

## Tribus et mesures

24 mars 2011

### 1 Introduction

Nous avons vu que pour construire l'intégrale de Riemann d'une fonction  $f: [a, b] \rightarrow \mathbb{R}$ , on découpe l'**intervalle**  $[a, b]$  **de définition de**  $f$  en sous-intervalles (*subdivisions*), de façon à pouvoir approcher  $f$  par des fonctions en escalier. Pour les fonctions en escalier, la définition de l'intégrale se fait naturellement, et  $f$  est intégrable, au sens de Riemann, si on peut l'encadrer par des fonctions en escalier  $\varphi$  et  $\psi : \varphi \leq f \leq \psi$ , dont la différence des intégrales est aussi petite que l'on veut. Avec cette méthode, il n'est possible d'avoir comme fonctions intégrables que des fonctions qui *n'oscillent pas trop sur les sous-intervalles de*  $[a, b]$ . Si la fonction oscille trop, sur des intervalles de longueur arbitrairement petite, elle ne sera pas Riemann-intégrable; par exemple, la *fonction de Dirichlet*, qui vaut 1 sur les rationnels et 0 sur les irrationnels (en d'autres termes, la *fonction indicatrice* de  $\mathbb{Q}$ ), n'est pas intégrable au sens de Riemann sur  $[a, b]$ . Riemann a d'ailleurs montré (ce n'est pas très difficile) que les fonctions Riemann-intégrables sont exactement les fonctions qui sont continues *presque partout* : l'ensemble des points où  $f$  est continue est, pour tout  $\varepsilon > 0$ , contenu dans la réunion d'une suite d'intervalles dont la somme des longueurs est  $\leq \varepsilon$ .

L'idée de Lebesgue (publiée dans une note aux Comptes-rendus de l'Académie des Sciences en 1902) est de remplacer ce découpage par un découpage de l'**ensemble des valeurs de**  $f$ .

Si  $J$  est un intervalle de  $\mathbb{R}$ , on lui associe :

$$\{x \in [a, b]; f(x) \in J\} = f^{-1}(J).$$

En d'autres termes, on *regroupe* les valeurs de  $f$  qui tombent dans  $J$ , au lieu de simplement regarder ce que l'on obtient au fur et à mesure quand on va de  $a$  à  $b$ . C'est un peu la différence qu'il y a entre Windows (du moins dans les premières versions) et Linux lors de l'enregistrement des données : Windows les enregistre comme elles viennent, alors que Linux les regroupe.

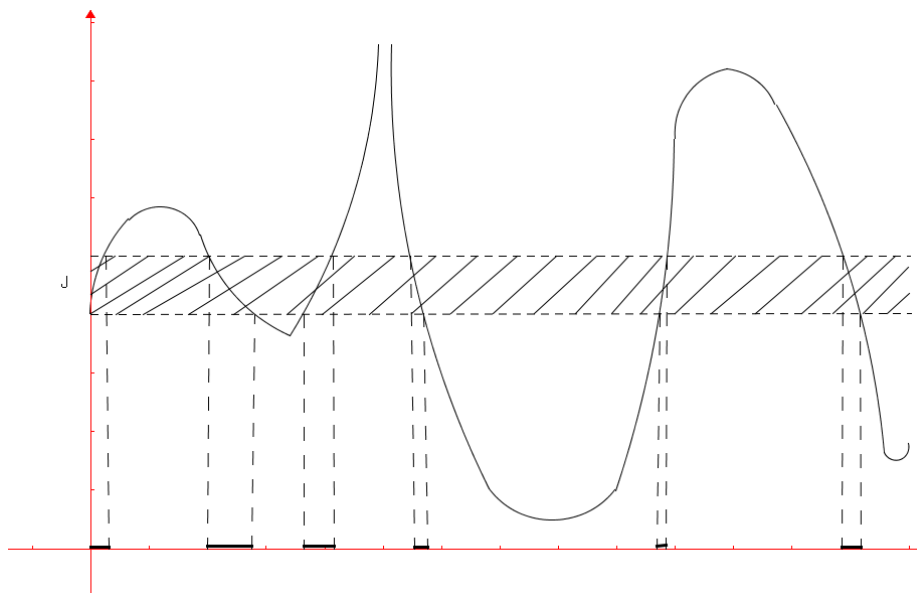


FIGURE 1 – Construction de Lebesgue

Ensuite, si  $(J_n)_{n \geq 1}$  est un découpage de  $\mathbb{R}$  en intervalles deux-à-deux disjoints, de longueur tendant vers 0, et si  $y_n \in J_n$ , on utilisera, comme valeur approchée de l'intégrale que l'on veut définir, la somme des produits de  $y_n$  par la "mesure" de  $f^{-1}(J_n)$ .

L'*avantage* de cette méthode est que l'on obtient ainsi un découpage beaucoup plus fin de  $[a, b]$ , et donc une **meilleure approximation** pour  $f$ ; en effet, *la longueur des intervalles  $J_n$ , que l'on choisit, contrôle l'approximation que l'on aura pour  $f$ .* Il en résulte que l'on pourra intégrer beaucoup plus de fonctions qu'avec l'intégrale de Riemann : les fonctions intégrables n'ont pas besoin de varier peu sur les sous-intervalles de  $[a, b]$ .

L'*inconvenient* est que  $f^{-1}(J_n)$  n'est plus, en général, un intervalle, ni même une réunion dénombrable d'intervalles (cela peut être une partie arbitraire de  $\mathbb{R}$ , en choisissant convenablement  $f$ ), et qu'il faut savoir attribuer une "mesure" à  $f^{-1}(J_n)$ . Or, si l'on admet l'*axiome du choix*, on peut montrer qu'il n'est pas possible d'attribuer une "mesure" (avec des propriétés convenables) à *toutes* les parties de  $[a, b]$ . On est donc amené à devoir préciser *quelles sont les parties auxquelles on pourra attribuer une mesure* : ce sera celles que l'on appellera les *parties mesurables*. Les fonctions  $f$  pour lesquelles  $f^{-1}(J)$  est mesurable pour tout intervalle  $J$  seront appelées les *fonctions mesurables*; c'est avec celles-ci que l'on pourra travailler.

Cet inconvénient est largement compensé par l'*avantage, décisif*, suivant : on n'a plus besoin d'être sur un intervalle  $[a, b]$  de  $\mathbb{R}$  : dès que sur un ensemble  $S$ , on sait préciser les parties mesurables et leur associer une mesure, on pourra inté-

grer des fonctions. Cela donne ainsi une théorie *très générale*, pouvant s'appliquer à un nombre très varié de situations.

## 2 Tribus de parties

### 2.1 Dénombrabilité

Il est essentiel, pour tout ce qui concerne la théorie de la mesure, de savoir distinguer ce qui est dénombrable de ce qui ne l'est pas.

Faisons quelques rappels.

**Définition 2.1** On dit qu'un ensemble  $E$  est dénombrable s'il existe une application surjective  $\sigma: \mathbb{N} \rightarrow E$

$$n \mapsto x_n = \sigma(n).$$

En d'autres termes, on peut écrire  $E = \{x_n; n \in \mathbb{N}\}$  comme les éléments (pas nécessairement tous distincts) d'une suite  $(x_n)_{n \in \mathbb{N}}$ .

**Exemples.**

- Tout ensemble fini est dénombrable.
- $\mathbb{N}, \mathbb{N}^2, \mathbb{Z}, \mathbb{Q}$  sont dénombrables.
- Toute partie d'un ensemble dénombrable est dénombrable.
- Si pour tout entier  $n \geq 1$ , l'ensemble  $A_n$  est dénombrable, alors  $\bigcup_{n \geq 1} A_n$  est dénombrable. Cela reste vrai si l'on remplace la suite d'ensembles  $(A_n)_{n \geq 1}$  par une famille dénombrable  $(A_i)_{i \in I}$ , c'est-à-dire avec  $I$  dénombrable.
- Tout produit d'un nombre fini d'ensembles dénombrables est dénombrable.
- $\mathbb{R}$  n'est pas dénombrable;  $\{0, 1\}^{\mathbb{N}}, \mathbb{N}^{\mathbb{N}}$  ne sont pas dénombrables.

Tout ensemble dénombrable pouvant être indexé par  $\mathbb{N}$ , on conviendra d'appeler *suite* toute famille (de fonctions, de parties d'un ensemble, etc.) indexée par un ensemble dénombrable (bien que l'on ne tiendra pas compte, la plupart du temps, de la relation d'ordre sur  $\mathbb{N}$ ).

### 2.2 Tribus

Commençons par donner une notation très pratique, qui sera constamment utilisée par la suite.

**Notation.** Pour tout ensemble  $S$  et  $A \subseteq S$ , on note

$$A^c = S \setminus A$$

le complémentaire de  $A$  dans  $S$ .

Cela suppose bien sûr qu'il n'y ait pas d'ambiguïté sur l'ensemble "contenant"  $S$ . Dans le cas contraire, on gardera la notation  $S \setminus A$ .

On utilisera aussi la terminologie suivante : si  $S$  est un ensemble et  $\mathcal{C} \subseteq \mathcal{P}(S)$  est un sous-ensemble de parties de  $S$ , on dira que  $\mathcal{C}$  est une *classe de parties* de  $S$ , ou encore une *famille de parties* de  $S$ .

### 2.2.1 Introduction

Avant de donner une définition formelle, essayons de la justifier un peu.

Plaçons-nous sur  $\mathbb{R}$ . Il est naturel de demander que les intervalles soient mesurables. D'autre part, comme on souhaite avoir des théorèmes de convergence pour les intégrales de *suites* de fonctions, il est naturel de demander que la réunion et l'intersection d'une suite de parties mesurables soit encore mesurable. On dit que l'ensemble des parties mesurables est *stable par réunion et intersection dénombrables*.

Considérons alors la *plus petite* classe de parties de  $\mathbb{R}$  contenant les intervalles et stable par réunion et intersection dénombrables. Il y a bien une plus petite telle classe ; en effet, si l'on considère l'ensemble  $\mathfrak{E}$  de toutes les classes de ce type, cet ensemble n'est pas vide : il contient  $\mathcal{P}(\mathbb{R})$ , et l'intersection  $\mathcal{M}_0$  des classes de  $\mathfrak{E}$ , c'est-à-dire l'ensemble des parties  $A \subseteq \mathbb{R}$  telles que  $A \in \mathcal{M}$ , pour tout  $\mathcal{M} \in \mathfrak{E}$ , contient encore les intervalles et est stable par réunion et intersection dénombrables ; donc  $\mathcal{M}_0 \in \mathfrak{E}$ , et c'est évidemment le plus petit élément de  $\mathfrak{E}$ .

$\mathcal{M}_0$  est ce que l'on appellera dans quelques instants la *tribu borélienne* de  $\mathbb{R}$ .

Voyons quelles sont les propriétés de  $\mathcal{M}_0$ .

1)  $\mathcal{M}_0$  contient tous les *ouverts* de  $\mathbb{R}$ . En effet, tout ouvert de  $\mathbb{R}$  est la réunion d'une suite (finie ou infinie) d'intervalles ouverts deux-à-deux disjoints (ses composantes connexes, pour être un peu pédant).

2)  $\mathcal{M}_0$  contient tous les *fermés* de  $\mathbb{R}$ . En effet, si  $F$  est un fermé, l'épaississement  $\Omega_n = \{x \in \mathbb{R} ; \text{dist}(x, F) < 2^{-n}\}$  est un ouvert et  $F = \bigcap_{n \geq 1} \Omega_n$ .

3)  $\mathcal{M}_0$  est, par définition, stable par réunion et intersection dénombrables.

4)  $\mathcal{M}_0$  est stable par passage au complémentaire : si  $A \in \mathcal{M}_0$ , alors  $A^c \in \mathcal{M}_0$ .  
Pour voir cela, considérons :

$$\widetilde{\mathcal{M}}_0 = \{A \subseteq \mathbb{R} ; A^c \in \mathcal{M}_0\}.$$

Alors, les relations :

$$\bigcap_{n \geq 1} A_n = \left( \bigcup_{n \geq 1} A_n^c \right)^c \quad \text{et} \quad \bigcup_{n \geq 1} A_n = \left( \bigcap_{n \geq 1} A_n^c \right)^c$$

montrent que  $\widetilde{\mathcal{M}}_0$  est stable par réunion et intersection dénombrables. D'autre part, puisque le complémentaire d'un ouvert est un fermé et que le complémentaire d'un fermé est un ouvert,  $\widetilde{\mathcal{M}}_0$  contient les ouverts et les fermés. Il contient donc tous les intervalles : pour ceux qui sont ouverts et ceux qui sont fermés, c'est clair ; pour ceux qui sont semi-ouverts, cela vient du fait qu'ils sont intersection d'une suite d'intervalles ouverts, et de la stabilité de  $\widetilde{\mathcal{M}}_0$  par intersection dénombrable. Donc  $\widetilde{\mathcal{M}}_0 \in \mathfrak{E}$ . Comme  $\mathcal{M}_0$  est la plus petite classe de  $\mathfrak{E}$ , on a  $\mathcal{M}_0 \subseteq \widetilde{\mathcal{M}}_0$ . Ainsi, si  $A \in \mathcal{M}_0$ , on a  $A \in \widetilde{\mathcal{M}}_0$ , ce qui, par définition, signifie que  $A^c \in \mathcal{M}_0$ .

On voit ainsi apparaître le fait que  $\mathcal{M}_0$  doit être stable par passage au complémentaire.

## 2.2.2 Définition

Donnons maintenant la définition d'une tribu de parties.

**Définition 2.2** Soit  $S$  un ensemble et  $\mathcal{T} \subseteq \mathcal{P}(S)$  une famille de parties de  $S$ . On dit que  $\mathcal{T}$  est une **tribu de parties** de  $S$  si l'on a les trois propriétés suivantes :

- ( $T_1$ )  $S \in \mathcal{T}$  et  $\emptyset \in \mathcal{T}$  ;
- ( $T_2$ )  $A \in \mathcal{T} \implies A^c \in \mathcal{T}$  : **stabilité par passage au complémentaire** ;
- ( $T_3$ ) si  $A_n \in \mathcal{T}$  pour tout  $n \geq 1$ , alors  $\bigcup_{n \geq 1} A_n \in \mathcal{T}$  : **stabilité par réunion dénombrable**.

On dit aussi que  $\mathcal{T}$  est une  $\sigma$ -algèbre, ou plus exactement  $\sigma$ -algèbre de Boole, de parties de  $S$ .

**Remarques.** 1) Dans la condition ( $T_1$ ), il suffirait, grâce à la condition ( $T_2$ ), de demander l'une des deux appartenances, puisque  $S = \emptyset^c$  et  $\emptyset = S^c$ , mais c'est plus clair ainsi, et cela se vérifiera de toute façon toujours de manière évidente.

2) La condition ( $T_3$ ) de stabilité par réunion dénombrable s'exprime aussi de la façon suivante : Pour tout ensemble d'indices  $J$  **dénombrable**, on a :

$$A_j \in \mathcal{T}, \forall j \in J \quad \implies \quad \bigcup_{j \in J} A_j \in \mathcal{T}.$$

Cette condition entraîne la **stabilité par réunion finie** : pour tout entier  $N \geq 1$ , on a :

$$A_1, \dots, A_N \in \mathcal{T} \quad \implies \quad A_1 \cup \dots \cup A_N \in \mathcal{T}$$

(prendre  $A_n = A_N$  si  $n \geq N$ ).

Il faut faire attention que si  $J$  est un ensemble d'indices *non* dénombrable,  $\bigcup_{j \in J} A_j$  n'a aucune raison d'être dans  $\mathcal{T}$  si les  $A_j \in \mathcal{T}$ .

3) On n'a pas demandé la stabilité par intersection dénombrable, car elle se déduit de la définition :

**Proposition 2.3** Soit  $\mathcal{T}$  une tribu de parties de  $S$ . Alors  $\mathcal{T}$  est **stable par intersection dénombrable** :

$$A_n \in \mathcal{T}, \forall n \geq 1 \quad \implies \quad \bigcap_{n \geq 1} A_n \in \mathcal{T}.$$

**Preuve.** Si  $A_n \in \mathcal{T}$ ,  $n \geq 1$ , alors  $A_n^c \in \mathcal{T}$ , par ( $T_2$ ) ; donc  $\bigcup_{n \geq 1} A_n^c \in \mathcal{T}$ , par ( $T_3$ ). Alors, de nouveau par ( $T_2$ ) :

$$\bigcap_{n \geq 1} A_n = \left( \bigcup_{n \geq 1} A_n^c \right)^c \in \mathcal{T}. \quad \square$$

**Remarque.** Comme dans le cas de la réunion, que cela entraîne la **stabilité par intersection finie** : pour tout entier  $N \geq 1$ , on a

$$A_1, \dots, A_N \in \mathcal{T} \quad \implies \quad A_1 \cap \dots \cap A_N \in \mathcal{T}.$$

On notera que l'on aurait pu prendre la stabilité par intersection dénombrable dans la définition (au lieu de  $(T_3)$ ), et on en aurait déduit, grâce à  $(T_2)$ , la stabilité par réunion dénombrable.

**Définition 2.4** *Étant donnée une tribu  $\mathcal{T}$  de parties sur un ensemble  $S$ , les éléments  $A \in \mathcal{T}$  sont appelées les **parties mesurables** (relativement à  $\mathcal{T}$ ).*

**Définition 2.5** *On appelle **espace mesurable** le couple  $(S, \mathcal{T})$  formé d'un ensemble  $S$  et d'une tribu  $\mathcal{T}$  de parties de  $S$ .*

### 2.2.3 Exemples de tribus

Les familles suivantes sont des tribus de parties de l'ensemble  $S$  :

1.  $\mathcal{T} = \mathcal{P}(S)$ , l'ensemble de *toutes* les parties de  $S$ ; on l'appelle la *tribu discrète*;
2.  $\mathcal{T} = \{\emptyset, S\}$ , formée uniquement de l'ensemble vide et de l'ensemble total  $S$ ; on l'appelle la *tribu grossière*;
3. Soit  $(S_n)_{n \geq 1}$  une *partition dénombrable* de  $S$  (pouvant être finie) :

$$\begin{cases} S_n \cap S_p = \emptyset & \text{si } n \neq p, \\ S = \bigcup_{n \geq 1} S_n. \end{cases}$$

Alors

$$\mathcal{T} = \left\{ \bigcup_{n \in I} S_n ; I \subseteq \mathbb{N}^* \right\}$$

est une tribu de parties de  $S$ , dite *engendrée par la partition  $(S_n)_{n \geq 1}$* , comme on le vérifiera à titre d'exercice.

Sur un ensemble  $S$  dénombrable, toute tribu est de ce type.

**Remarque importante.** En fait ce sont essentiellement les seuls cas de tribus dont on peut décrire *tous* les éléments. En général, une tribu contient énormément d'éléments (d'où le nom de *tribu*), que l'on ne peut donc pas tous décrire. **Cela n'a pas d'importance** ; ce qu'il faut retenir, c'est que :

On a juste besoin de savoir que certaines parties de  $S$  sont dans la tribu  $\mathcal{T}$ , et que l'on reste dans  $\mathcal{T}$  quand on itère sur les éléments de  $\mathcal{T}$  les opérations de complémentation et de réunion et intersection dénombrables.

Cela va être notamment le cas pour la tribu borélienne de  $\mathbb{R}$ .

### 2.2.4 Tribu borélienne

Soit  $(X, d)$  un espace métrique (ou plus généralement  $X$  un espace topologique). Il est alors naturel de définir les parties mesurables en tenant compte de la topologie de  $X$ , et notamment de demander que les ouverts et les fermés soient mesurables.

Rappelons d'abord que l'ensemble  $\mathcal{O}_X$  des *ouverts* de  $X$  a les propriétés suivantes :

(O<sub>1</sub>)  $X \in \mathcal{O}_X$  et  $\emptyset \in \mathcal{O}_X$  ;

(O<sub>2</sub>) (stabilité par **réunion arbitraire**) pour **tout** ensemble d'indices  $J$  :

$$\Omega_j \in \mathcal{O}_X, \forall j \in J \quad \Longrightarrow \quad \bigcup_{j \in J} \Omega_j \in \mathcal{O}_X;$$

(O<sub>3</sub>) (stabilité par **intersection finie**) pour tout entier  $N \geq 1$  :

$$\Omega_1, \dots, \Omega_N \in \mathcal{O}_X \quad \Longrightarrow \quad \Omega_1 \cap \dots \cap \Omega_N \in \mathcal{O}_X.$$

La famille de parties  $\mathcal{O}_X$  n'est (sauf cas très particuliers) *pas* une tribu, car elle n'est

\* *pas stable par intersection dénombrable* : par exemple, dans  $\mathbb{R}$ , l'intervalle semi-ouvert  $[0, 1[$ , qui n'est pas ouvert (et pas fermé non plus), est l'intersection de la suite d'intervalles ouverts  $] -\frac{1}{n}, 1[$  ;

\* *pas stable par complémentation* : le complémentaire d'un ouvert, c'est-à-dire un fermé, n'est en général pas ouvert.

On est donc amené à définir les ensembles mesurables comme éléments d'une tribu *contenant*  $\mathcal{O}_X$ . Il est naturel de prendre la plus petite possible. L'existence d'une plus petite tribu est donnée par la proposition suivante.

**Proposition 2.6** *Soit  $S$  un ensemble et  $\mathcal{C} \subseteq \mathcal{P}(S)$  une famille de parties de  $S$ . Alors il existe une plus petite tribu contenant  $\mathcal{C}$ . On dit que c'est la tribu engendrée par  $\mathcal{C}$  ; on la note  $\sigma(\mathcal{C})$ .*

**Preuve.** Cela résulte des deux observations suivantes, déjà faites dans l'Introduction, § 2.2.1, que l'on vérifie facilement.

1) Si  $(\mathcal{T}_j)_{j \in J}$  est une famille arbitraire de tribus de parties de  $S$ , alors  $\mathcal{T} = \bigcap_{j \in J} \mathcal{T}_j$  est encore une tribu de parties de  $S$ .

2) Si maintenant  $(\mathcal{T}_j)_{j \in J}$  est la famille de *toutes* les tribus de parties de  $S$  qui contiennent  $\mathcal{C}$  (il y en a au moins une :  $\mathcal{P}(S)$ ), alors la tribu  $\mathcal{T} = \bigcap_{j \in J} \mathcal{T}_j$  contient  $\mathcal{C}$ . C'est la plus petite car toute tribu contenant  $\mathcal{C}$  est l'une des  $\mathcal{T}_j$ , donc contient  $\mathcal{T}$ , puisque  $\mathcal{T} \subseteq \mathcal{T}_j$ .  $\square$

**Définition 2.7** Si  $X$  est un espace topologique, on appelle **tribu borélienne** de  $X$  la tribu  $\sigma(\mathcal{O}_X)$  engendrée par l'ensemble  $\mathcal{O}_X$  des ouverts de  $X$ . On la note  $\mathcal{Bor}(X)$ .

Les parties mesurables pour la tribu borélienne de  $X$  s'appellent les **boréliens**, ou parties boréliennes de  $X$ .

Par définition, on a donc :  $\mathcal{Bor}(X) = \sigma(\mathcal{O}_X)$ .

Comme les tribus sont stables par passage au complémentaire et que les fermés sont les complémentaires des ouverts, **tout fermé est borélien**. En fait :

**Proposition 2.8** La tribu borélienne de  $X$  est la plus petite tribu contenant tous les fermés de  $X$ .

**Preuve.** Soit  $\mathcal{T}$  cette tribu. Comme  $\mathcal{Bor}(X)$  est une tribu contenant les fermés, on a  $\mathcal{T} \subseteq \mathcal{Bor}(X)$ . Inversement, la tribu  $\mathcal{T}$  contient les complémentaires des fermés, c'est-à-dire les ouverts ; elle contient donc la tribu  $\mathcal{Bor}(X)$  engendrée par les ouverts.  $\square$

**Remarque.** La tribu borélienne va contenir, en général, énormément de parties, que l'on ne peut pas, en général, toutes décrire. En particulier, elle contient :

- les intersections d'un ouvert et d'un fermé ; par exemple, dans  $\mathbb{R}$ , les intervalles semi-ouverts ;
- Dans  $\mathbb{R}$ , tout intervalle est un borélien de  $\mathbb{R}$  ;
- les intersections dénombrables d'ouverts, que l'on appelle les  $G_\delta$  ;
- les réunions dénombrables de fermés, que l'on appelle les  $F_\sigma$  ;
- les réunions dénombrables de  $G_\delta$ , appelées  $G_{\delta\sigma}$ , les intersections dénombrables de  $F_\sigma$ , appelées  $F_{\sigma\delta}$ , etc. On peut continuer ainsi indéfiniment, et même *transfiniment*, à condition de ne le faire que dénombrablement.

Toutefois, la tribu borélienne de  $\mathbb{R}$  n'est pas si énorme que l'on pourrait le croire : on peut montrer que  $\text{card } \mathcal{Bor}(\mathbb{R}) = \text{card } \mathbb{R}$  ; donc  $\mathcal{Bor}(\mathbb{R})$  a bien moins d'éléments que  $\mathcal{P}(\mathbb{R})$ .

Il est à noter que :

**Proposition 2.9** Si  $X$  est un espace topologique séparé, en particulier si  $X$  est un espace métrique, alors toute partie dénombrable de  $X$  est borélienne. En particulier,  $\mathbb{N}$  et  $\mathbb{Z}$  sont des boréliens de  $\mathbb{R}$ , et  $\mathbb{Q}$  est un borélien de  $\mathbb{R}$ .

**Preuve.**  $X$  étant séparé, les singletons  $\{x\}$  sont fermés. Or pour toute partie dénombrable  $D$  de  $X$ , on peut écrire  $D = \bigcup_{x \in D} \{x\}$ .  $\square$

Pour terminer, on peut remarquer que la définition de la tribu borélienne de  $\mathbb{R}^N$  est faite à partir des ouverts de  $\mathbb{R}^N$ . Mais on ne connaît pas non plus tous les ouverts de  $\mathbb{R}^N$ . Il est donc important de pouvoir obtenir la tribu borélienne comme celle engendrée par une classe d'ensembles, connus, plus petite. C'est l'objet du résultat suivant.

**Proposition 2.10** *La tribu borélienne de  $\mathbb{R}^N$  est engendrée par l'ensemble des pavés ouverts  $\prod_{1 \leq k \leq N} ]a_k, b_k[$ , lorsque  $a_k, b_k$  parcourent  $\mathbb{R}$ , avec  $a_k < b_k$  (intervalles ouverts  $]a, b[$ , avec  $a, b \in \mathbb{R}$ ,  $a < b$ , dans le cas  $N = 1$ ).*

**Preuve.** Soit  $\mathcal{T}$  cette tribu. Comme les pavés ouverts sont des ouverts, ils sont boréliens; donc  $\mathcal{T} \subseteq \mathcal{Bor}(\mathbb{R}^N)$ .

Inversement, tout ouvert  $\Omega$  de  $\mathbb{R}^N$  est une réunion *dénombrable* de tels pavés; on peut même prendre des pavés ouverts avec  $a_k, b_k \in \mathbb{Q}$ ; donc  $\Omega \in \mathcal{T}$ . On en déduit que  $\mathcal{Bor}(\mathbb{R}^N) \subseteq \mathcal{T}$ , d'où l'égalité.  $\square$

Dans ce cours, un espace topologique sera toujours muni de sa tribu borélienne.

Ce sera en particulier le cas pour  $\mathbb{R}$ .

## 3 Applications mesurables

### 3.1 Définitions

**Définition 3.1** *Soit  $(S, \mathcal{T})$  et  $(S', \mathcal{T}')$  deux espaces mesurables. On dit qu'une application  $f: S \rightarrow S'$  est une application mesurable si :*

$$f^{-1}(A') \in \mathcal{T}, \quad \forall A' \in \mathcal{T}'.$$

**Définition 3.2** *Si  $X$  et  $X'$  sont deux espaces topologiques, on dit que l'application  $f: X \rightarrow X'$  est une application borélienne si elle est mesurable lorsque l'on munit  $X$  et  $X'$  de leur tribu borélienne.*

**Définition 3.3** *Soit  $S$  un ensemble. Pour toute partie  $A \subseteq S$ , on note  $\mathbb{1}_A$  l'application  $\mathbb{1}_A: S \rightarrow \mathbb{R}$  définie par*

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

On l'appelle la **fonction indicatrice** de  $A$

On trouve souvent la terminologie *fonction caractéristique* de  $A$ . Il vaut mieux utiliser le terme *fonction indicatrice*, même s'il n'y a en fait pas de risque de confusion, car la fonction caractéristique  $\mathbb{1}_a$ , en Probabilités, a une toute autre signification.

**Proposition 3.4** *La fonction indicatrice  $\mathbb{1}_A$  de  $A$  est mesurable si et seulement si  $A$  est une partie mesurable.*

**Preuve.** 1) Supposons  $A$  mesurable. Pour tout borélien  $B$  de  $\mathbb{R}$ , on a :

$$(\mathbb{1}_A)^{-1}(B) = \begin{cases} \emptyset & \text{si } 0 \notin B \text{ et } 1 \notin B; \\ A & \text{si } 0 \notin B \text{ et } 1 \in B; \\ A^c & \text{si } 0 \in B \text{ et } 1 \notin B; \\ S & \text{si } 0 \in B \text{ et } 1 \in B. \end{cases}$$

Dans tous les cas  $(\mathbb{1}_A)^{-1}(B) \in \mathcal{T}$ , et donc  $\mathbb{1}_A$  est mesurable

2) Supposons  $\mathbb{1}_A$  mesurable. Alors  $A = (\mathbb{1}_A)^{-1}(\{1\}) \in \mathcal{T}$ .  $\square$

### 3.2 Critères de mesurabilité

En général, la tribu  $\mathcal{T}'$  étant très grosse, la mesurabilité d'une fonction serait pratiquement impossible si l'on ne disposait pas d'un critère permettant de ne tester la fonction que sur une sous-classe de la tribu  $\mathcal{T}'$ . Le résultat suivant est donc d'une **importance capitale**.

**Théorème 3.5** *Soit  $(S, \mathcal{T})$  et  $(S', \mathcal{T}')$  deux espaces mesurables.*

*Soit  $f: S \rightarrow S'$  une application. S'il existe  $\mathcal{C}' \subseteq \mathcal{P}(S')$  telle que :*

*a)  $\sigma(\mathcal{C}') = \mathcal{T}'$  ;*

*b)  $f^{-1}(C') \in \mathcal{T}$  pour tout  $C' \in \mathcal{C}'$ ,*

*l'application  $f$  est mesurable.*

Pour prouver ce théorème, on utilisera plusieurs lemmes.

Auparavant, donnons la notation suivante.

**Notation.** Si  $f: S \rightarrow S'$  est une application, on note, pour toute famille de parties  $\mathcal{C}' \subseteq \mathcal{P}(S')$  :

$$f^{-1}(\mathcal{C}') = \{f^{-1}(C') ; C' \in \mathcal{C}'\}.$$

La condition b) du théorème s'écrit alors :

$$f^{-1}(\mathcal{C}') \subseteq \mathcal{T}.$$

Le premier lemme dit que l'image réciproque d'une tribu par une application est encore une tribu dans l'espace de départ.

**Lemme 3.6** *Si  $f: S \rightarrow S'$  est une application et si  $\mathcal{T}'$  est une tribu de parties de  $S'$ , alors  $f^{-1}(\mathcal{T}')$  est une tribu de parties de  $S$ .*

On notera que  $f$  est alors mesurable quand on munit  $S$  de la tribu  $f^{-1}(\mathcal{T}')$  et  $S'$  de  $\mathcal{T}'$ . Il est de plus clair que  $f^{-1}(\mathcal{T}')$  est la *plus petite* tribu de parties de  $S$  pour laquelle  $f$  soit mesurable, lorsque  $S'$  est muni de  $\mathcal{T}'$ .

Rappelons d'abord :

**Rappel.** Pour toute application  $f: S \rightarrow S'$ , on a, pour tout  $A' \subseteq S'$  et pour toute famille  $(A'_i)_{i \in I}$  de parties de  $S'$  :

$$\begin{cases} f^{-1}\left(\bigcup_{i \in I} A'_i\right) = \bigcup_{i \in I} f^{-1}(A'_i) \\ f^{-1}\left(\bigcap_{i \in I} A'_i\right) = \bigcap_{i \in I} f^{-1}(A'_i) \\ f^{-1}(S' \setminus A') = S \setminus f^{-1}(A') \\ f^{-1}(\emptyset) = \emptyset \quad \text{et} \quad f^{-1}(S') = S. \end{cases}$$

La preuve du lemme s'en déduit aussitôt.  $\square$

Par contre, l'image directe d'une tribu par une application n'est pas, en général, une tribu. Pour avoir une tribu dans l'espace d'arrivée, il faut faire une petite acrobatie.

**Lemme 3.7** Soit  $f: S \rightarrow S'$  une application, et  $\mathcal{T}$  une tribu de parties de  $S$ . Alors

$$\boxed{\mathcal{T}_f = \{A' \subseteq S'; f^{-1}(A') \in \mathcal{T}\}}$$

est une tribu de parties de  $S'$ .

C'est la plus grande tribu de parties de  $S'$  pour laquelle  $f$  soit mesurable quand on muni  $S$  de  $\mathcal{T}$ . On dit que c'est la tribu engendrée par  $f$ ; on la note aussi  $\sigma(f)$ .

**Preuve.** (T<sub>1</sub>) On a  $f^{-1}(\emptyset) = \emptyset \in \mathcal{T}$  et  $f^{-1}(S') = S \in \mathcal{T}$ ; donc  $\emptyset \in \mathcal{T}_f$  et  $S \in \mathcal{T}_f$ .

(T<sub>2</sub>) Si  $A' \in \mathcal{T}_f$ , alors  $f^{-1}(A') \in \mathcal{T}$ ; donc  $f^{-1}(S' \setminus A') = S \setminus f^{-1}(A') \in \mathcal{T}$ , ce qui signifie que  $f^{-1}(S' \setminus A') \in \mathcal{T}_f$ .

(T<sub>3</sub>) Si, pour tout  $n \geq 1$ , on a  $A'_n \in \mathcal{T}_f$ , c'est que l'on a, pour tout  $n \geq 1$ ,  $f^{-1}(A'_n) \in \mathcal{T}$ ; donc  $f^{-1}(\bigcup_{n \geq 1} A'_n) = \bigcup_{n \geq 1} f^{-1}(A'_n) \in \mathcal{T}$ , c'est-à-dire que  $\bigcup_{n \geq 1} A'_n \in \mathcal{T}_f$ .

Que ce soit la plus grande tribu de parties de  $S'$  pour laquelle  $f$  soit mesurable est clair.  $\square$

Le dernier lemme est très important et servira plusieurs fois par la suite.

**Lemme 3.8 (LEMME de TRANSFERT)** Soit  $f: S \rightarrow S'$  une application. Pour toute classe  $\mathcal{C}' \subseteq \mathcal{P}(S')$ , on a :

$$\boxed{\sigma[f^{-1}(\mathcal{C}')] = f^{-1}[\sigma(\mathcal{C}')]}.$$

**Preuve.** 1) Par le Lemme 3.6,  $f^{-1}[\sigma(\mathcal{C}')]$  est une tribu de parties de  $S$ . Comme elle contient  $f^{-1}(\mathcal{C}')$ , elle contient aussi  $\sigma[f^{-1}(\mathcal{C}')]$ .

2) Réciproquement, utilisons le Lemme 3.7, avec  $\mathcal{T} = \sigma[f^{-1}(\mathcal{C}')]$  : comme :

$$A' \in \mathcal{C}' \quad \implies \quad f^{-1}(A') \in \sigma[f^{-1}(\mathcal{C}')],$$

on a  $\mathcal{C}' \subseteq \mathcal{T}_f$ . Donc  $\sigma(\mathcal{C}') \subseteq \mathcal{T}_f$ ; cela s'écrit :

$$A' \in \sigma(\mathcal{C}') \quad \implies \quad f^{-1}(A') \in \mathcal{T} = \sigma[f^{-1}(\mathcal{C}')],$$

c'est-à-dire  $f^{-1}[\sigma(\mathcal{C}')] \subseteq \sigma[f^{-1}(\mathcal{C}')]$ . □

**Preuve du Théorème 3.5.** Les hypothèses s'écrivent :

$$\begin{cases} a) \sigma(\mathcal{C}') = \mathcal{T}' \\ b) f^{-1}(\mathcal{C}') \subseteq \mathcal{T}; \end{cases}$$

le Lemme de transfert donne donc :

$$f^{-1}(\mathcal{T}') = f^{-1}[\sigma(\mathcal{C}')] = \sigma[f^{-1}(\mathcal{C}')] \subseteq \sigma(\mathcal{T}) = \mathcal{T},$$

ce qui signifie que  $f$  est mesurable. □

Une conséquence **très importante** est la suivante.

**Théorème 3.9** *Si  $X$  et  $X'$  sont deux espaces topologiques, toute application continue  $f: X \rightarrow X'$  est borélienne.*

Ce résultat donne une classe très vaste d'applications mesurables. Toutes celles que l'on rencontre en pratique sont de ce type, ou s'en déduisent.

**Preuve.** Soit  $\mathcal{O}_X$  et  $\mathcal{O}_{X'}$ , l'ensemble des ouverts de  $X$  et  $X'$ , respectivement. Dire que  $f$  est continue s'exprime par :

$$f^{-1}(\mathcal{O}_{X'}) \subseteq \mathcal{O}_X.$$

On a donc, en particulier  $f^{-1}(\mathcal{O}_{X'}) \subseteq \text{Bor}(X)$ ; il suffit alors d'appliquer le Théorème 3.5, avec  $\mathcal{T} = \text{Bor}(X)$ ,  $\mathcal{T}' = \text{Bor}(X')$ , et  $\mathcal{C}' = \mathcal{O}_{X'}$ , puisque  $\sigma(\mathcal{O}_{X'}) = \text{Bor}(X')$ . □

Nous avons vu (Proposition 2.10) que la tribu borélienne de  $\mathbb{R}^N$  est engendrée par les pavés ouverts; on obtient donc :

**Corollaire 3.10** *Soit  $(S, \mathcal{T})$  un espace mesurable; l'application  $f: S \rightarrow \mathbb{R}^N$  est mesurable si, et seulement si,*

$$f^{-1}\left(\prod_{k=1}^N ]a_k, b_k[ \right) \in \mathcal{T}, \quad \forall a_k < b_k.$$

### 3.3 Sous-espaces

**Proposition 3.11** Soit  $(S, \mathcal{T})$  un espace mesurable et  $R \subseteq S$ . Alors

$$\mathcal{T}_R = \{A \cap R; A \in \mathcal{T}\}$$

est une tribu de parties de  $R$ .

On l'appelle la tribu induite sur  $R$  par  $\mathcal{T}$ , ou tribu-trace de  $\mathcal{T}$  sur  $R$ .

De plus, lorsque  $R \in \mathcal{T}$ , on a :

$$\mathcal{T}_R = \{A \in \mathcal{T}; A \subseteq R\} = \{A \subseteq R; A \in \mathcal{T}\}.$$

**Preuve.** C'est immédiat. □

**Remarque.** Soit  $j: R \rightarrow S$  l'injection canonique. Pour tout  $A \subseteq S$ , on a  $j^{-1}(A) = A \cap R$ . Il résulte donc de la définition de  $\mathcal{T}_R$  que  $j$  est *mesurable*, quand on munit  $R$  de la tribu trace  $\mathcal{T}_R$ , et que  $\mathcal{T}_R$  est la plus petite tribu de parties de  $R$  pour laquelle  $j$  soit mesurable.

Il est important d'avoir le résultat de compatibilité suivant.

**Théorème 3.12** Soit  $X$  un espace topologique, et  $Y \subseteq X$ , muni de la topologie induite. Alors :

$$\boxed{\mathcal{B}or(Y) = [\mathcal{B}or(X)]_Y}.$$

En d'autres termes :

$$\boxed{\text{les boréliens de } Y \text{ sont les traces sur } Y \text{ des boréliens de } X}.$$

**Preuve.** 1) Par définition, l'ensemble des ouverts de  $Y$  est :

$$\mathcal{O}_Y = \{\Omega \cap Y; \Omega \in \mathcal{O}_X\};$$

donc  $\mathcal{O}_Y \subseteq [\mathcal{B}or(X)]_Y$ . Il en résulte que  $\mathcal{B}or(Y) \subseteq [\mathcal{B}or(X)]_Y$ .

2) Réciproquement, soit  $j: Y \rightarrow X$  l'injection canonique. Elle est continue :

$$j^{-1}(\Omega) = Y \cap \Omega \in \mathcal{O}_Y, \quad \forall \Omega \in \mathcal{O}_X;$$

donc, par le Théorème 3.9, elle est borélienne :

$$j^{-1}(\mathcal{B}or(X)) \subseteq \mathcal{B}or(Y).$$

Cela s'exprime par :

$$A \cap Y = j^{-1}(A) \in \mathcal{B}or(Y), \quad \forall A \in \mathcal{B}or(X),$$

soit  $[\mathcal{B}or(X)]_Y \subseteq \mathcal{B}or(Y)$ . □

### 3.4 Mesurabilité des applications à valeurs réelles

Lorsque l'on prend des limites, ou des bornes supérieures ou inférieures, on aboutit souvent à des valeurs  $-\infty$  ou  $+\infty$ . Il est donc naturel de s'intéresser, dès le départ, à des fonctions à valeurs, non pas seulement dans  $\mathbb{R}$ , mais dans  $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ . On appellera donc *fonction à valeurs réelles* toute fonction prenant ses valeurs dans  $\overline{\mathbb{R}}$ .

**Rappel : Topologie de  $\overline{\mathbb{R}}$ .** La topologie induite sur  $\overline{\mathbb{R}}$  est la topologie usuelle de  $\mathbb{R}$ . Il suffit de préciser quels sont les voisinages de  $-\infty$  et  $+\infty$ . On dira que  $V$  est un voisinage de  $+\infty$  s'il existe  $A > 0$  ( $A \in \mathbb{R}_+^*$ ) tel que  $]A, +\infty[ \subseteq V$ . De même pour  $-\infty$ .

Par le Théorème 3.12, si  $B \subseteq \overline{\mathbb{R}}$  est un borélien de  $\overline{\mathbb{R}}$ , alors  $B_0 = B \cap \mathbb{R}$  est un borélien de  $\mathbb{R}$ . En fait, comme on n'ajoute que deux points à  $\mathbb{R}$  pour obtenir  $\overline{\mathbb{R}}$ , on a la situation simple suivante.

**Proposition 3.13** *L'ensemble  $B \subseteq \overline{\mathbb{R}}$  est un borélien de  $\overline{\mathbb{R}}$  si et seulement si il existe un borélien  $B_0$  de  $\mathbb{R}$  tel que :*

$$B = B_0 \text{ ou } B = B_0 \cup \{-\infty\} \text{ ou } B = B_0 \cup \{+\infty\} \text{ ou } B = B_0 \cup \{-\infty, +\infty\}.$$

**Corollaire 3.14** *L'ensemble  $\mathcal{I}_0$  de tous les intervalles de  $\overline{\mathbb{R}}$  du type :*

$$] \alpha, \beta[, \quad [-\infty, \beta[, \quad ] \alpha, +\infty[, \quad \text{pour } -\infty < \alpha < \beta < +\infty,$$

*engendre la tribu borélienne de  $\overline{\mathbb{R}}$  :  $\sigma(\mathcal{I}_0) = \text{Bor}(\overline{\mathbb{R}})$ .*

**Preuve.** L'ensemble de tous les intervalles ouverts  $] \alpha, \beta[$ , pour  $-\infty < \alpha < \beta < +\infty$ , engendre la tribu borélienne de  $\mathbb{R}$  (Proposition 2.10). Donc  $\sigma(\mathcal{I}_0)$  contient  $\text{Bor}(\mathbb{R})$ . D'autre part,  $\sigma(\mathcal{I}_0)$  contient  $\{-\infty\} = \bigcap_{n \geq 1} [-\infty, -n[$  et  $\{+\infty\} = \bigcap_{n \geq 1} ]n, +\infty[$ . La Proposition 3.13 donne le résultat.  $\square$

Le critère de mesurabilité suivant est très pratique.

**Théorème 3.15** *Soit  $(S, \mathcal{F})$  un espace mesurable et  $f: S \rightarrow \overline{\mathbb{R}}$  une application à valeurs réelles. Alors  $f$  est mesurable si et seulement si :*

$$\{f < a\} \in \mathcal{F} \quad \forall a \in \mathbb{R}.$$

Ce théorème s'applique bien sûr en particulier au cas des fonctions  $f: S \rightarrow \mathbb{R}$  à valeurs dans  $\mathbb{R}$ . On a utilisé la notation suivante, très pratique (qui vient, historiquement, de l'usage en Probabilités).

**Notation.**

$$\{f < a\} = \{x \in S; f(x) < a\}.$$

Autrement dit :  $\{f < a\} = f^{-1}([-\infty, a[)$ .

Plus généralement, si  $f: S \rightarrow S'$ , on note, pour toute partie  $A' \subseteq S'$  :

$$\boxed{\{f \in A'\} = \{x \in S; f(x) \in A'\}},$$

c'est-à-dire  $\{f \in A'\} = f^{-1}(A')$ .

On peut remplacer les intervalles ouverts par des intervalles fermés.

**Corollaire 3.16** *L'application  $f: S \rightarrow \overline{\mathbb{R}}$  est mesurable si et seulement si :*

$$\{f \leq a\} \in \mathcal{F} \quad \forall a \in \mathbb{R}.$$

**Preuve.** On pourrait faire la même preuve que celle qui sera faite ci-dessous du Théorème 3.15, mais il suffit de remarquer que :

$$\{f < a\} = \bigcup_{n \geq 1} \{f \leq a - \frac{1}{n}\}$$

pour que obtenir la condition est suffisante. Qu'elle soit nécessaire est évident, puisque  $[-\infty, a]$  est un fermé de  $\overline{\mathbb{R}}$ .  $\square$

**Preuve du Théorème 3.15.** 1) La condition est évidemment nécessaire car  $[-\infty, a[$  est un ouvert de  $\overline{\mathbb{R}}$ .

2) Réciproquement, soit :

$$\mathcal{I} = \{[-\infty, a[; a \in \mathbb{R}\}.$$

L'hypothèse s'écrit :  $f^{-1}(\mathcal{I}) \subseteq \mathcal{F}$ .

Considérons l'ensemble d'intervalles  $\mathcal{I}_0$  du Corollaire 3.14; on a  $\mathcal{I} \subseteq \mathcal{I}_0$ .

D'autre part, on a :

$$] \alpha, \beta[ = [-\infty, \beta[ \setminus [-\infty, \alpha] = [-\infty, \beta[ \bigcap_{n \geq 1} [-\infty, \alpha + \frac{1}{n}[ \in \sigma(\mathcal{I})$$

et :

$$] \alpha, +\infty[ = \overline{\mathbb{R}} \setminus \bigcap_{n \geq 1} [-\infty, \alpha + \frac{1}{n}[ \in \sigma(\mathcal{I});$$

donc  $\mathcal{I}_0 \subseteq \sigma(\mathcal{I})$ , et donc  $\sigma(\mathcal{I}_0) \subseteq \sigma(\mathcal{I})$ .

Il résulte du Corollaire 3.14 que l'on a  $\text{Bor}(\overline{\mathbb{R}}) \subseteq \sigma(\mathcal{I})$ , et donc l'égalité, puisque l'inclusion inverse est évidente.  $\square$

**Corollaire 3.17** *Si  $I$  est un intervalle de  $\mathbb{R}$ , alors toute application **monotone**  $f: I \rightarrow \overline{\mathbb{R}}$  est borélienne.*

**Preuve.** Si  $f$  est monotone, alors  $\{f < a\}$  est un sous-intervalle de  $I$  (donc un borélien) : si  $f(x_1) < a$  et  $f(x_2) < a$  et  $x_1 \leq x \leq x_2$ , alors (en supposant par exemple  $f$  décroissante), on a  $f(x_1) \geq f(x) \geq f(x_2)$ ; donc  $f(x) < a$ .  $\square$

Pour avoir un résultat plus général, nous utiliserons un lemme permettant de dire qu'une fonction est mesurable si elle se décompose en morceaux mesurables. Rappelons d'abord la définition suivante.

**Définition 3.18** Soit  $(S, \mathcal{F})$  un espace mesurable; on appelle partition mesurable dénombrable de  $S$ , toute suite, finie ou infinie,  $(S_n)_{n \geq 1}$  de parties de  $S$  telle que :

- (a)  $S_n \in \mathcal{F}, \forall n \geq 1$ ;
- (b)  $S_n \cap S_k = \emptyset$ , si  $n \neq k$ ;
- (c)  $S = \bigcup_{n \geq 1} S_n$ .

**Lemme 3.19 (Lemme de recollement)** Soit  $(S, \mathcal{F})$  et  $(S', \mathcal{F}')$  deux espaces mesurables,  $(S_n)_{n \geq 1}$  une partition mesurable dénombrable de  $S$ , et soit  $f: S \rightarrow S'$  une application. Si, pour chaque  $n \geq 1$ , la restriction  $f_n = f|_{S_n}: S_n \rightarrow S'$  de  $f$  à  $S_n$  est mesurable, alors  $f$  est mesurable.

On peut aussi exprimer le Lemme de recollement de la façon suivante :

**Lemme 3.20 (Lemme de recollement, variante)** Soit  $(S_n)_{n \geq 1}$  une partition mesurable dénombrable de  $S$ , et  $f_n: S_n \rightarrow S'$  des applications mesurables; si l'on pose :

$$f(x) = f_n(x) \quad \text{si } x \in S_n,$$

alors  $f$  est mesurable.

**Preuve.** Pour tout  $A' \in \mathcal{F}'$ , on a :

$$f^{-1}(A') = \bigcup_{n \geq 1} f_n^{-1}(A') \in \mathcal{F},$$

puisque  $f_n^{-1}(A') \in \mathcal{F}_{S_n}$  et que  $\mathcal{F}_{S_n} \subseteq \mathcal{F}$  (car  $S_n \in \mathcal{F}$ ). □

Rappelons,  $I$  étant un intervalle de  $\mathbb{R}$ , qu'une fonction  $f: I \rightarrow \mathbb{R}$ , est *continue par morceaux*, respectivement *monotone par morceaux*, s'il existe une suite, finie ou infinie, de points de  $I: \dots < x_k < x_{k+1} < \dots < x_0 < x_1 < x_2 < \dots < x_l < \dots$  tels que  $f$  soit continue, respectivement monotone, sur chaque intervalle ouvert  $]x_j, x_{j+1}[$ .

On obtient :

**Proposition 3.21** Toute fonction  $f: I \rightarrow \mathbb{R}$  définie sur un intervalle de  $\mathbb{R}$  qui est continue par morceaux, ou qui est monotone par morceaux, est borélienne.

Noter qu'aux points  $x_k$  de la subdivision, la restriction de  $f$  à  $\{x_k\}$  est trivialement mesurable.  $\square$

**Remarque.** Cet énoncé est d'un très grand intérêt pratique, car, essentiellement, **toute** fonction que l'on peut écrire *explicitement*, c'est-à-dire à l'aide des fonctions usuelles, est continue par morceaux.

La plupart du temps, la justification de la mesurabilité se résumera donc à dire que la fonction considérée est continue, ou continue par morceaux.

Comme pour la continuité, pour les fonctions à valeurs dans  $\mathbb{R}^N$ , on a mesurabilité si et seulement si chaque coordonnée l'est.

**Proposition 3.22** *Soit  $(S, \mathcal{T})$  un espace mesurable et  $f_1, \dots, f_N: S \rightarrow \mathbb{R}$   $N$  applications. Alors l'application  $f = (f_1, \dots, f_N): S \rightarrow \mathbb{R}^N$  est mesurable si et seulement si chacune des fonctions  $f_1, \dots, f_N$  est mesurable.*

**Preuve.** 1) Supposons d'abord  $f_1, \dots, f_N$  mesurables. On a :

$$\begin{aligned} f^{-1}\left(\prod_{k=1}^N ]a_k, b_k[ \right) &= \{x \in S; f_1(x) \in ]a_1, b_1[, \dots, f_N(x) \in ]a_N, b_N[ \} \\ &= \bigcap_{k=1}^N f_k^{-1}(]a_k, b_k[) \in \mathcal{T}; \end{aligned}$$

donc  $f$  est mesurable, puisque  $\mathcal{B}or(\mathbb{R}^N)$  est engendrée par les pavés ouverts (Proposition 2.10).

2) Réciproquement, si  $f$  est mesurable, alors :

$$f_k^{-1}(]a, b[) = f^{-1}(\mathbb{R} \times \dots \times ]a, b[ \times \dots \times \mathbb{R}) \in \mathcal{T},$$

où  $]a, b[$  est le  $k^{\text{ème}}$  facteur, et donc  $f_k$  est mesurable, toujours grâce à la Proposition 2.10.  $\square$

### 3.5 Opérations sur les applications mesurables

Commençons par le résultat général facile suivant.

**Proposition 3.23** *Soit  $(S_1, \mathcal{T}_1)$ ,  $(S_2, \mathcal{T}_2)$  et  $(S_3, \mathcal{T}_3)$  trois espaces mesurables. Si  $f: S_1 \rightarrow S_2$  et  $g: S_2 \rightarrow S_3$  sont mesurables, alors  $g \circ f: S_1 \rightarrow S_3$  est mesurable.*

**Preuve.** Il suffit d'écrire :

$$(g \circ f)^{-1}(\mathcal{T}_3) = f^{-1}[g^{-1}(\mathcal{T}_3)] \subseteq f^{-1}(\mathcal{T}_2) \subseteq (\mathcal{T}_1)$$

pour avoir le résultat.  $\square$

Nous allons nous focaliser ensuite sur les fonctions à valeurs réelles. Commençons par un rappel sur l'**arithmétique de  $\overline{\mathbb{R}}$** .

**Rappel.** On prolonge l'addition et la multiplication de  $\mathbb{R}$ , de façon *commutative*, en posant :

- $\forall a \in \mathbb{R} : \begin{aligned} (+\infty) + a &= a + (+\infty) = +\infty; \\ (-\infty) + a &= a + (-\infty) = -\infty; \end{aligned}$
- $\forall a \in \mathbb{R}_+^* : \begin{aligned} a \times (+\infty) &= (+\infty) \times a = +\infty; \\ a \times (-\infty) &= (-\infty) \times a = -\infty; \end{aligned}$
- $\forall a \in \mathbb{R}_-^* : \begin{aligned} a \times (+\infty) &= (+\infty) \times a = -\infty; \\ a \times (-\infty) &= (-\infty) \times a = +\infty, \end{aligned}$
- $\begin{aligned} (+\infty)(+\infty) &= +\infty; \\ (-\infty)(-\infty) &= +\infty; \\ (+\infty)(-\infty) &= (-\infty)(+\infty) = -\infty. \end{aligned}$

On notera que  $(+\infty) + (-\infty)$  et  $(-\infty) + (+\infty)$  ne sont **pas définis**. On définit par contre  $1/(+\infty) = 1/(-\infty) = 0$ . On notera que l'on ne peut toujours pas diviser par 0. Par contre, on conviendra de définir le produit de l'infini par 0 :

**CONVENTION** On conviendra que  $\boxed{0 \times (\pm\infty) = (\pm\infty) \times 0 = 0}$ .

Cette convention s'impose en Théorie de la mesure : il est naturel d'attribuer une aire nulle à une droite, qui est un rectangle de longueur infinie et de largeur nulle.

**Proposition 3.24** Soit  $(S, \mathcal{F})$  un espace mesurable et  $f, g: S \rightarrow \overline{\mathbb{R}}$  des applications mesurables. Alors les applications suivantes sont mesurables :

- 1)  $\alpha f$ , pour tout  $\alpha \in \mathbb{R}$ ;
- 2)  $f + g$ , lorsqu'elle existe;
- 3)  $1/f$ , lorsqu'elle existe;
- 4)  $fg$ .

**Remarque.** •  $f + g$  n'est pas définie s'il existe  $x \in S$  tel que  $f(x) = +\infty$  et  $g(x) = -\infty$ , ou bien tel que  $f(x) = -\infty$  et  $g(x) = +\infty$ .

- $1/f$  n'existe pas si  $f$  s'annule.

On peut toutefois se restreindre à l'ensemble  $S_1$  (qui est mesurable) sur lequel  $f + g$ , respectivement  $1/f$ , est définie.

- Par contre,  $\alpha f$  et  $fg$  sont toujours définies, grâce à la convention ci-dessus.

On obtient immédiatement, si l'on ne considère que des fonctions à valeurs finies (c'est-à-dire dans  $\mathbb{R}$ , et pas dans  $\overline{\mathbb{R}}$ ) :

**Corollaire 3.25** L'ensemble des fonctions mesurables  $f: S \rightarrow \mathbb{R}$  est un sous-espace vectoriel de l'espace vectoriel de toutes les fonctions de  $S$  dans  $\mathbb{R}$ .

**Preuve de la Proposition 3.24.**

- 1) • pour  $\alpha = 0$ , c'est évident car  $\alpha f \equiv 0$ .
- Pour  $\alpha > 0$ , on a, pour tout  $a \in \mathbb{R}$  :

$$\{\alpha f < a\} = \{f < a/\alpha\} \in \mathcal{F}.$$

- Pour  $\alpha < 0$ , on a, pour tout  $a \in \mathbb{R}$  :

$$\{\alpha f < a\} = \{f > a/\alpha\} \in \mathcal{T}.$$

2) Soit  $a \in \mathbb{R}$ . On veut montrer que  $\{f + g < a\} \in \mathcal{T}$ .

Or :

$$\begin{aligned} f(x) + g(x) < a &\iff f(x) < a - g(x) \\ &\iff \exists r \in \mathbb{Q} : f(x) < r < a - g(x); \end{aligned}$$

donc :

$$\{f + g < a\} = \bigcup_{r \in \mathbb{Q}} [\{f < r\} \cap \{g < a - r\}] \in \mathcal{T},$$

parce que  $\mathbb{Q}$  est dénombrable.

3) Il suffit que remarquer que l'on a, d'abord :  $\{1/f < 0\} = \{f < 0\} \in \mathcal{T}$ ; ensuite, pour  $a > 0$  :

$$\{1/f < a\} = \{f < 0\} \cup \{f > 1/a\} \in \mathcal{T};$$

et que, pour  $a > 0$  :

$$\{1/f < a\} = \{f < 0\} \cap \{f > 1/a\} \in \mathcal{T}.$$

4) On va le faire en plusieurs étapes.

a) Si  $f$  est mesurable, alors  $f^2$  est aussi mesurable.

En effet, si  $a \leq 0$ , on a  $\{f^2 < a\} = \emptyset \in \mathcal{T}$ , et si  $a > 0$ , alors  $\{f^2 < a\} = \{-\sqrt{a} < f < \sqrt{a}\} \in \mathcal{T}$ .

b) Si  $f$  est mesurable et  $A \in \mathcal{T}$ , alors  $f \cdot \mathbb{1}_A$  est mesurable.

En effet, comme  $(f \cdot \mathbb{1}_A)(x) = \begin{cases} f(x) & \text{si } x \in A, \\ 0 & \text{si } x \notin A \end{cases}$ , on a, d'une part, lorsque  $a \leq 0$  :  $\{f \cdot \mathbb{1}_A < a\} = \{f < a\} \cap A \in \mathcal{T}$ , et, d'autre part, si  $a > 0$ ,  $\{f \cdot \mathbb{1}_A < a\} = (\{f < a\} \cap A) \cup A^c \in \mathcal{T}$ .

Remarquons que l'on pourrait dire les choses de la façon suivante : la restriction de  $f \cdot \mathbb{1}_A$  à  $A^c$  est nulle, donc mesurable, et la restriction de  $f \cdot \mathbb{1}_A$  à  $A$  est égale à la restriction de  $f$  à  $A$ , c'est-à-dire aussi à  $f \circ j_A$ , où  $j_A : A \rightarrow S$  est l'injection canonique; comme  $f \circ j_A$  est mesurable, cette restriction aussi; le Lemme de recollement permet de conclure que  $f \cdot \mathbb{1}_A$  est mesurable.

c) Posons  $F = \{x \in S; |f(x)| < +\infty \text{ et } |g(x)| < +\infty\}$ ; c'est une partie mesurable, et, par le b),  $f_1 = f \cdot \mathbb{1}_F$  et  $g_1 = g \cdot \mathbb{1}_F$  sont mesurables. Comme elles sont à valeurs dans  $\mathbb{R}$ ,  $f_1 + g_1$  est définie, et est mesurable, par le 2). Il résulte alors du a), et du 2), puisque l'on a des fonctions à valeurs dans  $\mathbb{R}$ , que :

$$f_1 g_1 = \frac{1}{2} [(f_1 + g_1)^2 - f_1^2 - g_1^2]$$

est mesurable.

Maintenant, pour  $x \in F^c$ , on a :

- $f(x)g(x) = -\infty$  si

$$x \in (\{f = -\infty\} \cap \{g > 0\}) \cup (\{f = +\infty\} \cap \{g < 0\}) \\ \cup (\{f < 0\} \cap \{g = +\infty\}) \cup (\{f > 0\} \cap \{g = -\infty\});$$

- $f(x)g(x) = 0$  si  $x \in (\{f = \pm\infty\} \cap \{g = 0\}) \cup \{f = 0\} \cap \{g = \pm\infty\}$ ;
- $f(x)g(x) = +\infty$  si

$$x \in (\{f = -\infty\} \cap \{g < 0\}) \cup (\{f = +\infty\} \cap \{g > 0\}) \\ \cup (\{f > 0\} \cap \{g = +\infty\}) \cup (\{f < 0\} \cap \{g = -\infty\}).$$

Le Lemme de recollement permet donc de dire que  $f$  est mesurable.  $\square$

**Remarque.** Lorsque les fonctions  $f$  et  $g$  sont à valeurs dans  $\mathbb{R}$  (comme c'est d'ailleurs le cas de  $f_1$  et  $g_1$ ), il est possible de faire une preuve plus simple pour 2) et 4), en disant que  $f + g$  et  $fg$  sont, respectivement, les composées des applications mesurables  $(f, g): S \rightarrow \mathbb{R}^2$  et des applications continues, donc boréliennes,  $(u, v) \in \mathbb{R}^2 \mapsto u + v \in \mathbb{R}$  et, respectivement,  $(u, v) \in \mathbb{R}^2 \mapsto uv \in \mathbb{R}^2$ .

Le grand avantage de la mesurabilité, par rapport, par exemple, à la continuité, est que l'on garde cette mesurabilité quand on fait des opérations sur des familles *dénombrables*, pouvant être infinies, de fonctions.

**Proposition 3.26** *Si  $f_n: S \rightarrow \overline{\mathbb{R}}$ ,  $n \geq 1$ , est une suite de fonctions mesurables, alors  $\inf_{n \geq 1} f_n$  et  $\sup_{n \geq 1} f_n$  sont mesurables.*

On notera que, même si les  $f_n$  sont à valeurs dans  $\mathbb{R}$ ,  $\inf_{n \geq 1} f_n$  et  $\sup_{n \geq 1} f_n$  peuvent prendre des valeurs infinies.

**Preuve.** Il suffit d'écrire que, pour tout  $a \in \mathbb{R}$ , on a :

$$\left\{ \inf_{n \geq 1} f_n < a \right\} = \bigcup_{n \geq 1} \{f_n < a\} \in \mathcal{F}$$

et que :

$$\left\{ \sup_{n \geq 1} f_n \leq a \right\} = \bigcap_{n \geq 1} \{f_n \leq a\} \in \mathcal{F}$$

(noter que cette dernière égalité ne serait pas valable avec  $< a$ ).  $\square$

En prenant  $f_1 = f$  et  $f_n = g$  pour  $n \geq 2$ , on obtient, en particulier :

**Corollaire 3.27** *Si  $f, g: S \rightarrow \overline{\mathbb{R}}$  sont mesurables, alors  $\sup(f, g)$  et  $\inf(f, g)$  sont mesurables.*

La notation suivante sera très souvent utilisée.

**Notation.** Si  $f: S \rightarrow \overline{\mathbb{R}}$  est une application, on pose :

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

En d'autres termes :

$$f^+(x) = \begin{cases} f(x) & \text{si } f(x) \geq 0; \\ 0 & \text{si } f(x) \leq 0, \end{cases}$$

et :

$$f^-(x) = \begin{cases} -f(x) = |f(x)| & \text{si } f(x) \leq 0; \\ 0 & \text{si } f(x) \geq 0, \end{cases}$$

Ce sont des fonctions **positives**, et l'on a :

$$\boxed{f = f^+ - f^-} \quad \text{et} \quad \boxed{|f| = f^+ + f^-}.$$

$f^+$  s'appelle la *partie positive* de  $f$ , et  $f^-$  sa *partie négative* (ce dernier terme étant mal choisi, puisque  $f^-$  est positive).

**Corollaire 3.28** Si  $f: S \rightarrow \overline{\mathbb{R}}$  est mesurable, alors  $f^+$ ,  $f^-$  et  $|f|: S \rightarrow \overline{\mathbb{R}}_+$  sont mesurables.

**Remarque.**  $|f|$  peut être mesurable sans que  $f$  le soit.

**Exemple.** Soit  $A \notin \mathcal{T}$ , et  $f = \mathbb{1}_A - \mathbb{1}_{A^c}$  (c'est-à-dire que  $f(x) = 1$  si  $x \in A$  et  $f(x) = -1$  si  $x \in A^c$ ). Alors  $|f| \equiv \mathbb{1}$  est mesurable, mais  $f$  ne l'est pas :  $\{f = 1\} = A$  n'est pas mesurable.

**Notation.** Etant donnée une suite *croissante* de nombres  $a_n \in \overline{\mathbb{R}}$ , on notera  $\lim_{n \rightarrow \infty}^{\uparrow} a_n$  sa limite, la flèche  $\uparrow$  étant là pour préciser que la suite est **croissante**.

De même, on notera  $\lim_{n \rightarrow \infty}^{\downarrow} a_n$  la limite d'une suite **décroissante**  $(a_n)_{n \geq 1}$ .

On applique la même notation pour les suites de fonctions  $f_n: S \rightarrow \overline{\mathbb{R}}$ .

**Rappel.** Si  $(a_n)_{n \geq 1}$  est une suite de nombres ( $a_n \in \overline{\mathbb{R}}$ ), on définit sa *limite supérieure* et sa *limite inférieure* par :

$$\begin{aligned} \limsup_{n \rightarrow \infty} a_n &= \lim_{n \rightarrow \infty}^{\downarrow} \left[ \sup_{k \geq n} f_k \right] = \inf_{n \geq 1} \left[ \sup_{k \geq n} f_k \right] \\ \liminf_{n \rightarrow \infty} a_n &= \lim_{n \rightarrow \infty}^{\uparrow} \left[ \inf_{k \geq n} f_k \right] = \sup_{n \geq 1} \left[ \inf_{k \geq n} f_k \right]. \end{aligned}$$

**Corollaire 3.29** Si  $f_n: S \rightarrow \overline{\mathbb{R}}$ ,  $n \geq 1$ , est une suite de fonctions mesurables, alors  $\liminf_{n \rightarrow \infty} f_n$  et  $\limsup_{n \rightarrow \infty} f_n$  sont mesurables.

### 3.6 Fonctions étagées

Les fonctions étagées vont tenir, vis-à-vis de l'intégrale de Lebesgue, le rôle que les fonctions en escalier ont pour l'intégrale de Riemann. Ce sont les fonctions qui sont constantes, non plus sur des intervalles, mais sur les ensembles, beaucoup plus généraux, que sont les ensembles mesurables. Ce seront donc les fonctions fondamentales pour l'intégration.

**Définition 3.30** Soit  $(S, \mathcal{T})$  un espace mesurable. On dit qu'une fonction  $f: S \rightarrow \mathbb{R}$  est **étagée** si elle est mesurable et si l'ensemble  $f(S)$  de ses valeurs est une partie finie de  $\mathbb{R}$ .

On fera attention que les fonctions étagées prennent leurs valeurs dans  $\mathbb{R}$ , et pas dans  $\overline{\mathbb{R}}$ .

Toute fonction indicatrice d'un ensemble mesurable est étagée ; plus généralement, toute combinaison linéaire (finie) de fonctions indicatrices d'ensembles mesurables est étagée ; par exemple, si  $I$  est un intervalle de  $\mathbb{R}$ , toute fonction  $f: I \rightarrow \mathbb{R}$  en escalier est étagée ; par contre,  $\mathbb{1}_{\mathbb{Q}}$  est une fonction étagée qui n'est pas en escalier. Nous allons voir qu'en fait les fonctions étagées sont exactement les combinaisons linéaires de fonctions indicatrices d'ensembles mesurables. En effet, toute fonction étagée possède une **représentation canonique** : si l'on écrit  $f(S) = \{a_1, \dots, a_n\}$ , avec  $a_1, \dots, a_n$  **distincts**, alors l'ensemble :

$$S_k = \{f = a_k\}$$

(c'est-à-dire  $S_k = f^{-1}(\{a_k\})$ ) est mesurable ( $S_k \in \mathcal{T}$ ), car  $f$  est mesurable, et l'on a :

$$f = \sum_{k=1}^n a_k \mathbb{1}_{S_k} = \sum_{a \in f(S)} a \mathbb{1}_{\{f=a\}} .$$

Il est important d'avoir cette représentation canonique car une même fonction étagée peut s'écrire de nombreuses façons comme combinaison linéaire de fonctions indicatrices d'ensembles mesurables.

La construction de l'intégrale de Riemann se fait en définissant d'abord l'intégrale des fonctions en escalier, ce qui se fait facilement, et naturellement ; on peut ensuite intégrer les fonctions qui s'approchent convenablement par les fonctions en escalier. La démarche pour l'intégrale de Lebesgue sera la même : on définira d'abord, de façon évidente, l'intégrale des fonctions étagées, et l'on cherchera à intégrer les fonctions qui peuvent être approchées par des fonctions étagées. Comme on dispose de beaucoup de parties mesurables, c'est beaucoup plus facile que d'approcher, dans le cas de fonctions définies sur un intervalle de  $\mathbb{R}$ , les fonctions par des fonctions en escalier ; en fait *toutes* les fonctions mesurables peuvent être approchées par des fonctions étagées. C'est l'objet du théorème, évidemment fondamental, suivant.

**Théorème 3.31 (Théorème fondamental d'approximation)**

Soit  $f: S \rightarrow \overline{\mathbb{R}}_+ = [0, +\infty]$  une application mesurable positive. Alors il existe une suite croissante d'applications  $\varphi_n: S \rightarrow \mathbb{R}_+$  étagées positives dont la limite est  $f$ .

De plus, si  $f$  est bornée, la convergence est uniforme.

La suite  $(\varphi_n)_{n \geq 1}$  vérifie donc, pour tout  $x \in S$  :

$$1) 0 \leq \varphi_1(x) \leq \varphi_2(x) \leq \dots \leq \varphi_n(x) \leq \dots \leq f(x);$$

$$2) \lim_{n \rightarrow \infty} \uparrow \varphi_n(x) = f(x).$$

Précisons que, dans la dernière partie,  $f$  est supposée bornée : il existe  $M \in \mathbb{R}_+^*$  tel que  $0 \leq f(x) \leq M$  pour tout  $x \in S$ , et pas seulement à valeurs finies.

**Corollaire 3.32** Pour toute application mesurable  $f: S \rightarrow \overline{\mathbb{R}}$  ou  $\mathbb{C}$ , il existe une suite de fonctions étagées  $\psi_n: S \rightarrow \mathbb{R}$  ou  $\mathbb{C}$  telle que  $\lim_{n \rightarrow \infty} \psi_n(x) = f(x)$ , pour tout  $x \in S$ .

De plus, si  $f$  est bornée, la convergence est uniforme.

**Preuve.** Il suffit d'appliquer le théorème à  $f^+$  et à  $f^-$  si  $f$  est à valeurs réelles, puis à  $\operatorname{Re} f$  et  $\operatorname{Im} f$  lorsque  $f$  est à valeurs complexes.  $\square$

**Preuve du théorème.** Pour  $n \geq 1$ , on partage l'intervalle  $[0, n]$  en  $n2^n$  intervalles de longueur  $1/2^n$ , et l'on pose :

$$S_{n,k} = \begin{cases} \left\{ \frac{k}{2^n} \leq f < \frac{k+1}{2^n} \right\} & \text{si } 0 \leq k \leq n2^n - 1, \\ \{f \geq n\} & \text{si } k = n2^n. \end{cases}$$

Les parties  $S_{n,k}$ ,  $0 \leq k \leq n2^n$ , sont mesurables et forment une partition de  $S$ . On définit :

$$\varphi_n = \sum_{k=0}^{n2^n} \frac{k}{2^n} \mathbb{1}_{S_{n,k}},$$

c'est-à-dire que :

$$\varphi_n(x) = \frac{k}{2^n} \quad \text{si } x \in S_{n,k}, \quad 0 \leq k \leq n2^n.$$

Chaque fonction  $\varphi_n$  est étagée positive, et :

1) Par construction, on a  $0 \leq \varphi_n \leq f$ ; il reste à voir que  $\varphi_n \leq \varphi_{n+1}$ . Soit  $x \in S$ .

a) Si  $x \in S_{n,k} = \left\{ \frac{k}{2^n} \leq f < \frac{k+1}{2^n} \right\}$ , avec  $0 \leq k \leq n2^n - 1$ , on a  $\varphi_n(x) = \frac{k}{2^n}$ ; mais :

$$\frac{k}{2^n} \leq f(x) < \frac{k+1}{2^n} \iff \frac{2k}{2^{n+1}} \leq f(x) < \frac{2k+2}{2^{n+1}},$$

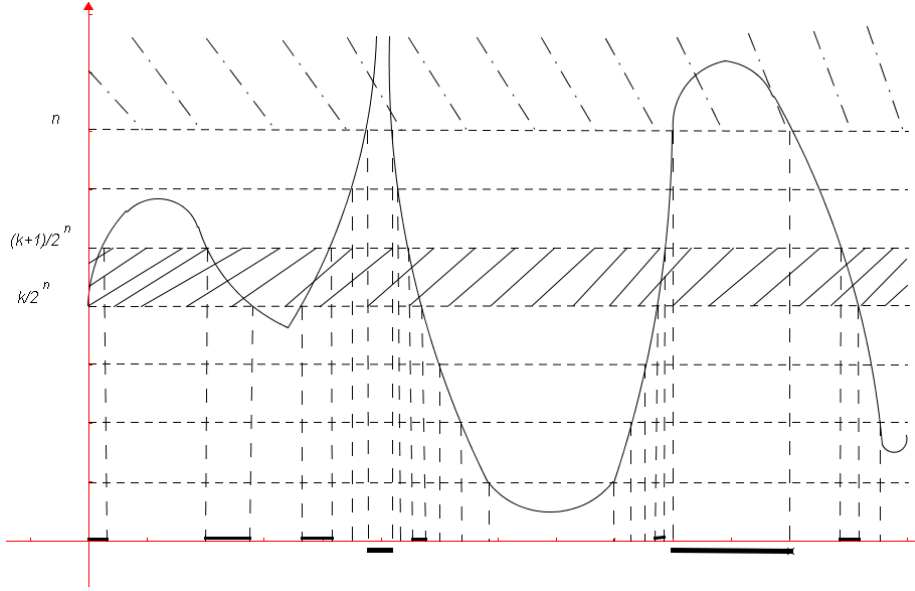


FIGURE 2 – Approximation des fonctions mesurables positives

(autrement dit :  $S_{n,k} = S_{n+1,2k} \cup S_{n+1,2k+1}$ ); donc :

- si  $\frac{2k}{2^{n+1}} \leq f(x) < \frac{2k+1}{2^{n+1}}$ , on a  $\varphi_{n+1}(x) = \frac{2k}{2^{n+1}} = \frac{k}{2^n} = \varphi_n(x)$ ;
- si  $\frac{2k+1}{2^{n+1}} \leq f(x) < \frac{2k+2}{2^{n+1}}$ , on a  $\varphi_{n+1}(x) = \frac{2k+1}{2^{n+1}} > \frac{k}{2^n} = \varphi_n(x)$ .

b) Si  $x \in S_{n,n2^n} = \{f \geq n\}$ , on a  $\varphi_n(x) = n$ , et alors :

- si  $n \leq f(x) < n+1$ , il existe  $l$ , avec  $n2^{n+1} \leq l < (n+1)2^{n+1} - 1$ , tel que  $\frac{l}{2^{n+1}} \leq f(x) < \frac{l+1}{2^{n+1}}$ , de sorte que  $\varphi_{n+1}(x) = \frac{l}{2^{n+1}} \geq n = \varphi_n(x)$ ;
- si  $f(x) \geq n+1$ , alors  $\varphi_{n+1}(x) = n+1 > n = \varphi_n(x)$ .

2) Montrons que  $\varphi_n(x)$  tend vers  $f(x)$ . Soit  $x \in S$ .

a) Si  $f(x) < +\infty$ , il existe  $n_0 \geq 1$  tel que  $f(x) < n_0$ ; alors, pour  $n \geq n_0$ , on a  $0 \leq f(x) - \varphi_n(x) \leq \frac{1}{2^n}$ , et on a bien  $\lim_{n \rightarrow \infty} \varphi_n(x) = f(x)$ .

b) Si  $f(x) = +\infty$ , alors  $f(x) \geq n$  pour tout  $n \geq 1$ ; donc, pour tout  $n \geq 1$ ,  $\varphi_n(x) = n$ , et  $\lim_{n \rightarrow \infty} \varphi_n(x) = +\infty = f(x)$ .

Pour terminer, il ne reste plus qu'à voir qu'il y a convergence uniforme lorsque  $f$  est bornée. Mais si  $0 \leq f(x) \leq M$  pour tout  $x \in S$ , alors, pour tout  $\varepsilon > 0$ , si l'on choisit un entier  $n_0 > M$  tel que  $\frac{1}{2^{n_0}} \leq \varepsilon$ , on a, pour  $n \geq n_0$  :

$$0 \leq f(x) - \varphi_n(x) \leq \frac{1}{2^n} \leq \frac{1}{2^{n_0}} \leq \varepsilon$$

ce qui termine la preuve. □

## 4 Mesures positives

### 4.1 Définition - Exemples

Maintenant que l'on a précisé quelles étaient les propriétés des parties que l'on souhaitait pouvoir mesurer (les *parties mesurables*), nous allons spécifier la mesure que l'on entend leur donner. On ne peut bien sûr pas associer simplement à chaque ensemble mesurable une valeur qui représentera sa mesure : il y a des relations de compatibilité à assurer. Elles vont en fait se résumer à une seule condition : la *propriété de  $\sigma$ -additivité*.

**Définition 4.1** Soit  $(S, \mathcal{F})$  un espace mesurable. On appelle **mesure positive** sur  $(S, \mathcal{F})$  (ou, plus simplement, sur  $S$ ), toute application

$$m: \mathcal{F} \rightarrow \overline{\mathbb{R}}_+ = [0, +\infty]$$

telle que :

- 1)  $m(\emptyset) = 0$ ;
- 2) pour toute **suite** (c'est-à-dire famille dénombrable)  $(A_n)_{n \geq 1}$  de parties mesurables  $A_n \in \mathcal{F}$  **deux-à-deux disjointes**, on ait :

$$m\left(\bigcup_{n \geq 1} A_n\right) = \sum_{n=1}^{\infty} m(A_n) .$$

On appelle cette condition la **propriété de  $\sigma$ -additivité de la mesure**.

Rappelons que les  $A_n$ ,  $n \geq 1$ , sont deux-à-deux disjointes si pour tous les entiers  $n, k \geq 1$ , on a  $A_n \cap A_k = \emptyset$  lorsque  $k \neq n$ .

Noter qu'il est naturel d'admettre  $+\infty$  comme valeur possible : par exemple, un demi-plan est d'aire infinie ;  $\mathbb{R}$  est de longueur infinie. Par contre on demande que l'ensemble vide soit de mesure nulle pour éliminer la possibilité que  $m$  soit constamment égale à  $+\infty$ , cas qui est bien sûr sans intérêt (et si  $m(\emptyset) < +\infty$ , la seule possibilité est que  $m(\emptyset) = 0$  : prendre tous les  $A_n = \emptyset$  dans la propriété de  $\sigma$ -additivité).

On fera attention que la mesure positive  $m$ , bien que dite mesure positive sur  $S$ , est définie, comme cela est naturel, sur la tribu  $\mathcal{F}$  des parties mesurables, et non pas sur  $S$  lui-même.

**Remarque.** La  $\sigma$ -additivité contient en particulier la *propriété d'additivité simple* : pour tout  $N \geq 1$ , si les  $N$  ensembles mesurables  $A_1, \dots, A_N$  sont deux-à-deux disjointes, alors :

$$m(A_1 \cup \dots \cup A_N) = m(A_1) + \dots + m(A_N)$$

(il suffit de prendre  $A_{N+1} = A_{N+2} = \dots = \emptyset$ ). En particulier, si  $A$  et  $B$  sont deux parties mesurables disjointes :  $A \cap B = \emptyset$ , alors  $m(A \cup B) = m(A) + m(B)$ . Inversement, cette propriété d'additivité pour deux parties mesurables disjointes entraîne l'additivité pour un nombre *fini* de parties mesurables deux-à-deux disjointes, par récurrence.

**Définition 4.2** On appelle **espace mesuré** tout triplet  $(S, \mathcal{F}, m)$  formé d'un ensemble  $S$ , d'un tribu  $\mathcal{F}$  de parties de  $S$ , et d'une mesure positive  $m$  sur l'espace mesurable  $(S, \mathcal{F})$ .

### Exemples de mesures positives

a. Soit  $S$  un ensemble (non vide) et  $a \in S$  un point fixé. On définit  $\delta_a : \mathcal{P}(S) \rightarrow \{0, 1\} \subseteq \overline{\mathbb{R}}_+$  par :

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

On vérifie immédiatement que  $\delta_a$  est une mesure positive sur  $(S, \mathcal{P}(S))$ , que l'on appelle la **mesure de Dirac** en  $a$ . On dit aussi *masse de Dirac* en  $a$ .

b. Soit  $S$  un ensemble; on définit  $c : \mathcal{P}(S) \rightarrow \overline{\mathbb{N}} \subseteq \overline{\mathbb{R}}_+$  par :

$$c(A) = \begin{cases} \text{card}(A) & \text{si } A \text{ est fini} \\ 0 & \text{si } A \text{ est infini} \end{cases}.$$

Alors  $c$  est une mesure positive sur  $(S, \mathcal{P}(S))$  : la propriété de  $\sigma$ -additivité est connue si seul un nombre fini de parties sont non vides, et si une infinité de  $A_n$  sont non vides, alors la réunion est infinie (car les parties sont deux-à-deux disjointes), et l'on a  $c(\bigcup_{n \geq 1} A_n) = +\infty = \sum_{n=1}^{\infty} c(A_n)$ , puisque pour une infinité de  $A_n$  on a  $c(A_n) \geq 1$ . On dit que  $c$  est la **mesure de comptage**, ou *mesure de dénombrement*, sur  $S$ .

Une des difficultés dans la théorie de la mesure est que, puisque, comme on l'a déjà dit, on ne peut pas, en général, décrire toutes les parties mesurables, on ne pourra pas, en général, expliciter complètement une mesure positive. L'*existence* même de mesures positives sur un espace mesurable donné ne va pas de soi. On verra dans les Annexes, à la fin du cours, comment, d'une part, construire la mesure de Lebesgue, et, d'autre part, comment le *Théorème de représentation de Riesz* permet de construire des mesures positives, sur les espaces localement compacts, à partir de *formes linéaires positives* sur l'espace des fonctions continues à support compact. Ce sont des théorèmes longs et assez délicats à prouver; aussi, nous **admettrons** l'existence de la *mesure de Lebesgue*, qui est l'exemple le plus important de mesure positive.

### c. Mesure de Lebesgue.

**Théorème 4.3** *Il existe une unique mesure positive  $\lambda$  sur  $\mathbb{R}$ , muni de sa tribu borélienne, appelée la mesure de Lebesgue sur  $\mathbb{R}$ , telle que :*

$$\lambda(]a, b[) = b - a, \quad \forall a, b \in \mathbb{R}, a < b.$$

L'unicité sera montrée un peu plus tard, dans le chapitre "Intégration sur un espace produit".

La mesure de Lebesgue est donc la mesure positive qui prolonge à toute la tribu borélienne de  $\mathbb{R}$  la "mesure" naturelle des intervalles (ici les intervalles ouverts bornés, mais on verra que cela reste vrai pour tous les intervalles) qu'est leur **longueur**. On n'a pas de définition explicite de la mesure de tous les boréliens (de toute façon, on ne connaît pas tous les boréliens), mais ce n'est pas vraiment gênant ; on pourra faire beaucoup de choses en sachant simplement les trois choses suivantes :

- 1) La mesure de Lebesgue  $\lambda$  est une **mesure positive** (avec toutes les propriétés des mesures positives que l'on va voir) ;
- 2) les intervalles ont pour mesure leur **longueur** ;
- 3)  $\lambda$  est la **seule** mesure positive sur  $\mathbb{R}$  ayant cette propriété.

On a énoncé l'existence, et l'unicité, de la mesure de Lebesgue sur  $\mathbb{R}$ , de façon séparée, pour que cela soit plus parlant ; mais on a la même chose sur  $\mathbb{R}^N$ , pour toute dimension  $N \geq 1$ .

**Théorème 4.4** *Il existe une unique mesure positive  $\lambda_N$  sur  $\mathbb{R}^N$ , muni de sa tribu borélienne, appelée la mesure de Lebesgue sur  $\mathbb{R}^N$ , telle que :*

$$\lambda_N \left( \prod_{k=1}^N ]a_k, b_k[ \right) = \prod_{k=1}^N (b_k - a_k)$$

pour tous  $a_k, b_k \in \mathbb{R}$ ,  $a_k < b_k$ ,  $1 \leq k \leq N$ .

Sauf mention expresse du contraire, quand on parlera de mesure sur  $\mathbb{R}$ , ou sur  $\mathbb{R}^N$ , il s'agira toujours de la mesure de Lebesgue.

## 4.2 Construction de mesures positives

A partir de mesures positives existantes, on peut en fabriquer d'autres.

**Proposition 4.5** Soit  $(S, \mathcal{F})$  un espace mesurable.

1) Si  $m_1$  et  $m_2$  sont deux mesures positives sur  $(S, \mathcal{F})$ , alors  $m_1 + m_2$  en est aussi une.

Plus généralement, si  $(m_n)_{n \geq 1}$  est une suite de mesures positives sur  $(S, \mathcal{F})$ , alors  $\sum_{n=1}^{\infty} m_n$  en est aussi une.

2) Si  $m$  est une mesure positive sur  $(S, \mathcal{F})$ , alors  $am$  en est aussi une, pour tout réel positif  $a \in \mathbb{R}_+$ .

**Preuve.** C'est immédiat.  $\square$

Par exemple, la mesure de comptage  $c$  sur  $\mathbb{N}$  est égale à la somme  $\sum_{n=0}^{\infty} \delta_n$  des mesures de Dirac en  $n$  sur  $\mathbb{N}$ .

**Proposition 4.6** Soit  $(S_1, \mathcal{F}_1)$  et  $(S_2, \mathcal{F}_2)$  deux espaces mesurables, et soit  $\varphi: S_1 \rightarrow S_2$  une application mesurable.

Si  $m_1$  est une mesure positive sur  $(S_1, \mathcal{F}_1)$ , on définit  $m_2: \mathcal{F}_2 \rightarrow \overline{\mathbb{R}}_+$  par :

$$m_2(A_2) = m_1[\varphi^{-1}(A_2)], \quad \forall A_2 \in \mathcal{F}_2.$$

Alors  $m_2$  est une mesure positive sur  $(S_2, \mathcal{F}_2)$ , appelée la **mesure-image** de  $m_1$  par  $\varphi$ . On la note  $m_2 = \varphi(m_1)$ , ou aussi  $(m_1)_\varphi$ .

Cette notion de mesure-image est fondamentale en Probabilité : c'est elle qui définit ce que l'on appelle la *loi de probabilité d'une variable aléatoire*.

**Preuve.** 1) On a  $m_2(\emptyset) = m_1[\varphi^{-1}(\emptyset)] = m_1(\emptyset) = 0$ .

2) Si l'on a des  $A_n \in \mathcal{F}_2$ ,  $n \geq 1$ , deux-à-deux disjoints, alors  $\varphi^{-1}(A_n) \in \mathcal{F}_1$ , car  $\varphi$  est mesurable, et ils sont aussi deux-à-deux disjoints car, pour  $k \neq n$  :

$$\varphi^{-1}(A_n) \cap \varphi^{-1}(A_k) = \varphi^{-1}(A_n \cap A_k) = \varphi^{-1}(\emptyset) = \emptyset;$$

donc :

$$\begin{aligned} m_2\left(\bigcup_{n \geq 1} A_n\right) &= m_1\left[\varphi^{-1}\left(\bigcup_{n \geq 1} A_n\right)\right] = m_1\left[\bigcup_{n \geq 1} \varphi^{-1}(A_n)\right] \\ &= \sum_{n=1}^{\infty} m_1[\varphi^{-1}(A_n)] = \sum_{n=1}^{\infty} m_2(A_n), \end{aligned}$$

ce qui prouve que  $m_2$  est une mesure positive.  $\square$

**Proposition 4.7** Soit  $(S, \mathcal{F})$  un espace mesurable, et soit  $R \in \mathcal{F}$  une partie mesurable, muni de la tribu-trace  $\mathcal{F}_R = \{A \in \mathcal{F}; A \subseteq R\}$ . Alors la restriction  $m_R = m|_{\mathcal{F}_R}: \mathcal{F}_R \rightarrow \overline{\mathbb{R}}_+$  de  $m$  à  $\mathcal{F}_R \subseteq \mathcal{F}$  est une mesure positive sur  $(R, \mathcal{F}_R)$ , appelée la *mesure-trace de  $m$  sur  $R$* .

**Preuve.** C'est évident.  $\square$

**Définition 4.8** Si  $B$  est un borélien de  $\mathbb{R}^N$ , la mesure-trace sur  $B$  de la mesure de Lebesgue sur  $\mathbb{R}^N$  s'appelle la *mesure de Lebesgue sur  $B$* .

### 4.3 Propriétés des mesures positives

**Proposition 4.9** *Toute mesure positive sur  $(S, \mathcal{F})$  est croissante sur  $\mathcal{F}$  :*

$$(\forall A, B \in \mathcal{F}) \quad \boxed{A \subseteq B \implies m(A) \leq m(B)}.$$

**Preuve.** Notons d'abord que, quelques soient  $A, B \subseteq S$  :

$$\boxed{A, B \in \mathcal{F} \implies B \setminus A = B \cap A^c \in \mathcal{F}},$$

et que  $A \cap (B \setminus A) = \emptyset$ . Lorsque  $A \subseteq B$ , on a  $B = A \cup (B \setminus A)$ ; on obtient donc  $m(B) = m(A \cup (B \setminus A)) = m(A) + m(B \setminus A) \geq m(A)$ .  $\square$

**Remarque.** On a donc toujours  $\boxed{m(A) \leq m(S)}$ , quelle que soit la partie mesurable  $A$ .

**Définition 4.10** *On dit qu'une mesure positive  $m$  est finie si elle est à valeurs finies :  $m(A) < +\infty$  pour tout  $A \in \mathcal{F}$ .*

Autrement dit,  $m: \mathcal{F} \rightarrow \mathbb{R}_+$  est à valeurs dans  $\mathbb{R}_+ = [0, +\infty[$ , et pas dans  $\overline{\mathbb{R}}_+ = [0, +\infty]$ . D'après la remarque précédente, cela arrive si et seulement si  $\boxed{m(S) < +\infty}$ , et l'on a alors :

$$m(A) \leq m(S) < +\infty, \quad \forall A \in \mathcal{F}.$$

C'est pourquoi on dit aussi que les mesures positives finies des *mesures positives bornées*.

Il y a des mesures positives qui ne sont pas bornées (il y en a même beaucoup), mais les plus intéressantes sont celles qui ne prennent pas trop souvent la valeurs  $+\infty$ . Cela conduit à la notion suivante.

**Définition 4.11** *Soit  $m$  une mesure positive sur  $(S, \mathcal{F})$ . On dit qu'elle est  $\sigma$ -finie s'il existe une suite de parties mesurables  $S_n, n \geq 1$ , telle que :*

- 1)  $S = \bigcup_{n \geq 1} S_n$ ;
- 2)  $m(S_n) < +\infty$  pour tout  $n \geq 1$ .

**Exemples.**

- a. La mesure de Dirac  $\delta_a$  est finie :  $\delta_a(S) = 1 < +\infty$ .
- b. La mesure de comptage sur  $S$  est :
  - finie si et seulement si  $S$  est fini;

•  $\sigma$ -finie si et seulement si  $S$  est dénombrable.

c. La *mesure de Lebesgue sur  $\mathbb{R}$*  (ou sur  $\mathbb{R}^N$ ) est  $\sigma$ -finie, car  $\mathbb{R} = \bigcup_{n \geq 1} ]-n, n[$  et que, pour tout  $n \geq 1$ ,  $\lambda(]-n, n[) = 2n < +\infty$ .

Par contre, elle n'est pas finie :  $\lambda(\mathbb{R}) = +\infty$  ; en effet,  $\lambda(\mathbb{R}) \geq \lambda(]-n, n[) = 2n$  pour tout  $n \geq 1$ .

d. Si  $B$  est un borélien *borné* de  $\mathbb{R}$  (ou de  $\mathbb{R}^N$ ) : il existe  $M \in \mathbb{R}_+$  tel que  $B \subseteq ]-M, M[$ , alors la mesure de Lebesgue sur  $B$  est bornée (car  $\lambda(B) \leq \lambda(]-M, M[) = 2M < +\infty$ ).

Par contre, un borélien de  $\mathbb{R}$  peut avoir une mesure de Lebesgue finie sans pour autant être borné.

Le théorème suivant donne les deux propriétés les plus importantes, en dehors de la  $\sigma$ -additivité de la définition, des mesures positives. Auparavant, donnons une notation.

**Notation.** Lorsque l'on a une suite **croissante** d'ensembles  $A_n$ ,  $n \geq 1$ , on notera :

$$\lim_{n \rightarrow \infty}^{\uparrow} A_n = \bigcup_{n \geq 1} A_n,$$

(la flèche vers le haut indiquant que la suite est croissante). De même, si  $(A_n)_{n \geq 1}$  est une suite **décroissante**, on note :

$$\lim_{n \rightarrow \infty}^{\downarrow} A_n = \bigcap_{n \geq 1} A_n.$$

**Théorème 4.12** *Toute mesure positive  $m$  sur un espace mesurable  $(S, \mathcal{F})$  vérifie :*

1) la **propriété de limite croissante** : si  $(A_n)_{n \geq 1}$  est une **suite croissante de parties mesurables**, alors :

$$m\left(\lim_{n \rightarrow \infty}^{\uparrow} A_n\right) = \lim_{n \rightarrow \infty}^{\uparrow} m(A_n);$$

2) la **propriété de sous- $\sigma$ -additivité** : si  $(A_n)_{n \geq 1}$  est une **suite quelconque de parties mesurables**, on a :

$$m\left(\bigcup_{n \geq 1} A_n\right) \leq \sum_{n=1}^{\infty} m(A_n).$$

La propriété de limite croissante sera à la base d'un des théorèmes les plus importants, et l'un des plus utilisés, pour l'intégrale de Lebesgue : le *Théorème de convergence monotone*.

**Preuve.** 1) Posons :

$$\begin{cases} B_1 &= A_1 \\ B_n &= A_n \setminus A_{n-1}, \quad \text{pour } n \geq 2. \end{cases}$$

Les ensembles  $B_n$  sont mesurables et :

$$B_1 \cup \dots \cup B_n = A_n ;$$

donc :

$$\bigcup_{k \geq 1} B_k = \bigcup_{k \geq 1} A_k.$$

De plus, les  $B_k$ ,  $k \geq 1$ , sont *deux-à-deux disjoints*. Donc :

$$\begin{aligned} m\left(\lim_{k \rightarrow \infty} \uparrow A_k\right) &= m\left(\bigcup_{k \geq 1} B_k\right) = \sum_{k=1}^{\infty} m(B_k) = \lim_{n \rightarrow \infty} \uparrow \sum_{k=1}^n m(B_k) \\ &= \lim_{n \rightarrow \infty} \uparrow m\left(\bigcup_{k=1}^n B_k\right) = \lim_{n \rightarrow \infty} \uparrow m(A_n). \end{aligned}$$

2) Si les ensembles mesurables  $A_n$ ,  $n \geq 1$ , sont arbitraires, posons :

$$\begin{cases} B_1 &= A_1 \\ B_n &= A_n \setminus (A_1 \cup \dots \cup A_{n-1}), \quad \text{pour } n \geq 2. \end{cases}$$

Les  $B_n$  sont mesurables et l'on a  $B_n \subseteq A_n$ ,

$$\bigcup_{n \geq 1} B_n = \bigcup_{n \geq 1} A_n.$$

et les  $B_n$ ,  $n \geq 1$ , sont *deux-à-deux disjoints*. Donc :

$$m\left(\bigcup_{n \geq 1} A_n\right) = m\left(\bigcup_{n \geq 1} B_n\right) = \sum_{n=1}^{\infty} m(B_n) \leq \sum_{n=1}^{\infty} m(A_n),$$

puisque  $B_n \subseteq A_n$  entraîne  $m(B_n) \leq m(A_n)$ . □

Dans le cas d'une suite *décroissante* d'ensembles mesurables, on n'a pas en général de propriété de continuité pour la mesure. C'est toutefois le cas lorsque la mesure est *finie*.

**Proposition 4.13** Si  $\boxed{m(S) < +\infty}$ , alors, pour toute suite décroissante de parties mesurables  $(A_n)_{n \geq 1}$ , on a :

$$\boxed{m\left(\lim_{n \rightarrow \infty} \downarrow A_n\right) = \lim_{n \rightarrow \infty} \downarrow m(A_n)}.$$

**Remarque.** Cela reste vrai s'il existe  $n_0 \geq 1$  tel que  $m(A_{n_0}) < +\infty$ , en remplaçant  $S$  par  $A_{n_0}$ , et en remplaçant la suite  $(A_n)_{n \geq 1}$  par la suite  $(A_n)_{n \geq n_0}$ . Mais c'est faux en général si  $m(S) = +\infty$  et si tous les  $A_n$  sont de mesure infinie.

**Exemple.** Prenons la mesure de Lebesgue sur  $S = \mathbb{R}$  et  $A_n = ]n, +\infty[$ . La suite  $(A_n)_{n \geq 1}$  est décroissante et  $\lambda(\bigcap_{n \geq 1} A_n) = \lambda(\emptyset) = 0$ , alors que  $\lambda(A_n) = +\infty$  pour tout  $n \geq 1$ ; en effet, pour tout  $j \geq n$ , on a  $\lambda(A_n) \geq \lambda(]n, j]) = j - n$ .

Pour prouver cette proposition, montrons d'abord un lemme.

**Lemme 4.14** *Si les parties  $A$  et  $B$  sont mesurables et que l'on a  $A \subseteq B$  et*

$m(B) < +\infty$ , alors :

$$m(B \setminus A) = m(B) - m(A).$$

**Preuve.** On a  $A \cap (B \setminus A) = \emptyset$  et  $B = A \cup (B \setminus A)$ ; on obtient donc, par additivité :  $m(B) = m(A) + m(B \setminus A)$ .  $\square$

**Preuve de la proposition.** Comme la suite  $(A_n)_{n \geq 1}$  est décroissante, la suite des complémentaires  $(A_n^c)_{n \geq 1}$  est croissante; donc, en utilisant le lemme :

$$\begin{aligned} m\left(\bigcap_{n \geq 1} A_n\right) &= m\left(S \setminus \bigcup_{n \geq 1} A_n^c\right) = m(S) - m\left(\bigcup_{n \geq 1} A_n^c\right) \\ &= m(S) - \lim_{n \rightarrow \infty} \uparrow m(A_n^c) = m(S) - \lim_{n \rightarrow \infty} \uparrow [m(S) - m(A_n)] \\ &= \lim_{n \rightarrow \infty} \downarrow m(A_n); \end{aligned}$$

car on peut simplifier par  $m(S)$  puisque toutes les valeurs écrites sont réelles (c'est-à-dire  $< +\infty$ ).  $\square$

**Remarque.** Il résulte du Théorème 4.12, que, dans la définition des mesures  $\sigma$ -finies, on peut, *au choix*, prendre les parties de mesure finie  $S_n$ ,  $n \geq 1$ , soit *deux-à-deux disjointes*, soit *telles que la suite  $(S_n)_{n \geq 1}$  soit croissante*.

En effet, si  $(S_n)_{n \geq 1}$  est la suite de la définition, posons  $S'_1 = S_1$ , et, pour  $n \geq 2$ ,  $S'_n = S_n \setminus (S_1 \cup \dots \cup S_{n-1})$ ; alors les  $S'_n$  sont mesurables, de réunion  $S$ , de mesure finie (car  $S'_n \subseteq S_n$ ), et les  $S'_n$ ,  $n \geq 1$ , sont deux-à-deux disjoints. D'autre part, si l'on pose  $S''_n = S_1 \cup \dots \cup S_n$ , alors les  $S''_n$  sont mesurables, de réunion  $S$ , de mesure finie :  $m(S''_n) \leq m(S_1) + \dots + m(S_n) < +\infty$ , et la suite  $(S''_n)_{n \geq 1}$  est croissante.

#### 4.4 Quelques propriétés de la mesure de Lebesgue sur $\mathbb{R}$

Ces propriétés sont aussi vraies pour la mesure de Lebesgue sur  $\mathbb{R}^N$ , au lieu de  $\mathbb{R}$ , mais on ne les énoncera que pour  $\mathbb{R}$ , par souci de simplicité d'écriture.

Une propriété importante de la mesure de Lebesgue est qu'elle *ne charge pas les points* de  $\mathbb{R}$  (on dit aussi que c'est une *mesure diffuse*) :

**Proposition 4.15** Pour tout  $x \in \mathbb{R}$ , on a  $\lambda(\{x\}) = 0$ .

**Preuve.** Pour tout  $n \geq 1$ , on a  $\{x\} \subseteq ]x - \frac{1}{n}, x + \frac{1}{n}[$ ; donc :

$$0 \leq \lambda(\{x\}) \leq \lambda\left(]x - \frac{1}{n}, x + \frac{1}{n}[ \right) = \frac{2}{n};$$

on obtient bien  $\lambda(\{x\}) = 0$ . □

**Remarque.** Cette propriété n'est pas vraie pour toutes les mesures positives; par exemple, pour la mesure de Dirac en  $a \in \mathbb{R}$ , on a  $\delta_a(\{a\}) = 1$ .

**Corollaire 4.16** Pour toute partie dénombrable  $D$  de  $\mathbb{R}$ , on a  $\lambda(D) = 0$ .

En particulier :  $\lambda(\mathbb{Q}) = 0$ .

Comme  $\mathbb{N}$  et  $\mathbb{Z}$  sont dénombrables, on a aussi  $\lambda(\mathbb{N}) = 0$  et  $\lambda(\mathbb{Z}) = 0$ , mais on peut aussi le déduire du fait que  $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q}$ .

**Preuve.** Ecrivons  $D = \{d_n; n \geq 1\} = \bigcup_{n \geq 1} \{d_n\}$ . La sous- $\sigma$ -additivité donne :

$$0 \leq \lambda(D) \leq \sum_{n=1}^{\infty} \lambda(\{d_n\}) = \sum_{n=1}^{\infty} 0 = 0;$$

donc  $\lambda(D) = 0$ . □

**Remarque.** Il existe par contre des boréliens *non dénombrables* de  $\mathbb{R}$  de mesure (de Lebesgue) nulle.

On peut montrer que c'est, par exemple, le cas de l'ensemble de Cantor  $K$  (voir le Chapitre 4).

**Corollaire 4.17** Pour tout intervalle  $I$  de  $\mathbb{R}$ , la mesure de Lebesgue  $\lambda(I)$  de  $I$  est égale à sa longueur.

**Preuve.** 1) Si  $I$  est un intervalle borné. Si  $a$  et  $b$  ( $-\infty < a < b < +\infty$ ) sont ses extrémités, on a  $]a, b[ \subseteq I \subseteq [a, b]$ ; donc :

$$b - a = \lambda(]a, b[) \leq \lambda(I) \leq \lambda([a, b]),$$

et comme

$$\lambda([a, b]) = \lambda(\{a\} \cup ]a, b[ \cup \{b\}) = \lambda(\{a\}) + \lambda(]a, b[) + \lambda(\{b\}) = 0 + (b - a) + 0 = b - a,$$

on obtient  $\lambda(I) = b - a$ .

2) Si  $I$  est un intervalle infini. Alors sa longueur est  $+\infty$ . D'autre part :

- si  $I = \mathbb{R}$ , on a vu que  $\lambda(\mathbb{R}) = +\infty$ ;
- si  $I = a, +\infty[$  ou  $I = ]-\infty, b]$ , alors  $I \supseteq ]a, b[$ , donc  $\lambda(I) \geq (b - a)$ , pour tout  $b > a$ ; par conséquent  $\lambda(I) = +\infty$ ;

- si  $I = ] - \infty, a[$  ou  $I = ] - \infty, a]$ , on a de même  $\lambda(I) = +\infty$ .  $\square$

Notons que l'on retrouve ainsi (mais on triche en faisant cela, car on cache le théorème de construction de la mesure de Lebesgue, et notamment la preuve de la  $\sigma$ -additivité pour la longueur des intervalles, preuve qui nécessite l'usage du Théorème de Borel-Lebesgue, disant que les intervalles fermés bornés de  $\mathbb{R}$  sont compacts, théorème dont la preuve utilise en fait, implicitement, *via* l'existence des bornes supérieures des parties majorées non vides de  $\mathbb{R}$ , la non-dénombrabilité de  $\mathbb{R}$ ) que tout intervalle de  $\mathbb{R}$ , de longueur strictement positive, et  $\mathbb{R}$  lui-même en particulier, est non dénombrable.

**Théorème 4.18** *Pour tout ouvert non vide  $\Omega$  de  $\mathbb{R}$ , on a  $\lambda(\Omega) > 0$ .*

**Preuve.** Puisque  $\Omega$  n'est pas vide, il existe un point  $x_0 \in \Omega$ . Comme  $\Omega$  est ouvert, il existe  $\varepsilon > 0$  tel que  $]x_0 - \varepsilon, x_0 + \varepsilon[ \subseteq \Omega$ . Donc  $\lambda(\Omega) \geq \lambda(]x_0 - \varepsilon, x_0 + \varepsilon[) = 2\varepsilon > 0$ .  $\square$

**Corollaire 4.19** *Si  $B$  est un borélien de  $\mathbb{R}$  tel que  $\lambda(B) = 0$ , alors  $B^c$  est dense dans  $\mathbb{R}$ .*

**Preuve.** Soit  $\overset{\circ}{B}$  l'intérieur de  $B$ ; c'est un ouvert, et  $\lambda(\overset{\circ}{B}) = 0$ , puisque  $\overset{\circ}{B} \subseteq B$ . Donc, par le Théorème 4.18,  $\overset{\circ}{B} = \emptyset$ . Alors  $\overline{B^c} = [\overset{\circ}{B}]^c = \mathbb{R}$ .  $\square$

Une des propriétés les plus importantes de la mesure de Lebesgue est qu'elle est invariante par translation. On peut en effet montrer que c'est la *seule* mesure positive invariante par translation (si on la normalise en demandant que l'intervalle  $]0, 1[$  soit de mesure 1).

**Théorème 4.20** *La mesure de Lebesgue est invariante par translation : pour tout borélien  $B$  de  $\mathbb{R}$ , on a, pour tout  $a \in \mathbb{R}$  :*

$$\lambda(B + a) = \lambda(B).$$

**Preuve.** Rappelons que  $B + a = \{x + a; x \in B\}$ .

Pour  $a \in \mathbb{R}$  fixé, posons :

$$\mu(B) = \lambda(B + a), \quad \forall B \in \mathcal{B}or(\mathbb{R}).$$

Alors  $\mu$  est une mesure positive sur  $\mathbb{R}$  (c'est la mesure-image de  $\lambda$  par la translation  $t_a: x \in \mathbb{R} \mapsto x - a \in \mathbb{R}$ , qui est une application continue, donc borélienne), et l'on a :

$$\mu(] \alpha, \beta [) = \lambda(] \alpha + a, \beta + a [) = ((\beta + a) - (\alpha + a)) = \beta - \alpha, \quad \forall \alpha < \beta.$$

Par la propriété d'unicité de la mesure de Lebesgue, on a  $\mu = \lambda$ . □

**Théorème 4.21** *La mesure de Lebesgue est invariante par symétrie : pour tout borélien  $B$  de  $\mathbb{R}$ , on a :*

$$\lambda(-B) = \lambda(B).$$

**Preuve.** Posons, pour tout borélien  $B$  de  $\mathbb{R}$ ,  $\nu(B) = \lambda(-B)$ . On définit ainsi une mesure positive sur  $\mathbb{R}$  (c'est la mesure-image de  $\lambda$  par l'application continue  $s: x \in \mathbb{R} \mapsto -x \in \mathbb{R}$ ), et :

$$\nu(] \alpha, \beta [) = \lambda(] -\beta, -\alpha [) = ((-\alpha) - (-\beta)) = \beta - \alpha$$

pour tous  $\alpha < \beta$ , et l'unicité de la mesure de Lebesgue assure que  $\nu = \lambda$ . □