CHAPITRE 2

INTEGRATION ET MESURES

La théorie de l'intégrale de Lebesgue, ou théorie de la mesure, est un outil d'intégration beaucoup plus puissant que la théorie de l'intégrale de Riemann et pas tellement plus récent (vers 1860 pour Riemann, vers 1905 pour Lebesgue). Il n'y aurait donc aucune raison de faire l'impasse sur la théorie de la mesure, si ce n'était sa plus grande difficulté. Mais cette relative difficulté concerne surtout l'entrée dans la théorie et non son utilisation : une fois établis les résultats fondamentaux (convergence monotone, convergence dominée, espaces L^1 et L^2, Lebesgue-Fubini), le maniement de l'outil devient simple et permet d'accéder à un domaine très vaste. La théorie de la mesure est un prérequis indispensable, notamment, pour la théorie des distributions et en même temps donne un avant-goût de celle-ci. Ceci est particulièrement vrai pour le mode d'exposition choisi ici (définition "fonctionnelle" de la mesure, par opposition à la définition "ensembliste", souvent utilisée mais, somme toute, peut-être plus abstraite).

Dans la Section **(1.3.1)**, nous avons introduit les fonctions continues à support compact définies sur un espace métrique X et à valeurs dans le corps des scalaires $\mathbf{F}$ ($\mathbf{R}$ ou $\mathbf{C}$) [1]. Nous noterons $\mathsf{K} = \mathsf{K}(X)$ (ou $\mathsf{K}_{\mathbf{R}}(X)$ et $\mathsf{K}_{\mathbf{C}}(X)$ s'il y a risque de confusion) le sous-espace vectoriel de C (X, $\mathbf{F}$) constitué de toutes ces fonctions; l'espace K est muni de la norme de la convergence uniforme (mais on peut vérifier qu'il n'est pas complet pour cette norme si X n'est pas compact). Nous appellerons *espace de base* X l'espace métrique sur lequel sont définies les fonctions que l'on veut intégrer.

Dans la Section facultative **(1.3.3)**, nous avons admis deux résultats qui seront utilisés dans la partie facultative de la partie **2** de ce chapitre, et qui sont valables si X est, de plus, séparable. Dans tout ce chapitre, nous supposerons donc que l'espace de base X est un espace métrique séparable et localement compact.

En pratique, l'espace de base X *sera un espace* $\mathbf{R}^n$ *ou un sous-ensemble fermé, ou ouvert, d'un espace* $\mathbf{R}^n$*, ou encore une partie quelconque d'un espace discret* $\mathbf{N}^n$ [2].

2.1 Définition d'une mesure, exemples et premières propriétés

(2.1.1) *Définition d'une mesure. Limite simple d'une suite de mesures*

(a) Supposons d'abord que l'espace de base X soit *compact*. Dans ce cas, l'espace K s'identifie à l'espace C de toutes les fonctions continues à valeurs scalaires et, de plus, est un espace de Banach lorsqu'on le munit de la norme de la convergence uniforme $\|\ \|_\infty$ **(1.2.4)**. On appelle alors *mesure sur* X toute *forme linéaire continue sur* K, donc tout élément du dual K'. Pour qu'une forme linéaire μ sur K soit une mesure, il faut et il suffit donc **(1.2.2** (i)**)** qu'il existe un réel positif a tel que l'on ait, pour toute fonction $f \in \mathsf{K}$, $|\mu(f)| \le a \, \|f\|_\infty$.

---

[1] Nous avons montré en option, Proposition **(1.3.1)** (iii), qu'il en existe beaucoup lorsque X est *localement compact* - cas incluant les espaces $\mathbf{R}^n$ et leurs parties fermées et leurs parties ouvertes, **(1.1.7)** (viii).

[2] En effet, les espaces $\mathbf{R}^n$ et toutes leurs parties sont séparables **(1.3.3)**. Les espaces $\mathbf{N}^n$ le sont aussi, puisqu'ils sont dénombrables. De plus un *espace discret* X (comme $\mathbf{N}^n$) est caractérisé par le fait que, pour chaque point x de X, le singleton $\{x\}$ est un voisinage de x. La topologie correspondante est $\mathsf{P}(X)$ tout entier et peut être définie par la distance $d(x, y) = 1$ si $x = y$ et 0 si $x \neq y$. Toute partie d'un espace discret est donc un espace discret. Il est clair qu'un singleton est un ensemble compact, donc tout espace discret est localement compact.

(b) Abandonnons maintenant l'hypothèse supplémentaire que l'espace de base X soit compact. On appelle *mesure sur* X toute *forme linéaire* μ *sur* K qui vérifie la propriété suivante : pour *tout* compact K de X, il existe un réel positif a_K tel que l'on ait, pour toute fonction $f \in \mathsf{K}$ *dont le support est inclus dans* K, $|\mu(f)| \leq a_K \| f \|_\infty$ (on note K(X ; K) l'ensemble de ces fonctions). Cette propriété signifie que la restriction de μ au sous-espace K (X ; K) de K est continue, et ceci doit être vrai quel que soit le compact K de X.

On note aussi $\mu(f) = \int_X f(x)\, d\mu(x) = \int_X f\, d\mu$, notation qui va être justifiée très bientôt.

(i) *Proposition*. Si (μ_n) est une suite de mesures qui converge simplement, i.e. si pour toute fonction f de K, $\mu_n(f)$ tend vers une limite $\mu(f)$, alors μ est une mesure sur X.

Soit K un compact de X, nous devons montrer que la restriction de μ à K (X ; K) est une forme linéaire continue. Montrons d'abord que K (X ; K) est un espace de Banach : si une suite (f_n) de fonctions continues à support inclus dans K est une suite de Cauchy pour $\| \ \|_\infty$, alors elle admet une limite f qui est continue et bornée **(1.2.4** (ii)**)**; on a évidemment $f(x) = 0$ si $x \notin$ K, donc Supp (f) $\subset$ K et par conséquent $f \in \mathsf{K}$ (X ; K) puisque toute partie fermée d'un compact est compacte **(1.1.7** (vii)**)**. Mais puisque K (X ; K) est un espace de Banach, nous sommes dans les conditions d'application du théorème **(1.3.4** (iv)**)** qui dit justement que l'application $f \to \mu(f)$ de K (X ; K) dans $\mathbf{F}$ est linéaire et continue, ce que nous devions montrer.

(2.1.2) *Exemples de mesures*

(i) Soit x un point quelconque de l'espace de base, et posons $\delta_x(f) = f(x)$. La linéarité est évidente et l'on a $|\delta_x(f)| \leq \| f \|_\infty$ pour toute fonction $f \in \mathsf{K}$ (et en fait pour toute fonction bornée), indépendamment du fait que Supp(f) soit ou non dans un compact fixé à l'avance. On dit que δ_x est la *mesure de Dirac au point x*.

(ii) Plus généralement, soient (a_n) une suite de points de X et (p_n) une suite de réels (les *poids affectés aux* a_n) telles que, pour tout compact K de X, la série $\Sigma_k\, p_{n_k}$, formée avec les poids correspondant aux points a_{n_k} de la suite (a_n) qui sont dans K, soit absolument convergente (peu importe l'ordre (n_k) choisi pour ces points, puisqu'on demande que la série soit absolument convergente). Si $f \in \mathsf{K}$(X ; K), la série $\Sigma_n\ p_n\ f(a_n)$ se réduit à la série $\Sigma_k\ p_{n_k}\ f(a_{n_k})$, qui est absolument convergente à cause de la majoration :

$$\left| \sum_{k=1}^{N} p_{n_k}\ f(a_{n_k}) \right| \leq \sum_{k=1}^{N} \left| p_{n_k} \right| \left| f(a_{n_k}) \right| \leq \|f\|_\infty \sum_{k=1}^{N} \left| p_{n_k} \right| \leq \|f\|_\infty \sum_{k=1}^{\infty} \left| p_{n_k} \right|. \qquad (1)$$

En posant $\mu(f) = \Sigma_n\ p_n\ f(a_n)$, on définit donc une forme linéaire sur K, et cette forme linéaire est une mesure, à cause de la même majoration. Si l'on remplace la suite (a_n) par une suite finie, on obtient une combinaison linéaire de mesures de Dirac (un barycentre si $p_n \geq 0$ et $\Sigma_n\ p_n = 1$).

(iii) Nous savons déjà intégrer, de la façon la plus classique, les fonctions scalaires f d'une variable réelle qui sont définies et continues sur un intervalle compact [a, b] de $\mathbf{R}$: cela consiste à poser $\int_a^b f(x)\, dx = F(b) - F(a)$ où F est une primitive de f ; et l'existence d'une primitive se démontre à partir des primitives des fonctions en escalier, et du fait que toute fonction continue sur [a, b] est limite uniforme de fonctions en escalier. En particulier, nous savons donc intégrer toute fonction scalaire continue à support compact définie sur **R**, car la valeur $\int_a^b f(x)\,dx$ est évidemment indépendante de l'intervalle compact $[a, b]$ contenant Supp(f), et on la note $\int_{\mathbf{R}} f(x)\; dx$. En posant, pour $f \in \mathsf{K}(\mathbf{R})$, $\lambda(f) = \int_{\mathbf{R}} f(x)\; dx$, on définit une forme linéaire λ sur $\mathsf{K}(\mathbf{R})$, et cette forme linéaire est une mesure car on a

$$|\lambda(f)| = |\int_{\mathbf{R}} f(x)\; dx\;| \leq \int_{\mathbf{R}} |f(x)|\; dx \leq (b - a)\; \|f\|_\infty \;\text{ si }\; [a, b] \supset \mathrm{Supp}(f). \tag{2}$$

On dit que λ est la *mesure de Lebesgue* sur **R**. On peut donc noter $\int_{\mathbf{R}} f(x)\; dx = \int_{\mathbf{R}} f\, d\lambda$.

(iv) Soient μ une mesure sur X et g une fonction continue quelconque de X dans **F**. Pour toute fonction f de K, la fonction fg est dans K (car Supp(fg)$\subset$ Supp(f)$\cap$Supp(g)) et l'on a pour tout compact K de X et toute fonction f de K(X ; K) : $\| fg \|_\infty \leq \| f \|_\infty . \| g_{|K} \|_\infty$. Si $|\mu(f)| \leq a_K \| f \|_\infty$ pour toute fonction f de K(X ; K), on a donc aussi

$$|\mu(fg)| \leq a_K \| g_{|K} \|_\infty \; \| f \|_\infty,$$

donc la forme linéaire ν définie par

$$\nu(f) = \mu(fg)\;,\text{ ou } \int_X f\, d\nu = \int_X fg\; d\mu \tag{3}$$

est une mesure sur X. On dit que ν est une *mesure de base* μ, et plus précisément que ν est la *mesure de densité g par rapport à* μ.

(2.1.3) *Mesures positives*

Dans la définition **(2.1.1)**, le corps des scalaires peut être aussi bien **R** que **C**, et on parle alors de mesures *réelles* (resp. *complexes*), qui sont donc des formes linéaires sur $\mathsf{K}_{\mathbf{R}}$(X) (resp. sur $\mathsf{K}_{\mathbf{C}}$(X)) telles que, pour tout compact K de X, leur restriction au sous-espace $\mathsf{K}_{\mathbf{R}}$(X ; K) (resp. $\mathsf{K}_{\mathbf{C}}$(X ; K)) soit continue.

On dit qu'une mesure complexe est *réelle* (!) si $\mu(f)$ est un réel lorsque f prend ses valeurs dans **R**. Ceci est cohérent car (α) par définition, toute mesure complexe réelle donne par restriction aux fonctions réelles une mesure réelle tout court, et (β) toute mesure réelle tout court μ_0 est la restriction aux fonctions réelles d'une mesure complexe réelle μ : il suffit de poser $\mu(f + ig) = \mu_0(f) + i\;\mu_0(g)$ pour f et g réelles. On peut donc parler d'une mesure réelle sans préciser autrement.

On va s'intéresser surtout à celles des mesures réelles qui, de plus, sont *positives*, i.e. vérifient $\mu(f) \geq 0$ si $f \geq 0$. C'est évidemment le cas des mesures de Dirac et de la mesure de Lebesgue. La positivité dispense de vérifier la continuité des restrictions :

(i) *Proposition*. Si une forme linéaire μ sur $\mathsf{K}_{\mathbf{R}}$(X) vérifie $\mu(f) \geq 0$ si $f \geq 0$, c'est une mesure (positive).

Preuve: Toute fonction réelle f s'écrit $f = f^+ - f^-$, où $f^+ = \sup(f, 0)$ et $f^- = -\inf(f, 0) = \sup(-f, 0)$ sont positives; on a si f est bornée : $\| f \|_\infty = \sup(\| f^+\|_\infty,\; \| f^-\|_\infty)$. A cause de la linéarité de μ, il nous suffira donc de montrer qu'une majoration $|\mu(f)| \leq a_K \| f \|_\infty$ existe pour les fonctions *positives* de $\mathsf{K}_{\mathbf{R}}$(X ; K), K étant un compact. Il existe dans K une fonction g qui prend ses valeurs dans [0, 1] et qui est égale à 1 dans K **(1.3.1** (iii)**)**. On a alors si $f \geq 0$ et Supp(f) $\subset$ K : $f \leq g \| f \|_\infty$, d'où par linéarité et puisque μ est positive : $0 \leq \mu(f) \leq \mu(g) \| f \|_\infty$, cqfd.

2.2 Prolongement d'une mesure positive aux fonctions intégrables

C'est cette partie qui constitue le "gros morceau à avaler" pour accéder à la théorie. Mais seuls le principe de la construction et quelques résultats cruciaux obtenus dans cette construction sont à retenir et feront partie du programme, le reste étant facultatif et en petits caractères.

(2.2.1) *Intégrale supérieure (resp. inférieure) de certaines fonctions semi-continues inférieurement (resp. supérieurement) par rapport à une mesure positive*

Nous allons prolonger toute mesure positive μ de façon à pouvoir intégrer f par rapport à μ (i.e. définir $\mu(f)$) pour des fonctions f beaucoup moins régulières que les fonctions continues, et ce prolongement aura la propriété très agréable que l'intégrale de la limite d'une suite croissante (ou d'une suite décroissante) soit la limite de l'intégrale de la suite. Pour étendre l'intégration depuis les fonctions continues à support compact, on peut justement penser à un passage à une limite croissante. Ceci oblige à considérer d'abord des fonctions f qui sont effectivement le sup d'une suite croissante (g_n) de fonctions continues à support compact (et qui sont donc aussi telles que $\forall x \in$ X, $g_n(x)$ *tend vers* $f(x)$ *dans* $\overline{\mathbf{R}}$: si le sup est fini, c'est une limite dans $\mathbf{R}$ et sinon, $g_n(x)$ tend vers $+\infty$). Or, il se trouve que c'est le cas de toutes les fonctions *s.c.i.* qui sont *minorées par une fonction de* K **(1.3.3** (ii)**)**. Nous noterons I l'ensemble de ces fonctions. Rappelons que $f \to -f$ intervertit les fonctions s.c.i. et s.c.s.; par suite, toute fonction *s.c.s. majorée par une fonction de* K est l'inf (donc la limite ponctuelle dans $\overline{\mathbf{R}}$) d'une suite décroissante de fonctions de K, nous noterons S l'ensemble de ces fonctions. Sauf mention du contraire, les valeurs des fonctions sont dans $\overline{\mathbf{R}}$ (exception : les fonctions de K sont, par hypothèse, des fonctions de X dans $\mathbf{R}$; de plus, les fonctions de K sont bornées comme fonctions continues à support compact).

Définition. Si $f \in \mathsf{I}$, on appelle *intégrale supérieure* de f, le nombre (réel ou égal à $+\infty$)

$$\mu^*(f) = \int^* f\, d\mu = \sup\ \{\mu(g);\ g \in \mathsf{K},\ g \le f\}. \tag{1}$$

Ce nombre ne serait pas défini si f n'était pas minorée par une fonction de K, puisqu'on prendrait la borne supérieure de l'ensemble vide! Et ce ne sont pas toutes les fonctions s.c.i. f qui sont minorées par une fonction de K, car ceci implique en particulier que f ne prenne pas la valeur $-\infty$ et que f soit positive en dehors d'un compact, à savoir K=Supp(g) si $g \le f$ avec $g \in \mathsf{K}$. Par contre, c'est évidemment le cas de toutes les fonctions $f \in \mathsf{K}$ et, bien sûr, on a alors $\mu^*(f) = \mu(f)$. Enfin, il résulte de la définition (1) de μ^* et de la positivité de μ que $\mu^*(f) \le \mu^*(h)$ si $f \le h$, lorsque f et h sont dans I.

De même, si $f \in \mathsf{S}$, on appelle *intégrale inférieure* de f, le nombre (réel ou égal à $-\infty$)

$$\mu_*(f) = \int_* f\, d\mu = \inf\ \{\mu(g);\ g \in \mathsf{K},\ g \ge f\}. \tag{2}$$

(i) *Proposition*. Soit (f_n) une suite croissante de fonctions de I. On a

$$\mu^*(\sup_n f_n) = \sup_n \mu^*(f_n) = \lim_{n \to \infty} \mu^*(f_n). \tag{3}$$

Posons $f = \sup_n f_n$, qui est s.c.i. et minorée par toute fonction de K qui minore un f_n ; ainsi f est dans I, de sorte que le premier membre est au moins défini. Le deuxième membre et le troisième sont sûrement définis et égaux puisque $(\mu^*(f_n))$ est une suite croissante et que nous acceptons que son sup puisse être $+\infty$. Puisque $f \geq f_n$ on a de même $\mu^*(f) \geq \mu^*(f_n)$ pour tout n. Il est donc clair que $\mu^*(f) \geq \sup_n \mu^*(f_n)$, c'est l'autre inégalité qui reste à obtenir.

Montrons que le résultat est vrai si f_n et f sont dans K. Comme on a $f_1 \leq f_n \leq f$, les compacts K_n = Supp (f_n) et K = Supp (f) sont tous inclus dans un même compact, à savoir K' = $K_1 \cup$K. Donc l'hypothèse faite et le théorème de Dini **(1.2.4** (iii)**)** entraînent que f_n tend uniformément vers f. Comme on peut alors enlever les étoiles dans (3), le résultat découle donc de l'existence d'un nombre a tel que $|\mu(g)| \leq a \| g \|_\infty$ si $g \in$ K et Supp$(g) \subset$ K' [en d'autres termes, il découle de la continuité de la restriction de μ à K(X ; K') pour la norme $\| \;\; \|_\infty$].

Montrons maintenant que $\mu^*(f) \leq \sup_n \mu^*(f_n)$ dans le cas général, i.e., vue la définition (1) de $\mu^*(f)$, que si $u \in$ K et $u \leq f$, on a $\mu^*(u) \leq \sup_n \mu^*(f_n)$. Pour tout n, il existe une suite croissante (g_{mn}) de fonctions de K telle que $f_n = \sup_m g_{mn}$ **(1.3.3** (ii)**)**. Posons $h_n = \sup_{p,q \leq n} g_{pq}$ ("procédé diagonal" souvent utilisé). On a $h_n \leq f_n$ car $g_{pq} \leq g_{nq} \leq f_q \leq f_n$ si $p \leq n$ et $q \leq n$, mais $\sup_n h_n = f$. Puisque $u \leq f$, on a $u = \sup_n k_n$ avec $k_n = \inf(u, h_n)$. En effet, on a $k_n \leq u$ donc déjà $\sup_n k_n \leq u$. Et s'il existait un $x \in$ X tel que $\sup_n k_n(x) < u(x)$, on aurait $k_n(x) < u(x)$ pour tout n puisque la suite (k_n) est croissante comme la suite (h_n). On aurait donc $k_n(x) = h_n(x)$ pour tout n, d'où $\sup_n k_n(x) = f(x)$ puisque $\sup_n h_n = f$, mais ceci entraîne $f(x) < u(x)$, ce qui est une contradiction. De $u = \sup_n k_n$, on déduit $\mu^*(u) = \sup_n \mu^*(k_n)$ puisque u et les k_n sont dans K . Comme $k_n \leq h_n \leq f_n$, on en déduit $\mu^*(u) \leq \sup_n \mu^*(h_n) \leq \sup_n \mu^*(f_n)$, cqfd.

Soient f_1 et f_2 deux fonctions de I. La somme $f_1 + f_2$ est partout définie, car aucune des deux ne prend la valeur $-\infty$, donc $f_1 + f_2$ est s.c.i. **(1.3.2** (iii)**)**; de plus $f_1 + f_2 \geq g_1 + g_2$ si $f_1 \geq g_1$ et $f_2 \geq g_2$, donc $f_1 + f_2$ est dans I. On a l'additivité pour l'intégrale supérieure des fonctions de I :

(ii) *Proposition*. Si $f_1 \in$ I et $f_2 \in$ I, alors $\mu^*(f_1 + f_2) = \mu^*(f_1) + \mu^*(f_2)$.

Soient $(g_{n\,1})$ et $(g_{n\,2})$ des suites croissantes de fonctions de K qui tendent vers f_1 et f_2, ce qui entraîne que la suite croissante $(g_{n\,1} + g_{n\,2})$ tend vers $f_1 + f_2$. Alors, par (i), $\mu^*(g_{n\,1})$ tend vers $\mu^*(f_1)$, $\mu^*(g_{n\,2})$ tend vers $\mu^*(f_2)$ et $\mu^*(g_{n\,1} + g_{n\,2})$ tend vers $\mu^*(f_1 + f_2)$. On a donc puisque $g_{n\,1}$ et $g_{n\,2}$ sont des fonctions de K : $\mu(g_{n\,1}) + \mu(g_{n\,2}) = \mu^*(g_{n\,1}) + \mu^*(g_{n\,2}) \to \mu^*(f_1) + \mu^*(f_2)$ et $\mu(g_{n\,1} + g_{n\,2}) = \mu^*(g_{n\,1} + g_{n\,2}) \to \mu^*(f_1 + f_2)$. Mais comme $\mu(g_{n\,1}) + \mu(g_{n\,2}) = \mu(g_{n\,1} + g_{n\,2})$, c'est fini.

L'additivité s'étend aux sommes finies, donc on a $\mu^*\left(\sum_{n=1}^{N} f_n\right) = \sum_{n=1}^{N} \mu^*(f_n)$ si les f_n sont dans I. Dans le cas où les f_n sont positives, la suite des sommes partielles est une suite croissante et en appliquant (i), on obtient donc :

(iii) *Proposition*. Soit (f_n) une suite de fonctions ≥ 0 de I. Alors $\mu^*\left(\sum_{n=1}^{\infty} f_n\right) = \sum_{n=1}^{\infty} \mu^*(f_n)$.

Bien entendu, on a les résultats correspondant à (i), (ii) et (iii) pour l'intégrale inférieure des fonctions de S. Il n'y a même pas besoin de transposer les démonstrations car on a, d'après les définitions, $(f \in$ I$) \Leftrightarrow (-f \in$ S), et $\mu^*(f) = -\mu_*(-f)$ si $f \in$ I. Pour (i), il faut évidemment remplacer "croissante" par "décroissante" et "sup" par "inf".

(2.2.2) *Intégrales supérieure et inférieure de fonctions quelconques par rapport à une mesure positive*

Soit donc f une fonction absolument quelconque de X dans $\overline{\mathbf{R}}$. Il existe des fonctions g de I qui majorent f, même s'il n'y avait que la fonction constante $g(x) \equiv +\infty$ (par contre il n'existe pas toujours une fonction de I qui minore f, puisque ceci implique par transitivité, ici aussi, que f ne prenne pas la valeur $-\infty$ et que f soit positive en dehors d'un compact). On appelle *intégrale supérieure* de f, le nombre

$$\mu^*(f) = \int^* f\, d\mu = \inf\{\mu^*(g);\, g\in \mathsf{I},\, g \geq f\} \qquad (1)$$

qui, cette fois, peut prendre n'importe quelle valeur dans $\overline{\mathbf{R}}$. Il est clair que cette définition coïncide avec la précédente lorsque $f\in \mathsf{I}$ et donc que $\mu^*(f) = \mu(f)$ lorsque $f\in \mathsf{K}$. Rappelons que l'on a $\mu^*(f_1) \leq \mu^*(f_2)$ si $f_1 \leq f_2$ lorsque f_1 et f_1 sont des fonctions de I. Le lecteur vérifiera que la définition (1) transfère cette propriété ***très importante*** à *toutes* les fonctions de X dans $\overline{\mathbf{R}}$:

On a $\mu^*(f_1) \leq \mu^*(f_2)$ si $f_1 \leq f_2$.

Le même raisonnement par "transport" montre aussi que $\mu^*(af) = a\,\mu^*(f)$ pour tout réel a.

Si f_1 et f_2 sont deux fonctions de X dans $\overline{\mathbf{R}}$ telles que leur somme soit partout définie, et si g_1 et g_2 sont deux fonctions de I telles que $g_1 \geq f_1$ et $g_2 \geq f_2$, alors $g_1 + g_2 \geq f_1 + f_2$ et $g_1 + g_2 \in \mathsf{I}$, donc **(2.2.1** (ii)**)** :

$$\mu^*(g_1 + g_2) = \mu^*(g_1) + \mu^*(g_2) \;\geq\; \mu^*(f_1 + f_2). \qquad (\$)$$

Prenant l'inf par rapport à g_1, nous obtenons $\mu^*(f_1) + \mu^*(g_2) \;\geq\; \mu^*(f_1 + f_2)$ si la somme a un sens, ce qui sera vrai (puisque $\mu^*(g_2) > -\infty$) sauf dans le cas où ($\mu^*(f_1) = -\infty$ *et* $\mu^*(g_2) = +\infty$) - cas qui peut être exclus à condition qu'on n'ait pas ($\mu^*(f_1) = -\infty$ *et* $\mu^*(f_2) = +\infty$). Prenant enfin l'inf par rapport à g_2, nous obtenons, en excluant de même les cas d'indétermination : $\mu^*(f_1) + \mu^*(f_2) \;\geq\; \mu^*(f_1 + f_2)$. On peut par exemple vérifier que ($) entraîne bien $\mu^*(f_1 + f_2) = -\infty$ si $\mu^*(f_1)\in \mathbf{R}$ et $\mu^*(f_2) = -\infty$. Donc :

(i) *Proposition*. Si f_1 et f_2 sont deux fonctions de X dans $\overline{\mathbf{R}}$ telles que leur somme soit partout définie et que la somme $\mu^*(f_1) + \mu^*(f_2)$ soit définie, on a

$$\mu^*(f_1 + f_2) \leq \mu^*(f_1) + \mu^*(f_2). \qquad (2)$$

(ii) ***Proposition***. *Soit* (f_n) *une suite croissante de fonctions de* X *dans* $\overline{\mathbf{R}}$ *telle que* $\mu^*(f_n) > -\infty$ *pour n assez grand. Alors*

$$\mu^*(\sup_n f_n) = \sup_n \mu^*(f_n) = \lim_{n\to\infty} \mu^*(f_n). \qquad (3)$$

Ceci est l'analogue de **(2.2.1** (i)**)** pour les fonctions quelconques. Posons $f = \sup_n f_n$. Comme pour **(2.2.1** (i)**)**, la chose à démontrer est que $\mu^*(f) \leq \sup_n \mu^*(f_n)$. L'hypothèse entraîne que la suite ($\mu^*(f_n)$) est croissante; on peut donc supposer que $\mu^*(f_n) > -\infty$ a lieu pour tout n.

Pour tout $\varepsilon > 0$, on va construire une suite croissante (g_n) de fonctions de I telles que $f_n \leq g_n$ et $\mu^*(g_n) \leq \mu^*(f_n) + \varepsilon$. En posant $g = \sup_n g_n$, qui est dans I, il en résultera par **(2.2.1** (i)**)** que $\mu^*(g) = \sup_n \mu^*(g_n) \leq \sup_n \mu^*($

$f_n) + \varepsilon$, mais d'autre part $g = \sup_n g_n \geq f = \sup_n f_n$, donc on aura $\mu^*(f) \leq \mu^*(g) \leq \sup_n \mu^*(f_n) + \varepsilon$, ainsi ce sera terminé puisque ε est arbitraire.

Pour construire la suite (g_n), on note que, d'après la définition (1) de $\mu^*(f_n)$, il existe dans **I** une fonction h_n telle que $f_n \leq h_n$ et $\mu^*(f_n) \leq \mu^*(h_n) \leq \mu^*(f_n) + \varepsilon / 2^n$. On pose alors
$g_n = \sup(h_1, ..., h_n)$: la suite (g_n) est une suite croissante de fonctions de **I** et l'on a $f_n \leq g_n$. Montrons par récurrence que $\mu^*(g_n) \leq \mu^*(f_n) + \varepsilon(1 - 1/2^n)$. C'est vrai pour $n = 1$ puisque $g_1 = h_1$. Passons de n à $n + 1$: on a $g_{n+1} = \sup(g_n, h_{n+1})$ et $f_n \leq \inf(g_n, h_{n+1})$ d'où $\mu^*(g_{n+1}) = \sup(\mu^*(g_n), \mu^*(h_{n+1}))$ et $\mu^*(f_n) \leq \inf(\mu^*(g_n), \mu^*(h_{n+1}))$, et par suite
$\mu^*(g_{n+1}) + \mu^*(f_n) \leq \sup(\mu^*(g_n), \mu^*(h_{n+1})) + \inf(\mu^*(g_n), \mu^*(h_{n+1})) = \mu^*(g_n) + \mu^*(h_{n+1})$, i.e.
$\mu^*(g_{n+1}) \leq (\mu^*(g_n) - \mu^*(f_n)) + \mu^*(h_{n+1})$. D'où par l'hypothèse de récurrence et par construction de h_{n+1} : $\mu^*(g_{n+1}) \leq \varepsilon(1 - 1/2^n) + \mu^*(f_{n+1}) + \varepsilon/2^{n+1} = \mu^*(f_{n+1}) + \varepsilon(1 - 1/2^{n+1})$. C'est fini.

Comme dans la section précédente, on déduit de (i) que $\mu^*\left(\sum_{n=1}^{N} f_n\right) \leq \sum_{n=1}^{N} \mu^*(f_n)$ si les f_n sont des fonctions de X dans $\overline{\mathbf{R}}$ telle que $\mu^*(f_n) > -\infty$. Dans le cas où les f_n sont positives, cette condition est vérifiée et la suite des sommes partielles $\Sigma_n f_n$ est une suite croissante. En appliquant (ii), on obtient alors :

(iii) *Proposition*. Soit (f_n) une suite de fonctions ≥ 0 de X dans $\overline{\mathbf{R}}$. Alors on a :

$$\mu^*\left(\sum_{n=1}^{\infty} f_n\right) \leq \sum_{n=1}^{\infty} \mu^*(f_n).$$

Un autre moyen d'obtenir une suite croissante à partir d'une suite (f_n) de fonctions ≥ 0 est de considérer la suite (g_n) avec $g_n = \inf_{m \geq n} (f_m)$. Cette suite croissante tend simplement vers la fonction $\underline{\lim} f_n$ **(1.3.2)**. En appliquant (ii), nous obtenons donc $\mu^*(\underline{\lim} f_n) = \lim_{n \to \infty} \mu^*(g_n)$. On a, pour tout n, $\mu^*(g_n) \leq \inf_{m \geq n} \mu^*(f_m) = a_n$, et la suite (a_n) est décroissante et tend vers $\underline{\lim}\, \mu^*(f_n)$. Nous aboutissons donc au résultat suivant :

(iv) *Lemme de Fatou*. Si (f_n) est une suite de fonctions ≥ 0 de X dans $\overline{\mathbf{R}}$, on a :

$$\mu^*(\underline{\lim} f_n) \leq \underline{\lim}\, \mu^*(f_n).$$

Symétriquement, on appelle *intégrale inférieure* d'une fonction quelconque f de X dans $\overline{\mathbf{R}}$, le nombre suivant de $\overline{\mathbf{R}}$:

$$\mu_*(f) = \int_* f\, d\mu = \sup\,\{\mu_*(g);\, g\in \mathsf{S},\, g \leq f\}. \tag{4}$$

Là encore, on obtient en comparant (1) et (4) : $\mu_*(f) = -\,\mu^*(-f)$ (quelle que soit f), ce qui permet de transposer instantanément les résultats démontrés pour l'intégrale supérieure, à commencer par le fait que $\mu_*(f_1)\leq \mu_*(f_2)$ si $f_1\leq f_2$ (non, il n'y a pas d'inversion! vérifier). Et aussi le fait que la définition se réduit à celle qui a déjà été donnée, dans le cas où $f\in \mathsf{S}$. Les adjectifs "inférieure" et "supérieure" sont justifiés :

(v) *Proposition*. Pour toute fonction f de X dans $\overline{\mathbf{R}}$, on a $\mu_*(f) \leq \mu^*(f)$.

Il suffit de montrer que, si $h\in \mathsf{I}$ et $k\in \mathsf{S}$ sont deux fonctions telles que $k \leq f \leq h$, on a $\mu_*(k) \leq \mu^*(h)$, ou $\mu^*(h) + \mu^*(-k) \geq 0$. Mais comme $h\in \mathsf{I}$ et $-k\in \mathsf{I}$, on a $\mu^*(h) + \mu^*(-k) = \mu^*(h - k)$ **(2.2.1** (ii)**)** et puisque $h - k \geq 0$, on a $\mu^*(h - k) \geq 0$.

Lorsque nous saurons intégrer des fonctions générales de X dans $\mathbf{R}$, nous pourrons définir ce qu'on appelle la *mesure d'une partie* de X. Pensons d'abord au cas X = $\mathbf{R}$: la longueur l(I) d'un intervalle (ouvert, disons, mais peu importe) I=]a,b[de $\mathbf{R}$ est aussi l'intégrale ordinaire de la fonction caractéristique **(1.3.2)** χ_A de cet intervalle, i.e. l'intégrale de χ_I pour la mesure de Lebesgue sur $\mathbf{R}$:

$$l(\mathrm{I}) = b - a = \int_a^b 1\,dx = \int_{\mathbf{R}} \chi_{\mathrm{I}}(x)\,dx = \int_{\mathbf{R}} \chi_{\mathrm{I}}\,d\lambda = \lambda(\chi_{\mathrm{I}}). \tag{5}$$

De même, nous aurons bientôt prolongé une mesure positive quelconque μ à une classe de fonctions suffisamment générale pour contenir les fonctions caractéristiques de beaucoup de parties A de X (même des parties positivement horribles). Guidés par (5), nous poserons alors $\mu(\mathrm{A}) = \mu(\chi_A)$. Pour l'instant, nous ne savons définir qu'une intégrale supérieure et une intégrale inférieure pour de telles fonctions. Nous définissons donc, pour une partie quelconque A de X, une "*mesure extérieure*":

$$\mu^*(\mathrm{A}) = \mu^*(\chi_A), \tag{6}$$

et une "*mesure intérieure*":

$$\mu_*(\mathrm{A}) = \mu_*(\chi_A). \tag{7}$$

Bien entendu, on a $\mu_*(\mathrm{A}) \leq \mu^*(\mathrm{A})$.

En écrivant le 4ème signe = dans (5), nous avons implicitement admis que le prolongement de la mesure de Lebesgue à des fonctions non continues comme χ_I coïncidera encore avec l'intégrale de Riemann dans les cas où celle-ci a un sens. Ce que nous pouvons déjà vérifier, c'est que la mesure extérieure de I=]a, b[, $\lambda^*(\chi_I)$, vaut bien $b - a$. La fonction caractéristique de l'ouvert I, χ_I, est s.c.i. **(1.3.2** (ii)**)**, donc nous pouvons utiliser la définition (1) de **(2.2.1)**. Si $g\in \mathsf{K}$ et $g \leq \chi_I$, on a évidemment $\lambda(g) \leq b - a$ donc $\lambda^*(\chi_I) \leq b - a$. Inversement, si $\varepsilon > 0$ est tel que $a' = a + \varepsilon$ et $b' = b - \varepsilon$ soient dans I, il existe $g\in \mathsf{K}$ telle que $0 \leq g \leq 1$, $g(x) = 1$ si $x\in [a', b']$ et Supp(g) $\subset$ I **(1.3.1** (iii)**)**. On a alors $\lambda(g) \geq b' - a' = b - a - 2\varepsilon$, donc $\lambda^*(\chi_I) \geq b - a - 2\varepsilon$ et, comme ε peut être pris arbitrairement petit, on a bien $\lambda^*(\chi_I) = b - a$.

(2.2.3) *Fonctions et ensembles négligeables*

Nous allons introduire une notion très importante : celle de propriété vraie "presque partout", et notamment les fonctions définies presque partout. "Presque partout (pour la mesure positive μ)" veut dire : "sur le complémentaire d'un ensemble négligeable (pour μ)".

Définition. On dit qu'une fonction f de X dans $\overline{\mathbf{R}}$ est *négligeable* (pour μ) si $\mu^*(|f|) = 0$. On dit qu'une partie A de X est *négligeable* (pour μ) si χ_A est négligeable, i.e. si $\mu^*(A) = 0$.
Toute partie A d'un ensemble μ-négligeable N est μ-négligeable car $0 \leq \chi_A \leq \chi_N$.

(i) *Proposition.* Si (f_n) est une suite de fonctions négligeables et ≥ 0, $\sum_{n=1}^{\infty} f_n$ est négligeable.

C'est une conséquence immédiate de **(2.2.2** (iii)**)**.

(ii) *Proposition*. Une réunion dénombrable d'ensembles négligeables est négligeable.

Soit (N_k) une suite de parties μ-négligeables de X et soit N la réunion de cette suite. Il est facile de vérifier que $\chi_N \leq \sum_{k=1}^{\infty} \chi_{N_k}$; le résultat découle donc de (i).

Par exemple, pour tout point $x \in \mathbf{R}$, l'ensemble $\{x\}$ est λ-négligeable car, $\forall \varepsilon > 0$, cet ensemble est inclus dans l'intervalle $I_\varepsilon =]x-\varepsilon, x+\varepsilon[$, donc $\lambda^*(\{x\}) \leq \lambda^*(I_\varepsilon) = 2\varepsilon$. On en déduit par (ii) que *toute partie dénombrable de* $\mathbf{R}$ *est négligeable pour la mesure de Lebesgue sur* $\mathbf{R}$ *(i.e. a une longueur nulle*). Corollaire : *aucun intervalle* I *non réduit à un point de* $\mathbf{R}$ *n'est dénombrable* (puisque $\lambda^*(I) \geq \lambda^*(]a, b[) = b - a \neq 0$ si I, de bornes a et b, n'est pas réduit à un point). En particulier, $\mathbf{R}$ *n'est pas dénombrable*. De plus, le cas de l'ensemble $\mathbf{Q}$ fournit l'exemple d'un ensemble dense dans $\mathbf{R}$ et de mesure nulle.

(iii) ***Proposition.*** *Pour qu'une fonction f de* X *dans* $\overline{\mathbf{R}}$ *soit négligeable, il faut et il suffit que* $f = 0$ *presque partout* (i.e. qu'il existe un ensemble N tel que $\mu(N) = 0$ et que $f(x) = 0$ si $x \notin N$).

On peut supposer que $f \geq 0$. Supposons d'abord que $f = 0$ presque partout. L'ensemble P $= \{x \in X ; f(x) > 0\}$ est négligeable, comme partie de l'ensemble N de l'hypothèse. Si $f(x) \neq 0$, alors $x \in P$, donc on a $f \leq \sup_n n\, \chi_P$. Comme on peut utiliser **(2.2.2** (ii)**)** (la suite $(n\chi_P)$ étant croissante), il en résulte $0 \leq \mu^*(f) \leq \sup_n \mu^*(n\, \chi_P) = \sup_n n\, \mu^*(\chi_P) = 0$, donc $\mu^*(f) = 0$. Réciproque : si $\mu^*(f) = 0$, considérons les ensembles $A_n = \{x \in X ;\ f(x) \geq 1/n\}$ (pour $n \geq 1$). Pour tout n, la définition de A_n entraîne que $f \geq (1/n)\, \chi_{A_n}$, d'où $0 = \mu^*(f) \geq (1/n)\, \mu^*(A_n) \geq 0$, ainsi tous les A_n sont négligeables. Donc P, qui est leur réunion, est négligeable, cqfd.

(iv) *Proposition.* Si la fonction f de X dans $\overline{\mathbf{R}}$ est telle que $\mu^*(f) < +\infty$, alors on a presque partout $f(x) < +\infty$, autrement dit l'ensemble $A_f = \{x \in X ; f(x) = +\infty\}$ est négligeable (pour μ).

Comme $\mu^*(f) = \inf \{\mu^*(g); g\in \mathsf{I}, g \geq f\}$, l'hypothèse entraîne qu'il existe une fonction $g\in \mathsf{I}$ telle que $g \geq f$ et $\mu^*(g) < +\infty$. Si la proposition est vraie pour les fonctions de I, on aura alors $\mu^*(A_g) = 0$, donc $\mu^*(A_f) = 0$ puisque $A_f \subset A_g$. Nous pouvons donc supposer que $f\in \mathsf{I}$. Montrons que l'hypothèse entraîne alors que $\mu^*(f^+) < +\infty$ [voir **(2.1.3**(i)**)** pour les définitions de f^+ et f^-]. Puisque $f\in \mathsf{I}$, il existe $g\in \mathsf{K}$ telle que $f \geq g$, d'où $f^- \leq g^-$ et ainsi $\mu^*(f^-) \leq \mu^*(g^-) \leq \mu^*(|g|)= \mu(|g|) < +\infty$ (car $|g|\in \mathsf{K}$). Comme $f^+ = f + f^-$, on a donc bien $\mu^*(f^+) \leq \mu^*(f) + \mu^*(f^-) < +\infty$. Ceci nous permet de supposer en plus que $f \geq 0$. Posant $A= A_f$, on a alors, pour tout $n \geq 0$, $f \geq n\,\chi_A$, d'où $\mu^*(f) \geq n\,\mu^*(A)$ et donc, puisque $\mu^*(f) < +\infty$, $\mu^*(A) = 0$, cqfd.

(v) ***Théorème***. *Si f et g :X $\rightarrow \overline{\mathbf{R}}$ sont égales presque partout, alors $\mu^*(f) = \mu^*(g)$.*

Montrons d'abord qu'on peut supposer $f \leq g$: nous admettons donc provisoirement que le théorème soit vrai lorsque $f \leq g$, et nous traitons le cas général. Posons $h = \sup(f, g)$. L'hypothèse signifie que l'ensemble N = $\{x\in X ; f(x) \neq g(x)\}$ est μ-négligeable, et l'on a $f = g = h$ sur X \ N. Nous pouvons donc appliquer le théorème à f et h, qui sont égales presque partout et vérifient $f \leq h$, et nous obtenons ainsi $\mu^*(h) = \mu^*(f)$. Mais nous pouvons aussi bien appliquer le théorème à g et h, ce qui donne $\mu^*(h) = \mu^*(g)$. Donc $\mu^*(f) = \mu^*(g)$, ainsi nous pouvons effectivement supposer $f \leq g$. Soit alors k la fonction égale à $+\infty$ sur N et à 0 sur X \ N, et soit u une fonction de I telle que que $u \geq f$. La somme $u + k$ est donc définie partout dans X (car u, étant dans I, ne peut pas prendre la valeur $-\infty$) et l'on a $g \leq u + k$ (car l'inégalité est vraie dans N aussi bien que dans X \ N). Puisque N est négligeable, on a $\mu^*(k) = 0$ d'après (iii), donc on peut appliquer **(2.2.2** (i)), ce qui donne $\mu^*(g) \leq \mu^*(u) + \mu^*(k) = \mu^*(u)$. De par la définition de $\mu^*(f)$ [Eq. (1) de **(2.2.2)**], cela entraîne $\mu^*(g) \leq \mu^*(f)$. Mais comme $f \leq g$, on a aussi $\mu^*(f) \leq \mu^*(g)$. La démonstration est terminée.

Le théorème (v) a une importance considérable, car il montre qu'on peut changer arbitrairement les valeurs d'une fonction sur un ensemble négligeable sans changer son intégrale supérieure (ni, évidemment, son intégrale inférieure). Certains ensembles négligeables, comme $\mathbf{Q}$ pour la mesure de Lebesgue, ne sont pas si petits que ça! Notons aussi que f et g sont ici des fonctions absolument quelconques de X dans $\overline{\mathbf{R}}$. Nous commençons à entrevoir la puissance de la théorie. En définitive, seules vont intervenir les *classes d'équivalence* $\widetilde{f}$ des fonctions de X dans $\overline{\mathbf{R}}$ pour la relation "*$f = g$ presque partout*". Ceci permet de considérer, sans aucun dommage pour ce qui concerne l'intégration, des fonctions h qui ne sont définies que sur le complémentaire d'un ensemble négligeable : on pourra en effet prolonger h de façon *arbitraire* en une fonction partout définie f , à cause de (v). On dira de telles fonctions h qu'elles sont *définies presque partout*.

Exercice.
1) Vérifier que la relation "$f = g$ presque partout" est bien une relation d'équivalence.
2) Montrer qu'on peut définir la somme et le produit de deux classes d'équivalence $\widetilde{f}$ et $\widetilde{g}$.
3) Montrer qu'on peut définir une relation d'ordre entre classes d'équivalence $\widetilde{f}$.

(2.2.4) *Fonctions intégrables*

Définition. On dit qu'une *fonction f* de X dans $\overline{\mathbf{R}}$ est *intégrable* (pour μ) si l'on a $\mu^*(f) = \mu_*(f)$ et que ce nombre est *fini*.

On pose alors $\mu(f) = \mu^*(f) = \mu_*(f)$, et on appelle ce nombre *l'intégrale de f par rapport à μ*; on le note aussi $\mu(f) == \int_X f(x)\, d\mu(x) = \int_X f\, d\mu$. Ceci est cohérent car, lorsque $f\in \mathsf{K}$, l'intégrale de f par rapport à μ coïncide bien sûr avec la valeur prise en f par la forme linéaire μ.

(i) *Proposition*. [f intégrable] $\Leftrightarrow$ [$\forall\varepsilon>0\ \exists g=g_\varepsilon\in \mathsf{S}\ \ \exists h=h_\varepsilon\in \mathsf{I} : g\leq f\leq h$ et $\mu^*(h)\leq\varepsilon+\mu_*(g)$].

($\Leftarrow$) : La condition entraîne d'abord que $\mu_*(g)$ et $\mu^*(h)$ sont finis (puisqu'on a toujours $\mu_*(g) < +\infty$ et $\mu^*(h) > -\infty$) puis, d'après les définitions (1) et (4) de **(2.2.2)**, que $\mu^*(f)$ et $\mu_*(f)$ sont finis. On a aussi : $\forall\varepsilon>0$ $\mu^*(f)\leq\mu^*(h_\varepsilon)\leq\varepsilon+\mu_*(g_\varepsilon)\leq\varepsilon+\mu_*(f)$, d'où $\mu^*(f)=\mu_*(f)$.

($\Rightarrow$) : se vérifie immédiatement à partir des définitions.

(ii) *Proposition*. Si f est intégrable, il existe une suite décroissante (h_n) de fonctions de I et une suite croissante (g_n) de fonctions de S telles que $g_n\leq f\leq h_n$ et vérifiant $\lim\limits_{n\to\infty}\mu^*(g_n)=\lim\limits_{n\to\infty}\mu^*(h_n)=\mu(f)$.

Pour tout n, il existe par définition une fonction k_n de I telle que $f\leq k_n$ et $\mu^*(f)\leq\mu^*(k_n)\leq\mu^*(f)+1/n$, et il suffit de prendre $h_n=\inf(k_1,\ldots,k_n)$. On procède de même pour la suite (g_n).

(iii) *Proposition*. [f intégrable] $\Leftrightarrow$ [$\forall\varepsilon>0\ \exists u\in\mathsf{K}\ \ \mu^*(|f-u|)\leq\varepsilon$].

($\Leftarrow$) : On a $f\leq|f-u|+u$ et $u\leq|f-u|+f$ et il n'y a pas d'indétermination (vérifier), donc on applique **(2.2.2** (i)**)** : $\mu^*(f)\leq\mu^*(|f-u|)+\mu^*(u)$ et $\mu^*(u)\leq\mu^*(|f-u|)+\mu^*(f)$, c'est à dire $|\mu^*(f)-\mu^*(u)|\leq\mu^*(|f-u|)\leq\varepsilon$. De même $|\mu_*(f)-\mu_*(u)|=|-\mu^*(-f)+\mu^*(-u)|\leq\mu^*(|f-u|)\leq\varepsilon$. Ceci montre d'abord que $\mu^*(f)$ et $\mu_*(f)$ sont finis, puisque $\mu(u)=\mu^*(u)=\mu_*(u)$ est fini. A nouveau parce que $\mu^*(u)=\mu_*(u)$, cela montre aussi que $|\mu^*(f)-\mu_*(f)|\leq 2\varepsilon$, ainsi f est intégrable.

($\Rightarrow$) : Puisque f est intégrable, soient $g\in\mathsf{S}$ et $h\in\mathsf{I}$ telles que $g\leq f\leq h$ et $\mu^*(h)-\mu_*(g)=\mu^*(h-g)\leq\varepsilon$ (i). Comme $h\in\mathsf{I}$, il existe $k\in\mathsf{K}$ telle que $k\leq h$ et $\mu(k)\geq\mu^*(h)-\varepsilon$, soit encore $\mu^*(h-k)\leq\varepsilon$ .On a alors $\mu^*(|f-k|)\leq\mu^*(|f-g|)+\mu^*(|h-g|)+\mu^*(|h-k|)$
$\leq\mu^*(h-g)+\mu^*(h-g)+\mu^*(h-k)\leq 3\varepsilon$, d'où la conclusion.

Grâce à **(2.2.3** (v)**)**, on peut, dans la définition d'une fonction intégrable f et dans le résultat qui précède, supposer que f n'est définie que *presque partout*, en la prolongeant arbitrairement sur l'ensemble μ-négligeable N en-dehors duquel elle est définie (par exemple, $f(x) := 0$ si $x\in$ N).

(iv) ***Théorème***. *L'ensemble* L^1(X, μ) *des fonctions μ-intégrables à valeurs dans* $\overline{\mathbf{R}}$ *définies presque partout dans* X *est un espace vectoriel sur* $\mathbf{R}$, *et l'application* $f\mapsto\int_X f\,d\mu=\mu(f)$ *est une **forme linéaire** sur* L^1(X, μ) *qui vérifie la condition de positivité :* $\mu(f)\geq 0$ si $f\geq 0$.

Quand il n'y a pas de risque de confusion, on note aussi $\mathsf{L}^1(\mu)$ ou même L^1 au lieu de L^1(X, μ). Si l'on veut marquer qu'il s'agit des fonctions réelles, on note aussi $\mathsf{L}^1_{\mathbf{R}}$(X, μ), ou $\mathsf{L}^1_{\mathbf{R}}(\mu)$, ou $\mathsf{L}^1_{\mathbf{R}}$.

Démonstration : comme on a $\mu^*(\alpha f)=\alpha\mu^*(f)$ et $\mu_*(\alpha f)=\alpha\mu_*(f)$, le produit d'une fonction intégrable par un scalaire est une fonction intégrable et l'on a $\mu(\alpha f)=\alpha\mu(f)$; tout comme μ^* et μ_* , μ vérifie la condition de positivité. Le reste est plus technique.

Le fait que la somme $f + g$ soit intégrable lorsque f et g le sont, résulte facilement de la caractérisation (iii) et de ce que la somme de deux fonctions de K est une fonction de K. Il ne reste à prouver que l'additivité de μ. Grâce à la proposition (iv) de **(2.2.3)**, on sait que les fonctions intégrables f et g sont finies presque partout, et le théorème **(2.2.3** (v)**)** permet donc de remplacer $f(x)$ ou $g(x)$ par 0 (par exemple) aux points x de X où $f(x)$ ou $g(x)$ est infini. Ainsi, la somme $f + g$ est définie presque partout et on peut la considérer comme définie partout. On peut alors appliquer **(2.2.2** (i)**)** et ainsi écrire : $\mu(f + g) = \mu^*(f + g) \leq \mu^*(f) + \mu^*(g) = \mu(f) + \mu(g)$. Le résultat dual de **(2.2.2** (i)**)** donne de même $\mu(f + g) = \mu_*(f + g) \geq \mu_*(f) + \mu_*(g) = \mu(f) + \mu(g)$ (vérifier l'inégalité). La démonstration est achevée.

(v) *Proposition*. Si f et g sont intégrables, alors $|f|$, f^+, f^-, $\sup(f, g)$ et $\inf(f, g)$ sont intégrables.

Si f est intégrable, l'intégrabilité de $|f|$ résulte de ce que $||f| - |u|| \leq |f - u|$ et du critère (iii). Puis :
On a $f^+ = (f + |f|)/2$ et $f^- = (-f + |f|)/2$, donc f^+ et f^- sont intégrables à cause de (iv). De même, $\sup(f, g) = (f + g + |f - g|)/2$ et $\inf(f, g) = (f + g - |f - g|)/2$ (vérifier) sont intégrables.

(vi) *Exercice*. Montrer que $|\int_X f\, d\mu| \leq \int_X |f|\, d\mu$.

(vii) *Proposition*. Soit h (resp. g) une fonction de I (resp. de S). Pour que h (resp. g) soit intégrable, il faut et il suffit que $\mu^*(h) < +\infty$ (resp. $\mu_*(h) > -\infty$).

Il est évident que la condition est nécessaire. Inversement, supposons que $h \in$ I et $\mu^*(h) < +\infty$. Pour tout $\varepsilon > 0$, il existe une fonction $k \in$ K telle que $k \leq h$ et $\mu(k) \geq \mu^*(h) - \varepsilon$, ou
$\mu^*(h) - \mu_*(k) = \mu^*(h - k) = \mu^*(|h - k|) \leq \varepsilon$, donc h est intégrable par (iii). De même pour g.

Nous avons ainsi mené à son terme le processus de prolongement de la mesure μ. Le reste de la théorie sera moins aride (les résultats vont pleuvoir).

(2.2.5) *Ensembles intégrables*

Définition. On dit qu'une *partie* A de X est *intégrable* (pour μ) si $\mu^*(A) = \mu_*(A)$, ce nombre étant *fini*, donc si la fonction χ_A est intégrable. On pose alors $\mu(A) = \mu^*(A) = \mu_*(A)$ et on appelle ce nombre réel la *mesure* de l'ensemble intégrable $A \subset X$.

Puisque $\chi_A \equiv 0$, l'ensemble vide est un ensemble intégrable de mesure nulle : $\mu(\emptyset) = 0$. De façon générale, les ensembles intégrables de mesure nulle sont exactement les ensembles négligeables; en effet, $\mu^*(A) = 0$ entraîne $\mu_*(A) = 0$ puisque $0 \leq \mu_*(A) \leq \mu^*(A)$.

(i) *Proposition*. Si A et B sont des ensembles intégrables, la réunion $A \cup B$, l'intersection $A \cap B$ et la différence $A \backslash B$ sont intégrables et l'on a $\mu(A \cup B) \leq \mu(A) + \mu(B)$, avec égalité si et seulement si $A \cap B$ est négligeable.

Il est facile de vérifier que $\chi_{A \cap B} = \inf(\chi_A, \chi_B)$, $\chi_{A \cup B} = \sup(\chi_A, \chi_B)$ et $\chi_{A \backslash B} = \chi_A - \chi_{A \cap B}$, ce qui montre que $A \cup B$, $A \cap B$ puis $A \backslash B$ sont intégrables **(2.2.4** (v)**)**, et l'inégalité provient de ce que $\chi_{A \cup B} \leq \chi_A + \chi_B$. Le cas d'égalité résulte de ce que $\chi_{A \cup B} + \chi_{A \cap B} = \chi_A + \chi_B$ et de **(2.2.4** (iv)**)**.

(ii) *Proposition*. Tout ensemble compact est intégrable. Pour qu'un ensemble ouvert A soit intégrable, il faut et il suffit que $\mu^*(A) < +\infty$.

Si A est fermé, alors χ_A est s.c.s. car $\chi_{E\setminus A} = 1 - \chi_A$ est s.c.i. **(1.3.2** (ii)**)**. Si A est compact, alors A est fermé et il existe de plus une fonction ϕ à support compact telle que $\phi \equiv 1$ sur A **(1.3.1** (iii)**)**, donc $\chi_A \leq \phi$ et ainsi $\chi_A \in \mathsf{S}$. Or $\mu^*(\chi_A) \geq 0 > -\infty$, donc χ_A est intégrable **(2.2.4** (vii)**)**, ainsi A est intégrable. Si A est ouvert, χ_A est s.c.i. et $\chi_A \geq 0$, donc $\chi_A \in \mathsf{I}$ et le résultat découle encore de **(2.2.4** (vii)**)**.

En particulier, ce résultat achève de montrer que la mesure de Lebesgue d'un intervalle borné quelconque (a, b) de $\mathbf{R}$ est bien $b - a$ (cf. **(2.2.2)** et **(2.2.3)**).

(iii) *Proposition.* Pour qu'une partie A soit intégrable, il faut et il suffit que, pour tout $\varepsilon > 0$, il existe un compact $K=K_\varepsilon$ et un ouvert $U=U_\varepsilon$ tels que $K \subset A \subset U$ et $\mu^*(U \setminus K) \leq \varepsilon$.

Comme on a $\mu^*(U \setminus K) = \mu^*(U) - \mu_*(K)$ et que K est intégrable (ii), la condition entraîne que $\mu^*(U) < \infty$ et donc que U est intégrable (ii). La condition entraîne alors :
$\forall \varepsilon > 0,\ \mu_*(A) \leq \mu^*(A) \leq \mu^*(U_\varepsilon) \leq \mu_*(K_\varepsilon) + \varepsilon \leq \mu_*(A) + \varepsilon$, d'où $\mu_*(A) = \mu^*(A) < \infty$: elle est donc suffisante. La nécessité est un peu laborieuse, nous l'admettrons.

2.3 Les résultats fondamentaux de la théorie

Nous continuons, sauf mention expresse du contraire, à parler d'une mesure *positive* μ sur X.

(2.3.1) *Théorèmes de convergence*

(i) ***Théorème de convergence monotone de Lebesgue***.

Soit (f_n) *une suite croissante de fonctions intégrables. Pour que sa limite ponctuelle* $f = \sup_n f_n$ *soit intégrable, il faut et il suffit que* $\sup_n \int_X f_n \, d\mu < +\infty$ *et l'on a alors*

$$\int_X f \, d\mu = \sup_n \int_X f_n \, d\mu = \lim_{n\to\infty} \int_X f_n \, d\mu . \tag{1}$$

Puisque nous avons ici une suite croissante de fonctions *intégrables*, nous pouvons appliquer **(2.2.2** (ii)**)**, d'où $\mu^*(f) = \lim \mu^*(f_n) = \lim \mu(f_n) = \sup \mu(f_n)$, qui est un réel ou $+\infty$. Si f est intégrable, il est donc clair que le sup est fini et qu'on a (1). Inversement, si $\sup \mu(f_n) < +\infty$, comme $\mu^*(f) \geq \mu_*(f) \geq \mu_*(f_n) = \mu^*(f_n)$ et comme $\mu^*(f_n) \to \mu^*(f)$, nous aurons $\mu_*(f) = \mu^*(f) \in \mathbf{R}$, donc f est intégrable : la démonstration est terminée. Il faut noter qu'on n'exige aucune condition de régularité (sauf l'intégrabilité qui, comme on le verra, n'impose qu'une régularité extrêmement faible), ni aucune uniformité de la convergence.

(ii) *Proposition*. Soit (f_n) une suite quelconque de fonctions intégrables. Pour que $f = \sup_n f_n$ soit intégrable, il faut et il suffit qu'il existe une une fonction g de X dans $[0, +\infty]$ telle que $\mu^*(g) < +\infty$ et que pour tout n on ait presque partout $f_n \leq g$.

Si f est intégrable, alors $|f|$ est intégrable **(2.2.4** (v)**)** et l'on a partout $f_n \leq f \leq |f|$, donc la condition est nécessaire. Inversement, si g est comme dans l'énoncé, soit $g_n = \sup(f_1,..., f_n)$. On a $g_n \leq g$ et la suite (g_n) est croissante, donc $\sup\limits_{n\geq 1} g_n$ est intégrable par (i), mais c'est aussi $\sup\limits_{n\geq 1} f_n$.

(iii) *Lemme*. Soit (f_n) une suite quelconque de fonctions intégrables, et supposons qu'il existe une fonction $h \geq 0$ vérifiant $\mu^*(h) < +\infty$ et telle que, pour tout n, on ait presque partout $-h \leq f_n \leq h$. Alors $\underline{\lim}\ f_n$ et $\overline{\lim}\, f_n$ sont intégrables et l'on a

$$\int \underline{\lim}\ f_n\, d\mu \leq \underline{\lim} \int f_n\, d\mu \leq \overline{\lim} \int f_n\, d\mu \leq \int \overline{\lim}\, f_n\, d\mu \qquad (2).$$

Posons $g_n = \inf\limits_{m\geq n} f_m$ et $h_n = \sup\limits_{m\geq n} f_m$. Chaque fonction h_n est intégrable grâce à (ii), puisque les f_m sont intégrables et vérifient $f_m \leq h$ presque partout. De même, chaque fonction g_n est intégrable. La suite (g_n) tend en croissant vers $\underline{\lim}\ f_n$ et $\mu(g_n) \leq \inf\limits_{m\geq n} \mu(f_m) \leq \mu^*(h) < +\infty$. Le théorème de convergence monotone entraîne donc que $\underline{\lim}\ f_n$ est intégrable et que
$\int \underline{\lim}\ f_n\, d\mu = \lim\limits_{n\to\infty} \mu(g_n) \leq \lim\limits_{n\to\infty} \inf\limits_{m\geq n} \mu(f_m) = \underline{\lim} \int f_n\, d\mu$. De même, $\overline{\lim}\, f_n$ est intégrable et l'on a l'inégalité de droite dans (2). Enfin l'inégalité $\underline{\lim} \int f_n\, d\mu \leq \overline{\lim} \int f_n\, d\mu$ est évidente.

(iv) ***Théorème de convergence dominée de Lebesgue***

Soit (f_n) une suite de fonctions intégrables, tendant presque partout vers une limite f, et supposons qu'il existe une fonction $h \geq 0$ vérifiant $\mu^*(h) < +\infty$ et telle que, pour tout n, on ait presque partout $|f_n| \leq h$. Alors f est intégrable et l'on a

$$\int_X f\, d\mu = \lim_{n\to\infty} \int_X f_n\, d\mu \quad \text{et} \quad \lim_{n\to\infty} \int_X |f_n - f|\, d\mu = 0. \qquad (3)$$

Pour obtenir la première égalité (3), nous appliquons le lemme, qui nous donne l'intégrabilité de $f = \underline{\lim}\, f_n = \overline{\lim}\, f_n$ et le résultat, puisque les membres extrêmes de (2) sont ici égaux.
La deuxième égalité (3) se déduit de la première, appliquée à la suite $g_n = |f_n - f|$; en effet, g_n est intégrable, on a presque partout $|g_n| \leq h + |f|$ (or $|f|$ est intégrable puisque f l'est, donc la somme $h+|f|$ est intégrable), enfin la suite (g_n) tend presque partout vers 0.

Ici encore, on n'impose aucune uniformité de la convergence, ni aucune régularité en dehors de l'intégrabilité, et on obtient une convergence en moyenne, souvent très utile. Ce théorème est sans doute le résultat le plus utile de la théorie de Lebesgue. Voyons quelques conséquences faciles, déjà assez impressionnantes :

(v) ***Proposition***. *Soit (f_n) une suite de fonctions intégrables. Si la série $\Sigma_n \int_X |f_n|\, d\mu$ est convergente, alors la série $\Sigma_n f_n(x)$ est presque partout absolument convergente, sa somme $f(x) = \Sigma_n f_n(x)$ (définie presque partout) est une fonction intégrable et l'on a*

$$\int_X \sum_{n=1}^{\infty} f_n(x)\, d\mu(x) = \sum_{n=1}^{\infty} \int_X f_n(x)\, d\mu(x).$$

Démonstration : *Exercice*.

(vi) ***Proposition***. *Soient* I *un intervalle ouvert de* $\mathbf{R}$ *et* $f : (t, x) \to f(t, x)$ *une application de* I×X *dans* $\overline{\mathbf{R}}$ *telle que :*

(a) pour tout $t \in$ I, *la fonction* $x \to f(t, x)$ *soit intégrable;*

(b) pour presque tout $x \in$ X, *la fonction* $t \to f(t, x)$ *soit (finie et) dérivable dans* I *;*

(c) il existe une fonction intégrable positive h telle que

$$\forall\ t \in \mathrm{I},\ \text{on ait } |\partial f(t, x)/\partial t| \leq h(x) \text{ pour presque tout } x \in \mathrm{X}\,.$$

Alors l'intégrale dépendant du paramètre t, $F(t) = \int_{\mathrm{X}} f(t, x)\, d\mu(x)$, *est dérivable dans* I *et :*

$$F'(t) = \int_{\mathrm{X}} (\partial f(t, x)/\partial t)\, d\mu(x).$$

Démonstration : *Exercice*.

(vii) ***Théorème***. *Soit* (A_n) *une suite d'ensembles intégrables.*

(a) L'intersection $A = \bigcap_n A_n$ *est un ensemble intégrable, et si la suite* (A_n) *est décroissante on a* $\mu(A) = \inf \mu(A_n)$.

(b) Si tous les A_n *sont contenus dans un même ensemble intégrable C, ou si la somme* $\sum_{n=1}^{\infty} \mu(A_n)$ *est finie, alors la réunion* $B = \bigcup_n A_n$ *est un ensemble intégrable et dans le deuxième cas on a:*

$$\mu(B) \leq \sum_{n=1}^{\infty} \mu(A_n),$$

avec l'égalité si les A_n *sont deux à deux disjoints.*

Pour (a), il suffit d'appliquer le théorème de convergence monotone à la suite des fonctions $f_n = -\chi_{A_1 \cap \ldots \cap A_n}$, qui tend simplement en croissant vers $-\chi_A$ (vérifier); lorsque la suite (A_n) est décroissante, on a $\mu(\chi_{A_1 \cap \ldots \cap A_n}) = \mu(\chi_{A_n}) = \inf(\mu(A_1), \ldots, \mu(A_n))$. Pour la première partie de (b), on applique le théorème de convergence dominée à $f_n = \chi_{A_1 \cup \ldots \cup A_n}$, compte-tenu de la majoration $0 \leq f_n \leq \chi_C$. Pour la deuxième partie de (b), on applique le théorème de convergence monotone à $f_n = \chi_{A_1 \cup \ldots \cup A_n}$, compte-tenu de ce que $\int_X f_N\, d\mu = \mu(A_1 \cup \ldots \cup A_N) \leq \sum_{n=1}^{N} \mu(A_n)$ **(2.2.5** (i)**)**, avec l'égalité si les A_n sont deux à deux disjoints.

(2.3.2) *Fonctions et ensembles mesurables. Définition "ensembliste" de la mesure*

Il est parfois utile de considérer des ensembles A (resp. des fonctions f) "mesurables", qui sont un peu plus généraux (resp. générales) que les ensembles et les fonctions intégrables, non en ce
qu'ils ou elles sont moins réguliers (régulières) mais en ce que la valeur $\mu(A)$ ou $\mu(f) = \int_X f\, d\mu$ n'est *pas nécessairement définie* (en mathématiques, "mesurable" ne veut donc pas dire "qui a

une mesure", ce qui est un peu regrettable). Plus exactement, un *ensemble* mesurable ou une fonction mesurable *positive* auront bien toujours une mesure, mais celle-ci pourra être égale à $+\infty$. Mais une fonction mesurable quelconque n'aura pas en général une intégrale (une mesure), de même qu'une série à termes quelconques n'a pas en général une somme, alors qu'une série à termes positifs a toujours une somme (éventuellement égale à $+\infty$).

L'intérêt principal qu'il y a à considérer des ensembles mesurables vient de la proposition (i) ci-dessous, qui entraîne que tous les ensembles que l'on peut être amené à considérer dans la pratique (des mathématiques actuelles, et a fortiori de toute autre discipline...) sont mesurables. De même, toutes les fonctions que l'on rencontre en pratique

sont mesurables. La condition d'intégrabilité se réduit alors au critère (vi) ci-dessous (les remarques de ce paragraphe ne sont pas rigoureuses, donc pas utilisables dans une preuve).

(a) *Définition*. On dit qu'une *partie* A de X est *mesurable* [pour la mesure positive μ] si, pour tout compact K, l'ensemble $K \cap A$ est intégrable. (On dit aussi que A est μ-mesurable).

En raison de **(2.2.5** (i)**)**, tout ensemble intégrable est mesurable; en raison de **(2.2.5** (ii)**)**, tout ensemble compact est donc mesurable, et l'ensemble X tout entier est mesurable.

(i) ***Proposition***.

1) *Le complémentaire d'un ensemble mesurable est mesurable;*
2) *Toute réunion dénombrable d'ensembles mesurables est mesurable;*
3) *Toute intersection dénombrable d'ensembles mesurables est mesurable;*
4) *Tout ensemble ouvert est mesurable, ainsi que tout ensemble fermé.*

Si A est mesurable, alors pour tout compact K, l'ensemble $(X \backslash A) \cap K = K \backslash (K \cap A)$ est intégrable **(2.2.5** (i)**)**, donc $X \backslash A$ est mesurable. Soit (A_n) une suite d'ensembles mesurables et A la réunion (resp. l'intersection) de cette suite. Pour tout compact K, l'ensemble $K \cap A$ est la réunion (resp. l'intersection) des ensembles $B_n = K \cap A_n$, et est donc intégrable par le théorème **(2.3.1** (vii)**)**, donc A est mesurable. Si F est un fermé et si K est un compact, l'intersection compacte $K \cap F$ est un ensemble intégrable à cause de **(2.2.5** (ii)**)**, donc F est mesurable. Il en résulte que les ouverts sont mesurables, à cause de 1).

Un ensemble non vide A de parties d'un ensemble X *quelconque* possédant les propriétés 1) à 3) (où l'on remplace "mesurable" par "appartenant à A ", ou "est mesurable" par "appartient à A ") est appelé une *tribu sur* X. Revenons au cas où X est un "*espace de base*" pour l'intégration, donc un espace métrique localement compact et séparable. A cause du fait que les ouverts sont mesurables (pour la mesure positive quelconque μ) et que les ensembles μ-mesurables forment une tribu, il existe un très grand nombre de parties de X qui sont μ-mesurables pour *toute* mesure positive μ sur X : ce sont les parties obtenues à partir des ouverts par l'application (qu'on a le droit d'itérer autant qu'on le souhaite) des opérations 1) à 3). Ces parties sont appelées les *boréliens* de X et leur ensemble forme une tribu B, qui est donc toujours incluse dans la tribu A_μ des ensembles μ-mesurables pour une certaine mesure μ.

Si A est mesurable et non intégrable, alors on peut dire que A *est de mesure infinie*. En effet, X est réunion d'une suite croissante (K_n) de compacts, donc la suite des $a_n = \mu(K_n \cap A) = \int_X \chi_{K_n \cap A}\, d\mu$ est croissante. Si elle était bornée, la fonction χ_A serait intégrable par le théorème de convergence monotone, ainsi A serait intégrable : donc $\mu^*(A)$, qui est égal à sup $\mu(K_n \cap A)$ **(2.2.2** (ii)**)**, est égal à $+\infty$. On a de même, pour tout n, $\mu_*(A) \geq \mu_*(K_n \cap A) = \mu(K_n \cap A)$, donc $\mu_*(A) = +\infty = \mu^*(A)$. Posons donc $\mu(A) = +\infty$ pour tout ensemble mesurable et non intégrable. Il est facile de vérifier, grâce à ce qui précède et au théorème (vii) de la section précédente, qu'on a maintenant pour toute suite (A_n) d'ensembles mesurables :

$$\mu(\bigcup_{n\geq 1} A_n) \leq \sum_{n=1}^{\infty} \mu(A_n), \text{ avec l'égalité si les } A_n \text{ sont deux à deux disjoints.} \qquad (1)$$

(b) ***Définition ensembliste de la mesure***. Soient X un ensemble quelconque et A une tribu sur X. On appelle *mesure positive* sur X (muni de la tribu A), toute application de A dans $\mathbf{R}_+ \cup\{+\infty\}$ qui ne prend pas seulement la valeur $+\infty$ et vérifie la propriété (1) ci-dessus (lorsque les A_n sont dans A).

La théorie de l'intégration peut également être construite à partir de cette définition plus générale et plus abstraite, mais nous avons choisi un autre mode d'exposition.

(c) *Définition*. On dit qu'une *application f* de X dans un espace métrique Y est *mesurable* si, pour tout compact $K \subset X$ et tout $\varepsilon > 0$, il existe un compact $K' \subset K$ tel que $\mu(K \setminus K') \leq \varepsilon$ et que la restriction de f à K' soit *continue*.

Il est clair que si $f : X \to Y$ est mesurable et si $g : Y \to Z$ est continue, $g_{\mathrm{o}} f$ est mesurable.

(ii) *Proposition*. Soient f et g deux applications mesurables de X dans un espace normé Y. Alors $f + g$ et $\alpha.f$ sont mesurables. Si $Y = \overline{\mathbf{R}}$, alors sup (f, g) et inf (f, g) sont mesurables.

Il est clair que, pour tout scalaire α, $\alpha.f$ est mesurable si f l'est. Soient K un compact, $\varepsilon > 0$, K '' et K ''' des compacts inclus dans K tels que $f_{|K''}$ et $g_{|K'''}$ soient continues et que $\mu(K \setminus K'') \leq \varepsilon$ et $\mu(K \setminus K''') \leq \varepsilon$. Posons $K' = K'' \cap K'''$, qui est un compact. Les restrictions de f et g à K ' sont continues, donc il en est de même de la restriction de $f + g$ à K ' et, si $Y = \overline{\mathbf{R}}$, des restrictions à K ' de sup (f, g) et inf (f, g). On a $K \setminus K' = (K \setminus K'') \cup (K \setminus K''')$ d'où $\mu(K \setminus K') \leq 2\varepsilon$. Ainsi $f + g$ est mesurable et, si $Y = \overline{\mathbf{R}}$, sup (f, g) et inf (f, g) sont mesurables.

(iii) *Proposition*. Si f est mesurable et si $g = f$ presque partout, alors g est mesurable.

Soit N l'ensemble μ-négligeable des points x de X où $f(x) \neq g(x)$. Si K et K' sont des compacts, avec $K' \subset K$ et $\mu(K \setminus K') \leq \varepsilon$, la restriction $f_{|K'}$ étant continue, considérons l'ensemble $A = K' \cap N$. Il s'agit d'un ensemble intégrable et l'on a $\mu(A) = \mu(K')$ **(2.2.5** (i)**)**. Il existe donc, pour tout $\varepsilon > 0$, un compact $K'' \subset K'$ tel que $\mu(K' \setminus K'') \leq \varepsilon$ **(2.2.5** (iii)**)**. On a alors $\mu(K \setminus K'') \leq 2\varepsilon$, et la restriction de g à K '' est continue, d'où la conclusion puisque ε est arbitraire.

(d) *Définition*. On dit qu'une *application définie presque partout h* de X dans un espace métrique Y est *mesurable* si *toute* application f de X dans Y et égale à h presque partout est mesurable. D'après la proposition (iii), il suffit qu'il existe *une* telle application f.

(iv) *Théorème* (admis).

1) Soit (f_n) une suite d'applications mesurables de X dans un espace métrique Y, et supposons que, pour presque tout x, la suite $f_n(x)$ tende vers une limite $f(x)$. Alors f est mesurable.

2) Si f et g sont des applications mesurables de X dans $\overline{\mathbf{R}}$, et si $f.g$ (resp. f/g) est partout définie, alors $f.g$ (resp. f/g) est mesurable.

Dans l'énoncé 1), on peut supposer que les f_n ne sont définies que presque partout, puisqu' une réunion dénombrable d'ensembles négligeables est négligeable **(2.2.3** (ii)**)**.

(v) *Corollaire*. Soit (f_n) une suite d'applications mesurables de X dans $\overline{\mathbf{R}}$. Alors $\sup\limits_{n\geq 1} f_n$, $\inf\limits_{n\geq 1} f_n$, $\overline{\lim} f_n$ et $\underline{\lim} f_n$ sont mesurables.

Car $\sup\limits_{n\geq 1} f_n = \lim\limits_{n\to\infty} g_n$ avec $g_n = \sup(f_1,\ldots,f_n)$, puis $\overline{\lim} f_n = \lim\limits_{n\to\infty} \sup\limits_{m\geq n} f_m$.

(vi) ***Théorème*** (admis). *Pour qu'une application f de* X *dans* $\overline{\mathbf{R}}$ *soit intégrable, il faut et il suffit que f soit mesurable et que* $\mu^*(|f|) < +\infty$.

Remarque heuristique : en pratique, on ne rencontrera que des fonctions f mesurables et l'on n'aura donc qu'à vérifier que $\mu^*(|f|) < +\infty$, généralement en majorant f par une fonction intégrable.

(vii) *Proposition*. Si A est un ensemble mesurable, alors la fonction χ_A est mesurable.

Soient K un compact et $\varepsilon > 0$, nous cherchons un compact $K' \subset K$, tel que $\mu(K \setminus K') \leq \varepsilon$ et que la restriction de χ_A à K ' soit continue. L'ensemble $A\cap K$ est intégrable, donc il existe un ouvert U et un compact K '' tels que $K'' \subset (A\cap K) \subset U$ et $\mu(U \setminus K'') \leq \varepsilon$ (dessin recommandé). Il existe alors $f \in \mathsf{K}$ telle que $\mathrm{Supp}(f) \subset U$ et $f=1$ sur K '' , ce qui implique que f coïncide avec χ_A sur le compact $K' = K'' \cup (K \setminus U)$. Or on a par construction $(K \setminus K') \subset (U \setminus K'')$ (vérifier) : c'est fini.

(viii) *Proposition (Intégration sur un sous-ensemble mesurable)*. Si f est une fonction intégrable et si A est un ensemble mesurable, alors la fonction $f\chi_A$ est intégrable. En effet, $f\chi_A$ est mesurable [(iv) 2)] et l'on a $\mu^*(|f|\chi_A) \leq \mu^*(|f|) < +\infty$. On pose

$$\mu(f\chi_A) = \int_A f\, d\mu\,.$$

(2.3.3) *Espace* L^1 *des classes de fonctions intégrables*

Nous avons déjà introduit l'espace $\mathsf{L}^1 = \mathsf{L}^1(X, \mu)$ des fonctions μ-intégrables définies presque partout à valeurs dans $\overline{\mathbf{R}}$ **(2.2.4)** et montré que l'intégrale : $f \mapsto \mu(f) = \int_X f\, d\mu$, est une forme linéaire positive sur cet espace vectoriel de fonctions. L'application $f \mapsto \mu(|f|)$ est une *semi*-norme **(1.3.4** (c)**)** sur L^1 puisqu'on a $|f+g| \leq |f| + |g|$ et donc $\mu(|f+g|) \leq$

$\mu(|f| + |g|) = \mu(|f|) + \mu(|g|)$ pour f et $g \in \mathsf{L}^1$. Mais ce n'est *pas une norme* car $\mu(|f|)$ est nul dès que f est nulle *presque partout*. C'est un cas particulier du fait que l'intégrale ne change pas lorsque l'on remplace une fonction f de X dans $\overline{\mathbf{R}}$ par une fonction g égale à f presque partout dans X, en d'autres termes ne dépend que de la classe $\widetilde{f}$ de f pour la relation d'équivalence "$f = g$ μ-presque partout". Il est donc *commode* d'introduire l'espace vectoriel, noté $\mathrm{L}^1 = \mathrm{L}^1(\mathrm{X}, \mu)$, des *classes des fonctions de* L^1 pour cette relation (L^1 est l'espace quotient $\mathsf{L}^1 / \mathsf{N}$, où N est l'espace vectoriel des fonctions définies presque partout et nulles presque partout). Par construction, l'application $\widetilde{f} \mapsto \|\widetilde{f}\|_1 = \int_{\mathrm{X}} |f| d\mu = \mu(|f|)$ est bien définie (i.e. la valeur $\|\widetilde{f}\|_1$ ne dépend pas du représentant f de $\widetilde{f}$) et cette fois c'est une *norme* sur L^1 (puisque toute fonction nulle presque partout est équivalente à la fonction nulle!).

(i) ***Théorème***. *L'espace* L^1, *muni de la norme* $\|\ \|_1$, *est un espace de Banach*, autrement dit : *si* (f_n) *est une suite de fonctions intégrables telle que* $\int_{\mathrm{X}} |f_n - f_m| d\mu$ *tende vers 0 quand m et n tendent vers l'infini, il existe une fonction intégrable f telle que* $\int_{\mathrm{X}} |f_n - f| d\mu \to 0$. *De plus, il existe une suite extraite* (f_{n_k}) *qui tend presque partout vers f.*

Soit (f_n) une suite de représentants d'une suite de Cauchy $(\widetilde{f}_n)$ d'éléments de L^1. Ceci veut dire que les f_n sont des fonctions intégrables et qu'il existe pour chaque $\varepsilon>0$ un entier $N(\varepsilon)$ (que nous pouvons bien considérer comme une fonction univoque de ε, par exemple en prenant le plus petit qui convienne) tel que l'on ait $\int_{\mathrm{X}} |f_n - f_m| d\mu \le \varepsilon$ lorsque $n \ge N(\varepsilon)$ et $m \ge N(\varepsilon)$. Construisons par récurrence une suite extraite (f_{n_k}) telle que

$$\int_{\mathrm{X}} \left| f_{n_k} - f_{n_l} \right| d\mu \le 1/2^k \quad \text{pour } l \ge k. \quad (1)$$

En partant de $n_1 = N(1/2)$, il suffit de prendre $n_{k+1} = \max(N(1/2^{k+1}), n_k + 1)$. Nous pouvons alors appliquer la proposition **(2.3.1** (v)**)** à la suite des fonctions $g_k = f_{n_{k+1}} - f_{n_k}$; il en résulte que la série $\Sigma_k\, g_k(x)$ est presque partout absolument convergente, ce qui signifie notamment que la suite (f_{n_k}) tend presque partout vers une limite $f = f_{n_1} + G$, où G est la somme de la série, et que f est une fonction intégrable. Passons à la limite $l \to \infty$ dans l'inégalité (1). Pour cela, nous observons que

$$\left| f_{n_k} - f_{n_l} \right| = \left| \sum_{p=k}^{l-1} f_{n_p} - f_{n_{p+1}} \right| \le \sum_{p=k}^{\infty} \left| f_{n_p} - f_{n_{p+1}} \right| = h$$

et que h est une fonction intégrable, à nouveau grâce à **(2.3.1** (v)**)** et à l'inégalité (1). Le théorème de convergence dominée nous donne donc
$\int_{\mathrm{X}} \left| f_{n_k} - f \right| d\mu = \lim_{l\to\infty} \int_{\mathrm{X}} \left| f_{n_k} - f_{n_l} \right| d\mu \le 1/2^k$, ce qui prouve que la suite $(\widetilde{f}_{n_k})$ converge vers $\widetilde{f}$ dans l'espace normé L^1. Avec **(1.1.6** (ii)**)**, nous en déduisons que $\widetilde{f}_n \to \widetilde{f}$ dans L^1, cqfd.

La proposition démontrée en **(2.2.4** (iii)**)** peut se réexprimer ainsi :

(ii) *Proposition*. L'ensemble des classes des fonctions continues à support compact est dense dans L^1, ou encore: toute fonction intégrable est limite en moyenne de fonctions continues à support compact.

Si $f = g + i\,h$ (avec g et h réelles) est une fonction mesurable à valeurs *complexes* (finies), il est clair d'après la définition **(2.3.2** (c)**)** que g et h sont des fonctions mesurables (à valeurs finies). Lorsque g et h sont de plus intégrables, on dit que f est intégrable et l'on pose $\mu(f) = \mu(g) + i\,\mu(h)$. Il est facile de vérifier, d'après **(2.2.4** (iv)**)**, que l'ensemble des fonctions complexes intégrables est un espace vectoriel sur $\mathbf{C}$, noté $\mathcal{L}^1_{\mathbf{C}}(X, \mu)$, et que l'application $f \to \mu(f)$, ainsi définie, est une forme linéaire sur $\mathcal{L}^1_{\mathbf{C}}(X, \mu)$. De par la définition de $\mu(f)$, on a $\mathrm{R}(\mu(f)) = \mu(\mathrm{R}(f)) = \mu(g)$.

(iii) *Proposition.* Si f est intégrable, le module $|f|$ est intégrable et l'on a $|\mu(f)| \leq \mu(|f|)$.

Si $f = g + i\,h$ (avec g et h réelles) est intégrable, alors par hypothèse g et h sont intégrables, donc $|f| = (g^2 + h^2)^{1/2} \leq \sqrt{2}\sup(|g|, |h|)$ est intégrable d'après **(2.2.4** (v)**)**. De plus, il existe un nombre complexe z, de module 1, tel que $z\,\mu(f) = |\mu(f)|$, ce qui est réel, d'où $z\,\mu(f) = \mu(zf) = \mathrm{R}(\mu(zf)) = \mu(\mathrm{R}(zf))$. Mais $\mathrm{R}(zf) \leq |f|$ puisque $|z| = 1$, donc $\mu(\mathrm{R}(zf)) \leq \mu(|f|)$, c'est-à-dire $|\mu(f)| \leq \mu(|f|)$, cqfd.

(2.3.4) *Espace* L^2 *des classes de fonctions de carré intégrable*

Définition. On note $\mathcal{L}^2 = \mathcal{L}^2(X, \mu)$, et plus précisément $\mathcal{L}^2_{\mathbf{R}}(X, \mu)$ (ou $\mathcal{L}^2_{\mathbf{C}}(X, \mu)$) s'il y a risque de confusion, l'ensemble des fonctions *mesurables* réelles (ou complexes) f dont le *carré* (ou le carré du module), $|f|^2$, est une fonction intégrable. On dit que $\mathcal{L}^2$ est l'ensemble des *fonctions de carré intégrable* réelles (ou complexes). Dans le cas complexe, on demande que $f(x)$ soit bien un nombre complexe (donc *fini*), pour tout x dans X.

En toute rigueur, $|f|^2$ peut être intégrable sans que f soit mesurable. Dans ce cas, $f \notin \mathcal{L}^2$, donc "fonction de carré intégrable" tout court est un abus de langage (cet abus n'est pas dangereux, car les fonctions non mesurables sont des monstres qu'on ne croise jamais).

(i) *Proposition*. Si f et g sont des fonctions réelles (resp. complexes) de carré intégrable, alors fg (resp. $|fg|$) est définie presque partout et intégrable.

En effet, f et g sont finies presque partout **(2.2.3** (iv)**)** donc fg (resp. $|fg|$) est définie presque partout et équivalente à une fonction mesurable **(2.3.2** (iv)**)** et l'on a $|fg| \leq (|f|^2 + |g|^2)/2$, d'où $\mu^*(|fg|) \leq [\mu^*(|f|^2) + \mu^*(|g|^2)]/2 < \infty$, ce qui prouve que fg (resp. $|fg|$) est intégrable **(2.3.2** (vi)**)**.

(ii) *Proposition*. $\mathcal{L}^2$ est un espace vectoriel (réel ou complexe selon le cas).

Il est clair que $\alpha f \in \mathcal{L}^2$ si $f \in \mathcal{L}^2$ et α est un scalaire. Si $f \in \mathcal{L}^2$ et $g \in \mathcal{L}^2$, alors $f + g$ est mesurable **(2.3.2** (ii)**)** et $|f + g|^2 \leq (|f| + |g|)^2 \leq |f|^2 + |g|^2 + 2\,|fg|$, donc $|f + g|^2$ est intégrable grâce à la proposition (i) et à **(2.3.2** (vi)**)**.

Définition. Soit $\mathrm{L}^2 = \mathrm{L}^2(X, \mu)$ l'ensemble des classes des fonctions de $\mathcal{L}^2$ (pour la relation "$f = g$ presque partout"). Si $f \in \mathcal{L}^2$ et $g \in \mathcal{L}^2$, posons

$$(f \mid g) = \int_X f(x)\,\overline{g(x)}\,d\mu(x). \qquad (1)$$

(où $\bar{z}$ est le complexe conjugué, $\bar{z} = z$ si $z \in \mathbf{R}$). Ceci définit une forme bilinéaire symétrique sur $\mathcal{L}^2 \times \mathcal{L}^2$ (resp. une forme sesquilinéaire hermitienne si on considère des fonctions à valeurs complexes [3]). La forme correspondante sur $L^2 \times L^2$: $(\widetilde{f}\ ,\ \widetilde{g}) \to (f \mid g)$, est bien définie (le produit $(f \mid g)$ ne change pas si on modifie f et g sur un ensemble négligeable) et c'est un *produit scalaire* : d'une part on a toujours $(f \mid f) \geq 0$ et d'autre part $(f \mid f) = 0$ signifie que $\int_X |f|^2\, d\mu = 0$ et entraîne donc que $f = 0$ presque partout, i.e. $\widetilde{f} = 0$.

(iii) *Inégalité de Cauchy-Schwarz pour les fonctions de carré intégrable*. Si f et g sont des fonctions de carré intégrable, alors $f.g$ est intégrable et l'on a

$$|(f \mid g)| = \left|\int_X f.\overline{g}\, d\mu\right| \leq \int_X |f.g|\, d\mu \leq \sqrt{\int_X |f|^2 d\mu}\ \sqrt{\int_X |g|^2 d\mu} \tag{2}$$

L'inégalité dépend seulement du fait que l'application $(f\ ,\ g) \to (f \mid g)$ est une forme bilinéaire symétrique (ou une forme sesquilinéaire hermitienne pour les fonctions à valeurs dans **C**) *positive*, i.e. $(f \mid f) \geq 0$ quelle que soit f. Elle est donc élémentaire dans le cas réel.

Traitons le cas moins simple d'une forme définie un espace vectoriel *complexe*, ce qui fournira l'inégalité pour les fonctions à valeurs complexes. La sesquilinéarité donne, pour tout nombre complexe z :

$$(f + zg \mid f + z\,g) = (f \mid f) + \overline{z}\ (f \mid g) + z\,(g \mid f)\ + z\,\overline{z}\ (g \mid g).$$

Prenons alors $z(t) = t\,(f \mid g)$ avec t réel, de façon à faire apparaître des modules. Nous obtenons en utilisant le fait que la forme est hermitienne :

$$z\,(g \mid f) = t\ |(f \mid g)|^2 = \overline{z}\ (f \mid g),\quad z\,\overline{z}\ = t^2\,|(f \mid g)|^2.$$

Puisque la forme est positive, le trinôme du deuxième degré

$$P(t) = (f + z(t)\,g \mid f + z(t)\,g) = |(f \mid g)|^2\,(g \mid g)\,t^2 + 2t\,|(f \mid g)|^2 + (f \mid f)$$

a des coefficients *réels* et il est toujours ≥ 0, donc le discriminant réduit est ≤ 0, i.e.:

$|(f \mid g)|^4 - |(f \mid g)|^2 (g \mid g)\,(f \mid f) \leq 0$, d'où $|(f \mid g)|^2 \leq (g \mid g)\,(f \mid f)$, ce qui dans le cas présent se réécrit

$$|(f \mid g)| = \left|\int_X f.\overline{g}\ d\mu\right| \leq \sqrt{\int_X |f|^2 d\mu}\ \sqrt{\int_X |g|^2 d\mu}\,.$$

En appliquant cette inégalité aux fonctions $f' = |f|$ et $g' = |g|$, qui sont dans L^2, on obtient

$$(\,|f| \mid |g|\,) = \int_X |f|.|g|\ d\mu \leq \sqrt{\int_X |f|^2 d\mu}\ \sqrt{\int_X |g|^2 d\mu}\,.$$

Enfin, l'inégalité $\left|\int_X f.\overline{g}\ d\mu\right| \leq \int_X |f.g|\, d\mu$ est un cas particulier de **(2.3.3** (iii)**)**.

---

[3] Soit E un espace vectoriel sur **C.** Une application Φ : ExE $\to$ **C** est une *forme sesquilinéaire sur* E si l'on a $\forall z \in \mathbf{C}\ \forall x \in E\ \forall x' \in E\ \forall y \in E \qquad \Phi(z.x + x', y) = z\ \Phi(x, y) + \Phi(x', y)$ et $\Phi(y, z.x + x') = \overline{z}\ \ \Phi(y, x) + \Phi(y, x')$.

On dit qu' une forme sesquilinéaire Φ sur E est *hermitienne* si $\forall x \in E\ \forall y \in E\ \ \Phi(y, x) = \overline{\Phi(x, y)}$

(iv) ***Théorème***. *L'espace* L^2, *muni de la norme* $\| \widetilde{f} \|_2 = |(f \mid f)|^{1/2}$ *associée au produit scalaire* $(\widetilde{f}, \widetilde{g}) \to (f \mid g)$, *est complet, on dit que c'est un espace de Hilbert.* Autrement dit : *si* (f_n) *est une suite de fonctions de carré intégrable telle que* $\int_X |f_n - f_m|^2 \, d\mu$ *tende vers* 0 *quand m et n tendent vers l'infini, il existe une fonction de carré intégrable f telle que* $\int_X |f_n - f|^2 \, d\mu \to 0$. *De plus, il existe une suite extraite* (f_{n_k}) *qui tend presque partout vers f.*

La démonstration est très similaire à celle pour L^1, nous ne la ferons pas.

(v) *Inégalité de Cauchy-Schwarz pour les fonctions mesurables*

Théorème. Soient deux fonctions mesurables *f* et *g* de X dans $\overline{\mathbf{R}}$ ou dans $\mathbf{C}$. Si le produit $|f|.|g|$ a un sens presque partout, on a avec la convention $0 \leq 0.\infty$:

$$\int_X^* |f|.|g| d\mu \quad \leq \sqrt{\int_X^* |f|^2 d\mu} \; \sqrt{\int_X^* |g|^2 d\mu} \; , \qquad (3)$$

le cas d'indétermination $0.\infty$ à droite impliquant que le membre de gauche est nul.

Traitons en premier lieu le cas ou l'une des intégrales du second membre est nulle, ce qui signifie que le deuxième membre est, soit nul, soit le produit indéterminé $0.\infty$. Si par exemple $\mu^*(|f|^2) = 0$, on a $f = 0$ sur le complémentaire d'un ensemble négligeable N **(2.2.3** (iii)**)**. Puisque le produit $|f|.|g|$ a par hypothèse un sens sur le complémentaire d'un ensemble négligeable N', on a $|f|.|g| = 0$ sur $X \backslash (N \cup N')$ i.e. presque partout et donc le membre de gauche est nul. Si maintenant l'une des intégrales du deuxième membre est infinie et que l'autre soit non nulle, le deuxième membre est infini et il n'y a rien à démontrer.

Nous pouvons donc supposer que $\mu^*(|f|^2)$ et $\mu^*(|g|^2)$ sont finies et non nulles, ce qui entraîne que *f* et *g* sont dans L^2 : nous sommes donc ramenés au cas étudié dans (iii).

(2.3.5) *Produits de mesures. Théorème de Lebesgue-Fubini*

Nous allons voir que la théorie de Lebesgue donne une réponse simple et générale à la question de l'existence des intégrales doubles et de l'interversion de l'ordre des intégrations.

(i) *Théorème*

Soient X et Y deux espaces de base, μ une mesure sur X, ν une mesure sur Y. Il existe une et une seule mesure π sur X×Y telle que l'on ait pour tout couple de fonctions $f \in \mathsf{K}(X)$, $g \in \mathsf{K}(Y)$:

$$\int_{X \times Y} f(x)\, g(y)\, d\pi(x,y) = \left(\int_X f(x)\; d\mu(x)\right)\left(\int_Y g(y)\; d\nu(y)\right). \qquad (1)$$

Dans cet énoncé, μ et ν sont des mesures quelconques (pas nécessairement positives). Il est clair que l'égalité (1) est le minimum de ce qu'on peut demander à une mesure sur X×Y pour qu'on puisse l'appeler *mesure produit de* μ *par* ν (ce qui est effectivement le nom donné à la mesure π). Il est donc agréable qu'il n'existe qu'une seule mesure ayant cette propriété.

Nous allons démontrer cette unicité... en admettant un autre théorème. Une mesure ρ sur X×Y est une forme linéaire sur K(X×Y) telle que, pour tout compact M de X×Y, la restriction de ρ à l'espace K(X×Y; M) soit continue **(2.1.1)**. Or, tout compact de X×Y est contenu dans un compact de la forme K×L où K est un compact de X et L un compact de Y. L'unicité de la mesure π vérifiant (1) résulte alors du théorème suivant, dont la démonstration est un peu technique :

(ii) *Théorème* (admis). Soient K un compact de X et L un compact de Y. L'ensemble des combinaisons linéaires de fonctions h de la forme $h(x, y) = f(x)\, g(y)$ avec $f \in$ K(X; K) et $g \in$ K(Y; L), est dense dans K(X×Y; K×L).

Autrement dit, étant donnée une fonction H continue sur X×Y et dont le support est un compact inclus dans K×L, il existe pour tout $\varepsilon>0$ des scalaires $\alpha_1,...,\alpha_k$ et des fonctions $f_n \in$ K(X; K) et $g_n \in$ K(Y; L) $(1 \leq n \leq k)$ tels que les fonctions $h_n \in$ K(X×Y; K×L), définies par $h_n(x, y) = f_n(x)\, g_n(y)$, vérifient $\left\| H - \sum_{n=1}^{k} \alpha_n\, h_n \right\|_\infty \leq \varepsilon$. Si π est une mesure vérifiant (1), on a $\pi\,(h_n) = \mu\,(f_n)\,\nu(g_n)$. Soit a la norme de la restriction de π à K(X×Y; K×L). En utilisant l'inégalité précédente, on obtient donc $\left| \pi\,(H) - \sum_{n=1}^{k} \alpha_n\, \mu\,(f_n)\,\nu(g_n) \right| \leq a\varepsilon$, d'où l'unicité de π puisque ε est arbitraire.

Il reste à démontrer qu'il existe effectivement une mesure π sur X×Y qui vérifie (1), i.e. il reste à définir *une* façon de calculer l'intégrale double d'une fonction h continue dans X×Y et de support compact. Nous pouvons par exemple commencer à intégrer par rapport à y :

(iii) *Proposition*. Si h est une fonction de K(X×Y), l'intégrale partielle $H(x) = \int_Y h(x, y)\, d\nu(y)$ est une fonction continue (à support compact). En posant $\pi(h) = \int_X d\mu(x) \int_Y h(x, y)\, d\nu(y)$, on définit une mesure sur X×Y vérifiant (1).

Cette proposition n'a rien de surprenant et nous n'en ferons pas la démonstration assez facile (il faut, comme ci-dessus, considérer un compact "carré" M = K×L de X×Y). L'égalité (1) est évidente, compte-tenu de la définition de π. Notons aussi qu'une fois acquise la continuité de H, il est immédiat que π est bien une *mesure* (positive) lorsque μ et ν sont des mesures *positives*, car la linéarité et la positivité sont évidentes, et suffisent à garantir la continuité de la restriction de π à tout espace K(X×Y; M) **(2.1.3** (i)**)**.

Puis à cause de l'unicité de π, *l'interversion des intégrations* est licite pour les fonctions de K(X×Y) : en intégrant d'abord par rapport à x, nous obtenons tout aussi bien une mesure vérifiant (1), or il n'y en a qu'une seule...

On peut ensuite itérer l'opération et ainsi définir le *produit de n mesures*. En particulier, le produit des n mesures de Lebesgue sur chaque espace **R** facteur de $\mathbf{R}^n = \mathbf{R} \times ... \times \mathbf{R}$ s'appelle la mesure de Lebesgue sur $\mathbf{R}^n$ et se note souvent encore λ.

Le cas le plus important est celui où μ et ν sont des mesures *positives*. La mesure π peut alors, comme toute mesure positive, se prolonger en une forme linéaire sur un espace plus vaste que K(X×Y), à savoir l'espace L^1(X×Y, π). Le calcul de l'intégrale $\pi(h)$ pour une fonction intégrable quelconque $h \in L^1$(X×Y) s'effectue de la même façon que pour une fonction de K(X×Y), i.e. exactement comme celui d'une bonne vieille intégrale double. La différence, c'est qu'il y a très peu de précautions à prendre :

(iv) ***Théorème de Lebesgue-Fubini*** *(admis)*

Soient μ une mesure positive sur X, *ν une mesure positive sur* Y *et π leur mesure produit.*

(α) *Si h est une fonction π-<u>intégrable</u> sur* X×Y *alors, pour μ-presque tout x∈*X, *la fonction* $y \to h(x,y)$ *est ν-intégrable, l'intégrale partielle* $H(x) = \int_Y h(x, y)\, d\nu(y)$ *est une fonction μ-intégrable et l'on a* $\int_{X\times Y} h(x, y)\, d\pi(x, y) = \int_X d\mu(x) \int_Y h(x, y)\, d\nu(y)$. *De même on a* $\int_{X\times Y} h(x, y)\, d\pi(x, y) = \int_Y d\nu(y) \int_X h(x, y)\, d\mu(x)$, *i.e. on peut intervertir l'ordre des intégrations.*

(β) *Si* $h \geqq 0$ *est une fonction π-<u>mesurable</u> sur* X×Y, *l' application* $x \to \int^*_Y h(x, y)\, d\nu(y)$ *est μ-mesurable et l'on a* $\int^*_{X\times Y} h(x, y)\, d\pi(x, y) = \int^*_X d\mu(x) \int^*_Y h(x, y)\, d\nu(y)$. *De même on a* $\int^*_{X\times Y} h(x, y)\, d\pi(x, y) = \int^*_Y d\nu(y) \int^*_X h(x, y)\, d\mu(x)$, *i.e. on peut intervertir l'ordre des intégrations. Pour que h soit π-intégrable, il faut et il suffit donc que l'un de ces trois nombres égaux soit fini.*

A cause de ce résultat, on note aussi $\int_{X\times Y} h(x, y)\, d\pi(x, y) = \iint_{X\times Y} h(x, y)\, d\mu(x)\, d\nu(y)$. En pratique la mesurabilité sera toujours vraie. La seule chose qui est illicite est donc de calculer une intégrale double pour une fonction qui ne garde pas un signe constant, sans vérifier que l'intégrale double de la valeur absolue (calculée comme on le souhaite) est finie. Bien entendu, le théorème s'étend aux intégrales multiples, par associativité.

(2.3.6) *Changement de variables et intégrale de Lebesgue*

Théorème (admis). *Soient λ la mesure de Lebesgue sur* $\mathbf{R}^n$, U *et* U' *deux ouverts de* $\mathbf{R}^n$, *et* F : $x = (x_j) \mapsto y = (y_i) = F(x) = (F_i((x_j)))$ *une bijection de* U *sur* U' *telle que les dérivées partielles* $J_{ij} = \dfrac{\partial F_i}{\partial x_j}$ *existent et soient continues et que le jacobien* $J = \det\left(J_{ij}\right)$ *ne s'annule pas sur* U. *Alors, pour toute fonction* $f \in \mathsf{L}^1(\mathrm{U}', \lambda)$, *la fonction* $g = |J|.f \circ F$ *appartient à* $\mathsf{L}^1(U, \lambda)$ *et l'on a :*

$$\int_{\mathrm{U}'} f(y)\, d\lambda(y) = \int_{\mathrm{U}} |J(x)|\, f(F(x))\, d\lambda(x).$$

Ceci se note aussi

$$\int_{\mathrm{U}'} f(y)\, dV(y) = \int_{\mathrm{U}} |J(x)|\, f(F(x))\, dV(x)$$

($V=\lambda$ étant effectivement la mesure du volume ou "hypervolume" sur $\mathbf{R}^n$). Cette notation est avantageuse s'il y a aussi des intégrales de surface, pour lesquelles on remplace V par S (mesure définie sur une surface ou hypersurface; la définition générale relève de la géométrie différentielle). On rencontre aussi la notation

$$\int_{\mathrm{U}'} f(y)\, dy = \int_{\mathrm{U}} |J(x)|\, f(F(x))\, dx.$$

Bien entendu, on peut également utiliser la notation classique $dx = dx_1 \ldots dx_n$.