1. THÉORIE GÉNÉRALE DE L'INTÉGRATION

1.1. Espaces mesurables

DÉFINITION 1.1.1. Soit X un ensemble. On[1] considère $\mathcal{M} \subset \mathcal{P}(X)$ une famille de parties de X. On dit que $\mathcal{M}$ est une tribu sur X si[2]

$$\begin{array}{llcl} (a) & A \in \mathcal{M} & \Longrightarrow & A^c \in \mathcal{M}, \\ \text{(b)} & (A_n \in \mathcal{M})_{n \in \mathbb{N}} & \Longrightarrow & \cup_{n \in \mathbb{N}} A_n \in \mathcal{M}, \\ \text{(c)} & X \in \mathcal{M}. & & \end{array}$$

Le couple $(X, \mathcal{M})$ est appelé espace mesurable.

DÉFINITION 1.1.2. Soient $(X_1, \mathcal{M}_1), (X_2, \mathcal{M}_2)$ deux espaces mesurables et $f : X_1 \to X_2$ une application. On dit que f est mesurable si pour tout $A_2 \in \mathcal{M}_2,\ f^{-1}(A_2) \in \mathcal{M}_1$. On note cette propriété symboliquement par $f^{-1}(\mathcal{M}_2) \subset \mathcal{M}_1$.

Les propriétés (a),(b) de la définition 1.1.1 impliquent immédiatement qu'une tribu est stable par intersection dénombrable. Par ailleurs la propriété (c) est une conséquence de (a), (b), pourvu que $\mathcal{M}$ soit non vide.

Dans la suite nous appellerons *dénombrable* tout ensemble D équipotent à une partie de $\mathbb{N}$, i.e. tel qu'il existe une injection de D dans $\mathbb{N}$. Si D est fini non vide, il peut être mis en bijection avec (i.e. est "équipotent à") $\{1, \ldots, n\}$ où n est le nombre d'éléments de D. Si D est infini (i.e. non fini) dénombrable, alors il est équipotent à $\mathbb{N}$: en effet, on peut considérer D comme une partie de $\mathbb{N}$, puis définir

$$d_1 = \min D,\ d_2 = \min D \backslash \{d_1\}, \ldots, d_{k+1} = \min D \backslash \{d_1, \ldots, d_k\}.$$

Comme l'ensemble D est infini et que $\mathbb{N}$ est bien ordonné (i.e. toute partie non vide possède un plus petit élément) cette définition a un sens pour tout entier $k \geq 1$. Maintenant si $d \in D$, on peut trouver un entier k tel que $d_k \leq d < d_{k+1}$ car la suite d_k est strictement croissante et on ne peut avoir $d_k < d < d_{k+1}$ sans contredire la définition de d_{k+1}, ce qui prouve que $d = d_k$ et que l'ensemble D est $\{d_k\}_{k \in \mathbb{N}}$ équipotent à $\mathbb{N}$. On montre facilement que $\mathbb{N}^* = \mathbb{N} \backslash \{0\}, 2\mathbb{N}, 2\mathbb{N}+1, \mathbb{Z}, \mathbb{N} \times \mathbb{N}$ sont équipotents à $\mathbb{N}$. Pour obtenir cette dernière propriété, il suffit de remarquer que

$$(p, q) \in \mathbb{N} \times \mathbb{N} \mapsto 2^p(2q+1) \in \mathbb{N}^*$$

est bijective[3]. Comme l'ensemble des nombres rationnels $\mathbb{Q}$ s'injecte dans $\mathbb{Z} \times \mathbb{Z}$, on déduit des remarques précédentes que $\mathbb{Q}$ est dénombrable ainsi que $\mathbb{Q}^d$ (d entier ≥ 1). On verra

[1] On notera $\mathcal{P}(X)$ l'ensemble des parties de X.

[2] On note ici A^c le complémentaire de A dans X : $A^c = \{x \in X,\ x \notin A\}$.

[3] La démonstration est laissée au lecteur.

dans les exercices qu'en revanche l'ensemble des nombres réels $\mathbb{R}$ est non dénombrable parce qu' équipotent à l'ensemble des parties de $\mathbb{N}$ (on montre facilement –cf.exercice 1.2.– que pour tout ensemble X, il n'y a pas de surjection de X sur $\mathcal{P}(X)$).

On peut donner quelques exemples simples de tribus. De manière générale, si X est un ensemble, $\{\emptyset, X\}$ est une tribu sur X ainsi que $\mathcal{P}(X)$. En outre, si $(A_k)_{1\le k\le n}$ est une partition de X (les A_k sont non vides, deux à deux disjoints, de réunion X), l'ensemble

$$\mathcal{M} = \{\cup_{k\in J} A_k\}_{J\subset\{1,\dots,n\}}$$

est une tribu sur X ; en effet, en posant $B(J) = \cup_{k\in J} A_k$, on trouve que $B(J)^c = B(J^c)$, ce qui donne la stabilité par complémentarité et la stabilité par réunion est évidente. Notons également que le cardinal de $\mathcal{M}$ (son nombre d'éléments) est 2^n puisque $\mathcal{M}$ est en bijection avec les parties de $\{1,\dots,n\}$. On se reportera à l'exercice 1.7. pour le cas d'une partition dénombrable.

Il est important de noter que si $(\mathcal{M}_i)_{i\in I}$ est une famille de tribus sur X, $\mathcal{M} = \cap_{i\in I}\mathcal{M}_i$ est aussi une tribu sur X: si $(A_n)_{n\in\mathbb{N}}$ est une suite de $\mathcal{M}$, donc de $\mathcal{M}_i$ pour tout i dans I alors $\cup_{n\in\mathbb{N}} A_n$ appartient à $\mathcal{M}_i$ pour tout $i \in I$, donc à $\mathcal{M}$. Le passage au complémentaire s'écrit de la même manière (et X appartient à $\mathcal{M}$ car à tous les $\mathcal{M}_i$). Comme une tribu sur X est incluse dans $\mathcal{P}(X)$, ceci permet de donner la définition suivante.

DÉFINITION 1.1.3. Soit X un ensemble et $\mathcal{F} \subset \mathcal{P}(X)$. On pose

$$\mathcal{M}(\mathcal{F}) = \bigcap_{\substack{\mathcal{M} \text{ tribu sur } X\\ \text{contenant } \mathcal{F}}} \mathcal{M}$$

et on dit que $\mathcal{M}(\mathcal{F})$ est la tribu engendrée par $\mathcal{F}$ (ou bien la plus petite tribu contenant $\mathcal{F}$).

LEMME 1.1.4. *Soient $(X_1, \mathcal{M}_1), (X_2, \mathcal{M}_2)$ des espaces mesurables avec $\mathcal{M}_2 = \mathcal{M}(\mathcal{F})$ et $f : X_1 \to X_2$ une application. Pour que f soit mesurable, il suffit (et il faut) que $f^{-1}(\mathcal{F}) \subset \mathcal{M}_1$.*

DÉMONSTRATION. On pose

$$\mathcal{N} = \{B \in \mathcal{M}_2, f^{-1}(B) \in \mathcal{M}_1\}$$

qui est une tribu sur X_2, car si $B \in \mathcal{N}$, $f^{-1}(B^c) = \big(f^{-1}(B)\big)^c \in \mathcal{M}_1$. D'autre part si $(B_n)_{n\in\mathbb{N}}$ est une suite de $\mathcal{N}$, $f^{-1}(\cup_{n\in\mathbb{N}} B_n) = \cup_{n\in\mathbb{N}} f^{-1}(B_n) \in \mathcal{M}_1$. En outre $X_2 \in \mathcal{N}$. Donc $\mathcal{N}$ est une tribu qui contient $\mathcal{F}$ si $f^{-1}(\mathcal{F}) \subset \mathcal{M}_1$. Ceci implique que

$$\mathcal{M}_2 = \mathcal{M}(\mathcal{F}) \subset \mathcal{N} \subset \mathcal{M}_2 \Longrightarrow \mathcal{M}_2 = \mathcal{N}. \quad \square$$

LEMME 1.1.5. *Soit* $(X, \mathcal{M})$ *un espace mesurable et* $f : X \to Y$ *une application. Alors l'ensemble* $\mathcal{N} = \{B \subset Y, f^{-1}(B) \in \mathcal{M}\}$ *est une tribu sur* Y.

DÉMONSTRATION. Si $B, B_n \in \mathcal{N}$, $f^{-1}(B^c) = f^{-1}(B)^c \in \mathcal{M}$ et

$$f^{-1}(\cup_n B_n) = \cup_n f^{-1}(B_n) \in \mathcal{M}.$$

Comme $Y \in \mathcal{N}$, on obtient le résultat. Notons que $\mathcal{N}$ est la plus grande tribu sur Y qui rende mesurable l'application f : si $(Y, \widetilde{\mathcal{N}})$ est un espace mesurable tel que f soit mesurable, alors si $B \in \widetilde{\mathcal{N}}$, la mesurabilité de f implique $f^{-1}(B) \in \mathcal{M}$, d'où $B \in \mathcal{N}$. □

LEMME 1.1.6. *On considère des espaces mesurables* $(X, \mathcal{M})$, $(Y, \mathcal{N})$, $(Z, \mathcal{T})$ *et*

$$X \xrightarrow{f} Y \xrightarrow{g} Z$$

des applications mesurables. Alors $g \circ f$ *est mesurable.*

DÉMONSTRATION. Pour $C \in \mathcal{T}$, $(g \circ f)^{-1}(C) = f^{-1}\big(g^{-1}(C)\big) \in \mathcal{M}$ car $g^{-1}(C) \in \mathcal{N}$ (g est mesurable) et f mesurable. □

Nous avons utilisé une identité sur les images réciproques qui est simple car

$$(1.1.1) \quad \begin{aligned} (g \circ f)^{-1}(C) &= \{x \in X, g\big(f(x)\big) \in C\} \\ &= \{x \in X, f(x) \in g^{-1}(C)\} = f^{-1}\big(g^{-1}(C)\big). \end{aligned}$$

1.2. Espaces mesurables et espaces topologiques

DÉFINITION 1.2.1. Soit X un ensemble. Une famille $\mathcal{O}$ de parties de X est une topologie sur X si

$$\begin{array}{lccc} \text{(a)} & O_i \in \mathcal{O} \text{ pour } i \in I & \Longrightarrow & \cup_{i \in I} O_i \in \mathcal{O}, \\ \text{(b)} & O_1, O_2 \in \mathcal{O} & \Longrightarrow & O_1 \cap O_2 \in \mathcal{O}, \\ \text{(c)} & \emptyset,\ X \in \mathcal{O}. & & \end{array}$$

Autrement dit la famille $\mathcal{O}$ est stable par réunion quelconque et par intersection finie[4]. Le couple $(X, \mathcal{O})$ est appelé espace topologique.

On appelle fermé dans un espace topologique un ensemble dont le complémentaire est ouvert. On définit également l'intérieur d'un sous-ensemble d'un espace topologique comme la réunion des ouverts qu'il contient, l'adhérence d'un sous-ensemble d'un espace

[4]On peut remarquer que la stabilité par réunion quelconque implique pour $I = \emptyset$ que l'ensemble vide est élément de $\mathcal{O}$. En outre, la stabilité par intersection finie implique pour un ensemble d'indice vide que $X \in \mathcal{O}$. Par suite l'axiome (c) ci-dessus est une conséquence de (a) et (b).

topologique comme l'intersection des fermés qui le contiennent. En notant $\bar{A}$ l'adhérence de A et $\mathring{A}$ son intérieur, on a

$$(1.2.1)\qquad (\mathring{A})^c = \Big[\bigcup_{\Omega \text{ ouvert } \subset A} \Omega\Big]^c = \bigcap_{\Omega \text{ ouvert } \subset A} \Omega^c = \bigcap_{F \text{ fermé } \supset A^c} F = \overline{A^c},$$

de sorte que si la frontière de A est définie comme $\bar{A}\backslash\mathring{A}$ et notée ∂A, on a

$$(1.2.2)\qquad \partial A = \bar{A} \cap (\mathring{A})^c = \bar{A} \cap \overline{A^c}.$$

On vérifie aussi facilement à partir de la définition que

$$(1.2.3)\qquad \text{intérieur}(A \cap B) = \mathring{A} \cap \mathring{B}, \qquad \text{adhérence}(A \cup B) = \bar{A} \cup \bar{B}.$$

En revanche, seules les inclusions suivantes sont vérifiées en général (et non les égalités)[5],

$$(1.2.4)\qquad \text{intérieur}(A \cup B) \supset \mathring{A} \cup \mathring{B}, \qquad \text{adhérence}(A \cap B) \subset \bar{A} \cap \bar{B}.$$

On dit que V est un voisinage d'un point x, s'il contient un ouvert contenant x. On montre facilement à partir des définitions qu'un ouvert est caractérisé par le fait qu'il est voisinage de tous ses points, qu'une intersection finie de voisinages d'un point x est encore un voisinage de x, qu'un ensemble contenant un voisinage de x est encore un voisinage de x.

Un exemple très important d'espace topologique est celui des espaces métriques, i.e. des espaces sur lesquels on a défini une distance, c'est à dire une application $d : X \times X \to \mathbb{R}_+$ telle que

$$(1.2.5)\qquad \begin{cases} d(x,y) = 0 \Longleftrightarrow x = y & \text{(séparation)},\\ d(x,y) = d(y,x) & \text{(symétrie)},\\ d(x,z) \le d(x,y) + d(y,z) & \text{(inégalité triangulaire)}.\end{cases}$$

Un ouvert dans un espace métrique est alors défini comme une réunion de boules "ouvertes" $B(x,r) = \{y \in X, d(y,x) < r\}$ ($x \in X, r > 0$ donnés). Pour montrer que cela donne une topologie, il suffit de remarquer que si $x \in B(x_1,r_1) \cap B(x_2,r_2)$ alors $B(x,r) \subset B(x_1,r_1) \cap B(x_2,r_2)$ avec $r = \min\big(r_1 - d(x,x_1), r_2 - d(x,x_2)\big)$ car

$$d(y,x) < r \Longrightarrow \begin{cases} d(y,x_1) \le d(y,x) + d(x,x_1) < r + d(x,x_1) \le r_1,\\ d(y,x_2) \le d(y,x) + d(x,x_2) < r + d(x,x_2) \le r_2.\end{cases}$$

[5]En prenant dans $\mathbb{C}$ l'intersection des demi-plans $\pm \operatorname{Re} z > 0$, on trouve un contre-exemple à la seconde égalité. Pour la première, il suffit de considérer en passant au complémentaire $\pm \operatorname{Re} z \ge 0$.

Par conséquent une intersection finie de boules ouvertes est réunion de boules ouvertes, ce qui prouve que l'ensemble défini ci-dessus est une topologie.

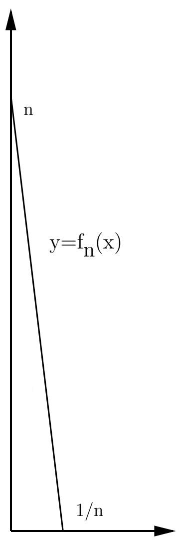
Les espaces $\mathbb{R}^d$ sont des espaces métriques pour la topologie définie par la distance euclidienne. De manière générale, tout espace vectoriel E sur $\mathbb{C}$ ou $\mathbb{R}$ muni d'une norme, i.e. d'une application

$$N : E \longrightarrow \mathbb{R}_+,\ \text{telle que}\ \begin{cases} N(x) = 0 \Longleftrightarrow x = 0 \\ N(\alpha x) = |\alpha| N(x),\ \text{pour}\ \alpha \in \mathbb{C}, x \in E, \\ N(x+y) \leq N(x) + N(y), \end{cases} \tag{1.2.6}$$

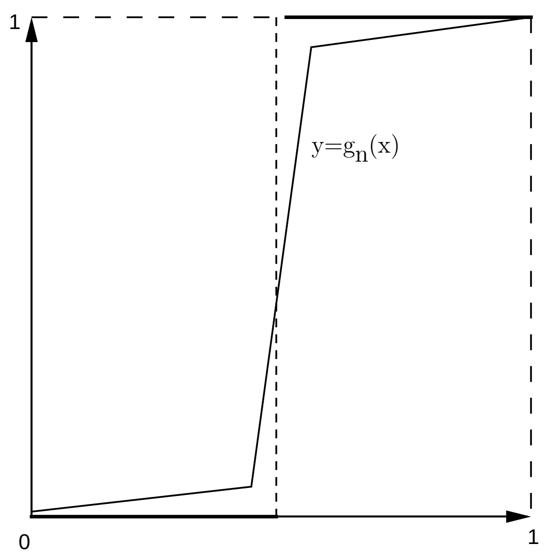
est un espace métrique pour la distance $N(x-y)$. C'est le cas par exemple de l'espace $C^0([0,1],\mathbb{R})$ des fonctions continues définies sur $[0,1]$ à valeurs réelles muni de la norme $\|f\|_\infty = \sup_{x\in[0,1]} |f(x)|$. Rappelons que sur $\mathbb{R}^d$, toutes les normes sont équivalentes[6]. L'espace $C^0([0,1],\mathbb{R})$ muni de la norme $\|\cdot\|_\infty$ définie ci-dessus est normé et complet[7]: on dit que c'est un espace de Banach. On peut remarquer que les normes sur $C^0([0,1],\mathbb{R})$ ne sont pas toutes équivalentes ; on peut considérer la norme (on vérifie facilement les axiomes des normes)

$$\|f\|_1 = \int_0^1 |f(t)| dt. \tag{1.2.7}$$

La suite f_n ci-dessous à gauche est bornée pour $\|\cdot\|_1$ et pas pour $\|\cdot\|_\infty$. Par ailleurs, la suite de fonctions continues g_n (à droite) est de Cauchy pour $\|\cdot\|_1$ et converge pour $\|\cdot\|_1$ vers la fonction discontinue $\mathbf{1}_{[1/2,1]}$.



SUITE f_n



SUITE g_n

[6]Les normes N_1 et N_2 sur E sont dites équivalentes s'il existe une constante $C > 0$ telle que pour tout $x \in E$, on ait $C^{-1}N_1(x) \leq N_2(x) \leq CN_1(x)$.

[7]Un espace vectoriel normé est dit complet si toute suite de Cauchy converge.

Nous pouvons remarquer qu'en général une topologie n'est pas stable par complémentarité : par exemple sur la droite réelle le complémentaire de l'ouvert $]0,+\infty[$ est $]-\infty,0]$ qui n'est pas ouvert car ne contient pas de boule ouverte contenant 0. En fait, un espace topologique est dit connexe (intuitivement d'un seul tenant) si les seuls ensembles à la fois ouverts et fermés sont le vide et l'espace entier. La situation la plus intéressante pour nous en théorie de l'intégration est donc de considérer un espace topologique et d'examiner la tribu engendrée par sa topologie.

DÉFINITION 1.2.2. Soit $(X,\mathcal{O})$ un espace topologique. La tribu de Borel sur X est la tribu engendrée par $\mathcal{O}$.

Il faut se rendre compte que cette définition, bien que parfaitement claire, nous donne peu de renseignements effectifs sur ce qu'est un borélien (élément de la tribu de Borel). Par exemple une réunion dénombrable de fermés (qu'on appellera un F_σ) est un borélien, comme par exemple $\mathbb{Q}$ l'ensemble des nombres rationnels, réunion dénombrable de points. Son complémentaire est une intersection dénombrable d'ouverts (on dit que c'est un G_δ): l'ensemble des nombres irrationnels est un G_δ de la droite réelle. Certains ensembles peuvent être à la fois F_σ et G_δ comme $[0,1]$ qui est fermé (donc F_σ) et G_δ car

$$[0,1] = \cap_{n\geq 1}]-\frac{1}{n}, 1+\frac{1}{n}[.$$

Néanmoins l'ensemble des irrationnels (un G_δ d'après ce qui précède) n'est pas un F_σ sinon on aurait pour des fermés F_n, $\mathbb{R}\backslash\mathbb{Q} = \cup_n F_n$. Comme $\mathbb{R}\backslash\mathbb{Q}$ ne contient aucun intervalle (car $\mathbb{Q}$ est dense dans $\mathbb{R}$) les F_n sont tous d'intérieurs vides. Finalement on pourrait écrire $\mathbb{R}$ comme réunion dénombrable de fermés d'intérieur vide. Or on pourra voir dans les exercices que le théorème de Baire assure que dans un espace métrique complet (comme c'est le cas pour $\mathbb{R}$), une réunion dénombrable de fermés d'intérieurs vides est aussi d'intérieur vide. L'égalité précédente décrivant les irrationnels comme F_σ est donc absurde.

PROPOSITION 1.2.3. *On considère $(X_1,\mathcal{O}_1)$, $(X_2,\mathcal{O}_2)$ des espaces topologiques et $\mathcal{B}_j$ les tribus boréliennes associées. Si $f : X_1 \to X_2$ est continue alors elle est mesurable.*

DÉMONSTRATION. La continuité signifie que $f^{-1}(\mathcal{O}_2) \subset \mathcal{O}_1$. Comme $\mathcal{B}_2$ est engendrée par $\mathcal{O}_2$, que $\mathcal{O}_1 \subset \mathcal{B}_1$, le lemme 1.1.4 donne que f est mesurable. □

Notons que la continuité de f en un point x_1 signifie que pour tout V_2 voisinage de $f(x_1)$, il existe un voisinage V_1 de x_1 tel que $f(V_1) \subset V_2$. Comme tout voisinage d'un point contient par définition un voisinage ouvert, on peut remplacer dans la propriété précédente le mot "voisinage" par "voisinage ouvert".

La continuité sur l'ensemble X_1 est la continuité en chaque point. Si f est continue sur X_1 et V_2 est un ouvert de X_2, si $x_1 \in f^{-1}(V_2)$, il existe un ouvert $V_1 \ni x_1$ tel que $f(V_1) \subset V_2$, d'où $V_1 \subset f^{-1}\big(f(V_1)\big) \subset f^{-1}(V_2)$, et $f^{-1}(V_2)$ est ouvert car voisinage de tous ses points. Réciproquement, si l'image réciproque par f de tout ouvert est ouverte, alors f est continue : si $x_1 \in X_1$, $x_2 = f(x_1)$ et V_2 est un voisinage ouvert de x_2, alors $V_1 = f^{-1}(V_2)$ est un ouvert de X_1 qui contient x_1. Il vient $f(V_1) = f\big(f^{-1}(V_2)\big) \subset V_2$, soit la continuité de f.

Il existe des fonctions continues en un seul point comme

$$f : \mathbb{R} \to \mathbb{R}, \qquad f(x) = \begin{cases} x & \text{si } x \text{ rationnel} \\ -x & \text{si } x \text{ irrationnel} \end{cases}, \qquad \text{continue seulement en } 0. \tag{1.2.8}$$

On peut démontrer (cf. exercices) que l'ensemble des points de discontinuité d'une fonction de $\mathbb{R}$ dans $\mathbb{R}$ est un F_σ, et que pour tout ensemble A qui est un F_σ, on peut trouver une fonction dont les points de discontinuité sont exactement A. Cela implique en particulier qu'il n'existe pas de fonction de $\mathbb{R}$ dans $\mathbb{R}$ dont les points de discontinuité soient les irrationnels. En revanche, on construit facilement (cf.exercices) une fonction discontinue sur $\mathbb{Q}$ continue sur $\mathbb{Q}^c$:

$$f(x) = \begin{cases} 1 & \text{si } x = 0, \\ 1/q & \text{si } x = p/q, p \in \mathbb{Z}^*, q \in \mathbb{N}^*, \text{ fraction irréductible,} \\ 0 & \text{si } x \text{ irrationnel.} \end{cases} \tag{1.2.9}$$

On peut remarquer par ailleurs que les ouverts de $\mathbb{R}$ sont des F_σ et plus précisément des réunions dénombrables d'intervalles fermés. Si U est un ouvert non vide de $\mathbb{R}$ et $x \in U$, il existe $\rho > 0$ tel que $]x-\rho, x+\rho[\subset U$; comme $\mathbb{Q}$ est dense dans $\mathbb{R}$, ceci implique qu'il existe $p, q \in \mathbb{Q}$ tels que $x - \rho < p < x < q < x + \rho$, et donc $[p, q] \subset U$. L'ouvert U est donc réunion d'intervalles fermés d'extrémités rationnelles. Or l'application

$$[p, q] \mapsto (p, q)$$

est injective de l'ensemble $\mathcal{Q}$ des intervalles fermés d'extrémités rationnelles dans $\mathbb{Q}\times\mathbb{Q}$ qui est équipotent à $\mathbb{N}$. Donc $\mathcal{Q}$ est (infini) et équipotent à une partie de $\mathbb{N}$, donc équipotent à $\mathbb{N}$. Ceci donne le résultat cherché. De manière générale on a en toute dimension le

LEMME 1.2.4. *Soit d un entier. Il existe une famille dénombrable de pavés compacts de $\mathbb{R}^d$ (produit d'intervalles compacts) telle que tout ouvert soit réunion(nécessairement dénombrable) d'une sous-famille de ces pavés. On peut choisir ces pavés comme produit d'intervalles compacts d'extrémités rationnelles.*

DÉMONSTRATION. Comme tout voisinage de $x = (x_1, \dots, x_d)$ dans $\mathbb{R}^d$ contient un cube

$$\{y \in \mathbb{R}^d, \max_{1 \le j \le d} |y_j - x_j| < \rho\}$$

avec $\rho > 0$, on peut trouver $p_j, q_j \in \mathbb{Q}$ tels que

$$x_j - \rho < p_j < x_j < q_j < x_j + \rho.$$

Par conséquent, tout voisinage de x contient un pavé compact d'extrémités rationnelles $\prod_{1\le j\le d}[p_j, q_j]$, et par suite tout ouvert est réunion de tels pavés. De même que dans la discussion qui précède le lemme, en utilisant la dénombrabilité de $\mathbb{Q}^{2d}$ on obtient le résultat du lemme. $\square$

La tribu des boréliens de $\mathbb{R}^d$ est donc engendrée par les pavés compacts ($\mathcal{M}$(pavés compacts) contient les ouverts, donc les boréliens et les pavés compacts étant des boréliens $\mathcal{M}$(pavés compacts) $\subset$ Boréliens) et comme $[p,q] = \cap_{n\ge 1}]p - 1/n, q + 1/n[$, cette tribu est aussi engendrée par les pavés ouverts. En outre, dans $\mathbb{R}$, comme

$$[p,q] = [p,+\infty[\cap] - \infty, q] = [p,+\infty[\cap]q,+\infty[^c = \cap_{n\ge 1}]p - 1/n, +\infty[\cap]q,+\infty[^c,$$

la tribu des boréliens sur $\mathbb{R}$ est engendrée par les intervalles $(]a,+\infty[)_{a\in\mathbb{R}}$ et donc aussi par les intervalles $(]-\infty,a])_{a\in\mathbb{R}}$ ou bien (comme $]a,+\infty[= \cup_{n\ge 1}[a + 1/n, +\infty[$) par les intervalles $([a,+\infty[)_{a\in\mathbb{R}}$ et donc aussi par les intervalles $(]-\infty,a[)_{a\in\mathbb{R}}$. En utilisant le lemme 1.1.4. pour vérifier la mesurabilité de $f : X \to \mathbb{R}$ il suffit de vérifier la mesurabilité des $f^{-1}([a,+\infty[)$.

Par exemple si X est une partie mesurable de $\mathbb{R}$ et f est monotone sur X alors elle est mesurable. En effet, si f est croissante, $b \in \mathbb{R}$ avec $A = f^{-1}(]b,+\infty[) \neq \emptyset$, on a

$$A = \cup_{a\in A}\big([a,+\infty[\cap X\big)$$

car si $X \ni x \ge a \in A$ on a $f(x) \ge f(a) > b$ et donc $x \in A$. Par suite, on a

$$]\inf A, +\infty[\cap X \subset A \subset [\inf A, +\infty[\cap X,$$

car c'est vrai si $\inf A = -\infty$ et si $\inf A > -\infty$, $\inf A$ est réel car $A \neq \emptyset$ et pour tout $\epsilon > 0$ on peut trouver $a_\epsilon \in A$ tel que

$$\inf A \le a_\epsilon < \inf A + \epsilon \qquad \Longrightarrow \qquad]\inf A, +\infty[\cap X \subset \cup_{\epsilon>0}[a_\epsilon, +\infty[\cap X \subset A.$$

Que $\inf A$ appartienne à A ou non, on trouve que $A =]\inf A, +\infty[\cap X$ ou bien $A = [\inf A, +\infty[\cap X$, mesurable dans les deux cas.

THÉORÈME 1.2.5. *Soient* $(X, \mathcal{M})$, $(Y, \mathcal{N})$ *des espaces mesurables et* $\mathbb{R}^d$ *muni de la tribu des boréliens. Soient* $u_1, \ldots, u_d$ *des applications mesurables de* X *dans* $\mathbb{R}$ *et* $\Phi : \mathbb{R}^d \to Y$ *mesurable. Alors l'application*

$$\begin{array}{ccc} X & \to & Y \\ x & \mapsto & \Phi\big(u_1(x), \ldots, u_d(x)\big) \end{array}$$

est mesurable. En particulier $f : X \to \mathbb{C}$ *est mesurable si (et seulement si)* $\operatorname{Re} f, \operatorname{Im} f$ *le sont. Alors* $|f|$ *est aussi mesurable. Si* $f, g : X \to \mathbb{C}$ *sont mesurables, c'est aussi le cas de* $f + g, fg$. *En outre si* $A \in \mathcal{M}$ *la fonction indicatrice de* A, *notée* $\mathbf{1}_A$ *(qui vaut 1 sur* A *et 0 ailleurs) est mesurable.*

DÉMONSTRATION. D'après le lemme de composition 1.1.6, il suffit d'examiner la mesurabilité de $V(x) = \big(u_1(x), \ldots, u_d(x)\big)$ de X dans $\mathbb{R}^d$. D'après le lemme 1.2.4. et le lemme 1.1.4, il suffit de vérifier que l'image réciproque par V d'un pavé compact de $\mathbb{R}^d$ est dans $\mathcal{M}$. Pour cela on remarque

$$V^{-1}\Big(\prod_{1 \le j \le d} [a_j, b_j]\Big) = \bigcap_{1 \le j \le d} u_j^{-1}([a_j, b_j]) \in \mathcal{M}$$

par mesurabilité des u_j. Les autres affirmations du théorème sont des conséquences immédiates de ce qui précède. $\square$

On obtient facilement la généralisation suivante du théorème 1.2.5.

THÉORÈME 1.2.5'. *Soient* $(X, \mathcal{M})$, $(Y, \mathcal{N})$ *des espaces mesurables et* T *un espace métrique séparable muni de la tribu des boréliens. Soient* $u_1, \ldots, u_d$ *des applications mesurables de* X *dans* T *et* $\Phi : T^d \to Y$ *mesurable. Alors l'application*

$$\begin{array}{ccc} X & \to & Y \\ x & \mapsto & \Phi\big(u_1(x), \ldots, u_d(x)\big) \end{array}$$

est mesurable.

DÉMONSTRATION. D'après le lemme de composition 1.1.6, il suffit d'examiner la mesurabilité de $V(x) = \big(u_1(x), \ldots, u_d(x)\big)$ de X dans T^d. D'après le lemme 1.1.4, il suffit de vérifier que l'image réciproque par V d'un produit de boules ouvertes de T est dans $\mathcal{M}$; en effet, si Ω est un ouvert de T^d et $x = (x_1, \ldots, x_d) \in \Omega$, il existe $r_1, \ldots, r_d$ strictement positifs (que l'on peut supposer rationnels) tels que le produit de boules ouvertes

$$B(x_1, r_1) \times \cdots \times B(x_d, r_d) \ni x$$

soit inclus dans Ω. Si $\mathbb{D}$ est une partie dénombrable dense de T, on peut trouver $y_1, \dots, y_d \in \mathbb{D}$ avec $\operatorname{dist}(x_j, y_j) < r_j/2$. Alors la boule $B(y_j, r_j/2)$ vérifie

$$x_j \in B(y_j, r_j/2) \subset B(x_j, r_j)$$

car $\operatorname{dist}(z, y_j) < r_j/2$ implique $\operatorname{dist}(z, x_j) \le \operatorname{dist}(z, y_j) + \operatorname{dist}(y_j, x_j) < r_j/2 + r_j/2$ d'où $z \in B(x_j, r_j)$. Par suite, l'ouvert Ω est réunion de produits

$$B(y_1, \rho_1) \times \cdots \times B(y_d, \rho_d), \qquad y_j \in \mathbb{D}, \rho_j \in \mathbb{Q}.$$

Il y a une surjection de $\mathbb{D}^d \times \mathbb{Q}^d$ sur l'ensemble $\mathcal{P}$ des ces parties et $\mathcal{P}$ est donc dénombrable. Considérons

$$V^{-1}\big(\prod_{1 \le j \le d} B(y_j, \rho_j)\big) = \bigcap_{1 \le j \le d} u_j^{-1}(B(y_j, \rho_j)) \in \mathcal{M}$$

par mesurabilité des u_j. Les autres affirmations du théorème sont des conséquences immédiates de ce qui précède. $\square$

LEMME 1.2.6. *Soit $(X, \mathcal{M})$ un espace mesurable. Si $f : X \to \mathbb{C}$ est mesurable, il existe une fonction mesurable $\alpha : X \to \mathbb{C}$ avec $|\alpha| \equiv 1$, telle que $f = \alpha |f|$.*

DÉMONSTRATION. L'ensemble $E = f^{-1}(\{0\})$ est élément de $\mathcal{M}$ par mesurabilité de f et $\mathbf{1}_E$ est mesurable. On remarque que $f(x) + \mathbf{1}_E(x)$ est toujours non nul (et vaut 1 pour $x \in E$, $f(x) \ne 0$ sinon). Posons

$$\alpha(x) = \frac{f(x) + \mathbf{1}_E(x)}{|f(x) + \mathbf{1}_E(x)|}.$$

La fonction α est mesurable comme composée des fonctions mesurables

$$\begin{array}{ccccc} X & \overset{\text{mesurable}}{\rightarrow} & \mathbb{C}^* & \overset{\text{continue}}{\rightarrow} & \mathbf{S}^1 \\ x & \mapsto & f(x) + \mathbf{1}_E(x) = t & \mapsto & t/|t| \end{array},$$

et on a $f(x) + \mathbf{1}_E(x) = \alpha(x)|f(x) + \mathbf{1}_E(x)|$, d'où pour $x \notin E$, $f(x) = \alpha(x)|f(x)|$ et pour $x \in E$, $f(x) = 0 = \alpha(x)|f(x)|$. $\square$

Remarquons que si $(X, \mathcal{M})$ est un un espace mesurable et $A \subset X$, l'ensemble

$$\mathcal{M}_A = \{M \cap A\}_{M \in \mathcal{M}} \tag{1.2.10}$$

est une tribu sur $\mathcal{M}$, dite tribu trace, rendant l'injection canonique mesurable. Si en outre $A \in \mathcal{M}$, $\mathcal{M}_A = \{M \in \mathcal{M}, M \subset A\}$.

DÉFINITION 1.2.7. On appelle droite achevée l'espace topologique noté $\overline{\mathbb{R}}$ obtenu en adjoignant à $\mathbb{R}$ deux éléments, notés $-\infty, +\infty$. La topologie contient les ouverts de $\mathbb{R}$ et les ensembles

$$]a, +\infty[\cup\{+\infty\}, \qquad \{-\infty\}\cup]-\infty, a[$$

(que l'on notera respectivement $]a, +\infty]$ et $[-\infty, a[$). La relation d'ordre sur $\overline{\mathbb{R}}$ fait de $-\infty$ le plus petit élément et de $+\infty$ le plus grand élément. Cet ordre est compatible avec la topologie car les ouverts sont réunions d'intervalles.

On montre facilement que $\overline{\mathbb{R}}$ est homéomorphe à l'intervalle [0,1] (i.e. on peut trouver une bijection bicontinue ψ de $\overline{\mathbb{R}}$ sur [0,1]), par exemple en prolongeant par continuité la fonction tangente à l'intervalle $[-\pi/2, \pi/2]$ et la fonction Arctg à $\overline{\mathbb{R}}$ (en outre les intervalles fermés bornés de $\mathbb{R}$ sont trivialement homéomorphes via un homéomorphisme croissant). Cet homéomorphisme respecte les relations d'ordre i.e. est croissant On remarque aussi que toute suite monotone (x_n) de $\overline{\mathbb{R}}$ est convergente car $\psi(x_n)$ est monotone dans [0,1] donc convergente (car croissante majorée ou décroissante minorée) et la continuité de ψ^{-1} assure le résultat. Comme $\overline{\mathbb{R}}$ est compact, toute partie non vide possède une borne supérieure et une borne inférieure.

DÉFINITION 1.2.8. Soit $(x_n)_{n\in\mathbb{N}}$ une suite de $\overline{\mathbb{R}}$. Les suites

$$(\inf_{k\geq n} x_k)_{n\in\mathbb{N}} \text{ et } (\sup_{k\geq n} x_k)_{n\in\mathbb{N}}$$

sont monotones (croissante pour la première, décroissante pour l'autre). On pose

$$\begin{aligned} \liminf x_n &= \lim_{n\to+\infty}\big(\inf_{k\geq n} x_k\big) &= \sup_{n\in\mathbb{N}}\big(\inf_{k\geq n} x_k\big), \\ \limsup x_n &= \lim_{n\to+\infty}\big(\sup_{k\geq n} x_k\big) &= \inf_{n\in\mathbb{N}}\big(\sup_{k\geq n} x_k\big). \end{aligned}$$

PROPOSITION 1.2.9. *Soit $(x_n)_{n\in\mathbb{N}}$ une suite de $\overline{\mathbb{R}}$. Alors* $\liminf x_n$ *(resp.*$\limsup x_n$*) est la plus petite (resp. plus grande) valeur d'adhérence de la suite. On a* $\liminf x_n \leq \limsup x_n$*) et l'égalité a lieu si et seulement si la suite converge vers cette valeur.*

DÉMONSTRATION. Par l'homéomorphisme strictement croissant ψ, on peut se ramener au cas d'une suite $(x_n)_{n\in\mathbb{N}}$ de $[0,1]$. Si y est une valeur d'adhérence de la suite, i.e. la limite d'une suite extraite $(x_{n_k})_{k\in\mathbb{N}}$ $(n_0 < n_1 < n_2 < \dots < n_k < n_{k+1} < \dots)$, alors

$$y \underset{k\to+\infty}{\longleftarrow} x_{n_k} \leq \sup_{l\geq n_k} x_l \underset{k\to+\infty,}{\longrightarrow} \limsup x_n$$

la seconde limite provenant de l'extraction d'une suite convergente. On obtient par passage à la limite que $y \leq \limsup x_n$ et de même $y \geq \liminf x_n$. En outre, $\limsup x_n$ est une valeur d'adhérence car pour tout $\epsilon > 0, N \geq 1$ on peut trouver $n_\epsilon \geq N$ avec

$$\limsup x_n = \inf_n (\sup_{k \geq n} x_k) \leq \sup_{k \geq n_\epsilon} x_k < \limsup x_n + \epsilon,$$

et donc

$$\limsup x_n - \eta = \inf_n (\sup_{k \geq n} x_k) - \eta \leq \sup_{k \geq n_\epsilon} x_k - \eta < x_{n(\epsilon,\eta)} \leq \sup_{k \geq n_\epsilon} x_k < \limsup x_n + \epsilon.$$

Ceci démontre que pour tous ϵ, η strictement positifs, et tout $N \geq 1$, on peut trouver $n(\epsilon, \eta) \geq n_\epsilon \geq N$ avec

$$\limsup x_n - \eta < x_{n(\epsilon,\eta)} < \limsup x_n + \epsilon,$$

ce qui donne le résultat. Si la lim sup et la lim inf sont égales à l alors

$$l \underset{n \to +\infty}{\longleftarrow} \inf_{k \geq n} x_k \leq x_n \leq \sup_{k \geq n} x_k \underset{n \to +\infty,}{\longrightarrow} l$$

ce qui implique que x_n tend vers l. Si en revanche, $\liminf x_n < \limsup x_n$, la suite possède au moins deux valeurs d'adhérence distinctes et ne peut converger. Ceci achève la démonstration de la proposition. □

PROPOSITION 1.2.10. *L'addition et la multiplication des réels se prolongent*[8] *par continuité respectivement sur* $(\overline{\mathbb{R}} \times \overline{\mathbb{R}}) \backslash \{(+\infty, -\infty), (-\infty, +\infty)\}$ *et* $(\overline{\mathbb{R}} \times \overline{\mathbb{R}}) \backslash \{(0, \pm\infty), (\pm\infty, 0)\}$. *Soient* $(x_n), (y_n)$ *des suites de* $\overline{\mathbb{R}}$ *telles que* $x_n + y_n$, $\liminf x_n + \liminf y_n$ *et* $\limsup x_n + \limsup y_n$ *aient un sens. Alors on a les inégalités*[9]

$$\liminf x_n + \liminf y_n \leq \liminf (x_n + y_n) \leq \limsup (x_n + y_n) \leq \limsup x_n + \limsup y_n.$$

[8] Autrement dit $x+y$ a un sens pour $x \in \overline{\mathbb{R}}, y \in \overline{\mathbb{R}}$, pourvu que l'on évite la "forme indéterminée" $+\infty - \infty$. Idem pour le produit et $0.\infty$. La terminologie forme indéterminée se justifie par le fait qu'il n'y a pas de prolongement par continuité de l'addition sur $\mathbb{R}$ à $\overline{\mathbb{R}}$: si un tel prolongement existait, en considérant $x_n = -n + l, y_n = n$, on aurait pour toutes les valeurs du paramètre réel l

$$l = \lim(x_n + y_n) = \lim x_n + \lim y_n = +\infty - \infty.$$

[9] Les égalités sont fausses en général: prendre par exemple $x_n = (-1)^n/2, y_n = (-1)^{n+1}$.

DÉMONSTRATION. Supposons tout d'abord que les suites x_n et y_n soient bornées dans $\mathbb{R}$. Pour $k \geq n$, on a $x_k + y_k \leq \sup_{l\geq n} x_l + \sup_{l\geq n} y_l$ et par conséquent $\sup_{k\geq n}(x_k + y_k) \leq \sup_{l\geq n} x_l + \sup_{l\geq n} y_l$. Par suite, on a en passant à la limite pour $n \to +\infty$,

$$\limsup\,(x_n + y_n) \leq \limsup x_n + \overline{\lim} y_n.$$

En notant que

$$\liminf(-x_n) = \sup_n(\inf_{k\geq n}(-x_n)) = \sup_n(-\sup_{k\geq n} x_n) = -\inf_n(\sup_{k\geq n} x_n) == -\limsup x_n,$$

on trouve le résultat. Il reste à examiner les cas où les suites ne sont pas bornées dans $\mathbb{R}$; cette vérification est laissée au lecteur. □

Nous aurons besoin du résultat suivant dans la suite du cours.

LEMME 1.2.11. *Soit $(a_{k,l})_{k\in\mathbb{N},l\in\mathbb{N}}$ une suite double de $\overline{\mathbb{R}}_+$. Alors*

$$\sum_k\Big(\sum_l a_{kl}\Big) = \sum_l\Big(\sum_k a_{kl}\Big). \qquad \textit{Cette somme est notée simplement } \sum_{k,l} a_{kl}.$$

DÉMONSTRATION. Notons tout d'abord qu'on a vu que les séries de $\overline{\mathbb{R}}_+$ convergeaient vers leur supremum[10]. Par suite, on a pour tous K, L,

$$\sigma = \sum_k\Big(\sum_l a_{kl}\Big) = \sum_k \sup_{L\geq 0}\Big(\sum_{0\leq l\leq L} a_{kl}\Big) = \sup_K \sum_{0\leq k\leq K}\Big[\sup_{L\geq 0}\Big(\sum_{0\leq l\leq L} a_{kl}\Big)\Big] \geq$$
$$\sum_{0\leq k\leq K}\Big[\sup_{L\geq 0}\Big(\sum_{0\leq l\leq L} a_{kl}\Big)\Big] \geq \sum_{0\leq k\leq K}\Big[\sum_{0\leq l\leq L} a_{kl}\Big] = \sum_{0\leq l\leq L}\Big[\sum_{0\leq k\leq K} a_{kl}\Big].$$

D'où il vient, pour tout L,

$$\sigma \geq \sum_{0\leq l\leq L}\Big[\sum_k a_{kl}\Big] \Longrightarrow \sigma \geq \sum_l\Big(\sum_k a_{kl}\Big),$$

ce qui donne le résultat en intervertissant les rôles de k et l. □

REMARQUE 1.2.12. L'addition des réels se prolonge par continuité à $\overline{\mathbb{R}}_+ \times \overline{\mathbb{R}}_+$; elle est donc associative, commutative, avec élément neutre 0. La multiplication des réels ne se

[10] En particulier, cela donne un sens aux sommes de l'énoncé.

prolonge pas par continuité à $\overline{\mathbb{R}}_+ \times \overline{\mathbb{R}}_+$ mais seulement à $\overline{\mathbb{R}}_+^* \times \overline{\mathbb{R}}_+^*$. On pourrait adopter la convention $0 \cdot \infty = 0$ et[11] on vérifie alors facilement que cette nouvelle multiplication est associative, commutative, avec élément neutre 1 sur $\overline{\mathbb{R}}_+^*$ et distributive par rapport à l'addition. On pourra également consulter la remarque 1.3.4 ci-dessous.

1.3. Structure des fonctions mesurables

PROPOSITION 1.3.1. *Soit $(X, \mathcal{M})$ un espace mesurable et $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables de X dans $\overline{\mathbb{R}}$. Alors les fonctions $\sup f_n$, $\inf f_n$, $\limsup f_n$, $\liminf f_n$ sont mesurables. En particulier, la limite simple d'une suite de fonctions mesurables est mesurable.*

DÉMONSTRATION. Soit en effet $g = \sup f_n$, i.e. $g(x) = \sup_{n\in\mathbb{N}} f_n(x)$. Pour $a \in \mathbb{R}$, on a

$$g^{-1}(]a, +\infty]) = \cup_{n\in\mathbb{N}} f_n^{-1}(]a, +\infty]).$$

En effet, on a puisque $a \in \mathbb{R}$

$$\sup_{n\in\mathbb{N}} f_n(x) = g(x) > a \Longleftrightarrow \exists n_0 \in \mathbb{N}, \text{tel que } f_{n_0}(x) > a.$$

Par conséquent $g^{-1}(]a, +\infty]) \in \mathcal{M}$. En vertu du lemme 1.1.4, c'est suffisant car la tribu des boréliens $\mathcal{B}_{\overline{\mathbb{R}}}$ de $\overline{\mathbb{R}}$ est égale à $\mathcal{T}$, tribu engendrée par les $]a, +\infty]$. Evidemment, comme $]a, +\infty]$ est un ouvert de $\overline{\mathbb{R}}$, on a

$$\mathcal{T} \subset \mathcal{B}_{\overline{\mathbb{R}}}.$$

Si ω est un ouvert de $\mathbb{R}$, il est réunion dénombrable d'intervalles fermés, d'après le lemme 1.2.4 ; or on a

$$[\alpha, \beta] = [-\infty, \beta] \cap [\alpha, +\infty] =]\beta, +\infty]^c \cap (\cap_{n\geq 1}]\alpha - 1/n, +\infty]).$$

Par suite, $\omega \in \mathcal{T}$. Si en outre ω est un ouvert de $\overline{\mathbb{R}}$ contenant $+\infty$ (resp. $-\infty$), il est réunion d'un ouvert de $\mathbb{R}$ et d'un ensemble $]a, +\infty]$ (resp.$[-\infty, a[$). Comme $[-\infty, a[^c = [a, +\infty]$ qui est élément de $\mathcal{T}$ d'après ce qui précède, on trouve que $\omega \in \mathcal{T}$, qed. Donc $g = \sup f_n$ est mesurable. En outre, les identités

$$\inf f_n = -\sup(-f_n), \qquad \limsup f_n = \inf_n(\sup_{k\geq n} f_k), \qquad \liminf f_n = \sup_n(\inf_{k\geq n} f_k)$$

assurent le résultat de la proposition. $\square$

[11] Cette convention souvent adoptée fait référence à une vision "potentielle" de l'infini. L'infini est quelque chose que l'on atteint après un processus d'approximation ; comme par exemple le produit $0n = 0$ pour tout n, il est naturel d'adopter cette convention. Cette vision de l'infini s'oppose à une conception "actuelle" de l'infini dans laquelle l'infini serait dès le départ présent. En théorie de la mesure, la convention mentionnée se justifie également par le fait que l'intégration d'une fonction nulle sur un ensemble quelconque (fut-il de mesure infinie) doit donner 0.

DEFINITION 1.3.2. Soit $(X, \mathcal{M})$ un espace mesurable. Une fonction s est dite étagée sur X si $s : X \longrightarrow [0, +\infty[$ est mesurable et ne prend qu'un nombre fini de valeurs.

Notons que si $\alpha_1, \dots, \alpha_m$ sont les valeurs distinctes prises par s, en posant

$$A_k = s^{-1}(\{\alpha_k\}) = \{x \in X, s(x) = \alpha_k\}, \qquad \text{noté} \qquad \{s = \alpha_k\},$$

on voit que les $(A_k)_{1 \le k \le m}$ forment une partition de X. Par conséquent

$$s(x) = \sum_{1 \le k \le m} \alpha_k \mathbf{1}_{A_k}(x).$$

THÉORÈME 1.3.3. *Soit $(X, \mathcal{M})$ un espace mesurable. Soit $f : X \longrightarrow \overline{\mathbb{R}}_+ = [0, +\infty]$ une application mesurable. Il existe une suite de fonctions étagées $(s_k)_{k \ge 1}$ telle que*

(a) $0 \le s_1 \le s_2 \le \dots \le s_k \le s_{k+1} \le \dots \le f$,
(b) $\forall x \in X,\ \lim_k s_k(x) = f(x)$,
(c) *la limite est uniforme si f est bornée, i.e.* $\sup_{x \in X} |f(x) - s_k(x)| \to 0$, *pour* $k \to +\infty$.

DÉMONSTRATION. Supposons tout d'abord que $0 \le f \le 1$. Posons[12]

$$s_k(x) = 2^{-k} E(2^k f(x)). \tag{1.3.1}$$

La fonction s_k ne prend qu'un nombre fini de valeurs car $0 \le 2^k f \le 2^k$. En outre, on a

$$2^k s_k \le 2^k f < 2^k s_k + 1 \Longrightarrow 0 \le f - s_k < 2^{-k} \tag{1.3.2}$$

et par conséquent s_k tend uniformément vers f. De plus, en multipliant (1.3.2) par 2 et en écrivant (1.3.1) pour $k+1$, on obtient

$$\mathbb{N} \ni 2^{k+1} s_k \le 2^{k+1} f, \qquad 2^{k+1} s_{k+1} = E(2^{k+1} f).$$

En utilisant la caractérisation de la partie entière, il vient

$$2^{k+1} s_k \le 2^{k+1} s_{k+1}, \text{ i.e. } s_k \le s_{k+1},$$

ce qui prouve que la suite s_k est monotone croissante. Par ailleurs, chaque fonction s_k est mesurable, comme composée de fonctions mesurables[13]. Si $0 \le f \le M$, pour un $M > 0$ réel, on applique le résultat précédent à f/M. Revenons au cas $0 \le f \le 1$. Posons

$$\tilde{s}_k = s_k - 2^{-k} E(f).$$

[12] $E(t)$ désigne la partie entière de t, i.e. l'unique entier tel que $E(t) \le t < E(t) + 1$.

[13] La fonction partie entière est mesurable car $E^{-1}([a, +\infty[) = [a, +\infty[$ si a est entier et sinon $E^{-1}([a, +\infty[) = [E(a) + 1, +\infty[$.

Si $f(x) < 1$, on a $s_k(x) = \tilde{s}_k(x)$. Si $f(x) = 1$, on a $s_k(x) - 2^{-k} = \tilde{s}_k(x)$. Dans les deux cas les suites $(\tilde{s}_k(x))_{k\in\mathbb{N}}$ sont croissantes de limite $f(x)$ et $0 \le \tilde{s}_k(x) < 1$. Or on peut identifier $\overline{\mathbb{R}}_+$ à $[0,1]$, par exemple à l'aide de l'homéomorphisme[14]

$$\varphi : \overline{\mathbb{R}}_+ \longrightarrow [0,1], \qquad \varphi(x) = x/(1+x), \qquad \varphi^{-1}(y) = y/(1-y)$$

On peut considérer

$$X \overset{f}{\longrightarrow} \overline{\mathbb{R}}_+ \overset{\varphi}{\longrightarrow} [0,1] \overset{\varphi^{-1}}{\longrightarrow} \overline{\mathbb{R}}_+.$$

En utilisant ce qui précède, on trouve une suite t_k étagée à valeurs dans [0,1[, croissante de limite $\varphi \circ f$. Par suite, $\varphi^{-1} \circ t_k$ est étagée (en particulier à valeurs finies car t_k est à valeurs < 1) et de limite f. La suite $\varphi^{-1} \circ t_k$ est croissante car t_k l'est et φ^{-1} est monotone croissante. Ceci achève la démonstration du théorème. $\square$

REMARQUE 1.3.4. Si f et g sont des fonctions mesurables de X dans $\overline{\mathbb{R}}_+$, alors $f+g$ est bien défini et mesurable. On peut utiliser le théorème 1.2.5' et la mesurabilité de l'application (continue)

$$\begin{array}{ccc} \overline{\mathbb{R}}_+ \times \overline{\mathbb{R}}_+ & \longrightarrow & \overline{\mathbb{R}}_+ \\ (\alpha,\beta) & \mapsto & \alpha+\beta \end{array}$$

De manière analogue, l'application symétrique (non continue) M

$$\begin{array}{ccc} \overline{\mathbb{R}}_+ \times \overline{\mathbb{R}}_+ & \longrightarrow & \overline{\mathbb{R}}_+ \\ (\alpha,\beta) & \mapsto & \alpha\cdot\beta \end{array}$$

prolongeant par continuité sur $\overline{\mathbb{R}}_+^* \times \overline{\mathbb{R}}_+^*$ la multiplication sur $\mathbb{R}_+^* \times \mathbb{R}_+^*$ et définissant $0.\infty = 0$ est Borel-mesurable. Pour $a \in \overline{\mathbb{R}}_+$, l'ensemble $E_a = \{(x,y) \in \overline{\mathbb{R}}_+ \times \overline{\mathbb{R}}_+, M(x,y) > a\}$ est inclus dans $\overline{\mathbb{R}}_+^* \times \overline{\mathbb{R}}_+^*$ sur lequel M est continue. Par conséquent E_a est un ouvert de $\overline{\mathbb{R}}_+^* \times \overline{\mathbb{R}}_+^*$ donc un borélien de $\overline{\mathbb{R}}_+ \times \overline{\mathbb{R}}_+$. Utilisant le théorème 1.2.5', ceci implique que $f \cdot g$ est mesurable.

1.4. Mesures positives

DEFINITION 1.4.1. Soit $(X, \mathcal{M})$ un espace mesurable. Une mesure positive sur X est une application $\mu : \mathcal{M} \to \overline{\mathbb{R}}_+$ telle que

(a) $\mu(\emptyset) = 0$,

(b) si $(A_k)_{k\in\mathbb{N}}$ est une suite d'éléments de $\mathcal{M}$ deux à deux disjoints[15]

$$\mu(\cup_{k\in\mathbb{N}} A_k) = \sum_{k\in\mathbb{N}} \mu(A_k).$$

[14] Application bijective bicontinue.

[15] pour $k \neq l$, $A_k \cap A_l = \emptyset$

La propriété (b) s'appelle la σ-additivité ou additivité dénombrable.

Donnons tout de suite quelques exemples simples.
(1.4.1) Si X est un ensemble fini, muni de la tribu $\mathcal{P}(X)$, on peut considérer la mesure μ_0 définie par

$$\mu_0(A) = \operatorname{Card} A.$$

(1.4.2) Si X est un ensemble fini non vide, muni de la tribu $\mathcal{P}(X)$, on peut considérer la *probabilité*[16] μ_1 définie par

$$\mu_1(A) = \operatorname{Card} A / \operatorname{Card} X.$$

(1.4.3) Si X est un ensemble quelconque, muni de la tribu $\mathcal{P}(X)$, on définit *la mesure de comptage* sur X par

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

Pour ce dernier cas, le lecteur pourra vérifier l'axiome (b).
(1.4.4) Si X est un ensemble non vide, muni de la tribu $\mathcal{P}(X)$, si $a \in X$ on peut considérer δ_a, la mesure de Dirac en a définie par

$$\delta_a(A) = \begin{cases} 1 & \text{si } a \in A, \\ 0 & \text{si } a \notin A. \end{cases}$$

On peut remarquer que formellement au moins, $\mu = \sum_{k\in\mathbb{N}} \delta_k$. Le lecteur pourra donner un sens à $\sum_{k\in\mathbb{N}} \nu_k$ où ν_k est une suite de mesures positives définies sur une tribu $\mathcal{M}$.
(1.4.5) Mesure de Borel sur $\mathbb{R}$. On considère $\mathcal{B}_{\mathbb{R}}$ la tribu de Borel sur $\mathbb{R}$. On cherche à construire une mesure positive définie sur $\mathcal{B}_{\mathbb{R}}$ telle que, pour $a \le b$ on ait

$$\mu([a,b]) = b - a = \mu(]a,b]).$$

Il est facile de construire μ sur les réunions finies d'intervalles disjoints. Bien que $\mathcal{B}_{\mathbb{R}}$ soit engendré par les intervalles, le problème de l'extension de μ à $\mathcal{B}_{\mathbb{R}}$ est difficile et constitue l'un des buts du cours.
(1.4.6) Mesure avec densité ν par rapport à la mesure de Borel sur $\mathbb{R}$. Etant donnée une fonction ν continue positive sur $\mathbb{R}$, on cherche à construire une mesure positive définie sur $\mathcal{B}_{\mathbb{R}}$ telle que, pour $a \le b$ on ait

$$\mu_\nu([a,b]) = \int_a^b \nu(t)dt,$$

[16] Une mesure positive vérifiant $\mu_1(X) = 1$.

où l'intégrale utilisée est celle de Riemann. Il est facile de construire μ_ν sur les réunions finies d'intervalles disjoints. C'est la version avec densité de l'exemple précédent pour lequel on avait $\nu \equiv 1$.
(1.4.7) Mesure de Borel sur $\mathbb{R}^d$. On considère $\mathcal{B}_{\mathbb{R}^d}$ la tribu de Borel sur $\mathbb{R}^d$. L'un des buts du cours est de construire une mesure positive, définie sur $\mathcal{B}_{\mathbb{R}^d}$, telle que, pour $a_j \le b_j$ réels, on ait

$$\mu(\prod_{1\le j\le d}[a_j, b_j]) = \prod_{1\le j\le d}(b_j - a_j).$$

C'est le version d-dimensionnelle de l'exemple (1.4.5).
(1.4.8) Probabilité de Cauchy sur $\mathbb{R}$ de paramètre $\alpha > 0$. C'est la mesure positive de densité

$$\frac{1}{\pi}\frac{\alpha}{\alpha^2+t^2}.$$

On remarque en effet que $\int_{\mathbb{R}} \frac{1}{\pi}\frac{\alpha}{\alpha^2+t^2}dt = \Big[\frac{1}{\pi}\arctan(t/\alpha)\Big]_{-\infty}^{+\infty} = 1$. On définit la fonction de répartition F de la probabilité μ sur $\mathbb{R}$ comme

$$F(t) = \mu(-\infty, t[.$$

On peut montrer que F est croissante (c'est donné par l'additivité), tend vers 0 en $-\infty$, tend vers 1 en $+\infty$ et est continue à gauche. Dans le cas de la probabilité de Cauchy, la fonction de répartition est

$$F(t) = \frac{1}{\pi}\arctan(\frac{t}{\alpha}) + \frac{1}{2}.$$

(1.4.9) Probabilité de Laplace-Gauss de moyenne m, d'écart-type σ, de densité

$$\frac{1}{\sigma\sqrt{2\pi}}\exp -\frac{(x-m)^2}{2\sigma^2}.$$

On remarque en effet que

$$\int_{\mathbb{R}} \exp -\frac{(x-m)^2}{2\sigma^2}dx = \sigma\sqrt{2\pi},$$

et

$$\int_{\mathbb{R}} x\exp -\frac{(x-m)^2}{2\sigma^2}\frac{dx}{\sigma\sqrt{2\pi}} = m, \qquad \int_{\mathbb{R}} (x-m)^2\exp -\frac{(x-m)^2}{2\sigma^2}\frac{dx}{\sigma\sqrt{2\pi}} = \sigma^2.$$

(1.4.10) Probabilité de Bernoulli de paramètre $p \in [0,1]$: $p\delta_0 + (1-p)\delta_1$.
(1.4.11) Probabilité binomiale de paramètres n entier ≥ 1 et $p \in [0,1]$,

$$\mu = \sum_{0\le k\le n} C_n^k p^k (1-p)^{n-k}\delta_k.$$

On peut par exemple considérer μ comme une mesure positive sur l'ensemble $\{0,1,\dots,n-1,n\}$ muni de la tribu de ses parties, et telle que $\mu(A)=\sum_{k\in A} C_n^k p^k(1-p)^{n-k}$.
(1.4.12) Probabilité de Poisson de paramètre $\lambda>0$ donnée par

$$e^{-\lambda}\sum_{k\in\mathbb{N}}\frac{\lambda^k}{k!}\delta_k.$$

On peut par exemple considérer $\mathbb{N}$ muni de la tribu de ses parties et poser

$$\mu(A)=e^{-\lambda}\sum_{k\in A}\frac{\lambda^k}{k!}.$$

(1.4.13) Un exemple plus abstrait est donné par la notion de mesure image. Si $(X,\mathcal{M},\mu)$ est un espace mesuré, où μ est une mesure positive et $f:X\longrightarrow Y$ est une application, l'ensemble

$$\mathcal{N}=\{B\subset Y, f^{-1}(B)\in\mathcal{M}\}$$

est une tribu sur Y (lemme 1.1.5), en sorte que f est mesurable. On définit alors la mesure image, $f_*(\mu)$ pour $B\in\mathcal{N}$

$$f_*(\mu)(B)=\mu(f^{-1}(B)). \tag{1.4.14}$$

Il est facile de vérifier que $f_*(\mu)$ satisfait les axiomes de la définition 1.4.1. On peut remarquer également que

$$(g\circ f)_*=g_*\circ f_*. \tag{1.4.15}$$

PROPOSITION 1.4.2. *Soit $(X,\mathcal{M},\mu)$ un espace mesuré où μ est une mesure positive. Alors pour $A,B,A_k\in\mathcal{M}$,*

(a) *$A\subset B$ implique $\mu(A)\le\mu(B)$.*
(b) *Si $(A_k)_{k\in\mathbb{N}}$ est une suite croissante de $\mathcal{M}$, $A=\cup_{k\in\mathbb{N}}A_k$, alors $\mu(A_k)\uparrow\mu(A)$ dans $\overline{\mathbb{R}}_+$.*
(c) *Si $(A_k)_{k\in\mathbb{N}}$ est une suite décroissante de $\mathcal{M}$, si $\mu(A_0)<\infty$, $A=\cap_{k\in\mathbb{N}}A_k$, alors $\mu(A_k)\downarrow\mu(A)$ dans $\mathbb{R}_+$.*

En outre les propriétés de la définition 1.4.1 *sont équivalentes à $\mu(\emptyset)=0$,*(b) *ci-dessus et $\mu(A\cup B)=\mu(A)+\mu(B)$, pour $A,B\in\mathcal{M}$ disjoints.*

DÉMONSTRATION. On a en effet la réunion disjointe $B=(B\backslash A)\cup A$, qui implique $\mu(B)=\mu(B\backslash A)+\mu(A)\ge\mu(A)$, ce qui donne (a). En posant $A_{-1}=\emptyset$, on trouve par récurrence[17] sur k

$$A_k=\cup_{0\le l\le k}(A_l\cap A_{l-1}^c).$$

[17]C'est vrai pour $k=0$ et $A_{k+1}=(A_{k+1}\cap A_k^c)\cup\overbrace{(A_{k+1}\cap A_k)}^{=A_k,\ \text{car } A_j\uparrow}=(A_{k+1}\cap A_k^c)\cup A_k$.

Par conséquent, on obtient

$$A = \cup_{k \geq 0} A_k = \cup_{k \geq 0}(A_k \cap A_{k-1}^c).$$

Pour $k \neq l$ (disons $k > l$), comme la suite A_j est croissante, on a

$$(A_k \cap A_{k-1}^c) \cap (A_l \cap A_{l-1}^c) = A_k \cap A_l \cap A_{k-1}^c \cap A_{l-1}^c = A_l \cap A_{k-1}^c \subset A_{k-1}^c \cap A_{k-1} = \emptyset.$$

Par suite, en utilisant le (b) de la définition 1.4.1, il vient

$$\mu(A) = \sum_{k \geq 0} \mu(A_k \cap A_{k-1}^c) = \lim_{n \to \infty} \sum_{0 \leq k \leq n} \mu(A_k \cap A_{k-1}^c) = \lim_{n \to \infty} \mu(A_n),$$

soit le résultat (b). Démontrons maintenant le point (c). On a

$$A_0 \backslash A = A_0 \cap (\cup_{k \geq 0} A_k^c) = \cup_{k \geq 0} \underbrace{(A_0 \cap A_k^c)}_{\text{croissant de } k}.$$

En appliquant la propriété (b) que l'on vient de prouver, on obtient $\mu(A_0 \cap A_k^c) \nearrow \mu(A_0 \backslash A)$. Pour chaque k, on a

$$+\infty > \mu(A_0) = \mu(A_k) + \mu(A_k^c \cap A_0),$$

ce qui implique que $\mu(A_0), \mu(A_k), \mu(A_k^c \cap A_0)$ sont des réels[18]; on a donc

$$\mu(A_k) = \mu(A_0) - \mu(A_k^c \cap A_0) \searrow \mu(A_0) - \mu(A_0 \backslash A) = \mu(A),$$

ce qui donne (c). Il reste à prouver l'équivalence des propriétés mentionnées avec la définition. Evidemment si μ est une mesure positive, ces propriétés sont satisfaites, comme on vient de le démontrer. Réciproquement, on doit démontrer le (b) de la définition 1.4.1. Soit $(A_k)_{k \in \mathbb{N}}$ une suite d'éléments disjoints de $\mathcal{M}$. D'après le (b) de la proposition 1.4.2, on a en utilisant l'additivité finie[19]

$$\sum_{0 \leq k \leq n} \mu(A_k) = \mu(\cup_{0 \leq k \leq n} A_k) \nearrow \mu(\cup_{k \geq 0} A_k), \text{ i.e. } \sum_{k \geq 0} \mu(A_k) = \mu(\cup_{k \geq 0} A_k). \quad \square$$

[18]C'est ici que l'on utilise l'hypothèse $\mu(A_0) < +\infty$. Celle-ci n'est pas superflue comme le montre l'exemple de la mesure de comptage sur $\mathbb{N}$ (exemple (1.4.3) de ce paragraphe) et de la suite décroissante $A_k = [k, +\infty[\cap \mathbb{N}$: on a pour chaque k, $\mu(A_k) = +\infty$ et $\mu(\cap_{k \geq 0} A_k) = \mu(\emptyset) = 0$.

[19]Celle-ci se démontre trivialement par récurrence à partir de l'additivité pour 2 éléments disjoints de $\mathcal{M}$: $\mu(\cup_{0 \leq k \leq n+1} A_k) = \mu(\cup_{0 \leq k \leq n} A_k) + \mu(A_{n+1}) = \sum_{0 \leq k \leq n} \mu(A_k) + \mu(A_{n+1})$.

REMARQUE 1.4.3. Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive, $(A_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathcal{M}$. Alors on a

$$\mu(\cup_{n\in\mathbb{N}} A_n) \leq \sum_{n\in\mathbb{N}} \mu(A_n). \tag{1.4.3}$$

En effet, en considérant la suite croissante $B_n = \cup_{0\leq k\leq n} A_k$, on peut appliquer la propriété (b) de la proposition 1.4.2, qui implique

$$\mu(\cup_{n\in\mathbb{N}} A_n) = \mu(\cup_{n\in\mathbb{N}} B_n) = \sup_{n\in\mathbb{N}} \mu(B_n) \leq \sup_{n\in\mathbb{N}} \sum_{0\leq k\leq n} \mu(A_k) = \sum_{n\in\mathbb{N}} \mu(A_n),$$

l'inégalité $\mu(B_n) \leq \sum_{0\leq k\leq n} \mu(A_k)$ se démontrant trivialement par récurrence sur n.

1.5. Intégration de fonctions positives

Le but du lemme suivant est de définir " l'intégrale par rapport à μ " des fonctions étagées définies en 1.3.2. On se convaincra aisément du caractère très naturel de la définition en gardant présent à l'esprit le fait que les ensembles de la tribu $\mathcal{M}$ peuvent être passablement compliqués (penser aux boréliens du type $F_\sigma, G_\delta, G_{\delta\sigma}, F_{\sigma\delta}, \dots$)[20]. C'est d'ailleurs un point fondamental de la théorie de Lebesgue qui pour intégrer une fonction f définie sur X, ne cherche pas à découper l'ensemble de définition de la fonction (par exemple en sous-intervalles si X est un intervalle de $\mathbb{R}$) mais découpe l'ensemble d'arrivée d'une manière qui dépend de f.

LEMME 1.5.1. *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Soit s une fonction étagée (cf.déf.1.3.2), i.e. une fonction mesurable $s : X \to [0, +\infty[$ ne prenant qu'un nombre fini de valeurs distinctes (réelles positives) $\alpha_1, \dots, \alpha_m$; on a ainsi $s = \sum_{1\leq j\leq m} \alpha_j \mathbf{1}_{A_j}$, $A_j = s^{-1}(\{\alpha_j\})$. On pose[21] alors*

$$I(s) = \sum_{\substack{1\leq j\leq m \\ \alpha_j>0}} \alpha_j \mu(A_j), \tag{1.5.1}$$

[20]Les F_σ sont les réunions dénombrables de fermés, les G_δ les intersections dénombrables d'ouverts, les $G_{\delta\sigma}$ sont les réunions dénombrables de G_δ, les $F_{\sigma\delta}$ sont les intersections dénombrables de F_σ. C'est Hausdorff qui a introduit cette terminologie et étudié le premier ce type de sous-ensembles. La lettre σ symbolise la réunion dénombrable, et δ l'intersection dénombrable. On obtient une suite de classes, $F_{\sigma\delta\sigma}, G_{\delta\sigma\delta}, \dots$

[21]Noter que cette définition ne fait intervenir que des produits de réels strictement positifs (les α_j) avec des éléments de $\overline{\mathbb{R}}_+$. En outre, la cohérence de la définition est assurée par le fait que la décomposition précédente de s est canonique en ce sens que les α_j et donc les A_j sont fonctions de s. La condition $I(0) = 0$ est en fait une conséquence de (1.5.1), puisque pour $s = 0$, on somme sur un ensemble vide d'indices. Par ailleurs, on aurait pu écrire simplement $I(s) = \sum_{1\leq j\leq m} \alpha_j \mu(A_j)$ en utilisant la convention $0.\infty=0$. Il nous a paru plus prudent d'éviter cette convention (discontinue), au prix d'un alourdissement des notations.

et $I(0) = 0$. On a, pour s, t étagées et $\lambda > 0$

$$I(s) = \sup_{\substack{\sigma \text{étagée} \\ 0 \leq \sigma \leq s}} I(\sigma), \quad I(s+t) = I(s) + I(t), \quad I(\lambda s) = \lambda I(s). \tag{1.5.2}$$

DÉMONSTRATION[22]. Tout d'abord si les fonctions σ, s sont étagées telles que $\sigma \leq s$ (i.e. pour tout $x \in X$, $\sigma(x) \leq s(x)$), on a l'écriture canonique

$$\sigma = \sum_{1 \leq k \leq n} \beta_k \mathbf{1}_{B_k}, \quad s = \sum_{1 \leq j \leq m} \alpha_j \mathbf{1}_{A_j},$$

où $(B_k)_{1 \leq k \leq n}$ et $(A_j)_{1 \leq j \leq m}$ sont des partitions des X. D'après la définition, on obtient

$$I(\sigma) = \sum_{\substack{1 \leq k \leq n \\ \beta_k > 0}} \beta_k \mu(B_k) = \sum_{\substack{1 \leq k \leq n, 1 \leq j \leq m \\ \beta_k > 0, B_k \cap A_j \neq \emptyset}} \beta_k \mu(B_k \cap A_j).$$

En remarquant que $B_k \cap A_j \neq \emptyset$ implique que $\beta_k \leq \alpha_j$ (car pour $x \in B_k \cap A_j, \beta_k = \sigma(x) \leq s(x) = \alpha_j$), et donc $\alpha_j > 0$ si $\beta_k > 0$, il vient

$$I(\sigma) \leq \sum_{\substack{1 \leq k \leq n, 1 \leq j \leq m \\ \alpha_j > 0}} \alpha_j \mu(B_k \cap A_j) = \sum_{\substack{1 \leq j \leq m \\ \alpha_j > 0}} \alpha_j \mu(A_j) = I(s),$$

ce qui prouve le premier résultat. Pour le second, on remarque tout d'abord que si s, t sont étagées, alors la fonction $s + t$ est mesurable comme somme de fonctions mesurables et par suite étagée car ne prenant qu'un nombre fini de valeurs ≥ 0. En prenant l'écriture canonique de s et t, on a

$$s = \sum_{1 \leq j \leq m} \alpha_j \mathbf{1}_{A_j}, \quad t = \sum_{1 \leq k \leq n} \beta_k \mathbf{1}_{B_k}, \quad \text{et donc} \quad s + t = \sum_{\substack{1 \leq j \leq m \\ 1 \leq k \leq n}} (\alpha_j + \beta_k) \mathbf{1}_{A_j \cap B_k}.$$

Les ensembles $A_j \cap B_k$ sont mesurables et deux à deux disjoints ($A_j \cap B_k \cap A_{j'} \cap B_{k'} = \emptyset$ si $j \neq j'$ ou $k \neq k'$), et comme

$$X = \cup_{1 \leq j \leq m} A_j = \cup_{\substack{1 \leq j \leq m \\ 1 \leq k \leq n}} (A_j \cap B_k),$$

la collection $(A_j \cap B_k)_{\substack{1 \leq j \leq m, 1 \leq k \leq n \\ A_j \cap B_k \neq \emptyset}}$ forme une partition de X. Comme $\mathbf{1}_\emptyset = 0$, on obtient

$$s + t = \sum_{\substack{1 \leq j \leq m, 1 \leq k \leq n \\ A_j \cap B_k \neq \emptyset}} (\alpha_j + \beta_k) \mathbf{1}_{A_j \cap B_k}. \tag{1.5.3}$$

[22]Cette démonstration est simple mais assez fastidieuse et peut sans doute être omise par le lecteur abordant ces lignes pour la première fois.

Si les $\alpha_j + \beta_k$ sont distincts, la formule (1.5.3) donne l'écriture canonique de $s+t$ et

$$I(s+t) = \sum_{\substack{1\le j\le m, 1\le k\le n \\ A_j\cap B_k\neq\emptyset, \alpha_j+\beta_k>0}} (\alpha_j+\beta_k)\mu(A_j\cap B_k). \tag{1.5.4}$$

Si les $\alpha_j+\beta_k$ ne sont pas distincts et prennent les valeurs distinctes strictement positives $\gamma_1, \dots, \gamma_p$, il faut réécrire (1.5.3) sous la forme

$$s+t = \sum_{1\le l\le p} \gamma_l \sum_{\substack{1\le j\le m, 1\le k\le n \\ A_j\cap B_k\neq\emptyset, \alpha_j+\beta_k=\gamma_l}} \mathbf{1}_{A_j\cap B_k}.$$

Il vient

$$\begin{aligned} I(s+t) &= \sum_{1\le l\le p} \gamma_l \mu\Big(\cup_{\substack{1\le j\le m, 1\le k\le n \\ A_j\cap B_k\neq\emptyset, \alpha_j+\beta_k=\gamma_l}} (A_j\cap B_k)\Big) \\ &= \sum_{1\le l\le p} \gamma_l \sum_{\substack{1\le j\le m, 1\le k\le n \\ A_j\cap B_k\neq\emptyset, \alpha_j+\beta_k=\gamma_l}} \mu(A_j\cap B_k) \\ &= \sum_{\substack{1\le j\le m, 1\le k\le n \\ A_j\cap B_k\neq\emptyset, \alpha_j+\beta_k>0}} (\alpha_j+\beta_k)\mu(A_j\cap B_k), \end{aligned}$$

soit la formule (1.5.4) finalement valide dans tous les cas. En outre, on a

$$\begin{aligned} I(s)+I(t) &= \sum_{\substack{1\le j\le m \\ \alpha_j>0}} \alpha_j\mu(A_j) + \sum_{\substack{1\le k\le n \\ \beta_k>0}} \beta_k\mu(B_k) \\ &= \sum_{\substack{1\le j\le m, 1\le k\le n \\ \alpha_j>0}} \alpha_j\mu(A_j\cap B_k) + \sum_{\substack{1\le j\le m, 1\le k\le n \\ \beta_k>0}} \beta_k\mu(A_j\cap B_k) \end{aligned}$$

On remarque que, avec $\mu_{jk} = \mu(A_j\cap B_k)$,

$$\begin{aligned} &\sum_{\alpha_j>0} \alpha_j\mu_{jk} + \sum_{\beta_k>0} \beta_k\mu_{jk} \\ &\quad= \sum_{\alpha_j>0,\beta_k>0} \alpha_j\mu_{jk} + \sum_{\alpha_j>0,\beta_k>0} \beta_k\mu_{jk} + \sum_{\alpha_j>0,\beta_k=0} \alpha_j\mu_{jk} + \sum_{\alpha_j=0,\beta_k>0} \beta_k\mu_{jk} \\ &\quad= \sum_{\alpha_j>0,\beta_k>0} (\alpha_j+\beta_k)\mu_{jk} + \sum_{\alpha_j>0,\beta_k=0} (\alpha_j+\beta_k)\mu_{jk} + \sum_{\alpha_j=0,\beta_k>0} (\alpha_j+\beta_k)\mu_{jk} \\ &\quad= \sum_{\alpha_j+\beta_k>0} (\alpha_j+\beta_k)\mu_{jk}. \end{aligned}$$

Ceci donne

$$I(s) + I(t) = \sum_{\substack{1 \le j \le m, 1 \le k \le n \\ \alpha_j + \beta_k > 0,\ A_j \cap B_k \neq \emptyset}} (\alpha_j + \beta_k)\mu(A_j \cap B_k) = I(s+t).$$

Finalement pour $\lambda > 0$ et s étagée, on a

$$I(\lambda s) = I(\lambda \sum_{1 \le j \le m} \alpha_j \mathbf{1}_{A_j}) = \sum_{\substack{1 \le j \le m \\ \alpha_j > 0}} \lambda \alpha_j \mu(A_j) = \lambda I(s),$$

ce qui achève la démonstration du lemme. $\square$

Ce lemme permet de définir l'intégrale des fonctions mesurables de $X \to [0, +\infty] = \overline{\mathbb{R}}_+$.

DÉFINITION 1.5.2. Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Soit $f : X \to [0, +\infty]$ une fonction mesurable. On pose[23]

$$\int_X f d\mu = \sup_{\substack{s \text{ étagée} \\ 0 \le s \le f}} I(s)$$

Remarquons que d'après le lemme 1.5.1, pour f étagée, on a $\int_X f d\mu = I(f)$.

REMARQUE. On peut revenir sur les exemples du paragraphe 4 et examiner comment se fait le passage de la mesure des ensembles à l'intégration des fonctions.
(1.5.5) Soit $X = \{x_1, \dots, x_n\}$ muni de la tribu $\mathcal{P}(X)$ avec $\mu_0(A) = \operatorname{Card} A$. On a

$$\int_X f d\mu_0 = \int_X \sum_{1 \le j \le n} f(x_j) \mathbf{1}_{\{x_j\}} d\mu_0 = f(x_1) + \dots + f(x_n).$$

(1.5.6) Soit $X = \{x_1, \dots, x_n\}$ muni de la tribu $\mathcal{P}(X)$ avec $\mu_1(A) = \operatorname{Card} A / \operatorname{Card} X$. On a

$$\int_X f d\mu_1 = \frac{f(x_1) + \dots + f(x_n)}{n}.$$

(1.5.7) Si $X = \{x_1, \dots, x_n\}$ muni de la tribu $\mathcal{P}(X)$, μ la mesure de densité ν par rapport à μ_0, on a

$$\int_X f d\mu = \sum_{1 \le j \le n} f(x_j) \nu_j.$$

En particulier si les réels positifs ν_j satisfont $\sum \nu_j = 1$, la mesure μ est une probabilité sur X.

[23]La notation $\int_X f(x) d\mu(x)$ est également utilisée ainsi que $\int_X f(x) \mu(dx)$.

(1.5.8) Soit $(X = (x_i)_{i\in I}, \mathcal{P}(X))$ muni de la mesure de comptage. On a

$$\int_X f d\mu = \sum_{i\in I} f(x_i) = \sup_{J \text{ fini} \subset I} \sum_{i\in J} f(x_i).$$

(1.5.9) Soit $(X, \mathcal{P}(X))$ un ensemble non vide et $a \in X$. Si μ est la masse de Dirac en a

$$\int_X f d\mu = f(a)$$

(1.5.10) Pour la mesure de Borel m sur $\mathbb{R}^d$, dont la construction reste à faire, on notera l'intégrale par le même symbole que l'intégrale de Riemann,

$$\int_{\mathbb{R}^d} f(x) dx$$

et on verra que cette intégrale coïncide avec celle de Riemann, pour $f \in C_c^0(\mathbb{R}^d)$. Notons également qu'on aura alors

$$\int_{\mathbb{R}} \mathbf{1}_{\mathbb{Q}}(x) dx = m(\mathbb{Q}) = 0$$

(1.5.11) Si μ est une mesure de densité ν par rapport à la mesure de Borel,

$$\int_{\mathbb{R}^d} f d\mu = \int_{\mathbb{R}^d} f(x)\nu(x) dx$$

de sorte que $d\mu(x) = \nu(x)dx$ et on peut considérer symboliquement que $\mu'(x) = \nu(x)$ ce qui est à l'origine de la notation $\int f(x)d\mu(x) = \int f(x)\nu(x)dx$.

(1.5.12) Si $(X, \mathcal{M}, \mu)$ est un espace mesuré où μ est une mesure positive et si $\Phi : X \to Y$ est une application, on a vu (exemple (13), §1.4) que l'on pouvait construire un espace mesuré $(Y, \mathcal{N}, \nu)$ où $\nu = \Phi_*(\mu)$ est la mesure image de μ. Si $g : Y \to \overline{\mathbb{R}}_+$ est mesurable, alors $g \circ \Phi$ l'est aussi et

$$\int_Y g d\nu = \int_X (g \circ \Phi) d\mu$$

car pour $g = \beta \mathbf{1}_B$

$$\int_Y g d\nu = \beta\nu(B) = \beta\mu(\Phi^{-1}(B)) = \int_X \beta \mathbf{1}_{(\Phi^{-1}(B))} d\mu = \int_X \beta(\mathbf{1}_B \circ \Phi) d\mu$$
$$= \int_X (g \circ \Phi) d\mu,$$

et le résultat par linéarité.

PROPOSITION 1.5.3. *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Soient $f, g : X \to \overline{\mathbb{R}}_+$ des fonctions mesurables, $A, B \in \mathcal{M}$ et $c > 0$ un réel. On pose*

$$\int_A f d\mu = \int_X \underbrace{f \cdot \mathbf{1}_A}_{f_A} d\mu, \quad \text{avec} \quad f_A(x) = \begin{cases} f(x) & \text{si } x \in A, \\ 0 & \text{si } x \notin A. \end{cases} \tag{1.5.13}$$

Les propriétés suivantes sont vérifiées.

(a) $0 \le f \le g \Longrightarrow \int_X f d\mu \le \int_X g d\mu$ *et* $A \subset B \Longrightarrow \int_A f d\mu \le \int_B f d\mu$.
(b) $\int_X c f d\mu = c \int_X f d\mu$
(c) $\mu(A) = 0 \Longrightarrow \int_A f d\mu = 0$, *même si* $f \equiv +\infty$.
(d) *Si s est une fonction étagée, $E \in \mathcal{M}$, on pose $\lambda_s(E) = \int_E s d\mu$. Alors λ_s est une mesure positive définie sur $\mathcal{M}$.*

DÉMONSTRATION. (a) est une conséquence de la définition 1.5.2 et la seconde partie vient de $f_A \le f_B$. (b) vient de la définition 1.5.2 et de (1.5.2):

$$\int_X c f d\mu = \sup_{s \text{ étagée} \le c f} I(s) = \sup_{s \text{ étagée} \le c f} I\left(\frac{cs}{c}\right) = c \sup_{\frac{s}{c} \text{ étagée} \le f} I\left(\frac{s}{c}\right) = c \int_X f d\mu.$$

Pour (c), on considère s étagée $\le f_A$. On a l'inclusion $A^c \subset \{s = 0\}$, de sorte que si $s = \sum_{\substack{1 \le j \le m \\ \alpha_j > 0}} \alpha_j \mathbf{1}_{A_j}$ est l'écriture canonique de s et $\alpha_i \neq 0$, $A_i \subset A$ et donc $\mu(A_i) = 0$, d'où $I(s) = 0$ et $\int_A f d\mu = 0$. Pour (d) on remarque d'abord que $\lambda(\emptyset) = \int_\emptyset s d\mu = 0$ d'après (c) et $\mu(\emptyset) = 0$. Soient $(E_j)_{j \ge 0}$ une suite d'éléments de $\mathcal{M}$ deux à deux disjoints et $s = \sum_{1 \le k \le m} \alpha_k \mathbf{1}_{A_k}$ une fonction étagée. On a avec $E = \cup_{j \ge 0} E_j$, d'après le lemme 1.5.1, la définition 1.5.2 et le lemme 1.2.11

$$\begin{aligned} \lambda(E) = \int_X s_E d\mu = \sum_{1 \le k \le m} \alpha_k \mu(A_k \cap E) &= \sum_{1 \le k \le m} \alpha_k \Big(\sum_{j \ge 0} \mu(A_k \cap E_j) \Big) \\ &= \sum_{j \ge 0} \sum_{1 \le k \le m} \alpha_k \mu(A_k \cap E_j) = \sum_{j \ge 0} \lambda(E_j). \quad \square \end{aligned}$$

1.6. Théorèmes de convergence

Dans le paragraphe précédent, nous avons pu définir $\int_X f d\mu$, l'intégrale par rapport à μ d'une fonction mesurable à valeurs dans $\overline{\mathbb{R}}_+$. Nous verrons ici que pour f mesurable de X dans $\mathbb{C}$, telle que $\int_X |f| d\mu < \infty$, il est facile de définir $\int_X f d\mu$. Nous allons pouvoir aborder la partie la plus intéressante de la théorie de l'intégration (élaborée par Henri Lebesgue en 1902 dans sa thèse à Nancy) et en particulier quelques théorèmes de convergence.

Typiquement, nous souhaitons prouver que sous une hypothèse assez faible de convergence d'une suite de fonctions f_n vers f on obtient la convergence de $\int_X f_n d\mu$ vers $\int_X f d\mu$ (en tout état de cause, la convergence exigée des f_n est beaucoup plus faible que la convergence uniforme). L'un des grands mérites de la théorie de Lebesgue est de dégager un espace fonctionnel, l'espace des fonctions intégrables, que l'on peut munir d'une norme qui en fait un espace de Banach. Le premier théorème de convergence que nous démontrons est dû à Beppo Levi.

Théorème 1.6.1. Théorème de convergence monotone, dit de Beppo-Levi. *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Soit $(f_n)_{n\geq 0}$ une suite de fonctions mesurables de $X \to \overline{\mathbb{R}}_+$. On suppose que,*

$$\forall x \in X,\ f_n(x) \nearrow f(x), \quad \textit{i.e. } f_n \textit{ tend simplement en croissant vers } f.$$

Alors f est mesurable et

$$\lim_{n\to\infty} \int_X f_n d\mu = \sup_{n\geq 0} \int_X f_n d\mu = \int_X f d\mu.$$

Remarquons que l'hypothèse de convergence est réduite à la convergence simple. Bien entendu sans l'hypothèse additionnelle de croissance, le résultat est faux.[24]

démonstration. D'après la proposition 1.3.1, $\sup f_n$ est mesurable et de la proposition 1.5.3 (a) il vient que la suite $(\int_X f_n d\mu)_{n\in\mathbb{N}}$ est croissante et majorée par $\int_X f d\mu$. Par conséquent

$$\lim_{n\to\infty} \int_X f_n d\mu = \sup_{n\in\mathbb{N}} \int_X f_n d\mu \leq \int_X f d\mu. \tag{1.6.1}$$

Il reste donc à prouver l'inégalité inverse. Soient $1 \geq \epsilon > 0$ et s une fonction étagée telle que $0 \leq s \leq f$. Considérons

$$E_n = \{x \in X, (1-\epsilon)s(x) \leq f_n(x)\}.$$

L'ensemble E_n est mesurable car s et f_n sont mesurables et donc $f - (1-\epsilon)s$ (qui a un sens car s prend des valeurs finies) l'est aussi : on a $E_n = (f_n - (1-\epsilon)s)^{-1}(\overline{\mathbb{R}}_+)$. De

[24]Prendre par exemple sur $[0,1]$, $f_n(x) = \begin{cases} xn^3, & \text{pour } 0 \leq x \leq 1/n, \\ 2n^2 - xn^3, & \text{pour } 1/n \leq x \leq 2/n \\ 0 & \text{ailleurs.} \end{cases}$. La suite de fonctions continues f_n tend simplement vers 0, néanmoins $\int_0^1 f_n(x)dx \to \infty$.

plus, la croissance des f_n implique celle des E_n. On a en outre $X = \cup_{n\in\mathbb{N}} E_n$ car, si x_0 appartenait à E_n^c pour tout $n \in \mathbb{N}$, on aurait

$$+\infty > (1-\epsilon)s(x_0) > f_n(x_0)(\geq 0)$$

et donc $s(x_0) \in]0, +\infty[$ et

$$f(x_0) = \sup_n f_n(x_0) \leq (1-\epsilon)s(x_0) < s(x_0) \leq f(x_0)$$

ce qui est impossible. Par suite, on a d'après la proposition 1.5.3(d)(λ_s est une mesure), la proposition 1.4.2(b)(convergence croissante pour les mesures d'ensembles) et la proposition 1.5.3(b)(homogénéité)

$$\int_{E_n} (1-\epsilon)s d\mu = \lambda_{(1-\epsilon)s}(E_n) \nearrow \lambda_{(1-\epsilon)s}(X) = \int_X (1-\epsilon)s d\mu = (1-\epsilon)I(s). \tag{1.6.2}$$

Or comme $(1-\epsilon)s \cdot \mathbf{1}_{E_n} \leq f_n \cdot \mathbf{1}_{E_n} \leq f_n$, la proposition 1.5.3(a) implique les inégalités

$$\int_{E_n} (1-\epsilon)s d\mu \leq \int_{E_n} f_n d\mu \leq \int_X f_n d\mu. \tag{1.6.3}$$

On déduit de $(1.6.1-3)$

$$(1-\epsilon)I(s) = \lim_n \int_{E_n} (1-\epsilon)s d\mu \leq \sup_n \int_X f_n d\mu,$$

et par conséquent

$$(1-\epsilon)\int_X f d\mu = (1-\epsilon) \sup_{s \text{ étagée} \leq f} I(s) \leq \lim_n \int_X f_n d\mu \leq \int_X f d\mu,$$

ceci pour tout $\epsilon > 0$. En prenant le supremum sur $\epsilon > 0$, on obtient le résultat[25]. □

Corollaire 1.6.2. *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Soit $(f_n)_{n\geq 0}$ une suite de fonctions mesurables de $X \longrightarrow \overline{\mathbb{R}}_+$. On pose $S(x) = \sum_{n\geq 0} f_n(x)$. Alors S est mesurable positive et*

$$\int_X S d\mu = \sum_{n\geq 0} \int_X f_n d\mu.$$

[25] Ceci est vrai même si $\int_X f d\mu = +\infty$ car tous les termes de la dernière inégalité sont alors $+\infty$.

DÉMONSTRATION. La mesurabilité de S est une conséquence de la proposition 1.3.1 puisque d'une part

$$S_n(x) = \sum_{0 \le k \le n} f_k(x) \nearrow S(x)$$

et la mesurabilité d'une somme finie de fonctions à valeurs dans $\overline{\mathbb{R}}_+$ est assurée par la mesurabilité (due à la continuité) de

$$\begin{array}{ccc} \overline{\mathbb{R}}_+ \times \overline{\mathbb{R}}_+ & \longrightarrow & \overline{\mathbb{R}}_+ \\ (\alpha, \beta) & \mapsto & \alpha + \beta \end{array}.$$

On applique alors le théorème de Beppo Levi ci-dessus et on obtient

$$\int_X S d\mu = \sup_{n \ge 0} \int_X S_n d\mu. \tag{1.6.4}$$

Or on a

$$\int_X S_n d\mu = \int_X \sum_{0 \le k \le n} f_k d\mu = \sum_{0 \le k \le n} \int_X f_k d\mu, \tag{1.6.5}$$

la seconde égalité étant due au lemme 1.6.3 ci-dessous. En admettant provisoirement ce lemme, on note que $(1.6.4 - 5)$ impliquent le résultat du corollaire.

LEMME 1.6.3. *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Soient $f_1, \dots, f_N$ des fonctions mesurables de $X \longrightarrow \overline{\mathbb{R}}_+$. Alors $f_1 + \cdots + f_N$ est mesurable et $\int_X (f_1 + \cdots + f_N) d\mu = \int_X f_1 d\mu + \cdots + \int_X f_N d\mu$*

DÉMONSTRATION. Par récurrence sur N, il suffit de prouver cela pour $N = 2$. Soient donc f_1, f_2 satisfaisant les hypothèses du lemme et, utilisant le théorème 1.3.3, soient $s_k^{(1)}, s_k^{(2)}$ des fonctions étagées $0 \le s_k^{(j)} \nearrow f_j$, $j = 1, 2$. Du théorème de Beppo Levi, il vient

$$\int_X s_k^{(j)} d\mu \nearrow \int_X f_j d\mu. \tag{1.6.6}$$

Par suite, du lemme 1.5.1 et du théorème de Beppo Levi, il vient

$$\int_X s_k^{(1)} d\mu + \int_X s_k^{(2)} d\mu = \int_X (s_k^{(1)} + s_k^{(2)}) d\mu \nearrow \int_X (f_1 + f_2) d\mu,$$

ce qui avec $(***)$ donne le résultat du lemme. $\square$

LEMME 1.6.4. LEMME DE FATOU. *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Soit $(f_n)_{n\ge 0}$ une suite de fonctions mesurables de $X \to \overline{\mathbb{R}}_+$. On a alors l'inégalité*

$$\int_X (\liminf_n f_n) d\mu \le \liminf_n \left(\int_X f_n d\mu \right).$$

DÉMONSTRATION. Notons tout d'abord que l'énoncé a bien un sens, puisque la proposition 1.3.1 assure la mesurabilité de $\liminf f_n$ (par ailleurs à valeurs dans $\overline{\mathbb{R}}_+$). Rappelons que l'on a $\liminf f_n = \sup_{n\in\mathbb{N}} (\inf_{k\ge n} f_k)$, de sorte qu'en posant $g_n = \inf_{k\ge n} f_k$, on trouve que g_n est mesurable et telle que $0 \le g_n \nearrow \liminf f_n$. On peut alors appliquer le théorème de Beppo Levi qui donne

$$\int_X g_n d\mu \nearrow \int_X (\liminf f_n) d\mu. \tag{1.6.7}$$

De la proposition 1.5.3(a), il vient $\int_X g_n d\mu = \int_X (\inf_{k\ge n} f_k) d\mu \le \int_X f_n d\mu$, ce qui implique[26] $\liminf \int_X g_n d\mu \le \liminf \int_X f_n d\mu$ et donc d'après (1.6.7) le résultat du lemme. □

PROPOSITION 1.6.5. *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Soit $\nu : X \longrightarrow \overline{\mathbb{R}}_+$ une application mesurable. Pour $E \in \mathcal{M}$, on définit $\lambda_\nu(E) = \int_E \nu d\mu$. Alors λ_ν est une mesure positive définie sur $\mathcal{M}$ et si $f : X \longrightarrow \overline{\mathbb{R}}_+$ est mesurable*

$$\int_X f d\lambda = \int_X f \cdot \nu \, d\mu$$

où $f \cdot \nu$ est la fonction mesurable[27] définie par la convention $0.\infty = 0$. On notera naturellement $d\lambda = \nu d\mu$ et on dira que $d\lambda$ est la mesure de densité ν par rapport à $d\mu$.

DÉMONSTRATION. On a trivialement $\lambda_\nu(\emptyset) = \int_\emptyset \nu d\mu = 0$ d'après la propriété 1.5.3(c). En outre si $(A_j)_{j\in\mathbb{N}}$ est une suite d'éléments deux à deux disjoints de $\mathcal{M}$, on a d'après le corollaire 1.6.2

$$\lambda_\nu(\cup_{j\ge 0} A_j) = \int_{\cup_{j\ge 0} A_j} \nu d\mu = \int_X \sum_{j\ge 0} \nu \cdot \mathbf{1}_{A_j} d\mu = \sum_{j\ge 0} \int_X \nu \cdot \mathbf{1}_{A_j} d\mu = \sum_{j\ge 0} \lambda_\nu(A_j),$$

ce qui donne la première partie de la proposition. Si la fonction f est étagée, on a $f = \sum_{1\le j\le m} \alpha_j \mathbf{1}_{A_j}$ et on peut supposer que les α_j sont des réels > 0. Dans ces conditions, on a

$$\int_X f d\lambda_\nu = \sum_{1\le j\le m} \alpha_j \lambda_\nu(A_j) = \sum_{1\le j\le m} \alpha_j \int_X \nu \cdot \mathbf{1}_{A_j} d\mu,$$

[26] On utilise ici que pour des suites $(x_n), (y_n)$ de $\overline{\mathbb{R}}$, les inégalités $\forall n, x_n \le y_n$ impliquent $\liminf x_n \le \liminf y_n$. C'est une trivialité car, pour $l \ge n$, $\inf_{k\ge n} x_k \le x_l \le y_l$ donc $\inf_{k\ge n} x_k \le \inf_{k\ge n} y_k$ et $\lim_n (\inf_{k\ge n} x_k) \le \lim_n (\inf_{k\ge n} y_k)$.

[27] Voir la remarque 1.3.4.

et en utilisant le lemme 1.6.3, il vient

$$\int_X f d\lambda_\nu = \int_X \sum_{1\le j\le m} \alpha_j \mathbf{1}_{A_j} \cdot \nu d\mu = \int_X f \cdot \nu d\mu,$$

soit le résultat pour f étagée. Dans le cas général, on utilise le théorème d'approximation 1.3.3 et le théorème de Beppo Levi (théorème 1.6.1) qui donnent avec des fonctions étagées (s_k) tendant simplement en croissant vers f

$$\int_X f d\lambda_\nu \overset{\text{B.Levi}}{=} \sup_k \int_X s_k d\lambda_\nu \overset{s_k \text{ étagée}}{=} \sup_k \int_X s_k \cdot \nu d\mu \overset{\text{B.Levi}}{=} \int_X f \cdot \nu d\mu \quad \square$$

Noter que l'on a utilisé subrepticement (dernière égalité) que $\sup_k(s_k \cdot \nu) = (\sup_k s_k) \cdot \nu$, ce qui est évident sauf si $\nu(x) = +\infty, \sup_k s_k(x) = 0$ ou bien $\nu(x) = 0, \sup_k s_k(x) = +\infty$. Dans ce dernier cas, on obtient bien 0 ainsi que dans le premier car tous les $s_k(x)$ sont nécessairement nuls.

DEFINITION 1.6.6. Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Soit $f : X \to \mathbb{C}$ une application mesurable. On dira que f appartient à $\mathcal{L}^1(\mu)$ si

$$\int_X |f| d\mu < +\infty.$$

On pose alors

$$\int_X f d\mu = \int_X (\operatorname{Re} f)_+ d\mu - \int_X (\operatorname{Re} f)_- d\mu + i \int_X (\operatorname{Im} f)_+ d\mu - i \int_X (\operatorname{Im} f)_- d\mu,$$

ce qui a un sens car les intégrales $\int_X (\operatorname{Re} f)_\pm d\mu$, $\int_X (\operatorname{Im} f)_\pm d\mu$ sont majorées (proposition 1.5.3(a)) par $\int_X |f| d\mu$ qui est finie.

PROPOSITION 1.6.7. *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Alors $\mathcal{L}^1(\mu)$ est un espace vectoriel sur $\mathbb{C}$ et $f \mapsto \int_X f d\mu$ est une forme linéaire sur cet espace.*

DÉMONSTRATION. Soient $f, g \in \mathcal{L}^1(\mu)$ et $\alpha, \beta \in \mathbb{C}$. Alors $\alpha f + \beta g$ est une fonction mesurable (th.1.2.5) et comme $|\alpha f + \beta g| \le |\alpha||f| + |\beta||g|$, la proposition 1.5.3(a,b) et le lemme 1.6.3 assurent $\alpha f + \beta g \in \mathcal{L}^1(\mu)$. Si $f = f_1 + i f_2, g = g_1 + i g_2$ est la décomposition en parties réelle et imaginaire, on a

$$\operatorname{Re} \int_X (f+g) d\mu = \int_X (f_1 + g_1)_+ d\mu - \int_X (f_1 + g_1)_- d\mu. \tag{1.6.8}$$

Or on a

$$\operatorname{Re}(f+g) = (f_1+g_1)_+ - (f_1+g_1)_- = f_1+g_1 = (f_1)_+ - (f_1)_- + (g_1)_+ - (g_1)_-,$$

et par conséquent

$$(f_1+g_1)_+ + (f_1)_- + (g_1)_- = (f_1)_+ + (g_1)_+ + (f_1+g_1)_-,$$

ce qui montre en appliquant le lemme1.6.3 que

$$\int_X (f_1+g_1)_+ d\mu + \int_X (f_1)_- d\mu + \int_X (g_1)_- d\mu = \int_X (f_1)_+ d\mu + \int_X (g_1)_+ d\mu + \int_X (f_1+g_1)_- d\mu,$$

et par suite de (1.6.8) il vient

$$\operatorname{Re}\int_X (f+g)d\mu = \int_X (f_1)_+ d\mu + \int_X (g_1)_+ d\mu - \int_X (f_1)_- d\mu + \int_X (g_1)_- d\mu = \int_X \operatorname{Re} f d\mu + \int_X \operatorname{Re} g d\mu.$$

Comme un raisonnement analogue fournit

$$\operatorname{Im}\int_X (f+g)d\mu = \int_X (f_2)_+ d\mu + \int_X (g_2)_+ d\mu - \int_X (f_2)_- d\mu + \int_X (g_2)_- d\mu = \int_X \operatorname{Im} f d\mu + \int_X \operatorname{Im} g d\mu,$$

il vient

$$\int_X (f+g)d\mu = \int_X \operatorname{Re} f d\mu + \int_X \operatorname{Re} g d\mu + i\int_X \operatorname{Im} f d\mu + i\int_X \operatorname{Im} g d\mu. \tag{1.6.9}$$

Or d'après la définition 1.6.6, on a

$$\int_X f d\mu = \int_X \operatorname{Re} f d\mu + i\int_X \operatorname{Im} f d\mu,$$

et donc (1.6.9) implique

$$\int_X (f+g)d\mu = \int_X f d\mu + \int_X g d\mu.$$

Par ailleurs, si $\alpha = \alpha_1 + i\alpha_2$ est un nombre complexe, d'après ce qui précède, on a

$$\int_X \alpha f d\mu = \int_X \alpha_1 f_1 d\mu - \int_X \alpha_2 f_2 d\mu + \int_X i\alpha_1 f_2 d\mu + \int_X i\alpha_2 f_1 d\mu.$$

Or pour α_1, f_1 réels, la définition 1.6.6 et la proposition 1.5.3 (b) donnent (en discutant sur le signe de α_1) $\int_X \alpha_1 f_1 d\mu = \alpha_1 \int f_1 d\mu$. Il nous reste donc à démontrer $\int_X i f_1 d\mu = i\int f_1 d\mu$, ce qui est une conséquence immédiate de la définition 1.6.6. Ceci achève la démonstration de la proposition 1.6.7.

THÉORÈME 1.6.8. THÉORÈME DE CONVERGENCE DOMINÉE DE LEBESGUE[28]. *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables de X dans $\mathbb{C}$ telle que les propriétés suivantes soient vérifiées.*

(1) Convergence simple. *Pour tout $x \in X$, $\lim_{n\to\infty} f_n(x) = f(x)$.*

(2) Domination. *Il existe $g : X \to \overline{\mathbb{R}}_+$ mesurable, vérifiant $\int_X g d\mu < +\infty$, telle que*

$$\forall n \in \mathbb{N}, \forall x \in X, \qquad |f_n(x)| \leq g(x).$$

Alors la fonction f est mesurable et l'on a $\int_X |f| d\mu < +\infty$ et

$$\lim_{n\to\infty} \int_X |f - f_n| d\mu = 0, \quad \textit{ce qui implique} \quad \lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu.$$

DÉMONSTRATION. La mesurabilité de f est une conséquence de la proposition 1.3.1. En outre, la proposition 1.5.3 (a) implique

$$\int_X |f_n| d\mu \leq \int_X g d\mu < +\infty.$$

Le lemme de Fatou (lemme 1.6.4) fournit alors

$$\int_X |f| d\mu = \int_X \liminf_n |f_n| d\mu \leq \liminf_n \int_X |f_n| d\mu \leq \int_X g d\mu < +\infty.$$

Par ailleurs, l'inégalité $|f_n - f| \leq 2g$ et le lemme de Fatou impliquent

$$\int_X 2g d\mu = \int_X \liminf_n \big(2g - |f_n - f|\big) d\mu \leq \liminf_n \int_X \big(2g - |f_n - f|\big) d\mu \leq \int_X 2g d\mu < +\infty.$$

Du lemme 1.6.3, il vient alors

$$\int_X \big(2g - |f_n - f|\big) d\mu + \int_X |f_n - f| d\mu = \int_X 2g d\mu \leq \liminf_n \int_X \big(2g - |f_n - f|\big) d\mu,$$

et par suite, on obtient

$$\limsup_n \int_X |f_n - f| d\mu \leq \liminf_n \int_X \big(2g - |f_n - f|\big) d\mu + \limsup_n - \left[\int_X \big(2g - |f_n - f|\big) d\mu\right] = 0,$$

car la suite numérique $\int_X \big(2g - |f_n - f|\big) d\mu$ est bornée. $\square$

[28]On verra plus bas une version un peu plus générale tenant compte des ensembles de mesure nulle

1.7. Espace $L^1(\mu)$ et ensembles de mesure nulle

La proposition suivante introduit la notion de propriété vraie presque partout dans une espace mesuré $(X, \mathcal{M}, \mu)$. On écrira en abrégé μ-pp pour μ presque partout.

PROPOSITION 1.7.1. *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Soient $f, g : X \to \overline{\mathbb{R}}_+$ des applications mesurables.*
(a) *Alors $\int_X f d\mu = 0$ équivaut à $f = 0, \mu$-pp, i.e. $\mu\big(\{x \in X, f(x) \neq 0\}\big) = 0$.*
(b) *Si $f \leq g$, μ-pp i.e. $\mu\big(\{x \in X, f(x) > g(x)\}\big) = 0$, alors, on a l'inégalité*

$$\int_X f d\mu \leq \int_X g d\mu.$$

(c) *Si $f = g$, μ-pp i.e. $\mu\big(\{x \in X, f(x) \neq g(x)\}\big) = 0$, alors $\int_X f d\mu = \int_X g d\mu$.*
(d) *Si $\int_X f d\mu < +\infty$, alors $f < +\infty$, μ-pp, i.e. $\mu\big(\{x \in X, f(x) = +\infty\}\big) = 0$.*

DÉMONSTRATION. Pour (a), si $\int_X f d\mu = 0$, on pose, pour $k \geq 1$ entier, $F_k = \{f \geq 1/k\}$. La suite F_k est croissante mesurable et $\cup_{k \geq 1} F_k = \{f > 0\}$. De la proposition 1.4.2, il vient $\mu(F_k) \nearrow \mu(\{f > 0\})$ lorsque $k \to +\infty$. Or on a

$$\mu(F_k) = \int_X \mathbf{1}(f \geq 1/k) d\mu \overbrace{\leq}^{\text{Prop.1.5.3(a)}} \int_X k \cdot f d\mu \overbrace{=}^{\text{Prop.1.5.3(b)}} k \int_X f d\mu = 0 \Longrightarrow \mu(\{f > 0\}) = 0.$$

Réciproquement, si $\mu(E) = 0$ avec $E = \{f > 0\}$, comme $f = f \cdot \mathbf{1}_E$, il vient

$$\int_X f d\mu = \int_X f \cdot \mathbf{1}_E d\mu = \int_E f d\mu = 0, \quad \text{d'après la Proposition 1.5.3 (c).}$$

En particulier, pour $f \in \mathcal{L}^1(\mu)$, on a

$$\int_X |f| d\mu = 0 \Longrightarrow f = 0, \mu - pp. \tag{1.7.1}$$

Démontrons (b). Considérons l'ensemble de mesure nulle $E = \{x \in X, f(x) > g(x)\}$. On a

$$f = f \cdot \mathbf{1}_E + f \cdot \mathbf{1}_{E^c}, \quad g = g \cdot \mathbf{1}_E + g \cdot \mathbf{1}_{E^c}, \tag{1.7.2}$$

et $f \cdot \mathbf{1}_{E^c} \leq g \cdot \mathbf{1}_{E^c}$. Compte-tenu de la proposition 1.5.3(a) et du lemme 1.6.3, il suffit de prouver

$$\int_X f \cdot \mathbf{1}_E dm = 0 = \int_X g \cdot \mathbf{1}_E dm$$

ce qui est effectivement le cas car $\int_X f \cdot \mathbf{1}_E dm = \int_E f d\mu = 0$, d'après la proposition 1.5.3(c). En utilisant (1.7.2) pour $E = \{x \in X, f(x) \neq g(x)\}$, le lemme 1.6.3 et la proposition 1.5.3 (c), on démontre (c). Pour obtenir (d), on pose $E = \{f = +\infty\}$, et l'on remarque que $\mu(E) > 0$ implique, pour tout entier $n \geq 1$,

$$\int_X f d\mu \geq \int_E f d\mu \geq n \int_E d\mu = n\mu(E),$$

ce qui donne $\int_X f d\mu = +\infty$. $\square$

Définition 1.7.2. Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. L'espace $L^1(\mu)$ est le quotient de $\mathcal{L}^1(\mu)$ (cf. définition 1.6.6) par la relation d'équivalence d'égalité μ-pp ($f \sim g$ signifie $\mu(\{x \in X, f(x) \neq g(x)\}) = 0$).

Remarquons que $L^1(\mu)$ est un espace vectoriel comme quotient de $\mathcal{L}^1(\mu)$ par le sous-espace vectoriel $\{f \in \mathcal{L}^1(\mu), f \sim 0\}$.[29] Par ailleurs, l'application linéaire $f \mapsto \int_X f d\mu$ définie sur $\mathcal{L}^1(\mu)$ passe au quotient, i.e. ne dépend que de la classe d'équivalence de f.[30] De même, si $f, g \in \mathcal{L}^1(\mu)$ sont à valeurs réelles et

$$f \leq g \ \mu\text{-pp, alors} \int_X f d\mu \leq \int_X g d\mu. \tag{1.7.3}$$

Ceci est une conséquence immédiate de la proposition 1.6.7 et du fait que

$$g - f \sim (g - f)\mathbf{1}_{N^c} \geq 0, \quad \text{avec } \mu(N) = 0,$$

donnant le résultat (1.7.3) en utilisant la proposition 1.7.1.

Théorème 1.7.3. *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive.*
(a) L'application de $L^1(\mu)$ dans $\mathbb{C}$ définie par $f \mapsto \int_X f d\mu$ est une forme linéaire.
(b) L'application de $L^1(\mu)$ dans $\mathbb{R}_+$ définie par $f \mapsto \int_X |f| d\mu = \|f\|_{L^1(\mu)}$ est une norme et pour $f \in L^1(\mu)$

$$\left| \int_X f d\mu \right| \leq \int_X |f| d\mu. \tag{1.7.4}$$

démonstration. Le (a) est démontré par la remarque précédente, et pour la même raison, l'application de (b) est bien définie sur l'espace quotient $L^1(\mu)$. Si $f \in \mathcal{L}^1(\mu)$

[29] Si $f_1, f_2 \in \mathcal{L}^1(\mu)$ sont nulles respectivement sur N_1^c, N_2^c où $\mu(N_j) = 0$, alors pour $\alpha_1, \alpha_2 \in \mathbb{C}$, on a $\alpha_1 f_1 + \alpha_2 f_2 = 0$ sur $(N_1 \cup N_2)^c$ donc μ-pp car $\mu(N_1 \cup N_2) = 0$.

[30] En effet, si $f \sim 0$ on a $\int_X f d\mu = \int_X (\operatorname{Re} f)_+ d\mu - \int_X (\operatorname{Re} f)_- d\mu + i \int_X (\operatorname{Im} f)_+ d\mu - i \int_X (\operatorname{Im} f)_+ d\mu = 0$ d'après la proposition 1.7.1(e).

vérifie $\|f\|_{L^1(\mu)} = 0$, la proposition 1.7.1(a) implique $f \sim 0$, i.e. $f = 0$ dans $L^1(\mu)$. La proposition 1.5.3(b) donne l'homogénéité de cette application, tandis que l'inégalité triangulaire découle de $\|f+g\|_{L^1(\mu)} = \int_X |f+g|d\mu \le \int_X(|f|+|g|)d\mu = \|f\|_{L^1(\mu)} + \|g\|_{L^1(\mu)}$. Pour terminer, démontrons (1.7.4). Posons

$$z = \int_X f d\mu = |z|e^{i\theta}.$$

On a, en utilisant la proposition 1.6.7, la définition 1.6.6 et (1.7.3),

$$\left|\int_X f d\mu\right| = \mathrm{Re}\left(e^{-i\theta}\int_X f d\mu\right) = \mathrm{Re}\int_X e^{-i\theta} f d\mu = \int_X \mathrm{Re}(e^{-i\theta}f)d\mu \le \int_X |e^{-i\theta}f|d\mu = \int_X |f|d\mu.$$

□

THÉORÈME 1.7.4. THÉORÈME DE CONVERGENCE DOMINÉE DE LEBESGUE. *Soit $(X, \mathcal{M}, \mu)$ un espace mesuré où μ est une mesure positive. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables de X dans $\mathbb{C}$ telle que les propriétés suivantes soient vérifiées.*

(1) Convergence simple. *Pour μ-presque tout*[31] $x \in X$, $\lim_{n\to\infty} f_n(x) = f(x)$.

(2) Domination. *Il existe $g : X \to \overline{\mathbb{R}}_+$ mesurable, vérifiant $\int_X g d\mu < +\infty$, telle que*

$$\forall n \in \mathbb{N}, \textit{pour } \mu\textit{-presque tout}^{32}\ x \in X, \qquad |f_n(x)| \le g(x).$$

Alors la fonction f est[33] *mesurable et l'on a $\int_X |f|d\mu < +\infty$ et*

$$\lim_{n\to\infty}\int_X |f - f_n|d\mu = 0, \quad \textit{ce qui implique} \quad \lim_{n\to\infty}\int_X f_n d\mu = \int_X f d\mu.$$

DÉMONSTRATION. En tenant compte des notes en bas de page, on pose

$$B = N \cup \cup_{n\in\mathbb{N}} M_n, \quad (\text{on a } B \in \mathcal{M} \text{ et } \mu(B) = 0), \quad \widetilde{f}(x) = \lim_n f_n(x)\mathbf{1}_{B^c}(x).$$

La suite $\widetilde{f_n} = \mathbf{1}_{B^c} f_n$ vérifie les hypothèses du théorème 1.6.8. Par suite, on obtient $\widetilde{f} \in \mathcal{L}^1(\mu)$ et

$$\lim_{n\to+\infty}\int_X |\widetilde{f_n} - \widetilde{f}|d\mu = 0.$$

Comme on a $|f - f_n| = |\widetilde{f} - \widetilde{f_n}| + |f - f_n|\mathbf{1}_B$ et $f = \widetilde{f} + f\mathbf{1}_B$ avec $\mu(B) = 0$, il vient de la proposition 1.5.3(c) que $f \in \mathcal{L}^1(\mu)$ et le résultat

$$\lim_{n\to+\infty}\int_X |f_n - f|d\mu = 0.$$

[31] Il existe $N \in \mathcal{M}$, tel que $\mu(N) = 0$ et $\forall x \in N^c$ la suite $(f_n(x))_{n\in\mathbb{N}}$ est convergente de limite $f(x)$.

[32] $\forall n \in \mathbb{N}, \exists M_n \in \mathcal{M}$ avec $\mu(M_n) = 0$ tel que $\forall x \in M_n^c$, $|f_n(x)| \le g(x)$.

[33] On définit $f(x) = \mathbf{1}_{N^c}(x)\lim_n f_n(x)$.

Remarque. Ce théorème peut s'exprimer de manière plus élégante et synthétique en disant simplement que si $(f_n)_{n\in\mathbb{N}}$ est une suite de $L^1(\mu)$ convergeant simplement vers f avec une condition de domination $|f_n| \leq g \in L^1(\mu)$, alors f_n converge vers f dans l'espace $L^1(\mu)$. En résumé, pour une suite f_n de $L^1(\mu)$,

$$\left.\begin{array}{c} f_n \xrightarrow{\text{simplement}} f \\ \text{et} \\ |f_n| \leq g \in L^1(\mu) \end{array}\right\} \quad \text{implique} \qquad f_n \xrightarrow{L^1(\mu)} f. \tag{1.7.5}$$

Notes

Passons tout d'abord en revue les quelques noms de mathématiciens rencontrés au fil de ce chapitre. Des détails supplémentaires peuvent être obtenus sur le web, par exemple à l'adresse

`http:// www-history.mcs.st-andrews.ac.uk/history/Mathematicians/`

Le baron Augustin **Cauchy**, mathématicien français (est-il besoin de le rappeler?) vécut de 1789 à 1857 et est l'un des fondateurs de l'analyse. Stefan **Banach** (1892-1945) mathématicien polonais, est l'un des pères de l'analyse fonctionnelle. Emile **Borel** (1871-1956), mathématicien et homme politique français, est l'un des fondateurs de la théorie de la mesure. René **Baire** (1874-1932) est un mathématicien français ; le théorème du même nom constitue le fondement de l'analyse fonctionnelle. Le lecteur aura noté l'orthographe du patronyme **Bernoulli**, qui comporte un seul " i ". Les frères Jacques (1654-1705) et Jean (1667-1748) Bernoulli ainsi que Daniel (1700-1782), fils de Jean, ont vécu à Bâle et ont contribué au développement du calcul intégral (Jacques), de la mécanique (Jean) et de la théorie cinétique des gaz (Daniel). La contribution de Jacques (qui est celui cité dans l'exemple (10) du §1.4) au calcul des probabilités est également célèbre à cause de la *Loi des grands nombres* sur laquelle nous aurons l'occasion de revenir. Paul **Dirac** (1902-1984) est un physicien britannique, dont les travaux sont fondateurs de la mécanique quantique. Pierre-Simon **Laplace** (1749-1827) est un mathématicien et astronome français. Carl-Friedrich **Gauss** (1777-1855) est le plus grand mathématicien allemand de son époque. Denis **Poisson** (1781-1840) est un mathématicien français. Henri **Lebesgue** (1875-1941) a formulé la théorie moderne de la mesure en 1901, généralisant en particulier l'intégrale de Riemann. Bernhard **Riemann** (1826-1866) est un mathématicien allemand aux contributions très variées, de la théorie des nombres à l'analyse mathématique. Felix **Hausdorff** (1869-1942) est un mathématicien allemand, fondateur de la topologie générale. Beppo **Levi** (1875-1961) est un mathématicien italien, professeur à l'université de Gênes et également grand spécialiste de géométrie algébrique ; il fut contraint en 1938 à l'exil en Argentine par les lois antisémites du fascisme mussolinien. Il y a actuellement un Institut Beppo Levi de recherche mathématique dans la ville argentine de Rosario. Pour en savoir plus, on pourra consulter le site (italien)

`http://www.math.unifi.it/matematicaitalia`

Pierre **Fatou** (1878-1929) est un mathématicien français, auteur du lemme démontré ci-dessus, qui est l'un des fondements de la théorie de la mesure.