# Leçon 2

## Rappels de la théorie de la mesure : mesure

Cette leçon introduit la notion de mesure sur un espace mesurable, dont les propriétés étendent les notions et intuitions de longueur, surface ou volume dans les espaces euclidiens classiques.

Ainsi que présenté dans la première leçon, un espace mesurable $(X, \mathcal{A})$ est composé d'un ensemble $X$ muni d'une $\sigma$-algèbre $\mathcal{A}$ de parties (de $X$).

# 1 Mesure

La notion fondamentale de mesure sur un espace mesurable est décrite dans la définition suivante.

**Définition 1** (Mesure)**.** *Une mesure sur un espace mesurable $(X, \mathcal{A})$ est une application $\mu : \mathcal{A} \to [0, \infty]$ telle que $\mu(\emptyset) = 0$ et si $A_n \in \mathcal{A}$, $n \in \mathbb{N}$, sont disjoints deux à deux,*

$$\mu\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) = \sum_{n\in\mathbb{N}} \mu(A_n).$$

Pour rappel, $\infty$ est entendu comme $+\infty$. La mesure $\mu$ est une application définie sur la $\sigma$-algèbre $\mathcal{A}$, même si dans bien des cas le langage glissera en « $\mu$ mesure sur $X$ » lorsque la $\sigma$-algèbre $\mathcal{A}$ sur $X$ sera clairement sous-entendue (par exemple la tribu borélienne sur $\mathbb{R}$ ou $\mathbb{R}^d$).

Il est à noter que dans l'égalité de cette définition, les deux membres peuvent être infinis, le premier parce que l'ensemble $\bigcup_{n\in\mathbb{N}} A_n$ peut être de mesure infinie, le second parce que la série (à termes positifs) peut diverger (ou l'un des $A_n$ être déjà de mesure infinie).

La définition précédente est plus précisément celle d'une *mesure positive*, il existe en effet des mesures de signe quelconque (*signées*), voire *complexes*.

Un espace mesurable $(X, \mathcal{A})$ muni d'une mesure $\mu$ est appelé *espace mesuré* $(X, \mathcal{A}, \mu)$.

Si $\mu(X) < \infty$, on dit que $\mu$ est finie, et si $\mu(X) = 1$, $\mu$ est une *mesure de probabilité*, ou plus simplement une *probabilité*. La mesure $\mu$ est dite *$\sigma$-finie* s'il existe une famille $A_n$, $n \in \mathbb{N}$, (qui peut être choisie croissante) d'éléments de $\mathcal{A}$ telle que $\bigcup_{n\in\mathbb{N}} A_n = X$ et $\mu(A_n) < \infty$ pour tout $n \in \mathbb{N}$.

Les premières propriétés d'une mesure sont rassemblées dans l'énoncé suivant.

**Proposition 2** (Propriétés d'une mesure)**.** *Soit $\mu$ une mesure sur un espace mesurable $(X, \mathcal{A})$ ;*

*a) (Croissance) Si $A, B \in \mathcal{A}$, $A \subset B$, alors $\mu(A) \leq \mu(B)$.*

*b) (Sous-additivité) Si $A_n \in \mathcal{A}$, $n \in \mathbb{N}$, est une famille quelconque, alors $\mu\big(\bigcup_{n\in\mathbb{N}} A_n\big) \leq \sum_{n\in\mathbb{N}} \mu(A_n)$.*

*c) (Convergence monotone croissante) Si $A_n \in \mathcal{A}$, $n \in \mathbb{N}$, est une famille croissante ($A_n \subset A_{n+1}$ pour tout $n$), alors $\mu\big(\bigcup_{n\in\mathbb{N}} A_n\big) = \lim_{n\to\infty} \mu(A_n)$.*

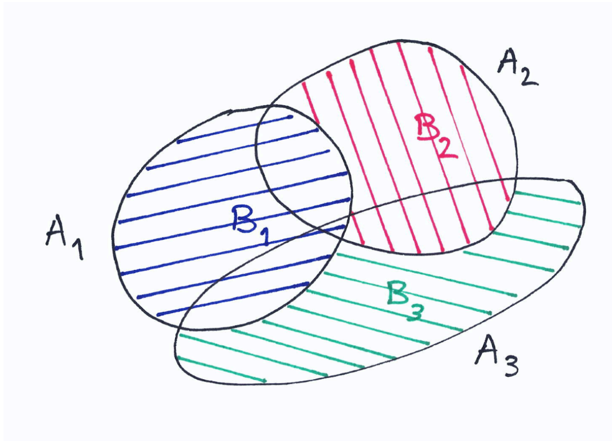
*d) (Convergence monotone décroissante) Si $A_n \in \mathcal{A}$, $n \in \mathbb{N}$, est une famille décroissante ($A_n \supset A_{n+1}$ pour tout $n$), et s'il existe $n_0 \in \mathbb{N}$ tel que $\mu(A_{n_0}) < \infty$, alors $\mu\big(\bigcap_{n\in\mathbb{N}} A_n\big) = \lim_{n\to\infty} \mu(A_n)$.*

Quelques éléments de démonstration. Pour a), il suffit de considérer la suite $A_n \in \mathcal{A}$, $n \in \mathbb{N}$, constituée des éléments disjoints $A_1 = A$, $A_2 = B \setminus A$, $A_n = \emptyset$ pour $n \geq 3$, de sorte que

$$\mu(B) \,=\, \mu(A_1 \cup A_2) \,=\, \mu(A_1) + \mu(A_2) \,\geq\, \mu(A_1) \,=\, \mu(A).$$

La démonstration du point b) est instructive car elle s'appuie sur une construction intéressante en soi (qui sera souvent utilisée par la suite). Il s'agit en effet de passer d'une réunion quelconque à une réunion disjointe. À cet effet, définir par récurrence la suite $B_n$, $n \in \mathbb{N}$, par $B_1 = A_1$, $B_2 = A_2 \cap B_1^c$, $B_3 = A_3 \cap (B_1 \cup B_2)^c$ et ainsi de suite, soit, pour tout $n \geq 2$,

$$B_n \,=\, A_n \cap \bigg(\bigcup_{k=1}^{n-1} B_k\bigg)^c.$$




Les ensembles $B_n$, $n \in \mathbb{N}$, sont éléments de $\mathcal{A}$, vérifient $B_n \subset A_n$, $n \in \mathbb{N}$, et sont disjoints deux à deux, de même réunion que la suite $A_n$, $n \in \mathbb{N}$. En effet, par construction (récurrence), $\bigcup_{k=1}^n B_k = \bigcup_{k=1}^n A_k$ pour tout $n \geq 1$, ce qui conduit à la dernière assertion. Pour vérifier la disjonction deux à deux, si $x \in B_k \cap B_n$ pour $k < n$, alors $x \in A_k$ (car $x \in B_k$), mais dans le même temps, comme

$$x \in B_n = A_n \cap \left( \bigcup_{k=1}^{n-1} B_k \right)^c = A_n \cap \left( \bigcup_{k=1}^{n-1} A_k \right)^c,$$

il s'avère que $x \notin A_k$ car $k < n$. Donc nécessairement $B_k \cap B_n = \emptyset$. En vertu de cette construction,

$$\mu\left( \bigcup_{n \in \mathbb{N}} A_n \right) = \mu\left( \bigcup_{n \in \mathbb{N}} B_n \right) = \sum_{n \in \mathbb{N}} \mu(B_n) \leq \sum_{n \in \mathbb{N}} \mu(A_n)$$

puisque, pour chaque $n$, $B_n \subset A_n$, et donc $\mu(B_n) \leq \mu(A_n)$.

Des raisonnements similaires conduisent aux assertions c) et d).

Deux petites propriétés souvent utiles peuvent être ajoutées à la liste de la proposition : pour tous $A, B \in \mathcal{A}$,

$$\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B),$$

et si $\mu(X) < \infty$,

$$\mu(A^c) = \mu(X) - \mu(A)$$

(en particulier, si $\mu$ est une mesure de probabilité, $\mu(A^c) = 1 - \mu(A)$).

Il est à noter que le point d) de la proposition précédente nécessite une hypothèse supplémentaire par rapport au point correspondant c) pour une famille croissante. Sur $X = \mathbb{N}$ muni de la $\sigma$-algèbre des parties $\mathcal{A} = \mathcal{P}(\mathbb{N})$, considérer la mesure $\mu$, dite de *comptage*, qui à toute partie $A$ de $\mathbb{N}$ associe son cardinal,

$$\mu(A) \;=\; \mathrm{Card}(A).$$

Si pour tout $n \in \mathbb{N}$, $A_n$ est l'ensemble des entiers $k \geq n$, les $A_n$, $n \in \mathbb{N}$, forment une suite décroissante d'éléments de $\mathcal{P}(\mathbb{N})$ dont l'intersection est réduite à l'ensemble vide (il n'existe pas d'entier plus grand que tous les entiers), et la conclusion d) de la Proposition 2 est alors en défaut tout simplement parce que $\mu(A_n) = \infty$ pour tout $n$.

En examinant la définition d'une mesure, deux exemples immédiats, mais d'intérêt limité, sont fournis par la mesure $\mu(A) = 0$ pour tout $A \in \mathcal{A}$ (*mesure nulle*), et la mesure $\mu(A) = \infty$ pour tout $A \in \mathcal{A}$ distinct de l'ensemble vide.

Un exemple plus intéressant, et en fait fondamental, est fourni par la *mesure de Dirac*[1] $\mu = \delta_x$ en un point $x \in X$ définie pour tout $A \in \mathcal{A}$ par

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

Il est instructif d'observer que $\delta_x(A) = \mathbb{1}_A(x)$. C'est une mesure de probabilité. Il est naturel de considérer cette mesure comme une mesure *atomique* affectant la *masse* 1 au singleton $\{x\}$, sous réserve toutefois que $\{x\} \in \mathcal{A}$ (ce qui sera la cas dans toutes les illustrations).

La mesure de comptage sur $\mathbb{N}$ peut s'interpréter comme $\sum_{n\in\mathbb{N}} \delta_n$, affectant à tout entier la masse 1, puisque

$$\bigg(\sum_{n\in\mathbb{N}} \delta_n\bigg)(A) \;=\; \sum_{n\in A} \delta_n \;=\; \mathrm{Card}(A), \quad A \in \mathcal{P}(\mathbb{N}).$$

1. Paul Dirac, mathématicien et physicien britannique (1902–1984).

C'est une mesure $\sigma$-finie. Une mesure de comptage peut être définie sur n'importe quel ensemble fini ou dénombrable.

Les paragraphes suivants seront l'occasion de construire de nouvelles mesures à partir d'opérations simples.

# 2 Ensemble négligeable et propriété presque partout

Un ensemble négligeable est essentiellement un ensemble de mesure nulle. La définition précise est un peu plus compliquée.

**Définition 3** (Ensemble négligeable)**.** *Soit $(X, \mathcal{A}, \mu)$ un espace mesuré ; une partie $N \subset X$ (donc pas nécessairement mesurable) est dite négligeable (par rapport à $\mu$), ou $\mu$-négligeable, s'il existe $A \in \mathcal{A}$ tel que $N \subset A$ et $\mu(A) = 0$.*

Il est souvent commode de compléter une $\sigma$-algèbre $\mathcal{A}$ afin qu'elle contienne les ensembles négligeables par rapport à une mesure $\mu$ donnée. Cette complétion s'effectue en posant

$$\mathcal{A}_\mu = \{A \cup N\,;\, A \in \mathcal{A},\, N \ \mu\text{-négligeable}\},$$

qui définit une $\sigma$-algèbre sur $X$ contenant les ensembles $\mu$-négligeables (et contenant $\mathcal{A}$) : c'est la $\sigma$-algèbre *complétée* de $\mathcal{A}$ par rapport à $\mu$.

La propriété *presque partout* est naturellement associée aux ensembles négligeables. Une propriété sur $(X, \mathcal{A}, \mu)$ est dite avoir lieu *$\mu$-presque partout* si l'ensemble où elle n'a pas lieu est $\mu$-négligeable. Par exemple, on dira qu'une fonction mesurable $f : (X, \mathcal{A}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ est nulle $\mu$-presque partout si l'ensemble $\{x \in X; f(x) \neq 0\}$ est $\mu$-négligeable (en fait, $f$ n'a pas besoin d'être mesurable).

# 3 Opérations sur les mesures

La définition d'une mesure se prête à des opérations simples. Il est clair que si $\mu$ et $\nu$ sont deux mesures sur un espace mesurable $(X, \mathcal{A})$, $\alpha\mu + \beta\nu$, pour $\alpha$ et $\beta$ des réels positifs ou nuls, est encore une mesure.

Cette opération s'étend à des sommes dénombrables de mesures à travers une propriété de stabilité par convergence monotone (celle-ci est en fait au cœur de la théorie). Si $\mu_n$, $n \in \mathbb{N}$, est une suite de mesures sur $(X, \mathcal{A})$, croissante au sens où $\mu_n(A) \leq \mu_{n+1}(A)$ pour tout $n \in \mathbb{N}$ et tout $A \in \mathcal{A}$, alors l'application

$$A \in \mathcal{A} \to \lim_{n\to\infty} \mu_n(A) \in [0, \infty]$$

est encore une mesure. Pour s'en convaincre, soit $A_k$, $k \in \mathbb{N}$, une famille d'éléments de $\mathcal{A}$ disjoints deux à deux de réunion $A$ ; par définition, et comme les $\mu_n$ sont des mesures,

$$\mu(A) = \lim_{n\to\infty} \mu_n(A) = \lim_{n\to\infty} \sum_{k\in\mathbb{N}} \mu_n(A_k).$$

(Toutes les limites sont à entendre ici à valeurs dans $[0, \infty]$.) Comme les $\mu_n$, $n \in \mathbb{N}$, sont croissantes,

$$\mu(A) = \sup_{n\in\mathbb{N}} \sum_{k\in\mathbb{N}} \mu_n(A_k) \leq \sum_{k\in\mathbb{N}} \sup_{n\in\mathbb{N}} \mu_n(A_k) = \sum_{k\in\mathbb{N}} \mu(A_k).$$

Par ailleurs, pour tout entier $K$,

$$\mu(A) \geq \lim_{n\to\infty} \sum_{k=1}^{K} \mu_n(A_k) = \sum_{k=1}^{K} \lim_{n\to\infty} \mu_n(A_k) = \sum_{k=1}^{K} \mu(A_k)$$

Comme $K$ est arbitraire, $\mu(A) \geq \sum_{k\in\mathbb{N}} \mu(A_k)$, et donc $\mu(A) = \sum_{k\in\mathbb{N}} \mu(A_k)$ avec le point précédent. Il en résulte que $\mu$ est bien une mesure. La mesure de comptage $\sum_{n\in\mathbb{N}} \delta_n$ sur les parties de $\mathbb{N}$ illustre par exemple cette propriété.

En conséquence de cette propriété, si $\mu_n$, $n \in \mathbb{N}$, est une suite de mesures sur $(X, \mathcal{A})$, alors $\sum_{n\in\mathbb{N}} \mu_n$ est encore une mesure (Exercice 2).

D'après ce qui précède, si $\mu$ est une mesure sur $(X, \mathcal{A})$, $\alpha\mu$, $\alpha \geq 0$, est encore une mesure. Si donc $\mu$ est finie de mesure totale $0 < \mu(X) < \infty$, la mesure $\frac{1}{\mu(X)}\,\mu$ est une mesure de probabilité. Cette normalisation est très commode, et sera souvent réalisée.

Si $B \in \mathcal{A}$, une nouvelle mesure notée $\mu_B$ peut être définie sur l'espace mesurable $(B, \mathcal{A} \cap B)$ en posant $\mu_B(C) = \mu(C)$ pour $C \in \mathcal{A} \cap B$. C'est la *mesure $\mu$ restreinte à $B$*. Réciproquement, si $\nu$ est une mesure sur $(B, \mathcal{A} \cap B)$, elle pourra être étendue en une mesure $\mu$ sur $(X, \mathcal{A})$ en posant $\mu(A) = \nu(A \cap B)$, $A \in \mathcal{A}$. (L'extension de $\mu_B$ est la mesure $\mu(A \cap B)$, $A \in \mathcal{A}$.) Ainsi, toute mesure définie sur un sous-ensemble mesurable peut être identifiée à une mesure sur l'espace tout entier ; par exemple une mesure sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ définit une mesure sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (prendre $B = \mathbb{N} \subset \mathbb{R}$).

Si $0 < \mu(B) < \infty$, la restriction précédente $\mu_B$ peut être normalisée en une mesure de probabilité

$$\frac{1}{\mu(B)}\,\mu_B(A) \;=\; \frac{\mu(A \cap B)}{\mu(B)}\,, \quad A \in \mathcal{A}.$$

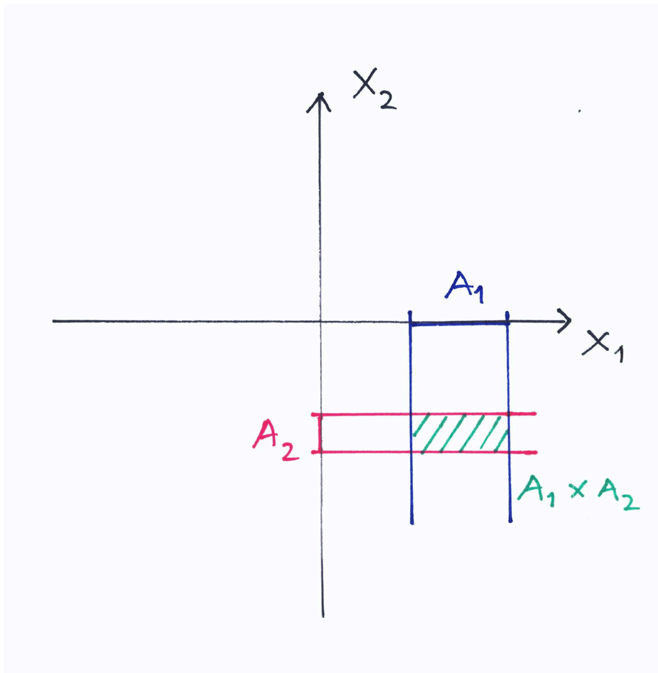
C'est alors la *mesure de probabilité conditionnelle* de $A$ sachant $B$, notée $\mu(A \,|\, B)$, $A \in \mathcal{A}$.

# 4 Mesure produit

Une autre opération sur les mesures est le produit. Si $\mu_1$ et $\mu_2$ sont deux mesures sur des espaces mesurables respectifs $(X_1, \mathcal{A}_1)$ et $(X_2, \mathcal{A}_2)$, la mesure produit (avec la notation tensorielle) $\mu = \mu_1 \otimes \mu_2$ est définie sur la $\sigma$-algèbre produit $\mathcal{A} = \mathcal{A}_1 \otimes \mathcal{A}_2$ sur l'espace produit $X = X_1 \times X_2$ par la règle

$$\mu(A) \;=\; \mu_1 \otimes \mu_2(A_1 \times A_2) \;=\; \mu_1(A_1)\,\mu_2(A_2)$$

pour tout pavé (ou rectangle) $A = A_1 \times A_2$, $A_1 \in \mathcal{A}_1$, $A_2 \in \mathcal{A}_2$.




La description de $\mu_1 \otimes \mu_2$ doit ensuite être étendue à la $\sigma$-algèbre produit $\mathcal{A} = \mathcal{A}_1 \otimes \mathcal{A}_2$ engendrée par les pavés. La construction rigoureuse s'appuie sur le théorème de Carathéodory présenté en complément, et nécessite de supposer les deux mesures $\mu_1$ et $\mu_2$ $\sigma$-finies.

Un exemple simple de mesure produit est le produit de deux copies de la mesure de comptage sur $\mathbb{N}$, produisant la mesure de comptage sur $\mathbb{N}^2 = \mathbb{N} \times \mathbb{N}$, affectant à tout couple d'entiers $(m, n)$ la masse 1.

# 5 Mesure image

La notion de mesure image, ou de mesure « transportée », ou « poussée », par une fonction mesurable est une notion fondamentale en mathématique, et en probabilité en particulier.

**Définition 4** (Mesure image). *Étant donné un espace mesuré $(X, \mathcal{A}, \mu)$, et $f : (X, \mathcal{A}) \to (E, \mathcal{B})$ une fonction mesurable, la mesure image de $\mu$ par $f$ est l'application $\mu_f : \mathcal{B} \to [0, \infty]$ définie par*

$$\mu_f(B) \,=\, \mu\big(f^{-1}(B)\big), \quad B \in \mathcal{B}.$$

La mesure image $\mu_f$ est aussi notée $\mu \circ f^{-1}$, ou encore $f_\# \mu$. C'est un exercice instructif de vérifier que $\mu_f$ est une mesure sur $(E, \mathcal{B})$. La mesure image $\mu_f$ est une mesure sur l'espace d'arrivée $(E, \mathcal{B})$ dont la définition est l'unique possible respectant la propriété de mesurabilité. En effet, la propriété de mesurabilité de $f$ exprime précisément que si $B \in \mathcal{B}$, alors $f^{-1}(B) \in \mathcal{A}$, qui peut ainsi être mesuré par $\mu$.

Deux exemples élémentaires sont déjà importants. Dans ces deux exemples, $\mu$ est une mesure de probabilité sur $(X, \mathcal{A})$ $(\mu(X) = 1)$.

Si $f$ est une fonction constante, i.e. $f(x) = e$ pour un $e \in E$, alors $\mu_f = \delta_e$. Noter qu'il suffirait que $f$ soit constante $\mu$-presque partout.

Si $f = \mathbb{1}_A : (X, \mathcal{A}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ est la fonction indicatrice d'un élément $A \in \mathcal{A}$, alors $\mu_f$ est la combinaison

$$\mu_f \,=\, \mu(A^c)\, \delta_0 + \mu(A)\, \delta_1.$$

En effet, si $B \in \mathcal{B}(\mathbb{R})$ contient 0 sans contenir 1, alors

$$f^{-1}(B) \,=\, \big\{x \in X\,; f(x) \in B\big\} \,=\, \big\{x \in X\,; \mathbb{1}_A(x) = 0\big\} \,=\, A^c,$$

et si $B$ contient 1 sans contenir 0, alors $f^{-1}(B) = A$. Plus simplement, considérer $f = \mathbb{1}_A$ à valeurs dans $\{0, 1\}$ avec la $\sigma$-algèbre des parties constituée de $\emptyset, \{0\}, \{1\}, \{0, 1\}$, de sorte que $\mu_f$ est concentrée sur deux points 0 et 1, et observer que $f^{-1}(\{0\}) = A^c$ et $f^{-1}(\{1\}) = A$.

Ce dernier exemple montre qu'il peut être utile de bien identifier l'image de la fonction mesurable $f$, qui constituera le support de $\mu_f$, même si un espace plus grand, $\mathbb{R}$ par exemple, pourra toujours être considéré (par extension, une mesure sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ est aussi une mesure sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$).

# 6 Construction de mesure

Alors que la théorie de la mesure et de l'intégration de Lebesgue se développe relativement sans peine une fois le cadre et les propriétés fondamentales posées, un point délicat est celui de la construction de mesures répondant aux intuitions naturelles de longueur, surface ou volume. La mesure de comptage, sur un ensemble fini ou dénombrable, ne pose pas de difficulté particulière de définition. Plus généralement, une mesure $\mu$ sur un ensemble fini ou dénombrable $X$, muni de la $\sigma$-algèbre $\mathcal{A} = \mathcal{P}(X)$ des parties de $X$, est une somme pondérée de masses de Dirac en les points de $X$, qui peut se représenter sous la forme

$$\mu \;=\; \sum_{x \in X} \mu\big(\{x\}\big)\, \delta_x.$$

En effet, chaque masse de Dirac est une mesure, et une combinaison linéaire à coefficients positifs, finie ou dénombrable, est encore une mesure. Pour toute partie $A$ de $X$,

$$\mu(A) \;=\; \sum_{x \in A} \mu\big(\{x\}\big),$$

en particulier pour $A = \{x\}$. Une mesure de comptage revient à considérer tous les poids $\mu(\{x\})$ égaux à 1.

Mais une mesure sur les boréliens de $\mathbb{R}$ pour laquelle la mesure d'un intervalle $[a, b]$, $a < b$ (borné) est égale à sa longueur $b - a$ demande à être construite avec plus de précision. Cette mesure, une fois bien définie, s'appelle la *mesure de Lebesgue* sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, notée le plus souvent $\lambda$. C'est bien entendu un exemple central de mesure, même si, une fois construite, son utilisation ne diffère pas des autres mesures.

Ce paragraphe a pour but de décrire, sans entrer dans les détails techniques, les idées conduisant à la mesure de Lebesgue. L'objectif est donc de définir une mesure $\lambda$ sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ telle que si $]a, b]$, $a < b$, est un intervalle $\lambda(]a, b]) = b - a$ (le choix de la forme de l'intervalle est justifié plus loin).

Soit $\mathcal{C}$ l'algèbre (de Boole) constituée des réunions finies d'intervalles, de type quelconque noté pour plus de simplicité $(a,b)$ dans la suite [2], que l'on peut toujours supposées disjointes. Si $A = \bigcup_i (a_i, b_i)$ est un élément de $\mathcal{C}$, il n'y a pas de difficulté à poser

$$\lambda(A) \;=\; \sum_i (b_i - a_i)$$

($\lambda(\emptyset) = 0$). L'application $A \in \mathcal{C} \to \lambda(A)$ ainsi définie est *simplement additive* au sens où si $A$ et $B$ sont deux éléments de $\mathcal{C}$ disjoints, alors $\lambda(A \cup B) = \lambda(A) + \lambda(B)$ (et idem par itération pour toute réunion disjointe finie). Pour faire de $\lambda$ une mesure, il convient d'atteindre la $\sigma$-additivité (sur une $\sigma$-algèbre contenant, et issue, de $\mathcal{C}$). Le théorème de Carathéodory présenté en complément répond à ce souhait, et permet d'étendre la mesure simplement additive $\lambda$ en une mesure $\sigma$-additive sur la $\sigma$-algèbre engendrée par $\mathcal{C}$ qui est par définition la tribu borélienne $\mathcal{B}(\mathbb{R})$. Son application requiert cependant la vérification d'un critère qui fait appel à des arguments de compacité (sur $\mathbb{R}$).

La mesure ainsi construite est la mesure de Lebesgue $\lambda$ sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Si $B$ est un borélien de $\mathbb{R}$, typiquement un intervalle, il sera souvent question dans la suite de la mesure de Lebesgue sur $B$ muni de la tribu trace (restriction), notée également $\lambda$ si le contexte le permet.

Le schéma esquissé ci-dessous s'applique de façon identique à la construction des mesures dite de *Stieltjes* [3]. Si $F : \mathbb{R} \to \mathbb{R}$ une fonction croissante, le développement précédent s'applique de la même manière pour définir une mesure $\lambda^F$ sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ telle que, pour tous $a < b$,

$$\begin{aligned}
\lambda^F\big(]a,b]\big) \;&=\; F(b+) - F(a+),\\
\lambda^F\big([a,b]\big) \;&=\; F(b+) - F(a-),\\
\lambda^F\big([a,b[\big) \;&=\; F(b-) - F(a-),\\
\lambda^F\big(]a,b[\big) \;&=\; F(b-) - F(a+),
\end{aligned}$$

2. Rappeler que la notation $(a,b)$, propre à ces leçons, ne doit pas être confondue avec la notation anglo-saxonne pour l'intervalle ouvert $]a,b[$.

3. Thomas Joannes Stieltjes, mathématicien néerlandais (1856–94).

où $F(t\pm)$ sont les limites à droite et à gauche de $F$ au point $t \in \mathbb{R}$. (Si $F$ est définie sur un intervalle $I$ de $\mathbb{R}$, il convient d'adapter la définition de $F(t\pm)$ aux bords de l'intervalle.)

La mesure de Lebesgue correspond au simple choix de la fonction $F(t) = t$, $t \in \mathbb{R}$. Toutefois, alors que cette dernière ne distingue pas la mesure d'un intervalle $(a, b)$ selon sa forme, en fait $\lambda(\{t\}) = 0$ pour tout $t \in \mathbb{R}$, une mesure de Stieltjes $\lambda^F$ est éventuellement associée à une fonction croissante ayant des points de discontinuité, ou points de saut ; par exemple, si $a$ est un tel point,

$$\lambda^F\big(\{a\}\big) \;=\; \lambda^F\big(]a, a]\big) \;=\; F(a+) - F(a-) \;>\; 0.$$

La mesure de Lebesgue $\lambda$ sur $\mathcal{B}(\mathbb{R})$ se tensorise en une mesure produit

$$\lambda \otimes \cdots \otimes \lambda$$

sur $\mathcal{B}(\mathbb{R}^d)$, $d \geq 1$, notée $\lambda^d$, ou simplement $\lambda$ si le contexte est sans ambiguïté. La mesure de Lebesgue $\lambda^d$ possède deux propriétés remarquables : elle est invariante par translation, à savoir pour tout $x \in \mathbb{R}^d$ et tout borélien $B$, $\lambda(x + B) = \lambda(B)$ (c'est une propriété caractéristique, voir Exercice 8) ; elle est également invariante par transformation orthogonale, c'est-à-dire que, pour toute transformation orthogonale $O : \mathbb{R}^d \to \mathbb{R}^d$ (linéaire, inversible avec $O^{-1} = {}^{\top}O$)) et tout borélien $B$, $\lambda(OB) = \lambda(B)$. Cette propriété est la clef de la formule du changement de variable (cf. Leçon 5).

La complétion de la tribu borélienne $\mathcal{B}(\mathbb{R})$ par rapport à la mesure de Lebesgue est la *tribu de Lebesgue*, notée $\mathcal{L}(\mathbb{R})$ (ou $\mathcal{L}(I)$ sur un intervalle $I$ de $\mathbb{R}$). Même s'il est agréable de couvrir de façon mesurable les ensembles négligeables, la tribu de Lebesgue n'est pas nécessairement toujours la plus commode, par exemple il n'est pas vrai que $\mathcal{L}(\mathbb{R}^2) = \mathcal{L}(\mathbb{R}) \otimes \mathcal{L}(\mathbb{R})$ (ce défaut est lié à l'existence de parties qui ne sont pas mesurables pour la tribu de Lebesgue). Ces leçons ne considèrent que les tribus boréliennes.

# 7 Complément : Théorème de Carathéodory

**Théorème 5** (Théorème de Carathéodory [4]). *Soit $\mu$ une mesure simplement additive sur une algèbre $\mathcal{C}$ de parties d'un ensemble $X$ ; si pour toute suite croissante $A_n$, $n \in \mathbb{N}$, d'éléments de $\mathcal{C}$ telle que $A = \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{C}$, il est vrai que $\mu(A) = \lim_{n\to\infty} \mu(A_n)$, alors $\mu$ se prolonge en une mesure $\sigma$-additive sur la $\sigma$-algèbre $\mathcal{A} = \sigma(\mathcal{C})$ engendrée par $\mathcal{C}$. Si $\mu$ est $\sigma$-finie (sur $\mathcal{C}$), le prolongement est unique (et $\sigma$-fini). De plus, pour tout $A \in \mathcal{A}$ tel que $\mu(A) < \infty$ et tout $\varepsilon > 0$, il existe $B \in \mathcal{C}$ tel que $\mu(A\Delta B) \leq \varepsilon$.*

La démonstration de l'existence dans ce théorème est un peu délicate, et s'appuie sur la notion de *mesure extérieure*. L'unicité découle du théorème des classes monotones de la Leçon 1. Par $\sigma$-additivité, il suffit de supposer $\mu$ finie. Si alors $\mu_1$ et $\mu_2$ sont deux extensions,

$$\mathcal{M} = \big\{A \in \sigma(\mathcal{C})\,;\, \mu_1(A) = \mu_2(A)\big\}$$

est une classe monotone, qui contient $\mathcal{C}$. Donc, par le théorème des classes monotones, $\mathcal{M} = \sigma(\mathcal{C})$, exprimant que $\mu_1 = \mu_2$.

Le critère de l'énoncé est équivalent à la $\sigma$-additivité de $\mu$ sur $\mathcal{C}$. Comme déjà mentionné, l'application de ce théorème à la construction de la mesure de Lebesgue, et plus généralement des mesures de Stieltjes, nécessite précisément la vérification de ce critère. Un peu de topologie est alors nécessaire.

Une famille $\mathcal{F}$ d'un ensemble $X$ est dite semi-compacte si pour toute suite $F_n$, $n \in \mathbb{N}$, d'élements de $\mathcal{F}$ telle que $\bigcap_{n\in\mathbb{N}} F_n = \emptyset$, il existe un entier $N$ tel que $F_1 \cap \cdots \cap F_N = \emptyset$. Il est classique que la famille des parties compactes d'un espace topologique (par exemple $\mathbb{R}$) est semi-compacte.

Dans le cadre de l'énoncé du théorème de Carathéodory, supposer l'existence d'une famille semi-compacte $\mathcal{F}$ contenue dans l'algèbre $\mathcal{C}$ telle que, pour tout

4. Constantin Carathéodory, mathématicien grec (1873–1950).

$A \in \mathcal{C}$,

$$\mu(A) \;=\; \sup\big\{\mu(F)\,;\; F \subset A, F \in \mathcal{F}, \mu(F) < \infty\big\}.$$

Alors il peut être vérifié que $\mu$ est $\sigma$-additive sur $\mathcal{C}$.

Pour l'exemple de la mesure de Lebesgue $\lambda$, ce raisonnement se développe pour la famille $\mathcal{F}$, semi-compacte, des réunions finies d'intervalles fermés bornés. Si $A$ est un intervalle, il est aisé de constater que

$$\lambda(A) \;=\; \sup\big\{\lambda\big([a,b]\big)\,;\; [a,b] \subset A\big\},$$

d'où la propriété précédente découle. Un schéma identique s'applique pour la construction des mesures de Stieltjes associées à une fonction croissante $F$.

## Exercices

(*Une étoile * désignera une question de difficulté supérieure.*)

**Exercice 1.** Soit $\mu$ une mesure sur un espace mesurable $(X, \mathcal{A})$.

a) Soit $B \in \mathcal{A}$ ; démontrer que l'application $A \in \mathcal{A} \mapsto \mu(B \cap A)$ est une mesure sur $(X, \mathcal{A})$.

b) Soient $A, B \in \mathcal{A}$ tels que $A \subset B$ et $\mu(A) = \mu(B) < \infty$. Démontrer que $\mu(A \cap C) = \mu(B \cap C)$ pour tout $C \in \mathcal{A}$.

**Exercice 2.** Soit $\mu_n$, $n \in \mathbb{N}$, une suite de mesures sur un espace mesurable $(X, \mathcal{A})$.

a) Pour tout $A \in \mathcal{A}$, poser $\mu(A) = \sum_{n \in \mathbb{N}} \mu_n(A)$. Démontrer que $\mu$ est une mesure sur $(X, \mathcal{A})$.

b) Il est à présent supposé que $\mu_n(X) = 1$ pour tout $n \in \mathbb{N}$. Soit $p_n$, $n \in \mathbb{N}$, une suite de réels positifs telle que $\sum_{n \in \mathbb{N}} p_n = 1$. Pour tout $A \in \mathcal{A}$, poser $\mu(A) = \sum_{n \in \mathbb{N}} p_n \mu_n(A)$. Vérifier que $\mu$ est une mesure de probabilité sur $(X, \mathcal{A})$.

c) Considérer les mesures $\mu_1 = \sum_{n \geq 1} \delta_n$ et $\mu_2 = \sum_{n \geq 1} n\delta_n$. Pour chacune de ces mesures, calculer la mesure des ensembles : $A_k = [k, k + 1 + \frac{1}{k^2}]$, $k \geq 1$ ; $\bigcup_{k=1}^N A_k$ ; $\bigcup_{k \geq 1} A_k$ ; $\bigcap_{k=1}^N A_k$ ; $\bigcap_{k \geq 1} A_k$ $(N \geq 1)$.

**Exercice 3.** Si $A_n$, $n \in \mathbb{N}$, est une famille de parties mesurables d'un espace mesuré $(X, \mathcal{A}, \mu)$ telle que $\sum_{n \in \mathbb{N}} \mu(A_n) < \infty$, montrer que $\mu(A) = 0$ où $A = \bigcap_{m \in \mathbb{N}} \bigcup_{n \geq m} A_n$ (c'est une partie du *lemme de Borel-Cantelli*[5] qui sera développé dans la Leçon 15).

En déduire que si $f_n$, $n \in \mathbb{N}$, est une suite de fonctions mesurables sur $(X, \mathcal{A}, \mu)$ telle que $\sum_{n \in \mathbb{N}} \mu(|f_n| \geq \varepsilon) < \infty$ pour tout $\varepsilon > 0$, alors $f_n \to 0$ $\mu$-presque partout.

5. Francesco Paolo Cantelli, mathématicien italien (1875–1966).

**Exercice 4** (*Théorème d'Egorov*[6]). Soit $(X, \mathcal{A}, \mu)$ un espace mesuré et soit $B \in \mathcal{A}$ tel que $\mu(B) < \infty$. Soient $f_n$, $n \in \mathbb{N}$, et $f$ des fonctions mesurables sur $(X, \mathcal{A}, \mu)$ telles que $f_n(x) \to f(x)$ pour tout $x \in B$.

a) Pour $\varepsilon > 0$, poser

$$G_n(\varepsilon) = \big\{x \in B\,; \big|f_n(x) - f(x)\big| > \varepsilon\big\}, \quad n \in \mathbb{N},$$

$H_m = \bigcup_{n \geq m} G_n(\varepsilon)$, $H = \bigcap_{m \in \mathbb{N}} H_m$. Identifier $H$ et démontrer que

$$\lim_{n \to \infty} \mu(H_n) \,=\, 0.$$

En déduire que pour tout $\eta > 0$, il existe $G(\eta, \varepsilon) \in \mathcal{A}$ et $m_0 \in \mathbb{N}$ tel que $\mu(G(\eta, \varepsilon)) \leq \eta$ et $|f_n(x) - f(x)| \leq \varepsilon$ pour tout $x \in B \setminus G(\eta, \varepsilon)$ et $n \geq m_0$.

b*) Démontrer que pour tout $\eta > 0$, il existe $G(\eta) \in \mathcal{A}$ tel que $\mu(G(\eta)) \leq \eta$ et $f_n$, $n \in \mathbb{N}$, converge uniformément vers $f$ sur $B \setminus G(\eta)$.

*Remarque.* Il est essentiel dans la démonstration que $\mu(B) < \infty$.

**Exercice 5** (*Support d'une mesure*). Soit $\mu$ une mesure sur un espace métrique $(X, d)$ muni de la tribu des boréliens $\mathcal{B}(X)$ ; on appelle support de $\mu$, noté $\mathrm{Supp}(\mu)$, l'ensemble des $x \in X$ pour lesquels $\mu(B(x, \varepsilon)) > 0$ pour tout $\varepsilon > 0$, où $B(x, \varepsilon)$ est la boule (ouverte) de centre $x$ et de rayon $\varepsilon$. Vérifier que $\mathrm{Supp}(\mu)$ est un fermé de $(X, d)$.

**Exercice 6.** Soit $f : (X, \mathcal{A}) \to (E, \mathcal{B})$ une fonction mesurable, et soit $\mu$ une mesure sur $(X, \mathcal{A})$. Démontrer que l'image $\mu_f$ de $\mu$ par $f$ est une mesure sur $(E, \mathcal{B})$.

Si $X = \mathbb{R}$ muni de la tribu des boréliens, et $f$ est la fonction partie entière, $f(x) = \lfloor x \rfloor$, $x \in \mathbb{R}$, décrire la mesure image de la mesure de Lebesgue $\lambda$ par $f$.

6. Dmitri Fiodorovitch Egorov, mathématicien russe et soviétique (1869–1931).

**Exercice 7.** Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction continue telle que $f = 0$ presque partout (pour la mesure de Lebesgue $\lambda$ sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$) ; démontrer que $f(x) = 0$ pour tout $x \in \mathbb{R}$.

**Exercice 8** (*Unicité de la mesure de Lebesgue*). Cet exercice a pour but de vérifier l'assertion suivante : toute mesure $\mu$ sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ invariante par translation et telle que $0 < \mu(]0,1]^d) < \infty$ est un multiple de la mesure de Lebesgue $\lambda = \lambda^d$ sur $\mathcal{B}(\mathbb{R}^d)$.

a) Soit $I = ]a_1, b_1] \times \cdots \times ]a_d, b_d]$ où les $a_\ell, b_\ell$, $\ell = 1, \ldots, d$, sont rationnels ; montrer qu'il existe $M, k \in \mathbb{N}$ et des points $x^{(j)} \in \mathbb{R}^d$ tels que

$$I = \bigcup_{j=1}^{k} \left( x^{(j)} + \left]0, \tfrac{1}{M}\right]^d \right)$$

où la réunion est disjointe.

b) Par l'invariance par translation de $\lambda$ et $\mu$, en déduire que $\mu(I) = \mu(]0,1]^d)\lambda(I)$. Conclure.

**Exercice 9** (*Ensemble de Cantor*[7]). Poser $E_0 = [0,1]$. Enlever le tiers central (ouvert) de $E_0$ pour obtenir $E_1 = I_1^1 \cup I_1^2$. Répéter la même opération pour $I_1^1$ et $I_1^2$ pour obtenir $E_2 = I_2^1 \cup I_2^2 \cup I_2^3 \cup I_2^4$, et ainsi de suite.

a) Démontrer que chaque $E_n$, $n \in \mathbb{N}$, est compact. En déduire que $C = \bigcap_{n \in \mathbb{N}} E_n$ est non vide et compact (l'ensemble $C$ est appelé *l'ensemble de Cantor*).

b) Montrer que $\lambda(C) = 0$ (où $\lambda$ est la mesure de Lebesgue restreinte à $[0,1]$).

c) Démontrer que $C$ ne contient aucun intervalle ouvert (et donc que l'intérieur de $C$ est vide).

d) Montrer que $C \setminus \{1\}$ comprend les sommes de toutes les séries $x = \sum_{n=1}^{\infty} \frac{x_n}{3^n}$ où $x_n \in \{0, 2\}$. En déduire que $C$ n'est pas dénombrable, en fait a la même cardinalité que $[0,1]$ (pourtant $\lambda(C) = 0 \neq 1 = \lambda([0,1])$).

7. Georg Cantor, mathématicien allemand (1845–1918).