Théorie de l’intégration

Licence de Mathématiques, 3ème année

Table des matières

Chapitre 1

Vocabulaire

1. Dénombrabilité

1.1. Ensembles dénombrables. Un ensemble D est dit **dénombrable** si ou bien $D = \varnothing$, ou bien on peut énumérer les éléments de D comme les termes d'une suite, *i.e.* écrire $D = \{x_0, x_1, x_2, \dots\} = \{x_n;\ n \in \mathbb{N}\}$. Si $D \neq \varnothing$, il revient au même de dire qu'il existe une *surjection* $s : \mathbb{N} \to D$ de $\mathbb{N}$ sur D (on prend alors $x_n := s(n)$).

Exemple 1. Par définition, $\mathbb{N}$ est dénombrable.

Exemple 2. $\mathbb{Z}$ est dénombrable car $\mathbb{Z} = \{0, 1, -1, 2, -2, \dots\}$.

Remarque 1. Tout ensemble *fini* est dénombrable : si $D = \{a_1, \dots, a_N\}$, on peut écrire $D = \{a_1, \dots, a_N, a_N, a_N, \dots\}$.

Remarque 2. En fait, un ensemble D est dénombrable si et seulement si il est fini ou en bijection avec $\mathbb{N}$.

Démonstration. Il suffit de montrer que tout ensemble dénombrable infini D est en bijection avec $\mathbb{N}$. En écrivant $D = \{x_0, x_1, x_2, \dots\}$, on définit une bijection $\sigma : \mathbb{N} \to D$ en posant $\sigma(0) = x_0$ et $\sigma(n) = x_{i(n)}$ pour $n \geqslant 1$, où $i(n)$ est le plus petit entier $i > i(n-1)$ tel que $x_i \neq x_j$ pour tout $j < i$ (vérifier). □

Remarque 3. Tout sous-ensemble d'un ensemble dénombrable est dénombrable.

Démonstration. Soit D' une partie d'un ensemble dénombrable D. Si $D' = \varnothing$, il n'y a rien à démontrer ; on suppose donc $D' \neq \varnothing$ et on choisit un point $a \in D'$. Si $D = \{x_0, x_1, x_2, \dots\}$, on peut alors énumérer D' comme $D' = \{x'_0, x'_1, x'_2, \dots\}$, où $x'_n = x_n$ si $x_n \in D'$ et $x'_n = a$ si $x_n \notin D'$. □

Remarque 4. Si on peut écrire $D = \{x_i;\ i \in I\}$ avec un ensemble d'indices I dénombrable, alors D est dénombrable.

Démonstration. C'est évident : comme I est dénombrable, on peut écrire $I = \{i_0, i_1, i_2, \dots\}$, et donc $D = \{z_0, z_1, z_2, \dots\}$ où $z_n := x_{i_n}$. □

PROPOSITION 1.1. *L'ensemble* $\mathbb{N} \times \mathbb{N}$ *est dénombrable.*

Démonstration. On peut énumérer $\mathbb{N}\times\mathbb{N}$ en "serpentant" le long des anti-diagonales

$$\Delta_i := \{(m, n);\ m + n = i\}\ ,\ i \in \mathbb{N}\,;$$

autrement dit, en écrivant

$$\mathbb{N} \times \mathbb{N} = \{(0,0), (1,0), (0,1), (0,2), (1,1), (2,0), (3,0), (2,1), \dots\}\,.$$

□

COROLLAIRE 1.2. *Tout* produit fini *d'ensembles dénombrables est dénombrable : si* $D_1, \dots, D_N$ *sont dénombrables, alors* $D_1 \times \cdots \times D_N$ *aussi.*

Démonstration. Par une récurrence instantanée, il suffit de traiter le cas $N = 2$; et ce cas découle immédiatement de la proposition puisque $D_1 \times D_2$ s'énumère par $\mathbb{N} \times \mathbb{N}$ si D_1 et D_2 sont dénombrables. □

Corollaire 1.3. *$\mathbb{Q}$ est dénombrable.*

Démonstration. C'est clair puisque $\mathbb{Q} = \left\{\frac{p}{q};\ (p,q) \in \mathbb{Z} \times \mathbb{N}^*\right\}$. □

Corollaire 1.4. *Toute* réunion dénombrable *d'ensembles dénombrables est un ensemble dénombrable : si $(A_i)_{i\in I}$ est une famille d'ensembles dénombrables indexée par un ensemble I lui même dénombrable, alors $A = \bigcup_{i\in I} A_i$ est dénombrable.*

Démonstration. Pour tout $i \in I$, on *choisit* une énumération de A_i par $\mathbb{N}$, disons $A_i = \{x_{i,n};\ n \in \mathbb{N}\}$. Alors $A = \{x_{i,n};\ (i,n) \in I \times \mathbb{N}\}$, donc A est dénombrable car $I \times \mathbb{N}$ l'est. □

Exemple 3. $\mathbb{R}$ n'est pas dénombrable.

Démonstration. Il existe de nombreuses preuves de ce fait essentiel. Voici une des plus expéditives : si $x \in \mathbb{R}$, alors $\mathbb{R}\backslash\{x\}$ est un ouvert dense de $\mathbb{R}$ car $\mathbb{R}$ n'a pas de point isolé ; par conséquent, si $D = \{x_n;\ n \in \mathbb{N}\}$ est une partie dénombrable de $\mathbb{R}$, alors $\mathbb{R}\backslash D = \bigcap_{n\in\mathbb{N}} \mathbb{R}\backslash\{x_n\}$ est une intersection dénombrable d'ouverts denses, et est donc encore dense dans $\mathbb{R}$ par le *théorème de Baire* ; en particulier, $\mathbb{R}\backslash D \neq \varnothing$. □

Exemple 4. L'ensemble $\{0,1\}^{\mathbb{N}}$ de toutes les suites infinies de 0 et de 1 n'est pas dénombrable.

Démonstration. Soit $D = \{x_n;\ n \in \mathbb{N}\}$ une partie dénombrable quelconque de $\{0,1\}^{\mathbb{N}}$. Il s'agit de montrer que $D \neq \{0,1\}^{\mathbb{N}}$, autrement dit de trouver $x \in \{0,1\}^{\mathbb{N}}$ qui soit différent de tous les x_n. Si on écrit chaque x_n sous la forme

$$x_n = (x_n(0), x_n(1), x_n(2), \dots),$$

il suffit de définir $x = (x(0), x(1), x(2), \dots)$ de sorte que x diffère de x_n sur la coordonnée d'indice n, pour tout $n \in \mathbb{N}$. Autrement dit, on pose pour tout $n \in \mathbb{N}$:

$$x(n) = \begin{cases} 1 & \text{si } x_n(n) = 0, \\ 0 & \text{si } x_n(n) = 1. \end{cases}$$

□

Remarque. Cette méthode de démonstration, dûe à Cantor, s'appelle la **méthode de la diagonale**. La raison est la suivante : si on écrit les suites $x_0, x_1, \dots$ en ligne, on obtient le tableau infini

$$\begin{array}{cccc} x_0(0) & x_0(1) & x_0(2) & \dots \\ x_1(0) & x_1(1) & x_1(2) & \dots \\ x_2(0) & x_2(1) & x_2(2) & \dots \\ \dots & \dots & \dots & \ddots \end{array}$$

La suite $x = (x(0), x(1), x(2), \dots)$ de la preuve précédente est la suite obtenue en changeant tous les 0 en 1 et tous les 1 en 0 dans la *diagonale* de ce tableau.

Exercice 1. Soit $\varphi : \{0,1\}^{\mathbb{N}} \to \mathbb{R}$ l'application définie par

$$\varphi(\alpha) = \sum_{n=0}^{\infty} \frac{\alpha_n}{3^n} \qquad \text{pour } \alpha = (\alpha_0, \alpha_1, \alpha_2, \dots) \in \{0,1\}^{\mathbb{N}}.$$

Montrer que si $\alpha \neq \alpha'$, alors $|\varphi(\alpha) - \varphi(\alpha')| \geqslant \frac{1}{2\times 3^{n_0}}$, où n_0 est le premier entier tel que $\alpha_{n_0} \neq \alpha'_{n_0}$. Conclure que φ est injective, et en déduire une autre preuve du fait que $\mathbb{R}$ n'est pas dénombrable.

Exercice 2. Montrer que l'ensemble de tous les polynômes à coefficients rationnels est dénombrable, et en déduire qu'il existe des nombres réels x qui ne sont racines d'aucune équation $P(x) = 0$, où P est un polynôme non nul à coefficients rationnels. (De tels nombres x sont dits **transcendants**.)

1.2. Espaces métriques séparables. Un espace métrique Ω est dit **séparable** s'il existe un ensemble $D \subseteq \Omega$ à la fois dénombrable et dense dans Ω. Par exemple, $\mathbb{R}$ est séparable car $\mathbb{Q}$ est dense dans $\mathbb{R}$; et $\mathbb{R}^N$ est séparable pour tout $N \geqslant 1$, car $\mathbb{Q}^N$ est dense dans $\mathbb{R}^N$.

Proposition 1.5. (propriété de Lindelöf)
Soit (Ω, d) un espace métrique séparable. Si $(O_i)_{i\in I}$ est une famille d'ouverts de Ω, alors il existe un ensemble dénombrable *d'indices $I_0 \subseteq I$ tel que $\bigcup_{i\in I_0} O_i = \bigcup_{i\in I} O_i$.*

Démonstration. Soit $D \subseteq \Omega$ un ensemble dénombrable dense, et soit

$$\Lambda := \left\{(z,n) \in D \times \mathbb{N}^*;\ \exists i \in I\ :\ B(z, 1/n) \subseteq O_i\right\},$$

où $B(z, 1/n)$ est la boule ouverte de centre z et de rayon $1/n$.

Pour tout $(z,n) \in \Lambda$, choisissons un indice $i(z,n) \in I$ tel que $B(z,1/n) \subseteq O_{i(z,n)}$. Alors $I_0 := \{i(z,n);\ (z,n) \in \Lambda\}$ est dénombrable car Λ l'est. On va voir que I_0 convient.

Soit $x \in \bigcup_{i\in I} O_i$ quelconque. Choisissons $i \in I$ tel que $x \in O_i$ et $r > 0$ tel que $B(x,r) \subseteq O_i$. Soit également $n \in \mathbb{N}^*$ tel que $2/n < r$. Comme D est dense dans Ω, on peut trouver $z \in D$ tel que $d(z,x) < 1/n$. Alors $B(z,1/n) \subseteq B(x,2/n)$ par l'inégalité triangulaire ; donc $B(z,1/n) \subseteq B(x,r) \subseteq O_i$. Par définition de Λ, on voir donc que $(z,n) \in \Lambda$. Donc $i_0 := i(z,n) \in I_0$; et comme $d(x,z) < 1/n$, on a $x \in B(z,1/n) \subseteq O_{i_0}$. Ainsi, pour tout $x \in \bigcup_{i\in I} O_i$, on peut trouver $i_0 \in I_0$ tel que $x \in O_{i_0}$. □

Corollaire 1.6. *Si $\Omega_1, \dots, \Omega_N$ sont des espaces métriques séparables, alors tout ouvert de $\Omega := \Omega_1 \times \cdots \times \Omega_N$ est réunion dénombrable de* ***produits d'ouverts****, i.e d'ensembles de la forme $O_1 \times \cdots \times O_N$, où O_j est ouvert dans Ω_j pour $j = 1, \dots, N$.*

Démonstration. Soit O un ouvert de Ω. Pour tout $x \in O$, on peut trouver un produit d'ouverts O_x tel que $x \in O_x$ et $O_x \subseteq O$. On a alors $O = \bigcup_{x\in O} O_x$, d'où le résultat par la proposition appliquée à la famille $(O_x)_{x\in O}$. □

Corollaire 1.7. *Tout ouvert de $\mathbb{R}^N$, $N \geqslant 1$ est réunion dénombrable de* ***pavés ouverts****, c'est à dire d'ensembles de la forme $P = I_1 \times \cdots \times I_N$, où $I_1, \dots, I_N$ sont des intervalles ouverts bornés de $\mathbb{R}$.*

Démonstration. Si Ω est un ouvert de $\mathbb{R}^N$ alors, pour tout $x \in \Omega$, on peut trouver un pavé ouvert P_x tel que $x \in P_x$ et $P_x \subseteq \Omega$; donc le résultat est à nouveau une conséquence immédiate de la propriété de Lindelöf. □

Exercice 1. Montrer qu'un espace métrique est séparable si et seulement si il vérifie la propriété suivante : pour tout $\varepsilon > 0$, on peut trouver un ensemble dénombrable $D_\varepsilon \subseteq \Omega$ tel que $\Omega = \bigcup_{z\in D_\varepsilon} B(z,\varepsilon)$. En déduire que tout espace métrique *compact* est séparable.

Exercice 2. Montrer qu'un espace métrique est séparable *si et seulement* si il possède la propriété de Lindelöf.

2. La “droite numérique achevée”

Par définition, la droite numérique achevée est l'ensemble $\overline{\mathbb{R}}$ obtenu en rajoutant à $\mathbb{R}$ deux points notés ∞ et $-\infty$:

$$\overline{\mathbb{R}} = \{-\infty\} \cup \mathbb{R} \cup \{\infty\}\,.$$

2.1. Extension à $\overline{\mathbb{R}}$ de l'ordre de $\mathbb{R}$. On étend l'ordre usuel de $\mathbb{R}$ à $\overline{\mathbb{R}}$ en décrétant qu'on a

$$-\infty < x < \infty \qquad \text{pour tout } x \in \mathbb{R}\,.$$

Les *intervalles* de $\overline{\mathbb{R}}$ sont définis de manière évidente. En particulier, on a

$$\overline{\mathbb{R}} = [-\infty, \infty]\,.$$

L'intérêt de cette extension de l'ordre est que maintenant *tout ensemble non vide $A \subseteq \overline{\mathbb{R}}$ possède une borne supérieure et une borne inférieure dans $\overline{\mathbb{R}}$*. Pour un ensemble (non vide) $A \subseteq \mathbb{R}$, on a $\sup A < \infty$ si et seulement si A est *majoré*, et $\inf A > -\infty$ si et seulement si A est *minoré.*

Exercice. Soit $I \subseteq \overline{\mathbb{R}}$. Montrer que I est un intervalle si et seulement si la propriété suivante a lieu : pour tous $u, v \in I$ vérifiant $u < v$, on a $]u, v[\subseteq I$.

2.2. Convergence des suites. On dit qu'une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de $\overline{\mathbb{R}}$ **converge dans $\overline{\mathbb{R}}$** si

- ou bien $x_n \in \mathbb{R}$ à partir d'un certain rang et (x_n) converge dans $\mathbb{R}$ au sens usuel ;
- ou bien $x_n \to \infty$, ce qui s'écrit $\forall A \in]0, \infty[\ \exists N\ \forall n \geqslant N\ :\ x_n > A$;
- ou bien $x_n \to -\infty$, ce qui s'écrit $\forall A \in]0, \infty[\ \exists N\ \forall n \geqslant N\ :\ x_n < -A$.

Exercice. Soit $(x_n)_{n\in\mathbb{N}} \subseteq \overline{\mathbb{R}}$ et soit $l \in \overline{\mathbb{R}}$. Montrer que $x_n \to l$ si et seulement si

$$\forall \alpha, \beta \text{ tels que } \alpha < l < \beta \qquad \text{on a} \quad \alpha < x_n < \beta \text{ à partir d'un certain rang}\,.$$

La preuve du théorème suivant est laissée en exercice.

THÉORÈME 2.1. *Toute suite* croissante $(x_n) \subseteq \overline{\mathbb{R}}$ *converge dans $\overline{\mathbb{R}}$ vers sa borne supérieure, et tout suite* décroissante *converge vers sa borne inférieure.*

Notons enfin que, comme pour les suites de nombres réels, les inégalités *larges* se conservent à la limite : si $x_n \to l$ et si $x_n \leqslant \beta$ à partir d'un certain rang, alors $l \leqslant \beta$; et de même, si $x_n \to l$ et si $x_n \geqslant \alpha$ à partir d'un certain rang, alors $l \geqslant \alpha$.

2.3. Topologie de $\overline{\mathbb{R}}$. Il est très facile de définir une distance raisonnable sur $\overline{\mathbb{R}}$, de la façon suivante. Soit $\phi : \mathbb{R} \to]-a, a[$ une bijection continue strictement croissante de $\mathbb{R}$ sur un intervalle *ouvert borné* $]-a, a[\subseteq \mathbb{R}$. (Par exemple, on peut prendre $\phi(x) = \arctan(x)$ et $a = \pi/2$.) On prolonge ϕ en une bijection de $\overline{\mathbb{R}}$ sur $[-a, a]$ en posant $\phi(\infty) = a$ et $\phi(-\infty) = -a$; et pour $x, y \in \overline{\mathbb{R}}$, on pose alors

$$d_\phi(x, y) := |\phi(y) - \phi(x)|\,.$$

LEMME 2.2. *Si $\phi : \mathbb{R} \to]-a, a[$ est une bijection continue strictement croissante, alors d_ϕ est une distance sur $\overline{\mathbb{R}}$ et les suites convergentes pour d_ϕ sont les suites convergentes au sens défini plus haut.*

Démonstration. Le fait que d_ϕ soit une distance est un exercice facile, qui utilise uniquement le fait que la fonction ϕ est injective.

Soit $(x_n) \subseteq \overline{\mathbb{R}}$ convergeant vers $l \in \overline{\mathbb{R}}$ au sens de la section précédente. Si $l \in \mathbb{R}$, alors $x_n \to l$ au sens usuel, donc $\phi(x_n) \to \phi(l)$ par continuité de ϕ, et donc $d_\phi(x_n, l) \to 0$. Si $l = \infty$, alors $\phi(x_n) \to a$ et donc $d_\phi(x_n, \infty) = |\phi(x_n) - a| \to 0$; et de même si $l = -\infty$.

Inversement, si $(x_n) \subseteq \overline{\mathbb{R}}$ converge vers l au sens de d_ϕ, on montre que $x_n \to l$ au sens de la section précédente en utilisant *la continuité de* ϕ^{-1}. □

Dans la suite, on munit $\overline{\mathbb{R}}$ de la topologie définie par n'importe quelle distance compatible avec la convergence des suites (*i.e.* vérifiant la conclusion du lemme précédent).

THÉORÈME 2.3. *L'espace* $\overline{\mathbb{R}}$ *est* compact.

Démonstration. Soit (x_n) une suite quelconque d'éléments de $\overline{\mathbb{R}}$. Alors ou bien (x_n) admet une sous-suite qui tend vers ∞, ou bien (x_n) admet une sous-suite qui tend vers $-\infty$, ou bien $x_n \in \mathbb{R}$ à partir d'un certain rang n_0 et la suite $(x_n)_{n \geqslant n_0}$ est bornée, auquel cas (x_n) possède une sous-suite convergente d'après le théorème de Bolzano-Weierstrass. Dans tous les cas, (x_n) admet une sous-suite qui converge dans $\overline{\mathbb{R}}$. □

Exercice. Montrer que si A est une partie de $\overline{\mathbb{R}}$, alors $\sup A$ et $\inf A$ appartiennent à $\overline{A}$, l'adhérence de A dans $\overline{\mathbb{R}}$.

2.4. Limites supérieures et limites inférieures. Si $(x_n)_{n \in \mathbb{N}}$ est une suite d'éléments de $\overline{\mathbb{R}}$, on note $\overline{\lim}\, x_n$ et $\underline{\lim}\, x_n$ les éléments de $\overline{\mathbb{R}}$ définis par

$$\overline{\lim}\, x_n := \inf_{m \in \mathbb{N}} \sup_{n \geqslant m} x_n \qquad \text{et} \qquad \underline{\lim}\, x_n = \sup_{m \in \mathbb{N}} \inf_{n \geqslant m} x_n\,.$$

On dit que $\overline{\lim}\, x_n$ est la **limite supérieure** de la suite (x_n), et que $\underline{\lim}\, x_n$ est sa **limite inférieure**. Cette terminologie sera expliquée par le lemme 2.4 ci-après.

Remarque 1. Supposons que (x_n) soit une suite de nombres réels. On voit facilement qu'on a $\overline{\lim}\, x_n < \infty$ si et seulement si la suite (x_n) est *majorée* par un nombre réel, et qu' on a et $\underline{\lim}\, x_n > -\infty$ si et seulement si (x_n) est *minorée*. Par conséquent, $\overline{\lim}\, x_n$ et $\underline{\lim}\, x_n$ appartiennent tous les deux à $\mathbb{R}$ si et seulement si la suite (x_n) est *bornée.*

Remarque 2. On a toujours $\underline{\lim}\, x_n \leqslant \overline{\lim}\, x_n$.

Démonstration. Si on pose $L_m := \sup_{n \geqslant m} x_n$ et $l_m := \inf_{n \geqslant m} x_n$, alors L_m décroit vers $\overline{\lim}\, x_n$ et l_m croit vers $l := \underline{\lim}\, x_n$. D'où le résultat puisque $l_m \leqslant L_m$ pour tout $m \in \mathbb{N}$. □

LEMME 2.4. *Si* $(x_n)_{n \in \mathbb{N}}$ *est une suite d'éléments de* $\overline{\mathbb{R}}$*, alors* $\overline{\lim}\, x_n$ *et* $\underline{\lim}\, x_n$ *sont des valeurs d'adhérence de* (x_n) *dans* $\overline{\mathbb{R}}$*. Plus précisément,* L *est la* plus grande *valeur d'adhérence de* (x_n) *dans* $\overline{\mathbb{R}}$*, et* $\underline{\lim}\, x_n$ *est la* plus petite *valeur d'adhérence de* (x_n) *dans* $\overline{\mathbb{R}}$*.*

Démonstration. Rappelons d'abord le fait général suivant.

FAIT. Si (z_n) est une suite dans un espace métrique K, alors l'ensemble des valeurs d'adhérence de (z_n) est égal à

$$\bigcap_{k \in \mathbb{N}} \overline{\{z_n;\ n \geqslant k\}}\,.$$

Maintenant, posons $L := \overline{\lim}\, x_n$ et $l := \underline{\lim}\, x_n$. Posons également $L_m := \sup_{n\geqslant m} x_n$ et $l_m := \inf_{n\geqslant m} x_n$. Alors $L_m \to L$ et $l_m \to l$.

Si $k \in \mathbb{N}$, on a pour tout $m \geqslant k$: $L_m \in \overline{\{x_n;\ n \geqslant m\}} \subseteq \overline{\{x_n;\ n \geqslant k\}}$. En faisant tendre m vers l'infini, on en déduit que $L = \lim L_m$ appartient à $\overline{\{x_n;\ n \geqslant k\}}$ pour tout $k \in \mathbb{N}$, et donc que L est une valeur d'adhérence de (x_n) d'après le Fait. De même, l est une valeur d'adhérence de (x_n).

Soit maintenant a une valeur d'adhérence quelconque de (x_n) dans $\overline{\mathbb{R}}$, et soit (n_k) une suite strictement croissante d'entiers telle que $x_{n_k} \to a$. Si $m \in \mathbb{N}^*$ est fixé, alors $n_k \geqslant m$ pour k assez grand, et donc $x_{n_k} \leqslant L_m = \sup_{n\geqslant m} x_n$. En faisant tendre k vers l'infini, on obtient $a \leqslant L_m$ pour tout $m \in \mathbb{N}$, et donc $a \leqslant \inf_m L_m = L$. On montre de même que $a \geqslant l$. Ainsi $l \leqslant a \leqslant L$ pour toute valeur d'adhérence a de (x_n), ce qui prouve la deuxième partie du lemme. □

Corollaire 2.5. *Une suite $(x_n) \subseteq \overline{\mathbb{R}}$ converge dans $\overline{\mathbb{R}}$ si et seulement si $\overline{\lim}\, x_n = \underline{\lim}\, x_n$.*

Démonstration. Comme $\overline{\mathbb{R}}$ est compact, la suite (x_n) converge si et seulement si elle a exactement une valeur d'adhérence dans $\overline{\mathbb{R}}$, *i.e* $\overline{\lim}\, x_n = \underline{\lim}\, x_n$. □

Exercice 1. Montrer que les inégalités larges sont préservées par "passage à la limsup" et "passage à la liminf".

Exercice 2. Montrer que si (a_n) et (b_n) sont deux suites bornés de nombres réels, alors

$$\underline{\lim}\, a_n + \underline{\lim}\, b_n \leqslant \underline{\lim}\, (a_n + b_n) \leqslant \underline{\lim}\, a_n + \overline{\lim}\, b_n \leqslant \overline{\lim}\, (a_n + b_n) \leqslant \overline{\lim}\, a_n + \overline{\lim}\, b_n\,.$$

En déduire que si la suite (a_n) est convergente, alors

$$\overline{\lim}\, (a_n + b_n) = \lim a_n + \overline{\lim}\, b_n \quad \text{et} \quad \underline{\lim}\, (a_n + b_n) = \lim a_n + \underline{\lim}\, b_n\,.$$

Exercice 3. Soit (x_n) une suite bornée de nombres réels, et soit $\alpha \in [0, 1[$. On suppose que $x_n + \alpha\, x_{2n} \to 1$ quand $n \to \infty$. Montrer que si on pose $l := \underline{\lim}\, x_n$ et $L := \overline{\lim}\, x_n$, alors $l + \alpha L \geqslant 1 \geqslant L + \alpha l$; puis montrer que (x_n) converge et trouver sa limite.

2.5. Extension des opérations de $\mathbb{R}$ à $\overline{\mathbb{R}}$. Il n'est pas difficile d'étendre l'addition et la multiplication de $\mathbb{R}$ à $\overline{\mathbb{R}}$, en faisant tout de même un peu attention.

Pour l'addition, on adopte les conventions suivantes :

$$a + \infty = \infty + a = \infty \qquad \text{si } a \in \mathbb{R} \cup \{\infty\},$$
$$a + (-\infty) = -\infty = -\infty + a \qquad \text{si } a \in \mathbb{R} \cup \{-\infty\}\,.$$

Comme il est naturel :

$$\infty + (-\infty) \quad \text{et} - \infty + \infty \quad \textit{ne sont pas définis}\,.$$

Pour la multiplication, on convient évidemment que

$$a \times \infty = \infty = \infty \times a \quad \text{et} \quad a \times (-\infty) = -\infty = (-\infty) \times a \qquad \text{si } a > 0\,,$$
$$a \times \infty = -\infty = \infty \times a \quad \text{et} \quad a \times (-\infty) = \infty = (-\infty) \times a \qquad \text{si } a < 0\,.$$

Enfin, et on verra que ceci est *particulièrement important*, on convient que

$$0 \times \infty = 0 = \infty \times 0\,.$$

Exercice 1. Montrer que l'addition est continue sur $[0, \infty] \times [0, \infty]$, mais que la multiplication n'est *pas* continue.

Exercice 2. Montrer que si (x_n) est une suite d'éléments de $\overline{\mathbb{R}}$, alors

$$\overline{\lim}\,(-x_n) = -\underline{\lim}\,x_n \quad \text{et} \quad \underline{\lim}\,(-x_n) = -\overline{\lim}\,x_n\,.$$

Exercice 3. Montrer que (a_n) et (b_n) sont deux suites *minorées* d'éléments de $]-\infty,\infty]$ ou bien deux suites *majorées* d'éléments de $[-\infty,\infty[$, alors $\underline{\lim}\,(a_n+b_n) \leqslant \underline{\lim}\,a_n + \overline{\lim}\,b_n \leqslant \overline{\lim}\,(a_n+b_n)$.

3. Somme d'une famille de nombres positifs

Dans cette section et dans tous les chapitres suivants, on appellera **nombre positif** tout élément de $[0,\infty]$.

3.1. Somme d'une *série*. Si $(x_i)_{i\in\mathbb{N}}$ est une suite de nombres positifs, on pose

$$\sum_{i=0}^{\infty} x_i := \lim_{n\to\infty} \sum_{i=0}^{n} x_i\,.$$

Cette définition a un sens car la suite des sommes partielles $S_n := \sum_{i=0}^n x_i$ est *croissante*, donc admet une limite dans $[0,\infty]$.

Si les x_i sont des nombres *réels* positifs, alors

$$\sum_{i=0}^{\infty} x_i < \infty \iff \text{la série } \sum x_i \text{ converge au sens usuel}\,.$$

3.2. Somme d'une famille quelconque. Si $(x_i)_{i\in I}$ est une famille de nombres positifs indexée par un ensemble $I \neq \varnothing$ quelconque, on pose

$$\sum_{i\in I} x_i := \sup\left\{\sum_{i\in F} x_i;\ F\subseteq I\,,\ F \text{ fini}\right\}.$$

Il s'agit d'un nombre positif bien défini, car tout ensemble non vide $A \subseteq [0,\infty]$ possède une borne supérieure dans $[0,\infty]$.

Exemple. Si $I=\mathbb{N}$, alors $\sum\limits_{i\in\mathbb{N}} x_i = \sum\limits_{i=0}^{\infty} x_i = \lim\limits_{n\to\infty} \sum\limits_{i=0}^{n} x_i\,.$

Démonstration. Pour tout $n\in\mathbb{N}$, on a $\sum_{i=0}^n x_i = \sum_{i\in[0,n]} x_i \leqslant \sum_{i\in\mathbb{N}} x_i$ par définition de $\sum_{i\in\mathbb{N}} x_i$; donc $\sum_{i=0}^\infty x_i = \lim_{n\to\infty}\sum_{i=0}^n x_i \leqslant \sum_{i\in I} x_i$. Inversement, tout ensemble fini $F\subseteq\mathbb{N}$ est contenu dans $[0,n]$ pour un certain n, donc $\sum_{i\in F} x_i \leqslant \sum_{i=0}^n x_i \leqslant \sum_{i=0}^\infty x_i$; d'où $\sum_{i\in\mathbb{N}} x_i \leqslant \sum_{i=0}^\infty x_i$ en prenant le "sup en F". □

Exercice 1. Montrer que si $\sigma : I \to I$ est une *permutation* de I, *i.e.* une bijection de I sur I, alors $\sum_{i\in I} x_{\sigma(i)} = \sum_{i\in I} x_i$.

Exercice 2. Montrer que si A et B sont deux parties de I telles que $A\cap B=\varnothing$, alors $\sum_{i\in A\cup B} x_i = \sum_{i\in A} x_i + \sum_{i\in B} x_i$.

Le lemme suivant semble évident, mais demande quand même une preuve.

Lemme 3.1. (additivité)
Si $(x_i)_{i\in I}$ et $(y_i)_{i\in I}$ sont deux familles de nombres positifs indexées par le même ensemble I, alors $\sum_{i\in I}(x_i+y_i) = \sum_{i\in I} x_i + \sum_{i\in I} y_i$.

Démonstration. On a pour tout ensemble fini $F \subseteq I$:

$$\sum_{i\in F}(x_i+y_i)=\sum_{i\in F}x_i+\sum_{i\in F}y_i\leqslant\sum_{i\in I}x_i+\sum_{i\in I}y_i\,;$$

d'où $\sum_{i\in I}(x_i+y_i)\leqslant\sum_{i\in I}x_i+\sum_{i\in I}y_i$ en prenant le "sup en F".

Inversement, si $F, F' \subseteq I$ sont des ensembles finis quelconques, alors

$$\sum_{i\in F}x_i+\sum_{i\in F'}y_i\leqslant\sum_{i\in F\cup F'}x_i+\sum_{i\in F\cup F'}y_i=\sum_{i\in F\cup F'}(x_i+y_i)\leqslant\sum_{i\in I}(x_i+y_i)\,;$$

d'où $\sum_{i\in I}x_i+\sum_{i\in I}y_i\leqslant\sum_{i\in I}(x_i+y_i)$ par "double passage au sup" (d'abord en F, puis en F' ; ou le contraire). $\square$

Voici maintenant un résultat qui nous sera extrêmement utile.

Proposition 3.2. (convergence monotone pour les sommes)
Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions positives, $f_n : I \to [0,\infty]$. On suppose que la suite (f_n) est croissante *(i.e. $f_n(i) \leqslant f_{n+1}(i)$ pour tout n et pour tout $i \in I$), et on pose $f(i) := \lim_{n\to\infty} f_n(i)$ (qui existe dans $[0,\infty]$). Alors*

$$\sum_{i\in I}f(i)=\lim_{n\to\infty}\sum_{i\in I}f_n(i)\,.$$

Autrement dit : pour une suite croissante *de fonctions* positives*, on a toujours le droit d'écrire*

$$\sum_{i\in I}\lim_{n\to\infty}f_n(i)=\lim_{n\to\infty}\sum_{i\in I}f_n(i)\,.$$

Démonstration. Notons d'abord que la limite $\lim_{n\to\infty}\sum_{i\in I}f_n(i)$ existe dans $[0,\infty]$ car $u_n := \sum_{i\in I}f_n(i)$ est une suite croissante de nombres positifs.

Comme la suite (f_n) croit vers f, on a $f_n \leqslant f$ et donc

$$\sum_{i\in I}f_n(i)\leqslant\sum_{i\in I}f(i)\qquad\text{pour tout } n\in\mathbb{N}\,;$$

d'où $\lim_{n\to\infty}\sum_{i\in I}f_n(i)\leqslant\sum_{i\in I}f(i)$.

Inversement, on a pour tout ensemble fini $F \subseteq I$:

$$\sum_{i\in F}f(i)=\lim_{n\to\infty}\sum_{i\in F}f_n(i)\leqslant\lim_{n\to\infty}\sum_{i\in I}f_n(i)\,;$$

d'où $\sum_{i\in I}f(i)\leqslant\lim_{n\to\infty}\sum_{i\in I}f_n(i)$ en prenant le sup en F. $\square$

Corollaire 3.3. (interversion série-somme)
Si $(x_{i,k})_{(i,k)\in I\times\mathbb{N}}$ est une famille double de nombres positifs indexée par $I \times \mathbb{N}$, alors

$$\sum_{i\in I}\left(\sum_{k=0}^{\infty}x_{i,k}\right)=\sum_{k=0}^{\infty}\left(\sum_{i\in I}x_{i,k}\right).$$

Démonstration. On applique la proposition aux fonctions $f_n : I \to [0,\infty]$ définies par $f_n(i) = \sum_{k=0}^{n} x_{i,k}$. La suite (f_n) est croissante car les $x_{i,k}$ sont positifs, et $f_n(i) \to$

$f(i) := \sum_{k=0}^{\infty} x_{i,k}$ pour tout $i \in I$. Donc

$$\begin{aligned}
\sum_{i\in I}\left(\sum_{k=0}^{\infty} x_{i,k}\right) &= \lim_{n\to\infty}\sum_{i\in I}\left(\sum_{k=0}^{n} x_{i,k}\right) && \text{par convergence monotone}\\
&= \lim_{n\to\infty}\sum_{k=0}^{n}\left(\sum_{i\in I} x_{i,k}\right) && \text{par additivité}\\
&= \sum_{k=0}^{\infty}\left(\sum_{i\in I} x_{i,k}\right).
\end{aligned}$$

□

Exercice 1. Pour $s > 1$, on pose $\zeta(s) := \sum\limits_{k=1}^{\infty} \frac{1}{k^s}$. Déterminer $\lim_{s\to 1^+} \zeta(s)$.

Exercice 2. Montrer que si (f_n) est une suite de fonctions positives sur un ensemble I, alors

$$\sum_{i\in I} \underline{\lim}\, f_n(i) \leqslant \underline{\lim} \sum_{i\in I} f_n(i)\,.$$

Exercice 3. Montrer que si $(x_\lambda)_{\lambda\in\Lambda}$ est une famille de nombres positifs et si $(\Lambda_i)_{i\in I}$ est une partition de Λ, alors

$$\sum_{\lambda\in\Lambda} x_\lambda = \sum_{i\in I}\left(\sum_{\lambda\in\Lambda_i} x_\lambda\right).$$

En déduire que si $(x_{i,j})_{(i,j)\in I\times J}$ est une "famille double" de nombres positifs, alors

$$\sum_{i\in I}\left(\sum_{j\in J} x_{i,j}\right) = \sum_{(i,j)\in I\times J} x_{i,j} = \sum_{i\in J}\left(\sum_{i\in I} x_{i,j}\right).$$

4. Notations et terminologie "ensemblistes"

• Pour tout ensemble Ω, on notera $\mathcal{P}(\Omega)$ l'ensemble de toutes les parties de Ω. Donc, "$A \in \mathcal{P}(\Omega)$" est synonyme de "$A \subseteq \Omega$".

• Si $A \in \mathcal{P}(\Omega)$, on note A^c ou $\Omega\backslash A$ le complémentaire de A dans Ω.

• Si $A, B \subseteq \Omega$, on pose $B\backslash A := B \cap A^c = B \cap (\Omega\backslash A)$. Cette notation ne nécessite pas d'avoir $A \subseteq B$.

• Si $(A_i)_{i\in I}$ est une famille de parties de Ω, on note $\bigcup_{i\in I} A_i$ la réunion de tous les A_i et $\bigcap_{i\in I} A_i$ leur intersection. Lorsque $I = \mathbb{N}$, on écrit aussi $\bigcup_{i=0}^{\infty} A_i$ et $\bigcap_{i=0}^{\infty} A_i$.

Il est *très important* d'avoir parfaitement compris que le symbole de réunion $\bigcup$ correspond au quantificateur "existentiel" $\exists$, et que le symbole d'intersection $\bigcap$ correspond au quantificateur "universel" $\forall$:

$$x \in \bigcup_{i\in I} A_i \iff \exists i \in I \;:\; x \in A_i\,,$$

$$x \in \bigcap_{i\in I} A_i \iff \forall i \in I \;:\; x \in A_i\,.$$

Exercice. Soit (f_n) une suite de fonctions, $f_n : \Omega \to \mathbb{R}$. Écrire avec des unions et des intersections l'ensemble $A := \{x \in \Omega;\ f_n(x) \to 0 \text{ quand } n \to \infty\}$.

• Si $f : \Omega \to \Omega'$ est une application d'un ensemble Ω dans un ensemble Ω', on pose pour tout $A \subseteq \Omega'$:

$$f^{-1}(A) := \{x \in \Omega;\ f(x) \in A\}\,.$$

Cette notation ne nécessite pas que l'application f soit bijective. On dit que $f^{-1}(A)$ est l'**image réciproque** de A par l'application f.

Il est important de se souvenir que, contrairement aux images directes, les images réciproques se comportent parfaitement bien par rapport aux opérations ensemblistes :

$$\begin{aligned} f^{-1}(A^c) &= f^{-1}(A)^c\,,\\ f^{-1}\left(\bigcup_{i\in I} A_i\right) &= \bigcup_{i\in I} f^{-1}(A_i)\,,\\ f^{-1}\left(\bigcap_{i\in I} A_i\right) &= \bigcap_{i\in I} f^{-1}(A_i)\,. \end{aligned}$$

• Soit Ω un ensemble. Pour tout ensemble $A \subseteq \Omega$, on note $\mathbf{1}_A : \Omega \to \{0,1\}$ la **fonction indicatrice de** A, qui est définie par

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

Remarque. L'application $A \mapsto \mathbf{1}_A$ est une bijection de $\mathcal{P}(\Omega)$ sur $\{0,1\}^\Omega$, l'ensemble de toutes les applications de Ω dans $\{0,1\}$. En particulier, $\mathcal{P}(\mathbb{N})$ est en bijection avec $\{0,1\}^{\mathbb{N}}$.

Exercice 1. Montrer que si Ω un ensemble quelconque, alors il n'existe pas de surjection de Ω sur $\mathcal{P}(\Omega)$. (Pour toute application $s = \Omega \to \mathcal{P}(\Omega)$, on pourra considérer l'ensemble $A := \{x \in \Omega;\ x \notin s(x)\}$.) En déduire une preuve du fait que $\{0,1\}^{\mathbb{N}}$ n'est pas dénombrable ; puis montrer que cette preuve est en fait exactement la même que celle utilisant la méthode de la diagonale.

Exercice 2. Soit Ω un ensemble. Montrer que si $A, B \subseteq \Omega$, alors

$$\begin{aligned} \mathbf{1}_{A\cap B} &= \mathbf{1}_A\,\mathbf{1}_B\,,\\ \mathbf{1}_{A\cup B} &= \mathbf{1}_A + \mathbf{1}_B - \mathbf{1}_A\mathbf{1}_B\,. \end{aligned}$$

Montrer également que si on pose $A\Delta B := (A\backslash B) \cup (B\backslash A)$, alors

$$\mathbf{1}_{A\Delta B} = |\mathbf{1}_A - \mathbf{1}_B|\,.$$

Exercice 3. Soit Ω un ensemble, et soit (A_n) une suite de parties de Ω. On pose

$$\overline{\lim}\, A_n := \bigcap_{m\in\mathbb{N}} \bigcup_{n\geqslant m} A_n \quad \text{et} \quad \overline{\lim}\, A_n := \bigcup_{m\in\mathbb{N}} \bigcap_{n\geqslant m} A_n\,.$$

Montrer qu'on a

$$\mathbf{1}_{\overline{\lim}\, A_n} = \overline{\lim}\,\mathbf{1}_{A_n} \quad \text{et} \quad \mathbf{1}_{\underline{\lim}\, A_n} = \underline{\lim}\,\mathbf{1}_{A_n}\,.$$

Exercice 4. Soit Ω un ensemble fini, et soient $A_1, \dots, A_N$ des parties de Ω. Exprimer la fonction $\mathbf{1}_{\Omega\backslash(A_1\cup\cdots\cup A_N)}$ à l'aide des fonctions $\mathbf{1}_{A_i}$, et en déduire la formule suivante (en observant que $\#B = \sum_{x\in\Omega} \mathbf{1}_B(x)$ pour tout $B \subseteq \Omega$) :

$$\#(A_1 \cup \cdots \cup A_N) = \sum_{k=1}^{N} (-1)^{k-1} \sum_{I\in\binom{N}{k}} \#A_I\,,$$

où $\binom{N}{k}$ est l'ensemble de toutes les parties de $\{1, \ldots, N\}$ à k éléments et $A_I = \bigcap_{i \in I} A_i$.

Chapitre 2

Tribus et mesures

1. Pourquoi ?

Ce chapitre étant assez abstrait, il n'est peut-être pas inutile d'essayer de "motiver" les notions qui vont y être introduites. On va le faire en expliquant que l'on tombe très vite sur d'assez gros ennuis lorsqu'on essaye de parler un peu sérieusement de la *mesure des aires.*

1.1. Les aires, ce n'est pas si simple. Tout le monde a une idée intuitive de ce qu'est l'aire d'une partie raisonnablement simple du plan, disons un polygône ou un disque ; et au vu de cette "familiarité sensible", il ne semble a priori pas très choquant de parler de l'aire d'une partie *quelconque* du plan.

Mais en y réfléchissant un peu, les choses ne sont pas si simples. Pour une partie *vraiment* quelconque du plan, on ne voit *a priori* pas très bien quel sens donner au mot "aire".

On peut essayer de s'en tirer "axiomatiquement", c'est à dire en faisant une liste de propriétés que devrait "évidemment" vérifier l'aire, quelle que soit la définition qu'on en donnerait. Quoi qu'il arrive, l'aire d'un ensemble $A \subseteq \mathbb{R}^2$ doit être un nombre positif (éventuellement égal à ∞, par exemple si $A = \mathbb{R}^2$). Ensuite, une liste raisonnable de propriétés pourrait être la suivante :

(i) L'ensemble vide "n'a pas d'aire" : $\mathrm{aire}(\emptyset) = 0$.
(ii) *Les aires s'ajoutent* : si (A_k) est une suite finie ou infinie de parties du plan deux à deux disjointes, alors $\mathrm{aire}\left(\bigcup_k A_k\right) = \sum_k \mathrm{aire}(A_k)$.
(iii) L'aire est *invariante par déplacements* : si deux ensembles A et A' se déduisent l'un de l'autre par une translation ou une rotation, alors $\mathrm{aire}(A) = \mathrm{aire}(A')$.
(iv) Pour tout rectangle non trivial $R \subseteq \mathbb{R}^2$, on a $0 < \mathrm{aire}(R) < \infty$.

Se pose alors le problème suivant : comment peut-on définir mathématiquement l'aire de sorte que ces propriétés soient vérifiées ? Et c'est là que les ennuis commencent :

Proposition 1.1. *Avec les règles les plus usuelles du jeu des mathématiques, il n'existe pas d'application* $A \mapsto \mathrm{aire}(A)$ *définie sur l'ensemble de toutes les parties de* $\mathbb{R}^2$ *et vérifiant les propriétés* (i) *à* (iv).

Autrement dit, les propriétés (i) à (iv) empêchent d'attribuer une aire à *toutes* les parties du plan : quelle que soit la définition choisie, il y aura toujours des ensembles suffisamment biscornus pour faire que ces 4 propriétés se contredisent entre elles.

Preuve de la Proposition 1.1. Précisons d'abord un peu ce qu'on entend par "les règles les plus usuelles du jeu des mathématiques" : il s'agit de quelques postulats de base, qu'on appelle les **axiomes de la théorie des ensembles**, dont la plupart des mathématiciens professionnels pensent qu'il est vraiment très problématique de les remettre en cause si on veut faire des mathématiques. Autrement dit, des règles avec lesquelles une grande majorité de mathématiciens acceptent de jouer au jeu des mathématiques, parce qu'ils considèrent que sans elles, ils ne pourraient rien faire du tout ; et qui suffisent à construire toutes les mathématiques. (En réalité, les choses sont un peu plus compliquées ; *cf* la Remarque 3 plus bas.)

Supposons qu'il existe une application $A \mapsto \text{aire}(A)$ définie sur toutes les parties de $\mathbb{R}^2$, à valeurs dans $[0, \infty]$ et vérifiant les propriétés (i) à (iv). Il s'agit d'obtenir une contradiction.

Pour que les notations soient les moins lourdes possibles, il va être commode d'identifier $\mathbb{R}^2$ à $\mathbb{C}$.

Notons Ω le "disque unité" de $\mathbb{C} = \mathbb{R}^2$ privé de 0 :

$$\Omega := \{u \in \mathbb{C};\ 0 < |u| < 1\} .$$

Posons également $\mathbb{T} := \{\gamma \in \mathbb{C};\ |\gamma| = 1\}$. Géométriquement, $\mathbb{T}$ s'identifie au groupe des rotations de $\mathbb{R}^2 = \mathbb{C}$ centrées en 0 : un nombre $\gamma = e^{i\theta} \in \mathbb{T}$ s'identifie à la rotation r_γ d'angle θ. Pour $u \in \mathbb{C}$, on écrira $\gamma \cdot u$ au lieu de $r_\gamma(u)$ (autrement dit, $\gamma \cdot u = \gamma u\,!$) ; et pour tout ensemble $E \subseteq \Omega$, on pose $\gamma \cdot E := \{\gamma \cdot u;\ u \in E\}$.

Comme les rotations préservent les longueurs, Ω est stable par n'importe quelle rotation. Autrement dit, le groupe $\mathbb{T}$ "agit" sur Ω.

Soit maintenant $\Gamma \subseteq \mathbb{T}$ l'ensemble des "rotations rationnelles" :

$$\Gamma := \{e^{i\theta};\ \theta \in 2\pi\mathbb{Q}\} .$$

Visiblement, Γ est un sous-groupe de $\mathbb{T}$. Donc on définit une *relation d'équivalence* sur Ω en décrétant que

$$u \sim v \quad \text{si et seulement si} \quad \exists \gamma \in \Gamma \ :\ \gamma \cdot u = v .$$

Notons $\mathfrak{C}$ l'ensemble des classes d'équivalence pour la relation $\sim$. Voici le point clé : grâce à une des règles du jeu des mathématiques, il est possible de **choisir** un point u_C dans chaque classe d'équivalence $C \in \mathfrak{C}$. Si on pose $E := \{u_C;\ C \in \mathfrak{C}\}$, alors E rencontre chaque classe d'équivalence $C \in \mathfrak{C}$ en exactement un point (à savoir u_C). On en déduit que E possède les propriétés suivantes :

(a) $\bigcup_{\gamma\in\Gamma} \gamma \cdot E = \Omega$;

(b) les ensembles $\gamma \cdot E$, $\gamma \in \Gamma$ sont deux à deux disjoints.

(La propriété (a) est évidente puisque E rencontre chaque classe d'équivalence. Pour (b), supposons que $\gamma \cdot E \cap \gamma' \cdot E \neq \varnothing$ avec $\gamma, \gamma' \in \Gamma$. Soient $u, u' \in E$ tels que $\gamma \cdot u = \gamma \cdot u'$. Alors $u \sim u'$, donc $u = u'$ car E rencontre chaque classe en *au plus* 1 point. Ainsi $\gamma u = \gamma' u$, et donc $\gamma = \gamma'$ car $u \neq 0$.)

Comme Γ est **dénombrable**, on doit avoir par (a), (b) et (ii) :

$$\text{aire}(\Omega) = \sum_{\gamma\in\Gamma} \text{aire}(\gamma \cdot E) .$$

Mais par (iii), on a aussi $\text{aire}(\gamma \cdot E) = \text{aire}(E)$ pour tout $\gamma \in \Gamma$. Comme Γ est **infini**, on en déduit que la somme $\sum_{\gamma\in\Gamma} \text{aire}(\gamma \cdot E)$ vaut ou bien 0 (si $\text{aire}(E) = 0$), ou bien ∞ (si $\text{aire}(E) > 0$). Ainsi, $\text{aire}(\Omega) = 0$ ou ∞. Pour conclure, on applique alors le fait suivant.

FAIT. Si $A \subseteq \mathbb{R}^2$ est *borné*, alors $\text{aire}(A) < \infty$; et si $A \subseteq \mathbb{R}^2$ est *d'intérieur non vide*, alors $\text{aire}(A) > 0$.

Preuve du Fait. Il découle de (ii) que l'application $A \mapsto \text{aire}(A)$ est *croissante* : si $A \subseteq A'$, alors $\text{aire}(A) \leqslant \text{aire}(A')$. (En effet A est la réunion disjointe de A' et de $A \backslash A'$, et donc $\text{aire}(A) = \text{aire}(A') + \text{aire}(A \backslash A') \geqslant \text{aire}(A')$.) Par (iv), on en déduit immédiatement le résultat : si $A \subseteq \mathbb{R}^2$ est borné, alors A est contenu dans un rectangle R, donc $\text{aire}(A) \leqslant \text{aire}(R) < \infty$; et si A est d'intérieur non vide, alors A contient un rectangle R non trivial, donc $\text{aire}(A) \geqslant \text{aire}(R) > 0$. □

Par le Fait, on a $0 < \text{aire}(\Omega) < \infty$, ce qui est la contradiction recherchée. □

Remarque 1. La contradiction est venue du fait que l'ensemble E *admet une aire*, puisqu'on a supposé qu'on pouvait attribuer une aire à *toutes* les parties du plan, y compris les plus bizarres.

Remarque 2. La règle du jeu autorisant à choisir un point dans chaque classe d'équivalence porte le nom d'**axiome du choix**. La morale de la démonstration précédente est que l'axiome du choix permet de "construire" des ensembles vraiment très étranges.

Remarque 3. Certains mathématiciens considèrent que l'axiome du choix n'est pas acceptable car il génère trop de pathologies.

Cependant, Gödel a montré en 1936 que si les mathématiques sans axiome du choix ne sont pas logiquement contradictoires, alors les mathématiques avec axiome du choix ne le sont pas non plus. (Bien entendu, personne ne sait si les mathématiques sont effectivement non contradictoires...) Par conséquent, si on accepte les axiomes "consensuels" de la théorie des ensembles, alors on ne pourra jamais montrer avec les règles de logique usuelles qu'on peut attribuer une aire à toutes les parties du plan, à moins de rajouter un axiome supplémentaire (qui sera forcément incompatible avec l'axiome du choix).

De façon tout à fait remarquable, ceci est essentiellement possible : Solovay a montré en 1970 que modulo un axiome supplémentaire stipulant l'existence d'ensembles "vraiment très gros", axiome considéré comme très plausible par les spécialistes de théorie des ensembles, il n'est pas contradictoire de supposer qu'*on peut* attribuer une aire à toutes les parties du plan. On peut même conserver une version faible de l'axiome du choix, qu'on appelle l'**axiome du choix dénombrable** (cet axiome autorise à effectuer simultanément une infinité de choix, pourvu que cette infinité reste dénombrable).

Remarque 4. On pourrait arguer que la propriété d'"additivité dénombrable" (ii) est trop forte, et qu'on pourrait se contenter d'exiger l'additivité *finie* : si $A_0, \dots, A_n$ sont deux à deux disjoints, alors $\text{aire}(A_0 \cup \cdots \cup A_n) = \text{aire}(A_0) + \cdots + \text{aire}(A_n)$. En notant (ii') la propriété d'additivité finie, il se passe alors une chose très étrange. D'une part, Banach a montré (en utilisant l'axiome du choix !) qu'il *est possible* d'attribuer une aire à toutes les parties du plan de façon à ce que (i), (ii'), (iii) et (iv) soient satisfaites ; mais d'autres part, Hausdorff a montré que ceci est impossible en dimension 3 : on ne peut pas attribuer un *volume* à toutes les parties de $\mathbb{R}^3$ de sorte que la fonction volume soit finiment additive et invariante par rotations, et que le volume du cube unité soit fini et non nul. (Pour en savoir plus, le "mot clé" est **paradoxe de Banach-Tarski**.)

1.2. Pour sortir de l'impasse. Comme il a été dit, la Proposition 1.1 signifie que les propriétés (i) à (iv) des aires empêchent d'attribuer une aire à *toutes* les parties du plan.

Pour autant, ce résultat ne doit pas être considéré comme une catastrophe absolue : la Proposition 1.1 n'exclut pas que l'on puisse définir l'aire de parties éventuellement très compliquées du plan, mais dit simplement que si on veut éviter de gros ennuis, il faut accepter que certains ensembles n'aient pas d'aire. Autrement dit, on doit *restreindre* la famille $\mathcal{B}$ des parties du plan auxquelles on peut attribuer une aire.

Cependant, au vu de (ii) il serait bon que $\mathcal{B}$ soit *stable par réunions dénombrables*, autrement dit que pour toute suite (A_k) d'éléments de $\mathcal{B}$, l'ensemble $\bigcup_k A_k$ appartienne encore à $\mathcal{B}$. De plus, la propriété (ii) suggère aussi que si $B \subseteq \mathbb{R}^2$ admet une aire *finie* et si $A \subseteq B$ admet une aire, alors $B \backslash A$ doit admettre une aire, égale à aire(B) − aire(A). Ainsi, il est raisonnable d'imposer que $B \backslash A \in \mathcal{B}$ dès que $A \subseteq B$ sont dans $\mathcal{B}$ et aire$(B) < \infty$. Mais si on veut lister *a priori* les propriétés de la famille $\mathcal{B}$ (sans faire référence à l'aire), la condition aire$(B) < \infty$ n'a pas lieu d'être ; donc on "doit" imposer que $B \backslash A \in \mathcal{B}$ pour tous $A, B \in \mathcal{B}$ vérifiant $A \subseteq B$. Comme le plan $\mathbb{R}^2$ admet évidemment une aire (égale à ∞), *i.e.* $\mathbb{R}^2 \in \mathcal{B}$, on voit que la famile $\mathcal{B}$ doit être *stable par complémentation* : si $A \in \mathcal{B}$, alors $A^c \in \mathcal{B}$.

Ces considérations nous amènent à la notion de *tribu*, que l'on peut définir de façon complètement abstraite : étant donné un ensemble $\Omega \neq \varnothing$, une tribu de parties de Ω est une famille de parties de Ω stable par réunions dénombrables et par complémentation. Cela étant précisé, on peut définir, également de façon abstraite, la notion de *mesure* : une mesure sur une tribu $\mathcal{B} \subseteq \mathcal{P}(\Omega)$ est une fonction associant à tout ensemble $A \in \mathcal{B}$ un nombre positif $\mu(A)$, de sorte que les propriétés (i) et (ii) des aires soient satisfaites. Ce sont les deux définitions de base de ce chapitre.

2. Définitions

2.1. Tribus de parties d'un ensemble.

DÉFINITION 2.1. *Soit Ω un ensemble non vide, et soit $\mathcal{B}$ une famille non vide de parties de Ω. On dit que la famille $\mathcal{B}$ est une* ***tribu*** *(de parties de Ω), ou encore une* ***σ-algèbre****, si elle vérifie les propriétés de* stabilité *suivantes.*

- $(\neg)$ *$\mathcal{B}$ est* ***stable par complémentation*** *: si $E \in \mathcal{B}$ alors $E^c \in \mathcal{B}$.*
- (σ) *$\mathcal{B}$ est* ***stable par réunions dénombrables*** *: pour toute famille dénombrable $(E_i)_{i \in I}$ d'éléments de $\mathcal{B}$ (i.e. l'ensemble d'indices I est dénombrable), l'ensemble $\bigcup_{i \in I} E_i$ appartient à $\mathcal{B}$.*

Remarque 1. Toute tribu $\mathcal{B} \subseteq \mathcal{P}(\Omega)$ doit contenir $\varnothing$ et Ω. En effet, soit $E_0 \in \mathcal{B}$ quelconque. Alors $\Omega = E_0 \cup E_0^c$, donc $\Omega \in \mathcal{B}$ par (i) et (ii) ; et donc $\varnothing = \Omega^c \in \mathcal{B}$.

Remarque 2. En raison de $(\neg)$, on peut remplacer la condition (σ) par

- (δ) $\mathcal{B}$ est stable par *intersections* dénombrables.

Remarque 3. On peut énoncer (σ) sous la forme suivante : pour toute *suite* $(E_n)_{n \in \mathbb{N}}$ d'éléments de $\mathcal{B}$, on a $\bigcup_{n \in \mathbb{N}} E_n \in \mathcal{B}$. En effet, cette formulation est équivalente à "(σ) avec un ensemble d'indices I infini dénombrable" ; mais on peut toujours compléter une famille finie $(E_0, \ldots, E_n)$ d'éléments de $\mathcal{B}$ en une suite infinie ayant la même réunion, par exemple en posant $E_i = E_0$ pour $i > n$.

Exemple 1. $\mathcal{B} := \mathcal{P}(\Omega)$ est une tribu, qui est visiblement la *plus grande* tribu de parties de Ω.

Exemple 2. $\mathcal{B} := \{\varnothing, \Omega\}$ est une tribu. C'est la *plus petite* tribu de parties de Ω.

Exemple 3. Supposons que Ω soit un espace métrique, et notons $\mathcal{B}$ la famille de tous les ouverts de Ω. Alors $\mathcal{B}$ vérifie (σ), mais ce n'est a priori pas une tribu car $(\neg)$ n'a aucune raison d'être vérifiée : le complémentaire d'un ouvert n'est en général pas un ouvert.

Exemple 4. Prenons $\Omega = \mathbb{R}$, et soit $\mathcal{B}$ la famille de toutes les parties A de $\mathbb{R}$ qui sont réunions finies d'intervalles deux à deux disjoints. Alors $\mathcal{B}$ vérifie $(\neg)$ (faire un dessin) ; mais $\mathcal{B}$ ne vérifie pas (σ) et n'est donc pas une tribu : si par exemple on pose $A_n := [3n, 3n+1]$ pour $n \in \mathbb{N}$, alors les A_n sont dans $\mathcal{B}$ mais $A := \bigcup_{n\in\mathbb{N}} A_n$ n'est pas une réunion finie d'intervalles (exercice).

DÉFINITION 2.2. *Un* ***espace mesurable*** *est une paire* $(\Omega, \mathcal{B})$, *où* Ω *est un ensemble non vide et* $\mathcal{B}$ *une tribu de parties de* Ω.

Remarque. Si l'espace mesurable $(\Omega, \mathcal{B})$ est clairement spécifié par le contexte, on appelera **ensemble mesurable** tout élément de la tribu $\mathcal{B}$.

2.2. Mesures : définition et exemples.

DÉFINITION 2.3. *Soit* $(\Omega, \mathcal{B})$ *un espace mesurable. Une* ***mesure sur*** $(\Omega, \mathcal{B})$ *est une application* $\mu : \mathcal{B} \to [0, \infty]$ *vérifiant les propriétés suivantes.*

(i) ***Non trivialité*** *:* $\mu(\varnothing) = 0$.

(ii) ***Additivité dénombrable*** *: pour toute suite* $(A_k)_{k\in\mathbb{N}}$ *d'éléments de* $\mathcal{B}$ deux à deux disjoints, *on a*

$$\mu\left(\bigcup_{k=0}^{\infty} A_k\right) = \sum_{k=0}^{\infty} \mu(A_k)\,.$$

Il est essentiel de bien saisir tout de suite ce dont on parle : une mesure μ est par définition une fonction qui à un *ensemble* $A \subseteq \Omega$ (supposé mesurable) associe un *nombre positif* $\mu(A)$, éventuellement égal à ∞. (De façon générale, une fonction qui à un ensemble associe un nombre s'appelle une **fonction d'ensembles**.)

Remarque 1. Les propriétés (i) et (ii) entrainent l'**additivité finie** : si $A_0, \dots, A_n \in \mathcal{B}$ sont deux à deux disjoints, alors $\mu(A_0 \cup \cdots \cup A_n) = \mu(A_0) + \cdots + \mu(A_n)$.

Démonstration. Il suffit de compléter $(A_0, \dots, A_n)$ en une suite infinie $(A_k)_{k\in\mathbb{N}}$ en posant $A_k = \varnothing$ pour $k > n$, et d'appliquer (ii) en tenant compte du fait que $\mu(\varnothing) = 0$. □

Remarque 2. On aurait pu éviter la Remarque 1 en écrivant (ii) sous la forme suivante : *pour toute famille dénombrable* $(A_i)_{i\in I}$ *d'ensembles mesurables deux à deux disjoints, on a* $\mu\left(\bigcup_{i\in I} A_i\right) = \sum_{i\in I} \mu(A_i)$.

Exemple 1. Soit $a \in \Omega$. On définit une mesure δ_a sur $(\Omega, \mathcal{P}(\Omega))$ en posant pour tout $A \subseteq \Omega$:

$$\delta_a(A) = \begin{cases} 1 & \text{si } A \ni a, \\ 0 & \text{si } A \not\ni a. \end{cases}$$

La mesure δ_a s'appelle la **masse de Dirac au point** a.

Démonstration. On a par définition $\delta_a(A) = \mathbf{1}_A(a)$ pour tout $A \subseteq \Omega$. Donc il suffit d'observer que si (A_k) est une suite de parties de Ω deux à deux disjointes, alors

$$\mathbf{1}_{\bigcup_{k=0}^{\infty} A_k} = \sum_{k=0}^{\infty} \mathbf{1}_{A_k}\,.$$

□

Exemple 2. On définit une mesure μ_c sur sur $(\Omega, \mathcal{P}(\Omega))$ en posant pour tout $A \subseteq \Omega$:

$$\mu_c(A) = \begin{cases} \infty & \text{si } A \text{ est infini,} \\ \#A & \text{si } A \text{ est fini.} \end{cases}$$

La mesure μ_c s'appelle la **mesure de comptage** sur Ω.

Démonstration. Notons d'abord le fait suivant, dont la preuve est laissée en exercice.

Fait. On a $\mu_c(A) = \sum_{x\in\Omega} \delta_x(A)$ pour tout $A \subseteq \Omega$.

On a évidemment $\mu_c(\varnothing) = 0$. Si maintenant $(A_k)_{k\in\mathbb{N}}$ est une suite de parties de Ω deux à deux disjointes, alors

$$\begin{aligned} \mu_c\left(\bigcup_{k=0}^{\infty} A_k\right) &= \sum_{x\in\Omega} \delta_x\left(\bigcup_{k=0}^{\infty} A_k\right) && \text{par le Fait} \\ &= \sum_{x\in\Omega}\left(\sum_{k=0}^{\infty} \delta_x(A_k)\right) && \text{car les } \delta_a \text{ sont des mesures} \\ &= \sum_{k=0}^{\infty}\left(\sum_{x\in\Omega} \delta_x(A_k)\right) && \text{par interversion série-somme} \\ &= \sum_{k=0}^{\infty} \mu_c(A_k) && \text{à nouveau par le Fait.} \end{aligned}$$

□

Remarque. Si $(\mu_i)_{i\in I}$ est une famille quelconque de mesures sur $(\Omega, \mathcal{B})$, on définit une mesure sur $(\Omega, \mathcal{B})$ en posant pour tout $A \in \mathcal{B}$:

$$\mu(A) = \sum_{i\in I} \mu_i(A)\,.$$

(La preuve est identique à celle faite pour la mesure de comptage.) La mesure μ s'appelle la **somme** des mesures μ_i, et se note $\sum_{i\in I} \mu_i$. Par exemple, on a $\mu_c = \sum_{x\in\Omega} \delta_x$.

Exemple 3. Une **mesure discrète** sur Ω est une mesure sur $(\Omega, \mathcal{P}(\Omega))$ de la forme

$$\mu = \sum_{i\in I} p_i \delta_{x_i}\,,$$

où $(x_i)_{i\in I}$ est une famille dénombrable de points de Ω et $p_i \in [0, \infty]$. Si Ω est lui même dénombrable, alors *toute* mesure sur $(\Omega, \mathcal{P}(\Omega))$ est de ce type.

Démonstration. Si Ω est dénombrable et si μ est une mesure sur $(\Omega, \mathcal{P}(\Omega))$, alors on a par additivité dénombrable :

$$\mu(A) = \mu\left(\bigcup_{x\in A} \{x\}\right) = \sum_{x\in A} \mu(\{x\}) = \sum_{x\in\Omega} \mu(\{x\})\, \delta_x(A)$$

pour tout ensemble $A \subseteq \Omega$. Donc $\mu = \sum_{x\in\Omega} p_x\, \delta_x$, où $p_x = \mu(\{x\})$. □

Exemple 4. Soit X une "variable aléatoire" prenant ses valeurs dans un certain ensemble Ω. Admettons que pour certaines parties A de Ω, on sache déterminer la "probabilité" que X appartienne à A (à supposer que ceci ait un sens), que l'on note

$\mathbb{P}(X \in A)$. Si $\mathcal{B}$ est la famille de parties A de Ω pour lesquelles on sait déterminer $\mathbb{P}(X \in A)$, et si on pose $\mathbb{P}_X(A) := \mathbb{P}(X \in A)$ pour $A \in \mathcal{B}$, alors les “axiomes de la théorie des probabilités” postulent que $\mathcal{B}$ est une tribu de parties de Ω et que $\mathbb{P}_X$ est une mesure sur $(\Omega, \mathcal{B})$ vérifiant $\mathbb{P}_X(\Omega) = 1$. Une mesure de ce type s'appelle une **distribution de probabilité** sur Ω.

Exemple 5. Intuitivement, on définit une “mesure sur $\mathbb{R}^2$” en posant pour $A \subseteq \mathbb{R}^2$:

$$\mu(A) = \text{aire}\,(A)\,.$$

On a vu plus haut que cette “définition” est assez problématique. Plus précisément, on doit faire face aux deux questions suivantes :

- Comment définit-on mathématiquement l'“aire” d'une partie du plan ?
- À quelles parties du plan peut-on attribuer une aire ? Autrement dit : qui sont les ensembles “mesurables”, i.e. quelle tribu $\mathcal{B}$ considérer ?

Ces deux questions sont bien entendu étroitement liées. On donnera une réponse sous forme de “boite noire” à la fin de ce chapitre, et on y reviendra en détail au Chapitre 10.

3. Propriétés élémentaires des mesures

3.1. Reformulation de la définition. Le lemme suivant montre que l'additivité dénombrable est en fait la conjonction de deux propriétés : l'additivité “finie”, et la “continuité par en dessous”.

LEMME 3.1. *Soit $(\Omega, \mathcal{B})$ un espace mesurable, et soit $\mu : \mathcal{B} \to [0, \infty]$. Alors μ est une mesure si et seulement si elle possède les propriétés suivantes.*

(i) ***Non trivialité*** *: $\mu(\varnothing) = 0$.*

(ii') ***Additivité finie*** *: si $A_0, \dots, A_n \in \mathcal{B}$ sont deux à deux disjoints, alors*

$$\mu(A_0 \cup \cdots \cup A_n) = \mu(A_0) + \cdots + \mu(A_n)\,.$$

(iii) ***Continuité par en dessous*** *: si $(B_n)_{n\in\mathbb{N}}$ est une suite* croissante *d'éléments de $\mathcal{B}$ (i.e. $B_n \subseteq B_{n+1}$ pour tout n), alors*

$$\mu\left(\bigcup_{n=0}^{\infty} B_n\right) = \lim_{n\to\infty} \mu(B_n)\,.$$

Démonstration. Supposons que μ soit une mesure. Alors (i) et (ii') sont vraies ; il s'agit de vérifier (iii). Soit $(B_n)_{n\in\mathbb{N}}$ une suite croissante d'éléments de $\mathcal{B}$. Posons $A_0 := B_0$, et

$$A_k := B_k \backslash (B_0 \cup \cdots \cup B_{k-1}) \quad \text{pour } k \geqslant 1\,.$$

Par définition, les A_k sont mesurables et deux à deux disjoints, et on a pour tout $n \in \mathbb{N}$:

$$A_0 \cup \cdots \cup A_n = B_0 \cup \cdots \cup B_n \tag{3.1}$$

$$= B_n \qquad \text{par croissance de la suite } (B_n). \tag{3.2}$$

Comme μ est une mesure, on obtient donc

$$\begin{aligned}\mu\left(\bigcup_{n=0}^{\infty} B_n\right) &= \sum_{k=0}^{\infty}\mu(A_k) \qquad \text{par } (3.1)\\ &= \lim_{n\to\infty}\sum_{k=0}^{n}\mu(A_k)\\ &= \lim_{n\to\infty}\mu(B_n) \qquad \text{par } (3.2)\,.\end{aligned}$$

Supposons maintenant que μ vérifie (i), (ii') et (iii). Soit $(A_k)_{k\in\mathbb{N}}$ une suite d'ensembles mesurables deux à deux disjoints. Pour $n \in \mathbb{N}$, posons

$$B_n := A_0 \cup \cdots \cup A_n\,.$$

Alors (B_n) est une suite croissante d'ensembles mesurables, et on a par (ii')

$$\mu(B_n) = \sum_{k=0}^{n}\mu(A_k)$$

pour tout $n \in \mathbb{N}$. Donc

$$\begin{aligned}\mu\left(\bigcup_{k=0}^{\infty} A_k\right) &= \mu\left(\bigcup_{n=0}^{\infty} B_n\right)\\ &= \lim_{n\to\infty}\mu(B_n) \qquad \text{par (iii)}\\ &= \sum_{k=0}^{\infty}\mu(A_k)\,;\end{aligned}$$

ce qui prouve que μ est une mesure. □

3.2. Autres propriétés importantes. Dans ce qui suit, $(\Omega, \mathcal{B})$ est un espace mesurable.

PROPOSITION 3.2. *Toute mesure μ sur $(\Omega, \mathcal{B})$ possède les propriétés suivantes.*

(1) ***Croissance*** *: si $A, B \in \mathcal{B}$ et $A \subseteq B$, alors $\mu(A) \leqslant \mu(B)$.*

(2) ***Propriété de valuation*** *: $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$ pour tous $A, B \in \mathcal{B}$.*

(3) ***Sous-additivité dénombrable*** *: pour toute famille dénombrable $(E_i)_{i\in I}$ d'ensembles mesurables (pas forcément disjoints), on a*

$$\mu\left(\bigcup_{i\in I} E_i\right) \leqslant \sum_{i\in I}\mu(E_i)\,.$$

Démonstration. (1) L'ensemble B est la réunion disjointe de A et de $B\backslash A$, donc $\mu(B) = \mu(A) + \mu(B\backslash A)$ par additivité finie, et donc $\mu(B) \geqslant \mu(A)$.

(2) On a $A = (A \cap B) \cup (A\backslash(A \cap B))$, $B = (A \cap B) \cup (B\backslash(A \cap B))$ et $A \cup B = (A\backslash(A \cap B)) \cup (A \cap B) \cup (B\backslash(A \cap B))$, où les réunions sont disjointes. Donc, par additivité finie :

$$\begin{aligned}\mu(A \cup B) + \mu(A \cap B) &= \mu(A\backslash(A \cap B)) + \mu(A \cap B) + \mu(B\backslash(A \cap B)) + \mu(A \cap B)\\ &= \mu(A) + \mu(B).\end{aligned}$$

(3) Par (2), on a $\mu(A \cup B) \leqslant \mu(A) + \mu(B)$ pour tous $A, B \in \mathcal{B}$; d'où, par une récurrence immédiate, $\mu\left(\bigcup_{i\in I} E_i\right) \leqslant \sum_{i\in I} \mu(E_i)$ pour toute famille *finie* d'ensembles mesurables E_i. Soit maintenant $(E_i)_{i\in\mathbb{N}}$ une suite infinie d'ensembles mesurables. D'après ce qui vient d'être dit, on a $\mu\left(\bigcup_{i=0}^n E_i\right) \leqslant \sum_{i=0}^n \mu(E_i)$ pour tout $n \in \mathbb{N}$; d'où, par la propriété de continuité par en dessous :

$$\mu\left(\bigcup_{i=0}^{\infty} E_i\right) = \lim_{n\to\infty} \mu\left(\bigcup_{i=0}^{n} E_i\right) \leqslant \sum_{i=0}^{\infty} \mu(E_i)\,.$$

□

Corollaire 3.3. *Si $A, B \in \mathcal{B}$ vérifient $\mu(A \cap B) = 0$, alors $\mu(A \cup B) = \mu(A) + \mu(B)$.*

Démonstration. C'est évident par (2). □

Proposition 3.4. *Soit μ une mesure sur $(\Omega, \mathcal{B})$. Si $A, B \in \mathcal{B}$ vérifient $A \subseteq B$* et si de plus *$\mu(A) < \infty$, alors $\mu(B \backslash A) = \mu(B) - \mu(A)$.*

Démonstration. Comme B est la réunion disjointe de A et de $B \backslash A$, on a $\mu(B) = \mu(A) + \mu(B\backslash A)$; et *on peut soustraire* $\mu(A)$ car $\mu(A) < \infty$. □

Proposition 3.5. *Soit μ une mesure sur $(\Omega, \mathcal{B})$. Si (C_n) est une suite* décroissante *d'ensembles mesurables* et si de plus *$\mu(C_n) < \infty$ pour au moins un $n \in \mathbb{N}$, alors*

$$\mu\left(\bigcap_{n=0}^{\infty} C_n\right) = \lim_{n\to\infty} \mu(C_n)\,.$$

Démonstration. Soit $n_0 \in \mathbb{N}$ tel que $\mu(C_{n_0}) < \infty$. Comme la suite d'ensembles (C_n) est décroissante, on a $\bigcap_{n\geqslant 0} C_n = \bigcap_{n\geqslant n_0} C_n$; donc on peut en fait supposer que $n_0 = 0$, i.e. $\mu(C_0) < \infty$. Soit $C := \bigcap_n C_n$.

Si on pose $B_n := C_0 \backslash C_n$, alors la suite (B_n) est croissante et $\bigcup_n B_n = C_0 \backslash C$. Par la propriété de continuité par en dessous, on a donc

$$\mu(C_0 \backslash C) = \lim_{n\to\infty} \mu(C_0 \backslash C_n)\,.$$

De plus, comme $\mu(C) \leqslant \mu(C_0) < \infty$ et $\mu(C_n) \leqslant \mu(C_0) < \infty$, on a le droit d'écrire $\mu(C_0\backslash C) = \mu(C_0) - \mu(C)$ et $\mu(C_0 \backslash C_n) = \mu(C_0) - \mu(C_n)$. On voit donc que $\mu(C_0) - \mu(C_n) \to \mu(C_0) - \mu(C)$, d'où le résultat en simplifiant par $\mu(C_0)$ (ce qui est possible car $\mu(C_0) < \infty$). □

Corollaire 3.6. *Si (C_n) est une suite décroissante d'ensembles mesurables telle que $\mu(C_0) < \infty$ et $\bigcap_0^{\infty} C_n = \varnothing$, alors $\mu(C_n) \to 0$ quand $n \to \infty$.*

Remarque. Ce résultat est faux si on ne suppose pas que les C_n sont de mesure finie. Par exemple, si on prend pour μ la mesure de comptage sur $\mathbb{N}$, alors les intervalles $[n, \infty[$ décroissent vers $\varnothing$ mais $\mu([n, \infty[) = \infty$ pour tout $n \in \mathbb{N}$.

4. Ensembles boréliens

4.1. Tribu engendrée par une famille d'ensemble. Si Ω est un ensemble non vide, il est *évident* que l'intersection d'une famille quelconque de tribus de parties de Ω est encore une tribu (vérifier). Par ailleurs, si $\mathcal{C}$ est une famille quelconque de parties de Ω, il y a toujours au moins une tribu de parties de Ω qui contient $\mathcal{C}$, à savoir $\mathcal{P}(\Omega)$. Cela justifie la définition suivante.

DÉFINITION 4.1. *Soit Ω un ensemble non vide, et soit $\mathcal{C}$ une famille de parties de Ω. La* ***tribu engendrée par*** *$\mathcal{C}$ est la tribu intersection de toutes les tribus de parties de Ω contenant $\mathcal{C}$; autrement dit, la* ***plus petite*** *tribu de parties de Ω contenant $\mathcal{C}$. Cette tribu se note $\sigma(\mathcal{C})$.*

Cette définition est très abstraite ; mais elle est extrêmement facile à utiliser. Le mode d'emploi est le suivant : par définition, la tribu $\sigma(\mathcal{C})$ est caractérisée par les 2 propriétés suivantes :

- $\sigma(\mathcal{C})$ est une tribu contenant $\mathcal{C}$;
- pour toute tribu $\mathcal{B} \subseteq \mathcal{P}(\Omega)$ contenant $\mathcal{C}$, on a $\sigma(\mathcal{C}) \subseteq \mathcal{B}$.

Par conséquent, si on veut montrer que *tous les éléments de* $\sigma(\mathcal{C})$ possèdent une certaine propriété (P), il suffit de vérifier

- que tout ensemble $C \in \mathcal{C}$ possède la propriété (P) ;
- que la famille de toutes les parties B de Ω possédant la propriété (P) est une tribu.

C'est *toujours* de cette façon qu'on utilise la définition.

De manière plus "mécanique", on peut aussi dire les choses comme suit : étant donné deux familles $\mathcal{C}_1$ et $\mathcal{C}_2$ de parties de Ω,

$$\text{si } \mathcal{C}_1 \subseteq \sigma(\mathcal{C}_2), \quad \text{alors } \sigma(\mathcal{C}_1) \subseteq \sigma(\mathcal{C}_2)\,.$$

Remarque. L'aspect réellement troublant de notion de tribu engendrée est que la définition de $\sigma(\mathcal{C})$ ne semble donner aucun moyen de *décrire explicitement* les éléments de $\sigma(\mathcal{C})$ à l'aide des éléments de $\mathcal{C}$, comme on peut le faire par exemple dans un espace vectoriel E pour les éléments de l'espace vectoriel engendré par une partie C. En fait, il est d'une certaine façon possible de donner une procédé "explicite" de construction des éléments de $\sigma(\mathcal{C})$ à partir des éléments de $\mathcal{C}$; mais décrire ce procédé nous entrainerait dans le monde des **ordinaux dénombrables** et de la **récurrence transfinie**, ce qui ne serait sans doute pas très raisonnable. On se contentera donc de considérer les éléments de $\sigma(\mathcal{C})$ "collectivement", et non "individuellement" ; avec le petit effort que cela nécessite pour faire taire une éventuelle gène.

4.2. Tribu borélienne. Si Ω est un espace métrique, la **tribu borélienne de** Ω est la tribu engendrée par les ouverts de Ω ; autrement dit, la plus petite tribu de parties de Ω contenant tous les ouverts de Ω. Cette tribu sera notée $\mathcal{B}(\Omega)$. En notant $\mathcal{O}(\Omega)$ la famille de tous les ouverts de Ω, on a ainsi

$$\mathcal{B}(\Omega) = \sigma\big(\mathcal{O}(\Omega)\big)\,.$$

Les éléments de la tribu $\mathcal{B}(\Omega)$ sont appelés les **boréliens** de l'espace métrique Ω.

Remarque. $\mathcal{B}(\Omega)$ est également la tribu engendrée par les *fermés.*

Démonstration. En notant $\mathcal{F}(\Omega)$ la famille de tous les fermés de Ω, on a $\mathcal{F}(\Omega) \subseteq \sigma\big(\mathcal{O}(\Omega)\big)$ car tout fermé $C \subseteq \Omega$ est le complémentaire d'un ouvert. Donc $\sigma\big(\mathcal{F}(\Omega)\big) \subseteq \sigma\big(\mathcal{O}(\Omega)\big) = \mathcal{B}(\Omega)$. De même, on a $\mathcal{O}(\Omega) \subseteq \sigma\big(\mathcal{F}(\Omega)\big)$ et donc $\mathcal{B}(\Omega) = \sigma\big(\mathcal{O}(\Omega)\big) \subseteq \sigma\big(\mathcal{F}(\Omega)\big)$. □

Exemples. Les ouverts et les fermés sont boréliens. Tout ensemble *dénombrable* est borélien, car c'est une réunion dénombrable de singletons (qui sont fermés).

Remarque. Une **mesure borélienne** sur Ω est une mesure sur $(\Omega, \mathcal{B}(\Omega))$.

4.3. Boréliens d'un sous-espace. Si Ω est un espace métrique, alors tout ensemble $E \subseteq \Omega$ est lui-même un espace métrique, et possède donc sa propre tribu borélienne $\mathcal{B}(E)$. La proposition suivante montre que les boréliens de E se décrivent très simplement à l'aide des boréliens du "gros" espace Ω.

PROPOSITION 4.2. *Soit Ω un espace métrique, et soit $E \subseteq \Omega$. Un ensemble $A \subseteq E$ est un borélien de E si et seulement si il est de la forme*

$$A = B \cap E \quad \textit{où } B \textit{ est un borélien de } \Omega\,.$$

Démonstration. La preuve est d'autant plus claire qu'on la rend la plus abstraite possible. Pour toute famille $\mathcal{C}$ de parties de Ω, posons

$$\mathcal{C}_E := \{C \cap E;\ C \in \mathcal{C}\} \subseteq \mathcal{P}(E)\,.$$

Avec cette notation, il s'agit de montrer qu'on a

$$\mathcal{B}(E) = \mathcal{B}(\Omega)_E\,.$$

Tout repose sur le fait suivant.

FAIT. Pour toute famille $\mathcal{C} \subseteq \mathcal{P}(\Omega)$, on a $\sigma(\mathcal{C}_E) = \sigma(\mathcal{C})_E$.

Preuve du Fait. On vérifie d'abord (exercice facile) que $\sigma(\mathcal{C})_E$ est une tribu de parties de E car $\sigma(\mathcal{C})$ est une tribu de parties de Ω. Comme $\mathcal{C}_E \subseteq \sigma(\mathcal{C})_E$ puisque $\mathcal{C} \subseteq \sigma(\mathcal{C})$, on a donc $\sigma(\mathcal{C}_E) \subseteq \sigma(\mathcal{C})_E$. Ensuite, on vérifie (autre exercice facile) que la famille $\mathcal{C}^E := \{A \subseteq \Omega;\ A \cap E \in \sigma(\mathcal{C}_E)\}$ est une tribu de parties de Ω car $\sigma(\mathcal{C}_E)$ est une tribu de parties de E. Comme $\mathcal{C} \subseteq \mathcal{C}^E$ par définition, on a donc $\sigma(\mathcal{C}) \subseteq \mathcal{C}^E$; autrement dit $\sigma(\mathcal{C})_E \subseteq \sigma(\mathcal{C}_E)$. □

Maintenant, notons $\mathcal{O}(E)$ la famille de tous les ouverts de E, et $\mathcal{O}(\Omega)$ la famille de tous les ouverts de Ω. On *sait* (propriété de la topologie induite) qu'on a

$$\mathcal{O}(E) = \{O \cap E;\ O \in \mathcal{O}(\Omega)\}\,;$$

autrement dit :

$$\mathcal{O}(E) = \mathcal{O}(\Omega)_E\,.$$

Par le Fait appliqué à $\mathcal{C} := \mathcal{O}(\Omega)$, on en déduit

$$\mathcal{B}(E) = \sigma(\mathcal{O}(E)) = \sigma(\mathcal{O}(\Omega))_E = \mathcal{B}(\Omega)_E\,.$$

□

COROLLAIRE 4.3. *Si $E \subseteq \Omega$ est borélien, alors un ensemble $A \subseteq E$ est borélien* dans E *si et seulement si il est borélien* dans Ω.

Remarque. Si $(\Omega, \mathcal{B})$ est un espace mesurable quelconque et si $E \subseteq \Omega$, alors la tribu

$$\mathcal{B}_E := \{B \cap E;\ B \in \mathcal{B}\} \subseteq \mathcal{P}(E)$$

s'appelle la **tribu trace** induite par $\mathcal{B}$ sur E.

4.4. Produits de boréliens.

Proposition 4.4. *Si $\Omega_1, \dots, \Omega_N$ sont des espaces métriques et si $A_1, \dots, A_N$ sont des boréliens de $\Omega_1, \dots, \Omega_N$ respectivement, alors $A_1 \times \cdots \times A_N$ est borélien dans $\Omega = \Omega_1 \times \cdots \times \Omega_N$.*

Démonstration. Par une récurrence immédiate, il suffit de le faire pour $N = 2$.
Si $A_1 \subseteq \Omega_1$ et $A_2 \subseteq \Omega_2$, alors

$$A_1 \times A_2 = (A_1 \times \Omega_2) \cap (\Omega_1 \times A_2)\,;$$

donc on est ramené à montrer que si A_1 et A_2 sont boréliens, alors $A_1 \times \Omega_2$ et $\Omega_1 \times A_2$ sont boréliens dans $\Omega_1 \times \Omega_2$.

C'est un "jeu de tribus" typique. Si on note $\mathcal{B}_1$ la famille de tous les ensembles $A \subseteq \Omega_1$ tels que $A \times \Omega_1 \in \mathcal{B}(\Omega_1 \times \Omega_2)$, il est très facile de vérifier que $\mathcal{B}_1$ *est une tribu* de parties de Ω_1. De plus $\mathcal{B}_1$ *contient tous les ouverts*, car si $O \subseteq \Omega_1$ est ouvert alors $O \times \Omega_2$ est ouvert dans $\Omega_1 \times \Omega_2$ et donc borélien. Donc $\mathcal{B}_1$ contient $\mathcal{B}(\Omega_1)$; autrement dit, $A_1 \times \Omega_2$ est borélien dans $\Omega_1 \times \Omega_2$ pour tout borélien $A_1 \subseteq \Omega_1$. On montre de même que $\Omega_1 \times A_2$ est borélien pour tout borélien $A_2 \subseteq \Omega_2$. □

4.5. Boréliens de $\mathbb{R}$, $\mathbb{R}^N$ et $\overline{\mathbb{R}}$. Pour la suite, il sera important de savoir que certains sous-ensembles très simples de $\mathbb{R}$, $\mathbb{R}^N$ et $\overline{\mathbb{R}}$ sont boréliens ; et aussi que les tribus boréliennes de ces espaces sont engendrées par d'autres familles d'ensembles que la famille des ouverts.

4.5.1. *Boréliens de $\mathbb{R}$.*

Proposition 4.5. *Tous les intervalles de $\mathbb{R}$ sont des boréliens*

Démonstration. Le résultat est clair pour les intervalles ouverts et les intervalles fermés, puisque ce sont des ouverts et des fermés. Il reste donc à traiter le cas d'un intervalle semi-ouvert, disons de la forme $[a, b[$ avec $-\infty < a < b \leqslant \infty$; mais le résultat est clair dans ce cas aussi car $[a, b[= [a, \infty[\cap]-\infty, b[$ est l'intersection d'un ouvert et d'un fermé. □

Proposition 4.6. *Les intervalles de $\mathbb{R}$ engendrent la tribu borélienne de $\mathbb{R}$. Plus précisément, $\mathcal{B}(\mathbb{R})$ est engendrée par l'une quelconque des familles suivantes :*

(i) *la famille des intervalles ouverts bornés ;*
(ii) *la famille des intervalles compacts ;*
(iii) *la famille des intervalles semi-ouverts de la forme $[\alpha, \beta[$, $\alpha, \beta \in \mathbb{R}$;*
(iii') *la famille des intervalles semi-ouverts de la forme $]\alpha, \beta]$, $\alpha, \beta \in \mathbb{R}$;*
(iv) *la famille des intervalles de la forme $]\alpha, \infty[$, $\alpha \in \mathbb{R}$;*
(v) *la famille des intervalles de la forme $[\alpha, \infty[$, $\alpha \in \mathbb{R}$;*
(iv') *la famille des intervalles de la forme $]-\infty, \beta[$, $\beta \in \mathbb{R}$;*
(v') *la famille des intervalles de la forme $]-\infty, \beta]$, $\beta \in \mathbb{R}$.*

Démonstration. (i) On sait que que tout ouvert de $\mathbb{R}$ est réunion dénombrable d'intervalles ouverts bornés (*cf* le Corollaire 1.7 du Chapitre 1). Donc la tribu engendrée par les intervalles ouverts bornés contient tous les ouverts, et donc est égale à la tribu borélienne.

(ii) Par (i), il suffit de montrer que tout intervalle ouvert borné $]a, b[$ appartient à la tribu engendrée par les intervalles compacts. Si on choisit une suite strictement décroissante (α_n) et une suite strictement croissante (β_n) telles que $\alpha_n \to a$ et $\beta_n \to b$, alors $]a,b[= \bigcup_n [\alpha_n, \beta_n]$; d'où le résultat.

(iii) Par (i), il suffit de montrer que tout intervalle ouvert borné $]a,b[$ appartient à la tribu engendrée par les intervalles semi-ouverts bornés $[\alpha,\beta[$; ce qui se voit en écrivant $]a,b[= \bigcup_n [\alpha_n, b[$, où est une suite strictement décroissante tendant vers a. La preuve de (iii') est identique.

(iv) Par (iii'), il suffit de montrer que tout intervalle semi-ouvert $]a,b]$ appartient à la tribu engendrée par les intervalles de la forme $]\alpha,\infty[$; ce qui est clair car $]a,b] =]a,\infty[\cap (\mathbb{R}\setminus]b,\infty[)$. Les preuves de (v), (iv') et (v') sont identiques.

□

4.5.2. *Boréliens de $\mathbb{R}^N$.* Dans $\mathbb{R}^N$, les ensembles qui vont jouer le rôle des intervalles seront les **pavés**, c'est à dire les ensembles $P \subseteq \mathbb{R}^N$ de la forme

$$P = I_1 \times \cdots \times I_N ,$$

où $I_1, \dots, I_N$ sont des intervalles bornés de $\mathbb{R}$. Un pavé $P = I_1 \times \cdots \times I_N$ sera dit **ouvert** si les intervalles I_j sont ouverts, **compact** si les I_j sont compacts, et **semi-ouvert** si les I_j sont semi-ouverts de la forme $[\alpha,\beta[$.

Ainsi, un pavé dans $\mathbb{R}$ est un intervalle borné ; un pavé dans $\mathbb{R}^2$ est un rectangle à côtés parallèles aux axes de coordonnées ; et un pavé dans $\mathbb{R}^3$ est un parallélépipède rectangle dont les arêtes sont parallèles aux axes de coordonnées.

Proposition 4.7. *Tous les pavés de $\mathbb{R}^N$ sont des boréliens.*

Démonstration. Ce sont des produits de boréliens de $\mathbb{R}$. □

Proposition 4.8. *Les pavés de $\mathbb{R}^N$ engendrent la tribu borélienne de $\mathbb{R}^N$. Plus précisément, $\mathcal{B}(\mathbb{R}^N)$ est engendrée par l'une quelconque des familles suivantes : les pavés ouverts ; les pavés compacts ; les pavés semi-ouverts.*

Démonstration. On sait (Corollaire 1.7 du Chapitre 1) que tout ouvert de $\mathbb{R}^N$ est réunion dénombrables de pavés ouverts ; donc la tribu engendrée par les pavés ouverts contient tous les ouverts, et est donc égale à $\mathcal{B}(\mathbb{R}^N)$.

Comme tout intervalle ouvert $I \subseteq \mathbb{R}$ est réunion dénombrable d'intervalles compacts J_n (*cf* la preuve de la Proposition 4.6), on voit que tout pavé ouvert $P = I_1 \times \cdots \times I_N$ est réunion dénombrable de pavés compacts : en écrivant

$$I_k = \bigcup_{n\in\mathbb{N}} J_{k,n} \qquad \text{pour } k = 1, \dots, N$$

où les $J_{k,n}$ sont des intervalles compacts, on a

$$P = I_1 \times \cdots \times I_N = \bigcup_{n_1,\dots,n_N \in \mathbb{N}} J_{1,n_1} \times \cdots \times J_{N,n_N} .$$

Donc la tribu engendrée par les pavés compacts contient tous les pavés ouverts, et est donc égale à $\mathcal{B}(\mathbb{R}^N)$. On montre de même que les pavés semi-ouverts engendrent $\mathcal{B}(\mathbb{R}^N)$. □

Remarque. On montre de la même façon que si Ω est un ouvert quelconque de $\mathbb{R}^d$, alors $\mathcal{B}(\Omega)$ est (par exemple) engendrée par les pavés compacts contenus dans Ω.

Exercice 4.9. Soient $\Omega_1, \dots, \Omega_N$ des espaces métriques *séparables*. Montrer que la tribu borélienne de $\Omega_1 \times \cdots \times \Omega_N$ est engendrée par les **produits d'ouverts**, c'est-à-dire les ensembles de la forme $A = O_1 \times \cdots \times O_N$, où O_k est ouvert dans Ω_k pour $k = 1, \dots, N$.

4.5.3. *Boréliens de* $\overline{\mathbb{R}}$. Il est très facile de décrire les boréliens de $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ à l'aide des boréliens de $\mathbb{R}$.

PROPOSITION 4.10. *Pour* $A \subseteq \overline{\mathbb{R}}$*, les propriétés suivante sont équivalentes :*

(1) A *est un borélien de* $\overline{\mathbb{R}}$ *;*

(2) $A \cap \mathbb{R}$ *est un borélien de* $\mathbb{R}$ *;*

(3) A *est de la forme* B *ou* $B \cup \{\infty\}$ *ou* $B \cup \{-\infty\}$ *ou* $B \cup \{-\infty, \infty\}$*, où* B *est un borélien de* $\mathbb{R}$*.*

Démonstration. L'implication (1) $\Longrightarrow$ (2) découle de la Proposition 4.2. L'implication (2) $\Longrightarrow$ (3) est immédiate. Et l'implication (3) $\Longrightarrow$ (1) vient du fait que tout borélien de $\mathbb{R}$ est borélien dans $\overline{\mathbb{R}}$ par la Proposition 4.2 (car $\mathbb{R}$ est ouvert dans $\overline{\mathbb{R}}$, donc borélien), et que $\{-\infty\}$ et $\{\infty\}$ sont boréliens dans $\overline{\mathbb{R}}$ (car fermés). $\square$

PROPOSITION 4.11. *La tribu borélienne de* $\overline{\mathbb{R}}$ *est engendrée par les intervalles de la forme* $]\alpha, \infty]$*,* $\alpha \in \mathbb{R}$*, et aussi par les intervalles de la forme* $[-\infty, \beta[$*,* $\beta \in \mathbb{R}$*.*

Démonstration. Soit $\mathcal{B}$ la tribu engendrée par les intervalles $]\alpha, \infty]$, $\alpha \in \mathbb{R}$. La tribu $\mathcal{B}$ contient $\{\infty\}$ car $\{\infty\} = \bigcap_{n \in \mathbb{N}}]n, \infty]$. Donc $\mathcal{B}$ contient tous les intervalles $]\alpha, \infty[$, $\alpha \in \mathbb{R}$ car $]\alpha, \infty[=]\alpha, \infty] \backslash \{\infty\}$; et par conséquent $\mathcal{B}$ contient tous les boréliens de $\mathbb{R}$ par la Proposition 4.6, car $\mathcal{B} \cap \mathcal{P}(\mathbb{R})$ est une tribu de parties de $\mathbb{R}$. Enfin, $\mathcal{B}$ contient également $\{-\infty\}$ car $\{-\infty\} = \bigcap_{n \in \mathbb{N}} (\overline{\mathbb{R}} \backslash]-n, \infty])$. Donc $\mathcal{B} = \mathcal{B}(\overline{\mathbb{R}})$ par la Proposition 4.10. $\square$

5. La mesure de Lebesgue sur $\mathbb{R}^N$

5.1. Un résultat d'unicité. Rappelons qu'une **mesure borélienne** sur un espace métrique Ω est par définition une mesure sur $(\Omega, \mathcal{B}(\Omega))$. Le lemme suivant signifie qu'une mesure borélienne sur un ouvert de $\mathbb{R}^N$ est entièrement déterminée par ses valeurs sur les pavés, pour peu que ces valeurs soient finies.

LEMME 5.1. *Soit* Ω *un ouvert de* $\mathbb{R}^N$*. Si* μ_1 *et* μ_2 *sont deux mesures boréliennes sur* Ω *telles que* $\mu_1(P) = \mu_2(P) < \infty$ *pour tout pavé compact* $P \subseteq \Omega$*, alors* $\mu_1 = \mu_2$*. La même conclusion vaut si on remplace "pavé compact" par "pavé ouvert" ou par "pavé semi-ouvert".*

On démontrera ce lemme dans un chapitre ultérieur, sous une forme plus générale. Pour le moment, on se contentera de dire que le résultat n'est pas très choquant "puisque les pavés engendrent la tribu borélienne".

Remarque. Le lemme devient faux si on ne suppose pas que les mesures μ_1 et μ_2 prennent des valeurs finies sur les pavés. Par exemple, soient μ_1 et μ_2 les mesures (boréliennes) sur $\mathbb{R}$ définies par

$$\mu_1(A) = \begin{cases} 0 & \text{si} \quad A = \varnothing, \\ \infty & \text{si} \quad A \neq \varnothing; \end{cases} \quad \text{et} \quad \mu_2(A) = \begin{cases} 0 & \text{si} \quad A \text{ est dénombrable}, \\ \infty & \text{si} \quad A \text{ n'est pas dénombrable}. \end{cases}$$

(Vérifier que ce sont effectivement des mesures...)

On a $\mu_1 \neq \mu_2$ car par exemple $\mu_1(\{0\}) = \infty$ et $\mu_2(\{0\}) = 0$; mais $\mu_1(I) = \mu(I)$ pour tout intervalle ouvert $I \subseteq \mathbb{R}$, car un intervalle ouvert non vide est non dénombrable.

Exercice. Soient μ_1 et μ_2 deux mesures boréliennes sur $\mathbb{R}$ telles que $\mu_1(I) = \mu_2(I)$ pour tout intervalle compact $I \subseteq \mathbb{R}$. Montrer qu'on a $\mu_1(J) = \mu_2(J)$ pour tout intervalle *ouvert* J, et en déduire que $\mu_1(O) = \mu_2(O)$ pour tout ouvert $O \subseteq \mathbb{R}$. (Utiliser une description "bien connue" des ouverts de $\mathbb{R}$.) Montrer que si on impose de plus $\mu_1(I), \mu_2(I) < \infty$ pour tout intervalle borné I, alors $\mu_1(C) = \mu_2(C)$ pour tout fermé $C \subseteq \mathbb{R}$. (Commencer par le cas où C est *compact*.)

5.2. Définition de la mesure de Lebesgue. Pour tout intervalle $I \subseteq \mathbb{R}$, on notera $|I|$ la longueur de I.

Si $P = I_1 \times \cdots \times I_N$ est un pavé de $\mathbb{R}^N$, on note $|P|$ le **volume** (N-dimensionnel) de P :

$$|P| := |I_1| \times \cdots \times |I_N|\,.$$

La remarque suivante est évidente, mais très importante.

Remarque. Si $P \subseteq \mathbb{R}^p$ et $Q \subseteq \mathbb{R}^q$ sont des pavés, alors $P \times Q$ est un pavé de $\mathbb{R}^p \times \mathbb{R}^q = \mathbb{R}^{p+q}$, et on a

$$|P \times Q| = |P|\,|Q|\,.$$

On peut maintenant définir la mesure de Lebesgue.

Théorème-Définition. *Il existe une unique mesure borélienne λ_N sur $\mathbb{R}^N$ telle que $\lambda_N(P) = |P|$ pour tout pavé $P \subseteq \mathbb{R}^N$. On dit que λ_N est la **mesure de Lebesgue** sur $\mathbb{R}^N$.*

Remarque 1. L'unicité de la mesure de Lebesgue est une conséquence immédiate du Lemme 5.1 (qui, lui, n'est pas immédiat !). La preuve de l'existence sera faite au Chapitre 10. Cependant, il est possible de donner tout de suite une formule : pour tout borélien $A \subseteq \mathbb{R}^N$, on a

$$\lambda_N(A) = \inf \sum_{k=0}^{\infty} |P_k|\,,$$

où la borne inférieure est prise sur toutes les suites de pavés (P_k) telles que $A \subseteq \bigcup_{k\in\mathbb{N}} P_k$.

Remarque 2. L'intuition est bien entendu que la mesure de Lebesgue λ_2 mesure *les aires*, et que λ_3 mesure *les volumes*. Quant à λ_1, elle mesure *les longueurs*.

5.3. Deux propriétés fondamentales. On démontre ici deux propriétés très importantes de la mesure de Lebesgue, qu'on utilisera constamment par la suite.

5.3.1. *Invariance par translations.* Pour tout ensemble $A \subseteq \mathbb{R}^N$ et pour tout $u \in \mathbb{R}^N$, notons $A + u$ le "translaté de A par u" :

$$A + u = \{x + u;\ x \in A\}\,.$$

A ce stade, le fait suivant n'est pas complètement évident :

Fait. Si $u \in \mathbb{R}^N$, alors $A + u$ est borélien pour tout borélien $A \subseteq \mathbb{R}^N$.

On peut le démontrer "à la main" en vérifiant que la famille des $A \subseteq \mathbb{R}^N$ tels que $A + u \in \mathcal{B}(\mathbb{R}^N)$ est une tribu contenant les ouverts ; mais on peut aussi attendre le chapitre suivant pour des résultats beaucoup plus généraux qui ne sont pas plus difficiles à démontrer.

Lemme 5.2. *La mesure de Lebesgue λ_N est **invariante par translations** : pour tout borélien $A \subseteq \mathbb{R}^N$ et pour tout $u \in \mathbb{R}^N$, on a*

$$\lambda_N(A + u) = \lambda_N(A)\,.$$

Démonstration. Soit $u \in \mathbb{R}^N$, et définissons une fonction d'ensembles $\mu : \mathcal{B}(\mathbb{R}^N) \to [0, \infty]$ par

$$\mu(A) = \lambda_N(A + u)\,.$$

On vérifie très facilement que μ est une mesure : il suffit juste d'observer que si (A_k) est une suite d'ensembles deux à deux disjoints, alors les $A_k + u$ sont disjoints et $\left(\bigcup_k A_k\right) + u = \bigcup_k (A_k + u)$.

Il s'agit de montrer qu'en fait $\mu = \lambda_N$; et pour cela, il suffit de vérifier qu'on a $\mu(P) = |P|$ pour tout pavé $P \subseteq \mathbb{R}^N$, par définition de la mesure de Lebesgue.

Écrivons $P = I_1 \times \cdots \times I_N$ et $u = (u_1, \dots, u_N)$. Alors $P + u$ est aussi un pavé : $P + u = \widetilde{I}_1 \times \cdots \times \widetilde{I}_N$, où $\widetilde{I}_j = I_j + u_j$. Comme $|\widetilde{I}_j| = |I_j|$ pour tout j, on a donc

$$\mu(P) = |P + u| = |\widetilde{I}_1| \cdots |\widetilde{I}_N| = |I_1| \cdots |I_N| = |P|\,.$$

$\square$

Remarque 1. On verra plus tard que la mesure de Lebesgue λ_N est en fait invariante par toutes les **isométries euclidiennes** de $\mathbb{R}^N$ (translations, rotations, symétries centrales, symétries orthogonales ; ...). L'invariance par rotations entraine par exemple que si $R \subseteq \mathbb{R}^2$ est un rectangle *quelconque* de $\mathbb{R}^2$ (dont les côtés ne sont peut-être pas parallèles aux axes de coordonnées), alors $\lambda_2(R)$ est l'aire de R au sens "naïf" du terme, *i.e.* le produit des longueurs de deux côtés adjacents.

Remarque 2. On verra également que la mesure de Lebesgue est la *seule* mesure borélienne sur $\mathbb{R}^N$ invariante par translations telle que $\lambda_N(\mathbf{P}) = 1$, où $\mathbf{P}$ est le "pavé semi-ouvert unité" $[0, 1[^N$.

Exercice 1. Montrer que la mesure de Lebesgue est effectivement invariante par symétries centrales. (Raisonner comme dans la preuve du Lemme 5.2.)

Exercice 2. Soit $h : \mathbb{R}^2 \to \mathbb{R}^2$ une homothétie de rapport c. Montrer que pour tout borélien $A \subseteq \mathbb{R}^2$, on a $\lambda_2(h(A)) = c^2 \lambda_2(A)$. (On ne demande pas de montrer que $h(A)$ est borélien.)

5.3.2. *Propriété de produit.* En termes savants (*cf* Chapitre 10), le lemme suivant signifie que la mesure de Lebesgue sur $\mathbb{R}^{p+q} = \mathbb{R}^p \times \mathbb{R}^q$ est la "mesure produit" des mesures de Lebesgue sur $\mathbb{R}^p$ et sur $\mathbb{R}^q$.

Lemme 5.3. *Si $A \in \mathcal{B}(\mathbb{R}^p)$ et $B \in \mathcal{B}(\mathbb{R}^q)$, alors $A \times B \in \mathcal{B}(\mathbb{R}^{p+q})$ et*

$$\lambda_{p+q}(A \times B) = \lambda_p(A)\, \lambda_q(B)\,.$$

Démonstration. On a déjà vu que $A \times B \in \mathcal{B}(\mathbb{R}^p \times \mathbb{R}^q) = \mathcal{B}(\mathbb{R}^{p+q})$ si $A \in \mathcal{B}(\mathbb{R}^p)$ et $B \in \mathcal{B}(\mathbb{R}^q)$; il s'agit ici de démontrer la formule pour $\lambda_{p+q}(A \times B)$. La preuve se fait en 2 étapes. (C'est le "principe du fusil à 2 coups".)

Étape 1. On a $\lambda_{p+q}(A \times B) = \lambda_p(A)\, \lambda_q(B)$ pour tout *pavé* $B \subseteq \mathbb{R}^q$ et pour tout borélien $A \subseteq \mathbb{R}^p$.

Démonstration. Fixons un pavé $B \subseteq \mathbb{R}^q$, et définissons deux fonctions d'ensembles $\mu_1 : \mathcal{B}(\mathbb{R}^p) \to [0, \infty]$ et $\mu_2 : \mathcal{B}(\mathbb{R}^p) \to [0, \infty]$ par

$$\mu_1(A) = \lambda_{p+q}(A \times B) \quad \text{et} \quad \mu_2(A) = \lambda_p(A)\, \lambda_q(B)\,.$$

On vérifie sans difficulté que μ_1 et μ_2 sont des mesures : c'est évident pour μ_2, qui est à un facteur près la mesure de Lebesgue λ_p ; et pour μ_1, il suffit de remarquer que si

(A_k) est une suite d'ensemble deux à deux disjoints, alors $\left(\bigcup_k A_k\right) \times B$ est la réunion disjointe des $A_k \times B$. De plus, si A est un pavé, alors $A \times B$ est un pavé ; donc

$$\mu_1(A) = |A \times B| = |A|\,|B| = \lambda_p(A)\,\lambda_q(B) = \mu_2(A) < \infty\,.$$

D'après le Lemme d'unicité 5.1, on a donc $\mu_1 = \mu_2$, ce qui est le résultat souhaité. □

Étape 2. On a $\lambda_{p+q}(A \times B) = \lambda_p(A)\,\lambda_q(B)$ pour tout borélien $A \subseteq \mathbb{R}^p$ tel que $\lambda_p(A) < \infty$ et pour tout borélien $B \subseteq \mathbb{R}^q$.

Démonstration. Cette fois, on fixe un borélien $A \subseteq \mathbb{R}^p$ tel que $\lambda_p(A) < \infty$, et on définit deux mesures $\mu_1, \mu_2 : \mathcal{B}(\mathbb{R}^q) \to [0, \infty]$ par

$$\mu_1(B) = \lambda_{p+q}(A \times B) \quad \text{et} \quad \mu_2(B) = \lambda_p(A)\,\lambda_q(B)\,.$$

Par l'Étape 1, on a $\mu_1(B) = \mu_2(B) < \infty$ pour tout pavé $B \subseteq \mathbb{R}^q$; donc $\mu_1 = \mu_2$ d'après le Lemme 5.1, ce qui est le résultat souhaité. □

Soient maintenant $A \in \mathcal{B}(\mathbb{R}^p)$ et $B \in \mathcal{B}(\mathbb{R}^q)$ quelconques. Si on pose $E_n := [-n, n]^p$ et $A_n = A \cap E_n$, alors la suite $(A_n)_{n\in\mathbb{N}}$ est croissante et $A = \bigcup_n A_n$; donc $A \times B$ est la réunion croissante des $A_n \times B$. De plus, on a $\lambda_p(A_n) \leqslant \lambda_p(E_n) < \infty$ pour tout n, donc $\lambda_{p+q}(A_n \times B) = \lambda_p(A_n)\,\lambda_q(B)$ d'après l'Étape 2. Ainsi,

$$\lambda_{p+q}(A \times B) = \lim_{n\to\infty} \lambda_{p+q}(A_n \times B) = \lim_{n\to\infty} \lambda_p(A_n)\,\lambda_q(B) = \lambda_p(A)\,\lambda_q(B)\,.$$

□

Exercice. La fin de la preuve a utilisé le fait suivant : si (a_n) est une suite croissante de nombres positifs de limite $a \in [0, \infty]$ et si $b \in [0, \infty]$, alors $a_n b \to ab$. Vérifier que cela est bien vrai (on a besoin de la convention $\infty \times 0 = 0$).

6. Ensembles négligeables ; propriétés vraies presque partout

6.1. Définitions. Dans tout ce qui suit, $(\Omega, \mathcal{B}, \mu)$ est un espace mesuré.

Définition 6.1. *On dit qu'un ensemble $N \subseteq B$ est μ-**négligeable** s'il existe un ensemble mesurable B tel que $N \subseteq B$ et $\mu(B) = 0$.*

Remarque 1. La définition n'impose pas que N soit mesurable.

Remarque 2. Un ensemble *mesurable* N est négligeable si et seulement si $\mu(N) = 0$. (C'est évident par la propriété de croissance.)

Remarque 3. La famile des ensembles μ-négligeables est **héréditaire** : si N est μ-négligeable, alors tout ensemble $N' \subseteq N$ est μ-négligeable.

Remarque 4. La μ-négligeabilité *dépend très fortement de la mesure μ considérée.*

Exemple 1. Soit $a \in \Omega$, et soit δ_a la masse de Dirac au point a. Alors un ensemble $N \subseteq \Omega$ est δ_a-négligeable si et seulement si $N \not\ni a$.

Exemple 2. Le seul ensemble négligeable pour la mesure de comptage μ_c est $N = \varnothing$.

Définition 6.2. *Soit $P(x)$ une propriété dépendant de $x \in \Omega$. On dit que $P(x)$ a lieu μ-**presque partout** si l'ensemble des $x \in \Omega$* ne vérifiant pas *$P(x)$ est μ-négligeable.*

Par exemple, étant donné $a \in \Omega$, on a "$x = a$" δ_a-presque partout ; tandis que "presque partout relativement à la mesure de comptage μ_c" est une manière compliquée de dire "partout".

La proposition suivante est très importante.

Proposition 6.3. *Toute réunion dénombrable d'ensembles μ-négligeables est μ-négligeable.*

Démonstration. Soit $(N_i)_{i\in I}$ une famille dénombrable d'ensembles négligeables, et soit $N := \bigcup_{i\in I} N_i$. Pour tout $i \in I$, on peut trouver $B_i \in \mathcal{B}$ tel que $N_i \subseteq B_i$ et $\mu(B_i) = 0$. Alors $B := \bigcup_{i\in I} \in \mathcal{B}$ et $\mu(B) \leqslant \sum_{i\in I} \mu(B_i) = 0$ par sous-additivité dénombrable ; donc $\mu(B) = 0$. Comme $N \subseteq B$, cela prouve que N est μ-négligeable. $\square$

Corollaire 6.4. *Toute* conjonction dénombrable *de propriétés vraies μ-presque partout est encore vraie μ-presque partout : si $(P_i)_{i\in I}$ est une famille dénombrable de propriétés vraies μ-presque partout, alors la propriété "$\forall i \in I \ : \ P_i(x)$" est elle aussi vraie μ-presque partout.*

Démonstration. C'est évident : si on pose $N_i := \{x \in \Omega$ ne vérifiant pas $P_i(x)\}$ et $N := \{x \in \Omega$ ne vérifiant pas "$\forall i \in I \ : \ P_i(x)$"$\}$, alors les N_i sont μ-négligeables et on a $N = \bigcup_{i\in I} N_i$. $\square$

Exercice 1. Montrer que tout ensemble λ_N-négligeable dans $\mathbb{R}^N$ est d'intérieur vide.

Exercice 2. Soit $I \subseteq \mathbb{R}$ un intervalle non trivial. Montrer que si deux fonction *continues* $f, g : I \to \mathbb{R}$ sont égales λ_1-presque partout, alors elles sont égales partout.

6.2. Exemples d'ensembles Lebesgue-négligeables. Il est bon de disposer d'un "stock" d'ensembles dont on est certain qu'ils sont négligeables pour la mesure de Lebesgue. On se contentera de quelques exemples rudimentaires.

Exemple 1. Tout ensemble *dénombrable* $D \subseteq \mathbb{R}$ est λ_1-négligeable.

Démonstration. Par la Proposition 6.3, il suffit de montrer que tout singleton $\{a\}$ est λ_1-négligeable ; ce qui est évident par définition de λ_1 puisque $\{a\}$ est l'intervalle "dégénéré" $[a, a]$ et donc $\lambda_1(\{a\}) = |[a, a]| = 0$. $\square$

Remarque 1. Comme $\mathbb{R}$ n'est pas λ_1-négligeable, on en déduit immédiatement que $\mathbb{R}$ n'est pas dénombrable. (Évidemment cette preuve n'est pas du tout élémentaire, car elle repose sur l'existence de la mesure de Lebesgue.)

Remarque 2. Cet exemple ne donne qu'un critère très grossier : il existe des quantités de parties λ_1-négligeables de $\mathbb{R}$ qui *ne sont pas* dénombrables.

Exercice. Pour tout intervalle compact (non vide) $I \subseteq \mathbb{R}$, notons I_0 et I_1 les deux sous-intervalles de I de longueur $\frac{1}{3}|I|$ contenant respectivement l'extrémité gauche de I et l'extrémité droite de I. Par "induction", on définit des intervalles I_s pour toute suite finie s de 0 et de 1 (éventuellement vide) de la façon suivante : $I_\varnothing := [0, 1]$, puis $I_{s0} = (I_s)_0$ et $I_{s1} = (I_s)_1$. L'**ensemble triadique de Cantor** est par définition

$$K := \bigcap_{n=0}^{\infty} \bigcup_{s\in\{0,1\}^n} I_s\,.$$

(i) Montrer que si $\alpha = (\alpha_0, \alpha_1, \dots)$ est une suite *infinie* de 0 et de 1, alors $\bigcap_{n=0}^{\infty} I_{\alpha_0\cdots\alpha_n}$ est réduite à un point x_α, et qu'on a $x_\alpha \neq x_\beta$ si $\alpha \neq \beta$. En déduire de K n'est pas dénombrable.

(ii) Calculer $\lambda_1\left(\bigcup_{s\in\{0,1\}^n} I_s\right)$ pour tout $n \in \mathbb{N}$, et en déduire que $\lambda_1(K) = 0$.

Exemple 2. Soit N un entier au moins égal à 2. Si $\gamma : I \to \mathbb{R}^N$ est une application de classe $\mathcal{C}^1$ d'un intervalle $I \subseteq \mathbb{R}$ dans $\mathbb{R}^N$, alors $\Gamma := \gamma(I)$ est λ_N-négligeable.

Démonstration. On sait que l'intervalle I est réunion dénombrable d'intervalles *compacts* I_n, $n \in \mathbb{N}$. Alors $\gamma(I) = \bigcup_n \gamma(I_n)$, donc il suffit de montrer que les $\gamma(I_n)$ sont négligeables. On s'est donc ramené au cas où l'intervalle I est compact ; disons $I = [a,b]$. Dans ce cas, Γ est compact (par continuité de γ), donc certainement borélien. (Dans le cas général, Γ est réunion dénombrable de compacts, et donc encore borélien.)

Comme $\gamma : [a,b] \to \mathbb{R}^N$ est de classe $\mathcal{C}^1$, sa dérivée γ' est *bornée* sur le compact $[a,b]$; donc γ est *lipschitzienne* (pour n'importe quelle norme sur $\mathbb{R}^N$), d'après l'inégalité des accroissements finis. Fixons une constante K telle que

$$\forall s,t \in [a,b] \; : \; \|\gamma(t) - \gamma(s)\|_\infty \leqslant K\,|t-s|\,.$$

Soit $R \in \mathbb{N}^*$ quelconque, et soit $a = s_0 < s_1 < \cdots < s_R = b$ la subdivision de $[a,b]$ obtenue en découpant $[a,b]$ en R intervalles de longueur $\eta := \frac{b-a}{R}\cdot$ On a

$$\Gamma = \gamma([a,b]) = \bigcup_{i=0}^{R-1} \gamma([s_i, s_{i+1}]) := \bigcup_{i=0}^{R-1} \Gamma_i\,.$$

Si $t \in [s_i, s_{i+1}]$ alors $\|\gamma(t) - \gamma(s_i)\|_\infty \leqslant K\,|t - s_i| \leqslant K\,\eta$. Donc $\gamma(t)$ appartient à la boule $\overline{B}_{\|\cdot\|_\infty}(\gamma(s_i), K\eta)$ pour tout $t \in [s_i, s_{i+1}]$, autrement dit

$$\Gamma_i \subseteq \overline{B}_{\|\cdot\|_\infty}(\gamma(s_i), K\eta) := P_i\,.$$

Mais comme on utilise la norme $\|\cdot\|_\infty$, les P_i sont des “pavés cubiques” de côté $2K\eta$. (En écrivant $\gamma(s_i) = (x_1, \dots, x_N)$, on a $P_i = [x_1 - \eta, x_1 + \eta] \times \cdots \times [x_N - \eta, x_N + \eta]$.) En particulier, $\lambda_N(P_i) = (2K\eta)^N$. Comme $\Gamma_i \subseteq P_i$, on a donc $\lambda_N(\Gamma_i) \leqslant (2\eta)^N$ pour $i = 0, \dots, N-1$, d'où

$$\begin{aligned} \lambda_N(\Gamma) &= \lambda_N\left(\bigcup_{i=0}^{R-1} \Gamma_i\right) \\ &\leqslant \sum_{i=0}^{R-1} \lambda_N(\Gamma_i) \\ &\leqslant R \times (2K\eta)^N\,. \end{aligned}$$

Comme $\eta = \frac{b-a}{R}$, on obtient donc une majoration de la forme

$$\lambda_N(\Gamma) \leqslant R \times \frac{C}{R^N} = \frac{C}{R^{N-1}}\,,$$

où C est une constante indépendante de R. Ceci étant vrai pour tout entier $R \geqslant 1$ (et comme on suppose que $N \geqslant 2$), on en déduit $\lambda_N(\Gamma) = 0$. □

Exemple 3. Si $N \geqslant 2$, alors toute *droite* $L \subseteq \mathbb{R}^N$ est λ_N-négligeable.

Démonstration. C'est un cas très particulier de l'exemple précédent. En effet, si $\overrightarrow{u}$ est un vecteur directeur de L et si on fixe un point $p \in L$, alors on peut écrire $L = \gamma(\mathbb{R})$ où $\gamma : \mathbb{R} \to \mathbb{R}^N$ est l'application $\mathcal{C}^1$ définie par $\gamma(t) = p + t\,\overrightarrow{u}$. □

Exercice. Montrer que si $A \subseteq \mathbb{R}^2$ est un polygone ou un disque fermé, alors $\lambda_2(A) = \lambda_2(\mathring{A})$.

Remarque 6.5. Il n'est pas difficile de généraliser un peu l'Exemple 2. Soient $N, d \in \mathbb{N}^*$ avec $d \leqslant N$. Convenons de dire qu'un ensemble $\Sigma \subseteq \mathbb{R}^N$ est **d-paramétrable** s'il est réunion dénombrable d'ensembles de la forme $\Sigma = \Phi(A)$, où $A \subseteq \mathbb{R}^d$ est réunion dénombrable de pavés compacts P_n et Φ est de classe $\mathcal{C}^1$ sur chaque P_n. En adaptant la preuve de l'Exemple 2, on montre que *si $d < N$, alors tout ensemble d-paramétrable $\Sigma \subseteq \mathbb{R}^N$ est λ_N-négligeable.*

Exemple 4. Tout *sous espace affine* $E \subseteq \mathbb{R}^N$ de dimension $d < N$ est négligeable.

Démonstration. Soit $a \in E$, et soit $(\overrightarrow{u_1}, \dots, \overrightarrow{u_d})$ une base du sous-espace vectoriel de $\mathbb{R}^N$ associé à E. Alors $E = \Phi(\mathbb{R}^d)$, où $\Phi : \mathbb{R}^d \to \mathbb{R}^N$ est l'application (de classe $\mathcal{C}^1$) définie par $\phi(t_1, \dots, t_d) = a + \sum_1^d t_i \overrightarrow{u_i}$; donc E est d-paramétrable. □

7. Mesure de Lebesgue et aire intuitive

Tout le monde a une idée intuitive de ce qu'est l'aire d'une partie raisonnablement simple du plan. Par exemple, l'aire d'un rectangle quelconque (à côtés pas forcément parallèles aux axes de coordonnées) est égale au produit des longueurs de deux côtés adjacents, et l'aire d'un triangle est égale à la moitié du produit d'une base quelconque par la hauteur correspondante. (On apprend aussi que l'aire d'un disque de rayon R est égale à πR^2...)

Cela étant, maintenant qu'on a introduit la mesure de Lebesgue, on peut dire que *par définition*, l'aire d'un borélien $A \subseteq \mathbb{R}^2$ est égale à $\lambda_2(A)$. La moindre des choses est évidemment de montrer sans tricher qu'il n'y a pas de conflit entre ces deux notions d'aire ("aire λ_2" et "aire intuitive").

Appelons **figure simple** tout ensemble $\Delta \subseteq \mathbb{R}^2$ qui peut se décomposer comme

$$\Delta = \Delta_1 \cup \cdots \cup \Delta_N,$$

où les Δ_i sont des triangles pleins (intérieur plus frontière) **presque disjoints**, i.e. d'intérieurs deux à deux disjoints.

L'aire intuitive d'une telle figure simple Δ sera notée aire(Δ). On a par définition

$$\text{aire}(\Delta) = \text{aire}(\Delta_1) + \cdots + \text{aire}(\Delta_N)\,,$$

pour n'importe quelle décomposition de Δ en triangles presque disjoints Δ_i. (Il faudrait vérifier que la somme $\sum \text{aire}(\Delta_i)$ est indépendante de la décomposition de Δ en triangles presque disjoints Δ_i ; ce qui est faisable "à la main", mais loin d'être évident.)

On peut maintenant énoncer le résultat de "non conflit" entre les deux notions d'aire disponibles pour les figures simples.

Proposition 7.1. *Si $\Delta \subseteq \mathbb{R}^2$ est une figure simple, alors $\lambda_2(\Delta) = \text{aire}(\Delta)$.*

Démonstration. On utilisera constamment le fait suivant.

Fait. Si Δ et Δ' sont des figures simples presque disjointes, alors $\lambda_2(\Delta \cup \Delta') = \lambda_2(\Delta) + \lambda_2(\Delta')$.

Preuve du Fait. Comme Δ et Δ' sont presque disjointes, $\Delta \cap \Delta'$ est une réunion finie de segments (éventuellement vide ou réduite à un nombre fini de points). Donc $\lambda_2(\Delta \cap \Delta') = 0$; d'où le résultat par le Corollaire 3.3. □

Cas 1. Cas où Δ est un triangle rectangle dont les côtés de l'angle droit sont parallèles aux axes de coordonnées.

Soient a et b les longueurs des côtés de l'angle droit, et Δ' le triangle symétrique de Δ par rapport au milieu de l'hypoténuse de Δ. Alors $P := \Delta \cup \Delta'$ est un rectangle à côtés parallèles aux axes de coordonnées, donc $\lambda_2(P) = ab$. De plus, Δ et Δ' sont presque disjoints, et on a $\lambda_2(\Delta') = \lambda_2(\Delta)$ car la mesure de Lebesgue est *invariante* *par symétries centrales* (*cf* l'Exercice 1 après le Lemme 5.2). Donc $\lambda_2(P) = \lambda_2(\Delta) + \lambda_2(\Delta') = 2\lambda_2(\Delta)$, et donc $\lambda_2(\Delta) = \frac{ab}{2} = \text{aire}(\Delta)$.

Cas 2. Cas où Δ est un *rectangle* quelconque.

C'est le cas le plus compliqué, qui en fait est équivalent à l'*invariance par rotations* de la mesure λ_2. Comme on ne peut pas tricher en utilisant cette invariance par rotations, il faut faire les choses à la main. Bien entendu, on suppose que les côtés de Δ ne sont pas parallèles aux axes de coordonnées. On notera a et b les longueurs de deux côtés adjacents de Δ.

On peut compléter Δ en un grand rectangle R à côtés parallèles aux axes de coordonnées, en construisant sur les côtés de Δ des triangles rectangles $\Delta_1, \ldots, \Delta_4$ à côtés de l'angle droit parallèles aux axes de coordonnées. Comme Δ et les Δ_i sont presque disjoints, on a

$$\lambda_2(R) = \lambda_2(\Delta) + \sum_{i=1}^{4} \lambda_2(\Delta_i)\,. \tag{7.1}$$

Les triangles Δ_i vont par paires. Notons x, y les longueurs des côtés de l'angle droit pour les 2 triangles dont l'hypoténuse est de longueur a, et u, v pour ceux dont l'hypoténuse est de longueur b. Quitte à permuter x, y et u, v, les longueurs A et B des côtés du grand rectangle R sont alors données par

$$A = x + u \quad \text{et} \quad B = y + v\,.$$

Comme R et les Δ_i ont des côtés parallèles aux axes de coordonnées, on a $\lambda_2(R) = (x+u)(y+v)$ par définition de λ_2, et $\lambda_2(\Delta_i) = \frac{xy}{2}$ ou $\frac{uv}{2}$ d'après le Cas 1. Donc l'équation (7.1) devient

$$\begin{aligned} \lambda_2(\Delta) &= (x+u)(y+v) - 2\left(\frac{xy}{2} + \frac{uv}{2}\right) \\ &= xv + yu\,. \end{aligned}$$

Par ailleurs, on a d'après le théorème de Pythagore :

$$x^2 + y^2 = a^2 \quad \text{et} \quad u^2 + v^2 = b^2\,. \tag{7.2}$$

Enfin, en considérant les angles des triangles Δ_i, on voit qu'on a aussi

$$\frac{x}{y} = \frac{v}{u}\,\cdot \tag{7.3}$$

En élevant au carré l'identité $\lambda_2(\Delta) = xv + yu$ et en utilisant (7.3) et (7.2), on obtient

$$\begin{aligned} \lambda_2(\Delta)^2 &= x^2v^2 + 2xuyv + y^2u^2 \\ &= x^2v^2 + (xu)^2 + (yv)^2 + y^2u^2 \\ &= x^2(u^2+v^2) + y^2(v^2+u^2) \\ &= (x^2+y^2)b^2 = a^2b^2\,. \end{aligned}$$

Donc $\lambda_2(\Delta) = ab = \text{aire}(\Delta)$.

Cas 3. Cas où Δ est un triangle rectangle quelconque.

Dans ce cas, on obtient la formule attendue en utilisant le Cas 2 et en raisonnant comme dans le Cas 1.

CAS 4. Cas où Δ est un triangle quelconque.

Notons A, B, C les sommets de Δ, et H le pied de la hauteur issue de A. Si H appartient au segment $[BC]$, alors Δ est la réunion presque disjointe des triangles rectangles ABH et AHB ; donc $\lambda_2(\Delta) = \lambda_2(ABH) + \lambda_2(AHC)$, et par le Cas 3 :

$$\lambda_2(\Delta) = \frac{1}{2}\,BH \times AH + \frac{1}{2}\,HC \times AH = \frac{1}{2}\,(BH + HC) \times AH = \frac{1}{2}\,BC \times AH = \text{aire}(\Delta)\,.$$

Si H est extérieur au segment $[BC]$, on a par exemple $B \in [HC]$. En considérant le triangle AHC, on obtient comme précédemment

$$\lambda_2(\Delta) = \frac{1}{2}\,(HC - HB) \times AH = \frac{1}{2}\,BC \times AH = \text{aire}(\Delta)\,.$$

CAS 5. Cas général.

La figure simple Δ se décompose en triangles presque disjoints $\Delta_1, \ldots, \Delta_N$, et on a alors $\lambda_2(\Delta) = \lambda_2(\Delta_1) + \cdots + \lambda_2(\Delta_N) = \text{aire}(\Delta)$, d'après le Cas 4. □

Exercice. Montrer que si $D \subseteq \mathbb{R}^2$ est un disque ouvert de rayon R, alors $\lambda_2(D) = \pi R^2$. (On pourra commencer par calculer, pour tout entier $n \geqslant 2$, l'aire d'un polygone régulier à 2^n côtés inscrit dans D.)

Chapitre 3

Applications mesurables

1. Définition et critères de mesurabilité

DÉFINITION 1.1. *Soient* $(\Omega, \mathcal{B})$ *et* $(\Omega', \mathcal{B}')$ *deux espaces mesurables. On dit qu'une application* $f : \Omega \to \Omega'$ *est* $(\mathcal{B}, \mathcal{B}')$***-mesurable*** *si*

$$\forall A \in \mathcal{B}' \ : \ f^{-1}(A) \in \mathcal{B}\,.$$

Il y une similitude formelle évidente avec la caractérisation bien connue (ou si on préfère, la définition) de la continuité : une application $f : \Omega \to \Omega'$ entre deux espaces métriques est continue si et seulement si $f^{-1}(O)$ est *ouvert* dans Ω pour tout *ouvert* $O \subseteq \Omega'$. On verra cependant que la mesurabilité est une notion beaucoup plus "souple" que la continuité, car elle est préservée par beaucoup plus d'opérations.

Remarque 1. Comme "$(\mathcal{B}, \mathcal{B}')$-mesurable" n'est pas très joli à l'oeil, on emploiera systématiquement les raccourcis suivants.

- Lorsque Ω' est un espace métrique, on dit que $f : \Omega \to \Omega'$ est $\mathcal{B}$**-mesurable** si elle est $(\mathcal{B}, \mathcal{B}(\Omega'))$-mesurable, où $\mathcal{B}(\Omega')$ est la tribu borélienne de Ω'.
- Lorsque Ω et Ω' sont tous les deux des espaces métriques, on dit que $f : \Omega \to \Omega'$ est **borélienne** si elle est $(\mathcal{B}(\Omega), \mathcal{B}(\Omega'))$-mesurable.

Dans le même esprit, au lieu d'écrire "$f : \Omega \to \Omega'$ est $(\mathcal{B}, \mathcal{B}')$-mesurable", on écrira souvent "$f : (\Omega, \mathcal{B}) \to (\Omega', \mathcal{B}')$ est mesurable".

Remarque 2. En réalité, il arrivera aussi très souvent qu'on dise simplement "mesurable" au lieu de "$(\mathcal{B}, \mathcal{B}')$-mesurable" si les tribus avec lesquelles on joue sont clairement identifiées. Ainsi, on peut dire qu'*une application* f *est mesurable si et seulement si* $f^{-1}(B)$ *est mesurable pour tout ensemble mesurable* B

Voici un exemple évident, mais quand même important.

Exemple. Toute application *constante* est mesurable, quelles que soient les tribus au départ et à l'arrivée.

Démonstration. Si $f : \Omega \to \Omega'$ est constante, $f(x) \equiv \alpha$, alors, pour tout ensemble $A \subseteq \Omega'$, l'ensemble $f^{-1}(A)$ est soit vide (si $\alpha \notin A$), soit égal à Ω (si $\alpha \in A$). Dans les deux cas, $f^{-1}(A)$ est mesurable. □

La proposition suivante dit que pour vérifier la mesurabilité d'une certaine application, il n'est pas nécessaire de "tester" sur *tous* les éléments de la tribu $\mathcal{B}'$ à l'arrivée, mais seulement sur tous les ensembles appartenant à une certaine famille de parties engendrant $\mathcal{B}'$. C'est un résultat *très utile* dans la pratique.

PROPOSITION 1.2. *Soient* $(\Omega, \mathcal{B})$ *et* $(\Omega', \mathcal{B}')$ *deux espaces mesurables, et soit* $f : \Omega \to \Omega'$*. On suppose qu'il existe une famille* $\mathcal{C} \subseteq \mathcal{B}'$ *telle que* $\sigma(\mathcal{C}) = \mathcal{B}'$ *et* $\forall C \in \mathcal{C} \ :$ $f^{-1}(C) \in \mathcal{B}$*. Alors* f *est* $(\mathcal{B}, \mathcal{B}')$*-mesurable.*

Démonstration. Posons $\mathcal{B}_f := \{A \subseteq \Omega';\ f^{-1}(A) \in \mathcal{B}\}$. Comme $\mathcal{B}$ est une tribu, on vérifie sans difficulté que $\mathcal{B}_f$ est une tribu de parties de Ω' (exercice ; on dit que $\mathcal{B}_f$ est la tribu **image** de $\mathcal{B}$ par f). Comme $\mathcal{C} \subseteq \mathcal{B}_f$ par hypothèse, on a donc $\mathcal{B}' = \sigma(\mathcal{C}) \subseteq \mathcal{B}_f$; autrement dit : $f^{-1}(A) \in \mathcal{B}$ pour tout $A \in \mathcal{B}'$. $\square$

COROLLAIRE 1.3. *Si Ω et Ω' sont des espaces métriques, alors toute application* continue $f : \Omega \to \Omega'$ *est borélienne.*

Démonstration. On applique la proposition en prenant pour $\mathcal{C}$ la famille de tous les ouverts de Ω'. Comme f est continue, on sait que $f^{-1}(O)$ est ouvert dans Ω (donc borélien) pour tout $O \in \mathcal{C}$, d'où le résultat puisque $\sigma(\mathcal{C}) = \mathcal{B}(\Omega')$ par définition de la tribu borélienne. $\square$

Exemple. Soit $u \in \mathbb{R}^N$. L'application $\tau_u : \mathbb{R}^N \to \mathbb{R}^N$ définie par $\tau_u(x) = x - u$ est continue, donc borélienne ; et si $A \subseteq \mathbb{R}^N$, alors $\tau_u^{-1}(A) = A + u$. On en déduit que si $A \subseteq \mathbb{R}^N$ est borélien, alors $A + u$ aussi.

Exercice 1. Soit $h : \mathbb{R}^2 \to \mathbb{R}^2$ une homothétie. Montrer que $h(A)$ est borélien pour tout borélien $A \subseteq \mathbb{R}^2$.

Exercice 2. Soient $\Omega_1, \dots, \Omega_N$ des espaces métriques. Montrer *en utilisant le Corollaire 1.3* que si $A_1, \dots, A_N$ sont boréliens dans $\Omega_1, \dots, \Omega_N$ respectivement, alors $A_1 \times \cdots \times A_N$ est borélien dans $\Omega_1 \times \cdots \times \Omega_N$. (Considérer les "projections canoniques" $\pi_j : \Omega_1 \times \cdots \times \Omega_N \to \Omega_j$.)

Exercice 3. Soit (f_n) une suite d'applications mesurables d'un espace mesurable $(\Omega, \mathcal{B})$ dans un espace *métrique complet* Ω'. On note E l'ensemble des points $x \in \Omega$ tels que la suite $(f_n(x))$ converge dans Ω'. Montrer que E est mesurable.

NOTATIONS. Pour toute fonction $f : \Omega \to \overline{\mathbb{R}}$ et pour $\alpha, \beta \in \mathbb{R}$, on pose

$$\{f > \alpha\} := \{x \in \Omega;\ f(x) > \alpha\} \quad \text{et} \quad \{f < \beta\} := \{x \in \Omega;\ f(x) < \beta\}.$$

COROLLAIRE 1.4. *Soit $(\Omega, \mathcal{B})$ un espace mesurable. Pour une fonction $f : \Omega \to \overline{\mathbb{R}}$, les propriétés suivantes sont équivalentes.*

(1) *f est $\mathcal{B}$-mesurable ;*

(2) *$\{f > \alpha\} \in \mathcal{B}$ pour tout $\alpha \in \mathbb{R}$;*

(2') *$\{f < \beta\} \in \mathcal{B}$ pour tout $\beta \in \mathbb{R}$.*

Démonstration. Comme $\{f < \alpha\} = f^{-1}(]\alpha, \infty])$ et $\{f < \beta\} = f^{-1}([-\infty, \beta])$, il suffit d'appliquer la proposition avec $\mathcal{C} := \{]\alpha, \infty];\ \alpha \in \mathbb{R}\}$ ou $\mathcal{C} = \{[-\infty, \beta[;\ \beta \in \mathbb{R}\}$. Dans les deux cas, la famille $\mathcal{C}$ engendre $\mathcal{B}(\overline{\mathbb{R}})$ d'après la Proposition 4.11 du Chapitre 2. $\square$

COROLLAIRE 1.5. *Si I est un intervalle de $\mathbb{R}$, toute fonction* monotone $f : I \to \mathbb{R}$ *est borélienne.*

Démonstration. Comme f est monotone, les ensembles $\{f > \alpha\}$ sont des *intervalles* (vérifier), donc des boréliens de I. $\square$

COROLLAIRE 1.6. *Soit $(\Omega, \mathcal{B})$ un espace mesurable, et soit $f : \Omega \to \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_N}$. On écrit $f(x) = (f_1(x), \dots, f_N(x))$, avec $f_j(x) \in \mathbb{R}^{m_j}$ pour $j = 1, \dots, N$. Alors f est $\mathcal{B}$-mesurable si et seulement ses "composantes" $f_1, \dots, f_N$ le sont.*

Démonstration. Suposons que f soit $\mathcal{B}$-mesurable. Pour $j \in \{1, \dots, N\}$ et pour $A \subseteq \mathbb{R}^{m_j}$ borélien, on a

$$f_j^{-1}(A) = f^{-1}\left(\mathbb{R}^{m_1} \times \cdots \times A \times \cdots \times \mathbb{R}^{m_N}\right),$$

et donc $f_j^{-1}(A) \in \mathcal{B}$ puisque f est $\mathcal{B}$-mesurable et $\mathbb{R}^{m_1} \times \cdots \times A \times \cdots \times \mathbb{R}^{m_N}$ est borélien dans $\mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_j} \times \cdots \times \mathbb{R}^{m_N}$. Ceci étant vrai pour tout borélien $A \subseteq \mathbb{R}^{m_j}$, on voit donc que f_j est $\mathcal{B}$-mesurable, pour tout $j \in \{1, \dots, N\}$.

Inversement, supposons que $f_1, \dots, f_N$ soient $\mathcal{B}$-mesurables. Pour montrer que f est $\mathcal{B}$-mesurable, il suffit de vérifier qu'on a $f^{-1}(P) \in \mathcal{B}$ pour tout *pavé* $P \subseteq \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_N} = \mathbb{R}^{m_1 + \cdots + m_N}$, car les pavés engendrent la tribu borélienne de $\mathbb{R}^{m_1 + \cdots + m_N}$.

Soit P un pavé quelconque de $\mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_N}$. On peut écrire $P = P_1 \times \cdots \times P_N$, où P_j est un pavé de $\mathbb{R}^{m_j}$ pour $j = 1, \dots, N$. Avec ces notations, on voit que $f(x) \in P$ si et seulement $f_j(x) \in P_j$ pour $j = 1, \dots, N$; autrement dit

$$f^{-1}(P) = f_1^{-1}(P_1) \cap \cdots \cap f_N^{-1}(P_N).$$

Comme les f_j sont $\mathcal{B}$-mesurables, on en déduit $f^{-1}(P) \in \mathcal{B}$. □

Corollaire 1.7. *Une fonction $f : \Omega \to \mathbb{C}$ est mesurable si et seulement si* $\mathrm{Re}(f)$ *et* $\mathrm{Im}(f)$ *le sont.*

Démonstration. C'est immédiat en identifiant $\mathbb{C}$ à $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$. □

Exercice 1.8. Soient $\Omega_1, \dots, \Omega_N$ des espaces métriques *séparables.* Montrer qu'une application $f : \Omega \to \Omega_1 \times \cdots \times \Omega_N$ est $\mathcal{B}$-mesurable *si et seulement si* ses composantes $f_1, \dots, f_N$ le sont.

La proposition suivante est un résultat de "recollement" très utile. Rappelons la notion de la **tribu trace** induite sur un sous-ensemble : si $(\Omega, \mathcal{B})$ est un espace mesurable et si $E \subseteq \Omega$, alors la tribu trace $\mathcal{B}_E \subseteq \mathcal{P}(E)$ est définie par

$$\mathcal{B}_E := \{B \cap E;\ B \in \mathcal{B}\}.$$

Proposition 1.9. *Soient $(\Omega, \mathcal{B})$ et $(\Omega', \mathcal{B}')$ deux espaces mesurables, et soit $(\Omega_i)_{i \in I}$ une famille* dénombrable *d'éléments de $\mathcal{B}$ telle que $\bigcup_{i \in I} \Omega_i = \Omega$. Une fonction $f : \Omega \to \Omega'$ est $(\mathcal{B}, \mathcal{B}')$-mesurable si et seulement si sa restriction à chaque Ω_i est $(\mathcal{B}_{\Omega_i}, \mathcal{B}')$-mesurable.*

Démonstration. Remarquons d'abord qu'on a $\mathcal{B}_{\Omega_i} \subseteq \mathcal{B}$ pour tout $i \in I$, car $\Omega_i \in \mathcal{B}$.

Supposons que f soit $(\mathcal{B}, \mathcal{B}')$-mesurable. Pour tout ensemble $A \subseteq \Omega'$ et pour tout $i \in I$, on a $\left(f_{|\Omega_i}\right)^{-1}(A) = f^{-1}(A) \cap \Omega_i$. Donc $\left(f_{|\Omega_i}\right)^{-1}(A) \in \mathcal{B}_{\Omega_i}$ si $A \in \mathcal{B}'$, car $f^{-1}(A) \in \mathcal{B}$; autrement dit, $f_{|\Omega_i}$ est $(\mathcal{B}_{\Omega_i}, \mathcal{B}')$-mesurable. (Ici, il n'est pas nécessaire de supposer que Ω_i est mesurable.)

Inversement, si $f_{|\Omega_i}$ est $(\mathcal{B}_{\Omega_i}, \mathcal{B}')$-mesurable pour tout $i \in I$ alors, comme $\Omega = \bigcup_{i \in I} \Omega_i$ et que l'ensemble d'indices I est dénombrable, on a pour tout $A \in \mathcal{B}'$:

$$f^{-1}(A) = \bigcup_{i \in I} \left(\Omega_i \cap f^{-1}(A)\right) = \bigcup_{i \in I} \underbrace{\left(f_{|\Omega_i}\right)^{-1}(A)}_{\in \mathcal{B}_{\Omega_i} \subseteq \mathcal{B}} \in \mathcal{B}.$$

□

Remarque. La preuve précédente a montré que si E est une partie quelconque de Ω et si $f : \Omega \to \Omega'$ est $(\mathcal{B}, \mathcal{B}')$-mesurable, alors $f_{|E}$ est $(\mathcal{B}_E, \mathcal{B}')$-mesurable.

Corollaire 1.10. *Avec les notations précédentes, on suppose que* Ω *et* Ω' *sont des espaces métriques et que les* Ω_i *sont des boréliens de* Ω*. Alors* $f : \Omega \to \Omega'$ *est borélienne si et seulement si sa restriction à chaque* Ω_i *est borélienne.*

Démonstration. C'est une conséquence immédiate de la proposition, car $\mathcal{B}_{\Omega_i} = \mathcal{B}(\Omega_i)$ pour tout $i \in I$ d'après la Proposition 4.2 du Chapitre 2. $\square$

Corollaire 1.11. *Soient* $(\Omega, \mathcal{B})$ *et* $(\Omega', \mathcal{B}')$ *deux espaces mesurables, et soit* $f : \Omega \to \Omega'$*. S'il existe une partition dénombrable* $(\Omega_i)_{i\in I}$ *de* Ω *en ensembles mesurables telle que* f *soit* constante *sur chaque* Ω_i*, alors* f *est mesurable.*

Démonstration. C'est évident par la proposition : la restriction de f à chaque Ω_i est constante, donc $(\mathcal{B}_{\Omega_i}, \mathcal{B}')$-mesurable. $\square$

Corollaire 1.12. *Soient* Ω *et* Ω' *deux espaces métriques, et soit* $f : \Omega \to \Omega'$*. Si l'ensemble des points de discontinuité de* f *est* dénombrable*, alors* f *est borélienne.*

Démonstration. Notons $D = \{a_i;\ i \in \mathbb{N}^*\}$ l'ensemble des points de discontinuité de f. Posons également $\Omega_0 = \Omega \backslash D$. Alors Ω_0 est borélien car D l'est (tout ensemble dénombrable dans un espace métrique est borélien), et la restriction de f à Ω_0 est continue (donc borélienne) puisque Ω_0 est l'ensemble des points de continuité de f. De plus, si on pose $\Omega_i = \{a_i\}$ pour $i \geqslant 1$, alors les Ω_i sont boréliens (!) et la restriction de f à chaque Ω_i est constante (donc borélienne ! !). On peut donc appliquer le corollaire 1.10. $\square$

Remarque. Ce dernier résultat montre en particulier que si I est un intervalle de $\mathbb{R}$, alors toute fonction $f : I \to \mathbb{C}$ *continue par morceaux* est borélienne.

Exercice 1. Soient Ω et Ω' deux espaces métriques, et soit $(\Omega_i)_{i\in I}$ une famille de parties de Ω telle que $\bigcup_{i\in I} \Omega_i = \Omega$. On suppose ou bien que tous les Ω_i sont ouverts, ou bien que tous les Ω_i sont fermés et que la famille (Ω_i) est finie. Montrer qu'une application $f : \Omega \to \Omega'$ est continue si et seulement si sa restriction à chaque Ω_i est continue.

Exercice 2. Soiet Ω et Ω' deux espaces métriques. Montrer que pour toute fonction $f : \Omega \to \Omega'$, l'ensemble des points de continuité de f est un borélien de Ω ; plus précisément, une intersection dénombrable d'ouverts. (On pourra commencer par vérifier que f est continue en un certain point x si et seulement si pour tout $\varepsilon > 0$, on peut trouver un voisinage ouvert V de x tel que $\forall u, v \in V\ :\ d(f(u), f(v)) < \varepsilon$.)

2. Propriétés de stabilité

Le résultat suivant est à peu près évident, mais tout à fait essentiel (comme l'est le résultat correspondant pour les fonctions continues).

Proposition 2.1. *Si* $f : (\Omega, \mathcal{B}) \to (\Omega', \mathcal{B}')$ *et* $g : (\Omega', \mathcal{B}') \to (\Omega'', \mathcal{B}'')$ *sont mesurables, alors* $g \circ f : (\Omega, \mathcal{B}) \to (\Omega'', \mathcal{B}'')$ *est mesurable. Avec des mots :* la composée de deux fonctions mesurables est une application mesurable.

Démonstration. Si $B \in \mathcal{B}''$, alors

$$(g \circ f)^{-1}(B) = \underbrace{f^{-1}\big(\underbrace{g^{-1}(B)}_{\in \mathcal{B}'}\big)}_{\in \mathcal{B}} \in \mathcal{B}\,.$$

$\square$

Corollaire 2.2. *Si $f, g : (\Omega, \mathcal{B}) \to \mathbb{C}$ sont mesurables, alors $f + g$ et fg sont mesurables, et f/g est mesurable si g ne s'annule pas.*

Démonstration. Soient $\Phi : \Omega \to \mathbb{C} \times \mathbb{C}$ et $S : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ définies par $\Phi(x) := (f(x), g(x))$ et $S(u, v) = u + v$. Alors Φ est $\mathcal{B}$-mesurable d'après le Corollaire 1.6, et S est continue donc borélienne. Donc $f + g = S \circ \Phi$ est $\mathcal{B}$-mesurable par composition. On montre de même que fg et f/g sont mesurables. $\square$

Remarque 2.3. Ce dernier résultat est vrai également pour des fonctions à valeurs dans $[0, \infty]$: si $f, g : \Omega \to [0, \infty]$ sont mesurables, alors $f + g$ et fg le sont également. Plus généralement, si f et g sont mesurables à valeurs dans $\overline{\mathbb{R}}$, alors $f+g$ est mesurable dès lors qu'elle est bien définie en tout point, et fg est mesurable.

Démonstration. Pour $f + g$, la preuve est essentiellement la même que plus haut. Voici cependant les détails : l'application "somme" S est bien définie et continue sur $\Lambda := \big(\overline{\mathbb{R}} \times \overline{\mathbb{R}}\big) \backslash \{(\infty, -\infty); (-\infty, \infty)\}$, qui est ouvert dans $\overline{\mathbb{R}} \times \overline{\mathbb{R}}$, donc borélien, et la fonction $\Phi = (f, g)$ est borélienne et prend ses valeurs dans Λ car $f + g$ supposée définie en tout point ; donc $f + g = S \circ \Phi$ est borélienne par composition.

Pour fg, il faut faire attention car le produit n'est pas continu sur $\overline{\mathbb{R}} \times \overline{\mathbb{R}}$. On peut par exemple procéder comme suit. Si on pose $\Omega_0 := \{x \in \Omega;\ f(x) = 0 \text{ ou } g(x) = 0\}$, alors Ω_0 et $\Omega_1 := \Omega \backslash \Omega_0 = \{x \in \Omega;\ f(x) \neq 0 \text{ et } g(x) \neq 0\}$ sont mesurables car f et g le sont. La restriction de fg à Ω_0 est identiquement nulle, donc mesurable ; et la restriction de fg à Ω_1 est mesurable "par composition" car le produit est continu sur $\big(\overline{\mathbb{R}} \backslash \{0\}\big) \times \big(\overline{\mathbb{R}} \backslash \{0\}\big)$. Donc fg est mesurable "par recollement". $\square$

Corollaire 2.4. *Les fonctions mesurables à valeurs complexes forment un espace vectoriel (et même une* algèbre*).*

Démonstration. C'est immédiat par le corollaire 2.2. $\square$

On va voir maintenant que la mesurabilité est préservée par "enveloppe supérieure ou inférieure dénombrable". Dans tout ce quit, si $(f_i)_{i \in I}$ est une famille de fonctions définies sur un même ensemble Ω, à valeurs dans $\overline{\mathbb{R}}$, on note $\sup_i f_i$ et $\inf_i f_i$ les fonctions définies par

$$\Big(\sup_i f_i\Big)(x) = \sup_{i \in I} f_i(x) \quad \text{et} \quad \Big(\inf_i f_i\Big)(x) = \inf_{i \in I} f_i(x)\,.$$

Proposition 2.5. *Si $(f_i)_{i \in I}$ est une famille* dénombrable *de fonctions mesurables, $f_i : (\Omega, \mathcal{B}) \to \overline{\mathbb{R}}$, alors les fonctions $\sup_i f_i$ et $\inf_i f_i$ sont mesurables.*

Démonstration. Soit $f := \sup_i f_i$. Par définition, si $\alpha \in \mathbb{R}$, alors l'équivalence suivante a lieu pour tout $x \in \Omega$:

$$f(x) > \alpha \iff \exists i \in I \ :\ f_i(x) > \alpha\,.$$

Avec les notations du Corollaire 1.4, on a donc

$$\{f > \alpha\} = \bigcup_{i \in I} \{f_i > \alpha\}\,.$$

Ainsi, tous les ensembles $\{f > \alpha\}$ sont mesurables, car ce sont des réunions *dénombrables* d'ensembles mesurables ; et donc $f = \sup_i f_i$ est mesurable d'après le Corollaire 1.4.

On montre de même que la fonction $\inf_i f_i$ est mesurable en considérant les ensembles $\{\inf_i f_i < \alpha\}$; ou bien "par composition" en observant que $\inf_i f_i = -\sup_i(-f_i)$ et en appliquant le cas déjà traité aux fonctions $-f_i$. $\square$

Exercice. Montrer que si $(f_i)_{i\in I}$ est une famille quelconque de fonctions *continues* sur un espace métrique Ω, à valeurs dans $\overline{\mathbb{R}}$, alors les fonctions $\sup_i f_i$ et $\inf_i f_i$ sont boréliennes.

COROLLAIRE 2.6. *Si $f, g : \Omega \to \overline{\mathbb{R}}$ sont mesurables, alors les fonctions* $\max(f,g)$ *et* $\min(f,g)$ *sont mesurables*

Une conséquence très importante de la Proposition 2.5 est le fait que la mesurabilité est préservé par "passage à la limsup et à la liminf" (ce qui est évidemment faux pour la continuité).

COROLLAIRE 2.7. *Si $(f_n)_{n\in\mathbb{N}}$ est une suite d'applications mesurables, $f_n : (\Omega, \mathcal{B}) \to \overline{\mathbb{R}}$, alors les fonctions $\overline{\lim}\, f_n$ et $\underline{\lim}\, f_n$ sont mesurables.*

Démonstration. C'est évident d'après la proposition puisqu'on a

$$\overline{\lim}\, f_n = \inf_{n\in\mathbb{N}} \sup_{m\geqslant n} f_m \quad \text{et} \quad \underline{\lim}\, f_n = \sup_{n\in\mathbb{N}} \inf_{m\geqslant n} f_m\,.$$

□

COROLLAIRE 2.8. *Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables sur $(\Omega, \mathcal{B})$, à valeurs dans $\overline{\mathbb{R}}$ ou dans $\mathbb{C}$. On suppose que la suite (f_n) convergence* simplement *vers une fonction $f : \Omega \to \overline{\mathbb{R}}$* ou $\mathbb{C}$*. Alors f est mesurable. Avec des mots :* toute limite simple d'une suite de fonctions mesurables est une fonction mesurable.

Démonstration. Lorsque les f_n sont à valeurs dans $\overline{\mathbb{R}}$, le résultat est évident par le corollaire précédent puisqu'on a $f = \overline{\lim}\, f_n$. Pour des fonctions f_n à valeurs complexes, on applique le "cas $\overline{\mathbb{R}}$" aux fonctions $\mathrm{Re}(f_n)$ et $\mathrm{Im}(f_n)$, en se souvenant qu'une fonction à valeur complexe est mesurable si et seulement si sa partie réelle et sa partie imaginaire le sont. □

Remarque. En fait, toute limite simple d'une suite d'applications mesurables $f_n : \Omega \to \Omega'$ à valeurs dans un espace métrique Ω' quelconque est mesurable. Ce n'est pas complètement évident (on ne peut pas utiliser de limsup ou de liminf...). Un moyen de le voir est de vérifier que si $f_n \to f$ simplement et si C est un fermé de Ω', alors l'équivalence suivante a lieu pour tout $x \in \Omega$:

$$f(x) \in C \iff \forall \varepsilon > 0\ \exists N\ \forall n \geqslant N\ :\ \mathrm{dist}(f_n(x), C) \leqslant \varepsilon\,.$$

De là, il n'est pas difficile d'écrire $f^{-1}(C)$ comme intersection dénombrable de réunions dénombrables d'intersections dénombrables d'ensembles mesurables. Donc $f^{-1}(C)$ est mesurable pour tout fermé $C \subseteq \Omega'$; et donc f est mesurable puisque les fermés engendrent la tribu borélienne de Ω'.

Exercice. Soit Ω un espace métrique. Montrer que pour tout point $a \in \Omega$, on peut trouver une fonction continue $\phi_a : \Omega \to \mathbb{R}$ telle que $\phi_a(a) = 1$ et $0 \leqslant \phi_a(x) < 1$ pour $x \neq a$, et en déduire que si la topologie de Ω n'est pas discrète, alors il existe une suite de fonctions continues $f_n : \Omega \to \mathbb{R}$ qui converge simplement vers une fonction f non continue. (Considérer des fonctions de la forme $f_n(x) = \phi(x)^n$.)

3. Fonctions étagées

3.1. Mesurabilité d'une indicatrice. Rappelons que si Ω est un ensemble, on note $\mathbf{1}_E$ la fonction indicatrice d'une partie E de Ω :

$$\mathbf{1}_E(x) = \begin{cases} 1 & \text{si } x \in E \\ 0 & \text{si } x \notin E \end{cases}$$

On considérera $\mathbf{1}_E$ comme une fonction de Ω dans $\mathbb{R}$ (autrement dit, l'espace d'arrivée est $\mathbb{R}$ et pas $\{0,1\}$). Il sera bon d'avoir toujours en tête le fait suivant :

LEMME 3.1. *Soit $(\Omega, \mathcal{B})$ un espace mesurable, et soit $E \subseteq \Omega$. Alors $\mathbf{1}_E$ est $\mathcal{B}$-mesurable si et seulement si $E \in \mathcal{B}$.*

Démonstration. Si $\mathbf{1}_E$ est mesurable, alors $E = (1_E)^{-1}(\{1\}) \in \mathcal{B}$. Inversement, si $E \in \mathcal{B}$ alors $\mathbf{1}_E$ est mesurable “par recollement”. □

Remarque. Si on ne veut pas utiliser la Proposition 1.9, on peut aussi montrer directement que $\mathbf{1}_E$ est mesurable si E l'est : pour tout $A \subseteq \mathbb{R}$, l'ensemble $(\mathbf{1}_E)^{-1}(A)$ est égal soit à Ω (si $0 \in A$ et $1 \in A$), soit à E (si $0 \notin B$ et $1 \in B$), soit à E^c (si $0 \in A$ et $1 \notin A$), soit à $\varnothing$ (si $0 \notin A$ et $1 \notin A$) ; dans tous les cas $(\mathbf{1}_E)^{-1}(A)$ est mesurable.

3.2. Définition et caractérisations des fonctions étagées. Dans ce qui suit, $(\Omega, \mathcal{B})$ est un espace mesurable.

DÉFINITION 3.2. *Une fonction $\varphi : \Omega \to \mathbb{C}$ est dite **$\mathcal{B}$-étagée** si elle est $\mathcal{B}$-mesurable et ne prend qu'un nombre fini de valeurs. On notera $\mathcal{E}(\Omega, \mathcal{B})$, ou simplement $\mathcal{E}(\Omega)$, l'ensemble de toutes les fonctions $\mathcal{B}$-étagées $\varphi : \Omega \to \mathbb{C}$.*

Remarque 1. Lorsque la tribu $\mathcal{B}$ est claire dans le contexte, on dira simplement “étagée” au lieu de “$\mathcal{B}$-étagée”. En particulier, une fonction étagée sur un espace métrique Ω est une fonction *borélienne* ne prenant qu'un nombre fini de valeurs.

Remarque 2. Une précision importante : les fonctions étagées sont par définition *à valeurs numériques* : elles ne prennent pas les valeurs ∞ et $-\infty$.

Exemple 1. Si $[a, b]$ est un intervalle compact de $\mathbb{R}$, alors toute fonction *en escalier* sur $[a, b]$ est étagée.

Démonstration. Rappelons qu'une fonction $\varphi : [a, b] \to \mathbb{R}$ est dite en escalier s'il existe une subdivision $(s_0, \dots, s_N)$ de $[a, b]$ (i.e. $a = s_0 < \dots < s_N = b$) telle que φ soit constante sur chaque intervalle $]x_i, x_{i+1}[$, $0 \leqslant i \leqslant N - 1$. Une telle fonction φ est continue par morceaux, donc borélienne ; et comme φ ne prend qu'un nombre fini de valeurs, elle est donc étagée. □

Exemple 2. La fonction $\mathbf{1}_{\mathbb{Q}}$ est étagée sur $\mathbb{R}$, mais elle n'est en escalier sur aucun intervalle non trivial $[a, b]$.

Démonstration. Comme $\mathbb{Q}$ est un borélien de $\mathbb{R}$, la fonction $\mathbf{1}_{\mathbb{Q}}$ est borélienne d'après le lemme 3.1, donc étagée puisqu'elle ne prend que deux valeurs. Mais comme tout intervalle non trivial contient des rationnels *et* des irrationnels, $\mathbf{1}_{\mathbb{Q}}$ n'est constante sur aucun intervalle ouvert non vide, donc ne peut pas être en escalier sur un intervalle $[a, b]$ non trivial. □

LEMME 3.3. *Pour $\varphi : \Omega \to \mathbb{C}$, les propriétés suivantes sont équivalentes :*

(1) *φ est étagée ;*

(2) *il existe une partition finie $(\Omega_1, \dots, \Omega_N)$ de Ω en ensembles mesurables telle que φ est constante sur chaque Ω_i ;*

(3) *il existe $E_1, \dots, E_n \in \mathcal{B}$ deux à deux disjoints et $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ tels que*

$$\varphi = \sum_{i=1}^{n} \alpha_i \, \mathbf{1}_{E_i} ;$$

(4) *il existe* $E_1, \dots, E_n \in \mathcal{B}$ *et* $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ *tels que* $\varphi = \sum\limits_{i=1}^{n} \alpha_i \, \mathbf{1}_{E_i}$.

Démonstration. (1) $\Longrightarrow$ (2) supposons φ étagée, et notons $\alpha_1, \dots, \alpha_N$ les valeurs prises par φ (les α_i sont supposés deux à deux distinctes). Si on pose $\Omega_i := \{\varphi = \alpha_i\} = \{x \in \Omega;\ \varphi(x) = \alpha_i\}$, on obtient une partition de Ω, en ensembles mesurables car φ est mesurable ; et par définition φ est constante sur chaque Ω_i.

L'implication (2) $\Longrightarrow$ (3) est évidente : si (2) est vérifié, on peut prendre $n := N$ et $E_i := \Omega_i$ pour $i = 1, \dots, n$; le nombre α_i étant la valeur constante de φ sur $E_i = \Omega_i$. L'implication (3) $\Longrightarrow$ (4) est bien sûr évidente... et l'implication (4) $\Longrightarrow$ (1) aussi : si $\varphi = \sum_{i=1}^n \alpha_i \, \mathbf{1}_{E_i}$, alors φ est mesurable car les $\mathbf{1}_{E_i}$ le sont d'après le Lemme 3.1, et φ ne prend qu'un nombre fini de valeurs. □

Corollaire 3.4. $\mathcal{E}(\Omega)$ *est l'espace vectoriel engendré par les fonctions indicatrices d'ensembles mesurables.*

Démonstration. Il est clair *par définition* que $\mathcal{E}(\Omega)$ est un espace vectoriel, car les fonctions mesurables à valeurs complexes forment un espace vectoriel. Le fait que $\mathcal{E}(\Omega)$ soit engendré par les fonctions indicatrices est exactement ce que dit la condition (4). □

Remarque. En général, il y a plusieurs façons d'écrire une fonction étagée sous la forme $\sum_1^n \alpha_i \, \mathbf{1}_{E_i}$. Par exemple, si on prend $\Omega = \mathbb{R}$, alors $\mathbf{1}_{[0,1]} = \mathbf{1}_{[0,1/2[} + \mathbf{1}_{[1/2,1]} = 2\,\mathbf{1}_{[-1,1]} - \mathbf{1}_{[0,1]} - 2\,\mathbf{1}_{[-1,0[}$. Cependant, il y a une **écriture canonique** : en notant $\alpha_1, \dots, \alpha_n$ les valeurs distinctes prises par φ, on a

$$\varphi = \sum_{i=1}^{n} \alpha_i \, \mathbf{1}_{\{\varphi = \alpha_i\}} \, .$$

3.3. Approximation par des fonctions étagées. Le résultat suivant est crucial pour toute la théorie de l'intégration.

Théorème 3.5. *Soit* $(\Omega, \mathcal{B})$ *un espace mesurable. Pour toute fonction mesurable* $f : \Omega \to [0, \infty]$, *on peut trouver une suite* croissante *de fonctions étagées positives* (φ_n) *qui converge simplement vers* f *(i.e.* $\varphi_n(x) \to f(x)$ *pour tout* $x \in \Omega$*). On dit qu'une telle suite* (φ_n) *est une* ***suite approximante*** *pour* f.

Démonstration. On définit d'abord des fonctions ψ_n, $n \in \mathbb{N}^*$ de la façon suivante :

$$\psi_n(x) = \left\{ \begin{array}{ccc} 0 & \text{si} & 0 \leqslant f(x) < \frac{1}{n}, \\ \frac{1}{n} & \text{si} & \frac{1}{n} \leqslant f(x) < \frac{2}{n}, \\ \vdots & & \\ \frac{n-1}{n} & \text{si} & \frac{n-1}{n} < f(x) \leqslant 1, \\ 1 & \text{si} & 1 \leqslant f(x) < 1 + \frac{1}{n} \\ 1 + \frac{1}{n} & \text{si} & 1 + \frac{1}{n} \leqslant f(x) < 1 + \frac{2}{n}, \\ \vdots & & \\ 1 + \frac{n-1}{n} & \text{si} & 1 + \frac{n-1}{n} \leqslant f(x) < 2, \\ 2 & \text{si} & 2 \leqslant f(x) < 2 + \frac{1}{n}, \\ \vdots & & \\ n - 1 + \frac{n-1}{n} & \text{si} & n - 1 + \frac{n-1}{n} \leqslant f(x) < n, \\ n & \text{si} & f(x) \geqslant n. \end{array} \right.$$

De manière plus condensée :

$$\psi_n(x) = \begin{cases} n & \text{si} \qquad f(x) \geqslant n, \\ i + \frac{k}{n} & \text{si} \quad i + \frac{k}{n} \leqslant f(x) < i + \frac{k+1}{n} \end{cases} \quad \text{avec } 0 \leqslant i, k \leqslant n-1.$$

Les ψ_n ne prennent qu'un nombre fini de valeurs, et elles sont mesurables “par recollement” car pour $n \in \mathbb{N}^*$ fixé, tous les ensembles

$$\Omega_{i,k} := \left\{ x \in \Omega;\ i + \frac{k}{n} \leqslant f(x) < i + \frac{k+1}{n} \right\} = f^{-1}\left(\left[i + \frac{k}{n}, i + \frac{k+1}{n} \right[\right)$$

sont mesurables et ψ_n est constante sur chacun de ces ensembles ; donc les ψ_n sont étagées. Il est également évident qu'on a $0 \leqslant \psi_n(x) \leqslant f(x)$ pour tout n et pour tout $x \in \Omega$.

De plus $\psi_n(x)$ tend vers $f(x)$ quand $n \to \infty$, pour tout $x \in \Omega$. En effet, si $f(x) = \infty$ alors $\psi_n(x) = n$ pour tout $n \geqslant 1$, et donc $\psi_n(x) \to \infty = f(x)$; et si $f(x) < \infty$, on peut trouver un entier $N(x)$ tel que $f(x) < N(x)$, et on a alors $0 \leqslant f(x) - u_n(x) \leqslant \frac{1}{n}$ pour tout $n \geqslant N(x)$.

La suite (ψ_n) n'a aucune raison d'être croissante ; mais si on pose

$$\varphi_n(x) = \max(\psi_1(x), \dots, \psi_n(x)) \,,$$

alors les φ_n sont toujours étagées, la suite (φ_n) *est* croissante, et elle converge également vers f car $\psi_n \leqslant \varphi_n \leqslant f$ pour tout $n \geqslant 1$. □

Remarque 3.6. Si la fonction f est à valeurs réelles et *bornée*, alors la suite (φ_n) définie dans la preuve du théorème converge en fait *uniformément* vers f.

Démonstration. Avec les notations de la preuve du théorème, l'entier $N(x)$ peut être choisi *indépendemment* de $x \in \Omega$. Par conséquent, la suite (ψ_n) converge uniformément ver f, et donc (φ_n) aussi puisque $\psi_n \leqslant \varphi_n \leqslant f$. □

Chapitre 4

Intégration abstraite, et moins abstraite

Dans tout ce chapitre, $(\Omega, \mathcal{B}, \mu)$ est un espace mesuré fixé une fois pour toutes.

1. Intégrale des fonctions étagées positives

1.1. Définition et propriétés essentielles. Dans tout ce qui suit, on notera $\mathcal{E}^+(\Omega)$ l'ensemble des fonctions étagées *positives* sur Ω :

$$\mathcal{E}^+(\Omega) = \{\varphi : (\Omega, \mathcal{B}) \to \mathbb{R};\ \varphi \text{ étagée} \geqslant 0\}\,.$$

On devrait en fait écrire $\mathcal{E}^+(\Omega, \mathcal{B})$; mais on ne le fera pas, pour alléger la notation.

Définition 1.1. *Soit $\varphi \in \mathcal{E}^+(\Omega)$, d'écriture canonique*

$$\varphi = \sum_{i=1}^{N} \alpha_i \mathbf{1}_{\{\varphi = \alpha_i\}}\,.$$

*L'**intégrale de** φ **sur** Ω **par rapport à** μ est le nombre $\int_\Omega \varphi\, d\mu \in [0, \infty]$ défini par*

$$\int_\Omega \varphi\, d\mu = \sum_{i=1}^{N} \alpha_i\, \mu\left(\{\varphi = \alpha_i\}\right).$$

Autres notations. On peut aussi écrire $\int_\Omega \varphi(x)\, d\mu(x)$; ou $\int_\Omega \varphi$ si on veut une notation la plus courte possible.

Exemple. Supposons que Ω soit un ensemble fini, et soit μ_c la mesure de comptage sur Ω. Alors toute fonction $\varphi : \Omega \to \mathbb{R}^+$ est étagée, et on a

$$\int_\Omega \varphi\, d\mu_c = \sum_{x \in \Omega} \varphi(x)\,.$$

Une autre manière d'écrire la même chose : si $c_1, \dots, c_N$ sont des nombres réels positifs, alors

$$\sum_{i=1}^{N} c_i = \int_{\{1, \dots, N\}} \varphi\, d\mu_c\,,$$

où $\varphi : \{1, \dots, N\} \to \mathbb{R}^+$ est la fonction définie par $\varphi(i) = c_i$ pour $i = 1, \dots, N$.

Démonstration. La preuve consiste à formaliser l'observation suivante. Si on a dans son porte-monnaie une certain nombre de pièces, de différentes valeurs, il y a au moins deux façons de compter son argent : on peut additionner une à une les valeurs des diffèrentes pièces ; mais on peut aussi commencer par regrouper les pièces selon leur valeur, puis compter le nombre de pièces ayant une valeur donnée, multiplier par la valeur en question, et tout ajouter.

Soient $\alpha_1, \dots, \alpha_N$ les valeurs distinctes prises par φ, et pour $i = 1, \dots, N$ posons $\Omega_i := \{\varphi = \alpha_i\}$. L'écriture canonique de φ est donc $\varphi = \sum_1^N \alpha_i \mathbf{1}_{\Omega_i}$, de sorte que $\int_\Omega \varphi\, d\mu_c = \sum_1^N \alpha_i\, \#\Omega_i$.

Comme les Ω_i forment une partition de Ω, on a

$$\begin{aligned}\sum_{x\in\Omega}\varphi(x) &= \sum_{i=1}^N\sum_{x\in\Omega_i}\underbrace{\varphi(x)}_{\alpha_i}\\ &= \sum_{i=1}^N \alpha_i\,\#\Omega_i\\ &= \int_\Omega \varphi\,d\mu_c\,.\end{aligned}$$

□

Remarque 1. Pour tout ensemble mesurable $A\subseteq\Omega$, on a $\int_\Omega \mathbf{1}_A\,d\mu=\mu(A)$.

Remarque 2. A cause de la convention $0\times\infty=0$, on a $\int_\Omega 0\,d\mu=0$.

Remarque 3. On peut très bien avoir $\int_\Omega\varphi\,d\mu=\infty$; et on peut aussi avoir $\int_\Omega\varphi\,d\mu=0$ même si $\varphi\neq 0$. Par exemple, si $\Omega=\mathbb{R}$ et $\mu=\lambda_1$, alors $\int_\mathbb{R}\mathbf{1}\,d\lambda_1=\lambda_1(\mathbb{R})=\infty$ et $\int_\mathbb{R}\mathbf{1}_\mathbb{Q}\,d\lambda_1=\lambda_1(\mathbb{Q})=0$.

Exercice. Montrer qu'on a $\int_\Omega\varphi\,d\mu<\infty$ si et seulement si $\mu\left(\{x;\ \varphi(x)\neq 0\}\right)<\infty$, et $\int_\Omega\varphi\,d\mu=0$ si et seulement si $\mu\left(\{x;\ \varphi(x)\neq 0\}\right)=0$.

Lemme 1.2. *L'intégrale des fonctions étagées positives possède les propriétés suivantes.*

(1) ***Homogénéité** : si $\varphi\in\mathcal{E}^+(\Omega)$ et si $\lambda\in\mathbb{R}^+$, alors $\int_\Omega(\lambda\varphi)=\lambda\int_\Omega\varphi$.*

(2) ***Additivité** : si $\varphi,\psi\in\mathcal{E}^+(\Omega)$, alors $\int(\varphi+\psi)=\int\varphi+\int\psi$.*

Démonstration. La partie (1) est évidente (exercice). En revanche, la partie (2) ne l'est pas.

Introduisons la terminologie suivante : si $\varphi\in\mathcal{E}^+(\Omega)$, on appellera **partition adaptée à** φ toute partition *finie* $(\Omega_\lambda)_{\lambda\in\Lambda}$ de Ω en ensembles mesurables telle que φ soit constante sur chaque Ω_λ.

Fait 1. Soit $\varphi\in\mathcal{E}^+(\Omega)$. Si $(\Omega_\lambda)_{\lambda\in\Lambda}$ est une partition adaptée à φ et si on note c_λ la valeur constante de φ sur Ω_λ, alors $\int_\Omega\varphi\,d\mu=\sum_{\lambda\in\lambda}c_\lambda\,\mu(\Omega_\lambda)$.

Preuve du Fait 1. Soient $\alpha_1,\dots,\alpha_N$ les valeurs distinctes prises par φ. Pour $i=1,\dots,N$, posons $\Lambda_i=\{\lambda\in\Lambda;\ c_\lambda=\alpha_i\}$. Alors l'écriture canonique de φ est

$$\varphi=\sum_{i=1}^N\alpha_i\,\mathbf{1}_{E_i},\quad \text{où } E_i=\bigcup_{\lambda\in\Lambda_i}\Omega_\lambda\,.$$

On a donc

$$\begin{aligned}
\int_\Omega \varphi\, d\mu &= \sum_{i=1}^N \alpha_i\, \mu\left(\bigcup_{\lambda\in\Lambda_i} \Omega_\lambda\right)\\
&= \sum_{i=1}^N \alpha_i \sum_{\lambda\in\Lambda_i} \mu(\Omega_\lambda) \quad \text{car les } \Omega_\lambda \text{ sont disjoints}\\
&= \sum_{i=1}^N \sum_{\lambda\in\Lambda_i} c_\lambda\, \mu(\Omega_\lambda) \quad \text{car } \alpha_i = c_\lambda \text{ si } \lambda\in\Lambda_i\\
&= \sum_{\lambda\in\Lambda} c_\lambda\, \mu(\Omega_\lambda) \quad \text{car } (\Lambda_1,\dots,\Lambda_N) \text{ est une partition de } \Lambda.
\end{aligned}$$

□

FAIT 2. Si $\varphi, \psi \in \mathcal{E}^+(\Omega)$, on peut trouver une partition $(\Omega_\lambda)_{\lambda\in\Lambda}$ adaptée à la fois à φ et à ψ.

Preuve du Fait 2. Soient $(A_i)_{i\in I}$ une partition adaptée à φ, et $(B_j)_{j\in J}$ une partition adaptée à ψ. Si on pose $\Omega_{i,j} := A_i \cap B_j$, alors les $\Omega_{i,j}$ sont mesurables, deux-à-deux disjoints, et on a $\bigcup_{i,j} \Omega_{i,j} = \Omega$. Les $\Omega_{i,j}$ non vides forment donc une partition de Ω, qui est visiblement adaptée à φ et à ψ. □

On peut maintenant démontrer (2). Soient $\varphi, \psi \in \mathcal{E}^+(\Omega)$, et soit $(\Omega_\lambda)_{\lambda\in\Lambda}$ une partition adaptée à la fois à φ et à ψ donnée par le Fait 2. Pour $\lambda\in\Lambda$, notons c_λ et d_λ les valeurs constantes de φ et ψ sur Ω_λ. Alors $\varphi+\psi \equiv c_\lambda + d_\lambda$ sur Ω_λ, $\lambda\in\Lambda$, donc on a d'après le Fait 1 :

$$\begin{aligned}
\int_\Omega (\varphi+\psi) &= \sum_{\lambda\in\Lambda} (c_\lambda + d_\lambda)\, \mu(\Omega_\lambda)\\
&= \sum_{\lambda\in\Lambda} c_\lambda\, \mu(\Omega_\lambda) + \sum_{\lambda\in\Lambda} d_\lambda\, \mu(\Omega_\lambda)\\
&= \int_\Omega \varphi + \int_\Omega \psi\,.
\end{aligned}$$

□

COROLLAIRE 1.3. *Si $\varphi \in \mathcal{E}^+(\Omega)$, alors, quelle que soit la manière d'écrire*

$$\varphi = \sum_{i=1}^n \alpha_i\, \mathbf{1}_{E_i} \qquad \textit{avec } E_i \in \mathcal{B} \textit{ et } \alpha_i \in \mathbb{R}^+,$$

on a

$$\int_\Omega \varphi = \sum_{i=1}^n \alpha_i\, \mu(E_i)\,.$$

Démonstration. Par le lemme, on a $\int_\Omega \varphi = \sum_{i=1}^n \int_\Omega \alpha_i\, \mathbf{1}_{E_i} = \sum_{i=1}^n \alpha_i\, \mu(E_i)$. □

COROLLAIRE 1.4. *L'intégrale $\int_\Omega \varphi$ est **croissante** par rapport à φ : Si $\varphi_1, \varphi_2 \in \mathcal{E}^+(\Omega)$ et si $\varphi_1 \leqslant \varphi_2$, alors $\int_\Omega \varphi_1 \leqslant \int_\Omega \varphi_2$.*

Démonstration. La fonction $\varphi := \varphi_2 - \varphi_1$ est étagée (les fonctions étagées forment un espace vectoriel), et elle est positive par hypothèse ; autrement dit $\varphi \in \mathcal{E}^+(\Omega)$. Comme $\varphi_2 = \varphi_1 + \varphi$, on a donc $\int_\Omega \varphi_2 = \int_\Omega \varphi_1 + \int_\Omega \varphi \geqslant \int_\Omega \varphi_1$. □

Exercice. Soit $[a,b]$ un segment de $\mathbb{R}$. Montrer que si $\varphi : [a,b] \to \mathbb{R}^+$ est *en escalier*, alors

$$\int_{[a,b]} \varphi \, d\lambda_1 = \int_a^b \varphi(x)\, dx \quad \text{(intégrale au sens “usuel”)}\,.$$

1.2. Intégrale sur un sous-ensemble. Rappelons que si $E \subseteq \Omega$, on note $\mathcal{B}_E \subseteq \mathcal{P}(E)$ la *tribu trace* induite par $\mathcal{B}$ sur E,

$$\mathcal{B}_E = \{B \cap E;\ B \in \mathcal{B}\}\,.$$

Si E est *mesurable*, on a $\mathcal{B}_E \subseteq \mathcal{B}$, et donc en fait

$$\mathcal{B}_E = \mathcal{B} \cap \mathcal{P}(E) = \{A \in \mathcal{B};\ A \subseteq E\}\,.$$

On peut donc considérer la restriction μ_E de la mesure μ à $\mathcal{B}_E$, et on obtient ainsi un espace mesuré $(E, \mathcal{B}_E, \mu_E)$.

On a vu que si Si $\varphi : \Omega \to \mathbb{R}$ est mesurable, alors sa restriction $\varphi_{|E}$ à E est $\mathcal{B}_E$-mesurable (voir la remarque suivant la Proposition 1.9 du Chapitre 3). En particulier, si $\varphi \in \mathcal{E}^+(\Omega)$, alors $\varphi_{|E} \in \mathcal{E}^+(E, \mathcal{B}_E)$; donc on peut définir l'intégrale $\int_E (\varphi_{|E})\, d\mu_E$. Pour alléger les notations, on adoptera la convention suivante.

Convention. Si $\varphi \in \mathcal{E}^+(\Omega)$, on écrit $\int_E \varphi\, d\mu$ au lieu de $\int_E (\varphi_{|E})\, d\mu_E$.

Le fait suivant sera d'usage constant.

Fait 1.5. *Si $\varphi \in \mathcal{E}^+(\Omega)$, alors on a pour tout ensemble mesurable $E \subseteq \Omega$:*

$$\int_E \varphi\, d\mu = \int_\Omega \mathbf{1}_E\, \varphi\, d\mu\,.$$

Démonstration. C'est essentiellement évident, mais casse-pieds à écrire. Si $\varphi = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i}$ avec $A_i \in \mathcal{B}$, alors $\varphi_{|E} = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i \cap E}$ où les $\mathbf{1}_{A_i\cap E}$ sont considérées comme des fonctions *sur* E, et $\mathbf{1}_E\, \varphi = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i \cap E}$ où les $\mathbf{1}_{A_i\cap E}$ sont considérées comme des fonctions *sur* Ω ; donc

$$\int_E \varphi\, d\mu = \int_E (\varphi_{|E})\, d\mu_E = \sum_{i=1}^n \alpha_i\, \mu_E(A_i \cap E) = \sum_{i=1}^n \alpha_i\, \mu(A_i \cap E) = \int_\Omega \mathbf{1}_E\, \varphi\, d\mu\,.$$

$\square$

Exercice 1. Montrer que si $\varphi \in \mathcal{E}^+(\Omega)$, alors l'intégrale $\int_E \varphi$ est croissante par rapport à E : si $E_1 \subseteq E_2$, alors $\int_{E_1} \varphi \leqslant \int_{E_2} \varphi$.

Exercice 2. Montrer que si $A, B \in \mathcal{B}$ vérifient $A \cap B = \varnothing$ et si $\varphi \in \mathcal{E}^+(\Omega)$, alors

$$\int_{A\cup B} \varphi = \int_A \varphi + \int_B \varphi\,.$$

2. Intégrale des fonctions mesurables positives

2.1. Définitions et remarques. Une **fonction mesurable positive** est par définition une fonction mesurable à valeurs dans $[0,\infty]$ (*la valeur ∞ est autorisée*). On notera $\mathfrak{M}^+(\Omega)$ l'ensemble des fonctions mesurables positives sur Ω :

$$\mathfrak{M}^+(\Omega) = \{f : (\Omega, \mathcal{B}) \to [0,\infty] \text{ mesurables}\}\,.$$

Comme pour les fonctions étagées, on devrait en fait écrire $\mathfrak{M}^+(\Omega, \mathcal{B})$.

La définition de l'intégrale d'une fonction mesurable positive est très vite donnée :

DÉFINITION 2.1. *Pour toute fonction* $f \in \mathfrak{M}^+(\Omega)$*, on pose*

$$\int_\Omega f\,d\mu := \sup\left\{\int_\Omega \varphi\,d\mu;\ \varphi \in \mathcal{E}^+(\Omega)\ ,\ 0 \leqslant \varphi \leqslant f\right\}.$$

On peut également écrire $\int_\Omega f(x)\,d\mu(x)$*, ou* $\int_\Omega f$*.*

Remarque 1. Par définition, $\int_\Omega f$ est un *nombre positif*, i.e. $\int_\Omega f \in [0,\infty]$.

Remarque 2. Si f est *étagée* positive, on retrouve la définition donnée dans la section précédente.

Démonstration. C'est clair par *croissance* de l'intégrale pour les fonctions étagées positives (Corollaire 1.4). □

Remarque 3. L'intégrale $\int_\Omega f$ est **croissante** par rapport à f : si $f, g \in \mathfrak{M}^+(\Omega)$ et si $f \leqslant g$, alors $\int_\Omega f \leqslant \int_\Omega g$.

Démonstration. Si $f \leqslant g$, alors il y a plus de fonctions étagés positives minorant g que de fonctions étagées positives minorant f, donc le sup apparaissant dans la définition est plus grand pour g que pour f. □

EXERCICE 2.2. Montrer que si $f \in \mathfrak{M}^+(\Omega)$, alors $\int_\Omega(\lambda f) = \lambda \int_\Omega f$ pour tout $\lambda \in \mathbb{R}^+$.

2.2. Le "Théorème de convergence monotone". Le résultat suivant est certainement le le plus important de toute la théorie.

THÉORÈME 2.3. (convergence monotone)
Soit $(f_n)_{n\in\mathbb{N}}$ *une suite de fonctions mesurables positives,* $f_n : (\Omega, \mathcal{B}) \to [0,\infty]$*. On suppose que la suite* (f_n) *est* croissante*, i.e.* $f_n(x) \leqslant f_{n+1}(x)$ *pour tout* n *et pour tout* $x \in \Omega$*. Alors, en posant* $f(x) := \lim_{n\to\infty} f_n(x)$ *(qui existe dans* $[0,\infty]$*), la fonction* $f : \Omega \to [0,\infty]$ *est mesurable et on a*

$$\int_\Omega f\,d\mu = \lim_{n\to\infty}\int_\Omega f_n\,d\mu\,.$$

Autrement dit, on peut toujours permuter *la limite et l'intégrale dès lors qu'on a affaire à une suite* croissante *de fonctions* positives *:*

$$\int_\Omega (\lim f_n)\,d\mu = \lim_{n\to\infty}\int_\Omega f_n\,d\mu\,.$$

Démonstration. On a besoin du fait suivant, qui sera généralisé plus bas (Lemme 2.8).

FAIT. Soit φ une fonction *étagée* positive sur Ω. Si (B_n) est une suite croissante d'ensembles mesurables telle que $\bigcup_0^\infty B_n = \Omega$, alors $\int_{B_n}\varphi\,d\mu \to \int_\Omega \varphi\,d\mu$ quand $n\to\infty$.

Preuve du Fait. Écrivons $\varphi = \sum_{i=1}^N \alpha_i\,\mathbf{1}_{E_i}$, où les E_i sont des ensembles mesurables et les α_i sont des constantes positives. Alors $\int_\Omega \varphi\,d\mu = \sum_{i=1}^N \alpha_i\mu(E_i)$. De plus, on a $\int_{B_n}\varphi\,d\mu = \int_\Omega \mathbf{1}_{B_n}\varphi\,d\mu$, et $\mathbf{1}_{B_n}\varphi = \sum_{i=1}^N \alpha_i\mathbf{1}_{B_n\cap E_i}$; donc

$$\int_{B_n}\varphi\,d\mu = \sum_{i=1}^N \alpha_i\mu(B_n\cap E_i)\,.$$

Mais pour $i \in \{1,\dots,N\}$ fixé, la suite $(B_n\cap E_i)$ croit vers E_i. Donc $\mu(B_n\cap E_i)\to\mu(E_i)$ pour $i = 1,\dots,N$ quand $n\to\infty$, et donc $\int_{B_n}\varphi\,d\mu \to \sum_{i=1}^N \alpha_i\mu(E_i) = \int_\Omega\varphi\,d\mu$. □

Démontrons maintenant le Théorème de convergence monotone.

La fonction f est mesurable en tant que limite simple de fonctions mesurables ; et bien sûr f est positive.

Comme la suite (f_n) est croissante, la suite des intégrales $\left(\int_\Omega f_n\right)$ est croissante et admet donc une limite dans $[0,\infty]$. De plus, on a $f_n \leqslant f$ pour tout n puisque (f_n) croit vers f, donc $\int_\Omega f_n \leqslant \int_\Omega f$ pour tout n et donc

$$\lim_{n\to\infty} \int_\Omega f_n\, d\mu \leqslant \int_\Omega f\, d\mu\,.$$

Le point délicat est l'inégalité inverse $\int_\Omega f\, d\mu \leqslant \lim_{n\to\infty} \int_\Omega f_n\, d\mu$.

Par définition de l'intégrale $\int_\Omega f\, d\mu$, il s'agit de voir qu'on a

$$\int_\Omega \varphi\, d\mu \leqslant \lim_{n\to\infty} \int_\Omega f_n\, d\mu$$

pour toute fonction φ *étagée* vérifiant $0 \leqslant \varphi \leqslant f$. Fixons une telle fonction φ.

Soit $\varepsilon \in\,]0,1[$ quelconque. Pour $n \in \mathbb{N}$, posons

$$B_n := \{f_n \geqslant (1-\varepsilon)\,\varphi\} = \big\{x \in \Omega;\; f_n(x) \geqslant (1-\varepsilon)\,\varphi(x)\big\}.$$

Les B_n sont des ensembles mesurables car $B_n = \{g_n \geqslant 0\}$, où $g_n := f_n - (1-\varepsilon)\,\varphi$ est une fonction mesurable bien définie à valeurs dans $\overline{\mathbb{R}}$, en tant que somme d'une fonction mesurable à valeurs dans $\overline{\mathbb{R}}$ et d'une fonction mesurable *à valeurs réelles* (voir la remarque après le Corollaire 2.2 du Chapitre 3). De plus, la suite (B_n) est croissante car la suite (f_n) l'est. Enfin, on a

$$\bigcup_{n=0}^{\infty} B_n = \Omega\,.$$

Soit en effet $x \in \Omega$ quelconque. Si $\varphi(x) = 0$ alors $x \in B_0$ car $f_0 \geqslant 0$. Si $\varphi(x) > 0$, alors $\varphi(x) > (1-\varepsilon)\,\varphi(x)$, donc $f(x) > (1-\varepsilon)\,\varphi(x)$ puisque $f \geqslant \varphi$, et donc on peut trouver un entier n tel que $f_n(x) > (1-\varepsilon)\,\varphi(x)$ (d'où $x \in B_n$) puisque $f_n(x) \to f(x)$.

D'après le Fait, on en déduit qu'on a

$$\int_\Omega \varphi\, d\mu = \lim_{n\to\infty} \int_{B_n} \varphi\, d\mu\,.$$

Mais par définition de B_n, on a $f_n \geqslant (1-\varepsilon)\,\varphi$ sur B_n, et donc $\mathbf{1}_{B_n} f_n \geqslant (1-\varepsilon)\,\mathbf{1}_{B_n}\varphi$. *A fortiori* $f_n \geqslant (1-\varepsilon)\,\mathbf{1}_{B_n}\varphi$, et donc $\int_\Omega f_n\, d\mu \geqslant \int_\Omega (1-\varepsilon)\,\mathbf{1}_{B_n}\varphi\, d\mu = (1-\varepsilon)\int_{B_n} \varphi\, d\mu$. Ainsi,

$$\int_{B_n} \varphi\, d\mu \leqslant \frac{1}{1-\varepsilon} \int_\Omega f_n\, d\mu \qquad \text{pour tout } n \in \mathbb{N}.$$

En faisant tendre n vers l'infini, on en déduit

$$\int_\Omega \varphi\, d\mu \leqslant \frac{1}{1-\varepsilon} \lim_{n\to\infty} \int_\Omega f_n\, d\mu \qquad \text{pour tout } \varepsilon \in\,]0,1[\,;$$

d'où $\int_\Omega \varphi\, d\mu \leqslant \lim_{n\to\infty} \int_\Omega f_n\, d\mu$ en faisant maintenant tendre ε vers 0. □

On reviendra au Chapitre 6 sur les conséquences "pratiques" du théorème de convergence monotone. Pour le moment, on va juste en déduire quelques faits indispensable pour la suite du présent chapitre.

Corollaire 2.4. *Si $f \in \mathfrak{M}^+(\Omega)$, alors $\int_\Omega f = \lim_{n\to\infty} \int_\Omega \varphi_n$, quelle que soit la suite approximante $(\varphi_n) \subseteq \mathcal{E}^+(\Omega)$ pour f.*

Corollaire 2.5. *L'intégrale des fonctions mesurables positives possède les propriétés suivantes.*

(1) ***Homogénéité*** *: si* $f \in \mathfrak{M}^+(\Omega)$*, alors* $\int_\Omega (\lambda f) = \lambda \int_\Omega f$ *pour tout* $\lambda \in [0, \infty]$.

(2) ***Additivité*** *: si* $f, g \in \mathfrak{M}^+(\Omega)$*, alors* $\int_\Omega (f+g) = \int_\Omega f + \int_\Omega g$.

Démonstration. (1) On a déjà vu que le résultat est vrai si $\lambda < \infty$ (Exercice 2.2) ; il reste donc à traiter le cas $\lambda = \infty$. Pour $n \in \mathbb{N}$, posons $f_n := n\,f$. La suite (f_n) est croissante car f est positive, et $f_n(x) \to \infty \times f(x)$ pour tout $x \in \Omega$, grâce à la convention $\infty \times 0 = 0$. Par convergence monotone et le cas $\lambda < \infty$, on a donc

$$\int_\Omega (\infty \times f) = \lim_{n\to\infty} \int_\Omega (n\,f) = \lim_{n\to\infty} n \int_\Omega f = \infty \times \int_\Omega f\,,$$

où on a à nouveau utilisé la convention $\infty \times 0 = 0$.

(2) Soient $(\varphi_n), (\psi_n) \subseteq \mathcal{E}^+(\Omega)$ des suites approximantes pour f et g. Alors $(\varphi_n + \psi_n)$ est une suite approximante pour $f+g$, donc

$$\begin{aligned} \int_\Omega (f+g) &= \lim_{n\to\infty} \int_\Omega (\varphi_n + \psi_n) \\ &= \lim_{n\to\infty} \left(\int_\Omega \varphi_n + \int_\Omega \psi_n \right) \\ &= \int_\Omega f + \int_\Omega g\,, \end{aligned}$$

où on a utilisé à la deuxième ligne l'additivité de l'intégrale *pour les fonctions étagées.* □

Remarque 2.6. Il est très important de retenir le principe de la démonstration précédente : si on veut démontrer une certaine propriété concernant l'intégrale d'une fonction positive quelconque, *il suffit de le faire pour une fonction* étagée ; la conclusion sera alors "gratuite" par approximation et convergence monotone.

Maintenant qu'on sait que l'intégrale est additive, on peut réappliquer le théorème de convergence monotone pour obtenir un résultat d'"interversion série-intégrale" optimal :

Corollaire 2.7. *Si* $(u_k)_{k\in\mathbb{N}}$ *est une suite quelconque de fonctions mesurables positives, alors*

$$\int_\Omega \left(\sum_{k=0}^\infty u_k \right) d\mu = \sum_{k=0}^\infty \int_\Omega u_k\, d\mu\,.$$

Démonstration. La fonction $f := \sum_0^\infty u_k$ est bien définie (à valeurs dans $[0,\infty]$), et on a $f = \lim_{n\to\infty} f_n$ où $f_n := \sum_{k=0}^n u_k$. La suite (f_n) est croissante car les u_k sont positives, donc on a par convergence monotone *et additivité de l'intégrale* :

$$\int_\Omega f = \lim_{n\to\infty} \int_\Omega f_n = \lim_{n\to\infty} \sum_{k=0}^n \int_\Omega u_k = \sum_{k=0}^\infty \int_\Omega u_k\,.$$

□

2.3. Intégrale sur un sous-ensemble. On adopte la même convention que pour les fonctions étagées : si $E \subseteq \Omega$ et si $f \in \mathfrak{M}^+(\Omega)$, on écrit $\int_E f\, d\mu$ au lieu de $\int_E (f_{|E})\, d\mu_E$. On vérifie qu'on a encore

$$\int_E f\, d\mu = \int_\Omega \mathbf{1}_E\, f\, d\mu\,.$$

Remarque. Si $f \in \mathfrak{M}^+(\Omega)$, alors l'intégrale $\int_E f\, d\mu$ est croissante par rapport à E : si $E_1 \subseteq E_2$, alors $\int_{E_1} f \leqslant \int_{E_2} f$.

Démonstration. C'est évident par croissance de l'intégrale par rapport à la fonction intégrée, puisque $\mathbf{1}_{E_1} f \leqslant \mathbf{1}_{E_2} f$. (On utilise bien entendu le fait que la fonction f est *positive*.) □

LEMME 2.8. *Si $f \in \mathfrak{M}^+(\Omega)$, alors on définit une mesure sur $(\Omega, \mathcal{B})$ en posant*

$$\nu_f(A) := \int_A f\, d\mu\,.$$

On dit que ν_f est la ***mesure de densité*** *f par rapport à la mesure μ.*

Démonstration. On a $\nu_f(\varnothing) = 0$ car $\mathbf{1}_\varnothing f = 0$. Si $(A_k)_{k\in\mathbb{N}}$ est une suite d'ensembles mesurables deux à deux disjoints, alors

$$\mathbf{1}_{\bigcup_{k=0}^\infty A_k} = \sum_{k=0}^\infty \mathbf{1}_{A_k};$$

donc

$$\nu_f\left(\bigcup_{k=0}^\infty A_k\right) = \int_\Omega \left(\sum_{k=0}^\infty \mathbf{1}_{A_k} f\right) d\mu = \sum_{k=0}^\infty \int_\Omega \mathbf{1}_{A_k} f = \sum_{k=0}^\infty \nu_f(A_k).$$

□

COROLLAIRE 2.9. *Si $f \in \mathfrak{M}^+(\Omega)$ et si $A, B \subseteq \Omega$ sont des ensembles mesurables tels $A \cap B = \varnothing$, alors*

$$\int_{A\cup B} f = \int_A f + \int_B f\,.$$

Exercice 1. Soit $f \in \mathfrak{M}^+(\Omega)$. Montrer que pour toute fonction $g \in \mathfrak{M}^+(\Omega)$, on a $\int_\Omega g\, d\nu_f = \int_\Omega gf\, d\mu$.

Exercice 2. Démontrer le Corollaire 2.9 en utilisant uniquement la définition de l'intégrale. (Commencer par vérifier que pour une fonction $\varphi \in \mathcal{E}^+(\Omega)$, il revient au même de dire qu'on a $\varphi \leqslant \mathbf{1}_{A\cup B}\, f$ ou que φ peut s'écrire sous la forme $\varphi = \varphi_1 + \varphi_2$, avec $\varphi_1, \varphi \in \mathcal{E}^+(\Omega)$ vérifiant $\varphi_1 \leqslant \mathbf{1}_A\, f$ et $\varphi_2 \leqslant \mathbf{1}_B\, f$.)

Le lemme suivant sera souvent utilisé. Il est d'une grande importance en théorie des probabilités.

LEMME 2.10. (inégalité de Markov)
Si $f \in \mathfrak{M}^+(\Omega)$, on a pour tout nombre réel $\alpha > 0$:

$$\mu\Big(\{f \geqslant \alpha\}\Big) \leqslant \frac{1}{\alpha} \int_{\{f\geqslant\alpha\}} f\, d\mu \leqslant \frac{1}{\alpha} \int_\Omega f\, d\mu\,.$$

Démonstration. Posons $A = \{f \geqslant \alpha\}$. Par définition de A, on a $f \geqslant \alpha\,\mathbf{1}$ sur A, et donc (par croissance de l'intégrale)

$$\int_A f\,d\mu \geqslant \int_A \alpha\,\mathbf{1}\,d\mu = \alpha\,\mu(A)\,;$$

d'où la première inégalité. La seconde est évidente par croissance de $\int_E f$ par rapport à E. □

Exercice. On suppose que $\mu(\Omega) < \infty$. Soit (α_n) une suite de réels positifs telle que, pour une certaine constante $C > 0$, la série $\sum e^{-C\lambda_n}$ est convergente. Montrer que si $u : \Omega \to \mathbb{R}$ est une fonction mesurable bornée, alors $\sum_0^\infty \mu\big(\{|u| \geqslant \alpha_n\}\big) < \infty$. (Commencer par observer qu'on a $\int_\Omega e^{|u|}\,d\mu < \infty$.)

2.4. Rôle des ensembles de mesure nulle. On va voir ici que pour tout ce qui concerne l'intégration, les ensembles négligeables "ne comptent pas". (C'est précisément ce qui les rend importants.) Pour des raisons esthétiques, on écrira souvent "presque partout" au lieu de "μ-presque partout".

Lemme 2.11. *Pour toute fonction $f \in \mathfrak{M}^+(\Omega)$, on a l'équivalence*

$$\int f\,d\mu = 0 \iff f(x) = 0 \ \text{ presque partout}\,.$$

Démonstration. Supposons qu'on ait $f(x) = 0$ μ-presque partout, et posons $E := \{f = 0\} = \{x;\ f(x) = 0\}$. L'ensemble E est mesurable, et on a $\mathbf{1}_E f = 0$ par définition. Donc $f = \mathbf{1}_{\Omega\setminus E}\, f$, et par conséquent

$$\int_\Omega f = \int_\Omega \mathbf{1}_{\Omega\setminus E} f\,d\mu \leqslant \left(\int_\Omega \infty \times \mathbf{1}_{\Omega\setminus E}\right) d\mu = \infty \times \mu(\Omega\setminus E) = 0\,.$$

Inversement, supposons qu'on ait $\int_\Omega f\,d\mu = 0$. Pour tout $n \in \mathbb{N}^*$, l'inégalité de Markov (Lemme 2.10 avec $\alpha = 1/n$) donne

$$\mu\Big(\{f \geqslant 1/n\}\Big) \leqslant n \times \int_\Omega f\,d\mu = 0\,.$$

Donc $\mu\Big(\{f \geqslant 1/n\}\Big) = 0$ pour tout $n \in \mathbb{N}^*$, autrement dit $f(x) < 1/n$ presque partout. Comme $f(x) = 0 \iff \forall n \in \mathbb{N}^* \ :\ f(x) < 1/n$ (puisque $f \geqslant 0$), on a donc $f(x) = 0$ presque partout □

Remarque. On ne peut pas espérer avoir $f(x) = 0$ *partout* si $\int_\Omega f\,d\mu = 0$: par exemple, $\int_\mathbb{R} \mathbf{1}_\mathbb{Q}\,d\lambda_1 = 0$. De plus, il est essentiel de prendre garde au fait que le lemme n'est valable que pour des fonctions *positives.*

Exercice. Soit I un intervalle (non trivial) de $\mathbb{R}$. Montrer que si $f : I \to \mathbb{R}^+$ est une fonction *continue* positive vérifiant $\int_I f\,d\lambda_1 = 0$, alors $f = 0$.

Corollaire 2.12. *Si $N \in \mathcal{B}$ vérifie $\mu(N) = 0$, alors $\int_N f\,d\mu = 0$ pour toute $f \in \mathfrak{M}^+(\Omega)$, et donc $\int_\Omega f\,d\mu = \int_{\Omega\setminus N} f\,d\mu$.*

Démonstration. La première égalité est claire puisque $\int_N f\,d\mu = \int_\Omega \mathbf{1}_N f\,d\mu$ et $\mathbf{1}_N f = 0$ μ-presque partout ; et la deuxième en découle par additivité relativement au domaine d'intégration : $\int_\Omega = \int_N + \int_{\Omega\setminus N}$. □

Corollaire 2.13. *Si $f, g \in \mathfrak{M}^+(\Omega)$ vérifient $f(x) = g(x)$ presque partout, alors $\int_\Omega f\,d\mu = \int_\Omega g\,d\mu$.*

Démonstration. Soit $E := \{x \in \Omega;\ f(x) = g(x)\}$. L'ensemble E est mesurable car f et g le sont (exercice), et on a $\mu(\Omega \backslash E) = 0$. Donc $\int_\Omega f = \int_E f$ et $\int_\Omega g = \int_E g$ par le corollaire précédent, d'où le résultat puisque $f \equiv g$ sur E. □

COROLLAIRE 2.14. *Si $[a,b]$ est un segment de $\mathbb{R}$, alors on a pour toute fonction borélienne $f : [a,b] \to [0,\infty]$:*

$$\int_{[a,b]} f\, d\lambda_1 = \int_{]a,b]} f\, d\lambda_1 = \int_{[a,b[} f\, d\lambda_1 = \int_{]a,b[} f\, d\lambda_1 \,.$$

Démonstration. Cela découle de l'additivité de l'intégrale par rapport au domaine d'intégration et du fait que les singletons sont λ_1-négligeables : par exemple,

$$\int_{[a,b]} f\, d\lambda_1 = \int_{\{a\}} f\, d\lambda_1 + \int_{]a,b[} f\, d\lambda_1 + \int_{\{b\}} f\, d\lambda_1 = \int_{]a,b[} f\, d\lambda_1 \,.$$

□

LEMME 2.15. *Si $f \in \mathfrak{M}^+(\Omega)$ vérifie $\int_\Omega f\, d\mu < \infty$, alors $f(x) < \infty$ μ-presque partout.*

Démonstration. Soit $N := \{f = \infty\} = \{x;\ f(x) = \infty\}$. L'ensemble N est mesurable et on a

$$\infty > \int_\Omega f\, d\mu \geqslant \int_N f\, d\mu = \int_\Omega \infty \times \mathbf{1}_N\, d\mu = \infty \times \mu(N)\,;$$

donc la seule possibilité pour $\mu(N)$ est de valoir 0. □

Remarque. A nouveau, le "presque partout" est indispensable : il n'y a aucune raison qu'on ait $f(x) < \infty$ *partout.* Par ailleurs, le lemme dit "si... alors", et pas "si et seulement si" : on peut évidemment avoir $f(x) < \infty$ partout et cependant $\int_\Omega f\, d\mu = \infty$. Par exemple, $\int_\mathbb{R} \mathbf{1}\, d\lambda_1 = \infty$.

3. Intégrale des fonctions à valeurs réelles ou complexes

3.1. Fonctions intégrables et définition de l'intégrale.

3.1.1. *Fonction à valeurs réelles.* Pour définir l'intégrale des fonctions à valeurs réelles, l'idée est de se ramener aux cas des fonctions positives en écrivant une fonction $f : \Omega \to \mathbb{R}$ sous la forme

$$f = v - u\,,$$

où les fonctions u et v sont positives.

Il y a une manière "canonique" de faire cela : pour toute fonction $f : \Omega \to \mathbb{R}$, on écrit

$$f = f^+ - f^-\,,$$

où f^+ et f^- sont les fonctions définies par

$$f^+(x) = \begin{cases} f(x) & \text{si } f(x) \geqslant 0, \\ 0 & \text{si } f(x) < 0\,; \end{cases} \quad \text{et} \quad f^-(x) = \begin{cases} 0 & \text{si } f(x) \geqslant 0, \\ -f(x) & \text{si } f(x) < 0. \end{cases}$$

Autrement dit :

$$f^+ = \max(f, 0) \quad \text{et} \quad f^- = \max(-f, 0) = -\min(f, 0)\,.$$

Les propriétés à retenir sont les suivantes : f^+et f^- sont des fonctions *positives à valeurs finies*, et on a

$$f = f^+ - f^- \quad \text{et} \quad |f| = f^+ + f^-\,.$$

EXERCICE 3.1. Montrer qu'on a $f^+ \times f^- = 0$, et que l'écriture $f = f^+ - f^-$ est la *seule* écriture possible de f sous la forme $f = v - u$ avec $u, v \geqslant 0$ vérifiant $uv = 0$.

Si $f : \Omega \to \mathbb{R}$ est mesurable, alors f^+ , f^- et $|f|$ sont des fonctions *mesurables* positives, donc $\int_\Omega f^+\, d\mu$, $\int_\Omega f^-\, d\mu$ et $\int_\Omega |f|$ sont des nombres positifs bien définis, éventuellement égaux à ∞.

DÉFINITION 3.2. *Soit $f : (\Omega, \mathcal{B}) \to \mathbb{R}$ mesurable. On dit que f est* ***intégrable sur*** Ω ***par rapport à*** μ *si on a $\int_\Omega f^+\, d\mu < \infty$ et $\int_\Omega f^-\, d\mu < \infty$. Dans ce cas, on pose*

$$\int_\Omega f\, d\mu := \int_\Omega f^+\, d\mu - \int_\Omega f^-\, d\mu\,.$$

On peut aussi écrire $\int_\Omega f(x)\, d\mu(x)$, ou $\int_\Omega f$.

Remarque 1. Si f est intégrable, alors $\int_\Omega f\, d\mu$ est un *nombre réel* bien défini. Autrement dit : l'intégrale d'une fonction intégrable *ne peut pas* valoir ∞ ou $-\infty$

Remarque 2. Si f est une fonction *positive*, alors $f = f^+$ et $f^- = 0$, donc f est intégrable au sens précédent si et seulement si on a $\int_\Omega f\, d\mu < \infty$, et dans ce cas $\int_\Omega f\, d\mu$ au sens de la définition est bien égal à $\int_\Omega f\, d\mu$ au sens de la section précédente. Ainsi, on voit que *les fonctions positives ne sont pas toutes "intégrables"*, même si l'intégrale d'une fonction positive existe toujours dans $[0, \infty]$. La terminologie "intégrable" est donc assez malheureuse. Il faudrait utiliser un autre mot, mais il est difficile d'aller contre un usage si bien établi. (Cependant, on rencontre parfois le mot **sommable** comme synonyme de "intégrable".)

Exemple 1. La fonction constante $\mathbf{1}$ est intégrable par rapport à μ si et seulement si la mesure μ est finie, i.e. $\mu(\Omega) < \infty$.

C'est évident puisque $\int_\Omega \mathbf{1}\, d\mu = \mu(\Omega)$.

Exemple 2. Une fonction *étagée* $\varphi : \Omega \to \mathbb{R}$ (pas forcément positive !) est intégrable par rapport à μ si et seulement si $\mu\big(\{\varphi \neq 0\}\big) < \infty$.

Démonstration. Remarquons d'abord que comme les intégrales $\int_\Omega \varphi^+$ et $\int_\Omega \varphi^-$ sont des nombres positifs, elles sont toutes les deux finies si et seulement si $\int_\Omega \varphi^+ + \int_\Omega \varphi^- < \infty$.

Notons $\alpha_1, \dots, \alpha_N$ les valeurs strictement positives prises par la fonction φ, et $-\alpha_{N+1}, \dots, -\alpha_{N+M}$ les valeurs strictement négatives (on a donc $\alpha_i > 0$ pour $i = 0, \dots, N+M$). Posons $\Omega_i := \{\varphi = \alpha_i\}$ pour $1 \leqslant i \leqslant N$, et $\Omega_i := \{\varphi = -\alpha_i\}$ pour $N < i \leqslant N+M$. Alors $\{\varphi \neq 0\} = \bigcup_{i=1}^{N+M} \omega_i$. De plus, on a $\varphi^+ = \sum_{i=1}^N \alpha_i\, \mathbf{1}_{\Omega_i}$ et $\varphi^- = \sum_{i=N+1}^M \alpha_i \mathbf{1}_{\Omega_i}$ donc $\int_\Omega \varphi^+\, d\mu + \int_\Omega \varphi^-\, d\mu = \sum_{i=1}^N \alpha_i\, \mu(\Omega_i)$. Comme tous les α_i sont strictement positifs, cette somme est finie si et seulement si $\mu(\Omega_i) < \infty$ pour tout i ; et comme $\mu\big(\{\varphi \neq 0\}\big) = \mu\left(\bigcup_{i=1}^{N+M} \Omega_i\right) = \sum_{i=1}^{N+M} \mu(\Omega_i)$, il revient au même de dire qu'on a $\mu\big(\{\varphi \neq 0\}\big) < \infty$. $\square$

3.1.2. *Fonctions à valeurs complexes.* Pour définir l'intégrale d'une fonction f à valeurs complexes, on se ramène évidemment au cas des fonctions à valeurs réelles en considérant séparément la partie réelle et la partie imaginaire de f.

DÉFINITION 3.3. *On dit qu'une fonction mesurable $f : \Omega \to \mathbb{C}$ est intégrable sur Ω par rapport à μ si les fonctions* $\mathrm{Re}(f)$ *et* $\mathrm{Im}(f)$ *le sont. Dans ce cas, on pose*

$$\int_\Omega f\, d\mu = \int_\Omega \mathrm{Re(f)}\, d\mu + i \int_\Omega \mathrm{Im(f)}\, d\mu\,.$$

Comme il se doit, on peut aussi écrire $\int_\Omega f(x)\,d\mu(x)$ *ou* $\int_\Omega f$.

Autrement dit, si $f = u + iv$ avec u, v à valeurs réelles, alors

$$\int_\Omega f = \int_\Omega (u+iv) = \int_\Omega u + i\int_\Omega v\,;$$

et on a ainsi (ce qui n'a l'air de rien mais qu'il ne faut pas oublier)

$$\mathrm{Re}\left(\int_\Omega f\right) = \int_\Omega \mathrm{Re}(f) \quad \text{et} \quad \mathrm{Im}\left(\int_\Omega f\right) = \int_\Omega \mathrm{Im}(f)\,.$$

Exercice. Montrer u'une fonction *étagée* $\varphi : \Omega \to \mathbb{C}$ est intégrable sur Ω si et seulement si elle est de la forme $\varphi = \sum_{i=1}^N \alpha \mathbf{1}_{E_i}$ avec $\mu(E_i) < \infty$ pour $i = 1, \dots, N$.

3.1.3. *Caractérisation de l'intégrabilité.* Les définitions précédentes ne sont pas très "opérationnelles" : pour montrer qu'une fonction $f : \Omega \to \mathbb{C}$ est intégrable, il faudrait en théorie écrire $f = u + iv$ avec u, v à valeurs réelles, puis montrer séparement que les 4 fonctions u^+, u^-, v, v^- ont une intégrale finie. Heureusement, il y a un moyen beaucoup plus efficace de vérifier l'intégrabilité d'une fonction.

Proposition 3.4. *Une fonction mesurable* $f : \Omega \to \mathbb{C}$ *est intégrable par rapport à* μ *si et seulement si on a*

$$\int_\Omega |f|\,d\mu < \infty\,.$$

Démonstration. Considérons d'abord le cas d'une fonction f à valeurs réelles. Dans ce cas, f est intégrable si et seulement si $\int_\Omega f^+ + \int_\Omega f^- < \infty$ car $\int_\Omega f^+$ et $\int_\Omega f^+$ sont des nombres positifs. Mais comme $|f| = f^+ + f^-$, on a $\int_\Omega f^+ + \int_\Omega f^- = \int_\Omega |f|$ par additivité de l'intégrale des fonctions positives, d'où le résultat dans ce cas.

Si f est à valeurs complexes, on écrit $f = u + iv$ avec u et v à valeurs réelles. On a alors $|u| \leqslant |f|$ et $|v| \leqslant |f|$; donc $\int_\Omega |u| \leqslant \int_\Omega |f|$ et $\int_\Omega |v| \leqslant \int_\Omega |f|$ par croissance de l'intégrale des fonctions positives, et donc u et v sont intégrables (i.e. f est intégrable) si $\int_\Omega |f| < \infty$. Inversement, si f est intégrable alors $\int_\Omega |f| \leqslant \int_\Omega (|u|+|v|) = \int_\Omega |u| + \int_\Omega |v| < \infty$, où on a utilisé la croissance et l'additivité de l'intégrale des fonctions positives. □

Corollaire 3.5. *Si la mesure* μ *est* finie, *i.e.* $\mu(\Omega) < \infty$, *alors toute fonction mesurable* bornée $f : \Omega \to \mathbb{C}$ *est intégrable par rapport à* μ.

Démonstration. On a $\int_\Omega |f|\,d\mu \leqslant \int_\Omega \|f\|_\infty\,d\mu = \|f\|_\infty \times \mu(\Omega) < \infty$. □

Corollaire 3.6. *Si* $I \subseteq \mathbb{R}$ *est un* intervalle borné *de* $\mathbb{R}$, *alors toute fonction borélienne* bornée $f : I \to \mathbb{C}$ *est intégrable par rapport à (la restriction de) la mesure de Lebesgue* λ_1.

Démonstration. C'est évident puisque $\lambda_1(I) = |I| < \infty$. □

Exercice. Soit $f : \Omega \to \mathbb{R}$ mesurable. Montrer qu'on a $\int_\Omega |f| = \int_\Omega f^+ + \int_\Omega f^-$ *sans utiliser* l'additivité de l'intégrale des fonctions positives.

3.2. Intégrabilité sur un sous-ensemble. Comme on s'y attend, on dit qu'une fonction mesurable $f : \Omega \to \mathbb{C}$ est intégrable *sur un ensemble* $E \subseteq \Omega$ par rapport à μ si la fonction $f_{|E}$ est inégrable par rapport à μ_E (la restrictions de mesure μ à la tribu trace $\mathcal{B}_E$). Par la Proposition 3.4 appliquée à $f_{|E}$, il revient au même de dire qu'on a

$$\int_E |f|\,d\mu < \infty\,.$$

En particulier, comme $\int_E |f| \leqslant \int_\Omega |f|$, on en déduit immédiatement le fait suivant :

FAIT. Si $f : \Omega \to \mathbb{C}$ est intégrable sur Ω, alors elle est intégrable sur *tout* ensemble mesurable $E \subseteq \Omega$.

Comme $\int_E |f|\, d\mu = \int_\Omega |\mathbf{1}_E\, f|\, d\mu$, on en déduit aussi que f est intégrable sur E si et seulement si la fonction $\mathbf{1}_E\, f$ est intégrable sur Ω ; et on vérifie très facilement qu'on a alors

$$\int_E f\, d\mu = \int_\Omega \mathbf{1}_E\, f\, d\mu\,.$$

3.3. Propriétés de l'intégrale. Dans toute la suite, on notera $\mathcal{L}^1 = \mathcal{L}^1(\Omega, \mathcal{B}, \mu)$ l'ensemble de toutes les fonctions mesurables $f : \Omega \to \mathbb{C}$ intégrables par rapport à μ. Par la Proposition 3.4, on a donc

$$\mathcal{L}^1 = \mathcal{L}^1(\Omega, \mathcal{B}, \mu) := \left\{ f : (\Omega, \mathcal{B}) \to \mathbb{C} \text{ mesurables; } \int_\Omega |f|\, d\mu < \infty \right\}.$$

Il faut faire tout de suite la remarque suivante :

FAIT 3.7. *$\mathcal{L}^1$ est un* espace vectoriel.

Démonstration. Si $f, g \in \mathcal{L}^1$ et $\lambda \in \mathbb{C}$, alors les fonctions $f+g$ et λf sont mesurables. Par croissance et additivité de l'intégrale des fonctions positives, on a

$$\int_\Omega |f+g| \leqslant \int_\Omega (|f| + |g|) = \int_\Omega |f| + \int_\Omega |g| < \infty\,;$$

donc $f + g \in \mathcal{L}^1$. De même, on a par homogénéité : $\int_\Omega |\lambda f| = |\lambda| \times \int_\Omega |f| < \infty$, donc $\lambda f \in \mathcal{L}^1$.

□

THÉORÈME 3.8. *L'intégrale possède les propriétés suivantes.*

(1) ***Linéarité** : si $f, g \in \mathcal{L}^1$ et $\lambda \in \mathbb{C}$, alors*

$$\int_\Omega (f+g) = \int_\Omega f + \int_\Omega g \quad \text{et} \quad \int_\Omega (\lambda f) = \lambda \int_\Omega f\,.$$

(2) ***Croissance** : si $f, g \in \mathcal{L}^1$ sont* à valeurs réelles *et si $f \leqslant g$, alors $\int_\Omega f \leqslant \int_\Omega g$.*

(3) ***Majoration du module** : si $f \in \mathcal{L}^1$, alors*

$$\left| \int_\Omega f\, d\mu \right| \leqslant \int_\Omega |f|\,.$$

Démonstration. (1) (i) Montrons que $\int_\Omega (f+g) = \int_\Omega f + \int_\Omega g$.

Si f et g sont réelles, on écrit $f = f^+ - f^-$ et $g = g^+ - g^-$. On a alors deux façons d'écrire $f+g$, à savoir $(f^+ + g^+) - (f^- + g^-) = f + g = (f+g)^+ - (f+g)^-$. Donc

$$(f+g)^+ + f^- + g^- = (f+g)^- + f^+ + g^+\,.$$

Par additivité de l'intégrale *pour les fonctions positives*, on en déduit

$$\int_\Omega (f+g)^+ + \int_\Omega f^- + \int_\Omega g^- = \int_\Omega (f+g)^- + \int_\Omega f^+ + \int_\Omega g^+;$$

et on obtient ainsi

$$\begin{aligned}\int_\Omega (f+g) = \int_\Omega (f+g)^+ - \int_\Omega (f+g)^- &= \left(\int_\Omega f^+ + \int_\Omega +g^+\right) - \left(\int_\Omega f^- + \int_\Omega g^-\right)\\ &= \left(\int_\Omega f^+ - \int_\Omega f^-\right) + \left(\int_\Omega g^+ - \int_\Omega g^-\right)\\ &= \int_\Omega f + \int_\Omega g\,.\end{aligned}$$

Si f et g sont à valeurs complexes, on applique le "cas réel" aux parties réelles et imaginaires de f et g, ce qui donne immédiatement le résultat (exercice).

(ii) Montrons maintenant qu'on a $\int_\Omega (\lambda f) = \lambda \int_\Omega f$ pour tout $\lambda \in \mathbb{C}$.

On sait déjà que la formule est valable si $f \geqslant 0$ et si $\lambda \in \mathbb{R}^+$.

Si f est réelle de signe quelconque et $\lambda \in \mathbb{R}^+$, alors $(\lambda f)^+ = \lambda f^+$ et $(\lambda f)^- = \lambda f^-$; donc $\int_\Omega (\lambda f) = \int_\Omega \lambda f^+ - \int_\Omega \lambda f^- = \lambda\left(\int_\Omega f^+ - \int_\Omega f^-\right) = \lambda \int_\Omega f$, où a utilisé le "cas positif".

Si f est réelle et $\lambda \in \mathbb{R}^-$, alors $\int_\Omega \lambda f + \int_\Omega (-\lambda f) = \int_\Omega ((\lambda f) + (-\lambda f)) = 0$ par additivité ; donc $\int_\Omega (\lambda f) = -\int_\Omega (-\lambda f)$, et donc $\int_\Omega (\lambda f) = \lambda \int_\Omega f$ puisque $\int_\Omega (-\lambda f) = (-\lambda)\int_\Omega f$ par le cas précédent.

En écrivant $f = u + iv$, on en déduit immédiatement que la formule est encore valable pour f à valeurs complexes et $\lambda \in \mathbb{R}$.

Pour conclure, il suffit maintenant de voir que si f est à valeurs complexes, alors $\int_\Omega (if) = i \int_\Omega f$; ce qui n'est pas difficile : si $f = u + iv$, alors $if = -v + iu$, donc $\int_\Omega (if) = \int_\Omega (-v) + i \int_\Omega u = -\int_\Omega v + i\int_\Omega u = i\left(\int_\Omega u + i \int_\Omega v\right) = i \int_\Omega f$.

(2) Par linéarité, $\int_\Omega g - \int_\Omega f = \int_\Omega (g-f) \geqslant 0$ puisque $g \geqslant f$.

(3) (i) Si f est à valeurs réelles, la preuve est immédiate : on a $-|f| \leqslant f \leqslant |f|$, donc

$$-\int_\Omega |f| \leqslant \int_\Omega f \leqslant \int_\Omega |f|$$

par croissance et linéarité ; autrement dit $\left|\int_\Omega f\right| \leqslant \int_\Omega |f|$.

(ii) Pour une fonction f à valeurs complexes, ce n'*est pas* complètement immédiat. L'idée est de choisir un nombre complexe ω tel que $|\omega| = 1$ et $\left|\int_\Omega f\right| = \omega \int_\Omega f$. Comme $\left|\int_\Omega f\right|$ est un nombre réel, on a alors (en utilisant la linéarité de l'intégrale)

$$\left|\int_\Omega f\right| = \mathrm{Re}\left(\omega \int_\Omega f\right) = \mathrm{Re}\left(\int_\Omega \omega f\right) = \int_\Omega \mathrm{Re}(\omega f)\,;$$

d'où $\left|\int_\Omega f\right| \leqslant \int_\Omega |f|$ par croissance de l'intégrale des fonctions à valeurs réelles, car $\mathrm{Re}(\omega f) \leqslant |\omega f| = |f|$. □

Corollaire 3.9. *Si $\varphi : \Omega \to \mathbb{C}$ est une fonction* étagée *intégrable, alors, quelle que soit la façon d'écrire $\varphi = \sum_{i=1}^n \alpha_i \mathbf{1}_{E_i}$ avec $E_i \in \mathcal{B}$ vérifiant $\mu(E_i) < \infty$ et $\alpha_i \in \mathbb{C}$, on a*

$$\int_\Omega \varphi = \sum_{i=1}^n \alpha_i\, \mu(E_i)\,.$$

Démonstration. Comme $\mu(E_i) < \infty$ pour tout i, les fonctions $\mathbf{1}_{E_i}$ sont dans $\mathcal{L}^1$. Donc le résultat est clair par linéarité de l'intégrale. □

Corollaire 3.10. *Si $f \in \mathcal{L}^1$ et si $A, B \subseteq \Omega$ sont deux ensembles mesurables tels que $A \cap B = \varnothing$, alors $\int_{A\cup B} f = \int_A f + \int_B f$.*

Démonstration. Comme $A \cap B = \varnothing$, on a $\mathbf{1}_{A\cup B} = \mathbf{1}_A + \mathbf{1}_B$; donc

$$\int_{A\cup B} f = \int_\Omega \mathbf{1}_{A\cup B}\, f = \int_\Omega (\mathbf{1}_A\, f + \mathbf{1}_B\, f) = \int_\Omega \mathbf{1}_A f + \int_\Omega \mathbf{1}_B f = \int_A f + \int_B f\,,$$

où on a utilisé la linéarité de l'intégrale. □

Exercice. Soit $f \in \mathcal{L}^1$. Montrer qu'on a $\left|\int_\Omega f\, d\mu\right| = \int_\Omega |f|$ si et seulement si il existe une constante $\omega \in \mathbb{C}$ de module 1 telle que $f(x) = \omega\, |f(x)|$ μ-presque partout. Dans le cas où f est à valeurs réelles, montrer qu'on a $\left|\int_\Omega f\, d\mu\right| = \int_\Omega |f|$ si et seulement si $f(x) \geqslant 0$ μ-presque partout ou $f(x) \leqslant 0$ μ-presque partout.

3.4. Extension des définitions. Dans certaines situations, il peut arriver qu'on ait affaire à une fonction f dont on a bien envie d'écrire l'intégrale, mais qui n'est peut-être pas mesurable. On va voir ici qu'il est facile de contourner cette difficulté.

LEMME 3.11. *Soient $f_1, f_2 : \Omega \to \mathbb{C}$ ou $[0,\infty]$ deux fonctions mesurables. On suppose qu'on a $f_1(x) = f_2(x)$ μ-presque partout.*

(1) *Si f_1 et f_2 sont positives, alors $\int_\Omega f_1\, d\mu = \int_\Omega f_2\, d\mu$.*
(2) *f_1 est intégrable par rapport à μ si et seulement si f_2 l'est ; et dans ce cas, on a $\int_\Omega f_1\, d\mu = \int_\Omega f_2\, d\mu$.*

Démonstration. La partie (1) a déjà été vue. Comme $|f_1| = |f_2|$ presque partout, on a $\int_\Omega |f_1|\, d\mu = \int_\Omega |f_2|\, d\mu$, donc f_1 est intégrable par rapport à μ si et seulement si f_2 l'est. Dans ce cas, on a $\left|\int_\Omega f_1\, d\mu - \int_\Omega f_2\, d\mu\right| = \left|\int_\Omega (f_1 - f_2)\, d\mu\right| \leqslant \int_\Omega |f_1 - f_2|\, d\mu = 0$, la dernière égalité venant du fait que la fonction mesurable positive $|f_1 - f_2|$ est nulle presque partout. □

Ce lemme justifie la définition suivante. Convenons d'appeler **fonction μ-presque partout définie** toute fonction f à valeurs complexes définie sur un ensemble $E \subseteq \Omega$ tel que $\Omega\backslash E$ est négligeable. On dira qu'une fonction μ-presque partout définie est **μ-mesurable** si elle est μ-presque partout égale à une fonction mesurable.

DÉFINITION 3.12. *Soit $f : \Omega \to \mathbb{C}$ ou $[0,\infty]$ une fonction μ-presque partout définie.*

(1) *Si f est μ-mesurable et positive, on pose $\int_\Omega f\, d\mu = \int_\Omega \widetilde{f}\, d\mu$, où $\widetilde{f}$ est n'importe quelle fonction mseurable positive égale à f presque partout.*
(2) *On dit que f est **μ-intégrable au sens de Lebesgue** si f est μ-presque partout égale à une fonction mesurable $\widetilde{f} : \Omega \to \mathbb{C}$ intégrable par rapport à μ. On pose alors $\int_\Omega f\, d\mu = \int_\Omega \widetilde{f}\, d\mu$, où $\widetilde{f}$ est n'importe quelle fonction intégrable par rapport à μ égale à f presque partout.*

On notera $\mathcal{L}^1(\Omega,\mu)$ l'ensemble de toutes les fonctions μ-intégrables $f : \Omega \to \mathbb{C}$:

$$\mathcal{L}^1(\Omega,\mu) := \{f : \Omega \to \mathbb{C};\ \mu\text{-intégrable}\}\,.$$

Remarque. La référence à la tribu $\mathcal{B}$ a disparu, car $\mathcal{L}^1(\Omega,\mu)$ de dépend que de la mesure $\mu = \lambda_1$.

La preuve de la proposition suivante est laissée en exercice. Notons d'abord que les fonctions μ-presque partout définies forment de manière évidente un espace vectoriel : si f et g sont chacune μ-presque partout définie, alors la propriété "f et g sont toutes les deux définies au point x" est vraie μ-presque partout, donc $f + g$ est μ-presque partout définie.

Proposition 3.13. *$\mathcal{L}^1(\Omega,\mu)$ est un espace vectoriel, et les propriétés de l'intégrale restent les mêmes.*

L'exercice suivant est important pour vérifier qu'on a bien compris ce dont il est question.

Exercice 3.14. Soit (f_n) une suite de fonctions μ-mesurables, $f_n : \Omega \to \mathbb{C}$. On suppose que (f_n) converge μ-presque partout vers une fonction $f : \Omega \to \mathbb{C}$ presque partout définie. Montrer que f est μ-mesurable.

4. Dirac, comptage et Riemann

4.1. Intégration par rapport à une masse de Dirac. Soit Ω un ensemble non vide quelconque, et soit $a \in \Omega$ fixé. Rappelons que la masse de Dirac δ_a au point a est la mesure définie sur $\mathcal{B} = \mathcal{P}(\Omega)$ par

$$\delta_a(A) = \begin{cases} 1 & \text{si } A \ni a \\ 0 & \text{si } A \not\ni a \end{cases} \qquad \text{pour tout ensemble } A \subseteq \Omega\,.$$

Proposition 4.1. *Pour toute fonction $f : \Omega \to [0,\infty]$, on a*

$$\int_\Omega f\, d\delta_a = f(a)\,.$$

Démonstration. Soit $f : \Omega \to [0,\infty]$ quelconque. Comme "$x = a$" δ_a-presque partout, on a $f(x) = f(a)$ δ_a-presque partout, i.e. $f = f(a)\,\mathbf{1}$ δ_a-presque partout, donc $\int_\Omega f\, d\delta_a = \int_\Omega f(a)\,\mathbf{1}\, d\delta_a = f(a) \times \delta_a(\Omega) = f(a)$. ☐

Corollaire 4.2. Toute *fonction $f : \Omega \to \mathbb{C}$ est intégrable par rapport à δ_a, et on a*

$$\int_\Omega f\, d\delta_a = f(a)\,.$$

Démonstration. Si $f : \Omega \to \mathbb{C}$ est quelconque, on a (par la proposition appliqué à la fonction positive $|f|$) $\int_\Omega |f|\, d\delta_a = |f(a)| < \infty$; donc toute fonction $f : \Omega \to \mathbb{C}$ est intégrable par rapport à δ_a.

Notons $\mathcal{F}(\Omega)$ l'espace vectoriel de toutes les fonctions $f : \Omega \to \mathbb{C}$. Les applications $f \mapsto \int_\Omega f\, d\delta_a$ et $f \mapsto f(a)$ sont des *formes linéaires* sur $\mathcal{F}(\Omega)$; et par la proposition, ces formes linéaires coïncident sur les fonctions positives. Comme les fonctions positives engendrent l'espace vectoriel $\mathcal{F}(\Omega)$, on a donc bien $\int_\Omega f\, d\delta_a = f(a)$ pour toute fonction $f : \Omega \to \mathbb{C}$. ☐

4.2. Intégration par rapport à la mesure de comptage. Soit Ω un ensemble non vide quelconque, et soit μ_c la mesure de comptage sur Ω,

$$\mu_c(A) = \begin{cases} \infty & \text{si } A \text{ est infini} \\ \#A & \text{si } A \text{ est fini} \end{cases}$$

On a observé au chapitre 2 que

$$\mu(A) = \sum_{x\in\Omega} \delta_x(A) \quad \text{pour tout ensemble } A \subseteq \Omega\,.$$

Proposition 4.3. *Pour toute fonction $f : \Omega \to [0,\infty]$, on a*

$$\int_\Omega f\, d\mu_c = \sum_{x\in\Omega} f(x)\,.$$

Démonstration. Montrons d'abord que la formule est valable pour toute fonction *étagée* positive φ,

$$\varphi = \sum_{i=1}^{N} \alpha_i \, \mathbf{1}_{E_i} \quad \text{avec } \alpha_i \in \mathbb{R}^+ \,.$$

C'est un calcul direct :

$$\begin{aligned}
\int_\Omega \varphi \, d\mu_c &= \sum_{i=1}^{N} \alpha_i \, \mu_c(E_i) \\
&= \sum_{i=1}^{N} \alpha_i \left(\sum_{x \in \Omega} \delta_x(E_i) \right) \\
&= \sum_{x \in \Omega} \left(\sum_{i=1}^{N} \alpha_i \, \delta_x(E_i) \right) \\
&= \sum_{x \in \Omega} \int_\Omega \varphi \, d\delta_x \\
&= \sum_{x \in \Omega} \varphi(x) \quad \text{par la Proposition 4.1} \ .
\end{aligned}$$

Soit maintenant $f : \Omega \to [0, \infty]$ quelconque, et soit $(\varphi_n)_{n \in \mathbb{N}}$ une suite approximante pour f (i.e. une suite croissante de fonctions étagées positives tendant vers f). On a

$$\begin{aligned}
\int_\Omega f \, d\mu_c &= \lim_{n \to \infty} \int_\Omega \varphi_n \, d\mu_c \quad \text{par convergence monotone} \\
&= \lim_{n \to \infty} \sum_{x \in \Omega} \varphi_n(x) \quad \text{par le ``cas étagé''} \\
&= \sum_{x \in \Omega} f(x) \quad \text{par convergence monotone } \textit{pour les sommes} \,.
\end{aligned}$$

□

Corollaire 4.4. *Une fonction $f : \mathbb{N} \to \mathbb{C}$ est intégrable sur $\mathbb{N}$ par rapport à la mesure de comptage μ_c si et seulement si la série $\sum f(n)$ est* absolument convergente, *et on a alors*

$$\int_{\mathbb{N}} f \, d\mu_c = \sum_{n=0}^{\infty} f(n) \,.$$

Démonstration. Par la Proposition appliquée à la fonction positive $|f|$, on a

$$\int_{\mathbb{N}} |f| \, d\mu_c = \sum_{n \in \mathbb{N}} |f(n)| = \sum_{n=0}^{\infty} |f(n)| \,;$$

donc f est intégrable par rapport à μ_c si et seulement si $\sum f(n)$ est absolument convergente. La formule souhaitée se déduit alors du "cas positif" et de la linéarité des applications $f \mapsto \sum_0^\infty f(n)$ et $f \mapsto \int_{\mathbb{N}} f \, d\mu_c$. □

Remarque 4.5. On obtient de la même façon des résultats analogue pour n'importe quelle mesure discrète

$$\mu = \sum_{i \in \mathbb{N}} p_i \, \delta_{x_i} \,,$$

où $(x_i)_{i\in\mathbb{N}}$ est une suite de points de Ω et les p_i sont des réels positifs. En particulier, dans le cas où $\Omega = \mathbb{R}$, la fonction $x \mapsto x$ est intégrable par rapport à μ si et seulement si la série $\sum p_i\, x_i$ est absolument convergente, et on a alors

$$\int_{\mathbb{R}} x\, d\mu(x) = \sum_{i=0}^{\infty} p_i\, x_i\,,$$

On peut interpréter ce résultat de la façons suivante. Soit X une "variable aléatoire discrète" à valeurs réelles, i.e. une variable aléatoire réelle dont l'ensemble des valeurs possibles est dénombrable. Notons $x_0, x_1, \dots$ les valeurs possibles de X, et posons $p_i := \mathbb{P}(X = x_i) = \mathbb{P}_X(\{x_i\})$. On dit que la variable aléatoire X *admet une espérance* si la série (éventuellement finie) $\sum x_i\, \mathbb{P}(X = x_i) = \sum p_i\, x_i$ est absolument convergente ; et dans ce cas, l'**espérance** de la variable aléatoire X est le nombre réel $\mathbb{E}(X)$ défini par

$$\mathbb{E}(X) := \sum_{i\geqslant 0} x_i\, \mathbb{P}(X = x_i) = \sum_{i\geqslant 0} p_i\, x_i\,.$$

Ainsi, on voit que X admet une espérance si et seulement si la fonction $x \mapsto x$ est intégrable par rapport à la mesure $\mathbb{P}_X$ (la loi de X), et qu'on a alors

$$\mathbb{E}(X) = \int_{\mathbb{R}} x\, d\mathbb{P}_X(x)\,.$$

4.3. Intégrale au sens de Riemann. Pour les fonctions définies sur un segment $[a, b]$ de $\mathbb{R}$, on connaissait déjà une théorie de l'intégration, à savoir l'intégrale *au sens de Riemann.* Il est évidemment indispensable de préciser le lien entre les deux théories qu'on a maintenant sous la main (intégrale au sens de Riemann et intégrale par rapport à la mesure de Lebesgue λ_1) : par exemple, il serait préoccupant que l'intégrale au sens de Lebesgue d'une fonction continue sur un segment $[a, b]$ ne soit pas égale à son intégrale au sens de Riemann ; et il serait également un peu béta d'avoir élaboré une théorie quand même plus compliquée que celle de Riemann pour s'apercevoir qu'elle s'applique en fait à moins de fonctions.

Rappelons rapidement la définition de l'intégrale au sens de Riemann. Dans tout ce qui suit, $[a, b]$ est un intervalle fermé borné de $\mathbb{R}$.

• Une fonction $\varphi : [a, b] \to \mathbb{R}$ est dite **en escalier** s'il existe une subdivision $a = s_0 < \cdots < s_N = b$ de l'intervalle $[a, b]$ telle que φ soit constante sur chaque intervalle ouvert $J_i :=]s_i, s_{i+1}[$, $0 \leqslant i \leqslant N - 1$:

$$\varphi(x) \equiv \alpha_i \quad \text{sur } J_i =]s_i, s_{i+1}[\,.$$

On pose alors

$$\int_a^b \varphi(x)\, dx = \sum_{i=0}^{N-1} \alpha_i\, |J_i| = \sum_{i=0}^{N-1} \alpha_i\, (s_{i+1} - s_i)\,.$$

Il faut vérifier que cette somme est bien indépendante de la subdivision "adaptée" à φ ; ce qui n'est pas très difficile mais demande quand même un peu de soin (*cf* le preuve du Lemme 1.2).

La signification géométrique de la formule est claire : en faisant un dessin, on voit que $\int_a^b \varphi(x)\, dx$ est l'**aire algébrique** déterminée par le graphe de φ entre a et b (l'aire au dessus de l'axe des abscisses est comptée positivement, et celle en dessous de l'axe des abscisses est comptée négativement).

• Pour toute fonction *bornée* $f : [a, b] \to \mathbb{R}$, on pose

$$\underline{\int_a^b} f := \sup \left\{ \int_a^b \varphi(x)\, dx;\ \varphi \text{ en escalier, } \varphi \leqslant f \right\},$$

et

$$\overline{\int_a^b} f := \inf \left\{ \int_a^b \varphi(x)\, dx;\ \psi \text{ en escalier, } \psi \geqslant f \right\}.$$

Ce sont des nombres réels bien définis précisément parce que la fonction f est bornée.

On dit que la fonction f est **intégrable au sens de Riemann** sur $[a, b]$ si on a $\underline{\int_a^b} f = \overline{\int_a^b} f$, et on note alors $\int_a^b f(x)\, dx$ cette valeur commune :

$$\underline{\int_a^b} f = \int_a^b f(x)\, dx = \overline{\int_a^b} f\,.$$

Exemple 1. Toute fonction *continue* $f : [a, b] \to \mathbb{R}$ est intégrable au sens de Riemann sur $[a, b]$; plus généralement, toute fonction *continue par morceaux* est intégrable au sens de Riemann.

Exemple 2. Toute fonction *monotone* sur $[a, b]$ est intégrable au sens de Riemann.

Exemple 3. La fonction $\mathbf{1}_{\mathbb{Q}}$ n'est intégrable au sens de Riemann sur aucun intervalle $[a, b]$ non trivial. (En revanche, elle est intégrable par rapport à la mesure de Lebesgue λ_1 sur $[a, b]$, avec $\int_{[a,b]} \mathbf{1}_{\mathbb{Q}}\, d\lambda_1 = 0$.)

Démonstration. C'est l'exemple classique qu'on donne en 1ère année. Le point clé est que tout intervalle ouvert $J \neq \varnothing$ contient à la fois des rationnels et des irrationnels ; autrement dit, $\mathbf{1}_{\mathbb{Q}}$ prend la valeur 1 et la valeur 0 sur tout intervalle ouvert non trivial. On en déduit que si ψ est une fonction en escalier sur $[a, b]$ telle que $\psi \geqslant \mathbf{1}_{\mathbb{Q}}$, alors $\psi \geqslant 1$ sur chacun des intervalles ouverts $J_0, \dots, J_{N-1}$ définis par une subdivision adaptée $a = s_0 < \cdots < s_N = b$, et donc $\int_a^b \psi(x)\, dx \geqslant \sum_{i=0}^{N-1} 1 \times |J_i| = 1 \times (b - a) = b - a$. Donc $\overline{\int_a^b} \mathbf{1}_{\mathbb{Q}} \geqslant (b - a) > 0$. On montre de même qu'on a $\underline{\int_a^b} \mathbf{1}_{\mathbb{Q}} \leqslant 0 \times (b - a) = 0$; donc $\overline{\int_a^b} \mathbf{1}_{\mathbb{Q}} \neq \underline{\int_a^b} \mathbf{1}_{\mathbb{Q}}$, et $\mathbf{1}_{\mathbb{Q}}$ n'est pas intégrable au sens de Riemann sur $[a, b]$. □

Le lien entre l'intégrale au sens de Riemann et l'intégrale au sens de Lebesgue est donné par la proposition suivante.

PROPOSITION 4.6. *Si $f : [a, b] \to \mathbb{R}$ est intégrable au sens de Riemann, alors elle est λ_1-intégrable au sens de Lebesgue sur $[a, b]$ et on a*

$$\int_{[a,b]} f\, d\lambda_1 = \int_a^b f(x)\, dx\,.$$

Démonstration. Démontrons d'abord que tout se passe pour le mieux pour les fonctions en escalier.

FAIT. Si $\varphi : [a, b] \to \mathbb{R}$ est en escalier, alors φ est intégrable par rapport à λ_1 et $\int_{[a,b]} \varphi\, d\lambda_1 = \int_a^b \varphi(x)\, dx$.

Démonstration. La fonction φ est borélienne et bornée, donc intégrable par rapport à λ_1 sur $[a, b]$ car l'intervalle $[a, b]$ est borné (Corollaire 3.6).

Soit $a = s_0 < \cdots < s_N = b$ une subdivision de $[a,b]$ adaptée à ϕ,

$$\varphi(t) \equiv \alpha_i \quad \text{sur } J_i =]s_i, s_{i+1}[\quad , \quad 0 \leqslant i \leqslant N-1\,.$$

On a alors

$$\varphi = \sum_{i=0}^{N-1} \alpha_i \, \mathbf{1}_{J_i} + \sum_{i=0}^{N} \varphi(\alpha_i)\, \mathbf{1}_{\{s_i\}}\,;$$

et donc

$$\begin{aligned}
\int_{[a,b]} \varphi \, d\lambda_1 &= \sum_{i=0}^{N-1} \alpha_i \, \lambda_1(J_i) + \sum_{i=0}^{N} \varphi(\alpha_i)\, \lambda_1(\{s_i\}) \\
&= \sum_{i=0}^{N-1} \alpha_i \, |J_i| \qquad \text{car } \lambda_1(\{s_i\}) = 0 \text{ pour tout } i \\
&= \int_a^b \varphi(x)\, dx\,.
\end{aligned}$$

□

Soit maintenant $f : [a,b] \to \mathbb{R}$ intégrable au sens de Riemann. Par définition de $\underline{\int_a^b} f$ et $\overline{\int_a^b} f$, on peut trouver deux suites (φ_n) et (ψ_n) de fonctions en escalier avec $\varphi_n \leqslant f \leqslant \psi_n$ pour tout $n \in \mathbb{N}$, telles que $\int_a^b \varphi_n(x)\, dx \to \underline{\int_a^b} f$ et $\int_a^b \psi_n(x)\, dx \to \overline{\int_a^b} f$. Comme f est supposée intégrable au sens de Riemann, on a donc

$$\lim_{n\to\infty} \int_a^b \varphi_n(x)\, dx = \int_a^b f(x)\, dx = \lim_{n\to\infty} \int_a^b \psi_n(x)\, dx\,;$$

et donc, par le Fait :

$$\lim_{n\to\infty} \int_{[a,b]} \varphi_n \, d\lambda_1 = \int_a^b f(x)\, dx = \lim_{n\to\infty} \int_{[a,b]} \psi_n \, d\lambda_1\,. \tag{4.1}$$

Posons alors $\varphi := \sup_{n\geqslant 0} \varphi_n$ et $\psi := \inf_{n\geqslant 0} \psi_n$. Les fonctions φ et ψ sont *boréliennes* d'après la Proposition 2.5 du Chapitre 3. De plus, on a $\varphi_0 \leqslant \varphi \leqslant f \leqslant \psi \leqslant \psi_0$, donc φ et ψ sont (à valeurs finies et) bornés sur $[a,b]$. En particulier φ et ψ sont intégrables sur $[a,b]$ par rapport à λ_1 (Corollaire 3.6) ; et il faut garder en tête on a

$$\varphi \leqslant f \leqslant \psi\,.$$

Par croissance de l'intégrale de Lebesgue et comme $\varphi_n \leqslant \varphi \leqslant \psi \leqslant \psi_n$, on a

$$\int_{[a,b]} \varphi_n \, d\lambda_1 \leqslant \int_{[a,b]} \varphi \, d\lambda_1 \leqslant \int_{[a,b]} \psi \, d\lambda_1 \leqslant \int_{[a,b]} \psi_n \, d\lambda_1$$

pour tout $n \in \mathbb{N}$. Comme $\int_{[a,b]} \varphi_n \, d\lambda_1$ et $\int_{[a,b]} \psi_n \, d\lambda_1$ tendent toutes les deux vers $\int_a^b f(x)\, dx$ par (4.1), on en déduit

$$\int_{[a,b]} \varphi \, d\lambda_1 = \int_a^b f(x)\, dx = \int_{[a,b]} \psi \, d\lambda_1\,.$$

Par linéarité de l'intégrale, on a donc $\int_{[a,b]} (\psi - \varphi)\, d\lambda_1 = 0$; et comme $\psi - \varphi \geqslant 0$, cela entraine que $\psi - \varphi = 0$ λ_1-presque partout. Donc $f = \varphi = \psi$ λ_1-presque partout puisque $\varphi \leqslant f \leqslant \psi$. Ainsi, la fonction f est λ_1-intégrable sur $[a, b]$, avec

$$\int_a^b f\, d\lambda_1 = \int_{[a,b]} \varphi\, d\lambda_1 = \int_a^b f(x)\, dx\,.$$

□

Remarque 1. La proposition précédente, combinée au contre-exemple $\mathbf{1}_{\mathbb{Q}}$, montre que *l'intégrale au sens de Lebesgue est une extension* stricte *de l'intégrale au sens de Riemann* : toute intégrale de Riemann est une intégrale de Lebesgue, mais il y a plus de fonctions intégrables au sens de Lebesgue que de fonctions intégrables au sens de Riemann. Il y en a même "énormément plus", puisque *n'importe quelle* fonction λ_1-mesurable bornée sur un intervalle $[a, b]$ est intégrable au sens de Lebesgue sur $[a, b]$.

D'un point de vue très pragmatique, cela signifie ceci : d'une part, si on a affaire à une fonction sympathique $f : [a, b] \to \mathbb{R}$ (par exemple, continue par morceaux ou monotone), alors on peut jouer avec son intégrale comme on l'a toujours fait puisque cette intégrale est la même que celle définie par la théorie de Riemann ; et d'autre part, la collection de jouets s'est considérablement enrichie puisque des fonctions fort peu sympathiques comme $\mathbf{1}_{\mathbb{Q}}$ en font maintenant partie.

Remarque 2. La proposition *ne dit pas* que toute fonction intégrable au sens de Riemann est borélienne, ce qui n'est pas vrai. C'est une des raisons pour lesquelles il est très commode d'avoir étendu la définition de l'intégrabilité à des fonctions non-nécessairement "mesurables" (i.e. boréliennes dans le cas présent).

Remarque 3. Il est naturel de se demander s'il existe une caractérisation intéressante des fonctions intégrables au sens de Riemann. C'est effectivement le cas : une fonction bornée $f : [a, b] \to \mathbb{R}$ est intégrable au sens de Riemann si et seulement si *l'ensemble de ses points de discontinuité est négligeable.* On démontrera ce très beau résultat (dû à ... Lebesgue) un peu plus tard. Pour le moment, on observera juste que cette caractérisation explique de façon particulièrement limpide "pourquoi" la fonction $\mathbf{1}_{\mathbb{Q}}$ n'est intégrable au sens de Riemann sur aucun intervalle $[a, b]$ non trivial : cette fonction est en effet *discontinue en tout point.* A l'inverse, on déduit immédiatement de ce résultat qu'une fonction bornée $f : [a, b] \to \mathbb{R}$ dont l'ensemble des points de discontinuité est *dénombrable* est intégrable au sens de Riemann ; ce qui englobe évidemment le cas des fonctions continues par morceaux, mais aussi celui des fonctions monotones (exercice intéressant).

Exercice. Notons μ la mesure de Lebesgue λ_1. Soit Ω un borélien de $\mathbb{R}$ tel que $\mu(\Omega) < \infty$, et soit $f : \Omega \to \mathbb{R}$ une fonction borélienne *bornée.* On pose $m := \inf_\Omega f$ et $M := \sup_\Omega f$. À toute subdivision pointée $\Sigma = ((J_0, y_0), \ldots, (J_{N_1}, y_{N-1}))$ de l'intervalle $[m, M]$ (*i.e.* $(J_0, \ldots, J_{N-1})$ est un découpage de $[m, M]$ en intervalles fermés et $y_k \in J_k$), on associe la **somme de Lebesgue**

$$L(f, \Sigma) := \sum_{k=0}^{N-1} \mu(\{f \in J_k\})\, y_k\,.$$

Montrer que $L(f, \Sigma)$ tend vers $\int_\Omega f\, d\mu$ quand le "pas" de la subdivision Σ tend vers 0.

Chapitre 5

Intégration sur un intervalle de $\mathbb{R}$

1. Cadre et notations

Dans ce chapitre, la lettre I désignera toujours un intervalle de $\mathbb{R}$, de nature quelconque (ouvert, fermé, semi-ouvert), borné ou non. Le mot "intégrable" signifiera toujours "λ_1-intégrable" au sens de la Définition 3.12 du Chapitre 4, où λ_1 est la mesure de Lebesgue. (En réalité, on ne considérera que des fonctions *boréliennes*. Le point important est qu'on intégre *par rapport à la mesure de Lebesgue*.)

Si $f : I \to \mathbb{C}$ est une fonction ou bien boélienne positive, ou bien intégrable sur I, on écrira $\int_I f(x)\,dx$ au lieu de $\int_I f\,d\lambda_1$. Ceci est en accord avec ce qu'on a vu à la fin du chapitre précédent, à savoir que la théorie de Lebesgue est une extension de la théorie de Riemann. Comme la notation "Riemannienne" $f(x)\,dx$ est sans doute plus familière que la notation "Lebesguienne" $f\,d\lambda_1$, on préférera la première à la seconde, même si l'intervalle d'intégration I n'est pas fermé borné. Bien entendu, on peut aussi écrire $\int_I f(t)\,dt$ ou $\int_I f(u)\,du$...

Si $I = (a,b)$ avec $-\infty \leqslant a < b \leqslant \infty$, on écrira en fait généralement $\int_a^b f(x)\,dx$ plutôt que $\int_{(a,b)} f(x)\,dx$; et on pourra même écrire $\int_a^b f$ si cela semble judicieux. (Ces notations ne sont pas ambigues, car du point de vue de l'intégration la nature de l'intervalle I n'a aucune importance ; *cf* le Corollaire 2.14 du Chapitre 4.)

On dira qu'une fonction $f : I \to \mathbb{C}$ est **localement intégrable** sur I si elle est intégrable sur tout intervalle compact $[\alpha,\beta] \subseteq I$. Par exemple (et ceci est particulièrement important), toute fonction *continue* est localement intégrable sur I.

La "relation de Chasles"

$$\int_u^w f = \int_u^v f + \int_v^w f$$

est valable pour toute fonction $f : I \to \mathbb{C}$ borélienne positive ou localement intégrable, et pour $u,v,w \in I$ vérifiant $u < v < w$. (Cela vient du fait que $[u,v] = [u,w[\cup [w,v]$ et donc $\int_{[u,v]} = \int_{[u,w[} + \int_{[w,v]}$.)

Si $u,v \in I$ avec $u < v$, on pose

$$\int_v^u f(x)\,dx := -\int_u^v f(x)\,dx\,.$$

Avec cette convention, la relation de Chasles est valable pour tous $u,v,w \in I$, quel que soit l'ordre dans lequels les points sont rangés, *à condition* que la fonction f soit localement intégrable sur I. On a besoin de supposer f localement intégrable pour éviter les blagues du style $\infty - \infty$.

Exercice. Soit $f : I \to \mathbb{C}$ une fonction borélienne. Montrer que f est localement intégrable si et seulement si, pour tout $t \in I$, on peut trouver un voisinage V de t dans I tel que f est intégrable sur V.

2. Des choses *très* importantes à ne pas oublier

Le résultat sans conteste le plus important concernant les "intégrales pour enfants" est ce qu'on appelle le **théorème fondamental de l'analyse** :

THÉORÈME 2.1. *Soit I un intervalle de $\mathbb{R}$ et soit $f : I \to \mathbb{C}$ une fonction* continue. *Soit également $x_0 \in I$, et soit $F : I \to \mathbb{C}$ la fonction définie par*

$$F(x) = \int_{x_0}^{x} f(t)\, dt\,.$$

Alors F est de classe $\mathcal{C}^1$ et $F' = f$.

Démonstration. Comme f est continue, il suffit de montrer que F est *dérivable* en tout point, avec $F' = f$.

Fixons un point $x \in I$. Si $h \in \mathbb{R}\backslash\{0\}$ est tel que $x + h \in I$, alors on a d'une part

$$F(x+h) - F(x) = \int_x^{x+h} f(t)\, dt \quad \text{d'après la relation de Chasles};$$

et d'autre part

$$f(x) = \frac{1}{h} \times h\, f(x) = \frac{1}{h}\int_x^{x+h} f(x)\, dt\,.$$

Par linéarité de l'intégrale, on en déduit :

$$\frac{F(x+h) - F(x)}{h} - f(x) = \frac{1}{h}\int_x^{x+h} \big(f(t) - f(x)\big)\, dt\,.$$

En notant $I_x(h)$ est l'intervalle fermé d'extrémités x et $x + h$, on a donc

$$\begin{aligned} \left|\frac{F(x+h) - F(x)}{h} - f(x)\right| &= \frac{1}{|I_x(h)|}\left|\int_{I_x(h)} \big(f(t) - f(x)\big)\, dt\right| \\ &\leqslant \frac{1}{|I_x(h)|}\int_{I_x(h)} \big|f(t) - f(x)\big|\, dt \\ &\leqslant \sup_{t \in I_x(h)} |f(t) - f(x)|\,, \end{aligned}$$

d'où on déduit immédiatement, par continuité de f au point x, que $\frac{F(x+h)-F(x)}{h}$ tend vers $f(x)$ quand $h \to 0$. $\square$

COROLLAIRE 2.2. *Toute fonction continue sur un intervalle $I \subseteq \mathbb{R}$ possède des primitives.*

C'est évident par le théorème ; mais très loin de l'être si on ne dispose pas de l'outil "intégrale" !

COROLLAIRE 2.3. *Si $f = I \to \mathbb{C}$ est continue et si $a, b \in I$, alors*

$$\int_a^b f(t)\, dt = \big[F\big]_a^b := F(b) - F(a)\,,$$

où F est n'importe quelle primitive de f sur I.

Démonstration. La quantité $\big[F\big]_a^b$ ne dépend pas de la primitive F choisie, car deux primitives quelconques de f sur l'intervalle I *diffèrent d'une constante* (ce qui n'est pas évident) ; et si on prend $F(x) := \int_{x_0}^x f(t)\, dt$ comme dans le théorème, alors $[F]_a^b = \int_{x_0}^b f - \int_{x_0}^a f = \int_a^b f$ par la relation de Chasles. $\square$

Corollaire 2.4. *Si $F : I \to \mathbb{C}$ est de classe $\mathcal{C}^1$, alors on a pour tous $u, v \in I$:*

$$F(v) - F(u) + \int_u^v F'(t)\, dt\,.$$

Démonstration. On applique le corollaire précédent à $f := F'$! □

Corollaire 2.5. (formule d'intégration par parties)
Si u et v sont deux fonctions de classe $\mathcal{C}^1$ sur I, alors on a pour tous $a, b \in I$:

$$\int_a^b u'v = \big[uv\big]_a^b - \int_a^b uv'\,.$$

Démonstration. C'est clair par le corollaire précédent appliqué à $F := uv$: on a $F' = u'v + uv'$ et donc $[F]_a^b = \int_a^b u'v + \int_a^b uv'$. □

Bien que très facile à démontrer, la formule d'intégration par parties est *extraordinairement utile*, dans des situations très variées. Sans exagérer, on peut dire qu'une fois sur deux, intégrer par parties règle les difficultés qu'on peut avoir avec une intégrale.

Exercice 1. Déterminer les primitives de $f(t) = \log(t)$ sur $]0, \infty[$.

Exercice 2. Calculer *par parties* $\int_0^1 \frac{dt}{1+t^2}$ en écrivant $\frac{1}{1+t^2} = 1 \times \frac{1}{1+t^2}$, et en déduire la valeur de l'intégrale $\int_0^1 \frac{dt}{(1+t^2)^2}$.

Exercice 3. Soit $f : [a, b] \to \mathbb{C}$ une fonction de classe $\mathcal{C}^1$ sur un segment $[a, b] \subseteq \mathbb{R}$, et soit $\phi : [a, b] \to \mathbb{R}$ de classe $\mathcal{C}^2$. Soit également $g : \mathbb{R} \to \mathbb{C}$ une fonction continue. On suppose qu'on a $\phi'(t) \neq 0$ pour tout $t \in [a, b]$, et que $G(x) := \int_a^x g(t)\, dt$ est *borné* sur $\mathbb{R}$. Montrer qu'on a

$$\lim_{|\lambda| \to \infty} \int_a^b f(t) g(\lambda\phi(t))\, dt = 0\,.$$

(On pourra écrire $f(t)g(\lambda\phi(t)) = \frac{f(t)}{\phi'(t)} \times g(\lambda\phi(t))\phi'(t)$.)

Exercice 4. Montrer que pour tout $\lambda > 0$ et pour tout $n \in \mathbb{N}$, on a

$$\int_0^\infty t^n e^{-\lambda t} dt = \frac{1}{\lambda^{n+1} n!}\cdot$$

Exercice 5. Pour $n \in \mathbb{N}$, on pose $I_n := \int_0^{\pi/2} (\sin t)^n dt$. (Les I_n sont les très célèbres **intégrales de Wallis**.) Trouver une relation de récurrence entre I_{n+2} et I_n, et en déduire des expressions de I_{2k} et I_{2k+1} à base de factorielles. Calculer également $u_n := I_{n+1} I_n$ pour tout $n \geqslant 0$, puis déterminer un équivalent simple de I_n quand $n \to \infty$.

Corollaire 2.6. (formule de changement de variable)
Si $\phi : [a, b] \to \mathbb{R}$ est une fonction de classe $\mathcal{C}^1$ *sur un intervalle* fermé borné $[a, b]$, *et si f est une fonction* continue *sur un intervalle I contenant $\phi([a, b])$, alors*

$$\int_a^b f(\phi(t))\, \phi'(t)\, dt = \int_{\phi(a)}^{\phi(b)} f(x)\, dx\,.$$

Démonstration. Si on pose $F(u) := \int_{\phi(a)}^u f(x)\, dx$ et $G(t) := F(\phi(t))$, alors F est de classe $\mathcal{C}^1$ sur I avec $F' = f$ et G est de classe $\mathcal{C}^1$ sur $[a, b]$ avec $G'(t) = F'(\phi(t))\, \phi'(t) = f(\phi(t))\, \phi'(t)$; donc

$$\int_a^b f(\phi(t))\, \phi'(t)\, dt = \big[G\big]_a^b = \big[F \circ \phi\big]_a^b = \big[F\big]_{\phi(a)}^{\phi(b)} = \int_{\phi(a)}^{\phi(b)} f(x)\, dx\,.$$

□

Remarque. La formule de changement de variable semble un peu rébarbative ; mais comme chacun sait, elle est très utile et son mode d'emploi est simple avec un peu d'habitude. Il y a en fait *deux* façons d'utiliser cette formule.

- Si on doit calculer une intégrale $\int_\alpha^\beta f(x)\,dx$ et si, pour une raison ou pour une autre, il semble judicieux de "changer de variable d'intégration" en posant $x = \phi(t)$ pour une certaine fonction ϕ, on peut le faire à condition d'écrire $dx = \phi'(t)dt$ et de bien changer les bornes d'intégrations.
- En sens inverse, si on doit calculer une intégrale de la forme $\int_a^b g(t)\,dt$ et si, dans l'expression de $g(t)$, apparait de façon claire un terme x de la forme $\phi(t)$ qui permet d'écrire g de façon plus simple, on peut prendre x comme nouvelle variable d'intégration à condition d'exprimer $g(t)dt$ uniquement en fonction de x et de dx. C'est possible si la fonction ϕ est un "vrai" changement de variable, i.e. est une bijection de $[a,b]$ sur $\phi([a,b])$ (ce qui revient à dire que ϕ est strictement monotone) : on peut alors écrire $t = \psi(x)$ où $\psi = \phi^{-1}$, et $dt = \psi'(x)dx$. Concrètement : on écrit $x = \phi(t)$, $dx = \phi'(t)dt$, de sorte que $g(t)dt = \frac{g(t)}{\phi'(t)}dx := h(t)dx$, et il faut se débrouiller pour exprimer $h(t)$ en fonction de x.

En réalité, la véritable difficulté est de *deviner* quel changement de variable peut se révéler utile.

Exemple 1. Si on veut calculer l'intégrale

$$I := \int_{-1}^1 \sqrt{1-x^2}\,dx\,,$$

il semble assez naturel de poser $x = \cos t$, et donc $dx = -\sin t\,dt$. On a $x = -1$ pour $t = \pi$ et $x = 1$ pour $t = 0$; et $\sqrt{1-x^2} = \sqrt{1-\cos^2 t} = \sqrt{\sin^2 t} = \sin t$ si $t \in [0,\pi]$, car $\sin t \geqslant 0$ sur $[0,\pi]$. Donc

$$\int_{-1}^1 \sqrt{1-x^2}\,dx = \int_\pi^0 \sin t \times (-\sin t\,dt) = \int_0^\pi \sin^2 t\,dt\,.$$

On peut alors terminer le calcul en utilisant la formule trigonométrique

$$\sin^2 t = \frac{1-\cos(2t)}{2}\,,$$

ce qui donne

$$I = \int_0^\pi \frac{1-\cos(2t)}{2}\,dt = \frac{1}{2}\left[t - \frac{1}{2}\sin(2t)\right]_0^\pi = \frac{\pi}{2}\cdot$$

On aurait pu avoir l'idée saugrenue de dire que $x = \cos t$ vaut -1 pour $t = 21\pi$ et 1 pour $t = 48\pi$. Dans ce cas, il aurait fallu écrire

$$\int_{-1}^1 \sqrt{1-x^2}\,du = \int_{21\pi}^{48\pi} \sqrt{\sin^2 t}\,\sin t\,dt = \int_{21\pi}^{48\pi} |\sin t|\,\sin t\,dt\,,$$

et faire bien attention au signe de $\sin t$...

Exercice 1. Expliquer la valeur de I en faisant un dessin.

Exercice 2. Calculer l'intégrale $\int_0^1 x\sqrt{1-x^2}\,dx$.

Exemple 2. Si on veut calculer l'intégrale

$$I := \int_{\pi/3}^{\pi/2} \frac{dt}{\sin t}\,,$$

on peut utiliser le changement de variable $x := \tan(t/2)$. On a alors

$$dx = \frac{1}{2}(1+\tan^2(t/2))dt = \frac{1}{2}(1+x^2)dt\,,$$

ce qui donne

$$dt = \frac{2dx}{1+x^2}\,;$$

et $\sin t = 2\sin(t/2)\cos(t/2) = 2\tan(t/2)\cos^2(t/2)$, autrement dit

$$\sin t = \frac{2x}{1+x^2}\cdot$$

Comme $x = \tan(\pi/6) = \frac{1}{\sqrt{3}}$ si $t = \pi/3$ et $x = \tan(\pi/4) = 1$ si $t = \pi/2$, on obtient ainsi

$$I = \int_{1/\sqrt{3}}^{1} \frac{1+x^2}{2x} \times \frac{2dx}{1+x^2} = \int_{1/\sqrt{3}}^{1} \frac{dx}{x} = \frac{\log(3)}{2}\cdot$$

Exercice 1. Calculer les intégrale $\int_0^{\pi/2} \frac{dt}{2+\sin t}$ et $\int_0^{\pi/2} \frac{dt}{2+\cos t}\cdot$

Exercice 2. Calculer $\int_0^1 \frac{dx}{(1+x^2)^2}$ en posant $x = \tan t$.

3. Passage à la limite dans les bornes d'intégration

Le résultat suivant est intuitivement évident, mais en fait il ne l'est pas tant que ça.

PROPOSITION 3.1. *Soit I un intervalle de la forme $[a,b[$, avec $a < b \leqslant \infty$. Si $f : [a,b[\to \mathbb{C}$ est une fonction borélienne ou bien positive, ou bien intégrable sur $[a,b[$, alors*

$$\int_a^b f(t)\,dt = \lim_{\beta\to b^-} \int_a^\beta f(t)\,dt\,,$$

où la limite existe dans $[0,\infty]$ si f est positive, et dans $\mathbb{C}$ si f est intégrable sur $[a,b[$.

Démonstration. Commençons par traiter le cas où $f \geqslant 0$. Alors $\int_a^\beta f$ est une fonction croissante de β, donc $\lim_{\beta\to b^-}\int_a^\beta f$ existe dans $[0,\infty]$. Soit $(\beta_n) \subseteq [a,\beta[$ une suite croissante tendant vers β. On a

$$\int_a^{\beta_n} f = \int_a^b \mathbf{1}_{[a,\beta_n[}\, f := \int_a^b f_n.$$

Les fonctions f_n forment une suite *croissante* car la suite (β_n) est croissante et $f \geqslant 0$; et de plus $f_n(t) \to f(t)$ pour tout $t \in [a,b[$, car tout point $t \in [a,b[$ appartient à $[a,\beta_n[$ pour n assez grand. Par convergence monotone, on en déduit que $\int_a^{\beta_n} f \to \int_a^b f$, et donc $\int_a^b f = \lim_{\beta\to b^-}\int_a^\beta f$.

Le cas où f est intégrable se déduit très facilement du "cas positif" en écrivant $f = u + iv = (u^+ - u^-) + i(v^+ - v^-)$ et en appliquant le cas positif aux fonctions u^+, u^-, v^+, v^-. (Écrire les détails.) $\square$

Remarque 3.2. On a évidemment un résultat analogue pour un intervalle de la forme $I =]a, b]$; et donc, un résultat du même type pour un intervalle *ouvert* $I =]a, b[$, en faisant *deux passages à la limite* : si $f :]a, b[\to \mathbb{C}$ est une fonction borélienne ou bien positive, ou bien intégrable sur $]a, b[$, alors

$$\int_a^b f(t)\,dt = \lim_{\alpha\to a^+}\lim_{\beta\to b^-}\int_\alpha^\beta f(t)\,dt = \lim_{\beta\to b^-}\lim_{\alpha\to a^+}\int_\alpha^\beta f(t)\,dt.$$

Corollaire 3.3. *Si $f :]a, b[\to \mathbb{C}$ est une fonction continue, ou bien positive ou bien intégrable sur $[a, b[$, alors on a le droit d'écrire*

$$\int_a^b f(t)\,dt = [F]_a^b$$

où F est une primitive quelconque de f sur $]a, b[$, à condition d'interpréter $F(b)$ comme $\lim_{\beta\to b^-} F(\beta)$, qui existe dans $] - \infty, \infty]$ si f est positive et dans $\mathbb{C}$ si f est intégrable, et $F(a)$ comme $\lim_{\alpha\to a^+} F(\alpha)$, qui existe dans $[-\infty, \infty[$ si f est positive et dans $\mathbb{C}$ si f est intégrable.

Démonstration. Supposons d'abord $f \geqslant 0$. Si on fixe α tel que $a < \alpha < b$, on a $\int_\alpha^\beta f(t)\,dt = [F]_\alpha^\beta = F(\beta) - F(\alpha)$ pour tout $\alpha < \beta < b$, et $\int_\alpha^\beta f(t)\,dt$ tend vers $\int_\alpha^b f(t)\,dt$ quand $\beta \to b^-$. Donc $F(\beta)$ a une limite $F(b) \in]-\infty, \infty]$ quand $\beta \to b^-$, à savoir $F(b) = F(\alpha) + \int_\alpha^b f(t)\,dt$ (qui ne dépend pas de α). De même, on montre que $F(\alpha)$ a un limite $F(a) \in [-\infty, \infty[$ quand $\alpha \to a^+$. On peut donc faire $\alpha \to a$ et $\beta \to b$ dans la formule $\int_\alpha^\beta f(t)\,dt = [F]_\alpha^\beta = F(\beta) - F(\alpha)$ pour obtenir le résultat souhaité. (Il n'y a pas d'indétermination pour $[F]_a^b = F(b) - F(a)$ car $F(b) > -\infty$ et $F(a) < \infty$.)

La preuve est la même dans le cas "intégrable. □

Corollaire 3.4. *Soit $\phi :]a, b[\to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ et* monotone. *Si $f : J \to \mathbb{C}$ est une fonction continue sur un intervalle J contenant $\phi(]a, b[)$, et si f est ou bien positive ou bien intégrable sur $\phi(]a, b[)$, alors on a le droit d'écrire*

$$\int_{\phi(a)}^{\phi(b)} f(x)\,dx = \int_a^b f(\phi(t))\,\phi'(t)dt,$$

en interprétant $\phi(b)$ et $\phi(a)$ comme $\lim_{\beta\to b^-}\phi(\beta)$ et $\lim_{\alpha\to a^+}$, qui existent dans $\overline{\mathbb{R}}$.

Démonstration. Les limites $\lim_{\alpha\to a^+}\phi(\beta)$ et $\lim_{\beta\to b^-}\phi(\beta)$ existent dans $\overline{\mathbb{R}}$ par monotonie de ϕ.

(i) Supposons f positive et ϕ croissante. Par la formule de changement de variable "classique", on a

$$\int_{\phi(\alpha)}^{\phi(\beta)} f(x)\,dx = \int_a^\beta f(\phi(t))\,\phi'(t)dt \quad \text{pour tous } a < \alpha < \beta < b;$$

et comme f et $(f \circ \phi)\phi'$ sont des fonctions positives, on en déduit immédiatement le résultat par la Proposition.

(i') Si f est positive et ϕ décroissante, la preuve est la même (en mettant des signes "$-$" aux bons endroits).

(ii) Si f est intégrable sur $\phi(]a, b[)$, on commence par appliquer le "cas positif" à $|f|$ pour montrer que $(f \circ \phi)$ est intégrable sur $]a, b[$; puis on réapplique le cas positif aux parties positives et négatives de $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$. Les détails sont laissés en exercice. □

Remarque 3.5. Ce qui vient d'être fait pour un intervalle de la forme $I = [a, b[$ vaut évidemment pour un intervalle de la forme $I =]a, b]$; et par conséquent, cela vaut encore pour un intervalle *ouvert* $I =]a, b[$, en faisant *deux passages à la limite.*

Exemple 1. Soit $\alpha > 0$. Pour tout $A > 0$, on a $\int_A^\infty \frac{dt}{t^\alpha} < \infty$ si et seulement si $\alpha > 1$; et pour tout $\varepsilon > 0$, on a $\int_0^\varepsilon \frac{dt}{t^\alpha} < \infty$ si et seulement si $\alpha < 1$.

Démonstration. Si $\alpha > 1$, alors

$$\int_A^\infty \frac{dt}{t^\alpha} = \left[-\frac{1}{\alpha - 1}\frac{1}{t^{\alpha-1}}\right]_A^\infty = \frac{1}{\alpha-1}\frac{1}{A^{\alpha-1}} < \infty\,;$$

si $\alpha = 1$, alors

$$\int_1^\infty \frac{dt}{t^\alpha} = \int_A^\infty \frac{dt}{t} = [\log(t)]_A^\infty = \infty\,;$$

et si $\alpha < 1$, alors

$$\int_A^\infty \frac{dt}{t^\alpha} = \left[\frac{1}{1-\alpha}\, t^{1-\alpha}\right]_A^\infty = \infty\,.$$

La preuve est la même pour l'intégrale entre 0 et ε. □

Exemple 2. Pour calculer l'intégrale

$$I := \int_0^\infty \frac{dt}{\sqrt{t}(1+t)}\,,$$

on peut utiliser le changement de variable $x = \sqrt{t}$, car $\phi(t) = \sqrt{t}$ est de classe $\mathcal{C}^1$ et strictement monotone sur $]0, \infty[$. On a $dx = \frac{dt}{2\sqrt{t}}$, donc

$$\int_0^\infty \frac{dt}{\sqrt{t}(1+t)} = \int_0^\infty \frac{2dx}{1+x^2} = [2\arctan(x)]_0^\infty = \pi\,.$$

4. Critères d'intégrabilité

Par définition (ou presque), une fonction borélienne $f : I \to \mathbb{C}$ est intégrable sur I si et seulement si on a $\int_I |f(t)|\, dt < \infty$. On en déduit aussitôt le fait suivant, qu'on peut appeler le **principe de comparaison**.

Fait 4.1. *Soit $f : I \to \mathbb{C}$ une fonction borélienne. S'il existe une fonction borélienne $g : I \to \mathbb{R}^+$ telle que $\int_I g(t)\, dt < \infty$ et $|f(t)| \leqslant g(t)$ pour tout $t \in I$ (ou seulement presque partout), alors f est intégrable sur I.*

En particulier :

Exemple 4.2. Si l'intervalle I est *borné*, alors toute fonction borélienne *bornée* $f : I \to \mathbb{C}$ est intégrable sur I.

Démonstration. Si $|f(t)| \leqslant M$ sur I, on peut prendre $g(t) \equiv M$ dans le critère de comparaison. □

Bien entendu, ceci est *complètement faux* si l'intervalle I n'est pas borné : les fonctions constantes sont intégrables sur I seulement lorsque I est borné.

Le principe de comparaison est une tautologie. En voici une version légérement différente, plus directement applicable dans certains cas. Rappelons qu'une fonction est dite *localement intégrable* si elle est intégrable sur tout intervalle compact $[\alpha, \beta] \subseteq I$.

Fait 4.3. *Soit $f : I \to \mathbb{C}$ une fonction borélienne localement intégrable sur un intervalle I de la forme $I = [a, b[$, où $a < b \leqslant \infty$. On suppose qu'il existe une fonction borélienne positive g, intégrable sur un certain intervalle $[\beta, b[$ avec $\beta < b$, telle que $|f(t)| = O(g(t))$ quand $t \to b^-$. Alors f est intégrable sur $[a, b[$.*

Démonstration. Par hypothèse, on peut trouver β_0 avec $\beta \leqslant \beta_0 < b$ et une constante $C < \infty$ tels que $|f(t)| \leqslant C\, g(t)$ sur $[\beta_0, b[$. On a donc

$$\begin{aligned}
\int_a^b |f(t)|\, dt &= \int_a^{\beta_0} |f(t)|\, dt + \int_{\beta_0}^b |f(t)|\, dt \\
&\leqslant \int_a^{\beta_0} |f(t)|\, dt + C \int_{\beta_0}^b g(t)\, dt \\
&< \infty \qquad \text{car } f \text{ est intégrable sur } [a, \beta_0] \text{ et } \textstyle\int_{\beta_0}^b g(t)\, dt < \infty\,.
\end{aligned}$$

□

Remarque 1. On a bien sûr un énoncé analogue pour un intervalle I de la forme $I =]a, b]$, avec $-\infty \leqslant a < b$.

Remarque 2. Pour un intervalle *ouvert* $I =]a, b[$, on peut montrer qu'un fonction $f :]a, b[\to \mathbb{C}$ est intégrable sur $]a, b[$ en fixant $x_0 \in]a, b[$ et en appliquant deux fois le critère précédent, sur $]a, x_0]$ et sur $[x_0, b[$.

Voici un exemple typique et *très important* d'utilisation du "principe de comparaison modifié".

Exemple 4.4. (critère d'intégrabilité "de Riemann")

(1) Soit $A < \infty$, et soit $f : [A, \infty[\to \mathbb{C}$ une fonction localement intégrable (par exemple continue). S'il existe $\alpha > 1$ tel que $|f(t)| = o\left(\frac{1}{t^\alpha}\right)$ quand $t \to \infty$, alors f est intégrable sur $[A, \infty[$.

(2) Soit $\varepsilon > 0$, et soit $f :]0, \varepsilon] \to \mathbb{C}$ une fonction localement intégrable. S'il existe $\alpha < 1$ tel que $|f(t)| = o\left(\frac{1}{t^\alpha}\right)$ quand $t \to 0^+$, alors f est intégrable sur $]0, \varepsilon]$.

Démonstration. Il suffit d'observer que la fonction $t \mapsto \frac{1}{t^\alpha}$ est intégrable sur $[1, \infty[$ si $\alpha > 1$, et intégrable sur $]0, 1]$ si $\alpha < 1$. □

Exercice. Montrer que $f(t) := \frac{\log(t)}{\sqrt{t}}\, e^{-t^{1/4}}$ est intégrable sur $]0, \infty[$.

5. Intégrales "généralisées"

Il existe une croyance malheureusement fort répandue selon laquelle "on ne peut pas traiter les intégrales généralisées avec l'intégrale de Lebesgue, alors qu'on peut le faire avec l'intégrale de Riemann". Ceci est absurde : la définition d'une intégrale généralisée est exactement la même dans les deux théories ; et bien entendu on peut traiter davantage d'intégrales avec la théorie de Lebesgue.

Définition 5.1. *Soit I un intervalle de la forme $[a, b[$ avec $-\infty < a < b \leqslant \infty$, et soit $f : [a, b[\to \mathbb{C}$ localement intégrable. Si la limite $\lim_{\beta \to b^-} \int_a^\beta f(t)\, dt$ existe dans $\mathbb{C}$, on note cette limite $\int_a^b f(t)\, dt$ et on dit que $\int_a^b f(t)\, dt$* ***existe en tant qu'intégrale généralisée****, ou bien que* ***l'intégrale $\int_a^b f(t)\, dt$ est convergente****.*

Remarque 1. Il y a une analogie évidente avec les séries. Le rôle des sommes partielles $\sum_{k=0}^{n} u_k$ est joué par les "intégrales partielles" $\int_a^\beta f(t)\,dt$.

Remarque 2. On peut bien entendu donner une définition analogue pour une fonction $f :]a,b] \to \mathbb{C}$, où $-\infty \leqslant a < b < \infty$. De même, on peut définir l'intégrale généralisée d'une fonction f (localement intégrable) définie sur un intervalle *ouvert* $]a,b[$, avec $-\infty \leqslant a < b \leqslant \infty$: on fixe un point $x_0 \in]a,b[$, et on dit que $\int_a^b f(t)\,dt$ existe si $\int_a^{x_0}$ et $\int_{x_0}^b f(t)\,dt$ existent séparément. (Il faut vérifier que ceci ne dépend pas du point x_0, ce qui est très facile.)

Faux exemple. Si la fonction f est **intégrable** sur $[a,b[$, alors l'intégrale $\int_a^b f(t)\,dt$ est convergente : c'est ce que dit la Proposition 3.1. Mais dans ce ce cas, on n'a pas du tout affaire à une intégrale "généralisée" : $\int_a^b f(t)\,dt$ est une intégrale "ordinaire", par rapport à la mesure de Lebesgue. Si on veut à tout prix utiliser une terminologie Riemannienne, on peut cependant dire que la fonction f (supposée localement intégrable) est intégrable sur $[a,b[$ si et seulement si l'intégrale $\int_a^b f(t)\,dt$ est **absolument convergente**.

Lorsqu'on n'est pas dans le cadre "intégrable", l'outil principal pour montrer qu'une certaine intégrale généralisée existe est le

Critère de Cauchy. Soit $f : [a,b[\to \mathbb{C}$ une fonction localement intégrable. L'intégrale $\int_a^b f(t)\,dt$ existe en tant qu'intégrale généralisée si et seulement si

$$\int_X^Y f(t)\,dt \to 0 \quad \text{quand } X,Y \to b^- . \tag{5.1}$$

Démonstration. Si on pose $F(x) := \int_a^x f(t)\,dt$, alors dire que $\int_a^b f(t)\,dt$ existe signifie que $F(x)$ admet une limite quand $x \to b^-$. D'après le *critère de Cauchy pour les fonctions*, il revient au même de dire que $F(Y) - F(X) \to 0$ quand $X,Y \to b^-$; autrement dit que (5.1) est vérifiée. □

Le critère de Cauchy signifie que pour montrer l'existence d'une intégrale généralisée $\int_a^b f(t)\,dt$, il faut être capable de majorer en module des intégrale du type $\int_X^Y f(t)\,dt$. La proposition suivante donne deux cas importants où cela est possible.

Proposition 5.2. (critères d'Abel)
Soient $f,g : [a,\infty[\to \mathbb{C}$, avec f continue et g de classe $\mathcal{C}^1$. On suppose de plus que la fonction g' est intégrable sur $[a,\infty[$, i.e.

$$\int_a^\infty |g'(t)|\,dt < \infty .$$

Dans chacun des deux cas suivants, l'intégrale $\int_a^\infty f(t)g(t)\,dt$ est convergente.

(1) *L'intégrale $\int_a^\infty f(t)\,dt$ existe.*

(2) *$g(t) \to 0$ quand $t \to \infty$, et il existe une constante C telle que*

$$\forall X \geqslant a \ : \ \left| \int_a^X f(t)\,dt \right| \leqslant C .$$

Démonstration. Remarquons d'abord que l'hypothèse d'intégrabilité de g' entraine que la fonction g est *bornée* sur $[a,\infty[$: en effet, on a

$$|g(x)| = \left| g(a) + \int_a^x g'(t)\ dt \right| \leqslant |g(a)| + \int_a^x |g'(t)|\, dt \leqslant M := |g(a)| + \int_a^\infty |g'(t)|\, dt$$

pour tout $x \in [a,\infty[$.

(1) Supposons que $\int_a^\infty f(t)\, dt$ existe, et posons pour $x \geqslant a$:

$$F(x) := \int_x^\infty f(t)\, dt = \int_a^\infty f(t)\, dt - \int_a^x f(t)\, dt\,.$$

Comme f est suposée continue, la fonction F est de classe $\mathcal{C}^1$ sur $[a,\infty[$ et $F' = -f$. *De plus*, $F(x) \to 0$ quand $x \to b^-$ car $F(x) = \int_a^\infty f(t)\, dt - \int_a^x f(t)\, dt$; et en particulier, la fonction F est bornée sur $[a,\infty[$.

Si $a \leqslant X < Y < \infty$, une intégration par parties donne

$$\begin{aligned} \int_X^Y f(t)\, g(t)\, dt &= [-Fg]_X^Y + \int_X^Y F(t) g'(t)\, dt \\ &= F(X)g(X) - F(Y)g(Y) + \int_X^Y F(t)\, g'(t)\, dt\,; \end{aligned}$$

et on en déduit

$$\left| \int_X^Y f(t)\, g(t)\, dt \right| \leqslant \Big(|F(X)| + |F(Y)| \Big) \|g\|_\infty + \|F\|_\infty \times \int_X^\infty |g'(t)|\, dt\,.$$

Comme $F(x) \to 0$ quand $x \to \infty$ et comme $\int_X^\infty |g'| \to 0$ quand $X \to \infty$ d'après la Proposition 3.1 (puisque $\int_X^\infty |g'| = \int_a^\infty |g'| - \int_a^X |g'|$), on voit donc que le critère de Cauchy est vérifié.

(2) On intégre à nouveau par parties, mais cette fois en considérant

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

qui est de classe $\mathcal{C}^1$ avec $F' = f$. L'hypothèse faite sur f signifie que la fonction F est *bornée* sur $[a,\infty[$.

En intégrant par parties, on trouve pour $a \leqslant X < Y < \infty$:

$$\begin{aligned} \int_X^Y f(t)\, g(t)\, dt &= [F\, g]_X^Y - \int_X^Y F(t)\, g'(t)\, dt \\ &= F(Y)g(Y) - F(X)g(X) - \int_X^Y F(t)\, g'(t)\, dt\,. \end{aligned}$$

On en déduit

$$\left| \int_X^Y f(t)\, g(t)\, dt \right| \leqslant \Big(|g(X)| + |g(Y)| \Big) \|F\|_\infty + \|F\|_\infty \times \int_X^\infty |g'(t)|\, dt\,;$$

et donc le critère de Cauchy est vérifié, car $g(t) \to 0$ quand $t \to \infty$ et $\int_X^\infty |g'| \to 0$ quand $X \to \infty$ par la Proposition 3.1. $\square$

Corollaire 5.3. *Soit* $f : [a, \infty[\to \mathbb{C}$ *continue, et soit* $g : [a, \infty[\to \mathbb{R}$ *de classe* $\mathcal{C}^1$. *On suppose que* g *est* positive et décroissante, *avec* $\lim_{t\to\infty} g(t) = 0$, *et qu'il existe une constante* C *telle que*

$$\forall X \geqslant a \ : \ \left| \int_a^X f(t)\, dt \right| \leqslant C\,.$$

Alors l'intégrale $\int_a^\infty f(t)g(t)\, dt$ *est convergente.*

Démonstration. D'après (2), il suffit de vérifier que g' est intégrable sur $[a, \infty[$; ce qui est facile : comme g est décroissante, on a $|g'(t)| = -g'(t)$ pour tout t, donc

$$\int_a^\infty |g'(t)|\, dt = \int_a^\infty (-g'(t))\, dt = [-g(t)]_a^\infty = g(a) < \infty\,.$$

□

Remarque. La proposition est en fait valable en supposant seulement f localement intégrable (pas forcément continue). La preuve est essentiellement la même, mais il faut préciser ce qu'on entend par "intégration par parties" dans ce cas. (C'est possible, mais on ne le fera pas.) De même, le corollaire est valable sans suposer g de classe $\mathcal{C}^1$; mais dans ce cas, la preuve classique utilise ce qu'on appelle la *2ème formule de la moyenne* ; dont on ne parlera pas.

Exemple 5.4. Pour tout $\alpha > 0$, l'intégrale $\int_1^\infty \frac{e^{it}}{t^\alpha}\, dt$ est convergente ; et donc les intégrales $\int_1^\infty \frac{\cos t}{t^\alpha}\, dt$ et $\int_1^\infty \frac{\sin t}{t^\alpha}\, dt$ le sont aussi.

Démonstration. On applique le "2ème critère d'Abel" avec $f(t) := e^{it}$ et $g(t) := \frac{1}{t^\alpha}$, qui est positive et décroit vers 0 quand $t \to \infty$. Pour tout $X \geqslant 0$, on a

$$\left| \int_1^X e^{it}\, dt \right| = \left| \frac{1}{i} \left(e^{iX} - e^{i1} \right) \right| \leqslant 2\,;$$

donc le critère s'applique bel et bien. □

Exercice. Montrer que $\frac{\sin t}{t}$ n'est *pas intégrable* sur $[1, \infty[$. (On pourra par exemple observer que $|\sin t| \geqslant \sin^2 t$ et utiliser une formule trigonométrique ; ou bien découper $[\pi, \infty[$ en intervalles de longueur π.)

Remarque. Il existe également des critères d'Abel *pour les séries*. Le plus classique est le suivant : *Soient* (a_k) *une suite de nombres complexes et* (b_k) *une suite de nombres réels décroissant vers* 0. *Si les sommes partielle* $S_n := \sum_{k=0}^n a_k$ *restent bornées, alors la série* $\sum a_k b_k$ *est convergente.* Plus généralement, au lieu de supposer que (b_k) est positive et décroit vers 0, on peut autoriser les b_k à être complexes, à condition de supposer que $b_k \to 0$ et qu'on a $\sum_0^\infty |b_{k+1} - b_k| < \infty$. La preuve repose sur une "intégration par parties discrète", c'est-à-dire une **transformation d'Abel**.

Exercice. Soit $\alpha > 0$. Montrer que pour tout $\theta \in \mathbb{R}\backslash 2\pi\mathbb{Z}$, la série $\sum \frac{e^{in\theta}}{n^\alpha}$ est convergente.

6. Comparaison série-intégrale

Il existe une sorte de règle empirique selon laquelle, pour une fonction $f : [0, \infty[\to \mathbb{C}$ à peu près sympathique, il est plus facile de déterminer si l'intégrale $\int_0^\infty f(t)\, dt$ converge que de déterminer si la série $\sum f(k)$ converge. C'est assez facile à comprendre : les intégrales se calculent souvent bien avec des primitives tandis que les sommes se calculent rarement bien ; une intégration par parties, c'est très simple et on ne se trompe pas, alors que pour une transformation d'Abel on se trompe dans les indices ; ...

Malheureusement, dans certaines situations c'est la question de la convergence de la série $\sum f(k)$ qu'il faut régler. Il est donc souhaitable de disposer de conditions assez générales permettant d'affirmer que les deux problèmes sont en fait équivalents. On va démontrer ici deux résultats de ce type.

PROPOSITION 6.1. *Soit $a \in \mathbb{N}$, et soit $f : [a, \infty[\to \mathbb{R}$ une fonction* positive et décroissante.

(1) *Si on pose $u_n := \sum_{k=a}^n f(k) - \int_a^n f(t)\,dt$ pour $n \geqslant a$, alors u_n admet une limite finie l quand l'entier n tend vers l'infini, et on a $l \geqslant f(a) - \int_a^{a+1} f(t)\,dt \geqslant 0$.*
(2) *La série $\sum f(k)$ est convergente si et seulement si $\int_a^\infty f(t)\,dt < \infty$.*
(3) *En cas de* divergence, *on a $\sum_{k=a}^n f(k) \sim \int_a^n f(t)\,dt$ quand $n \to \infty$.*

Démonstration. Remarquons d'abord que la fonction f est localement intégrable car elle est monotone, donc borélienne et bornée sur tout intervalle compact $[\alpha, \beta] \subseteq [a, \infty[$.

(1) L'idée est d'écrire $\int_a^n f(t)\,dt = \sum\limits_{k=a}^{n-1} \int_k^{k+1} f(t)\,dt$ (relation de Chasles), de sorte que

$$u_n = \sum_{k=a}^{n-1} \left(f(k) - \int_k^{k+1} f(t)\,dt \right) + f(n)\,. \tag{6.1}$$

Pour tout entier $k \geqslant a$, on a $f(k+1) \leqslant f(t) \leqslant f(k)$ si $k \leqslant t \leqslant k+1$ car la fonction f est décroissante ; d'où en intégrant entre k et $k+1$:

$$f(k+1) \leqslant \int_k^{k+1} f(t)\,dt \leqslant f(k) \quad \text{pour tout } k \geqslant a\,.$$

On en déduit d'une part que tous les termes de la somme apparaissant dans (6.1) sont positifs, et donc (en ne gardant que $k = a$)

$$u_n \geqslant f(a) - \int_a^{a+1} f(t)\,dt \geqslant 0\,;$$

et d'autre part que

$$u_{n+1} - u_n = f(n+1) - \int_n^{n+1} f(t)\,dt \leqslant 0\,.$$

Ainsi, la suite (u_n) est décroissante et minorée par $c(a) := f(a) - \int_a^{a+1} f(t)\,dt$; donc u_n admet une limite $l \geqslant c(a) \geqslant 0$ quand $n \to \infty$.

La partie (2) est évidente : d'après (1), la série $\sum f(k)$ converge si et seulement si $\int_a^n f(t)\,dt$ admet une limite finie quand l'entier n tend vers l'infini ; et comme f est positive, cela signifie exactement qu'on a $\int_a^\infty f(t)\,dt < \infty$. Enfin, (3) est également évident : si la série $\sum f(k)$ diverge, alors $\sum_{k=a}^n f(k)$ et $\int_a^n f(t)\,dt$ tendent tous les deux vers ∞ car $f \geqslant 0$, et sont donc équivalents car leur différence reste bornée d'après (1). $\square$

COROLLAIRE 6.2. *Pour $n \in \mathbb{N}^*$, posons $H_n := \sum_{k=1}^n \frac{1}{k}$. Alors $H_n - \log(n)$ admet une limite finie $\gamma > 0$ quand $n \to \infty$. (Le nombre γ s'appelle la* ***constante d'Euler****.) En particulier, $H_n \sim \log(n)$ quand $n \to \infty$.*

Démonstration. On applique la proposition avec $a := 1$ et $f(t) := \frac{1}{t}\cdot$ On a bien $\gamma > 0$, et même $\gamma \geqslant 1 - \int_1^2 \frac{dt}{t} = 1 - \log(2)$. $\square$

Exercice. Déterminer pour quelles valeurs de $\alpha > 0$ la série $\sum \frac{1}{k\log(k)^\alpha}$ est convergente.

Le deuxième résultat de "comparaison série-intégrale" qu'on va établir est valable pour des fonctions f non nécessairement positives.

Proposition 6.3. *Soit $a \in \mathbb{N}$ et soit $f : [a, \infty[\to \mathbb{C}$. On suppose que f est de classe $\mathcal{C}^1$ avec $\int_a^\infty |f'(t)|\,dt < \infty$, et que $f(t) \to 0$ quand $t \to \infty$. Alors la série $\sum f(k)$ est convergente si et seulement si l'intégrale $\int_a^\infty f(t)\,dt$ est convergente.*

Démonstration. Faisons d'abord deux remarques très simples.

Fait. L'intégrale $\int_a^\infty f(t)\,dt$ est convergente si et seulement si $\int_a^n f(t)\,dt$ a une limite dans $\mathbb{C}$ quand l'*entier* $n \in \mathbb{N}$ tend vers l'infini ; et la série $\sum f(k)$ est convergente si et seulement si $\sum_{k=a}^{n-1} f(k)$ a une limite quand $n \to \infty$.

Preuve du Fait. Pour tout nombre réel $X \geqslant a$, on a

$$\int_a^X f(t)\,dt = \int_a^{[X]} f(t)\,dt + \int_{[X]}^X f(t)\,dt\,.$$

De plus

$$\left|\int_{[X]}^X f(t)\,dt\right| \leqslant \int_{[X]}^X |f(t)|\,dt \leqslant \sup_{t \geqslant [X]} |f(t)|\,;$$

et donc $\int_{[X]}^X f(t)\,dt \to 0$ quand $X \to \infty$ car f tend vers 0 à l'infini. Cela prouve la première partie du Fait. La deuxième partie est évidente. $\square$

Par le Fait, il suffit maintenant de voir que

$$u_n := \int_a^n f(t)\,dt - \sum_{k=a}^{n-1} f(k)$$

admet une limite dans $\mathbb{C}$ quand l'entier n tend vers l'infini.

Écrivons

$$u_n = \sum_{k=a}^{n-1}\left(\int_k^{k+1} f(t)\,dt - f(k)\right) := \sum_{k=a}^{n-1} v_k\,.$$

Si on pose $F(x) := \int_a^x f(t)\,dt$, alors la fonction F est de classe $\mathcal{C}^2$ car f est supposée $\mathcal{C}^1$, avec $F' = f$ et $F'' = f'$. D'après la formule de Taylor, on a donc

$$\begin{aligned} v_k &= \int_k^{k+1} f(t)\,dt - f(k) \\ &= F(k+1) - F(k) - F'(k) \times ((k+1) - k) \\ &= \int_k^{k+1} (k+1-t)F''(t)\,dt = \int_k^{k+1} (k+1-t)f'(t)\,dt\,. \end{aligned}$$

On en déduit

$$|v_k| \leqslant \int_k^{k+1} \left|(k+1-t)f'(t)\right| dt \leqslant \int_k^{k+1} |f'(t)|\,dt \quad \text{pour tout } k \geqslant a\,;$$

et par conséquent

$$\begin{aligned}\sum_{k=a}^{\infty}|v_k| &\leqslant \sum_{k=a}^{\infty}\int_k^{k+1}|f'(t)|\,dt\\ &= \sum_{k=a}^{\infty}\int_{[k,k+1[}|f'(t)|\,dt\\ &= \int_a^{\infty}|f'(t)|\,dt<\infty\,,\end{aligned}$$

où on a utilisé la fait que l'application $A\mapsto\int_A|f'(t)|\,dt$ est une *mesure* (borélienne) sur $[a,\infty[$ (*cf* l'exercice 2.8 du Chapitre 4). Si on préfère, on peut écrire $\sum_k\int_k^{k+1}|f'| = \sum_k\int_{[a,\infty[}\mathbf{1}_{[k,k+1[}|f'| = \int_{[a,\infty[}\sum_k\mathbf{1}_{[k,k+1[}|f'| = \int_{[a,\infty[}|f'|$.

Ainsi, la série $\sum v_k$ est (absolument) convergente ; autrement dit la *suite* (u_n) est convergente. □

Exercice. Déterminer la nature des séries $\sum_{k\geqslant 1}\frac{\cos(\log(k))}{k}$ et $\sum_{k\geqslant 1}\frac{\sin\sqrt{k}}{k}$.

7. Une formule de changement de variable "générale"

La formule de changement de variable $\int_a^b f(\phi(t))\,\phi'(t)\,dt=\int_{\phi(a)}^{\phi(b)}f(x)\,dx$ a été établie pour des fonctions f supposées *continues*. On va voir maintenant que cette restriction n'est pas nécessaire, si on demande au "changement de variable" ϕ d'en être vraiment un, i.e. d'être injectif.

Proposition 7.1. *Soit $\phi: I\to\mathbb{R}$ une fonction de classe $\mathcal{C}^1$ strictement monotone sur un intervalle $I=(a,b)$. On pose $\phi(a):=\lim_{t\to a^+}\phi(t)$ et $\phi(b):=\lim_{t\to b^-}\phi(t)$.*

(1) *Pour toute fonction* borélienne positive *f sur $I':=\phi(I)$, on a*

$$\int_{\phi(a)}^{\phi(b)}f(x)\,dx=\int_a^b f(\phi(t))\,\phi'(t)\,dt\,. \tag{7.1}$$

(2) *Si $f: I'\to\mathbb{C}$ est borélienne, alors f est intégrable sur $I'=\phi(I)$ si et seulement si la fonction $(f\circ\phi)\,\phi'$ est intégrable sur I, et dans ce cas la formule (7.1) est encore valable.*

Démonstration. Comme d'habitude, le cas "intégrable" se déduit du cas "positif" ; donc on va se contenter de démontrer (1).

Comme ϕ est continue et strictement monotone, on sait que I' est un intervalle et que ϕ est une bijection de I sur I'. On peut bien entendu supposer ϕ croissante, de sorte que $I'=(\phi(a),\phi(b))$ et $\phi'\geqslant 0$. On commence par se simplifier la vie grâce au fait suivant.

Fait 1. Il suffit de démontrer (7.1) dans le cas où la fonction f est une *fonction indicatrice.*

Preuve du Fait 1. Si la formule est vraie pour toutes les fonctions indicatrices, elle est vraie pour toutes les fonctions *étagées* positives, par "linéarité" des 2 membres de (7.1) par rapport à f. Si maintenant f est une fonctions borélienne positive quelconque, alors f est limite d'une suite croissante de fonctions étagées positives (f_n) ; et comme la formule est vraie pour chaque f_n, elle l'est également pour f par convergence monotone. □

On s’est donc ramené a démontrer que (7.1) est vraie pour toute fonction f de la forme $\mathbf{1}_A$, où A est un borélien de $I' = (\phi(a), \phi(b))$; autrement dit, qu’on a

$$\int_{\phi(a)}^{\phi(b)} \mathbf{1}_A(x)\, dx = \int_a^b \mathbf{1}_A(\phi(t))\, \phi'(t)\, dt\, .$$

Comme en fait $\int_{\phi(a)}^{\phi(b)} \mathbf{1}_A(x)\, dx$ est une manière compliquée d’écrire $\lambda_1(A)$ pour $A \subseteq I' = (\phi(a), \phi(b))$, il s’agit donc de montrer qu’on a

$$\int_a^b \mathbf{1}_A(\phi(t))\, \phi'(t)\, dt = \lambda_1(A) \qquad \text{pour tout borélien } A \subseteq I'\, . \tag{7.2}$$

Soit $\mu : \mathcal{B}(I') \to [0, \infty]$ la fonction d’ensembles définie par

$$\mu(A) = \int_a^b \mathbf{1}_A(\phi(t))\, \phi'(t)\, dt\, .$$

Fait 2. μ est une mesure.

Preuve du Fait 2. On a évidemment $\mu(\varnothing) = 0$. Pour l’additivité dénombrable, soit $(A_k)_{k\in\mathbb{N}}$ une suite de boréliens de I' deux à deux disjoints. On a

$$\mathbf{1}_{\bigcup_0^\infty A_k} = \sum_{k=0}^{\infty} \mathbf{1}_{A_k}\, ,$$

et on en déduit

$$\begin{aligned} \mu\left(\bigcup_{k=0}^{\infty} A_k\right) &= \int_a^b \left(\sum_{k=0}^{\infty} \mathbf{1}_{A_k}(\phi(t))\phi'(t)\right) dt \\ &= \sum_{k=0}^{\infty} \int_a^b \mathbf{1}_{A_k}(\phi(t))\phi'(t)\, dt \\ &= \sum_{k=0}^{\infty} \mu(A_k)\, . \end{aligned}$$

À la deuxième ligne, on a appliqué le théorème d’interversion série-intégrale (Corollaire 2.7 du Chapitre 4) aux fonctions $u_k(t) := \mathbf{1}_{A_k}(\phi(t))\phi'(t)$ qui sont positives car $\phi' \geqslant 0$. □

Pour achever la preuve de la proposition, il s’agit de montrer que $\mu(A) = \lambda_1(A)$ pour tout borélien $A \subseteq I'$; et pour cela, il suffit de vérifier qu’on a $\mu(J) = |J|$ pour tout *intervalle compact* $J \subseteq I'$, d’après le lemme d’unicité 5.1 du Chapitre 2.

Soit $J = [c, d] \subseteq I' = (\phi(a), \phi(b))$ un intervalle compact quelconque. Comme ϕ est une bijection de I sur I', on peut écrire

$$\begin{aligned}
\mu(J) &= \int_a^b \mathbf{1}_{[c,d]}(\phi(t))\, \phi'(t)\, dt \\
&= \int_a^b \mathbf{1}_{[\phi^{-1}(c),\phi^{-1}(d)]}(t)\, \phi'(t)\, dt \\
&= \int_{\phi^{-1}(c)}^{\phi^{-1}(d)} \phi'(t)\, dt \\
&= \phi(\phi^{-1}(d)) - \phi(\phi^{-1}(c)) \\
&= d - c = |J|\,.
\end{aligned}$$

Donc la preuve est terminée. □

Exercice 1. Soit $I \subseteq \mathbb{R}$ un intervalle de milieu m. Pour $x \in I$, on note x^* le symétrique de x par rapport à m. Montrer que si $f : I \to \mathbb{C}$ est une fonction intégrable “impaire”, i.e. vérifiant $f(x^*) = -f(x)$ pour tout $x \in I$, alors $\int_I f(x)\, dx = 0$.

Exercice 2. Soit $f : \mathbb{R} \to \mathbb{C}$ une fonction localement intégrable et T-périodique, pour un certain $T > 0$. Montrer que la valeur de l'intégrale $\int_I f(t)\, dt$ est la même pour tous les intervalles I vérifiant $|I| = T$.

Chapitre 6

Théorèmes de convergence et applications

1. Les deux résultats principaux

1.1. Convergence “monotone”. Dans cette section $(\Omega, \mathcal{B}, \mu)$ est un espace mesuré. Le théorème suivant a été démontré au chapitre 4.

Théorème 1.1. (théorème de convergence monotone)
Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables positives, $f_n : (\Omega, \mathcal{B}) \to [0, \infty]$. On suppose que la suite (f_n) est croissante, *i.e. $f_n(x) \leqslant f_{n+1}(x)$ pour tout n et pour tout $x \in \Omega$. Alors, en posant $f(x) := \lim_{n\to\infty} f_n(x)$ (qui existe dans $[0, \infty]$), la fonction $f : \Omega \to [0, \infty]$ est mesurable et on a*

$$\int_\Omega f\, d\mu = \lim_{n\to\infty} \int_\Omega f_n\, d\mu\,.$$

Le corollaire suivant est particulièrement agréable. Il signifie qu'*on a toujours le droit* de permuter une somme infinie Σ et une intégrale $\int$ si on a affaire à des fonctions *positives*. Autrement dit : pour des séries de fonctions positives, il n'y a aucun état d'âme à avoir : tout est autorisé.

Corollaire 1.2. (interversion série-intégrale, cas “positif”)
Si $(u_k)_{k\in\mathbb{N}}$ est une suite quelconque de fonctions mesurables positives, alors

$$\int_\Omega \left(\sum_{k=0}^\infty u_k\right) d\mu = \sum_{k=0}^\infty \int_\Omega u_k\, d\mu\,.$$

Démonstration. On applique le théorème aux fonctions $f_n := \sum_{k=0}^n u_k$. $\square$

Exemple. Comme application du théorème d'interversion série-intégrale, on va établir l'incroyable formule suivante :

$$\int_0^1 x^{-x} dx = \sum_{n=1}^\infty n^{-n}\,.$$

Notons d'abord que les deux membres de la formule à obtenir ont un sens : le premier est l'intégrale d'une fonction borélienne positive, et le deuxième est la somme d'une famille de nombres positifs. (En réalité, pour la somme c'est franchement de la mauvaise foi : la série $\sum n^{-n}$ est manifestement convergente car $n^{-n} = o(2^{-n})$ quand $n \to \infty$.)

Le point de départ est d'écrire $x^{-x} = e^{-x\log(x)}$ et de développer l'exponentielle en série entière ; ce qui donne

$$x^{-x} = \sum_{k=0}^\infty \frac{\big(-x\log(x)\big)^k}{k!}\,.$$

Comme $-x\log(x) \geqslant 0$ sur $]0,1[$, on peut appliquer le théorème d'interversion série-intégrale aux fonctions $u_k(x) := \frac{(-x\log(x))^k}{k!}$. On obtient ainsi

$$\int_0^1 x^{-x} dx = \sum_{k=0}^{\infty} \int_0^1 \frac{\big(-x\log(x)\big)^k}{k!}\, dx := \sum_{k=0}^{\infty} J_k\,.$$

Il reste alors à calculer les intégrales J_k. Pour y voir plus clair, on effectue le changement de variable $u := -\log(x)$, *i.e.* $x = e^{-u}$. On a $dx = -e^{-u}du$, et donc

$$J_k = \frac{1}{k!}\int_0^\infty u^k e^{-(k+1)u} du\,.$$

Pour tout $\lambda > 0$, l'intégrale $\int_0^\infty u^k e^{-\lambda u} du$ se calcule en intégrant k fois par parties : elle vaut $\frac{k!}{\lambda^{k+1}}\cdot$ Donc $J_k = \frac{1}{(k+1)^{k+1}} = (k+1)^{-(k+1)}$, et donc

$$\int_0^1 x^{-x} dx = \sum_{k=0}^{\infty} (k+1)^{-(k+1)} = \sum_{n=1}^{\infty} n^{-n}.$$

Exercice. Montrer qu'on a $\displaystyle\int_0^1 \frac{\log(x)}{x-1}\, dx = \sum_{n=1}^{\infty} \frac{1}{n^2}\cdot$

Le théorème de convergence monotone a été énoncé pour une *suite* de fonctions. Cependant, on peut aussi en donner une version "continue" :

Corollaire 1.3. (convergence monotone, version continue)
Soit $(f_\lambda)_{\lambda\in\Lambda}$ une famille de fonctions mesurables positives indexée par un intervalle $\Lambda \subseteq \mathbb{R}$, et soit λ_0 une des bornes de l'intervalle Λ. On suppose que $f_\lambda(x)$ croit quand λ croit ou décroit vers λ_0 (selon que λ_0 est la borne de droite ou de gauche de l'intervalle Λ). Alors on a le droit d'écrire

$$\lim_{\lambda\to\lambda_0} \int_\Omega f_\lambda(x)\, d\mu(x) = \int_\Omega \left(\lim_{\lambda\to\lambda_0} f_\lambda(x)\right) d\mu(x)\,.$$

Démonstration. Supposons par exemple que λ_0 soit la borne de gauche de l'intervalle Λ. La condition de monotonie entraine que $f(x) := \lim_{\lambda\to\lambda_0} f_\lambda(x)$ est bien défini pour tout $x \in \Omega$, et que $\lim_{\lambda\to\lambda_0}\int_\Omega f_\lambda$ existe dans $[0,\infty]$. De plus, ces deux limites s'obtiennent en faisant tendre λ vers 0 le long de n'importe quelle suite décroissante $(\lambda_n)_{n\geqslant 1}$ tendant vers λ_0. Il suffit donc d'appliquer le théorème de convergence monotone à la *suite* de fonctions (f_{λ_n}). □

Exemple. En appliquant cette version du théorème de convergence à $f_\lambda(t) := \frac{e^{-\lambda t}}{t}$, on voit qu'on a

$$\lim_{\lambda\to 0^+} \int_1^\infty \frac{e^{-\lambda t}}{t} = \int_1^\infty \frac{dt}{t} = \infty\,.$$

Voici enfin une autre conséquence très utile du théorème de convergence monotone, qui servira de manière essentielle dans la preuve du théorème de convergence dominée (voir la section suivante).

Corollaire 1.4. (lemme de Fatou)
Si (h_n) est une suite quelconque de fonctions mesurables positives sur Ω, alors

$$\int_\Omega \left(\underline{\lim}\, h_n\right) d\mu \leqslant \underline{\lim} \int_\Omega h_n\, d\mu\,.$$

Démonstration. Par définition,

$$\underline{\lim}\, h_n = \sup_{n\in\mathbb{N}} \inf_{m\geqslant n} h_m = \lim_{n\to\infty} \inf_{m\geqslant n} h_m := \lim_{n\to\infty} f_n\,.$$

La suite (f_n) est visiblement croissante, donc on a par convergence monotone

$$\begin{aligned}
\int_\Omega \left(\underline{\lim}\, h_n\right) d\mu &= \lim_{n\to\infty} \int_\Omega f_n\, d\mu \\
&= \sup_{n\in\mathbb{N}} \int_\Omega f_n\, d\mu \\
&\leqslant \sup_{n\in\mathbb{N}} \inf_{m\geqslant n} \int_\Omega h_m\, d\mu \qquad \text{car } f_n \leqslant h_m \text{ pour tout } m \geqslant n \\
&\leqslant \underline{\lim} \int_\Omega h_m\, d\mu\,.
\end{aligned}$$

□

Exercice. Soit (f_n) une suite de fonctions intégrables par rapport à μ. On suppose que (f_n) converge simplement vers une fonction $f : \Omega \to \mathbb{C}$, et qu'il existe une constante $M < \infty$ telle $\int_\Omega |f_n|\, d\mu \leqslant M$ pour tout $n \in \mathbb{N}$. Montrer que f est intégrable par rapport à μ.

1.2. Convergence "dominée". Dans cette section, $(\Omega, \mathcal{B}, \mu)$ est toujours un espace mesuré. Pour alléger les énoncés, on convient que le mot "intégrable" signifiera toujours "μ-intégrable" au sens de la définition 3.12 du Chapitre 4, et que "presque partout" signifiera "μ-presque partout".

Le théorème suivant est un résultat de "passage à la limite sous le signe intégrale" extraordinairement efficace.

Théorème 1.5. (théorème de convergence dominée)
Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions intégrables, $f_n : \Omega \to \mathbb{C}$, et soit $f : \Omega \to \mathbb{C}$ une fonction presque partout définie. On fait les hypothèses suivantes.

(i) $f_n(x) \to f(x)$ *presque partout.*

(ii) ***Hypothèse de domination** : il existe une fonction mesurable $g : \Omega \to \mathbb{R}^+$,* indépendante de n, *telle que*

$$\int_\Omega g\, d\mu < \infty \qquad \text{et} \qquad \forall n \in \mathbb{N}\ :\ |f_n(x)| \leqslant g(t) \quad \textit{presque partout}.$$

Alors la fonction f est intégrable, et $\int_\Omega |f_n - f|\, d\mu \to 0$ quand $n \to \infty$. En particulier :

$$\int_\Omega f\, d\mu = \lim_{n\to\infty} \int_\Omega f_n\, d\mu\,.$$

Démonstration. La deuxième partie de la conclusion découle de la première car $\left|\int_\Omega f_n - \int f\right| \leqslant \int_\Omega |f_n - f|$. On se concentre donc sur cette première partie.

Cas 1. On suppose que les f_n sont $\mathcal{B}$-mesurables, convergent partout vers f, et qu'on a $|f_n| \leqslant g$ partout.

Dans ce cas, f est partout définie et $\mathcal{B}$-mesurable. De plus, on a par passage à la limite $|f(x)| \leqslant g(x)$ pour tout $x \in \Omega$; donc f est intégrable sur Ω.

L'idée est d'appliquer le *Lemme de Fatou* aux fonctions h_n définies par

$$h_n(x) := 2g(x) - |f_n - f(x)|\,.$$

Les h_n sont mesurables car les f_n le sont, et elles sont bien *positives* car $|f_n - f| \leqslant |f_n| + |f| \leqslant 2g$. De plus $h_n(x) \to 2g(x)$ pour tout $x \in \Omega$, et donc $\underline{\lim}\, h_n = 2g$. Par Fatou, on a donc

$$\begin{aligned}
\int_\Omega 2g\, d\mu &\leqslant \underline{\lim} \int_\Omega h_n\, d\mu \\
&= \underline{\lim} \left[\int_\Omega 2g\, d\mu - \int_\Omega |f_n - f|\, d\mu \right] \\
&= \int_\Omega 2g\, d\mu + \underline{\lim} \left(- \int_\Omega |f_n - f|\, d\mu \right) \\
&= \int_\Omega 2g\, d\mu - \overline{\lim} \int_\Omega |f_n - f|\, d\mu\, .
\end{aligned}$$

Comme $\int_\Omega 2g\, d\mu < \infty$, on peut simplifier par $\int_\Omega 2g\, d\mu$, et on obtient ainsi

$$\overline{\lim} \int_\Omega |f_n - f|\, d\mu \leqslant 0\, ,$$

autrement dit $\lim_{n\to\infty} \int_\Omega |f_n - f|\, d\mu = 0$.

Cas 2. Cas général.

On va se ramener au Cas 1 en jouant avec le “presque partout”. Ce n'est pas particulièrement palpitant, mais il faut quand même le faire.

Pour tout $n \in \mathbb{N}$, on peut trouver une fonction $\mathcal{B}$-mesurable $\widetilde{f}_n$ intégrable par rapport à μ telle que $f_n = \widetilde{f}_n$ presque partout. Considérons alors la propriété $P(x)$ (dépendant de $x \in \Omega$) définie comme suit :

$$\begin{aligned}
P(x) \text{ est vraie} \quad \Longleftrightarrow \quad & \forall n \in \mathbb{N}\ :\ f_n(x) \text{ est bien défini et } f_n(x) = \widetilde{f}_n(x) \\
& \text{et } \forall n \in \mathbb{N}\ :\ |f_n(x)| \leqslant g(x) \\
& \text{et } f(x) \text{ est bien défini} \\
& \text{et } f_n(x) \to f(x) \quad \text{quand } n \to \infty.
\end{aligned}$$

La propriété (P) est la conjonction d'une famille dénombrable de propriétés vraies presque partout ; donc $P(x)$ est vraie presque partout. Soit $\widetilde{\Omega}$ un ensemble mesurable tel que $\mu(\Omega\backslash\widetilde{\Omega}) = 0$ et $P(x)$ est vraie pour tout $x \in \widetilde{\Omega}$. Par définition de $\widetilde{\Omega}$, les fonctions $(f_n)_{|\widetilde{\Omega}} : \widetilde{\Omega} \to \mathbb{C}$ sont $\mathcal{B}_{\widetilde{\Omega}}$-mesurables (car égales aux $(\widetilde{f}_n)_{|\widetilde{\Omega}}$), partout dominées en module par $g_{|\widetilde{\Omega}}$ (qui est $\mathcal{B}_{\widetilde{\Omega}}$-mesurable et intégrable sur $\widetilde{\Omega}$), et convergent vers $f_{|\widetilde{\Omega}}$ en tout point de $\widetilde{\Omega}$. Par le Cas 1, on en déduit que $f_{|\widetilde{\Omega}}$ est $\mathcal{B}_{\widetilde{\Omega}}$-mesurable, intégrable sur $\widetilde{\Omega}$, et que $\int_{\widetilde{\Omega}} |f_n - f|\, d\mu \to 0$. Si on définit $\widetilde{f} : \Omega \to \mathbb{C}$ par $\widetilde{f}(x) = f(x)$ sur $\widetilde{\Omega}$ et $\widetilde{f}(x) = 0$ sur $\Omega\backslash\widetilde{\Omega}$, alors $\widetilde{f}$ est $\mathcal{B}$-mesurable par recollement, intégrable sur Ω car $\int_\Omega |\widetilde{f}|\, d\mu = \int_{\widetilde{\Omega}} |f|\, d\mu < \infty$, et $\widetilde{f} = f$ presque partout puisque $\mu(\Omega\backslash\widetilde{\Omega}) = 0$; donc la fonction f est μ-intégrable. Enfin, comme $\mu(\Omega\backslash\widetilde{\Omega}) = 0$, on a $\int_\Omega |f_n - f|\, d\mu = \int_{\widetilde{\Omega}} |f_n - f|\, d\mu$ et donc $\int_\Omega |f_n - f|\, d\mu \to 0$. □

La démonstration qu'on vient de donner, basée sur le lemme de Fatou, est devenue la “preuve standard” du théorème de convergence dominée ; et c'est certainement la plus courte possible. Il est cependant intéressant d'en donner une autre, qui par certains aspects est plus naturelle.

Deuxième preuve du théorème de convergence dominée. Il suffit de traiter le cas où les f_n sont mesurables, convergent vers f partout, et $|f_n| \leqslant g$ partout *cf* plus haut).

Cas 1. On suppose que $\mu(\Omega) < \infty$ et que la suite (f_n) est *uniformément bornée.*

Dans ce cas, tout repose sur le fait suivant.

Fait. La suite (f_n) converge **en mesure** vers f : pour tout $\varepsilon > 0$, on a

$$\lim_{n\to\infty} \mu\Big(\{|f_n - f| \geqslant \varepsilon\}\Big) = 0\,.$$

Preuve du Fait. Soit $\varepsilon > 0$. Pour $n \in \mathbb{N}$, posons

$$A_{\varepsilon,n} := \{x \in \Omega;\ \exists k \geqslant n\ :\ |f_k(x) - f(x)| \geqslant \varepsilon\}\,.$$

Les $A_{\varepsilon,n}$ sont mesurables car les f_n le sont, la suite $(A_{\varepsilon,n})_{n\in\mathbb{N}}$ est visiblement *décroissante*, et on a $\bigcap_n A_{\varepsilon,n} = \varnothing$ car $f_n(x) \to f(x)$ pour tout $x \in \Omega$. Comme on suppose que $\mu(\Omega) < \infty$, on a donc $\lim_{n\to\infty} \mu(A_{\varepsilon,n}) = 0$; d'où le résultat puisque $\{|f_n - f| \geqslant \varepsilon\} \subseteq A_{\varepsilon,n}$. □

Choisissons maintenant une constante M telle que $|f_n(x)| \leqslant M$ pour tout n et pour tout $x \in \Omega$.

Pour $\varepsilon > 0$ donné, on a

$$\begin{aligned}
\int_\Omega |f_n - f|\,d\mu &= \int_{\{|f_n - f| < \varepsilon\}} |f_n - f|\,d\mu + \int_{\{|f_n - f| \geqslant \varepsilon\}} |f_n - f|\,d\mu \\
&\leqslant \varepsilon\,\mu\Big(\{|f_n - f| < \varepsilon\}\Big) + \int_{\{|f_n - f| \geqslant \varepsilon\}} 2M\,d\mu \\
&\leqslant \varepsilon\,\mu(\Omega) + 2M\,\mu\Big(\{|f_n - f| \geqslant \varepsilon\}\Big)\,.
\end{aligned}$$

Par le Fait, on en déduit

$$\overline{\lim} \int_\Omega |f_n - f|\,d\mu \leqslant \varepsilon\,\mu(\Omega) \quad \text{pour tout } \varepsilon > 0;$$

et donc $\lim_{n\to\infty} \int_\Omega |f_n - f|\,d\mu = 0$.

Cas 2. Cas général.

Comme $|f_n - f| = 0$ sur l'ensemble $\{g = 0\}$ (car $|f_n - f| \leqslant 2g$), on peut supposer, quitte à remplacer Ω par $\Omega_g := \{g > 0\}$, qu'on a

$$g(x) > 0 \quad \text{sur } \Omega\,.$$

On va se ramener au Cas 1 *en changeant de mesure.* Soit $\mu_g : \mathcal{B} \to [0, \infty]$ la fonction d'ensembles définie par

$$\mu_g(A) = \int_A g\,d\mu\,.$$

On sait que μ_g est une mesure sur $(\Omega, \mathcal{B})$ (c'est l'Exercice 2.8 du Chapitre 4) ; et de plus, μ_g est une mesure *finie* car g est intégrable sur Ω. On a besoin du fait suivant.

Fait. Pour toute fonction mesurable positive h sur Ω, on a

$$\int_\Omega h\,d\mu_g = \int_\Omega hg\,d\mu\,.$$

Preuve du Fait. Pour une fonction h *étagée*, c'est immédiat par définition de μ_g et par "linéarité" de l'intégrale pour les fonctions positives ; et le cas d'une fonction mesurable $h \geqslant 0$ quelconque s'en déduit par convergence monotone, en considérant une suite approximante pour h. □

Si on pose maintenant

$$\widetilde{f}_n := \frac{f_n}{g} \quad \text{et} \quad \widetilde{f} := \frac{f}{g}\,,$$

alors $\widetilde{f}_n \to \widetilde{f}$ et $|\widetilde{f}_n| \leqslant \widetilde{g} := \mathbf{1}$ sur Ω. D'après le Cas 1 appliqué *avec la mesure* μ_g, on a donc

$$\lim_{n\to\infty} \int_\Omega |\widetilde{f}_n - \widetilde{f}|\, d\mu_g = 0\,;$$

autrement dit, par le Fait :

$$\lim_{n\to\infty} \int_\Omega |f_n - f|\, d\mu = 0\,.$$

□

Remarque. Il est facile de formuler une version "continue" du théorème de convergence dominée, concernant une famille de fonctions intégrables $(f_\lambda)_{\lambda\in\Lambda}$ paramétrée par un intervalle $\Lambda \subseteq \mathbb{R}$ et où on fait tendre le paramètre λ vers une des bornes λ_0 de Λ. De façon précise : *Si $f_\lambda(x) \to f(x)$ presque partout quand $\lambda \to \lambda_0$ et s'il existe une fonction intégrable g telle que, pour tout $\lambda \in \Lambda$, on ait $|f_\lambda(x)| \leqslant g(x)$ presque partout, alors f est intégrable et $\int_\Omega f\, d\mu = \lim_{\lambda\to\lambda_0} \int_\Omega f_\lambda$.*

Démonstration. Si on applique le théorème à la *suite* (f_{λ_n}), pour n'importe quelle suite $(\lambda_n)_{n\geqslant 1}$ tendant vers λ_0, on voit que f est intégrable et $\int_\Omega f = \lim \int_\Omega f_{\lambda_n}$. Comme ceci ne dépend pas de la suite (λ_n) tendant vers λ_0, on en déduit que $\int_\Omega f_\lambda \to \int_\Omega f$ quand $\lambda \to \lambda_0$. □

Corollaire 1.6. (théorème de convergence bornée)
Suposons que la mesure μ soit finie, *i.e. $\mu(\Omega) < \infty$. Si (f_n) est une suite de fonctions intégrables convergeant presque partout vers une fonction $f : \Omega \to \mathbb{C}$ et s'il existe une* constante *$M < \infty$* indépendante de *n telle que $\forall n \in \mathbb{N}\ :\ |f_n(x)| \leqslant M$ presque partout, alors f est intégrable et $\int_\Omega f\, d\mu = \lim_{n\to\infty} \int_\Omega f_n\, d\mu$.*

Démonstration. Le théorème de convergence dominée s'applique avec $g(x) :\equiv M$, qui est intégrable car $\mu(\Omega) < \infty$. □

Remarque. Comme cas *très particulier* du théorème de convergence bornée, on obtient une généralisation du théorème standard de passage à la limite sous le signe intégral sous une hypothèse de convergence *uniforme* : si (f_n) est une suite de fonctions μ-mesurables bornées sur un segment $[a, b]$ convergeant uniformément vers une fonction $f : [a, b] \to \mathbb{C}$, alors f est intégrable et $\int_a^b f_n \to \int_a^b f$. Mais bien entendu, ce résultat est très facile à démontrer directement (exercice).

Exemple. On sait qu'on a pour tout $x \in [0, 1[$:

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

Si on intègre formellement cette identité entre 0 et 1, on obtient

$$\log(2) = \int_0^1 \frac{dx}{1+x} = \sum_{k=0}^{\infty} (-1)^k \int_0^1 x^k dx = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1}\,.$$

Ce calcul formel peut se justifier à l'aide du théorème de convergence dominée. Pour $n \in \mathbb{N}$, soit $f_n : [0,1[\to \mathbb{R}$ la fonction définie par

$$f_n(x) := \sum_{k=0}^{n} (-1)^k x^k\,.$$

Alors $f_n(x) \to \frac{1}{1+x}$ pour tout $x \in [0,1[$, *avec de plus*

$$0 \leqslant f_n(x) \leqslant 1 \qquad \text{pour tout } n \in \mathbb{N}\,.$$

En effet, on a $f_{2m}(x) = f_{2m-1}(x)+x^{2m} \geqslant f_{2m-1}(x)$ pour tout $m \geqslant 1$; et en écrivant $f_{2m}(x) = 1-(x-x^2)-\cdots-(x^{2m-1}-x^{2m})$ et $f_{2m-1}(x) = (1-x)+\cdots+(x^{2m-2}-x^{2m-1})$, on voit qu'on a $0 \leqslant f_{2m-1} \leqslant f_{2m}(x) \leqslant 1$.

Donc le théorème de convergence bornée s'applique (puisqu'on est sur un intervalle de longueur finie), et on obtient ainsi

$$\int_0^1 \frac{dx}{1+x} = \lim_{n\to\infty} \int_0^1 f_n(x)\,dx\,,$$

ce qui est le résultat souhaité.

Il est important de noter que $f_n(x)$ *ne tend pas uniformément* vers $\frac{1}{1+x}$ sur $[0,1[$ quand $n \to \infty$; donc on ne peut pas s'en sortir en appliquant un théorème rudimentaire à base de convergence uniforme.

Remarque. En réalité, on peut obtenir la formule $\sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} = \log(2)$ sans appliquer aucun théorème (!) Il suffit pour cela d'écrire que pour tout $n \in \mathbb{N}$ et pour $x \in [0,1[$, on a

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

les deux extrêmes étant d'ailleurs égaux pour $x = 1$ aussi. En intégrant entre 0 et 1, on en déduit

$$\log(2) = \sum_{k=0}^{n} \frac{(-1)^k}{k+1} + (-1)^{n+1} \int_0^1 \frac{x^{n+1}}{1+x}\,dx\,.$$

L'intégrale apparaissant au membre de droite tend vers 0 quand $n \to \infty$, car elle se majore en valeur absolue par $\int_0^1 x^{n+1} dx = \frac{1}{n+2}$; d'où le résultat.

Exercice. Montrer que pour tout nombre réel $\theta \notin 2\pi\mathbb{Z}$, on a

$$\int_0^1 \frac{e^{i\theta}}{1-e^{i\theta}x}\,dx = \sum_{k=1}^{\infty} \frac{e^{ik\theta}}{k}\,.$$

En déduire la formule

$$\sum_{k=1}^{\infty} \frac{\cos(k\theta)}{k} = -\log\bigl|2\sin\bigl(\theta/2\bigr)\bigr|\,.$$

Corollaire 1.7. (interversion série-intégrale, cas "sommable")
Soit (u_k) *une suite de fonctions mesurables,* $u_k : \Omega \to \mathbb{C}$*. On suppose qu'on a*

$$\sum_{k=0}^{\infty} \int_{\Omega} |u_k|\, d\mu < \infty\,.$$

Alors la série $\sum u_k(x)$ *est presque partout absolument convergente, la fonction (presque partout définie)* $\sum_0^\infty u_k$ *est intégrable, et on a*

$$\int_{\Omega} \left(\sum_{k=0}^{\infty} u_k \right) d\mu = \sum_{k=0}^{\infty} \int_{\Omega} u_k\, d\mu\,.$$

Démonstration. Soit $g : \Omega \to [0, \infty]$ la fonction (mesurable) définie par

$$g(x) := \sum_{k=0}^{\infty} |u_k(x)|\,.$$

Par interversion série-intégrale (cas "positif"), on a $\int_\Omega g\, d\mu = \sum_0^\infty \int_\Omega |u_k|\, d\mu < \infty$. Donc $g(x) < \infty$ presque partout ; autrement dit la série $\sum u_k(x)$ est presque partout absolument convergente. Ainsi, la fonction $f := \sum_0^\infty u_k$ est bien définie presque partout.

Si on pose $f_n(x) := \sum_{k=0}^n u_k(x)$, alors $f_n(x) \to f(x)$ presque partout, et $|f_n(x)| \leqslant \sum_{k=0}^n |u_k(x)| \leqslant \sum_0^\infty |u_k(x)| = g(x)$ pour tout n et pour tout $x \in \Omega$. Comme $\int_\Omega g\, d\mu < \infty$, on peut donc appliquer le théorème de convergence dominée : la fonction f est intégrable et $\int_\Omega \widetilde{f} = \lim_{n\to\infty} \int_\Omega f_n$, autrement dit

$$\int_{\Omega} \left(\sum_{k=0}^{\infty} u_k \right) = \lim_{n\to\infty} \sum_{k=0}^{n} \int_{\Omega} f_n = \sum_{k=0}^{\infty} \int_{\Omega} u_k\,.$$

□

Remarque. Dans les calculs concrets, on sait en général dès le début que la série $\sum u_k(x)$ converge partout (ou presque partout). Le sens du corollaire est le suivant : *on a le droit* d'intervertir $\sum$ et $\int$ *à condition d'avoir vérifié* qu'on a $\sum_0^\infty \int_\Omega |u_k| < \infty$; ou, ce qui revient au même par interversion série-intégrale dans le cas "positif", qu'on a $\int_\Omega \sum_0^\infty |u_k| < \infty$, .

Exemple. Comme application du théorème d'interversion série-intégrale, on va établir la formule suivante :

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)(2k+2)} = \frac{\pi}{4} - \frac{\log(2)}{2}\cdot$$

Pour $k \in \mathbb{N}$, soit $u_k : [0, 1[\to \mathbb{R}$ la fonction définie par

$$u_k(x) = (-1)^k x^{2k} (1 - x)\,.$$

Pour tout $x \in [0, 1[$, la série $\sum u_k(x)$ est absolument convergente car $|u_k(x)| \leqslant x^{2k}$, et on a

$$\sum_{k=0}^{\infty} u_k(x) = (1 - x) \sum_{k=0}^{\infty} (-x^2)^k = \frac{1 - x}{1 + x^2}\cdot$$

De plus, on a

$$\begin{aligned}\sum_{k=0}^{\infty}\int_0^1 |u_k(x)|\,dx &= \sum_{k=0}^{\infty}\int_0^1 (1-x)x^{2k}\,dx\\ &= \sum_{k=0}^{\infty}\left(\frac{1}{2k+1}-\frac{1}{2k+2}\right)=\sum_{k=0}^{\infty}\frac{1}{(2k+1)(2k+2)}<\infty\,.\end{aligned}$$

Par interversion série-intégrale, on peut donc écrire $\int_0^1\sum_0^\infty u_k=\sum_0^\infty\int_0^1 u_k$, autrement dit

$$\begin{aligned}\int_0^1\frac{1-x}{1+x^2}\,dx &= \sum_{k=0}^{\infty}(-1)^k\int_0^1 x^{2k}(1-x)\,dx\\ &= \sum_{k=0}^{\infty}(-1)^k\left(\frac{1}{2k+1}-\frac{1}{2k+2}\right)\\ &= \sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)(2k+2)}\,.\end{aligned}$$

Pour conclure, il suffit donc de vérifier que l'intégrale figurant au membre de gauche vaut $\frac{\pi}{4}-\frac{\log(2)}{2}$; ce qui est facile :

$$\begin{aligned}\int_0^1\frac{1-x}{1+x^2}\,dx &= \int_0^1\frac{dx}{1+x^2}-\int_0^1\frac{x}{1+x^2}\,dx\\ &= \big[\arctan(x)\big]_0^1-\left[\frac12\log(1+x^2)\right]_0^1\\ &= \frac{\pi}{4}-\frac{\log(2)}{2}\cdot\end{aligned}$$

Remarque. Le Corollaire 1.7 est assez "brutal" : il peut arriver que l'on ait le droit d'intervertir une somme et une intégrale sans que la condition de "sommabilité" $\sum_0^\infty\int_\Omega|u_k|<\infty$ soit satisfaite. C'est par exemple le cas avec $u_k(x)=(-1)^kx^k$ sur $[0,1[$, où la condition de sommabilité n'est pas satisfaite car $\sum_0^\infty\int_0^1|u_k(x)|\,dx=\sum_0^\infty\frac{1}{k+1}=\infty$, mais où on a quand même le droit d'intervertir la somme et l'intégrale grâce au théorème de convergence dominée (*cf* plus haut).

Exercice. Soit f une fonction intégrable sur $\mathbb{R}$. Montrer que pour presque tout $x\in\mathbb{R}$, la série $\sum_{k\in\mathbb{Z}}f(x+k)$ est absolument convergente, et que la fonction presque partout définie $f_{\rm per}(x):=\sum_{-\infty}^{\infty}f(x+k)$ est 1-périodique et intégrable sur $[0,1[$.

Comme cas particulier du théorème d'interversion série-intégrale, on obtient un résultat très utile concernant les séries.

Corollaire 1.8. (convergence dominée pour les séries)
Soit (f_n) une suite de fonctions définies sur $\mathbb{N}$, et soit $f:\mathbb{N}\to\mathbb{C}$. On fait les hypothèses suivantes :

(i) *$f_n(i)\to f(i)$ pour tout $i\in\mathbb{N}$;*
(ii) *il existe une fonction $g:\mathbb{N}\to\mathbb{R}^+$ telle que $\sum_0^\infty g(i)<\infty$ et $|f_n(i)|\leqslant g(i)$ pour tout n et pour tout $i\in\mathbb{N}$.*

Alors la série $\sum f(i)$ *est absolument convergente et*

$$\sum_{i=0}^{\infty} f(i) = \lim_{n\to\infty} \sum_{i=0}^{\infty} f_n(i)\,.$$

Démonstration. On applique le théorème de convergence dominée à la mesure de comptage μ_c sur $\Omega := \mathbb{N}$. □

Exercice 1. Démontrer directement le théorème de convergence dominée pour les séries.

Exercice 2. En utilisant le formule du binôme et le théorème de convergence dominée pour les séries, montrer que pour tout $z \in \mathbb{C}$, on a

$$e^z = \lim_{n\to\infty} \left(1 + \frac{z}{n}\right)^n .$$

Voici enfin un dernier résultat sur les séries, très bien connu mais qu'on peut voir comme une application du théorème d'interversion série-intégrale. Rappelons que pour toute "suite double" de nombres positifs $(a_{k,l})_{(k,l)\in\mathbb{N}\times\mathbb{N}}$, on a

$$\sum_{k=0}^{\infty} \left(\sum_{l=0}^{\infty} a_{k,l}\right) = \sum_{(k,l)\in\mathbb{N}\times\mathbb{N}} a_{k,l} = \sum_{l=0}^{\infty} \left(\sum_{k=0}^{\infty} a_{k,l}\right) .$$

(Que les deux sommes doubles soient égales est une conséquence du théorème d'interversion série-somme démontré au Chapitre 1 ; et qu'elles soient égales à $\sum_{\mathbb{N}\times\mathbb{N}} a_{k,l}$ est un exercice du même chapitre.)

Corollaire 1.9. (séries doubles absolument convergentes)
Soit $(u_{k,l})_{(k,l)\in\mathbb{N}\times\mathbb{N}}$ *une "suite double" de nombres complexes. Si on a*

$$\sum_{(k,l)\in\mathbb{N}\times\mathbb{N}} |u_{k,l}| < \infty\,,$$

alors on a le droit d'écrire

$$\sum_{k=0}^{\infty} \left(\sum_{l=0}^{\infty} u_{k,l}\right) = \sum_{l=0}^{\infty} \left(\sum_{k=0}^{\infty} u_{k,l}\right) .$$

Démonstration. L'expression "on a le droit d'écrire" signifie que toutes les séries considérées convergent (en fait, convergent absolument, ce qui est clair d'après l'hypothèse) et que les deux sommes doubles sont égales. Cela étant précisé, le résultat est une conséquence immédiate du théorème d'interversion série-intégrale appliqué à la mesure de comptage μ_c sur $\Omega := \mathbb{N}$ et aux fonctions $u_k : \mathbb{N} \to \mathbb{C}$ définies par $u_k(l) := u_{k,l}$: le 1er membre de l'égalité à démontrer est égal à $\sum_0^\infty \int_{\mathbb{N}} u_k\, d\mu_c$, et le second membre à $\int_{\mathbb{N}} (\sum_0^\infty u_k)\, d\mu_c$. □

Exercice. Donner deux autres preuves du Corollaire 1.9 : en utilisant le théorème de convergence dominée pour les séries, ou bien directement à partir du cas "positif".

2. Caractérisation de l'intégrabilité au sens de Riemann

Dans cette section, on va utiliser le théorème de convergence dominée pour établir la caractérisation des fonctions intégrables au sens de Riemann mentionnée à la fin du Chapitre 5. (Le théorème de convergence dominée n'intervient en fait que dans l'une des deux implications du théorème suivant.)

THÉORÈME 2.1. *Une fonction* bornée $f : [a, b] \to \mathbb{R}$ *est intégrable au sens de Riemann si et seulement si l'ensemble de ses points de discontinuité est négligeable.*

Démonstration. Dans ce qui suit on notera $D(f)$ l'ensemble des points de discontinuité de f *appartenant à l'intervalle ouvert* $]a, b[$. Comme l'ensemble à deux éléments $\{a, b\}$ est négligeable, il suffit de montrer que f est Riemann-intégrable si et seulement si $D(f)$ est négligeable.

Pour tout intervalle non trivial $J \subseteq [a, b]$, on notera $\mathrm{osc}(f, J)$ l'**oscillation de** f **sur** J :

$$\mathrm{osc}(f, J) := \sup_J f - \inf_J f\,.$$

On aura besoin du fait suivant, dont la preuve est laissée en exercice.

FAIT 1. Pour un point $t \in [a, b]$, les propriétés suivantes sont équivalentes :

(i) t est un point *de continuité* de f ;

(ii) $\mathrm{osc}(f, J) \to 0$ quand $J \to \{t\}$, où "$J \to \{t\}$" est une abréviation signifiant "$J \ni t$ et $|J| \to 0$".

Si de plus $t \in]a, b[$, ces propriétés sont aussi équivalentes à

(iii) pour tout $\varepsilon > 0$, on peut trouver $J \ni t$ ouvert tel que $\mathrm{osc}(f, J) < \varepsilon$.

Après cette remarque préliminaire, on peut maintenant commencer la preuve du théorème.

(1) Supposons que la fonction f soit intégrable au sens de Riemann, et montrons que $D(f)$ est négligeable.

Pour $\varepsilon > 0$, notons $D_\varepsilon(f)$ l'ensemble des points $t \in]a, b[$ vérifiant la propriété suivante :

$$\mathrm{osc}(f, J) \geqslant \varepsilon \quad \text{pour tout intervalle ouvert } J \ni t\,.$$

Autrement dit : $]a, b[\setminus D_\varepsilon(f)$ est la réunion de tous les intervalles ouverts $J \subseteq]a, b[$ tels que $\mathrm{osc}(f, J) < \varepsilon$. En particulier, $]a, b[\setminus D_\varepsilon(f)$ est un ouvert de $\mathbb{R}$, ce qui prouve que $D_\varepsilon(f)$ est borélien.

Par le Fait 1, on a

$$D(f) = \bigcup_{\varepsilon > 0} D_\varepsilon(f) = \bigcup_{n \in \mathbb{N}^*} D_{1/n}(f)\,.$$

Pour conclure, il suffit donc de montrer que

$$\lambda_1(D_\varepsilon(f)) = 0 \quad \text{pour tout } \varepsilon > 0\,.$$

Tout repose sur le fait suivant.

FAIT 2. Soient φ et ψ des fonctions en escalier telles que $\varphi \leqslant f \leqslant \psi$. Soit également $a = s_0 < \cdots < s_N = b$ une subdivision de $[a, b]$ adaptée à φ et à ψ, et notons $J_0, \ldots, J_{N-1}$ les intervalles ouverts déterminés par cette subdivision. Pour $i = 0, \ldots, N-1$, on a

$$\int_{J_i} \psi - \int_{J_i} \varphi \geqslant \mathrm{osc}(f, J_i) \times |J_i|\,.$$

Preuve du Fait 2. Comme φ et ψ sont *constantes* sur J_i et comme $\varphi \leqslant f \leqslant \psi$, on a $\varphi \leqslant \inf_{J_i} f$ et $\psi \geqslant \sup_{J_i} f$ sur J_i. Donc $\psi - \phi \geqslant \sup_{J_i} f - \inf_{J_i} f = \mathrm{osc}(f, J_i)$ sur J_i, et donc

$$\int_{J_i} \psi - \int_{J_i} \varphi = \int_{J_i} (\psi - \phi) \geqslant \mathrm{osc}(f, J_i) \times |J_i|\,.$$

□

Fixons maintenant $\varepsilon > 0$. Pour $\eta > 0$ donné, on peut trouver deux fonctions en escalier φ et ψ telles que

$$\varphi \leqslant f \leqslant \psi \quad \text{et} \quad \int_a^b \psi - \int_a^b \varphi \leqslant \eta\,.$$

Soit $a = s_0 < \cdots < s_N = b$ une subdivision de $[a,b]$ adaptée à φ et à ψ, et soient $J_0, \dots, J_{N-1}$ les intervalles ouverts déterminés par cette subdivision. Notons Λ l'ensemble des indices $i \in \{0, \dots, N-1\}$ tels que $J_i \cap D_\varepsilon(f) \neq \varnothing$. Alors

$$D_\varepsilon(f) \subseteq \{s_1, \dots, s_{N-1}\} \cup \bigcup_{i\in\Lambda} J_i\,. \tag{2.1}$$

De plus, on a *par définition* de $D_\varepsilon(f)$:

$$\operatorname{osc}(f, J_i) \geqslant \varepsilon \quad \text{pour tout } i \in \Lambda\,.$$

Par le Fait 2, on en déduit

$$|J_i| \leqslant \frac{1}{\varepsilon}\int_{J_i} (\psi - \phi) \quad \text{pour tout } i \in \Lambda\,.$$

En revenant à (2.1) et comme $\lambda_1(\{s_1, \dots, s_{N-1}\}) = 0$, on obtient alors

$$\begin{aligned}
\lambda_1(D_\varepsilon(f)) &\leqslant \lambda_1\left(\bigcup_{i\in\Lambda} J_i\right)\\
&= \sum_{i\in\Lambda} |J_i|\\
&\leqslant \frac{1}{\varepsilon}\sum_{i\in\Lambda}\int_{J_i}(\psi-\varphi)\\
&\leqslant \frac{1}{\varepsilon}\sum_{i=0}^{N-1}\int_{J_i}(\psi-\varphi) \qquad \text{car } \psi - \varphi \geqslant 0\\
&= \frac{1}{\varepsilon}\int_a^b(\psi-\varphi) \leqslant \frac{\eta}{\varepsilon}\,.
\end{aligned}$$

Ceci étant vrai pour tout $\eta > 0$, on en déduit $\lambda_1(D_\varepsilon(f)) = 0$ en faisant tendre η vers 0.

(2) Supposons que l'ensemble des points de discontinuité de f soit négligeable, et montrons que f est Riemann-intégrable.

Pour $n \in \mathbb{N}^*$, découpons l'intervalle $[a, b[$ en n intervalles semi-ouverts $J_{1,n}, \dots ; J_{n,n}$ de longueur $h_n := \frac{b-a}{n}$; explicitement $J_{i,n} := [a + (i-1)h_n, a + ih_n[$ pour $i = 1, \dots, n$. Posons

$$M_{i,n} := \sup_{J_{i,n}} f \quad \text{et} \quad m_{i,n} := \inf_{J_{i,n}} f \qquad (1 \leqslant i \leqslant n)\,,$$

et définissons deux fonction en escalier φ_n et ψ_n comme suit :

$$\varphi_n(t) = \begin{cases} m_{1,n} & \text{si} \quad t \in J_{1,n} \\ m_{2,n} & \text{si} \quad t \in J_{2,n} \\ \vdots & \\ m_{n,n} & \text{si} \quad t \in J_{n,n} \\ f(b) & \text{si} \quad t = b \end{cases} \qquad \text{et} \qquad \psi_n(t) = \begin{cases} M_{1,n} & \text{si} \quad t \in J_{1,n} \\ M_{2,n} & \text{si} \quad t \in J_{2,n} \\ \vdots & \\ M_{n,n} & \text{si} \quad t \in J_{n,n} \\ f(b) & \text{si} \quad t = b \end{cases}$$

Autrement dit : $\varphi_n(b) = f(b) = \psi_n(b)$ et

$$\varphi_n(t) = m_{i,n} \quad \text{et} \quad \psi_n(t) = M_{i,n} \quad \text{pour } t \in J_{i,n}\ , \quad 1 \leqslant i \leqslant n\,.$$

On a par définition $\varphi_n \leqslant f \leqslant \psi_n$. Pour montrer que f est Rieman-intégrable, il suffit donc de montrer que $\int_a^b \psi_n - \int_a^b \varphi_n \to 0$ quand $n \to \infty$. Le point clé est le fait suivant.

Fait 3. $\varphi_n(t) \to f(t)$ et $\psi_n(t) \to f(t)$ en tout point de continuité f.

Preuve du Fait 3. Soit $t \in [a,b]$ un point de continuité de f. Si $t = b$, alors $\varphi_n(t) = f(t)$ pour tout n donc il n'y a rien à démontrer. Supposons donc que $t \in [a,b[$. Pour tout $n \in \mathbb{N}^*$, il existe alors un unique $i = i_n(t) \in \{1,\dots,n\}$ tel que $t \in J_{i,n}$. Si on pose $J_n(t) := J_{i_n(t),n}$, alors $|J_n(t)| = h_n \to 0$, et donc $J_n(t) \to \{t\}$ quand $n \to \infty$. Par le Fait 1 et comme t est un point de continuité de f, on en déduit que $\mathrm{osc}(f, J_n(t)) \to 0$; autrement dit (par définition de φ_n et ψ_n) que $\psi_n(t) - \varphi_n(t) \to 0$. Comme $\varphi_n(t) \leqslant f(t) \leqslant \psi_n(t)$, on voit ainsi que $\varphi_n(t)$ et $\psi_n(t)$ tendent tous les deux vers $f(t)$. $\square$

Il est maintenant facile de conclure. Par le Fait 3 et comme on suppose que l'ensemble des points de discontinuité de f est négligeable, les suites (φ_n) et (ψ_n) tendent vers f *presque partout*, et donc $\psi_n - \varphi_n \to 0$ presque partout. De plus, on a $|\psi_n(t) - \varphi_n(t)| \leqslant 2\|f\|_\infty$ pour tout n et pour tout $t \in [a,b]$. Par le théorème de convergence bornée, on en déduit que $\int_a^b(\psi_n - \varphi_n) \to 0$, ce qui termine la démonstration. $\square$

Exercice 1. Montrer que le produit de deux fonctions intégrables au sens de Riemann sur $[a,b]$ est intégrable au sens de Riemann. Plus généralement, montrer que si $f_1,\dots,f_N : [a,b] \to \mathbb{R}$ sont intégrables au sens de Riemann et si $\Phi : \mathbb{R}^N \to \mathbb{R}$ est une fonction continue, alors $f(x) := \Phi(f_1(x),\dots,f_N(x))$ est intégrable au sens de Riemann sur $[a,b]$.

Exercice 2. Montrer que toute limite uniforme de fonctions intégrables au sens de Riemann est intégrable au sens de Riemann.

Exercice 3. Soit $f : [a,b] \to \mathbb{R}$ une fonction bornée. Montrer que si f est intégrable au sens de Riemann sur tout intervalle compact $[\alpha,\beta] \subseteq]a,b[$, alors elle est intégrable au sens de Riemann sur $[a,b]$.

3. Fonctions définies par des intégrales

Comme exemples très importants d'application du théorème de convergence dominée, on va maintenant démontrer deux résultats concernant la continuité et la dérivabilité d'une “intégrale à paramètre”, i.e. d'une fonction f de la forme

$$f(x) = \int_\Omega F(x,t)\ d\mu(t)\,,$$

où x appartient à un certain “espace de paramètres” X et F est une fonction de 2 variables, $F : X \times \Omega \to \mathbb{C}$.

Dans ce qui suit, $(\Omega, \mathcal{B}, \mu)$ est comme d'habitude un espace mesuré.

3.1. Continuité. Le "théorème de continuité pour les intégrales à paramètres" s'énonce comme suit.

THÉORÈME 3.1. *Soit $F : X \times \Omega \to \mathbb{C}$, où X est un espace métrique. Soit également $a \in X$. On fait les hypothèses suivantes.*

(i) *Pour tout $x \in X$ fixé, la fonction $t \mapsto F(x,t)$ est intégrable sur Ω.*
(ii) *Pour presque tout $t \in \Omega$, la fonction $x \mapsto F(x,t)$ est continue au point a.*
(iii) ***Hypothèse de domination** : pour tout compact $K \subseteq X$, on peut majorer $|F(x,t)|$ pour $x \in K$ de la façon suivante : $|F(x,t)| \leqslant g_K(t)$, où $g_K(t)$ est* indépendant *de $x \in K$ et la fonction $g_K : \Omega \to \mathbb{R}^+$ est intégrable sur Ω.*

Alors la fonction $f : X \to \mathbb{C}$ définie par

$$f(x) := \int_\Omega F(x,t)\, d\mu(t)$$

est continue au point a.

Démonstration. La fonction f est bien définie par (i). Pour montrer la continuité, fixons une suite $(x_n)_{n\in\mathbb{N}} \subseteq X$ tendant vers a. Il s'agit de montrer que $f(x_n) \to f(a)$.

Par (ii), on sait que $F(x_n,t)$ tend vers $F(a,t)$ presque partout. De plus, $K := \{a\} \cup \{x_n;\ n \in \mathbb{N}\}$ est un *compact* de X car $x_n \to a$. Donc, par (iii), on peut trouver une fonction g_K intégrable sur Ω telle que $\forall n\ \forall t \in \Omega\ :\ |F(x_n,t)| \leqslant g_K(t)$. Par le théorème de convergence dominée appliqué à la suite de fonctions $F_n(t) := F(x_n,t)$, on en déduit que $\int_\Omega F(x_n,t)\, d\mu(t) \to \int_\Omega F(a,t)\, d\mu(t)$, i.e. $f(x_n) \to f(a)$. □

Remarque. Dans (i), il suffit en fait de supposer que la fonction $t \mapsto F(x,t)$ est *μ-mesurable* pour tout $x \in X$ fixé, car si on applique (iii) avec $K := \{x\}$, on obtient $|F(x,t)| \leqslant g_{\{x\}}(t)$, d'où l'intégrabilité de $t \mapsto F(x,t)$. Cependant, si on veut montrer la continuité d'une intégrale à paramètre f, il parait plus naturel de commencer par montrer que la fonction f est bien définie, i.e. que (i) est vérifiée sous la forme donnée plus haut.

COROLLAIRE 3.2. *Soit $F : X \times \Omega \to \mathbb{C}$, où X est un espace métrique. On suppose que $F(x,t)$ est mesurable par rapport à $t \in \Omega$ (pour tout $x \in X$ fixé) et continu par rapport à $x \in X$ (pour tout $t \in \Omega$ fixé),* et de plus *que pour tout compact $K \subseteq X$, il existe une fonction g_K intégrable sur Ω telle que $|F(x,t)| \leqslant g_K(t)$ pour tout $x \in K$ et pour tout $t \in \Omega$. Alors la formule $f(x) := \int_\Omega F(x,t)\, d\mu(t)$ définit une fonction continue sur X.*

Comme cas très particulier de ce résultat, on retrouve le théorème de continuité "pour enfants" :

COROLLAIRE 3.3. *Soit $F : X \times [a,b] \to \mathbb{C}$, où X est un espace métrique et $[a,b]$ est un intervalle compact de $\mathbb{R}$. On suppose que la fonction F est* continue sur $[a,b] \times X$. *Alors la formule $f(x) = \int_a^b F(x,t)\, dt$ définit une fonction continue sur X.*

Démonstration. Comme $F(x,t)$ est évidemment continu par rapport à $x \in X$ et borélien par rapport à $t \in [a,b]$, il suffit de vérifier l'hypothèse de domination. Si K est un compact de X, alors $K \times [a,b]$ est compact, donc la fonction continue F est *bornée* sur $K \times [a,b]$, disons $|F(x,t)| \leqslant M$. La constante M est intégrable sur $[a,b]$ car $[a,b]$ est de longueur finie, donc on peut prendre $g_K(t) := M$. □

Exercice. Démontrer directement le Corollaire 3.3.

Exemple 1. Si $\varphi : \mathbb{R} \to \mathbb{C}$ est une fonction intégrable sur $\mathbb{R}$, alors la formule

$$f(z) := \int_{\mathbb{R}} \frac{\varphi(t)}{t-z}\, dt$$

définit une fonction continue sur $\mathbb{C}\backslash\mathbb{R}$.

Démonstration. La fonction $F : (\mathbb{C}\backslash\mathbb{R}) \times \mathbb{R} \to \mathbb{C}$ définie par $F(z,t) := \frac{\varphi(t)}{t-z}$ est borélienne par rapport à $t \in \mathbb{R}$ et continue par rapport à $z \in \mathbb{C}\backslash\mathbb{R}$. De plus, si $z \in \mathbb{C}\backslash\mathbb{R}$, alors $\delta(z) := \mathrm{dist}(z, \mathbb{R})$ est *strictement positif*, et on a $|F(z,t)| \leqslant \frac{|\varphi(t)|}{\delta(z)}$ pour tout $t \in \mathbb{R}$; donc la fonction $t \mapsto F(z,t)$ est intégrable sur $\mathbb{R}$ (car φ est intégrable), autrement dit $f(z)$ est bien défini.

Si K est un compact de $\mathbb{C}\backslash\mathbb{R}$, alors on peut trouver $\delta_K > 0$ tel que $\delta(z) \geqslant \delta_K$ pour tout $z \in K$: cela vient du fait que la fonction $z \mapsto \delta(z)$ est continue sur le compact K, donc admet un *minimum* sur K, minimum qui est *strictement* positif puisque $\delta(z) > 0$ pour tout $z \in K$. (Ce genre d'"argument de compacité" doit devenir un réflexe.) Pour $z \in K$ et $t \in \mathbb{R}$, on a alors $|F(z,t)| \leqslant \frac{\varphi(t)}{\delta_K} := g_K(t)$. La fonction g_K est intégrable sur $\mathbb{R}$ car φ l'est, et indépendante de $z \in K$. Donc l'hypothèse de domination (iii) dans le Théorème 3.1 est satisfaite. □

Remarque. Cet exemple illustre bien le "mode d'emploi" du Théorème 3.1. Pour montrer que $f(z)$ est bien défini pour tout $z \in \mathbb{C}\backslash\mathbb{R}$, on a besoin de vérifier que la fonction $t \mapsto F(z,t)$ est intégrable, ce qui se fait en majorant $|F(z,t)|$. La majoration obtenu à ce stade dépend de z. Ensuite, pour vérifier la condition (iii) du théorème, il faut s'affranchir de la dépendance en z lorsque z varie dans un compact $K \subseteq \mathbb{C}\backslash\mathbb{R}$. Autrement dit, on reprend la majoration obtenue pour $|F(z,t)|$, et on essaye de faire disparaitre z en utilisant le fait que z est astreint à rester dans un compact.

Exemple 2. Si $f : \mathbb{R} \to \mathbb{C}$ est une fonction borélienne *intégrable sur* $\mathbb{R}$, alors la formule

$$\widehat{f}(x) := \int_{\mathbb{R}} f(t) e^{-ixt} dt$$

définit une fonction $\widehat{f}$ continue sur $\mathbb{R}$. La fonction $\widehat{f}$ est la **transformée de Fourier** de f.

Démonstration. La fonction $F : \mathbb{R} \times \mathbb{R} \to \mathbb{C}$ définie par $F(x,t) := f(t)e^{-ixt}$ est borélienne par rapport à t continue par rapport à x. De plus, on a $|F(x,t)| = |f(t)|$, donc la fonction $t \mapsto F(x,t)$ est intégrable sur $\mathbb{R}$ pour tout x fixé ; autrement dit $\widehat{f}$ est bien définie. Enfin, pour n'importe quel compact $K \subseteq \mathbb{R}$, on peut prendre comme "dominante intégrable" $g_K(t) := |f(t)|$. Donc $\widehat{f}$ est continue par le Théorème 3.1. □

Exemple 3. Soit $f : [0, \infty[\to \mathbb{C}$ une fonction borélienne. On suppose que pour tout $\lambda > 0$, la fonction $t \mapsto f(t)e^{-\lambda t}$ est intégrable sur $[0, \infty[$. Alors la fonction $\mathcal{L}f$ définie par

$$\mathcal{L}f(\lambda) := \int_0^\infty f(t) e^{-\lambda t} dt$$

est continue sur $]0, \infty[$. La fonction $\mathcal{L}f$ est la **transformée de Laplace** de f.

Démonstration. La fonction $F :]0, \infty[\times [0, \infty[\to \mathbb{C}$ définie par $F(\lambda, t) = f(t)e^{-\lambda t}$ est intégrable par rapport à t pour tout $\lambda > 0$ fixé, et continue par rapport à λ pour tout t fixé.

Si K est un compact de $]0,\infty[$, on peut trouver $\varepsilon_K > 0$ tel que $\forall \lambda \in K\ :\ \lambda \geqslant \varepsilon_K$. Pour $\lambda \in K$ et $t \in [0,\infty[$, on a donc $|F(\lambda,t)| \leqslant |f(t)|\, e^{-\varepsilon_K t}$. La fonction $g_K(t) := |f(t)|\, e^{-\varepsilon_K t}$ est intégrable sur $[0,\infty[$ par hypothèse sur f, donc l'hypothèse de domination (iii) dans le Théorème 3.1 est satisfaite. $\square$

Exemple 4. Soit I un intervalle de $\mathbb{R}$, et soit $x_0 \in I$. Si $g : I \to \mathbb{C}$ est localement intégrable, alors la fonction G définie par

$$G(x) := \int_{x_0}^{x} g(t)\, dt$$

est continue sur I. (Bien sûr, si g est *continue*, alors G est même de classe $\mathcal{C}^1$.)

Démonstration. La fonction G est bien définie par hypothèse sur g. Fixons $a \in I$ et montrons que G est continue au point a. Comme $G(x) = \int_{x_0}^{a} f(t)\, dt + \int_a^x f(t)\, dt$, il suffit en fait de démontrer que $G_a(x) := \int_a^x f(t)\, dt$ est continue au point a. On va se contenter de montrer la continuité *à droite* au point a. (La preuve est la même pour la continuité à gauche.)

On a pour $x \geqslant a$:

$$G_a(x) = \int_I \mathbf{1}_{[a,x]}(t) f(t)\, dt := \int_I F(x,t)\, dt\,.$$

Comme $F(x,t) = f(t)\mathbf{1}_{[a,\infty[}(t)\, \mathbf{1}_{[t,\infty[}(x)$, on voit que pour $t \in I$ fixé, le seul point de discontinuité possible de la fonction $x \mapsto F(x,t)$ est le point $x = t$ (il se peut que cette fonction soit continue car $f(t)$ pourrait valoir 0). Par conséquent, la fonction $x \mapsto F(x,t)$ est continue (à droite) au point a pour tout $t \neq a$, et donc pour *presque tout* $t \in I$. Donc la condition (ii) du Théorème 3.1 est satisfaite.

Soit K un compact de I. Alors $K \cup \{a\}$ est aussi un compact de I, donc on peut trouver $\alpha_K, \beta_K \in I$ tels que $K \cup \{a\} \subseteq [\alpha_K, \beta_K]$. On a alors $[a,x] \subseteq [\alpha_K, \beta_K]$ pour tout $x \in K$, donc $|F(x,t)| = \mathbf{1}_{[a,x]}(t)\, |f(t)| \leqslant \mathbf{1}_{[\alpha_K,\beta_K]}(t)|f(t)|$ pour $x \in K$ et pour tout $t \in I$. Comme f est localement intégrable, la fonction $g_K := \mathbf{1}_{[\alpha_K,\beta_K]}|f|$ est intégrable sur I, donc l'hypothèse de domination (iii) du Théorème 3.1 est également satisfaite. $\square$

Remarque. Ce dernier exemple est un peu plus subtil que les deux précédents, car on a vraiment besoin de la latitude offerte par le “presque partout” dans la condition (ii) du Théorème 3.1 : quoi qu'on fasse, pour $a \in I$ donné, la fonction $x \mapsto F(x,t) = f(t)1_{[a,\infty[}(t)\, 1_{[a,x]}(t)$ n'est a priori pas continue au point a pour *tout* $t \in I$, mais seulement pour $t \neq a$.

3.2. Dérivabilité. Le “théorème de dérivabilité pour les intégrales à paramètres” s'énonce comme suit.

Théorème 3.4. *Soit $F : I \times \Omega \to \mathbb{C}$, où I est un intervalle de $\mathbb{R}$. On fait les hypothèses suivantes.*

(i) *Pour tout $x \in I$, la fonction $t \mapsto F(x,t)$ est intégrable sur Ω ;*
(ii) *Pour tout $t \in \Omega$, la fonction $x \mapsto F(x,t)$ est dérivable sur I ;*
(iii) ***Hypothèse de domination** : pour tout compact $K \subseteq X$, on peut majorer $|\frac{\partial F}{\partial x}(x,t)|$ pour $x \in K$ de la façon suivante : $|\frac{\partial F}{\partial x}(x,t)| \leqslant g_K(t)$, où $g_K(t)$ est indépendant de $x \in K$ et la fonction $g_K : \Omega \to \mathbb{R}^+$ est intégrable sur Ω.*

Alors la fonction $f : I \to \mathbb{C}$ définie par

$$f(x) := \int_\Omega F(x,t)\, d\mu(t)$$

est dérivable sur I, et on peut dériver sous l'intégrale :

$$f'(x) = \int_\Omega \frac{\partial F}{\partial x}(x,t)\, d\mu(t)\,.$$

Remarque 1. Si dans (ii) on suppose que la fonction $x \mapsto F(x,t)$ est *de classe* $\mathcal{C}^1$ sur I (pour tout $t \in \Omega$), alors on peut conclure que la fonction f est de classe $\mathcal{C}^1$: cela découle de la formule pour $f'(x)$ et du théorème de *continuité* des intégrales à paramètres.

Remarque 2. Le théorème de dérivabilité ressemble énormément au théorème de continuité. La différence principale est que dans l'hypothèse de domination, ce n'est pas $|F(x,t)|$ qu'on majore mais $\left|\frac{\partial F}{\partial x}(x,t)\right|$. Ce n'est pas du tout surprenant, si on pense par exemple au théorème bien connu permettant de conclure à la dérivabilité de la limite d'une suite de fonctions dérivables (f_n) : l'hypothèse importante dans ce théorème est la convergence uniforme *de la suite des dérivées* (f_n').

Preuve du Théorème 3.4. La fonction f est bien définie par (i). De plus, si $x \in I$ alors, en appliquant (iii) au compact $K := \{x\}$, on voit que la fonction $t \mapsto \frac{\partial F}{\partial x}(x,t)$ est intégrable sur Ω.

Soit $a \in I$ fixé. Pour montrer que f est dérivable en a avec la bonne formule pour $f'(a)$, il suffit de vérifier que pour toute suite $(x_n) \subseteq I$ tendant vers a (avec $x_n \neq a$), on a

$$\lim_{n\to\infty} \frac{f(x_n) - f(a)}{x_n - a} = \int_\Omega \frac{\partial F}{\partial x}(a,t)\, d\mu(t)\,.$$

Par définition de f, on a

$$\frac{f(x_n) - f(a)}{x_n - a} = \int_\Omega \frac{F(x_n,t) - F(a,t)}{x_n - a}\, d\mu(t)\,.$$

Par (ii), $\frac{F(x_n,t)-F(a,t)}{x_n-a} \to \frac{\partial F}{\partial x}(a,t)$ pour tout $t \in \Omega$. Il s'agit donc d'intervertir une limite et une intégrale ; ce qu'on va bien entendu faire à l'aide du théorème de convergence dominée.

Comme $x_n \to a$, on peut trouver un intervalle compact $K \subseteq I$ contenant a et tous les x_n. Comme $\left|\frac{\partial F}{\partial x}(x,t)\right| \leqslant g_K(t)$ pour tout $x \in K$, on a par l'*inégalité des accroissements finis* appliquée à la fonction $x \mapsto F(x,t)$:

$$\left|\frac{F(x_n,t) - F(a,t)}{x_n - a}\right| \leqslant g_K(t)$$

pour tout n et pour tout $t \in \Omega$. Par convergence dominée, on obtient donc

$$\lim_{n\to\infty} \int_\Omega \frac{F(x_n,t) - F(a,t)}{x_n - a}\, d\mu(t) = \int_\Omega \frac{\partial F}{\partial x}(a,t)\, d\mu(t)\,,$$

ce qui termine la preuve du théorème. $\square$

Comme cas très particulier du Théorème 3.4, on retrouve le théorème de dérivabilité “pour enfants” :

COROLLAIRE 3.5. *Soit $[a,b]$ un intervalle compact de $\mathbb{R}$, et soit $F : I \times [a,b] \to \mathbb{C}$. On suppose que la fonction F est continue sur $I \times [a,b]$, que F possède en tout point (x,t) une dérivée partielle $\frac{\partial F}{\partial x}(x,t)$ par rapport à la 1ère variable, et que la fonction $\frac{\partial F}{\partial x}$ est continue sur $I \times [a,b]$. Alors la formule $f(x) = \int_a^b F(x,t)\, dt$ définit une fonction de classe $\mathcal{C}^1$ sur I, avec $f'(x) = \int_a^b \frac{\partial F}{\partial x}(x,t)\, dt$.*

Démonstration. Pour tout $x \in I$, la fonction $t \mapsto F(x,t)$ est continue sur l'intervalle compact $[a,b]$, donc intégrable sur $[a,b]$. Pour tout $t \in [a,b]$, la fonction $x \mapsto F(x,t)$ est dérivable sur I par hypothèse. Enfin, si K est un compact de I, alors la fonction continue $\frac{\partial F}{\partial x}$ est *bornée* sur le compact $K \times [a,b]$, disons $\left|\frac{\partial F}{\partial x}\right| \leqslant M$; et comme les constantes sont intégrables sur l'intervalle borné $[a,b]$, on en déduit que la condition de domination (iii) est satisfaite avec $g_K(t) := M$. Donc le théorème de dérivation s'applique (*cf* la Remarque 1 plus haut). □

Corollaire 3.6. *Soit $F : U \times \Omega \to \mathbb{C}$, où U est un ouvert de $\mathbb{R}^N$. On suppose que*

(1) *$F(x,t)$ est intégrable par rapport à $t \in \Omega$ (pour $x \in U$ fixé) et de classe $\mathcal{C}^1$ par rapport à $x \in U$ (pour $t \in \Omega$ fixé) ;*
(2) *pour tout compact $K \subseteq U$, il existe des fonctions $g_{K,1}, \dots, g_{K,N}$ intégrables sur Ω telles que pour $j = 1, \dots, N$ on ait*

$$\forall x \in K \ \forall t \in \Omega \ : \ \left|\frac{\partial F}{\partial x_j}(x,t)\right| \leqslant g_{K,j}(t)\,.$$

Alors $f(x) := \int_\Omega F(x,t)\, d\mu(t)$ est de classe $\mathcal{C}^1$ sur U, avec pour $j = 1, \dots, N$:

$$\frac{\partial f}{\partial x_j}(x) = \int_\Omega \frac{\partial F}{\partial x_j}(x,t)\, d\mu(t)\,.$$

Démonstration. Par le théorème, on voit que $\frac{\partial f}{\partial x_j}$ existe sur U pour $j = 1, \dots, N$, avec $\frac{\partial f}{\partial x_j}(x) = \int_\Omega \frac{\partial F}{\partial x_j}(x,t)\, d\mu(t)$. Par le théorème de continuité, cette formule montre que les fonctions $\frac{\partial f}{\partial x_j}$ sont continues sur U ; donc f est de classe $\mathcal{C}^1$. □

Exemple 1. Si $f : \mathbb{R} \to \mathbb{C}$ est une fonction intégrable vérifiant

$$\int_{\mathbb{R}} |t|^N\, |f(t)|\, dt < \infty$$

pour un certain entier $N \geqslant 1$, alors sa transformée de Fourier

$$\widehat{f}(x) = \int_{\mathbb{R}} f(t) e^{-ixt} dt$$

est de classe $\mathcal{C}^N$ sur $\mathbb{R}$, avec pour $k = 1, \dots, N$:

$$\widehat{f}^{(k)}(x) = (-i)^k \int_{\mathbb{R}} t^k f(t) e^{-ixt} dt = (-i)^k \widehat{(t^k f)}(x)\,.$$

Démonstration. On a besoin du fait suivant ;

Fait. Pour tout $k \in \{1, \dots, N\}$ la fonction $t \mapsto t^k f(t)$ est intégrable sur $\mathbb{R}$.

Démonstration. [*Preuve du Fait*] Cette fonction est intégrable sur $[-1,1]$ car f l'est et t^k est borné sur $[-1,1]$. De plus, on a $|t^k f(t)| \leqslant |t^N f(t)|$ si $|t| \geqslant 1$, donc $t^k f(t)$ est aussi intégrable sur $\{|t| \geqslant 1\}$. Au total, $t^k f(t)$ est intégrable sur $\mathbb{R}$. □

Si on pose $F(x,t) := f(t)e^{-ixt}$, alors $F(x,t)$ est intégrable par rapport à $t \in \mathbb{R}$ et de classe $\mathcal{C}^N$ par rapport à x, avec $\frac{\partial^k F}{\partial x^k}(x,t) = (-i)^k\, t^k f(t)\, e^{-ixt}$ pour tout $k \in \{1, \dots, N\}$. On a $\left|\frac{\partial^k F}{\partial x^k}(x,t)\right| = |t^k f(t)|$, fonction intégrable sur $\mathbb{R}$ d'après le Fait, et indépendante de x ; donc on peut appliquer N fois le Corollaire 3.6 : $\widehat{f}$ est de classe $\mathcal{C}^N$ avec

$$\widehat{f}^{(k)}(x) = \int_{\mathbb{R}} (-i)^k\, t^k f(t) e^{-ixt} dt = (-i)^k\, \widehat{(t^k f)}(x) \qquad \text{pour } k = 1, \dots, N\,.$$

□

Exemple 2. Si $f : [0, \infty[\to \mathbb{C}$ est une fonction borélienne admettant une transformée de Laplace, i.e. telle que la fonction $t \mapsto f(t)e^{-\lambda t}$ est intégrable sur $[0, \infty[$ pour tout $\lambda > 0$, alors sa transformée de Laplace

$$\mathcal{L}f(\lambda) = \int_0^\infty f(t)e^{-\lambda t}dt$$

est de classe $\mathcal{C}^\infty$ sur $]0, \infty[$, avec pour tout $k \in \mathbb{N}$:

$$(\mathcal{L}f)^{(k)}(\lambda) = (-1)^k \int_0^\infty t^k f(t)e^{-\lambda t}dt\,.$$

Démonstration. Cette fois, on a besoin du fait suivant.

Fait. Pour tout $\varepsilon > 0$ et pour tout $k \in \mathbb{N}^*$, la fonction $t \mapsto t^k f(t)e^{-\varepsilon t}$ est intégrable sur $\mathbb{R}$.

Preuve du Fait. Fixons $\varepsilon > 0$, et choisissons α tel que $0 < \alpha < \varepsilon$. On a

$$\left|t^k f(t)e^{-\varepsilon t}\right| = \left|f(t)e^{-\alpha t}\right| \times |t^k e^{-(\varepsilon-\alpha)t}|\,.$$

De plus, $f(t)e^{-\alpha t}$ est intégrable sur $[0, \infty[$ car f admet une transformée de Laplace, et la fonction $t \mapsto t^k e^{-(\varepsilon-\alpha)t}$ est *bornée* sur $[0, \infty[$ car elle est continue et tend vers 0 à l'infini (puisque $\varepsilon - \alpha > 0$). Donc $t^k f(t)e^{-\varepsilon t}$ est intégrable sur $[0, \infty[$. □

Si on pose $F(\lambda, t) := f(t)e^{-\lambda t}$, alors $F(\lambda, t)$ est intégrable par rapport à $t \in [0, \infty[$ et de classe $\mathcal{C}^\infty$ par rapport à $\lambda \in]0, \infty[$, avec $\frac{\partial^k F}{\partial \lambda^k}(\lambda, t) = (-1)^k t^k f(t)e^{-\lambda t}$ pour tout $k \in \mathbb{N}^*$. Si K est un compact de $]0, \infty[$, on peut trouver $\varepsilon_K > 0$ tel que $\forall \lambda \in K \;:\; \lambda \geqslant \varepsilon_K$. Pour $\lambda \in K$, on a alors $\left|\frac{\partial^k F}{\partial \lambda^k}(\lambda, t)\right| \leqslant t^k |f(t)| e^{-\varepsilon_K t}$, fonction intégrable sur $[0, \infty[$ d'après le Fait et indépendante de $\lambda \in K$. Donc on peut appliquer le Corollaire 3.6 une infinité de fois. □

Remarque 3.7. Les hypothèses du théorème de dérivabilité sont un peu rigides, car elles ne laissent aucune place au "presque partout". Voici une version plus souple, qui couvre effectivement des cas où le théorème ne s'applique pas tel quel (*cf* l'exercice ci-après).

Soit $F : I \times \Omega \to \mathbb{C}$ telle que $f(x) = \int_\Omega F(x, t)\, d\mu(t)$ est bien défini pour tout $x \in I$, et soit $a \in I$. On fait les hypothèses suivantes :

(ii') *pour presque tout $t \in \Omega$, la fonction $x \mapsto F(x, t)$ est dérivable en a ;*

(iii') *pour tout compact $K \subseteq I$, il existe une fonction intégrable $C_K : \Omega \to \mathbb{R}^+$ telle que, pour presque tout $t \in \Omega$, la fonction $x \mapsto F(x, t)$ est $C_K(t)$-lipschitzienne sur K.*

Alors on peut conclure que f est dérivable en a, avec la bonne formule pour $f'(a)$.

Démonstration. Il suffit essentiellement de recopier la preuve du Théorème 3.4. (Les détails sont laissés en exercice.) □

Exercice. Soit $\phi : \mathbb{R} \to \mathbb{C}$ une fonction intégrable telle que $t\phi(t)$ est également intégrable sur $\mathbb{R}$. Montrer que la formule $f(x) := \int_\mathbb{R} |x - t|\, \phi(t)\, dt$ définit une fonction de classe $\mathcal{C}^1$ sur $\mathbb{R}$.

4. Deux exemples d'utilisation des théorèmes

4.1. Transformée de Fourier d'une gaussienne. Soit $g : \mathbb{R} \to \mathbb{R}$ la fonction "gaussienne" définie par

$$g(t) := e^{-t^2/2}.$$

Comme g est continue (donc localement intégrable) et $g(t) = o(1/t^2)$ en $\pm\infty$, la fonction g est intégrable sur $\mathbb{R}$. Le résultat suivant montre que g est, à une constante près, sa propre transformée de Fourier ; ou, si on préfère, un *vecteur propre* de la transformation de Fourier.

Proposition 4.1. *On a $\widehat{g} = \sqrt{2\pi}\, g$; autrement dit,*

$$\int_{-\infty}^{\infty} e^{-t^2/2} e^{-ixt} dt = \sqrt{2\pi}\, e^{-x^2/2} \quad \textit{pour tout } x \in \mathbb{R}\,.$$

Démonstration. La fonction $t \mapsto te^{-t^2/2}$ est intégrable sur $\mathbb{R}$, car elle est continue et $o(1/t^2)$ en $\pm\infty$. Donc $\widehat{g}$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$, avec

$$\widehat{g}\,'(x) = -i \int_{-\infty}^{\infty} te^{-t^2/2} e^{-ixt} dt\,.$$

Maintenant, on intègre par parties, en remarquant que $te^{-t^2/t}$ est au signe près la dérivée de $e^{-t^2/2}$. Pour tout $T > 0$, on a

$$\int_{-T}^{T} te^{-t^2/2} e^{-ixt} dt = \Big[-e^{-t^2/2} e^{-ixt}\Big]_{-T}^{T} + \int_{-T}^{T} e^{-t^2/2} \times (-ixe^{-ixt})\, dt\,.$$

Quand $T \to \infty$, le "crochet" tend vers 0 car $e^{-t^2/2} \to 0$ en $\pm\infty$ et $|e^{-ixt}| = 1$; et l'intégrale apparaissant au second membre tend vers $-ix \int_{-\infty}^{\infty} e^{-t^2/2} e^{-ixt} dt = -ix\, \widehat{g}(x)$. On obtient donc

$$\widehat{g}\,'(x) = -i \times \big(-ix\, \widehat{g}(x)\big) = -x\, \widehat{g}(x)\,.$$

Ainsi, $\widehat{g}$ est solution de l'équation différentielle $y'(x) = -x\, y(x)$; et donc

$$\widehat{g}(x) = C\, e^{-x^2/2}\,,$$

pour une certaine constante C. On a $C = \widehat{g}(0) = \int_{-\infty}^{\infty} e^{-t^2/2} dt$, et "on sait bien" que cette intégrale est égale à $\sqrt{2\pi}$ (on verra une preuve de ce fait au Chapitre 7). □

Exercice. Pour tout $\lambda > 0$, déterminer la transformée de Fourier de $g_\lambda(t) := e^{-\lambda t^2}$.

4.2. Calcul de l'intégrale $\int_0^\infty \frac{\sin t}{t}\, dt$. On a vu au Chapitre 4 que l'intégrale $\int_1^\infty \frac{\sin t}{t}\, dt$ existe en tant qu'intégrale généralisée. Comme de plus la fonction $f(t) := \frac{\sin t}{t}$ se prolonge par continuité en 0 (avec $f(0) = 1$) et est donc localement intégrable sur $[0, \infty[$, on en déduit que l'intégrale $\int_0^\infty \frac{\sin t}{t}\, dt$ existe également. Dans cette section, on va calculer la valeur de cette intégrale.

On va utiliser le résultat général suivant concernant la transformée de Laplace.

Proposition 4.2. *Soit $f : [0, \infty[\to \mathbb{C}$ une fonction continue admettant une transformée de Laplace. On suppose que $\int_0^\infty f(t)\, dt$ existe en tant qu'intégrale généralisée. Alors*

$$\mathcal{L}f(\lambda) \to \int_0^\infty f(t)\, dt \quad \textit{quand } \lambda \to 0^+\,.$$

Démonstration. On notera I l'intégrale $\int_0^\infty f(t)\,dt$.

Soit $F : [0,\infty[\to \mathbb{C}$ définie par $F(x) = \int_0^x f(t)\,dt$. Comme f est continue, F est de classe $\mathcal{C}^1$ avec $F' = f$. De plus, $F(x) \to I$ quand $x \to \infty$. En particulier, la fonction F est *bornée* sur $[0,\infty[$.

Soit $\lambda > 0$. Pour tout $X \geqslant 0$, une intégration par parties donne

$$\begin{aligned}\int_0^X f(t)e^{-\lambda t}dt &= \Big[F(t)e^{-\lambda t}\Big]_0^X + \lambda\int_0^X F(t)e^{-\lambda t}dt\\ &= F(X)e^{-\lambda X} + \lambda\int_0^X F(t)e^{-\lambda t}dt\,.\end{aligned}$$

Comme F est bornée, $F(X)e^{-\lambda X} \to 0$ quand $X \to \infty$ et la fonction $t \mapsto F(t)e^{-\lambda t}$ est intégrable sur $[0,\infty[$. En faisant tendre X vers l'infini, on obtient donc

$$\begin{aligned}\mathcal{L}f(\lambda) &= \lambda\int_0^\infty F(t)e^{-\lambda t}dt\\ &= \int_0^\infty F\left(\frac{u}{\lambda}\right)e^{-u}du\,,\end{aligned}$$

pour tout $\lambda > 0$.

Comme $F(x) \to I$ quand $x \to \infty$, on a $\lim_{\lambda\to 0^+} F(u/\lambda) = I$ pour tout $u > 0$. De plus, $|F(u/\lambda)e^{-u}| \leqslant \|F\|_\infty e^{-u}$ pour tout $\lambda > 0$ et pour tout u. Comme e^{-u} est intégrable sur $]0,\infty[$, on peut donc appliquer le theéorème de convergence dominée (dans sa version "continue") pour conclure que

$$\mathcal{L}f(\lambda) \xrightarrow{\lambda\to 0^+} \int_0^\infty I\times e^{-u}du = I\int_0^\infty e^{-u}du = I\,.$$

□

Remarque. La proposition 4.2 pourrait s'appeler *théorème d'Abel pour les transformées de Laplace*, car c'est un analogue "continu" du théorème d'Abel "bien connu" pour les séries entières.

De façon précise, si on fait le parallèle

$$\text{(intégrale)} \longleftrightarrow \text{(série)} \qquad \text{et} \qquad (e^{-\lambda}, \lambda > 0) \longleftrightarrow (x \in]0,1[\,),$$

une transformée de Laplace $\int_0^\infty f(t)e^{-\lambda t}dt = \int_0^\infty f(t)(e^{-\lambda})^t dt = \int_0^\infty f(t)x^t dt$ pour $\lambda > 0$ est l'analogue continu de la somme d'une série entière $\sum_0^\infty c_k x^k$ pour $x < 1$.

L'existence de l'intégrale $\int_0^\infty f(t)\,dt = \int_0^\infty f(t)e^{-0t}$ correspond à la convergence de la série $\sum c_k = \sum c_k 1^k$. Donc la Proposition 4.2, i.e. l'énoncé "$\int_0^\infty f(t)e^{-\lambda t}dt \to \int_0^\infty f(t)\,dt$ quand $\lambda \to 0^+$si $\int_0^\infty f(t)dt$ existe" correspond à l'énoncé "$\sum_0^\infty c_k x^k \to \sum_0^\infty c_k$ quand $x \to 1^-$ si la série $\sum c_k$ est convergente", i.e. au théorème d'Abel.

Corollaire 4.3. *On a* $\displaystyle\int_0^\infty \frac{\sin t}{t}\,dt = \frac{\pi}{2}\cdot$

Démonstration. On applique la proposition avec $f(t) := \frac{\sin t}{t}$, qui est continue sur $[0,\infty[$ avec $f(0) = 1$. La fonction f est bornée sur $[0,\infty[$, donc elle admet une transformée de Laplace ; et on a vu que $I := \int_0^\infty \frac{\sin t}{t}\,dt$ existe.

On sait que la fonction $\mathcal{L}f$ est de classe $\mathcal{C}^1$ sur $]0,\infty[$, avec

$$(\mathcal{L}f)'(t) = -\int_0^\infty tf(t)\,dt = -\int_0^\infty e^{-\lambda t}\sin t\,dt\,.$$

Cette dernière intégrale est la partie imaginaire de $\int_0^\infty e^{-\lambda t} e^{it} dt = \int_0^\infty e^{-(\lambda - i)t} dt$, qui vaut $\frac{1}{\lambda - i}\cdot$ Donc

$$(\mathcal{L}f)'(\lambda) = -\mathrm{Im}\left(\frac{1}{\lambda - i}\right) = -\frac{1}{1+\lambda^2} \qquad \text{pour tout } \lambda > 0\,;$$

et par conséquent

$$\mathcal{L}f(\lambda) = -\arctan(\lambda) + C\,,$$

pour une certaine constante C.

Comme $\mathcal{L}f(\lambda) \to I$ quand $\lambda \to 0^+$ d'après la proposition et comme $\arctan(0) = 0$, on voit que $C = I$.

Mais par ailleurs, on a aussi

$$\lim_{\lambda \to \infty} \mathcal{L}f(\lambda) = 0\,,$$

ce qui se voit en appliquant le théorème de convergence dominée, ou simplement en écrivant $|\mathcal{L}f(\lambda)| \leqslant \int_0^\infty |f(t)| e^{-\lambda t} dt \leqslant \int_0^\infty e^{-\lambda t} dt = \frac{1}{\lambda}\cdot$ Comme $\arctan(\lambda) \to \frac{\pi}{2}$ quand $\lambda \to \infty$, on en déduit que $C = \frac{\pi}{2}$, d'où finalement $I = \frac{\pi}{2}\cdot$ $\square$

Chapitre 7

Intégration dans $\mathbb{R}^N$

1. Cadre et notations

Dans ce chapitre, on va intégrer des fonctions *boréliennes*, positives ou à valeurs complexes, sur des parties boréliennes de $\mathbb{R}^N$, $N \geqslant 1$, par rapport à la mesure de Lebesgue λ_N.

Le mot "intégrable" signifiera toujours "intégrable par rapport à λ_N". Une fonction f est donc intégrable sur $\Omega \subseteq \mathbb{R}^N$ si et seulement si elle est borélienne et

$$\int_\Omega |f|\, d\lambda_N < \infty\,.$$

En particulier, si $\Omega \subseteq \mathbb{R}^N$ est un borélien *borné*, alors toute fonction borélienne *bornée* $f : \Omega \to \mathbb{C}$ est intégrable sur Ω, car $\lambda_N(\Omega) < \infty$. C'est le cas par exemple si f est continue et si Ω est compact.

Comme d'habitude, l'intégrale d'une fonction f sur un ensemble $\Omega \subseteq \mathbb{R}^N$ peut se noter

$$\int_\Omega f\, d\lambda_N, \quad \int_\Omega f(x)\, d\lambda_N(x) \quad \text{ou} \quad \int_\Omega f\,.$$

Cependant, on utilisera aussi les notations

$$\int_\Omega f(x)\, dx \quad \text{ou} \quad \int_\Omega f(x_1, \dots, x_N)\, dx_1 \cdots dx_N\,.$$

Enfin, pour $N = 2$ et $N = 3$, on écrira plutôt

$$\int_\Omega f(x, y)\, dxdy \quad \text{et} \quad \int_\Omega f(x, y, z)\, dxdydz\,.$$

Remarque. Si $p, q \in \mathbb{N}^*$, on identifiera constamment $\mathbb{R}^{p+q}$ avec $\mathbb{R}^p \times \mathbb{R}^q$. On rappelle la **propriété de produit** établie au Chapitre 2 : si $A \subseteq \mathbb{R}^p$ et $B \subseteq \mathbb{R}^q$, alors

$$\lambda_{p+q}(A \times B) = \lambda_p(A) \times \lambda_q(B)\,.$$

2. Interprétation géométrique de l'intégrale

Si f est une fonction positive définie sur un ensemble $\Omega \subseteq \mathbb{R}^N$, on notera $\mathrm{SG}(f, \Omega)$ le **sous-graphe de f au dessus de Ω** :

$$\mathrm{SG}(f, \Omega) := \{(x, \tau) \in \mathbb{R}^N \times \mathbb{R};\ x \in \Omega \ \text{et}\ 0 \leqslant \tau < f(x)\}\,.$$

Par définition, $\mathrm{SG}(f, \Omega)$ est donc une partie de $\mathbb{R}^N \times \mathbb{R} = \mathbb{R}^{N+1}$.

Lemme 2.1. *Si $\Omega \subseteq \mathbb{R}^N$ est borélien et si $f : \Omega \to [0, \infty]$ est borélienne, alors $\mathrm{SG}(f, \Omega)$ est un borélien de $\mathbb{R}^{N+1}$.*

Démonstration. Soit $\Phi : \Omega \times \mathbb{R} \to [0,\infty] \times \mathbb{R}$ définie par $\Phi(x,\tau) := (f(x),\tau)$. L'application Φ est borélienne car ses composantes le sont ; et on a

$$\mathrm{SG}(f,\Omega) = \Phi^{-1}(A)\,,$$

où $A := \{u,v) \in [0,\infty] \times \mathbb{R};\ 0 \leqslant v < u\}$. L'ensemble A est borélien dans $[0,\infty] \times \mathbb{R}$ car c'est l'intersection d'un ouvert et d'un fermé, à savoir $\{(u,v);\ v < u\}$ et $\{(u,v);\ v \geqslant 0\}$. Donc $\mathrm{SG}(f,\Omega) = \Phi^{-1}(A)$ est borélien dans $\Omega \times \mathbb{R}$, et donc dans $\mathbb{R}^{N+1}$ puisque Ω est borélien dans $\mathbb{R}^N$. □

La proposition suivante donne en particulier un sens précis à la phrase "l'intégrale, c'est l'aire sous le graphe".

Proposition 2.2. *Soit Ω un borélien de $\mathbb{R}^N$. Pour toute fonction borélienne $f : \Omega \to [0,\infty]$, on a*

$$\int_\Omega f\,d\lambda_N = \lambda_{N+1}(\mathrm{SG}(f,\Omega))\,.$$

En particulier, si $\Omega \subseteq \mathbb{R}$, alors $\int_\Omega f(x)\,dx$ est l'aire *sous le graphe de f ; et si $\Omega \subseteq \mathbb{R}^2$, alors $\int_\Omega f(x,y)\,dxdy$ est* le volume *sous le graphe de f.*

Démonstration. Soit $f : \Omega \to [0,\infty]$ une fonction borélienne positive. On va distinguer 2 cas.

Cas 1. Supposons que f soit *étagée.*

Écrivons $f = \sum_{i=1}^n \alpha_i \mathbf{1}_{E_i}$, où $\alpha_i \in \mathbb{R}^+$ et les E_i sont des boréliens formant une partition de Ω. On a alors

$$\int_\Omega f\,d\lambda_n = \sum_{i=1}^n \alpha_i\,\lambda_N(E_i)\,.$$

Par ailleurs, on voit en faisant un dessin que $\mathrm{SG}(f,\Omega)$ est la réunion disjointe des $E_i \times [0,\alpha_i[$. Donc

$$\begin{aligned}
\lambda_{N+1}(\mathrm{SG}(f,\Omega)) &= \sum_{i=1}^n \lambda_{N+1}\big(E_i \times [0,\alpha_i[\big)\\
&= \sum_{i=1}^n \lambda_N(E_i)\,\lambda_1([0,\alpha_i[) \qquad \text{par la propriété de produit}\\
&= \sum_{i=1}^n \alpha_i\,\lambda_N(E_i) = \int_\Omega f\,d\lambda_N\,.
\end{aligned}$$

Cas 2. Cas général.

Soit (φ_n) une suite approximante de fonctions étagées positives pour f. Comme $f = \sup_n \varphi_n$, l'équivalence suivante a lieu pour tout $(x,\tau) \in \Omega \times \mathbb{R}$:

$$f(x) > \tau \iff \exists n \in \mathbb{N}\ :\ \varphi_n(x) > \tau\,.$$

Autrement dit, on a

$$\mathrm{SG}(f,\Omega) = \bigcup_{n=0}^\infty \mathrm{SG}(\varphi_n,\Omega)\,.$$

Comme la suite des sous-graphes $\mathrm{SG}(\varphi_n,\Omega)$ est croissante (car la suite des fonctions φ_n l'est), on en déduit

$$\begin{aligned}\lambda_{N+1}(\mathrm{SG}(f,\Omega)) &= \lim_{n\to\infty}\lambda_{N+1}\Big(\mathrm{SG}(\varphi_n,\Omega)\Big)\\ &= \lim_{n\to\infty}\int_\Omega \varphi_n\,d\lambda_N \qquad \text{par le Cas 1}\\ &= \int_\Omega f\,d\lambda_N \qquad \text{par convergence monotone}\,.\end{aligned}$$

□

Remarque. On peut utiliser cet exemple pour calculer... l'aire d'un disque.

Soit en effet $R>0$ et soit $D\subseteq\mathbb{R}^2$ le disque ouvert de centre $0=(0,0)$ et de rayon R :

$$D := \{(x,y)\in\mathbb{R}^2;\ x^2+y^2<R^2\}\,.$$

Si on pose $D^+ := D\cap\{(x,y);\ x\geqslant 0\}$ et $D^- := D\cap\{(x,y);\ x\leqslant 0\}$, alors $D=D^+\cup D^-$ et $D^+\cap D^- =]-R,R[\times\{0\}$. Comme $\lambda_2\big(]-R,R[\times\{0\}\big)=0$ (car $]-R,R[\times\{0\}$ est contenu dans une droite), on a donc $\lambda_2(D)=\lambda_2(D^+)+\lambda_2(D^-)$. De plus $\lambda_2(D^+)=\lambda_2(D^-)$ par symétrie, donc $\lambda_2(D)=2\lambda_2(D^+)$. Il suffit donc de calculer $\lambda_2(D^+)$.

Si on note $f:]-R,R[\to\mathbb{R}$ la fonction définie par $f(x):=\sqrt{R^2-x^2}$, alors on a visiblement $D^+=\mathrm{SG}(f,\,]-R,R[)$. Donc

$$\lambda_2(D^+)=\int_{-R}^R f(x)\,dx=\int_{-R}^R\sqrt{R^2-x^2}\,dx\,.$$

En posant $x=Rt$, de sorte que $dx=Rdt$ et $\sqrt{R^2-x^2}=R\sqrt{1-t^2}$, on obtient $\int_{-R}^R\sqrt{R^2-x^2}\,dx=R^2\int_{-1}^1\sqrt{1-t^2}\,dt$. L'intégrale $\int_0^1\sqrt{1-t^2}\,dt$ a été calculée au Chapitre 5 : elle vaut fort heureusement $\frac{\pi}{2}\cdot$ Ainsi, $\lambda_2(D^+)=\frac{\pi R^2}{2}$, et donc $\lambda_2(D)=\pi R^2$.

3. Intégrales itérées

3.1. Coupes et volumes. Dans cette section, p et q sont des entiers strictement positifs. Comme indiqué plus haut, on identifiera constamment $\mathbb{R}^p\times\mathbb{R}^q$ avec $\mathbb{R}^{p+q}$.

NOTATIONS. Soit $\Omega\subseteq\mathbb{R}^p\times\mathbb{R}^q=\mathbb{R}^{p+q}$.

- Pour tout $x\in\mathbb{R}^p$, on note Ω_x la **coupe de Ω à l'abscisse** x :
$$\Omega_x := \{y\in\mathbb{R}^q;\ (x,y)\in\Omega\}\subseteq\mathbb{R}^q\,.$$
- Pour tout $y\in\mathbb{R}^p$, on note Ω^y la **coupe de Ω à l'ordonnée** y :
$$\Omega^y := \{x\in\mathbb{R}^p;\ (x,y)\in\Omega\}\subseteq\mathbb{R}^p\,.$$

EXEMPLE. Supposons que Ω soit de la forme $\Omega=A\times B$, avec $A\subseteq\mathbb{R}^p$ et $B\subseteq\mathbb{R}^q$. Alors

$$\Omega_x=\begin{cases}B & \text{si } x\in A,\\ \varnothing & \text{si } x\notin A;\end{cases}\qquad\text{et}\qquad \Omega^y=\begin{cases}A & \text{si } y\in B,\\ \varnothing & \text{si } y\notin B.\end{cases}$$

LEMME 3.1. *$\Omega\subseteq\mathbb{R}^p\times\mathbb{R}^q$ est borélien, alors toutes les coupes Ω_x et Ω^y sont des ensembles boréliens.*

Démonstration. Si $x \in \mathbb{R}^p$, alors $\Omega_x = \Phi_x^{-1}(\Omega)$, où $\Phi_x : \mathbb{R}^q \to \mathbb{R}^p \times \mathbb{R}^q$ est l'application $y \mapsto (x, y)$; donc Ω_x est borélienne car Ω est borélien et Φ_x est continue. On montre de même que les coupes Ω^y sont boréliennes. $\square$

LEMME 3.2. *Si $\Omega \subseteq \mathbb{R}^p \times \mathbb{R}^q$ est borélien, alors les applications $x \mapsto \lambda_q(\Omega_x)$ et $y \mapsto \lambda_p(\Omega^y)$ sont boréliennes, respectivement de $\mathbb{R}^p$ dans $[0, \infty]$ et de $\mathbb{R}^q$ dans $[0, \infty]$.*

Bien que d'allure très innocente, ce lemme n'est *pas du tout* évident. Par souci d'efficacité, on va l'admettre à ce stade. On le démontrera au Chapitre 10, dans un cadre plus général.

Le résultat suivant porte parfois le nom de **Principe de Cavalieri**. Il entraine en particulier que si A et B sont deux parties de $\mathbb{R}^2$ dont les sections horizontales ou verticales ont la même longueur ($\lambda_1(A_x) = \lambda_1(B_x)$ pour tout x, ou bien $\lambda_1(A^y) = \lambda_1(B^y)$ pour tout y), alors A et B ont la même aire ; et que si A et B sont des parties de $\mathbb{R}^3$ dont les coupes parallèles à un certain plan de coordonnées ont la même aire, alors A et B ont le même volume.

PROPOSITION 3.3. *Pour tout borélien $\Omega \subseteq \mathbb{R}^p \times \mathbb{R}^q$, on a*

$$\lambda_{p+q}(\Omega) = \int_{\mathbb{R}^p} \lambda_q(\Omega_x)\, dx = \int_{\mathbb{R}^q} \lambda_p(\Omega^y)\, dy\,.$$

Démonstration. On peut se contenter de montrer l'une des deux égalités ; par exemple que $\lambda_{p+q}(\Omega) = \int_{\mathbb{R}^p} \lambda_q(\Omega_x)\, dx$ pour tout borélien $\Omega \subseteq \mathbb{R}^{p+q}$.

FAIT. La fonction d'ensembles $\mu : \mathcal{B}(\mathbb{R}^{p+q}) \to [0, \infty]$ définie par

$$\mu(A) = \int_{\mathbb{R}^p} \lambda_q(A_x)\, dx$$

est une mesure (borélienne) sur $\mathbb{R}^{p+q}$.

Preuve du Fait. On a $\mu(\emptyset) = 0$ car $\emptyset_x = \emptyset$ pour tout $x \in \mathbb{R}^p$.

Soit $(A_k)_{k\in\mathbb{N}}$ une suite de boréliens de $\mathbb{R}^{p+q}$ deux à deux disjoints. On a pour tout $x \in \mathbb{R}^p$:

$$\left(\bigcup_{k=0}^{\infty} A_k \right)_x = \bigsqcup_{k=0}^{\infty} (A_k)_x\,,$$

où le symbole $\bigsqcup$ signifie "réunion disjointe" ; donc

$$\lambda_q \left[\left(\bigcup_{k=0}^{\infty} A_k \right)_x \right] = \sum_{k=0}^{\infty} \lambda_q\big((A_k)_x\big)\,.$$

On en déduit

$$\begin{aligned} \mu\left(\bigcup_{k=0}^{\infty} A_k \right) &= \int_{\mathbb{R}^p} \left(\sum_{k=0}^{\infty} \lambda_q\big((A_k)_x\big) \right) dx \\ &= \sum_{k=0}^{\infty} \int_{\mathbb{R}^p} \lambda_q\big((A_k)_x\big)\, dx \qquad \text{par interversion série-intégrale} \\ &= \sum_{k=0}^{\infty} \mu(A_k)\,. \end{aligned}$$

$\square$

Il s'agit de montrer qu'en fait $\mu = \lambda_{p+q}$; et il suffit pour cela de vérifier qu'on a $\mu(P) = |P|$ pour tout *pavé* $P \subseteq \mathbb{R}^{p+q}$.

Soit P un pavé quelconque de $\mathbb{R}^{p+q}$, et écrivons $P = A \times B$, où A et B sont des pavés de $\mathbb{R}^p$ et de $\mathbb{R}^q$ respectivement. Alors $P_x = B$ si $x \in A$ et $P_x = \varnothing$ si $x \notin A$; donc

$$\lambda_q(P_x) = \begin{cases} \lambda_q(B) & \text{si } x \in A, \\ 0 & \text{si } x \notin A. \end{cases}$$

On en déduit

$$\begin{aligned} \mu(P) = \int_A \lambda_q(B)\, dx &= \int_A \lambda_q(B)\, d\lambda_p \\ &= \lambda_q(B)\, \lambda_p(A) \\ &= |A|\,|B| = |P|\,. \end{aligned}$$

□

Remarque. La Proposition 2.2 est une conséquence formelle du principe de Cavalieri. En effet, si f est une fonction borélienne $\geqslant 0$ sur $\Omega \subseteq \mathbb{R}^N$ et si $x \in \mathbb{R}^N$, alors $\mathrm{SG}(f,\Omega)_x = \varnothing$ si $x \notin \Omega$, et $\mathrm{SG}(f,\Omega)_x = [0, f(x)[$ si $x \in \Omega$. Donc $\lambda_1\big(\mathrm{SG}(f,\Omega)_x\big) = 0$ si $x \notin \Omega$ et $\lambda_1\big(\mathrm{SG}(f,\Omega)_x\big) = f(x)$ si $x \in \Omega$; et donc, par Cavalieri :

$$\lambda_{N+1}\big(\mathrm{SG}(f,\Omega)\big) = \int_{\mathbb{R}^N} \lambda_1\big(\mathrm{SG}(f,\Omega)_x\big)\, dx = \int_\Omega f(x)\, dx.$$

Exemple 1. Aire d'un disque (!!)

Soit $R > 0$, et soit $\overline{D} \subseteq \mathbb{R}^2$ le disque fermé de centre $0 = (0,0)$ et de rayon R :

$$\overline{D} = \{(x,y) \in \mathbb{R}^2;\ x^2 + y^2 \leqslant R^2\}\,.$$

On va calculer $\lambda_2(\overline{D})$ *en utilisant la Proposition 3.3* . Ici, on est dans $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ donc $p = 1 = q$.

Si $x \in \mathbb{R}$, alors $\overline{D}_x = \varnothing$ si $x \notin [-R,R]$, et $\overline{D}_x = \big[-\sqrt{R^2 - x^2}, \sqrt{R^2 - x^2}\,\big]$ si $x \in [-R,R]$; donc

$$\lambda_1(\overline{D}_x) = \begin{cases} 0 & \text{si } x \notin [-R,R], \\ 2\sqrt{R^2 - x^2} & \text{si } x \in [-R,R]. \end{cases}$$

Par conséquent,

$$\begin{aligned} \lambda_2(\overline{D}) &= \int_{-R}^{R} 2\sqrt{R^2 - x^2}\, dx \\ &= \int_{-1}^{1} 2R\sqrt{1 - u^2}\, R du \\ &= \cdots 2R^2 \times \frac{\pi}{2} \qquad \text{(calcul déjà fait).} \end{aligned}$$

Exemple 2. Volume d'une boule dans $\mathbb{R}^3$.

Soit $R > 0$, et soit B la boule euclidienne fermée de rayon R (centrée en 0) dans $\mathbb{R}^3$:

$$B = \{(x,y,z) \in \mathbb{R}^3;\ x^2 + y^2 + z^2 \leqslant R^2\}\,.$$

On se doute que $\lambda_3(B) = \frac{4}{3}\pi R^3$; mais encore faut-il le montrer proprement. On va le faire grâce à la Proposition 3.3.

On écrit $\mathbb{R}^3 = \mathbb{R}^2 \times \mathbb{R}$, et donc $(x, y, z) = ((x, y), z))$. Si $z \in \mathbb{R}$, alors $B^z = \varnothing$ si $z \notin [-R, R]$, et B^z est un disque de rayon $\sqrt{R^2 - z^2}$ si $z \in [-R, R]$; donc

$$\lambda_2(B^z) = \begin{cases} 0 & \text{si } z \notin [-R, R], \\ \pi(R^2 - z^2) & \text{si } z \in [-R, R]. \end{cases}$$

Par conséquent,

$$\begin{aligned} \lambda_3(B) &= \int_{-R}^{R} \pi(R^2 - z^2)\, dz \\ &= \pi \left(R^2 \times 2R - \left[\frac{z^3}{3} \right]_{-R}^{R} \right) \\ &= \pi \left(2R^3 - \frac{2R^3}{3} \right) \\ &= \frac{4}{3} \pi R^3 . \end{aligned}$$

Exercice 1. Soient $R, h > 0$. Calculer le volume du cône $C \subseteq \mathbb{R}^3$ de hauteur h basé sur le disque de centre 0 et de rayon R dans le plan des xy.

Exercice 2. Montrer que si $f : \mathbb{R} \to \mathbb{R}$ est une fonction borélienne, alors le graphe de f est λ_2-négligeable.

Exercice 3. Calculer le volume d'une boule euclidienne de rayon R dans $\mathbb{R}^4$.

3.2. Le Théorème de Fubini. Dans cette section, p et q sont toujours des entiers strictement positifs. Le point générique de $\mathbb{R}^{p+q} = \mathbb{R}^p \times \mathbb{R}^q$ se note

$$u = (x, y) \qquad \text{avec } x \in \mathbb{R}^p \text{ et } y \in \mathbb{R}^q .$$

LEMME 3.4. *Soit Ω un borélien de $\mathbb{R}^{p+q} = \mathbb{R}^p \times \mathbb{R}^q$. Si $f : \Omega \to [0, \infty]$ est une fonction borélienne, alors les applications $x \mapsto \int_{\Omega_x} f(x, y)\, dy$ et $y \mapsto \int_{\Omega^y} f(x, y)\, dx$ sont boréliennes, respectivement sur $\mathbb{R}^p$ et sur $\mathbb{R}^q$.*

Démonstration. On se concentre uniquement sur la fonction $x \mapsto \int_{\Omega_x} f(x, y)\, dy$, que l'on note Φ_f. Selon une routine désormais familière, on traite d'abord le cas d'une fonction f étagée, et on conclut par convergence monotone.

CAS 1. Cas où la fonction f est étagée.

On écrit $f = \sum_{i=1}^{N} \alpha_i \mathbf{1}_{E_i}$, où $\alpha_i \geqslant 0$ et les E_i sont des boréliens de Ω. Si $x \in \mathbb{R}^p$, alors

$$\begin{aligned} \int_{\Omega_x} f(x, y)\, dy &= \sum_{i=1}^{N} \alpha_i \int_{\Omega_x} \mathbf{1}_{E_i}(x, y)\, dy \\ &= \sum_{i=1}^{N} \alpha_i \int_{\Omega_x} \mathbf{1}_{(E_i)_x}(y)\, d\lambda_q(y) \\ &= \sum_{i=1}^{N} \alpha_i\, \lambda_q\big((E_i)_x\big). \end{aligned}$$

D'après le lemme 3.2, les applications $x \mapsto \lambda_q\big((E_i)_x\big)$ sont boréliennes ; donc Φ_f est borélienne.

Cas 2. Cas général.

Soit (φ_n) une suite approximante pour f. Par convergence monotone, on a

$$\int_{\Omega_x} f(x,y)\,dy = \lim_{n\to\infty} \int_{\Omega_x} \varphi_n(x,y)\,dy \qquad \text{pour tout } x\in\mathbb{R}^p\,;$$

donc Φ_f est borélienne par le Cas 1. □

Voici maintenant le très fameux **théorème de Fubini**.

Théorème 3.5. *Soit Ω un borélien de $\mathbb{R}^{p+q} = \mathbb{R}^p\times\mathbb{R}^q$, et soit f une fonction borélienne sur Ω, positive ou à valeurs complexes.*

(1) ***Cas "positif"** : si f est* positive*, alors*

$$\int_\Omega f\,d\lambda_{p+q} = \int_{\mathbb{R}^p}\left(\int_{\Omega_x} f(x,y)\,dy\right)dx = \int_{\mathbb{R}^q}\left(\int_{\Omega^y} f(x,y)\,dx\right)dy\,.$$

(2) ***Cas "intégrable"** : si f est* intégrable *sur Ω (par rapport à λ_{p+q}), alors* on a le droit d'écrire

$$\int_\Omega f\,d\lambda_{p+q} = \int_{\mathbb{R}^p}\left(\int_{\Omega_x} f(x,y)\,dy\right)dx = \int_{\mathbb{R}^q}\left(\int_{\Omega^y} f(x,y)\,dx\right)dy\,.$$

Remarque. La signification précise de "on a le droit d'écrire" dans (2) est la suivante :

- pour *presque tout* $x\in\mathbb{R}^p$, la fonction $y\mapsto f(x,y)$ est intégrable sur Ω_x par rapport à λ_q ;
- pour *presque tout* $y\in\mathbb{R}^q$, la fonction $x\mapsto f(x,y)$ est intégrable sur Ω^p par rapport à λ_p ;
- Les fonctions presque partout définies

$$\Phi(x) := \int_{\Omega_x} f(x,y)\,dy \qquad \text{et} \qquad \Psi(y) := \int_{\Omega^y} f(x,y)\,dx$$

sont respectivement λ_p-intégrable sur $\mathbb{R}^p$ et λ_q-intégrable sur $\mathbb{R}^q$, et on a

$$\int_{\mathbb{R}^p}\Phi(x)\,dx = \int_\Omega f\,d\lambda_{p+q} = \int_{\mathbb{R}^q}\psi(y)\,dy\,.$$

Corollaire 3.6. *Soient $A\subseteq\mathbb{R}^p$ et $B\subseteq\mathbb{R}^q$ des ensembles boréliens. Si f est une fonction borélienne positive ou intégrable définie sur $A\times B\subseteq\mathbb{R}^{p+q}$, alors*

$$\int_{A\times B} f = \int_A\left(\int_B f(x,y)\,dy\right)dx = \int_B\left(\int_A f(x,y)\,dy\right)dx\,.$$

Démonstration. On applique le théorème avec $\Omega := A\times B$. Comme $\Omega_x = \varnothing$ si $x\notin A$ et $\Omega_x = B$ si $x\in A$, on a $\int_{\Omega_x} f(x,y)\,dy = 0$ si $x\notin A$ et $\int_{\Omega_x} f(x,y)\,dy = \int_B f(x,y)\,dy$ si $x\in A$. Donc $\int_{\mathbb{R}^p}\left(\int_{\Omega_x} f(x,y)\,dy\right)dx = \int_A\left(\int_B f(x,y)\,dy\right)dx$; et de même $\int_{\mathbb{R}^q}\left(\int_{\Omega^y} f(x,y)\,dx\right)dy = \int_B\left(\int_A f(x,y)\,dy\right)dx$. □

Avant de donner la preuve du théorème de Fubini, il est bon d'expliciter son "mode d'emploi" et de donner quelques exemples.

Mode d'emploi de Fubini. Le théorème de Fubini permet *sous certaines hypothèses* de permuter des intégrales.

- Si la fonction f qu'on considère est *positive*, il n'y a rien à vérifier : tous les calculs sont permis.
- Si la fonction f n'est pas positive, *il faut vérifier* qu'elle est intégrable sur le domaine Ω considéré, i.e. qu'on a $\int_\Omega |f| < \infty$. Comme $|f| \geqslant 0$, on peut le faire... en appliquant le théorème de Fubini (cas "positif").

Exemple 1. Calcul de l'intégrale $I := \displaystyle\int_{[0,\infty[\times[0,\infty[} \frac{x}{(1+x^2+y^2)^2}\, dxdy\,.$

La fonction à intégrer est borélienne positive sur $\Omega := [0,\infty[\times[0,\infty[$, donc il n'y a pas d'états d'âme à avoir. On a

$$\begin{aligned} I &= \int_0^\infty \left(\int_0^\infty \frac{x}{(1+x^2+y^2)^2}\, dx\right) dy \\ &= \int_0^\infty \left[-\frac{1}{2}\frac{1}{1+x^2+y^2}\right]_{x=0}^{x=\infty} dy \\ &= \frac{1}{2}\int_0^\infty \frac{dy}{1+y^2} \\ &= \frac{\pi}{4}\cdot \end{aligned}$$

Exemple 2. Calcul de l'intégrale $I := \displaystyle\int_\Omega \frac{x}{y}\, dxdy$, où

$$\Omega := \{(x,y)\in\mathbb{R}^2;\ x\geqslant 1\,, 1\leqslant y\leqslant x^2 \ \text{ et } \ x+y\leqslant 6\}\,.$$

Cet exemple ne présente aucun intérêt, hormis le fait qu'il oblige à *dessiner le domaine* Ω et à identifier les coupes. La fonction à intégrer est positive sur Ω, donc on peut appliquer Fubini sans se poser de questions.

En faisant soigneusement le dessin, on constate qu'on a

$$\Omega_x = \begin{cases} \varnothing & \text{si } x<1 \text{ ou } x>5, \\ [1,x^2] & \text{si } 1\leqslant x\leqslant 2, \\ [1,6-x] & \text{si } 2\leqslant x\leqslant 5. \end{cases}$$

On a donc

$$\begin{aligned} I &= \int_1^2 \left(\int_1^{x^2} \frac{x}{y}\, dy\right) dx + \int_2^5 \left(\int_1^{6-x} \frac{x}{y}\, dy\right) dx \\ &= \int_1^2 x\log(x^2)\, dx + \int_2^5 x\log(6-x)\, dx \\ &= 2\int_1^2 x\log(x)\, dx + \int_2^5 x\log(6-x)\, dx\,. \end{aligned}$$

La première intégrale se calcule par parties en intégrant x et en dérivant $\log(x)$; et la deuxième se calcule elle aussi par parties, ce qui se voit peut-être mieux en écrivant $\int_2^5 x\log(6-x)\, dx = \int_1^4 (6-u)\log(u)\, du$. Au total, on trouve $I = \frac{135}{4}\log(2) - 18$ (sauf très probable erreur de calcul).

Exemple 3. Montrons (à nouveau) qu'on a pour tout $\lambda > 0$

$$\int_0^\infty \frac{\sin x}{x}\, e^{-\lambda x}\, dx = \frac{\pi}{2} - \arctan(\lambda)\,.$$

La fonction $x \mapsto \frac{\sin x}{x} e^{-\lambda x}\, dx$ est intégrable sur $]0,\infty[$ car $\frac{\sin x}{x}$ est borné (par 1), donc $I(\lambda) := \int_0^\infty \frac{\sin x}{x}\, e^{-\lambda x}\, dx$ est bien définie pour tout $\lambda > 0$. Dans la suite, on fixe $\lambda > 0$.

Le *point de départ* est d'écrire que pour tout $x > 0$, on a

$$e^{-\lambda x} = x \int_\lambda^\infty e^{-xy} dy\,.$$

Donc $\frac{\sin x}{x} = \int_\lambda^\infty \sin x\, e^{-xy} dy$, et donc

$$I(\lambda) = \int_0^\infty \left(\int_\lambda^\infty \sin x\, e^{-xy} dy \right) dx\,.$$

Si on admet provisoirement qu'on peut permuter les deux intégrales, on obtient

$$\begin{aligned} I(\lambda) &= \int_\lambda^\infty \left(\int_0^\infty \sin x\, e^{-yx} dx \right) dy \\ &= \int_\lambda^\infty \mathrm{Im} \left(\int_0^\infty e^{(i-y)x} dx \right) dy \\ &= \int_\lambda^\infty \frac{dy}{1+y^2} \qquad \text{(calcul déjà fait)} \\ &= \frac{\pi}{2} - \arctan(\lambda)\,. \end{aligned}$$

Il reste à justifier qu'on avait bien le droit de permuter les deux intégrales. D'après le théorème de Fubini, il suffit de vérifier que la fonction $f(x,y) := \sin x\, e^{-xy}$ est intégrable sur $]0,\infty[\times[\lambda,\infty[$; autrement dit, par "Fubini positif", qu'on a

$$\int_0^\infty \left(\int_\lambda^\infty |\sin x|\, e^{-xy} dy \right) dx < \infty \quad \text{ou encore} \quad \int_\lambda^\infty \left(\int_0^\infty |\sin x|\, e^{-xy} dx \right) dy < \infty$$

Il ne faut surtout pas majorer $|\sin x|$ par 1 (vérifier que ça ne mène à rien). Pour éviter cette tentation, on va intégrer d'abord en y, puis en x : on obtient

$$\begin{aligned} \int_0^\infty \left(\int_\lambda^\infty |\sin x|\, e^{-xy} dy \right) dx &= \int_0^\infty |\sin x| \left(\int_\lambda^\infty e^{-yx} dy \right) dx \\ &= \int_0^\infty \frac{|\sin x|}{x}\, e^{-\lambda x} dx \\ &\leqslant \int_0^\infty e^{-\lambda x} dx < \infty\,. \end{aligned}$$

Il est temps à présent de donner la

Preuve du théorème de Fubini. Soit f une fonction borélienne sur $\Omega \subseteq \mathbb{R}^p \times \mathbb{R}^q$, positive ou intégrable par rapport à λ_{p+q}.

(1) Supposons $f \geqslant 0$. Par le principe de Cavalieri, on a

$$\int_\Omega f\, d\lambda_{p+q} = \lambda_{p+q+1}\big(\mathrm{SG}(f,\Omega)\big) = \int_{\mathbb{R}^p} \lambda_{q+1}\big(\mathrm{SG}(f,\Omega)_x\big)\, dx.$$

De plus, on vérifie sans difficulté que si $x \in \mathbb{R}^p$, alors

$$\mathrm{SG}(f,\Omega)_x = \mathrm{SG}(f_x,\Omega_x),$$

où $f_x : \Omega_x \to [0,\infty]$ est la fonction définie par $f_x(y) = f(x,y)$.

Donc, à nouveau par Cavalieri,

$$\begin{aligned}\int_\Omega f\,d\lambda_{p+q} &= \int_{\mathbb{R}^p} \lambda_{q+1}\big(\mathrm{SG}(f_x,\Omega_x)\big)\,dx\\ &= \int_{\mathbb{R}^p}\left(\int_{\Omega_x} f_x\,d\lambda_q\right)dx\\ &= \int_{\mathbb{R}^p}\left(\int_{\Omega_x} f(x,y)\,dy\right)dx.\end{aligned}$$

On montre de même que $\int_{\mathbb{R}^q}\big(\int_{\Omega^y} f(x,y)\,dx\big)\,dy = \int_\Omega f\,d\lambda_{p+q}$.

(2) Supposons f intégrable.

Par linéarité, on peut supposer que la fonction f est à valeurs réelles, ce qui va permettre d'appliquer le cas "positif" aux fonctions f^+ et f^-. On doit montrer que pour presque tout $x\in\mathbb{R}^p$, la fonction $y\mapsto f(x,y)$ est intégrable sur Ω_x, et que la fonction presque partout définie $\Phi(x) := \int_{\Omega_x} f(x,y)\,dy$ est λ_p-intégrable sur $\mathbb{R}^p$, avec $\int_{\mathbb{R}^p}\Phi(x)\,dx = \int_\Omega f$. (La deuxième "intégrale itérée" se traite évidemment de la même façon.)

Par (1), on a

$$\int_{\mathbb{R}^p}\left(\int_{\Omega_x}|f(x,y)|\,dy\right)dx = \int_\Omega |f|\,d\lambda_{p+q} < \infty\,.$$

Autrement dit, si on pose $\Theta(x) := \int_{\Omega_x}|f(x,y)|\,dy$, alors la fonction (*borélienne*) positive Θ vérifie $\int_{\mathbb{R}^p}\Theta(x)\,dx<\infty$. En particulier, on a $\Theta(x)<\infty$ presque partout ; donc

$$\Phi(x) := \int_{\Omega_x} f(x,y)\,dy \quad \text{est bien défini presque partout}\,.$$

De façon précise, Φ est bien définie en tout point de l'ensemble borélien

$$E := \{x\in\mathbb{R}^p;\ \Theta(x)<\infty\};$$

et pour tout $x\in E$, on a

$$\Phi(x) = \int_{\Omega_x} f^+(x,y)\,dy - \int_{\Omega_x} f^-(x,y)\,dy\,.$$

Cette formule et la Proposition 3.4 montrent que la fonction Φ est borélienne sur E. On la prolonge en une fonction borélienne $\widetilde{\Phi}$ sur $\mathbb{R}^p$ tout entier en posant $\widetilde{\Phi}(x)=0$ pour $x\notin E$. La fonction $\widetilde{\Phi}$ vérifie $|\widetilde{\Phi}(x)|\leqslant\Theta(x)$ pour tout $x\in\mathbb{R}^p$ (c'est évident si $x\notin E$, et si $x\in E$ on a $|\widetilde{\Phi}(x)| = |\Phi(x)| = \left|\int_{\Omega_x} f(x,y)\,dy\right| \leqslant \int_{\Omega_x}|f(x,y)|\,dy = \Theta(x)$). Donc $\widetilde{\Phi}$ est intégrable sur $\mathbb{R}^p$, et par conséquent Φ est λ_p-intégrable.

Il reste à montrer qu'on a

$$\int_{\mathbb{R}^p}\Phi(x)\,dx = \int_\Omega f\,;$$

ce qui se fait en appliquant (1) aux fonctions positives et intégrables f^+ et f^- et en se souvenant que $\lambda_p(\mathbb{R}^p \backslash E) = 0$:

$$\begin{aligned}
\int_{\mathbb{R}^p} \Phi(x)\, dx &= \int_E \Phi(x)\, dx \\
&= \int_E \left(\int_{\Omega_x} f^+(x,y)\, dy - \int_{\Omega_x} f^-(x,y)\, dy \right) dx \\
&= \int_E \left(\int_{\Omega_x} f^+(x,y)\, dy \right) dx - \int_E \left(\int_{\Omega_x} f^-(x,y)\, dy \right) dx \\
&= \int_{\mathbb{R}^p} \left(\int_{\Omega_x} f^+(x,y)\, dy \right) dx - \int_{\mathbb{R}^p} \left(\int_{\Omega_x} f^-(x,y)\, dy \right) dx \\
&= \int_\Omega f^+ - \int_\Omega f^- \qquad \text{par (1)} \\
&= \int_\Omega f\,.
\end{aligned}$$

□

Exercice. Soit $f :]0,1[\times]0,1[$ la fonction définie par

$$f(x,y) := \frac{x^2 - y^2}{(x^2+y^2)^2}\,.$$

Vérifier qu'on a

$$f(x,y) = \frac{\partial}{\partial y}\left(\frac{y}{x^2+y^2}\right) = -\frac{\partial}{\partial x}\left(\frac{x}{x^2+y^2}\right),$$

puis montrer que les deux intégrales itérées $\int_0^1 \left(\int_0^1 f(x,y)\, dy\right) dx$ et $\int_0^1 \left(\int_0^1 f(x,y)\, dx\right) dy$ existent mais *ne sont pas égales.* A-t-on pour autant découvert une contradiction dans les mathématiques ?

4. Changements de variables

4.1. Ré-écriture de la formule de changement de variable dans $\mathbb{R}$. La formule de changement de variable "générale" démontrée au Chapitre 5 peut se ré-écrire de la façon suivante.

PROPOSITION 4.1. *Soit $\phi : I \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ et strictement monotone sur un intervalle $I \subseteq \mathbb{R}$. Pour toute fonction borélienne f positive ou intégrable sur $\phi(I)$, on a*

$$\int_{\phi(I)} f(x)\, dx = \int_I f(\phi(u))\, |\phi'(u)| du\,. \tag{4.1}$$

Démonstration. Écrivons $I = (a,b)$.
Si ϕ est croissante, alors $\phi(I) = (\phi(a), \phi(b))$ et $\phi' \geqslant 0$; donc

$$\int_I f(\phi(u))\, |\phi'(u)| du = \int_a^b f(\phi(u))\, \phi'(u) du = \int_{\phi(a)}^{\phi(b)} f(x)\, dx = \int_{\phi(I)} f(x)\, dx.$$

Si ϕ est décroissante, alors $\phi(I) = (\phi(b), \phi(a))$ et $\phi' \leqslant 0$; donc

$$\int_I f(\phi(u))\, |\phi'(u)| du = \int_a^b f(\phi(u)) \times (-\phi'(u)) du = -\int_{\phi(a)}^{\phi(b)} f(x)\, dx = \int_{\phi(I)} f(x)\, dx\,.$$

$\square$

L'intérêt de la ré-écriture est qu'il est très facile de donner un sens à (4.1) pour des fonctions de *plusieurs* variables : il suffit de remplacer $\phi'(u)$ par le *déterminant jacobien* $J_\phi(u)$ pour obtenir un "candidat formule de changement de variable" dans $\mathbb{R}^N$.

4.2. Énoncé de la formule de changement de variables dans $\mathbb{R}^N$. Rappelons qu'un **difféomorphisme** entre deux ouverts $\Omega, \Omega' \subseteq \mathbb{R}^N$ est une bijection $\Phi : \Omega \to \Omega'$ telle que Φ et Φ^{-1} sont de classe $\mathcal{C}^1$. Par exemple, si I est un intervalle ouvert de $\mathbb{R}$ et si $\phi : I \to \mathbb{R}$ est une fonction de classe $\mathcal{C}^1$ telle que ϕ' ne s'annule jamais, alors ϕ est un difféomorphisme de I sur l'intervalle ouvert $I' := \phi(I)$.

Si $\Phi : \Omega \to \mathbb{R}^N$ est une application de classe $\mathcal{C}^1$ sur un ouvert $\Omega \subseteq \mathbb{R}^N$, on note $J_\Phi(u)$ le **déterminant jacobien** de Φ en un point $u \in \Omega$, i.e. le déterminant de la matrice jacobienne de Φ au point u. En écrivant $\Phi(u) = (\Phi_1(u), \dots, \Phi_N(u))$, on a donc

$$J_\Phi(u) = \det\left(\partial_j \Phi_i(u)\right)_{1 \leqslant i,j \leqslant N} .$$

Le fait suivant, très utile dans la pratique, est une conséquence du *théorème d'inversion locale.*

Fait. Une bijection $\Phi : \Omega \to \Omega'$ entre deux ouverts de $\mathbb{R}^N$ est un difféomorphisme si et seulement si Φ est de classe $\mathcal{C}^1$ avec $J_\Phi(u) \neq 0$ pour tout $u \in \Omega$. Plus généralement, si $\Phi : \Omega \to \mathbb{R}^N$ est de classe $\mathcal{C}^1$ et injective, avec $J_\Phi(u) \neq 0$ pour tout $u \in \Omega$, alors $\Omega' := \Phi(\Omega)$ est un ouvert de $\mathbb{R}^N$ et Φ est un difféomorphisme de Ω sur Ω'.

Théorème 4.2. (formule de changement de variables)
Soit $\Phi : \Omega \to \Omega'$ un difféomorphisme entre deux ouverts de $\mathbb{R}^N$. Pour tout borélien $A \subseteq \Omega$ et pour toute fonction borélienne f positive ou intégrable *sur $\Phi(A)$, on a*

$$\int_{\Phi(A)} f(x)\, dx = \int_A f(\Phi(u))\, |J_\Phi(u)|\, du .$$

Remarque 1. L'ensemble $\Phi(A)$ est bien un borélien de $\mathbb{R}^N$ car $\Phi(A) = \left(\Phi^{-1}\right)^{-1}(A)$ et l'application Φ^{-1} est continue (donc borélienne).

Remarque 2. Dans le cas "intégrable", l'énoncé précis est le suivant : si f est intégrable sur $\Phi(A)$, alors $f(\Phi(u))|J_\Phi(u)|$ est intégrable sur A et la formule est vraie.

Corollaire 4.3. *Si $\Phi : \Omega \to \Omega'$ est un difféomorphisme entre deux ouverts de $\mathbb{R}^N$, alors on a pour tout borélien $A \subseteq \Omega$:*

$$\lambda_N\big(\Phi(A)\big) = \int_A |J_\Phi(u)|\, du .$$

Démonstration. On applique le théorème avec $f := \mathbf{1}$:

$$\lambda_N\big(\Phi(A)\big) = \int_{\Phi(A)} \mathbf{1}(x)\, dx = \int_A \mathbf{1}(\Phi(u))\, |J_\Phi(u)| du = \int_A |J_\phi(u)|\, du .$$

$\square$

Mode d'emploi de la formule. Supposons qu'on veuille calculer une intégrale

$$I = \int_{\mathcal{D}} f(x_1, \dots, x_N)\, dx_1 \cdots dx_N ,$$

où $\mathcal{D}$ est un certain "domaine" dans $\mathbb{R}^N$. On peut procéder comme suit.

(i) On *devine* (peu importe comment !) qu'il peut être avantageux d'introduire de nouvelles variables $u_1, \dots, u_N$.

(ii) On exprime $x = (x_1, \dots, x_N)$ en fonction de $u = (u_1, \dots, u_N)$, et on vérifie qu'on a bien un difféomorphisme $x = \Phi(u)$ défini sur un ouvert Ω tel que $\Phi(\Omega) \supseteq \mathcal{D}$.

(iii) On écrit $x = \Phi(u)$ et $dx = |J_\Phi(u)|du$, i.e.

$$dx_1 \cdots dx_N = |J_\Phi(u_1, \dots, u_N)|du_1 \cdots du_N \, .$$

(iv) On trouve le domaine $\widetilde{\mathcal{D}}$ tel que $\Phi(\widetilde{\mathcal{D}}) = \mathcal{D}$.

(v) On écrit la formule de changement de variables

$$\int_{\mathcal{D}} f(x_1, \dots, x_N)\, dx_1 \cdots dx_N = \int_{\widetilde{\mathcal{D}}} f(u_1, \dots, u_N)\, |J_\Phi(u_1, \dots, u_N)| du_1 \cdots du_N \, ,$$

en n'oubliant pas de remplacer $\mathcal{D}$ par $\widetilde{\mathcal{D}}$ et de mettre un valeur absolue autour du déterminant jacobien.

Exemple 1. Soit $\mathcal{D} = \{(x,y) \in \mathbb{R}^2;\ x+1 \leqslant y \leqslant x+2 \text{ et } -1 \leqslant x+2y \leqslant 2\}$. On veut calculer l'intégrale $I := \int_{\mathcal{D}} xy\, dxdy$.

On devine qu'il peut être utile d'introduire les variables

$$u = y - x \quad \text{et} \quad v = x + 2y \, .$$

La raison est assez claire : par définition de $\mathcal{D}$,

$$(x,y) \in \mathcal{D} \iff (u,v) \in \widetilde{\mathcal{D}} := [1,2] \times [-1,2] \, ,$$

et le domaine $\widetilde{\mathcal{D}}$ est visiblement "plus simple" que $\mathcal{D}$.

Pour exprimer (x,y) en fonction de (u,v), on inverse le système

$$\begin{cases} u &= & y - x \\ v &= & x + 2y \end{cases},$$

ce qui donne

$$(x,y) = \left(\frac{v-2u}{3}, \frac{u+v}{3} \right) := \Phi(u,v) \, .$$

Il est évident que Φ est un difféomorphisme (linéaire) de $\mathbb{R}^2$ sur $\mathbb{R}^2$ (l'inverse est $\Phi^{-1}(x,y) = (u,v) = (y-x, x+2y)$!) ; et on a

$$J_\Phi(u,v) = \det \begin{pmatrix} -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} \end{pmatrix} = -\frac{1}{3} \qquad \text{pour tout } (u,v) \in \mathbb{R}^2.$$

D'après la formule de changement de variables, on a donc

$$\begin{aligned} \int_{\mathcal{D}} xy\, dxdy &= \int_{[1,2]\times[-1,2]} \left(\frac{v-2}{3} \right) \left(\frac{u+v}{3} \right) \times \frac{1}{3} dudv \\ &= \frac{1}{27} \int_1^2 \left(\int_{-1}^2 (v^2 - uv - 2u^2)\, dv \right) du \qquad \text{par Fubini.} \end{aligned}$$

Le calcul ne présente alors plus aucune difficulté. On obtient finalement $I = -\frac{29}{108}$ (sauf erreur de calcul).

Exercice. Calculer directement I à l'aide du théorème de Fubini, sans utiliser la formule de changement de variables.

Exemple 2. Calcul de l'intégrale

$$I := \int_{]0,1[\times]0,1[} \frac{dxdy}{1-x^2y^2}\cdot$$

Soit $\Omega = \{(u,v)\in\mathbb{R}^2;\ u>0\,,\ v>0\,,\ u+v<\pi/2\}$ et soit $\Phi = \Omega \to \mathbb{R}^2$ l'application définie par

$$\Phi(u,v) = \left(\frac{\sin u}{\cos v}, \frac{\sin v}{\cos u}\right)\,.$$

Il n'est pas trop difficile de montrer (c'est un bon exercice) que Φ est une bijection de Ω sur le carré $]0,1[\times]0,1[$, dont la réciproque est donnée par

$$\Phi^{-1}(x,y) = \left(\arccos\sqrt{\frac{1-x^2}{1-x^2y^2}}\ ,\arccos\sqrt{\frac{1-y^2}{1-x^2y^2}}\right)\,.$$

(En revanche, ne surtout pas demander comment on a pu deviner ce changement de variables...)

Les formules montrent que Φ Φ et Φ^{-1} sont de classe $\mathcal{C}^1$, *i.e.* Φ est un difféomorphisme ; et pour $(u,v)\in\Omega$, on a (en posant $(x,y)=\Phi(u,v)$)

$$\begin{aligned} J_\Phi(u,v) &= \det\begin{pmatrix} \frac{\cos u}{\cos v} & \frac{\sin u \sin v}{\cos^2 v} \\ \frac{\sin v \sin u}{\cos^2 u} & \frac{\cos v}{\cos u}\end{pmatrix} \\ &= 1-\tan^2 u\tan^2 v \\ &= 1-x^2y^2\,. \end{aligned}$$

D'après la formule de changement de variables, on a donc

$$I = \int_{\Phi(\Omega)} \frac{dxdy}{1-x^2y^2} = \int_\Omega dudv = \lambda_2(\Omega)\,;$$

autrement dit

$$I = \frac{\pi^2}{8}\cdot$$

Exercice. Utiliser l'exemple précédent pour calculer $\sum_{k=0}^\infty \frac{1}{(2k+1)^2}$; puis calculer $\sum_{n=1}^\infty \frac{1}{n^2}$.

Remarque 4.4. Bien entendu (comme en dimension 1), on peut aussi utiliser la formule de changement de variables "en sens inverse" : on doit au départ calculer une intégrale $\int_\Omega g(u_1,\dots,u_N)\,du_1\cdots du_N$, et on introduit de nouvelles variables $x_1,\dots,x_N$ telles que $(x_1,\dots,x_N)=\Phi(u_1,\dots,u_N)$ pour une certaine fonction Φ. Il faut alors exprimer $(u_1,\dots,u_N)$ en fonction de $(x_1,\dots,x_N)$, autrement dit déterminer le difféomorphisme Φ^{-1}, et appliquer la formule de changement de variable à Φ^{-1}. En pratique, cela veut dire : se débrouiller pour tout exprimer en fonction de $x_1,\dots,x_N$ et $dx_1,\dots,dx_N$ dans $g(u_1,\dots,u_N)du_1\cdots du_N$.

Exercice. Calculer l'intégrale $I := \int_{\{u>v>0\}}(u^4-v^4)e^{-(u+v)^2}dudv$ en posant $(x,y) = (u^2+v^2, 2uv)$.

4.3. Intégration en coordonnées polaires. Le cas particulier suivant de la formule de changement de variables est d'une très grande utilité.

PROPOSITION 4.5. *Soit $\mathcal{D}$ un borélien de $\mathbb{R}^2$, et soit $p = (x_0, y_0) \in \mathbb{R}^2$. Pour toute fonction borélienne f positive ou intégrable sur $\mathcal{D}$, on a*

$$\int_{\mathcal{D}} f(x,y)\, dxdy = \int_{\widetilde{\mathcal{D}}} f\big(x_0 + r\cos\theta, y_0 + r\sin\theta\big)\, rdrd\theta\,,$$

où $\widetilde{\mathcal{D}} := \big\{(r,\theta) \in \mathbb{R}^2;\ r > 0\ ,\ 0 < \theta < 2\pi \text{ et } (x_0 + r\cos\theta, y_0 + r\sin\theta) \in \mathcal{D}\big\}$.

Démonstration. Soit $\Phi :]0,\infty[\times\,]0,2\pi[\to \mathbb{R}^2$ l'application définie par

$$\Phi(r,\theta) := (x_0 + r\cos\theta, y_0 + r\sin\theta)\,.$$

Il est clair que Φ est une bijection de classe $\mathcal{C}^1$ de $\Omega :=]0,\infty[\times\,]0,2\pi[$ sur l'ensemble de tous les $(x,y) \in \mathbb{R}^2$ s'écrivant $(x_0 + r\cos\theta, y_0 + r\sin\theta)$ avec $r > 0$ et $\theta \neq 0\ [2\pi]$; autrement dit sur $\Omega' := \mathbb{R}^2 \backslash \Delta$, où Δ est la demi-droite d'origine (x_0, y_0) parallèle à l'axe des abscisses et dirigée vers la droite. En particulier, Ω' est un ouvert de $\mathbb{R}^2$.

On a

$$J_\Phi(r,\theta) = \det\begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} = r > 0 \qquad \text{pour tout } (r,\theta) \in \Omega.$$

Donc Φ est un difféomorphisme ; et comme $\Phi\big(\widetilde{\mathcal{D}}\big) = \mathcal{D} \cap \Omega'$ par définition de $\widetilde{\mathcal{D}}$, la formule de changement de variables s'écrit

$$\int_{\mathcal{D}\cap\Omega'} f(x,y)\, dxdy = \int_{\widetilde{\mathcal{D}}} f\big(x_0 + r\cos\theta, y_0 + r\sin\theta\big)\, rdrd\theta\,.$$

Enfin, $\mathcal{D}\backslash\Omega'$ est *négligeable* pour la mesure de Lebesgue λ_2, car il est contenu dans la demi-droite Δ. Donc on peut écrire $\int_{\mathcal{D}}$ au lieu de $\int_{\mathcal{D}\cap\Omega'}$. $\square$

Remarque. La seule chose à retenir pour intégrer correctement en coordonnées polaires est qu'on a

$$dxdy = rdrd\theta\,.$$

Si de plus le "centre d'intégration" est $(x_0, y_0) = (0,0)$, alors il faut aussi se souvenir qu'on a

$$r^2 = x^2 + y^2\,.$$

En général, c'est ce qui doit alerter : si on voit apparaitre x^2+y^2 dans une intégrale, on *doit penser* au coordonnées polaires.

EXEMPLE 4.6. Comme application de la formule d'intégration en coordonnées polaires, on va calculer l'**intégrale de Gauss**

$$I := \int_{-\infty}^{\infty} e^{-t^2/2} dt\,.$$

L'astuce est de calculer ... I^2. On a

$$\begin{aligned}
I^2 &= \left(\int_{-\infty}^{\infty} e^{-x^2/2}dx\right) \times \left(\int_{-\infty}^{\infty} e^{-y^2/2}dy\right) \\
&= \int_{\mathbb{R}^2} e^{-\frac{1}{2}(x^2+y^2)}dxdy \qquad \text{par Fubini “positif”} \\
&= \int_{]0,\infty[\times]0,2\pi[} e^{-r^2/2}rdrd\theta \\
&= \int_0^{2\pi} \left(\int_0^{\infty} re^{-r^2/2}\,dr\right) d\theta \\
&= \int_0^{2\pi} \left[-e^{-r^2/2}\right]_{r=0}^{r=\infty} d\theta \\
&= 2\pi\,.
\end{aligned}$$

Et ainsi :

$$I = \sqrt{2\pi}\,.$$

4.4. Volumes et transformations affines. Rappelons qu'une application $\Phi : \mathbb{R}^N \to \mathbb{R}^N$ est dite **affine** si elle est de la forme

$$\Phi(u) = a + L(u)\,,$$

où $a \in \mathbb{R}^N$ est fixé et L est une application linéaire. (On dit que L est l'**application linéaire associée à** Φ.) De manière équivalente, Φ est affine si elle est de la forme $\Phi(u) = a + Mu$, où M est une matrice $N \times N$. Par exemple, les *translations* sont des applications affines ($L = Id$), et toute application *linéaire* est affine ($a = 0$). De même, les projections et les symétries (pas forcément orthogonales) sont des applications affines.

Le **déterminant** d'une application affine Φ est le déterminant de l'application linéaire associée : si $\Phi(u) = a + L(u) = a + Mu$, alors

$$\det(\Phi) = \det(L) = \det(M)\,.$$

Une application affine $\Phi(u) = a + L(u) = a + Mu$ est un difféomorphisme si et seulement l'application linéaire associée est inversible. De plus, la matrice jacobienne de Φ en tout point $u \in \mathbb{R}^N$ est égale à M, et donc

$$J_\Phi(u) = \det(\Phi)\,.$$

D'après la formule de changement de variables, on en déduit immédiatement le résultat suivant.

Proposition 4.7. *Si $\Phi : \mathbb{R}^N \to \mathbb{R}^N$ est un difféomorphisme affine, alors*

$$\lambda_N(\Phi(A)) = |\det(\Phi)| \times \lambda_N(A) \qquad \textit{pour tout borélien } A \subseteq \mathbb{R}^N\,. \tag{4.2}$$

Autrement dit, Φ multiplie tous les volumes par le facteur constant $|\det(\Phi)|$.

Démonstration. On applique le Corollaire 4.3 :

$$\lambda_N(\Phi(A)) = \int_A |J_\Phi(u)|\,d\lambda_N(u) = |\det(\Phi)| \times \lambda_N(A)\,.$$

□

Vu l'importance de ce résultat, on va maintenant en donner une peuve “directe” n'utilisant la formule de changement de variables générale.

Preuve directe de la Proposition 4.7. Comme la mesure de Lebesgue est invariante par translations, il suffit de démontrer (4.2) pour toute transforamtion *linéaire* Φ. Autrement dit, il s'agit de montrer que pour toute matrice inversible $M \in GL_N(\mathbb{R})$ et pour tout borélien $A \subseteq \mathbb{R}^N$, on a

$$\lambda_N(M \cdot A) = |\det(M)|\, \lambda_N(A), \tag{4.3}$$

où $M \cdot A = \{Mu;\ u \in A\}$.

On va pour cela utiliser des matrices particulières : les matrices dites **de transvection**. Par définition, une matrice de transvection est une matrice T de la forme

$$T = Id + \alpha\, E_{i,j},$$

où $\alpha \in \mathbb{R}$, $i \neq j$ et $E_{i,j}$ est la matrice dont tous les coefficients sont nuls sauf celui d'indice (i, j) qui vaut 1. Notons qu'une telle matrice T est de déterminant 1. L'intérêt de ces matrices tient au fait suivant.

Fait 1. Toute matrice $M \in GL_N(\mathbb{R})$ de déterminant 1 est produit de matrices de transvections. Par conséquent, toute matrice $M \in GL_N(\mathbb{R})$ est produit de matrices de transvactions et d'une matrice diagonale.

Preuve du Fait 1. Le premier point se démontre sans trop de difficultés en utilisant des opérations élémentaires sur les lignes et les colonnes ; *cf* n'importe quel livre d'algèbre linéaire. Pour le deuxième point, il suffit d'observer que toute matrice $M \in GL_N(\mathbb{R})$ peut s'écrire $M = DS$, où D est diagonale et $\det(S) = 1$; par exemple, on peut prendre pour D la matrice dont le premier coefficient diagonal vaut $1/\det(M)$ et tous les autres valent 1 (une matrice ce type est ce qu'on appelle une **matrice de dilatation**). □

Fait 2. La formule (4.3) est vraie pour toute matrice de transvection.

Preuve du Fait 2. Soit T une matrice de transvection. Comme $\det(T) = 1$, on doit montrer qu'on a $\lambda_N(T \cdot A) = \lambda_N(A)$ pour tout borélien $A \subseteq \mathbb{R}^d$. Si on pose $\mu(A) = \lambda_N(T \cdot A)$, on voit sans difficulté que μ est une mesure borélienne sur $\mathbb{R}^N$. Donc, il s'agit de montrer qu'on a

$$\lambda_N(T \cdot P) = |P| \qquad \text{pour tout pavé } P \subseteq \mathbb{R}^N.$$

Pour simplifier les écritures, on va supposer que ($N \geqslant 2$ et)

$$T = Id + \alpha E_{1,2},$$

où $\alpha \in \mathbb{R}$. Si $u = (x, y, u_3 \dots, u_N) \in \mathbb{R}^N$, on a donc

$$Tu = \begin{pmatrix} x + \alpha y \\ y \\ \vdots \\ u_N \end{pmatrix}$$

En écrivant $P = I \times J \times Q$ où I, J sont des intervalles et Q est un pavé de $\mathbb{R}^{N-2}$, on en déduit que

$$T \cdot P = A \times Q,$$

où

$$A = \{(x + \alpha y, y);\ (x, y) \in I \times J\} \subseteq \mathbb{R}^2.$$

D'après le principe de Cavalieri, on a

$$\begin{aligned}\lambda_2(A) &= \int_J \lambda_1(A^y)\,dy \\ &= \int_J \lambda_1(I+\alpha y)\,dy \\ &= |I|\,|J|,\end{aligned}$$

car $\lambda_1(I+\alpha y) = \lambda_1(I) = |I|$ pour tout $y \in \mathbb{R}$. Par conséquent,

$$\lambda_N(T\cdot P) = \lambda_2(A)\,\lambda_{N-2}(Q) = |I|\,|J|\,|Q| = |P|.$$

□

Fait 3. La formule (4.3) est vraie pour toute matrice diagonale.

Preuve du Fait 3. On procède comme pour le Fait 2 en introduisant la mesure μ définie par $\mu(A) = \mu(D\cdot A)$, où D est la matrice diagonale considérée. Les détails constituent un bon exercice. □

Fait 4. Si la formule (4.3) est vraie pour deux matrices M_1, M_2, alors elle est vraie pour $M = M_1M_2$.

Preuve du Fait 4. Si $A \subseteq \mathbb{R}^N$, alors

$$\lambda_N(M\cdot A) = \lambda_N\big(M_1\cdot(M_2\cdot A)\big) = \det(M_1)\,\lambda_N(M_2\cdot A) = \det(M_1)\det(M_2)\,\lambda_N(A);$$

d'où le résultat puisque $\det(M_1)\det(M_2) = \det(M)$. □

La Proposition 4.7 découle immédiatement des faits précédents : si $M \in GL_N(\mathbb{R})$, on écrit $M = DT_1\cdots T_K$ où D est diagonale et les T_k sont des matrices de transvection ; la formule (4.3) est vraie pour D et chacune des T_k par les Faits 2 et 3, donc elle est vraie pour M d'après le Fait 4. □

Exemple 1. Si $ABCD$ est un parallèlogramme dans $\mathbb{R}^2$, alors l'aire de $ABCD$ est égale à $|\det(\overrightarrow{AB},\overrightarrow{AD})|$.

Démonstration. Un point M appartient au parallèlogramme $ABCD$ si et seulement si on peut écrire

$$\overrightarrow{AM} = \lambda\,\overrightarrow{AB} + \mu\,\overrightarrow{AD}$$

avec $0 \leqslant \lambda \leqslant 1$ et $0 \leqslant \mu \leqslant 1$.

En d'autres termes, le parallèlogramme $ABCD$ est l'image du pavé $[0,1]\times[0,1]$ par l'application $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ définie par

$$\Phi(\lambda,\mu) := A + \lambda\,\overrightarrow{AB} + \mu\,\overrightarrow{AD}\,.$$

L'application Φ est un difféomorphisme affine de $\mathbb{R}^2$ sur $\mathbb{R}^2$ dont l'application linéaire associée envoie $e_1 = (1,0)$ sur $\overrightarrow{AB}$ et $e_2 = (0,1)$ sur $\overrightarrow{AD}$. On a donc

$$\det(\Phi) = \det(\overrightarrow{AB},\overrightarrow{AD})\,;$$

et par conséquent

$$\lambda_2(ABCD) = |\det(\overrightarrow{AB},\overrightarrow{AD})| \times \lambda_2([0,1]\times[0,1]) = |\det(\overrightarrow{AB},\overrightarrow{AD})|\,.$$

□

Exemple 2. Aire d'une ellipse "pleine".

Soient $a, b > 0$. On veut calculer l'aire du domaine $\mathcal{E}$ entouré par une ellipse Γ de "demi-axes" a et b. On suppose que Γ est centrée en $(0,0)$ et que ses axes de symétrie sont les axes de coordonnées, de sorte que

$$\mathcal{E} = \left\{(x,y) \in \mathbb{R}^2;\ \frac{x^2}{a^2} + \frac{y^2}{b^2} \leqslant 1\right\}.$$

On voit tout de suite qu'il est avantageux de poser $(u,v) = \left(\frac{x}{a}, \frac{y}{b}\right)$, car

$$(x,y) \in \mathcal{E} \iff u^2 + v^2 \leqslant 1\,.$$

Le domaine $\mathcal{E}$ correspond donc au disque $\widetilde{\mathcal{E}}$ de centre $(0,0)$ et de rayon 1, qui est clairement "plus simple" que $\mathcal{E}$.

On a $(x,y) = (au, bv) = \Phi(u,v)$. L'application Φ est un difféomorphisme linéaire, avec

$$\det(\Phi) = \det\begin{pmatrix} a & 0 \\ 0 & b\end{pmatrix} = ab\,.$$

Par conséquent, $\lambda_2(\mathcal{E}) = ab \times \lambda_2(\widetilde{\mathcal{E}})$, autrement dit

$$\lambda_2(\mathcal{E}) = \pi\, ab\,.$$

4.4.1. *Invariance par isométrie.* Le résultat suivant est une propriété fondamentale de la mesure de Lebesgue. Rappelons qu'une application $\Phi : \mathbb{R}^N \to \mathbb{R}^N$ est une **isométrie euclidienne** si elle préserve la distance euclidienne :

$$\|\Phi(v) - \Phi(u)\|_2 = \|v - u\|_2 \qquad \text{pour tous } u, v \in \mathbb{R}^N\,,$$

où $\|\cdot\|_2$ est la norme euclidienne sur $\mathbb{R}^N$. Par exemple, les translations, les rotations et les symétries orthogonales sont des isométries.

THÉORÈME 4.8. *La mesure de Lebesgue est **invariante par isométries** : si Φ est une isométrie euclidienne de $\mathbb{R}^N$, alors $\lambda_N(\Phi(A)) = \lambda_N(A)$ pour tout borélien $A \subseteq \mathbb{R}^N$.*

Pour la preuve, on a besoin d'un lemme intéressant en lui même.

LEMME 4.9. *Toute isométrie euclidienne $\Phi : \mathbb{R}^N \to \mathbb{R}^N$ est affine.*

Démonstration. On va utiliser une voie qui n'est pas la plus courte possible, mais qui a l'avantage d'être praticable avec d'autres normes que la norme euclidienne.

FAIT 1. Si $\Phi : \mathbb{R}^N \to \mathbb{R}^N$ est une isométrie, alors Φ *conserve les milieux* : on a

$$\Phi\left(\frac{u+v}{2}\right) = \frac{\Phi(u) + \Phi(v)}{2} \qquad \text{pour tous } u, v \in \mathbb{R}^N\,.$$

Preuve du Fait 1. De façon générale, le milieu m d'un segment $[x,y] \subseteq \mathbb{R}^N$ est *caractérisé* par les égalités purement métriques

$$\|m - x\| = \frac{1}{2}\,\|y - x\| = \|y - m\|\,.$$

Maintenant, si $u, v \in \mathbb{R}^N$ et si on pose $m' := \Phi\left(\frac{u+v}{2}\right)$, alors

$$\|m' - \Phi(u)\| = \left\|\Phi\left(\frac{u+v}{2}\right) - \Phi(u)\right\| = \left\|\frac{u+v}{2} - u\right\| = \frac{1}{2}\,\|v - u\|\,,$$

et de même

$$\|\Phi(v) - m'\| = \frac{1}{2}\,\|v - u\|\,;$$

donc $m' = \Phi\left(\frac{u+v}{2}\right)$ est le milieu de $[\Phi(u), \Phi(v)]$. □

Fait 2. Si $f : \mathbb{R}^N \to \mathbb{R}^N$ est *continue* et conserve les milieux, alors f est affine.

Démonstration. Quitte à remplacer f par $f - f(0)$ (qui vérifie les mêmes hypothèses que f), on peut supposer que $f(0) = 0$. Dans ce cas, on doit montrer que f est linéaire.

L'hypothèse faite sur f est qu'on a

$$f\left(\frac{u+v}{2}\right) = \frac{f(u)+f(v)}{2} \qquad \text{pour tous } u, v \in \mathbb{R}^N.$$

En prenant $u := x \in \mathbb{R}^N$ et $v := 0$ et comme $f(0) = 0$, on obtient $f(x/2) = f(x)/2$ pour tout $x \in \mathbb{R}^N$. Donc $f\left(\frac{u+v}{2}\right) = \frac{1}{2} f(u+v)$, et on obtient ainsi

$$f(u+v) = f(u) + f(v) \qquad \text{pour tous } u, v \in \mathbb{R}^N.$$

Autrement dit, l'application f est *additive*.

Il reste alors à montrer qu'on a $f(\lambda u) = \lambda f(u)$ pour tout $u \in \mathbb{R}^N$ et pour tout scalaire $\lambda \in \mathbb{R}$; ce qui est un exercice très classique. En utilisant l'additivité, on montre d'abord que cela est vrai pour $\lambda = n \in \mathbb{N}$ (récurrence immédiate). Comme de plus $f(-x) = -f(x)$ pour tout $x \in \mathbb{R}^N$ (car $0 = f(0) = f(x + (-x)) = f(x) + f(-x)$), on en déduit qu'on a $f(nu)$ pour tout $n \in \mathbb{Z}$. De là, on montre que $f(ru) = rf(u)$ pour tout *rationnel* $r = \frac{p}{q}$ en écrivant $pf(u) = f(pu) = f(qru) = qf(ru)$. Enfin, on utilise la *continuité* de f pour montrer que l'identité $f(\lambda u) = \lambda f(u)$ est vraie pour tout $\lambda \in \mathbb{R}$. □

La preuve du lemme est maintenant terminée : si Φ est une isométrie de $\mathbb{R}^N$, alors Φ conserve les milieux par le Fait 1, donc Φ est affine par le Fait 2 car l'isométrie Φ est (évidemment) continue. □

Preuve du Théorème 4.8. Soit $\Phi : \mathbb{R}^N \to \mathbb{R}^N$ une isométrie euclidienne. Par le Lemme 4.9, Φ est affine et donc de la forme $\Phi(u) = a + Mu$ pour une certaine matrice M. La matrice M vérifie $\|Mu\|_2 = \|\Phi(u) - \Phi(0)\|_2 = \|u\|_2$ pour tout $u \in \mathbb{R}^N$; donc M est une matrice *orthogonale*, et donc $\det(\Phi) = \det(M) = \pm 1$. D'où le résultat par la Proposition 4.7. □

Remarque. Soit E un sous-espace vectoriel de $\mathbb{R}^N$, de dimension $d \geqslant 1$. En choisissant une base orthonormée $\mathbf{e} = (e_1, \dots, e_d)$ de E, on peut identifier E à $\mathbb{R}^d$, et donc définir une "mesure de Lebesgue associée à la base $\mathbf{e}$" sur E. Le théorème 4.8 montre que cette mesure ne dépend en fait pas de la base $\mathbf{e}$. On peut donc la noter λ_E, et l'appeler *la mesure de Lebesgue sur E*. Plus généralement, on peut définir une mesure de Lebesgue sur tout sous-espace *affine* de $\mathbb{R}^N$ (i.e. un translaté d'un sous-espace vectoriel de $\mathbb{R}^N$).

4.5. Un peu plus sur l'invariance par translations. On sait depuis le Chapitre 2 que la mesure de Lebesgue λ_N est invariante par translations, et on vient de voir qu'elle est en fait invariante par toutes les isométries de $\mathbb{R}^N$. On va maintenant montrer que l'invariance par translations *caractérise à elle seule* la mesure de Lebesgue, à une constante de normalisation près. (En particulier, l'invariance par translations entraine l'invariance par n'importe quelle isométrie, ce qui n'a rien d'évident *a priori*.)

Proposition 4.10. *Si μ est une mesure borélienne sur $\mathbb{R}^N$ invariante par translations et telle que $\mu([0,1[^N) < \infty$, alors μ est multiple de la mesure de Lebesgue : il existe une constante $C < \infty$ telle que $\mu(A) = C\,\lambda_N(A)$ pour tout borélien $A \subseteq \mathbb{R}^N$.*

Démonstration. On va montrer que $\mu = C\,\lambda_N$, avec $C = \mu([0,1[^N)$. D'après le lemme d'unicité (Lemme 5.1 du Chapitre 2), il suffit de prouver qu'on a $\mu(P) = C\,\lambda_N(P)$ pour tout pavé $P \subseteq \mathbb{R}^N$.

On commence par le cas où P est un *pavé rationnel semi-ouvert*, c'est-à-dire un pavé semi-ouvert $P = [a_1, b_1[\times \cdots \times [a_N, b_N[$ où les a_i et les b_i sont rationnels. En réduisant tout au même dénominateur, on voit qu'on peut trouver un entier $K \geqslant 1$ et des entiers p_i, q_i tels que $a_i = \frac{p_i}{K}$ et $b_i = \frac{q_i}{K}$ pour tout $i \in \{1, \dots, N\}$.

Chaque intervalle $[a_i, b_i[$ est ainsi la réunion de $m_i := q_i - p_i$ intervalles semi-ouverts de longueur $1/K$. En faisant un dessin (dans le plan), on se convainc alors sans peine que le pavé P est réunion de $m := m_1 \cdots m_N = (q_1 - p_1) \cdots (q_N - p_N)$ translatés deux-à-deux disjoints du pavé $[0, \frac{1}{K}[^N$. Comme μ est une mesure invariante par translations, on a donc $\mu(P) = m \times \mu([0, \frac{1}{K}[^N)$.

Par ailleurs, le pavé "unité" $[0,1[^N$ est réunion de K^N translatés deux à deux disjoints du pavé $[0, \frac{1}{K}[^N$. Donc $\mu([0,1[^N) = K^N \times \mu\big([0, \frac{1}{K}[^N\big)$; autrement dit : $\mu\big([0, \frac{1}{K}[^N\big) = \frac{C}{K^N}\cdot$

On obtient ainsi

$$\begin{aligned} \mu(P) &= m \times \frac{C}{K^N} \\ &= C \times \frac{(q_1 - p_1) \cdots (q_N - p_N)}{K^N} \\ &= C\,|P|\,, \end{aligned}$$

ce qui est la conclusion souhaitée.

On passe ensuite au cas des pavés *compacts*. Comme tout intervalle fermé borné $[a, b] \subseteq \mathbb{R}$ est l'intersection d'une suite décroissante d'intervalles semi-ouverts $[a_n, b_n[$ à extrémités rationnelles (exercice), on voit que tout pavé compact P est l'intersection d'une suite décroissante de pavés rationnels semi-ouverts $(P_n)_{n\in\mathbb{N}}$. Comme μ et λ_N sont des mesures et comme $\mu(P_0) = \lambda_N(P_0) < \infty$, on a $\mu(P) = \lim_{n\to\infty} \mu(P_n)$ et $\lambda_N(P) = \lim_{n\to\infty} \lambda_N(P_n)$; donc $\mu(P) = C\,\lambda_N(P)$ pour tout pavé compact P, d'après le cas précédent.

Pour le cas général, on utilise le fait qu'un pavé P quelconque est réunion d'une suite croissante de pavés compact et on applique le cas précédent. (En réalité, cette 3ème étape n'est pas nécessaire par le lemme d'unicité tel qu'on l'a énoncé au Chapitre 2.) □

Exercice. Utiliser la Proposition 4.10 pour redémontrer le fait que la mesure de Lebesgue est invariante par isométries. (Se ramener au cas d'une isométrie *linéaire* Φ, et considérer la fonction d'ensembles μ définie par $\mu(A) = \lambda_N(\Phi(A))$.)

4.6. Preuve de la formule de changement de variables. Pour démontrer la formule de changement de variables, il suffit (comme d'habitude) de traiter le cas des fonctions *positives*.

Pour abréger, on écrira (CV) au lieu de "formule de changement de variables", et on dira que (CV) *est vraie pour un difféomorphisme* $\Phi : \Omega \to \Omega'$ si on a

$$\int_{\Phi(A)} f(x)\,dx = \int_A f(\Phi(u))\,|J_\Phi(u)|\,du$$

pour tout borélien $A \subseteq \Omega$ et pour toute fonction borélienne $f : \Phi(A) \to [0, \infty]$.

Notre but est de montrer que (CV) est vraie pour *tout* difféomorphisme $\Phi : \Omega \to \Omega'$ entre deux ouverts quelconques $\Omega, \Omega' \subseteq \mathbb{R}^N$, quel que soit l'entier $N \geqslant 1$. On va le faire en procédant *par récurrence sur la dimension* N.

Le "lemme simplificateur" suivant sera très utile.

Lemme 4.11. *Pour un difféomorphisme $\Phi : \Omega \to \Omega'$, les propriétés suivantes sont équivalentes :*

(1) (CV) *est vraie pour* Φ *;*

(2) *pour tout borélien* $A \subseteq \Omega$,
$$\lambda_N(\Phi(A)) = \int_A |J_\Phi(u)|\, du\,;$$

(3) *pour tout point* $a \in \Omega$, *on peut trouver un voisinage ouvert* U_a *de* a (*avec* $U_a \subseteq \Omega$) *tel que*
$$\lambda_N(\Phi(A)) = \int_A |J_\Phi(u)|\, du \quad \textit{pour tout borélien } A \subseteq U_a\,.$$

Démonstration. L'implication (1) $\implies$ (2) est évidente : il suffit de prendre $f = \mathbf{1}$ dans (CV) ; et l'implication (2) $\implies$ (3) est encore plus évidente (si (2) est vraie, on peut prendre $U_a := \Omega$ pour tout $a \in \Omega$). Quant à l'implication (2) $\implies$ (1), elle a essentiellement déjà été prouvée quand on a démontré la formule de changement de variable "générale" en dimension $N = 1$ (*cf* Chapitre 5) : si (2) est vraie, alors (CV) est vraie pour les fonctions étagées positives par "linéarité", et donc pour toutes les fonctions boréliennes positives par convergence monotone.

Il reste à montre que (3) entraine (2) ; supposons donc (3) vérifiée. Pour tout $a \in \Omega$, choisissons un voisinage ouvert U_a de a comme dans (3). Alors $\Omega = \bigcup_{a\in\Omega} U_a$; donc, par la *propriété de Lindelöf*, on peut trouver un ensemble dénombrable $\Omega_0 \subseteq \Omega$ tel que $\bigcup_{a\in\Omega_0} U_a = \Omega$. Autrement dit, on peut trouver une suite d'ouverts $(U_k)_{k\in\mathbb{N}}$ tels que
$$\bigcup_{k\in\mathbb{N}} U_k = \Omega \quad \text{et} \quad \forall k \in \mathbb{N}\ \forall A \subseteq\ U_k\ :\ \lambda_N(\Phi(A)) = \int_A |J_\Phi(u)|\, du\,.$$

Soit alors A un borélien quelconque de Ω. Posons (un peu comme d'habitude) $A_0 := U_0$ et
$$A_k := U_k \backslash (U_0 \cup \cdots \cup U_{k-1}) \qquad \text{pour } k \geqslant 1\,.$$

Les A_k sont des boréliens deux à deux disjoints, et $\bigcup_{k\in\mathbb{N}} A_k = A$. Comme Φ est injective, $\Phi(A)$ est donc la réunion *disjointe* des $\Phi(A_k)$:
$$\Phi(A) = \bigsqcup_{k=0}^{\infty} \Phi(A_k)\,.$$

On en déduit

$$\begin{aligned}
\lambda_N(A) &= \sum_{k=0}^{\infty} \Phi(A_k) \\
&= \sum_{k=0}^{\infty} \int_{A_k} |J_\Phi(u)|\, du \qquad \text{car } A_k \subseteq U_k \text{ pour tout } k \\
&= \int_{\bigcup_{k=0}^{\infty} A_k} |J_\Phi(u)|\, du \qquad \text{car les } A_k \text{ sont disjoints} \\
&= \int_A |J_\Phi(u)|\, du\,.
\end{aligned}$$

Donc (2) est vérifiée. □

Dans la preuve de (CV), on va considérer des difféomorphismes de deux types particuliers.

Définition 4.12. *Soient V et V' deux ouverts de $\mathbb{R}^N$, $N \geqslant 1$.*

(1) *On dira qu'un difféomorphisme $\Sigma : V \to V'$ est une* ***permutation de coordonnées*** *si Σ est de la forme $\Sigma(x_1, \dots, x_N) = (x_{\sigma(1)}, \dots, x_{\sigma(N)})$, où σ est une permutation de l'ensemble $\{1, \dots, N\}$.*

(2) *Étant donné $j \in \{1, \dots, N\}$, on dira qu'un difféomorphisme $\Phi : V \to V'$* ***préserve la j-ième coordoonnée*** *si, pour tout $v \in V$, la j-ième coordonnée de $\Phi(v)$ est égale à celle de v ; autrement dit, si Φ est de la forme*

$$\Phi(v_1, \dots, v_j, \dots, v_N) = \big(\Phi_1(v_1, \dots, v_N), \dots, v_j, \dots, \Phi_N(v_1, \dots, v_N)\big)\,.$$

Par exemple, le difféomorphisme $\Sigma : \mathbb{R}^3 \to \mathbb{R}^3$ défini par $\Sigma(x, y, z) = (y, z, x)$ est une permutation de coordonnés ; et le difféomorphisme $\Phi : \mathbb{R}^3 \to \mathbb{R}^3$ défini par $\Phi(x, y, z) = (x, x + y, y + z)$ préserve la 1ère coordonnée.

Le fait suivant dit que ces difféomorphismes particuliers suffisent à "reconstruire localement" *tous* les difféomorphismes.

Fait 1. Soit $\Phi : \Omega \to \Omega'$ un difféomorphisme entre deux ouverts *de* $\mathbb{R}^{N+1}$. Pour tout $a \in \Omega$, on peut trouver un voisinage ouvert U_a de a tel que

$$\forall u \in U_a \;:\; \Phi(u) = \Phi^1 \circ \Phi^2 \circ \Sigma(u)\,,$$

où Σ est une permutation de coordonnées, Φ^2 est un difféomorphisme préservant les N premières coordonnées (défini sur $\Sigma(U_a)$), et Φ^1 est un difféomorphisme préservant la $(N + 1)$-ième coordonnée (défini sur $\Phi^1 \circ \Sigma(U_a)$).

Démonstration. Écrivons $\Phi(u) = \begin{pmatrix} \Phi_1(u) \\ \vdots \\ \Phi_{N+1}(u) \end{pmatrix}$ et fixons un point $a \in \Omega$. Comme Φ est un difféomorphisme, la matrice jacobienne $\mathrm{Jac}_\Phi(a)$ est inversible, et en particulier sa dernière ligne n'est pas identiquement nulle. On peut donc trouver $j \in \{1, \dots, N+1\}$ tel que $\frac{\partial \Phi_{N+1}}{\partial u_j}(a) \neq 0$. Par continuité de $\frac{\partial \Phi_{N+1}}{\partial u_j}$, on peut alors trouver un pavé ouvert U_a contenant a tel que $\frac{\partial \Phi_{N+1}}{\partial u_j}(u) \neq 0$ pour tout $u \in U_a$.

Notons σ la permutation de $\{1, \dots, N + 1\}$ qui échange j et $N + 1$, et Σ la permutation de coordonnées associée. Posons également $V := \Sigma(U_a)$ et définissons une application $\Psi : V \to \mathbb{R}^{N+1}$ par $\Psi = \Phi \circ \Sigma^{-1}$. Si on écrit $\Psi = \begin{pmatrix} \Psi_1 \\ \vdots \\ \Psi_{N+1} \end{pmatrix}$, alors on a

$$\forall v \in V \ : \ \frac{\partial \Psi_{N+1}}{\partial v_{N+1}}(v) \neq 0\,,$$

par définition de σ. De plus, Ψ est un difféomorphisme de V sur $\Psi(V)$.

Soit maintenant $\Phi^2 : V \to \mathbb{R}^{N+1}$ l'application définie par

$$\Phi^2(v) = \begin{pmatrix} v_1 \\ \vdots \\ v_N \\ \Psi_{N+1}(v) \end{pmatrix}$$

L'application Φ^2 est de classe $\mathcal{C}^1$, et il n'est pas difficile de voir que Φ^2 est injective sur V. En effet, $V = \Sigma(U_a)$ est un pavé ouvert car Σ est une permutation de coordonnés. Écrivons $V = \Pi \times I$, où Π est un pavé de $\mathbb{R}^N$ et I est un intervalle ouvert. Pour $(v_1, \dots, v_n) \in \Pi$ fixé, la fonction $t \mapsto \frac{\partial \Psi_{N+1}}{\partial v_{N+1}}(v_1, \dots, v_N, t)$ ne s'annule pas sur I par le choix de V. Donc la fonction $t \mapsto \Psi_{N+1}(v_1, \dots, v_N, t)$ est injective sur I d'après le théorème des accroissements finis ; et vu la forme de Φ^2, on en déduit immédiatement que Φ^2 est injective sur V. De plus, par définition de Φ^2 on a $J_{\Phi^2}(v) = \frac{\partial \Psi_{N+1}}{\partial v_{N+1}}(v)$, et donc $J_{\Phi^2}(v) \neq 0$ pour tout $v \in V$. Par conséquent, $V' := \Phi^2(V)$ est un ouvert et Φ^2 est un difféomorphisme de V sur V' ; et bien sûr Φ^2 préserve les N premières coordonnées.

Notons Φ^1 le difféomorphisme $\Psi \circ (\Phi^2)^{-1}$ (défini sur V'). Par définition, on a $\Phi^1(\Phi^2(v)) = \Psi(v)$ pour tout $v = (v_1, \dots, v_{N+1}) \in V$, ce qui s'écrit

$$\Phi^1(v_1, \dots, v_N, \Psi_{N+1}(v)) = (\Psi_1(v), \dots, \Psi_N(v), \Psi_{N+1}(v))\,.$$

Ainsi, on voit que pour tout $v' = \Phi^2(v) \in V'$, la $(N+1)$-ième coordonnée de $\Phi^1(v')$ est égale à celle de v' ; autrement dit, Φ^1 préserve la $(N+1)$-ième coordonnée.

Au total, on a obtenu $\Phi \circ \Sigma^{-1} = \Psi = \Phi^1 \circ \Phi^2$ sur $V = \Sigma(U)$, i.e. $\Phi = \Phi^1 \circ \Phi^2 \circ \Sigma$ sur U, où Φ^1, Φ^2 et Σ ont les propriétés requises. □

Pour ce qui nous occupe, l'intérêt du résultat de "décomposition" qu'on vient de démontrer tient au fait suivant.

Fait 2. Si (CV) est vraie pour $\Phi^1 : \Omega \to \Omega'$ et $\Phi^2 : \Omega' \to \Omega''$, alors elle est vraie pour $\Phi := \Phi^2 \circ \Phi^1$.

Démonstration. Soit A un borélien quelconque de Ω. Si (CV) est vraie pour Φ^1 et Φ^2, on peut écrire

$$\begin{aligned} \lambda_N(\Phi(A)) &= \lambda_N(\Phi^2(\Phi^1(A))) \\ &= \int_{\Phi^1(A)} |J_{\Phi^2}(v)|\, dv \\ &= \int_A |J_{\Phi^2}(\Phi^1(u))|\, |J_{\Phi^1}(u)|\, du\,. \end{aligned}$$

De plus, comme la matrice jacobienne de $\Phi = \Phi^2 \circ \Phi^1$ en un point $u \in \Omega$ est donnée par $\mathrm{Jac}_\Phi(u) = \mathrm{Jac}_{\Phi^2}(\Phi^1(u))\, \mathrm{Jac}_{\Phi^1}(u)$, on a

$$J_{\Phi^2}(\Phi^1(u))\, J_{\Phi^1}(u) = J_\Phi(u)\,;$$

d'où finalement

$$\lambda_N(\Phi(A)) = \int_A |J_\Phi(u)|\, du\,.$$

Ainsi, (CV) est vraie pour Φ d'après le Lemme 4.11. □

Les Faits 1 et 2 disent essentiellement que pour démontrer (CV) en général (*i.e.* pour un difféomorphisme Φ quelconque), il suffit de le faire pour les permutations de coordonnées et pour les difféomorphismes préservant certaines coordonnées. Le cas des permutations de coordonnées est "direct" :

Fait 3. (CV) est vraie pour toute permutation de coordonnées.

Démonstration. Soit $\Sigma : V \to V'$ une permutation de coordonnées, et soit σ la permutation de $\{1,\dots,N\}$ associée. Le difféomorphisme Σ est la restriction à V de l'application linéaire $\widetilde{\Sigma} : \mathbb{R}^N \to \mathbb{R}^N$ définie par $\widetilde{\Sigma}(e_j) = e_{\sigma^{-1}(j)}$ pour $j = 1,\dots,N$, où $(e_1,\dots,e_N)$ est la base canonique de $\mathbb{R}^N$. La matrice de $\widetilde{\Sigma}$ est une "matrice de permutation" ; donc $\det(\widetilde{\Sigma}) = \pm 1$, et donc $|J_\Sigma(u)| = |\det(\widetilde{\Sigma})| = 1$ pour tout $u \in V$. D'après le Lemme 4.11, il suffit donc vérifier qu'on a $\lambda_N(\Sigma(A)) = \lambda_N(A)$ pour tout borélien $A \subseteq V$; et comme l'application $A \mapsto \lambda_N(\Sigma(A))$ est une mesure borélienne sur V (exercice), il suffit de vérifier qu'on a $\lambda_N(\Sigma(P)) = |P|$ pour tout pavé $P \subseteq V$. Mais dans ce cas tout est clair : si $P \subseteq V$ est un pavé, $P = I_1 \times \cdots \times I_N$, alors $\Sigma(P)$ est le pavé $P_\sigma := I_{\sigma(1)} \times \cdots \times I_{\sigma(N)}$, qui a visiblement le même volume que P. □

On peut maintenant donner la preuve de la formule de changement de variables, en procédant comme il a été dit par récurrence sur la dimension N.

Le cas $N = 1$ est une conséquence très facile de la Proposition 4.1. Soit $\Phi : \Omega \to \Omega'$ un difféomorphisme entre deux ouverts de $\mathbb{R}$. Pour tout point $a \in \Omega$, on peut trouver un intervalle ouvert I tel que $a \in I \subseteq \Omega$. Alors Φ est strictement monotone sur I. Pour tout borélien $A \subseteq I$, on a d'après la Proposition 4.1 :

$$\lambda_1(\Phi(A)) = \int_{\Phi(I)} \mathbf{1}_{\Phi(A)}(x)\, dx = \int_I \mathbf{1}_{\Phi(A)}(\Phi(u))\, |\Phi'(u)|\, du = \int_A |\Phi'(u)|\, du\,.$$

Donc (CV) est vraie pour Φ d'après le Lemme 4.11.

Pour le passage de N à $N+1$, supposons avoir établi que (CV) est vraie en dimension N, et fixons un difféomorphisme $\Phi : \Omega \to \Omega'$ entre deux ouvert de $\mathbb{R}^{N+1}$.

D'après le Lemme 4.11, il suffit de vérifier que pour tout point $a \in \Omega$, on peut trouver un voisinage ouvert U_a de a tel que

$$\lambda_{N+1}(A) = \int_A |J_\Phi(u)|\, du \qquad \text{pour tout borélien } A \subseteq U_a. \tag{4.4}$$

On fixe donc un point $a \in \Omega$.

Par le Fait 1, on peut trouver un voisinage ouvert U_a de a, noté simplement U dans la suite, sur lequel le difféomorphisme Φ peut se "décomposer" sous la forme $\Phi = \Phi^2 \circ \Phi^1 \circ \Sigma$, où Σ est une permutation de coordonnées, Φ^1 est un difféomorphisme préservant la première coordonnée, et Φ^2 est un difféomorphisme préservant la dernière coordonnée. Compte tenu des Faits 2 et 3, on voit donc qu'il suffit d'établir (4.4) sous l'hypothèse additionnelle que Φ *préserve la première ou la dernière coordonnée.*

Supposons par exemple que Φ préserve la dernière coordonnée (le calcul serait le même si Φ préservait la première coordonnée). En écrivant le "point générique" de

$\mathbb{R}^{N+1}$ sous la forme $u = (v, t)$ où $v \in \mathbb{R}^N$ et $t \in \mathbb{R}$, on a alors

$$\Phi(u) = \Phi(v, t) = (\Psi(v, t), t) \tag{4.5}$$

pour tout $u = (v, t) \in U$, où Ψ est une application de classe $\mathcal{C}^1$ de U dans $\mathbb{R}^N$.

Rappelons la notation suivante : pour $A \subseteq \mathbb{R}^{N+1}$ et $t \in \mathbb{R}$, on pose

$$A^t := \{v \in \mathbb{R}^N;\ (v, t) \in A\}\,.$$

Pour tout $t \in \mathbb{R}$ tel que $U^t \neq \varnothing$, on notera $\Psi_t : U^t \to \mathbb{R}^N$ l'application définie par

$$\Psi_t(v) = \Psi(v, t)\,.$$

Vu la forme de Φ, on voit immédiatement que l'application Ψ_t est injective, et qu'on a

$$J_{\Psi_t}(v) = J_\Phi(v, t)$$

pour tout $v \in U^t$. En particulier J_{Ψ_t} ne s'annule pas sur V, de sorte que $\Psi_t(U^t)$ est un ouvert de $\mathbb{R}^N$ et Ψ_t est un difféomorphisme de U^t sur $\Psi_t(U^t)$.

Soit maintenant A un borélien quelconque de U. D'après le principe de Cavalieri (Proposition 3.3), on a

$$\lambda_{N+1}(\Phi(A)) = \int_{\mathbb{R}} \lambda_N(\Phi(A)^t)\, dt\,.$$

De plus, par (4.5) on peut écrire

$$\begin{aligned}
\Phi(A)^t &= \{v' \in \mathbb{R}^N;\ \exists (v, s) \in A\ :\ \Phi(v, s) = (v', t)\}\\
&= \{v' \in \mathbb{R}^N;\ \exists (v, s) \in A\ :\ (\Psi(v, s), s) = (v', t)\}\\
&= \{v' \in \mathbb{R}^N;\ \exists v \in A^t\ :\ v' = \Psi(v, t)\}\\
&= \Psi_t(A^t)\,.
\end{aligned}$$

On en déduit

$$\begin{aligned}
\lambda_{N+1}(\Phi(A)) &= \int_{\mathbb{R}} \lambda_N(\Psi_t(A^t))\, dt\\
&= \int_{\mathbb{R}} \left(\int_{A^t} |J_{\Psi_t}(v)|\, dv\right) dt\,,
\end{aligned}$$

où on a appliqué l'hypothèse de récurrence aux difféomorphismes Ψ_t.

Comme $J_{\Psi_t}(v) = J_\Phi(v, t)$ pour tout $(v, t) \in U$, on obtient donc finalement

$$\begin{aligned}
\lambda_{N+1}(\Phi(A)) &= \int_{\mathbb{R}} \left(\int_{A^t} |J_\Phi(v, t)|\, dv\right) dt\\
&= \int_A |J_\Phi(u)|\, du \quad \text{par Fubini}\,.
\end{aligned}$$

Ceci achève la récurrence et la preuve de la formule de changement de variables.

Chapitre 8

Espaces L^p

Dans tout le chapitre, $(\Omega, \mathcal{B}, \mu)$ est un espace mesuré abstrait. Les expressions "intégrable" et "presque partout" signifieront toujours "intégrable par rapport à μ" et "μ-presque partout", respectivement.

1. Semi-normes $\|\cdot\|_p$ et espaces $\mathcal{L}^p$

1.1. Le cas $1 \leqslant p < \infty$. Dans ce qui suit, on fixe un nombre p tel que $1 \leqslant p < \infty$.

Définition 1.1. *Pour toute fonction mesurable $f : \Omega \to \mathbb{C}$ ou $[0, \infty]$, on pose*

$$\|f\|_p := \left(\int_\Omega |f(t)|^p \, d\mu(t) \right)^{1/p},$$

avec la convention $\infty^{1/p} = \infty$. On pose aussi

$$\begin{aligned} \mathcal{L}^p(\Omega, \mathcal{B}, \mu) &:= \{ f : \Omega \to \mathbb{C} \text{ mesurable; } \|f\|_p < \infty \} \\ &= \left\{ f : \Omega \to \mathbb{C} \text{ mesurable; } \int_\Omega |f|^p \, d\mu < \infty \right\}. \end{aligned}$$

Remarque 1. Pour abréger, on écrira souvent $\mathcal{L}^p$ au lieu de $\mathcal{L}^p(\Omega, \mathcal{B}, \mu)$. Mais on pourra aussi écrire $\|\cdot\|_{\mathcal{L}^p(\Omega,\mathcal{B},\mu)}$ au lieu de $\|\cdot\|_p$.

Remarque 2. Par définition, si $f : \Omega \to \mathbb{C}$ est mesurable, alors

$$f \in \mathcal{L}^p \iff |f|^p \in \mathcal{L}^1.$$

Remarque 3. Évidemment, "$f \in \mathcal{L}^1$" signifie que f est *intégrable* sur Ω.

Cas particuliers. (1) Si Ω est un borélien de $\mathbb{R}^N$, la notation $\mathcal{L}^p(\Omega)$ signifiera toujours $\mathcal{L}^p(\Omega, \mathcal{B}(\Omega), \lambda_N)$.

(2) Pour tout ensemble $I \neq \varnothing$, l'espace $\mathcal{L}^p(I, \mathcal{P}(I), \mu_c)$ se notera $\ell^p(I)$. Pour toute fonction $f : I \to \mathbb{C}$ ou $[0, \infty]$, on a

$$\|f\|_p = \left(\sum_{i \in I} |f(i)|^p \right)^{1/p},$$

et donc

$$\ell^p(I) = \left\{ f : I \to \mathbb{C};\ \sum_{i \in I} |f(i)|^p < \infty \right\}.$$

(3) Si $I = \{1, \dots, N\}$, alors $\ell^p(I)$ s'identifie à $\mathbb{C}^N$; et pour $x = (x_1, \dots, x_N) \in \mathbb{C}^N$ on a

$$\|x\|_p = \left(\sum_{i=1}^N |x_i|^p \right)^{1/p};$$

(4) si $I = \mathbb{N}$, alors $\ell^p(\mathbb{N})$ est l'ensemble de toutes les *suites* $x = (x_i)_{i\in\mathbb{N}} \in \mathbb{C}^{\mathbb{N}}$ vérifiant
$$\sum_{i=0}^{\infty} |x_i|^p < \infty\,.$$

DÉFINITION 1.2. *Soit X un espace vectoriel sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Une* ***semi-norme*** *sur X est une application $x \mapsto \|x\|$ de X dans $\mathbb{R}$ vérifiant*

- $\|x\| \geqslant 0$ *pour tout* $x \in X$ *;*
- $\|\lambda x\| = |\lambda|\,\|x\|$ *pour* $x \in X$ *et* $\lambda \in \mathbb{K}$ *;*
- $\|x + y\| \leqslant \|x\| + \|y\|$ *pour* $x, y \in X$.

Autrement dit, une semi-norme est une norme à laquelle on n'impose pas que $\|x\| = 0$ seulement pour $x = 0$.

PROPOSITION 1.3. *$\mathcal{L}^p = \mathcal{L}^p(\Omega, \mathcal{B}, \mu)$ est un espace vectoriel et $\|\cdot\|_p$ est une semi-norme sur $\mathcal{L}^p$. De plus, pour $f \in \mathcal{L}^p$ on a l'équivalence $\|f\|_p = 0 \iff f(t) = 0$ presque partout.*

Démonstration. (i) Comme $\|f\|_p^p = \int_\Omega |f|^p\,d\mu$ et $|f|^p \geqslant 0$, on a $\|f\|_p = 0$ si et seulement si $|f(t)|^p = 0$ presque partout, *i.e.* $f(t) = 0$ presque partout.

(ii) Il est évident qu'on a $\|f\|_p \geqslant 0$ pour toute $f \in \mathcal{L}^p$, et aussi que $\|\lambda f\|_p = |\lambda|\,\|f\|_p$ pour tout scalaire $\lambda \in \mathbb{C}$:
$$\|\lambda f\|_p = \left(\int_\Omega |\lambda f(t)|^p\,d\mu(t)\right)^{1/p} = \left(|\lambda|^p \int_\Omega |f|^p\,d\mu\right)^{1/p} = |\lambda|\,\|f\|_p\,.$$

(iii) Pour l'inégalité triangulaire $\|f+g\|_p \leqslant \|f\|_p + \|g\|_p$, on a besoin du fait suivant.

FAIT. L'ensemble $B := \{u \in \mathcal{L}^p;\ \|u\|_p \leqslant 1\}$ est *convexe* : si $u, v \in B$ et $\lambda \in [0,1]$, alors $(1-\lambda)u + \lambda v \in B$.

Preuve du Fait. Le point clé est que la fonction $z \mapsto |z|^p$ est convexe sur $\mathbb{C}$, car $z \mapsto |z|$ est convexe et $x \mapsto x^p$ est convexe croissante sur $\mathbb{R}^+$ (ne pas oublier que $p \geqslant 1$).

Si $u, v \in B$ et $\lambda \in [0,1]$, alors on a par convexité de $z \mapsto |z|^p$:
$$\big|(1-\lambda)u(t) + \lambda v(t)\big|^p \leqslant (1-\lambda)\,|u(t)|^p + \lambda\,|v(t)|^p \quad \text{pour tout } t \in \Omega\,;$$
d'où en intégrant :
$$\begin{aligned}\int_\Omega \big|(1-\lambda)u(t) + \lambda v(t)\big|^p\,d\mu(t) &\leqslant (1-\lambda)\int_\Omega |u(t)|^p\,d\mu(t) + \lambda\int_\Omega |v(t)|^p\,d\mu(t)\\ &\leqslant (1-\lambda) + \lambda \qquad \text{car } u, v \in B\,.\end{aligned}$$

Ainsi $\big\|(1-\lambda)u + \lambda v\big\|_p^p \leqslant 1$ et donc $(1-\lambda)u + \lambda v \in B$. □

Soient maintenant $f, g \in \mathcal{L}^p$: on veut montrer que $\|f+g\|_p \leqslant \|f\|_p + \|g\|_p$.

Si $\|f\|_p = 0$, alors $f = 0$ presque partout, donc $f + g = g$ presque partout et donc $\|f+g\|_p = \|g\|_p \leqslant \|f\|_p + \|g\|_p$. De même si $\|g\|_p = 0$.

Supposons $\|f\|_p > 0$ et $\|g\|_p > 0$. Par (ii), il s'agit de montrer que
$$\left\|\frac{f+g}{\|f\|_p + \|g\|_p}\right\|_p \leqslant 1\,.$$

Les fonctions
$$u := \frac{f}{\|f\|_p} \qquad \text{et} \qquad v := \frac{g}{\|g\|_p}$$

sont bien définies, et par (ii) on a $\|u\|_p = 1 = \|v\|_p = 1$. De plus,

$$\begin{aligned} \frac{f+g}{\|f\|_p + \|g\|_p} &= \frac{\|f\|_p}{\|f\|_p + \|g\|_p}\, u + \frac{\|g\|_p}{\|f\|_p + \|g\|_p}\, v \\ &= (1-\lambda)u + \lambda v \qquad \text{avec } \lambda := \tfrac{\|g\|_p}{\|f\|_p + \|g\|_p} \in [0,1]\,. \end{aligned}$$

Par le Fait, on en déduit le résultat souhaité :

$$\left\| \frac{f+g}{\|f\|_p + \|g\|_p} \right\|_p \leqslant 1\,.$$

(iv) Maintenant que l'inégalité triangulaire est établie, il est évident que si $f, g \in \mathcal{L}^p$ alors $f + g \in \mathcal{L}^p$; et comme $\lambda f \in \mathcal{L}^p$ pour tout $\lambda \in \mathbb{C}$ si $f \in \mathcal{L}^p$ (par (ii)), on en conclut que $\mathcal{L}^p$ est un espace vectoriel. □

Remarque 1. L'inégalité triangulaire $\|f+g\|_p \leqslant \|f\|_p + \|g\|_p$ s'appelle l'**inégalité de Minkowski**. Elle est en fait valable quelles que soient les fonctions mesurables $f, g : \Omega \to \mathbb{C}$ ou $[0,\infty]$, appartenant ou non à $\mathcal{L}^p$ (car elle est évidente si $\|f\|_p = \infty$ ou $\|g\|_p = \infty$). Explicitement, elle s'écrit

$$\left(\int_\Omega |f+g|^p\, d\mu\right)^{1/p} \leqslant \left(\int_\Omega |f|^p\, d\mu\right)^{1/p} + \left(\int_\Omega |g|^p\, d\mu\right)^{1/p}.$$

Comme cas particulier, elle contient l'inégalité de Minkowski *pour les sommes* : si $(x_i)_{i\in I}$ et $(y_i)_{i\in I}$ sont deux familles de nombres complexes indexées par le même ensemble I, alors

$$\left(\sum_{i\in I} |x_i + y_i|^p\right)^{1/p} \leqslant \left(\sum_{i\in I} |x_i|^p\right)^{1/p} + \left(\sum_{i\in I} |y_i|^p\right)^{1/p}.$$

Remarque 2. Pour $p = 1$, l'inégalité de Minkowski est *évidente* car $|f+g| \leqslant |f| + |g|$.

Exercice. Montrer que $\mathcal{L}^p$ est un espace vectoriel *sans utiliser* l'inégalité de Minkowski. (Montrer d'abord que si $f, g \in \mathcal{L}^p$, alors $\max(|f|, |g|) \in \mathcal{L}^p$.)

1.2. Le cas $p = \infty$. La définition de l'espace $\mathcal{L}^\infty(\Omega, \mathcal{B}, \mu)$ est un peu différente de celle des autres espaces $\mathcal{L}^p$. Tout repose sur le lemme suivant.

Lemme 1.4. *Pour toute fonction $f : \Omega \to \mathbb{C}$ ou $[0,\infty]$, il existe une* plus petite *constante $C \leqslant \infty$ telle que $|f(t)| \leqslant C$ presque partout.*

Démonstration. Il existe au moins une constante $C \leqslant \infty$ telle que $|f(t)| \leqslant C$ presque partout, à savoir $C := \infty$. Donc

$$K := \inf\{C \leqslant \infty;\ |f(t)| \leqslant C \text{ presque partout}\}$$

est un nombre positif bien défini, éventuellement égal à ∞. Il suffit visiblement de montrer qu'on a $|f(t)| \leqslant K$ presque partout.

Par définition de K, on peut trouver une suite $(C_n) \subseteq [0,\infty]$ telle que

$$C_n \to K \qquad \text{et} \qquad \forall n \in \mathbb{N} \ :\ \Big(|f(t)| \leqslant C_n \ \text{ presque partout}\Big).$$

Comme une conjection dénombrable de propriétés vraies presque partout est encore vraie presque partout, on a

$$\Big(\forall n \in \mathbb{N} \ :\ |f(t)| \leqslant C_n\Big) \ \text{ presque partout}\,;$$

et donc $|f(t)| \leqslant K$ presque partout puisque $C_n \to K$. □

DÉFINITION 1.5. *Pour toute fonction mesurable $f : \Omega \to \mathbb{C}$ ou $[0,\infty]$, on note $\|f\|_\infty$ la plus petite constante $C \leqslant \infty$ telle que $|f(t)| \leqslant C$ presque partout ; et on pose*

$$\mathcal{L}^\infty = \mathcal{L}^\infty(\Omega, \mathcal{B}, \mu) := \{f : \Omega \to \mathbb{C} \text{ mesurable; } \|f\|_\infty < \infty\}\,.$$

Ainsi, pour une fonction mesurable $f : \Omega \to \mathbb{C}$, on a l'équivalence

$$f \in \mathcal{L}^\infty \iff \exists C < \infty \text{ telle que } |f(t)| \leqslant C \text{ presque partout ;}$$

et pour toute constante $C \leqslant \infty$, on a l'équivalence

$$\|f\|_\infty \leqslant C \iff |f(t)| \leqslant C \text{ presque partout}\,.$$

En particulier, si $f : \Omega \to \mathbb{C}$ est mesurable et *bornée*, alors $f \in \mathcal{L}^\infty$ et

$$\|f\|_\infty \leqslant \sup\{|f(t)|;\ t \in \Omega\}\,.$$

Les fonctions de $\mathcal{L}^\infty(\Omega, \mathcal{B}, \mu)$ sont dites **essentiellement bornées** (relativement à la mesure μ).

Exercice. Montrer qu'une fonction mesurable $f : \Omega \to \mathbb{C}$ appartient à $\mathcal{L}^\infty$ si et seulement si elle est presque partout égale à une fonction mesurable bornée.

CAS PARTICULIERS. (1) Si Ω est un borélien de $\mathbb{R}^N$, la notation $\mathcal{L}^\infty(\Omega)$ est synonyme de $\mathcal{L}^\infty(\Omega, \mathcal{B}(\Omega), \lambda_N)$.

(2) Si I est un ensemble non vide, l'espace $\mathcal{L}^\infty(I, \mathcal{P}(I), \mu_c)$ se note $\ell^\infty(I)$. Comme "presque partout relativement à la mesure de comptage μ_c" veut simplement dire "partout", $\ell^\infty(I)$ est l'ensemble de toutes les fonctions *bornées* $f : I \to \mathbb{C}$, et on a pour toute fonction $f : I \to \mathbb{C}$:

$$\|f\|_\infty = \sup\{|f(t)|;\ t \in I\}\,.$$

(3) Si $I = \{1, \dots, N\}$ alors $\ell^\infty(I)$ s'identifie à $\mathbb{C}^N$, et pour $x = (x_1, \dots, x_N) \in \mathbb{C}^N$ on a

$$\|x\|_\infty = \max\big(|x_1|, \dots, |x_N|\big)\,.$$

(4) Si $I = \mathbb{N}$, alors $\ell^\infty(\mathbb{N})$ est l'ensemble de toutes les suites bornées $x = (x_i) \in \mathbb{C}^\mathbb{N}$, et pour toute suite $x = (x_i) \in \mathbb{C}^\mathbb{N}$ on a

$$\|x\|_\infty = \sup_{i \in \mathbb{N}} |x_i|\,.$$

Exemple 1. Soit $\Omega =]0,1]$ muni de la mesure $\mu = \delta_{1/2}$. La fonction $f :]0,1] \to \mathbb{R}$ définie par $f(t) = \frac{1}{t}$ n'est pas bornée (!), mais elle est dans $\mathcal{L}^\infty(]0,1], \mathcal{P}(]0,1]), \delta_{1/2})$ avec $\|f\|_{\mathcal{L}^\infty} = |f(1/2)| = 2$.

Démonstration. C'est clair car "pour $\delta_{1/2}$-presque tout $t \in]0,1]$" est synonyme de "pour $t = 1/2$". □

Exemple 2. Soit I un intervalle de $\mathbb{R}$ non trivial. Si $f : I \to \mathbb{C}$ est une fonction *continue*, alors $\|f\|_\infty = \sup\{|f(t)|;\ t \in I\}$. Donc $f \in \mathcal{L}^\infty(I)$ si et seulement si f est bornée, et $\|f\|_\infty$ a son sens "habituel".

Démonstration. La preuve repose sur le fait suivant (déjà mentionné au Chapitre 2).

FAIT 1.6. *Si E est un borélien de I tel que $\lambda_1(I \backslash E) = 0$, alors E est* dense *dans I.*

Preuve du Fait. Si E n'était pas dense dans I, alors $E \backslash I$ serait d'intérieur non vide dans I, donc contiendrait un intervalle J non trivial ; on aurait donc $\lambda_1(I \backslash E) \geqslant \lambda_1(J) = |J| > 0$, ce qui est exclu. □

Soit maintenant $f : I \to \mathbb{C}$ continue. On a vu que $\|f\|_\infty \leqslant \sup\{|f(t)|;\ t \in I\}$. Inversement, on a $|f(t)| \leqslant \|f\|_\infty$ presque partout par définition de $\|f\|_\infty$. Donc l'ensemble $E := \{t \in I;\ |f(t)| \leqslant \|f\|_\infty\}$ est dense dans I par le Fait. Mais E est également *fermé* dans I car f est supposée continue. Donc $E = I$, autrement dit $|f(t)| \leqslant \|f\|_\infty$ pour *tout* $t \in I$, et ainsi $\sup\{|f(t)|;\ t \in I\} \leqslant \|f\|_\infty$. □

PROPOSITION 1.7. *$\mathcal{L}^\infty = \mathcal{L}^\infty(\Omega, \mathcal{B}, \mu)$ est un espace vectoriel et $\|\cdot\|_\infty$ est une semi-norme sur $\mathcal{L}^\infty$. De plus, pour $f \in \mathcal{L}^\infty$ on a l'équivalence $\|f\|_\infty = 0 \iff f(t) = 0$ presque partout.*

La preuve (facile) est laissée en exercice.

1.3. Extension des définitions. On peut définir $\|f\|_p$ pour une fonction $f : \Omega \to \mathbb{C}$ ou $[0, \infty]$ seulement supposée *μ-mesurable*, *i.e.* presque partout définie et presque partout égale à une fonction mesurable : $\|f\|_p := \|\widetilde{f}\|_p$, pour tout fonction mesurable $\widetilde{f}$ égale à f presque partout.

2. Les espaces L^p, $1 \leqslant p \leqslant \infty$

2.1. Définition des L^p. Pour $p \in [1, \infty]$, on définit l'espace $L^p = L^p(\Omega, \mu)$ comme suit : une fonction *presque partout définie* $f : \Omega \to \mathbb{C}$ appartient à L^p si elle est presque partout égale à une fonction de $\mathcal{L}^p(\Omega, \mathcal{B}, \mu)$. Autrement dit, f appartient à L^p si et seulement si elle est μ-mesurable et $\|f\|_p < \infty$.

Le symbole d'égalité "=" doit être interprété au sens suivant :

$$f = g \text{ dans } L^p \iff f(t) = g(t) \text{ presque partout}.$$

Avec cette interprétation, on voit que si $f \in L^p$, alors

$$\|f\|_p = 0 \iff f = 0 \text{ dans } L^p\,;$$

et on déduit immédiatement le fait suivant :

FAIT 2.1. *$L^p = L^p(\Omega, \mu)$ est un espace vectoriel, et $\|\cdot\|_p$ est une* norme *sur L^p.*

Remarque. Ce n'est pas la définition standard. Normalement, pour faire les choses "proprement", il faut définir L^p comme étant le **quotient** de l'espace $\mathcal{L}^p$ par le sous-espace vectoriel $\mathcal{N}(\mu) := \{f;\ f(t) = 0 \text{ presque partout}\} = \{f;\ \|f\|_p = 0\}$, muni de la norme $\|\cdot\|_p$ induite par passage au quotient. Mais il n'est pas du tout clair que la définition "propre" soit plus parlante que celle qu'on vient de donner...

2.2. Complétude. Le résultat suivant porte le nom de **théorème de Riesz-Fischer**.

THÉORÈME 2.2. *Pour tout $p \in [1, \infty]$, l'espace vectoriel normé $(L^p, \|\cdot\|_p)$ est* complet*. Autrement dit : L^p est un* espace de Banach.

Pour la preuve, on a besoin d'un lemme bien connu.

LEMME 2.3. *Soit $(Z, \|\cdot\|)$ un espace vectoriel normé. On suppose que pour toute suite $(z_k) \subseteq Z$ telle que $\sum_0^\infty \|z_k\| < \infty$, la série $\sum z_k$ converge dans Z. Alors Z est complet.*

Remarque. Il est encore mieux connu que la réciproque du lemme est vraie... et beaucoup plus importante.

Preuve du lemme. Soit $(x_n)_{n\in\mathbb{N}}$ une suite de Cauchy dans Z. On veut montrer que (x_n) converge dans Z ; et comme (x_n) est de Cauchy, il suffit de prouver que (x_n) *possède une sous-suite convergente.*

Pour tout $k \in \mathbb{N}$, on peut trouver un entier n_k tel que

$$\forall p, q \geqslant n_k \ : \ \|x_p - x_q\| \leqslant 2^{-k} ;$$

et on peut de plus supposer que la suite (n_k) est strictement croissante.

Si on pose $z_0 := x_{n_0}$ et $z_k := x_{n_k} - x_{n_{k-1}}$ pour $k \geqslant 1$, alors $\|z_k\| \leqslant 2^{-(k-1)}$ pour tout $k \geqslant 1$, et donc $\sum_0^\infty \|z_k\| < \infty$. Par hypothèse sur Z, on en déduit que la série $\sum z_k$ converge dans Z ; et comme $\sum_{i=0}^k z_i = x_{n_k}$ pour tout $k \in \mathbb{N}$, cela signifie que la *suite* (x_{n_k}) est convergente. Ainsi, la suite de Cauchy (x_n) possède une sous-suite convergente, et donc converge dans Z. □

Preuve du Théorème 2.2. On va utiliser le fait suivant.

Fait. Si $(u_k)_{k\in\mathbb{N}}$ est une suite de fonctions mesurables positives sur Ω, alors

$$\left\| \sum_{k=0}^\infty u_k \right\|_p \leqslant \sum_{k=0}^\infty \|u_k\|_p .$$

Preuve du Fait. (i) Supposons d'abord que $p < \infty$. Si on pose $U := \sum_{k=0}^\infty u_k$ et $U_n := \sum_{k=0}^n u_k$ pour $n \in \mathbb{N}$, alors la suite $((U_n)^p)_{n\in\mathbb{N}}$ est croissante et tend vers U^p en tout point. Par convergence monotone, on a donc

$$\int_\Omega \left(\sum_{k=0}^\infty u_k \right)^p = \lim_{n\to\infty} \int_\Omega \left(\sum_{k=0}^n u_k \right)^p ;$$

d'où, en prenant les "puissances $1/p$-ièmes" :

$$\begin{aligned} \left\| \sum_{k=0}^\infty u_k \right\|_p &= \lim_{n\to\infty} \left\| \sum_{k=0}^n u_k \right\|_p \\ &= \underline{\lim} \left\| \sum_{k=0}^n u_k \right\|_p \\ &\leqslant \underline{\lim} \sum_{k=0}^n \|u_k\|_p = \sum_{k=0}^\infty \|u_k\|_p \quad \text{par Minkowski} . \end{aligned}$$

(ii) Supposons maintenant que $p = \infty$. Par définition de $\| \cdot \|_\infty$, on a

$$\forall k \in \mathbb{N} \ : \ \Big(u_k(t) \leqslant \|u_k\|_\infty \ \text{presque partout} \Big) ;$$

et donc en fait (par l'argument habituel)

$$\Big(\forall k \in \mathbb{N} \ : \ u_k(t) \leqslant \|u_k\|_\infty \Big) \ \text{presque partout} .$$

On en déduit

$$0 \leqslant \sum_{k=0}^\infty u_k(t) \leqslant \sum_{k=0}^\infty \|u_k\|_\infty \quad \text{presque partout} ,$$

et donc $\left\| \sum_0^\infty u_k \right\|_\infty \leqslant \sum_0^\infty \|u_k\|_\infty$.

□

Montrons maintenant que L^p est complet à l'aide du Lemme 2.3. Soit $(f_k)_{k\in\mathbb{N}}$ une suite d'éléments de L^p telle que $\sum_0^\infty \|f_k\|_p < \infty$. Il s'agit de montrer que la série $\sum f_k$ converge dans L^p.

On peut supposer que les f_k sont (partout définies et) mesurables. Par le Fait appliqué aux fonctions $u_k := |f_k|$, on a

$$\left\| \sum_{k=0}^\infty |f_k| \right\|_p \leqslant \sum_{k=0}^\infty \big\| |f_k| \big\|_p = \sum_{k=0}^\infty \|f_k\|_p < \infty\,.$$

En particulier, on a

$$\sum_{k=0}^\infty |f_k(t)| < \infty \qquad \text{presque partout}\,.$$

Soit alors $F : \Omega \to \mathbb{C}$ la fonction définie presque partout par

$$F(t) := \sum_{k=0}^\infty f_k(t)\,.$$

La fonction F est μ-mesurable car les f_k le sont ; et on a $|F(t)| \leqslant \sum_0^\infty |f_k(t)|$ pour tout $t \in \Omega$. Donc $\|F\|_p \leqslant \left\|\sum_0^\infty |f_k|\right\|_p < \infty$, et donc $F \in L^p$.

Si $n \in \mathbb{N}$, alors $F(t) - \sum_{k=0}^n f_k(t) = \sum_{k=n+1}^\infty f_k(t)$ presque partout. On a donc

$$\left| F(t) - \sum_{k=0}^n f_k(t) \right| \leqslant \left| \sum_{k=n+1}^\infty f_k(t) \right| \leqslant \sum_{k=n+1}^\infty |f_k(t)| \quad \text{presque partout}\,,$$

et donc

$$\begin{aligned} \left\| F - \sum_{k=0}^n f_k \right\|_p &\leqslant \left\| \sum_{k=n+1}^\infty |f_k| \right\|_p \\ &\leqslant \sum_{k=n+1}^\infty \|f_k\|_p \quad \text{par le Fait}\,. \end{aligned}$$

Par conséquent $\left\|F - \sum_{k=0}^n f_k\right\|_p \to 0$ quand $n \to \infty$; autrement dit la série $\sum f_k$ converge dans L^p et a pour somme F. □

Exercice. Montrer directement que les espaces $\ell^\infty(\mathbb{N})$ et $\ell^1(\mathbb{N})$ sont complets.

2.3. Le cas "hilbertien" $p = 2$. Le cas $p = 2$ est particulièrement important car l'espace L^2 est canoniquement muni d'une structure d'*espace de Hilbert.*

Rappelons d'abord quelques faits de base sur les espaces munis d'un produit scalaire.

• On dit qu'un $\mathbb{C}$-espace vectoriel normé $(H, \|\cdot\|)$ est une espace **préhilbertien** si la norme $\|\cdot\|$ dérive d'un produit scalaire, *i.e.* il existe un produit scalaire $\langle\cdot\,,\cdot\rangle$ sur H tel que

$$\forall x \in H \;:\; \|x\|^2 = \langle x, x\rangle\,.$$

• Comme on est sur un espace vectoriel complexe, la définition d'un produit scalaire est la suivante : $\langle x, y\rangle$ est linéaire par rapport à x et anti-linéaire par rapport à y (ou le contraire...) ; et $\langle x, x\rangle \in \mathbb{R}^+$ pour tout $x \in H$. On a alors nécessairement $\langle y, x\rangle = \overline{\langle x, y\rangle}$ pour tous $x, y \in H$ (exercice).

• Dans toute espace préhilbertien $(H, \|\cdot\|)$, on a l'**inégalité de Cauchy-Schwarz**

$$|\langle x, y\rangle| \leqslant \|x\|\,\|y\| \quad \text{pour tous } x, y \in H\,.$$

• Un **espace de Hilbert** est un espace préhilbertien *complet.*

Pour montrer que $L^2 = L^2(\Omega, \mathcal{B}, \mu)$ est un espace de Hilbert, il faut définir le produit scalaire ; et pour cela, on a besoin du lemme suivant.

Lemme 2.4. *Si $u, v \in L^2$, alors $uv \in L^1$.*

Démonstration. On a $|uv| \leqslant \frac{1}{2}(|u|^2+|v|^2)$ car $|u|^2+|v|^2-2|u|\,|v| = (|u|-|v|)^2 \geqslant 0$; donc $\int_\Omega |uv| \leqslant \frac{1}{2}\left(\int_\Omega |u|^2 + \int_\Omega |v|^2\right) < \infty$. □

Remarque. Ce lemme contient en particulier le résultat suivant : *si (u_n) et (v_n) sont deux suites de nombres complexes telles que $\sum_0^\infty |u_n|^2 < \infty$ et $\sum_0^\infty |v_n|^2 < \infty$, alors $\sum_0^\infty |u_n v_n| < \infty$.*

Théorème 2.5. *L'espace $L^2 = L^2(\Omega, \mathcal{B}, \mu)$est un espace de Hilbert, dont le produit scalaire est défini par*

$$\langle f, g\rangle_{L^2} := \int_\Omega f\,\overline{g}\,d\mu \quad \textit{pour } f, g \in L^2\,.$$

Démonstration. Par le lemme, $\langle f, g\rangle$ est bien défini pour toutes $f, g \in L^2$; et il est évident que $\langle\,\cdot\,,\,\cdot\,\rangle_{L^2}$ est un produit scalaire et qu'on a $\langle f, f\rangle_{L^2} = \|f\|_{L^2}^2$ pour toute $f \in L^2$. □

Corollaire 2.6. (Cauchy-Schwarz)
Si $f, g \in L^2$, alors

$$\left|\int_\Omega f(t)\overline{g}(t)\,d\mu(t)\right| \leqslant \left(\int_\Omega |f(t)|^2\,d\mu(t)\right)^{1/2} \left(\int_\Omega |g(t)|^2\,d\mu(t)\right)^{1/2} \;;$$

et en fait

$$\left|\int_\Omega |f(t)g(t)|\,d\mu(t)\right| \leqslant \left(\int_\Omega |f(t)|^2\,d\mu(t)\right)^{1/2} \left(\int_\Omega |g(t)|^2\,d\mu(t)\right)^{1/2} \,.$$

Démonstration. La première inégalité est simplement Cauchy-Schwarz pour le produit scalaire $\langle\,\cdot\,,\,\cdot\,\rangle_{L^2}$; et la deuxième est Cauchy-Schwarz appliqué aux fonctions $|f|$ et $|g|$, qui sont dans L^2. □

Corollaire 2.7. *Si $(u_i)_{i\in I}$ et $(v_i)_{i\in I}$ sont deux familles de nombres complexes, alors*

$$\sum_{i\in I} |u_i v_i| \leqslant \left(\sum_{i\in I} |u_i|^2\right)^{1/2} \left(\sum_{i\in I} |v_i|^2\right)^{1/2} \,.$$

Exercice. Soit H un espace préhilbertien, et soient $x, y \in H$ avec $y \neq 0$. Montrer que si on pose $u := \left\langle x, \frac{y}{\|y\|}\right\rangle \frac{y}{\|y\|}$, alors $u - x$ est orthogonal à y, et donc à u. En déduire l'inégalité de Cauchy-Schwarz pour x et y.

2.4. Convergence dans L^p et convergence presque partout. La proposition suivante est la "version L^p" du théorème de convergence dominée.

PROPOSITION 2.8. *Supposons $p < \infty$. Soit (f_n) une suite de fonctions de L^p, et soit $f : \Omega \to \mathbb{C}$. On suppose que $f_n(t) \to f(t)$ presque partout et qu'il existe une fonction $g \in L^p$ telle que $\forall n \in \mathbb{N}\ :\ |f_n(t)| \leqslant g(t)$ presque partout. Alors $f \in L^p$ et $f_n \to f$ en norme L^p.*

Démonstration. On a $|f(t)|^p = \lim |f_n(t)|^p \leqslant g(t)^p$ presque partout, donc $|f|^p$ est intégrable (car g^p l'est) *i.e.* $f \in L^p$. Ensuite, on applique le théorème de convergence dominée aux fonctions $\widetilde{f}_n(t) := |f_n(t) - f(t)|^p$ pour montrer que $\int_\Omega |f_n - f|^p \to 0$, autrement dit $\|f_n - f\|_p^p \to 0$. □

Remarque. Il n'y a pas de théorème de convergence dominée dans L^∞. Par exemple, les fonctions $f_n := \mathbf{1}_{[n,\infty[}$ sont dans $L^\infty(\mathbb{R})$, convergent en tout point vers 0 et sont dominées par la fonction $\mathbf{1} \in L^\infty$, mais elles ne convergent certainement pas vers 0 en norme L^∞.

Exercice 1. Montrer qu'en général, la convergence presque partout vers 0 d'une suite $(f_n) \subseteq L^p$ n'entraine pas sa convergence en norme L^p.

Exercice 2. Soit (f_n) une suite d'éléments de L^1, et soit $f \in L^1$. On suppose que $f_n \to f$ presque partout *et de plus* que $\|f_n\|_1 \to \|f\|_1$. Montrer que $f_n \to f$ en norme L^1. (Appliquer le lemme de Fatou aux fonctions $h_n(t) := |f_n| + |f| - |f_n - f|$.)

Le théorème de convergence dominée dit que la convergence presque partout entraine la convergence en norme L^p (pour $p < \infty$) si elle est associée à une hypothèse de domination. Inversement, la proposition suivante montre que la convergence en norme L^p n'est "pas si loin" d'entrainer la convergence presque partout.

PROPOSITION 2.9. *Supposons $p < \infty$. Si $(f_n) \subseteq L^p$ est une suite convergeant en norme L^p vers une fonction $f \in L^p$, alors (f_n) admet une* sous-suite *(f_{n_k}) qui tend vers f presque partout.*

Démonstration. Comme $f_n \to f$ en norme L^p, on peut trouver une suite strictement croissante d'entiers (n_k) telle que $\|f_{n_k} - f\|_p \leqslant 2^{-k}$ pour tout $k \in \mathbb{N}$. On a alors

$$\int_\Omega \left(\sum_{k=0}^\infty |f_{n_k} - f|^p \right) = \sum_{k=0}^\infty \int_\Omega |f_{n_k} - f|^p = \sum_{k=0}^\infty \|f_{n_k} - f\|_p^p \leqslant \sum_{k=0}^\infty 2^{-pk} < \infty\,.$$

En particulier, $\sum_0^\infty |f_{n_k}(t) - f(t)|^p < \infty$ presque partout ; et donc $|f_{n_k}(t) - f(t)| \to 0$ presque partout. □

Remarque. En général, la convergence en norme L^p n'entraine pas la convergence presque partout. Par exemple, soit (I_n) une suite de sous-intervalles de $I := [0,1]$ telle que $|I_n| \to 0$ et telle que tout point $t \in [0,1]$ appartienne à une infinité d'intervalles I_n (vérifier qu'une telle suite (I_n) existe). Si on choisit un suite α_n tendant vers l'infini telle que $\alpha_n\,|I_n| \to 0$ et si on pose $f_n := \alpha_n\,\mathbf{1}_{I_n}$, alors $f_n \to 0$ en norme L^1, mais (f_n) ne tend vers 0 en aucun point car on a $\sup_n |f_n(t)| = \infty$ pour tout $t \in [0,1]$.

Exercice. Montrer que la convergence en norme L^∞ entraine la convergence presque partout.

Voici enfin un dernier résultat utile, qu'il serait tentant d'appeler "lemme de Fatou dans L^p", ou bien "semi-continuité inférieure des normes L^p". (En fait, ce résultat a déjà été implicitement utilisé dans la preuve de la complétude de L^p.)

Proposition 2.10. *Si (f_n) est une suite de fonctions mesurables, $f_n : \Omega \to \mathbb{C}$ ou $[0,\infty]$ convergeant presque partout vers une fonction $f : \Omega \to \mathbb{C}$ ou $[0,\infty]$, alors on a pour tout $p \in [1,\infty]$:*

$$\|f\|_p \leqslant \underline{\lim}\, \|f_n\|_p\,.$$

Démonstration. Si $p < \infty$, on applique le lemme de Fatou aux fonctions $|f_n|^p$: cela donne

$$\int_\Omega \underline{\lim}\, |f_n|^p \leqslant \underline{\lim} \int_\Omega |f_n|^p\,,$$

autrement dit $\|f\|_p^p \leqslant \underline{\lim}\, \|f_n\|_p^p$.

Supposons maintenant $p = \infty$. Par définition de la norme L^∞ et le jeu habituel avec le presque partout, on a

$$\Big(|f(t)| = \lim |f_n(t)| = \underline{\lim}\, |f_n(t)| \text{ et } \forall n \in \mathbb{N} \,:\, |f_n(t)| \leqslant \|f_n\|_\infty\Big) \quad \text{presque partout}\,.$$

On en déduit aussitôt qu'on a $|f(t)| \leqslant \underline{\lim}\, \|f_n\|_\infty$ presque partout, autrement dit $\|f\|_\infty \leqslant \underline{\lim}\, \|f_n\|_\infty$. □

3. L'inégalité de Hölder

3.1. L'inégalité et quelques conséquences. Si $p \in [1,\infty]$, on appelle **exposant conjugué** de p l'unique $q \in [1,\infty]$ tel que

$$\frac{1}{p} + \frac{1}{q} = 1\,.$$

Par exemple, si $p = 1$ alors $q = \infty$; si $p = \infty$ alors $q = 1$; et si $p = 2$ alors $q = 2$. Il faut également se souvenir qu'on a $1 < p < \infty$ si et seulement si $1 < q < \infty$.

Exercice. Montrer que si p, q sont des exposants conjugués avec $1 < p, q < \infty$, alors

$$(q-1)p = q \quad \text{et} \quad (p-1)q = p\,.$$

Théorème 3.1. (inégalité de Hölder)
Soit $p \in [1,\infty]$, et soit q l'exposant conjugué. Pour toutes fonctions mesurables $f, g : \Omega \to [0,\infty]$ ou $\mathbb{C}$, on a

$$\|fg\|_1 \leqslant \|f\|_p\, \|g\|_q\,.$$

Si $1 < p, q < \infty$, cette inégalité s'écrit

$$\int_\Omega |fg|\, d\mu \leqslant \left(\int_\Omega |f|^p\, d\mu\right)^{1/p} \left(\int_\Omega |g|^q\, d\mu\right)^{1/q}.$$

Remarque. Pour $p = 2 = q$, on retrouve l'inégalité de Cauchy-Schwarz.

Preuve du Théorème 3.1. Si $p = 1$, $q = \infty$, on a $|f(t)g(t)| \leqslant |f(t)| \times \|g\|_\infty$ presque partout, et donc $\|fg\|_1 = \int_\Omega |f(t)g(t)|\, d\mu(t) \leqslant \|g\|_\infty \int_\Omega |f|\, d\mu = \|g\|_\infty\, \|f\|_1$. De même si $p = \infty$, $q = 1$.

Supposons maintenant $1 < p, q < \infty$. On a besoin du fait suivant.

Fait. Si $a, b \in [0,\infty]$, alors $ab \leqslant \frac{1}{p}\, a^p + \frac{1}{q}\, b^q$.

Preuve du Fait. L'inégalité est évidente si $a = \infty$ ou $b = \infty$ (car $\frac{1}{p}\, a^p + \frac{1}{q}\, b^q = \infty$ dans ce cas), et également si $a = 0$ ou $b = 0$ (car $ab = 0$ dans ce cas). On suppose donc qu'on a $0 < a, b < \infty$, ce qui permet de "passer aux logarithmes". Par *concavité* de la fonction logarithme et comme $\frac{1}{p} + \frac{1}{q} = 1$, on a

$$\log\left(\frac{1}{p}\, a^p + \frac{1}{q}\, b^q\right) \geqslant \frac{1}{p}\log(a^p) + \frac{1}{q}\log(b^q) = \log(a) + \log(b) = \log(ab)\,;$$

ce qui est l'inégalité souhaitée. □

Soient maintenant f, g deux fonctions mesurables sur Ω (à valeurs dans $\mathbb{C}$ ou $[0, \infty]$).

Si $\|f\|_p = 0$ ou $\|g\|_q = 0$, l'inégalité de Hölder est évidente car on a dans ce cas $f(t) = 0$ presque partout ou $g(t) = 0$ presque partout, et donc $fg = 0$ presque partout. On suppose donc qu'on a $\|f\|_p > 0$ et $\|g\|_q > 0$.

Les fonctions $u := \frac{f}{\|f\|_p}$ et $v := \frac{g}{\|g\|_q}$ vérifient $\|u\|_p = 1 = \|v\|_q$. Par le fait, on a

$$|u(t)v(t)| \leqslant \frac{1}{p}\,|u(t)|^p + \frac{1}{q}\,|v(t)|^q \quad \text{pour tout } t \in \Omega\,;$$

d'où en intégrant :

$$\begin{aligned}
\int_\Omega |uv|\, d\mu &\leqslant \frac{1}{p}\int_\Omega |u|^p\, d\mu + \frac{1}{q}\int_\Omega |v|^q\, d\mu \\
&= \frac{1}{p}\|u\|_p^p + \frac{1}{q}\|v\|_q^q \\
&= \frac{1}{p} + \frac{1}{q} = 1\,.
\end{aligned}$$

Par définition de u et v, cela s'écrit $\int_\Omega \frac{|fg|}{\|f\|_p\|g\|_q}\, d\mu \leqslant 1$; d'où

$$\|fg\|_1 = \int_\Omega |fg|\, d\mu \leqslant \|f\|_p\|g\|_q\,.$$

□

Exercice. Déduire l'inégalité de Minkowski $\|f+g\|_p \leqslant \|f\|_p + \|g\|_p$ de l'inégalité de Hölder. (En supposant $1 < p < \infty$, écrire $|f+g| = |f+g|^{p-1}|f+g|$, puis majorer $|f+g|$ et utiliser deux fois Hölder.)

Corollaire 3.2. *Si $f \in L^p$ et $g \in L^q$, alors $fg \in L^1$.*

Corollaire 3.3. *Si $u = (u_i)_{i\in I}$ et $v = (v_i)_{i\in I}$ sont deux familles de nombres complexes, alors*

$$\sum_{i\in I}|u_iv_i| \leqslant \left(\sum_{i\in I}|u_i|^p\right)^{1/p}\left(\sum_{i\in I}|v_i|^q\right)^{1/q}.$$

En particulier, si $u \in \ell^p(I)$ et $v \in \ell^q(I)$, alors $uv \in \ell^1(I)$.

Corollaire 3.4. *On suppose que la mesure μ est* finie, *i.e. $\mu(\Omega) < \infty$.*

(1) *La famille des espaces L^p est décroissante : si $1 \leqslant p_1 \leqslant p_2 \leqslant \infty$, alors $L^{p_2} \subseteq L^{p_1}$. En particulier, on a $L^\infty \subseteq L^p \subseteq L^1$ pour tout $p \in [1, \infty]$.*

(2) *Si de plus $\mu(\Omega) = 1$, alors la famille des normes L^p est croissante : si $1 \leqslant p_1 \leqslant p_2 \leqslant \infty$, alors $\|\cdot\|_{p_1} \leqslant \|\cdot\|_{p_2}$. En particulier, on a $\|\cdot\|_1 \leqslant \|\cdot\|_p \leqslant \|\cdot\|_\infty$ pour tout $p \in [1, \infty]$.*

Démonstration. (1) Supposons $p_1 < p_2 < \infty$. Pour toute fonction mesurable $f : \Omega \to \mathbb{C}$, l'inégalité de Hölder appliquée avec $p := \frac{p_2}{p_1}$ donne

$$\begin{aligned}\int_\Omega |f|^{p_1}\,d\mu &= \int_\Omega |f|^{p_1} \times \mathbf{1}\,d\mu \\ &\leqslant \left(\int_\Omega (|f|^{p_1})^p\,d\mu\right)^{1/p} \left(\int_\Omega \mathbf{1}^q\,d\mu\right)^{1/q} \\ &= \left(\int_\Omega |f|^{p_2}\,d\mu\right)^{p_1/p_2} \mu(\Omega)^{1-\frac{p_1}{p_2}}.\end{aligned}$$

En élevant à la puissance $1/p_1$, on obtient

$$\|f\|_{p_1} \leqslant \|f\|_{p_2} \times \mu(\Omega)^{\frac{1}{p_1}-\frac{1}{p_2}}. \tag{3.1}$$

Comme $\mu(\Omega) < \infty$, on en déduit immédiatement que si $\|f\|_{p_2} < \infty$ alors $\|f\|_{p_1} < \infty$; autrement dit : $f \in L^{p_2} \implies f \in L^{p_1}$.

Supposons maintenant $p_1 < p_2 = \infty$. Dans ce cas, on a

$$\|f\|_{p_1} = \left(\int_\Omega |f(t)|^p\,d\mu(t)\right)^{1/p} \leqslant \|f\|_\infty \times \mu(\Omega)^{1/p_1}; \tag{3.2}$$

et donc $f \in L^\infty \implies f \in L^{p_1}$.

(2) Le résultat est clair par (3.1) et (3.2). □

Remarque. Pour les espaces $\ell^p(\mathbb{N})$, la situation est exactement inverse : la famille des normes $\|\cdot\|_{\ell^p(\mathbb{N})}$ est *décroissante* (si $p_1 \leqslant p_2$, alors $\|\cdot\|_{\ell^{p_1}} \geqslant \|\cdot\|_{\ell^{p_2}}$), et par conséquent la suite des espaces $\ell^p(\mathbb{N})$ est croissante. En particulier, on a $\|\cdot\|_\infty \leqslant \|\cdot\|_p \leqslant \|\cdot\|_1$ et donc $\ell^1(\mathbb{N}) \subseteq \ell^p(\mathbb{N}) \subseteq \ell^\infty(\mathbb{N})$ pour tout $p \in [1,\infty]$.

Démonstration. On commence par observer que si $x = (x_n) \in \mathbb{C}^\mathbb{N}$, alors $|x_n| \leqslant \|x\|_p$ pour tout n et pour tout $p \in [1,\infty]$. On en déduit que si $p_1 \leqslant p_2$ et $\|x\|_{p_1} = 1$, alors $|x_n|^{p_2} \leqslant |x_n|^{p_1}$ pour tout $n \in \mathbb{N}$ et donc $\|x\|_{p_2} \leqslant 1$; ce qui entraine la conclusion souhaitée “par homogénéité”. Les détails sont laissés en exercice. □

Exercice 1. Montrer qu'il n'y a aucune inclusion entre espaces $L^p(\mathbb{R})$. (Considérer des fonctions de la forme $f(t) := \frac{1}{t^\alpha}\mathbf{1}_{[1,\infty[}(t)$ ou $f(t) := \frac{1}{t^\alpha}\mathbf{1}_{]0,1]}(t)$.)

Exercice 2. On suppose que $\mu(\Omega) = 1$. Montrer que si $f \in L^\infty$, alors $\|f\|_\infty = \lim_{p\to\infty}\|f\|_p$. (Observer que pour tout $\varepsilon > 0$, l'ensemble $B_\varepsilon := \{|f| \geqslant (1-\varepsilon)\|f\|_\infty\}$ est de mesure strictement positive, puis minorer $\int_{B_\varepsilon} |f|^p\,d\mu$ pour tout $p < \infty$.)

Corollaire 3.5. *Soit I un intervalle de $\mathbb{R}$. Si $f : I \to \mathbb{C}$ appartient à $L^p(I)$ pour un certain $p \in [1,\infty]$, alors f est intégrable sur tout intervalle borné $J \subseteq I$. En particulier, f est localement intégrable.*

Démonstration. Si $J \subseteq I$ est un intervalle borné, alors $f_{|J} \in L^p(J) \subseteq L^1(J)$; donc f est intégrable sur J. □

Exercice. Soit I un intervalle de $\mathbb{R}$, et soit $a \in \overline{I}$. Montrer que si $f \in L^p(I)$ pour un certain $p > 1$, alors la fonction $F(x) := \int_a^x f(t)\,dt$ est uniformément continue sur I. (Montrer qu'en fait $|F(y) - F(x)| \leqslant C\,|y-x|^\alpha$ pour certaines constantes $C < \infty$ et $\alpha > 0$.)

Corollaire 3.6. *Si $g \in L^q$, alors la formule $\Phi_g(f) := \int_\Omega fg\,d\mu$ définit une* forme linéaire continue *sur L^p, avec $\|\Phi_g\| \leqslant \|g\|_q$.*

Démonstration. Par l'inégalité de Hölder, la forme linéaire Φ_g est bien définie et on a $|\Phi_g(f)| \leqslant \int_\Omega |fg| \leqslant \|g\|_q \, \|f\|_p$ pour toute $f \in L^p$; autrement dit, Φ_g est continue avec $\|\Phi_g\| \leqslant \|g\|_q$. $\square$

Remarque 1. Si $p > 1$, on a en fait $\|\Phi_g\| = \|g\|_q$ pour toute fonction $g \in L^q$. Si $p = 1$, *i.e.* $q = \infty$, on a $\|\Phi_g\| = \|g\|_\infty$ pour toute $g \in L^\infty$ *à condition de supposer* que l'espace mesuré $(\Omega, \mathcal{B}, \mu)$ est **semi-fini**, ce qui signifie que pour tout ensemble mesurable B tel que $\mu(B) > 0$, on peut trouver un ensemble mesurable $A \subseteq B$ tel que $0 < \mu(A) < \infty$.

Démonstration. (i) Supposons $p > 1$, donc $q < \infty$. Fixons $g \in L^q$, et montrons qu'on a $\|\Phi_g\| \geqslant \|g\|_q$. Pour cela, il suffit de trouver une fonction $f \in L^p$ non nulle telle que $\Phi_g(f) = \|g\|_q \, \|f\|_p$. Évidemment, on peut supposer $g \neq 0$.

L'idée est de choisir la fonction f de sorte que $fg = |g|^q$. Cela ne laisse pas trop le choix : on doit poser

$$f(x) := \frac{|g(x)|^q}{g(x)} \quad \text{si } g(x) \neq 0\,.$$

Si $g(x) = 0$, on pose $f(x) := 0$. Par définition de f, on a ainsi

$$|f| = |g|^{q-1}\,,$$

avec la convention bizarre $0^0 = 0$ si $q = 1$.

Si $1 < p < \infty$, alors $|f|^p = |g|^{(q-1)p} = |g|^q$ car $(q-1)p = q$. Donc $\int_\Omega |f|^p = \int_\Omega |g|^q = \|g\|_q^q$, de sorte que $f \in L^p$ avec $\|f\|_p = \|g\|_q^{q/p}$. De plus, $\Phi_g(f) = \int_\Omega fg = \int_\Omega |g|^q = \|g\|_q^q$; ce qui s'écrit encore $\Phi_g(f) = \|g\|_q \times \|g\|_q^{q-1} = \|g\|_q \|g\|^{q/p}$ car $q - 1 = q/p$, autrement dit $\Phi_g(f) = \|g\|_q \, \|f\|_p$.

Si $p = \infty$, donc $q = 1$, alors $|f| = |g|^0$ avec la convention $0^0 = 0$; autrement dit, $|f(t)| = 1$ si $g(t) \neq 0$ et $f(t) = 0$ si $g(t) = 0$. Donc $f \in L^\infty$ avec $\|f\|_\infty = 1$; et $\Phi_g(f) = \int_\Omega |g| = \|g\|_1$.

(ii) Supposons maintenant que $p = 1$, $q = \infty$, et que l'espace mesuré $(\Omega, \mathcal{B}, \mu)$ soit semi-fini. Soit $g \in L^\infty$ quelconque, que l'on peut évidemment supposer non nulle.

Pour tout $\varepsilon \in]0,1[$, on a $\mu\big(\{|g| \geqslant (1-\varepsilon)\|g\|_\infty\}\big) > 0$ par définition de $\|g\|_\infty$. Comme l'espace mesuré est semi-fini, on peut donc trouver un ensemble mesurable A_ε tel que

$$|g| \geqslant (1-\varepsilon)\|g\|_\infty \ \text{ sur } A_\varepsilon \quad \text{et} \quad 0 < \mu(A_\varepsilon) < \infty\,.$$

Posons alors $f_\varepsilon := \mathbf{1}_{A_\varepsilon} \frac{|g|}{g}$, avec la convention $0 \times \frac{|0|}{0} = 0$. La fonction f_ε est dans L^1 car $|f| = \mathbf{1}_{A_\varepsilon}$ et $\mu(A_\varepsilon) < \infty$. On a $\|f\|_1 = \mu(A_\varepsilon) > 0$, donc $f \neq 0$; et

$$\Phi_g(f_\varepsilon) = \int_\Omega \mathbf{1}_{A_\varepsilon} |g| = \int_{A_\varepsilon} |g| \, d\mu \geqslant (1-\varepsilon)\|g\|_\infty \, \mu(A_\varepsilon) = (1-\varepsilon)\|g\|_\infty \, \|f_\varepsilon\|_1$$

par définition de A_ε. On en déduit $\|\Phi_g\| \geqslant (1-\varepsilon)\|g\|_\infty$ pour tout $\varepsilon > 0$, et donc $\|\Phi_g\| \geqslant \|g\|_\infty$. $\square$

Remarque 2. On peut montrer (mais on ne le fera pas ici) que si $1 < p < \infty$, alors *toute* forme linéaire continue sur L^p est de la forme Φ_g. Autrement dit, *le dual de L^p s'identifie isométriquement à L^q*. C'est encore vrai pour $p = 1$ si l'espace mesuré est “raisonnable”, mais c'est faux en général pour $p = \infty$.

Exercice. On suppose que l'espace mesuré $(\Omega, \mathcal{B}, \mu)$ n'est pas semi-fini. Soit B un ensemble mesurable vérifiant $\mu(B) > 0$ tel que $\mu(A) = 0$ pour tout ensemble mesurable $A \subseteq B$ de mesure finie. Montrer que si on prend $g := \mathbf{1}_B \in L^\infty$, alors $\Phi_g = 0$ (et donc $\|\Phi_g\| < \|g\|_\infty$).

3.2. Raisonnement par dualité. Soit $p \in [1, \infty]$, et soit q l'exposant conjugué. Si l'espace mesuré $(\Omega, \mathcal{B}, \mu)$ est semi-fini, la Remarque 1 suivant le Corollaire 3.6 (en échangeant les rôles de p et q) dit que pour toute fonction $f \in L^p$, la forme linéaire $\Phi_f : L^q \to \mathbb{C}$ définie par $\Phi_f(g) = \int_\Omega fg$ a une norme exactement égale à $\|f\|_p$. Autrement dit, si $f \in L^p$ alors

$$\|f\|_p = \sup \left\{ \left| \int_\Omega fg \right| ;\ g \in L^q\ ,\ \|g\|_q \leqslant 1 \right\} .$$

Comme de plus $\left|\int_\Omega fg\right| \leqslant \int_\Omega |fg| \leqslant \|f\|_p$ si $g \in L^q$ et $\|g\|_q \leqslant 1$, cela peut se ré-écrire comme suit :

$$\|f\|_p = \sup \left\{ \int_\Omega |fg|;\ g \in L^q\ ,\ \|g\|_q \leqslant 1 \right\} . \tag{3.3}$$

Ceci n'est *a priori* valable que pour une fonction f dont on sait qu'elle appartient à L^p. Mais le lemme suivant montre qu'en fait il n'est pas nécessaire de supposer que $f \in L^p$.

Lemme 3.7. *Supposons que l'espace mesuré $(\Omega, \mathcal{B}, \mu)$ soit **σ-fini**, ce qui signifie qu'il existe une suite croissante (Ω_n) d'ensembles mesurables telle que $\Omega = \bigcup_n \Omega_n$ et $\mu(\Omega_n) < \infty$ pour tout $n \in \mathbb{N}$. Alors (3.3) est valable pour toute fonction mesurable $f : \Omega \to \mathbb{C}$ ou $[0, \infty]$.*

Démonstration. Remarquons d'abord que l'espace mesuré $(\Omega, \mathcal{B}, \mu)$ est semi-fini : si $B \in \mathcal{B}$ vérifie $\mu(B) > 0$, alors $A_n := B \cap \Omega_n$ vérifie $\mu(A_n) < \infty$, et on a $\mu(A_n) > 0$ pour n assez grand car $\mu(A_n) \to \mu(B)$.

(i) Si $\|f\|_p < \infty$, le résultat est acquis. En effet, on a $|f(t)| < \infty$ presque partout, donc on peut supposer que f est à valeurs finies quitte à la redéfinir sur un ensemble de mesure nulle. Alors $f \in L^p$. Comme $(\Omega, \mathcal{B}, \mu)$ est semi-fini, la norme de la forme linéaire Φ_f agissant sur L^q est exactement égale à $\|f\|_p$, ce qui donne (3.3).

(ii) Supposons maintenant que $\|f\|_p = \infty$. Il s'agit de montrer que le sup apparaissant dans (3.3) vaut ∞. On va distinguer deux cas.

Cas 1. $|f(t)| < \infty$ presque partout.

Quitte à redéfinir f sur un ensemble de mesure nulle, on peut dans ce cas supposer que f est à valeurs finies.

Pour $n \in \mathbb{N}$, posons

$$f_n := \mathbf{1}_{B_n} f\ , \qquad \text{où } B_n := \Omega_n \cap \{|f| \leqslant n\} .$$

Comme les ensembles forment une suite croissante et qu'on a $\bigcup_n B_n = \Omega$ (car $|f(t)| < \infty$ pour tout $t \in \Omega$), on voit que $f_n(t) \to f(t)$ pour tout $t \in \Omega$. Par la Proposition 2.10, on en déduit qu'on a $\|f\|_p \leqslant \underline{\lim}\, \|f_n\|_p$, et donc

$$\lim_{n\to\infty} \|f_n\|_p = \infty .$$

Par ailleurs, les fonctions f_n *sont dans L^p* : c'est évident si $p = \infty$ car f_n est bornée ($|f_n| \leqslant n$), et clair aussi si $p < \infty$ car f_n est bornée et nulle en dehors de Ω_n, qui est de mesure finie. Par le cas déjà traité, on a donc

$$\|f_n\|_p = \sup \left\{ \int_\Omega |f_n g|;\ g \in L^q\ ,\ \|g\|_q \leqslant 1 \right\} \qquad \text{pour tout } n \in \mathbb{N} .$$

Mais $|f| \geqslant |f_n|$ pour tout $n \in \mathbb{N}$ par définition de f_n ; donc on obtient

$$\begin{aligned}\sup\left\{\int_\Omega |fg|;\ g\in L^q\ ,\ \|g\|_q\leqslant 1\right\} &\geqslant \sup_{n\in\mathbb{N}}\sup\left\{\int_\Omega |f_n g|;\ g\in L^q\ ,\ \|g\|_q\leqslant 1\right\}\\ &= \sup_{n\in\mathbb{N}}\|f_n\|_p\\ &= \infty\,.\end{aligned}$$

Cas 2. f est à valeurs dans $[0,\infty]$ et $\mu\big(\{f=\infty\}\big)>0$.

Dans ce cas, comme l'espace mesuré $(\Omega,\mathcal{B},\mu)$ est semi-fini, on peut trouver un ensemble mesurable E tel que $0<\mu(E)<\infty$ et $f(x)\equiv\infty$ sur E. Si on pose $g:=\frac{\mathbf{1}_E}{\mu(E)^{1/q}}$, alors $g\in L^q$ avec $\|g\|_q=1$ car $\mu(E)>0$, et $\int_\Omega fg=\frac{1}{\mu(E)^{1/q}}\int_E f\,d\mu=\infty$ car $f\equiv\infty$ sur E et $\mu(E)<\infty$. Donc le sup apparaissant au membre de droite de (3.3) vaut ∞. □

L'intérêt du lemme est qu'il permet de faire des raisonnements "par dualité", via la conséquence immédiate suivante :

Corollaire 3.8. *Supposons que $(\Omega,\mathcal{B},\mu)$ soit σ-fini. Soit $f:\Omega\to\mathbb{C}$ ou $[0,\infty]$ une fonction mesurable, et soit $C\leqslant\infty$. On suppose qu'on a*

$$\int_\Omega |fg|\leqslant C\,\|g\|_q \qquad \textit{pour toute fonction } g\in L^q\,.$$

Alors on peut conclure que $\|f\|_p\leqslant C$.

Comme illustration, on va démontrer un résultat très utile qu'on appelle parfois la *forme intégrale de l'inégalité de Minkowski.* Il y a des versions plus générale, mais celle là sera suffisante.

Corollaire 3.9. *Supposons que Ω soit un borélien de $\mathbb{R}^N$. Soit $\Phi:\Omega\times J\to[0,\infty]$ une fonction borélienne, où J est un intervalle de $\mathbb{R}$. Posons $\Phi_t(x):=\Phi(x,t)$, et*

$$h(x):=\int_J \Phi_t(x)\,dt=\int_J\Phi(x,t)\,dt\,.$$

Pour tout $p\in[1,\infty]$, on a

$$\|h\|_p\leqslant\int_J\|\Phi_t\|_p\,dt\,.$$

Démonstration. Remarquons d'abord qu'on a bien affaire à un espace mesuré σ-fini car Ω est réunion dénombrable de boréliens bornés.

Par dualité, il suffit de montrer que pour toute fonction $g\in L^q$ (où q est l'exposant conjugué de p) on a

$$\int_\Omega |h(x)g(x)|\,dx\leqslant\|g\|_q\times\int_J\|\Phi_t\|_p\,dt\,.$$

C'est un calcul totalement transparent :

$$\begin{aligned}
\int_\Omega |h(x)g(x)|\,dx &= \int_\Omega |g(x)|\left(\int_J \Phi(x,t)\,dt\right)dx\\
&= \int_J\left(\int_\Omega \Phi_t(x)|g(x)|\,dx\right)dt \qquad \text{par Fubini}\\
&\leqslant \int_J \|\Phi_t\|_p\|g\|_q\,dt \qquad \text{par Hölder}\\
&= \|g\|_q \times \int_J \|\Phi_t\|_p\,dt\,.
\end{aligned}$$

□

Voici une jolie conséquence de l'inégalité "intégrale" de Minkowski, qu'on appelle souvent l'**inégalité de Hardy**. (Il y a en fait beaucoup d'"inégalités de Hardy".)

Exemple 3.10. Soit p vérifiant $1 < p < \infty$. Pour toute fonction $f \in L^p(]0,\infty[)$, on note $Tf :]0,\infty[\to \mathbb{C}$ la fonction définie par

$$Tf(x) := \frac{1}{x}\int_0^x f(t)\,dt\,.$$

Alors $Tf \in L^p$ et $\|Tf\|_p \leqslant q\,\|f\|_p$, où q est l'exposant conjugué de p.

Démonstration. La fonction Tf est bien définie car f est intégrable sur tout intervalle borné $I \subseteq]0,\infty[$ par le Corollaire 3.5 ; et de plus, Tf est une fonction *continue* (*cf* Chapitre 6 ; voir aussi l'exercice après le Corollaire 3.5) et donc borélienne.

Par changement de variable, on a

$$\begin{aligned}
Tf(x) &= \int_0^x f(t)\,\frac{dt}{x}\\
&= \int_0^1 f(xu)\,du\\
&:= \int_0^1 \Phi_u(x)\,du\,.
\end{aligned}$$

D'après la forme intégrale de l'inégalité de Minkowski, on en déduit

$$\|Tf\|_p \leqslant \int_0^1 \|\Phi_u\|_p\,du\,.$$

Maintenant, on a pour $u \in]0,1[$:

$$\begin{aligned}
\|\Phi_u\|_p^p &= \int_0^\infty |f(xu)|^p dx\\
&= \frac{1}{u}\int_0^\infty |f(t)|^p\,dt\\
&= \frac{1}{u}\,\|f\|_p^p\,.
\end{aligned}$$

Donc $\|\Phi_u\|_p = \frac{1}{u^{1/p}}\,\|f\|_p$ pour tout $u \in]0,1[$, et donc

$$\int_0^1 \|\Phi_u\|_p\,du = \|f\|_p \int_0^1 \frac{du}{u^{1/p}} = \frac{1}{1-\frac{1}{p}}\,\|f\|_p = q\,\|f\|_p\,.$$

Ainsi, $Tf \in L^p$ avec $\|Tf\|_p \leqslant q\,\|f\|_p$, comme annoncé. □

3.2.1. *Autre preuve de l'inégalité intégrale de Minkowski.* Comme si tout cela ne suffisait pas, on va donner maintenant une autre démonstration de l'inégalité intégrale de Minkowski, qui ne repose pas explicitement sur une raisonnement par dualité.

Soit donc $\Phi : \Omega \times J \to [0, \infty]$ une fonction borélienne (où Ω est un borélien de $\mathbb{R}^N$ et J un intervalle de $\mathbb{R}$), et soit

$$h(x) = \int_J \Phi_t(x)\, dt = \int_J \Phi(x,t)\, dt\,.$$

Il s'agit de montrer qu'on a

$$\|h\|_p \leqslant \int_J \|\Phi_t\|_p\, dt \qquad \text{pour tout } p \in [1, \infty]\,.$$

Si $p = 1$, le résultat est immédiat par le théorème de Fubini, et l'inégalité est même une *égalité* :

$$\|h\|_1 = \int_\Omega h(x)\, dx = \int_\Omega \left(\int_I \Phi_t(x)\, dt\right) dx = \int_I \left(\int_\Omega \Phi_t(x)\, dx\right) dt = \int_I \|\Phi_t\|_1\, dt\,.$$

Dans la suite, on suppose donc que $p > 1$. On va distinguer trois cas.

Cas 1. $1 < p < \infty$ et $0 < \|h\|_p < \infty$.

Dans ce cas, l'idée est d'écrire $h(x) = h(x)^{p-1}h(x)$ et d'appliquer Hölder de façon judicieuse. On a

$$\begin{aligned}
\|h\|_p^p &= \int_\Omega h(x)^{p-1}h(x)\, dx \\
&= \int_\Omega h(x)^{p-1}\left(\int_I \Phi_t(x)\, dt\right) dx \\
&= \int_I \left(\int_\Omega \Phi_t(x)h(x)^{p-1}\, dx\right) dt \qquad \text{par Fubini} \\
&= \leqslant \int_I \left(\int_\Omega \Phi_t(x)^p dx\right)^{1/p} \left(\int_\Omega h(x)^{(p-1)q} dx\right)^{1/q} \qquad \text{par Hölder} \\
&= \int_I \|\Phi_t\|_p \left(\int_\Omega h(x)^p dx\right)^{1/q} dt \qquad \text{car } (p-1)q = p \\
&= \|h\|_p^{p/q} \int_I \|\Phi_t\|_p\, dt\,.
\end{aligned}$$

Comme $0 < \|h\|_p < \infty$, on en déduit *en divisant par* $\|h\|_p^{p/q}$:

$$\|h\|_p^{p(1-\frac{1}{q})} \leqslant \int_I \|\Phi_t\|_p\, dt\,;$$

autrement dit $\|h\|_p \leqslant \int_J \|\Phi_t\|_p\, dt$.

Cas 2. $1 < p < \infty$ et $\|h\|_p = \infty$.

Dans ce cas, il faut montrer qu'on a aussi $\int_I \|\Phi_t\|_p\, dt = \infty$.

Supposons d'abord qu'on ait $h(x) < \infty$ presque partout. Soit (Ω_n) une suite croissante de boréliens bornés telle que $\Omega = \bigcup_{n=0}^\infty \Omega_n$, et posons

$$h_n := \mathbf{1}_{A_n}\, h\,, \qquad \text{où} \quad A_n := \Omega_n \cap \{h \leqslant n\}\}\,.$$

La suite (h_n) est croissante, et comme on suppose que $h(x) < \infty$ presque partout, $h_n(x) \to h(x)$ presque partout. Par le théorème de convergence monotone appliqué aux h_n^p, on a donc

$$\lim_{n\to\infty} \|h_n\|_p = \|h\|_p = \infty\,.$$

De plus, on a

$$h_n(x) = \int_I \Phi_{t,n}(x)\, dt\,, \qquad \text{où} \quad \Phi_{t,n}(x) := \mathbf{1}_{A_n}(x)\Phi_t(x)\,;$$

et enfin, on a aussi $\|h_n\|_p < \infty$, car h_n est bornée sur A_n et A_n est *de mesure finie* car borné. Par le Cas 1, on en déduit

$$\|h_n\|_p \leqslant \int_I \|\mathbf{1}_{A_n}\Phi_t\|_p\, dt \leqslant \int_I \|\Phi_t\|_p\, dt \quad \text{pour tout } n \in \mathbb{N}\,,$$

et donc $\int_I \|\Phi_t\|_p\, dt = \infty$.

Supposons maintenant *qu'on n'ait pas* $h(x) < \infty$ presque partout, autrement dit que l'ensemble $\{h = \infty\}$ soit de mesure strictement positive. Comme Ω est réunion d'une suite de boréliens bornés Ω_n, l'un des $\Omega_n \cap \{h = \infty\}$ est de mesure strictement positive *et finie*. Ainsi, il existe un ensemble borélien $E \subseteq \Omega$ tel que

$$h(x) \equiv \infty \ \text{ sur } \ E \quad \text{et} \quad 0 < \mu(E) < \infty\,.$$

où on a noté μ la mesure de Lebesgue λ_N.

Par l'inégalité de Hölder, on a pour tout $t \in I$:

$$\int_E \Phi_t(x)\, dx \leqslant \left(\int_E \Phi_t(x)^p\, dx\right)^{1/p} \times \mu(E)^{1/q} \leqslant \mu(E)^{1/q}\, \|\Phi_t\|_p\,.$$

Donc

$$\begin{aligned}
\mu(E)^{1/q} \int_I \|\Phi_t\|_p\, dt \quad &\geqslant \quad \int_I \left(\int_E \Phi_t(x)\, dx\right) dt \\
&= \quad \int_E \left(\int_I \Phi_t(x)\, dt\right) dx \\
&= \quad \int_E h(x)\, dx \\
&= \quad \infty \qquad \text{car } h(x) \equiv \infty \text{ sur } E \text{ et } \mu(E) > 0\,.
\end{aligned}$$

D'où $\int_I \|\Phi_t\|_p\, dt = \infty$ car $\mu(E) < \infty$.

Cas 3. $p = \infty$.

Dans ce cas, il faut montrer qu'on a $h(x) \leqslant \int_I \|\Phi_t\|_\infty\, dt$ presque partout.

Introduisons l'ensemble

$$B := \big\{(x,t) \in \Omega \times I;\ \Phi_t(x) > \|\Phi_t\|_\infty\big\} \subseteq \mathbb{R}^N \times \mathbb{R}\,.$$

Par le théorème de Fubini, on a (en notant toujours $\mu = \lambda_N$)

$$\lambda_{N+1}(B) = \int_\Omega \lambda_1\big(B_x\big)\, dx = \int_I \mu(B^t)\, dt\,.$$

Mais pour tout $t \in I$ fixé, on a $\mu(B^t) = 0$ car $B^t = \{x;\ \Phi_t(x) > \|\Phi_t\|_\infty\}$ et $\Phi_t(x) \leqslant \|\Phi_t\|_\infty$ presque partout. Donc $\lambda_{N+1}(B) = 0$; et par conséquent

$$0 = \int_\Omega \lambda_1(B_x)\, dx = \int_\Omega \lambda_1\big(\{t;\ \Phi_t(x) > \|\Phi_t\|_\infty\}\big)\,.$$

Par conséquent, $\lambda_1\big(\{t;\ \Phi_t(x) > \|\Phi_t\|_\infty\}\big) = 0$ pour presque tout $x \in \Omega$. Autrement dit, pour presque tout $x \in \Omega$, on a

$$\Phi_t(x) \leqslant \|\Phi_t\|_\infty \qquad \text{pour presque tout } t \in I\,.$$

Donc, pour presque tout $x \in \Omega$,

$$h(x) = \int_I \Phi_t(x)\,dt \leqslant \int_I \|\Phi_t\|_\infty\,dt\,.$$

4. Résultats de densité

4.1. Fonctions étagées. Dans ce qui suit, l'espace mesuré $(\Omega, \mathcal{B}, \mu)$ est quelconque. On note $\mathcal{E}(\Omega)$ l'ensemble des fonctions étagées sur Ω, à valeurs complexes. Ainsi, une fonction $\varphi : \Omega \to \mathbb{C}$ appartient à $\mathcal{E}(\Omega)$ si et seulement si elle est de la forme

$$\varphi = \sum_{i=1}^{N} \alpha_i\, \mathbf{1}_{A_i}\,,$$

où les A_i sont des ensembles mesurables et $\alpha_i \in \mathbb{C}$.

Remarque 1. Toute fonction étagée est *bornée*, et donc $\mathcal{E}(\Omega) \subseteq L^\infty$. De plus, si $\varphi \in \mathcal{E}(\Omega)$ s'écrit $\varphi = \sum_1^N \alpha_i \mathbf{1}_{A_i}$ avec des A_i deux à deux disjoints vérifiant $\mu(A_i) > 0$, alors $\|\varphi\|_{L^\infty} = \max(|\alpha_1|, \dots, |\alpha_N|)$.

Démonstration. La première partie est évidente ; et la deuxième est laissée en exercice. □

Remarque 2. Soit $p \in [1, \infty[$. Si $\varphi \in \mathcal{E}(\Omega)$ s'écrit $\varphi = \sum_{i=1}^N \alpha_i \mathbf{1}_{A_i}$ avec des A_i deux à deux disjoints, alors $\|\varphi\|_p^p = \sum_{i=1}^N |\alpha_i|^p\, \mu(A_i)$. Par conséquent, une fonction étagée φ appartient à L^p si et seulement si $\mu(\{\varphi \neq 0\}) < \infty$.

Démonstration. La première partie est évidente puisque $|\varphi|^p = \sum_{i=1}^N |\alpha_i|^p \mathbf{1}_{A_i}$. Pour la deuxième, il suffit d'observer d'une part que si on pose $I := \{i;\ \alpha_i \neq 0\}$, alors $\|\varphi\|_p^p < \infty$ si et seulement si $\mu(A_i) < \infty$ pour tout $i \in I$ (ce qui revient à dire que $\sum_{i \in I} \mu(A_i) < \infty$ puisque I est fini), et d'autre part que l'ensemble $\{\varphi \neq 0\}$ est égal à $\bigcup_{i \in I} A_i$. □

LEMME 4.1. *Pour toute fonction mesurable $f : \Omega \to \mathbb{C}$, on peut trouver une suite (φ_n) de fonctions étagées telle que*

(i) *$\varphi_n(t) \to f(t)$ pour tout $t \in \Omega$;*

(ii) *$|\varphi_n(t)| \leqslant |f(t)|$ pour tout n et pour tout $t \in \Omega$.*

Si de plus la fonction f est bornée, on peut faire en sorte que (φ_n) converge uniformément vers f.

Démonstration. Pour une fonction f *positive*, le résultat a déjà été démontré au Chapitre 3, et on sait qu'on peut prendre les φ_n positives dans ce cas.

Supposons f à valeurs réelles, et écrivons $f = f^+ - f^-$. Par le cas “positif”, on peut trouver des suites $(\varphi_{n,+})$ et $(\varphi_{n,-})$ convenant pour f^+ et f^-. On a alors $\varphi_{n,+} \equiv 0$ sur $\Omega^- := \{f \leqslant 0\}$ car $0 \leqslant \varphi_{n,+} \leqslant f^+$ et $f^+ \equiv 0$ sur Ω^- ; et de même $\varphi_{n,-} \equiv 0$ sur $\Omega^+ := \{f \geqslant 0\}$. Donc $\varphi_{n,+}\varphi_{n,-} = 0$. Par conséquent, si on pose $\varphi_n := \varphi_{n,+} - \varphi_{n,-}$, alors $(\varphi_n)^\pm = \varphi_{n,\pm}$ (*cf* l'Exercice 3.1 du Chapitre 4). Donc $|\varphi_n| = \varphi_n^+ + \varphi_n^+ = \varphi_{n,+} + \varphi_{n,-} \leqslant f^+ + f^- = |f|$; et $\varphi_n \to f^+ - f^- = f$, avec convergence uniforme si f est bornée (car f^+ et f^- sont alors bornées).

Pour une fonction f à valeurs complexes, on applique le cas “réel” aux fonctions $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$. Les détails sont laissés en exercice. □

THÉORÈME 4.2. *Les fonctions étagées appartenant à L^p sont denses dans L^p, pour tout $p \in [1, \infty]$.*

Démonstration. (i) Supposons $p < \infty$. Soit $f \in L^p$ quelconque, et soit (φ_n) une “suite approximante” de fonctions étagées pour f comme dans le lemme. Comme $|\varphi_n| \leqslant |f| \in L^p$, les fonctions φ_n sont dans L^p ; et la suite (φ_n) converge vers f en norme L^p par convergence dominée dans L^p.

(ii) Supposons maintenant $p = \infty$. Soit $f \in L^\infty$ quelconque. Il existe alors une fonction mesurable *bornée* $\widetilde{f} : \Omega \to \mathbb{C}$ telle que $\widetilde{f} = f$ presque partout.

Soit (φ_n) une suite de fonctions étagées (donnée par le lemme) convergeant uniformément vers $\widetilde{f}$ sur Ω, et posons $\varepsilon_n := \sup\{|\varphi_n(t) - \widetilde{f}(t)|;\ t \in \Omega\}$, de sorte que $\varepsilon_n \to 0$. Comme $f = \widetilde{f}$ presque partout, on a $|\varphi_n(t) - f(t)| = |\varphi_n(t) - \widetilde{f}(t)| \leqslant \varepsilon_n$ presque partout ; donc $\|\varphi_n - f\|_\infty \leqslant \varepsilon_n$, et donc $\varphi_n \to f$ pour la norme L^∞. □

4.2. Fonctions continues à support compact. Dans cette section, Ω est un *ouvert de* $\mathbb{R}^N$. On dira qu'une fonction $\varphi : \Omega \to \mathbb{C}$ est **à support compact** s'il existe un compact $K \subseteq \Omega$ tel que $\varphi(t) \equiv 0$ *en dehors de* K. On notera $\mathcal{C}_{00}(\Omega)$ l'ensemble de toutes les fonctions continues à support compact $\varphi : \Omega \to \mathbb{C}$.

Il faut prendre garde au fait que le compact K dans la définition *dépend de la fonction* φ, et que K doit être entièrement contenu dans Ω. Par exemple, la fonction $\varphi :]-1, 1[\to \mathbb{R}$ définie par $\varphi(t) := 1 - t^2$ n'*appartient pas* à $\mathcal{C}_{00}(\,]-1,1[\,)$; mais elle appartient à $\mathcal{C}_{00}(\mathbb{R})$ si on la prolonge par 0 en dehors de $]-1, 1[$.

Exercice. Montrer que $\mathcal{C}_{00}(\Omega)$ est un espace vectoriel.

Remarque. On a $\mathcal{C}_{00}(\Omega) \subseteq L^p(\Omega)$ pour tout $p \in [1, \infty]$.

Démonstration. Soit $\varphi \in \mathcal{C}_{00}(\Omega)$, et soit $K \subseteq \Omega$ un compact tel que $\varphi(t) \equiv 0$ en dehors de K. Comme φ est continue, elle est bornée sur le compact K, donc bornée sur Ω puisqu'elle est nulle en dehors de K ; donc $\varphi \in L^\infty(\Omega)$. Si $p < \infty$, alors

$$\int_\Omega |\varphi|^p = \int_K |\varphi|^p \, d\lambda_N \leqslant \|\varphi\|_\infty^p \times \lambda_N(K) < \infty \quad \text{car } K \text{ est borné},$$

donc $\varphi \in L^p(\Omega)$. □

THÉORÈME 4.3. *Si $p < \infty$, alors $\mathcal{C}_{00}(\Omega)$ est dense dans $L^p(\Omega)$.*

Ce résultat n'est pas du tout évident. On a besoin de deux lemmes.

LEMME 4.4. (régularité de la mesure de Lebesgue)
Pour tout borélien $A \subseteq \mathbb{R}^N$ tel que $\lambda_N(A) < \infty$ et pour tout $\varepsilon > 0$, on peut trouver un compact E et un ouvert U tels que

$$E \subseteq A \subseteq U \quad \text{et} \quad \lambda_N(U \setminus E) < \varepsilon\,.$$

Démonstration. Fixonas A et $\varepsilon > 0$. On procède en deux étapes.

FAIT 1. Il existe un ouvert U tel que $A \subseteq U$ et $\lambda_N(U \setminus A) < \varepsilon/2$.

Preuve du Fait 1. Par définition de la mesure de Lebesgue, on peut trouver une suite de pavés $(P_k)_{k\geqslant 1}$ telle que

$$A \subseteq \bigcup_{k=1}^{\infty} P_k \qquad \text{et} \qquad \sum_{k=1}^{\infty} |P_k| < \lambda_N(A) + \frac{\varepsilon}{4}\cdot$$

Pour tout $k \geqslant 1$, on peut ensuite trouver un pavé *ouvert* Q_k tel que

$$P_k \subseteq Q_k \qquad \text{et} \qquad |Q_k| \leqslant |P_k| + 2^{-k}\,\frac{\varepsilon}{4}\cdot$$

Si on pose $U := \bigcup_{k=1}^{\infty} Q_k$, alors U est ouvert, $A \subseteq U$, et

$$\lambda_N(U) \leqslant \sum_{k=1}^{\infty} |Q_k| \leqslant \sum_{k=1}^{\infty} |P_k| + \frac{\varepsilon}{4}\sum_{k=1}^{\infty} 2^{-k} < \lambda_N(A) + \frac{\varepsilon}{2};$$

d'où $\lambda_N(U\backslash A) < \varepsilon/2$. □

Fait 2. Il existe un compact E tel que $E \subseteq A$ et $\lambda_N(A\backslash E) < \varepsilon/2$.

Preuve du Fait 2. On observe d'abord qu'il existe un pavé compact P tel que $\lambda(A\backslash(P\cap A)) < \varepsilon/4$. En effet, soit $(P_n)_{n\geqslant 0}$ une suite de croissante de pavés compacts telle que $\bigcup_{n\geqslant 0} P_n : \mathbb{R}^N$. La suite $(P_n\cap A)$ croit vers A, donc $\lambda_N(P_n\cap A)$ tend vers $\lambda_N(A)$, et par conséquent $\lambda_N(A\backslash(P_n\cap A)) = \lambda_N(A) - \lambda_N(P_n\cap A)$ tend vers 0 ; donc on peut prendre $P = P_n$ pour n assez grand.

Maintenant, on applique le Fait 1 à $A' = P\backslash A$, avec $\varepsilon/4$ au lieu de $\varepsilon/2$: cela fournit un ouvert U tel que

$$P\backslash A \subseteq U \qquad \text{et} \qquad \lambda_N\big(U\backslash(P\backslash A)\big) < \varepsilon/4.$$

Si on pose $E := P\backslash U$, alors E est compact (intersection du compact P et du fermé $\mathbb{R}^N\backslash U$), et $E \subseteq A$ car $E\backslash A = (P\backslash U)\backslash A = (P\backslash A)\backslash U = \varnothing$. De plus, $(P\cap A)\backslash E = (P\cap A)\backslash(P\backslash U) \subseteq U\backslash(P\backslash A)$, et donc $\lambda_N\big((P\cap A)\backslash E\big) < \varepsilon/4$. Comme $A\backslash E = A\backslash(P\cap A)\cup(P\cap A)\backslash E$, on a donc

$$\lambda_N(A\backslash E) \leqslant \lambda_N\big(A\backslash(P\cap A)\big) + \lambda_N\big((P\cap A)\backslash E\big) < \frac{\varepsilon}{2}\cdot$$

□

La preuve du lemme est mantenant terminée : si U et E sont donnés par les Faits 1 et 2, alors $E \subseteq A \subseteq U$ et $\lambda_N(U\backslash E) = \lambda_N(U\backslash A) + \lambda_N(A\backslash) < \varepsilon$. □

Remarque. On démontrera une version beaucoup plus générale de ce lemme au Chapitre 10.

Lemme 4.5. *Si U est un ouvert de $\mathbb{R}^N$ et si $E \subseteq \Omega$ est un compact tel que $E \subseteq U$, alors on peut trouver une fonction $\varphi \in \mathcal{C}_{00}(\Omega)$ telle que $\varphi(t) \equiv 1$ sur E et $\varphi(t) \equiv 0$ en dehors de U, avec de plus $0 \leqslant \varphi \leqslant 1$.*

Démonstration. Quitte à remplacer U par $U\cap\Omega$ (qui est ouvert et contient E), on peut supposer que $U \subseteq \Omega$.

Comme E est compact et U ouvert, on peut trouver un ouvert borné W tel que

$$E \subseteq W \quad \text{et} \quad \overline{W} \subseteq U\,.$$

(Il suffit de prendre $W := \{x \in \mathbb{R}^N ; \mathrm{dist}(x, E) < \varepsilon\}$ pour $\varepsilon > 0$ assez petit.)

On a alors

$$\mathrm{dist}(x, E) + \mathrm{dist}(x, \mathbb{R}^N\backslash W) > 0 \quad \text{pour tout } x \in \mathbb{R}^N\,.$$

En effet : comme $E \subseteq W$, on a ou bien $x \notin E$ ou bien $x \notin \mathbb{R}^N \backslash W$, donc ou bien $\operatorname{dist}(x, E) > 0$ ou bien $\operatorname{dist}(x, \mathbb{R}^N \backslash W) > 0$ car E et $\mathbb{R}^N \backslash W$ sont des fermés de $\mathbb{R}^N$.

On peut donc définir $\varphi : \Omega \to \mathbb{R}$ par

$$\varphi(x) := \frac{\operatorname{dist}(x, \mathbb{R}^N \backslash W)}{\operatorname{dist}(x, E) + \operatorname{dist}(x, \mathbb{R}^N \backslash W)} .$$

La fonction φ est continue, avec $0 \leqslant \varphi \leqslant 1$. On a $\varphi(x) = 1$ si $x \in E$ car $\operatorname{dist}(x, E) = 0$; et $\varphi(x) = 0$ si $x \notin W$, donc $\varphi \equiv 0$ en dehors de U car $W \subseteq U$. Enfin, φ est nulle en dehors de $K := \overline{W}$ qui est compact et contenu dans U, donc dans Ω ; donc $\varphi \in \mathcal{C}_{00}(\Omega)$. □

Preuve du Théorème 4.3. Comme on sait que les fonctions étagées sont denses dans $L^p(\Omega)$, il suffit de montrer que l'adhérence de $\mathcal{C}_{00}(\Omega)$ contient toutes les fonctions étagées appartenant à L^p. De plus, $\mathcal{C}_{00}(\Omega)$ est un espace vectoriel, donc $\overline{\mathcal{C}_{00}(\Omega)}^{L^p}$ aussi. Il suffit donc de montrer que $\overline{\mathcal{C}_{00}(\Omega)}^{L^p}$ contient toutes les *fonctions indicatrices* appartenant à $L^p(\Omega)$; autrement dit (comme $p < \infty$) toutes les fonctions de la forme $\mathbf{1}_A$ avec $A \subseteq \Omega$ et $\lambda_N(A) < \infty$.

On fixe donc un borélien $A \subseteq \Omega$ vérifiant $\lambda_N(A) < \infty$, et $\varepsilon > 0$. Il s'agit de trouver une fonction $\varphi \in \mathcal{C}_{00}(\Omega)$ telle que $\|\varphi - \mathbf{1}_A\|_p < \varepsilon$.

Par le Lemme 4.4, on peut trouver un compact E et un ouvert U tels que $E \subseteq A \subseteq U$ et $\lambda_N(U \backslash E) < \varepsilon$; et par le Lemme 4.5, on peut trouver une fonction $\varphi \in \mathcal{C}_{00}(\Omega)$ telle que $\varphi \equiv 1$ sur E et $\varphi \equiv 0$ en dehors de U, avec $0 \leqslant \varphi \leqslant 1$. Quitte à remplacer U par $U \cap \Omega$, on peut aussi supposer que $U \subseteq \Omega$.

Par définition de φ, on a $\mathbf{1}_E \leqslant \varphi \leqslant \mathbf{1}_U$; et on a aussi $\mathbf{1}_E \leqslant \mathbf{1}_A \leqslant \mathbf{1}_U$ car $E \subseteq A \subseteq U$. Donc $|\varphi - \mathbf{1}_A| \leqslant \mathbf{1}_U - \mathbf{1}_E$. Mais $\mathbf{1}_U - \mathbf{1}_E = \mathbf{1}_{U \backslash E}$ car $E \subseteq U$, donc

$$|\varphi - \mathbf{1}_A| \leqslant \mathbf{1}_{U \backslash E} .$$

On en déduit

$$\begin{aligned} \|\varphi - \mathbf{1}_A\|_p^p &= \int_\Omega |\varphi - \mathbf{1}_A|^p \, d\lambda_N \\ &\leqslant \int_\Omega \mathbf{1}_{U \backslash E}^p \, d\lambda_N \\ &= \lambda_N(U \backslash E) < \varepsilon . \end{aligned}$$

C'est la conclusion souhaitée, avec $\varepsilon^{1/p}$ au lieu de ε. □

Exercice. Montrer que $\mathcal{C}_{00}(\Omega)$ *n'est pas dense* dans $L^\infty(\Omega)$. Plus précisément, montrer que la fonction constante $\mathbf{1}$ n'est pas dans l'adhérence de $\mathcal{C}_{00}(\Omega)$ pour la norme L^∞.

Chapitre 9

Convolution et transformation de Fourier

1. Translations et symétries sur $L^p(\mathbb{R})$

DÉFINITION 1.1. *Soit $a \in \mathbb{R}$. Pour toute fonction $f : \mathbb{R} \to \mathbb{C}$, on note $\tau_a f$ la fonction définie par*

$$\tau_a f\,(t) := f(t-a)\,.$$

*On dit que $\tau_a f$ est la **translatée de f par** a.*

LEMME 1.2. *Soit $a \in \mathbb{R}$. Si $f \in L^p(\mathbb{R})$, $1 \leqslant p \leqslant \infty$, alors $\tau_a f \in L^p(\mathbb{R})$ et $\|\tau_a f\|_p = \|f\|_p$.*

Démonstration. Si $p < \infty$, on a par changement de variable

$$\|\tau_a f\|_p^p = \int_{\mathbb{R}} |f(t-a)|^p dt = \int_{\mathbb{R}} |f(u)|^p du = \|f\|_p^p\,.$$

Le cas $p = \infty$ est laissé en exercice. □

THÉORÈME 1.3. (continuité des translations)
Si $1 \leqslant p < \infty$ et si $f \in L^p(\mathbb{R})$ est fixée, alors l'application $u \mapsto \tau_u f$ est uniformément continue *de $\mathbb{R}$ dans $L^p(\mathbb{R})$.*

Démonstration. Si $u, v \in \mathbb{R}$, alors

$$\begin{aligned} \|\tau_u f - \tau_v f\|_p^p &= \int_{\mathbb{R}} |f(t-u) - f(t-v)|^p dx \\ &= \int_{\mathbb{R}} |(f(x) - f(x-(v-u))|^p dx \\ &= \|\tau_h f - f\|_p^p \quad \text{où } h := v-u\,. \end{aligned}$$

Donc il suffit de montrer que l'application $h \mapsto \tau_h f$ est *continue en* 0.

CAS 1. On suppose f *continue à support compact.*

Soit $A < \infty$ tel que $f \equiv 0$ en dehors de $[-A, A]$. Si $|h| \leqslant 1$, alors $\tau_h f(t) = f(t-h) \equiv 0$ en dehors de $[-(A+1), A+1]$; et de même pour f. Donc

$$\|\tau_h f - f\|_p^p = \int_{-(A+1)}^{A+1} |f(t-h) - f(t)|^p dt\,.$$

Le second membre tend vers 0 quand $h \to 0$ par convergence dominée, car f est continue bornée et l'intervalle d'intégration est borné. Donc $\tau_h f \to f$ dans L^p.

CAS 2. Cas général.

Soit $f \in L^p(\mathbb{R})$ quelconque. Comme les fonctions continues à support compact sont denses dans $L^p(\mathbb{R})$ (car $p < \infty$), on peut trouver une suite $(\varphi_n) \subseteq \mathcal{C}_{00}(\mathbb{R})$ telle que $\varphi_n \to f$ dans L^p.

Pour $h \in \mathbb{R}$ et pour tout $n \in \mathbb{N}$, on a

$$\|\tau_h \varphi_n - \tau_h f\|_p = \|\tau_h(\varphi_n - f)\|_p = \|\varphi_n - f\|_p\,.$$

Donc $\tau_h\varphi_n \to \tau_h f$ quand $n \to \infty$, *uniformément par rapport à* $h \in \mathbb{R}$. Comme les applications $h \mapsto \tau_h\varphi_n$ sont continues en 0 par le Cas 1, on en déduit que l'application $h \mapsto \tau_h f$ est également continue en 0. □

Remarque. Il est bon de retenir le principe de la démonstration précédente, qui est très général et très efficace : on commence par traiter le cas d'une fonction f "régulière", puis on conclut "par approximation".

Exercice 1. Soit $f := \mathbf{1}_{[0,1]} \in L^\infty(\mathbb{R})$. Montrer que l'application $u \mapsto \tau_u f$ n'est *pas* continue de $\mathbb{R}$ dans $L^\infty(\mathbb{R})$.

Exercice 2. Soient X et Y deux espaces vectoriels normés, et soit (T_k) une suite d'applications linéaires continues de X dans Y. On suppose que la suite (T_k) est bornée, et que $T_k(z) \to 0$ pour tout $z \in D$, où D est une partie dense de X. Montrer que $T_k(x) \to 0$ pour tout $x \in D$. Que fait cet exercie ici ?

Définition 1.4. *Pour toute fonction $f : \mathbb{R} \to \mathbb{C}$, on notera f_- la fonction définie par $f_-(t) := f(-t)$. On dit que f_- est la* ***symétrisée*** *de la fonction f.*

La preuve du lemme suivant est laissée en exercice.

Lemme 1.5. *Si $f \in L^p(\mathbb{R})$, $1 \leqslant p \leqslant \infty$, alors $f_- \in L^p(\mathbb{R})$ et $\|f_-\|_p = \|f\|_p$.*

2. Produit de convolution

2.1. Définition. Soient $f, g : \Omega \to \mathbb{C}$ deux fonctions boréliennes. Pour tout $x \in \mathbb{R}$ *tel que la fonction $t \mapsto f(x-t)g(t)$ est intégrable sur* $\mathbb{R}$, on pose

$$f * g\,(x) := \int_{\mathbb{R}} f(x-t)g(t)\,dt\,.$$

La fonction $f * g$ (dont le domaine de définition est éventuellement vide) s'appelle la **convolée** des fonctions f et g.

Remarque 1. Comme $f(x-t) = f(-(t-x)) = f_-(t-x) = \big(\tau_x f_-\big)(t)$, on a

$$f * g\,(x) = \int_{\mathbb{R}} \big(\tau_x f_-\big)(t) g(t)\,dt$$

en tout point x où $f * g\,(x)$ est bien défini. (Cette formule assez vilaine est en fait très utile.)

Remarque 2. Si $f * g$ est bien définie en un certain point x, alors $g * f$ aussi et

$$f * g\,(x) = g * f\,(x)\,.$$

Démonstration. Pour tout $x \in \mathbb{R}$, on a $\int_{\mathbb{R}} |f(x-t)g(t)|\,dt = \int_{\mathbb{R}} |f(u)g(u-x)|\,du$ par changement de variable ; donc $f * g\,(x)$ est bien défini si et seulement si $g * f\,(x)$ l'est, et dans ce cas le même calcul donne $f * g\,(x) = g * f\,(x)$. □

2.2. Quelques cas où $f * g$ est bien définie. Dans ce qui suit, f et g sont deux fonctions boréliennes sur $\mathbb{R}$, à valeurs complexes.

PROPOSITION 2.1. *Si f est* localement intégrable *et si g est* dans L^∞ et à support compact, *alors $f * g\,(x)$ est bien défini pour tout $x \in \mathbb{R}$.*

Démonstration. Soit $A < \infty$ tel que $g \equiv 0$ en dehors de $[-A, A]$. Si $x \in \mathbb{R}$ est donné, alors $f(x-t)g(t) = 0$ si $t \notin [-A, A]$, et $|f(x-t)g(t)| \leqslant \|g\|_\infty \, |f(x-t)|$ pour presque tout $t \in \mathbb{R}$. Donc

$$\begin{aligned}
\int_{\mathbb{R}} |f(x-t)g(t)|\, dt &= \int_{-A}^{A} |f(x-t)g(t)|\, dt \\
&\leqslant \|g\|_\infty \int_{-A}^{A} |f(x-t)|\, dt \\
&= \|g\|_\infty \int_{x-A}^{x+A} |f(u)|\, du \\
&< \infty \qquad \text{car } f \text{ est localement intégrable.}
\end{aligned}$$

□

Exercice. Montrer que si f et g sont toutes les deux à support compact, alors $f * g$ est à support compact.

Exemple. Soient $a, b \in \mathbb{R}$ vérifiant $0 < a \leqslant b$. La convolée $\mathbf{1}_{[-a,a]} * \mathbf{1}_{[-b,b]}$ est une fonction “trapèze”, donnée par la formule suivante :

$$\mathbf{1}_{[-a,a]} * \mathbf{1}_{[-b,b]}\,(x) = \begin{cases} x + a + b & \text{si} \quad -(a+b) \leqslant x \leqslant a - b \\ 2a & \text{si} \quad a - b \leqslant x \leqslant b - a \\ a + b - x & \text{si} \quad b - a \leqslant x \leqslant a + b \\ 0 & \text{si} \quad |x| \geqslant a + b \end{cases}$$

(Si $a = b$, le trapèze devient un triangle...)

Démonstration. Par définition, on a

$$\begin{aligned}
\mathbf{1}_{[-a,a]} * \mathbf{1}_{[-b,b]}\,(x) &= \int_{\mathbb{R}} \mathbf{1}_{[-a,a]}(t)\mathbf{1}_{[-b,b]}(x-t)\, dt \\
&= \int_{\mathbb{R}} \mathbf{1}_{[-a,a]}(t)\mathbf{1}_{[x-b,x+b]}(t) \\
&= \big|[-a, a] \cap [x-b, x+b]\big|.
\end{aligned}$$

Après un début de migraine, on en déduit le résultat annoncé. □

PROPOSITION 2.2. *Si $f \in L^p(\mathbb{R})$ et $g \in L^q(\mathbb{R})$, où $p, q \in [1, \infty]$ sont des exposants conjugués, alors $f * g\,(x)$ est bien défini partout, et $f * g$ est une fonction* continue bornée, *avec*

$$\|f * g\|_\infty \leqslant \|f\|_p \, \|g\|_q \, .$$

Démonstration. Pour $x \in \mathbb{R}$ donné, on a $f(x-t)g(t) = \tau_x f_-(t)g(t)$, où $f_-(u) = f(-u)$. Comme $\tau_x f_- \in L^p$ et $g \in L^q$, il découle de l'inégalité de Hölder que $\big(\tau_x f_-\big)g \in L^1$; donc $f * g\,(x)$ est bien défini. De plus, on a (toujours par l'inégalité de Hölder)

$$|f * g\,(x)| = \left|\int_{\mathbb{R}} \big(\tau_x f_-\big)(t)g(t)\, dt\right| \leqslant \int_{\mathbb{R}} \big|\big(\tau_x f_-\big)(t)g(t)\big|\, dt = \big\|\big(\tau_x f_-\big)g\big\|_1 \leqslant \|\tau_x f_-\|_p \, \|g\|_q \, .$$

Donc $f * g$ est bornée, avec $\|f * g\|_\infty \leqslant \|f\|_p \, \|g\|_q$.

Pour la continuité, on introduit la forme linéaire $\Phi_g : L^p \to \mathbb{C}$ définie par

$$\Phi_g(h) := \int_{\mathbb{R}} h(t)g(t)\,dt\,.$$

Par Hölder, cette forme linéaire est continue (avec $\|\Phi_g\| \leqslant \|g\|_q$) ; et on a par définition :

$$f * g\,(x) = \int_{\mathbb{R}} \big(\tau_x f_-\big)(t)g(t)\,dt = \Phi_g\big(\tau_x f_-\big)\,.$$

Comme l'application $x \mapsto \tau_x f_-$ est continue de $\mathbb{R}$ dans L^p (continuité des translations), on en déduit par composition que $f * g$ est continue. □

Corollaire 2.3. *Si f est localement intégrable et si g est dans L^∞ et à support compact, alors $f * g$ est continue.*

Démonstration. Par la Proposition 2.1, $f * g$ est bien définie en tout point. Pour la continuité, il suffit de montrer que $f * g$ est continue sur tout intervalle $[-B, B]$, $B > 0$. On fixe donc $B > 0$; et on fixe également $A < \infty$ tel que $g \equiv 0$ en dehors de $[-A, A]$.

On a pour tout $x \in \mathbb{R}$:

$$f * g\,(x) = \int_{-A}^{A} f(x-t)g(t)\,dt\,.$$

Si de plus $x \in [-B, B]$, alors $t - x \in [-(A+B), A+B] := K$ pour tout $t \in [-A, A]$, et donc on peut écrire $f(x-t) = \big(\mathbf{1}_K f\big)(t-x)$. Par conséquent :

$$\begin{aligned} f * g\,(x) &= \int_{-A}^{A} \big(\mathbf{1}_K f\big)(t-x)g(t)\,dt \\ &= \int_{\mathbb{R}} \big(\mathbf{1}_K f\big)(t-x)g(t)\,dt \\ &= \big(\mathbf{1}_K f\big) * g\,(x)\,. \end{aligned}$$

Ainsi, on a $f * g \equiv \big(\mathbf{1}_K f\big) * g$ sur $[-B, B]$. Mais comme $\mathbf{1}_K f \in L^1$ (car f est localement intégrable et K est compact) et $g \in L^\infty$, la fonction $\big(\mathbf{1}_K f\big) * g$ est continue sur $\mathbb{R}$ par la Proposition 2.2 ; donc $f * g$ est continue sur $[-B, B]$. □

Exercice 1. Soit $A \subseteq \mathbb{R}$ un borélien de mesure de Lebesgue non nulle. En considérant la convolée $\mathbf{1}_A * \mathbf{1}_{-A}$, montrer que l'ensemble $A - A := \{x - y;\ x, y \in A\}$ est un voisinage de 0.

Exercice 2. Soient p, q des exposants conjugués tels que $1 < p, q < \infty$. Montrer que si $f \in L^p(\mathbb{R})$ et $g \in L^q(\mathbb{R})$, alors $f * g$ *tend vers* 0 *à l'infini.* (Commencer par le cas où f et g sont continues à support compact.)

Voici maintenant un cas d'existence un peu plus subtil.

Théorème 2.4. *Si $f \in L^1(\mathbb{R})$ et si $g \in L^p(\mathbb{R})$, $1 \leqslant p < \infty$, alors $f * g$ est bien définie* presque partout *et appartient à $L^p(\mathbb{R})$, avec*

$$\|f * g\|_p \leqslant \|f\|_1\, \|g\|_p\,.$$

Démonstration. (i) On va traiter à part le cas $p = 1$, qui est plus élémentaire (même si les idées sont les mêmes pour p quelconque).

Soit h la fonction borélienne positive définie par

$$h(x) := \int_{\mathbb{R}} |f(x-t)g(t)|\, dt\,.$$

Par le théorème de Fubini, on a

$$\begin{aligned}
\int_{\mathbb{R}} h(x)\, dx &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x-t)g(t)|\, dt \right) dx \\
&= \int_{\mathbb{R}} |g(t)| \underbrace{\left(\int_{\mathbb{R}} |f(x-t)|\, dx \right)}_{\|f\|_1} dt \\
&= \|f\|_1 \int_{\mathbb{R}} |g(t)|\, dt \\
&= \|f\|_1\, \|g\|_1\,.
\end{aligned}$$

En particulier, $\int_{\mathbb{R}} h(x)\, dx < \infty$. Donc $h(x) < \infty$ presque partout ; ce qui signifie que $f * g\,(x)$ est bien défini presque partout.

La fonction $f * g$ est λ_1-mesurable car l'ensemble $E := \{h < \infty\}$ est borélien et $f * g$ est borélienne sur E (*cf* la preuve du théorème de Fubini). De plus, on a

$$|f * g\,(x)| = \left| \int_{\mathbb{R}} f(x-t)g(t)\, dt \right| \leqslant \int_{\mathbb{R}} |f(x-t)g(t)|\, dt = h(x) \quad \text{presque partout;}$$

donc $f * g \in L^1$ avec $\|f * g\|_1 \leqslant \|h\|_1 = \|f\|_1\, \|g\|_1$.

(ii) Traitons maintenant le cas général, où $p \in [1, \infty[$ est quelconque.

Comme dans le cas $p = 1$, on considère la fonction borélienne positive h définie par

$$h(x) := \int_{\mathbb{R}} |f(x-t)g(t)|\, dt = \int_{\mathbb{R}} |f(t)g(t-x)|\, dt\,.$$

Par définition :

$$h(x) = \int_{\mathbb{R}} \Phi_t(x)\, dt \qquad \text{avec } \Phi_t(x) := |f(t)\, \tau_t g(x)|\,.$$

D'après la forme intégrale de l'inégalité de Minkowski (Corollaire 3.9 du Chapitre 8), on a donc

$$\begin{aligned}
\|h\|_p &\leqslant \int_{\mathbb{R}} \|\Phi_t\|_p\, dt \\
&= \int_{\mathbb{R}} |f(t)|\, \|\tau_t g\|_p\, dt \\
&= \|g\|_p \int_{\mathbb{R}} |f(t)|\, dt \\
&= \|f\|_1\, \|g\|_p\,.
\end{aligned}$$

En particulier $\|h\|_p < \infty$; donc $|h(x)| < \infty$ presque partout, *i.e.* $f * g\,(x)$ est bien défini presque partout. Enfin (comme dans (i)) $f*g$ est λ_1-mesurable et $|f*g\,(x)| \leqslant h(x)$ presque partout, donc $f * g \in L^p$ avec $\|f * g\|_p \leqslant \|h\|_p \leqslant \|f\|_1\, \|g\|_p$.

□

Exercice. Montrer que si $f, g, h \in L^1(\mathbb{R})$, alors $(f * g) * h = f * (g * h)$. (Expliquer d'abord pourquoi ceci a un sens, et quel sens.)

2.3. Régularisation. Le théorème suivant montre que le produit de convolution permet de "régulariser" des fonction éventuellement très discontinues.

THÉORÈME 2.5. *Soit $N \in \mathbb{N}^*$. Si $f : \mathbb{R} \to \mathbb{C}$ est localement intégrable et si $g : \mathbb{R} \to \mathbb{C}$ est de classe $\mathcal{C}^N$ à support compact, alors $f * g$ est de classe $\mathcal{C}^N$ sur $\mathbb{R}$, avec pour $k = 1, \ldots, N$:*

$$(f * g)^{(k)} = f * (g^{(k)}) .$$

Démonstration. Par une récurrence immédiate, il suffit de traiter le cas $N = 1$; soient donc f localement intégrable et g de classe $\mathcal{C}^1$ à support compact.

L'idée est d'écrire

$$f * g\,(x) = g * f\,(x) = \int_{\mathbb{R}} f(t)g(x-t)\,dt\,,$$

de façon à faire apparaitre le x dans la fonction régulière g et pas dans la fonction f, dont on ne sait rien.

Cela étant fait, on applique le théorème de dérivabilité pour les intégrales à paramètres. Si on pose $F(x,t) := f(t)g(x-t)$, alors $F(x,t)$ est intégrable par rapport à t pour tout x fixé (car f est localement intégrable et g est bornée à support compact), et de classe $\mathcal{C}^1$ par rapport à x, avec

$$\frac{\partial F}{\partial x}(x,t) = f(t)g'(x-t)\,.$$

Pour vérifier l'hypothèse de domination, fixons un compact $K \subseteq \mathbb{R}$, et $B_K > 0$ tel que $K \subseteq [-B_K, B_K]$. Fixons également $A > 0$ tel que $g \equiv 0$ en dehors de $[-A, A]$.

On a $g'(u) \equiv 0$ en dehors de $[-A, A]$. Donc, si $x \in K \subseteq [-B_K, B_K]$, alors $g'(x-t) = 0$ pour tout $t \notin [-(A + B_K), A + B_K] := J_K$. On en déduit

$$\left|\frac{\partial F}{\partial x}(x,t)\right| \leqslant g_K(t) \quad \text{pour tout } x \in K\,,$$

où la fonction g_K est définie par

$$g_K(t) := \begin{cases} 0 & \text{si } t \notin J_K\,, \\ \|g'\|_\infty\,|f(t)| & \text{si } t \in J_K\,. \end{cases}$$

La fonction g_K est intégrable sur $\mathbb{R}$ car f est intégrable sur l'intervalle *compact* J_K. Donc le théorème de dérivabilité s'applique : la fonction $f * g$ est de classe $\mathcal{C}^1$, avec

$$(f * g)'(x) = \int_{\mathbb{R}} f(t)g'(x-t)\,dt = g * f'(x) = f * g'(x)\,.$$

□

REMARQUE 2.6. Le théorème n'est pas vide : il existe bel et bien des fonctions de classe $\mathcal{C}^N$ à support compact. Plus précisément, pour tout intervalle ouvert borné non trivial $I \subseteq \mathbb{R}$, on peut trouver une fonction $\chi \in \mathcal{C}^\infty(\mathbb{R})$ positive, nulle en dehors de I et vérifiant $\chi(t) > 0$ sur I.

Démonstration. Soit $\theta : \mathbb{R} \to \mathbb{R}$ la fonction définie par $\theta(t) := 0$ si $t \leqslant 0$ et $\theta(t) := e^{-1/t}$ si $t > 0$. La fonction θ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}\backslash\{0\}$, et on vérifie très facilement que les dérivées $\theta^{(k)}(t)$ pour $t > 0$ sont de la forme $\theta^{(k)}(t) = P_k(1/t)e^{-1/t}$, où P_k est un polynôme. Donc $\theta^{(k)}(t) \to 0$ quand $t \to 0^+$ pour tout $k \in \mathbb{N}$, et par conséquent θ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ avec $\theta^{(k)}(0) = 0$ pour tout $k \in \mathbb{N}$.

Si on maintenant on écrit $I =]a, b[$ et si on pose $\chi(t) := \theta(t-a)\theta(b-t)$, alors χ est de classe $\mathcal{C}^\infty$, identiquement nulle en dehors de $I =]a, b[$, car $\theta(t-a) = 0$ si $t \leqslant a$ et $\theta(b-t) = 0$ si $t \geqslant b$, et on a $\chi(t) > 0$ sur I car $\theta(s) > 0$ pour tout $s > 0$. $\square$

Exercice. Montrer que pour tout intervalle non trivial $I \subseteq \mathbb{R}$ de centre m, on peut trouver une fonction $\chi \in \mathcal{C}^\infty(\mathbb{R})$ positive, nulle en dehors de I et vérifiant $\chi(t) \equiv 1$ au voisinage de m. (On pourra commencer par montrer, en primitivant une fonction $\mathcal{C}^\infty$ à support compact convenable, qu'il existe une fonction $\phi \in \mathcal{C}^\infty(\mathbb{R})$ telle que $\phi(t) = 0$ pour tout $t \leqslant 0$ et $\phi(t) = 1$ pour tout $t \geqslant 1$; puis montrer qu'il existe une fonction $\psi \in \mathcal{C}^\infty(\mathbb{R})$ identiquement nulle en dehors de $]0, 4[$ telle que $\psi(t) \equiv 1$ sur $[1, 3]$.)

2.4. Approximation par convolution.

Définition 2.7. *Soit $(k_n)_{n\in\mathbb{N}}$ une suite dans $L^1(\mathbb{R})$. On dit que (k_n) est une **suite de Dirac** si elle vérifie les propriétés suivantes.*

(i) $\int_{\mathbb{R}} k_n(t)\, dt = 1$ *pour tout* $n \in \mathbb{N}$.

(ii) (k_n) *est **bornée dans** L^1, i.e. il existe une constante M telle que $\|k_n\|_1 \leqslant M$ pour tout $n \in \mathbb{N}$;*

(iii) *Pour tout $\delta > 0$, on a* $\lim_{n\to\infty} \int_{\{|t|\geqslant\delta\}} |k_n(t)|\, dt = 0.$

Remarque. La condition (iii) signifie que "k_n se concentre autour de 0" quand $n \to \infty$.

Exemple. Soit $k \in L^1(\mathbb{R})$ vérifiant $\int_{\mathbb{R}} k(t)\, dt = 1$. Si (λ_n) est une suite de nombre réels strictement positifs *tendant vers* ∞, alors la suite (k_n) définie par

$$k_n(t) := \lambda_n\, k(\lambda_n t)$$

est une suite de Dirac.

Démonstration. Le changement de variable $u = \lambda_n t$ donne

$$\int_{\mathbb{R}} |k_n(t)|\, dt = \int_{\mathbb{R}} |k(\lambda_n u)|\, \lambda_n dt = \int_{\mathbb{R}} |k(u)|\, du = \|k\|_1\, .$$

Donc $k_n \in L^1(\mathbb{R})$ avec $\|k_n\|_1 = \|k\|_1$ pour tout $n \in \mathbb{N}$; en particulier, (ii) est vérifiée. Le même changement de variable $u = \lambda_n t$ donne $\int_{\mathbb{R}} k_n(t)\, dt = \int_{\mathbb{R}} k(u)\, du = 1$ pour tout $n \in \mathbb{N}$, *i.e.* (i) est vérifiée.

Pour vérifier (iii), fixons $\delta > 0$. On a

$$\int_{\{|t|\geqslant\delta\}} |k_n(t)|\, dt = \int_{\{|t|\geqslant\delta\}} |k_n(\lambda_n t)|\, \lambda_n dt = \int_{\{|u|\geqslant\lambda_n\delta\}} |k(u)|\, du\, .$$

Comme $\lambda_n\delta \to \infty$ et comme k est intégrable sur $\mathbb{R}$, l'intégrale

$$\int_{\{|u|\geqslant\lambda_n\delta\}} |k(u)|\, du = \int_{\mathbb{R}} |k(u)|\, du - \int_{-\lambda_n\delta}^{\lambda_n\delta} |k(u)|\, du$$

tend vers 0 quand $n \to \infty$; donc (iii) est vérifiée. $\square$

L'intérêt des suites de Dirac vient du théorème d'approximation suivant.

Théorème 2.8. *Soit $(k_n) \subseteq L^1(\mathbb{R})$ une suite de Dirac, et soit $f : \mathbb{R} \to \mathbb{C}$ borélienne.*

(1) *Si f est* bornée *et* continue *en un point* $x_0 \in \mathbb{R}$, *alors* $k_n * f(x_0) \to f(x_0)$ *quand* $n \to \infty$.
(2) *Si f est* bornée et uniformément continue, *alors* $k_n * f \to f$ uniformément.
(3) *Si* $f \in L^p(\mathbb{R})$ *avec* $1 \leqslant p < \infty$, *alors* $k_n * f \to f$ en norme L^p.

Remarque 1. Notons $\mathcal{C}_b(\mathbb{R})$ l'espace des fonctions continues bornées sur $\mathbb{R}$, muni de la norme $\|\cdot\|_\infty$. Toute mesure borélienne finie μ sur $\mathbb{R}$ définit une forme linéaire continue sur $\mathcal{C}_b(\mathbb{R})$, notée $f \mapsto \langle \mu, f\rangle$, par la formule

$$\langle \mu, f\rangle := \int_{\mathbb{R}} f\, d\mu\,.$$

Par ailleurs, toute fonction $g \in L^1(\mathbb{R})$ définit également une forme linéaire continue sur $\mathcal{C}_b(\mathbb{R})$, notée $f \mapsto \langle g, f\rangle$, par la formule

$$\langle g, f\rangle := \int_{\mathbb{R}} f(t)\, g(t) dt\,.$$

Avec cette notation, on a $g * f(0) = \int_{\mathbb{R}} g(t) f(-t)\, dt = \langle g, f_-\rangle$ pour toute $f \in \mathcal{C}_b(\mathbb{R})$. Donc, la partie (1) du théorème, appliquée avec $x_0 := 0$ et f_- au lieu de f, dit que

$$\langle k_n, f\rangle \to \langle \delta_0, f\rangle \quad \text{pour toute } f \in \mathcal{C}_b(\mathbb{R})\,.$$

Ainsi, la suite (k_n) “tend vers la masse de Dirac δ_0” en un sens bien précis ; d'où la terminologie “suite de Dirac”.

Remarque 2. En raison de (3), les suites de Dirac s'appellent souvent des **unités approchées** pour la convolution.

Exercice. Montrer que $L^1(\mathbb{R})$ n'a pas d'unité pour la convolution ; autrement dit, qu'il n'existe pas de fonction $k \in L^1(\mathbb{R})$ telle que $k * f = f$ pour toute $f \in L^1(\mathbb{R})$. (Montrer qu'une telle fonction k devrait vérifier $\int_{x-1}^{x+1} k(u)\, dt = \mathbf{1}_{[-1,1]}(x)$ presque partout, et utiliser le fait que la fonction $x \mapsto \int_{x-1}^{x+1} k(u)\, du$ est continue sur $\mathbb{R}$.)

Preuve du Théorème 2.8. Tout va découler assez facilement du fait suivant.

Fait. Si $\varphi : \mathbb{R} \to \mathbb{R}^+$ est borélienne bornée et vérifie $\lim\limits_{t\to 0} \varphi(t) = 0$, alors

$$\int_{\mathbb{R}} |k_n(t)|\, \varphi(t)\, dt \to 0 \quad \text{quand } n \to \infty\,.$$

Preuve du Fait. Pour pouvoir utiliser la propriété (iii) dans la définition d'une suite de Dirac, on coupe l'intégrale en deux : pour tout $\delta > 0$ et pour tout $n \in \mathbb{N}$, on a

$$\begin{aligned}
\int_{\mathbb{R}} |k_n(t)|\, \varphi(t)\, dt &= \int_{\{|t|<\delta\}} |k_n(t)|\, \varphi(t)\, dt + \int_{\{|t|\geqslant\delta\}} |k_n(t)|\, \varphi(t)\, dt \\
&\leqslant \sup_{|t|<\delta} \varphi(t) \int_{\{|t|<\delta\}} |k_n(t)|\, dt + \|\varphi\|_\infty \int_{\{|t|\geqslant\delta\}} |k_n(t)|\, dt \\
&\leqslant M \sup_{|t|<\delta} \varphi(t) + \|\varphi\|_\infty \int_{\{|t|\geqslant\delta\}} |k_n(t)|\, dt\,,
\end{aligned}$$

où la constante M est donnée par la propriété (ii) dans la définition d'une suite de Dirac. Comme $\int_{\{|t|\geqslant\delta\}} |k_n(t)|\, dt \to 0$ par (iii), on en déduit

$$\overline{\lim} \int_{\mathbb{R}} |k_n(t)| \varphi(t)\, dt \leqslant M \sup_{|t|<\delta} \varphi(t) \quad \text{pour tout } \delta > 0\,;$$

et donc $\int_{\mathbb{R}} |k_n(t)|\varphi(t)\, dt \to 0$ puisque $\varphi(t) \to 0$ quand $t \to 0$. □

On peut maintenant passer à la preuve du théorème.

(1) Remarquons d'abord que l'énoncé a un sens : comme $f \in L^\infty$ et $k_n \in L^1$, la convolée $k_n * f$ est bien définie *partout*, donc sa valeur en x_0 est parfaitement déterminée.

Le point clé est l'observation suivante : comme $\int_{\mathbb{R}} k_n = 1$, on peut écrire

$$f(x_0) = f(x_0) \times \int_{\mathbb{R}} k_n(t)\, dt\,,$$

de sorte que

$$\begin{aligned} k_n * f(x_0) - f(x_0) &= \int_{\mathbb{R}} k_n(t) f(x_0 - t)\, dt - f(x_0) \times \int_{\mathbb{R}} k_n(t)\, dt \\ &= \int_{\mathbb{R}} k_n(t)\big(f(x_0 - t) - f(x_0)\big)\, dt\,. \end{aligned}$$

On a donc

$$|k_n * f(x_0) - f(x_0)| \leqslant \int_{\mathbb{R}} |k_n(t)|\, \underbrace{|f(x_0 - t) - f(x_0)|}_{\varphi(t)}\, dt\,.$$

Le fonction φ est bornée car f l'est, et $\varphi(t) \to 0$ quand $t \to 0$ puisque f est continue en x_0 ; donc $k_n * f(x_0) \to f(x_0)$ par le Fait.

(2) La preuve est la même. Pour tout $x \in \mathbb{R}$, on a comme plus haut

$$\begin{aligned} |k_n * f(x) - f(x)| &\leqslant \int_{\mathbb{R}} |k_n(t)|\, |f(x - t) - f(x)|\, dt \\ &\leqslant \int_{\mathbb{R}} |k_n(t)| \times \underbrace{\sup\big\{|f(v) - f(u)|;\ |v - u| \leqslant |t|\big\}}_{\varphi(t)}\, dt\,. \end{aligned}$$

Le fonction φ est bornée car f l'est, et $\varphi(t) \to 0$ quand $t \to 0$ car f est uniformément continue. Donc l'intégrale de droite tend vers 0 par le Fait, et donc $k_n * f(x) - f(x) \to 0$ uniformément puisque la majoration ne dépend pas de x.

(3) Cette fois il faut être un peu prudent car $k_n * f$ n'est définie que *presque partout* ; mais l'idée est toujours la même.

On a pour presque tout $x \in \mathbb{R}$:

$$\begin{aligned} |k_n * f(x) - f(x)| &= \left| \int_{\mathbb{R}} \big(f(x - t) - f(x)\big) k_n(t)\, dt \right| \\ &\leqslant \int_{\mathbb{R}} |f(x - t) - f(x)|\, |k_n(t)|\, dt \\ &= \int_{\mathbb{R}} |k_n(t)|\, |\tau_t f(x) - f(x)|\, dt \\ &:= \int_{\mathbb{R}} \Phi_t(x)\, dt\,. \end{aligned}$$

D'après la forme intégrale de l'inégalité de Minkowski, on en déduit

$$\begin{aligned}\|k_n * f - f\|_p &\leqslant \int_{\mathbb{R}} \|\Phi_t\|_p \, dt \\ &= \int_{\mathbb{R}} |k_n(t)| \, \underbrace{\|\tau_t f - f\|_p}_{\varphi(t)} \, dt\end{aligned}$$

La fonction φ est bornée sur $\mathbb{R}$ car $\varphi(t) \leqslant 2\|f\|_p$, et $\varphi(t) \to 0$ quand $t \to 0$ par *continuité des translations* dans $L^p(\mathbb{R})$. Donc $\|k_n * f - f\|_p \to 0$ par le Fait. □

Corollaire 2.9. *Pour tout $p \in [1, \infty[$, les fonctions* de classe $\mathcal{C}^\infty$ à support compact *sont denses dans $L^p(\mathbb{R})$.*

Démonstration. Comme les fonctions *continues* à support compact sont denses dans $L^p(\mathbb{R})$ (car $p < \infty$), il suffit de montrer que l'adhérence de $\mathcal{C}_{00}^\infty(\mathbb{R})$ dans L^p contient $\mathcal{C}_{00}(\mathbb{R})$. On fixe donc une fonction $f \in \mathcal{C}_{00}(\mathbb{R})$, et on cherche une suite $(\varphi_n) \subseteq \mathcal{C}_{00}^\infty(\mathbb{R})$ qui converge vers f en norme L^p.

Soit $\chi \in \mathcal{C}_{00}^\infty(\mathbb{R})$ positive et non identiquement nulle. Alors $\int_{\mathbb{R}} \chi \neq 0$, donc on peut définir $k(t) := \frac{1}{\int_{\mathbb{R}} \chi} \chi(t)$. La fonction k est de classe $\mathcal{C}^\infty$ à support compact avec $\int_{\mathbb{R}} k(t) \, dt = 1$. Donc, si on pose $k_n(t) := n \, k(nt)$, alors (k_n) est une suite de Dirac formée de fonctions $\mathcal{C}^\infty$ à support compact.

Par le Théorème 2.8, $\varphi_n := k_n * f$ tend vers f en norme L^p. De plus, les φ_n sont de classe $\mathcal{C}^\infty$ par le Théorème 2.5. Enfin, chaque φ_n est *à support compact* car k_n et f le sont. (Si on choisit A et A_n tels que $f \equiv 0$ en dehors de $[-A, A]$ et $\varphi_n \equiv 0$ en dehors de $[-A_n, A_n]$, alors $k_n * f(x) = \int_{-A}^{A} k_n(x-t) f(t) \, dt$ vaut 0 en dehors de $[-A - A_n, A + A_n]$.) □

Exercice 1. Soit $f : \mathbb{R} \to \mathbb{C}$ localement intégrable. Montrer que si $I \subseteq \mathbb{R}$ est un intervalle borné de centre x et de longueur $2h$, alors

$$\int_I f(t) \, dt = \mathbf{1}_{[-h,h]} * f \, (x) \, .$$

En déduire que si on a $\int_I f(t) \, dt = 0$ pour tout intervalle borné $I \subseteq \mathbb{R}$, alors $f(x) = 0$ presque partout.

Exercice 2. Soit $f : \mathbb{R} \to \mathbb{C}$ localement intégrable, et soit $x_0 \in \mathbb{R}$. On pose

$$F(x) := \int_{x_0}^{x} f(t) \, dt \, .$$

Pour $h \in \mathbb{R}$ donné, exprimer $F(x + h) - F(x)$ comme une convolution ; puis montrer que de toute suite (h_n) de nombres réels non nuls tendant vers 0, on peut extraire une sous-suite (h_{n_k}) telle que

$$\frac{F(x + h_{n_k}) - F(x)}{h_{n_k}} \to f(x) \quad \text{presque partout} \, .$$

(En fait, il est inutile d'extraire une sous-suite car un résultat beaucoup plus fort est vrai : la fonction F est *dérivable* en presque tout point, avec $F' = f$ presque partout. C'est ce qu'on appelle le **théorème de dérivation de Lebesgue** ; qui est plus délicat à démontrer.)

3. Transformation de Fourier

3.1. Ce qu'on sait déjà. Si $f : \mathbb{R} \to \mathbb{C}$ est une fonction intégrable, sa **transformée de Fourier** est la fonction $\widehat{f} : \mathbb{R} \to \mathbb{C}$ définie par

$$\widehat{f}(x) := \int_{\mathbb{R}} f(t)\, e^{-ixt}\, dt\,.$$

On a déjà vu que $\widehat{f}$ est une fonction *continue bornée*, avec

$$\|\widehat{f}\|_\infty \leqslant \|f\|_1\,.$$

De plus, il est évident que l'application $f \mapsto \widehat{f}$ est *linéaire*. On peut résumer ces propriétés de la façon suivante.

Proposition 3.1. *Notons $\mathcal{C}_b(\mathbb{R})$ l'espace de toutes les fonctions continues bornées sur $\mathbb{R}$, et posons $\mathcal{F}f := \widehat{f}$ pour $f \in L^1(\mathbb{R})$. Alors la transformation de Fourier $\mathcal{F}$ est une application linéaire continue de $L^1(\mathbb{R})$ dans $\mathcal{C}_b(\mathbb{R})$, avec $\|\mathcal{F}\| \leqslant 1$.*

3.2. Propriétés formelles. La transformation de Fourer possède de nombreuses propriétés "formelles" agréables, qu'on va maintenant détailler.

Proposition 3.2. (Fourier et convolution)
La transformation de Fourier change le produit de convolution en produit ordinaire : si $f, g \in L^1(\mathbb{R})$, alors

$$\widehat{f * g} = \widehat{f}\,\widehat{g}\,.$$

Démonstration. On sait que $f * g$ est bien définie presque partout et appartient à $L^1(\mathbb{R})$. Un calcul formel donne

$$\begin{aligned}
\widehat{f * g}(x) &= \int_{\mathbb{R}} f * g\,(t) e^{-ixt}\, dt\\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(t-s) g(s)\, ds \right) e^{-ixt} dt\\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(t-s) e^{-ix(t-s)} g(s) e^{-ixs} ds \right) dt\\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(t-s) e^{-ix(t-s)} dt \right) g(s) e^{-ixs} ds \qquad \text{par Fubini}\\
&= \int_{\mathbb{R}} \underbrace{\left(\int_{\mathbb{R}} f(u) e^{-ixu} du \right)}_{\widehat{f}(x)} g(s) e^{-ixs} ds\\
&= \widehat{f}(x) \int_{\mathbb{R}} g(s) e^{-ixs} ds = \widehat{f}(x)\, \widehat{g}(x)\,.
\end{aligned}$$

Pour justifier l'utilisation du théorème de Fubini, il suffit de vérifier qu'on a

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(t-s) e^{-ix(t-s)} g(s) e^{-ixs}|\, ds \right) dt < \infty\,;$$

ce qui est facile en utilisant Fubini “positif” :

$$\begin{aligned}\int_{\mathbb{R}}\left(\int_{\mathbb{R}}|f(t-s)e^{-ix(t-s)}g(s)e^{-ixs}|\,ds\right)dt &= \int_{\mathbb{R}}\left(\int_{\mathbb{R}}|f(t-s)g(s)|\,ds\right)dt\\ &= \int_{\mathbb{R}}\left(\int_{\mathbb{R}}|f(t-s)|\,dt\right)|g(s)|\,ds\\ &= \|f\|_1\,\|g\|_1<\infty\,.\end{aligned}$$

□

Proposition 3.3. (Fourier et dérivation)
Si $f:\mathbb{R}\to\mathbb{C}$ est de classe $\mathcal{C}^N$, $N\geqslant 1$ et si $f,f',\dots,f^{(N)}$ sont dans $L^1(\mathbb{R})$ alors

$$\widehat{f^{(k)}}(x)=(ix)^k\,\widehat{f}(x)\quad \textit{pour tout } k\in\{0,\dots,N\}\,.$$

Démonstration. Par une récurrence immédiate, il suffit de traiter le cas $N=1$. On suppose donc que f est de classe $\mathcal{C}^1$ avec $f,f'\in L^1$, et on montre que $\widehat{f'}(x)=ix\,\widehat{f}(x)$.

Fait. On a $\lim\limits_{t\to\pm\infty}f(t)=0$.

Preuve du Fait. D'après le théorème fondamental de l'analyse, on peut écrire

$$f(t)=f(0)+\int_0^t f'(s)\,ds\,.$$

Donc $f(t)$ tend vers $l:=f(0)+\int_0^\infty f'(s)\,ds$ puisque f' est intégrable sur $[0,\infty[$. Si on avait $l\neq 0$, on pourrait trouver une constante $c>0$ (par exemple $c:=|l|/2$) et $t_0\in\mathbb{R}^+$ tels que $|f(t)|\geqslant c$ pour tout $t\geqslant t_0$; et on en déduirait $\int_{t_0}^\infty|f(t)|\,dt\geqslant\int_{t_0}^\infty c\,dt=\infty$, ce qui est exclu puique f est supposée intégrable sur $\mathbb{R}$. Donc $l=0$, *i.e.* $f(t)\to 0$ quand $t\to\infty$. On montre de même que $f(t)\to 0$ quand $t\to-\infty$. □

On peut maintenant calculer $\widehat{f'}(x)$ en intégrant par parties. Pour tout $A>0$, on a

$$\int_{-A}^{A}f'(t)e^{-ixt}dt=\left[f(t)e^{-ixt}\right]_{-A}^{A}+\int_{-A}^{A}f(t)\times(ixe^{-ixt})\,dt\,.$$

Le “crochet” tend vers 0 quand $A\to\infty$ d'après le Fait ; donc on obtient en faisant tendre A vers l'infini :

$$\widehat{f'}(x)=ix\int_{-\infty}^{\infty}f(t)e^{-ixt}dt=ix\,\widehat{f}(x)\,.$$

□

Corollaire 3.4. *Si $f:\mathbb{R}\to\mathbb{C}$ est de classe $\mathcal{C}^2$ à support compact, alors $\widehat{f}\in L^1(\mathbb{R})$.*

Démonstration. La fonction $\widehat{f}$ est continue, donc localement intégrable. Si $x\neq 0$, on a d'après le Fait :

$$|\widehat{f}(x)|=\frac{1}{x^2}\,|\widehat{f''}(x)|\,.$$

Donc $|\widehat{f}(x)|=O(1/x^2)$ quand $x\to\pm\infty$ car $\widehat{f''}$ est bornée, et par conséquent $\widehat{f}$ est intégrable sur $\mathbb{R}$. □

Exercice 3.5. On dit qu'une fonction $u:\mathbb{R}\to\mathbb{C}$ est **à décroissance rapide** si $t^k u(t)\to 0$ quand $t\to\pm\infty$, pour tout $k\in\mathbb{N}$. L'**espace de Schwartz**, noté $\mathcal{S}(\mathbb{R})$, est l'ensemble des fonctions $f:\mathbb{R}\to\mathbb{C}$ de classe $\mathcal{C}^\infty$ telles que f et toutes ses dérivées sont à décroissance rapide. Montrer que $\mathcal{S}(\mathbb{R})$ est “stable par Fourier” : si $f\in\mathcal{S}(\mathbb{R})$, alors $\widehat{f}\in\mathcal{S}(\mathbb{R})$.

Proposition 3.6. (Fourier et dilatations)
Soit $f \in L^1(\mathbb{R})$. Pour $\lambda > 0$, notons f_λ la fonction définie par

$$f_\lambda(t) := f(\lambda t)\,.$$

Alors $f_\lambda \in L^1(\mathbb{R})$ et

$$\widehat{f_\lambda}(x) = \frac{1}{\lambda}\,\widehat{f}\left(\frac{x}{\lambda}\right).$$

Démonstration. On a $\int_{\mathbb{R}} |f_\lambda(t)|\,dt = \int_{\mathbb{R}} |f(\lambda t)|\,dt = \int_{\mathbb{R}} |f(u)|\frac{du}{\lambda} = \frac{1}{\lambda}\,\|f\|_1 < \infty$; donc $f_\lambda \in L^1$. Ensuite,

$$\widehat{f_\lambda}(x) = \int_{\mathbb{R}} f(\lambda t)e^{-ixt}dt = \int_{\mathbb{R}} f(u)e^{-ix\frac{u}{\lambda}}\,\frac{du}{\lambda} = \frac{1}{\lambda}\,\widehat{f}\left(\frac{x}{\lambda}\right)\,.$$

□

Exercice. Exprimer $\widehat{\tau_a f}$ à l'aide de $\widehat{f}$ pour $a \in \mathbb{R}$.

3.3. Le "lemme de Riemann-Lebesgue". On sait que la transformée de Fourier d'une fonction intégrable est une fonction continue bornée. La proposition suivante en dit un peu plus.

Proposition 3.7. *Si $f \in L^1(\mathbb{R})$, alors $\widehat{f}(x) \to 0$ quand $x \to \pm\infty$.*

Démonstration. On va procéder "par approximation", comme dans la preuve du Théorème 1.3.

Cas 1. Cas où f est de classe $\mathcal{C}^1$ à support compact.

Dans ce cas, f' est intégrable sur $\mathbb{R}$ car continue à support compact. Par la Proposition 3.3, on a donc pour tout $x \neq 0$:

$$\widehat{f}(x) = \frac{1}{ix}\,\widehat{f'}(x)\,;$$

d'où le résultat car $\widehat{f'}$ est bornée sur $\mathbb{R}$.

Cas 2. Cas général.

Notons $\mathcal{C}_b(\mathbb{R})$ l'espace de toutes les fonctions continues bornées sur $\mathbb{R}$ muni de la norme $\|\cdot\|_\infty$, et $\mathcal{C}_0(\mathbb{R})$ le sous espace vectoriel de $\mathcal{C}_b(\mathbb{R})$ constitué par les fonctions tendant vers 0 à l'infini. Il n'est pas difficile de vérifier que $\mathcal{C}_0(\mathbb{R})$ est un sous-espace *fermé* de $\mathcal{C}_b(\mathbb{R})$ (exercice). Comme l'application $f \mapsto \widehat{f}$ est continue de $L^1(\mathbb{R})$ dans $\mathcal{C}_b(\mathbb{R})$, on en déduit que l'ensemble

$$\mathcal{D} := \{f \in L^1(\mathbb{R});\ \widehat{f} \in \mathcal{C}_0(\mathbb{R})\}$$

est une partie fermée de $L^1(\mathbb{R})$. Mais $\mathcal{D}$ est également *dense* dans $L^1(\mathbb{R})$ car il contient toutes les fonctions de classe $\mathcal{C}^1$ à support compact d'après le Cas 1. Donc $\mathcal{D} = L^1(\mathbb{R})$, ce qui est le résultat souhaité. □

3.4. La formule d'inversion. Le théorème qui suit dit essentiellement que si f est une fonction intégrable sur $\mathbb{R}$ dont la transformée de Fourier est elle aussi intégrable sur $\mathbb{R}$, alors f est la transfomée de Fourier de sa transformée de Fourier.

L'énoncé précis demande d'introduire la notation suivante : pour toute fonction $g \in L^1(\mathbb{R})$, on notera $\breve{g} : \mathbb{R} \to \mathbb{C}$ la fonction définie par

$$\breve{g}\,(t) := \int_{\mathbb{R}} g(x)\,e^{itx}\,dx\,.$$

(On dit parfois que $\breve{g}$ est la **transformée de Fourier inverse** de la fonction g.)

THÉORÈME 3.8. (formule d'inversion de Fourier)
Si $f \in L^1(\mathbb{R})$ est telle que $\widehat{f}$ appartient également à $L^1(\mathbb{R})$, alors

$$f = \frac{1}{2\pi}\,\breve{\widehat{f}} \quad \textit{presque partout}.$$

Autrement dit :

$$f(t) = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{f}(x)\, e^{itx}\, dx \quad \textit{presque partout}.$$

Démonstration. On a besoin de deux faits préliminaires.

FAIT 1. Si $f, g \in L^1(\mathbb{R})$, alors $g\widehat{f} \in L^1(\mathbb{R})$ et $\breve{\overline{g\widehat{f}}} = \breve{g} * f$.

Preuve du Fait 1. On a $g\widehat{f} \in L^1$ car $g \in L^1$ et $\widehat{f}$ est bornée. Ensuite, le calcul se fait tout seul :

$$\begin{aligned}
\breve{g\widehat{f}}\,(t) &= \int_{\mathbb{R}} g(x)\widehat{f}\,(x)\, e^{itx}\, dt \\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(s)e^{-ixs} ds\right) g(x)e^{itx} dx \\
&= \int_{\mathbb{R}} f(s)\left(\int_{\mathbb{R}} g(x)e^{ix(t-s)}\right) ds \qquad \text{par Fubini} \\
&= \int_{\mathbb{R}} f(s)\,\breve{g}\,(t-s)\, ds \\
&= f * \breve{g}\,(t) = \breve{g} * f\,(t)\,.
\end{aligned}$$

(La justification de l'emploi du théorème de Fubini est laissée en exercice.) □

FAIT 2. Il existe une suite $(g_n) \subseteq L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ vérifiant les propriétés suivantes :

(a) (g_n) est uniformément bornée et $g_n(x) \to 1$ pour tout $x \in \mathbb{R}$;
(b) $\breve{g}_n \in L^1(\mathbb{R})$ pour tout n, et la suite $\left(\frac{1}{2\pi}\breve{g}_n\right)$ est une suite de Dirac.

Preuve du Fait 2. Soit $g : \mathbb{R} \to \mathbb{R}$ la fonction gaussienne définie par

$$g(x) := e^{-x^2/2}.$$

La fonction g est intégrable sur $\mathbb{R}$, et on a vu au Chapitre 6 qu'on a

$$\breve{g}(t) = \int_{\mathbb{R}} e^{-x^2/2} e^{itx} dt = \sqrt{2\pi}\, e^{-t^2/2} \quad \text{pour tout } t \in \mathbb{R}\,.$$

En particulier, $\breve{g} \in L^1(\mathbb{R})$ et

$$\int_{\mathbb{R}} \breve{g}(t)\, dt = \sqrt{2\pi}\int_{\mathbb{R}} e^{-t^2/2}\, dt = 2\pi\,.$$

Posons alors

$$g_n(x) := g(\varepsilon_n x)\,,$$

où (ε_n) est n'importe quelle suite de réels strictement positifs tendant vers 0.

D'après la (version "Fourier inverse de" la) Proposition 3.6, on a

$$\breve{g}_n(t) = \frac{1}{\varepsilon_n}\,\breve{g}\left(\frac{t}{\varepsilon_n}\right)\,;$$

donc la suite $\left(\frac{1}{2\pi}\breve{g}_n\right)$ est une suite de Dirac puisque $\int_{\mathbb{R}} \frac{1}{2\pi}\,\breve{g} = 1$ et $1/\varepsilon_n \to \infty$.

De plus, la suite (g_n) est uniformément bornée car g est bornée, et $g_n(x) \to g(0) = 1$ pour tout $x \in \mathbb{R}$ car g est continue en 0. $\square$

On peut maintenant démontrer la formule d'inversion de Fourier. Soit $f \in L^1(\mathbb{R})$ telle que $\widehat{f} \in L^1$. Si la suite (g_n) est comme dans le Fait 2, on a d'après le Fait 1 :

$$\widecheck{g_n \widehat{f}} = \breve{g}_n * f \qquad \text{pour tout } n \in \mathbb{N}\,.$$

Comme (g_n) est uniformément bornée et tend simplement vers 1, et comme $\widehat{f} \in L^1$, on voit que

$$\widecheck{g_n \widehat{f}}\,(t) = \int_{\mathbb{R}} g_n(x)\widehat{f}(x)\, e^{itx}\, dt \to \int_{\mathbb{R}} \widehat{f}(x) e^{itx}\, dx \quad \text{pour tout } t \in \mathbb{R}\,, \tag{3.1}$$

d'après le théorème de convergence dominée.

De plus, d'après le Fait 2, $\breve{g}_n * f$ tend vers $2\pi f$ *en norme* L^1 ; donc on peut trouver une suite strictement croissante d'entiers (n_k) telle que

$$\breve{g}_{n_k} * f\,(t) \to 2\pi f(t) \qquad \text{presque partout}\,. \tag{3.2}$$

De (3.1) et (3.2), on déduit immédiatement le résultat souhaité :

$$f(t) = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{f}(x)\, e^{itx}\, dx \qquad \text{presque partout}\,.$$

$\square$

Corollaire 3.9. *La transformation de Fourier est* injective *sur* $L^1(\mathbb{R})$ *: si* $f_1, f_2 \in L^1(\mathbb{R})$ *vérifient* $\widehat{f_1} = \widehat{f_2}$*, alors* $f_1 = f_2$ *presque partout.*

Démonstration. Si $f \in L^1$ et $\widehat{f} = 0$, alors $f = 0$ presque partout d'après la formule d'inversion ; d'où le résultat par linéarité. $\square$

Corollaire 3.10. *Si* $f : \mathbb{R} \to \mathbb{C}$ *est une fonction* continue *intégrable sur* $\mathbb{R}$ *telle que* $\widehat{f} \in L^1(\mathbb{R})$*, alors*

$$f(t) = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{f}(x)\, e^{itx}\, dx \qquad \textit{pour } \text{tout } t \in \mathbb{R}\,.$$

Démonstration. On a $f = \frac{1}{2\pi} \widecheck{\widehat{f}}$ presque partout, donc *partout* car ces deux fonctions sont continues (exercice ; *cf* le Fait 1.6 du Chapitre 8). $\square$

Corollaire 3.11. *Si* $f \in L^1(\mathbb{R})$ *est telle que* $\widehat{f} \in L^1(\mathbb{R})$*, alors* f *est presque partout égale à une fonction continue.*

Exercice. Soit $f : \mathbb{R} \to \mathbb{R}$ la fonction (intégrable !) définie par $f(t) := e^{-t^4}$ si $t < 0$ et $f(t) := 2e^{-t^{10}}$ si $t \geqslant 0$. La transformée de Fourier de f est-elle intégrable sur $\mathbb{R}$?

Exemple. Comme illustration de la formule d'inversion de Fourier, on va montrer que pour tout $t \in \mathbb{R}$, on a

$$\int_{\mathbb{R}} \frac{e^{itx}}{1 + x^2}\, dx = \pi\, e^{-|t|}\,. \tag{3.3}$$

Soit $f : \mathbb{R} \to \mathbb{R}$ la fonction définie par

$$f(t) := e^{-|t|}\,.$$

La fonction f est clairement intégrable sur $\mathbb{R}$, et sa transformée de Fourier n'est pas difficile à calculer : on a

$$\begin{aligned}\widehat{f}(x) &= \int_{\mathbb{R}} e^{-|t|}\, e^{-ixt}\, dt \\ &= \int_{-\infty}^{0} e^{(1-ix)t}\, dt + \int_0^{\infty} e^{(-1-ix)t}\, dt \\ &= \frac{1}{1-ix} + \frac{1}{1+ix} \\ &= \frac{2}{1+x^2}\,.\end{aligned}$$

Donc $\widehat{f}$ est intégrable sur $\mathbb{R}$; et comme f est de plus continue, on a par la formule d'inversion :

$$f(t) = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{f}(x)\, e^{itx}\, dx \qquad \text{pour tout } t \in \mathbb{R}\,,$$

autrement dit (3.3).

Exercice 1. Montrer que la transformation de Fourier est une bijection de l'espace de Schwartz $\mathcal{S}(\mathbb{R})$ sur lui même. (Voir l'Exercice 3.5 pour la définition de $\mathcal{S}(\mathbb{R})$.)

Exercice 2. Montrer que $\mathcal{S}(\mathbb{R})$ est stable par convolution.

Exercice 3. Soit $P(z) = a_0 + a_1 z + \cdots + a_N z^N$ un polynôme non nul à coefficents complexes *sans racines réelles*. Montrer que pour toute fonction $g \in \mathcal{S}(\mathbb{R})$, il existe une unique fonction $f \in \mathcal{S}(\mathbb{R})$ telle que $P(D)f = g$, où $P(D)f = a_0 f + a_1 f' + \cdots + a_N f^{(N)}$.

Exercice 4. Dans cet exercice, on donne une autre preuve de la formule d'inversion pour les fonctions appartenant à l'espace de Schwartz $\mathcal{S}(\mathbb{R})$. On notera $\mathcal{F} : \mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R})$ la transformation de Fourier, et $\overline{\mathcal{F}}$ la transformation de Fourier "inverse". De plus, on notera D l'opérateur de dérivation, et M l'opérateur de multiplication par la fonction $t \mapsto t$.

(i) Soit $T : \mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R})$ un opérateur linéaire vérifiant $TM = MT$. Montrer que T est l'opérateur de multiplication par une certaine fonction ϕ, *i.e.* $Tf = \phi f$ pour toute $f \in \mathcal{S}(\mathbb{R})$. (Commencer par montrer que si $f \in \mathcal{S}(\mathbb{R})$ et si $a \in \mathbb{R}$ est fixé, on peut écrire $f = f(a)\mathbf{1} + (M - aId)g_a$ pour une certaine fonction $g \in \mathcal{S}(\mathbb{R})$.)

(ii) Trouver une relation entre les opérateurs $\mathcal{F}M$ et $D\mathcal{F}$ d'une part, et les opérateurs $\overline{\mathcal{F}}D$ et $M\overline{\mathcal{F}}$ d'autre part. Que peut-on en déduire pour l'opérateur $T = \overline{\mathcal{F}}\mathcal{F}$?

(iii) Montrer qu'il existe une fonction $g \in \mathcal{S}(\mathbb{R})$ ne s'annulant pas telle que la formule d'inversion est vraie pour g.

(iv) Démontrer la formule d'inversion pour toutes les fonctions de $\mathcal{S}(\mathbb{R})$.

3.5. La formule de Plancherel. Le théorème suivant est l'analogue de la *formule de Parseval* pour les séries de Fourier.

THÉORÈME 3.12. *Si $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, alors $\widehat{f} \in L^2(\mathbb{R})$ et on a la* ***formule de Plancherel***

$$\int_{\mathbb{R}} |\widehat{f}(x)|^2\, dx = 2\pi \int_{\mathbb{R}} |f(t)|^2\, dt\,.$$

Pour la preuve, on a besoin de deux lemmes.

LEMME 3.13. *Si $f, g \in L^1(\mathbb{R})$, alors $\overline{f}\,\check{g}$ et $\overline{\widehat{f}}\,g$ sont dans $L^1(\mathbb{R})$ et on a la* formule d'échange

$$\int_{\mathbb{R}} \overline{f}\,\check{g} = \int_{\mathbb{R}} \overline{\widehat{f}}\,g\,.$$

Démonstration. Les fonctions $\overline{f}\,\check{g}$ et $\overline{\widehat{f}}\,g$ sont dans L^1 car $\overline{f}$, g sont dans L^1 et $\check{g}$, $\overline{\widehat{f}}$ sont bornées. Ensuite, c'est un calcul direct avec le théorème de Fubini (la justification nécessaire est laissée en exercice) :

$$\begin{aligned}
\int_{\mathbb{R}} \overline{f}\,\check{g} &= \int_{\mathbb{R}} \overline{f(t)} \left(\int_{\mathbb{R}} g(x) e^{itx} dx \right) dt \\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \overline{f(t)} e^{ixt} dt \right) g(x)\, dx \\
&= \int_{\mathbb{R}} \overline{\left(\int_{\mathbb{R}} f(t) e^{-ixt} dt \right)}\, g(x)\, dx = \int_{\mathbb{R}} \overline{\widehat{f}}\, g\,.
\end{aligned}$$

□

LEMME 3.14. *Si $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, alors il existe une suite (φ_n) de fonctions de classe $\mathcal{C}^2$ à support compact telle que $\varphi_n \to f$ en norme L^1* et *en norme L^2.*

Démonstration. Posons d'abord $f_n := \mathbf{1}_{[-n,n]}\, f$. Alors $f_n(t) \to f(t)$ presque partout, et $|f_n| \leqslant |f|$ pour tout $n \in \mathbb{N}$. Comme $f \in L^1 \cap L^2$, on en déduit par convergence dominée (dans L^1 et dans L^2) que $f_n \to f$ en norme L^1 et en norme L^2.

Maintenant, soit $(k_r)_{r\in\mathbb{N}} \subseteq L^1(\mathbb{R})$ une suite de Dirac formée de fonctions de classe $\mathcal{C}^2$ à support compact. Pour $n \in \mathbb{N}$ fixé, on sait que $k_r * f_n \to f_n$ en norme L^1 et en norme L^2 quand $r \to \infty$; donc on peut trouver un entier r_n tel que $\|k_{r_n} * f_n\|_p \leqslant 2^{-n}$ pour $p = 1, 2$.

Comme $\|k_{r_n} * f_n - f\|_p \leqslant \|k_{r_n} * f_n - f_n\|_p + \|f_n - f\|_p$, on voit que $\varphi_n := k_{r_n} * f_n$ tend vers f en norme L^1 et en norme L^2 quand $n \to \infty$. De plus, chaque φ_n est de classe $\mathcal{C}^2$ à support compact car k_{r_n} est de classe $\mathcal{C}^2$ à support compact et f_n est à support compact. □

Preuve du Théorème 3.12. Soit $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. On va distinguer 2 cas.

CAS 1. On supose que $\widehat{f} \in L^1(\mathbb{R})$.

Dans ce cas, on a $f = \frac{1}{2\pi}\,\check{\widehat{f}}$ par la formule d'inversion. Donc

$$\begin{aligned}
\int_{\mathbb{R}} |f|^2 &= \frac{1}{2\pi} \int_{\mathbb{R}} \overline{f}\,\check{\widehat{f}} \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} \overline{\widehat{f}}\,\widehat{f} \qquad \text{par le Lemme 3.13 avec } g := \widehat{f} \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} |\widehat{f}|^2.
\end{aligned}$$

CAS 2. Cas général.

Soit (φ_n) une suite de fonctions de classe $\mathcal{C}^2$ à support compact convergeant vers f en norme L^1 et en norme L^2. (La suite (φ_n) est donnée par le Lemme 3.14.)

D'après le Corollaire 3.4, on sait que $\widehat{\varphi_n} \in L^1(\mathbb{R})$ pour tout $n \in \mathbb{N}$. D'après le Cas 1, on a donc

$$\|\widehat{\varphi_n}\|_2^2 = 2\pi \, \|\varphi_n\|_2^2 \, .$$

Comme de plus $\|\varphi_n\|^2 \to \|f\|_2^2$ puisque $\varphi_n \to f$ en norme L^2, on en déduit que

$$\|\widehat{\varphi_n}\|_2^2 \to 2\pi \, \|f\|^2 \, . \tag{3.4}$$

Par ailleurs, on a $\|\widehat{\varphi_q} - \widehat{\varphi_p}\|_2^2 = \|\widehat{\varphi_p - \varphi_q}\|_2^2 = \|\varphi_q - \varphi_p\|_2^2$ pour tous $p, q \in \mathbb{N}$, d'après le Cas 1 appliqué à $\varphi_q - \varphi_p$. Donc la suite $(\widehat{\varphi_n})$ est *de Cauchy* dans L^2 car (φ_n) est de Cauchy. Comme L^2 est complet, la suite $\widehat{\varphi_n}$ converge donc dans L^2 vers une certaine fonction g ; et par conséquent, $(\widehat{\varphi_n})$ admet une sous-suite qui tend vers g presque partout. Mais $\widehat{\varphi_n}(x) \to \widehat{f}(x)$ en tout point (et même uniformément) car $\varphi_n \to f$ en norme L^1 et $|\widehat{\varphi_n}(x) - \widehat{f}(x)| = |\widehat{(\varphi_n - f)}(x)| \leqslant \|\varphi_n - f\|_1$ pour tout $x \in \mathbb{R}$. Donc $g(x) = \widehat{f}(x)$ presque partout ; *et donc* $\widehat{\varphi_n} \to \widehat{f}$ *en norme* L^2. En particulier :

$$\|\widehat{\varphi_n}\|_2^2 \to \|\widehat{f}\|_2^2 \, . \tag{3.5}$$

En comparant (3.4) et (3.5), on obtient immédiatement la formule de Plancherel. □

Exercice 1. Déterminer la transformée de Fourier de la fonction $f := \mathbf{1}_{[-1,1]}$, et en déduire qu'on a

$$\int_{\mathbb{R}} \left(\frac{\sin x}{x} \right)^2 dx = \pi \, .$$

Exercice 2. Montrer que si $f \in L^1(\mathbb{R})$ est de classe $\mathcal{C}^1$ et si $f' \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, alors $\widehat{f} \in L^1(\mathbb{R})$.

Exercice 3. Pour tout $\lambda > 0$, on note $\mathcal{F}_\lambda$ la "transformation de Fourier normalisée par λ" : si $f \in L^1(\mathbb{R})$, alors

$$\mathcal{F}_\lambda f(x) := \int_{\mathbb{R}} f(t) e^{-i\lambda x t} dt \, .$$

On définit de même la transofrmation de Fourier inverse $\overline{\mathcal{F}}_\lambda$. Montrer que si on choisit convenablement λ (il y a exactement 1 choix possible), alors la formule d'inversion et la formule de Plancherel prennent les formes suivantes :

$$\overline{\mathcal{F}}_\lambda \mathcal{F}_\lambda f = f \quad \text{et} \quad \int_{\mathbb{R}} |\mathcal{F} f|^2 = \int_{\mathbb{R}} |f|^2 \, .$$

Montrer également que pour ce choix de λ, il existe une et une seule valeur de $\alpha > 0$ pour laquelle la fonction $g_\alpha(t) := e^{-\alpha t^2}$ vérifie $\mathcal{F}_\lambda g_\alpha = g_\alpha$.

Exercice 4. Montrer que si $f \in L^2(\mathbb{R})$, alors les fonctions f_A définies par $f_A(x) := \int_{-A}^{A} f(t) e^{-ixt} dt$ admettent une limite dans $L^2(\mathbb{R})$ quand $A \to \infty$, et que si on note $\mathcal{F} f$ cette limite, alors $\|\mathcal{F} f\|_2 = \sqrt{2\pi} \, \|f\|_2$. La fonction $\mathcal{F} f$ s'appelle la **transformée de Fourier-Plancherel** de f.

3.6. Analogie avec les séries de Fourier. Tout ce qui vient d'être raconté sur la transformation de Fourier peut être transcrit dans le cadre des *séries* de Fourier, avec des démonstrations essentiellement identiques. Détaillons un peu cela.

On considère l'espace $L^1_{2\pi}$ formé par toutes les fonctions $f : \mathbb{R} \to \mathbb{C}$ presque partout définies, 2π-périodiques et intégrables sur $[-\pi, \pi[$, muni de la norme

$$\|f\|_1 := \int_{-\pi}^{\pi} |f(t)| \, \frac{dt}{2\pi} \cdot$$

(Il s'agit bien d'une norme car on considère des fonctions 2π-périodiques : si $\|f\|_1 = 0$, alors $f = 0$ presque partout sur $[-\pi, \pi[$, donc presque partout sur $\mathbb{R}$ par périodicité. Au lieu de $[-\pi, \pi[$, on peut d'ailleurs prendre n'importe quel intervalle de longueur 2π.)

Les espaces $L^p_{2\pi}$, $1 \leqslant p \leqslant \infty$ sont définis de la même façon. Contrairement à ce qui se passe dans le cas de $L^p(\mathbb{R})$, la famille des espaces $L^p_{2\pi}$ est décroissante, et la suite des normes $\|\cdot\|_p$ est même croissante car on travaille sur un espace de mesure totale égale à 1 (l'intervalle $[-\pi, \pi[$ muni de la mesure de Lebesgue *normalisée* $\frac{dt}{2\pi}$). En particulier, toute fonction *continue* 2π-périodique appartient à $L^1_{2\pi}$.

La *convolée* de deux fonctions 2π-périodiques $f, g : \mathbb{R} \to \mathbb{C}$ est définie (en tout point x où la formule a un sens) par

$$f * g\,(x) = \int_{-\pi}^{\pi} f(x-t) g(t) \, \frac{dt}{2\pi} \cdot$$

C'est encore une fonction 2π-périodique : $f * g\,(x)$ est bien défini si et seulement si $f * g\,(x + 2\pi)$ l'est, et dans ce cas $f * g\,(x) = f * g\,(x + 2\pi)$.

Les suites de Dirac sont définies ainsi : une suite $(k_n)_{n\in\mathbb{N}} \subseteq L^1_{2\pi}$ est une suite de Dirac si elle est bornée dans $L^1_{2\pi}$ avec $\int_{-\pi}^{\pi} k_n(t)\, \frac{dt}{2\pi} \equiv 1$, et si pour tout δ vérifiant $0 < \delta < \pi$, on a

$$\lim_{n\to\infty} \int_{\delta<|t|<\pi} |k_n(t)| \, dt = 0 \,.$$

Exemple 3.15. Pour $r \in [0, 1[$, notons $P_r : \mathbb{R} \to \mathbb{C}$ la fonction continue 2π-périodique définie par

$$P_r(t) := \sum_{k=-\infty}^{\infty} r^{|k|} e^{ikt} \,.$$

(La famille $(P_r)_{0\leqslant r<1}$ s'appelle le **noyau de Poisson**.) Pour toute suite (r_n) tendant vers 1, la suite (P_{r_n}) est une suite de Dirac dans $L^1_{2\pi}$.

Démonstration. En intégrant terme à terme (ce qui est licite car la série définissant P_r converge normalement sur $\mathbb{R}$), on voit qu'on a

$$\int_{-\pi}^{\pi} P_r(t) \, \frac{dt}{2\pi} = \sum_{k=-\infty}^{\infty} r^{|k|} \int_{-\pi}^{\pi} e^{-ikt} \, \frac{dt}{2\pi} = 1 \,.$$

(Tous les termes de la somme sont nuls, sauf pour $k = 0$.)

De plus,

$$\begin{aligned} P_r(t) &= \sum_{k=1}^{\infty} r^k e^{-ikt} + 1 + \sum_{k=1}^{\infty} r^k e^{ikt} \\ &= \frac{re^{-it}}{1-re^{-it}} + 1 + \frac{re^{it}}{1-re^{it}} \\ &= \frac{1-r^2}{|1-re^{it}|^2} \qquad \text{après calcul}\,. \end{aligned}$$

On en déduit que P_r est une fonction *positive*, et donc que $\|P_r\|_1 = \int_{-\pi}^{\pi} P_r(t)\,\frac{dt}{2\pi} = 1$ pour tout r. On en déduit aussi sans difficulté que

$$\int_{\delta<|t|<\pi} |P_r(t)|\,dt \xrightarrow{r\to 1} 0 \qquad \text{pour tout } \delta \text{ vérifiant } 0<\delta<\pi\,;$$

d'où le résultat. □

Dans le cadre $L^1_{2\pi}$, le rôle de la transformée de Fourier est maintenant joué par la suite des **coefficients de Fourier** : si $f \in L^1_{2\pi}$, on pose

$$c_k(f) := \int_{-\pi}^{\pi} f(t)e^{-ikt}\,\frac{dt}{2\pi}\ ,\ k\in\mathbb{Z}\,.$$

La transformation de Fourier "inverse" est maintenant définie sur $\ell^1(\mathbb{Z})$, l'espace de toutes les suites $d = (d(k))_{k\in\mathbb{Z}}$ telles que $\sum_{k=-\infty}^{\infty} |d(k)| < \infty$: si $d = (d(k))_{k\in\mathbb{Z}} \in \ell^1(\mathbb{Z})$, alors $\check{d}$ est la fonction (continue) 2π-périodique définie par

$$\check{d}(t) := \sum_{k=-\infty}^{\infty} d(k)e^{ikt}\,.$$

La transformation de Fourier des fonctions périodiques possède des propriétés formelles analogues à la transformation de Fourier sur $L^1(\mathbb{R})$. En particulier :

- Si $f, g \in L^1_{2\pi}$, alors
$$c_k(f*g) = c_k(f)c_k(g) \quad \text{pour tout } k\in\mathbb{Z}\,;$$
- si $f:\mathbb{R}\to\mathbb{C}$ est 2π-périodique et de classe $\mathcal{C}^1$, alors
$$c_k(f) = \frac{1}{ik}\,c_k(f') \quad \text{pour tout } k\neq 0\,.$$

Le *lemme de Riemann-Lebesgue* est vrai pour les coefficients de Fourier : si $f \in L^1_{2\pi}$, alors

$$c_k(f)\to 0 \quad \text{quand } |k|\to\infty\,.$$

(On peut commencer par le montrer pour une fonction de classe $\mathcal{C}^1$, puis procéder par approximation.)

La **formule d'inversion de Fourier** s'énonce comme suit : *Si $f \in L^1_{2\pi}$ et si la suite $\widehat{f} := (c_k(f))_{k\in\mathbb{Z}}$ appartient à $\ell^1(\mathbb{Z})$, i.e. $\sum_{k=-\infty}^{\infty} |c_k(f)| < \infty$, alors $f = \check{\widehat{f}}$ presque partout ; autrement dit*

$$f(t) = \sum_{k=-\infty}^{\infty} c_k(f)e^{ikt} \quad \textit{presque partout}\,.$$

Si f est de plus continue, alors cette formule est vraie pour tout $t \in \mathbb{R}$.

On peut démontrer ce résultat exactement comme pour la transformation de Fourier, en utilisant les deux faits suivants.

Fait 1. Si $f \in L^1_{2\pi}$ et si $d = (d(k)) \in \ell^1(\mathbb{Z})$ alors $\widecheck{d\widehat{f}} = \check{d} * f$, *i.e.*

$$\sum_{k=-\infty}^{\infty} d(k) c_k(f) e^{ikt} = (\check{d} * f)(t) \qquad \text{pour tout } t \in \mathbb{R}\,.$$

Démonstration. C'est un calcul direct utilisant le théorème d'interversion série-intégrale. □

Fait 2. Il existe une suite (d_n) d'éléments de $\ell^1(\mathbb{Z})$, bornée dans $\ell^\infty(\mathbb{Z})$, telle que $d_n(k) \to 1$ pour tout $k \in \mathbb{Z}$ quand $n \to \infty$, et telle que la suite $(\check{d}_n)$ est une suite de Dirac dans $L^1_{2\pi}$.

Démonstration. Il suffit de prendre $d_n(k) := r_n^{|k|} = c_k(P_{r_n})$, où (P_r) est le noyau de Poisson et $r_n \to 1$. Par définition, on a $\check{d}_n = P_{r_n}$; d'où le résultat par l'Exemple 3.15. □

Une fois la formule d'inversion acquise, on peut adapter la preuve de la formule de Plancherel pour démontrer la **formule de Parseval** : *Si $f \in L^2_{2\pi}$, alors la suite $(c_k(f))$ appartient à $\ell^2(\mathbb{Z})$ et on a*

$$\sum_{k=-\infty}^{\infty} |c_k(f)|^2 = \int_{-\pi}^{\pi} |f(t)|^2 \, \frac{dt}{2\pi} \cdot$$

Remarque. Cette manière d'obtenir la formule de Parseval n'est pas la plus "standard". En général, on procède plutôt comme suit.

Rappelons d'abord que $L^2_{2\pi}$ est un espace de Hilbert, dont le produit scalaire est donné par

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(t)\overline{g}(t) \, \frac{dt}{2\pi} \cdot$$

Pour $k \in \mathbb{Z}$, notons $e_k \in L^2_{2\pi}$ la fonction définie par

$$e_k(t) := e^{ikt}\,.$$

On vérifie très facilement que la famille $(e_k)_{k\in\mathbb{Z}}$ est **orthonormale** dans $L^2_{2\pi}$, *i.e.* $\langle e_k, e_l \rangle = 0$ si $k \neq l$ et $\|e_k\|^2 = 1$ pour tout k. De plus, si $f \in L^2_{2\pi}$ on a par définition

$$c_k(f) = \langle f, e_k \rangle \quad \text{pour tout } k \in \mathbb{Z}\,.$$

On en déduit que si, pour $f \in L^2_{2\pi}$ et $N \in \mathbb{N}$ on pose

$$S_N f(t) := \sum_{k=-N}^{N} c_k(f) e^{ikt}\,,$$

alors $S_N f$ est le **projeté orthogonal** de f sur le sous espace $\mathcal{P}_N$ de $L^2_{2\pi}$ engendré par les vecteurs e_k pour $k \in \{-N, \dots, N\}$. (C'est clair puisque $S_N f = \sum_{-N}^{N} \langle f, e_k \rangle e_k$.)

Par les propriétés générales des espaces munis d'un produit scalaire (autrement dit : par le théorème de Pythagore), cela signifie que $S_N f$ est l'élement de $\mathcal{P}_N$ *le plus proche de f* au sens de la norme L^2 : on a

$$\|S_N f - f\|^2 = \operatorname{dist}(f, \mathcal{P}_N)\,.$$

Maintenant, le sous-espace $\mathcal{P}_N$ n'est rien d'autre que l'ensemble des **polynômes trigonométriques** de degré $\leqslant N$; et "tout le monde sait bien" que les polynômes trigonométriques sont *denses* dans $L^2_{2\pi}$. (C'est le point clé dans cette approche. Cela découle par exemple du "théorème de Weierstrass trigonométrique" : les polynômes trigonométriques sont dense dans $\mathcal{C}_{2\pi}$, l'espace des fonctions *continues* 2π-périodiques muni de la norme $\|\cdot\|_\infty$; donc ils sont aussi denses dans $\mathcal{C}_{2\pi}$ pour la norme L^2, qui est plus petite que la norme L^∞ ; d'où le résultat car $\mathcal{C}_{2\pi}$ est dense dans $L^2_{2\pi}$. Évidemment, il faut être capable de démontrer le théorème de Weierstrass trigonométrique... Cela se fait classiquement via le **théorème de Fejér** ; mais *cf* aussi l'Exercice 3 ci-dessous.)

On en déduit immédiatement que $\mathrm{dist}(f, \mathcal{P}_N) \to 0$ quand $N \to \infty$; autrement dit :

$$S_N f \to f \quad \text{en norme } L^2\,.$$

En particulier, $\|S_N f\|^2 \to \|f\|^2$; et comme

$$\|S_N f\|^2 = \sum_{k=-N}^{N} |c_k(f)|^2$$

par orthonormalité de la suite (e_k), cela donne la formule de Parseval.

Exercice 1. Montrer que si $f : \mathbb{R} \to \mathbb{C}$ est 2π-périodique, continue et de classe $\mathcal{C}^1$ par morceaux sur $[-\pi, \pi]$, alors $\widehat{f} \in \ell^1(\mathbb{Z})$, i. e. $\sum_{k=-\infty}^{\infty} |c_k(f)| < \infty$.

Exercice 2. Calculer $\sum_{k=1}^{\infty} \frac{\sin^2(k\lambda)}{k^2}$ pour tout $\lambda \in [-\pi, \pi]$, et en déduire $\sum_{n=1}^{\infty} \frac{1}{n^2}$.

Exercice 3. Pour $r \in [0, 1[$ et $N \in \mathbb{N}$, on pose

$$P_{r,N}(t) := \sum_{k=-N}^{N} r^{|k|} e^{ikt}\,.$$

Montrer que $P_{r,N}(t) \to P_r(t)$ uniformément sur $\mathbb{R}$ quand $N \to \infty$, et que $P_{r,N} \to P_r$ dans $L^p_{2\pi}$ pour tout $p < \infty$. *En déduire* que que les polynômes trigonométriques sont denses dans $L^p_{2\pi}$ pour tout $p < \infty$, et qu'ils sont également dense dans $\mathcal{C}_{2\pi}$, l'espace des fonctions *continues* 2π-périodiques muni de la norme $\|\cdot\|_\infty$.

Exercice 4. Montrer que l'application $f \mapsto (c_k(f))_{k\in\mathbb{Z}}$ est une isométrie bijective de $L^2_{2\pi}$ sur $\ell^2(\mathbb{Z})$.

Chapitre 10

Théorie de la mesure "sérieuse"

Ce chapitre pourrait s'intituler "bouchage de trous". En effet, il s'agit avant tout de démontrer un certains nombres de résultats qui ont été admis en cours de route (lemme d'unicité des mesures, existence de la mesure de Lebesgue, Lemme 3.2 du Chapitre 7, régularité de la mesure de Lebesgue). Comme cela ne coûte pas plus cher, on va se placer dans un cadre assez général ; ce qui explique le titre choisi...

1. Classes monotones

1.1. Le "théorème des classes monotones". Dans ce qui suit, Ω est un ensemble non vide.

DÉFINITION 1.1. *Soit $\mathcal{M}$ une famille de parties de Ω. On dit que $\mathcal{M}$ est une* ***classe monotone*** *si elle vérifie les propriétés suivantes.*

(o) $\Omega \in \mathcal{M}$.

(i) *Stabilité par* ***différences propres*** *: si $A, A' \in \mathcal{M}$ et $A \subseteq A'$, alors $A' \backslash A \in \mathcal{M}$.*

(ii) *Stabilité par* ***réunions dénombrables croissantes*** *: si $(A_n)_{n\in\mathbb{N}}$ est une suite croissante d'éléments de $\mathcal{M}$, alors $\bigcup_{n=0}^{\infty} A_n \in \mathcal{M}$.*

Remarque. Les propriétés (o) et (i) entrainent que toute classe monotone est stable par *complémentation.*

Exemple 1. Toute tribu est évidemment une classe monotone.

Exemple 2. Soit $\mathcal{B}$ une tribu de parties de Ω. Si μ_1 et μ_2 sont deux mesures *finies* sur $(\Omega, \mathcal{B})$ telles que $\mu_1(\Omega) = \mu_2(\Omega)$, alors $\mathcal{M} := \{A \in \mathcal{B};\ \mu_1(A) = \mu_2(A)\}$ est une classe monotone.

Démonstration. On a $\Omega \in \mathcal{M}$ par hypothèse. La stabilité par réunions dénombrables croissantes découle de la continuité par en dessous des mesures μ_1 et μ_2 ; et si $A, A' \in \mathcal{M}$ avec $A \subseteq A'$, alors $\mu_1(A'\backslash A) = \mu_1(A') - \mu_1(A) = \mu_2(A') - \mu_2(A) = \mu_2(A'\backslash A)$, donc $A'\backslash A \in \mathcal{M}$. (On a besoin de savoir que les mesures sont finies pour pouvoir soustraire $\mu_1(A)$ et $\mu_2(A)$.) $\square$

Exemple 3. Soient Ω_1 et Ω_2 deux espaces métriques. Pour $A \subseteq \Omega_1 \times \Omega_2$ et $x \in \Omega_1$, notons A_x la coupe de A à l'abscisse x. Si μ est une mesure borélienne *finie* sur Ω_2, alors la famille

$$\mathcal{M} := \big\{A \in \mathcal{B}(\Omega_1 \times \Omega_2);\ \text{la fonction } x \mapsto \mu(A_x) \text{ est borélienne sur } \Omega_1\big\}$$

est une classe monotone.

Démonstration. Ici, $\mathcal{M}$ est une famille de parties de $\Omega = \Omega_1 \times \Omega_2$. On a $\Omega \in \mathcal{M}$ car la fonction $x \mapsto \mu(\Omega_x)$ est constante (égale à $\mu(\Omega_2)$). Si (A_n) est une suite croissante d'éléments de $\mathcal{M}$ et si on pose $A := \bigcup_{n=0}^{\infty} A_n$, alors $\mu(A_x) = \lim_{n\to\infty} \mu\big((A_n)_x\big)$ car

A_x est la réunion croissante des $(A_n)_x$; donc la fonction $x \mapsto \mu(A_x)$ est borélienne car les fonctions $x \mapsto \mu\big((A_n)_x\big)$ le sont. Si $A, A' \in \mathcal{M}$ avec $A \subseteq A'$, alors on peut écrire $\mu\big((A'\backslash A)_x\big) = \mu(A'_x\backslash A_x) = \mu(A'_x) - \mu(A_x)$ car la mesure μ est finie ; donc la fonction $x \mapsto \mu\big((A'\backslash A)_x\big)$ est borélienne. $\square$

Le résultat suivant, qu'on appelle le **théorème des classes monotones**, est un outil technique extrêmement utile.

THÉORÈME 1.2. *Soit $\mathcal{C}$ une famille de parties de Ω. On suppose que $\mathcal{C}$ est stable par* intersections finies. *Si $\mathcal{M} \subseteq \mathcal{P}(\Omega)$ est une classe monotone contenant $\mathcal{C}$, alors $\mathcal{M}$ contient la tribu $\sigma(\mathcal{C})$ engendrée par $\mathcal{C}$.*

Démonstration. Notons $\mathcal{M}_0$ la classe monotone **engendrée par** $\mathcal{C}$, *i.e.* l'intersection de toutes les classes monotones contenant $\mathcal{C}$, ou encore la *plus petite* classe monotone contenant $\mathcal{C}$. On a par définition $\mathcal{C} \subseteq \mathcal{M}_0 \subseteq \mathcal{M}$; donc il suffit de montrer que $\mathcal{M}_0$ *est en fait une tribu* : on aura alors $\sigma(\mathcal{C}) \subseteq \mathcal{M}_0$, et donc $\sigma(\mathcal{C}) \subseteq \mathcal{M}$.

On sait déjà que $\mathcal{M}_0$ contient Ω et est stable par complémentation.

La stabilité par réunions dénombrables va se prouver en 3 étapes, dont les deux premières sont une nouvelle illustration du “principe du fusil à 2 coups”. On a besoin du fait suivant.

FAIT. Si $E \subseteq \Omega$, alors $\mathcal{M}_E := \{M \subseteq \Omega;\ A \cap E \in \mathcal{M}_0\}$ est une classe monotone.

Preuve du Fait. Si $M, M' \in \mathcal{M}_E$ et $M \subseteq M'$, alors

$$(M'\backslash M) \cap E = \underbrace{(M' \cap E)}_{\in \mathcal{M}_0} \backslash \underbrace{(M \cap E)}_{\in \mathcal{M}_0} \in \mathcal{M}_0$$

car $\mathcal{M}_0$ est une classe monotone ; donc $M'\backslash M \in \mathcal{M}_E$. De même, si (M_n) est une suite croissante d'éléments de $\mathcal{M}_E$, alors $\big(\bigcup_0^\infty M_n\big) \cap E = \bigcup_0^\infty (M_n \cap E) \in \mathcal{M}_0$ car $\mathcal{M}_0$ est une classe monotone ; donc $\bigcup_0^\infty M_n \in \mathcal{M}_E$. $\square$

ÉTAPE 1. Si $B \in \mathcal{C}$, alors $A \cap B \in \mathcal{M}_0$ pour tout $A \in \mathcal{M}_0$.

Démonstration. Par le Fait, la famille $\mathcal{M}_B = \{A \subseteq \Omega;\ A \cap B \in \mathcal{M}_0\}$ est une classe monotone. De plus, $\mathcal{M}_B$ contient $\mathcal{C}$ car $\mathcal{C}$ est stable par intersections finies : si $A \in \mathcal{C}$, alors $A \cap B \in \mathcal{C} \subseteq \mathcal{M}_0$. Donc $\mathcal{M}_B$ contient $\mathcal{M}_0$, ce qui est le résultat souhaité. $\square$

ÉTAPE 2. $\mathcal{M}_0$ est stable par intersections finies.

Démonstration. Il suffit de montrer qu'on a $A \cap B \in \mathcal{M}_0$ pour tous $A, B \in \mathcal{M}_0$. Pour cela, on fixe $A \in \mathcal{M}_0$ et on considère la famille $\mathcal{M}_A = \{B \subseteq \Omega;\ A \cap B \in \mathcal{M}_0\}$. Par le Fait, $\mathcal{M}_A$ est une classe monotone. De plus $\mathcal{M}_A$ contient $\mathcal{C}$ *par l'Étape 1.* Donc $\mathcal{M}_A$ contient $\mathcal{M}_0$. $\square$

ÉTAPE 3. $\mathcal{M}_0$ est stable par réunions dénombrables.

Démonstration. Soit $(A_n)_{n\in\mathbb{N}}$ une suite quelconque d'éléments de $\mathcal{M}_0$. Si on pose $B_n := A_0 \cup \cdots \cup A_n$, alors $B_n \in \mathcal{M}_0$ d'après l'Étape 2 car $B_n = \big(\bigcap_{k=0}^n A_k^c\big)^c$ et $\mathcal{M}_0$ est stable par complémentation. De plus, la suite (B_n) est *croissante* et $\bigcup_0^\infty B_n = \bigcup_0^\infty A_n$; donc $\bigcup_0^\infty A_n \in \mathcal{M}_0$ car $\mathcal{M}_0$ est une classe monotone. $\square$

Maintenant qu'on sait que $\mathcal{M}_0$ est une tribu, la preuve du théorème des classes monotones est achevée. $\square$

1.2. Unicité des mesures. Comme première application du théorème des classes monotones, on va démontrer une version générale du "lemme d'unicité" admis au Chapitre 2.

Théorème 1.3. *Soit $(\Omega, \mathcal{B})$ un espace mesurable, et soit $\mathcal{C}$ une famille de parties de Ω. On fait les hypothèses suivantes.*

(i) *$\mathcal{C}$ est stable par* ***intersections finies****.*

(ii) *$\sigma(\mathcal{C}) = \mathcal{B}$.*

(iii) *Il existe une suite $(E_k)_{k\in\mathbb{N}}$ d'éléments de $\mathcal{C}$ telle que $\Omega = \bigcup_{k=0}^{\infty} E_k$.*

Si μ_1 et μ_2 sont deux mesures sur $(\Omega, \mathcal{B})$ vérifiant $\mu_1(C) = \mu_2(C) < \infty$ pour tout $C \in \mathcal{C}$, alors $\mu_1 = \mu_2$.

Démonstration. On va distinguer deux cas.

Cas 1. Supposons que Ω appartienne à la famille $\mathcal{C}$.

Dans ce cas, les mesures μ_1 et μ_2 sont *finies*. Donc la famille

$$\mathcal{M} := \{A \in \mathcal{B};\ \mu_1(A) = \mu_2(A)\}$$

est une classe monotone (c'est l'Exemple 2 de la section précédente). Comme $\mathcal{C}$ contient Ω et est stable par intersections finies, et comme $\mathcal{C} \subseteq \mathcal{M}$ par hypothèse sur μ_1 et μ_2, le théorème des classes monotones dit que $\mathcal{M}$ contient la tribu $\sigma(\mathcal{C})$, qui est égale à $\mathcal{B}$; donc $\mathcal{M} = \mathcal{B}$, *i.e.* $\mu_1 = \mu_2$.

Cas 2. Cas général.

On peut appliquer le cas précédent pour tout $k \in \mathbb{N}$ aux restrictions $\mu_{1,k}$ et $\mu_{2,k}$ de μ_1 et μ_2 à E_k, où les E_k sont donnés par (iii).

Les mesures $\mu_{1,k}$ et $\mu_{2,k}$ sont définies sur $\mathcal{B}_k := \mathcal{B} \cap \mathcal{P}(E_k)$, et elles coïncident sur $\mathcal{C}_k := \mathcal{C} \cap \mathcal{P}(E_k)$. De plus, la famille $\mathcal{C}_k$ *engendre la tribu* $\mathcal{B}_k$: en effet, on a $\mathcal{C}_k = \mathcal{C}_{E_k} := \{E_k \cap C;\ C \in \mathcal{C}\}$ car $\mathcal{C}$ est stable par intersections finies ; donc $\sigma(\mathcal{C}_k) = \sigma(\mathcal{C}_{E_k}) = \sigma(\mathcal{C})_{E_k} = \mathcal{B}_{E_k} = \mathcal{B} \cap \mathcal{P}(E_k)$ (*cf* la preuve de la Proposition 4.2, Chapitre 2). Par le Cas 1, on a donc $\mu_{1,k} = \mu_{2,k}$; autrement dit

$$\mu_1(A) = \mu_2(A) \quad \text{pour tout } A \in \mathcal{B} \text{ vérifiant } A \subseteq E_k\,.$$

Soit maintenant $A \in \mathcal{B}$ quelconque. Si on pose $\Omega_0 := E_0$ et

$$\Omega_k := E_k \backslash \big(E_0 \cup \cdots \cup E_{k-1}\big) \quad \text{pour } k \geqslant 1,$$

alors $\Omega = \bigcup_0^\infty \Omega_k$ par (iii), et les Ω_k sont deux à deux disjoints. Donc A est la réunion disjointe des $A_k := A \cap \Omega_k$. De plus, comme $A_k \subseteq E_k$, on a $\mu_1(A_k) = \mu_2(A_k)$ pour tout k et donc

$$\mu_1(A) = \sum_{k=0}^{\infty} \mu_1(A_k) = \sum_{k=0}^{\infty} \mu_2(A_k) = \mu_2(A)\,.$$

□

Comme conséquence immédiate du Théorème 1.3, on obtient le "lemme d'unicité" du Chapitre 2 (Lemme 5.1).

Corollaire 1.4. *Soit Ω un ouvert de $\mathbb{R}^N$. Si μ_1, μ_2 sont deux mesures boréliennes sur Ω vérifiant $\mu_1(P) = \mu_2(P) < \infty$ pour tout pavé compact $P \subseteq \Omega$, alors $\mu_1 = \mu_2$.*

Démonstration. On applique le théorème à la famille $\mathcal{C}$ constituée par tous les pavés compacts $P \subseteq \Omega$. La famille $\mathcal{C}$ est stable par intersections finies et engendre la tribu borélienne de Ω. De plus, l'ouvert Ω est réunion dénombrable de pavés compacts, autrement dit (iii) est vérifiée. □

1.3. Un lemme de mesurabilité. Comme deuxième application du théorème des classes monotones, on va démontrer un lemme admis au Chapitre 7 (le Lemme 3.2). Rappelons que si $A \subseteq \mathbb{R}^p \times \mathbb{R}^q$, on note A_x la coupe de A à une certaine abscisse $x \in \mathbb{R}^p$ et A^y la coupe de A à une certaine ordonnée $y \in \mathbb{R}^q$.

Lemme 1.5. *Pour tout borélien $A \subseteq \mathbb{R}^p \times \mathbb{R}^q$, les fonctions $x \mapsto \lambda_q(A_x)$ et $y \mapsto \lambda_p(A^y)$ sont boréliennes, respectivement sur $\mathbb{R}^p$ et sur $\mathbb{R}^q$.*

Démonstration. On montre seulement que la fonction $x \mapsto \lambda_q(A_x)$ est borélienne, pour tout borélien $A \subseteq \mathbb{R}^p \times \mathbb{R}^q$.

Soit $(E_n)_{n\in\mathbb{N}}$ une suite croissante de boréliens bornés de $\mathbb{R}^q$ telle que $\bigcup_{n=0}^\infty E_n = \mathbb{R}^q$. Si $A \in \mathcal{B}(\mathbb{R}^p \times \mathbb{R}^q)$, alors $\lambda_q(A_x) = \lim_{n\to\infty} \lambda_q(A_x \cap E_n)$; donc il suffit de montrer que pour tout $n \in \mathbb{N}$, la fonction $x \mapsto \lambda_q(A_x \cap E_n)$ est borélienne.

Fixons $n \in \mathbb{N}$, et notons μ la mesure borélienne sur $\mathbb{R}^q$ définie par

$$\mu(B) := \lambda_q(B \cap E_n)\,.$$

Il s'agit de montrer que pour tout $A \in \mathcal{B}(\mathbb{R}^p \times \mathbb{R}^q)$, la fonction $x \mapsto \mu(A_x)$ est borélienne. Autrement dit, si on pose

$$\mathcal{M} := \{A \in \mathcal{B}(\mathbb{R}^p \times \mathbb{R}^q);\ \text{la fonction } x \mapsto \mu(A_x) \text{ est borélienne}\}\,,$$

il s'agit de montrer que $\mathcal{M} = \mathcal{B}(\mathbb{R}^p \times \mathbb{R}^q)$.

Comme E_n est borné, la mesure μ est *finie* ; donc la famille $\mathcal{M}$ est une classe monotone (c'est l'Exemple 3 de la section précédente). De plus $\mathcal{M}$ contient la famille

$$\mathcal{C} := \left\{C_1 \times C_2;\ C_1 \in \mathcal{B}(\mathbb{R}^p)\,,\, C_2 \in \mathcal{B}(\mathbb{R}^q)\right\}.$$

En effet, si $C = C_1 \times C_2 \in \mathcal{C}$, alors $\mu(C_x) = \mu(C_2)$ si $x \in C_1$ et $\mu(C_x) = 0$ si $x \notin C_1$, donc la fonction $x \mapsto \mu(C_x)$ est borélienne par recollement. Comme la famille $\mathcal{C}$ est visiblement stable par intersection finies, contient $\mathbb{R}^p \times \mathbb{R}^q$ et engendre la tribu borélienne de $\mathbb{R}^p \times \mathbb{R}^q$ (car $\mathcal{C}$ contient tous les pavés), le théorème des classes monotones permet de conclure que $\mathcal{M} = \mathcal{B}(\mathbb{R}^p \times \mathbb{R}^q)$. □

2. Régularité

Dans cette section, on démontre la “régularité” de la mesure de Lebesgue ; et en fait, la régularité de toute mesure borélienne sur $\mathbb{R}^N$ prenant des valeurs finies sur les compacts. Cela va découler du théorème suivant.

Théorème 2.1. *Soit Ω un espace métrique, et soit μ une mesure borélienne sur Ω. On suppose qu'il existe une suite d'ouverts $(\Omega_k)_{k\in\mathbb{N}}$ tels que $\mu(\Omega_k) < \infty$ pour tout k et $\bigcup_{k\in\mathbb{N}} \Omega_k = \Omega$. Alors, pour tout borélien $A \subseteq \Omega$ vérifiant $\mu(A) < \infty$ et pour tout $\varepsilon > 0$, on peut trouver un fermé $F \subseteq \Omega$ et un ouvert $U \subseteq \Omega$ tels que $F \subseteq A \subseteq U$ et $\mu(U\backslash F) < \varepsilon$.*

Démonstration. On dira qu'un borélien $A \subseteq \Omega$ est **régulier pour** μ s'il vérifie la conclusion du théorème : pour tout $\varepsilon > 0$, on peut trouver un fermé $F \subseteq \Omega$ et un ouvert $U \subseteq \Omega$ tels que $F \subseteq A \subseteq U$ et $\mu(U\backslash F) < \varepsilon$. Il s'agit donc de montrer que tout borélien $A \subseteq \Omega$ vérifiant $\mu(A) < \infty$ est régulier pour μ.

Cas 1. On suppose que la mesure μ est finie.

Dans ce cas, on veut montrer que tout borélien $A \subseteq \Omega$ est régulier pour μ. En notant $\mathcal{R}_\mu$ la famille des boréliens réguliers pour μ, il suffit de vérifier que $\mathcal{R}_\mu$ est une tribu et que tout ouvert de Ω est régulier pour μ.

Il est évident que $\Omega \in \mathcal{R}_\mu$ (pour tout $\varepsilon > 0$, on peut prendre $U = \Omega = F$), et facile de voir (exo) que $\mathcal{R}_\mu$ est stable par complémentation (si U et F conviennent pour A et un certain ε, alors F^c et U^c conviennent pour A^c et le même ε). Pour montrer la stabilité par réunions dénombrables, fixons une suite $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{R}_\mu$ et posons $A = \bigcup_{n=0}^\infty A_n$. Fixons également $\varepsilon > 0$. Comme les A_n sont réguliers, on peut pour tout $n \in \mathbb{N}$ choisir un ouvert U_n et un fermé F_n tels que $F_n \subseteq A_n \subseteq U_n$ et $\mu(U_n \backslash F_n) < 2^{-n-1}\varepsilon$. Posons $U = \bigcup_{n=0}^\infty U_n$ et $H = \bigcup_{n=0}^\infty F_n$. Alors $H \subseteq A \subseteq U$; et comme

$$U \backslash H = \left(\bigcup_{n=0}^\infty U_n\right) \backslash \left(\bigcup_{n=0}^\infty F_n\right) \subseteq \bigcup_{n=0}^\infty (U_n \backslash F_n),$$

on a

$$\mu(U\backslash H) \leqslant \sum_{n=0}^\infty \mu(U_n\backslash F_n) < \sum_{n=0}^\infty 2^{-n-1}\varepsilon = \varepsilon.$$

De plus, U est un ouvert de Ω (réunion d'ouverts) ; mais H n'est *a priori* pas un fermé. Cependant, si on pose $H_n = \bigcup_{k=0}^n F_k$ pour tout $n \in \mathbb{N}$, alors les H_n forment une suite croissante de fermés telle que $\bigcup_{n=0}^\infty H_n = H$. Donc $\mu(H_n) \to \mu(H)$; et donc, comme la mesure μ est finie, on peut trouver un entier N tel que $\mu(H\backslash H_N) = \mu(H) - \mu(H_N) < \eta := \varepsilon - \mu(U\backslash H)$. Si on pose $F = H_N$, alors $F \subseteq H \subseteq A$ et $\mu(U\backslash F) = \mu(U\backslash H) + \mu(H\backslash F) \leqslant \mu(U\backslash H) + \mu(H\backslash H_N) < \varepsilon$. On a donc bien montré que $A = \bigcup_{n=0}^\infty A_n$ est régulier pour μ.

Montrons maintenant que tout ouvert $O \subseteq \Omega$ est régulier pour μ. Le point clé (propre aux espaces métriques) est que l'ouvert O est *réunion dénombrable de fermés*. Par exemple, si on pose $F_n := \{x \in \Omega;\ \mathrm{dist}(x, O^c) \geqslant 2^{-n}\}$, alors on vérifie que les F_n forment une suite croissante de fermés telle que $\bigcup_{n=0}^\infty F_n = O$. Pour $\varepsilon > 0$ donné, on peut donc trouver un entier N tel que $\mu(O\backslash F_N) < \varepsilon$. Si on pose $U := O$ et $F := F_N$, on a alors $F \subseteq O \subseteq U$ et $\mu(U\backslash F) < \varepsilon$.

Cas 2. Cas général.

Soit $A \subseteq \Omega$ un borélien vérifiant $\mu(A) < \infty$, et soit $\varepsilon > 0$. Pour tout $k \in \mathbb{N}$, on peut appliquer le Cas 1 à la mesure finie μ_k définie par $\mu_k(B) = \mu(B \cap \Omega_k)$. On obtient ainsi un ouvert U_k et un fermé F_k tels que

$$F_k \subseteq A \subseteq U_k \qquad \text{et} \qquad \mu\big((U_k\backslash F_k) \cap \Omega_n\big) < 2^{-k-1}\varepsilon.$$

Posons alors $U := \bigcup_{k=0}^\infty (U_k \cap \Omega_k)$ et $H := \bigcup_{k=0}^\infty F_k$. Comme les $U_k \cap \Omega_k$ sont des ouverts, U est un ouvert de Ω ; et H est réunion dénombrable de fermés. De plus, on a

$$H = \bigcup_{k=0}^\infty F_k \subseteq A = \bigcup_{k=0}^\infty (A \cap \Omega_k) \subseteq \bigcup_{k=0}^\infty U_k \cap \Omega_k = U;$$

et

$$\begin{aligned}\mu(U\backslash H) &= \mu\left(\bigcup_{k=0}^{\infty}(U_k \cap \Omega_k)\backslash \bigcup_{k=0}^{\infty} F_k\right)\\ &\leqslant \mu\left(\bigcup_{k=0}^{\infty}(U_k \cap \Omega_k)\backslash F_k\right) = \mu\left(\bigcup_{k=0}^{\infty}(U_k\backslash F_n) \cap \Omega_k\right)\\ &< \sum_{k=0}^{\infty} 2^{-k-1}\varepsilon = \varepsilon.\end{aligned}$$

Enfin, comme H est réunion dénombrable de fermés et comme $\mu(H) \leqslant \mu(A) < \infty$, on peut trouver un fermé $F \subseteq H$ tel que $\mu(H\backslash F) < \eta := \varepsilon - \mu(U\backslash H)$ (exo). Alors $F \subseteq A \subseteq U$ et $\mu(U\backslash F) = \mu(U\backslash H) + \mu(H\backslash F) < \varepsilon$.

□

Corollaire 2.2. *Soit Ω un ouvert de $\mathbb{R}^N$, et soit μ une mesure borélienne sur Ω. On suppose qu'on a $\mu(K) < \infty$ pour tout compact $K \subseteq \Omega$. Alors, pour tout borélien $A \subseteq \Omega$ vérifiant $\mu(A) < \infty$ et pour tout $\varepsilon > 0$, on peut trouver un compact $E \subseteq \Omega$ et un ouvert $U \subseteq \Omega$ tels que $E \subseteq A \subseteq U$ et $\mu(U\backslash E) < \varepsilon$.*

Démonstration. On peut trouver une suite $(\Omega_k)_{k\in\mathbb{N}}$ d'ouverts bornés tels que $\overline{\Omega_k} \subseteq \Omega$ pour tout k et $\bigcup_{k=0}^{\infty} \Omega_k = \mathbb{R}^d$. Comme les $\overline{\Omega_k}$ sont des compacts de Ω, on a $\mu(\Omega_k) \leqslant \mu(\overline{\Omega_k}) < \infty$ pour tout k ; donc le théorème s'applique. Pour $\varepsilon > 0$ donné, on peut donc trouver un ouvert U de Ω et un fermé F de Ω tels que $F \subseteq A \subseteq U$ et $\mu(U\backslash F) < \varepsilon$. De plus, le fermé F de Ω est *réunion dénombrable de compacts*. En effet, Ω lui même est réunion dénombrable de compacts, disons $\Omega = \bigcup_{k\in\mathbb{N}} C_k$; donc on peut écrire $F = \bigcup_{k=0}^{\infty} F \cap C_k$, d'où la conclusion puisque chaque $F \cap C_k$ est compact (fermé dans le compact C_k). Comme $\mu(F) \leqslant \mu(A) < \infty$, on peut donc trouver un compact $E \subseteq F$ tel que $\mu(F\backslash E) < \eta := \varepsilon - \mu(U\backslash F)$ (exo). Alors $E \subseteq A \subseteq U$ et $\mu(U\backslash E) = \mu(U\backslash F) + \mu(F\backslash E) < \varepsilon$. □

Exercice. Soit Ω un espace métrique *complet* et *séparable*, et soit μ une mesure borélienne sur Ω. On suppose que tout point $x \in \Omega$ possède un voisinage ouvert V_x tel que $\mu(\Omega_x) < \infty$. Le but de l'exercice est de montrer que pour tout borélien $A \subseteq \Omega$ vérifiant $\mu(A) < \infty$ et pour tout $\varepsilon > 0$, on peut trouver un ouvert U et un compact E tels que $E \subseteq A \subseteq U$ et $\mu(U\backslash E) < \varepsilon$.

(i) Soit F un fermé de Ω. Montrer que pour tout $\eta > 0$ donné, on peut écrire F comme réunion dénombrable de fermés H_n de diamètre inférieur à η. En déduire que si $\mu(F) < \infty$ alors, pour tout $\eta > 0$ et pour tout $\alpha > 0$, on peut trouver un ensemble fermé $H \subseteq F$ tel que H est réunion finie d'ensembles de diamètre inférieur à η et $\mu(F\backslash H) < \alpha$.

(ii) Montrer que pour tout fermé $F \subseteq \Omega$ vérifiant $\mu(F) < \infty$ et pour tout $\varepsilon > 0$, on peut trouver un compact E tel que $E \subseteq F$ et $\mu(F\backslash E) < \varepsilon$. (Observer d'abord qu'un ensemble $E \subseteq \Omega$ est compact si et seulement si il est fermé et précompact.)

(iii) Démontrer le résultat souhaité.

3. Construction de mesures

3.1. Un théorème de prolongement. On va montrer ici que sous des hypothèses raisonnables, une fonction d'ensembles définie sur un certaine famille d'ensembles $\mathcal{C}$ peut se prolonger en une mesure définie sur la tribu $\sigma(\mathcal{C})$ engendrée par $\mathcal{C}$. Il faut malheureusement introduire une définition.

DÉFINITION 3.1. *Soit Ω un ensemble non vide, et soit $\mathcal{C}$ une famille de parties de Ω. On dit que la famille $\mathcal{C}$ est un* ***semi-anneau*** *si elle vérifie les propriétés suivantes :*

- *$\mathcal{C}$ contient $\varnothing$ et est stable par intersections finies ;*
- *Si $C, C' \in \mathcal{C}$, alors $C'\backslash C$ est réunion finie d'éléments de $\mathcal{C}$ deux à deux disjoints.*

EXEMPLE 3.2. Prenons $\Omega = \mathbb{R}^N$. Alors la famille de tous les pavés de $\mathbb{R}^N$ est un semi-anneau.

Démonstration. Il est évident que l'intersection de deux pavés est un pavé : si $P = I_1 \times \cdots \times I_N$ et $Q = J_1 \times \cdots \times J_N$, alors $P \cap Q = (I_1 \cap J_1) \times \cdots \times (I_N \cap J_N)$.

Pour montrer que la différence de deux pavés est réunion finie de pavés disjoints, on procède par récurrence sur la dimension N. Plus précisément, on va montrer que si P et P' sont des pavés de $\mathbb{R}^N$, alors $P'\backslash P$ est réunion d'au plus $2N$ pavés disjoints.

Pour $N = 1$, le résultat est clair : si I, I' sont des intervalles bornés, alors $I'\backslash I$ est effectivement réunion d'au plus 2 intervalles bornés disjoints (faire un dessin).

Supposons avoir établi le résultat pour un certain entier $N \geqslant 1$. Soient P et P' deux pavés de $\mathbb{R}^{N+1}$. Écrivons $P = Q \times I$ et $P' = Q' \times I'$, où Q, Q' sont des pavés de $\mathbb{R}^N$ et I, I' des intervalles bornés. Alors

$$P'\backslash P = \big((Q'\backslash Q) \times I\big) \bigsqcup \big((Q' \cap Q) \times (I'\backslash I)\big) .$$

Comme $Q'\backslash Q$ est réunion d'au plus $2N$ pavés disjoints par hypothèse de récurrence et comme $I'\backslash I$ est réunion d'au plus 2 intervalles bornés disjoints, on en déduit aussitôt que $P'\backslash P$ est réunion d'au plus $2N + 2 = 2(N + 1)$ pavés disjoints. ☐

THÉORÈME 3.3. *Soit Ω un ensemble non vide, et soit $\mathcal{C}$ un semi-anneau de parties de Ω. Soit également $\alpha : \mathcal{C} \to [0, \infty]$ une fonction d'ensembles définie sur $\mathcal{C}$. On fait les hypothèses suivantes.*

(i) *$\alpha(\varnothing) = 0$, et α est* ***finiment additive sur*** *$\mathcal{C}$: si $C_0, \dots, C_n \in \mathcal{C}$ sont deux à deux disjoints et si $C_0 \cup \cdots \cup C_n$* appartient à *$\mathcal{C}$, alors $\alpha(C_0 \cup \cdots \cup C_n) = \alpha(C_0) + \cdots + \alpha(C_n)$.*

(ii) *Si $C \in \mathcal{C}$ et si $(C_k)_{k\in\mathbb{N}}$ est une suite d'éléments de $\mathcal{C}$ telle que $C \subseteq \bigcup_{k=0}^{\infty} C_k$, alors*

$$\alpha(C) \leqslant \sum_{k=0}^{\infty} \alpha(C_k) .$$

Alors il existe une mesure μ sur $(\Omega, \sigma(\mathcal{C}))$ telle que $\mu(C) = \alpha(C)$ pour tout $C \in \mathcal{C}$. Cette mesure μ est donnée par la formule

$$\mu(A) := \inf \sum_{k=0}^{\infty} \alpha(C_k) ,$$

où la borne inférieure est prise sur toutes les suites (C_k) d'éléments de $\mathcal{C}$ telle que $A \subseteq \bigcup_{k=0}^{\infty} C_k$ (avec la convention $\inf \varnothing = \infty$*).*

Démonstration. Soit $\mu^* : \mathcal{P}(\Omega) \to [0, \infty]$ la fonction d'ensemble définie par

$$\mu^*(A) := \inf \left\{ \sum_{k=0}^{\infty} \alpha(C_k);\ (C_k) \subseteq \mathcal{C} \text{ et } A \subseteq \bigcup_{k=0}^{\infty} C_k \right\}.$$

Il est important de noter que μ^* est définie *pour toutes les parties de* Ω. Il s'agit de montrer que la restriction de μ^* à la tribu $\sigma(\mathcal{C})$ est une mesure, et qu'on a $\mu^*(C) = \alpha(C)$ pour tout $C \in \mathcal{C}$. Cela va se faire en plusieurs étapes, dont aucune n'est réellement difficile ; mais au total, la preuve est quand même assez longue.

Fait 1. On a $\mu^*(C) = \alpha(C)$ pour tout $C \in \mathcal{C}$.

Preuve du Fait 1. Soit $C \in \mathcal{C}$ quelconque. On a $C = \bigcup_{k=0}^{\infty} C_k$, avec $C_0 := C$ et $C_k := \emptyset \in \mathcal{C}$ pour $k \geqslant 1$. Comme $\alpha(\emptyset) = 0$, on en déduit $\mu^*(C) \leqslant \sum_{k=0}^{\infty} \alpha(C_k) = \alpha(C_0) = \alpha(C)$. (Ceci ne nécessite pas d'hypothèse sur la famille $\mathcal{C}$.) Réciproquement, l'hypothèse (ii) dit précisément que $\mu^*(C) \geqslant \alpha(C)$. □

Fait 2. La fonction d'ensembles μ^* possède les propriétés suivantes.
- *Non trivialité* : $\mu(\emptyset) = 0$.
- *Croissance* : $\mu^*(A) \leqslant \mu^*(A')$ si $A \subseteq A'$.
- *Sous-additivité dénombrable* : pour toute suite $(A_i)_{i\in\mathbb{N}}$ de parties de Ω, on a

$$\mu^* \left(\bigcup_{i\in\mathbb{N}} A_i \right) \leqslant \sum_{i\in\mathbb{N}} \mu^*(A_i)\,.$$

Preuve du Fait 2. Il est évident par définition que μ^* est croissante ; et on a $\mu(\emptyset) = 0$ par le Fait 1 puisque $\emptyset \in \mathcal{C}$.

Soit $(A_i)_{i\in\mathbb{N}}$ une suite de parties de Ω, et soit

$$A := \bigcup_{i\in\mathbb{N}} A_i\,.$$

Si $\mu^*(A_i) = \infty$ pour un certain i, alors l'inégalité à démontrer est évidente ; donc on suppose qu'on a $\mu^*(A_i) < \infty$ pour tout $i \in \mathbb{N}$.

Soit $\varepsilon > 0$. Par définition de μ^*, on peut, pour tout $i \in \mathbb{N}$, trouver une suite $(C_{k,i})_{k\in\mathbb{N}}$ d'éléments de $\mathcal{C}$ telle que

$$A_i \subseteq \bigcup_{k=0}^{\infty} C_{k,i} \quad \text{et} \quad \sum_{k=0}^{\infty} \alpha(C_{k,i}) \leqslant \mu^*(A_i) + 2^{-i}\varepsilon\,.$$

Comme $\mathbb{N} \times \mathbb{N}$ est dénombrable, on peut énumérer (sans répétition) les termes de la suite double $(C_{k,l})$ en une suite $(C_m)_{m\in\mathbb{N}}$. On a ainsi $A \subseteq \bigcup_{m=0}^{\infty} C_m$, et

$$\begin{aligned} \sum_{m=0}^{\infty} \alpha(C_m) &= \sum_{i=0}^{\infty} \left(\sum_{k=0}^{\infty} \alpha(C_{k,i}) \right) \\ &\leqslant \sum_{i=0}^{\infty} \big(\mu^*(A_i) + 2^{-i}\varepsilon \big) \\ &= \sum_{i=0}^{\infty} \mu^*(A_i) + 2\varepsilon\,. \end{aligned}$$

Par définition de μ^*, on en déduit

$$\mu^*(A) \leqslant \sum_{i=0}^{\infty} \mu^*(A_i) + 2\varepsilon \qquad \text{pour tout } \varepsilon > 0\,;$$

et donc $\mu^*(A) \leqslant \sum_{i=0}^{\infty} \mu^*(A_i)$. □

Remarque. De façon générale, une fonction d'ensemble $\mu^* : \mathcal{P}(\Omega) \to [0, \infty]$ non triviale, croissante et dénombrablement sous-additive s'appelle une **mesure extérieure** sur Ω.

Introduisons maintenant une définition cruciale.

DÉFINITION 3.4. *On dira qu'un ensemble $A \subseteq \Omega$ est* ***μ^*-mesurable au sens de Caratheodory*** *si on a*

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \backslash A) \quad \textit{pour tout } E \subseteq \Omega\,.$$

Et on notera $\mathcal{M}(\mu^*)$ la famille de tous les ensembles μ^*-mesurables.

Cette définition est difficilement compréhensible, et en tous cas semble assez artificielle ; mais c'est pourtant la notion de μ^*-mesurabilité qui est la clé de la preuve du théorème.

Remarque. Par le Fait 2, on a $\mu^*(E) = \mu^*((E \cap A) \bigcup (E \backslash A)) \leqslant \mu^*(E \cap A) + \mu^*(E \backslash A)$ pour tout $A, E \subseteq \Omega$. Donc un ensemble $A \subseteq \Omega$ est μ^*-mesurable si et seulement si

$$\mu^*(E) \geqslant \mu^*(E \cap A) + \mu^*(E \backslash A) \quad \text{pour tout } E \subseteq \Omega\,.$$

FAIT 3. Tous les éléments de $\mathcal{C}$ sont μ^*-mesurables.

Preuve du Fait 3. Soit $C \in \mathcal{C}$, et soit $E \subseteq \Omega$ quelconque. Il s'agit de montrer que $\mu^*(E) \geqslant \mu^*(E \cap C) + \mu^*(E \backslash C)$; et comme ceci est évident si $\mu^*(E) = \infty$, on suppose qu'on a $\mu^*(E) < \infty$.

Soit $\varepsilon > 0$. Par définition de μ^*, on peut trouver une suite (C_k) d'éléments de $\mathcal{C}$ telle que

$$E \subseteq \bigcup_{k=0}^{\infty} C_k \quad \text{et} \quad \sum_{k=0}^{\infty} \alpha(C_k) \leqslant \mu^*(E) + \varepsilon\,. \tag{3.1}$$

Fixons $k \in \mathbb{N}$. Comme $\mathcal{C}$ est un semi-anneau, on peut trouver $C^0, \dots, C^{n_k} \in \mathcal{C}$ deux à deux disjoints tels que $C_k \backslash C = C^0 \cup \dots \cup C^{n_k}$. On a alors

$$C_k = (C_k \cap C) \sqcup C^0 \sqcup \dots \sqcup C^{n_k}\,;$$

et donc, comme α est finiment additive sur $\mathcal{C}$:

$$\begin{aligned} \alpha(C_k) &= \alpha(C_k \cap C) + \sum_{i=0}^{n_k} \alpha(C_k \cap C^i) \\ &= \mu^*(C_k \cap C) + \sum_{i=0}^{n_k} \mu^*(C_k \cap C^i) \quad \text{par le Fait 1} \\ &\geqslant \mu^*(C_k \cap C) + \mu^*(C_k \backslash C) \quad \text{par le Fait 2}\,, \end{aligned}$$

puisque $C_k\backslash C = \bigcup_{i=0}^{n_k} C^i$. Ceci étant vrai pour tout $k \in \mathbb{N}$, on en déduit

$$\begin{aligned}\sum_{k=0}^{\infty}\alpha(C_k) &\geqslant \sum_{k=0}^{\infty}\mu^*(C_k\cap C)+\sum_{k=0}^{\infty}\mu^*(C_k\backslash C)\\ &\geqslant \mu^*(E\cap C)+\mu^*(E\backslash C)\quad\text{par le Fait 2}\,,\end{aligned}$$

car $E\cap C\subseteq \bigcup_{k=0}^{\infty}(C_k\cap C)$ et $E\backslash C\subseteq\bigcup_{k=0}^{\infty}(C_k\backslash C)$.

En revenant à (3.1), on obtient finalement

$$\mu^*(E\cap C)+\mu^*(E\backslash C)\leqslant \mu^*(E)+\varepsilon\quad\text{pour tout } \varepsilon>0\,,$$

et donc $\mu^*(E)\geqslant \mu^*(E\cap C)+\mu^*(E\backslash C)$. □

Le point clé est maintenant le lemme suivant.

Lemme 3.5. *La famille $\mathcal{M}(\mu^*)$ est une tribu, et la restriction de μ^* à $\mathcal{M}(\mu^*)$ est une mesure.*

Preuve du Lemme 3.5. On a besoin... de plusieurs étapes.

Étape 0. On a $\varnothing\in\mathcal{M}(\mu^*)$ et $\mu^*(\varnothing)=0$.

Démonstration. On a déjà vu que $\mu^*(\varnothing)=0$. Le fait que $\varnothing$ appartienne à $\mathcal{M}(\mu^*)$ découle formellement du Fait 3 puisque $\varnothing\in\mathcal{C}$, mais peut aussi se voir directement : pour tout $E\subseteq\Omega$, on a $\mu^*(E)=\mu^*(E)+0=\mu^*(E\backslash\varnothing)+\mu^*(E\cap\varnothing)$. □

Étape 1. $\mathcal{M}(\mu^*)$ est stable par passage au complémentaire.

Démonstration. C'est évident par définition. □

Étape 2. Si $A,B\in\mathcal{M}(\mu^*)$, alors $A\cup B\in\mathcal{M}(\mu^*)$ et $A\cap B\in\mathcal{M}(\mu^*)$. De plus, si $A\cap B=\varnothing$, alors $\mu^*\big(E\cap(A\cup B)\big)=\mu^*(E\cap A)+\mu^*(E\cap B)$ pour tout $E\subseteq\Omega$.

Démonstration. (a) Pour tout $E\subseteq\Omega$, on a d'une part

$$\begin{aligned}\mu^*\big(E\cap(A\cup B)\big) &= \mu^*\Big(\big(E\cap(A\cup B)\big)\cap A\Big)+\mu^*\Big(\big(E\cap(A\cup B)\big)\backslash A\Big)\\ &= \mu^*(E\cap A)+\mu^*\big((E\backslash A)\cap B\big)\,;\end{aligned}$$

et d'autre part

$$\mu^*\big(E\backslash(A\cup B)\big)=\mu^*\big((E\backslash A)\backslash B\big)\,.$$

Donc

$$\begin{aligned}\mu^*\big(E\cap(A\cup B)\big)+\mu^*\big(E\backslash(A\cup B)\big) &= \mu^*(E\cap A)+\mu^*\big((E\backslash A)\cap B\big)+\mu^*\big((E\backslash A)\backslash B\big)\\ &= \mu^*(E\cap A)+\mu^*(E\backslash A)\\ &= \mu^*(E)\,.\end{aligned}$$

Par conséquent, $A\cup B\in\mathcal{M}(\mu^*)$ pour tous $A,B\in\mathcal{M}(\mu^*)$; et comme $\mathcal{M}(\mu^*)$ est stable par passage au complémentaire, on en déduit que $A\cap B=\big(A^c\cup B^c\big)^c\in\mathcal{M}(\mu^*)$.

(b) Si de plus $A\cap B=\varnothing$, on a pour tout $E\subseteq\Omega$:

$$\begin{aligned}\mu^*\big(E\cap(A\cup B)\big) &= \mu^*\Big(\big(E\cap(A\cup B)\big)\cap A\Big)+\mu^*\Big(\big(E\cap(A\cup B)\big)\backslash A\Big)\\ &= \mu^*(E\cap A)+\mu^*(E\cap B)\,.\end{aligned}$$

□

Étape 3. Si $(A_k)_{k\in\mathbb{N}}$ est une suite d'ensemble μ^*-mesurables *deux à deux disjoints*, alors $\bigcup_{k=0}^\infty A_k \in \mathcal{M}(\mu^*)$ et

$$\mu^*\left(\bigcup_{k=0}^\infty A_k\right) = \sum_{k=0}^\infty \mu^*(A_k)\,.$$

Démonstration. On posera $A := \bigcup_{k=0}^\infty A_k$. Par l'Étape 2, on sait que $A_0\cup\cdots\cup A_n \in \mathcal{M}(\mu^*)$ pour tout $n\in\mathbb{N}$ et qu'on a

$$\mu^*\big(E\cap(A_0\cup\cdots\cup A_n)\big) = \sum_{k=0}^n \mu^*(E\cap A_k) \quad \text{pour tout } E\subseteq\Omega\,.$$

On en déduit :

$$\begin{aligned}
\mu^*(E) &= \mu^*\big(E\cap(A_0\cup\cdots\cup A_n)\big) + \mu^*\big(E\backslash(A_0\cup\cdots\cup A_n)\big)\\
&= \sum_{k=0}^n \mu^*(E\cap A_k) + \mu^*\big(E\backslash(A_0\cup\cdots\cup A_n)\big)\\
&\geqslant \sum_{k=0}^n \mu^*(E\cap A_k) + \mu^*(E\backslash A) \qquad \text{car } \mu^* \text{ est croissante}\,.
\end{aligned}$$

En faisant tendre n vers l'infini, on obtient

$$\mu^*(E) \geqslant \sum_{k=0}^\infty \mu^*(E\cap A_k) + \mu^*(E\backslash A) \qquad \text{pour tout } E\subseteq\Omega\,. \tag{3.2}$$

Par le Fait 2 et comme $\bigcup_k (E\cap A_k) = E\cap A$, cela entraine en particulier

$$\mu^*(E) \geqslant \mu^*(E\cap A) + \mu^*(E\backslash A)\,;$$

donc $A = \bigcup_{k=0}^\infty A_k \in \mathcal{M}(\mu^*)$. Enfin, en prenant $E := A$ dans (3.2), on obtient

$$\mu^*(A) \geqslant \sum_{k=0}^\infty \mu^*(A_k) + \mu(\varnothing)\,,$$

et donc $\mu^*(A) = \sum_{k=0}^\infty \mu^*(A_k)$, à nouveau par le Fait 2. □

Étape 4. $\mathcal{M}(\mu^*)$ est stable par réunion dénombrable.

Démonstration. Soit $(B_k)_{n\in\mathbb{N}}$ une suite quelconque d'ensemble μ^*-mesurables. Si on pose $A_0 := B_0$ et

$$A_k := B_k\backslash\big(B_0\cup\cdots\cup B_{k-1}\big) = B_k\cap\big(B_0\cup\cdots\cup B_{k-1}\big)^c \quad \text{pour } k\geqslant 1\,,$$

alors les A_k sont μ^*-mesurables par les Étapes 1 et 2, et on a $\bigcup_{k=0}^\infty A_k = \bigcup_{k=0}^\infty B_k$. Comme les A_k sont deux à deux disjoints, on en déduit que $\bigcup_{k=0}^\infty B_k \in \mathcal{M}(\mu^*)$ par l'Étape 3. □

La preuve du Lemme 3.5 est maintenant terminée : par les Étapes 0, 1 et 4, $\mathcal{M}(\mu^*)$ est une tribu ; et par l'Étape 3, la restriction de μ^* à $\mathcal{M}(\mu^*)$ est une mesure. □

Remarque. Dans la preuve du lemme, on n'a utilisé à aucun moment la *définition* de μ^* : la seule chose qui importe est que μ^* soit une mesure extérieure, *i.e.* vérifie les propriétés listées dans le Fait 2.

La preuve du Théorème 3.3 est maintenant achevée : par le Fait 3, la tribu $\mathcal{M}(\mu^*)$ contient $\mathcal{C}$, donc $\sigma(\mathcal{C})$; par le Lemme 3.5, on en déduit que la restriction μ de μ^* à $\sigma(\mathcal{C})$ est une mesure ; et par le Fait 1, on a $\mu(C) = \alpha(C)$ pour tout $C \in \mathcal{C}$. □

3.2. La mesure de Lebesgue. Le Théorème 3.3 va permettre de démontrer très facilement l'existence de la mesure de Lebesgue sur $\mathbb{R}^N$, $N \geqslant 1$. C'est heureux, car il eût été dommage de se fatiguer pour rien.

Prenons pour $\mathcal{C}$ le semi-anneau constitué par tous les pavés de $\mathbb{R}^N$, et soit $\alpha : \mathcal{C} \to \mathbb{R}^+$ la fonction d'ensembles définie par

$$\alpha(P) := |P| \quad \text{pour tout pavé } P \subseteq \mathbb{R}^N .$$

Pour montrer qu'il existe une mesure μ sur $\mathbb{R}^N$ vérifiant $\mu(P) = |P| = \alpha(P)$ pour tout pavé $P \subseteq \mathbb{R}^N$, il suffit de vérifier les deux propriétés suivantes.

(i) Si $P_0, \dots, P_n$ sont des pavés deux à deux disjoints et si $P_0 \cup \cdots \cup P_n$ *est un pavé*, alors $|P_0 \cup \cdots \cup P_n| = |P_0| + \cdots + |P_n|$.

(ii) Si P est un pavé et si $(P_k)_{k\in\mathbb{N}}$ est une suite de pavés telle que $P \subseteq \bigcup_{k=0}^{\infty} P_k$, alors

$$|P| \leqslant \sum_{k=0}^{\infty} |P_k| .$$

On aura pour cela besoin du lemme suivant, que l'on peut démontrer en utilisant ironiquement... l'intégrale des fonctions en escalier. (On pourrait aussi s'en passer.)

Lemme 3.6. *Soient $I_0, \dots, I_m, J_0, \dots, J_n$ des intervalles bornés de $\mathbb{R}$, et soient $\alpha_0, \dots, \alpha_m, \beta_0, \dots, \beta_n \in \mathbb{R}$. Si on a*

$$\sum_{l=0}^{m} \alpha_l \, \mathbf{1}_{I_l} \leqslant \sum_{k=0}^{n} \beta_k \, \mathbf{1}_{J_k} ,$$

alors

$$\sum_{l=0}^{m} \alpha_l \, |I_l| \leqslant \sum_{k=0}^{n} \beta_k \, |J_k| .$$

(Par conséquent, si $\sum_l \alpha_l \, \mathbf{1}_{I_l} = \sum_k \beta_k \, \mathbf{1}_{J_k}$ alors $\sum_l \alpha_l \, |I_l| = \sum_k \beta_k \, |J_k|$.)

Preuve du lemme 3.6. Choisissons un intervalle $[a, b]$ contenant tous les I_k et tous les J_l. Alors $\varphi := \sum_{l=0}^{m} \alpha_l \, \mathbf{1}_{I_l}$ et $\psi := \sum_{k=0}^{n} \beta_k \, \mathbf{1}_{J_k}$ sont des fonctions en escaliers sur $[a, b]$, et $\varphi \leqslant \psi$. Donc $\int_a^b \varphi(t)\, dt \leqslant \int_a^b \psi(t)\, dt$, ce qui est la conclusion souhaitée. □

Exercice. Soit Ω un ensemble non vide, et soit $\mathcal{C}$ un semi-anneau de parties de Ω. On se donne une fonction d'ensembles finiment additive sur $\mathcal{C}$, notée $C \mapsto |C|$, et on suppose que $|C| < \infty$ pour tout $C \in \mathcal{C}$. Enfin, on note $\mathcal{E}(\Omega, \mathcal{C})$ l'espace vectoriel engendré par les fonction $\mathbf{1}_C$, $C \in \mathcal{C}$. Montrer que la somme $\sum_{i=1}^{n} \alpha_i |C_i|$ ne dépend pas de l'écriture d'une fonction $\varphi \in \mathcal{E}(\Omega, \mathcal{C})$ sous la forme $\varphi = \sum_{i=1}^{n} \alpha_i \mathbf{1}_{C_i}$ avec $\alpha_i \in \mathbb{R}$ et $C_i \in \mathcal{C}$. (On pourra commencer par montrer que toute fonction $\varphi \in \mathcal{E}(\Omega, \mathcal{C})$ peut s'écrire sous la forme $\varphi = \sum_{k=1}^{N} \phi_k \mathbf{1}_{\Omega_k}$ avec $\phi_k \in \mathbb{R}$ et des $\Omega_k \in \mathcal{C}$ deux à deux disjoints.) En déduire une forme “abstraite” du Lemme 3.6.

Preuve de (i). Soient $P_0, \dots, P_n$ des pavés de $\mathbb{R}^N$ deux à deux disjoints tels que $P := P_0 \cup \cdots \cup P_n$ soit un pavé. On a par hypothèse

$$\mathbf{1}_P = \sum_{k=0}^{n} \mathbf{1}_{P_k} .$$

Si $N = 1$, le Lemme 3.6 donne immédiatement

$$|P| = \sum_{k=0}^{n} |P_k| .$$

Si $N = 2$, écrivons $P = I_1 \times I_2$ et $P_k = I_{k,1} \times I_{k,2}$ pour $k = 0, \dots n$. Alors

$$\mathbf{1}_{I_1}(t_1)\mathbf{1}_{I_2}(t_2) = \sum_{k=0}^{n} \mathbf{1}_{I_{k,1}}(t_1)\mathbf{1}_{I_{k,2}}(t_2) \quad \text{pour tous } t_1, t_2 \in \mathbb{R} .$$

En appliquant le Lemme 3.6 avec $m := 0$, $\alpha_0 := \mathbf{1}_{I_1}(t_1)$ et $n := n$, $\beta_k := \mathbf{1}_{I_{k,1}}(t_1)$, on obtient

$$\mathbf{1}_{I_1}(t_1)\,|I_2| = \sum_{k=0}^{n} \mathbf{1}_{I_{k,1}}(t_1)|I_{k,2}| \quad \text{pour tout } t_1 \in \mathbb{R} ;$$

d'où en appliquant à nouveau le Lemme 3.6 :

$$|I_2|\,|I_1| = \sum_{k=0}^{n} |I_{k,2}|\,|I_{k,1}| ,$$

ce qui est le résultat souhaité.

Pour un entier $N \geqslant 1$ quelconque, la preuve est identique en appliquant N fois le Lemme 3.6. $\square$

Preuve de (ii). Soit P un pavé de $\mathbb{R}^N$, et soit $(P_k)_{k\in\mathbb{N}}$ une suite de pavés telle que $P \subseteq \bigcup_{k=0}^{\infty} P_k$. On va distinguer 2 cas.

Cas 1. P est *compact* et les P_k sont *ouverts*.

Dans ce cas, le pavé compact P est recouvert par un nombre *fini* des pavés ouverts P_k : on peut trouver un entier n tel que $P \subseteq P_0 \cup \cdots \cup P_n$. On a alors

$$\mathbf{1}_P \leqslant \mathbf{1}_{P_0\cup\cdots\cup P_n} \leqslant \sum_{k=0}^{n} \mathbf{1}_{P_k} ;$$

et en raisonnant exactement comme dans la preuve de (i), on en déduit

$$|P| \leqslant \sum_{k=0}^{n} |P_k| \leqslant \sum_{k=0}^{\infty} |P_k| .$$

Cas 2. Cas général.

Soit $\varepsilon > 0$ quelconque. On peut trouver un pavé *compact* Q tel que $Q \subseteq P$ et $|Q| \geqslant |P| - \varepsilon$; et pour tout $k \in \mathbb{N}$, on peut trouver un pavé *ouvert* Q_k tel que $P_k \subseteq Q_k$ et $|Q_k| \leqslant |P_k| + 2^{-k}\varepsilon$. Alors $Q \subseteq \bigcup_{k=0}^{\infty} Q_k$; donc d'après le Cas 1 :

$$|Q| \leqslant \sum_{k=0}^{\infty} |Q_k| .$$

On obtient ainsi

$$|P| - \varepsilon \leqslant \sum_{k=0}^{\infty} \big(|P_k| + 2^{-k}\varepsilon\big) = \sum_{k=0}^{\infty} |P_k| + 2\varepsilon$$

pour tout $\varepsilon > 0$, et donc $|P| \leqslant \sum_0^\infty |P_k|$. □

3.3. Ensembles mesurables au sens de Lebesgue. Cette section est consacrée à un point "technique" qu'on a souvent tendance à passer sous silence. Il s'agit de dire qui sont, finalement, les parties de $\mathbb{R}^N$ que l'on peut "mesurer".

Notons λ^* la **mesure extérieure de Lebesgue** sur $\mathbb{R}^N$: par définition, $\lambda^* : \mathcal{P}(\mathbb{R}^N) \to [0, \infty]$ est la fonction d'ensembles définie par

$$\lambda^*(A) := \inf \left\{ \sum_{k=0}^{\infty} |P_k|;\ (P_k) \subseteq P \text{ et } A \subseteq \bigcup_{k=0}^{\infty} P_k \right\},$$

où $\mathcal{P}$ est la famille de tous les pavés de $\mathbb{R}^N$.

On dira qu'un ensemble $A \subseteq \mathbb{R}$ est **mesurable au sens de Lebesgue** s'il est λ^*-mesurable au sens de Caratheodory (Définition 3.4) ; autrement dit si on a

$$\lambda^*(E) = \lambda^*(E \cap A) + \lambda^*(E \backslash A) \quad \text{pour tout } E \subseteq \mathbb{R}^N.$$

Par le Lemme 3.5, on sait que la famille des ensembles mesurables au sens de Lebesgue est une tribu contenant tous les boréliens, et qu'on a $\lambda^*(B) = \lambda_N(B)$ pour tout *borélien* $B \subseteq \mathbb{R}$.

La proposition qui va suivre donne une description nettement plus intuitive des ensembles mesurables au sens de Lebesgue : un ensemble est mesurable si et seulement si il est "borélien à un ensemble négligeable près". Démontrons d'abord un lemme important.

Lemme 3.7. *Un ensemble $Z \subseteq \mathbb{R}$ est λ_N-négligeable si et seulement si $\lambda^*(Z) = 0$.*

Démonstration. Si Z est λ_N-négligeable, on peut trouver un borélien B tel que $Z \subseteq B$ et $\lambda_N(B) = 0$. Comme B est borélien, on sait que $\lambda_N(B) = \lambda^*(B)$; donc $\lambda^*(B) = 0$, et donc $\lambda^*(Z) = 0$ par croissance de λ^*.

Inversement, supposons que $\lambda^*(Z) = 0$. Par définition de λ^*, on peut, pour tout $m \in \mathbb{N}$, trouver une suite de pavés $(P_{k,m})_{k \in \mathbb{N}}$ telle que

$$Z \subseteq \bigcup_{k=0}^{\infty} P_{k,m} \quad \text{et} \quad \sum_{k=0}^{\infty} |P_{k,m}| \leqslant 2^{-m}.$$

Si on pose $B_m := \bigcup_{k=0}^\infty I_{k,m}$ et $B := \bigcap_{m=0}^\infty B_m$, alors les B_m sont boréliens, donc B également, et $Z \subseteq B$. De plus, on a par sous-additivité dénombrable de la mesure de Lebesgue λ_N :

$$\lambda_N(B_m) \leqslant \sum_{k=0}^{\infty} |I_{k,m}| \leqslant 2^{-m} \quad \text{pour tout } m \in \mathbb{N};$$

et donc $\lambda_N(B) = 0$ car $\lambda_N(B) \leqslant \lambda_N(B_m)$ pour tout m. Donc Z est λ_1-négligeable puisque $Z \subseteq B$. □

Corollaire 3.8. *Tout ensemble λ_N-négligeable est mesurable au sens de Lebesgue.*

Démonstration. Si Z est λ_N-négligeable, alors $\lambda^*(Z) = 0$. Donc on a pour tout $E \subseteq \mathbb{R}$:

$$\lambda^*(E \cap Z) + \lambda^*(E \backslash Z) = \lambda^*(E \backslash Z) \leqslant \lambda^*(E)\,;$$

et donc Z est λ^*-mesurable. □

Voici maintenant la description annoncée des ensembles mesurables. (Dans la condition (4) ci-dessous, Δ désigne la *différence symétrique* : $A\Delta\widetilde{A} = (A\backslash\widetilde{A}) \cup (\widetilde{A}\backslash A)$.)

Proposition 3.9. *Pour $A \subseteq \mathbb{R}^N$, les propriétés suivantes sont équivalentes :*

(1) *A est mesurable au sens de Lebesgue ;*

(2) *on peut trouver deux boréliens B et B' tels que $B' \subseteq A \subseteq B$ et $\lambda_N(B\backslash B') = 0$;*

(3) *on peut trouver un borélien B tel que $A \subseteq B$ et $B\backslash A$ est λ_N-négligeable ;*

(4) *on peut trouver un borélien $\widetilde{A}$ tel que $A\Delta\widetilde{A}$ est λ_N-négligeable.*

Démonstration. Les implications (2) $\implies$ (3) $\implies$ (4) sont évidentes. On va montrer que (4) entraine (2) (de sorte que (2), (3), (4) sont équivalentes), puis que (1) entraine (3) et que (2) entraine (1).

(4) $\implies$ (2). Supposons (4) vérifiée avec un certain borélien $\widetilde{A}$. On peut donc trouver un borélien E tel que

$$A\Delta\widetilde{A} = (A\backslash\widetilde{A}) \cup (\widetilde{A}\backslash A) \subseteq E \quad \text{et} \quad \lambda_N(E) = 0\,.$$

Si on pose $B := \widetilde{A} \cup E$ et $B' := \widetilde{A}\backslash E$, alors B et B' sont boréliens et $B' \subseteq A \subseteq B$ (vérifier). De plus, $B\backslash B' \subseteq E$ et donc $\lambda_N(B'\backslash B) = 0$. Par conséquent, (2) est vérifiée.

(1) $\implies$ (3). Supposons que A soit mesurable au sens de Lebesgue. On cherche un borélien B tel que $A \subseteq B$ et $\lambda^*(B\backslash A) = 0$. On va distinguer deux cas

Cas 1. On suppose que $\lambda^*(A) < \infty$.

Par définition de λ^*, on peut, pour tout $m \in \mathbb{N}$, trouver une suite de pavés $(P_{k,m})_{k\in\mathbb{N}}$ telle que

$$A \subseteq \bigcup_{k=0}^{\infty} P_{k,m} \quad \text{et} \quad \sum_{k=0}^{\infty} |P_{k,m}| \leqslant \lambda^*(A) + 2^{-m}\,,$$

Posons alors

$$B_m := \bigcup_{k=0}^{\infty} I_{k,m} \quad \text{et} \quad B := \bigcap_{m=0}^{\infty} B_m\,.$$

Par définition, les B_m et B sont boréliens et $A \subseteq B$. De plus, comme A est λ^*-mesurable et $\lambda^*(A) < \infty$, on a pour tout $m \in \mathbb{N}$:

$$\begin{aligned}
\lambda^*(B_m\backslash A) &= \lambda^*(B_m) - \lambda^*(A) \\
&\leqslant \sum_{k=0}^{\infty} \lambda^*(I_{k,m}) - \lambda^*(A) \quad \text{par sous-additivité dénombrable} \\
&= \sum_{k=0}^{\infty} |I_{k,m}| - \lambda^*(A) \leqslant 2^{-m}.
\end{aligned}$$

Donc $\lambda^*(B\backslash A) = 0$ puisque $B\backslash A \subseteq B_m\backslash A$ pour tout m ; et donc $B\backslash A$ est λ_N-négligeable d'après le Lemme 3.7.

Cas 2. Cas général.

Soit (Π^n) une suite de pavés telle que $\bigcup_n \Pi^n = \mathbb{R}^N$, et posons $A^n := A \cap \Pi^n$. Les A^n sont mesurables au sens de Lebesgue car A l'est et les Π^n sont boréliens (donc mesurables), et $A = \bigcup_n A^n$. De plus, $\lambda^*(A^n) \leqslant \lambda^*(J^n) = |J^n| < \infty$. Par le Cas 1, on peut donc trouver, pour tout $n \in \mathbb{N}$, un borélien B^n tel que $A^n \subseteq B^n$ et $B^n \backslash A^n$ est λ_N-négligeable. Alors $A \subseteq B := \bigcup_n B^n$, l'ensemble B est borélien, et $B\backslash A$ est λ_N-négligeable car $B\backslash A \subseteq \bigcup_n (B^n\backslash A^n)$.

(2) $\implies$ (1). Supposons (4) vérifiée avec des boréliens B et B'. Alors $A = B' \cup (A\backslash B')$; donc A est mesurable au sens de Lebesgue car B' est borélien (donc mesurable) et $A\backslash B' \subseteq B\backslash B'$ est λ_N-négligeable (donc mesurable par le Lemme 3.7). □

Exercice. Soit $A \subseteq \mathbb{R}^N$ un ensemble mesurable au sens de Lebesgue. Montrer qu'on peut trouver deux ensembles G et K tels que G est G_δ (intersection dénombrable d'ouverts), K est K_σ (réunion dénombrable de compacts), $K \subseteq A \subseteq G$ et $\lambda_N(G\backslash H) = 0$. (En particulier A est à la fois "G_δ modulo un ensemble négligeable" et "K_σ modulo un ensemble négligeable".)

4. Mesures produits et théorème de Fubini "général"

4.1. Tribus produits. Soient $(\Omega_1, \mathcal{B}_1), \dots, (\Omega_N, \mathcal{B}_N)$ des espaces mesurables, et soit $\Omega := \Omega_1 \times \cdots \times \Omega_N$. On appellera **pavé mesurable** tout ensemble $A \subseteq \Omega$ de la forme
$$A = A_1 \times \cdots \times A_N\,, \quad \text{avec } A_j \in \mathcal{B}_j \text{ pour } j = 1, \dots, N\,.$$
La **tribu produit** sur $\Omega = \Omega_1 \times \cdots \times \Omega_N$ est la tribu de parties de Ω engendrée par les pavés mesurables. Cette tribu se note $\mathcal{B}_1 \otimes \cdots \otimes \mathcal{B}_N$.

Exemple 4.1. Si $\Omega_1, \dots, \Omega_N$ sont des espaces métriques séparables, alors la tribu $\mathcal{B}(\Omega_1) \otimes \cdots \otimes \mathcal{B}(\Omega_N)$ est la tribu borélienne de $\Omega_1 \times \cdots \times \Omega_N$. En particulier, la tribu $\mathcal{B}(\mathbb{R}) \otimes \cdots \otimes \mathcal{B}(\mathbb{R})$ est la tribu borélienne de $\mathbb{R}^N$; et si $p, q \in \mathbb{N}^*$, alors $\mathcal{B}(\mathbb{R}^p) \otimes \mathcal{B}(\mathbb{R}^q) = \mathcal{B}(\mathbb{R}^{p+q})$.

Démonstration. On a déja vu que si $A_1, \dots, A_N$ sont des boréliens de $\Omega_1, \dots, \Omega_N$, alors $A_1 \times \cdots \times A_N$ est borélien dans $\Omega_1 \times \cdots \times \Omega_N$; autrement dit : tout "pavé borélien" est borélien dans $\Omega_1 \times \cdots \times \Omega_N$, et donc $\mathcal{B}(\Omega_1) \otimes \cdots \otimes \mathcal{B}(\Omega_N) \subseteq \mathcal{B}(\Omega_1 \times \cdots \times \Omega_N)$ par définition de la tribu produit. Inversement, $\mathcal{B}(\Omega_1) \otimes \cdots \otimes \mathcal{B}(\Omega_N)$ contient tous les *produits d'ouverts*, c'est à dire tous les ensembles de la forme $O = O_1 \times \cdots \times O_N$ avec O_j ouvert dans Ω_j pour $j = 1, \dots ; N$. Comme les Ω_j sont supposés séparables, on sait que ces produits d'ouverts engendrent la tribu borélienne de $\Omega_1 \times \cdots \times \Omega_N$ (c'est l'Exercice 4.9 du Chapitre 2) ; donc $\mathcal{B}(\Omega_1) \otimes \cdots \otimes \mathcal{B}(\Omega_N)$ contient $\mathcal{B}(\Omega_1 \times \cdots \times \Omega_N)$. □

Chaque fois qu'on aura affaire à des espaces mesurables $(\Omega_1 \mathcal{B}_1)$, ... ,$(\Omega_N, \mathcal{B}_N)$, on considérera que $\Omega_1 \times \cdots \times \Omega_N$ est "par défaut" muni de la tribu produit. L'exercice suivant est important.

Exercice 4.2. Soit $(\Omega, \mathcal{B})$ un espace mesurable, et soit $f : \Omega \to \Omega_1 \times \cdots \times \Omega_N$. On écrit $f(x) = (f_1(x), \dots, f_N(x))$. Montrer que f est mesurable si et seulement si ses composantes $f_1, \dots, f_N$ le sont.

4.2. Mesure produit. Dans cette section, $(\Omega_1, \mathcal{B}_1)$ et $(\Omega_2, \mathcal{B}_2)$ sont deux espaces mesurables. Comme on n'a que deux espaces, on parlera de **rectangles mesurables** plutôt que de pavés mesurables.

On rappelle qu'une mesure μ sur un espace mesurable $(\Omega, \mathcal{B})$ est dite **σ-finie** s'il existe une suite (E_n) d'ensembles mesurables telle que $\Omega = \bigcup_0^\infty E_n$ et $\mu(E_n) < \infty$ pour tout $n \in \mathbb{N}$. On peut alors toujours supposer que la suite (E_n) est *croissante*, quitte à remplacer E_n par $E_0 \cup \cdots \cup E_n$.

THÉORÈME 4.3. *Si μ_1 et μ_2 sont deux mesures σ-finies sur $(\Omega_1, \mathcal{B}_1)$ et $(\Omega_2, \mathcal{B}_2)$ respectivement, alors il existe une unique mesure μ sur $(\Omega_1 \times \Omega_2, \mathcal{B}_1 \otimes \mathcal{B}_2)$ telle que*

$$\mu(A_1 \times A_2) = \mu_1(A_1)\, \mu_2(A_2)$$

*pour tout rectangle mesurable $A_1 \times A_2 \subseteq \Omega_1 \times \Omega_2$. La mesure μ s'appelle la **mesure produit** des mesures μ_1 et μ_2, et se note $\mu_1 \otimes \mu_2$.*

Exemple. Si $p, q \in \mathbb{N}^*$, alors $\lambda_{p+q} = \lambda_p \otimes \lambda_q$.

Preuve du théorème. On notera $\mathcal{R}$ la famille de tous les rectangles mesurables $R = A_1 \times A_2 \subseteq \Omega_1 \times \Omega_2$; et pour tout $R = A_1 \times A_2 \in \mathcal{R}$, on posera

$$|R| = |A_1 \times A_2| := \mu_1(A_1)\, \mu_2(A_2)\,.$$

(1) L'*unicité* de la mesure produit découle du Théorème 1.3 appliqué à la famille

$$\mathcal{C} := \{C \in \mathcal{R};\ |C| < \infty\}\,.$$

La famille $\mathcal{C}$ est stable par intersections finies car $\mathcal{R}$ l'est et $|C \cap R| \leqslant |R|$ si $C \in \mathcal{C}$ et $R \in \mathcal{R}$ (cela montre en fait que $C \cap R \in \mathcal{C}$ dès que $C \in \mathcal{C}$ et $R \in \mathcal{R}$). Comme les mesures μ_1 et μ_2 sont σ-finies, on peut trouver deux suites croissantes $(E_{n,1})_{n\in\mathbb{N}} \subseteq \mathcal{B}_1$ et $(E_{n,2})_{n\in\mathbb{N}} \subseteq \mathcal{B}_2$ telles que $\Omega_i = \bigcup_{n=0}^\infty E_{n,i}$ pour $i = 1, 2$ et $\mu_i(E_{n,i}) < \infty$. Alors $E_n := E_{n,1} \times E_{n,2} \in \mathcal{C}$ pour tout $n \in \mathbb{N}$ et $\Omega_1 \times \Omega_2 = \bigcup_0^\infty E_n$. Enfin, $\mathcal{C}$ engendre la tribu $\mathcal{B}_1 \otimes \mathcal{B}_2$ car $\sigma(\mathcal{C})$ contient tous les rectangles mesurables : si $R \in \mathcal{R}$, alors $R = \bigcup_0^\infty C_n$, où $C_n := R \cap E_n \in \mathcal{C}$. Donc le Théorème 1.3 s'applique bel et bien.

(2) Pour l'*existence* de la mesure produit, on utilise le Théorème 3.3, en prenant cette fois pour $\mathcal{C}$ la famille $\mathcal{R}$ de *tous* les rectangles mesurables et pour α la fonction d'ensembles évidente : $\alpha(R) := |R|$ pour tout $R \in \mathcal{R}$.

Il n'est pas difficile de voir que la famille $\mathcal{C}$ est un semi-anneau. (Plus précisément, on montre que si $C, C' \in \mathcal{C}$, alors $C' \backslash C$ est réunion de 2 éléments de $\mathcal{C}$ disjoints : *cf* la preuve de l'Exemple 3.2.) Il s'agit de vérifier les propriétés (i) et (ii) requises pour appliquer le théorème.

(i) Soient $R_0, \dots, R_n \in \mathcal{R}$ deux à deux disjoints tels que $R := R_0 \cup \cdots \cup R_n$ appartienne encore à $\mathcal{R}$. Écrivons $R = A_1 \times A_2$ et $R_k = A_{k,1} \times A_{k,2}$ pour $k = 0, \dots, n$. Comme les R_k sont deux à deux disjoints, on a $\mathbf{1}_R = \sum_{k=0}^n \mathbf{1}_{R_k}$; autrement dit

$$\mathbf{1}_{A_1}(x)\mathbf{1}_{A_2}(y) = \sum_{k=0}^n \mathbf{1}_{A_{k,1}}(x)\, \mathbf{1}_{A_{k,2}}(y) \quad \text{pour tout } (x, y) \in \Omega_1 \times \Omega_2\,.$$

En fixant $x \in \Omega_1$ et en intégrant en $y \in \Omega_2$ par rapport à μ_2, on en déduit

$$\mathbf{1}_{A_1}(x)\, \mu_2(A_2) = \sum_{k=0}^n \mathbf{1}_{A_{k,1}}(x)\, \mu_2(A_{k,2}) \quad \text{pour tout } x \in \Omega_1\,;$$

d'où en intégrant par rapport à μ_1 :

$$\mu_2(A_2)\mu_1(A_1) = \sum_{k=0}^n \mu_2(A_{k,2})\mu_1(A_{k,1})\,,$$

autrement dit $|R| = \sum_0^n |R_k|$.

(ii) Soit $R \in \mathcal{R}$, et soit (R_k) une suite de rectangles mesurables telle que $R \subseteq \bigcup_0^\infty R_k$. On a cette fois $\mathbf{1}_R \leqslant \sum_{k=0}^\infty \mathbf{1}_{R_k}$; autrement dit, en écrivant $R = A_1 \times A_2$ et $R_k = A_{k,1} \times A_{k,2}$:

$$\mathbf{1}_{A_1}(x)\mathbf{1}_{A_2}(y) \leqslant \sum_{k=0}^\infty \mathbf{1}_{A_{k,1}}(x)\,\mathbf{1}_{A_{k,2}}(y) \quad \text{pour tout } (x,y) \in \Omega_1 \times \Omega_2\,.$$

En intégrant en y par rapport à μ_2 et *en utilisant le théorème de convergence monotone* (sous la forme “interversion série-intégrale”), on en déduit

$$\begin{aligned}\mathbf{1}_{A_1}(x)\,\mu_2(A_2) \quad &\leqslant \quad \int_{\Omega_2} \left(\sum_{k=0}^\infty \mathbf{1}_{A_{k,1}}(x)\,\mathbf{1}_{A_{k,2}}(y)\right) d\mu_2(y)\\ &= \quad \sum_{k=0}^\infty \mathbf{1}_{A_{k,1}}(x)\,\mu_2(A_{k,2}) \quad \text{pour tout } x \in \Omega_1\,;\end{aligned}$$

et donc, en intégrant maintenant par rapport à μ_1 :

$$\mu_2(A_2)\mu_1(A_1) \leqslant \sum_{k=0}^\infty \mu_2(A_{k,2})\mu_1(A_{k,1})\,,$$

autrement dit $|R| \leqslant \sum_0^\infty |R_k|$. □

Remarque. On montrerait de la même façon l’existence et l’unicité de la mesure produit de N mesures σ-finies $\mu_1, \dots, \mu_N$, notée $\mu_1 \otimes \cdots \otimes \mu_N$.

Exercice. Montrer que si $(\Omega_1, \mathcal{B}_1)$, $(\Omega_2, \mathcal{B}_2)$, $(\Omega_3, \mathcal{B}_3)$ sont 3 espaces mesurables, alors $(\mathcal{B}_1 \otimes \mathcal{B}_2) \otimes \mathcal{B}_3 = \mathcal{B}_1 \otimes \mathcal{B}_2 \otimes \mathcal{B}_3 = \mathcal{B}_1 \otimes (\mathcal{B}_2 \otimes \mathcal{B}_3)$; et que si μ_1, μ_2, μ_3 sont 3 mesures σ-finies, alors $(\mu_1 \otimes \mu_2) \otimes \mu_3 = \mu_1 \otimes \mu_2 \otimes \mu_3 = \mu_1 \otimes (\mu_2 \otimes \mu_3)$.

4.3. Coupes et théorème de Fubini. Dans cette section, on se donne deux espaces mesurés *σ-finis* $(\Omega_1, \mathcal{B}_1, \mu_1)$ et $(\Omega_2, \mathcal{B}_2, \mu_2)$.

Maintenant qu’on est assuré de l’existence et de l’unicité de la mesure produit $\mu_1 \otimes \mu_2$, on peut essentiellement recopier ce qui a été fait au Chapitre 7 pour obtenir un théorème de Fubini complètement général.

On rappelle que pour tout ensemble $A \subseteq \Omega_1 \times \Omega_2$ on note A_x et A^y les coupes de de A respctivement à l’abscisse $x \in \Omega_1$ et à l’ordonnée $y \in \Omega_2$:

$$\begin{aligned} A_x &:= \{y \in \Omega_2;\ (x,y) \in A\} \quad \text{pour } x \in \Omega_1\,,\\ A^y &:= \{x \in \Omega_1;\ (x,y) \in A\} \quad \text{pour } y \in \Omega_2\,.\end{aligned}$$

Lemme 4.4. *Si $A \subseteq \Omega_1 \times \Omega_2$ appartient à la tribu produit $\mathcal{B}_1 \otimes \mathcal{B}_2$, alors $A_x \in \mathcal{B}_2$ pour tout $x \in \Omega_1$ et $A^y \in \mathcal{B}_1$ pour tout $y \in \Omega_2$.*

Démonstration. Pour $x \in \Omega_1$ fixé, l’application $\Phi_x : \Omega_2 \to \Omega_1 \times \Omega_2$ définie par $\Phi_x(y) := (x,y)$ est mesurable car ses composantes le sont (c’est l’Exercice 4.2). Donc $A_x = \Phi_x^{-1}(A)$ est mesurable, *i.e.* appartient à $\mathcal{B}_2$. On montre de même que $A^y \in \mathcal{B}_1$ pour tout $y \in \Omega_2$. □

Lemme 4.5. *Pour tout $A \in \mathcal{B}_1 \otimes \mathcal{B}_2$, les fonctions $x \mapsto \mu_2(A_x)$ et $y \mapsto \mu_1(A^y)$ sont mesurables, respectivement sur Ω_1 et sur Ω_2.*

Démonstration. La preuve est une adaptation facile de celle du Lemme 1.5 : il suffit de remplacer partout “borélien(ne)” par “mesurable” et $\mathcal{B}(\mathbb{R}^p \times \mathbb{R}^q)$ par $\mathcal{B}_1 \otimes \mathcal{B}_2$. Les détails sont laissés en exercice. □

Dans le cadre “abstrait”, le **principe de Cavalieri** s'énonce comme suit. La preuve est identique à celle donnée au Chapitre 7 dans le cas $\mathbb{R}^p \times \mathbb{R}^q$.

PROPOSITION 4.6. *Pour tout ensemble mesurable $A \subseteq \Omega_1 \times \Omega_2$, on a*

$$(\mu_1 \otimes \mu_2)(A) = \int_{\Omega_1} \mu_2(A_x)\, d\mu_1(x) = \int_{\Omega_2} \mu_1(A^y)\, d\mu_2(y).$$

Comme dans le cas de $\mathbb{R}^N$, on en déduit que “l'intégrale, c'est le volume sous le graphe”.

COROLLAIRE 4.7. *Soit $(\Omega, \mathcal{B}, \mu)$ est un espace mesuré sigma-fini. Pour toute fonction mesurable $f : \Omega \to [0, \infty]$, on a*

$$\int_\Omega f\, d\mu = (\mu \otimes \lambda_1)\big(\mathrm{SG}(f, \Omega)\big).$$

Démonstration. Par Cavalieri, on a

$$(\mu \otimes \lambda_1)\big(\mathrm{SG}(f, \Omega)\big) = \int_\Omega \lambda_1\Big(\mathrm{SG}(f, \Omega)_x\Big) d\mu(x) = \int_\Omega f(x)\, d\mu(x).$$

□

Remarque. Si on ne suppose pas que les mesures μ_1 et μ_2 sont σ-finies, alors les deux intégrales itérées apparaissant dans le principe de Cavalieri peuvent très bien avoir un sens avec cependant $\int_{\Omega_1} \mu_2(A_x)\, d\mu_1(x) \neq \int_{\Omega_2} \mu_1(A^y)\, d\mu_2(y)$. Prenons par exemple $\Omega_1 = \mathbb{R} = \Omega_2$ avec $\mathcal{B}_1 := \mathcal{B}(\mathbb{R}) = \mathcal{B}_2$ et les mesures $\mu_1 := \lambda_1$ (la mesure de Lebesgue) et $\mu_2 := \mu_c$ (la mesure de comptage, qu'on restreint à $\mathcal{B}(\mathbb{R})$). Soit

$$\Delta := \{(x, y) \in \mathbb{R} \times \mathbb{R};\ x = y\}$$

la “diagonale” de $\mathbb{R} \times \mathbb{R}$. Comme Δ est fermée dans $\mathbb{R} \times \mathbb{R}$, c'est un ensemble borélien. Pour tous $x, y \in \mathbb{R}$, on a $\Delta_x = \{x\}$ et $\Delta^y = \{y\}$. Donc $\mu_2(\Delta_x) = \mu_c(\{x\}) = 1$ pour tout $x \in \mathbb{R}$ et $\mu_1(\Delta^y) = \lambda_1(\{y\}) = 0$ pour tout $y \in \mathbb{R}$. Ainsi,

$$\int_\mathbb{R} \mu_2(\Delta_x)\, d\mu_1(x) = \int_\mathbb{R} \mathbf{1}\, d\lambda_1 = \infty \qquad \text{et} \qquad \int_\mathbb{R} \mu_1(\Delta^y)\, d\mu_2(y) = \int_\mathbb{R} 0\, d\mu_c = 0\,.$$

Voici pour finir le **théorème de Fubini**, dont la preuve est à nouveau identique à celle donnée au Chapitre 7 dans le cas $\mathbb{R}^p \times \mathbb{R}^q$.

THÉORÈME 4.8. *Soit f une fonction mesurable sur $\Omega_1 \times \Omega_2$, positive ou à valeurs complexes.*

(1) *Si f est* positive*, alors*

$$\begin{aligned}\int_{\Omega_1 \times \Omega_2} f\, d(\mu_1 \otimes \mu_2) &= \int_{\Omega_1} \left(\int_{\Omega_2} f(x, y)\, d\mu_2(y) \right) d\mu_1(x) \\ &= \int_{\Omega_2} \left(\int_{\Omega_1} f(x, y)\, d\mu_1(x) \right) d\mu_2(y)\,.\end{aligned}$$

(2) *Si f est* intégrable *par rapport à $\mu_1 \otimes \mu_2$, alors* on a le droit d'écrire

$$\begin{aligned}\int_{\Omega_1 \times \Omega_2} f\, d(\mu_1 \otimes \mu_2) &= \int_{\Omega_1} \left(\int_{\Omega_2} f(x, y)\, d\mu_2(y) \right) d\mu_1(x) \\ &= \int_{\Omega_2} \left(\int_{\Omega_1} f(x, y)\, d\mu_1(x) \right) d\mu_2(y)\,.\end{aligned}$$

Démonstration. Il suffit de traiter le cas où $f \geqslant 0$. Pour cela, *en notant* λ la mesure de Lebesgue sur $\mathbb{R}$, on part de l'identité

$$\begin{aligned}\int_{\Omega_1 \times \Omega_2} f \, d(\mu_1 \otimes \mu_2) &= (\mu_1 \otimes \mu_2 \otimes \lambda)\big(\mathrm{SG}(f, \Omega_1 \times \Omega_2)\big) \\ &= \int_{\Omega_1} (\mu_2 \otimes \lambda)\Big(\mathrm{SG}(f, \Omega_1 \times \Omega_2)_x\Big)\, d\mu_1(x),\end{aligned}$$

et on conclut comme au Chapitre 7. □

Remarque. Énoncé sous cette forme, le théorème de Fubini contient à la fois le théorème de Fubini "classique" dans $\mathbb{R}^p \times \mathbb{R}^q$ et les théorèmes d'interversion série-intégrale (à condition de supposer qu'on a affaire à un espace mesuré σ-fini).

Exercice 1. Soit $(\Omega, \mathcal{B}, \mu)$ un espace mesuré σ-fini. Montrer que pour toute fonction mesurable $f : \Omega \to \mathbb{R}^+$ et pour toute fonction $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ continue sur $\mathbb{R}^+$, de classe $\mathcal{C}^1$ sur $]0, \infty[$ et vérifiant $\phi(0) = 0$, on a

$$\int_\Omega \phi(f(x))\, d\mu(x) = \int_0^\infty \phi'(t)\, \mu\big(\{f > t\}\big)\, dt\,.$$

Exercice 2. Montrer que l'hypothèse de σ-finitude est en fait inutile dans l'exercice précédent. (Considérer d'abord le cas d'une fonction f étagée.)

Exercice 3. Énoncer et démontrer une version générale de la forme intégrale de l'inégalité de Minkowski