduit l'assertion *(i)* en toute généralité. L'assertion *(ii)* résulte du fait que, si F, G sont respectivement des primitives de f, g sur $[a,b]$, on a :

$$\tfrac{d}{dx}(F(x) - G(x)) = f(x) - g(x) \leq 0$$

sauf peut-être pour un ensemble dénombrable de valeurs de x. La fonction $F - G$ est donc décroissante en vertu du théorème de la moyenne, ce qui démontre la première partie de l'assertion *(ii)*. La seconde partie s'ensuit immédiatement. L'assertion *(iii)* résulte du fait que, si F est une primitive de f sur $[a,b]$, on a :

$$F(b) - F(a) = (F(b) - F(c)) + (F(c) - F(a)).$$

Enfin, l'assertion *(iv)* est une reformulation du théorème 2 *(iii)*. ■

COROLLAIRE 2 (changement de variable). — *Soit* $\varphi : [\alpha, \beta] \to \mathbb{R}$ *une application de classe* C^1. *Posons* $a = \varphi(\alpha)$, $b = \varphi(\beta)$ *et notons* I *l'intervalle* $[a,b]$ *ou* $[b,a]$ *selon que* $a \leq b$ *ou* $b \leq a$. *Alors, pour toute fonction continue* $f : I \to E$ *à valeurs dans un espace de Banach* E *, on a :*

$$\int_a^b f(x)dx = \int_\alpha^\beta f(\varphi(t))\varphi'(t)dt.$$

En effet, si F est une primitive de f sur I, la fonction G définie par $G(t)=F(\varphi(t))$ est une primitive de la fonction $g(t)=f(\varphi(t))\varphi'(t)$ sur $[\alpha,\beta]$. La relation :

$$F(b)-F(a)=F(\varphi(\beta))-F(\varphi(\alpha))=G(\beta)-G(\alpha)$$

démontre alors le corollaire. ■

COROLLAIRE 3 (Intégration par parties). — *Soient E,F,G trois espaces de Banach et $B:E\times F\to G$ une application bilinéaire continue. Alors, si f,g sont des primitives de fonctions réglées sur $[a,b]$ à valeurs dans E et F respectivement, on a :*

$$\int_a^b B(f(x),g'(x))dx = B(f(b),g(b)) - B(f(a),g(a)) - \int_a^b B(f'(x),g(x))dx.$$

En effet, comme $x\to f'(x)$ et $x\to g'(x)$ sont des fonctions réglées et que B est continue, la fonction continue $x\to B(f(x),g(x))$ a pour dérivée au point x :

$$(1)\quad \tfrac{d}{dx}B(f(x),g(x))=B(f'(x),g(x))+B(f(x),g'(x)),$$

sauf peut-être pour des valeurs de x formant un ensemble dénombrable. Par ailleurs, sachant que f,g sont des fonctions continues à dérivées réglées, on vérifie facilement que $x\to B(f'(x),g(x))$ et $x\to B(f(x),g'(x))$ sont des fonctions réglées. Le corollaire résulte alors directement de la définition de l'intégrale et de *(1)*. ■

EXERCICES

EXERCICE 1. Soit $f:\,]0,1[\,\to\mathbb{R}$ la fonction définie par $f(x)=0$ si $0<x\leq\frac{1}{2}$ et $f(x)=1$ sinon. Montrer qu'il n'existe pas de fonction dérivable $F:\,]0,1[\,\to\mathbb{R}$ telle que $F'(x)=f(x)$ quel que soit $0<x<1$. Montrer que f admet cependant des primitives sur $]0,1[$, que l'on déterminera.

EXERCICE 2. Soit $f:[0,1]\to\mathbb{R}$ la fonction définie par $f(x)=0$ si x est rationnel et $f(x)=1$ sinon. Montrer que f admet une primitive sur $[0,1]$ et calculer son intégrale définie entre 0 et 1. La fonction f est-elle réglée sur $[0,1]$?

EXERCICE 3. Soit $f:[0,\pi]\to\mathbb{R}$ définie par :

$$f(x)=\frac{x\sin x}{1+\cos^2 x} \quad si \quad x\neq\frac{\pi}{2} \quad et \quad f(\tfrac{\pi}{2})=0 .$$

Déterminer une primitive de la fonction f sur $[0,\pi]$ et calculer $\int_0^\pi f(x)dx$.

EXERCICE 4. Soit $F:[a,b]\to\mathbb{R}$ une fonction dérivable et posons $f(x)=F'(x)$ pour tout $x\in[a,b]$.

a) On considère la fonction G définie sur $[a,b]$ par :

$$G(x)=\frac{F(x)-F(a)}{x-a} \quad si \quad x\neq a \quad et \quad G(a)=f(a) .$$

Montrer que, pour tout réel C compris entre $f(a)$ et $G(b)$, il existe $\alpha\in[a,b]$ tel que $G(\alpha)=C$. En déduire qu'il existe $c\in[a,b]$ tel que $f(c)=C$ (on pourra utiliser le théorème des accroissements finis).

b) On considère la fonction H définie sur $[a,b]$ par :

$$H(x)=\frac{F(b)-F(x)}{b-x} \quad si \quad x\neq b \quad et \quad H(b)=f(b) .$$

Montrer que, pour tout C compris entre $H(a)=G(b)$ et $f(b)$, il existe $c\in[a,b]$ tel que $f(c)=C$.

c) En déduire que la fonction $f(x)=F'(x)$ prend, entre a et b, toute valeur comprise entre $f(a)$ et $f(b)$.

d) Montrer que la fonction considérée à l'exercice 2 n'est pas le dérivée d'une fonction définie sur $[0,1]$.

EXERCICE 5. Soit $f:[0,1]\to\mathbb{R}$ définie par :

$$f(x)=signe(x\sin\tfrac{1}{x}) \text{ si } x\neq 0 \text{ et } f(0)=0 .$$

Montrer que f n'est pas réglée sur $[0,1]$.

EXERCICE 6. On dit qu'une fonction $f:[a,b]\to\mathbb{R}$ est à *variation bornée* s'il existe une constante $C\geq 0$ telle que l'on ait :

$$V(\sigma)=\sum_{i=0}^{n-1}|f(x_{i+1})-f(x_i)|\leq C$$

pour toute subdivision

$$\sigma : a=x_o<x_1<\ldots<x_{n-1}<x_n=b \text{ de } [a,b].$$

On pose alors $V(a,b)=Sup_\sigma V(\sigma)$, et on dit que $V(a,b)$ est la *variation* de f sur $[a,b]$.

a) Montrer que, si $f:[a,b]\to\mathbb{R}$ est croissante, elle est à variation bornée. En déduire qu'une différence de deux fonctions croissantes est à variation bornée.

b) Soit $f:[a,b]\to\mathbb{R}$ une fonction à variation bornée sur $[a,b]$. Montrer que, pour tout $x\in[a,b]$, la fonction f est à variation bornée sur $[a,x]$. Montrer que les fonctions $g(x)=V(a,x)$ et $h(x)=V(a,x)-f(x)$ sont croissantes. En déduire que f est la différence de deux fonctions croissantes.

c) Déduire de ce qui précède que toute fonction $f:[a,b]\to\mathbb{R}$ à variation bornée sur $[a,b]$ est réglée sur $[a,b]$.

EXERCICE 7. Calculer la limite de la suite :

$$u_n=\frac{1}{2n+1}+\frac{1}{2n+3}+\ldots+\frac{1}{4n-1}\quad(n\geq 1).$$

EXERCICE 8. Soit $f:[a,b]\to E$ une fonction réglée à valeurs dans un espace de Banach.

a) On suppose que l'image de f est incluse dans une boule fermée B de rayon $r>0$. Montrer que :

$$\frac{1}{b-a}\int_a^b f(x)dx\in B.$$

b) On suppose maintenant que f est continue et que $E=\mathbb{R}$. Montrer qu'il existe $c\in[a,b]$ tel que :

$$\frac{1}{b-a}\int_a^b f(x)dx=f(c).$$

EXERCICE 9. Soit $f:[a,b]\to E$ une fonction réglée à valeurs dans un espace de Banach. Montrer qu'il existe, pour tout $\varepsilon>0$, une fonction continue $g:[a,b]\to E$ telle que :

$$\int_a^b \| f(x)-g(x)\| dx \le \varepsilon$$

(on pourra se ramener au cas où f est en escalier).

EXERCICE 10. Soient $f:[a,b]\to E$ une fonction réglée et F une primitive de f sur $[a,b]$. Montrer que F admet en tout point x une dérivée à droite $F'_d(x)=f(x+0)$ et une dérivée à gauche $F'_g(x)=f(x-0)$.

EXERCICE 11. a) Soit $f:[a,b]\to\mathbb{R}$ une fonction réglée sur $[a,b]$. Montrer que, pour toute fonction continue de signe constant $g:[a,b]\to\mathbb{R}$, il existe $y\in[a,b]$ tel que l'on ait (*première formule de la moyenne*) :

$$\int_a^b f(x)g(x)dx = g(y)\int_a^b f(x)dx\,.$$

b) Soient $f:[a,b]\to\mathbb{R}$ une fonction continue et $g:[a,b]\to\mathbb{R}$ une fonction continûment dérivable dont la dérivée garde un signe constant. Montrer qu'il existe $y\in[a,b]$ tel que (*seconde formule de la moyenne*) :

$$\int_a^b f(x)g(x)dx = g(a)\int_a^y f(x)dx + g(b)\int_y^b f(x)dx\,.$$

(on pourra faire une intégration par partie et utiliser a)).

EXERCICE 12. Soit $f:\mathbb{R}\to E$ une fonction continue à valeurs dans un espace de Banach. Pour tout entier $n\ge 1$, on définit la fonction $f_n:\mathbb{R}\to E$ par :

$$f_n(x)=\frac{3n}{4}\int_{-\frac{1}{n}}^{\frac{1}{n}} f(x+t)(1-n^2t^2)dt\,.$$

Montrer que f_n est continûment dérivable sur $\mathbb{R}$ et que la suite $(f_1,f_2,\ldots)$ converge uniformément vers f sur tout intervalle fermé et borné de $\mathbb{R}$.

CHAPITRE II

L'INTÉGRALE DE RIEMANN

1. — Fonctions Riemann intégrables

1.1. Définition de l'intégrale de Riemann. — Dans un mémoire de 1854[1], Bernhard RIEMANN définit l'intégrale d'une fonction bornée $f:[a,b]\to\mathbb{R}$ comme la limite (lorsqu'elle existe) des sommes finies $\sum_i f(\xi_i)\Delta x_i$ quand $Sup_i \Delta x_i \to 0$. Chaque somme $\sum_i f(\xi_i)\Delta x_i$ est obtenue en divisant $[a,b]$ en un nombre fini d'intervalles $[x_i, x_{i+1}]$ de longueur Δx_i, puis en choisissant arbitrairement un point ξ_i dans chaque intervalle $[x_i, x_{i+1}]$; elle s'interprète comme l'intégrale d'une fonction en escalier égale à $f(\xi_i)$ sur $]x_i, x_{i+1}[$ (cf. fig. 1). Pour que ce procédé définisse l'intégrale de f, il faut évidemment que les sommes $\sum_i f(\xi_i)\Delta x_i$ convergent vers une limite

[1] Bernhard RIEMANN a défini l'intégrale qui porte son nom en 1854, dans les préliminaires d'un mémoire sur les séries trigonométriques.

quand $Sup_i \Delta x_i \to 0$, ce qu'impose RIEMANN. Détaillons cette construction.

Soit $f : [a,b] \to \mathbb{R}$ une fonction arbitraire. Par *subdivision riemannienne* de $[a,b]$, on désigne un couple (σ,ξ) formé d'une subdivision

$$\sigma : a = x_o < x_1 < \ldots < x_n = b$$

de $[a,b]$ et d'une suite $\xi = (\xi_o, \xi_1, \ldots, \xi_{n-1})$ de réels choisis de sorte que $\xi_i \in [x_i, x_{i+1}]$ quel que soit $i = 0,1,\ldots,n-1$. Le *pas* de la subdivision σ est par définition la quantité :

$$\delta(\sigma) = Max_{0 \le i \le n-1}(x_{i+1} - x_i).$$

A toute subdivision riemannienne (σ,ξ) de $[a,b]$, on associe la *somme de* RIEMANN :

$$S(f,\sigma,\xi) = \sum_{i=0}^{n-1}(x_{i+1} - x_i) f(\xi_i).$$

Fig.1. — Sommes de Riemann d'une fonction.

DÉFINITION 1. — *La fonction* $f : [a,b] \to \mathbb{R}$ *est dite intégrable au sens de* RIEMANN *sur* $[a,b]$ *si les sommes de* RIEMANN $S(f,\sigma,\xi)$ *tendent vers une limite* I *quand* $\delta(\sigma) \to 0$.

Si f est intégrable au sens de RIEMANN, la limite I (qui est unique) est appelée *intégrale de* RIEMANN de f sur $[a,b]$; on la note :

$$I = \int_a^b f(x)dx .$$

On désignera par $\mathfrak{R}^1([a,b])$ l'ensemble des fonctions réelles intégrables au sens de RIEMANN sur $[a,b]$. La fonction f appartient donc à $\mathfrak{R}^1([a,b])$ et a pour intégrale I s'il existe, pour tout $\varepsilon > 0$, un réel $\eta > 0$ tel que l'on ait, quelle que soit la subdivision riemannienne (σ,ξ) de $[a,b]$:

$$\delta(\sigma) \leq \eta \Rightarrow \left| S(f,\sigma,\xi) - I \right| \leq \varepsilon .$$

Comme $\mathbb{R}$ est un espace complet[2], la convergence des sommes de RIEMANN $S(f,\sigma,\xi)$ lorsque $\delta(\sigma) \to 0$ équivaut au fait qu'elles vérifient la propriété de CAUCHY :

$$\left| S(f,\sigma,\xi) - S(f,\sigma^*,\xi^*) \right| \to 0$$

quand $\delta(\sigma), \delta(\sigma^*) \to 0$. Ceci signifie qu'il existe, pour tout $\varepsilon > 0$, un réel $\eta > 0$ tel que l'on ait, quelles que soient les subdivisions riemanniennes (σ,ξ) et (σ^*,ξ^*) de $[a,b]$:

$$\delta(\sigma), \delta(\sigma^*) \leq \eta \Rightarrow \left| S(f,\sigma,\xi) - S(f,\sigma^*,\xi^*) \right| \leq \varepsilon .$$

Donnons maintenant des exemples de fonctions intégrables au sens de RIEMANN.

PROPOSITION 1. — *Toute fonction en escalier sur* $[a,b]$ *est intégrable au sens de* RIEMANN, *et son in-*

[2] Rappelons que toute suite de CAUCHY de $\mathbb{R}$ converge.

tégrale de RIEMANN *coïncide avec son intégrale définie entre* a *et* b.

En effet, soit $f:[a,b]\to\mathbb{R}$ une fonction en escalier associée à une subdivision de $[a,b]$ en n intervalles $[a_i,a_{i+1}]$ sur lesquels elle prend la valeur constante c_i. Pour toute subdivision riemannienne (σ,ξ) de $[a,b]$, on vérifie facilement (en considérant la subdivision formée des points de σ et des a_i) que l'on a :

$$\left|S(f,\sigma,\xi)-\sum_{i=0}^{n}(a_{i+1}-a_i)c_i\right|\leq 4n\|f\|_\infty\,\delta(\sigma).$$

Il s'ensuit que $S(f,\sigma,\xi)\to\sum_{i=0}^{n}(a_{i+1}-a_i)c_i$ quand $\delta(\sigma)\to 0$, ce qui prouve que f est RIEMANN intégrable, d'intégrale égale à $\sum_{i=0}^{n}(a_{i+1}-a_i)c_i$. ■

Plus généralement, toute fonction réglée sur $[a,b]$ est intégrable au sens de RIEMANN, comme nous le verrons plus loin. On notera qu'il existe des fonctions non intégrables au sens de RIEMANN et dont l'intégrale définie existe cependant.

C'est le cas par exemple de la fonction de DIRICHLET $\chi:[0,1]\to\mathbb{R}$ définie par $\chi(x)=0$ si x est rationnel et $\chi(x)=1$ sinon. La fonction χ admet pour primitive (au sens défini au chapitre I) la fonction $F(x)=x$; son intégrale est donc définie par la formule :

$$\int_o^1\chi(x)dx=F(1)-F(0)=1.$$

La fonction χ n'est cependant pas intégrable au sens de RIEMANN. On peut en effet associer à toute subdivision

$$\sigma: a=x_o<x_1<\ldots<x_n=b$$

de $[0,1]$ une subdivision riemannienne (σ,ξ) telle que tous les ξ_i soient rationnels, et une autre (σ,ξ^*) dont tous les ξ_i^* sont irrationnels. On a alors :

$$\left|S(\chi,\sigma,\xi)-S(\chi,\sigma,\xi^*)\right|=|0-1|=1,$$

ce qui prouve que les sommes de RIEMANN $S(\chi,\sigma,\xi)$ ne convergent pas lorsque $\delta(\sigma)\to 0$.

1.2. Propriétés élémentaires de l'intégrale de Riemann. — Une fonction intégrable au sens de RIEMANN sur un intervalle $[a,b]$ est nécessairement bornée sur cet intervalle. En effet, désignons par $\|f\|_\infty = Sup_{a\le x\le b}|f(x)|$ la norme de la convergence uniforme sur $[a,b]$; on a :

PROPOSITION 2. — *Soit* $f:[a,b]\to\mathbb{R}$ *une fonction intégrable au sens de Riemann sur* $[a,b]$. *Alors,* f *est bornée sur* $[a,b]$ *et on a :*

$$\left|\int_a^b f(x)dx\right| \le (b-a)\|f\|_\infty .$$

Comme f est intégrable au sens de RIEMANN, il existe en vertu du critère de CAUCHY une subdivision :

$$\sigma : a = x_o < x_1 < \ldots < x_n = b$$

de $[a,b]$ telle que l'on ait, quelle que soit la manière de choisir ξ et ξ^* pour former des subdivisions riemanniennes (σ,ξ) et (σ,ξ^*) de $[a,b]$:

$$(1)\quad \left|S(f,\sigma,\xi)-S(f,\sigma,\xi^*)\right| \le 1 .$$

Choisissons ξ de sorte que $\xi_j = x_j$ pour tout j. Pour i fixé et $x\in[x_i,x_{i+1}]$, choisissons ξ^* de sorte que $\xi_j^* = \xi_j$ si $j\ne i$ et $\xi_i^* = x$. La relation (1) s'écrit alors :

$$\left|(x_{i+1}-x_i)(f(x_i)-f(x))\right| \le 1 \ ;$$

elle implique que :

$$|f(x)| \le |f(x_i)| + \tfrac{1}{x_{i+1}-x_i} \le C = Max_i(|f(x_i)| + \tfrac{1}{x_{i+1}-x_i}).$$

Comme cette majoration est vraie pour tout $x \in [x_i, x_{i+1}]$, on a $|f(x)| \leq C$ pour tout $x \in [a,b]$ et f est bornée. De la majoration évidente :

$$|S(f,\sigma,\xi)| \leq (b-a)\|f\|_\infty$$

on déduit alors, en faisant tendre $\delta(\sigma)$ vers 0 :

$$\left|\int_a^b f(x)dx\right| \leq (b-a)\|f\|_\infty . \blacksquare$$

La proposition suivante énonce quelques propriétés élémentaires de l'intégrale de RIEMANN.

PROPOSITION 3. — *Soient* $a < b$. *Alors* $\mathfrak{R}^1([a,b])$ *est un sous-espace vectoriel de l'espace des fonctions bornées sur* $[a,b]$ *sur lequel l'intégrale de* RIEMANN $f \to \int_a^b f(x)dx$ *est une forme linéaire possédant les propriétés suivantes :*

(i) Pour toute $f \in \mathfrak{R}^1([a,b])$, *on a :*

$$\left\|\int_a^b f(x)dx\right\| \leq (b-a)\|f\|_\infty .$$

En particulier, la forme linéaire $f \to \int_a^b f(x)dx$ *est continue sur* $\mathfrak{R}^1([a,b])$ *muni de la norme de la convergence uniforme ;*

(ii) Si les fonctions f,g *sont intégrables au sens de* RIEMANN *et vérifient* $f(x) \leq g(x)$ *pour tout* x, *on a :*

$$\int_a^b f(x)dx \leq \int_a^b g(x)dx .$$

En particulier, si $m \leq f(x) \leq M$ *pour tout* x, *on a :*

$$m(b-a) \leq \int_a^b f(x)dx \leq M(b-a).$$

Soient f,g deux fonctions intégrables au sens de RIEMANN sur $[a,b]$ et $\lambda,\mu \in \mathbb{R}$. Pour toute subdivision riemannienne (σ,ξ) de $[a,b]$, on a :

$$S(\lambda f+\mu g,\sigma,\xi)=\lambda S(f,\sigma,\xi)+\mu S(g,\sigma,\xi),$$

et donc

$$S(\lambda f+\mu g,\sigma,\xi)\to\lambda\int_a^b f(x)dx+\mu\int_a^b g(x)dx$$

quand $\delta(\sigma)\to 0$. Ceci prouve que $\lambda f+\mu g$ est intégrable au sens de RIEMANN sur $[a,b]$ et que :

$$\int_a^b(\lambda f+\mu g)(x)dx=\lambda\int_a^b f(x)dx+\mu\int_a^b g(x)dx.$$

Ainsi, $\Re^1([a,b])$ est un sous-espace vectoriel de l'espace des fonctions bornées sur $[a,b]$, et l'intégrale de RIEMANN $f\to\int_a^b f(x)dx$ est une forme linéaire. La propriété *(i)* résulte de la proposition 2. Supposons enfin que les fonctions f,g de $\Re^1([a,b])$ vérifient $f(x)\le g(x)$ pour tout x. On en déduit que $S(f,\sigma,\xi)\le S(g,\sigma,\xi)$ pour toute subdivision riemannienne (σ,ξ) de $[a,b]$, d'où la propriété *(ii)* quand $\delta(\sigma)\to 0$. ∎

1.3. Limites des fonctions intégrables. — Une limite simple de fonctions RIEMANN intégrables n'est pas RIEMANN intégrable en général (car cette limite peut ne pas être bornée). En revanche, pour les limites uniformes de fonctions intégrables au sens de RIEMANN, on a :

THÉORÈME 1. — *Soit $f_n:[a,b]\to\mathbb{R}$ une suite de fonctions intégrables au sens de* RIEMANN *sur $[a,b]$. On suppose que les f_n convergent uniformément sur $[a,b]$ vers une fonction $f:[a,b]\to\mathbb{R}$. Alors f est intégrable au sens de* RIEMANN *sur $[a,b]$ et on a :*

$$\int_a^b f(x)dx=\lim_{n\to\infty}\int_a^b f_n(x)dx.$$

En effet, de la relation :

$$\left|\int_a^b f_n(x)dx - \int_a^b f_p(x)dx\right| \leq (b-a)\left\|f_n - f_p\right\|_\infty ,$$

on déduit que la suite des intégrales $I_n = \int_a^b f_n(x)dx$ est de Cauchy, donc converge vers une limite I. Pour toute subdivision riemannienne (σ,ξ) de $[a,b]$, on a :

$$\begin{aligned}\left|S(f,\sigma,\xi)-I\right| &\leq \left|S(f,\sigma,\xi)-S(f_n,\sigma,\xi)\right| \\ &+\left|S(f_n,\sigma,\xi)-I_n\right|+\left|I_n-I\right| \\ &\leq (b-a)\left\|f-f_n\right\|_\infty+\left|S(f_n,\sigma,\xi)-I_n\right|+\left|I_n-I\right|,\end{aligned}$$

de sorte qu'il existe, pour tout $\varepsilon > 0$, un entier $N \geq 1$ tel que l'on ait :

$$\left|S(f,\sigma,\xi)-I\right| \leq \tfrac{\varepsilon}{2}+\left|S(f_N,\sigma,\xi)-I_N\right|$$

quelle que soit (σ,ξ). Comme $S(f_N,\sigma,\xi) \to I_N$ quand $\delta(\sigma) \to 0$, il existe $\eta > 0$ tel que la condition $\delta(\sigma) \leq \eta$ implique $\left|S(f_N,\sigma,\xi)-I_N\right| \leq \frac{\varepsilon}{2}$. Mais alors, on a :

$$\delta(\sigma) \leq \eta \Rightarrow \left|S(f,\sigma,\xi)-I\right| \leq \tfrac{\varepsilon}{2}+\tfrac{\varepsilon}{2} = \varepsilon ,$$

ce qui prouve que f est RIEMANN intégrable et que son intégrale est égale à $I = \lim_{n\to\infty} \int_a^b f_n(x)dx$. ■

Du théorème 1, on déduit immédiatement :

COROLLAIRE 1. — *La somme* $u = \sum_{n=0}^{\infty} u_n$ *d'une série uniformément convergente de fonctions* RIEMANN *intégrables* $u_n : [a,b] \to \mathbb{R}$ *est intégrable au sens de* RIEMANN *sur* $[a,b]$ *et s'intègre terme à terme :*

$$\int_a^b u(x)dx = \sum_{n=0}^{\infty} \int_a^b u_n(x)dx .$$

Le théorème 1 implique que les fonctions réglées sont intégrables au sens de RIEMANN :

COROLLAIRE 2. — *Soit* $f : [a,b] \to \mathbb{R}$ *une fonction réglée sur* $[a,b]$. *Alors,* f *est intégrable au sens* *de* RIEMANN *sur* $[a,b]$ *et son intégrale de* RIEMANN *coïncide avec son intégrale définie entre* a *à* b.

La fonction réglée $f:[a,b]\to\mathbb{R}$ est RIEMANN intégrable comme limite uniforme d'une suite $(f_1, f_2,...)$ de fonctions en escalier (donc RIEMANN intégrables en vertu de la proposition 1). En outre, l'intégrale définie de f entre a et b est la limite des intégrales $\int_a^b f_n(x)dx$ (cf. chapitre I), tout comme l'intégrale de RIEMANN de f en vertu du théorème 1. ■

2. — Le critère d'intégrabilité de Riemann

L'intégrabilité (au sens de RIEMANN) d'une fonction réelle bornée sur un intervalle $[a,b]$ est directement reliée à la manière dont cette dernière oscille autour de chacune de ses valeurs. Un premier critère, dû à RIEMANN, établit qu'une fonction bornée $f:[a,b]\to\mathbb{R}$ est intégrable si et seulement si son oscillation moyenne est nulle. Nous exposons ci-dessous ce résultat.

2.1. Oscillation sur un ensemble. — Considérons une fonction bornée $f:[a,b]\to\mathbb{R}$. Pour toute partie V de $[a,b]$, on note respectivement par

$$m(f,V) = Inf_{x\in V}\, f(x) \text{ et } M(f,V) = Sup_{x\in V}\, f(x)$$

la borne inférieure et la borne supérieure de f sur V. La différence

$$\omega(f,V) = M(f,V) - m(f,V)$$

est appelée l'*oscillation* de f sur V. On posera :

$$m(f) = m(f,[a,b]),\ M(f) = M(f,[a,b]), \text{ et}$$

$$\Omega(f) = \omega(f,[a,b]) = M(f) - m(f).$$

On vérifie immédiatement que :

$$0 \leq \Omega(f) \leq 2Sup_{a \leq x \leq b} |f(x)| = 2\|f\|_\infty,$$

où $\|f\|_\infty = Sup_{a \leq x \leq b} |f(x)|$ désigne la norme uniforme de f sur $[a,b]$.

On notera pour la suite que, si V, W sont deux parties de $[a,b]$ telles que $W \subset V$, alors on a :

$$m(f,V) \leq m(f,W) \leq M(f,W) \leq M(f,V),$$

et donc :

$$0 \leq \omega(f,W) \leq \omega(f,V) \leq \Omega(f).$$

2.2. Sommes de Darboux. — Pour toute subdivision $\sigma : a = x_o < x_1 < ... < x_n = b$ du segment $[a,b]$, posons $m_i = m(f,[x_i, x_{i+1}])$, $M_i = M(f,[x_i, x_{i+1}])$ et $\omega_i = \omega(f,[x_i, x_{i+1}])$. Les sommes :

$$\Sigma_-(f,\sigma) = \sum_{i=0}^{n-1}(x_{i+1} - x_i)m_i,$$

$$\Sigma_+(f,\sigma) = \sum_{i=0}^{n-1}(x_{i+1} - x_i)M_i,$$

$$\omega(f,\sigma) = \sum_{i=0}^{n-1}(x_{i+1} - x_i)\omega_i,$$

sont appelées *somme de* DARBOUX *inférieure*, *somme de* DARBOUX *supérieure* et *oscillation moyenne* de f relatives à la subdivision σ.

On dira qu'une subdivision σ^* de $[a,b]$ est *plus fine* qu'une subdivision σ si les points de subdivision de σ sont des points de subdivision de σ^*. Étant donné deux subdivisions σ et σ^* de $[a,b]$, on notera $\sigma \cup \sigma^*$ la subdivision formée des points de σ et de

σ^*. Il est clair que $\sigma \cup \sigma^*$ est plus fine que σ et σ^*. Avec ces notations, on a :

PROPOSITION 4. — *Soit* $f : [a,b] \to \mathbb{R}$ *une fonction bornée.*

(i) Pour toute subdivision riemannienne (σ, ξ) *de* $[a,b]$*, on a :*

$$\Sigma_-(f,\sigma) \leq S(f,\sigma,\xi) \leq \Sigma_+(f,\sigma) \ ;$$

(ii) Si la subdivision σ^* *de* $[a,b]$ *est plus fine que la subdivision* σ*, on a :*

$$\Sigma_-(f,\sigma) \leq \Sigma_-(f,\sigma^*) \leq \Sigma_+(f,\sigma^*) \leq \Sigma_+(f,\sigma) \ ;$$

(iii) Quelles que soient les subdivisions riemanniennes (σ,ξ) *et* (σ^*,ξ^*) *de* $[a,b]$*, on a :*

$$\left| S(f,\sigma,\xi) - S(f,\sigma^*,\xi^*) \right| \leq \omega(f,\sigma) + \omega(f,\sigma^*) .$$

La propriété *(i)* est immédiate. Pour démontrer la propriété *(ii)*, il suffit de considérer le cas où σ^* est obtenue en ajoutant un point supplémentaire c à la subdivision σ. Si $c \in \left[x_i, x_{i+1}\right]$, l'inégalité $\Sigma_-(f,\sigma) \leq \Sigma_-(f,\sigma^*)$ se ramène alors à la relation :

$$\begin{aligned}(x_{i+1} - x_i)m(f,\left[x_i,x_{i+1}\right]) \leq{} & (c - x_i)m(f,\left[x_i,c\right]) \\ & + (x_{i+1} - c)m(f,\left[c,x_{i+1}\right]),\end{aligned}$$

qui résulte immédiatement du fait que $m(f,\left[x_i,x_{i+1}\right])$ est inférieur ou égal à $m(f,\left[x_i,c\right])$ et à $m(f,\left[c,x_{i+1}\right])$. On démontre de même l'inégalité $\Sigma_+(f,\sigma^*) \leq \Sigma_+(f,\sigma)$, d'où l'assertion *(ii)*. Pour démontrer l'assertion *(iii)*, considérons une subdivision riemannienne $(\sigma \cup \sigma^*, \zeta)$ de $[a,b]$. Comme la subdivision $\sigma \cup \sigma^*$ est plus fine que σ, le segment $\left[\Sigma_-(f,\sigma \cup \sigma^*), \Sigma_+(f,\sigma \cup \sigma^*)\right]$ est contenu dans le segment $\left[\Sigma_-(f,\sigma), \Sigma_+(f,\sigma)\right]$ en vertu de l'assertion *(ii)*. Il s'ensuit que $S(f,\sigma \cup \sigma^*,\zeta)$ appartient au segment $\left[\Sigma_-(f,\sigma),\Sigma_+(f,\sigma)\right]$, tout comme $S(f,\sigma,\xi)$ (en vertu de *(i)*), d'où :

$$(1)\quad \left|S(f,\sigma,\xi)-S(f,\sigma\cup\sigma^*,\zeta)\right|\leq\Sigma_+(f,\sigma)-\Sigma_-(f,\sigma) = \omega(f,\sigma).$$

Le même raisonnement montre que :

$$(2)\quad \left|S(f,\sigma\cup\sigma^*,\zeta)-S(f,\sigma^*,\xi^*)\right|\leq\omega(f,\sigma^*),$$

et l'assertion *(iii)* résulte alors immédiatement de (1) et (2) (via l'inégalité triangulaire). ■

Le lemme suivant permet de comparer $\omega(f,\sigma)$ et $\omega(f,\sigma^*)$ sans supposer que l'une ou l'autre des subdivisions σ,σ^* est plus fine que l'autre.

LEMME 1. — *Soient* $f:[a,b]\to\mathbb{R}$ *une fonction bornée et* σ_o *une subdivision de* $[a,b]$. *Notons* $N(\sigma_o)$ *le nombre de points de subdivision de* σ_o. *Alors on a, pour toute subdivision* σ *de* $[a,b]$ *:*

$$\omega(f,\sigma)\leq\omega(f,\sigma_o)+2N(\sigma_o)\Omega(f)\delta(\sigma).$$

Notons $x_{o,i}$ (resp. x_j) les points de subdivision de σ_o (resp. de σ) et $\omega_{o,i}$ (resp. ω_j) l'oscillation de f sur $\left[x_{o,i},x_{o,i+1}\right]$ (resp. sur $\left[x_j,x_{j+1}\right]$). Pour tous les intervalles $\left[x_j,x_{j+1}\right]$ contenus dans un même intervalle $\left[x_{o,i},x_{o,i+1}\right]$, on a $\omega_j\leq\omega_{o,i}$; par conséquent, la somme des quantités $(x_{j+1}-x_j)\omega_j$ correspondantes est majorée par $(x_{o,i+1}-x_{o,i})\omega_{o,i}$. Il s'ensuit que la somme des nombres $(x_{j+1}-x_j)\omega_j$ associés aux intervalles $\left[x_j,x_{j+1}\right]$ qui sont contenus dans un intervalle $\left[x_{o,i},x_{o,i+1}\right]$ est majorée par $\omega(f,\sigma_o)$. Les intervalles $\left[x_j,x_{j+1}\right]$ non contenus dans un intervalle $\left[x_{o,i},x_{o,i+1}\right]$ sont au plus au nombre de $2+2(N(\sigma_o)-2)\leq 2N(\sigma_o)$. Pour de tels intervalles, le nombre $(x_{j+1}-x_j)\omega_j$ est majoré par $\delta(\sigma)\Omega(f)$. La somme des quantités $(x_{j+1}-x_j)\omega_j$ associées aux intervalles $\left[x_j, x_{j+1}\right]$ non contenus dans un intervalle $\left[x_{o,i}, x_{o,i+1}\right]$ est donc majorée par $2N(\sigma_o)\Omega(f)\delta(\sigma)$, d'où le lemme 1. ■

2.3. Oscillation moyenne d'une fonction. — Soit $f : [a,b] \to \mathbb{R}$ une fonction bornée. On appelle *oscillation moyenne* de f la borne inférieure $\overline{\omega}(f)$ des oscillations $\omega(f,\sigma)$ lorsque σ parcourt l'ensemble des subdivisions de $[a,b]$. Montrons que cette oscillation moyenne est aussi la limite des $\omega(f,\sigma)$ quand $\delta(\sigma) \to 0$:

PROPOSITION 5. — *L'oscillation moyenne $\overline{\omega}(f)$ d'une fonction bornée $f : [a,b] \to \mathbb{R}$ est la limite des oscillations $\omega(f,\sigma)$ relatives aux subdivisions σ de $[a,b]$ quand $\delta(\sigma) \to 0$.*

Par définition de l'oscillation moyenne, il existe en effet pour tout $\varepsilon > 0$ une subdivision σ_o de $[a,b]$ telle que :

$$\overline{\omega}(f) \le \omega(f,\sigma_o) \le \overline{\omega}(f) + \tfrac{\varepsilon}{2}.$$

Choisissons $\eta > 0$ tel que $N(\sigma_o)\Omega(f)\eta \le \frac{\varepsilon}{4}$. D'après le lemme 1, on a pour toute subdivision σ telle que $\delta(\sigma) \le \eta$:

$$\begin{aligned}\overline{\omega}(f) \le \omega(f,\sigma) &\le \omega(f,\sigma_o) + 2N(\sigma_o)\Omega(f)\eta \\ &\le \tfrac{\varepsilon}{2} + \overline{\omega}(f) + \tfrac{\varepsilon}{2} = \overline{\omega}(f) + \varepsilon,\end{aligned}$$

d'où $\left|\overline{\omega}(f) - \omega(f,\sigma)\right| \le \varepsilon$. Ceci démontre la proposition 5. ■

Il résulte de la proposition 5 que l'oscillation moyenne d'une fonction bornée est nulle si et seulement si l'une des conditions équivalentes suivantes est vérifiée :

(i) Pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que l'on ait, pour toute subdivision σ de $[a,b]$:

$$\delta(\sigma) \leq \eta \Rightarrow \omega(f,\sigma) \leq \varepsilon \ ;$$

(ii) Pour tout $\varepsilon > 0$, il existe une subdivision σ de $[a,b]$ telle que $\omega(f,\sigma) \leq \varepsilon$.

2.4. Le critère de Riemann. — Avec les notations de la section précédente, le critère d'intégrabilité de RIEMANN s'énonce :

THÉORÈME 2. — *Une fonction réelle bornée $f : [a,b] \to \mathbb{R}$ est intégrable au sens de* RIEMANN *si et seulement si son oscillation moyenne $\overline{\omega}(f)$ est nulle.*

Supposons que f soit intégrable au sens de RIEMANN et montrons que $\overline{\omega}(f) = 0$. Pour tout $\varepsilon > 0$, il existe une subdivision

$$\sigma : a = x_o < x_1 < \ldots < x_n = b$$

de $[a,b]$ telle que l'on ait, quels que soient les choix de ξ et ξ' avec $\xi_i, \xi_i' \in [x_i, x_{i+1}]$:

$$(1) \quad \left| S(f,\sigma,\xi') - S(f,\sigma,\xi) \right| \leq \varepsilon .$$

Choisissons des suites $(\xi_\nu)_\nu$ et $(\xi'_\nu)_\nu$ telles que, pour tout $i = 0,1,\ldots,n-1$, on ait $f(\xi_{\nu,i}) \to m_i$ et $f(\xi'_{\nu,i}) \to M_i$ quand $\nu \to \infty$. De la relation (1), on déduit :

$$\left| \sum_{i=0}^{n-1} (x_{i+1} - x_i)(f(\xi'_{\nu,i}) - f(\xi_{\nu,i}) \right| \leq \varepsilon ,$$

d'où, en faisant tendre ν vers l'infini :

$$\omega(f,\sigma) = \sum_{i=0}^{n-1} (x_{i+1} - x_i)(M_i - m_i) \leq \varepsilon .$$

Comme $\varepsilon > 0$ est arbitraire, il s'ensuit que l'oscillation moyenne de f est nulle.

Inversement, si $\overline{\omega}(f)=0$, il existe pour tout $\varepsilon>0$ un réel $\eta>0$ tel que l'on ait pour toute subdivision σ de $[a,b]$ (cf. proposition 5) :

$$(2)\quad \delta(\sigma)\leq\eta\Rightarrow\omega(f,\sigma)\leq\tfrac{\varepsilon}{2}.$$

Mais alors, si (σ,ξ) et (σ^*,ξ^*) sont deux subdivisions riemanniennes de $[a,b]$ qui vérifient $\delta(\sigma),\delta(\sigma^*)\leq\eta$, on a en vertu de la proposition 4 *(iii)* et de la relation (2) :

$$\left|S(f,\sigma,\xi)-S(f,\sigma^*,\xi^*)\right|\leq\omega(f,\sigma)+\omega(f,\sigma^*)\leq\varepsilon,$$

ce qui prouve que f est intégrable au sens de RIEMANN sur $[a,b]$. ■

Ainsi, une fonction bornée $f:[a,b]\to\mathbb{R}$ est intégrable au sens de RIEMANN si et seulement si l'une des deux conditions équivalentes suivantes est réalisée :

(i) Pour tout $\varepsilon>0$, il existe $\eta>0$ tel que l'on ait, pour toute subdivision σ de $[a,b]$:

$$\delta(\sigma)\leq\eta\Rightarrow\omega(f,\sigma)\leq\varepsilon\ ;$$

(ii) Pour tout $\varepsilon>0$, il existe une subdivision σ de $[a,b]$ telle que $\omega(f,\sigma)\leq\varepsilon$.

On en déduit la caractérisation alternative suivante des fonctions intégrables au sens de RIEMANN :

COROLLAIRE 3. — *Une fonction $f:[a,b]\to\mathbb{R}$ est intégrable au sens de* RIEMANN *si et seulement s'il existe, pour tout $\varepsilon>0$, deux fonctions e_ε et g_ε en escalier sur $[a,b]$ vérifiant $e_\varepsilon\leq f\leq g_\varepsilon$ et*

$$\int_a^b\left[g_\varepsilon(x)-e_\varepsilon(x)\right]dx\leq\varepsilon.$$

Si f est Riemann intégrable sur $[a,b]$, il existe pour tout $\varepsilon > 0$ une subdivision $\sigma : a = x_o < x_1 < ... < x_n = b$ de $[a,b]$ telle que l'on ait (cf. théorème 2) :

$$(1) \quad \Sigma_+(f,\sigma) - \Sigma_-(f,\sigma) = \omega(f,\sigma) \leq \varepsilon .$$

Notons e_ε (resp. g_ε) la fonction en escalier égale à $m(f,[x_i,x_{i+1}])$ (resp. à $M(f,[x_i,x_{i+1}])$) sur $[x_i,x_{i+1}[$ et à $f(b)$ au point b. On a $e_\varepsilon \leq f \leq g_\varepsilon$ et, comme :

$$\Sigma_-(f,\sigma) = \int_a^b e_\varepsilon(x)dx, \quad \Sigma_+(f,\sigma) = \int_a^b g_\varepsilon(x)dx ,$$

la relation (1) implique que :

$$\int_a^b \left[g_\varepsilon(x) - e_\varepsilon(x)\right]dx \leq \varepsilon .$$

Inversement, supposons l'existence de deux fonctions en escalier e_ε et g_ε vérifiant les conditions du corollaire 3 et montrons que f est intégrable au sens de RIEMANN. Comme e_ε et g_ε sont bornées, la relation $e_\varepsilon \leq f \leq g_\varepsilon$ implique que f est bornée. Notons σ_ε une subdivision de $[a,b]$ associée à e_ε (i.e. e_ε est constante entre deux points consécutifs de σ_ε). Quitte à modifier e_ε aux points de σ_ε (par exemple en remplaçant la valeur de e_ε en ces points par celle de f) on peut supposer, sans changer la relation $e_\varepsilon \leq f$, que la borne inférieure de f sur chaque intervalle de la subdivision σ_ε est égale à la valeur prise par e_ε à l'intérieur de cet intervalle. L'intégrale définie de e_ε n'est pas modifiée, et on a :

$$\Sigma_-(e_\varepsilon,\sigma_\varepsilon) = \int_a^b e_\varepsilon(x)dx .$$

De même, il existe une subdivision σ_ε^* associée à g_ε (que l'on a éventuellement modifiée en certains points de cette subdivision sans changer la relation $f \leq g_\varepsilon$) telle que l'on ait :

$$\Sigma_+(g_\varepsilon,\sigma_\varepsilon^*) = \int_a^b g_\varepsilon(x)dx .$$

On peut donc choisir les subdivisions σ_ε et σ_ε^* de telle sorte que l'on ait :

$$\Sigma_+(g_\varepsilon,\sigma_\varepsilon^*)-\Sigma_-(e_\varepsilon,\sigma_\varepsilon)=\int_a^b g_\varepsilon(x)dx-\int_a^b e_\varepsilon(x)dx\leq\varepsilon\,.$$

Considérons alors la subdivision $\sigma=\sigma_\varepsilon\cup\sigma_\varepsilon^*$ de $[a,b]$. De la relation $e_\varepsilon\leq f$ on déduit que :

$$\Sigma_-(e_\varepsilon,\sigma_\varepsilon)\leq\Sigma_-(f,\sigma_\varepsilon)\leq\Sigma_-(f,\sigma)\,.$$

De même, on montre que :

$$\Sigma_+(f,\sigma)\leq\Sigma_+(f,\sigma_\varepsilon^*)\leq\Sigma_+(g,\sigma_\varepsilon^*)\,,$$

d'où l'on déduit finalement que :

$$\omega(f,\sigma)\leq\Sigma_+(g_\varepsilon,\sigma_\varepsilon^*)-\Sigma_-(e_\varepsilon,\sigma_\varepsilon)\leq\varepsilon\,.$$

Il résulte alors du théorème 2 que f est intégrable au sens de RIEMANN sur $[a,b]$.■

3. — Le critère d'intégrabilité de Lebesgue

Nous démontrons dans ce paragraphe qu'une fonction bornée $f:[a,b]\to\mathbb{R}$ est intégrable au sens de RIEMANN si et seulement si l'ensemble de ses points de discontinuité est négligeable. Ce critère, dû à LEBESGUE, relie l'intégrabilité d'une fonction à la « petitesse » de l'ensemble de ses discontinuités. Il est très utile en pratique, même s'il souligne surtout le caractère très particulier des fonctions intégrables au sens de RIEMANN.

3.1. Oscillation d'une fonction. — Rappelons tout d'abord la définition de l'oscillation d'une fonction bornée $f:[a,b]\to\mathbb{R}$ en un point $x\in[a,b]$.

DÉFINITION 2. — *Soient $f:[a,b]\to\mathbb{R}$ une fonction bornée et $x\in[a,b]$. On appelle oscillation de f au point x la borne inférieure*

$$\omega(f,x)=inf_V\,\omega(f,V)$$

des oscillations $\omega(f,V)$ de f sur V lorsque V parcourt l'ensemble des voisinages de x dans $[a,b]$.

On vérifie aisément que $\omega(f,x)$ coïncide avec la borne inférieure des $\omega(f,[a,b]\cap I)$ lorsque I parcourt l'ensemble des intervalles ouverts contenant x.

Par définition, la fonction f est continue au point x si et seulement si $\omega(f,x)=0$. Si f possède une limite à droite $f(x+0)$ et une limite à gauche $f(x-0)$ au point $x\in]a,b[$, on vérifie aisément que $\omega(f,x)$ est le plus grand des trois nombres suivants : $|f(x+0)-f(x-0)|$, $|f(x)-f(x+0)|$ et $|f(x)-f(x-0)|$.

La fonction $x\to\omega(f,x)$ n'est pas continue en général. Elle est cependant toujours semi-continue supérieurement, comme nous allons le voir. Rappelons qu'une fonction réelle $h:[a,b]\to\mathbb{R}$ est dite *semi-continue supérieurement* au point $x\in[a,b]$ s'il existe, pour tout réel $\varepsilon>0$, un voisinage V de x dans $[a,b]$ tel que :

$$y\in V\Rightarrow h(y)\leq h(x)+\varepsilon .$$

Si h est semi-continue supérieurement en tout point $x\in[a,b]$, on dit qu'elle est semi-continue supérieurement sur $[a,b]$. On vérifie immédiatement que la semi-continuité supérieure de h sur $[a,b]$ équivaut à l'une ou l'autre des deux assertions équivalentes suivantes :

(i) Pour tout $\lambda\in\mathbb{R}$, l'ensemble des $x\in[a,b]$ tels que $h(x)<\lambda$ est ouvert dans $[a,b]$;

(ii) Pour tout $\lambda\in\mathbb{R}$, l'ensemble des $x\in[a,b]$ tels que $h(x)\geq\lambda$ est fermé (dans $[a,b]$).

Avec ces rappels, on a :

PROPOSITION 6. — *Soit* $f:[a,b]\to\mathbb{R}$ *une fonction bornée. Alors, l'application* $x\to\omega(f,x)$ *est semi-continue supérieurement sur* $[a,b]$. *En particulier, pour tout* $\alpha\geq 0$, *l'ensemble des* $x\in[a,b]$ *tels que* $\omega(f,x)\geq\alpha$ *est fermé dans* $[a,b]$.

Il suffit de vérifier que, si $\lambda>\omega(f,x)$, il existe un voisinage V de x dans $[a,b]$ tel que $\lambda>\omega(f,y)$ quel que soit $y\in V$. Or, si $\lambda>\omega(f,x)$, il existe un voisinage W de x dans $[a,b]$ tel que $\omega(f,W)<\lambda$. Choisissons un voisinage ouvert V de x dans $[a,b]$ qui soit inclus dans W. Comme V est un voisinage de chacun de ses points, on a :

$$\omega(f,y)\leq\omega(f,V)\leq\omega(f,W)<\lambda,$$

pour tout $y\in V$, ce qui démontre que la fonction $x\to\omega(f,x)$ est semi-continue supérieurement. ∎

3.2. Ensembles négligeables. — On dit qu'un sous-ensemble $N\subset\mathbb{R}$ est *négligeable* (ou de mesure nulle) s'il existe, pour tout $\varepsilon>0$, une suite d'intervalles $I_n=[a_n,b_n]$, $n\geq 1$, dont la réunion contient N et dont la somme des longueurs est inférieure ou égale à ε :

$$N\subset\cup_{n\geq 1}I_n \text{ et } \textstyle\sum_{n\geq 1}b_n-a_n\leq\varepsilon.$$

On notera qu'un sous-ensemble négligeable $N\subset\mathbb{R}$ est nécessairement d'intérieur vide (s'il contenait un intervalle ouvert, ce dernier devrait être de longueur inférieur à tout $\varepsilon>0$, ce qui est absurde). Tout sous-ensemble fini de $[a,b]$ est clairement négligeable. Un sous-ensemble dénombrable de

$[a,b]$ est également négligeable en vertu de la proposition suivante :

PROPOSITION 7. — *Une réunion dénombrable d'ensembles négligeables est négligeable.*

Soit en effet $N_1, N_2, ...$ une suite d'ensembles négligeables et notons N leur réunion. Pour tout $\varepsilon > 0$ et tout entier $n \geq 1$, il existe une suite $(I_{n,k})_{k\geq 1}$ d'intervalles dont la réunion contient I_n et dont la somme des longueurs est $\leq \frac{\varepsilon}{2^n}$. Les intervalles $I_{n,k}$ $(n,k \geq 1)$ recouvrent alors N et vérifient :

$$\sum_{n,k\geq 1} |I_{n,k}| = \sum_{n\geq 1} \left(\sum_{k\geq 1} |I_{n,k}| \right) \leq \sum_{n\geq 1} \frac{\varepsilon}{2^n} = \varepsilon ,$$

où $|I|$ désigne la longueur de l'intervalle I, de sorte que N est négligeable. ■

On prendra garde au fait qu'une réunion non dénombrable d'ensembles négligeables n'est pas négligeable en général (sinon tout ensemble non vide de réels serait négligeable comme réunion de ses points). On notera également qu'il existe des ensembles négligeables qui ne sont pas dénombrables.

C'est par exemple le cas de l'ensemble triadique de CANTOR $K \subset [0,1]$, constitué des réels x de la forme :

$$x = \frac{a_1}{3} + \frac{a_2}{3^2} + \frac{a_3}{3^3} + ...$$

où les a_n sont tous égaux à 0 ou à 2. Cet ensemble s'obtient à partir du segment $[0,1]$ en itérant la construction qui consiste à diviser un intervalle fermé borné en trois segments égaux, et à enlever les points intérieurs à la partie du milieu. Ainsi, en retranchant de l'intervalle $[0,1]$ le segment ouvert $\left]\frac{1}{3}, \frac{2}{3}\right[$, on obtient un sous-ensemble $K_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$ formé de deux intervalles fermés dont la somme des longueurs est égale à $\frac{2}{3}$. En appliquant la construction indiquée à chacun des intervalles de K_1, on construit alors un sous-ensemble

$$K_2 = \left[0,\frac{1}{3^2}\right] \cup \left[\frac{2}{3^2},\frac{1}{3}\right] \cup \left[\frac{2}{3},\frac{7}{3^2}\right] \cup \left[\frac{8}{3^2},1\right] \subset K_1$$

qui est réunion de 4 intervalles fermés dont la somme des longueurs est égale à $\frac{4}{9} = \left(\frac{2}{3}\right)^2$. En itérant cette construction, on met en évidence une suite de fermés emboîtés :

$$[0,1] \supset K_1 \supset K_2 \supset ... \supset K_n \supset K_{n+1} \supset ...$$

dont l'intersection est exactement l'ensemble de CANTOR K. Comme l'ensemble de CANTOR K est contenu dans chaque K_n, qui est une réunion finie d'intervalles dont la somme des longueurs est égale à $\left(\frac{2}{3}\right)^n$, il est négligeable. En particulier, il est d'intérieur vide. Enfin, K a la puissance du continu[3], puisqu'il a même cardinal que $\{0,2\}^{\mathbb{N}}$.

Une propriété $P(x)$ portant sur les points d'un segment $[a,b]$ est dite *vraie presque partout* si l'ensemble des points x où elle n'est pas vérifiée est négligeable. Ainsi, une fonction presque partout continue est une fonction dont les points de discontinuité forment un ensemble négligeable.

3.3. La caractérisation de Lebesgue. — Henri LEBESGUE a donné une caractérisation très simple des fonctions intégrables au sens de RIEMANN :

THÉORÈME 3. — *Une fonction réelle bornée* $f:[a,b] \to \mathbb{R}$ *est intégrable au sens de* RIEMANN *si*

---

[3] Le développement dyadique « propre » d'un nombre réel compris entre 0 et 1 permet de mettre en bijection $[0,1]$ (qui a la puissance du continu) et l'ensemble formé des suites de 0 et de 1, qui a même cardinalité que $\{0,2\}^{\mathbb{N}}$.

et seulement si l'ensemble de ses points de discontinuité est négligeable.

Supposons que f soit intégrable au sens de RIEMANN et montrons que l'ensemble D des points où elle n'est pas continue est négligeable. Pour tout $\alpha > 0$, considérons l'ensemble $D_\alpha = \{x \in [a,b] \mid \omega(f,x) \geq \alpha\}$. Comme D est la réunion des ensembles $D_{\frac{1}{n}}$ où n parcourt la suite des entiers positifs, il suffit de montrer que D_α est négligeable pour tout $\alpha > 0$. Puisque f est intégrable au sens de RIEMANN, il existe pour tout $\varepsilon > 0$ une subdivision

$$\sigma : a = x_o < x_1 < \ldots < x_n = b$$

de $[a,b]$ telle que l'on ait :

$$\omega(f,\sigma) = \sum_{i=0}^{n-1} (x_{i+1} - x_i)\omega(f,[x_{i+1}, x_i]) \leq \alpha \tfrac{\varepsilon}{2}.$$

Montrons que $D_\alpha \setminus \{x_o, x_1, \ldots, x_n\}$ est inclus dans une réunion d'intervalles $[x_i, x_{i+1}]$ dont la somme des longueurs est $\leq \frac{\varepsilon}{2}$. Puisqu' ensemble fini de points est négligeable, on en déduira que D_α est inclus dans une réunion finie d'intervalles dont la somme des longueurs est $\leq \varepsilon$, et donc qu'il est négligeable. Notons Λ l'ensemble des entiers i tels que l'intervalle $[x_i, x_{i+1}]$ contienne un point de $D_\alpha \setminus \{x_o, x_1, \ldots, x_n\}$. Comme on a :

$$D_\alpha \setminus \{x_o, x_1, \ldots, x_n\} \subset \cup_{i \in \Lambda} [x_i, x_{i+1}],$$

il suffit de montrer que la somme des longueurs des intervalles $[x_i, x_{i+1}]$ pour $i \in \Lambda$ est $\leq \frac{\varepsilon}{2}$. Or, pour tout $i \in \Lambda$, l'intervalle $[x_i, x_{i+1}]$ est voisinage d'un point $x \in D_\alpha$ de sorte que l'on a :

$$\omega(f,[x_i, x_{i+1}]) \geq \omega(f,x) \geq \alpha.$$

Il s'ensuit que l'on a :

$$\begin{aligned}\alpha \sum_{i \in \Lambda} (x_{i+1} - x_i) &\leq \sum_{i \in \Lambda} (x_{i+1} - x_i)\omega(f,[x_{i+1}, x_i]) \\ &\leq \omega(f,\sigma) \leq \alpha \tfrac{\varepsilon}{2},\end{aligned}$$

ce qui montre que la somme $\sum_{i\in\Lambda}(x_{i+1}-x_i)$ des longueurs des intervalles $\left[x_i, x_{i+1}\right]$ pour $i\in\Lambda$ est $\leq\frac{\varepsilon}{2}$.

Supposons inversement que l'ensemble D des points de discontinuité de la fonction f soit négligeable et montrons que f est intégrable au sens de RIEMANN sur $\left[a,b\right]$. Pour tout $\varepsilon>0$, posons $\varepsilon'=\frac{\varepsilon}{\Omega(f)+b-a}$, où $\Omega(f)=\omega(f,\left[a,b\right])$. L'ensemble $D_{\varepsilon'}$ est négligeable (car il est inclus dans D). Il peut donc être recouvert par des intervalles ouverts $I_n=\left]\alpha_n,\beta_n\right[$ dont la somme des longueurs est $\leq\varepsilon'$. Comme $D_{\varepsilon'}$ est fermé (proposition 6) et borné, il peut être recouvert par un nombre fini d'intervalles I_n. Quitte à regrouper ces intervalles, on peut supposer qu'ils sont deux à deux disjoints et les noter $\left]\alpha_i,\beta_i\right[$ ($i=1,2,...,n$) avec $\beta_i<\alpha_{i+1}$. Si a appartient à $\left]\alpha_1,\beta_1\right[$, on remplacera cet intervalle par $\left[a,\beta_1\right[$, et on procédera de même pour b en remplaçant éventuellement $\left]\alpha_n,\beta_n\right[$ par $\left]\alpha_n,b\right]$. La somme des longueurs de tous ces intervalles est donc $\leq\varepsilon'$. En outre, le complémentaire F dans $\left[a,b\right]$ de la réunion de ces intervalles est un ensemble compact (c'est la réunion des intervalles $\left[\beta_i,\alpha_{i+1}\right]$). Cela étant posé, on a $\omega(f,x)<\varepsilon'$ pour tout $x\in F$. Il existe donc un voisinage $V(x)$ de x dans $\left[a,b\right]$ tel que $\omega(f,V)<\varepsilon'$. Par compacité de F, on peut le recouvrir par un nombre fini de voisinages $V(x)$, et obtenir ainsi une subdivision

$$\beta_i=y_o^i<y_1^i<...<y_{n(i)}^i=\alpha_{i+1}$$

de chaque segment $\left[\beta_i,\alpha_{i+1}\right]$ telle que l'on ait :

$$\omega(f,\left[y_j^i,y_{j+1}^i\right])\leq\varepsilon'$$

pour tout intervalle de cette subdivision. Notons alors $\sigma: a=x_o<x_1<...<x_m=b$ la subdivision de $\left[a,b\right]$ formée de tous les points α_i,β_i,y_j^i, et montrons que l'on a :

$$\omega(f,\sigma) = \sum_{j=0}^{m-1}(x_{j+1} - x_j)\omega(f,[x_j, x_{j+1}]) \le \varepsilon .$$

A cet effet, notons Λ l'ensemble des indices j pour lesquels l'intervalle $[x_j, x_{j+1}]$ est inclus dans F, et Λ' son complémentaire dans $\{0,1,...,m-1\}$. Si $j \in \Lambda$, on a (par construction) $\omega(f,[x_j, x_{j+1}]) < \varepsilon'$, et donc :

$$\sum_{j\in\Lambda}(x_{j+1} - x_j)\omega(f,[x_j, x_{j+1}]) \le \varepsilon'(b-a).$$

Si $j \in \Lambda'$, on a :

$$(x_{j+1} - x_j)\omega(f,[x_j, x_{j+1}]) \le (x_{j+1} - x_j)\Omega(f),$$

et donc :

$$\begin{aligned}\sum_{j\in\Lambda'}(x_{j+1} - x_j)\omega(f,[x_j, x_{j+1}]) &\\ &\le \Omega(f)\sum_{j\in\Lambda'}(x_{j+1} - x_j)\\ &\le \Omega(f)\sum_{i=1}^{n}(\beta_i - \alpha_i) \le \Omega(f)\varepsilon' .\end{aligned}$$

Il s'ensuit que :

$$\begin{aligned}\omega(f,\sigma) &= \sum_{j=0}^{m-1}(x_{j+1} - x_j)\omega(f,[x_j, x_{j+1}])\\ &\le \varepsilon'(b-a) + \Omega(f)\varepsilon' = \varepsilon,\end{aligned}$$

ce qui prouve que f est intégrable au sens de RIEMANN en vertu du théorème 2. ■

COROLLAIRE 4. — *Si $f : [a,b] \to \mathbb{R}$ est intégrable au sens de* RIEMANN, *la fonction $|f|$ l'est aussi et on a :*

$$\left|\int_a^b f(x)dx\right| \le \int_a^b |f(x)|dx .$$

Comme la fonction $|f|$ est continue aux points où f l'est, elle est RIEMANN intégrable en vertu du théorème 3. En outre, on a $|S(f,\sigma,\xi)| \le S(|f|,\sigma,\xi)$ pour toute subdivision riemannienne (σ,ξ) de $[a,b]$. Le corollaire s'en déduit par passage à la limite. ■

COROLLAIRE 5. — *Si* $f,g:[a,b]\to\mathbb{R}$ *sont intégrables au sens de* RIEMANN, *il en va de même des fonctions* $fg, Sup(f,g), Inf(f,g)$.

C'est une conséquence immédiate du théorème 3. ■

COROLLAIRE 6. — *Soient* $a<b<c$. *Une fonction* $f:[a,b]\to\mathbb{R}$ *est intégrable au sens de* RIEMANN *si et seulement si elle est intégrable sur* $[a,c]$ *et sur* $[c,b]$, *et on a :*

$$\int_a^b f(x)dx=\int_a^c f(x)dx+\int_c^b f(x)dx\,.$$

Nous laissons le soin au lecteur d'établir ce résultat, en utilisant le théorème 3 et en observant qu'une subdivision de $[a,c]$ et une subdivision de $[c,d]$ fournissent, par juxtaposition, une subdivision de $[a,b]$. ■

Soit $f:[a,b]\to\mathbb{R}$ une fonction intégrable au sens de RIEMANN. Pour tout $\alpha\in[a,b]$, on posera par convention : $\int_\alpha^\alpha f(x)dx=0$. Par ailleurs, si α,β sont deux points de $[a,b]$ tels que $\beta<\alpha$, on posera :

$$\int_\alpha^\beta f(x)dx=-\int_\beta^\alpha f(x)dx\ .$$

Avec ces notations, la relation de CHASLES pour les intégrales :

$$\int_\alpha^\beta f(x)dx=\int_\alpha^\gamma f(x)dx+\int_\gamma^\beta f(x)dx$$

est vérifiée quels que soient les points $\alpha,\beta,\gamma\in[a,b]$.

3.4. Intégrale de Riemann et primitives. — Soit $f:[a,b]\to\mathbb{R}$ une fonction intégrable au sens de

RIEMANN sur $[a,b]$. On appelle *intégrale indéfinie* de f l'une quelconque des fonctions :

$$F(x)=\int_a^x f(t)dt+C\ ,$$

où C est une constante réelle. Indiquons quelques propriétés de l'intégrale indéfinie.

PROPOSITION 7. — *Soit* $f:[a,b]\to\mathbb{R}$ *une fonction intégrable au sens de* RIEMANN *sur* $[a,b]$ *et posons, pour tout* $x\in[a,b]$ *:*

$$F(x)=\int_a^x f(t)dt\ .$$

(i) F *est continue sur* $[a,b]$ *;*

(ii) F *est dérivable en tout point* x *où* f *est continue, et on a* $F'(x)=f(x)$. *En particulier,* $F'(x)=f(x)$ *pour presque tout* $x\in[a,b]$.

La propriété *(i)* résulte immédiatement de l'inégalité :

$$\left|F(x)-F(y)\right|=\left|\int_y^x f(t)dt\right|\leq\|f\|_\infty\left|x-y\right|\ ,$$

où x,y sont deux points quelconques de $[a,b]$. Si f est continue au point x_o il existe, pour tout $\varepsilon>0$, un réel $\eta>0$ tel que l'on ait :

$$\left|x-x_o\right|\leq\eta\Rightarrow\left|f(x)-f(x_o)\right|\leq\varepsilon\ .$$

Pour $0<\left|x-x_o\right|\leq\eta$, on a alors :

$$\begin{aligned}\left|\frac{F(x)-F(x_o)}{x-x_o}-f(x_o)\right|&=\frac{1}{|x-x_o|}\left|\int_{x_o}^x (f(t)-f(x_o))dt\right|\\&\leq\frac{1}{|x-x_o|}\varepsilon\left|\int_{x_o}^x dt\right|\\&\leq\frac{\varepsilon|x-x_o|}{|x-x_o|}=\varepsilon,\end{aligned}$$

ce qui prouve que F est dérivable en x_o, de dérivée :

$$F'(x_o)=f(x_o)\ .$$

Comme les discontinuités d'une fonction RIEMANN intégrable constituent un ensemble négligeable, la proposition 7 est démontrée. ■

Ainsi, pour une fonction $f:[a,b]\to\mathbb{R}$ intégrable au sens de RIEMANN, la fonction $F(x)=\int_a^x f(t)dt$ est continue et vérifie $F'(x)=f(x)$ pour presque tout $x\in[a,b]$. La fonction F peut donc être considérée comme une primitive de f en un sens généralisé. On observera cependant que, si les fonctions $x\to F(x)+C^{ste}$ admettent $f(x)$ pour dérivée en presque tout point x, ce ne sont pas les seules fonctions continues à posséder cette propriété. En effet, on peut montrer l'existence de fonctions continues non constantes admettant presque partout une dérivée nulle. En ajoutant une telle fonction à F, on obtient une fonction continue dont la dérivée est presque partout égale à f, mais qui n'est pas égale à F à une constante près. L'intégrale de RIEMANN permet cependant d'exprimer une fonction dérivable à partir de sa dérivée, à condition que cette dernière soit bornée et presque partout continue. On a en effet :

PROPOSITION 8. — *Soit $F:[a,b]\to\mathbb{R}$ une fonction dérivable sur $[a,b]$. Si F' est bornée et presque partout continue, on a pour tout $x\in[a,b]$:*

$$F(x)=F(a)+\int_a^x F'(t)dt.$$

En effet, la fonction $f=F'$ est RIEMANN intégrable en vertu du critère de LEBESGUE. Par ailleurs, fixons $x\in[a,b]$ et subdivisons l'intervalle $[a,b]$ en n segments $[x_i,x_{i+1}]$ de longueur $\frac{x-a}{n}$. D'après le théorème des accroissements finis, il existe pour tout $i = 0,1,...,n-1$ un réel $\xi_i \in [x_i, x_{i+1}]$ tel que l'on ait :

$$F(x_{i+1}) - F(x_i) = (x_{i+1} - x_i) f(\xi_i).$$

On en déduit que :

$$\begin{aligned} F(x) - F(a) &= \sum_{i=0}^{n-1} F(x_{i+1}) - F(x_i) \\ &= \sum_{i=0}^{n-1} (x_{i+1} - x_i) f(\xi_i), \end{aligned}$$

d'où l'égalité désirée en faisant tendre n vers l'infini, puisque $f = F'$ est intégrable au sens de RIEMANN. ■

On notera qu'une fonction admettant une primitive au sens indiqué au chapitre I n'est pas nécessairement intégrable au sens de RIEMANN, même si elle est bornée. C'est par exemple le cas de la fonction de DIRICHLET. Par ailleurs, on peut montrer qu'il existe des fonctions $F : [a,b] \to \mathbb{R}$ dérivables, mais dont la dérivée n'est pas RIEMANN intégrable. L'intégrale de RIEMANN ne peut donc pas réellement être considérée comme l'opération inverse de la dérivation[4].

EXERCICES

Pour éviter toute confusion dans ce qui suit, rappelons qu'une fonction f admet une primitive F sur $[a,b]$ si $F'(x) = f(x)$ sauf peut-être pour un ensemble dénombrable de valeurs de x. Dans ce cas, l'intégrale définie $\int_a^x f(t)dt$ a un sens ; c'est par définition la différence $F(x) - F(a)$.

EXERCICE 1. On considère la fonction $f : [0,1] \to \mathbb{R}$ définie par $f(x) = 1$ si $\frac{1}{2n} < x \leq \frac{1}{2n-1}$ $(n = 1,2,...)$, et

[4] C'est précisément pour essayer de clarifier la question de la primitivation des fonctions bornées que LEBESGUE a introduit l'intégrale qui porte son nom.

$f(x)=0$ sinon. Montrer que f est RIEMANN intégrable sur $[0,1]$, mais qu'elle n'est pas réglée sur ce segment.

EXERCICE 2 *(Une fonction non RIEMANN intégrable, dont le module est intégrable).* On considère la fonction $f:[0,1]\to\mathbb{R}$ définie par $f(x)=1$ si x est rationnel et $f(x)=-1$ sinon.

a) Montrer que f n'est pas RIEMANN intégrable sur $[0,1]$, mais que $|f|$ l'est.

b) Montrer que l'intégrale définie $F(x)=\int_0^x f(t)dt$ a un sens pour tout $x\in[0,1]$ et donner sa valeur en fonction de x. Montrer que F est dérivable partout, mais que F' n'est pas égale à f.

EXERCICE 3 *(Une fonction réglée à ensemble de discontinuités partout dense).* Pour tout $x\geq 0$, on note (x) la différence entre x et l'entier le plus voisin. Considérons la fonction $f:[0,1]\to\mathbb{R}$ définie par[5] :

$$f(x)=\frac{(x)}{1}+\frac{(2x)}{4}+\frac{(3x)}{9}+\dots$$

a) Déterminer l'ensemble des discontinuités de f. Montrer que cet ensemble est dénombrable et partout dense dans $[0,1]$.

b) Montrer que f est réglée. En déduire que f est intégrable au sens de RIEMANN sur $[0,1]$.

EXERCICE 4 (*Une fonction RIEMANN intégrable qui n'est pas réglée*). Soit $K\subset[0,1]$ l'ensemble triadique de CANTOR. On considère la fonction $f:[0,1]\to\mathbb{R}$ égale à 1 en tout point de K et à 0 sinon. Montrer que les points de discontinuité de f sont exactement les points de K. En déduire que f est RIEMANN intégrable, mais qu'elle n'est pas réglée.

---

[5] Cet exemple est dû à RIEMANN.

EXERCICE 5. Soit $K \subset [0,1]$ l'ensemble triadique de CANTOR. Pour tout $x \in [0,1]$, on pose :

$$d(x,K) = Inf_{u \in K} |x-u| .$$

a) Montrer que la fonction $x \to d(x,K)$ est continue sur $[0,1]$.

b) On considère la fonction $f : [0,1] \to \mathbb{R}$ définie par $f(x) = d(x,K)$ si $x \notin K$ et $f(x) = 1$ si $x \in K$. Montrer que f est RIEMANN intégrable, mais qu'elle n'est pas réglée (on pourra montrer que K est l'ensemble des points de discontinuité de f).

EXERCICE 6. Montrer que la série de fonctions :

$$\sum_{n=0}^{+\infty} (-1)^n \frac{x^n}{n+1}$$

est uniformément convergente sur $[0,1]$. Calculer sa somme et montrer que l'on a :

$$1 - \frac{1}{2^2} + \frac{1}{3^2} - ... + \frac{(-1)^n}{(n+1)^2} + ... = \int_o^1 ln(1+x) \frac{dx}{x} .$$

EXERCICE 7. Montrer que, pour $0 \le r < 1$, la série de fonctions :

$$\sum_{n=1}^{+\infty} cos(nx) \frac{r^n}{n}$$

est uniformément convergente sur $[0,2\pi]$. En déduire que :

$$\int_o^{2\pi} ln(1 + r^2 - 2r \cos x) dx = 0 .$$

EXERCICE 8. Montrer que l'on a, pour toute fonction continue $f : [0,1] \to \mathbb{R}$:

$$\lim_{n \to +\infty} n \int_0^1 x^n f(x) dx = f(1) .$$

(On pourra se ramener au cas où $f(1) = 0$ et majorer chacune des intégrales $\int_o^a$ et $\int_a^1$ pour un choix convenable de a) .

EXERCICE 9 *(Une limite simple des fonctions RIEMANN intégrable qui n'est pas intégrable).* On range l'ensemble des rationnels compris entre 0 et 1 en une suite $(r_1, r_2, ...)$. Pour tout entier $n \geq 1$, on note f_n la fonction définie sur $[0,1]$ par :

$$f_n(x) = 1 \text{ si } x \in \{r_1, r_2, ..., r_n\} \text{ et } f_n(x) = 0 \text{ sinon.}$$

a) Montrer que les f_n sont intégrables au sens de RIEMANN sur $[0,1]$.

b) Montrer que la suite $(f_1, f_2, ...)$ est croissante et qu'elle converge en tout point vers une fonction f qui n'est pas RIEMANN intégrable sur $[0,1]$. La convergence est-elle uniforme ?

c) Montrer que la limite $\lim_{n\to\infty} \int_0^1 f_n(x)dx$ existe, que f possède une primitive F sur $[0,1]$, mais que :

$$\lim_{n\to\infty} \int_0^1 f_n(x)dx \neq F(1) - F(0).$$

EXERCICE 10 *(Exemple d'une fonction RIEMANN intégrable avec un ensemble dense de discontinuités).* On considère la fonction périodique de période 1 définie[6] sur $\mathbb{R}$ par :

$$e(x) = x \text{ si } -\tfrac{1}{2} < x < \tfrac{1}{2} \text{ et } e(\pm\tfrac{1}{2}) = 0.$$

a) Montrer que, pour tout $x \in \mathbb{R}$, la série $\sum_{n=1}^{\infty} \frac{e(nx)}{n^2}$ est convergente.

b) Pour $x \in [0,1]$, on pose $f(x) = \sum_{n=1}^{\infty} \frac{e(nx)}{n^2}$. Déterminer l'ensemble D des discontinuités de f. Montrer que D est dénombrable et partout dense dans $[0,1]$.

c) Montrer que f est RIEMANN intégrable sur $[0,1]$ et calculer son intégrale de RIEMANN.

EXERCICE 11 *(Une fonction dérivable à dérivée non RIEMANN intégrable).* Soit $F : [0,1] \to \mathbb{R}$ la fonction définie par $F(x) = x^2 \sin(\frac{1}{x^2})$ si $x \neq 0$ et $F(0) = 0$ sinon.

[6] La construction est due à RIEMANN.

Montrer que F est dérivable sur $[0,1]$, mais que sa dérivée F' n'est pas RIEMANN intégrable sur $[0,1]$ (on montrera que F' n'est pas bornée).

EXERCICE 12 *(Fonction continue, à dérivée nulle presque partout mais non constante).* Soit K l'ensemble triadique de Cantor, et notons K_n le fermé obtenu lors de sa construction après avoir enlevé successivement 2^n segments ouverts de $[0,1]$. Soit $f_n:[0,1]\to\mathbb{R}$ l'unique fonction continue définie en posant : *(i)* $f_n(0)=0$, $f_n(1)=1$; *(ii)* f_n est constante sur chaque intervalle ouvert du complémentaire de K_n et prend respectivement les valeurs $\frac{1}{2^n},\frac{2}{2^n},\ldots,\frac{2^n-1}{2^n}$ sur ces intervalles, énumérés de la gauche vers la droite ; *(iii)* f_n est affine sur les segments fermés de K_n.

a) Montrer que les fonctions f_n sont croissantes et convergent uniformément vers une fonction continue f sur $[0,1]$.

b) Montrer que f est croissante et que $f'=0$ sur le complémentaire de K dans $[0,1]$. En déduire l'existence de fonctions croissantes et continues sur $[0,1]$, dérivables presque partout et de dérivée nulle, mais qui ne sont pas constantes.

c) Calculer $\int_o^1 f(x)dx$ et $\int_o^1 xf(x)dx$.

EXERCICE 13 *(Une fonction RIEMANN intégrable qui n'admet pas de primitive).* Soit K l'ensemble triadique de Cantor, et notons K_n le fermé obtenu lors de la construction de K après avoir enlevé successivement 2^n segments ouverts de $[0,1]$. Considérons la fonction $f:[0,1]\to\mathbb{R}$ définie par $f(x)=1$ si $x\in K$ et $f(x)=0$ sinon.

a) Montrer que f est intégrable au sens de RIEMANN sur $[0,1]$.

On suppose désormais qu'il existe une fonction continue $F:[0,1]\to\mathbb{R}$ dérivable sauf sur un ensemble dénombrable D et qui vérifie $F'(x) = f(x)$ pour $x \notin D$. On pose $G(x) = F(x) - F(0)$. Pour tout entier $n \geq 1$, on note $G_n : [0,1] \to \mathbb{R}$ l'unique fonction continue définie en posant : *(i)* $G_n(0) = 0$, *(ii)* G_n est affine de pente 1 sur chaque intervalle fermé de K_n ; *(iii)* G_n est constante sur les intervalles du complémentaire de K_n.

b) Montrer que les fonctions G_n convergent uniformément sur $[0,1]$ vers la fonction nulle.

c) Montrer que la dérivée de $G - G_n$ est ≤ 0 sur le complémentaire d'un ensemble dénombrable. En déduire que $G(x) \leq G_n(x)$ pour tout x.

d) Montrer que F est constante, et que sa dérivée ne peut pas être égale à f en dehors d'un ensemble dénombrable.

e) Déduire de ce qui précède que f n'admet pas de primitive F au sens indiqué.

Thierry Fack

1. ESPACES MESURÉS

1. ESPACES MESURÉS

1.1- TRIBUS

1.1.1- Définition. Soit X un ensemble. On appelle tribu (ou σ-algèbre) sur X une famille $\mathcal{M}$ de parties de X vérifiant les propriétés suivantes :

(i) $X \in \mathcal{M}$;
(ii) si $A \in \mathcal{M}$, alors $A^c = X \setminus A \in \mathcal{M}$;
(iii) si $A_1, A_2, \ldots$ est une suite(*) d'éléments de $\mathcal{M}$, alors $A = \bigcup_{n \geq 1} A_n \in \mathcal{M}$.

1.1.2 - Espace mesurable. On appelle espace mesurable la donnée d'un couple $(X, \mathcal{M})$ où X est un ensemble et $\mathcal{M}$ une tribu sur X. Les éléments de $\mathcal{M}$ sont alors appelés les parties mesurables de X.

1.1.3 - Soit $\mathcal{M}$ une tribu sur un ensemble X. Alors, $\phi = X^c \in \mathcal{M}$. Par ailleurs, si $A_1, A_2, \ldots$ est une suite d'éléments de $\mathcal{M}$, la relation :

$$\bigcap_{n \geq 1} A_n = \Big(\bigcup_{n \geq 1} A_n^c\Big)^c$$

montre que $\bigcap_{n \geq 1} A_n \in \mathcal{M}$. Il s'ensuit que, si $A, B \in \mathcal{M}$, la différence ensembliste $A \setminus B = A \cap B^c$ appartient à $\mathcal{M}$. Ainsi, une tribu est stable par réunion dénombrable, intersection dénombrable, différence ensembliste et passage au complémentaire.

1.1.4. Exemples. Soit X un ensemble. La plus petite tribu sur X est la classe $\{\phi, X\}$, et la plus grande est la famille $\mathcal{P}(X)$ de toutes les parties de X. La classe des parties A de X qui sont dénombrable ou de complémentaire dénombrable constitue également une tribu sur X.

1.1.5- Tribu engendrée. Soient X un ensemble et $\mathcal{F}$

(*) Les suites peuvent être éventuellement finies.

une famille quelconque de sous-ensembles de X. L'intersection de toutes les tribus de X qui contiennent $\mathcal{F}$ est encore une tribu qui contient $\mathcal{F}$; la vérification est immédiate. Cette tribu est évidemment la plus petite tribu qui contient $\mathcal{F}$. On dit que c'est la <u>tribu engendrée par $\mathcal{F}$</u>.

<u>1.1.6- Exemples</u>. Soit X un ensemble. La tribu engendrée par les points de X contient nécessairement les parties dénombrables de X, ainsi que celles dont le complémentaire est dénombrable. Il s'ensuit que la tribu engendrée par les points de X est la tribu des parties de X qui sont dénombrables ou de complémentaire dénombrable.

Soient X un ensemble et $A_1, A_2, \ldots$ une partition dénombrable de X en ensembles non vides et deux à deux disjoints. Alors, la tribu engendrée par les A_i $(i = 1, 2, \ldots)$ est la classe des parties de X de la forme $\bigcup_{j \in J} A_j$, où J décrit l'ensemble $\mathcal{P}(\mathbb{N})$ de toutes les parties de $\mathbb{N}$.

<u>1.1.7- Tribu des Boréliens</u>. Soit X un espace topologique. On appelle <u>tribu des Boréliens</u> de X la tribu engendrée par les parties ouvertes de X. Les éléments de cette tribu sont appelés les <u>Boréliens</u> de X. Ainsi, toute partie ouverte, ou fermée, est borélienne.

La tribu des Boréliens de $\mathbb{R}^p$ est engendrée par les <u>pavés</u> de la forme

$$P = \prod_{i=1}^{p} (a_i, b_i)$$

où a_i, b_i sont des réels et où (a_i, b_i) désigne un intervalle de $\mathbb{R}$ d'origine a_i et d'extrémité b_i, qui peut aussi bien être ouvert que semi ouvert ou fermé. En effet, on a :

<u>1.1.8. Proposition</u>. <u>La tribu des Boréliens de $\mathbb{R}^p$ est</u>

engendrée par les pavés de la forme
$$P = \prod_{i=1}^{p} (a_i, b_i).$$

Démonstration. En utilisant la densité de $\mathbb{Q}^p$ dans $\mathbb{R}^p$, on montre aisément que tout ouvert non vide de $\mathbb{R}^p$ est réunion dénombrable de pavés ouverts de la forme
$$\prod_{i=1}^{p} \left] r_i - \frac{1}{k}, r_i + \frac{1}{k} \right[,$$
où les r_i sont des rationnels et où $k \in \mathbb{N} - \{0\}$. Il s'ensuit que la tribu engendrée par les pavés contient les ouverts de $\mathbb{R}^p$, donc contient la tribu des Boréliens. Par ailleurs, comme tout pavé est un Borélien de $\mathbb{R}^p$, la tribu des Boréliens contient la tribu engendrée par les pavés, d'où la proposition. ■

1.2. MESURES POSITIVES

1.2.1- Définition. Soient X un ensemble et $\mathcal{M}$ une tribu sur X. On appelle mesure positive sur $\mathcal{M}$ une fonction
$$\mu : \mathcal{M} \longrightarrow [0, +\infty]$$
qui vérifie les deux conditions suivantes :

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

(ii) $\mu\left(\bigcup_{n \geq 1} A_n\right) = \sum_{n \geq 1} \mu(A_n)$ pour toute suite $(A_1, A_2, \dots)$ de parties deux à deux disjointes de $\mathcal{M}$.

Lorsque $\mu(X) < +\infty$, on dit que la mesure μ est finie. Si $\mu(X) = 1$, on dit que μ est une mesure de probabilités. S'il existe une suite $(A_1, A_2, \dots)$ de parties de $\mathcal{M}$ telle que $\mu(A_n) < +\infty$ pour tout

$n \geqq 1$ et $X = \bigcup_{n \geqq 1} A_n$, on dit que la mesure μ est <u>σ-finie</u>.

La condition (i) a pour but d'écarter le cas pathologique où $\mu(A) = +\infty$ pour tout $A \in \mathcal{M}$. On aurait pu remplacer (i) par la condition:

(i') Il existe $A \in \mathcal{M}$ tel que $\mu(A) < +\infty$.

En effet, s'il existe $A \in \mathcal{M}$ tel que $\mu(A) < +\infty$, on a en vertu de (ii) :

$$\mu(A) = \mu(A \cup \emptyset) = \mu(A) + \mu(\emptyset),$$

d'où il résulte que $\mu(\emptyset) = 0$.

1.2.2. On appelle <u>espace mesuré</u> la donnée d'un triplet $(X, \mathcal{M}, \mu)$ où X est un ensemble, $\mathcal{M}$ une tribu sur X et μ une mesure positive sur $\mathcal{M}$. Les éléments de $\mathcal{M}$ sont appelés les <u>parties mesurables</u> de X. La donnée de $\mathcal{M}$ est souvent implicite ; on écrit alors (X, μ) au lieu de $(X, \mathcal{M}, \mu)$ et on dit que μ est une <u>mesure positive sur X</u> au lieu parler de mesure positive sur $\mathcal{M}$. Lorsque μ est finie (resp. σ-finie), on dit que $(X, \mathcal{M}, \mu)$ est un <u>espace mesuré fini</u> (resp. <u>σ-fini</u>). Lorsque $\mu(X) = 1$, on dit que $(X, \mathcal{M}, \mu)$ est un <u>espace de probabilités</u>.

Les espaces de probabilités fournissent de nombreux exemples d'espaces mesurés. Ces espaces décrivent souvent des expériences aléatoires. L'espace X représente alors l'ensemble des résultats possibles de l'expérience et $\mathcal{M}$ décrit les événements que l'on veut étudier. Par exemple, pour un jet de deux dés, on aura

$$X = \{(m, n) \in \mathbb{N} \times \mathbb{N} \, ; \ 1 \leqq m, n \leqq 6\}.$$

Le fait que le total des points obtenus est inférieur ou égal à huit est alors un <u>événement</u> $A \in \mathcal{M}$

donné par :

$$A = \{ (m,n) \in X \mid m+n \leqq 8 \}.$$

Dans beaucoup d'expériences aléatoires, la probabilité $\mu(A)$ d'un évènement $A \in \mathcal{M}$ est déterminée de manière empirique en répétant N fois l'expérience, et en notant $N(A)$ le nombre de fois où A est réalisé. La limite

$$\mu(A) = \lim_{N \to \infty} \frac{N(A)}{N},$$

à supposer qu'elle existe, fournit alors la probabilité $\mu(A) \in [0,1]$.

1.2.3- Exemples.

1.2.3.1. Mesure de dénombrement. Soit X un ensemble. Pour tout $A \subset X$, posons :

$$\mu(A) = \begin{cases} \mathrm{Card}(A) & \text{si } A \text{ est fini} \\ +\infty & \text{si } A \text{ est infini.} \end{cases}$$

On définit ainsi une mesure positive sur $\mathcal{P}(X)$ appelée mesure de dénombrement sur X.

1.2.3.2- Masse de Dirac en un point. Soient X un ensemble et $x_0 \in X$ un point de X. Pour tout sous-ensemble A de X, posons :

$$\mu(A) = \begin{cases} 1 & \text{si } x_0 \in A \\ 0 & \text{si } x_0 \notin A \end{cases}$$

On définit ainsi une mesure positive sur $\mathcal{P}(X)$ appelée masse de Dirac au point x_0.

1.2.3.3. Mesures discrètes. Soient X un ensemble, $a_1, a_2, \ldots$ une suite de points de X et $\alpha_1, \alpha_2, \ldots$ une suite de réels >0. Pour toute partie A de X, on pose

$$\mu(A) = \sum_{n \geq 1} \alpha_n \delta_{a_n}(A) = \sum_{n \geq 1,\, a_n \in A} \alpha_n$$

On définit ainsi une mesure positive μ sur X que l'on note :

$$\mu = \sum_{n \geq 1} \alpha_n \delta_{a_n}$$

1.2.3.4- Mesures Boréliennes. Soit X un espace topologique. On appelle <u>mesure Borélienne</u> sur X toute mesure positive sur la tribu des Boréliens de X qui vérifie :

$$\mu(K) < +\infty$$

pour tout compact K de X. La mesure de Lebesgue sur $\mathbb{R}^p$, dont la construction sera exposée plus loin, est une mesure Borélienne sur $\mathbb{R}^p$.

1.2.4 - Mesure induite. Soient $(X, \mathcal{M}, \mu)$ un espace mesuré et $A \in \mathcal{M}$ une partie mesurable de X. Posons $\mathcal{M}_A = \{ Z \in \mathcal{M} \mid Z \subset A \}$; on définit ainsi une tribu $\mathcal{M}_A$ sur A. La mesure μ induit alors une mesure positive μ_A sur A par la formule :

$$\mu_A(Z) = \mu(Z)$$

pour tout $Z \in \mathcal{M}_A$. On dit que μ_A est la <u>mesure induite</u> par μ sur A. Lorsque nous aurons construit la mesure de Lebesgue sur $\mathbb{R}^n$, la construction ci-dessus nous permettra de disposer automatiquement de la mesure de Lebesgue sur tout ouvert de $\mathbb{R}^n$.

Donnons maintenant quelques propriétés élémentaires des mesures.

1.2.5- Théorème. Soit $(X,\mathcal{M},\mu)$ un espace mesuré. On a :

(i) $A\subset B \Longrightarrow \mu(A)\leqq\mu(B)$ si $A,B\in\mathcal{M}$;

(ii) Pour toute suite croissante $A_1\subset A_2\subset\cdots$ d'éléments de $\mathcal{M}$, on a :

$$\mu\Big(\bigcup_{n\geqq 1} A_n\Big)=\lim_{n\to\infty}\mu(A_n)=\operatorname{Sup}_{n\geqq 1}\mu(A_n);$$

(iii) Pour toute suite décroissante $A_1\supset A_2\supset\cdots$ d'éléments de $\mathcal{M}$ telle que $\mu(A_1)<+\infty$, on a :

$$\mu\Big(\bigcap_{n\geqq 1} A_n\Big)=\lim_{n\to\infty}\mu(A_n)=\operatorname{Inf}_{n\geqq 1}\mu(A_n);$$

(iV) Pour toute suite $A_1,A_2,\ldots$ d'éléments de $\mathcal{M}$, on a :

$$\mu\Big(\bigcup_{n\geqq 1} A_n\Big)\leqq\sum_{n\geqq 1}\mu(A_n).$$

Démonstration.

(i) Si $A\subset B$, on a : $B=A\cup(B-A)$ avec $A\cap B-A=\emptyset$, d'où :

$$\mu(B)=\mu(A)+\mu(B-A)\geqq\mu(A).$$

(ii) Soit $A_1\subset A_2\subset\cdots$ une suite d'éléments de $\mathcal{M}$. On a :

$$\bigcup_{n\geqq 1}A_n=A_1\cup(A_2-A_1)\cup(A_3-A_2)\cup\cdots\cup(A_{n+1}-A_n)\cup\cdots$$

et, puisque A_1 et les $A_{n+1}-A_n$ sont des éléments de $\mathcal{M}$ deux à deux disjoints, on obtient :

$$\begin{aligned}\mu\Big(\bigcup_{n\geqq 1}A_n\Big)&=\mu(A_1)+\mu(A_2-A_1)+\cdots+\mu(A_{n+1}-A_n)+\cdots\\&=\lim_{n\to\infty}\mu(A_1)+\mu(A_2-A_1)+\cdots+\mu(A_n-A_{n-1})\\&=\lim_{n\to\infty}\mu(A_n).\end{aligned}$$

En remplaçant $\lim_{n\to\infty}$ par $\operatorname{Sup}_{n\geqq 1}$, on obtient (ii).

(iii) Posons $B_n = A_1 - A_n$. Alors, $B_1 \subset B_2 \subset \cdots$ et
$\bigcup_{n \geqq 1} B_n = A_1 - A$, où $A = \bigcap_{n \geqq 1} A_n$.

De (ii) on déduit :

$$\mu(A_1) - \mu(A) = \mu(A_1 - A) = \mu\Big(\bigcup_{n \geqq 1} B_n\Big) = \lim_{n \to \infty} \mu(B_n)$$
$$= \lim_{n \to \infty} \big(\mu(A_1) - \mu(A_n)\big) = \mu(A_1) - \lim_{n \to \infty} \mu(A_n),$$

d'où $\mu(A) = \lim_{n \to \infty} \mu(A_n)$. Comme la suite $(\mu(A_n))_{n \geqq 1}$ est décroissante, on a aussi :

$$\mu(A) = \operatorname{Inf}_{n \geqq 1} \mu(A_n).$$

(iv). Si $A, B \in \mathcal{M}$, on a : $A \cup B = A \cup (B - A)$ où A et $B - A$ sont disjoints, d'où :

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

On en déduit que, si $A_1, A_2, \ldots$ est une suite d'éléments de $\mathcal{M}$, on a :

$$\mu(A_1 \cup A_2 \cup \cdots \cup A_n) \leqq \mu(A_1) + \cdots + \mu(A_n).$$

Posons $B_n = A_1 \cup A_2 \cup \cdots \cup A_n$. Alors $B_1 \subset B_2 \subset \cdots$ et $\bigcup_{n \geqq 1} B_n = \bigcup_{n \geqq 1} A_n$, de sorte que l'on a, en vertu de (ii) :

$$\mu\Big(\bigcup_{n \geqq 1} A_n\Big) = \mu\Big(\bigcup_{n \geqq 1} B_n\Big) = \lim_{n \to \infty} \mu(B_n)$$
$$\leqq \lim_{n \to \infty} \big(\mu(A_1) + \cdots + \mu(A_n)\big) = \sum_{n \geqq 1} \mu(A_n).$$

■

1.2.6 - Ensembles de mesure nulle. Soit $(X, \mathcal{M}, \mu)$ un espace mesuré. On dit qu'un sous-ensemble N de X est μ-négligeable (ou de mesure nulle) s'il existe $A \in \mathcal{M}$ tel que $N \subset A$ et $\mu(A) = 0$. Les parties $N \in \mathcal{M}$ telles que $\mu(N) = 0$ sont donc en particulier μ-négligeables.

La réunion $N = N_1 \cup N_2 \cup \ldots$ d'une suite $(N_1, N_2, \ldots)$ de parties μ-négligeables est encore μ-négligeable. En effet, il existe pour tout $n \geqq 1$ une partie $A_n \in \mathcal{M}$ telle que :

$$N_n \subset A_n \quad \text{et} \quad \mu(A_n) = 0.$$

On en déduit que $N \subset A = \bigcup\limits_{n \geqq 1} A_n$ où :

$$\mu(A) = \mu\Big(\bigcup_{n \geqq 1} A_n\Big) \leqq \sum_{n \geqq 1} \mu(A_n) = 0,$$

ce qui prouve que N est μ-négligeable.

1.2.7- <u>Complétion d'un espace mesuré</u>. Soit $(X, \mathcal{M}, \mu)$ un espace mesuré. Les parties $A \in \mathcal{M}$ sont appelées mesurables. On pose :

1.2.7.1- <u>Définition</u>. <u>On dit qu'une partie $A \subset X$ est μ-mesurable si elle est de la forme</u>

$$\underline{A = B \cup N}$$

<u>où $B \in \mathcal{M}$ et où $N \subset X$ est μ-négligeable.</u>

On notera $\mathcal{M}_\mu$ la famille des parties μ-mesurables.

1.2.7.2- <u>Proposition</u>. <u>Soit $(X, \mathcal{M}, \mu)$ un espace mesuré. Alors, les parties μ-mesurables forment une tribu $\mathcal{M}_\mu$ qui contient $\mathcal{M}$.</u>

<u>Démonstration</u>. On a $\mathcal{M} \subset \mathcal{M}_\mu$, de sorte que $X \in \mathcal{M}_\mu$. Comme $\mathcal{M}$ est une tribu et qu'une réunion dénombrable d'ensembles μ-négligeables est μ-négligeable, $\mathcal{M}_\mu$ est stable par réunion dénombrable. Il suffit donc, pour terminer la démonstration, de montrer que si $A \in \mathcal{M}_\mu$, alors $A^c \in \mathcal{M}_\mu$. Ecrivons

$$A = B \cup N \quad \text{où } B \in \mathcal{M} \text{ et } N \subset Z \in \mathcal{M} \text{ avec } \mu(Z) = 0.$$

On a :

$$(B \cup N)^c = (B \cup Z)^c \cup \big((Z - N) \cap B^c\big) \text{ et,}$$

puisque $(B\cup Z)^c = B^c \cap Z^c \in \mathcal{M}$ et que $B^c \cap (Z-N) \subset Z$ avec $\mu(Z)=0$, i.e. $B^c \cap (Z-N)$ est μ-négligeable, on obtient que $(B\cup N)^c \in \mathcal{M}_\mu$. ■

Pour toute partie μ-mesurable $A = B\cup N$ ($B\in\mathcal{M}$, N μ-négligeable), posons :

$$\mu(A) = \mu(B).$$

On définit ainsi un nombre $\mu(A)\in[0,+\infty]$ qui ne dépend que de A, et pas de la décomposition $A = B\cup N$. En effet, si $A = B_1\cup N_1$ avec $B_1\in\mathcal{M}$ et N_1 μ-négligeable, il existe $Z_1\in\mathcal{M}$ tel que $N_1\subset Z_1$ et $\mu(Z_1)=0$, d'où :

$$B\subset A = B_1\cup N_1 \subset B_1\cup Z_1$$

et donc

$$\mu(B) \leqq \mu(B_1\cup Z_1) \leqq \mu(B_1)+\mu(Z_1) = \mu(B_1).$$

En échangeant les rôles de B et de B_1, on obtient

$$\mu(B_1)\leqq \mu(B)$$

d'où, finalement, $\mu(B)=\mu(B_1)$. Ainsi, $\mu(A)$ est défini sans ambiguïté pour $A\in\mathcal{M}_\mu$. On a alors :

<u>1.2.7.3 - Proposition</u>. <u>L'application $A\in\mathcal{M}_\mu \longmapsto \mu(A)\in[0,+\infty]$ est une mesure positive sur $\mathcal{M}_\mu$.</u>

<u>Démonstration</u>. Il suffit de montrer que, si $A_1, A_2, \ldots$ est une suite de parties μ-mesurables deux à deux disjointes, on a :

$$\mu\Big(\bigcup_{n\geqq 1} A_n\Big) = \sum_{n\geqq 1}\mu(A_n).$$

A cet effet, écrivons $A_n = B_n\cup N_n$ avec $B_n\in\mathcal{M}$ et N_n μ-négligeable. On a :

$$\bigcup_{n\geq 1} A_n = \Big(\bigcup_{n\geq 1} B_n\Big) \cup N$$

où $N = \bigcup_{n\geq 1} N_n$ est μ-négligeable et où les $B_n \in \mathcal{M}$ sont deux à deux disjointes. On a donc :

$$\mu\Big(\bigcup_{n\geq 1} A_n\Big) = \mu\Big(\bigcup_{n\geq 1} B_n\Big) = \sum_{n\geq 1} \mu(B_n) = \sum_{n\geq 1} \mu(A_n).$$

∎

Ainsi, pour tout espace mesuré $(X, \mathcal{M}, \mu)$, on a construit un espace mesuré $(X, \mathcal{M}_\mu, \mu)$ pour lequel les parties μ-négligeables appartiennent à la tribu des sous-ensembles μ-mesurables. On dit que $(X, \mathcal{M}_\mu, \mu)$ est le <u>complété</u> de l'espace $(X, \mathcal{M}, \mu)$.

1.3 - CONSTRUCTION DE LA MESURE DE LEBESGUE SUR $\mathbb{R}^p$.

<u>1.3.1 - Mesure des pavés</u>. Un <u>pavé</u> de $\mathbb{R}^p$ est par définition un produit d'intervalles bornés (a_i, b_i) $(1 \leq i \leq n)$:

$$P = \prod_{1\leq i\leq p} (a_i, b_i).$$

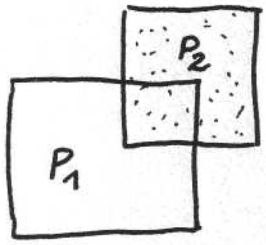
On ne précise pas ici si (a_i, b_i) est ouvert, fermé ou semi-ouvert. Le volume p-dimensionnel de P est par définition :

$$\lambda(P) = \prod_{1\leq i\leq p} (b_i - a_i).$$

Pour $p = 1$, un pavé est un intervalle borné (a, b), et $\lambda((a,b)) = b - a$ est la longueur de cet intervalle. On a vu (cf. 1.1.8) que les pavés engendrent la tribu Borélienne $\mathcal{B}$ de $\mathbb{R}^p$, et on se demande si la connaissance de la mesure des pavés permet de définir une mesure positive sur $\mathcal{B}$. A cet effet, on étend tout d'abord le volume p-dimensionnel aux réunions finies de pavés.

1.3.2. Mesure des réunions finies de pavés

Notons $\mathcal{R}$ l'ensemble des parties de $\mathbb{R}^p$ qui sont réunions finies de pavés de $\mathbb{R}^p$. On vérifie immédiatement qu'une réunion finie de pavés s'écrit toujours comme une réunion finie de pavés deux à deux disjoints.



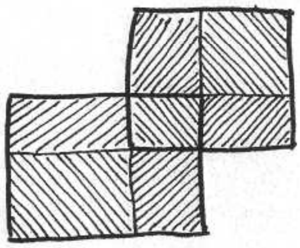


Fig. Décomposition de $P_1 \cup P_2$ en pavés disjoints dans le cas $p=2$.

La famille $\mathcal{R}$ n'est pas une tribu sur $\mathbb{R}^p$, mais elle a néanmoins une structure d'algèbre au sens suivant:

1.3.2.1- Définition. On appelle algèbre sur un ensemble X une famille $\mathcal{R}$ de parties de X vérifiant les conditions suivantes :

(i) $A, B \in \mathcal{R} \Longrightarrow A \cup B \in \mathcal{R}$;

(ii) $A, B \in \mathcal{R} \Longrightarrow A \cap B \in \mathcal{R}$;

(iii) $A, B \in \mathcal{R} \Longrightarrow A - B = A \cap B^c \in \mathcal{R}$.

Avec cette définition, on a:

1.3.2.2 - Proposition. La famille $\mathcal{R}$ des parties de $\mathbb{R}^p$ qui sont réunions finies de pavés de $\mathbb{R}^p$ constitue une algèbre sur $\mathbb{R}^p$.

Démonstration. Il est clair que $\mathcal{R}$ est stable par réunions finies et intersections finies. Pour démontrer que, si $A, B \in \mathcal{R}$, alors $A - B \in \mathcal{R}$, on se ramène au cas où A et B sont des pavés, pour lequel la conclusion est immédiate. ∎

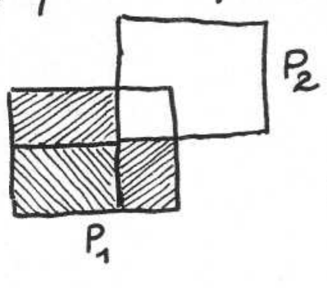


Fig. Décomposition de $P_1 \setminus P_2$ pour $p=2$.

1.3.2.3. Soit $A \in \mathcal{R}_0$. D'après ce qui précède, on sait que A s'écrit comme une réunion de pavés deux à deux disjoints :

$$A = P_1 \cup \cdots \cup P_n .$$

On pose alors :

$$\lambda(A) = \lambda(P_1) + \cdots + \lambda(P_n).$$

On vérifie que le nombre $\lambda(A)$ ainsi défini ne dépend que de A et pas de la décomposition de A en pavés deux à deux disjoints. Cette vérification, fastidieuse et sans intérêt, repose sur le fait que $\lambda(A)$ ne change pas lorsqu'on « raffine » la décomposition de A en pavés deux à deux disjoints. La figure ci-dessous illustre ce principe dans le cas $p=2$.

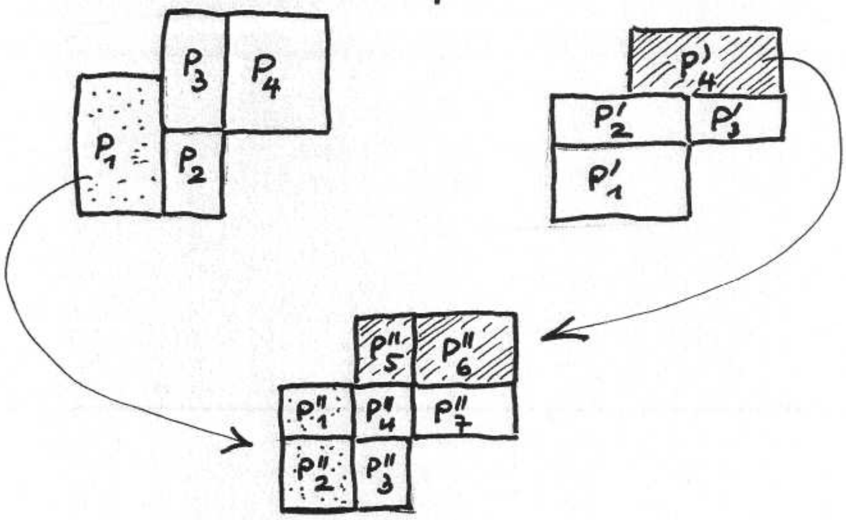


Fig. Raffinement des décompositions de A en réunions disjointes de pavés P_i et P'_j. Cas $p=2$.

Avec cette définition de $\lambda(A)$ pour $A \in \mathcal{R}_0$, il n'est pas difficile de démontrer que l'on a :

$$\lambda(A_1 \cup \cdots \cup A_n) = \lambda(A_1) + \cdots + \lambda(A_n)$$

si les $A_i \in \mathcal{R}_0$ sont deux à deux disjoints. L'application $A \longmapsto \lambda(A)$ possède en fait une propriété de σ-additivité qui en fait une mesure sur l'algèbre $\mathcal{R}_0$ au sens suivant :

1.3.2.4- Définition. Soient X un ensemble et $\mathcal{R}_0$ une algèbre de parties de X. On appelle mesure positive sur $\mathcal{R}_0$ une application $\mu : \mathcal{R}_0 \longrightarrow [0,+\infty]$ qui vérifie les propriétés suivantes :

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

(ii) $\mu(\bigcup_{n\geq 1} A_n)=\sum_{n\geq 1} \mu(A_n)$ si les A_n sont des éléments de $\mathcal{R}_0$ deux à deux disjoints tels que $\bigcup_{n\geq 1} A_n \in \mathcal{R}_0$.

Avec cette définition, on a :

1.3.2.5. Théorème. L'application qui associe à toute partie $A\subset \mathbb{R}^p$ qui est une réunion finie de pavés sa mesure $\lambda(A)$ est une mesure positive sur l'algèbre $\mathcal{R}_0$ des parties de $\mathbb{R}^p$ qui sont réunions finies de pavés de $\mathbb{R}^p$.

Démonstration. Comme $\lambda(\emptyset)=0$, il suffit de démontrer que, si $A_1, A_2, \dots$ est une suite d'éléments de $\mathcal{R}_0$ telle que $A_i \cap A_j = \emptyset$ pour $i\neq j$ et $\bigcup_{n\geq 1} A_n \in \mathcal{R}_0$, on a :

$$(1) \quad \lambda(\bigcup_{n\geq 1} A_n)=\sum_{n\geq 1} \lambda(A_n).$$

A cet effet, opérons d'abord quelques réductions.

Étape 1. On peut se ramener au cas où $A=\bigcup_{n\geq 1} A_n$ est un pavé.

Supposons en effet que (1) soit vraie lorsque A est un pavé, et montrons que (1) est toujours vraie. Si $A = P_1 \cup \dots \cup P_m$ est une décomposition de A en pavés disjoints, on a en vertu de l'hypothèse :

$$\lambda(\bigcup_{n\geq 1} A_n \cap P_i) = \sum_{n\geq 1} \lambda(A_n\cap P_i), \text{ soit :}$$

$$\lambda(A\cap P_i) = \sum_{n\geq 1} \lambda(A_n \cap P_i).$$

En sommant par rapport à i et en utilisant l'additivité finie de λ, on obtient :

$$\lambda(A) = \sum_{i=1}^{m} \lambda(A \cap P_i) = \sum_{i=1}^{m} \sum_{n \geqq 1} \lambda(A_n \cap P_i)$$
$$= \sum_{n \geqq 1} \left(\sum_{i=1}^{m} \lambda(A_n \cap P_i) \right) = \sum_{n \geqq 1} \lambda(A_n),$$

d'où (1) en toute généralité.

Étape 2. On peut se ramener au cas où les A_n et A sont des pavés.

Supposons en effet que (1) soit vraie lorsque les A_n et A sont des pavés, et montrons que (1) est toujours vraie. On peut supposer, sans perte de généralité, que A est un pavé P (étape 1). Écrivons alors chaque A_n comme une réunion finie de pavés P_n^i deux à deux disjoints :

$$A_n = P_n^1 \cup \dots \cup P_n^i \cup \dots \cup P_n^{\nu_n}.$$

Par hypothèse, on a :

$$\lambda\left(\bigcup_{n \geqq 1} A_n \right) = \lambda\left(\bigcup_{n,i} P_n^i \right) = \sum_{n,i} \lambda(P_n^i)$$
$$= \sum_{n \geqq 1} \left(\sum_{i=1}^{\nu_n} \lambda(P_n^i) \right) = \sum_{n \geqq 1} \lambda(A_n),$$

et (1) est vraie en toute généralité.

Étape 3. Démontrons (1) lorsque les $A_n = P_n$ sont des pavés deux à deux disjoints dont la réunion est un pavé $A = P$.

Comme on a : $P_1 \cup P_2 \cup \dots \cup P_n \subset P$ où les P_i sont des pavés deux à deux disjoints, on a :

$$\sum_{i=1}^{n} \lambda(P_i) = \lambda(P_1 \cup \dots \cup P_n) \leqq \lambda(P) \quad \text{et donc,}$$

en faisant tendre n vers $+\infty$:

$$\sum_{n\geqq 1} \lambda(P_n) \leqq \lambda(P).$$

Pour montrer l'inégalité inverse, fixons $\varepsilon>0$ et écrivons

$$P=\prod_{i=1}^{p}(a_i,b_i), \quad P_n=\prod_{i=1}^{p}(a_{n,i},b_{n,i}).$$

Fixons $\delta>0$ suffisamment petit pour que l'on ait :

$$a_i<a_i+\delta<b_i-\delta<b_i \quad \text{pour tout } i=1,\dots,p,$$

et posons

$$P_\delta=\prod_{i=1}^{p}[a_i+\delta, b_i-\delta].$$

Puisque $P=\bigcup_{n\geqq 1} P_n$, le pavé P_δ est recouvert par les pavés ouverts

$$P_{n,\varepsilon}=\prod_{i=1}^{p}\left]a_{n,i}-\frac{\varepsilon}{2^n}, b_{n,i}+\frac{\varepsilon}{2^n}\right[.$$

D'après la propriété de Borel-Lebesgue, il existe un entier $N\geqq 1$ tel que

$$P_\delta \subset \bigcup_{n=1}^{N} P_{n,\varepsilon}$$

Par sous-additivité finie de λ (*), on a :

$$\lambda(P_\delta)\leqq\lambda\Big(\bigcup_{n=1}^{N}P_{n,\varepsilon}\Big)\leqq\sum_{n=1}^{N}\lambda(P_{n,\varepsilon})$$

$$\leqq\sum_{n\geqq 1}\lambda(P_{n,\varepsilon}).$$

Or, puisque les P_n sont inclus dans P, les nombres

(*) Cette sous-additivité finie résulte immédiatement de la relation $\lambda(A\cup B)+\lambda(A\cap B)=\lambda(A)+\lambda(B)$, pour $A,B\in\mathcal{R}_0$.

$|a_{n,i}|$, $|b_{n,i}|$ sont bornés par une constante indépendante de n. On en déduit qu'il existe une constante $C > 0$ telle que l'on ait, pour $\varepsilon < 1$:

$$\prod_{i=1}^{p} \left(b_{n,i} - a_{n,i} + \frac{\varepsilon}{2^{n-1}}\right) \leq \prod_{i=1}^{p} (b_{n,i} - a_{n,i}) + C \frac{\varepsilon}{2^{n-1}}.$$

Il s'ensuit que :

$$\lambda(P_\delta) \leq \sum_{n \geq 1} \lambda(P_{n,\varepsilon}) \leq \sum_{n \geq 1} \lambda(P_n) + C\varepsilon \sum_{n \geq 1} \frac{1}{2^{n-1}}$$
$$= \sum_{n \geq 1} \lambda(P_n) + 2C\varepsilon.$$

En faisant tendre ε vers 0, on obtient :

$$\lambda(P_\delta) \leq \sum_{n \geq 1} \lambda(P_n).$$

Mais $\lambda(P_\delta) \xrightarrow[\delta \to 0]{} \lambda(P)$, de sorte que l'on obtient finalement :

$$\lambda(P) \leq \sum_{n \geq 1} \lambda(P_n),$$

et donc $\lambda(P) = \sum_{n \geq 1} \lambda(P_n)$. Ceci achève de montrer que λ est une mesure positive sur $\mathcal{R}$. ∎

1.3.2.6. On observera que $\mathbb{R}^p$ est réunion dénombrable d'une suite de parties de $\mathcal{R}$. En effet, les pavés

$$P_n = \underbrace{[-n,n] \times [-n,n] \times \cdots \times [-n,n]}_{p \text{ facteurs}}$$

ont pour réunion $\mathbb{R}^p$ et pour volume p-dimensionnel :

$$\lambda(P_n) = (2n)^p < +\infty.$$

Pour étendre la mesure λ sur $\mathcal{R}$ en une mesure positive sur les boréliens de $\mathbb{R}^p$, on utilisera le résultat suivant, dû à Hahn :

1.3.3 - Le théorème de Hahn.

1.3.3.1. Le théorème de Hahn permet d'étendre une mesure positive sur une algèbre de parties d'un espace X à la tribu engendrée par les parties de cette algèbre. Dans ce qui suit, nous désignerons par X un ensemble, par $\mathcal{R}$ une algèbre de parties de X, et par $\mathcal{B}$ la tribu engendrée par $\mathcal{R}$. On supposera donnée une mesure positive μ sur $\mathcal{R}$ vérifiant la propriété suivante :

(1) Il existe une suite $A_1, A_2, \dots$ d'éléments de $\mathcal{R}$ telle que $\mu(A_n) < +\infty$ pour tout $n \geqq 1$ et
$$X = \bigcup_{n \geqq 1} A_n.$$

Sous ces hypothèses, nous allons montrer que μ admet une unique extension en une mesure positive sur $\mathcal{B}$. D'après 1.3.2.5, les hypothèses sont vérifiées pour $X = \mathbb{R}^p$, $\mathcal{R}$ étant l'algèbre des parties de $\mathbb{R}^p$ qui sont réunions finies de pavés de $\mathbb{R}^p$, et $\mu = \lambda$ étant la mesure définie en 1.3.2.3. Avant de formuler le théorème de Hahn, introduisons la notion de mesure extérieure.

1.3.3.2 - **Mesure extérieure**. Soient X un ensemble, $\mathcal{R}$ une algèbre de parties de X et $\mu : \mathcal{R} \longrightarrow [0, +\infty]$ une mesure positive sur $\mathcal{R}$.

Définition. Pour toute partie $A \subset X$, on appelle mesure extérieure de A, et on note $\mu^*(A)$, le nombre
$$\mu^*(A) = \operatorname{Inf}\left\{ \sum_{n \geqq 1} \mu(A_n) \;\middle|\; A_n \in \mathcal{R} \text{ et } A \subset \bigcup_{n \geqq 1} A_n \right\} \in [0, +\infty].$$

La mesure extérieure $\mu^*(A)$ est définie pour toute partie $A \subset X$, même pour celles qui n'appartiennent pas à la tribu $\mathcal{B}$ engendrée par les parties de $\mathcal{R}$. Si $\mathcal{R}$ est une tribu, on vérifie aisément que l'on a :

$$\mu^*(A) = \text{Inf}\{\mu(B) \mid B \in \mathcal{R} \text{ et } A \subset B\}.$$

Par exemple, si X est un ensemble infini non dénombrable et si $\mathcal{R}$ est la tribu des parties de X qui sont dénombrables ou de complémentaire dénombrable, la formule

$$\mu(A) = \begin{cases} 0 & \text{si } A \text{ est dénombrable} \\ 1 & \text{si } A^c \text{ est dénombrable} \end{cases}$$

définit une mesure sur $\mathcal{R}$ (le vérifier) pour laquelle on a, quelle que soit la partie $A \subset X$:

$$\mu^*(A) = \begin{cases} 0 & \text{si } A \text{ est dénombrable} \\ 1 & \text{si } A \text{ n'est pas dénombrable} \end{cases}$$

Si $X = A_1 \cup A_2$ est une partition de X en deux sous-ensembles A_1 et A_2 non dénombrables tels que $A_1 \cap A_2 = \emptyset$, on a :

$$\mu^*(A_1 \cup A_2) = \mu^*(X) = 1 \quad \text{et}$$
$$\mu^*(A_1) + \mu^*(A_2) = 1 + 1 = 2,$$

de sorte que μ^* n'est certainement pas une mesure sur $\mathcal{P}(X)$. On a toutefois :

1.3.3.3. Proposition. Soient X un ensemble, $\mathcal{R}$ une algèbre de parties de X et $\mu : \mathcal{R} \to [0, +\infty]$ une mesure positive sur $\mathcal{R}$ vérifiant la condition (1). Alors on a :

(i) $\forall A \in \mathcal{R}$, $\mu^*(A) = \mu(A)$;

(ii) $\forall A, B \subset X$, on a : $A \subset B \Rightarrow \mu^*(A) \leq \mu^*(B)$;

(iii) Pour toute suite $A_1, A_2, \ldots$ de parties de X, on a :

$$\mu^*\Big(\bigcup_{n \geq 1} A_n\Big) \leq \sum_{n \geq 1} \mu^*(A_n);$$

(iv) Si $A \in \mathcal{R}$ est fixée, on a pour toute $E \subset A$ et toute $F \subset A^c$:
$$\mu^*(E \cup F) = \mu^*(E) + \mu^*(F).$$

Démonstration : (ii) est immédiate.

(i) Soit $A \in \mathcal{R}$. En prenant $A_1 = A$, $A_2 = A_3 = \dots = \emptyset$, on obtient que $\mu^*(A) \leq \mu(A)$. Si $A_1, A_2, \dots$ est une suite d'éléments de $\mathcal{R}$ telle que $A \subset \bigcup_{n \geq 1} A_n$, on a : $A = \bigcup_{n \geq 1} B_n$ avec $B_n = A_n \cap A \in \mathcal{R}$. En adaptant les assertions (ii) et (iv) du théorème 1.2.5 au cas d'une mesure sur une algèbre, on en déduit que :
$$\mu(A) = \lim_{n \to \infty} \mu(B_1 \cup \dots \cup B_n)$$
$$\leq \sum_{n \geq 1} \mu(B_n) \leq \sum_{n \geq 1} \mu(A_n),$$
et donc que
$$\mu(A) \leq \mu^*(A).$$

(iii) Soit $A_1, A_2, \dots$ une suite de parties de X, et montrons que
$$\mu^*\Big(\bigcup_{n \geq 1} A_n\Big) \leq \sum_{n \geq 1} \mu^*(A_n).$$
On peut supposer, sans perte de généralité, que $\sum_{n \geq 1} \mu^*(A_n) < +\infty$. Fixons $\varepsilon > 0$. Pour tout entier $n \geq 1$, il existe des parties $A_n^k \in \mathcal{R}$ telles que :
$$A_n \subset \bigcup_k A_n^k \quad \text{et} \quad \sum_k \mu(A_n^k) \leq \mu^*(A_n) + \frac{\varepsilon}{2^n}.$$
Alors, les A_n^k sont des éléments de $\mathcal{R}$

tels que $\bigcup_{n\geq 1} A_n \subset \bigcup_{n,k} A_n^k$ et qui vérifient :

$$\sum_{n,k} \mu(A_n^k) = \sum_{n\geq 1} \Big(\sum_k \mu(A_n^k)\Big) \leq \sum_{n\geq 1} \mu^*(A_n) + \frac{\varepsilon}{2^n}$$

$$= \sum_{n\geq 1} \mu^*(A_n) + \varepsilon \, .$$

On en déduit que

$\mu^*(\bigcup_{n\geq 1} A_n) \leq \sum_{n\geq 1} \mu^*(A_n) + \varepsilon$ et donc, en faisant tendre ε vers 0 :

$$\mu^*(\bigcup_{n\geq 1} A_n) \leq \sum_{n\geq 1} \mu^*(A_n).$$

(iv) Fixons $A \in \mathcal{R}_0$ et soient $E \subset A$, $F \subset A^c$. D'après (iii), on a :

$$\mu^*(E \cup F) \leq \mu^*(E) + \mu^*(F).$$

Montrons qu'il y a en fait égalité. A cet effet, on peut supposer sans perte de généralité que $\mu^*(E\cup F) < +\infty$. Soit $A_1, A_2, \dots$ une suite d'éléments de $\mathcal{R}_0$ telle que $E \cup F \subset \bigcup_{n\geq 1} A_n$. Posons $E_n = A \cap A_n$, $F_n = A_n \setminus A$.

On a : $E \subset \bigcup_{n\geq 1} E_n$, $F \subset \bigcup_{n\geq 1} F_n$ et, comme E_n, F_n appartiennent à $\mathcal{R}_0$, il vient :

$$\mu^*(E) + \mu^*(F) \leq \sum_{n\geq 1} \mu(E_n) + \sum_{n\geq 1} \mu(F_n)$$

$$= \sum_{n\geq 1} \mu(E_n) + \mu(F_n)$$

$$= \sum_{n\geq 1} \mu(E_n \cup F_n) = \sum_{n\geq 1} \mu(A_n).$$

On en déduit que :

$\mu^*(E) + \mu^*(F) \leq \mu^*(E \cup F)$, d'où (iv). ∎

1.3.3.4. Théorème (Hahn). Soient X un ensemble, $\mathcal{R}$ une algèbre de parties de X et $\mu : \mathcal{R} \to [0,+\infty]$ une mesure positive sur $\mathcal{R}$. On suppose qu'il existe une suite $A_1, A_2, \ldots$ d'éléments de $\mathcal{R}$ telle que :

$$X = \bigcup_{n \geq 1} A_n \quad \text{et} \quad \mu(A_n) < +\infty \text{ pour tout } n \geq 1.$$

Désignons par $\mathcal{B}$ la tribu engendrée par $\mathcal{R}$. Alors, on a :

(i) Il existe une unique mesure positive $\bar{\mu}$ sur $\mathcal{B}$ qui prolonge μ, i.e. telle que

$$\mu(A) = \bar{\mu}(A) \text{ quel que soit } A \in \mathcal{R};$$

(ii) La tribu $\mathcal{B}_{\bar{\mu}}$ des ensembles $\bar{\mu}$-mesurables de X coïncide avec la tribu des parties $A \subset X$ vérifiant la propriété suivante :

$$(\forall E \subset A)(\forall F \subset A^c), \text{ on a } \mu^*(E \cup F) = \mu^*(E) + \mu^*(F);$$

(iii) Une partie $N \subset X$ est $\bar{\mu}$-négligeable si et seulement si $\mu^*(N) = 0$, c'est à dire s'il existe, pour tout $\varepsilon > 0$, une suite $N_1, N_2, \ldots$ d'éléments de $\mathcal{R}$ telle que :

$$N \subset \bigcup_{n \geq 1} N_n \quad \text{et} \quad \sum_{n \geq 1} \mu(N_n) \leq \varepsilon.$$

Démonstration. Démontrons tout d'abord l'existence du prolongement $\bar{\mu}$ à $\mathcal{B}$. A cet effet, posons :

$$\mathcal{C} = \{ A \subset X \mid (\forall E \subset A)(\forall F \subset A^c), \mu^*(E \cup F) = \mu^*(E) + \mu^*(F) \}.$$

Lemme 1. $\mathcal{C}$ est une tribu sur X qui contient les éléments de $\mathcal{B}$ ainsi que les parties $N \subset X$ telles que $\mu^*(N) = 0$. En outre, l'application $\bar{\mu} : \mathcal{C} \to [0,+\infty]$ définie pour tout $A \in \mathcal{C}$ par

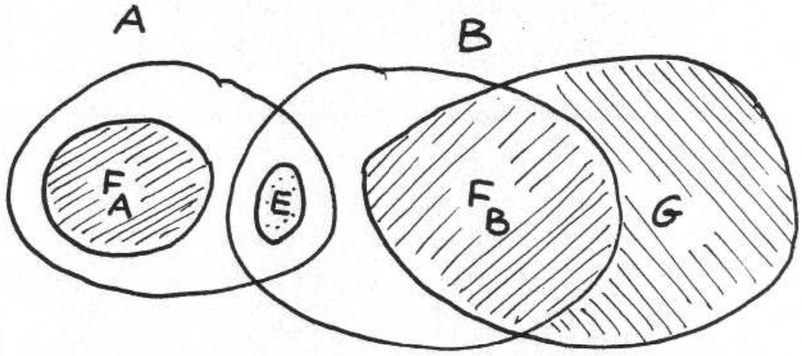
$$\bar{\mu}(A) = \mu^*(A)$$

est une mesure positive $\bar{\mu}$ sur $\mathcal{C}$ qui prolonge la mesure μ sur $\mathcal{R}$.

Démonstration. D'après la proposition 1.3.3.3 (iv), on sait déjà que $\mathcal{R} \subset \mathcal{C}$. Comme X est réunion dénombrable de parties $A_n \in \mathcal{R}$, il suffit, pour montrer que $\mathcal{T}$ est une tribu sur X, de montrer qu'elle est stable par réunions dénombrables et par passage au complémentaire.

Il est clair que, si $A \in \mathcal{T}$, alors $A^c \in \mathcal{T}$. Il nous suffit donc, pour prouver que $\mathcal{T}$ est une tribu, de montrer qu'elle est stable par réunions dénombrables. A cet effet, commençons par montrer que $\mathcal{T}$ est stable par intersections finies. Soient $A, B \in \mathcal{T}$ et considérons $E \subset A \cap B$, $F \subset (A \cap B)^c$. Posons :

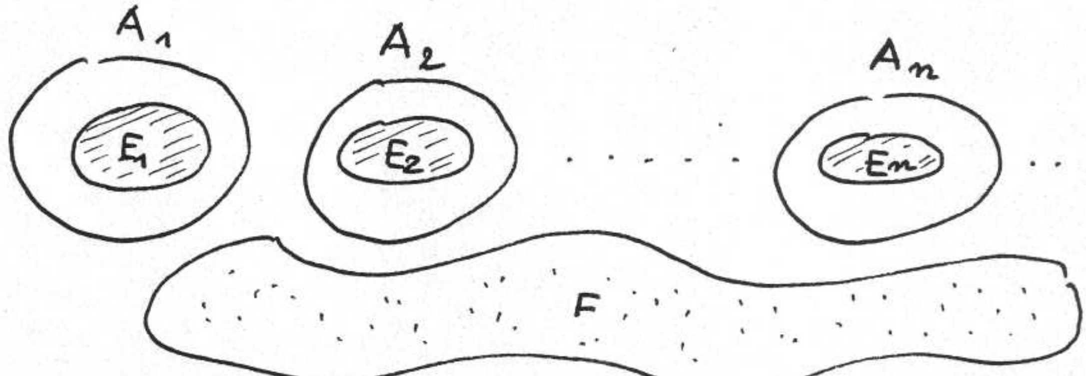
$$F_A = F \cap A, \quad F_B = F \cap B \quad \text{et} \quad G = F \setminus (A \cup B)$$



On a :

$$\mu^*(E \cup F) = \mu^*(E \cup F_A) + \mu^*(G \cup F_B) \qquad (\text{car } A \in \mathcal{T})$$
$$= \mu^*(F_A) + \mu^*(E) + \mu^*(F_B) + \mu^*(G) \quad (\text{car } B \in \mathcal{T})$$
$$\geqq \mu^*(E) + \mu^*(F) \qquad (\text{car } F = F_A \cup F_B \cup G)$$

Comme on a, par ailleurs, $\mu^*(E \cup F) \leqq \mu^*(E) + \mu^*(F)$, on en déduit que $\mu^*(E \cup F) = \mu^*(E) + \mu^*(F)$, ce qui prouve que $A \cap B \in \mathcal{T}$. Il s'ensuit que $\mathcal{T}$ est stable par intersections finies, et donc aussi par réunions finies (puisque $\mathcal{T}$ est stable par passage au complémentaire). On peut donc se limiter, pour prouver que $\mathcal{T}$ est stable par réunions dénombrables, à montrer que $\mathcal{T}$ est stable par réunions dénombrables disjointes. Soient donc $A_1, A_2, \dots$ une suite de parties $A_n \in \mathcal{T}$ qui sont deux à deux disjointes. Posons $A = \bigcup_{n \geq 1} A_n$ et soient $E \subset A$, $F \subset A^c$



On a, pour $n \geqq 1$ fixé, en posant $E_n = E \cap A_n$:

$$\begin{aligned}\mu^*(E \cup F) &\geqq \mu^*(E_1 \cup \ldots \cup E_n \cup F) \\ &= \mu^*(E_1) + \mu^*(E_2 \cup \ldots \cup E_n \cup F) \quad (\text{car } A_1 \in \mathcal{T}) \\ &= \mu^*(E_1) + \mu^*(E_2) + \mu^*(E_3 \cup \ldots \cup E_n \cup F) \quad (\text{car } A_2 \in \mathcal{T}) \\ &\vdots \\ &= \mu^*(E_1) + \mu^*(E_2) + \cdots + \mu^*(E_n) + \mu^*(F) \\ &\qquad (\text{car } A_n \in \mathcal{T}).\end{aligned}$$

En faisant tendre n vers $+\infty$, on en déduit que :

$$\mu^*(E \cup F) \geqq \sum_{n \geqq 1} \mu^*(E_n) + \mu^*(F) \geqq \mu^*(E) + \mu^*(F),$$

d'où $\mu^*(E \cup F) = \mu^*(E) + \mu^*(F)$, et donc $A \in \mathcal{T}$.

Nous avons donc ainsi démontré que $\mathcal{T}$ est une tribu, et elle contient $\mathcal{R}$ en vertu de 1.3.3.3 (iv). Comme $\mathcal{B}$ est engendrée par $\mathcal{R}$, on en déduit que $\mathcal{B} \subset \mathcal{T}$.

Considérons alors l'application $\bar{\mu} : \mathcal{T} \to [0, +\infty]$ définie par $\bar{\mu}(A) = \mu^*(A)$ pour $A \in \mathcal{T}$. Comme $\phi \in \mathcal{R}$, on a :

$$\bar{\mu}(\phi) = \mu(\phi) = 0.$$

Soit par ailleurs $A_1, A_2, \ldots$ une suite de parties de X appartenant à $\mathcal{T}$ et deux à deux disjointes. Le raisonnement ci-dessus, appliqué à $E = \bigcup_{n \geqq 1} A_n$ et $F = \phi$ donne alors $\mu^*(\bigcup_{n \geqq 1} A_n) \geqq \sum_{n \geqq 1} \mu^*(A_n)$ et, comme

$\mu^*(\bigcup_{n\geqq 1} A_n) \leqq \sum_{n\geqq 1} \mu^*(A_n)$ en vertu de 1.3.3.3 (iii),
on obtient finalement :

$$\mu^*(\bigcup_{n\geqq 1} A_n) = \sum_{n\geqq 1} \mu^*(A_n).$$

Ceci démontre que $\bar{\mu}$ est bien une mesure positive sur $\mathcal{C}$. Pour achever la démonstration du lemme 1, il reste à montrer que toute partie $N\subset X$ telle que $\mu^*(N)=0$ appartient à $\mathcal{C}$. Soit N une telle partie. Pour $E\subset N$ et $F\subset N^c$, on a :

$$\mu^*(F) \leqq \mu^*(E\cup F) \leqq \mu^*(E)+\mu^*(F)$$
$$\leqq \mu^*(N)+\mu^*(F) = \mu^*(F),$$

d'où $\mu^*(E\cup F)=\mu^*(F)=\mu^*(F)+\mu^*(E)$, ce qui prouve que $N\in\mathcal{C}$. ■

<u>Démonstration de (i)</u>. D'après le lemme 1, la restriction de $\bar{\mu}$ à $\mathcal{B}$ est une mesure positive sur $\mathcal{B}$ qui coïncide avec μ sur $\mathcal{R}$. Soit ν une autre mesure positive sur $\mathcal{B}$ qui coïncide avec μ sur $\mathcal{R}$, et montrons que $\nu=\bar{\mu}$. Par hypothèse, il existe une suite $A_1, A_2, \dots$ d'éléments de $\mathcal{R}$ telle que

$$X=\bigcup_{n\geqq 1} A_n \quad \text{et} \quad \mu(A_n)<+\infty \text{ pour tout } n\geqq 1.$$

Comme $\mathcal{R}$ est une algèbre, on peut sans perte de généralité supposer que les A_n sont deux à deux disjoints. Pour tout $A\in\mathcal{B}$, on a :

$$A=\bigcup_{n\geqq 1}(A\cap A_n)$$

où les $A\cap A_n\in\mathcal{B}$ sont deux à deux disjoints, d'où

$$\bar{\mu}(A)=\sum_{n\geqq 1}\bar{\mu}(A\cap A_n) \quad \text{et}$$

$$\nu(A)=\sum_{n\geqq 1}\nu(A\cap A_n).$$

Pour montrer que $\nu=\bar{\mu}$, il suffit de montrer que

$$\nu(A\cap A_n)=\bar{\mu}(A\cap A_n) \quad \text{pour tout } n\geqq 1 \text{ et tout } A\in\mathcal{B}.$$

On peut ainsi, pour montrer que $\nu(A)=\bar{\mu}(A)$ pour tout $A\in\mathcal{B}$, supposer que A est contenu dans $B\in\mathcal{R}$ avec $\mu(B)<+\infty$. Soit $A_1, A_2, \dots$ une suite de parties de X appartenant à $\mathcal{R}$ telle que

$$A\subset\bigcup_{n\geqq 1} A_n.$$

On a :

$$\nu(A)\leqq\nu\Big(\bigcup_{n\geqq 1} A_n\Big)\leqq\sum_{n\geqq 1}\nu(A_n)=\sum_{n\geqq 1}\mu(A_n),$$

d'où $\nu(A)\leqq\mu^*(A)=\bar{\mu}(A)$. De même, on a :

$$\nu(B-A)\leqq\bar{\mu}(B-A).$$

On en déduit, par addition de ces deux inégalités, que

$$\nu(A)+\nu(B-A)\leqq\bar{\mu}(A)+\bar{\mu}(B-A).$$

Or, le premier membre est égal à $\nu(B)=\mu(B)$, et le second membre aussi, de sorte que l'on a nécessairement

$$\nu(A)=\bar{\mu}(A)$$

et donc $\nu=\bar{\mu}$. Ceci achève la démonstration de (i).

<u>Démonstration de (iii)</u>. Si $\mu^*(N)=0$, il existe pour tout $\varepsilon>0$ une suite $N_1^\varepsilon, N_2^\varepsilon, \dots$ d'éléments de $\mathcal{R}$ telle que :

$$N\subset\bigcup_{n\geqq 1} N_n^\varepsilon \quad \text{et} \quad \sum_{n\geqq 1}\mu(N_n^\varepsilon)\leqq\varepsilon.$$

Posons $B^\varepsilon=\bigcup_{n\geqq 1} N_n^\varepsilon\in\mathcal{B}$. On a :

$$N\subset B^\varepsilon \quad \text{et} \quad \bar{\mu}(B^\varepsilon)=\bar{\mu}\Big(\bigcup_{n\geqq 1} N_n^\varepsilon\Big)\leqq\sum_{n\geqq 1}\mu(N_n^\varepsilon)\leqq\varepsilon.$$

Faisons $\varepsilon = \frac{1}{n}$. On construit ainsi une suite $B_1, B_2, \dots$ d'éléments de $\mathcal{B}$ telle que :

$$N \subset B_n \text{ pour tout } n \geqq 1 \text{ et } \bar{\mu}(B_n) \leqq \frac{1}{n}.$$

Quitte à remplacer cette suite par $B_1, B_1 \cap B_2, B_1 \cap B_2 \cap B_3, \dots$ on peut supposer que $B_1 \supset B_2 \supset \dots$. Comme $\bar{\mu}(B_1) < +\infty$, on a :

$$\bar{\mu}\left(\bigcap_{n \geqq 1} B_n\right) = \lim_{n \to \infty} \bar{\mu}(B_n) = 0 .$$

Posons $B = \bigcap_{n \geqq 1} B_n \in \mathcal{B}$; on a :

$$N \subset B \quad \text{et} \quad \bar{\mu}(B) = 0,$$

de sorte que N est $\bar{\mu}$-négligeable. Supposons inversement que N soit $\bar{\mu}$-négligeable. Alors, il existe $B \in \mathcal{B}$ tel que

$$N \subset B \quad \text{et} \quad \bar{\mu}(B) = 0,$$

d'où :

$$\mu^*(N) \leqq \mu^*(B) = \bar{\mu}(B) = 0.$$

Il s'ensuit que $\mu^*(N) = 0$, ce qui prouve (iii).

<u>Démonstration de (ii)</u>. D'après ce qui précède, $\mathcal{B}_{\bar{\mu}}$ est la tribu des parties de X de la forme $B \cup N$ où $B \in \mathcal{B}$ et où $\mu^*(N) = 0$. D'après le lemme 1, on a :

$$\mathcal{B}_{\bar{\mu}} \subset \mathcal{T}$$

Montrons l'inclusion inverse. Soit $A \in \mathcal{T}$ et montrons que $A \in \mathcal{B}_{\bar{\mu}}$. En utilisant le fait que X est réunion d'une suite croissante $A_1 \subset A_2 \subset \dots$ d'éléments de $\mathcal{R}$ telle que $\mu(A_n) < +\infty$ pour tout $n \geqq 1$, on se ramène à prouver que, si $A \in \mathcal{T}$ est contenu dans $B \in \mathcal{B}$ avec $\bar{\mu}(B) < +\infty$, alors $A \in \mathcal{B}_{\bar{\mu}}$.

Considérons alors $C = B - A$. En utilisant la définition de $\mu^*(C)$, on montre comme dans la preuve de (iii) qu'il existe une suite décroissante

$$E_1 \supset E_2 \supset \dots \supset E_n \supset E_{n+1} \supset \dots .$$

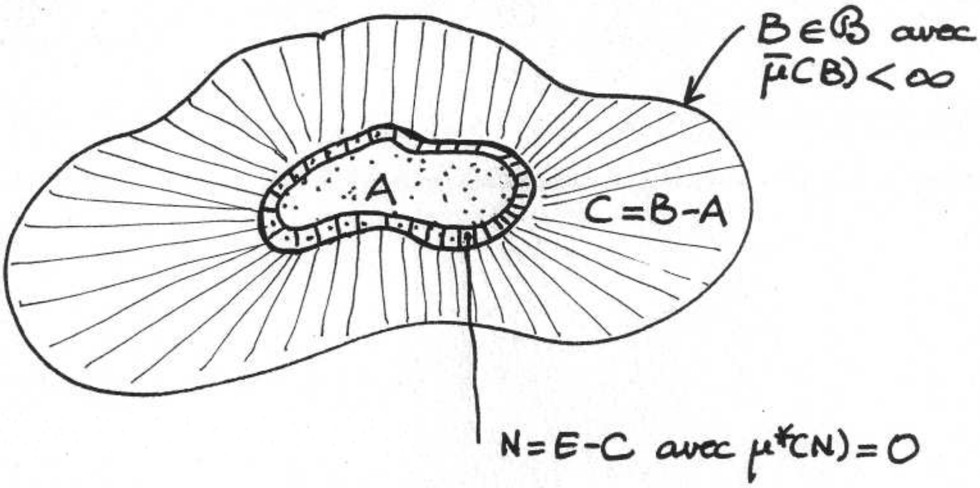
d'éléments de $\mathcal{B}$ vérifiant :

$$C \subset E_n \text{ pour tout } n \geqq 1, \ E_n \subset B, \text{ et}$$
$$\bar{\mu}(C) \leqq \mu(E_n) \leqq \bar{\mu}(C) + \frac{1}{n} \text{ pour tout } n \geqq 1.$$

Posons $E = \bigcap_{n \geqq 1} E_n \in \mathcal{B}$. On a $C \subset E$, d'où

$$\bar{\mu}(C) \leqq \bar{\mu}(E) = \lim_{n \to \infty} \mu(E_n) \leqq \bar{\mu}(C),$$

et donc $\bar{\mu}(C) = \bar{\mu}(E)$. Il s'ensuit que $E - C = N$ est $\bar{\mu}$-négligeable.



On a alors : $A = (B - E) \cup N$ avec $B - E \in \mathcal{B}$ et $\mu^*(N) = 0$, ce qui prouve que $A \in \mathcal{B}_{\bar{\mu}}$ et achève la démonstration du théorème de Hahn. ∎

1.3.3.5 - Mesure de Lebesgue sur $\mathbb{R}^p$ - Il résulte de la construction de la mesure de Lebesgue λ sur l'algèbre des parties qui sont réunions finies de pavés et du théorème de Hahn, l'existence d'une unique mesure positive λ sur la tribu des Boréliens de $\mathbb{R}^p$ qui vérifie :

$$\lambda\left(\prod_{i=1}^{p} [a_i, b_i[\right) = \prod_{i=1}^{p} (b_i - a_i).$$

Cette mesure λ s'étend en une mesure positive sur la tribu $\mathcal{B}_\lambda$ des parties de $\mathbb{R}^p$ qui sont réunion d'un Borélien et d'un ensemble λ-négligeable. Pour tout $A \in \mathcal{B}_\lambda$, on a :

$$\lambda(A) = \inf\left\{ \sum_{n\geq 1} \lambda(P_n) \;\middle|\; P_n \text{ est un pavé pour tout } n \text{ et } A \subset \bigcup_{n\geq 1} P_n \right\}.$$

En outre, un sous-ensemble $N \subset \mathbb{R}^p$ est λ-négligeable s'il existe, pour tout $\varepsilon > 0$, une suite $P_1, P_2, \dots$ de pavés telle que

$$N \subset \bigcup_{n\geq 1} P_n \quad \text{et} \quad \sum_{n\geq 1} \lambda(P_n) \leq \varepsilon.$$

On dit que la mesure λ sur la tribu $\mathcal{B}_\lambda$ des parties λ-mesurables est la <u>mesure de Lebesgue</u> de $\mathbb{R}^p$. D'après 1.2.4, cette mesure λ induit une mesure positive sur toute partie λ-mesurable A de $\mathbb{R}^p$, qui est appelée la <u>mesure de Lebesgue de A</u>, et que l'on note λ_A.

1.4- ENSEMBLES LEBESGUE-MESURABLES

<u>1.4.1</u>. Sur le segment fermé borné $I = [0,1]$ de $\mathbb{R}$, considérons la relation d'équivalence :

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

D'après l'axiome du choix, il existe un sous-ensemble $A \subset [0,1]$ qui rencontre chaque classe d'équivalence de la relation ci-dessus en un point et un seul.
Pour tout $r \in \mathbb{Q} \cap [-1,1]$, posons :

$$A_r = A + r = \{a + r \mid a \in A\}.$$

On définit ainsi des sous-ensembles $A_r \subset [-1,2]$, en nombre au plus dénombrable. Par ailleurs, ces A_r sont deux à deux disjoints. En effet, si $A_r \cap A_s \neq \emptyset$, il existe $a \in A$ et $b \in A$ tels que $a + r = b + s$, d'où $a - b = s - r \in \mathbb{Q}$ et donc $a \sim b$. Comme A rencontre chaque classe d'équivalence en un seul point, on en déduit que $a = b$ et donc que $r = s$.

1.4.2 - Proposition. L'ensemble A n'est pas Lebesgue-mesurable. En particulier, il n'est ni Borélien ni négligeable.

Démonstration. Notons tout d'abord que, si E est une réunion finie d'intervalles bornés, alors $E+r$ est une réunion finie d'intervalles bornés et on a :

$$\lambda(E+r) = \lambda(E).$$

Il s'ensuit immédiatement que $\lambda^*(E+r) = \lambda^*(E)$ pour tout $E \subset \mathbb{R}$.

Supposons par l'absurde que A soit Lebesgue mesurable. En utilisant la caractérisation des ensembles Lebesgue-mesurables donnée par la condition (ii) du théorème de Hahn, on montre aisément que $A_r = A + r$ est Lebesgue mesurable et vérifie $\lambda(A_r) = \lambda(A)$ pour tout $r \in \mathbb{Q}$. On a donc :

$$\lambda\Big(\bigcup_{r\in[-1,1]\cap\mathbb{Q}} A_r\Big) = \sum_{r\in[-1,1]\cap\mathbb{Q}} \lambda(A_r) = \sum_{r\in[-1,1]\cap\mathbb{Q}} \lambda(A)$$

Mais on a, par ailleurs :

$$[0,1] \subset \bigcup_{r\in[-1,1]\cap\mathbb{Q}} A_r \subset [-1,2],$$

d'où :

$$1 = \lambda([0,1]) \leqq \sum_{r\in[-1,1]\cap\mathbb{Q}} \lambda(A) \leqq \lambda([-1,2]) = 3$$

La première inégalité implique que $\lambda(A) > 0$. Mais alors, la seconde inégalité se réduit à

$$+\infty \leqq 3,$$

ce qui est absurde. Comme les Boréliens et les négligeables sont Lebesgue mesurables, la proposition est démontrée. ■

1.4.3. La tribu des Boréliens de $\mathbb{R}$ et celle des sous-ensembles Lebesgue-mesurables de $\mathbb{R}$ ne sont donc pas égales à $\mathcal{P}(\mathbb{R})$. Pour construire un ensemble non Lebesgue-mesurable, nous avons utilisé l'axiome du choix. On peut montrer que l'on ne peut construire un ensemble non mesurable de $\mathbb{R}$ sans faire appel à l'axiome du choix. En d'autres termes, sauf à le faire exprès, on ne manipulera dans la pratique aucun ensemble non mesurable au sens de Lebesgue. On peut donc considérer que la mesurabilité d'un sous-ensemble de $\mathbb{R}$ est toujours - sauf cas réellement exceptionnel - un fait acquis, même si la vérification de ce fait peut se révéler scholastiquement ardue pour le débutant.

Thierry Fack
octobre 2010

2 - FONCTIONS MESURABLES

2 - FONCTIONS MESURABLES

2.1 - DÉFINITIONS - EXEMPLES

2.1.1 - DÉFINITIONS

2.1.1.1 - Fonction mesurable - Soient $(X, \mathcal{M})$ et $(Y, \mathcal{P})$ deux espaces mesurables. On dit que $f : X \to Y$ est mesurable si, pour toute $A \in \mathcal{P}$, l'image réciproque

$$f^{-1}(A) = \{x \in X \mid f(x) \in A\}$$

appartient à $\mathcal{M}$.

Un espace topologique Y sera, sauf mention contraire, muni de la tribu $\mathcal{P} = \mathcal{B}_Y$ des Boréliens de Y. Ainsi, si $(X, \mathcal{M})$ est un espace mesurable et Y un espace topologique, on dit que $f : X \to Y$ est mesurable si, pour tout Borélien A de Y, on a :

$$f^{-1}(A) \in \mathcal{M}.$$

2.1.1.2 - Fonction μ-mesurable. Soient (X, μ) un espace mesuré et Y un espace topologique. On dit que $f : X \to Y$ est μ-mesurable si, pour tout Borélien A de Y, l'image réciproque $f^{-1}(A)$ est un sous-ensemble μ-mesurable de X.

On dira que deux fonctions $f, g : X \to Y$ sont égales μ-presque partout s'il existe un sous-ensemble μ-négligeable $N \subset X$ tel que $f(x) = g(x)$ quel que soit $x \notin N$. On a :

2.1.1.3 - Proposition - Soient (X, μ) un espace mesuré et Y un espace topologique. Alors, toute application $f : X \to Y$ qui est égale presque partout à une application μ-mesurable $g : X \to Y$ est elle-même μ-mesurable.

Démonstration. Soient $g : X \to Y$ une application μ-mesurable et $f : X \to Y$ qui coïncide avec g hors d'un sous-ensemble μ-négligeable $N \subset X$. Pour tout Borélien A de Y, on a :

$$f^{-1}(A) = [N^c \cap g^{-1}(A)] \cup [N \cap f^{-1}(A)]$$

où $N^c \cap g^{-1}(A)$ est μ-mesurable (car N^c et $g^{-1}(A)$ sont μ-mesurables), et où $N \cap f^{-1}(A)$ est μ-négligeable (car il est contenu dans N, qui est μ-négligeable). Il s'ensuit que $f^{-1}(A)$ est μ-mesurable, ce qui prouve que f est μ-mesurable. ■

2.1.1.4 - Fonctions numériques. Désignons par $\overline{\mathbb{R}}$ la droite numérique achevée obtenue en rajoutant les points $\pm\infty$ à $\mathbb{R}$. On munit $\overline{\mathbb{R}}$ de la relation d'ordre naturelle pour laquelle $-\infty$ est le plus petit élément et $+\infty$ le plus grand élément. On observera que l'addition de deux nombres est définie de manière évidente dans $\overline{\mathbb{R}}$, sauf pour le couple $(-\infty, +\infty)$. De même, le produit de deux nombres de $\overline{\mathbb{R}}$ est défini, sauf pour les couples $(0, +\infty)$ et $(0, -\infty)$.

On munit $\overline{\mathbb{R}}$ de la distance définie par :

$$d(x, y) = |\operatorname{Arctan}(x) - \operatorname{Arctan}(y)|,$$

où $\operatorname{Arctan}(\pm\infty) = \pm\frac{\pi}{2}$. Comme espace topologique, $\overline{\mathbb{R}}$ est homéomorphe à $[-\frac{\pi}{2}, \frac{\pi}{2}]$, muni de sa topologie canonique. Il s'ensuit que la tribu $\mathcal{B}_{\overline{\mathbb{R}}}$ des Boréliens de $\overline{\mathbb{R}}$ est engendrée par les intervalles $[a, +\infty]$ (resp. $]a, +\infty]$, $[-\infty, a]$, $[-\infty, a[$) où a varie dans $\mathbb{R}$.

Dans la suite, nous aurons souvent à considérer des fonctions à valeurs dans $\overline{\mathbb{R}}$, c'est à dire susceptibles de prendre les valeurs $\pm\infty$.

Définition. On appelle fonction numérique sur un ensemble X une application $f : X \to \overline{\mathbb{R}}$.

Les fonctions $f : X \to [0, +\infty]$ seront appelées, par extension, fonctions numériques positives. Il y a donc un sens à parler de fonctions numériques mesurables, ou μ-mesurables.

2.1.2 - CARACTÉRISATION DES FONCTIONS MESURABLES

2.1.2.1 - Proposition. Soient $(X, \mathcal{M})$, $(Y, \mathcal{P})$ deux espaces mesurables, et $\mathcal{P}_0$ une classe de parties de Y engendrant $\mathcal{P}$ comme tribu. Alors, pour toute application $f: X \to Y$, les conditions suivantes sont équivalentes :

(i) f est mesurable ;

(ii) $\forall A \in \mathcal{P}_0$, $f^{-1}(A) \in \mathcal{M}$.

Démonstration : (i) $\Rightarrow$ (ii) est évident. Montrons que (ii) $\Rightarrow$ (i). A cet effet, posons :

$$\mathcal{Q} = \{ A \subset Y \mid f^{-1}(A) \in \mathcal{M} \}.$$

Des relations $f^{-1}(A^c) = f^{-1}(A)^c$ et $f^{-1}(\bigcup_n A_n) = \bigcup_n f^{-1}(A_n)$, vraies quelles que soient les parties A et A_n de Y, on déduit immédiatement que $\mathcal{Q}$ est une tribu comme $\mathcal{M}$. Cette tribu contient $\mathcal{P}_0$ par hypothèse, donc elle contient la tribu engendrée par $\mathcal{P}_0$, d'où

$$\mathcal{P} \subset \mathcal{Q}.$$

Mais alors, pour tout $A \in \mathcal{P}$, $f^{-1}(A) \in \mathcal{M}$, ce qui démontre que f est mesurable. ■

2.1.2.2 - Corollaire. Soient (X, μ) un espace mesuré et $f: X \to \overline{\mathbb{R}}$ une fonction numérique. Les conditions suivantes sont équivalentes :

(i) f est μ-mesurable ;

(ii) Pour tout $a \in \mathbb{R}$, l'ensemble $E_a = \{ x \in X \mid f(x) \leqq a \}$ est μ-mesurable ;

(iii) Pour tout $a \in \mathbb{R}$, l'ensemble $E'_a = \{ x \in X \mid f(x) \geqq a \}$ est μ-mesurable ;

(iv) Pour tout $a \in \mathbb{R}$, l'ensemble $F_a = \{ x \in X \mid f(x) < a \}$ est μ-mesurable ;

(v) Pour tout $a \in \mathbb{R}$, l'ensemble $F'_a = \{ x \in X \mid f(x) > a \}$ est μ-mesurable.

Démonstration. C'est une conséquence immédiate de la proposition 2.1.2.1, compte tenu des remarques sur l'engendrement de la tribu des Boréliens de $\overline{\mathbb{R}}$ faites en 2.1.1.4. ■

2.1.3- EXEMPLES D'APPLICATIONS MESURABLES

2.1.3.1- Soient X et Y deux espaces topologiques. On dit qu'une application $f : X \to Y$ est Borélienne si elle est mesurable lorsqu'on munit X (resp. Y) de la tribu des Boréliens $\mathcal{B}_X$ (resp. $\mathcal{B}_Y$).

Si X est un espace topologique muni d'une mesure Borélienne μ, toute application Borélienne $f : X \to \overline{\mathbb{R}}$ est μ-mesurable.

2.1.3.2- Proposition. Soient X et Y deux espaces topologiques. Toute application continue $f : X \to Y$ est Borélienne.

Démonstration. Comme f est continue, on sait que pour tout ouvert U de Y, $f^{-1}(U)$ est un ouvert de X, donc un Borélien de X. Comme $\mathcal{B}_Y$ est engendré par les ouverts de Y, l'application f est Borélienne en vertu de la proposition 2.1.2.1. ■

2.1.3.3- Corollaire. Toute application $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ qui est presque partout égale à une fonction continue est Lebesgue-mesurable.

Démonstration. D'après la proposition 2.1.1.3, il suffit de montrer que toute application continue $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ est Lebesgue-mesurable. Or f est Borélienne en vertu de la proposition 2.1.3.2 et, comme tout Borélien de $\mathbb{R}^n$ est a fortiori Lebesgue-mesurable, f est Lebesgue-mesurable. ■

Ainsi, toute fonction $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ obtenue à partir d'une fonction continue $\tilde{f} : \mathbb{R}^p \to \overline{\mathbb{R}}$ en modifiant arbitrairement ses valeurs sur un ensemble Lebesgue-négligeable (par exemple sur $\mathbb{Q}^p \subset \mathbb{R}^p$) est Lebesgue-mesurable.

2.1.3.4. Fonctions caractéristiques. Soient (X, μ) un espace mesuré et $A \subset X$. La fonction caractéristique de A est par définition la fonction $\mathbb{1}_A : X \to \mathbb{R}$ définie par :

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

Pour tout $a \in \mathbb{R}$, on a :

$$E_a = \{x \in X \mid \mathbb{1}_A(x) \leqq a\} = \begin{cases} X & \text{si } a \geqq 1 \\ A^c & \text{si } 0 \leqq a < 1 \\ \phi & \text{si } a < 0 \end{cases}$$

de sorte que l'on a :

$$\mathbb{1}_A \text{ est } \mu\text{-mesurable} \iff A \text{ est } \mu\text{-mesurable.}$$

2.1.3.5 - Remarque. Considérons l'espace $\mathbb{R}$, muni de la mesure de Lebesgue λ. On a vu qu'il existe des parties $A \subset \mathbb{R}$ qui ne sont pas Lebesgue-mesurables. D'après 2.1.3.4, il existe donc des fonctions numériques positives $f : \mathbb{R} \to [0, +\infty]$ qui ne sont pas Lebesgue-mesurables. On montre cependant que l'on ne peut pas construire de fonction non mesurable $f : \mathbb{R} \to \overline{\mathbb{R}}$ sans faire appel à l'axiome du choix. En d'autres termes, sauf à le faire exprès en utilisant l'axiome du choix, toutes les fonctions $f : \mathbb{R} \to \overline{\mathbb{R}}$ que nous construirons en pratique seront, automatiquement Lebesgue mesurables. Le fait d'être Lebesgue-mesurable est donc une restriction très faible que l'on impose aux fonctions $f : \mathbb{R} \to \overline{\mathbb{R}}$.

2.2. THÉORÈMES DE STABILITÉ

2.2.1 - Composition des fonctions mesurables.

2.2.1.1- Proposition. Soient $(X,\mathcal{M})$, $(Y,\mathcal{P})$ et $(Z,\mathcal{Q})$ trois espaces mesurables. Si $f: X \to Y$ et $g: Y \to Z$ sont des applications mesurables, alors $g\circ f: X \to Z$ est mesurable.

Démonstration. Pour tout $A \in \mathcal{Q}$, $g^{-1}(A) \in \mathcal{P}$ car g est mesurable, et donc $f^{-1}(g^{-1}(A)) \in \mathcal{M}$ puisque f est également mesurable. Mais alors :

$$(g\circ f)^{-1}(A) = f^{-1}(g^{-1}(A)) \in \mathcal{M},$$

ce qui prouve que $g\circ f$ est mesurable. ■

2.2.1.2 - Corollaire. Soient (X,μ) un espace mesuré et $f: X \to \overline{\mathbb{R}}$ une fonction numérique μ-mesurable. Alors, pour toute fonction continue $\alpha : \overline{\mathbb{R}} \to \overline{\mathbb{R}}$, la fonction numérique $\alpha\circ f : X \to \overline{\mathbb{R}}$ est μ-mesurable.

Démonstration. Résulte immédiatement des propositions 2.1.3.2 et 2.2.1.1. ■

Par exemple, si $f: X \to [0,+\infty]$ est une fonction numérique positive μ-mesurable, l'application $\ln f : X \to \overline{\mathbb{R}}$ est μ-mesurable, puisque l'application $\ln : [0,+\infty] \to \overline{\mathbb{R}}$ définie par :

$$\ln(x) = \begin{cases} -\infty & \text{si } x = 0 \\ \ln(x) & \text{si } 0 < x < \infty \\ +\infty & \text{si } x = +\infty \end{cases}$$

est continue.

2.2.1.3 - Proposition - Soient (X,μ) un espace mesuré et $f: X \to \overline{\mathbb{R}}$ une fonction numérique μ-mesurable. Alors, l'application $|f| : X \to [0,+\infty]$ est μ-mesurable.

Démonstration. Considérons la fonction $\alpha : \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ définie par :

$$\alpha(x) = \begin{cases} +\infty & \text{si } x = +\infty \\ |x| & \text{si } -\infty < x < \infty \\ +\infty & \text{si } x = -\infty . \end{cases}$$

Cette fonction est continue et, puisque $|f| = \alpha\circ f$, la conclusion résulte du corollaire 2.2.1.2. ∎

On prendra garde au fait que $|f| : X \to [0,+\infty]$ peut être μ-mesurable sans que f le soit. Par exemple, si $A \subset [0,1]$ est un sous-ensemble non Lebesgue mesurable de $\mathbb{R}$, la fonction $f : \mathbb{R} \to \mathbb{R}$ définie par :

$$f(x) = \begin{cases} 1 & \text{si } x \in A \\ -1 & \text{si } x \in [0,1] \setminus A \\ 0 & \text{si } x \notin [0,1] \end{cases}$$

n'est pas Lebesgue-mesurable car $f^{-1}(\{1\}) = A$ n'est pas Lebesgue-mesurable. Pourtant,

$$|f| = \mathbb{1}_{[0,1]}$$

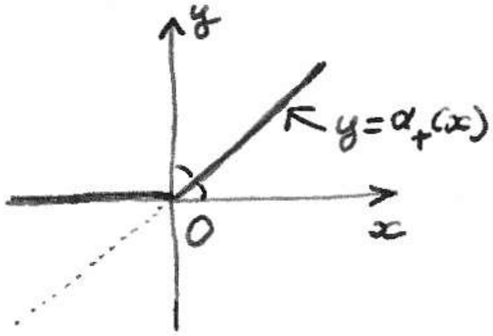
est Lebesgue-mesurable.

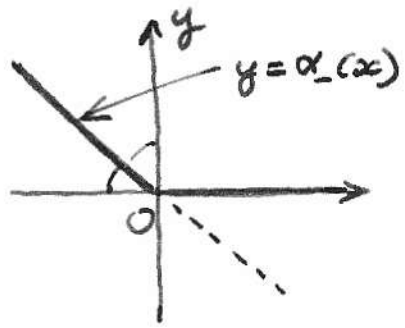
2.2.1.4- Décomposition canonique d'une fonction réelle.

Soient X un ensemble et $f : X \to \mathbb{R}$ une fonction réelle définie sur X. Pour étudier f, il est parfois utile de la décomposer en différence de deux fonctions positives. A cet effet, on pose

$$f^+ = \operatorname{Sup}(f, 0)$$
$$f^- = \operatorname{Sup}(-f, 0)$$

Posons $\alpha_+(x) = \operatorname{Sup}(x,0)$, $\alpha_-(x) = \operatorname{Sup}(-x,0)$; on définit ainsi deux fonctions continues de $\mathbb{R}$ dans $[0,+\infty[$, dont les graphes sont figurés ci-dessous :





On a : $f^+ = \alpha_+ \circ f$, $f^- = \alpha_- \circ f$, et on tire de la relation : $\alpha_+(x) - \alpha_-(x) = x$ pour tout $x \in \mathbb{R}$, la décomposition

$$f = f^+ - f^-$$

de f en différence des deux fonctions positives f^+ et f^-. On dit que c'est la <u>décomposition canonique</u> de f en différence de deux fonctions positives. De la relation :

$$|x| = \alpha_+(x) + \alpha_-(x) \quad \text{pour tout } x \in \mathbb{R},$$

on déduit que :

$$|f| = f^+ + f^-,$$

d'où $f^+ \leqq |f|$ et $f^- \leqq |f|$.

Si maintenant (X, μ) est un espace mesuré et si $f : X \to \mathbb{R}$ est μ-mesurable, les relations

$$f^\pm = \alpha_\pm \circ f$$

montrent, en utilisant le corollaire 2.2.1.2, que f^+ et f^- sont μ-mesurables.

2.2.2- Mesurabilité de la somme et du produit

Pour étudier la mesurabilité de la somme et du produit de deux fonctions μ-mesurables, on utilisera le résultat suivant, qui a son intérêt propre :

<u>2.2.2.1- Proposition.</u> <u>Soient (X,μ) un espace mesuré et $f, g : X \to \mathbb{R}$ deux fonctions réelles. Les conditions suivantes sont équivalentes :</u>

<u>(i) La fonction $h : X \to \mathbb{R}^2$ définie par :</u>

$$h(x) = (f(x), g(x))$$

<u>est μ-mesurable ;</u>

<u>(ii) Les fonctions f et g sont μ-mesurables.</u>

<u>Démonstration.</u> (i) $\Rightarrow$ (ii). Soit $p : \mathbb{R}^2 \to \mathbb{R}$ l'application définie par $p(x,y) = x$. On a :

$$f = p \circ h$$

et, comme p est continue, f est μ-mesurable si h est μ-mesurable en vertu des propositions 2.1.3.2 et

2.2.1.1. On montre de même que, si h est μ-mesurable, alors g est μ-mesurable, ce qui prouve l'implication (i) $\Rightarrow$ (ii)

(ii) $\Rightarrow$ (i) - Supposons que f et g soient μ-mesurables et montrons que h est μ-mesurable. Comme la tribu des Boréliens de $\mathbb{R}^2$ est engendrée par les rectangles de la forme

$$[a_1, a_2] \times [b_1, b_2] \qquad (a_1 < a_2,\ b_1 < b_2),$$

il suffit de montrer que l'ensemble

$$E = \{x \in X \mid h(x) \in [a_1, a_2] \times [b_1, b_2]\}$$

est μ-mesurable quels que soient les réels a_1, a_2, b_1, b_2 vérifiant $a_1 < a_2$ et $b_1 < b_2$. Or on a :

$$\begin{aligned} E &= \{x \in X \mid f(x) \in [a_1, a_2] \text{ et } g(x) \in [b_1, b_2]\} \\ &= f^{-1}([a_1, a_2]) \cap g^{-1}([b_1, b_2]) \end{aligned}$$

et, puisque f et g sont μ-mesurables, E est μ-mesurable comme intersection de deux sous-ensembles μ-mesurables de X. La proposition est donc démontrée. ∎

2.2.2.2. Proposition - Soient (X, μ) un espace mesuré et $f, g : X \to \mathbb{R}$ deux fonctions μ-mesurables. Alors, les applications $f+g$ et fg sont μ-mesurables.

Démonstration. L'application $f+g$ (resp. fg) est composée de l'application $x \mapsto (f(x), g(x))$, qui est μ-mesurable en vertu de la proposition 2.2.2.1, et de l'application continue $(x,y) \mapsto x+y$ (resp. $(x,y) \mapsto xy$). Il résulte alors des propositions 2.1.3.2 et 2.2.1.1 que $f+g$ (resp. fg) est μ-mesurable. ∎

Cette proposition s'étend immédiatement au cas où f, g sont des fonctions numériques positives, à condition de convenir que $0 \times (+\infty) = 0$.

2.2.2.3 - Fonctions simples. Soit (X,μ) un espace mesuré. On dit qu'une fonction $f : X \to \mathbb{R}$ est étagée si son image est un sous-ensemble fini de $\mathbb{R}$. Une fonction étagée μ-mesurable est dite simple. Toute fonction étagée $f : X \to \mathbb{R}$ se décompose canoniquement en une combinaison linéaire de fonctions caractéristiques. Soient en effet $a_1, a_2, \dots, a_n$ les valeurs distinctes prises par la fonction f, et posons :

$$A_i = \{x \in X \mid f(x) = a_i\} \qquad (1 \leq i \leq n).$$

On vérifie immédiatement que l'on a :

$$f = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}.$$

Si f est μ-mesurable, chaque $A_i = f^{-1}(\{a_i\})$ est μ-mesurable. Inversement, si chaque A_i est μ-mesurable, f est μ-mesurable en vertu de 2.1.3.4 et de 2.2.2.2. Ainsi, f est simple si et seulement si chaque A_i est μ-mesurable. Une fonction simple est donc une combinaison linéaire de fonctions caractéristiques de sous-ensembles μ-mesurables. On notera que, si f est positive, les coefficients a_i sont positifs.

2.2.3 - Limites simples de fonctions mesurables

2.2.3.1 - Limite simple et limite presque partout. Soient (X,μ) un espace mesuré et $f_n : X \to \overline{\mathbb{R}}$ une suite de fonctions numériques $(n = 1, 2, \dots)$. On dit que la suite $(f_n)_{n \geq 1}$ converge simplement vers la fonction $f : X \to \overline{\mathbb{R}}$ si, pour tout $x \in X$, $\lim_{n \to \infty} f_n(x) = f(x)$.

On dit que la suite $(f_n)_{n \geq 1}$ converge simplement μ-presque partout vers la fonction $f : X \to \overline{\mathbb{R}}$

s'il existe un sous-ensemble μ-négligeable $N \subset X$ tel que l'on ait, pour tout $x \notin N$, $f(x) = \lim_{n\to\infty} f_n(x)$.

Dans ce cas, $f(x)$ n'est déterminé par la suite des $f_n(x)$ que pour $x \notin N$. En d'autres termes, f n'est déterminée par les f_n qu'à un ensemble μ-négligeable près. Ceci n'est nullement gênant pour l'étude de la μ-mesurabilité de f, puisqu'une fonction égale μ-presque partout à une fonction μ-mesurable est μ-mesurable en vertu de la proposition 2.1.1.3.

Pour examiner la μ-mesurabilité d'une limite simple μ-presque partout d'une suite de fonctions μ-mesurables, on utilisera le résultat suivant :

<u>2.2.3.2 - Proposition</u>. <u>Soit (X,μ) un espace mesuré. Pour toute suite $f_n : X \to \overline{\mathbb{R}}$ de fonctions numériques μ-mesurables, les fonctions numériques $\operatorname{Sup}_n f_n$ et $\operatorname{Inf}_n f_n$ sont μ-mesurables.</u>

<u>Démonstration</u>. Posons $f = \operatorname{Sup}_n f_n$ et $g = \operatorname{Inf}_n f_n$. Pour tout $a \in \mathbb{R}$, on a :

$$\{x \in X \mid f(x) \leq a\} = \bigcap_n \{x \in X \mid f_n(x) \leq a\}$$
$$\{x \in X \mid g(x) < a\} = \bigcup_n \{x \in X \mid f_n(x) < a\}$$

et la μ-mesurabilité de f et g résulte immédiatement de 2.1.2.2. ■

<u>2.2.3.3 - Corollaire</u>. <u>Soient (X,μ) un espace mesuré et $f_n : X \to \overline{\mathbb{R}}$ $(n=1,2,\dots)$ une suite de fonctions μ-mesurables. Alors, la fonction numérique $\sum_{n\geq 1} |f_n|$ est μ-mesurable.</u>

<u>Démonstration</u>. On a :

$$\sum_{n\geq 1} |f_n| = \sup_{N\geq 1} \Big(\sum_{1\leq n\leq N} |f_n| \Big),$$

et la conclusion résulte de 2.2.1.3, 2.2.2.2 et 2.2.3.2. ■

Nous pouvons maintenant énoncer le résultat principal de cette section :

2.2.3.4 - Théorème - Soient (X,μ) un espace mesuré et $f_n : X \to \overline{\mathbb{R}}$ $(n=1,2,\dots)$ une suite de fonctions numériques μ-mesurables. On suppose que la suite $(f_n)_{n\geq 1}$ converge simplement μ-presque partout vers une fonction $f : X \to \overline{\mathbb{R}}$. Alors, f est μ-mesurable.

Démonstration. Soit $N \subset X$ un sous-ensemble μ-négligeable de X tel que $f_n(x) \to f(x)$ pour tout $x \notin N$. Quitte à remplacer les f_n et f par les fonctions $\tilde{f}_n$ et $\tilde{f}$ définies par :

$$\tilde{f}_n(x) = \begin{cases} f_n(x) & \text{si } x \notin N \\ 0 & \text{si } x \in N \end{cases}, \quad \tilde{f}(x) = \begin{cases} f(x) & \text{si } x \notin N \\ 0 & \text{si } x \in N, \end{cases}$$

on peut - en utilisant 2.1.1.3 - supposer que $f_n(x) \to f(x)$ pour tout $x \in X$. Mais alors, on a :

$$f(x) = \limsup_{n\to\infty} f_n(x) = \operatorname{Inf}_n \operatorname{Sup}(f_n(x), f_{n+1}(x), \dots),$$

et la conclusion résulte immédiatement de 2.2.3.2. ■

2.2.3.5. Corollaire - Soient (X,μ) un espace mesuré et $f_n : X \to \overline{\mathbb{R}}$ $(n=1,2,\dots)$ une suite de fonctions μ-mesurables. On suppose que la série $\sum_{n\geq 1} f_n(x)$ converge pour μ-presque tout $x \in X$. Alors, toute fonction $f : X \to \overline{\mathbb{R}}$ telle que $f(x) = \sum_{n\geq 1} f_n(x)$ pour μ-presque tout $x \in X$ est μ-mesurable.

Démonstration. Cela résulte immédiatement de 2.2.3.4. ■

Terminons cette section par un résultat qui relie la convergence simple μ-presque partout à la convergence uniforme sur des parties de mesure arbitrairement proche de la mesure de X, lorsque (X,μ) est un espace mesuré fini.

2.2.3.6 - Théorème (Egorov). Soient (X,μ) un espace mesuré et $f_n : X \to \overline{\mathbb{R}}$ une suite de fonctions numériques μ-mesurables. On suppose que la suite $(f_n)_{n\geq 1}$ converge simplement μ-presque partout vers une fonction $f : X \to \overline{\mathbb{R}}$. Alors, pour toute partie μ-mesurable $A \subset X$ telle que $\mu(A) < +\infty$ et tout $\varepsilon > 0$, il existe une partie μ-mesurable $A_\varepsilon \subset A$ telle que l'on ait :

(i) $\mu(A \setminus A_\varepsilon) \leqq \varepsilon$;
(ii) la suite $(f_n)_{n \geqq 1}$ converge uniformément vers f sur A.

Démonstration. Quitte à remplacer X par A, on peut supposer que $\mu(X) < +\infty$. Soit d la distance définissant la topologie de $\mathbb{R}$. Comme les f_n et f sont μ-mesurables et que l'application $(x,y) \mapsto d(x,y)$ est continue, l'application $x \mapsto d(f_n(x), f(x))$ est μ-mesurable. Il s'ensuit que, pour tout entier $q \geqq 1$, l'ensemble

$$E_{n,q} = \{x \in X \mid d(f_n(x), f(x)) > \tfrac{1}{q}\}$$

est μ-mesurable. Montrons que, pour tout entier $q \geqq 1$,

$$\mu\Big(\bigcap_{N \geqq 1} \bigcup_{n \geqq N} E_{n,q}\Big) = 0$$

Par hypothèse, il existe un sous-ensemble μ-négligeable $E \subset X$ tel que l'on ait, pour tout $x \notin E$, $d(f_n(x), f(x)) \to 0$ quand $n \to +\infty$. On a donc :

$$x \notin E \implies \forall q \geqq 1, \exists N \geqq 1 \text{ tel que } d(f_n(x), f(x)) \leqq \tfrac{1}{q} \text{ pour tout } n \geqq N.$$

Par contraposition, on a pour tout $x \in X$:

$$\Big(\exists q \geqq 1 \ \ \forall N \geqq 1 \ \ \exists n \geqq N \text{ tel que } d(f_n(x), f(x)) > \tfrac{1}{q}\Big) \implies x \in E,$$

soit

$$\bigcup_{q \geqq 1} \Big(\bigcap_{N \geqq 1} \bigcup_{n \geqq N} E_{n,q}\Big) \subset E,$$

et donc $\bigcap_{N \geqq 1} \bigcup_{n \geqq N} E_{n,q} \subset E$ pour tout entier $q \geqq 1$.

Comme $\mu(E) = 0$, on en déduit que

$$\mu\Big(\bigcap_{N \geqq 1} \bigcup_{n \geqq N} E_{n,q}\Big) = 0.$$

Mais les ensembles $F_{N,q} = \bigcup_{n \geqq N} E_{n,q}$ forment une suite décroissante et, comme $\mu(X) < +\infty$, on en déduit que $\mu(F_{N,q}) \to 0$ quand $N \to +\infty$.

Pour tout $\varepsilon > 0$ donné, on peut donc choisir, pour tout $q \geqq 1$, un entier $N(q) \geqq 1$ tel que l'on ait :

$$\mu(F_{N(q),q}) < \frac{\varepsilon}{2^q} .$$

Posons $X_\varepsilon = \left(\bigcup_{q \geqq 1} F_{N(q),q}\right)^c = \bigcap_{q \geqq 1} F^c_{N(q),q} = \bigcap_{q \geqq 1} \bigcap_{n \geqq N(q)} E^c_{n,q}$ -

On a :

$$\mu(X \setminus X_\varepsilon) = \mu\left(\bigcup_{q \geqq 1} F_{N(q),q}\right) \leqq \sum_{q \geqq 1} \frac{\varepsilon}{2^q} = \varepsilon ,$$ d'où

la condition (i). Pour terminer la démonstration, montrons que la suite $(f_n)_{n \geq 1}$ converge vers f uniformément sur X_ε. Or, si $x \in X_\varepsilon$, on a pour tout $q \geqq 1$:

$$d(f_n(x), f(x)) \leqq \frac{1}{q} \quad \text{quel que soit } n \geqq N(q),$$

ce qui prouve bien que

$$\sup_{x \in X_\varepsilon} d(f_n(x), f(x)) \longrightarrow 0 \quad \text{quand } n \to +\infty.$$

■

2.3. APPROXIMATION DES FONCTIONS MESURABLES PAR DES FONCTIONS SIMPLES.

2.3.1 - Soient (X,μ) un espace mesuré et $(f_1, f_2, \dots)$ une suite de fonctions simples $f_n : X \to \mathbb{R}$. Si la suite $(f_1, f_2, \dots)$ converge simplement vers $f : X \to \overline{\mathbb{R}}$, alors f est μ-mesurable en vertu de 2.2.2.3 et de 2.2.3.4. Inversement, on a :

2.3.2 - Proposition - Soient (X,μ) un espace mesuré et $f : X \to \mathbb{R}$ une fonction μ-mesurable. On a :

(i) Il existe une suite $(f_1, f_2, \dots)$ de fonctions simples qui converge simplement vers f ;

(ii) Si f est en outre positive, il existe une suite $(f_1, f_2, \dots)$ de fonctions simples vérifiant

$$0 \leqq f_1 \leqq f_2 \leqq \dots \leqq f$$

et qui converge simplement vers f ;

(iii) Si f est en outre bornée, il existe une suite $(f_1, f_2, \dots)$ de fonctions simples qui converge vers f

uniformément sur X.

Démonstration. (i) Pour tout entier $n \geqq 1$, posons :

$$E_{n,-\infty} = \{x \in X \mid f(x) < -n\}$$
$$E_{n,i} = \{x \in X \mid \tfrac{i}{n} \leqq f(x) < \tfrac{i+1}{n}\} \text{ pour } i = -n^2, -n^2+1, \dots, n^2-1$$
$$E_{n,+\infty} = \{x \in X \mid n \leqq f(x)\}.$$

On définit ainsi des sous-ensembles μ-mesurables de X qui sont deux à deux disjoints et dont la réunion est égale à X. Posons :

$$f_n = -n \mathbb{1}_{E_{n,-\infty}} + \sum_{i=-n^2}^{n^2-1} \frac{i}{n} \mathbb{1}_{E_{n,i}} + n \mathbb{1}_{E_{n,+\infty}}.$$

On définit ainsi une fonction simple sur X. Pour $x \in X$ tel que $f(x) \in \mathbb{R}$, il existe $N \geqq 1$ tel que $-N \leqq f(x) < N$, et on a, pour $n \geqq N$:

$$|f(x) - f_n(x)| \leqq \frac{1}{n},$$

d'où l'on déduit que $f_n(x) \longrightarrow f(x)$ quand $n \to +\infty$.
Si $f(x) = +\infty$, on a $f_n(x) = n$ pour tout $n \geqq 1$, et donc $f_n(x) \to +\infty = f(x)$ quand $n \to +\infty$.
Si $f(x) = -\infty$, on a $f_n(x) = -n$ pour tout $n \geqq 1$, et donc $f_n(x) \to -\infty = f(x)$ quand $n \to +\infty$. On a donc $f(x) = \lim_{n\to\infty} f_n(x)$ pour tout $x \in X$, d'où (i).

(ii) Supposons que $f(x) \geqq 0$ pour tout $x \in X$.
Pour tout entier i tel que $1 \leqq i \leqq n2^n$, posons

$$E_{n,i} = \{x \in X \mid \frac{i-1}{2^n} \leqq f(x) < \frac{i}{2^n}\} \text{ et}$$
$$E_{n,\infty} = \{x \in X \mid n \leqq f(x)\}.$$

Alors, la fonction

$$f_n = \sum_{i=1}^{n2^n} \frac{i-1}{2^n} \mathbb{1}_{E_{n,i}} + n \mathbb{1}_{E_{n,\infty}}$$

est simple. En outre, on a $f_n \leqq f_{n+1}$ pour tout $n \geqq 1$, et $f_n \leqq f$ pour tout $n \geqq 1$. Comme dans (i), on montre que $(f_n)_{n \geqq 1}$ converge simplement vers f, d'où l'assertion (ii).

(iii) Soit $f : X \to \overline{\mathbb{R}}$ une fonction μ-mesurable bornée. Alors, $f_+, f_- : X \to \mathbb{R}$ sont également bornées et, comme $f = f_+ - f_-$, il suffit de prouver que f_+ et f_- sont limite uniforme de fonctions simples. On est ainsi ramené au cas où f est positive et bornée. Or, dans ce cas, la construction utilisée en (ii) fournit une suite de fonctions simples f_n qui converge uniformément vers f sur X. ∎

3. FONCTIONS INTÉGRABLES

3. FONCTIONS INTÉGRABLES

3- FONCTIONS INTÉGRABLES

Dans tout, ce chapitre, on désigne par (X,μ) un espace mesuré et par $\mathcal{M}_\mu$ la tribu des sous-ensembles μ-mesurables de X. Rappelons qu'un sous-ensemble μ-négligeable est un sous-ensemble $N \in \mathcal{M}_\mu$ de X tel que $\mu(N) = 0$. Une fonction numérique sur X est une application $f : X \to \overline{\mathbb{R}}$; une telle fonction est dite positive si elle prend ses valeurs dans $\overline{\mathbb{R}}_+ = [0,+\infty]$. Enfin, deux fonctions numériques sur X sont dites égales μ-presque partout si elles coïncident en dehors d'un sous-ensemble μ-négligeable de X.

3.1- INTÉGRALE SUPÉRIEURE

3.1.1. Rappelons qu'une fonction simple sur X est une fonction numérique μ-mesurable $e : X \to \mathbb{R}$ dont l'image est un sous-ensemble fini $\{a_1, \dots, a_n\}$ de $\mathbb{R}$. Une telle fonction s'écrit canoniquement :

$$e = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}$$

où les sous-ensembles μ-mesurables

$$A_i = \{x \in X \mid e(x) = a_i\} = e^{-1}(\{a_i\})$$

sont deux à deux disjoints et ont pour réunion l'ensemble X. Lorsque tous les a_i sont positifs ou nuls, on dit que e est une fonction simple positive.

3.1.2. Définition. Soient (X,μ) un espace mesuré et $e : X \to [0,+\infty]$ une fonction simple positive, de la forme :

$$e = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}$$

où $a_1, \dots, a_n$ sont les valeurs distinctes de e. On appelle intégrale supérieure de e le nombre

$$\int_X e \, d\mu \in [0,+\infty]$$

défini par :

$$\int_X e\,d\mu = \sum_{i=1}^{n} a_i \mu(A_i).$$

Pour définir le produit de deux nombres de $\overline{\mathbb{R}}_+ = [0,+\infty]$ sans ambiguité, on utilise la convention $0.\infty = 0$. Lorsque $\mu(A_i) = +\infty$, on a donc :

$$a_i \mu(A_i) = \begin{cases} 0 & \text{si } a_i = 0 \\ +\infty & \text{si } 0 < a_i \end{cases}$$

3.1.3. Soient $e : X \to [0,+\infty]$ une fonction simple positive et $E \subset X$ un sous-ensemble μ-mesurable de X. Alors, la fonction $e \mathbb{1}_E$ égale à e sur E et nulle sur E^c est encore une fonction simple positive. On pose par définition :

$$\int_E e\,d\mu = \int_X e \mathbb{1}_E \,d\mu$$

Si $e = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}$, on a :

$$\int_E e\,d\mu = \sum_{i=1}^{n} a_i \mu(A_i \cap E),$$

comme on le vérifie aisément.

3.1.4. Lemme. Soit (X,μ) un espace mesuré.

(i) Pour toute fonction simple positive $e : X \to [0,+\infty]$, l'application $\nu : \mathcal{M}_\mu \to [0,+\infty]$ définie par :

$$\nu(E) = \int_E e\,d\mu$$

est une mesure positive sur $\mathcal{M}_\mu$;

(ii) Si $e, f : X \to [0,+\infty]$ sont deux fonctions simples, on a :

$$\int_X (e+f)\,d\mu = \int_X e\,d\mu + \int_X f\,d\mu.$$

<u>Démonstration</u>. (i) Soit $e = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}$ une fonction simple positive. Si $E = \emptyset$, la fonction simple $e\mathbb{1}_E$ est identiquement nulle, d'où il résulte que :

$$\nu(\emptyset) = \int_X 0 \, d\mu = 0.$$

Soit $(E_1, E_2, \dots)$ une suite de sous-ensembles μ-mesurables deux à deux disjoints de X et notons E leur réunion. On a :

$$\nu(E) = \sum_{i=1}^{n} a_i \mu(E \cap A_i) = \sum_{i=1}^{n} a_i \sum_{p \geq 1} \mu(E_p \cap A_i)$$
$$= \sum_{p \geq 1} \left(\sum_{i=1}^{n} a_i \mu(E_p \cap A_i) \right) = \sum_{p \geq 1} \nu(E_p),$$

ce qui achève de démontrer que ν est une mesure positive sur $\mathcal{M}_\mu$.

(ii) Soient $e = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}$ et $f = \sum_{j=1}^{p} b_j \mathbb{1}_{B_j}$ deux fonctions simples positives sur X. Pour montrer la relation

$$\int_X (e+f) \, d\mu = \int_X e \, d\mu + \int_X f \, d\mu,$$

nous allons montrer que l'on a, pour tout sous-ensemble μ-mesurable E de X :

$$\int_E (e+f) \, d\mu = \int_E e \, d\mu + \int_E f \, d\mu.$$

A cet effet, introduisons les mesures positives :

$$\nu(E) = \int_E (e+f) \, d\mu, \quad \nu'(E) = \int_E e \, d\mu, \quad \nu''(E) = \int_E f \, d\mu.$$

Pour montrer que

$$(1) \quad \nu(E) = \nu'(E) + \nu''(E)$$

pour tout $E \in \mathcal{M}_\mu$, observons que E est la réunion disjointe des $E_{ij} = E \cap A_i \cap B_j$ ($1 \leq i \leq n$ et $1 \leq j \leq p$). Comme ν, ν' et ν'' sont des mesures positives, il suffit, pour démontrer (1), de prouver que l'on a :

$$\nu(E_{i,j}) = \nu'(E_{i,j}) + \nu''(E_{i,j}) \quad \text{quels que soient } i,j.$$

Or on a :

$$\begin{aligned}\nu(E_{i,j}) &= (a_i + b_j)\mu(E \cap A_i \cap B_j) \\ &= a_i\,\mu(E\cap A_i \cap B_j) + b_j\,\mu(E \cap A_i \cap B_j) \\ &= \nu'(E_{i,j}) + \nu''(E_{i,j}),\end{aligned}$$

et (1) est démontrée. ■

3.1.5- Définition. Soient (X,μ) un espace mesuré et $f : X \to [0,+\infty]$ une fonction μ-mesurable positive. On appelle intégrale supérieure de f le nombre $\int_X f d\mu \in [0,+\infty]$ défini par :

$$\int_X f d\mu = \operatorname{Sup}\left\{ \int_X e\,d\mu \;\middle|\; e \text{ fonction simple, } 0 \leq e \leq f \right\}$$

Si $E \subset X$ est un sous-ensemble μ-mesurable de X, on pose (cf. 3.1.3) :

$$\int_E f d\mu = \int_X f \mathbb{1}_E \, d\mu ,$$

où $f\mathbb{1}_E : X \to [0,+\infty]$ est la fonction numérique positive μ-mesurable égale à f sur E et à 0 hors de E.

3.1.6. Proposition. Soient (X,μ) un espace mesuré et $f, g : X \to [0,+\infty]$ deux fonctions μ-mesurables positives. On a :

(i) $0 \leq f \leq g$ implique $\int_X f d\mu \leq \int_X g\, d\mu$;

(ii) Pour tout $\lambda \in [0,+\infty[$, on a :

$$\int_X \lambda f \, d\mu = \lambda \int_X f d\mu \, ;$$

(iii) $\int_X f d\mu + \int_X g d\mu \leqq \int_X (f+g) d\mu$;

(iV) Si f et g sont égales μ-presque partout, alors $\int_X f d\mu = \int_X g d\mu$

Démonstration. (i) et (ii) sont immédiates.

(iii) Soient e, s, $0 \leqq e \leqq f$, $0 \leqq s \leqq g$, deux fonctions simples positives. Alors, $e+s$ est une fonction simple positive qui vérifie :

$$0 \leqq e+s \leqq f+g.$$

On a donc :

$$\int_X (e+s) d\mu \leqq \int_X (f+g) d\mu,$$

d'où, en vertu de 3.1.4 (ii) :

$$\int_X e d\mu + \int_X s d\mu \leqq \int_X (f+g) d\mu.$$

En prenant le sup sur toute les fonctions simples positives $e \leqq f$, puis le sup sur toute les fonctions simples positives $s \leqq g$, on en déduit que :

$$\int_X f d\mu + \int_X g d\mu \leqq \int_X (f+g) d\mu.$$

(iV). Pour démontrer (iv), il suffit de prouver que l'on a, pour toute fonction μ-mesurable positive $f : X \longrightarrow [0,+\infty]$ et tout ensemble μ-négligeable $N \subset X$:

$$\int_X f d\mu = \int_{N^c} f d\mu.$$

Comme $f \mathbb{1}_{N^c} \leqq f$, il résulte immédiatement de (i) que $\int_{N^c} f d\mu \leqq \int_X f d\mu$. Montrons l'inégalité inverse.

A cet effet, soit e, $0 \leqq e \leqq f$ une fonction simple positive, de la forme

$$e = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}$$

Alors, on a :

$$\int_{N^c} e\,d\mu = \sum_{i=1}^{n} a_i\,\mu(A_i \cap N^c) = \sum_{i=1}^{n} a_i\,\mu(A_i) = \int_X e\,d\mu$$

car $\mu(A_i \cap N) \leqq \mu(N) = 0$. Comme $0 \leqq e\mathbb{1}_{N^c} \leqq f\mathbb{1}_{N^c}$, on en déduit que :

$$\int_X e\,d\mu = \int_X e\mathbb{1}_{N^c}\,d\mu \leqq \int_X f\mathbb{1}_{N^c}\,d\mu = \int_{N^c} f\,d\mu.$$

En prenant le sup sur les fonctions simples positives e telles que $e \leqq f$, on en déduit que :

$$\int_X f\,d\mu \leqq \int_{N^c} f\,d\mu \text{ , d'où } \int_X f\,d\mu = \int_{N^c} f\,d\mu. \qquad \blacksquare$$

On notera qu'il résulte de la propriété (iv) que, si $\mu(E) = 0$, alors $\int_E f\,d\mu = 0$, même si $f(x) = +\infty$ pour tout $x \in E$. On notera que l'on a en fait l'égalité dans la propriété (iii), ce qui signifie que l'intégrale supérieure est en fait additive. Nous allons montrer qu'elle est même dénombrablement additive, c'est à dire que l'on a :

$$\int_X \Big(\sum_{n \geqq 1} f_n\Big) d\mu = \sum_{n \geqq 1} \int_X f_n\,d\mu$$

pour toute suite $(f_1, f_2, \dots)$ de fonctions μ-mesurables positives. Cette propriété résulte du théorème suivant :

¶ 3.1.7. Théorème (de convergence monotone). Soient (X,μ) un espace mesuré et $(f_1, f_2, \dots)$ une suite croissante de fonctions μ-mesurables positives:
$$0 \leqq f_1(x) \leqq f_2(x) \leqq \cdots\cdots \leqq +\infty \text{ pour tout } x \in X.$$
Posons $f = \operatorname{Sup}_n f_n$. Alors f est une fonction μ-mesurable positive et on a:
$$\int_X f d\mu = \operatorname{Sup}_n \int_X f_n d\mu.$$

Démonstration. La mesurabilité de f résulte de la proposition 2.2.3.2. Comme on a par hypothèse $f_n \leqq f_{n+1}$ pour tout $n \geqq 1$, il résulte de 3.1.6 (i) que $\int_X f_n d\mu \leqq \int_X f_{n+1} d\mu$, et donc $\int_X f_n d\mu \xrightarrow[n \to +\infty]{} \operatorname{Sup}_n \int_X f_n d\mu$.

Posons $S = \operatorname{Sup}_n \int_X f_n d\mu$. De la relation $f_n \leqq f$ on déduit que $\int_X f_n d\mu \leqq \int_X f d\mu$ pour tout $n \geqq 1$, et donc que $S \leqq \int_X f d\mu$. Montrons que l'on a aussi
$$\int_X f d\mu \leqq S.$$
A cet effet, fixons une fonction simple positive e telle que $0 \leqq e \leqq f$. Pour tout λ tel que $0 < \lambda < 1$, posons:
$$E_n^\lambda = \{x \in X \mid f_n(x) \geqq \lambda e(x)\}$$
Comme les f_n et e sont μ-mesurables, on définit ainsi un sous-ensemble μ-mesurable $E_n^\lambda \subset X$. Comme la suite $(f_1, f_2, \dots)$ est croissante, on a:
$$E_1^\lambda \subset E_2^\lambda \subset \cdots \subset E_n^\lambda \subset E_{n+1}^\lambda \subset \cdots$$
En outre, on a: $X = \bigcup_{n \geqq 1} E_n^\lambda$. En effet, si $f(x) = 0$,

alors $e(x) = 0$ et $x \in E_1^\lambda$. Si $f(x) \neq 0$, on a

$$\lambda e(x) < f(x) \qquad \text{puisque } \lambda < 1,$$

et il existe un entier $n \geqq 1$ tel que :

$$\lambda e(x) \leqq f_n(x),$$

d'où $x \in E_n^\lambda$. On a donc bien $X = \bigcup_{n \geqq 1} E_n^\lambda$. Par ailleurs, pour $x \in E_n^\lambda$, on a :

$$\lambda e(x) \leqq f_n(x),$$

de sorte que :

$$\lambda e \, \mathbb{1}_{E_n^\lambda} \leqq f_n \mathbb{1}_{E_n^\lambda} \leqq f_n .$$

Il résulte alors de la proposition 3.1.6 que :

$$\lambda \int_{E_n^\lambda} e \, d\mu \leqq \int_X f_n \, d\mu \leqq S.$$

Comme l'application $E \longmapsto \nu(E) = \int_E e \, d\mu$ est une mesure positive sur $\mathcal{M}_\mu$ (cf. (E 3.1.4. i), on déduit de la relation

$$\lambda \nu(E_n^\lambda) \leqq S$$

et du fait que les $(E_n^\lambda)_{n \geqq 1}$ forment une suite croissante d'ensembles μ-mesurables de réunion X, que :

$$\lambda \nu(X) \leqq S.$$

En d'autres termes, on a :

$$\lambda \int_X e \, d\mu \leqq S \quad \text{pour tout } \lambda \in]0, 1[,$$

d'où, en faisant tendre λ vers 1 :

$$\int_X e \, d\mu \leqq S.$$

Puisque cette dernière relation est vraie pour toute fonction simple e telle que $0 \leqq e \leqq f$, on en déduit que $\int_X f \, d\mu \leqq S$, et le théorème est démontré.

■

3.1.8. Théorème. Soient (X,μ) un espace mesuré et $(f_1, f_2, \dots)$ une suite de fonctions μ-mesurables positives $f_n : X \longrightarrow [0,+\infty]$. Posons $f = \sum_{n=1}^{\infty} f_n$. Alors on a :

$$\int_X f d\mu = \sum_{n=1}^{\infty} \int_X f_n \, d\mu.$$

Démonstration. Montrons tout d'abord que, si $f, g : X \to [0,+\infty]$ sont deux fonctions μ-mesurables positives, on a :

$$\int_X (f+g) d\mu = \int_X f d\mu + \int_X g d\mu.$$

D'après la proposition 2.3.2, il existe une suite

$$0 \leqq e_1 \leqq e_2 \leqq \dots\dots \leqq f$$

de fonctions simples telle que $f(x) = \lim_{n \to +\infty} e_n(x)$ pour tout $x \in X$. De même, il existe une suite

$$0 \leqq s_1 \leqq s_2 \leqq \dots\dots \leqq g$$

de fonctions simples telle que $g(x) = \lim_{n \to +\infty} s_n(x)$ pour tout $x \in X$. On a alors :

$$0 \leqq e_1 + s_1 \leqq e_2 + s_2 \leqq \dots \leqq e_n + s_n \leqq \dots \leqq f+g$$

avec $(f+g)(x) = \lim_{n \to +\infty} (e_n + s_n)(x)$ pour tout $x \in X$. D'après le théorème de convergence monotone, on a :

$$\begin{aligned}
\int_X (f+g) d\mu &= \lim_{n \to \infty} \int_X (e_n + s_n) d\mu \\
&= \lim_{n \to \infty} \left(\int_X e_n d\mu + \int_X s_n d\mu \right) \quad (\text{cf. } 3.1.4 \text{ ii}) \\
&= \lim_{n \to \infty} \int_X e_n d\mu + \lim_{n \to \infty} \int_X s_n d\mu \\
&= \int_X f d\mu + \int_X g d\mu \quad (\text{cf. } 3.1.7).
\end{aligned}$$

Nous sommes maintenant en mesure de démontrer le théorème 3.1.8. A cet effet, posons $g_n = f_1 + \dots + f_n$.

D'après ce qui précède, on a :

$$\int_X g_n \, d\mu = \sum_{p=1}^{n} \int_X f_p \, d\mu .$$

En appliquant le théorème de convergence monotone à la suite

$$0 \leqq g_1 \leqq g_2 \leqq \cdots\cdots \leqq f = \sum_{n=1}^{\infty} f_n = \sup_n g_n ,$$

on obtient que

$$\int_X f d\mu = \sup_n \sum_{p=1}^{n} \int_X f_p d\mu = \sum_{n=1}^{\infty} \int_X f_n d\mu . \quad \blacksquare$$

3.1.9 - Proposition (Lemme de Fatou). Soient (X, μ) un espace mesuré et $(f_1, f_2, \ldots)$ une suite de fonctions μ-mesurables $f_n : X \longrightarrow [0, +\infty]$. On a :

$$\int_X (\liminf_{n \to +\infty} f_n) \, d\mu \leqq \liminf_{n \to +\infty} \int_X f_n \, d\mu .$$

Démonstration. Pour tout entier $k \geqq 1$, posons

$$g_k = \inf(f_k, f_{k+1}, \ldots) .$$

Comme on a $g_k \leqq f_k$ pour tout $k \geqq 1$, la croissance de l'intégrale supérieure implique que :

$$(1) \quad \int_X g_k \, d\mu \leqq \int_X f_k \, d\mu .$$

Comme les g_k forment une suite croissante de fonctions μ-mesurables positives, et que $g_k(x) \xrightarrow[k \to \infty]{} \liminf_{n \to \infty} f_n(x)$, on a :

$$\int_X g_k \, d\mu \xrightarrow[k \to +\infty]{} \int_X (\liminf_{n \to +\infty} f_n) \, d\mu .$$ Mais alors,

il résulte immédiatement de (1) que :

$$\int_X (\liminf_{n \to +\infty} f_n) d\mu \leqq \liminf_{n \to +\infty} \int_X f_n \, d\mu . \quad \blacksquare$$

On montre que l'inégalité stricte dans la relation ci-dessus peut se produire ; il n'y a donc pas égalité en général.

3.2. FONCTIONS INTÉGRABLES

3.2.1- Définition. Soit (X,μ) un espace mesuré. On dit qu'une fonction $f: X \to \overline{\mathbb{R}}$ est μ-intégrable si elle est μ-mesurable et si $\int_X |f| d\mu < +\infty$.

On notera $\mathcal{L}^1(X,\mu)$ l'ensemble des fonctions numériques μ-intégrables sur X. Il résulte immédiatement de la définition 3.2.1 que toute fonction μ-mesurable $f: X \to \overline{\mathbb{R}}$ qui est égale μ-presque partout à une fonction $g \in \mathcal{L}^1(X,\mu)$ est elle-même μ-intégrable.

Un sous-ensemble $A \subset X$ sera dit μ-intégrable si sa fonction caractéristique $\mathbb{1}_A$ est μ-intégrable, ce qui équivaut à dire que A est μ-mesurable et que $\mu(A) < +\infty$.

Donnons un premier exemple de fonction μ-intégrable.

¶ 3.2.2- Proposition- Soient (X,μ) un espace mesuré et $f: X \to \mathbb{R}$ une fonction μ-mesurable. On suppose qu'il existe un sous-ensemble μ-intégrable $A \subset X$ tel que l'on ait :

(i) $f(x) = 0$ pour tout $x \notin A$;

(ii) Il existe un réel $M \geqq 0$ tel que $|f(x)| \leqq M$ pour tout $x \in A$.

Alors, $f \in \mathcal{L}^1(X,\mu)$.

Démonstration. De (i) et (ii), on déduit que :

$$|f| \leqq M \mathbb{1}_A ,$$

d'où $\int_X |f| d\mu \leqq \int_X M \mathbb{1}_A d\mu = M\mu(A) < +\infty$. Comme f est μ-mesurable, on en déduit que $f \in \mathcal{L}^1(X,\mu)$. ■

3.2.3- Corollaire. Toute fonction continue $f: \mathbb{R}^n \to \mathbb{R}$ qui est nulle hors d'un pavé est intégrable pour la mesure de Lebesgue sur $\mathbb{R}^p$

Démonstration. Comme f est continue, elle est λ-mesurable en vertu du Corollaire 2.1.3.3. Soit P un pavé fermé tel que $f(x)=0$ pour $x\notin P$. Comme P est fermé et borné, la fonction continue f est bornée sur P. Enfin, on a $\lambda(P)<+\infty$, et il résulte de 3.2.2 que $f\in\mathcal{L}^1(\mathbb{R}^p,\lambda)$. ∎

3.2.4. Proposition. Soient (X,μ) un espace mesuré et $f\in\mathcal{L}^1(X,\mu)$ une fonction numérique μ-intégrable. Alors, l'ensemble des $x\in X$ tels que $|f(x)|=+\infty$ est μ-négligeable.

En d'autres termes, toute fonction μ-intégrable $f:X\to\overline{\mathbb{R}}$ est égale μ-presque partout à une fonction μ-intégrable $\tilde{f}:X\to\mathbb{R}$.

Démonstration. Soit $f\in\mathcal{L}^1(X,\mu)$ et posons

$$A=\{x\in X \mid |f(x)|=+\infty\}.$$

On a:

$$A=\bigcap_{n\geq 1} A_n,$$

où

$$A_n=\{x\in X \mid |f(x)|\geq n\}.$$

Comme f est μ-mesurable, A_n est un sous-ensemble μ-mesurable de X pour tout entier $n\geq 1$. On a:

$$A_1\supset A_2\supset\cdots\supset A_n\supset A_{n+1}\supset\cdots$$

et la relation $\mathbb{1}_{A_1}\leq|f|$ implique que $\mu(A_1)<+\infty$.

On a donc

$$\mu(A)=\lim_{n\to+\infty}\mu(A_n)$$

en vertu du théorème 1.2.5 (iii). Or on a, pour tout $n\geq 1$:

$$n\,\mathbb{1}_{A_n}\leq|f|, \text{ d'où}$$

$$n\,\mu(A_n)\leq\int_X|f|\,d\mu=C<+\infty.$$

On en déduit que:

$$\mu(A_n) \leq \frac{C}{n} \xrightarrow[n\to+\infty]{} 0 ,$$

et donc $\mu(A) = \lim_{n\to\infty} \mu(A_n) = 0$. ∎

3.2.5- Intégrale d'une fonction intégrable

3.2.5.1- Soit $f \in \mathcal{L}^1(X,\mu)$ une fonction numérique μ-intégrable. On a:

$$f = f^+ - f^-$$

où $f^+ = \sup(f,0)$ et $f^- = \sup(-f,0)$ sont des fonctions numériques positives μ-mesurables en vertu de 2.2.1.4. Comme on a

$$f^+ \leqq |f| \quad \text{et} \quad f^- \leqq |f|,$$

on en déduit que

$$\int_X f^{\pm} d\mu \leqq \int_X |f| d\mu < +\infty ,$$

et par conséquent f^+ et f^- sont μ-intégrables. On pose:

3.2.5.2- Définition. Pour toute fonction numérique μ-intégrable $f \in \mathcal{L}^1(X,\mu)$, on appelle intégrale de f le nombre $\int_X f d\mu$ défini par:

$$\int_X f d\mu = \int_X f^+ d\mu - \int_X f^- d\mu ,$$

où $f^+ = \sup(f,0)$ et $f^- = \sup(-f,0)$ sont respectivement la partie positive et la partie négative de f.

On a:

¶ 3.2.5.3- Théorème. Soit $f \in \mathcal{L}^1(X,\mu)$. Alors, $|f| \in \mathcal{L}^1(X,\mu)$, et on a:

$$\left| \int_X f d\mu \right| \leqq \int_X |f| d\mu$$

Démonstration. Si $f \in \mathcal{L}^1(X,\mu)$, il résulte immédiatement

de 2.2.1.3 et de la définition 3.2.1 que $|f| \in \mathcal{L}^1(X,\mu)$.
Par ailleurs, on a :

$$\left|\int_X f\,d\mu\right| = \left|\int_X f^+ d\mu - \int_X f^- d\mu\right| \leq \int_X f^+ d\mu + \int_X f^- d\mu = \int_X (f^+ + f^-)\,d\mu,$$

et comme $f^+ + f^- = |f|$, le théorème est démontré. ■

<u>3.2.5.4</u>. Pour toute $f \in \mathcal{L}^1(X,\mu)$, on pose :

$$\|f\|_1 = \int_X |f|\,d\mu .$$

On a :

<u>3.2.5.5 - Proposition</u>. <u>Soit</u> $f \in \mathcal{L}^1(X,\mu)$. <u>On a</u> :

$$\underline{\|f\|_1 = 0 \implies f = 0 \ \mu\text{-presque partout}}.$$

<u>Démonstration</u>. Pour tout entier $n \geq 1$, posons :

$$A_n = \left\{ x \in X \mid |f(x)| \geq \frac{1}{n} \right\}.$$

On définit ainsi une suite croissante de sous-ensembles μ-mesurables dont la réunion est l'ensemble :

$$A = \{ x \in X \mid f(x) \neq 0 \}.$$

On a donc :

$$\mu(A) = \lim_{n \to \infty} \mu(A_n).$$

Or on a, pour tout entier $n \geq 1$:

$$\frac{1}{n} 1\!\!1_{A_n} \leq |f| \quad , \text{ et donc}$$

$$\frac{1}{n}\mu(A_n) = \int_X \frac{1}{n} 1\!\!1_{A_n} d\mu \leq \int_X |f|\,d\mu = \|f\|_1 = 0.$$

Il s'ensuit que $\mu(A_n) = 0$ pour tout entier $n \geq 1$, et par conséquent $\mu(A) = \lim_{n\to\infty} \mu(A_n) = 0$. Ceci prouve que f est nulle μ-presque partout. ■

Démonstration - Si $f = g$ μ-presque partout (μ.p.p. en abrégé), on a $f^{\pm} = g^{\pm}$ μ.p.p., d'où

$$\int_X f^{\pm} d\mu = \int_X g^{\pm} d\mu \quad \text{en vertu de 3.1.6.}$$

Il s'ensuit que

$$\int_X f d\mu = \int_X g d\mu.$$

■

3.2.5.6. Une des propriétés importantes de l'intégrale des fonctions intégrables est sa linéarité. Il faut toutefois prendre garde au fait que la somme de deux fonctions numériques intégrables n'est pas toujours définie, car si $f(x) = +\infty$ et $g(x) = -\infty$, la somme $f(x)+g(x)$ n'a pas de sens. En fait, si $f \in \mathcal{L}^1(X,\mu)$ et $g \in \mathcal{L}^1(X,\mu)$, il existe en vertu de 3.2.4 un sous-ensemble μ-négligeable $N \subset X$ tel que $f(x) \in \mathbb{R}$ et $g(x) \in \mathbb{R}$ pour tout $x \in N^c$. La somme $f+g$ est alors définie μ-presque partout, en posant pour $x \notin N$:

$$(f+g)(x) = f(x) + g(x).$$

Cette circonstance nous conduira, au paragraphe suivant, à considérer les classes de fonctions μ-intégrables modulo l'égalité presque partout, afin de discuter la structure linéaire de l'espace des (classes de) fonctions μ-intégrables.

3.2.6. Approximation par des fonctions simples intégrables

3.2.6.1. Lemme. Soit (X,μ) un espace mesuré. Une fonction simple $e : X \longrightarrow \overline{\mathbb{R}}$ est μ-intégrable si et seulement si e est nulle hors d'un ensemble μ-intégrable et si $e(x) \in \mathbb{R}$ pour μ-presque tout $x \in X$.

Démonstration - Ecrivons $e = \sum_{i=1}^{n} a_i 1\!\!1_{A_i} + \infty 1\!\!1_{A_\infty} + (-\infty) 1\!\!1_{A_{-\infty}}$, où les a_i sont des réels non nuls. Comme e est

μ-mesurable, elle est μ-intégrable si et seulement si $|e| = \sum_{i=1}^{n} |a_i| 1\!\!1_{A_i} + \infty 1\!\!1_{A_\infty \cup A_{-\infty}}$ est μ-intégrable, i.e. si et seulement si

$$\sum_{i=1}^{n} |a_i| \mu(A_i) + \infty \mu(A_\infty \cup A_{-\infty}) < +\infty$$

Cette dernière condition équivaut à $\mu(A_\infty \cup A_{-\infty}) = 0$ et $\mu(A_i) < +\infty$ pour tout $i = 1, \dots, n$, d'où le lemme. ■

¶ 3.2.6.2- Théorème. Soient (X, μ) un espace mesuré et $f \in \mathcal{L}^1(X, \mu)$ une fonction numérique μ-intégrable. Il existe une suite $e_1, e_2, \dots$ de fonctions simples μ-intégrables telles que

$$\| f - e_n \|_1 = \int_X |f - e_n| d\mu \longrightarrow 0$$

quand $n \to +\infty$.

Démonstration. Supposons tout d'abord que f soit positive. Comme

$$\int_X f d\mu = \operatorname{Sup} \left\{ \int_X e \, d\mu \mid e \text{ simple et } 0 \leqq e \leqq f \right\}$$

il existe une suite $e_1, e_2, \dots$ de fonctions simples telles que $0 \leqq e_n \leqq f$ et $\int_X f d\mu = \lim_{n \to \infty} \int_X e_n d\mu$.

Comme $|e_n| \leqq f$ et que f est intégrable, il résulte de 3.1.6.(i) que e_n est μ-intégrable. Par ailleurs, on a:

$$\int_X f \, d\mu = \int_X (|f - e_n| + e_n) d\mu$$
$$= \int_X |f - e_n| d\mu + \int_X e_n d\mu$$

en vertu de 3.1.8, et donc

$$\int_X |f - e_n| d\mu = \int_X f d\mu - \int_X e_n d\mu \xrightarrow[n \to +\infty]{} 0 .$$

Quitte à modifier chaque e_n sur un ensemble μ-négligeable, on peut en outre supposer que e_n est à valeurs réelles pour tout entier $n \geqq 1$.

Si f est intégrable sans être positive, on décompose f sous la forme $f = f^+ - f^-$, où f^+ et f^- sont des fonctions positives μ-intégrables. D'après ce qui précède, il existe des suites $(e'_1, e'_2, \dots)$ et $(e''_1, e''_2, \dots)$ de fonctions simples intégrables, à valeurs réelles positives ou nulles, telles que :

$$\int_X |f^+ - e'_n| d\mu \longrightarrow 0 \text{ et } \int_X |f^- - e''_n| d\mu \longrightarrow 0.$$

Posons $e_n = e'_n - e''_n$ (ce qui est dépourvu d'ambiguïté puisque e'_n et e''_n sont à valeurs réelles). On définit ainsi pour tout entier $n \geqq 1$ une fonction simple, qui est intégrable en vertu du lemme 3.2.6.1. On a en outre :

$$|f - e_n| = |f^+ - e'_n - f^- + e''_n| \leqq |f^+ - e'_n| + |f^- - e''_n|$$

et donc :

$$\int_X |f - e_n| d\mu \leqq \int_X |f^+ - e'_n| d\mu + \int_X |f^- - e''_n| d\mu \xrightarrow[n \to +\infty]{} 0.$$

■

3.3. L'ESPACE $L^1(X, \mu)$

<u>3.3.1</u>. Considérons sur $\mathcal{L}^1(X,\mu)$ la relation d'équivalence :

$$f \sim g \iff f \text{ et } g \text{ sont égales } \mu\text{-presque partout.}$$

On note $L^1(X,\mu)$ le quotient de $\mathcal{L}^1(X,\mu)$ par cette relation d'équivalence. Un élément de $L^1(X,\mu)$ est donc une classe d'équivalence de fonctions de $\mathcal{L}^1(X,\mu)$; la classe de $f \in \mathcal{L}^1(X,\mu)$ sera notée $\dot{f} \in L^1(X,\mu)$. D'après la proposition 3.2.4, tout élément de $L^1(X,\mu)$ est de la forme $\dot{f}$, où $f \in \mathcal{L}^1(X,\mu)$

est une fonction numérique finie partout, c'est à dire telle que $f(x) \in \mathbb{R}$ pour tout $x \in X$.

Soit f une fonction numérique définie seulement μ-presque partout sur X, c'est à dire sur le complémentaire $X-N$ d'un sous-ensemble μ-négligeable N de X. Supposons que la fonction $\tilde{f} : X \longrightarrow \mathbb{R}$ définie par :

$$\tilde{f}(x) = \begin{cases} f(x) & \text{si } x \notin N \\ 0 & \text{si } x \in N \end{cases}$$

appartienne à $\mathcal{L}^1(X,\mu)$. Alors, tout autre prolongement de f à X en donnant des valeurs arbitraire sur N appartiendra à $\mathcal{L}^1(X,\mu)$, et tous ces prolongements seront dans la même classe de $L^1(X,\mu)$. On notera encore $\dot{f}$ cette classe, bien que f ne soit à proprement parler que définie μ-presque partout.

3.3.2. Soient $\dot{f}, \dot{g} \in L^1(X,\mu)$ et $\lambda \in \mathbb{R}$. On definit sans ambiguité la somme $\dot{f}+\dot{g}$ et le produit de $\dot{f}$ par le scalaire λ en posant :

$$\dot{f}+\dot{g} = (\dot{\widehat{f+g}})$$
$$\lambda \dot{f} = (\dot{\widehat{\lambda f}})$$

En effet, comme f et g sont finies μ-presque partout, la somme $f+g$ est définie μ-presque partout, c'est à dire sauf peut-être sur un sous-ensemble μ-négligeable N de X. Notons encore $f+g$ le prolongement obtenu en donnant la valeur 0 sur N. Alors, $f+g$ est μ-mesurable et vérifie :

$$|f+g| \leqq |f| + |g|,$$

d'où

$$\int_X |f+g| \, d\mu \leqq \int_X |f| \, d\mu + \int_X |g| \, d\mu < +\infty.$$

Ceci montre que $f+g \in \mathcal{L}^1(X,\mu)$, et on vérifie immédiatement que $(\widehat{f+g})$ ne depend pas du choix de f et de g dans $\dot{f}$ et $\dot{g}$. La somme $\dot{f}+\dot{g}$ est donc parfaitement définie comme élément de $L^1(X,\mu)$. De même, le produit $\lambda \cdot \dot{f}$ est bien défini comme élément de $L^1(X,\mu)$.

On vérifie immédiatement que, muni de ces lois, $L^1(X,\mu)$ est un espace vectoriel sur $\mathbb{R}$, le zéro de cet espace vectoriel étant la classe des fonctions qui sont nulles μ-presque partout.

On sait (cf. 3.2.5.5) que, si $f, g \in \mathcal{L}^1(X,\mu)$ sont égales μ-presque partout, alors

$$\int_X f\,d\mu = \int_X g\,d\mu.$$

On peut donc définir l'intégrale de $\dot{f} \in L^1(X,\mu)$ en posant

$$\int_X \dot{f}\,d\mu = \int_X f\,d\mu,$$

et poser (sans ambiguïté):

$$\|\dot{f}\|_1 = \int_X |f|\,d\mu.$$

Il résulte alors du théorème 3.2.5.3 que l'on a:

$$\left|\int_X \dot{f}\,d\mu\right| \leqq \int_X |\dot{f}|\,d\mu,$$

où $|\dot{f}| = \dot{\widehat{|f|}}$ pour $f \in \mathcal{L}^1(X,\mu)$.

¶ Dans la pratique, on commet l'abus de langage qui consiste à noter par la même lettre la fonction $f \in \mathcal{L}^1(X,\mu)$ et sa classe $\dot{f} \in L^1(X,\mu)$. L'intérêt du « point de vue $\mathcal{L}^1(X,\mu)$ » est que les éléments de $\mathcal{L}^1(X,\mu)$ sont des fonctions (et non des classes d'équivalence), mais l'intérêt de $L^1(X,\mu)$ est d'être un espace vectoriel. De même, il arrive fréquemment que l'on écrive $f = g$ au lieu de $f(x) = g(x)$ pour μ-presque tout $x \in X$.

Avec ces conventions, on a:

3.3.3 - Proposition. Soit (X,μ) un espace mesuré.

(i) Si $f, g \in \mathcal{L}^1(X,\mu)$ et si $\alpha, \beta \in \mathbb{R}$, on a:

$$\int_X (\alpha f + \beta g)\,d\mu = \alpha \int_X f\,d\mu + \beta \int_X g\,d\mu;$$

(ii) Si $f, g \in \mathcal{L}^1(X,\mu)$ et si $f \leqq g$ μ-presque partout, alors

$$\int_X f\,d\mu \leqq \int_X g\,d\mu;$$

(iii) Si $f, g \in \mathcal{L}^1(X,\mu)$, on a :

$$\|f+g\|_1 \leqq \|f\|_1 + \|g\|_1 .$$

Démonstration : (i). Si f, g sont des fonctions positives de $\mathcal{L}^1(X,\mu)$, on a :

$$\int_X (f+g)\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu$$

en vertu du théorème 3.1.8. Pour $f, g \in \mathcal{L}^1(X,\mu)$, on a donc :

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

d'où

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

et il résulte de ce qui précède que :

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

On en déduit que :

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

De même, on vérifie que $\int_X \alpha f\,d\mu = \alpha \int_X f\,d\mu$ pour tout $\alpha \in \mathbb{R}$, ce qui démontre (i).

(ii) Si $f \leqq g$, alors $g - f \geqq 0$ et ce qui précède implique que :

$$0 \leqq \int_X (g-f)\,d\mu = \int_X g\,d\mu - \int_X f\,d\mu ,$$

d'où $\int_X f\,d\mu \leqq \int_X g\,d\mu$. Le raisonnement s'étend sans peine au cas où les inégalités sont remplacées par des inégalités μ-presque partout.

(iii) On a :

$$\|f+g\|_1 = \int_X |f+g|\,d\mu \leqq \int_X (|f|+|g|)\,d\mu = \int_X |f|\,d\mu + \int_X |g|\,d\mu$$
$$= \|f\|_1 + \|g\|_1 .$$

■

3.3.4. Proposition. Soit (X,μ) un espace mesuré. Alors :
(i) L'application $\dot{f} \longmapsto \|f\|_1$ est une norme sur $L^1(X,\mu)$;
(ii) L'application $\dot{f} \longmapsto \int_X f\,d\mu$ est une forme linéaire continue sur $L^1(X,\mu)$ de norme $\leqq 1$.

Démonstration : (i) On a $\|\lambda f\|_1 = |\lambda|\,\|f\|_1$ pour toute $f \in \mathcal{L}^1(X,\mu)$ et tout $\lambda \in \mathbb{R}$ en vertu de 3.1.6 (i). Par ailleurs, on a $\|f+g\|_1 \leqq \|f\|_1 + \|g\|_1$ quelles que soient $f,g \in \mathcal{L}^1(X,\mu)$ en vertu de 3.3.3 (iii). Enfin, si $\|\dot{f}\| = 0$, alors f est égale à 0 μ-presque partout en vertu de 3.2.5.5 et donc $\dot{f} = 0$. Ceci prouve (i).

(ii) On a, pour toute $f \in \mathcal{L}^1(X,\mu)$:

$$\left|\int_X f\,d\mu\right| \leqq \int_X |f|\,d\mu = \|f\|_1$$

en vertu de 3.2.5.3, ce qui prouve que l'application linéaire $\dot{f} \longmapsto \int_X f\,d\mu$ est continue, de norme $\leqq 1$.

■

Le théorème suivant est un résultat clef de la théorie de l'intégration :

¶ 3.3.5 - THÉORÈME. Soit (X,μ) un espace mesuré. Si $(\dot{f}_1, \dot{f}_2, \ldots)$ est une suite de Cauchy dans $L^1(X,\mu)$ pour la norme $\|\cdot\|_1$, il existe une sous-suite $(f_{n_k})_k$ de la suite $(f_n)_n$ qui converge μ-presque partout vers une fonction $f \in \mathcal{L}^1(X,\mu)$. En outre, on a $\|f - f_{n_k}\|_1 \xrightarrow[k\to+\infty]{} 0$.

Démonstration. Puisque $\|f_n - f_p\|_1 \longrightarrow 0$ quand

$n, p \to +\infty$, on peut trouver une suite strictement croissante $(n_k)_k$ d'entiers ≥ 1 telle que :

$$\| f_{n_{k+1}} - f_{n_k} \|_1 \leq \frac{1}{2^k} \quad \text{pour tout } k \geq 0.$$

Posons $u_0 = f_{n_0}$ et $u_k = f_{n_k} - f_{n_{k-1}}$. On a, en vertu de 3.1.8 :

$$\int_X \sum_{k=0}^{\infty} |u_k| \, d\mu = \sum_{k=0}^{\infty} \int_X |u_k| \, d\mu = \| f_{n_0} \|_1 + \sum_{k=1}^{\infty} \| f_{n_k} - f_{n_{k-1}} \|_1$$

$$\leq \| f_{n_0} \|_1 + \sum_{k=1}^{\infty} \frac{1}{2^{k-1}} = \| f_{n_0} \|_1 + 2 < +\infty,$$

ce qui montre que la fonction $g = \sum_{k \geq 0} |u_k|$ est μ-intégrable. D'après 3.2.4, on a :

$$\sum_{k \geq 0} |u_k(x)| < +\infty \quad \text{pour } \mu\text{-presque tout } x \in X.$$

La série $\sum_{k \geq 0} u_k(x)$ est ainsi absolument convergente, donc convergente, pour μ-presque tout $x \in X$. Soit f la fonction numérique définie presque partout sur X par :

$$f(x) = \sum_{k=0}^{\infty} u_k(x).$$

Comme on a $|f(x)| \leq \sum_{k=0}^{\infty} |u_k(x)|$ pour μ-presque tout $x \in X$, et que $g = \sum_{k \geq 0} |u_k| \in \mathcal{L}^1(X, \mu)$, la fonction f est μ-intégrable. Par ailleurs, comme on a

$$u_0(x) + u_1(x) + \cdots + u_k(x) = f_{n_k}(x)$$

pour μ-presque tout $x \in X$, la suite $(f_{n_k})_k$ converge μ-presque partout vers $f \in \mathcal{L}^1(X, \mu)$.

Enfin, on a pour μ-presque tout $x \in X$:

$$|f(x) - f_{n_k}(x)| = \left| \sum_{n=k+1}^{\infty} u_n(x) \right| \leq \sum_{n=k+1}^{\infty} |u_n(x)|,$$

d'où :

$$\|f - f_{n_k}\|_1 = \int_X |f - f_{n_k}| d\mu \leqq \int_X \left(\sum_{p=k+1}^{\infty} |u_p|\right) d\mu$$
$$= \sum_{p=k+1}^{\infty} \int_X |u_p| d\mu = \sum_{p=k+1}^{\infty} \|f_{n_p} - f_{n_{p-1}}\|_1$$
$$\leqq \sum_{p=k+1}^{\infty} \frac{1}{2^{p-1}} = \frac{1}{2^{k-1}}.$$

Il s'ensuit que $\|f - f_{n_k}\|_1 \longrightarrow 0$ quand $k \to +\infty$, et le théorème est démontré. ■

3.3.6- COROLLAIRE 1 - Soit (X,μ) un espace mesuré. Alors :

(i) $L^1(X,\mu)$ est un espace vectoriel normé complet pour la norme $\|\cdot\|_1$;

(ii) Les (classes de) fonctions simples μ-intégrables forment un sous-espace vectoriel de $L^1(X,\mu)$ qui est partout dense dans $L^1(X,\mu)$ pour la norme $\|\cdot\|_1$-

Démonstration. (i) Soit $(\dot{f}_1, \dot{f}_2, \cdots)$ une suite de Cauchy dans $L^1(X,\mu)$. D'après le théorème 3.3.5, il existe $f \in \mathcal{L}^1(X,\mu)$ et une sous-suite $(f_{n_k})_k$ de $(f_n)_n$ telles que $\|\dot{f}_{n_k} - \dot{f}\|_1 \longrightarrow 0$ quand $k \to +\infty$. Comme la suite $(\dot{f}_n)_n$ est de Cauchy, on en déduit immédiatement que $\|\dot{f}_n - \dot{f}\|_1 \longrightarrow 0$ quand $n \to +\infty$, d'où (i).

(ii). Résulte immédiatement du théorème 3.2.6.2. ■

3.3.7- COROLLAIRE 2 . Soit (X,μ) un espace mesuré. Si $(\dot{f}_1, \dot{f}_2, \cdots)$ est une suite de $L^1(X,\mu)$ qui converge vers $\dot{f} \in L^1(X,\mu)$ pour la norme $\|\cdot\|_1$, il existe une sous-suite $(f_{n_k})_k$ qui converge μ-presque partout vers f.

Démonstration. Comme la suite $(\dot{f}_1, \dot{f}_2, \ldots)$ converge vers $\dot{f} \in L^1(X,\mu)$, elle est de Cauchy pour la norme $\|\cdot\|_1$. D'après le théorème 3.3.5, il existe une sous-suite $(f_{n_k})_k$ et $\dot{g} \in L^1(X,\mu)$ telles que $\|\dot{f}_{n_k} - \dot{g}\|_1 \to 0$ quand $k \to +\infty$ et $f_{n_k}(x) \to g(x)$ pour μ-presque tout $x \in X$. Comme $\|\dot{f}_{n_k} - \dot{f}\|_1 \to 0$ quand $k \to +\infty$, on a $\dot{f} = \dot{g}$ et donc $f(x) = g(x)$ pour μ-presque tout $x \in X$. Mais alors, $f_{n_k}(x) \to g(x) = f(x)$ pour μ-presque tout $x \in X$. ■

Montrons par un exemple que l'on ne peut pas espérer, sous les hypothèses du corollaire 2, que f_n converge vers f μ-presque partout. Soit $\mu = \lambda$ la mesure de Lebesgue sur $\mathbb{R}$ et considérons la suite $(f_n)_n$ des fonctions caractéristiques des intervalles $[0,1]$, $[0,\frac{1}{2}]$, $[\frac{1}{2},1]$, $[0,\frac{1}{3}]$, $[\frac{1}{3},\frac{2}{3}]$, $[\frac{2}{3},1]$, ... On a $\|f_n\|_1 \to 0$ quand $n \to +\infty$, mais $f_n(x)$ ne tend vers 0 en aucun point $x \in [0,1]$. Le corollaire 2 est ici facile à vérifier, car on peut prendre pour suite extraite la suite des fonctions caractéristiques des intervalles $[0,\frac{1}{n}]$.

3.3.8. Corollaire 3. Soit (X,μ) un espace mesuré. Soit $(f_n)_{n\geq 1}$ une suite de fonctions numériques μ-intégrables telles que $\sum_{n\geq 1} \int_X |f_n| d\mu < +\infty$. Alors, la série $\sum_{n\geq 1} f_n(x)$ est convergente pour μ-presque tout $x \in X$, sa somme $f(x)$ (définie μ-presque partout) est μ-intégrable et on a :

$$\|f - \sum_{n=1}^{N} f_n\|_1 \to 0 \text{ quand } N \to +\infty,$$

i.e. $f = \sum_{n\geq 1} f_n$ dans $L^1(X,\mu)$. En outre, on a :

$$\int_X f \, d\mu = \sum_{n\geq 1} \int_X f_n \, d\mu.$$

Démonstration. Comme $L^1(X,\mu)$ est un espace vectoriel normé complet, la condition $\sum_{n\geq 1} \int_X |f_n| d\mu < +\infty$, qui s'écrit $\sum_{n\geq 1} \|f_n\|_1 < +\infty$, implique que la série $\sum_{n\geq 1} f_n$ converge dans $L^1(X,\mu)$. Il existe donc $f\in L^1(X,\mu)$ telle que $\|f-\sum_{n=1}^N f_n\|_1 \longrightarrow 0$ quand $N\to+\infty$, et la continuité de la forme linéaire $f\longmapsto \int_X f d\mu$ sur $L^1(X,\mu)$ implique que

$$\int_X f d\mu = \lim_{N\to\infty} \int_X \sum_{n=1}^N f_n d\mu = \lim_{N\to\infty} \sum_{n=1}^N \int_X f_n d\mu$$
$$= \sum_{n\geq 1} \int_X f_n d\mu .$$

Par ailleurs, comme les fonctions $S_N = \sum_{n=1}^N f_n$ convergent vers f dans $L^1(X,\mu)$, il résulte du théorème 3.3.5 que, pour toute suite extraite $(S_{N_k})_k$, il existe une suite extraite $(S_{N_{k(\ell)}})_\ell$ qui converge μ-presque partout vers une fonction μ-intégrable g et telle que $\|S_{N_{k(\ell)}} - g\|_1 \longrightarrow 0$ quand $\ell\to+\infty$.
Comme $\|S_{N_{k(\ell)}} - f\|_1 \longrightarrow 0$ quand $\ell\to+\infty$, on a $f=g$ μ-presque partout, et par conséquent

$$S_{N_{k(\ell)}}(x) \longrightarrow f(x) \text{ pour } \mu\text{-presque tout } x\in X.$$

Comme la limite $f(x)$ est indépendante du choix de la suite extraite $(S_{N_k})_k$, on en déduit que

$$S_N(x) \longrightarrow f(x) \text{ quand } N\to+\infty \text{ pour}$$

μ-presque tout $x\in X$, et la série $\sum_{n\geq 1} f_n(x)$ est convergente de somme $f(x)$ pour μ-presque tout $x\in X$. On aurait pu éviter ce raisonnement utilisant une double extraction de sous-suites en remarquant que :

$$\int_X \Big(\sum_{n\geq 1} |f_n|\Big) d\mu = \sum_{n\geq 1} \int_X |f_n| d\mu < +\infty ,$$

de sorte que $\sum_{n\geq 1} |f_n(x)| < +\infty$ pour μ-presque tout

$x \in X$, ce qui implique que la série $\sum_{n \geq 1} f_n(x)$ converge vers $g(x)$ pour μ-presque tout $x \in X$. En utilisant le théorème 3.3.5, on montre alors comme ci-dessus que g est intégrable, et égale presque partout à f. ∎

3.4. THÉORÈMES DE CONVERGENCE

3.4.1. Soit $(f_1, f_2, \dots)$ une suite de fonctions μ-intégrables sur un espace mesuré (X, μ). On a vu (cf. 3.3.7) que, s'il existe une fonction $f \in \mathcal{L}^1(X, \mu)$ telle que

$$\| f - f_n \|_1 \longrightarrow 0 \quad \text{quand } n \to +\infty,$$

alors il existe une suite extraite $(f_{n_k})_k$ qui converge μ-presque partout vers f. Dans ce paragraphe, on étudie dans quelles conditions la limite μ-presque partout d'une suite de fonctions μ-intégrables est elle-même μ-intégrable.

En général, la limite μ-presque partout d'une suite de fonctions μ-intégrables n'est pas μ-intégrable. Par exemple, sur $\mathbb{R}$ muni de la mesure de Lebesgue $\mu = \lambda$, la suite des fonctions $f_n = \mathbb{1}_{[-n, n]} \in \mathcal{L}^1(\mathbb{R}, \lambda)$ converge partout vers la fonction $f(x) = 1$, qui n'est pas intégrable.

Par ailleurs, si $f_n(x) \to f(x)$ pour μ-presque tout $x \in X$, et si les f_n et f sont μ-intégrables, on aimerait savoir si $\int_X f \, d\mu = \lim_{n \to \infty} \int_X f_n \, d\mu$. En général, ce n'est pas le cas. Par exemple, la suite de fonctions λ-intégrables $f_n = n \, \mathbb{1}_{[0, \frac{1}{n}]}$ converge λ-presque partout sur $\mathbb{R}$ vers la fonction λ-intégrable $f = 0$, mais

$$\int_{\mathbb{R}} f_n \, d\lambda = 1 \underset{n \to +\infty}{\not\longrightarrow} \int_{\mathbb{R}} f \, d\lambda = 0.$$

Il nous faut donc imposer des conditions supplémentaires pour assurer que, si $f_n(x) \to f(x)$ pour μ-presque tout $x \in X$ et si $f_n \in \mathcal{L}^1(X, \mu)$, alors $f \in \mathcal{L}^1(X, \mu)$ et

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

3.4.2 - Théorème (Beppo-Lévi) - Soit (X, μ) un espace mesuré. Soit $(f_1, f_2, \dots)$ une suite croissante (resp. décroissante) de fonctions μ-intégrables, et supposons que les nombres $\int_X f_n d\mu$ sont bornés supérieurement (resp. inférieurement). Alors, $f = \lim f_n$ est μ-intégrable et $\| f_n - f \|_1 \to 0$ quand $n \to +\infty$. En particulier, on a :

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

Démonstration. Remarquons tout d'abord que, pour tout $x \in X$, $\lim_{n \to \infty} f_n(x)$ (éventuellement égal à $\pm\infty$) existe bien puisque la suite $(f_n(x))_n$ est monotone. Quitte à remplacer f_n par $-f_n$, on peut se limiter à traiter le cas où la suite $(f_n)_n$ est croissante. Dans ce cas, on a pour $n, p \to +\infty$, $n > p$:

$$\int_X |f_n - f_p| \, d\mu = \int_X (f_n - f_p) \, d\mu = \int_X f_n d\mu - \int_X f_p d\mu \to 0,$$

car la suite des nombres $\int_X f_n d\mu$ est croissante et majorée. La conclusion résulte alors du théorème 3.3.5. ■

¶ 3.4.3 - Théorème (de convergence dominée de Lebesgue) - Soit (X, μ) un espace mesuré. Soit $(f_1, f_2, \dots)$ une suite de fonctions μ-intégrables qui converge μ-presque partout vers une fonction f. Supposons qu'il existe une fonction μ-intégrable fixe g telle que :

$$|f_n| \leq g \quad \mu\text{-presque partout}.$$

Alors, f est μ-intégrable et $\| f_n - f \|_1 \to 0$ quand $n \to +\infty$. En particulier, on a :

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

Démonstration. Tout d'abord, comme les f_n sont μ-mesurables et convergent μ-presque partout vers f, la fonction f est μ-mesurable. Puisque

$$|f_n(x)| \leqq g(x) \quad \text{pour } \mu\text{-presque tout } x,$$

on en déduit que

$$|f(x)| \leqq g(x) \quad \text{pour } \mu\text{-presque tout } x.$$

Comme g est μ-intégrable, on a :

$$\int_X |f|\,d\mu \leqq \int_X g\,d\mu < +\infty,$$

et par conséquent f est μ-intégrable. Montrons que $\|f_n - f\|_1 \longrightarrow 0$ quand $n \to +\infty$. A cet effet, posons pour tout $k \geqq 1$:

$$F_k = \sup_{i,j \geqq k} |f_i - f_j|$$

On définit ainsi une fonction μ-mesurable positive, qui est μ-intégrable car

$$|f_i - f_j| \leqq |f_i| + |f_j| \leqq 2g \quad \text{pour tout couple } (i,j),$$

d'où $F_k \leqq 2g$. Comme g est μ-intégrable, la fonction F_k est μ-intégrable. La suite $(F_k)_k$ est une suite décroissante de fonctions positives μ-intégrables qui converge vers 0 μ-presque partout. En effet, comme $f_i, f_j \longrightarrow f$ μ-presque partout, $f_i - f_j \longrightarrow 0$ μ-presque partout quand $i,j \to +\infty$, d'où il résulte que $F_k \longrightarrow 0$ μ-presque partout. D'après le théorème de Beppo-Levi, $\int_X F_k\,d\mu \longrightarrow 0$ quand $k \to +\infty$. De la relation :

$$\int_X |f_{k+l} - f_k|\,d\mu \leqq \int_X F_k\,d\mu \xrightarrow[k \to +\infty]{} 0,$$

on déduit que la suite $(f_k)_k$ est de Cauchy dans $L^1(X,\mu)$, donc converge dans $L^1(X,\mu)$ vers une fonction $g \in L^1(X,\mu)$. D'après 3.3.7, il existe une sous-suite $(f_{k_l})_l$ qui converge

μ-presque partout vers g et, comme $f_k \to f$ μ-presque partout, on en déduit que $f = g$ μ-presque partout. Mais alors, $\|f_n - f\|_1 = \|f_n - g\|_1 \to 0$ quand $n \to +\infty$, ce qui achève la démonstration du théorème de convergence dominée de Lebesgue. ■

3.4.4- Corollaire. Soit (X, μ) un espace mesuré. Soit $(u_1, u_2, \ldots)$ une suite de fonctions μ-intégrables. On suppose que la série de fonctions $\sum_{k=1}^{\infty} u_k$ converge μ-presque partout, et que les fonctions $|\sum_{k=1}^{n} u_k|$ sont majorées par une fonction μ-intégrable indépendante de n. Alors, $\sum_{k=1}^{\infty} u_k$ est μ-intégrable et on a :

$$\int_X \left(\sum_{k=1}^{\infty} u_k\right) d\mu = \sum_{k=1}^{\infty} \int_X u_k \, d\mu.$$

Démonstration. Il suffit d'appliquer le théorème de convergence dominée de Lebesgue à la suite de fonctions μ-intégrables

$$f_n = \sum_{k=1}^{n} u_k$$

qui converge μ-presque partout vers $f = \sum_{k=1}^{\infty} u_k$ et qui sont majorées en module par une fonction μ-intégrable fixe. ■

Mentionnons, pour terminer ce paragraphe, un résultat de convergence qui est parfois utile :

3.4.5- Proposition. Soit (X, μ) un espace mesuré. Soit $(f_n)_n$ une suite de fonctions μ-intégrables qui converge μ-presque partout vers une fonction f. Supposons qu'il existe un réel $C \geq 0$ tel que :

$$\int_X |f_n| \, d\mu \leqq C \quad \text{pour tout } n.$$

Alors, f est μ-intégrable.

Démonstration. Tout d'abord, f est μ-mesurable comme limite μ-presque partout d'une suite de fonctions μ-mesurables (car μ-intégrables). Posons :

$$g_n = \inf_{k \geq 0} (|f_{n+k}|).$$

On définit ainsi une suite croissante de fonctions positives μ-mesurables, et qui sont intégrables puisque

$$\int_X g_n d\mu \leq \int_X |f_n| d\mu < +\infty.$$

Par ailleurs, comme $|f_k| \longrightarrow |f|$ μ-presque partout, on a $\operatorname{Sup}_n g_n = |f|$ μ-presque partout et donc

$$\int_X |f| d\mu = \operatorname{Sup}_n \int_X g_n d\mu \leqq \operatorname{Sup}_n \int_X |f_n| d\mu \leqq C.$$

Il s'ensuit que f est μ-intégrable. ■

3.5- INTÉGRALE DE RIEMANN ET INTÉGRALE DE LEBESGUE

3.5.1- Soit $f : [a,b] \longrightarrow \mathbb{R}$ une fonction réelle définie sur un intervalle fermé et borné $[a,b]$ de $\mathbb{R}$. Si f est bornée sur $[a,b]$, on associe à toute subdivision

$$\sigma : a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$$

du segment $[a,b]$ les sommes de Darboux inférieure $s(\sigma)$ et supérieure $S(\sigma)$ définies par :

$$s(\sigma) = \sum_{i=0}^{n-1} m_i (x_{i+1} - x_i), \quad S(\sigma) = \sum_{i=0}^{n-1} M_i (x_{i+1} - x_i),$$

où $m_i = \inf\{ f(x) \mid x \in [x_i, x_{i+1}]\}$ et
$M_i = \operatorname{Sup}\{ f(x) \mid x \in [x_i, x_{i+1}]\}$.

Rappelons que f est dite Riemann intégrable sur $[a,b]$ si elle est bornée et si les sommes $s(\sigma)$ et $S(\sigma)$ tendent vers une même limite I quand le pas $\delta(\sigma) = \operatorname{Max}_{0 \leq i < n} (x_{i+1} - x_i)$ de σ tend vers 0. Dans ce cas, la limite commune I est appelée l'intégrale de Riemann de f et notée

$$I = \int_a^b f(x) dx.$$

On a :

3.5.2 - Proposition. Soit $f : [a,b] \to \mathbb{R}$ une fonction. Les conditions suivantes sont équivalentes :

(i) f est Riemann intégrable sur $[a,b]$;

(ii) f est bornée sur $[a,b]$ et l'ensemble de ses points de discontinuité est négligeable pour la mesure de Lebesgue.

Nous admettons ce résultat, qui n'a qu'un intérêt purement historique. La plupart des fonctions Riemann intégrables rencontrées en analyse sont des fonctions réglées sur $[a,b]$, c'est à dire des fonctions $f : [a,b] \to \mathbb{R}$ qui admettent en tout point une limite à droite et une limite à gauche. Par exemple, les fonctions continues par morceaux ou les fonctions monotones sont réglées. On peut montrer qu'une fonction $f : [a,b] \to \mathbb{R}$ est réglée sur $[a,b]$ si et seulement si elle est limite uniforme sur $[a,b]$ de fonctions en escaliers. Ceci implique qu'une fonction réglée est bornée et que l'ensemble de ses points de discontinuité est au plus dénombrable. Une fonction réglée est donc Riemann-intégrable en vertu de la proposition 3.5.2.

3.5.3 - Théorème. Soit $f : [a,b] \to \mathbb{R}$ une fonction Riemann intégrable sur $[a,b]$. Alors, f est Lebesgue intégrable sur $[a,b]$ et son intégrale de Lebesgue coïncide avec son intégrale de Riemann :

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

Démonstration. Notons tout d'abord que, pour toute subdivision σ :

$$a = x_0 < x_1 < \dots < x_n = b$$

de $[a,b]$, les fonctions

$$\varphi_\sigma = \sum_{i=0}^{n-1} m_i \mathbb{1}_{J_i} \quad \text{et} \quad \psi_\sigma = \sum_{i=0}^{n-1} M_i \mathbb{1}_{J_i}$$

où $J_i = [x_i, x_{i+1}[$ pour $i = 0, 1, \dots, n-2$ et $J_{n-1} = [x_{n-1}, x_n]$,

$m_i = \mathrm{Inf}\{f(x) \mid x \in [x_i, x_{i+1}]\}$
$M_i = \mathrm{Sup}\{f(x) \mid x \in [x_i, x_{i+1}]\}$

sont des fonctions simples Lebesgue-intégrables qui vérifient

$$\int_{[a,b]} \varphi_\sigma d\lambda = s(\sigma) \; ; \; \int_{[a,b]} \psi_\sigma d\lambda = S(\sigma).$$

Pour tout entier $n \geqq 1$, notons σ_n la subdivision de $[a,b]$ en 2^n segments de longueur égale à $\frac{b-a}{2^n}$. Pour simplifier, notons $\varphi_n = \varphi_{\sigma_n}$, $\psi_n = \psi_{\sigma_n}$. Comme les subdivisions σ_n sont emboîtées, on a :

$$\varphi_n \leqq \varphi_{n+1} , \quad \psi_{n+1} \leqq \psi_n \quad \text{et} \quad \varphi_n \leqq f \leqq \psi_n .$$

Comme f est Riemann-intégrable sur $[a,b]$, on a :

$$s(\sigma_n), \; S(\sigma_n) \longrightarrow I = \int_a^b f(x)dx$$

quand $n \to +\infty$, et par conséquent :

$$\int_{[a,b]} \varphi_n d\lambda = s(\sigma_n) \leqq I, \qquad I \leqq S(\sigma_n) = \int_{[a,b]} \psi_n d\lambda .$$

Posons $\varphi = \mathrm{Sup}_n \varphi_n$, $\psi = \mathrm{Inf}_n \psi_n$. On a :

$$\varphi \leqq f \leqq \psi$$

et, en vertu du théorème de Beppo-Levi,

$$\int_{[a,b]} \varphi d\lambda = \lim_{n \to \infty} \int_{[a,b]} \varphi_n d\lambda = \lim_{n\to\infty} s(\sigma_n) = I$$
$$= \lim_{n\to\infty} S(\sigma_n) = \lim_{n\to\infty} \int_{[a,b]} \psi_n d\lambda = \int_{[a,b]} \psi d\lambda .$$

On en déduit que $\int_{[a,b]} (\psi - \varphi) d\lambda = 0$, d'où $\varphi = \psi$ presque partout, et donc $f = \varphi = \psi$ presque partout. Il s'ensuit que f est Lebesgue-intégrable (comme φ et ψ) et que

$$\int_{[a,b]} f\,d\lambda = \int_{[a,b]} \varphi\,d\lambda = I = \int_a^b f(x)\,dx,$$

ce qui démontre le théorème 3.5.3. ■

3.5.4. Intégrales généralisées. Soit $I = (\alpha, \beta)$ un intervalle non compact de $\mathbb{R}$ (soit I n'est pas borné, soit I est borné et non fermé). Soit $f : I \longrightarrow \mathbb{R}$ une fonction et supposons que la restriction de f à tout intervalle compact $[a,b]$ de I soit Riemann-intégrable. Lorsque la limite

$$\lim_{a \searrow \alpha,\, b \nearrow \beta} \int_a^b f(x)\,dx$$

existe, on dit que l'intégrale $\int_I f(x)\,dx$ est convergente et on pose :

$$\int_I f(x)\,dx = \lim_{a \searrow \alpha,\, b \nearrow \beta} \int_a^b f(x)\,dx.$$

Dans le cas contraire, on dit que l'intégrale $\int_I f(x)\,dx$ est divergente.

Si $\int_I |f(x)|\,dx$ est convergente, on dit que l'intégrale $\int_I f(x)\,dx$ est absolument convergente. Dans ce cas, le critère de Cauchy pour les intégrales montre que $\int_I f(x)\,dx$ est convergente. On prendra garde au fait que l'intégrale $\int_I f(x)\,dx$ peut être convergente sans être absolument convergente. Par exemple, l'intégrale $\int_0^{+\infty} \frac{\sin x}{x}\,dx$ est convergente (et de valeur $\frac{\pi}{2}$) mais pas absolument convergente.

Le théorème suivant fait le lien entre convergence absolue de l'intégrale $\int_I f(x)\,dx$ et Lebesgue-intégrabilité de f sur I :

3.5.5- Théorème. Soit I un intervalle non compact de $\mathbb{R}$. Pour toute fonction $f : I \to \mathbb{R}$ dont la restriction à tout intervalle compact $[a,b] \subset I$ est Riemann-intégrable, les conditions suivantes sont équivalentes :

(i) f est Lebesgue-intégrable sur I ;

(ii) L'intégrale $\int_I |f(x)|dx$ est convergente.

Lorsque l'une de ces conditions est réalisée, on a :

$$\int_I f(x)dx = \int_I f d\lambda .$$

Démonstration. (i) $\Rightarrow$ (ii). Si f est Lebesgue-intégrable sur I, alors $|f|$ est aussi Lebesgue-intégrable sur I et on a, pour tout intervalle compact $[a,b]$ inclus dans I :

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

d'où il résulte que $\int_I |f(x)| dx < +\infty$.

(ii) $\Rightarrow$ (i). Posons $I = (\alpha, \beta)$ et choisissons des suites $(a_n)_n$, $(b_n)_n$ de points de I tels que :

$$a_n \leqq b_n \;;\; a_n \searrow \alpha \text{ et } b_n \nearrow \beta .$$

Posons $f_n = f \mathbb{1}_{[a_n, b_n]}$. Comme f est Riemann-intégrable sur $[a_n, b_n]$, elle est Lebesgue-intégrable sur $[a_n, b_n]$ et f_n est Lebesgue-intégrable sur I. Les fonctions $|f_n|$ forment une suite croissante de fonctions Lebesgue-intégrables sur I qui converge simplement vers $|f|$.
En outre,

$$\int_I |f_n| d\lambda = \int_{[a_n, b_n]} |f| d\lambda = \int_{a_n}^{b_n} |f(x)| dx \leqq \int_I |f(x)| dx < +\infty ,$$

et le théorème de convergence monotone prouve que $|f|$ est intégrable au sens de Lebesgue sur I. Comme on a :

$$|f_n| \leqq |f| ,$$

le théorème de convergence dominée de Lebesgue implique que f est lebesgue-intégrable sur I et que

$$\int_I f d\lambda = \lim_{n \to +\infty} \int_I f_n d\lambda = \lim_{n\to+\infty} \int_{a_n}^{b_n} f(x)dx = \int_I f(x)dx.$$

■

3.5.6. Exemple. La fonction $\frac{1}{x^\alpha}$ est intégrable sur $[a,+\infty[$ (où $a>0$) si et seulement si $\alpha > 1$.

En effet, la relation

$$\int_a^n \frac{dx}{x^\alpha} = \begin{cases} \frac{1}{1-\alpha}\left(\frac{1}{n^{\alpha-1}} - \frac{1}{a^{\alpha-1}}\right) & \text{si } \alpha \neq 1 \\ \log n - \log a & \text{si } \alpha = 1 \end{cases}$$

montre que $\int_a^{+\infty} \frac{dx}{x^\alpha}$ est convergente si et seulement si $\alpha > 1$.

Il s'ensuit que $\frac{1}{x^\alpha}$ est lebesgue intégrable sur $[a,+\infty[$ si et seulement si $\alpha > 1$. Dans ce cas, on a :

$$\int_{[a,+\infty[} \frac{1}{x^\alpha} d\lambda(x) = \frac{1}{(\alpha-1)\, a^{\alpha-1}}.$$

On démontrerai de même que $\frac{1}{x^\alpha}$ est lebesgue-intégrable sur $[0,a]$ (avec $a>0$) si et seulement si $\alpha < 1$.

3.6. FONCTIONS DÉFINIES PAR DES INTÉGRALES

3.6.1. Soient (X,μ) un espace mesuré et Ω un espace métrique. On suppose donnée une fonction numérique

$$f : X \times \Omega \longrightarrow \overline{\mathbb{R}},$$

et on désigne par f_λ et f_x les applications partielles $x \longmapsto f_\lambda(x) = f(x,\lambda)$ et $\lambda \longmapsto f_x(\lambda) = f(x,\lambda)$.

Nous supposerons dans tout ce paragraphe que, pour tout $\lambda \in \Omega$, la fonction f_λ est μ-intégrable.

On définit alors une fonction $F : \Omega \to \mathbb{R}$ en posant :

$$F(\lambda) = \int_X f_\lambda(x)\, d\mu(x) = \int_X f(x,\lambda)\, d\mu(x).$$

On dit que F est une fonction définie par une intégrale. Dans ce qui suit, nous nous intéresserons à la continuité de F et (lorsque Ω est un ouvert de $\mathbb{R}^n$) à sa différentiabilité.

3.6.2. Théorème. Soient (X,μ) un espace mesuré, Ω un espace métrique et $f : X \times \Omega \to \mathbb{R}$ une fonction numérique telle que f_λ soit μ-intégrable pour tout $\lambda \in \Omega$. Supposons que les conditions suivantes soient réalisées :

(i) Pour μ-presque tout $x \in X$, la fonction f_x est continue de la variable λ au point $\lambda_0 \in \Omega$;

(ii) Il existe un voisinage V de λ_0 dans Ω et une fonction μ-intégrable fixe g telle que :

$$|f(x,\lambda)| \leq g(x) \quad \text{pour tout } \lambda \in V.$$

Alors, $F(\lambda) = \int f(x,\lambda)\, d\mu(x)$ est une fonction continue au point $\lambda = \lambda_0$.

Démonstration. Comme Ω est un espace métrique, il suffit de montrer que $F(\lambda_n) \to F(\lambda_0)$ pour toute suite $(\lambda_n)_n$ qui converge vers λ_0. Posons

$$f_n(x) = f(x,\lambda_n).$$

Pour μ-presque tout x, la fonction f_x est continue au point λ_0 et donc $f_n(x) = f(x,\lambda_n) = f_x(\lambda_n) \to f_x(\lambda_0) = f(x,\lambda_0)$ quand $n \to +\infty$. Par ailleurs, pour n assez grand, on a $\lambda_n \in V$ et donc :

$$|f_n(x)| = |f(x,\lambda_n)| \leq g(x).$$

D'après le théorème de convergence dominée de Lebesgue, on a :

$$F(\lambda_n) = \int_X f_n(x)\, d\mu(x) \longrightarrow \int_X f(x,\lambda_0)\, d\mu(x) = F(\lambda_0)$$

quand $n \to +\infty$, d'où le théorème. ∎

3.6.3. Théorème (dérivation sous le signe d'intégration).
Soient (X,μ) un espace mesuré et Ω un ouvert de $\mathbb{R}^n$ dont on note $\lambda = (\lambda_1, \dots, \lambda_n)$ le point courant. Soit $f : X \times \Omega \to \mathbb{R}$ une fonction numérique telle que f_λ soit μ-intégrable pour tout $\lambda \in \Omega$. On suppose que les conditions suivantes sont réalisées :

(i) Pour μ-presque tout $x \in X$, f_x possède une dérivée partielle par rapport à λ_j et $\lambda \mapsto \frac{\partial f}{\partial \lambda_j}(x,\lambda)$ est continue sur Ω ;

(ii) Il existe une fonction g_j μ-intégrable telle que l'on ait :

$$\left|\frac{\partial f}{\partial \lambda_j}(x,\lambda)\right| \leq g_j(x) \text{ pour tout } \lambda \in \Omega.$$

Alors, la fonction $F(\lambda) = \int_X f(x,\lambda)\,d\mu(x)$ possède une dérivée partielle par rapport à λ_j et l'application $\lambda \mapsto \frac{\partial F}{\partial \lambda_j}(\lambda)$ est continue sur Ω.

En particulier, si les conditions (i) et (ii) sont réalisées pour tout $j = 1, 2, \dots, n$, la fonction F est différentiable sur Ω.

Démonstration. Pour tout $\varepsilon > 0$, on a :

$$\frac{F(\lambda_1, \dots, \lambda_j + \varepsilon, \dots, \lambda_n) - F(\lambda)}{\varepsilon} = \int_X \frac{f(x; \lambda_1, \dots, \lambda_j + \varepsilon, \dots, \lambda_n) - f(x,\lambda)}{\varepsilon}\, d\mu(x)$$

D'après le théorème des accroissements finis (applicable pour μ-presque tout $x \in X$), il existe un réel θ_ε, $|\theta_\varepsilon| < \varepsilon$, tel que :

$$\frac{f(x; \lambda_1, \dots, \lambda_j + \varepsilon, \dots, \lambda_n) - f(x,\lambda)}{\varepsilon} = \frac{\partial f}{\partial \lambda_j}(x; \lambda_1, \dots, \lambda_j + \theta_\varepsilon, \dots, \lambda_n)$$

Si ε parcourt une suite $\varepsilon_p \to 0$, la fonction

$$h_p(x) = \frac{\partial f}{\partial \lambda_j}(x; \lambda_1, \dots, \lambda_j + \theta_{\varepsilon_p}, \dots, \lambda_n)$$

converge μ-presque partout vers $h(x) = \frac{\partial f}{\partial \lambda_j}(x,\lambda)$ car l'application $\lambda \mapsto \frac{\partial f}{\partial \lambda_j}(x,\lambda)$ est continue pour μ-presque tout

$x \in X$. Par ailleurs, on a :

$$|h_p(x)| = \left|\frac{\partial f}{\partial \lambda_j}(x; \lambda_1, \dots, \lambda_j + \theta_{\varepsilon_p}, \dots, \lambda_n)\right| \leq g_j(x)$$

pour μ-presque tout $x \in X$, où g_j est une fonction μ-intégrable fixe. D'après le théorème de convergence dominée de Lebesgue, la fonction $x \longmapsto \frac{\partial f}{\partial \lambda_j}(x, \lambda)$ (qui est définie pour μ-presque tout x) est μ-intégrable et on a :

$$\frac{F(\lambda_1, \dots, \lambda_j + \varepsilon_p, \dots, \lambda_n) - F(\lambda)}{\varepsilon_p} = \int_X h_p \, d\mu \xrightarrow[p \to \infty]{} \int_X \frac{\partial f}{\partial \lambda_j}(x, \lambda) \, d\mu(x),$$

ce qui prouve que F possède une dérivée partielle par rapport à λ_j en tout point $\lambda \in \Omega$ et que l'on a :

$$\frac{\partial F}{\partial \lambda_j}(\lambda) = \int_X \frac{\partial f}{\partial \lambda_j}(x, \lambda) \, d\mu(x).$$

La continuité de l'application $\lambda \longmapsto \frac{\partial F}{\partial \lambda_j}(\lambda)$ sur Ω résulte alors du théorème 3.6.2. ■

3.6.4. APPLICATION AU CALCUL DE $\int_0^\infty e^{-t^2} dt$. D'après le théorème 3.5.5, la fonction $t \longmapsto e^{-t^2}$ est intégrable au sens de Lebesgue sur $[0, +\infty[$. Posons

$$I = \int_0^\infty e^{-t^2} dt$$

et proposons-nous de calculer la valeur de I. A cet effet, considérons la fonction définie pour $x \geq 0$ par :

$$f(x) = \int_0^\infty \frac{e^{-xt^2}}{1+t^2} dt.$$

Comme on a $\left|\frac{e^{-xt^2}}{1+t^2}\right| \leq \frac{1}{1+t^2}$ et que la fonction $t \mapsto \frac{1}{1+t^2}$ est Lebesgue intégrable sur $[0, +\infty[$, la fonction $f(x)$ est bien définie, et elle est continue en vertu du théorème 3.6.2. Notons que l'on a : $f(0) = \int_0^\infty \frac{dt}{1+t^2} = \frac{\pi}{2}$.

En outre, puisque $\dfrac{e^{-xt^2}}{1+t^2} \longrightarrow 0$ quand $x \to +\infty$ pour tout $t>0$, il résulte du théorème de convergence dominée de Lebesgue que $\lim\limits_{x \to +\infty} f(x) = 0$. Pour $x>0$, on a : $\dfrac{\partial}{\partial x}\left(\dfrac{e^{-xt^2}}{1+t^2}\right) = \dfrac{-t^2 e^{-xt^2}}{1+t^2}$. En outre, pour $x \geqq a > 0$, on a :

$$\left|\frac{\partial}{\partial x}\left(\frac{e^{-xt^2}}{1+t^2}\right)\right| = \frac{t^2}{1+t^2} e^{-xt^2} \leqq e^{-at^2}$$

et, comme la fonction $t \longmapsto e^{-at^2}$ est intégrable sur $[0,+\infty[$, il résulte du théorème de dérivation sous le signe d'intégration que f est dérivable sur tout intervalle $]a,+\infty[$ avec $a>0$, donc est dérivable sur $]0,+\infty[$, de dérivée :

$$f'(x) = -\int_0^{+\infty} \frac{t^2}{1+t^2} e^{-xt^2} dt.$$

On a donc :

$$f(x) - f'(x) = \int_0^{+\infty} e^{-xt^2} dt \text{ pour } x>0.$$

Posons $u = \sqrt{x}\, t$ dans la dernière intégrale ; on obtient :

$$\int_0^{+\infty} e^{-xt^2} dt = \frac{1}{\sqrt{x}} \int_0^{+\infty} e^{-u^2} du = \frac{I}{\sqrt{x}}.$$

La fonction f est donc solution de l'équation différentielle

$$f - f' = \frac{I}{\sqrt{x}}.$$

On a donc

$$f(x) = C(x) e^{x}, \quad \text{où la fonction } C$$

vérifie : $-C'(x) e^{x} = \dfrac{I}{\sqrt{x}}$, et donc $C'(x) = -I \dfrac{e^{-x}}{\sqrt{x}}$.

On en déduit que $C(x) = -I \displaystyle\int \frac{e^{-x}}{\sqrt{x}} dx = -I \int \frac{e^{-v^2}}{v} 2v\, dv \quad (v^2 = x)$, soit $\quad C(x) = C - 2I\int_0^{\sqrt{x}} e^{-t^2}dt$, où C est une constante réelle.

Il s'ensuit que :

$$f(x) = \left(C - 2I\int_0^{\sqrt{x}} e^{-t^2}dt\right)e^{x} \quad \text{pour } x > 0,$$

et cette formule reste vraie par continuité pour $x=0$. Puisque $f(0) = \frac{\pi}{2}$, on a $C = \frac{\pi}{2}$ et donc, au total :

$$f(x) = \left(\frac{\pi}{2} - 2I\int_0^{\sqrt{x}} e^{-t^2}dt\right)e^{x}.$$

Puisque $f(x) \longrightarrow 0$ quand $x \to +\infty$, on a nécessairement

$$\frac{\pi}{2} - 2I\int_0^{\infty} e^{-t^2}dt = 0,$$

soit $\quad \frac{\pi}{2} - 2I^2 = 0$, d'où l'on tire :

$$\boxed{I = \frac{\sqrt{\pi}}{2}}$$

On peut retrouver rapidement ce résultat en remarquant que :

$$I^2 = \left(\int_0^{+\infty} e^{-x^2}dx\right)\left(\int_0^{+\infty} e^{-y^2}dy\right) = \iint_D e^{-(x^2+y^2)}dxdy$$

où D est le domaine de $\mathbb{R}^2$ défini par :

$$D = \{(x,y) \in \mathbb{R}^2 \mid x \geq 0, y \geq 0\}.$$

En utilisant les coordonnées polaires, on obtient :

$$I^2 = \int_0^{+\infty}\int_0^{\pi/2} e^{-\rho^2}\rho\, d\rho\, d\theta = \frac{\pi}{2}\int_0^{+\infty} e^{-\rho^2}\rho\, d\rho = \frac{\pi}{2}\int_0^{+\infty}\frac{e^{-u}}{2}du$$

$$= \frac{\pi}{4},$$

d'où $\quad I = \frac{\sqrt{\pi}}{2}$. Les règles de calcul des intégrales doubles utilisées ici ne ~~seront~~ cependant pleinement justifiées qu'au chapitre suivant.

4. INTÉGRATION SUR LES ESPACES PRODUITS

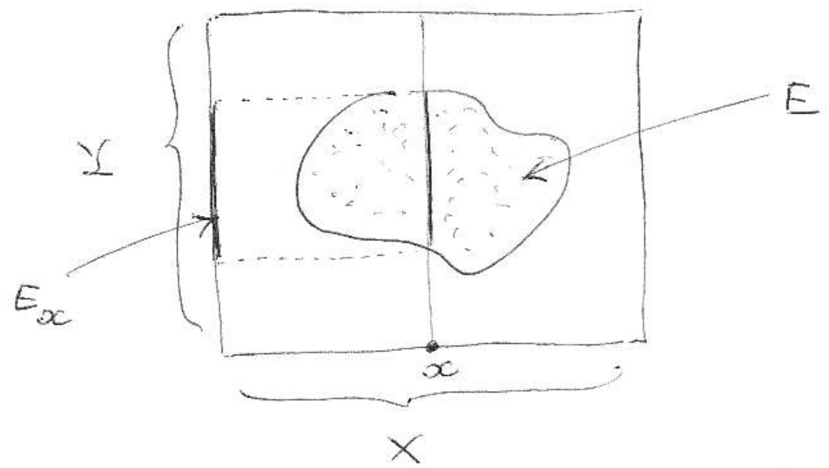
4. INTÉGRATION SUR LES ESPACES PRODUITS

4.1. MESURE PRODUIT

4.1.1. Soient $(X, \mathcal{M}, \mu)$ et $(Y, \mathcal{N}, \nu)$ deux espaces mesurés σ-finis. Un sous-ensemble de $X \times Y$ de la forme $A \times B$ avec $A \in \mathcal{M}$ et $B \in \mathcal{N}$ sera appelé un rectangle mesurable. On désignera par $\mathcal{M} \otimes \mathcal{N}$ la tribu sur $X \times Y$ engendrée par les rectangles mesurables.

4.1.2. Pour toute partie E de $X \times Y$ et tout $x \in X$, on pose :

$$E_x = \{ y \in Y \mid (x,y) \in E \}.$$



On dit que E_x est la coupe de E selon $x \in X$. Pour tout $y \in Y$, on définit la coupe de E selon y par :

$$E_y = \{ x \in X \mid (x,y) \in E \}.$$

Par exemple, si $E = A \times B$ où $A \subset X$ et $B \subset Y$, on a pour tout $x \in X$:

$$E_x = \begin{cases} \emptyset & \text{si } x \notin A \\ B & \text{si } x \in A. \end{cases}$$

4.1.3. Soit $f : X \times Y \longrightarrow Z$ une application du produit $X \times Y$ dans un ensemble Z. Pour $x \in X$, on définit l'application partielle $f_x : Y \longrightarrow Z$ par :

$$f_x(y) = f(x,y).$$

De manière analogue, on définit pour tout $y \in Y$ une application partielle $f_y : X \to Z$ par :

$$f_y(x) = f(x, y).$$

Une propriété importante de la tribu $\mathcal{M} \otimes \mathcal{N}$ est d'assurer la mesurabilité des coupes et des applications partielles. Plus précisément, on a :

4.1.4- Proposition. Soient $(X, \mathcal{M}, \mu)$ et $(Y, \mathcal{N}, \nu)$ deux espaces mesurés σ-finis. Alors, on a :

(i) Soit $E \in \mathcal{M} \otimes \mathcal{N}$. Pour tout $x \in X$, on a : $E_x \in \mathcal{N}$;

(ii) Soit $f : X \times Y \to Z$ une application mesurable de $X \times Y$ muni de la tribu $\mathcal{M} \otimes \mathcal{N}$ dans un espace topologique Z. Pour tout $x \in X$, l'application partielle $f_x : Y \to Z$ est mesurable quand on munit Y de la tribu $\mathcal{N}$.

Démonstration.

(i) Posons $\mathcal{C} = \{ E \subset X \times Y \mid E_x \in \mathcal{N} \text{ pour tout } x \in X \}$.
Comme $(X \times Y)_x = Y \in \mathcal{N}$ pour tout $x \in X$, on a : $X \times Y \in \mathcal{C}$. Si $E \in \mathcal{C}$ on a pour tout $x \in X$:

$$\left(\complement_{X \times Y} E \right)_x = \complement_Y E_x \in \mathcal{N},$$

et par conséquent $\complement_{X \times Y} E \in \mathcal{C}$. Enfin, si $E_1, E_2, \dots$ est une suite d'éléments de $\mathcal{C}$, on a pour tout $x \in X$:

$$\left(\bigcup_{n \geq 1} E_n \right)_x = \bigcup_{n \geq 1} (E_n)_x \in \mathcal{N},$$

et par conséquent $\bigcup_{n \geq 1} E_n \in \mathcal{C}$. On a ainsi montré que $\mathcal{C}$ est une tribu sur $X \times Y$. En outre, si $A \in \mathcal{M}$ et $B \in \mathcal{N}$, on a pour tout $x \in X$:

$$(A \times B)_x = \begin{cases} \emptyset & \text{si } x \notin A \\ B & \text{si } x \in A, \end{cases}$$

de sorte que $A \times B \in \mathcal{C}$. Comme $\mathcal{M} \otimes \mathcal{N}$ est la plus petite tribu sur $X \times Y$ contenant les rectangles mesurables, on en déduit que $\mathcal{M} \otimes \mathcal{N} \subset \mathcal{C}$, ce qui démontre (i).

(ii) Soit B un Borélien de Z. Comme $f : (X\times Y, \mathcal{M}\otimes\mathcal{N}) \to Z$ est mesurable, on a $f^{-1}(B) \in \mathcal{M}\otimes\mathcal{N}$. Pour tout $x\in X$, on a en vertu de (i) :

$$f_x^{-1}(B) = \left(f^{-1}(B)\right)_x \in \mathcal{N},$$

et par conséquent $f_x : (Y, \mathcal{N}) \to Z$ est mesurable. ∎

L'objet de la section 4.1 est de démontrer le résultat suivant :

¶ 4.1.5- Théorème - Soient $(X, \mathcal{M}, \mu)$ et $(Y, \mathcal{N}, \nu)$ deux espaces mesurés σ-finis.

(i) Il existe une unique mesure positive $\mu\otimes\nu$ sur $\mathcal{M}\otimes\mathcal{N}$ qui vérifie :

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

quels que soient $A\in\mathcal{M}$ et $B\in\mathcal{N}$;

(ii) Pour toute fonction $f : X\times Y \to [0,+\infty]$ $\mathcal{M}\otimes\mathcal{N}$-mesurable, on a :

$$\int_{X\times Y} f\, d(\mu\otimes\nu) = \int_X d\mu(x) \int_Y f(x,y)\, d\nu(y) = \int_Y d\nu(y) \int_X f(x,y)\, d\mu(x).$$

où les intégrales superposées sont des intégrales de fonctions mesurables ≥ 0.

On dit que $\mu\otimes\nu$ est le produit des mesures μ et ν.

La démonstration du théorème 4.1.5 repose en partie sur la notion de classe monotone, que nous développons ci-dessous.

4.1.6. Classes monotones

4.1.6.1- Définition. On appelle classe monotone sur un ensemble X une famille $\mathcal{A}$ de parties de X qui vérifie les deux conditions suivantes :

(i) Pour toute suite croissante $A_1 \subset A_2 \subset \cdots \subset A_n \subset A_{n+1} \subset \cdots$ d'éléments de $\mathcal{A}$, la réunion

$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{A} ;$$

(ii) Pour toute suite décroissante $A_1 \supset A_2 \supset \cdots \supset A_n \supset A_{n+1} \supset \cdots$ d'éléments de $\mathcal{A}$, on a :

$$\bigcap_{i=1}^{\infty} A_i \in \mathcal{A}.$$

Ainsi, toute tribu sur X est une classe monotone.

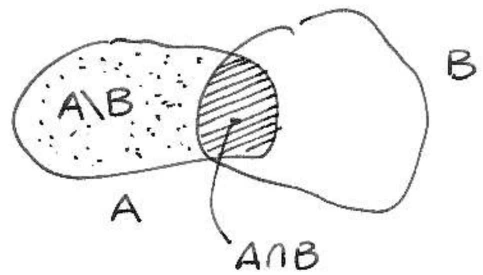
Rappelons qu'une algèbre de parties d'un ensemble X est une famille $\mathcal{R}$ de parties de X vérifiant les deux conditions suivantes :

(i) $A, B \in \mathcal{R} \Longrightarrow A \cup B \in \mathcal{R}$;
(ii) $A, B \in \mathcal{R} \Longrightarrow A \setminus B \in \mathcal{R}$.

Dans ce cas, on a alors :

(iii) $A, B \in \mathcal{R} \Longrightarrow A \cap B \in \mathcal{R}$

car $A \cap B = A \setminus (A \setminus B)$.



Pour démontrer le théorème 4.1.5, nous aurons besoin du lemme suivant :

4.1.6.2 - Lemme. Soient X un ensemble, $\mathcal{A}$ une classe monotone sur X et $\mathcal{R}$ une algèbre de parties de X. On suppose que $X \in \mathcal{R}$ et que $\mathcal{R} \subset \mathcal{A}$. Alors, on a :

$$\mathcal{R}^\sigma \subset \mathcal{A},$$

où $\mathcal{R}^\sigma$ désigne la tribu engendrée par $\mathcal{R}$.

Démonstration. Notons $\mathcal{R}^m$ la plus petite classe monotone contenant $\mathcal{R}$, c'est à dire l'intersection de toutes les classes monotones sur X contenant $\mathcal{R}$. Comme $\mathcal{A}$ est une classe monotone contenant $\mathcal{R}$, on a :

$$\mathcal{R}^m \subset \mathcal{A}$$

1ère étape. Pour prouver le lemme, il suffit de montrer que $\mathcal{R}^m$ est une algèbre. Tout d'abord, puisque $\mathcal{R}^m \supset \mathcal{R}$, on a $X \in \mathcal{R}^m$. Soient $A_1, A_2, \ldots$ des éléments de $\mathcal{R}^m$. Si $\mathcal{R}^m$ est une algèbre, on a :

$$B_n = A_1 \cup A_2 \cup \cdots \cup A_n \in \mathcal{R}^m$$

et, comme $\mathcal{R}^m$ est une classe monotone, on a :

$$\bigcup_{n \geq 1} A_n = B_1 \cup B_2 \cup \cdots \in \mathcal{R}^m.$$ Enfin, pour tout $A \in \mathcal{R}^m$, on a : $X \setminus A \in \mathcal{R}^m$ si $\mathcal{R}^m$ est une algèbre.

Il s'ensuit que, si $\mathcal{B}^m$ est une algèbre, c'est une tribu. Mais alors, comme $\mathcal{B}^m \supset \mathcal{B}$, on a :

$$\mathcal{B}^\sigma \subset \mathcal{B}^m \subset \mathcal{A},$$

et le lemme est démontré.

2ème étape. Montrons que $\mathcal{B}^m$ est une algèbre. A cet effet, posons pour tout $A \in \mathcal{B}^m$:

$$[A] = \{ B \in \mathcal{B}^m \mid A \cup B \in \mathcal{B}^m,\ A \setminus B \in \mathcal{B}^m,\ B \setminus A \in \mathcal{B}^m \}$$

On définit ainsi une classe $[A]$ de parties de X qui est monotone. En effet si $B_1 \subset B_2 \subset \cdots$ est une suite croissante d'éléments $B_n \in \mathcal{B}^m$, on a :

$$A \cup \Big(\bigcup_{n \geq 1} B_n \Big) = \bigcup_{n \geq 1} (A \cup B_n) \qquad \text{avec } A \cup B_n \in \mathcal{B}^m$$

et $A \cup B_n \subset A \cup B_{n+1}$. Comme $\mathcal{B}^m$ est une classe monotone, on a :

$$A \cup \Big(\bigcup_{n \geq 1} B_n \Big) \in \mathcal{B}^m$$

et, de même, $\displaystyle A \setminus \Big(\bigcup_{n \geq 1} B_n \Big) = \bigcap_{n \geq 1} A \setminus B_n \in \mathcal{B}^m,$

$$\Big(\bigcup_{n \geq 1} B_n \Big) \setminus A = \bigcup_{n \geq 1} (B_n \setminus A) \in \mathcal{B}^m,$$

ce qui prouve que $\displaystyle\bigcup_{n \geq 1} B_n \in \mathcal{B}^m$. De même, si $B_1 \supset B_2 \supset \cdots$ est une suite décroissante d'éléments $B_n \in \mathcal{B}^m$, on a :

$$\bigcap_{n \geq 1} B_n \in \mathcal{B}^m,$$

de sorte que $[A]$ est une classe monotone. Si $A \in \mathcal{B}$, alors on a :

$$\mathcal{B} \subset [A],$$

car $\mathcal{B}$ est une algèbre par hypothèse. Il s'ensuit que l'on a :

$$\mathcal{B}^m \subset [A],$$

car $[A]$ est une classe monotone contenant $\mathcal{B}$. On a donc établi que, pour tout $A \in \mathcal{B}$, on a :

$$B \in \mathcal{B}^m \Longrightarrow A \cup B \in \mathcal{B}^m,\ A \setminus B \in \mathcal{B}^m \text{ et } B \setminus A \in \mathcal{B}^m.$$

Pour $B \in \mathcal{B}^m$, on a donc

$\mathcal{B} \subset [B]$ d'où, puisque $[B]$ est une classe monotone :

$$\mathcal{B}^m \subset [B].$$

Mais alors, quel que soit $A \in \mathcal{B}^m$ et $B \in \mathcal{B}^m$, on a :

$A \cup B \in \mathcal{B}^m$, $A \setminus B \in \mathcal{B}^m$ et $B \setminus A \in \mathcal{B}^m$.

Ceci prouve que $\mathcal{B}^m$ est une algèbre, et achève la démonstration du lemme 4.1.6.2. ■

4.1.7. L'algèbre $\mathcal{M} \times \mathcal{N}$

On désigne par $\mathcal{M} \times \mathcal{N}$ la famille des parties de $X \times Y$ qui sont réunions finies de rectangles mesurables. On a :

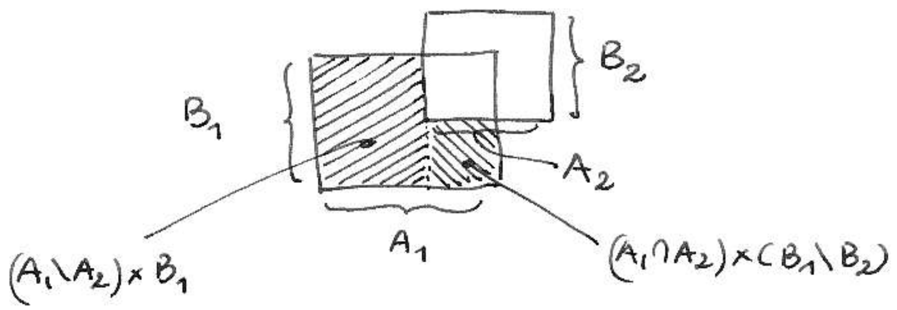
Lemme. $\mathcal{M} \times \mathcal{N}$ est une algèbre de parties de $X \times Y$ qui contient $X \times Y$. En outre, tout élément de $\mathcal{M} \times \mathcal{N}$ est une réunion finie de rectangles mesurables deux à deux disjoints.

Démonstration. $\mathcal{M} \times \mathcal{N}$ est clairement stable par réunions finies et contient $X \times Y$, qui est un rectangle mesurable. Pour montrer que $\mathcal{M} \times \mathcal{N}$ est stable par intersections finies, il suffit de démontrer que l'intersection de deux rectangles mesurables est encore un rectangle mesurable. Or ceci résulte immédiatement de la formule

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

Pour démontrer que, si $E, F \in \mathcal{M} \times \mathcal{N}$, alors $E \setminus F \in \mathcal{M} \times \mathcal{N}$, on peut se limiter en vertu de ce qui précède au cas où E et F sont des rectangles mesurables. Mais ceci résulte immédiatement de la formule :

$$(A_1 \times B_1) \setminus (A_2 \times B_2) = [(A_1 \setminus A_2) \times B_1] \cup [(A_1 \cap A_2) \times (B_1 \setminus B_2)]$$



Il résulte également de ce qui précède que la réunion de deux rectangles mesurables est réunion de cinq rectangles mesurables deux à deux disjoints. Par récurrence sur n, on en déduit immédiatement qu'une réunion de n rectangles mesurables s'écrit comme une réunion finie de rectangles mesurables deux à deux disjoints. ■

Comme $(X, \mathcal{M}, \mu)$ et $(Y, \mathcal{N}, \nu)$ sont des espaces mesurés σ-finis, il existe une suite $(A_n)_{n\geq 1}$ d'éléments de $\mathcal{M}$ vérifiant $\mu(A_n)<+\infty$ et une suite $(B_m)_{m\geq 1}$ d'éléments de $\mathcal{N}$ vérifiant $\nu(B_m)<+\infty$, telles que :

$$X = \bigcup_{n\geq 1} A_n, \qquad Y = \bigcup_{m\geq 1} B_m.$$

On a donc :

$$X\times Y = \bigcup_{n,m\geq 1} A_n \times B_m \quad \text{avec } \mu(A_n)<+\infty \text{ et } \nu(B_m)<+\infty.$$

¶ 4.1.8- Théorème. Soient $(X, \mathcal{M}, \mu)$ et $(Y, \mathcal{N}, \nu)$ deux espaces mesurés σ-finis. Soit $f : (X\times Y, \mathcal{M}\otimes\mathcal{N}) \longrightarrow [0,+\infty]$ une fonction mesurable. Alors, pour tout $x\in X$, l'application

$$f_x : y\in Y \longmapsto f_x(y) = f(x,y) \in [0,+\infty]$$

est $\mathcal{N}$-mesurable, et l'application

$$x\in X \longmapsto \int_Y f_x \, d\nu = \int_Y f(x,y)\, d\nu(y)$$

est $\mathcal{M}$-mesurable. De même, pour tout $y\in Y$, l'application

$$f_y : x\in X \longmapsto f_y(x) = f(x,y) \in [0,+\infty]$$

est $\mathcal{M}$-mesurable, et l'application

$$y\in Y \longmapsto \int_X f_y \, d\mu = \int_X f(x,y)\, d\mu(x)$$

est $\mathcal{N}$-mesurable.

Démonstration. D'après la proposition 4.1.4, pour tout $x\in X$ et tout $y\in Y$, les applications f_x et f_y sont respectivement $\mathcal{N}$-mesurable et $\mathcal{M}$-mesurable. Nous allons montrer que l'application $x\in X \longmapsto \int_Y f_x \, d\nu$ est $\mathcal{M}$-mesurable, la démonstration de la $\mathcal{N}$-mesurabilité de l'application $y\in Y \longmapsto \int_X f_y \, d\mu$ étant identique, à l'échange près des espaces mesurés $(X, \mathcal{M}, \mu)$ et $(Y, \mathcal{N}, \nu)$.

1ère étape - On peut supposer que $\mu(X)<+\infty$ et $\nu(Y)<+\infty$.

Soit en effet $(A_n)_{n\geq 1}$ (resp. $(B_n)_{n\geq 1}$) une suite croissante d'éléments de $\mathcal{M}$ (resp. $\mathcal{N}$) telle que $\mu(A_n) < +\infty$ (resp. $\nu(B_n) < +\infty$) et vérifiant :

$$X = \bigcup_{n \geq 1} A_n \quad (\text{resp. } Y = \bigcup_{n \geq 1} B_n).$$

Alors, les rectangles mesurables $A_n \times B_n$ forment une suite croissante d'éléments de $\mathcal{M} \otimes \mathcal{N}$ telle que

$$X \times Y = \bigcup_{n \geq 1} A_n \times B_n .$$

On a donc $f(x,y) = \sup_{n \geq 1} \left(f \mathbb{1}_{A_n \times B_n}(x,y) \right)$, où la suite des fonctions $f_n = f \mathbb{1}_{A_n \times B_n}$ est croissante. Comme les f_n sont $\mathcal{M} \otimes \mathcal{N}$-mesurables, les $(f_n)_x$ sont $\mathcal{N}$-mesurables pour tout $x \in X$, et on a en vertu du théorème de convergence monotone :

$$\int_Y f_x \, d\nu = \sup_{n \geq 1} \int_Y (f_n)_x \, d\nu .$$

Pour montrer que l'application $x \longmapsto \int_Y f_x \, d\nu$ est $\mathcal{M}$-mesurable, il suffit d'établir que, pour tout $n \geq 1$, l'application

$$x \in X \longmapsto \int_Y f(x,y) \mathbb{1}_{A_n}(x) \mathbb{1}_{B_n}(y) \, d\nu(y) = \mathbb{1}_{A_n}(x) \int_{B_n} f(x,y) \, d\nu(y)$$

est $\mathcal{M}$-mesurable, ou encore que l'application

$$x \in A_n \longmapsto \int_{B_n} f(x,y) \, d\nu(y)$$

est mesurable relativement à la tribu induite par $\mathcal{M}$ sur A_n. On peut donc, pour démontrer le théorème 4.1.8, remplacer X par A_n et Y par B_n, de sorte que l'on peut supposer que $\mu(X) < +\infty$ et $\nu(Y) < +\infty$.

<u>2ème étape</u>. <u>On peut supposer que $f = \mathbb{1}_E$ où $E \in \mathcal{M} \otimes \mathcal{N}$</u> (ainsi que $\mu(X) < +\infty$, $\nu(Y) < +\infty$). En effet, il existe en vertu de la proposition 2.3.2 (ii) une suite croissante

$$0 \leq f_1 \leq f_2 \leq \cdots$$

de fonctions mesurables positives telle que :

$$f = \sup_n f_n .$$

On en déduit que

$$\int_Y f_x \, d\nu = \sup_n \int_Y f_{n,x} \, d\nu ,$$

de sorte qu'il suffit de prouver que l'application $x \longmapsto \int_Y f_{n,x} \, d\nu$ est $\mathcal{M}$-mesurable pour tout entier n, pour en déduire la $\mathcal{M}$-mesurabilité de l'application $x \longmapsto \int_Y f_x \, d\nu$.

Or, pour tout $n \geqq 1$ fixé, on a :

$$f_n = \sum_{i=1}^{k} \lambda_i \mathbb{1}_{E_i}$$

où les λ_i sont des réels positifs et les E_i des éléments de $\mathcal{M} \otimes \mathcal{N}$ deux à deux disjoints. On a donc :

$$\int_Y f_{n,x} \, d\nu = \sum_{i=1}^{k} \lambda_i \int_Y (\mathbb{1}_{E_i})_x \, d\nu ,$$

ce qui ramène la $\mathcal{M}$-mesurabilité de l'application $x \longmapsto \int_Y f_{n,x} \, d\nu$ à la $\mathcal{M}$-mesurabilité des applications $x \longmapsto \int_Y (\mathbb{1}_{E_i})_x \, d\nu$.

On peut donc supposer, pour démontrer le théorème, que $f = \mathbb{1}_E$ où $E \in \mathcal{M} \otimes \mathcal{N}$.

<u>3ème étape. Pour tout $E \in \mathcal{M} \otimes \mathcal{N}$, l'application $x \longmapsto \int_Y \mathbb{1}_E(x,y) \, d\nu(y) = \nu(E_x)$ est $\mathcal{M}$-mesurable.</u>

Pour démontrer cette assertion, qui achèvera la preuve du théorème 4.1.8, nous supposerons (cf. 1ère étape) que $\mu(X) < +\infty$ et $\nu(Y) < +\infty$. A cet effet, posons :

$$\mathcal{A} = \{ E \in \mathcal{M} \otimes \mathcal{N} \mid x \longmapsto \nu(E_x) \text{ est } \mathcal{M}\text{-mesurable} \}.$$

On définit ainsi une classe monotone $\mathcal{A}$ sur $X \times Y$. En effet, si $E_1 \subset E_2 \subset \cdots$ est une suite croissante d'éléments de $\mathcal{A}$ et si $E = \bigcup_{n \geq 1} E_n$, on a pour tout $x \in X$:

$$E_{1,x} \subset E_{2,x} \subset \cdots \quad \text{et} \quad E_x = \bigcup_{n \geq 1} E_{n,x}.$$

D'après le théorème 1.2.5, on a :

$$\nu(E_x) = \sup_{n \geq 1} \nu(E_{n,x}),$$

et, comme chaque application $x \mapsto \nu(E_{n,x})$ est $\mathcal{M}$-mesurable puisque $E_n \in \mathcal{A}$, l'application $x \mapsto \nu(E_x)$ est $\mathcal{M}$-mesurable, et $E \in \mathcal{A}$.

De même, si $E_1 \supset E_2 \supset \cdots$ est une suite décroissante d'éléments de $\mathcal{A}$, l'intersection $E = \bigcap_{n \geq 1} E_n$ vérifie

$$E_x = \bigcap_{n \geq 1} E_{n,x} \quad \text{avec} \quad E_{1,x} \supset E_{2,x} \supset \cdots$$

et, puisque $\nu(Y) < +\infty$, on a :

$$\nu(E_x) = \operatorname{Inf}_{n \geq 1} \nu(E_{n,x}),$$

d'où il résulte que $x \mapsto \nu(E_x)$ est $\mathcal{M}$-mesurable, et donc que $E \in \mathcal{A}$.

Montrons que $\mathcal{A}$ contient l'algèbre $\mathcal{M} \times \mathcal{N}$. Tout d'abord, si $E = A \times B$ avec $A \in \mathcal{M}$ et $B \in \mathcal{N}$, on a pour $x \in X$:

$$\nu(E_x) = \mathbb{1}_A(x)\nu(B),$$

et par conséquent $x \mapsto \nu(E_x)$ est $\mathcal{M}$-mesurable. Si $E \in \mathcal{M} \times \mathcal{N}$, il existe une suite finie $A_1 \times B_1, \ldots, A_n \times B_n$ de rectangles mesurables deux à deux disjoints telle que

$$E = \bigcup_{i=1}^{n} A_i \times B_i.$$

Alors, pour tout $x \in E$, on a :

$$E_x = \bigcup_{i=1}^{n} (A_i \times B_i)_x, \quad \text{où les } (A_i \times B_i)_x \text{ sont}$$

deux à deux disjoints, et par conséquent :

$$\nu(E_x) = \sum_{i=1}^{n} \nu((A_i \times B_i)_x)$$

et la $\mathcal{M}$-mesurabilité de chaque application $x \mapsto \nu((A_i \times B_i)_x)$ implique que $x \mapsto \nu(E_x)$ est $\mathcal{M}$-mesurable, de sorte que $E \in \mathcal{A}$. Ainsi, $\mathcal{A}$ est une classe monotone qui contient l'algèbre $\mathcal{M} \times \mathcal{N}$ (laquelle contient $X \times Y$); d'après le lemme 4.1.6.2, $\mathcal{A}$ contient la tribu engendrée par $\mathcal{M}\times\mathcal{N}$, c'est à dire contient $\mathcal{M}\otimes\mathcal{N}$. On a donc

$$\mathcal{M}\otimes\mathcal{N}\subset\mathcal{A}\subset\mathcal{M}\otimes\mathcal{N},$$

d'où $\mathcal{A}=\mathcal{M}\otimes\mathcal{N}$. Ceci prouve que, pour tout $E\in\mathcal{M}\otimes\mathcal{N}$, l'application $x\mapsto\nu(E_x)$ est $\mathcal{M}$-mesurable, ce qui achève la démonstration du théorème 4.1.8. ■

Nous pouvons maintenant démontrer le théorème 4.1.5.

4.1.9 - Démonstration du théorème 4.1.5

Soient $(X,\mathcal{M},\mu)$ et $(Y,\mathcal{N},\nu)$ deux espaces mesurés σ-finis.

1ère étape. Montrons l'existence d'une mesure positive β sur $\mathcal{M}\otimes\mathcal{N}$ qui vérifie $\beta(A\times B)=\mu(A)\nu(B)$ quels que soient $A\in\mathcal{M}$ et $B\in\mathcal{N}$.

Pour toute partie $E\in\mathcal{M}\otimes\mathcal{N}$, posons:

$$\beta(E)=\int_X \nu(E_x)\,d\mu(x).$$

Comme l'application $x\mapsto\nu(E_x)$ est $\mathcal{M}$-mesurable positive en vertu du théorème 4.1.8, son intégrale sur X par rapport à la mesure μ est bien définie. L'application

$$\beta:\mathcal{M}\otimes\mathcal{N}\longrightarrow[0,+\infty]$$

vérifie $\beta(\emptyset)=0$, car $\nu(\emptyset_x)=\nu(\emptyset)=0$. Par ailleurs, si $E_1,E_2,\dots$ sont des parties deux à deux disjointes de $\mathcal{M}\otimes\mathcal{N}$, on a, en posant $E=\bigcup_{n\geq 1}E_n$:

$$\nu(E_x)=\nu\Big(\bigcup_{n\geq 1}E_{n,x}\Big)=\sum_{n\geq 1}\nu(E_{n,x}),$$

car les $E_{n,x}$ sont deux à deux disjoints. Il résulte alors du théorème 3.1.8 que:

$$\int_X\nu(E_x)\,d\mu(x)=\sum_{n\geq 1}\int_X\nu(E_{n,x})\,d\mu(x),\quad\text{soit}$$

$$\beta(E)=\sum_{n\geq 1}\beta(E_n).$$

Ceci prouve que β est une mesure positive sur $\mathcal{M}\otimes\mathcal{N}$.

Si $E=A\times B$ avec $A\in\mathcal{M}$ et $B\in\mathcal{N}$, on a:

$$\beta(A\times B)=\int_X \mathbb{1}_A(x)\nu(B)d\mu(x)=\nu(B)\mu(A),$$

de sorte que β vérifie la condition demandée.

<u>2ème étape. Unicité de $\mu\otimes\nu$.</u>

Notons $\mu\otimes\nu$ la mesure positive sur $\mathcal{M}\otimes\mathcal{N}$ définie par

$$\mu\otimes\nu(E)=\int_X \nu(E_x)\,d\mu(x),$$

et soit β une mesure positive sur $\mathcal{M}\otimes\mathcal{N}$ qui vérifie $\beta(A\times B)=\mu(A)\nu(B)$ quels que soient $A\in\mathcal{M}$ et $B\in\mathcal{N}$. Montrons que $\beta=\mu\otimes\nu$. A cet effet, observons que $\beta(E)=\mu\otimes\nu(E)$ pour tout $E\in\mathcal{M}\times\mathcal{N}$. En effet, si $E\in\mathcal{M}\times\mathcal{N}$, il existe une suite finie $A_1\times B_1,\dots,A_n\times B_n$ de rectangles mesurables deux à deux disjoints telle que:

$$E=\bigcup_{i=1}^{n}A_i\times B_i .$$

On a donc:

$$\beta(E)=\sum_{i=1}^{n}\beta(A_i\times B_i)=\sum_{i=1}^{n}\mu(A_i)\nu(B_i)=\sum_{i=1}^{n}(\mu\otimes\nu)(A_i\times B_i)$$
$$=(\mu\otimes\nu)(E).$$

Comme $\mathcal{M}\times\mathcal{N}$ est une algèbre de parties de $X\times Y$, qu'il existe une suite $(A_n\times B_n)_{n\geq 1}$ de parties de $\mathcal{M}\times\mathcal{N}$ telle que $\mu(A_n)<+\infty$, $\nu(B_n)<+\infty$ et

$$X\times Y=\bigcup_{n\geq 1}A_n\times B_n\quad,\quad A_n\times B_n\subset A_{n+1}\times B_{n+1},$$

avec $\beta(A_n\times B_n)=\mu(A_n)\nu(B_n)=(\mu\otimes\nu)(A_n\times B_n)<+\infty$, et qu'enfin $\mathcal{M}\otimes\mathcal{N}$ est la tribu engendrée par $\mathcal{M}\times\mathcal{N}$, le théorème de Hahn (cf. 1.3.3.4 (i)) implique que les mesures β et $\mu\otimes\nu$, qui coïncident sur $\mathcal{M}\times\mathcal{N}$, sont égales sur $\mathcal{M}\otimes\mathcal{N}$. Donc, $\beta=\mu\otimes\nu$, et l'unicité est démontrée.

<u>3ème étape. Pour toute fonction $f:X\times Y\to[0,+\infty]$ $\mathcal{M}\otimes\mathcal{N}$-mesurable, on a:</u> $\displaystyle\int_{X\times Y} f\,d(\mu\otimes\nu)=\int_X d\mu(x)\int_Y f(x,y)\,d\nu(y)$

Observons tout d'abord que $\int_X d\mu(x)\int_Y f(x,y)\,d\nu(y)$, qui est par définition l'intégrale par rapport à μ de la fonction $x \longmapsto \int_Y f(x,y)\,d\nu(y) = \nu(f_x)$, a un sens puisque la fonction positive $x \mapsto \nu(f_x)$ est $\mathcal{M}$-mesurable en vertu du théorème 4.1.8. Posons :

$$L(f) = \int_X d\mu(x)\int_Y f(x,y)\,d\nu(y)$$

($f : X\times Y \to [0,+\infty]$ est $\mathcal{M}\otimes\mathcal{N}$-mesurable). Si $(f_1, f_2, \dots)$ est une suite croissante de fonctions $\mathcal{M}\otimes\mathcal{N}$-mesurables positives, on a :

$$(1) \quad L(\operatorname{Sup}_n f_n) = \operatorname{Sup}_n L(f_n).$$

En effet, on a pour tout $x\in X$, en posant $f = \operatorname{Sup}_n f_n$:

$$f_x = \operatorname{Sup}_n f_{n,x} \quad \text{avec } f_{1,x} \leqq f_{2,x} \leqq \cdots,$$

et il résulte du théorème de convergence monotone que :

$$\nu^x(f) = \int_Y f_x\,d\nu = \operatorname{Sup}_n \int_Y f_{n,x}\,d\nu = \operatorname{Sup}_n \nu^x(f_n).$$

Comme les fonctions $x \mapsto \nu^x(f_n)$ sont $\mathcal{M}$-mesurables positives et forment une suite croissante, une nouvelle application du théorème de convergence monotone montre que :

$$\int_X \nu^x(f)\,d\mu(x) = \operatorname{Sup}_n \int_X \nu^x(f_n)\,d\mu(x),$$

soit $L(f) = \operatorname{Sup}_n L(f_n)$.

Si f est une fonction simple positive, elle peut s'écrire sous la forme

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

où $\lambda_i \in [0,+\infty]$ et où les E_i sont des éléments de $\mathcal{M}\otimes\mathcal{N}$ qui sont deux à deux disjoints. On a alors :

$$\int_{X\times Y} f\,d(\mu\otimes\nu) = \sum_{i=1}^n \lambda_i (\mu\otimes\nu)(E_i) = \sum_{i=1}^n \lambda_i \int_X \nu(E_{i,x})\,d\mu(x)$$

$$= \sum_{i=1}^n \lambda_i \int_X d\mu(x)\int_Y \mathbb{1}_{E_i}(x,y)\,d\nu(y)$$

$$= \int_X d\mu(x)\int_Y \Big(\sum_{i=1}^n \lambda_i \mathbb{1}_{E_i}(x,y)\Big)\,d\nu(y) = L(f).$$

Si $f : X\times Y \to [0,+\infty]$ est une fonction positive $\mathcal{M}\otimes\mathcal{N}$ - mesurable, il existe une suite $(f_1, f_2, \dots)$ de fonctions simples, $f_n \leq f_{n+1}$) telle que $f = \sup_n f_n$. On a alors, en vertu du théorème de convergence monotone :

$$\int_{X\times Y} f\, d(\mu\otimes\nu) = \sup_n \int_{X\times Y} f_n\, d(\mu\otimes\nu) = \sup_n L(f_n) = L(f),$$

où la dernière égalité résulte de (1). On a ainsi montré que l'intégrale (supérieure) de f est donnée par la formule:

$$\int_{X\times Y} f\, d(\mu\otimes\nu) = \int_X d\mu(x) \int_Y f(x,y)\, d\nu(y).$$

<u>4ème étape. Pour toute fonction $f : X\times Y \to [0,+\infty]$ $\mathcal{M}\otimes\mathcal{N}$-mesurable, on a</u> : $\displaystyle\int_X d\mu(x)\int_Y f(x,y)\,d\nu(y) = \int_Y d\nu(y)\int_X f(x,y)\,d\mu(x)$.

Considérons la mesure α sur $\mathcal{M}\otimes\mathcal{N}$ définie par:

$$\alpha(E) = \int_Y \mu(E_y)\, d\nu(y).$$

Comme dans la première étape, on démontre que α est une mesure positive sur $\mathcal{M}\otimes\mathcal{N}$, qui vérifie:

$$\alpha(A\times B) = \mu(A)\,\nu(B)$$

quels que soient $A\in\mathcal{M}$ et $B\in\mathcal{N}$. D'après la 2ème étape, on a $\alpha = \mu\otimes\nu$. Comme dans la 3ème étape, on démontre alors que l'on a, pour toute fonction f $\mathcal{M}\otimes\mathcal{N}$-mesurable ≥ 0 :

$$\int_{X\times Y} f\, d(\mu\otimes\nu) = \int_Y d\nu(y) \int_X f(x,y)\, d\mu(x),$$

en introduisant la fonction

$$K(f) = \int_Y d\nu(y) \int_X f(x,y)\, d\mu(x).$$

Il s'ensuit que $\displaystyle\int_X d\mu(x)\int_Y f(x,y)\,d\nu(y) = \int_Y d\nu(y)\int_X f(x,y)\,d\mu(x)$,

et le théorème 4.1.5 est démontré. ■

4.1.10- Produit de n mesures positives

4.1.10.1. Définition de $\mu_1 \otimes \cdots \otimes \mu_n$. Soient $(X_i, \mathcal{M}_i, \mu_i)$ $(1 \leq i \leq n)$ n espaces mesurés σ-finis. Notons $X = X_1 \times \cdots \times X_n$ le produit de ces espaces, et désignons par $\mathcal{M} = \mathcal{M}_1 \otimes \cdots \otimes \mathcal{M}_n$ la σ-algèbre sur X engendrée par les parties de la forme $A_1 \times \cdots \times A_n$, où $A_i \in \mathcal{M}_i$ pour tout i.

Par application répétée du théorème 4.1.5, on peut facilement établir les résultats suivants :

(i) Il existe une unique mesure positive $\mu = \mu_1 \otimes \cdots \otimes \mu_n$ sur $\mathcal{M} = \mathcal{M}_1 \otimes \cdots \otimes \mathcal{M}_n$ qui vérifie :

$$\mu(A_1 \times \cdots \times A_n) = \prod_{i=1}^{n} \mu_i(A_i)$$

quels que soient les $A_i \in \mathcal{M}_i$ $(i = 1, 2, \dots, n)$;

(ii) L'intégrale d'une fonction $\mathcal{M}$-mesurable positive $f : X \longrightarrow [0, +\infty]$ est donnée par la formule :

$$\int_X f \, d\mu = \int_{X_{\sigma(1)}} d\mu_{\sigma(1)}(x_{\sigma(1)}) \int_{X_{\sigma(2)}} d\mu_{\sigma(2)}(x_{\sigma(2)}) \cdots \int_{X_{\sigma(n)}} f(x_1, \dots, x_n) \, d\mu_{\sigma(n)}(x_{\sigma(n)})$$

où σ est une permutation quelconque de $\{1, 2, \dots, n\}$, et où les intégrales successives du second membre sont des intégrales de fonctions mesurables positives.

4.1.10.2 - Exemple. Soit λ_n la mesure de Lebesgue sur la tribu $\mathcal{B}_{\mathbb{R}^n}$ des Boréliens de $\mathbb{R}^n$. Alors, on a :

$$\lambda_n = \underbrace{\lambda_1 \otimes \lambda_1 \otimes \cdots \otimes \lambda_1}_{n \text{ facteurs.}}$$

En effet, $\mathcal{B}_{\mathbb{R}^n}$ coïncide avec $\underbrace{\mathcal{B}_{\mathbb{R}} \otimes \cdots \otimes \mathcal{B}_{\mathbb{R}}}_{n \text{ facteurs}}$, et λ_n vérifie bien la relation :

$$\lambda_n(A_1 \times \cdots \times A_n) = \lambda_1(A_1) \cdots \lambda_1(A_n), \quad A_i \in \mathcal{B}_{\mathbb{R}},$$

comme on peut le vérifier. D'après le résultat d'unicité de la mesure produit, on a :

$$\lambda_n = \underbrace{\lambda_1 \otimes \lambda_1 \otimes \cdots \otimes \lambda_1}_{n \text{ facteurs}},$$

et les fonctions Boréliennes positives sur $\mathbb{R}^n$ s'intègrent « par tranches », en utilisant uniquement l'intégrale de Lebesgue des fonctions d'une variable réelle.

4.2- THÉORÈME DE FUBINI

Soient $(X, \mathcal{M}, \mu)$ et $(Y, \mathcal{N}, \nu)$ deux espaces mesurés σ-finis, et notons $\mu\otimes\nu$ la mesure produit sur $X\times Y$. Les fonctions $\mu\otimes\nu$-intégrables sur $X\times Y$ sont des fonctions qui sont mesurables par rapport à la tribu complétée de $\mathcal{M}\otimes\mathcal{N}$, de sorte que, pour étudier les questions de $\mu\otimes\nu$-intégrabilité sur $X\times Y$, il nous faut étendre le théorème 4.1.8 et la propriété (ii) du théorème 4.1.5 au cas des fonctions $\mu\otimes\nu$-mesurables positives sur $X\times Y$. A cet effet, nous utiliserons les résultats suivants :

4.2.1- Proposition- Soient $(X, \mathcal{M}, \mu)$ et $(Y, \mathcal{N}, \nu)$ deux espaces mesurés σ-finis. Pour toute fonction numérique $\mu\otimes\nu$-mesurable $f : X\times Y \longrightarrow \overline{\mathbb{R}}$, il existe une fonction numérique $\mathcal{M}\otimes\mathcal{N}$-mesurable $g : X\times Y \longrightarrow \overline{\mathbb{R}}$ et un ensemble $N\in\mathcal{M}\otimes\mathcal{N}$ vérifiant $(\mu\otimes\nu)(N) = 0$, tels que l'on ait :

$$\forall (x,y)\notin N, \quad f(x,y) = g(x,y).$$

Démonstration. Quitte à considérer séparément les parties positive et négative f^+ et f^- de f, on peut sans perte de généralité supposer que f est une fonction numérique positive. Soit $(f_n)_{n\geqq 1}$ une suite de fonctions $\mu\otimes\nu$-mesurables simples et positives telle que

$$f(x,y) = \lim_{n\to\infty} f_n(x,y)$$

pour tout $(x,y)\in X\times Y$. Montrons qu'il existe, pour tout entier $n\geqq 1$, une fonction $g_n : X\times Y \to [0,+\infty]$ $\mathcal{M}\otimes\mathcal{N}$-mesurable et un sous-ensemble N_n de $X\times Y$, $N_n\in\mathcal{M}\otimes\mathcal{N}$, tels que

$$f_n(x,y) = g_n(x,y) \quad \text{quel que soit } (x,y)\notin N_n.$$

A cet effet, écrivons :

$$f_n = \sum_{i=1}^{k} \lambda_i \mathbb{1}_{E_i}$$

où les λ_i sont des réels $\geqq 0$ et où les E_i sont des ensembles $\mu\otimes\nu$-mesurables deux à deux disjoints. Pour tout i, il existe $F_i \in \mathcal{M}\otimes\mathcal{N}$ et Z_i $\mu\otimes\nu$-négligeable, disjoint de F_i, tels que :

$$E_i = F_i \cup Z_i .$$

On a :

$$f_n = \sum_{i=1}^{k} \lambda_i \mathbb{1}_{E_i} = \sum_{i=1}^{k} \lambda_i \mathbb{1}_{F_i} + \sum_{i=1}^{k} \lambda_i \mathbb{1}_{Z_i},$$

de sorte que, si l'on pose $g_n = \sum_{i=1}^{k} \lambda_i \mathbb{1}_{F_i}$, et si l'on note N_n un élément de $\mathcal{M}\otimes\mathcal{N}$ contenant l'ensemble $\mu\otimes\nu$-négligeable $\bigcup_{i=1}^{k} Z_i$ et vérifiant $(\mu\otimes\nu)(N_n) = 0$, on a :

$$f_n(x,y) = g_n(x,y) \quad \text{pour } (x,y) \notin N_n .$$

En outre, la fonction g_n est $\mathcal{M}\otimes\mathcal{N}$-mesurable. Posons $N = \bigcup_{n\geq 1} N_n$; on définit ainsi un élément de $\mathcal{M}\otimes\mathcal{N}$ qui vérifie $(\mu\otimes\nu)(N) = 0$. Soit $g = f\,\mathbb{1}_{N^c}$; on a $g(x,y) = f(x,y)$ pour $(x,y) \notin N$ et, comme

$$g(x,y) = \lim_{n\to\infty} (g_n \mathbb{1}_{N^c})(x,y)$$

pour tout $(x,y) \in X\times Y$, où g_n est $\mathcal{M}\otimes\mathcal{N}$-mesurable et $N \in \mathcal{M}\otimes\mathcal{N}$, on en déduit que g est $\mathcal{M}\otimes\mathcal{N}$-mesurable. La proposition est ainsi démontrée. ■

4.2.2. Proposition. *Soient $(X,\mathcal{M},\mu)$ et $(Y,\mathcal{N},\nu)$ deux espaces mesurés σ-finis. Soit $N \subset X\times Y$ un sous-ensemble $\mu\otimes\nu$-négligeable. Alors, N_x est ν-négligeable pour μ-presque tout x, et N_y est μ-négligeable pour ν-presque tout y.*

Démonstration - Soit $Z \in \mathcal{M}\otimes\mathcal{N}$ tel que $N \subset Z$ et $(\mu\otimes\nu)(Z) = 0$. D'après le théorème 4.1.5, on a :

$$(\mu\otimes\nu)(Z) = \int_X \nu(Z_x)\, d\mu(x) = 0,$$

où la fonction $x \longmapsto \nu(Z_x)$ est μ-mesurable positive. On en déduit que $\nu(Z_x) = 0$ pour μ-presque tout $x \in X$ et, comme $N_x \subset Z_x \in \mathcal{N}$, N_x est ν-négligeable pour μ-presque tout $x \in X$. De même, on montre que N_y est μ-négligeable pour ν-presque tout $y \in Y$. ■

Nous pouvons maintenant démontrer les analogues des théorèmes 4.1.8 et 4.1.5 (ii) pour les fonctions $\mu\otimes\nu$-mesurables.

4.2.3 - Théorème. Soient $(X, \mathcal{M}, \mu)$ et $(Y, \mathcal{N}, \nu)$ deux espaces mesurés σ-finis. Soit $f : X\times Y \longrightarrow [0, +\infty]$ une fonction $\mu\otimes\nu$-mesurable. Alors, on a :

(i) L'application f_x est ν-mesurable pour μ-presque tout $x\in X$, et l'application $x \longmapsto \int_Y f_x \, d\nu$, définie μ-presque partout, est μ-mesurable ;

(ii) L'application f_y est μ-mesurable pour ν-presque tout $y\in Y$, et l'application $y \longmapsto \int_X f_y \, d\mu$, définie ν-presque partout, est ν-mesurable ;

(iii)
$$\int_{X\times Y} f \, d(\mu\otimes\nu) = \int_X d\mu(x) \int_Y f_x \, d\nu = \int_Y d\nu(y) \int_X f_y \, d\mu .$$

On note parfois $\iint_{X\times Y} f\, d(\mu\otimes\nu)$ ou $\iint_{X\times Y} f(x,y)\, d\mu(x)\, d\nu(y)$ l'intégrale $\int_{X\times Y} f\, d(\mu\otimes\nu)$, et on parle de « l'intégrale double » de f sur $X\times Y$.

Démonstration (i) D'après la proposition 4.2.1, il existe un ensemble $N\in \mathcal{M}\otimes\mathcal{N}$ tel que $(\mu\otimes\nu)(N) = 0$ et tel que $g = f\,\mathbb{1}_{N^c}$ soit $\mathcal{M}\otimes\mathcal{N}$-mesurable. D'après la proposition 4.2.2, il existe un sous-ensemble $A\subset X$, A μ-négligeable, tel que N_x soit ν-négligeable pour tout $x\notin A$. Fixons $x\notin A$. Alors $f_x = g_x$ hors de N_x qui est ν-négligeable et, comme g_x est $\mathcal{N}$-mesurable, f_x est ν-mesurable et vérifie :

$$\int_Y f_x \, d\nu = \int_Y g_x \, d\nu .$$

D'après le théorème 4.1.8, la fonction $x \mapsto \int_Y g_x \, d\nu$ est $\mathcal{M}$-mesurable ; comme elle est égale μ-presque partout à la fonction (définie μ-presque partout) $x\mapsto \int_Y f_x\, d\nu$, cette dernière est μ-mesurable et on a :

$$\int_X d\mu(x) \int_Y f_x \, d\nu = \int_X d\mu(x) \int_Y g_x \, d\nu = \int_{X\times Y} g \, d(\mu\otimes\nu) = \int_{X\times Y} f\, d(\mu\otimes\nu) .$$

(ii) Se démontre de la même manière que (i). On a :

$$\int_{X\times Y} f\,d(\mu\otimes\nu) = \int_Y d\nu(y) \int_X f_y\,d\mu.$$

(iii) Résulte immédiatement de (i) et (ii). ∎

4.2.4- Théorème (Fubini). Soient $(X,\mathcal{M},\mu)$ et $(Y,\mathcal{N},\nu)$ deux espaces mesurés σ-finis et $f : X\times Y \longrightarrow \mathbb{R}$ une fonction $\mu\otimes\nu$-intégrable. Alors, on a :

(i) Pour μ-presque tout $x\in X$, la fonction $f_x : y \longmapsto f(x,y)$ est ν-intégrable, et la fonction (définie μ-presque partout) $x \longmapsto \int_Y f(x,y)\,d\nu(y)$ est μ-intégrable ;

(ii) Pour ν-presque tout $y\in Y$, la fonction $f_y : x \longmapsto f(x,y)$ est μ-intégrable, et la fonction (définie ν-presque partout) $y \longmapsto \int_X f(x,y)\,d\mu(x)$ est ν-intégrable ;

(iii) L'intégrale $\int_{X\times Y} f\,d(\mu\otimes\nu) = \iint_{X\times Y} f(x,y)\,d\mu(x)\,d\nu(y)$ de f sur $X\times Y$ par rapport à la mesure $\mu\otimes\nu$ est donnée par :

$$\iint_{X\times Y} f(x,y)\,d\mu(x)\,d\nu(y) = \int_X d\mu(x) \int_Y f(x,y)\,d\nu(y)$$
$$= \int_Y d\nu(y) \int_X f(x,y)\,d\mu(x).$$

Démonstration. Quitte à considérer f^+ et f^-, on peut supposer, sans perte de généralité, que f est positive. D'après le théorème 4.2.3, la fonction f_x est ν-mesurable pour μ-presque tout x, et la fonction $x \longmapsto \int_Y f_x\,d\nu$ est μ-mesurable. En outre, on a :

$$\int_X d\mu(x) \int_Y f(x,y)\,d\nu(y) = \iint_{X\times Y} f(x,y)\,d\mu(x)\,d\nu(y) < +\infty.$$

Il s'ensuit que $\int_Y f(x,y)\,d\nu(y) < +\infty$ pour μ-presque tout x, i.e. f_x est ν-intégrable pour μ-presque tout x. Par ailleurs, $x \longmapsto \int_Y f(x,y)\,d\nu(y)$ est μ-intégrable, et on a

$$\iint_{X\times Y} f(x,y)\,d\mu(x)\,d\nu(y) = \int_X d\mu(x) \int_Y f(x,y)\,d\nu(y).$$

En échangeant les rôles de X et de Y, on obtient (ii) et la totalité de (iii). ∎

4.2.5- Théorème (Fatou). Soient $(X, \mathcal{M}, \mu)$ et $(Y, \mathcal{N}, \nu)$ deux espaces mesurés σ-finis. Soit $f : X\times Y \longrightarrow \mathbb{R}$ une fonction $\mu\otimes\nu$-mesurable. On suppose que la fonction $y \longmapsto |f(x,y)|$ est ν-intégrable pour μ-presque tout x, et que la fonction $x \longmapsto \int_Y |f(x,y)|\,d\nu(y)$ est μ-intégrable. Alors, f est $\mu\otimes\nu$-intégrable, et les conclusions du théorème de Fubini sont vérifiées.

Démonstration. Comme la fonction $|f|$ est $\mu\otimes\nu$-mesurable, on a en vertu du théorème 4.2.3 :

$$\iint_{X\times Y} |f|\,d(\mu\otimes\nu) = \int_X d\mu(x) \int_Y |f(x,y)|\,d\nu(y),$$

et les hypothèses entraînent que $\iint_{X\times Y} |f|\,d(\mu\otimes\nu) < +\infty$.

La fonction $|f|$ est donc $\mu\otimes\nu$-intégrable, ce qui implique que f est $\mu\otimes\nu$-intégrable. ∎

4.2.6. Corollaire - Soient $(X, \mathcal{M}, \mu)$ et $(Y, \mathcal{N}, \nu)$ deux espaces mesurés σ-finis. Soient $f : X \longrightarrow \mathbb{R}$ et $g : Y \longrightarrow \mathbb{R}$ deux fonctions respectivement μ-intégrable et ν-intégrable. Alors, la fonction

$$f\otimes g : (x,y) \in X\times Y \longrightarrow f(x)g(y) \in \mathbb{R}$$

est $\mu\otimes\nu$-intégrable et on a :

$$\iint_{X\times Y} f(x)g(y)\,d\mu(x)\,d\nu(y) = \left(\int_X f(x)\,d\mu(x)\right)\left(\int_Y g(y)\,d\nu(y)\right).$$

Démonstration. La fonction $(x,y) \longmapsto f(x)$ est $\mu\otimes\nu$-mesurable car, pour tout $\lambda \in \mathbb{R}$, l'ensemble des $(x,y)\in X\times Y$ tels que $f(x) \leqq \lambda$ est de la forme $(A\cup N)\times Y$ où $A \in \mathcal{M}$ et N est μ-négligeable. Or

$$(A\cup N)\times Y = (A\times Y)\cup(N\times Y) \qquad \text{où } A\times Y \in \mathcal{M}\otimes\mathcal{N},$$

et où $N\times Y$ est $\mu\otimes\nu$-négligeable, de sorte que $(A\cup N)\times Y$ est $\mu\otimes\nu$-mesurable. La fonction $(x,y) \longmapsto f(x)$ est donc bien $\mu\otimes\nu$-mesurable. De même, la fonction

$(x,y) \longmapsto g(y)$ est $\mu\otimes\nu$-mesurable, et donc la fonction produit $(x,y)\longmapsto f(x)g(y)$ est $\mu\otimes\nu$-mesurable. D'après le théorème de Fatou, $f\otimes g$ est $\mu\otimes\nu$-intégrable, et le théorème de Fubini fournit l'égalité

$$\iint_{X\times Y} f(x)g(y)\,d\mu(x)\,d\nu(y) = \left(\int_X f\,d\mu\right)\left(\int_Y g\,d\nu\right). \blacksquare$$

4.2.7. Corollaire. Soient $(X,\mathcal{M},\mu)$ et $(Y,\mathcal{N},\nu)$ deux espaces mesurés σ-finis. Soit N une partie $\mu\otimes\nu$-mesurable de $X\times Y$. Si, pour μ-presque tout $x\in X$, la coupe N_x est ν-négligeable, alors N est $\mu\otimes\nu$-négligeable.

Démonstration : C'est une conséquence immédiate du théorème de Fubini. $\blacksquare$

4.2.8. Remarque. Il existe des fonctions $\mu\otimes\nu$-mesurables $f : X\times Y \to \mathbb{R}$ pour lesquelles les intégrales superposées $\int_X d\mu(x)\int_Y f(x,y)\,d\nu(y)$ et $\int_Y d\nu(y)\int_X f(x,y)\,d\mu(x)$ ont un sens, mais qui ne sont pas $\mu\otimes\nu$-intégrables. Considérons par exemple, sur le produit $[0,1]\times[0,1]$ muni de la mesure produit des mesures de Lebesgue sur $[0,1]$, la fonction f définie par :

$$f(x,y) = \frac{x^2-y^2}{(x^2+y^2)^2} \quad \text{si } (x,y)\neq(0,0), \text{ et}$$

$$f(0,0) = 0.$$

Cette fonction est mesurable, car elle est continue sauf au point $(0,0)$ qui est un ensemble négligeable. Pour $x\neq 0$, la fonction f_x est continue donc intégrable sur $[0,1]$, de primitive $y\longmapsto \frac{y}{x^2+y^2}$, de sorte que $\int_0^1 f_x\,dy = \frac{1}{1+x^2}$ pour $x\neq 0$. Cette dernière fonction, définie pour $x\neq 0$, est intégrable, et on a :

$$\int_0^1 dx \int_0^1 \frac{x^2-y^2}{(x^2+y^2)^2}\,dy = \frac{\pi}{4}.$$

On vérifie de même que $\int_0^1 dy\int_0^1 \frac{x^2-y^2}{(x^2+y^2)^2}\,dx = -\frac{\pi}{4}$.

Le théorème de Fubini implique alors que f n'est pas intégrable pour la mesure produit des mesures de Lebesgue sur $[0,1]\times[0,1]$.

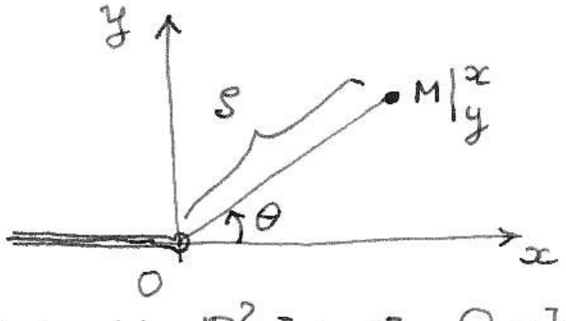
4.3 - THÉORÈME DU CHANGEMENT DE VARIABLE

4.3.1. Soit U un ouvert de $\mathbb{R}^n$ dont le point courant est noté y. Faire un changement de variable consiste à écrire $y = \phi(x)$, où $\phi : \Omega \to U$ est un difféomorphisme de classe C^1 d'un ouvert Ω de $\mathbb{R}^n$ sur l'ouvert U. Cela signifie que ϕ est un homéomorphisme de Ω sur U qui est de classe C^1 ainsi que $\phi^{-1} : U \to \Omega$. On note $D\phi(x)$ la différentielle de ϕ au point x ; c'est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^n$ qui est inversible, d'inverse $D(\phi^{-1})(\phi(x))$ en vertu de la règle de différentiation des fonctions composées. On notera $J_\phi(x) = \det(D\phi(x))$ le Jacobien de ϕ au point x.

Par exemple, le passage en coordonnées polaires consiste, pour tout point $(x,y) \in \mathbb{R}^2 -]-\infty, 0]$ à poser :

$$\begin{cases} x = \rho\cos\theta \\ y = \rho\sin\theta \end{cases}$$

où $\rho = \sqrt{x^2+y^2}$ et où θ est l'angle $(-\pi < \theta < \pi)$ que fait le vecteur de coordonnées (x,y) avec l'axe des x (cf. figure ci-contre)



Ici, $U = \mathbb{R}^2 -]-\infty, 0]$, $\Omega =]0, +\infty[\times]-\pi, \pi[$ et le changement de variable est le difféomorphisme $\phi : \Omega \to U$ défini par :

$$\phi(\rho, \theta) = (\rho\cos\theta, \rho\sin\theta).$$

On a :

a) La différentielle $D\phi(\rho,\theta)$ a pour matrice dans la base canonique de $\mathbb{R}^2$:

$$\begin{pmatrix} \frac{\partial\phi_1}{\partial\rho}(\rho,\theta), & \frac{\partial\phi_1}{\partial\theta}(\rho,\theta) \\ \frac{\partial\phi_2}{\partial\rho}(\rho,\theta), & \frac{\partial\phi_2}{\partial\theta}(\rho,\theta) \end{pmatrix} = \begin{pmatrix} \cos\theta & -\rho\sin\theta \\ \sin\theta & \rho\cos\theta \end{pmatrix}$$

où $\phi_1(\rho,\theta) = \rho\cos\theta$, $\phi_2(\rho,\theta) = \rho\sin\theta$;

b) Le Jacobien de ϕ au point (ρ,θ) est ici :

$$J_\phi(\rho,\theta) = \rho$$

Le but de cette section est d'exprimer l'intégrale $\int_U f(y)dy$ de toute fonction f intégrable sur U pour la mesure de Lebesgue dy, dans la nouvelle variable x qui est définie par le changement de variable $y = \phi(x)$.
A cet effet, on utilise le résultat suivant :

4.3.2- Théorème. Soient U, Ω deux ouverts de $\mathbb{R}^n$ et $\phi : \Omega \to U$ un C^1-difféomorphisme de Ω sur U. Soit $f : U \to \mathbb{R}$ une fonction numérique sur U. Alors, f est intégrable pour la mesure de Lebesgue sur U si et seulement si la fonction $x \longmapsto f(\phi(x))|J_\phi(x)|$ est intégrable pour la mesure de Lebesgue sur Ω, et on a :

$$\int_U f(y)dy = \int_\Omega f(\phi(x))|J_\phi(x)|dx.$$

Ainsi, dans le cas du passage en coordonnées polaires sur $\mathbb{R}^2$, une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ est intégrable pour la mesure de Lebesgue sur $\mathbb{R}^2$ si et seulement si elle est intégrable sur $\mathbb{R}^2 -]-\infty,0]$ (car $]-\infty,0]$ est de mesure nulle), et donc, d'après le théorème ci-dessus, si et seulement si la fonction $(\rho,\theta) \longmapsto f(\rho\cos\theta, \rho\sin\theta)\rho$ est intégrable pour la mesure de Lebesgue sur $]0,+\infty[\times]-\pi,\pi[$, et on a alors :

$$\iint_{\mathbb{R}^2} f(x,y)\,dxdy = \iint_{\mathbb{R}^2 -]-\infty,0]} f(x,y)\,dxdy = \int_0^{+\infty}\int_{-\pi}^{\pi} f(\rho\cos\theta, \rho\sin\theta)\rho\, d\rho\, d\theta$$

$$= \int_{-\pi}^{\pi}\int_0^{+\infty} f(\rho\cos\theta, \rho\sin\theta)\rho\, d\rho\, d\theta.$$

On dit parfois que « l'élément d'aire de $\mathbb{R}^2$ est $dxdy$ en coordonnées cartésiennes, et $\rho\, d\rho\, d\theta$ en coordonnées polaires ».

La démonstration du théorème 4.3.2 utilise le lemme suivant :

<u>4.3.3- Lemme</u> - <u>Pour tout Borélien B de $\mathbb{R}^n$ et tout isomorphisme linéaire T de $\mathbb{R}^n$, on a :</u>

$$\lambda_n(T(B)) = |\det(T)|\lambda_n(B),$$

<u>où λ_n désigne la mesure de Lebesgue sur $\mathbb{R}^n$.</u>

<u>Démonstration</u>. Montrons tout d'abord la formule indiquée lorsque B est un pavé P de $\mathbb{R}^n$. Comme $\lambda_n(a+P) = \lambda_n(P)$ et $\lambda_n(T(a+P)) = \lambda_n(T(a)+T(P)) = \lambda_n(T(P))$ pour tout $a \in \mathbb{R}^n$, on peut supposer sans perte de généralité que

$$P = \prod_{i=1}^{n} [0, a_i].$$

Comme $T(P)$ est le pavé n-dimensionnel engendré par les vecteurs $T(a_i e_i)$, où $(e_1, \dots, e_n)$ est la base canonique de $\mathbb{R}^n$, on a :

$$\begin{aligned}\lambda_n(T(P)) &= |\det(T(a_1e_1), \dots, T(a_ne_n))| \\ &= |\det(T)|\,|\det(a_1e_1, \dots, a_ne_n)| \\ &= |\det(T)|\,|a_1 \cdots a_n| \\ &= |\det(T)|\,\lambda_n(P),\end{aligned}$$

et la formule désirée est démontrée pour tout pavé P de $\mathbb{R}^n$. Soit maintenant B un Borélien de $\mathbb{R}^n$, et recouvrons-le par une réunion dénombrable de pavés P_k :

$$B \subset \bigcup_{k \geq 1} P_k .$$

On a : $\quad T(B) \subset \bigcup_{k\geq 1} T(P_k)$, d'où :

$$\lambda_n(T(B)) \leq \sum_{k\geq 1} \lambda_n(T(P_k)) = |\det(T)| \sum_{k \geq 1} \lambda_n(P_k).$$

Or $\lambda_n(B) = \inf\{\sum_{k\geq 1} \lambda_n(P_k) \mid B \subset \bigcup_{k\geq 1} P_k\ ;\ P_k = \text{pavé de } \mathbb{R}^n\}$, de sorte que l'on déduit de ce qui précède :

$$\lambda_n(T(B)) \leqq |\det(T)|\, \lambda_n(B).$$

En remplaçant dans cette formule B par $T(B)$ et T par T^{-1}, on obtient :

$$\lambda_n(B) = \lambda_n(T^{-1}(T(B))) \leqq |\det(T)^{-1}|\, \lambda_n(T(B)),$$

d'où $\quad |\det(T)|\, \lambda_n(B) \leqq \lambda_n(T(B))$ et donc :

$$\lambda_n(T(B)) = |\det(T)|\, \lambda_n(B). \quad \blacksquare$$

4.3.4- Démonstration du théorème 4.3.2.

Soit $\phi : \Omega \to U$ un C^1-difféomorphisme de Ω sur U.

1ère étape. Montrons que, pour tout pavé P borné de $\mathbb{R}^n$ tel que $\overline{P} \subset \Omega$, on a :

$$\lambda_n(\phi(P)) \leqq \int_P |J_\phi(x)|\,dx.$$

A cet effet, on peut supposer que $P = \prod_{i=1}^n [a_i, b_i]$ avec $a_i, b_i \in \mathbb{Q}$. En effet, si $P = \prod_{i=1}^n (\alpha_i, \beta_i)$ est un pavé borné tel que $\overline{P} = \prod_{i=1}^n [\alpha_i, \beta_i] \subset \Omega$, il existe pour tout i une suite croissante $(a_{m,i})_m$ de rationnels convergeant vers α_i et une suite décroissante de rationnels $(b_{m,i})_m$ convergeant vers β_i telles que

$$P_m = \prod_{i=1}^n [a_{m,i}, b_{m,i}] \subset \Omega.$$

Si $\lambda_n(\phi(P_m)) \leqq \int_{P_m} |J_\phi(x)|\,dx$, on en déduit que :

$$\lambda_n(\phi(P)) \leqq \lambda_n(\phi(P_m)) \leqq \int_{P_m} |J_\phi(x)|\,dx.$$

Comme $\overline{P} = \bigcap_m P_m$, on en déduit que

$$\lambda_n(\phi(P)) \leqq \lim_{m\to\infty} \int_{P_m} |J_\phi(x)|\,dx = \int_{\overline{P}} |J_\phi(x)|\,dx,$$

et donc

$$\lambda_n(\phi(P)) \leqq \int_P |J_\phi(x)|\,dx,$$ puisque $\overline{P} \setminus P$ est négligeable. On peut donc supposer que $P = \prod_{i=1}^n [a_i, b_i]$ avec $a_i, b_i \in \mathbb{Q}$. Dans ce cas, il existe un entier $q \geqq 1$ et des entiers $p_i \geqq 1$ tels que

$b_i - a_i = \frac{p_i}{q}$, de sorte que l'on peut diviser P en une réunion disjointe de $N = p_1 p_2 \cdots p_n$ petits pavés cubiques de côtés $\frac{1}{q}$. En divisant ces petits pavés cubiques en une réunion disjointe de pavés cubiques de côtés $\frac{1}{2^m q}$, on peut décomposer P en somme finie disjointe de petits pavés cubiques ayant tous une arête de même longueur s aussi petite que désiré.

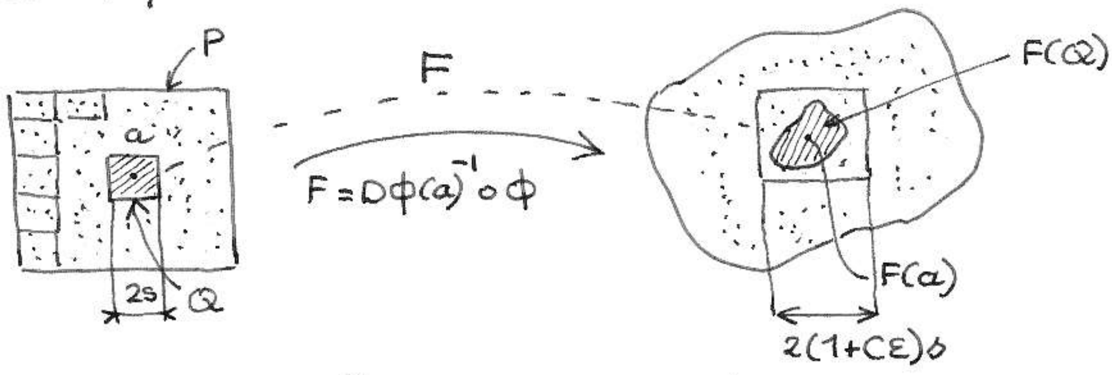
Fixons $P = \prod_{i=1}^{n} [a_i, b_i]$ avec $a_i, b_i \in \mathbb{Q}$. Comme l'application $x \longmapsto \|D\phi(x)^{-1}\|$ est continue sur P qui est compact, elle est bornée. Posons (*)

$$C = \sup_{x \in P} \|D\phi(x)^{-1}\|.$$

Soit $\varepsilon > 0$ fixé. Comme les fonctions $x \longmapsto D\phi(x)$ et $x \longmapsto J_\phi(x) = \det(D\phi(x))$ sont continues sur P qui est compact, elles sont uniformément continues et il existe $\delta > 0$ tel que l'on ait :

$$x, y \in P \text{ et } \|x - y\| \leqq \delta \Longrightarrow \begin{cases} \|D\phi(x) - D\phi(y)\| \leqq \varepsilon \\ \text{et} \\ |J_\phi(x) - J_\phi(y)| \leqq \varepsilon. \end{cases}$$

Décomposons P en une réunion disjointe de petits pavés cubiques Q ayant tous une arête de même longueur $2s$ avec $s \leqq \delta$. Considérons l'un de ces pavés cubiques Q, et notons a son centre.



Soit $F : \Omega \to \mathbb{R}^n$ la fonction définie par :

$$F(x) = D\phi(a)^{-1}(\phi(x))$$

et montrons que $F(Q)$ est contenue dans un cube de centre

(*) On pose, pour $x \in \mathbb{R}^n$, $\|x\| = \sup_{1 \leqq i \leqq n} |x_i|$. Si $T \in \mathcal{L}(\mathbb{R}^n)$ est un endomorphisme de $\mathbb{R}^n$, on pose :

$$\|T\| = \sup_{x \in \mathbb{R}^n, \|x\| \leqq 1} \|Tx\|.$$

$F(a)$ et d'arête $2(1+C\varepsilon)s$. A cet effet, notons que F est de classe C^1, de différentielle :

$$DF(x) = D\Phi(a)^{-1} D\Phi(x)$$
$$= D\Phi(a)^{-1}\left[D\Phi(x) - D\Phi(a)\right] + I .$$

Pour $x \in Q$, on a $\|x-a\| \leqq s \leqq \delta$, d'où $\|D\Phi(x) - D\Phi(a)\| \leqq \varepsilon$, et donc :

$$\|DF(x)\| \leqq \|D\Phi(a)^{-1}\| \, \|D\Phi(x) - D\Phi(a)\| + 1$$
$$\leqq C\varepsilon + 1 .$$

D'après le théorème des accroissements finis, appliqué à F sur le cube Q qui est convexe, on a :

$$\|F(x) - F(a)\| \leqq \|x-a\| \, (C\varepsilon+1) \leqq s\,(C\varepsilon+1)$$

quel que soit $x \in Q$. Il s'ensuit que $F(Q)$ est contenu dans le cube de centre $F(a)$ et d'arête $2s(C\varepsilon+1)$, d'où :

$$\lambda_n(F(Q)) \leqq (C\varepsilon+1)^n (2s)^n = (C\varepsilon+1)^n \lambda_n(Q)$$

Comme $F(Q)$ est l'image de $\Phi(Q)$ par l'application linéaire inversible $D\Phi(a)^{-1}$, il résulte du lemme 4.3.3 que

$$\lambda_n(F(Q)) = |\det(D\Phi(a))^{-1}| \, \lambda_n(\Phi(Q)),$$

et donc

$$\lambda_n(\Phi(Q)) \leqq (C\varepsilon+1)^n \, |J_\phi(a)| \, \lambda_n(Q)$$
$$= (C\varepsilon+1)^n \int_Q |J_\phi(a)| \, dx .$$

Or, pour $x \in Q$, on a $|J_\phi(x) - J_\phi(a)| \leqq \varepsilon$, de sorte que

$$|J_\phi(a)| \leqq |J_\phi(x)| + \varepsilon \quad \text{et donc}$$

$$\int_Q |J_\phi(a)| \, dx \leqq \int_Q |J_\phi(x)| \, dx + \varepsilon \lambda_n(Q).$$

Il s'ensuit que :

$$\lambda_n(\Phi(Q)) \leqq (C\varepsilon+1)^n \left[\int_Q |J_\phi(x)| \, dx + \varepsilon \lambda_n(Q) \right],$$

d'où l'on déduit, par sommation sur les petits cubes Q :

$$\lambda_n(\Phi(P)) \leqq (C\varepsilon+1)^n \left[\int_P |J_\Phi(x)|\,dx + \varepsilon\,\lambda_n(P) \right].$$

En faisant tendre ε vers 0, on obtient finalement :

$$\lambda_n(\Phi(P)) \leqq \int |J_\Phi(x)|\,dx,$$

ce qui achève la ~~P~~ première étape.

2ème étape. Montrons que l'on a, pour tout Borélien B de U : $\lambda_n(B) \leqq \int_{\Phi^{-1}(B)} |J_\Phi(x)|\,dx$.

Soit A un Borélien de Ω. Pour toute suite $(P_1, P_2, \ldots)$ de pavés de Ω telle que

$$A \subset \bigcup_{m \geqq 1} P_m,$$

on a en vertu de la 1ère étape :

$$(1) \qquad \lambda_n(\Phi(A)) \leqq \sum_{m\geqq 1} \lambda_n(\Phi(P_m)) \leqq \sum_{m \geqq 1} \int_{P_m} |J_\Phi(x)|\,dx.$$

Or on a, pour la mesure positive Borélienne ν sur Ω définie par

$$\nu(A) = \int_A |J_\Phi(x)|\,dx :$$

$$\nu(A) = \mathrm{Inf}\left\{ \sum_{m\geqq 1} \nu(P_m) \;\middle|\; A \subset \bigcup_{m\geqq 1} P_m \;;\; P_m = \text{pavé de } \Omega \right\}.$$

En faisant varier le recouvrement de A par les P_m dans (1), on obtient donc :

$$\lambda_n(\Phi(A)) \leqq \int_A |J_\Phi(x)|\,dx.$$

Si B est un Borélien de U, la relation ci-dessus appliquée au Borélien $A = \Phi^{-1}(B)$ de Ω nous donne :

$$\lambda_n(B) \leqq \int_{\Phi^{-1}(B)} |J_\Phi(x)|\,dx,$$

ce qui achève la seconde étape.

3ème étape. Montrons que l'on a, pour toute fonction f Borélienne positive sur U :

$$\int_U f(y)\,dy = \int_\Omega f(\phi(x))\,|J_\phi(x)|\,dx .$$

Pour tout Borélien B de U, on a d'après l'étape 2 :

$$\int_U 1\!\!1_B(y)\,dy = \lambda_n(B) \leqq \int_{\phi^{-1}(B)} |J_\phi(x)|\,dx = \int_\Omega 1\!\!1_B(\phi(x))\,|J_\phi(x)|\,dx .$$

Il s'ensuit que, pour toute fonction Borélienne positive e simple, on a :

$$\int_U e(y)\,dy \leqq \int_\Omega e(\phi(x))\,|J_\phi(x)|\,dx$$

Si $f : U \to [0,+\infty]$ est Borélienne positive, il existe une suite croissante $(e_n)_n$ de fonctions Boréliennes positives simples telle que $f = \sup_n e_n$. De la relation

$$\int_U e_n(y)\,dy \leqq \int_\Omega e_n(\phi(x))\,|J_\phi(x)|\,dx \leqq \int_\Omega f(\phi(x))\,|J_\phi(x)|\,dx,$$

on déduit alors :

$$\int_U f(y)\,dy = \sup_n \int_U e_n(y)\,dy \leqq \int_\Omega f(\phi(x))\,|J_\phi(x)|\,dx.$$

Echangeons alors les rôles de U et Ω, et remplaçons ϕ par ϕ^{-1}. La relation ci-dessus, appliquée à la fonction Borélienne positive $g(x) = f(\phi(x))\,|J_\phi(x)|\,dx$ donne alors :

$$\int_\Omega g(x)\,dx \leqq \int_U g(\phi^{-1}(y))\,|J_{\phi^{-1}}(y)|\,dy, \text{ soit :}$$

$$\int_\Omega f(\phi(x))\,|J_\phi(x)|\,dx \leq \int_U f(y)\,|\det D\phi(\phi^{-1}(y))|\,|\det(D\phi^{-1}(y))|\,dy$$

$$= \int_U f(y)\,dy$$

car $D\phi(\phi^{-1}(y))\,D\phi^{-1}(y) = D(\phi\circ\phi^{-1})(y) = I$. Il s'ensuit que

$$\int_U f(y)\,dy = \int_\Omega f(\phi(x))\,|J_\phi(x)|\,dx,$$ ce qui achève la démonstration de la $3^{\text{ème}}$ étape.

Fin de la démonstration du théorème 4.3.2

De l'étape 3, il résulte immédiatement qu'un Borélien B est négligeable pour la mesure de Lebesgue si et seulement si $\phi^{-1}(B)$ est négligeable pour la mesure de Lebesgue. Il s'ensuit immédiatement que $N \subset U$ est négligeable pour la mesure de Lebesgue si et seulement si $\phi^{-1}(N) \subset \Omega$ est négligeable. Il s'ensuit alors qu'une fonction positive f sur U est Lebesgue-mesurable si et seulement si la fonction $x \longmapsto f(\phi(x))\,|J_\phi(x)|$ est Lebesgue-mesurable, et on a alors :

$$\int_U f(y)\,dy = \int_\Omega f(\phi(x))\,|J_\phi(x)|\,dx.$$

En utilisant la décomposition $f = f^+ - f^-$ d'une fonction quelconque et le fait qu'une fonction $f \geqq 0$ est intégrable sur U ssi $\int_U f(y)\,dy < +\infty$, i.e. ssi $\int_\Omega f(\phi(x))\,|J_\phi(x)|\,dx < +\infty$, c'est à dire ssi la fonction $x \longmapsto f(\phi(x))\,|J_\phi(x)|$ est intégrable, on déduit aisément de (1) le théorème de changement de variable pour les fonctions intégrables. ■

4.3.5- Exemple. Cherchons à quelle condition sur α la fonction $f_\alpha(x,y) = \dfrac{1}{(x^2+y^2)^\alpha}$ est intégrable sur le disque $D_R = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 < R^2\}$ $(R > 0)$.

Par passage en coordonnées polaires, f_α est intégrable sur D_R si et seulement si la fonction

$$(\rho,\theta) \longmapsto \frac{1}{\rho^{2\alpha}}\,\rho = \rho^{1-2\alpha}$$

est intégrable sur $]0,R[\times]-\pi,\pi[$. Cette dernière fonction est intégrable sur $]0,R[\times]-\pi,\pi[$ si et seulement si la fonction $\rho \longmapsto \rho^{1-2\alpha}$ est intégrable sur $]0,R[$, c'est à dire si et seulement si $\alpha < 1$. Dans ce cas, on a :

$$\iint_{x^2+y^2<R^2} \frac{dx\,dy}{(x^2+y^2)^\alpha} = 2\pi\int_0^R \rho^{1-2\alpha}\,d\rho = \frac{2\pi}{2(1-\alpha)} \times \frac{1}{R^{2(\alpha-1)}} = \frac{\pi}{(1-\alpha)R^{2(\alpha-1)}}.$$

FIN DU COURS DE CALCUL INTÉGRAL