

Chapitre 3

Intégration de Lebesgue

Dans ce chapitre, nous allons nous intéresser à la notion d'intégrale dans sa forme moderne. Après une brève Introduction historique sur ce sujet en Section 3.1, nous définirons les outils élémentaires de la théorie de la mesure en Section 3.2, et nous les utiliserons en Section 3.3 pour construire *l'intégrale de Lebesgue* et étudier quelques unes de ses propriétés élémentaires. Nous verrons ensuite quelques résultats portant sur les intégrales multiples en Section 3.4, et nous finirons par une introduction aux espaces L^p en Section 3.5. Ces derniers sont d'une importance cruciale dans l'étude des équations (notamment aux dérivées partielles) issues de la physique.

3.1 Éléments historiques et intégrale de Riemann

Cette première partie est dédiée à une introduction intuitivo-historique de la notion d'intégrale et présente des éléments motivant la construction de l'intégrale de Lebesgue que nous détaillons dans les sections suivantes.

La notion d'intégration fait partie de ces objets un peu diffus des mathématiques qui ont connu un nombre important d'évolutions conceptuelles au cours du temps, et desquels il est subséquentement difficile de dater précisément l'apparition. Bien que l'idée de calculer les aires d'objets non polygonaux remonte à la Grèce antique, il semblerait que ce soit Leibniz (1646-1706) qui le premier fit apparaître cet outil de manière formelle en conjonction avec son calcul infinitésimal et l'introduction de la notion de dérivée. Cependant, on peut raisonnablement attribuer le premier formalisme intégral exhaustif – celui que certains d'entre vous ont sûrement déjà vu au lycée ou en classes préparatoires – au mathématicien allemand Bernhard Riemann (1826-1866). Sa construction relativement générale, repose sur l'idée d'approcher le graphe d'une fonction donnée par des rectangles de plus en plus fins.

Plus précisément, soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction “suffisamment régulière”¹ avec a et b deux réels. Pour n’importe quelle subdivision $\sigma = \{x_0, \dots, x_N\}$ de $[a, b]$ en N segments, on définit la somme élémentaire de Riemann $S_\sigma(f)$ associée à f par

$$S_\sigma(f) = \sum_{i=0}^{N-1} (x_{i+1} - x_i) f(x_i).$$

Cette quantité représente une approximation de l’aire sous le graphe de la fonction f par la somme des aires des rectangles de hauteur $f(x_i)$ construits à l’aide de la subdivision σ . Riemann définit alors l’intégrale de f entre a et b comme étant la limite lorsque le pas de subdivision tend vers 0 de $S_\sigma(f)$, et la note

$$\lim_{|\sigma| \downarrow 0} [S_\sigma(f)] = \int_a^b f(x) dx,$$

où $|\sigma| := \max_{1 \leq i \leq N-1} \{x_{i+1} - x_i\}$

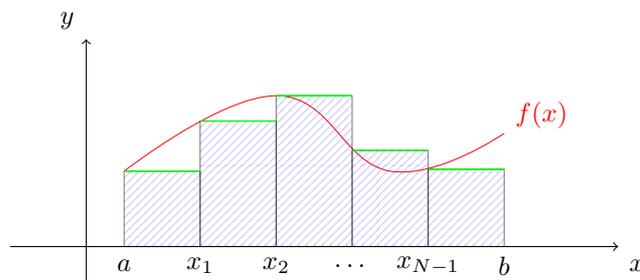


FIGURE 3.1 – *Illustration de la construction de l’intégrale de Riemann*

Cette définition mêle l’idée déjà ancienne d’approximation d’une courbe régulière par des polygones avec l’intuition du calcul infinitésimal à *la Leibniz*, où “ dx ” représente un petit élément de longueur par lequel on multiplie $f(x)$ pour obtenir une aire. Cependant, aussi esthétique soit-elle, cette construction souffre de certains défauts que nous allons survoler.

Limites de l’intégrale de Riemann

1) L’intégrale de Riemann se base sur l’idée d’approcher l’aire sous le graphe d’une fonction par des rectangles. Cependant, cette condition est assez restrictive et beaucoup de fonctions que l’on aimerait pouvoir intégrer ne possèdent pas des graphes assez réguliers pour que cette estimation soit valide.

1. Ici, suffisamment régulière signifie que l’on peut approximer raisonnablement le graphe de cette fonction par des rectangles.

2) La construction que nous avons ébauchée en dimension 1 et illustrée en Figure 3.1 s'appuie intégralement sur la structure linéaire de \mathbb{R} . Elle peut s'étendre à \mathbb{R}^d pour n'importe quel $d \geq 1$, mais il devient plus difficile de lui donner un sens rigoureux sur des espaces courbes (comme des sphères, des cylindres, etc...) ou, encore pire, de dimension infinie.

3) Enfin, l'intégrale de Riemann ne permet pas d'élaborer une bonne théorie d'interversion des opérations de limite et d'intégration. Il est en effet possible de construire des suites de fonction croissantes $(f_n)_{n \in \mathbb{N}}$, convergeant ponctuellement vers une fonction f , et dont la suite des intégrales au sens de Riemann convergent vers un réel différent de l'intégrale de f !

Ces "trous" présents dans la théorie de Riemann proviennent du fait que l'élément infinitésimal " dx " ne permet de compter que des longueurs de segments (ou des aires / volumes de pavés en dimensions supérieures), et qu'il ne définit donc pas une notion assez riche et flexible de ce qu'est la *taille d'un ensemble*. Les mathématiciens Henry Léon Lebesgue (1875-1941) et Émile Borel (1871-1956) comblèrent ce vide conceptuel en définissant une notion complète et rigoureuse de ce qu'est la *mesure d'un ensemble* et que nous allons explorer en Section 3.2. Nous verrons ensuite en Section 3.3 comment construire, en nous appuyant sur ces premiers résultats, une intégrale robuste et possédant nombre de bonnes propriétés qui porte maintenant le nom *d'intégrale de Lebesgue*.

3.2 Tribus et mesures

Dans cette partie, nous allons établir les définitions d'*ensembles mesurables* – les ensembles dont il sera possible de mesurer la taille – et de *mesures* sur ces ensembles. La "taille" de certains ensembles semble en effet difficile à définir, du moins en utilisant la notion infinitésimale de " dx ". Par exemple, quelle est la longueur² de l'ensemble $\mathbb{Q} \cap [0, 1]$? La définition de Riemann ne permet pas de répondre à cette question (le vérifier!), alors que la théorie de Lebesgue le permettra.

Nous allons commencer par introduire la première pierre de l'édifice qu'est la théorie de la mesure : la définition d'une *tribu*.

Définition 1 (Tribu ou σ -algèbre). Soit Ω un ensemble quelconque. On note $\mathcal{P}(\Omega)$ l'ensemble des parties de Ω . Une famille d'ensembles $\mathcal{A} \subset \mathcal{P}(\Omega)$ est appelée une **tribu**, ou **σ -algèbre**, si elle vérifie les trois propriétés suivantes :

2. Il est à noter que cette question n'est pas qu'une technicité gratuite de mathématiciens comme on aura l'occasion de le voir en section 3.5.

- (i) $\emptyset \in \mathcal{A}$ (contient l'ensemble vide).
 - (ii) Si $A \in \mathcal{A}$ alors $A^c \in \mathcal{A}$ (stabilité par complémentarité).
 - (iii) Si $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$ alors $\bigcup_n A_n \in \mathcal{A}$ (stabilité par union dénombrable).
- Le couple (Ω, \mathcal{A}) est appelé un **espace mesurable**.

Intuitivement, les tribus sont les familles d'ensembles qui sont laissées invariantes par les opérations élémentaires d'union et de complémentarité. **Il s'agit des familles d'ensembles sur lesquels on va définir les mesures.**

Exercices :

1. Soit \mathcal{A} une tribu de Ω . Montrer que $\Omega \in \mathcal{A}$ et que \mathcal{A} est stable par intersection dénombrable.
2. Montrer que les ensembles suivants sont des tribus : $\{\emptyset, \Omega\}$, $\{\emptyset, \Omega, A, A^c\}$ quel que soit $A \subset \Omega$.
3. Montrer que si \mathcal{A} et \mathcal{B} sont des tribus sur Ω , alors $\mathcal{A} \cap \mathcal{B}$ est aussi une tribu.

Certaines tribus sont bien entendu plus intéressantes que d'autres. La tribu la plus grossière $\{\emptyset, \Omega\}$ n'est pas fascinante puisqu'elle ne définit comme ensembles mesurables que l'ensemble vide \emptyset et Ω tout entier. Elle sert cependant de brique élémentaire pour construire des tribus plus intéressantes. Étant donné par exemple un ensemble $A \subset \Omega$, on peut construire $\{\emptyset, \Omega, A, A^c\}$ qui est la *plus petite tribu* (au sens de l'inclusion) contenant A . Cette idée se généralise à n'importe quelle famille d'ensembles de Ω : c'est la notion de *tribu engendrée*.

Définition 2 (Tribu engendrée). Soient Ω un ensemble et $E \subset \mathcal{P}(\Omega)$ un collection d'ensembles quelconque. On appelle **tribu engendrée** par E la plus petite tribu contenant E . On note cette dernière $\sigma(E)$.

Exercice. Montrer que $\sigma(E)$ existe quel que soit $E \subset \mathcal{P}(\Omega)$.

Lorsque l'on veut faire de l'analyse – par exemple étudier des fonctions –, les familles d'ensembles qui sont sans conteste les plus importantes sont les *topologies*, i.e. les *familles d'ouverts* de notre espace. Ce sont ces ensembles qui permettent de définir les notions de voisinage, de continuité, de convergence, de compacité, etc.... Nous rappelons ci-dessous la définition générale d'une topologie.

Définition 3 (Espace topologique). Soit Ω un ensemble. Une famille d'ensembles $\mathcal{T} \subset \mathcal{P}(\Omega)$ est appelée une **topologie** si elle vérifie les trois propriétés suivantes :

- (i) $\Omega \in \mathcal{T}$ (contient l'espace entier).
- (ii) Si $(\mathcal{O}_i)_{i \in I} \subset \mathcal{T}$ alors $\bigcup_i \mathcal{O}_i \in \mathcal{T}$ (stabilité par union quelconque).

(iii) Si $(\mathcal{O}_i)_{i=1}^n \subset \mathcal{T}$ alors $\bigcap_{i=1}^n \mathcal{O}_i \in \mathcal{T}$ (stabilité par intersection finie).

Les éléments de \mathcal{T} sont appelés des **ouverts**, et on dit que le couple (Ω, \mathcal{T}) est un **espace topologique**.

Nous travaillerons essentiellement dans des espaces topologiques tels que la droite réelle $(\mathbb{R}, \mathcal{O})$ munie de sa topologie naturelle \mathcal{O} , i.e. la topologie générée par la convergence associée à la norme Euclidienne. Il est alors raisonnable de considérer la tribu engendrée par les ouverts de notre espace. On appelle cette dernière la *tribu Borélienne*.

Définition 4 (Tribu Borélienne). Soit (Ω, \mathcal{T}) un espace topologique. La **tribu Borélienne** $\mathcal{B}(\Omega)$ sur Ω est la tribu engendrée par les ouverts, i.e. $\mathcal{B}(\Omega) = \sigma(\mathcal{T})$.

Il est bon de remarquer que par construction, la tribu Borélienne contient tous les ensembles ouverts et fermés de Ω , mais aussi toutes les unions dénombrables possibles d'ouverts et de fermés (qui ne sont potentiellement ni l'un ni l'autre).

Les deux résultats suivants, dont la preuve est laissée en exercice, offrent des éléments de réponse aux problèmes soulevés dans l'Introduction.

Exercices :

1. Montrer que la tribu Borélienne $\mathcal{B}(\mathbb{R})$ de \mathbb{R} muni de sa topologie usuelle \mathcal{O} (voir plus haut) est engendrée par les intervalles rationnels, i.e. $\mathcal{B}(\mathbb{R}) = \sigma(\{]a, b[\text{ t.q. } a, b \in \mathbb{Q}\})$.
2. Montrer que $[0, 1] \cap \mathbb{Q} \in \mathcal{B}(\mathbb{R})$.

Il est à noter que le premier point est extrêmement important, puisqu'il donne une caractérisation assez explicite des Boréliens sur \mathbb{R} qui ne se base que sur l'ensemble \mathbb{Q} qui a la bonne propriété d'être *dénombrable*.

Il est possible de montrer (mais c'est difficile) que $\mathcal{P}(\mathbb{R}) \neq \mathcal{B}(\mathbb{R})$, i.e. : **il existe des parties de \mathbb{R} qui ne sont pas mesurables**. Il existe une très belle illustration (que je vous encourage à aller voir !) de ce fait proposée en 1924 par Stefan Banach (1892-1945) et Alfred Tarski (1901-1983), connue sous le nom de *paradoxe de Banach-Tarski*.

Maintenant que nous avons défini les familles d'ensembles dont on est capable de mesurer la taille, introduisons la notion de *mesure* à proprement parler.

Définition 5 (Mesure). Soit (Ω, \mathcal{A}) un espace mesurable. On appelle **mesure positive** sur (Ω, \mathcal{A}) une application $\mu : \mathcal{A} \rightarrow [0, +\infty]$ telle que :

- (i) $\mu(\emptyset) = 0$.
- (ii) Pour tout $A, B \in \mathcal{A}$ tels que $A \cap B = \emptyset$ on a $\mu(A \cup B) = \mu(A) + \mu(B)$.

(iii) Pour toute suite croissante $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$, on a $\lim_{n \rightarrow \infty} \mu(A_n) = \mu(\bigcup_n A_n)$.
Le triplet $(\Omega, \mathcal{A}, \mu)$ est appelé un **espace mesuré**.

Exercices :

1. Montrer que pour tout $A, B \in \mathcal{A}$ on a $\mu(A \cup B) \leq \mu(A) + \mu(B)$.
2. Montrer que pour $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$, on a³

$$\mu\left(\bigcup_n A_n\right) \leq \sum_n \mu(A_n).$$

3. Montrer que pour $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$ tels que les A_n sont deux à deux disjoints,

$$\mu\left(\bigcup_n A_n\right) = \sum_n \mu(A_n).$$

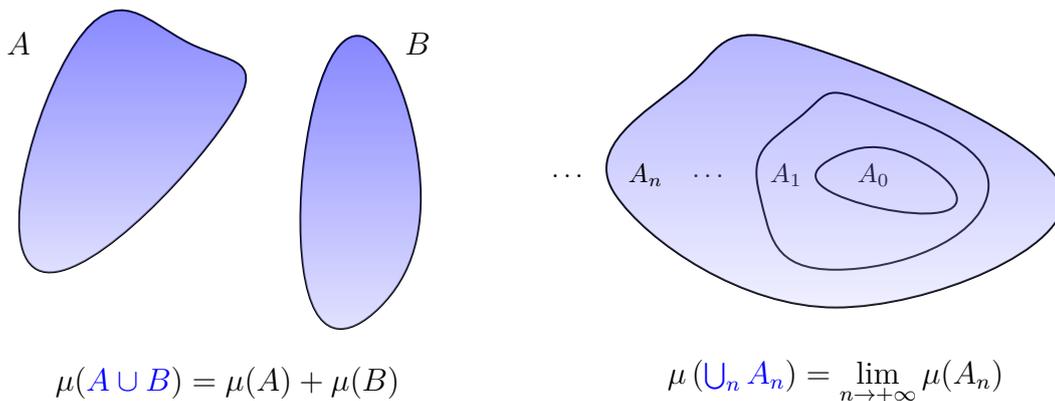


FIGURE 3.2 – Illustration des point (ii) et (iii) de la Définition 5

Nous allons maintenant donner quelques exemples de mesures que l'on rencontre régulièrement en analyse ou en théorie des probabilités.

Propriété 1 (Exemples de mesures). Les applications suivantes sont des mesures sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$.

1. Mesure de cardinalité $\mu : A \mapsto \text{card}(A)$.
2. Mesure de Dirac $\delta_a : A \mapsto 1$ si $a \in A$ et 0 sinon.
3. Mesure de comptage $\mu : A \mapsto \text{card}(A \cap \mathbb{N}^d)$.

3. Cette propriété est parfois utilisée à la place de (iii) dans la définition d'une mesure.

Remarque (Lien avec la théorie des probabilités). On appelle usuellement la mesure de tout l'espace $\mu(\Omega)$ la **masse** de μ . Une mesure de masse 1 est une mesure de probabilité (ou simplement une probabilité).

Cette terminologie vient du fait que toute mesure μ de masse totale égale à 1 sur Ω peut-être vue comme la loi $\mu = \mathcal{L}(X)$ d'une variable aléatoire X . Autrement dit, on a avec ces notation

$$\mathbb{P}(X \in A) = \mu(A),$$

pour toute variable aléatoire X de loi μ définie sur Ω .

Une mesure très importante dans le formalisme que nous venons d'introduire est celle qui correspond à la notion de longueur sur \mathbb{R} , de surface sur \mathbb{R}^2 , de volume sur \mathbb{R}^3 , etc... Il s'agit en somme de celle qui généralise la notion intuitive d'intégrale contre un volume que Riemann et Leibniz ont d'abord formalisé avec l'élément infinitésimal "dx". Il s'agit de la *mesure de Lebesgue*.

Théorème 6 (Mesure de Lebesgue). Il existe une unique mesure que l'on note \mathcal{L}^d définie sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ qui soit **positive, localement finie**⁴, **invariante par translation**⁵, et telle que pour toute collection $a_i, b_i \in \mathbb{R}$ avec $i \in \{1, \dots, d\}$, l'on ait

$$\mathcal{L}^d([a_1, b_1] \times \dots \times [a_d, b_d]) = \prod_{i=1}^d |b_i - a_i|.$$

Cette mesure est appelée la **mesure de Lebesgue**.

Remarque (Particularité des espaces \mathbb{R}^d et mesure de Lebesgue). La mesure de Lebesgue telle qu'elle est décrite ci-dessus est un objet fondamentalement dépendant de la structure linéaire et "plate" (sans courbure) des espaces \mathbb{R}^d . Bien qu'elle possède des analogues conceptuels (comme la mesure du volume) sur des espaces courbes de dimension finie, ces derniers deviennent rapidement très compliqués à construire dès que l'on sort du cas des variétés Riemanniennes. Qui plus est, il a été démontré qu'il n'existe pas de mesure analogue à la mesure de Lebesgue dans le cas où l'espace ambiant est de dimension infinie.

Nous allons terminer cette première partie du cours par l'introduction des notions d'ensembles négligeables et de μ -presque partout.

4. Heuristiquement, une mesure μ sur un espace mesurable (Ω, \mathcal{A}) est dite localement finie si pour tout point $x \in \Omega$, il existe un voisinage \mathcal{N}_x de x tel que $\mu(\mathcal{N}_x) < +\infty$.

5. Une mesure μ est invariante par translation si $\mu(A + v) = \mu(A)$ pour tout ensemble $A \in \mathcal{B}(\mathbb{R}^d)$ et vecteur $v \in \mathbb{R}$ où l'on rappelle que $A + v = \{a + v \text{ t.q. } a \in A\}$.

Définition 7 (Ensembles négligeables). Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré. Un ensemble $\mathcal{N} \in \mathcal{A}$ est dit **négligeable** par rapport à la mesure μ si $\mu(\mathcal{N}) = 0$. En l'absence d'ambiguïté par rapport à la mesure que l'on considère, on dira simplement que \mathcal{N} est un ensemble négligeable.

Propriété 2 (Mesure de Lebesgue des nombres rationnels). L'ensemble $\mathbb{Q} \cap [0, 1]$ des nombres rationnels compris entre 0 et 1 est négligeable par rapport à la mesure de Lebesgue \mathcal{L}^1 sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, i.e. $\mathcal{L}^1(\mathbb{Q} \cap [0, 1]) = 0$.

Exercice. Prouver la proposition précédente avec l'aide de l'indice suivant : comme $\mathbb{Q} \cap [0, 1]$ est un ensemble dénombrable, on peut l'écrire sous la forme d'une énumération $(q_n)_{n \in \mathbb{N}}$ de ses éléments. Conclure en comparant la mesure de Lebesgue de $\mathbb{Q} \cap [0, 1]$ avec celle de l'ensemble $\bigcup_n B(q_n, \epsilon 2^{-n})$ pour tout $\epsilon > 0$ où $B(x, r) =]x - r, x + r[$ désigne la boule centrée en x de rayon r .

Génériquement, une mesure μ sur un espace mesurable (Ω, \mathcal{A}) va admettre des ensembles négligeables. Ainsi, lorsque l'on va raisonner sur des objets (comme par exemple des fonctions) avec des outils construits à l'aide de mesures tels que les intégrales, on va naturellement pouvoir "négliger" ce qu'il advient de ces objets sur les ensembles de mesure nulle. Il s'agit de la notion de μ -presque partout que nous introduisons ci-dessous dans les cas particuliers de l'égalité et de la convergence.

Définition 8 (Égalité et convergence μ -presque partout). Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré. On dit que deux fonctions f et g définies sur Ω sont **égales μ -presque partout** (μ -p.p. en abrégé) si

$$\mu(\{x \in \Omega \text{ t.q. } f(x) \neq g(x)\}) = 0.$$

De la même façon, on dit qu'une suite de fonctions $(f_n)_{n \in \mathbb{N}}$ **converge μ -presque partout** vers une fonction f si

$$\mu(\{x \in \Omega \text{ t.q. } f_n(x) \not\rightarrow f(x)\}^c) = 0.$$

Nous sommes maintenant parés pour introduire et étudier l'intégrale au sens de Lebesgue ainsi que ses nombreuses propriétés.

3.3 Fonctions mesurables et intégrale de Lebesgue

Cette partie est dédiée à l'introduction et à l'étude des fonctions dites *mesurables*, ainsi qu'à la définition de l'intégrale de Lebesgue. Comme nous l'avons informellement esquissé dans l'introduction, la construction d'une intégrale est

forcément reliée à une notion d'intégrabilité, ou en d'autres termes de régularité suffisante pour être intégré. Dans le cas de l'intégrale de Riemann, il fallait que les fonctions aient des graphes appochables par des rectangles. Outre son caractère un peu "manuel", cette condition est en fait très restrictive et force les fonctions correspondantes à être continues par morceaux.

Dans le cadre de la théorie de l'intégrale de Lebesgue, nous allons voir que cette condition va se formuler de manière très naturelle à l'aide des outils que nous avons introduits en Section 3.2. Nous présentons ci-dessous la classe des fonctions dites *mesurables*, qui seront exactement les fonctions présentant suffisamment de régularité pour être intégrées au sens de Lebesgue.

Définition 9 (Fonction mesurable). *Une fonction $f : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ est dite **mesurable** si pour tout $B \in \mathcal{B}(\mathbb{R})$ on a $f^{-1}(B) \in \mathcal{A}$.*

Cette définition, probablement assez peu intuitive et effrayante à première vue, devient plus compréhensible si on la compare à la notion plus familière de *continuité* que nous rappellons ci-dessous.

Rappels (Définition générale de la continuité). *Soit (Ω, \mathcal{T}) un espace topologique et $(\mathbb{R}, \mathcal{O})$ la droite réelle munie de sa topologie usuelle. On dit qu'une fonction $f : \Omega \mapsto \mathbb{R}$ est **continue** si pour tout ouvert $O \in \mathcal{O}$ on a $f^{-1}(O) \in \mathcal{T}$. Autrement dit, une fonction est continue si et seulement si elle envoie la topologie d'arrivée sur celle de départ par image réciproque⁶.*

On peut donc voir la mesurabilité comme une sorte de relaxation de la continuité. Il s'agit aussi d'une notion de régularité – puisque l'on demande à f de stabiliser des ensembles par image réciproque –, mais où l'on impose la condition plus faible de transformer un ensemble Borélien en un ensemble mesurable.

Avant d'aller plus loin, nous allons introduire les concepts d'*infimum*, *supremum*, *limite inférieure* et *limite supérieure* qui s'avèreront utiles par la suite.

Rappels (Infimum et supremum). *Soit $(a_n)_{n \in \mathbb{N}} \subset \mathbb{R}$. On définit l'**infimum** et le **supremum** de la suite $(a_n)_{n \in \mathbb{N}}$, que l'on note respectivement $\inf_n a_n$ et $\sup_n a_n$, comme suit*

- ◇ $L = \inf_n a_n$ si et seulement si $L \leq a_n$ pour tout $n \in \mathbb{N}$ et si pour tout $\epsilon > 0$, il existe $N_\epsilon \in \mathbb{N}$ tel que $a_{N_\epsilon} \leq L + \epsilon$.
- ◇ $L = \sup_n a_n$ si et seulement si $L \geq a_n$ pour tout $n \in \mathbb{N}$ et si pour tout $\epsilon > 0$, il existe $N_\epsilon \in \mathbb{N}$ tel que $a_{N_\epsilon} \geq L - \epsilon$.

6. La notion de continuité dépend donc des topologies que l'on considère, de même que la notion de mesurabilité dépend des tribus mises en jeu.

Soit $(f_n)_{n \in \mathbb{N}}$ est une suite de fonctions à valeurs réelles définie sur un espace topologique Ω . On définit subséquentement les fonctions $\inf_n f_n$ et $\sup_n f_n$ par

$$\inf_n f_n : x \mapsto \inf_n f_n(x) \quad , \quad \sup_n f_n : x \mapsto \sup_n f_n(x).$$

Définition 10 (Limites inférieures et supérieures). Soit $(a_n)_{n \in \mathbb{N}} \subset \mathbb{R}$. On définit la **limite inférieure** de la suite $(a_n)_{n \in \mathbb{N}}$, dénotée $\liminf_{n \rightarrow +\infty} [a_n]$, par

$$\liminf_{n \rightarrow +\infty} [a_n] = \lim_{n \rightarrow +\infty} [\inf_{k \geq n} a_k].$$

Symétriquement, on définit la **limite supérieure** de la suite $(a_n)_{n \in \mathbb{N}}$ par

$$\limsup_{n \rightarrow +\infty} [a_n] = \lim_{n \rightarrow +\infty} [\sup_{k \geq n} a_k].$$

On vérifie aisément que $\liminf_n [a_n] > -\infty$ (respectivement $\limsup_n [a_n] < +\infty$) si et seulement si $(a_n)_{n \in \mathbb{N}}$ est minorée (respectivement majorée) par une constante. Par ailleurs, on a par définition que

$$\liminf_{n \rightarrow +\infty} [a_n] \leq \limsup_{n \rightarrow +\infty} [a_n],$$

avec égalité si et seulement si la suite $(a_n)_{n \in \mathbb{N}}$ admet une unique limite $a \in \mathbb{R}$.

Heuristiquement, les limites inférieures et supérieures sont égales à la plus petite et à la plus grande *valeur d'adhérence* - i.e. limites de sous-suites - d'une suite donnée. Nous illustrons cette idée dans l'exemple ci-dessous.

Exemple (Limites inférieure et supérieure de suites oscillantes).

Soit $(a_n)_{n \in \mathbb{N}} \subset \mathbb{R}$ la suite de terme général $a_n = \sin\left(\frac{n\pi}{2}\right) e^{-n}$. Pour tout $n \in \mathbb{N}$, on peut voir que

$$\begin{cases} \sup_{k \geq n} [a_k] = a_{p(n)} & \text{où } p(n) \in \{n, \dots, n+7\} \text{ est tel que } p(n) \equiv 2 \pmod{8}, \\ \inf_{k \geq n} [a_k] = a_{q(n)} & \text{où } q(n) \in \{n, \dots, n+7\} \text{ est tel que } q(n) \equiv 6 \pmod{8}, \end{cases}$$

comme nous l'illustrons en Figure 3.3. Étant donné que ces deux suites tendent vers 0, on retrouve bien le fait que les *liminf* et *limsup* sont égales à la limite de la suite, qui est elle-même 0.

Si maintenant l'on considère la suite $(b_n)_{n \in \mathbb{N}} \subset \mathbb{R}$ de terme général $b_n = \sin\left(\frac{n\pi}{2}\right)$, on a de manière assez évidente que

$$\sup_{k \geq n} b_k = 1 \quad , \quad \inf_{k \geq n} b_k = -1$$

pour tout $n \in \mathbb{N}$. Ainsi, $\liminf_n [b_n] = -1$ alors que $\limsup_n b_n = 1$: la suite ne converge pas, mais on retrouve bien sa plus petite et sa plus grande valeur d'adhérence $\{\pm 1\}$.

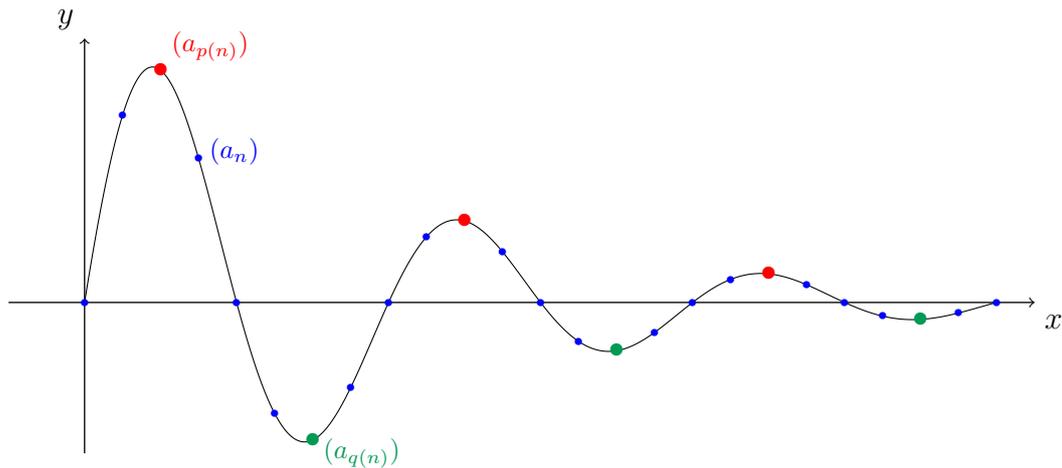


FIGURE 3.3 – Illustration des suites $(a_n)_{n \in \mathbb{N}}$, $(a_{p(n)})_{n \in \mathbb{N}}$ et $(a_{q(n)})_{n \in \mathbb{N}}$

Les notions de \liminf et \limsup sont très utiles car elles définissent des quantités finies dès lors que la suite correspondante est bornée. De plus, elles permettent de caractériser la convergence d'une suite vers une valeur unique lorsqu'elles sont égales.

Les opérations élémentaires d'infimum, supremum, \liminf et \limsup refont régulièrement surface dans la construction de l'intégrale de Lebesgue et dans l'énoncé de ses diverses propriétés. Ces dernières possèdent la propriété important de **préserver la mesurabilité** d'une suite de fonction comme nous l'illustrons dans le théorème ci-dessous.

Théorème 11 (Stabilité des fonctions mesurables) :

Soit (Ω, \mathcal{A}) un espace mesurable. Alors :

1. Si Ω est un espace topologique et $\mathcal{A} = \mathcal{B}(\Omega)$, toute fonction $f : \Omega \rightarrow \mathbb{R}$ continue sur Ω est mesurable.
2. Si f et g sont deux fonctions mesurables et $a \in \mathbb{R}$, alors les fonctions $f + g$, af , fg et f/g (si $g \neq 0$) sont mesurables.
3. Si $(f_n)_{n \in \mathbb{N}}$ est une suite de fonctions mesurables, alors les fonctions $\sup_n f_n$, $\inf_n f_n$, $\liminf_n f_n$ et $\limsup_n f_n$ sont mesurables.
4. La fonction indicatrice d'un ensemble $A \subset \Omega$ définie par

$$\mathbb{1}_A : \omega \mapsto \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{sinon} \end{cases}$$

est mesurable si et seulement si $A \in \mathcal{A}$.

La preuve de ce théorème se fait bien à la main, en vérifiant que pour chaque fonction indiquée ci-dessus, l'image réciproque d'un Borélien par cette dernière appartient bien à la tribu \mathcal{A} . Nous introduisons dans la prochaine définition une famille de fonctions qui vont s'avérer être cruciales dans la construction de l'intégrale de Lebesgue : les *fonctions étagées*.

Définition 12 (Fonctions étagées). Soient $a_1, \dots, a_n \in \mathbb{R}$ et $A_1, \dots, A_n \in \mathcal{A}$. Par le théorème précédent, la fonction

$$x \in \Omega \mapsto \sum_{i=1}^n a_i \mathbb{1}_{A_i}(x)$$

est mesurable. On appelle **fonction étagée** toute fonctions mesurable pouvant s'écrire sous cette forme.

De même que l'on avait commencé par définir l'intégrale de Riemann pour des fonctions constantes par morceaux sur des segments, nous allons maintenant définir l'intégrale de Lebesgue d'une fonction étagée.

Définition 13 (Intégrale de Lebesgue de fonctions étagées).

Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré et $f = \sum_{i=1}^n a_i \mathbb{1}_{A_i}$ une fonction étagée sur Ω . L'intégrale de Lebesgue de f par rapport à μ est définie par

$$\int f(x) d\mu(x) \equiv \int f d\mu = \sum_{i=1}^n a_i \mu(A_i).$$

Il est important de remarquer que l'intuition derrière cette construction ressemble à celle ébauchée dans l'introduction pour l'intégrale de Riemann. La différence centrale entre ces deux approches se traduit par le fait que l'on se laisse la possibilité de considérer des ensembles mesurables et des mesures quelconques, et pas uniquement des segments et l'infinitésimal " dx ".

Exercice. L'écriture d'une fonction étagée n'est pas unique. Montrer que cette définition ne dépend pas du choix de l'écriture. Autrement dit, montrer que si

$$f = \sum_{i=1}^n a_i \mathbb{1}_{A_i} = \sum_{i=1}^n b_i \mathbb{1}_{B_i},$$

alors

$$\int f d\mu = \sum_{i=1}^n a_i \mu(A_i) = \sum_{i=1}^n b_i \mu(B_i).$$

Ce qui va nous permettre de passer de cette première définition simple de l'intégrale de Lebesgue pour des fonctions étagées à sa définition pour des fonctions quelconque est un théorème d'approximation d'une fonction mesurable par des fonctions étagées que nous énonçons ci-dessous.

Théorème 14 (Approximation des fonctions mesurables positives). Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré. Pour toute fonction mesurable et positive $f : \Omega \rightarrow \mathbb{R}_+$, il existe deux suites $(a_n)_{n \in \mathbb{N}} \subset \mathbb{R}_+$ et $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$ tels que

$$f(x) = \lim_{n \rightarrow +\infty} \sum_{i=1}^n a_i \mathbb{1}_{A_i}(x)$$

pour μ -presque tout $x \in \Omega$.

Ce que ce théorème nous dit essentiellement, c'est que les fonctions mesurables sont suffisamment régulières pour être approchées par des fonctions constantes par morceaux, mais au prix de considérer des ensembles mesurables arbitraires qui sont génériquement bien plus tortueux que des segments.

Nous sommes maintenant prêts à introduire la notion d'intégrale au sens de Lebesgue pour des fonctions mesurables quelconques.

Définition 15 (Intégrale de Lebesgue). Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré et $f : \Omega \rightarrow \mathbb{R}_+$ une fonction mesurable positive. Soit $(f_n)_{n \in \mathbb{N}}$ une suite croissante de fonctions étagées positives convergeant μ -presque partout vers f . Alors, l'**intégrale de Lebesgue de f par rapport à μ** est définie⁷ par

$$\int f \, d\mu = \lim_{n \rightarrow +\infty} \int f_n \, d\mu.$$

Pour une fonction f mesurable quelconque, il existe deux fonctions positives f^+ et f^- telles que $f = f^+ - f^-$. Si $\int (f^+ + f^-) \, d\mu < +\infty$ (ou de manière équivalente si $\int |f| \, d\mu < +\infty$), on dit que la fonction f est μ -**intégrable**. Son intégrale est alors définie par

$$\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu.$$

Pour $A \in \mathcal{A}$, on définit l'intégrale de f par rapport à μ sur le **domaine A** par

$$\int_A f \, d\mu = \int \mathbb{1}_A f \, d\mu.$$

Il est important de remarquer que cette définition ne dépend pas du choix de la suite de fonctions étagées (voir l'exercice plus haut), et que l'intégrale d'une fonction positive peut tout à fait être égale à $+\infty$.

7. Il est à noter que la suite d'intégrales $(\int f_n \, d\mu)_{n \in \mathbb{N}} \subset \mathbb{R}_+$ est croissante et qu'elle admet donc une limite (potentiellement infinie).

Remarque (Intégrale de Lebesgue et découpe horizontale). *Une autre interprétation que l'on peut donner à l'intégrale au sens de Lebesgue est la suivante. Rappelons que l'intégrale de Riemann propose d'approcher l'aire présente sous le graphe d'une fonction par une découpe verticale avec des rectangles. A l'inverse, le Théorème 14 énoncé ci-dessus se démontre (croyez moi !) en construisant inductivement des ensembles de sur-niveaux de f . Il s'agit d'une découpe horizontale où l'on considère comme éléments de Ω les images réciproques d'intervalles par f .*

Ceci permet d'ailleurs de comprendre pourquoi l'on doit se contenter d'ensembles mesurables généraux pour nos fonctions étagées. En effet, f étant une fonction mesurable, tout ce que l'on peut dire des pré-images de ses sur-niveaux est qu'ils forment des ensembles eux aussi mesurables.

Le théorème suivant contient des résultats classiques de linéarité, croissance et sous-additivité de l'intégrale, dont on disposait déjà dans la théorie de Riemann.

Théorème 16 (Linéarité, croissance et sous-additivité de l'intégrale).

Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré et $f, g : \Omega \rightarrow \mathbb{R}$ deux fonctions intégrables par rapport à μ . Alors :

1. *Pour tout $a \in \mathbb{R}$, on a*

$$\int (af + g) d\mu = a \int f d\mu + \int g d\mu.$$

2. *Si $f \leq g$ μ -presque partout dans Ω , alors*

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

3. *On déduit des deux points précédents que*

$$\int |f + g| d\mu \leq \int |f| d\mu + \int |g| d\mu.$$

Idée de preuve. La démonstration se fait en démontrant d'abord les points 1,2 et 3 pour des fonctions étagées, puis par passage à la limite on obtient les résultats correspondants pour des fonctions mesurables quelconques. \square

Telle quelle, cette définition de la notion d'intégrale reste assez abstraite et difficile à manipuler pour le calcul effectif d'intégrales. Qui plus est, il serait intéressant de voir si en pratique, cette nouvelle notion d'intégrale coïncide avec celle plus familière de Riemann lorsque les deux sont bien définies. Le théorème ci-dessous apporte une réponse positive à cette question.

Théorème 17 (Lien avec l'intégrale de Riemann). Soit $f : \mathbb{R}^d \rightarrow \mathbb{R}$ une fonction mesurable dont les parties positives et négatives f^+ et f^- sont Riemann-intégrables. Alors f est \mathcal{L}^d -intégrable – ou Lebesgue-intégrable –, et les intégrales de Riemann et Lebesgue coïncident, i.e.

$$\int f \, dx = \int f \, d\mathcal{L}^d.$$

En l'absence d'ambiguïté, on notera parfois $\int f \, dx$ l'intégrale de la fonction f contre la mesure de Lebesgue \mathcal{L}^d .

Remarque. Attention, les fonctions dont les intégrales au sens de Riemann sont semi-convergentes ne sont pas Lebesgue-intégrables ! Par exemple, la fonction $f : x \mapsto \frac{\sin(x)}{1+x}$ n'est pas Lebesgue-intégrable sur la droite réelle.

Jusqu'à présent, nous avons réussi à corriger les deux premiers défauts de l'intégrale de Riemann relevés en Section 3.1, à savoir son incapacité à mesurer la taille de certains ensembles assez naturels tels que \mathbb{Q} , et la rigidité de sa construction.

Les trois théorèmes énoncés ci-dessous nous montrent que l'intégrale de Lebesgue est robuste vis à vis de la notion de passage à la limite, ce qui répond à la troisième critique que nous avons relevée dans l'Introduction.

Théorème 18 (Lemme de Fatou). Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré et $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions à valeurs réelles, mesurables et **positives**. Alors

$$\int \liminf_{n \rightarrow +\infty} f_n \, d\mu \leq \liminf_{n \rightarrow +\infty} \int f_n \, d\mu.$$

Ce théorème – dont la preuve est un peu technique – est extrêmement général et puissant. Il va nous permettre notamment de démontrer en quelques lignes les théorèmes bien connus de *convergence monotone* et *convergence dominée*.

Théorème 19 (Convergence Monotone). Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré et $(f_n)_{n \in \mathbb{N}}$ une suite **croissante** de fonctions mesurables **positives** sur Ω . Supposons qu'il existe une fonction $f : \Omega \rightarrow \mathbb{R}_+$ telle que $f_n \rightarrow f$ μ -presque partout. Alors,

$$\lim_{n \rightarrow \infty} \int f_n \, d\mu = \int f \, d\mu,$$

où $\int f \, d\mu$ est bien définie car f est mesurable et positive comme limite ponctuelle de fonctions mesurables positives.

Preuve. Puisque par hypothèse la suite $(f_n)_{n \in \mathbb{N}}$ est croissante et converge ponctuellement vers f , l'on a que

$$\int f_n \, d\mu \leq \int f \, d\mu$$

pour tout $n \in \mathbb{N}$. Par croissance de la suite de fonctions $(f_n)_{n \in \mathbb{N}}$, il vient que $(\int f_n \, d\mu)_{n \in \mathbb{N}}$ est une suite croissante de réels dont la limite (potentiellement infinie!) vérifie

$$\lim_{n \rightarrow +\infty} \int f_n \, d\mu \leq \int f \, d\mu.$$

Nous obtenons l'inégalité inverse en remarquant que

$$\int f \, d\mu = \int \lim_{n \rightarrow +\infty} f_n \, d\mu = \int \liminf_{n \rightarrow +\infty} f_n \, d\mu \leq \liminf_{n \rightarrow +\infty} \int f_n \, d\mu = \lim_{n \rightarrow +\infty} \int f_n \, d\mu,$$

en utilisant le Lemme de Fatou et le fait que pour toute suite convergente, la limite et la limite inférieure coïncident. Nous pouvons alors déduire de ces deux inégalités que

$$\lim_{n \rightarrow +\infty} \int f_n \, d\mu = \int f \, d\mu,$$

ce qui achève la preuve. □

Théorème 20 (Convergence dominée). *Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré et $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions mesurables convergeant μ -presque partout vers une fonction $f : \Omega \rightarrow \mathbb{R}$. Supposons qu'il existe une autre fonction mesurable $g : \Omega \rightarrow \mathbb{R}$ vérifiant*

$$\sup_n f_n \leq g \quad \mu\text{-presque partout et} \quad \int g \, d\mu < +\infty.$$

Alors, la fonction f est Lebesgue-intégrable et

$$\lim_{n \rightarrow +\infty} \int f_n \, d\mu = \int f \, d\mu.$$

En particulier, on a

$$\lim_{n \rightarrow +\infty} \int |f - f_n| \, d\mu = 0.$$

Preuve. Remarquons tout d'abord que puisque $\lim_{n \rightarrow +\infty} |f - f_n| = 0$ μ -presque partout, l'on a

$$\int 2g \, d\mu = \int \left(2g - \liminf_{n \rightarrow +\infty} |f - f_n| \right) \, d\mu = \int \liminf_{n \rightarrow +\infty} (2g - |f - f_n|) \, d\mu.$$

Or, on sait par hypothèse que $f_n \leq g$ pour tout $n \in \mathbb{N}$, mais aussi que $f \leq g$ puisque g majore le supremum des $(f_n)_{n \in \mathbb{N}}$ et que cette suite converge vers f .

Ainsi, la fonction $2g - |f - f_n|$ est positive, et nous pouvons appliquer le Lemme de Fatou à l'expression précédente, i.e.

$$\int \liminf_{n \rightarrow +\infty} (2g - |f - f_n|) \, d\mu \leq \liminf_{n \rightarrow +\infty} \int (2g - |f - f_n|) \, d\mu.$$

Par construction (le vérifier!), les limites inférieures et supérieures d'une suite de réels $(a_n)_{n \in \mathbb{N}}$ sont reliées par la relation

$$\liminf_{n \rightarrow +\infty} [-a_n] = - \limsup_{n \rightarrow +\infty} [a_n],$$

En combinant cette identité et les deux séries d'inégalités ci-dessus, on obtient que

$$\limsup_{n \rightarrow +\infty} \int |f - f_n| \, d\mu \leq 0.$$

Ceci associé avec le fait que la limite inférieure d'une suite positive soit positive nous permet d'obtenir que

$$0 \leq \liminf_{n \rightarrow +\infty} \int |f - f_n| \, d\mu \leq \limsup_{n \rightarrow +\infty} \int |f - f_n| \, d\mu \leq 0,$$

et donc de conclure. □

Le théorème de convergence dominée permet donc de transformer une convergence μ -presque partout en une convergence *uniforme* au sens des intégrales contre μ (à mettre en parallèle avec la notion de norme L^p que nous introduirons en Section 3.5). Ce dernier permet aussi d'obtenir les deux résultats suivants traitant de *continuité* et de *différentiabilité sous le signe intégrale* qui sont d'une très grande utilité en calcul intégral et en modélisation de manière générale.

Théorème 21 (Continuité par rapport à un paramètre). Soient $(\Omega, \mathcal{A}, \mu)$ un espace mesuré, $E \subset \mathbb{R}$ et $f : E \times \Omega \rightarrow \mathbb{R}$. Supposons que

1. pour tout $t \in E$, la fonction $x \mapsto f(t, x)$ est mesurable,
2. pour μ -presque tout $x \in \Omega$, la fonction $t \mapsto f(t, x)$ est continue en $t = a$ pour un certain $a \in E$,
3. il existe une fonction $g : \Omega \rightarrow \mathbb{R}_+$ intégrable et telle que $|f(t, x)| \leq g(x)$ pour tout $t \in E$ et μ -presque tout $x \in \Omega$.

Alors la fonction $F : t \mapsto \int f(t, x) \, d\mu(x)$ est continue au point $a \in E$.

Théorème 22 (Dérivation par rapport à un paramètre). Soient $(\Omega, \mathcal{A}, \mu)$ un espace mesuré, $T \subset \mathbb{R}$ **ouvert** et $f : T \times \Omega \rightarrow \mathbb{R}$. Supposons que

1. pour tout $t \in T$, la fonction $x \mapsto f(t, x)$ est **intégrable**,

2. pour μ -presque tout $x \in \Omega$, la fonction $t \mapsto f(t, x)$ est dérivable sur T ,
3. il existe une fonction $g : \Omega \rightarrow \mathbb{R}_+$ intégrable et telle que

$$\left| \frac{\partial f}{\partial t}(t, x) \right| \leq g(x),$$

pour tout $t \in T$ et μ -presque tout $x \in \Omega$.

Alors, pour tout $t \in T$, la fonction $x \mapsto \frac{\partial f}{\partial t}(t, x)$ est intégrable. De plus, la fonction $F : t \mapsto \int f(t, x) d\mu(x)$ est dérivable sur T avec

$$F'(t) = \int \frac{\partial f}{\partial t}(t, x) d\mu(x).$$

Exercice. Démontrer les deux théorèmes ci-dessus en utilisant le théorème de Convergence Dominée et les rappels suivants.

1. Une fonction $t \mapsto \phi(t)$ est continue en $a \in E$ si et seulement si $\phi(a_n) \rightarrow \phi(a)$ pour toute suite $(a_n)_{n \in \mathbb{N}} \in E$ convergeant vers a .
2. Une fonction $t \mapsto \phi(t)$ est dérivable en $a \in T$ si et seulement si la limite

$$\lim_{h \rightarrow 0} \left[\frac{\phi(a+h) - \phi(a)}{h} \right]$$

existe et est finie en a . Cette limite est alors notée $\phi'(a)$.

Nous en avons maintenant fini avec la construction de l'intégrale de Lebesgue et l'étude de ses propriétés élémentaires. Le lecteur désireux d'approfondir ces notions et de découvrir des outils plus en vogue dans l'analyse moderne pourra consulter un des ouvrages suivants : *Functions of Bounded Variations and Free Discontinuity Problems* d'Ambrosio, Fusco et Pallara qui est très complet mais assez difficile, ou alors *Measure Theory and Fine Properties of Functions* de Evans et Gariepy, qui est beaucoup plus accessible. Pour quelque chose d'orienté aussi un peu vers les probabilités, je recommande chaudement les notes de cours *Intégration, Probabilités et Processus Aléatoires* de Jean-François Le Galle.

3.4 Intégrales multiples

Nous avons dans la section précédente détaillé la construction de l'intégrale de Lebesgue sur un espace mesuré général $(\Omega, \mathcal{A}, \mu)$. Cependant, il peut arriver que nous ayons des informations supplémentaires sur l'espace Ω . Si l'on veut par exemple étudier une équation d'évolution (d'un profil de température, du mouvement d'un fluide, etc...) jusqu'à un horizon de temps T , on aura typiquement

$\Omega = [0, T] \times \mathbb{R}^d$, voire $[0, T] \times M$ pour un espace de configurations d'états M potentiellement plus compliqué.

On en vient alors à se poser naturellement la question suivante : pour une fonction $f : E \times F \rightarrow \mathbb{R}$ où (E, \mathcal{E}) et (F, \mathcal{F}) sont deux espaces mesurables donnés, quelle notion de mesurabilité doit-on utiliser ? En particulier, y a-t-il une tribu que l'on puisse construire sur $E \times F$ à partir de \mathcal{E} et \mathcal{F} ? Nous répondons par l'affirmative à cette première question avec l'introduction ci-dessous de la notion de *tribu produit*.

Définition 23 (Tribu produit). Soient (E, \mathcal{E}) et (F, \mathcal{F}) deux espaces mesurables. On appelle **tribu produit**, que l'on note $\mathcal{E} \otimes \mathcal{F}$, la tribu engendrée par les produits cartésiens d'ensembles de \mathcal{E} et \mathcal{F} , i.e.

$$\mathcal{E} \otimes \mathcal{F} = \sigma(\{A \times B \text{ t.q. } A \in \mathcal{E}, B \in \mathcal{F}\}).$$

Dans ce qui suit, on dira systématiquement d'une fonction $f : E \times F \rightarrow \mathbb{R}$ qu'elle est mesurable pour dire qu'elle est mesurable au sens de la tribu produit $\mathcal{E} \otimes \mathcal{F}$.

Si l'on prend le cas particulier de l'espace \mathbb{R}^{p+q} avec p, q entiers, il existe deux notions de tribu qu'il semblerait naturel de considérer : la tribu Borélienne $\mathcal{B}(\mathbb{R}^{p+q})$, i.e. la tribu engendrée par les ouverts de \mathbb{R}^{p+q} , et la tribu produit $\mathcal{B}(\mathbb{R}^p) \otimes \mathcal{B}(\mathbb{R}^q)$. Le théorème ci-dessous montre qu'en réalité ces deux tribus coïncident.

Théorème 24 (Tribu produit sur \mathbb{R}^{p+q}). Pour tous p et q entiers, l'on a que

$$\mathcal{B}(\mathbb{R}^{p+q}) = \mathcal{B}(\mathbb{R}^p) \otimes \mathcal{B}(\mathbb{R}^q).$$

Éléments de preuve. On a par définition de la notion de tribu produit que

$$\mathcal{B}(\mathbb{R}^p) \otimes \mathcal{B}(\mathbb{R}^q) = \sigma(\{A \times B \text{ t.q. } A \in \mathcal{B}(\mathbb{R}^p), B \in \mathcal{B}(\mathbb{R}^q)\}).$$

Par ailleurs, on a vu que la tribu Borélienne de \mathbb{R} était explicitement caractérisée par $\mathcal{B}(\mathbb{R}) = \sigma(\{]a, b[\text{ t.q. } a, b \in \mathbb{Q}\})$. Il se trouve que l'on a en fait

$$\mathcal{B}(\mathbb{R}^d) = \sigma(\{]a_1, b_1[\times \dots \times]a_d, b_d[\text{ t.q. } a_i, b_i \in \mathbb{Q} \text{ pour tout } i \in \{1, \dots, d\}\}).$$

En prenant $d = p + q$ et en jouant avec ces deux notions, on peut démontrer le résultat souhaité. \square

De la même façon, il est possible de définir une notion canonique de *mesure produit* sur un produit d'espaces mesurés.

Définition 25 (Mesure produit). Soient (E, \mathcal{E}, μ) et (F, \mathcal{F}, ν) deux espaces mesurés. Il existe une mesure sur $(E \times F, \mathcal{E} \otimes \mathcal{F})$, que l'on note $\mu \times \nu$, telle que

$$(\mu \times \nu)(A \times B) = \mu(A)\nu(B) \text{ pour tout } A \in \mathcal{E} \text{ et } B \in \mathcal{F}.$$

On dit que $\mu \times \nu$ est une **mesure produit**.

Il est à noter que cette mesure est unique dans un bon nombre de situations, par exemple lorsque μ et ν sont des mesures de probabilité⁸. On dispose d'une caractérisation explicite des produits de mesures de Lebesgue donnée dans le théorème ci-dessous.

Théorème 26 (Produits de mesures de Lebesgue). La mesure produit obtenue à partir de la mesure de Lebesgue sur \mathbb{R}^p et de la mesure de Lebesgue sur \mathbb{R}^q est la mesure de Lebesgue sur \mathbb{R}^{p+q} , i.e.

$$\mathcal{L}^p \times \mathcal{L}^q = \mathcal{L}^{p+q}.$$

Remarque. Le résultat du théorème précédent est intuitivement cohérent si l'on se rappelle que les tribus Boréliennes sont engendrées par des produits d'intervalles, et que la mesure de Lebesgue est elle aussi caractérisée de manière unique par son comportement sur ces ensembles produits.

Étant donné une fonction mesurable $f : E \times F \rightarrow \mathbb{R}$, on peut définir son intégrale au sens de Lebesgue contre des mesures données sur $(E \times F, \mathcal{E} \otimes \mathcal{F})$ grâce à la théorie développée dans les Sections 3.2 et 3.3. Dans le cas particulier où l'on considère une mesure produit $\mu \times \nu$, il est raisonnable de se demander si l'on peut obtenir l'intégrale de f contre $\mu \times \nu$ en intégrant d'abord f par rapport à μ et ensuite par rapport à ν , ou vice versa.

La réponse à cette question est oui, mais sous certaines hypothèses comme l'illustrent les théorèmes de *Tonelli* et *Fubini* énoncés ci-dessous.

Théorème 27 (Tonelli). Soient (E, \mathcal{E}, μ) et (F, \mathcal{F}, ν) deux espaces mesurés et $f : E \times F \rightarrow \mathbb{R}$ une fonction **mesurable** et **positive**. Alors,

$$\int f(x, y) d(\mu \times \nu)(x, y) = \int \left(\int f(x, y) d\mu(x) \right) d\nu(y) = \int \left(\int f(x, y) d\nu(y) \right) d\mu(x).$$

Il est à noter que le théorème de Tonelli ne requiert pas que les intégrales soient finies pour s'appliquer. Ce théorème d'intervention se généralise aux fonctions signées au prix d'une hypothèse supplémentaire d'intégrabilité.

⁸. Une condition assez générale pour que cette mesure soit unique est que les mesures μ et ν soient σ -finies, mais cette notion dépasse le cadre du cours.

Théorème 28 (Fubini). Soient (E, \mathcal{E}, μ) et (F, \mathcal{F}, ν) deux espaces mesurés et $f : E \times F \rightarrow \mathbb{R}$ une fonction **mesurable**. Si la fonction f est **intégrable par rapport à $\mu \times \nu$** , i.e. si

$$\int |f| d(\mu \times \nu) < +\infty,$$

alors

$$\int f(x, y) d(\mu \times \nu)(x, y) = \int \left(\int f(x, y) d\mu(x) \right) d\nu(y) = \int \left(\int f(x, y) d\nu(y) \right) d\mu(x).$$

Remarque. La condition d'intégrabilité $\int |f| d(\mu \times \nu) < +\infty$ à vérifier dans le théorème de Fubini peut se réécrire sous la forme

$$\int \left(\int |f(x, y)| d\mu(x) \right) d\nu(y) < +\infty \quad \text{ou encore} \quad \int \left(\int |f(x, y)| d\nu(y) \right) d\mu(x) < +\infty$$

grâce au théorème de Tonelli.

Exercices :

1. En utilisant ces théorèmes, donner deux critères sur une suite de fonctions $(f_n)_{n \in \mathbb{N}}$ pour que

$$\int \sum_{n \geq 1} f_n(x) dx = \sum_{n \geq 1} \int f_n(x) dx.$$

2. Montrer à l'aide des théorèmes ci-dessus que

$$\int_D xy d\mathcal{L}^2(x, y) = \frac{1}{8}$$

où D est le domaine contenu dans le triangle $\{(0, 0), (1, 0), (0, 1)\}$.

3. Montrer que la fonction $f : (x, y) \mapsto \frac{x^2 - y^2}{(x^2 + y^2)^2}$ ne satisfait pas les hypothèses du théorème de Fubini.

Nous concluons cette partie du cours par un autre théorème très classique et utile permettant de manipuler l'intégrale de Lebesgue sur \mathbb{R}^d : le *théorème de changement de variable*.

Théorème 29 (Changement de variable). Soient $f : \mathbb{R}^d \rightarrow \mathbb{R}$ une fonction mesurable et $U, V \subset \mathbb{R}^d$ deux ouverts tels qu'il existe un C^1 -difféomorphisme $\Phi : U \rightarrow V$. Alors si f est **positive** ou **Lebesgue-intégrable**, l'on a

$$\int_V f d\mathcal{L}^d = \int_U \det(J_\Phi) f \circ \Phi d\mathcal{L}^d$$

où J_Φ désigne la matrice Jacobienne de Φ .

Rappels (Difféomorphismes et matrice Jacobiennes).

- ◇ On dit que $\Phi : U \rightarrow V$ est un C^1 -difféomorphisme si Φ est une bijection de classe C^1 dont la réciproque $\Phi^{-1} : V \rightarrow U$ est aussi de classe C^1 .
- ◇ La matrice Jacobienne $J_\Phi : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ d'une application $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ est la matrice de ses dérivées partielles dans une base choisie, i.e.

$$J_\Phi : x \mapsto \left(\frac{\partial \Phi_i}{\partial x_j}(x) \right)_{1 \leq i, j \leq d}.$$

Exercice. Calculer la matrice Jacobienne du changement de variable très classique en mécanique du point $\Phi : (r, \theta) \mapsto (r \cos(\theta), r \sin(\theta))$.

Remarque (Interprétation géométrique du déterminant Jacobien).

Le terme $\det(J_\Phi)$ présent dans la formule de changement de variable peut-être interprété comme la quantité de distortion d'un petit élément de volume (au sens de la mesure de Lebesgue) lorsque l'on prend son image par Φ . Étant donné que l'intégrale contre la mesure de Lebesgue ambiante mesure des volumes et que l'écriture $f \circ \Phi$ ne rend compte que d'un changement de coordonnées, il est naturel que ce facteur d'échelle apparaisse dans la formule de changement de variables.

3.5 Espaces L^p

Dans cette dernière partie du cours, nous introduisons une famille d'espaces de fonctions qui sont centraux dans l'étude des équations issues de la physique et dont la construction est basée sur la notion d'intégrale au sens de Lebesgue : les espaces L^p .

Pour un espace mesuré (E, \mathcal{E}, μ) , nous avons vu grâce aux Théorèmes 11 et 16 que l'ensemble des fonctions μ -intégrables forment espace vectoriel, que nous allons noter $\mathcal{L}^1(E, \mathcal{E}, \mu)$. De même pour tout réel $p \in [1, +\infty[$, l'on note

$$\mathcal{L}^p(E, \mathcal{E}, \mu) = \left\{ f : E \rightarrow \mathbb{R} \text{ t.q. } f \text{ est mesurable et } \int |f|^p d\mu < \infty \right\}.$$

En l'absence d'ambiguïté, il est fréquent d'écrire

$$\mathcal{L}^p(E) = \mathcal{L}^p(E, \mu) = \mathcal{L}^p(E, \mathcal{E}, \mu).$$

L'intégrale de Lebesgue va nous permettre de construire une notion canonique et souple de mesure de distance sur ces espaces de fonctions. Nous rappelons ci-dessous les définitions des concepts de *semi-norme* et de *norme* sur un espace vectoriel.

Définition 30 (Normes et semi-normes). Soit V un \mathbb{R} -espace vectoriel quelconque. Une application $p : V \rightarrow [0, +\infty]$ est une **norme** sur V si elle satisfait les conditions suivantes :

- (i) Pour tout $\lambda \in \mathbb{R}$ et $v \in V$, on a $p(\lambda v) = |\lambda|p(v)$ (homogénéité positive).
- (ii) Pour tout $v, w \in V$, on a $p(v + w) \leq p(v) + p(w)$ (inégalité triangulaire).
- (iii) On a $p(v) = 0$ si et seulement si $v = 0$ (caractère défini).

Une application $p : V \rightarrow [0, +\infty]$ satisfaisant (i) et (ii) est appelée une **semi-norme**.

On dénotera génériquement dans la suite par $(V, \|\cdot\|_V)$ un espace vectoriel V muni d'une norme $\|\cdot\|_V$. Il est alors facile de définir une semi-norme sur les espaces $\mathcal{L}^p(E, \mathcal{E}, \mu)$ comme illustré dans la proposition suivante.

Propriété 3 (\mathcal{L}^p semi-norme). Soit (E, \mathcal{E}, μ) un espace mesuré. L'application

$$|\cdot|_{\mathcal{L}^p(\mu)} : f \mapsto \left(\int |f|^p d\mu \right)^{1/p}$$

définit une semi-norme sur l'espace $\mathcal{L}^p(E, \mathcal{E}, \mu)$.

Cette proposition se démontre aisément utilisant les résultats de la Section 3.3. Le fait que cette application ne soit qu'une semi-norme vient du fait que par construction, l'intégrale de Lebesgue ne capte pas le comportement d'une fonction sur des ensembles de mesure nulle. On sait par exemple que $\int_{[0,1]} \mathbb{1}_{\mathbb{Q}} d\mathcal{L}^1 = 0$ alors que $\mathbb{1}_{\mathbb{Q}}$ n'est pas la fonction nulle (elle prend même la valeur 1 une infinité de fois).

Afin d'obtenir un espace vectoriel normé, nous allons être obligés d'identifier les fonctions qui coïncident μ -presque partout. Nous rappelons à cet effet les notions de *relation d'équivalence* et d'*espace quotient*.

Rappels (Relation d'équivalence et quotient). Soit X un ensemble quelconque. Une **relation d'équivalence** \sim sur X est une loi de correspondance entre les éléments de X vérifiant les prérequis suivants.

- (i) Pour tout $x \in X$, $x \sim x$ (réflexivité).
- (ii) Pour tout $x, y \in X$, $x \sim y$ si et seulement si $y \sim x$ (symétrie).
- (iii) Pour tout $x, y, z \in X$, $x \sim y$ et $y \sim z$ impliquent $x \sim z$ (transitivité).

On appelle **classe d'équivalence** d'un élément $x \in X$ l'ensemble des $y \in X$ tels que $x \sim y$.

Étant donné un ensemble muni d'une relation d'équivalence (X, \sim) , on définit le **quotient** X/\sim de X par \sim comme étant l'ensemble des classes d'équivalence de X .

Dans notre cas, nous allons utiliser cette notion de famille de classes d'équivalence pour identifier les fonctions de $\mathcal{L}^p(E, \mathcal{E}, \mu)$ qui coïcident μ -presque partout : c'est la définition d'un espace L^p .

Définition 31 (Égalité presque partout et espaces L^p). Soit (E, \mathcal{E}, μ) un espace mesuré. La relation d'égalité μ -presque partout définit une relation d'équivalence \sim_μ sur l'espace $\mathcal{L}^p(E, \mathcal{E}, \mu)$, et l'on désigne par

$$L^p(E, \mathcal{E}, \mu) \equiv \mathcal{L}^p(E, \mathcal{E}, \mu) / \sim_\mu$$

l'espace quotient correspondant où l'on a identifié toutes les fonctions qui coïcident μ -presque partout. La restriction de la semi-norme $|\cdot|_{\mathcal{L}^p(\mu)}$ à l'espace $L^p(E, \mathcal{E}, \mu)$ définit une norme que l'on note $\|\cdot\|_{L^p(\mu)}$.

Ainsi, l'espace $(L^p(E, \mathcal{E}, \mu), \|\cdot\|_{L^p(\mu)})$ est un espace vectoriel normé de dimension infinie pour tout $p \in [1, +\infty[$. Nous allons voir dans le théorème ci-dessous qu'il possède en fait une structure *beaucoup* plus riche.

Théorème 32 (Riesz-Fischer). L'espace $(L^p(E, \mathcal{E}, \mu), \|\cdot\|_{L^p(\mu)})$ est un espace de Banach séparable pour tout $p \in [1, +\infty[$. Dans le cas particulier $p = 2$, nous avons que $(L^2(E, \mathcal{E}, \mu), \|\cdot\|_{L^2(\mu)})$ est un espace de Hilbert pour le produit scalaire

$$\langle \cdot, \cdot \rangle_{L^2(\mu)} : (f, g) \mapsto \int fg \, d\mu.$$

La notion de convergence dans $L^2(E, \mathcal{E}, \mu)$ est généralement appelée convergence en moyenne quadratique.

Rappels (Suites de Cauchy, espaces de Banach et espaces de Hilbert).

Soit $(V, \|\cdot\|_V)$ un espace vectoriel normé.

◇ Une suite $(v_n)_{n \in \mathbb{N}} \subset V$ est dite de Cauchy pour $\|\cdot\|_V$ si et seulement si

$$\lim_{n, p \rightarrow +\infty} \|v_n - v_p\|_V = 0,$$

i.e. les termes de $(v_n)_{n \in \mathbb{N}}$ tendent à se rapprocher de plus en plus.

◇ On dit que $(V, \|\cdot\|_V)$ est un espace de Banach si et seulement si toute suite de Cauchy pour $\|\cdot\|_V$ converge.

◇ On dit que $(V, \|\cdot\|_V)$ est un espace de Hilbert si c'est un espace de Banach dont la norme est induite par un produit scalaire, i.e.

$$\|\cdot\|_V \equiv \sqrt{\langle \cdot, \cdot \rangle_V} \quad \text{où } \langle \cdot, \cdot \rangle_V \text{ est un produit scalaire sur } V.$$

Définition 33 (Espace normé séparable). Soit $(V, \|\cdot\|_V)$ un espace normé. On dit de V qu'il est **séparable** s'il admet un sous ensemble $V_D \subset V$ qui est **dénombrable** et **dense** dans V i.e. pour tout $v \in V$, il existe une suite d'éléments $(v_n)_{n \in \mathbb{N}} \subset V_D$ telle que $\|v - v_n\|_V \rightarrow 0$.

Remarque (Importance de la séparabilité). La séparabilité est souvent une propriété cruciale pour construire de bons outils d'analyse fonctionnelle sur un espace. Il existe des espaces non séparables que l'on utilise, comme l'espace $L^\infty(E, \mathcal{E}, \mu)$ que l'on va introduire ci-dessous, mais ils ont généralement une structure bien moins riche topologiquement parlant et bien moins commode que les espaces qui le sont.

À ce stade, ça ne coûte d'ailleurs pas cher de remarquer que \mathbb{R} est un espace séparable avec comme sous ensemble dénombrable dense l'ensemble \mathbb{Q} des nombres rationnels!

Le lecteur attentif aura remarqué que dans ce qui précède, nous avons supposé systématiquement que p était fini. Il est cependant possible de définir l'espace L^p correspondant à la valeur $p = +\infty$ comme nous l'explicitons ci-dessous.

Définition 34 (Espace L^∞). Soit (E, \mathcal{E}, μ) un espace mesuré. On dit qu'une fonction f appartient à l'espace⁹ $L^\infty(E, \mathcal{E}, \mu)$ si f est mesurable et s'il existe une constante $M \geq 0$ telle que

$$\mu(\{x \in E \text{ t.q. } |f(x)| \geq M\}) = 0.$$

Cet espace est un espace de Banach lorsqu'on le munit de la norme

$$\|\cdot\|_{L^\infty(\mu)} : f \mapsto \inf \{M > 0 \text{ t.q. } \mu(\{x \in E \text{ t.q. } |f(x)| \geq M\}) = 0\}.$$

En revanche, $(L^\infty(E, \mathcal{E}, \mu), \|\cdot\|_{L^\infty(\mu)})$ n'est pas un espace séparable.

Nous allons à présent finir cette introduction à la théorie des espaces L^p par les énoncés deux résultats d'analyse fonctionnelle extrêmement utiles : les inégalités de *Cauchy-Schwartz* et de *Hölder*.

Théorème 35 (Inégalité de Cauchy-Schwartz). Soient (E, \mathcal{E}, μ) un espace mesuré et $f, g \in L^2(E, \mathcal{E}, \mu)$. Alors, on a que $|\langle f, g \rangle_{L^2(\mu)}| \leq \|f\|_{L^2(\mu)} \|g\|_{L^2(\mu)}$, ou plus explicitement

$$\left| \int fg \, d\mu \right| \leq \sqrt{\int |f|^2 \, d\mu} \sqrt{\int |g|^2 \, d\mu}.$$

9. Où l'on n'oublie pas de quotienter à nouveau par la relation d'équivalence "égalité μ -presque partout"!

Il est bon de remarquer que cette inégalité est exactement la même que celle que l'on connaissait dans le cas des espaces \mathbb{R}^d muni de leur produit scalaire Euclidien. Ce fait n'est pas si surprenant, étant donné que les espaces de Hilbert sont ceux qui de par leur structure ressemblent le plus à des espaces Euclidiens de dimensions finie.

L'inégalité de Hölder énoncée ci-dessous est un genre de généralisation de l'inégalité de Cauchy-Schwartz pour les espaces L^p avec $p \in [1, +\infty]$ quelconque.

Théorème 36 (Inégalité de Hölder). *Soit (E, \mathcal{E}, μ) un espace mesuré et $p, q \in [1, +\infty]$ deux réels tels que $\frac{1}{p} + \frac{1}{q} = 1$, où par convention l'on pose que $q = +\infty$ quand $p = 1$ et vice versa.*

Alors, pour toutes fonctions $f \in L^p(E, \mathcal{E}, \mu)$ et $g \in L^q(E, \mathcal{E}, \mu)$, on a que $fg \in L^1(E, \mathcal{E}, \mu)$ avec

$$\|fg\|_{L^1(\mu)} \leq \|f\|_{L^p(\mu)} \|g\|_{L^q(\mu)}$$

En particulier, $L^p(E, \mathcal{E}, \mu) \subset L^q(E, \mathcal{E}, \mu)$ pour tout $p < q$ dès lors que $\mu(E) < +\infty$.

Exercices de base – Intégration de Lebesgue

Exercice. *Lequel des ensembles suivants est une tribu sur \mathbb{R} ?*

1. $\{0, 1, 2\}$
2. $\{\mathbb{R}, \emptyset\}$
3. \mathbb{R}
4. $\{\mathbb{R}, \emptyset, [0, 1]\}$
5. $\{\mathbb{R}, \emptyset, [0, 1],] - \infty, 0[\cup]1; \infty[\}$.

Exercice. *Soit $(A_i)_{i \geq 1}$ une suite d'éléments distincts de la tribu \mathcal{A} . Laquelle des fonctions suivantes est une fonction étagée pour \mathcal{A} ?*

1. $13\mathbf{1}_{A_1}$
2. $4\mathbf{1}_{A_1 \cap A_3} - 12$
3. $\sum_{i=1}^5 3^i \mathbf{1}_{A_i}$
4. $\sum_{i=1}^{\infty} 2^{-i} \mathbf{1}_{A_i}$.

Exercice. *Soit λ la mesure de Lebesgue sur \mathbb{R} . On note $f = 4\mathbf{1}_{[0,2]} - 5\mathbf{1}_{[12;15]}$. Calculer $\int f d\lambda$.*

Exercice. *Soit $f = 4\mathbf{1}_{[0,1]}$ et $g = -5\mathbf{1}_{[2,4]}$. Calculer $\int f d\delta_2$ et $\int g d\delta_2$.*

Exercice. *Soit $\mu = \delta_1 + \delta_2$. On pose $f = \mathbf{1}_{[0,1]} + 3\mathbf{1}_{[2,4]}$. Calculer $\int f d\mu$.*

Exercice. *Soit $\mu = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_2$. On pose $f = \mathbf{1}_{[0,1]} + 3\mathbf{1}_{[2,4]}$. Calculer $\int f d\mu$.*

Exercice. *Soit $\mu = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_2$. On définit $f : \mathbb{R} \rightarrow \mathbb{R}$ telle que $\forall x \in \mathbb{R}, f(x) = x$. Calculer $\int f d\mu$.*

Exercice. *Soit $\mu = \lambda + \delta_2$ où λ désigne la mesure de Lebesgue. On pose $f = \mathbf{1}_{[0,1]} + 3\mathbf{1}_{[2,4]}$. Calculer $\int f d\mu$.*