Université des Sciences et Technologies de Lille
U.F.R. de Mathématiques Pures et Appliquées
Bât. M2, F-59655 Villeneuve d'Ascq Cedex

Intégration
Analyse de Fourier
Probabilités

Charles SUQUET

Licence de Mathématiques 2003–2004

Chapitre 1

Mesures

Ce chapitre introduit les mesures sur un ensemble abstrait et leurs propriétés générales. On discute aussi la construction de ces mesures sur $\mathbb{R}$ et $\mathbb{R}^d$. L'ensemble de définition d'une mesure est une *tribu* et ceci nous amène à étudier ce type de structure.

1.1 Prologue

Définition 1.1. *Soit* $\mathcal{F}$ *une* tribu *de parties de* Ω. *On appelle* mesure positive *sur* $(\Omega, \mathcal{F})$ *une application*

$$\mu : \mathcal{F} \longrightarrow [0, +\infty],$$

vérifiant

a) $\mu(\emptyset) = 0$ *;*

b) μ *est* σ*-additive : pour toute suite* $(A_i)_{i\in\mathbb{N}^*}$ *d'éléments de* $\mathcal{F}$ *deux à deux disjoints,*

$$\mu\Big(\bigcup_{i\in\mathbb{N}^*} A_i\Big) = \sum_{i=1}^{+\infty} \mu(A_i). \tag{1.1}$$

La définition de la structure de tribu est précisée ci-dessous. D'ores et déjà, en raison de (1.1), une tribu doit posséder l'ensemble vide et être stable par union dénombrable. Remarquons que l'union des A_i est invariante par permutation sur les indices i et aussi sous l'effet de regroupements en paquets des A_i. La cohérence de la définition de la σ-additivité utilise donc implicitement, mais fortement, les propriétés des séries à termes dans $\overline{\mathbb{R}}_+$. Si l'on voulait définir des mesures à signe ou à valeurs complexes, il faudrait ajouter dans la définition la condition de sommabilité de la famille des nombres $\mu(A_i)$. Dans ce cours, nous nous limiterons aux mesures à valeurs dans $\overline{\mathbb{R}}_+$ que nous désignerons désormais simplement par le mot mesure.

Dans la définition 1.1, la propriété importante est clairement b). On pourrait remplacer a) par $\mu(\emptyset) < +\infty$, car en prenant chaque A_i égal à l'ensemble vide dans (1.1), on en déduirait $\mu(\emptyset) = 0$.

Comme « fonction d'ensembles », une mesure n'est pas nécessairement définie sur la famille $\mathcal{P}(\Omega)$ de toutes les parties de Ω. Elle a un ensemble de définition qui est généralement plus petit, c'est la tribu $\mathcal{F}$. Pour $A \in \mathcal{F}$, $\mu(A)$ s'appelle mesure de l'ensemble A.

Si $\mu(\Omega) = 1$, la mesure μ est appelée *probabilité* sur $(\Omega, \mathcal{F})$ et la condition a) devient superflue (prendre $A_1 = \Omega$ et $A_i = \emptyset$ pour $i \geq 2$ dans (1.1)).

Définition 1.2. *Une famille $\mathcal{F}$ de parties de Ω est appelée* tribu *(ou σ-algèbre) sur Ω si elle*

a) possède l'ensemble vide : $\emptyset \in \mathcal{F}$;

b) est stable par complémentaire : $\forall A \in \mathcal{F},\ A^c \in \mathcal{F}$;

c) est stable par union dénombrable : $(\forall i \in \mathbb{N}^*,\ A_i \in \mathcal{F}) \Rightarrow \bigcup_{i \in \mathbb{N}^*} A_i \in \mathcal{F}$.

La définition d'une mesure présuppose la donnée d'une tribu $\mathcal{F}$. Par contre, on peut très bien définir une tribu sans parler de mesure, ou définir plusieurs mesures sur la même tribu.

1.2 Tribus

1.2.1 Exemples immédiats

Les deux exemples les plus simples de tribu sur un ensemble Ω sont $\mathcal{F}_0 := \{\emptyset, \Omega\}$ et $\mathcal{F}_1 = \mathcal{P}(\Omega)$, famille de tous les sous-ensembles de Ω. Ce sont des cas extrémaux puisqu'il est clair que toute tribu $\mathcal{F}$ sur Ω vérifie $\{\emptyset, \Omega\} \subset \mathcal{F} \subset \mathcal{P}(\Omega)$.

Si A est une partie de Ω, la famille

$$\mathcal{F} := \{\emptyset, A, A^c, \Omega\}$$

est une tribu sur Ω. En effet, elle vérifie par construction les propriétés a) et b) de la définition 1.2. Regardons la stabilité par union dénombrable en prenant une suite quelconque $(A_i)_{i\geq 1}$ d'éléments de $\mathcal{F}$ et en notant E leur réunion. Notons $\mathcal{H}$ l'ensemble de toutes les valeurs prises par la suite $(A_i)_{i\geq 1}$. Il y a 16 possibilités pour $\mathcal{H}$ (autant que de parties de $\mathcal{F}$, soit 2^4). On a seulement 4 cas pour la valeur de E :

- $\mathcal{H} = \{\emptyset\}$, $E = \emptyset$;
- $\mathcal{H} = \{\emptyset, A\}$ ou $\{A\}$, alors $E = A$;
- $\mathcal{H} = \{\emptyset, A^c\}$ ou $\{A^c\}$, alors $E = A^c$;
- dans tous les autres cas, $\mathcal{H}$ a pour éléments A et A^c simultanément ou Ω, alors $E = \Omega$.

Ainsi on trouve toujours pour E un élément de $\mathcal{F}$, qui est donc bien stable par union dénombrable.

En raison des propriétés a) et b), $\mathcal{F}$ est incluse dans toute tribu $\mathcal{G}$ ayant A comme élément. On l'appelle tribu engendrée par A. Rien que sur cet exemple qui est le plus simple possible en dehors des tribus $\mathcal{F}_0$ et $\mathcal{F}_1$ ci-dessus, on perçoit qu'il peut être très laborieux de vérifier qu'un ensemble même fini est une tribu. Les masochistes pourront s'exercer en déterminant manuellement la plus petite tribu $\mathcal{F}'$ contenant deux parties A et B de Ω telles que $A \neq B$ et $A \neq B^c$ ($\mathcal{F}'$ a 16 éléments et une fois qu'on les a trouvés, il reste à vérifier la stabilité par union dénombrable). Pour les autres, il est conseillé de résoudre cet exercice à l'aide de la proposition 1.7 ci-dessous.

1.2.2 Opérations ensemblistes dans une tribu

Une conséquence de la définition d'une tribu $\mathcal{F}$ est *grosso modo* qu'en effectuant des opérations ensemblistes finies ou dénombrables sur les éléments de $\mathcal{F}$, on ne sort jamais de $\mathcal{F}$. Cet énoncé est bien sûr trop flou pour pouvoir être démontré et doit seulement être considéré comme un guide pratique. En pratique, pour démontrer l'appartenance à $\mathcal{F}$, on dispose outre la définition, des propriétés suivantes.

Proposition 1.3. *Soit $\mathcal{F}$ une tribu sur l'ensemble Ω.*

a) $\mathcal{F}$ est stable par union finie ;

b) $\mathcal{F}$ est stable par intersections dénombrables ou finies ;

c) si $A \in \mathcal{F}$ et $B \in \mathcal{F}$, $A \setminus B$ et $A\Delta B$ sont dans $\mathcal{F}$.

Démonstration. Soit $A_1, \dots, A_n$ une suite finie d'éléments de $\mathcal{F}$. Pour vérifier l'appartenance à $\mathcal{F}$ de $A_1 \cup \dots \cup A_n$, il suffit de compléter cette suite en posant pour $i > n$, $A_i := \emptyset$ et d'appliquer la stabilité de $\mathcal{F}$ par union dénombrable.

Pour la stabilité par intersection dénombrable, il suffit de remarquer que pour toute suite $(A_i)_{i\in\mathbb{N}^*}$ d'éléments de $\mathcal{F}$,

$$\bigcap_{i\in\mathbb{N}^*} A_i = \left(\left(\bigcap_{i\in\mathbb{N}^*} A_i\right)^c\right)^c = \left(\bigcup_{i\in\mathbb{N}^*} A_i^c\right)^c,$$

et d'utiliser la stabilité de $\mathcal{F}$ par complémentaire et par union dénombrable. La stabilité par intersection finie s'en déduit en prenant tous les A_i égaux à Ω à partir d'un certain rang.

Pour le c), il suffit de rappeler que $A \setminus B = A \cap B^c$ et $A\Delta B = (A\cap B^c)\cup(A^c\cap B)$. □

Le résultat suivant est parfois utile pour vérifier pratiquement qu'une famille donnée est une tribu.

Proposition 1.4. *La famille non vide $\mathcal{F}$ de parties de Ω est une tribu sur Ω si et seulement si elle satisfait aux trois conditions :*

i) stabilité par complémentaire : $\forall A \in \mathcal{F}$, $A^c \in \mathcal{F}$;

ii) stabilité par intersections finies : $\forall A, B \in \mathcal{F}$, $A \cap B \in \mathcal{F}$;

iii) stabilité par réunion dénombrable disjointe *: pour toute suite $(E_i)_{i\in\mathbb{N}}$ d'éléments deux à deux disjoints de $\mathcal{F}$, $\bigcup_{i\in\mathbb{N}} E_i \in \mathcal{F}$.*

En raison de l'associativité de l'intersection : $(A\cap B)\cap C = A\cap B\cap C$, la stabilité pour l'intersection de *deux* éléments A et B quelconques de $\mathcal{F}$ est équivalente à la stabilité pour l'intersection d'un nombre fini $n \geq 2$ quelconque d'éléments de $\mathcal{F}$. En pratique pour vérifier la stabilité par intersection finie, il suffit donc de traiter le cas $n = 2$.

Démonstration. La condition i) est déjà dans la définition d'une tribu ; ii) est nécessaire par la prop. 1.3 b) ; iii) est un cas particulier de la stabilité par union dénombrable. Ainsi la nécessité des conditions i), ii) et iii) est claire et il nous reste à montrer leur suffisance.

Puisque $\mathcal{F}$ est non vide, il y a au moins un élément A dans $\mathcal{F}$, alors par i), $A^c \in \mathcal{F}$ et par ii), $A \cap A^c \in \mathcal{F}$, d'où $\emptyset \in \mathcal{F}$. Il reste donc à montrer la stabilité de $\mathcal{F}$ par union dénombrable quelconque. Auparavant deux remarques nous seront utiles. Si A et B sont dans $\mathcal{F}$, $B \setminus A = B \cap A^c$ aussi. En combinant i) et ii), on obtient la stabilité de $\mathcal{F}$ par réunion finie.

Soit $(A_i)_{i\in\mathbb{N}}$ une suite quelconque dans $\mathcal{F}$. Posons pour tout $n \in \mathbb{N}$,

$$B_n := \bigcup_{i=0}^{n} A_i, \quad B := \bigcup_{i\in\mathbb{N}} A_i.$$

Il est clair que les suites $(A_i)_{i\in\mathbb{N}}$ et $(B_n)_{n\in\mathbb{N}}$ ont même réunion, le problème est donc de vérifier l'appartenance de B à $\mathcal{F}$.

Admettons, pour l'instant, que pour tout $n \geq 1$, B_n vérifie la décomposition en union disjointe :

$$B_n = B_0 \cup \left(\bigcup_{i=1}^{n} (B_i \setminus B_{i-1}) \right).$$

En écrivant la réunion infinie des B_n à l'aide de cette décomposition et en « effaçant » toutes les répétitions des $B_i \setminus B_{i-1}$, on en déduit immédiatement que B vérifie la décomposition en union dénombrable disjointe :

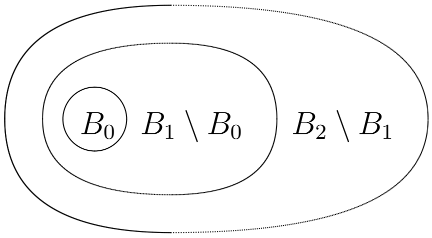
$$B = B_0 \cup \left(\bigcup_{i\in\mathbb{N}^*} (B_i \setminus B_{i-1}) \right).$$

Par stabilité de $\mathcal{F}$ par union finie, chaque B_n est dans $\mathcal{F}$. Alors chaque $B_i \setminus B_{i-1}$ est dans $\mathcal{F}$ (stabilité par différence). Enfin la réunion dénombrable disjointe des $B_i \setminus B_{i-1}$ est dans $\mathcal{F}$.

Ainsi pour compléter la preuve, il ne reste plus qu'à justifier la décomposition de B_n. Posons :

$$D_n = B_0 \cup \left(\bigcup_{i=1}^{n} (B_i \setminus B_{i-1}) \right).$$

Pour montrer que $B_n = D_n$, il suffit de montrer que $D_n \subset B_n$ et $B_n \subset D_n$. La première inclusion est évidente puisque pour tout $i \leq n$, $B_i \setminus B_{i-1} \subset B_i \subset B_n$. Pour prouver l'inclusion inverse, on note ω un élément *quelconque* de B_n et on montre que ω appartient à D_n. Soit $i_0 = i_0(\omega)$ le plus petit des indices i tels que $\omega \in B_i$. Comme cet ensemble d'indices contient au moins n, on a $0 \leq i_0 \leq n$. Si $i_0 = 0$, $\omega \in B_0$ et comme $B_0 \subset D_n$, $\omega \in D_n$. Si $i_0 \geq 1$, par la définition même de i_0, on a $\omega \in B_{i_0}$ et $\omega \notin B_{i_0-1}$, donc $\omega \in B_{i_0} \setminus B_{i_0-1}$ et comme $i_0 \leq n$, $B_{i_0} \setminus B_{i_0-1} \subset D_n$ donc $\omega \in D_n$. Le raisonnement précédent étant valable pour tout ω de B_n, on en déduit $B_n \subset D_n$.



$\square$

1.2.3 Génération de tribus

Nous examinons maintenant plusieurs méthodes de construction de tribus qui auront toutes leur utilité dans la suite de ce cours.

Commençons par un résultat négatif qu'il convient de connaître : si $\mathcal{F}_1$ et $\mathcal{F}_2$ sont des tribus sur Ω, $\mathcal{F}_1 \cup \mathcal{F}_2$ n'est en général pas une tribu. Voici un contre-exemple simple. Prenons $\Omega := \{1, 2, 3\}$ et pour $i = 1, 2$, $A_i := \{i\}$, $\mathcal{F}_i := \{\emptyset, A_i, A_i^c, \Omega\}$. La réunion de ces deux tribus est

$$\mathcal{F}_1 \cup \mathcal{F}_2 = \big\{\emptyset, \{1\}, \{2\}, \{2, 3\}, \{1, 3\}, \Omega\big\}.$$

On constate que A_1 et A_2 sont éléments de $\mathcal{F}_1 \cup \mathcal{F}_2$, alors que $A_1 \cup A_2 = \{1, 2\}$ n'est pas élément de $\mathcal{F}_1 \cup \mathcal{F}_2$.

La situation est bien plus agréable pour les intersections de tribus.

Proposition 1.5. *Si I est une famille* quelconque *d'indices et $(\mathcal{F}_i)_{i \in I}$ une famille de tribus sur Ω, alors $\bigcap_{i \in I} \mathcal{F}_i$ est une tribu sur Ω.*

Vouloir démontrer une proposition aussi évidente reviendrait simplement à expliquer sur cet exemple ce qu'est une intersection. Noter que l'on ne suppose pas I dénombrable. Cette proposition permet notamment de justifier la définition suivante (noter l'analogie formelle avec l'intersection d'espaces vectoriels et le sous-espace vectoriel engendré par une famille de vecteurs).

Définition 1.6. *Soit $\mathcal{S}$ une famille quelconque de parties de Ω ($\mathcal{S} \subset \mathcal{P}(\Omega)$). On appelle tribu engendrée par $\mathcal{S}$ et on note $\sigma(\mathcal{S})$, l'intersection de toutes les tribus sur Ω contenant $\mathcal{S}$. C'est la plus petite (au sens de l'inclusion) des tribus $\mathcal{F}$ telles que $\mathcal{S} \subset \mathcal{F}$.*

Au risque d'insister lourdement, notons comme conséquence pratique de cette définition *l'argument de minimalité* que nous utiliserons intensivement :

$$(\mathcal{S} \subset \mathcal{F} \text{ et } \mathcal{F} \text{ tribu}) \Rightarrow (\sigma(\mathcal{S}) \subset \mathcal{F}). \tag{1.2}$$

En général, on ne sait pas donner une description constructive d'une tribu infinie différente de $\mathcal{P}(\Omega)$. Le résultat suivant est une exception notable.

Proposition 1.7. *Si I est fini ou dénombrable et si $\Pi = \{A_i; i \in I\}$ est une* partition *de Ω, alors*

$$\sigma(\Pi) = \Big\{ \bigcup_{i \in J} A_i;\ J \subset I \Big\}.$$

Rappelons que par convention, une union indexée par l'ensemble vide est elle même vide. Une partition de Ω est une famille de parties dont aucune n'est vide, deux à deux disjointes et dont la réunion est Ω. La proposition 1.7 pourra être démontrée en T.D.

Nous examinons maintenant le « transport » de tribus par une application. Commençons par quelques précisions sur les images ensemblistes. Dans ce qui suit, on fixe deux ensembles E et F et une application *quelconque* $f : E \to F$. À partir de f, on va construire des applications opérant à 3 niveaux de structures. Nous noterons provisoirement ces niveaux par un indice avec la signification suivante :

0 : on opère sur des éléments de E ou F (niveau E, F),

1 : on opère sur des parties de E ou F, (niveau $\mathcal{P}(E)$, $\mathcal{P}(F)$),

2 : on opère sur des *familles* de parties de E ou F (niveau $\mathcal{P}\big(\mathcal{P}(E)\big)$, $\mathcal{P}\big(\mathcal{P}(F)\big)$).

Au niveau 0, on pose simplement $f_0 := f$. Ensuite au niveau 1, on définit les images ensemblistes directe et réciproque par

$$\begin{aligned} \forall A \in \mathcal{P}(E), \quad f_1(A) &:= \{f_0(x) \in F;\ x \in A\}; \\ \forall B \in \mathcal{P}(F), \quad f_1^{-1}(B) &:= \{x \in E;\ f_0(x) \in B\}. \end{aligned}$$

En particulier $f_1(\emptyset) = \emptyset$ et $f_1^{-1}(\emptyset) = \emptyset$. Il importe de noter que f_1^{-1} n'est pas une application réciproque au sens classique : $f_1^{-1}(B)$ est défini via f_0 pour tout $B \in \mathcal{P}(F)$ et f_1 peut très bien ne pas être bijective. Remarquons ici qu'on a toujours $f_1\big(f_1^{-1}(B)\big) \subset B$ et que l'inclusion peut être stricte (prendre f_0 non surjective et B non inclus dans $f_1(E)$).

Passons au niveau 2, où nous définirons seulement f_2^{-1} en posant pour toute famille $\mathcal{B}$ de parties de F (donc $\mathcal{B} \subset \mathcal{P}(F)$ ou encore $\mathcal{B} \in \mathcal{P}\big(\mathcal{P}(F)\big)$) :

$$f_2^{-1}(\mathcal{B}) := \{A \in \mathcal{P}(E);\ \exists B \in \mathcal{B}, A = f_1^{-1}(B)\} = \{f_1^{-1}(B); B \in \mathcal{B}\}.$$

Signalons ici le piège : $f_2^{-1}(\mathcal{B})$ n'est pas forcément égal à $\{A \in \mathcal{P}(E);\ f_1(A) \in \mathcal{B}\}$. Voici un contre-exemple. On prend $f = f_0 : [-\pi, \pi] \to \mathbb{R}$, $x \mapsto \sin x$ et $\mathcal{B} = \{[0, b],\ b \in]1, +\infty[\}$. Alors $f_2^{-1}(\mathcal{B}) = \{[0, \pi]\}$, mais $\{A \in \mathcal{P}([-\pi, \pi]); f_1(A) \in \mathcal{B}\} = \emptyset$.

Dans l'usage courant, que nous adoptons désormais, on utilise la même lettre f pour désigner f_0, f_1 ou f_2, le contexte permettant de comprendre à quel niveau on se situe. C'est assez facile pour peu que l'on respecte les conventions typographiques désignant par des minuscules italiques comme x, y, ω, les éléments de E ou F, par des majuscules italiques A, B, E, les sous-ensembles de E ou F et par des majuscules cursives $\mathcal{A}$, $\mathcal{B}$, $\mathcal{F}$, $\mathcal{P}$, les familles d'ensembles (familles de parties de E ou F).

Proposition 1.8. *Soient E et F des ensembles et $f : E \to F$ une application. L'inverse ensembliste f^{-1} vérifie :*

a) pour tout $B \subset F$, $f^{-1}(B^c) = \big(f^{-1}(B)\big)^c$;

b) pour toute famille $(B_i)_{i\in I}$ *de parties de* F, $f^{-1}\Bigl(\bigcup_{i\in I} B_i\Bigr) = \bigcup_{i\in I} f^{-1}(B_i)$.

c) pour toute famille $(B_i)_{i\in I}$ *de parties de* F, $f^{-1}\Bigl(\bigcap_{i\in I} B_i\Bigr) = \bigcap_{i\in I} f^{-1}(B_i)$.

Démonstration. Dans le a), les deux complémentaires utilisés n'ont pas la même signification : $B^c = F \setminus B$ tandis que $\bigl(f^{-1}(B)\bigr)^c = E \setminus f^{-1}(B)$. On vérifie a) ainsi :

$$\begin{aligned} f^{-1}(B^c) &= \{x \in E; f(x) \in B^c\} \\ &= \{x \in E; f(x) \notin B\} \\ &= \{x \in E; x \notin f^{-1}(B)\} \\ &= E \setminus f^{-1}(B) = \bigl(f^{-1}(B)\bigr)^c. \end{aligned}$$

Pour montrer b), on écrit

$$\begin{aligned} f^{-1}\Bigl(\bigcup_{i\in I} B_i\Bigr) &= \Bigl\{x \in E; f(x) \in \bigcup_{i\in I} B_i\Bigr\} \\ &= \{x \in E;\ \exists i \in I, f(x) \in B_i\} \\ &= \{x \in E;\ \exists i \in I, x \in f^{-1}(B_i)\} \\ &= \bigcup_{i\in I} f^{-1}(B_i). \end{aligned}$$

La vérification de c) est analogue en remplaçant « $\exists i$ » par « $\forall i$ ». □

Remarque 1.9. La situation est moins confortable pour l'image ensembliste directe. On pourra vérifier en exercice que $f\Bigl(\bigcup_{i\in I} A_i\Bigr) = \bigcup_{i\in I} f(A_i)$. Par contre l'image directe ne commute pas en général avec l'intersection. Comme contre-exemple, on peut prendre $E = \mathbb{R}$, $F = \mathbb{R}^+$, $f : x \mapsto x^2$, $A_1 = [-1, 0[$ et $A_2 =]0, 1]$. Ainsi $f(A_1 \cap A_2) = f(\emptyset) = \emptyset$, alors que $f(A_1) \cap f(A_2) =]0, 1]$. Elle ne commute pas davantage avec le passage au complémentaire, contre-exemple : $E = F = \mathbb{R}$, $f(x) = x^2$, $A =]0, +\infty[$, on trouve $f(A^c) = [0, +\infty[$ et $f(A)^c =]-\infty, 0]$.

Proposition 1.10. *Soient* E *et* F *des ensembles et* $f : E \to F$ *une application.*

i) Si $\mathcal{B}$ *est une tribu sur* F, $f^{-1}(\mathcal{B})$ *est une tribu sur* E.

ii) Si $\mathcal{A}$ *est une tribu sur* E, $\mathcal{T} := \{B \in \mathcal{P}(F);\ f^{-1}(B) \in \mathcal{A}\}$ *est une tribu sur* F. *On l'appelle tribu image de* $\mathcal{A}$ *par* f *et on la note* $f(\mathcal{A})$.

Preuve de i). L'ensemble vide est bien un élément de $f^{-1}(\mathcal{B})$, car $\mathcal{B}$ étant une tribu, $\emptyset \in \mathcal{B}$ et $f^{-1}(\emptyset) = \emptyset$. Pour vérifier la stabilité par complémentaire, soit A un élément quelconque de $f^{-1}(\mathcal{B})$. Cela signifie qu'il existe $B \in \mathcal{B}$ tel que $A = f^{-1}(B)$. Or $A^c = f^{-1}(B^c)$ et la tribu $\mathcal{B}$ est stable par complémentaire donc $B^c \in \mathcal{B}$ et ainsi $A^c \in f^{-1}(\mathcal{B})$. Enfin, pour vérifier la stabilité par union dénombrable, soit $(A_i)_{i\geq 1}$ une suite d'éléments de $f^{-1}(B)$. Il existe alors une suite $(B_i)_{i\geq 1}$ dans $\mathcal{B}$ telle que pour tout $i \geq 1$, $A_i = f^{-1}(B_i)$. On peut alors écrire

$$\bigcup_{i\geq 1} A_i = \bigcup_{i\geq 1} f^{-1}(B_i) = f^{-1}\Bigl(\bigcup_{i\geq 1} B_i\Bigr) \in f^{-1}(\mathcal{B}),$$

car la tribu $\mathcal{B}$ est stable par union dénombrable. □

Preuve de ii). Laissée en exercice. □

Proposition 1.11. *Soient E et F des ensembles, $f : E \to F$ une application et $\mathcal{S} \subset \mathcal{P}(F)$ une famille quelconque de parties de F. Alors*

$$f^{-1}\bigl(\sigma(\mathcal{S})\bigr) = \sigma\bigl(f^{-1}(\mathcal{S})\bigr). \tag{1.3}$$

Démonstration. On prouve l'égalité (1.3) en montrant l'inclusion dans les deux sens.

a) Vérification de $\sigma\bigl(f^{-1}(\mathcal{S})\bigr) \subset f^{-1}\bigl(\sigma(\mathcal{S})\bigr)$. De l'inclusion $\mathcal{S} \subset \sigma(\mathcal{S})$, on tire $f^{-1}(\mathcal{S}) \subset f^{-1}\bigl(\sigma(\mathcal{S})\bigr)$. Par la proposition 1.10 i), $f^{-1}\bigl(\sigma(\mathcal{S})\bigr)$ est une tribu. Par minimalité (voir (1.2)) on en déduit

$$\sigma\bigl(f^{-1}(\mathcal{S})\bigr) \subset f^{-1}\bigl(\sigma(\mathcal{S})\bigr). \tag{1.4}$$

b) Vérification de $f^{-1}\bigl(\sigma(\mathcal{S})\bigr) \subset \sigma\bigl(f^{-1}(\mathcal{S})\bigr)$. Posons

$$\mathcal{A} := \sigma\bigl(f^{-1}(\mathcal{S})\bigr), \qquad \mathcal{F} := \{B \in \mathcal{P}(F);\ f^{-1}(B) \in \mathcal{A}\}.$$

Par la proposition 1.10 ii), $\mathcal{F}$ est une tribu sur F (tribu image de $\mathcal{A}$ par f). On voit que

$$\mathcal{S} \subset \mathcal{F},$$

car pour tout $B \in \mathcal{S}$, $f^{-1}(B) \in f^{-1}(\mathcal{S}) \subset \sigma\bigl(f^{-1}(\mathcal{S})\bigr) = \mathcal{A}$. En invoquant à nouveau l'argument de minimalité (1.2), on en déduit que $\sigma(\mathcal{S}) \subset \mathcal{F}$, d'où

$$f^{-1}\bigl(\sigma(\mathcal{S})\bigr) \subset f^{-1}(\mathcal{F}).$$

Remarquons maintenant que $f^{-1}(\mathcal{F}) \subset \mathcal{A}$ en raison de la définition de $\mathcal{F}$ puisque tout $B \in \mathcal{F}$ est tel que $f^{-1}(B) \in \mathcal{A}$. Finalement on a

$$f^{-1}\bigl(\sigma(\mathcal{S})\bigr) \subset f^{-1}(\mathcal{F}) \subset \mathcal{A} = \sigma\bigl(f^{-1}(\mathcal{S})\bigr). \tag{1.5}$$

L'égalité (1.3) découle de (1.4) et (1.5). □

1.2.4 Le théorème π-λ

Pour montrer une propriété relative à une tribu $\mathcal{F}$, il est souvent plus commode de montrer cette propriété pour une famille $\mathcal{L}$ de parties de Ω plus grande que $\mathcal{F}$ et d'obtenir le résultat souhaité par restriction à $\mathcal{F}$. Un outil important dans cette perspective est le théorème de Dynkin, connu aussi sous le nom de théorème π-λ.

Définition 1.12. *On dit que $\mathcal{P}$ est une π-classe (ou π-système) de parties de Ω si elle est stable pour l'intersection finie : $\forall A, B \in \mathcal{P}$, $A \cap B \in \mathcal{P}$.*

Définition 1.13. *Une famille $\mathcal{L}$ de parties de Ω est appelée λ-classe (ou λ-système) si elle vérifie :*

a) $\Omega \in \mathcal{L}$ *;*

b) stabilité par différence propre : si $A, B \in \mathcal{L}$ et $A \subset B$, alors $B \setminus A \in \mathcal{L}$;

c) *stabilité par union dénombrable croissante : si* $(A_k)_{k\in\mathbb{N}}$ *est une suite dans* $\mathcal{L}$ *telle que* $\forall k \in \mathbb{N}$, $A_k \subset A_{k+1}$, *l'ensemble* $A := \bigcup_{k\in\mathbb{N}} A_k$ *est élément de* $\mathcal{L}$.

Toute tribu est à la fois une π-classe et une λ-classe (cf. prop. 1.3). La réciproque est vraie.

Proposition 1.14. *Si la famille* $\mathcal{F}$ *de parties de* Ω *est à la fois une* π*-classe et une* λ*-classe, c'est une tribu.*

Démonstration. $\mathcal{F}$ est stable par complémentaire puisque $\Omega \in \mathcal{F}$ (déf. 1.13-a)) et $A^c = \Omega \setminus A$ (déf. 1.13-b)). En particulier, $\emptyset = \Omega^c$ est élément de $\mathcal{F}$. En combinant la stabilité de $\mathcal{F}$ par intersection finie (comme π-classe) et la stabilité par complémentaire on obtient la stabilité par union finie. Soit $(B_k)_{k\in\mathbb{N}}$ une suite *quelconque* dans $\mathcal{F}$. Pour tout $n \in \mathbb{N}$, l'union finie $A_n := \cup_{0\leq k\leq n} B_k$ est dans $\mathcal{F}$. De plus la suite (A_n) est *croissante* pour l'inclusion et a pour réunion $A = \cup_{k\in\mathbb{N}} B_k$. En raison de déf. 1.13-c), $\cup_{k\in\mathbb{N}} B_k \in \mathcal{F}$, ce qui établit la stabilité de $\mathcal{F}$ par union dénombrable. $\mathcal{F}$ est donc bien une tribu. □

Théorème 1.15 (Dynkin). *Si une* λ*-classe* $\mathcal{L}$ *de parties de* Ω *contient une* π*-classe* $\mathcal{P}$, *elle contient aussi la tribu engendrée par* $\mathcal{P}$.

Démonstration. Il est facile de voir que l'intersection d'une familllle quelconque non vide de λ-classes de parties de Ω est encore une λ-classe de parties de Ω. De plus $\mathcal{P}(\Omega)$ est une λ-classe. La famille de toutes les λ-classes contenant $\mathcal{P}$ est donc non vide et on peut définir $\Lambda(\mathcal{P})$, λ-classe engendrée par $\mathcal{P}$, comme l'intersection de toutes les λ-classes contenant $\mathcal{P}$. On va montrer que $\Lambda(\mathcal{P}) = \sigma(\mathcal{P})$ et comme $\Lambda(\mathcal{P}) \subset \mathcal{L}$, ceci prouvera le théorème.

Comme $\sigma(\mathcal{P})$ est une λ-classe et contient $\mathcal{P}$, il est clair que $\Lambda(\mathcal{P}) \subset \sigma(\mathcal{P})$. Il reste à prouver l'inclusion inverse. Pour cela, il suffit de montrer que la λ-classe $\Lambda(\mathcal{P})$ est aussi une π-classe car alors elle sera une tribu (prop. 1.14) contenant $\mathcal{P}$, donc contenant aussi la tribu $\sigma(\mathcal{P})$.

Il s'agit donc de prouver que si A et B sont deux éléments quelconques de $\Lambda(\mathcal{P})$, $A \cap B \in \Lambda(\mathcal{P})$. Définissons pour tout $C \in \mathcal{P}(\Omega)$ la famille de parties

$$\mathcal{G}_C := \{D \in \mathcal{P}(\Omega);\ C \cap D \in \Lambda(\mathcal{P})\}.$$

Vérifions d'abord que si $C \in \Lambda(\mathcal{P})$, $\mathcal{G}_C$ est une λ-classe. Ω est bien élément de $\mathcal{G}_C$ puisque $C \cap \Omega = C \in \Lambda(\mathcal{P})$. Si $D, E \in \mathcal{G}_C$ et $D \subset E$, alors $(E \setminus D) \cap C = (E \cap C) \setminus (D \cap C)$ est bien élément de $\Lambda(\mathcal{P})$ en appliquant la propriété b) de la définition 1.13 avec $D \cap C$ et $E \cap C$. Ainsi, $E \setminus D$ satisfait la condition d'appartenance à $\mathcal{G}_C$. Ceci établit la stabilité de $\mathcal{G}_C$ par différence propre. Soit (D_n) une suite dans $\mathcal{G}_C$, croissante pour l'inclusion et D sa réunion. Alors $(C \cap D_n)$ est une suite dans $\Lambda(\mathcal{P})$, croissante et de réunion $C \cap D$. Comme $\Lambda(\mathcal{P})$ est une λ-classe, $C \cap D \in \Lambda(\mathcal{P})$, donc D est élément de $\mathcal{G}_C$ et ceci établit la stabilité de $\mathcal{G}_C$ par union dénombrable croissante.

Si $C \in \mathcal{P}$, alors $\mathcal{G}_C$ contient $\mathcal{P}$ puisque pour tout $D \in \mathcal{P}$, $C \cap D \in \mathcal{P} \subset \Lambda(\mathcal{P})$. Or $\mathcal{G}_C$ est une λ-classe, donc elle contient aussi $\Lambda(\mathcal{P})$. Nous venons ainsi d'établir que

$$\forall C \in \mathcal{P},\ \forall B \in \Lambda(\mathcal{P}), \quad C \cap B \in \Lambda(\mathcal{P}).$$

Ce résultat peut se réécrire : $\forall B \in \Lambda(\mathcal{P})$, $\mathcal{G}_B \supset \mathcal{P}$. Or $\mathcal{G}_B$ est une λ-classe, donc par minimalité elle contient aussi $\Lambda(\mathcal{P})$. L'inclusion $\mathcal{G}_B \supset \Lambda(\mathcal{P})$ vraie pour tout $B \in \Lambda(\mathcal{P})$ se traduit par : $\forall A \in \Lambda(\mathcal{P})$, $\forall B \in \Lambda(\mathcal{P})$, $A \cap B \in \Lambda(\mathcal{P})$. Ceci établit la stabilité de $\Lambda(\mathcal{P})$ par intersections finies et achève la preuve du théorème de Dynkin. □

1.2.5 Tribus boréliennes

Dans un espace métrique E, on peut définir les notions d'ensemble ouvert, fermé, de voisinage, d'intérieur, de frontière, etc. Toutes ces notions qui permettent de faire de l'analyse dans ce cadre se déduisent de la notion d'ouvert. On peut aussi faire de l'analyse dans un cadre plus général, celui d'*espace topologique* $(E, \mathcal{O})$. Dans un tel espace, on se donne *a priori* la famille $\mathcal{O}$ des ouverts appelée *topologie*. Une topologie $\mathcal{O}$ sur E est une famille de parties de E vérifiant :

a) $\emptyset \in \mathcal{O}$ et $E \in \mathcal{O}$;

b) $\mathcal{O}$ est stable par intersection finie ;

c) $\mathcal{O}$ est stable par union quelconque.

Tout espace métrique est aussi un espace topologique en prenant pour $\mathcal{O}$ la famille des unions quelconques de boules ouvertes (y compris l'ensemble vide, considéré comme union indexée par $\emptyset$). Une topologie sur E n'est pas une tribu, il lui manque la stabilité par complémentaire. Notons que c) est une condition plus forte que la stabilité par union dénombrable.

Définition 1.16. *Soit $(E, \mathcal{O})$ un espace topologique. La tribu borélienne de E est la tribu engendrée par la famille $\mathcal{O}$ des* ouverts *de E. On la note* $\mathrm{Bor}(E)$*, ou* $\mathcal{B}(E)$ *ou* $\mathcal{B}_E$*.*

En raison de la stabilité d'une tribu par passage au complémentaire, il est facile de voir que $\mathrm{Bor}(E)$ est aussi la tribu engendrée par les fermés de E. Dans la suite, nous nous intéressons plus particulièrement au cas $E = \mathbb{R}^d$, $d \geq 1$. Sauf mention explicite du contraire, $\mathbb{R}^d$ sera toujours supposé muni de sa topologie d'espace vectoriel[1].

Nous aurons besoin du résultat suivant.

Lemme 1.17.

i) Tout ouvert de $\mathbb{R}$ (resp. $\mathbb{R}^d$) est union au plus dénombrable d'intervalles ouverts (resp. de pavés ouverts) à extrémités rationnelles (resp. dans $\mathbb{Q}^d$).

ii) Tout ouvert de $\mathbb{R}$ est union au plus dénombrable d'intervalles ouverts deux à deux disjoints.

[1] On démontre (cf. cours d'Analyse Réelle) que sur un espace vectoriel de *dimension finie*, il existe une seule topologie séparée rendant continues l'addition des vecteurs et la multiplication d'un vecteur par un scalaire. Toutes les normes sur cet espace sont équivalentes et engendrent cette topologie. On est donc libre de choisir sur $\mathbb{R}^d$ les boules associées à la norme de son choix, la famille des ouverts correspondante et la tribu borélienne associée seront toujours les mêmes.

Preuve de i). Traitons d'abord le cas particulier où l'ouvert est un intervalle ouvert à extrémités réelles $a < b$. Il suffit de noter que

$$]a, b[= \bigcup_{n\in\mathbb{N}}]r_n, s_n[,$$

où $(r_n)_{n\in\mathbb{N}}$ et $(s_n)_{n\in\mathbb{N}}$ sont deux suites de rationnels telles que $r_n \downarrow a$, $s_n \uparrow b$ et pour tout n, $a < r_n < s_n < b$.

Soit V un ouvert quelconque de $\mathbb{R}$. Alors il est réunion d'intervalles ouverts et en exploitant le cas précédent, on peut l'écrire

$$V = \bigcup_{i\in I}]a_i, b_i[= \bigcup_{i\in I} \bigcup_{n\in\mathbb{N}}]r_{n,i}, s_{n,i}[= \bigcup_{(r,s)\in D}]r, s[,$$

où l'on a noté $D := \{(r_{n,i}, s_{n,i});\ i \in I, n \in \mathbb{N}\}$. L'introduction de cette indexation par D revient à nettoyer l'indexation par I et $\mathbb{N}$ en effaçant les répétitions (il peut y en avoir une infinité non dénombrable !) de chaque intervalle $]r, s[$ apparaissant dans la réunion de tous les $]r_{n,i}, s_{n,i}[$. Comme D est inclus dans $\mathbb{Q}^2$, il est au plus dénombrable et on a bien réussi à écrire V comme union au plus dénombrable d'intervalles à extrémités rationnelles. Le résultat est donc établi en dimension 1. La généralisation en dimension $d > 1$ est une simple adaptation du cas unidimensionnel et est laissée en exercice. □

Preuve de ii). Remarquons qu'un sous-ensemble E de $\mathbb{R}$ est un intervalle si et seulement si pour tous x, y dans E avec $x \leq y$, $[x, y] \subset E$. On part de la décomposition $V = \bigcup_{(r,s)\in D}]r, s[$, établie dans la preuve du i). On définit sur D la relation binaire :

$$(r, s) \sim (r', s') \quad \Leftrightarrow \quad \exists J \text{ intervalle inclus dans } V \text{ tel que }]r, s[\subset J \text{ et }]r's'[\subset J.$$

Il est clair qu'il s'agit d'une relation d'équivalence. Les classes $\{D_k; k \in K\}$ forment une partition de D, donc l'ensemble d'indices K est au plus dénombrable. Posant alors

$$\forall k \in K, \quad E_k := \bigcup_{(r,s)\in D_k}]r, s[,$$

on a clairement

$$V = \bigcup_{k\in K} E_k.$$

Par construction, chaque E_k est un ouvert. Vérifions que c'est un intervalle. Soient x et y quelconques dans E_k avec $x \leq y$. Alors il existe des couples de rationnels $(r, s) \in D_k$ et $(r', s') \in D_k$ tels que $x \in]r, s[$ et $y \in]r', s'[$. De plus comme ces deux couples sont dans la même classe D_k, il existe un intervalle J inclus dans V et contenant à la fois $]r, s[$ et $]r', s'[$. On en déduit que x et y sont dans cet intervalle J, donc aussi $[x, y] \subset J \subset V$. Pour conclure que E_k est un intervalle (nécessairement ouvert), il reste à vérifier que $[x, y]$ est inclus dans E_k. Si tel n'est pas le cas, il existe un $z \in]x, y[$ tel que $z \notin E_k$ et comme $z \in V$, nécessairement $z \in E_j$ pour une autre classe E_j disjointe de E_k. Dans ce cas il existe un couple de rationnels $(r'', s'') \in D_j$ tels que $z \in]r'', s''[$. Or $J_0 := J \cup]r'', s''[$ est un intervalle (comme réunion de deux intervalles ayant au moins un point commun z) et est inclus dans V. Comme J_0 contient à la fois $]r, s[$, $]r', s'[$ et $]r'', s''[$, les trois couples de rationnels correspondants sont dans la même classe, d'où la contradiction. Donc $z \in E_k$ et E_k est bien un intervalle. □

Proposition 1.18. *La tribu borélienne de $\mathbb{R}$ est engendrée par chacune des familles*
- *intervalles ouverts* $]a, b[$, $a, b \in \mathbb{R}$ *;*
- *intervalles fermés* $[a, b]$, $a, b \in \mathbb{R}$ *;*
- *intervalles semi-ouverts* $]a, b]$, $a, b \in \mathbb{R}$ *;*
- *demi-droites ouvertes* $]-\infty, a[$, $a \in \mathbb{R}$ *;*
- *demi-droites fermées* $]-\infty, a]$, $a \in \mathbb{R}$.

Démonstration. On se contentera de traiter le cas de la famille

$$\mathcal{I} := \{]a, b[;\ a, b \in \mathbb{R}, a < b\},$$

les autres cas étant laissés en exercice. Comme $\mathcal{I}$ est incluse dans la famille des ouverts de $\mathbb{R}$, elle est aussi incluse dans la tribu $\mathrm{Bor}(\mathbb{R})$, donc $\sigma(\mathcal{I}) \subset \mathrm{Bor}(\mathbb{R})$. Dans l'autre sens, l'ensemble des ouverts de $\mathbb{R}$ est inclus dans la tribu $\sigma(\mathcal{I})$ grâce au lemme 1.17 i). Donc $\mathrm{Bor}(\mathbb{R}) \subset \sigma(\mathcal{I})$. □

Proposition 1.19. *La tribu borélienne de $\mathbb{R}^d$ est engendrée par chacune des familles suivantes :*
- *les pavés ouverts* $\prod_{i=1}^d]a_i, b_i[$, $a_i, b_i \in \mathbb{R}$ *;*
- *les pavés fermés* $\prod_{i=1}^d [a_i, b_i]$, $a_i, b_i \in \mathbb{R}$ *;*
- *les pavés de la forme* $\prod_{i=1}^d]a_i, b_i]$, $a_i, b_i \in \mathbb{R}$ *;*
- *les hyperquadrants* $\prod_{i=1}^d]-\infty, a_i]$, $a_i \in \mathbb{R}$.

Démonstration. Le cas de la famille $\mathcal{P}$ des pavés ouverts se traite exactement comme celui des intervalles ouverts à la proposition 1.18. Laissant au lecteur le cas des pavés fermés et des hyperquadrants, nous traiterons seulement le cas de la famille $\mathcal{C}$ des pavés $\prod_{i=1}^d]a_i, b_i]$, qui nous sera utile dans la suite. On vérifie d'abord que tout pavé ouvert est élément de $\sigma(\mathcal{C})$, en écrivant par exemple

$$Q := \prod_{i=1}^d]a_i, b_i[= \bigcup_{n\geq 1} \prod_{i=1}^d]a_i, b_i - 1/n]. \tag{1.6}$$

Noter que si pour un i, $b_i \leq a_i$, les deux intervalles $]a_i, b_i[$ et $]a_i, b_i - 1/n]$ se réduisent à l'ensemble vide et Q est lui même vide. On déduit de (1.6) l'inclusion

$$\mathrm{Bor}(\mathbb{R}^d) = \sigma(\mathcal{P}) \subset \sigma(\mathcal{C}). \tag{1.7}$$

Dans l'autre sens, l'écriture de Q comme intersection dénombrable de pavés ouverts

$$Q := \prod_{i=1}^d]a_i, b_i] = \bigcap_{n\geq 1} \prod_{i=1}^d]a_i, b_i + 1/n[, \tag{1.8}$$

nous donne l'inclusion $\mathcal{C} \subset \sigma(\mathcal{P})$, d'où

$$\sigma(\mathcal{C}) \subset \sigma(\mathcal{P}) = \mathrm{Bor}(\mathbb{R}^d). \tag{1.9}$$

Au vu de (1.7) et (1.9), on a bien $\sigma(\mathcal{C}) = \mathrm{Bor}(\mathbb{R}^d)$. □

1.3 Mesures

1.3.1 Un peu de vocabulaire

On appelle *espace mesurable* un couple $(\Omega, \mathcal{F})$ où Ω est un ensemble et $\mathcal{F}$ une tribu sur Ω. Un tel espace mesurable peut être muni d'une (ou de plusieurs) mesure(s) (cf. définition 1.1). Si μ est une mesure sur $(\Omega, \mathcal{F})$, le triplet $(\Omega, \mathcal{F}, \mu)$ s'appelle *espace mesuré*.

On dit qu'une propriété (π) vérifiée par des éléments de Ω est $\mathcal{F}$ mesurable si $B := \{\omega \in \Omega; \ \omega \text{ vérifie } (\pi)\}$ est un élément de $\mathcal{F}$. Une propriété $\mathcal{F}$-mesurable (π) est dite vraie μ-presque partout (ou μ-presque sûrement si μ est une probabilité) si $\mu(B^c) = 0$. Dans le cas où μ est une probabilité, $\mu(B^c) = 0$ équivaut à $\mu(B) = 1$.

Une mesure μ est dite finie si $\mu(\Omega) < +\infty$. D'après la proposition 1.21 2) ci-dessous, pour tout $A \in \mathcal{F}$, $\mu(A) \leq \mu(\Omega)$. En conséquence, pour une mesure considérée comme fonction d'ensembles ayant pour domaine de définition $\mathcal{F}$, « finie » équivaut à « bornée ». Cette équivalence est bien sûr fausse pour une fonction quelconque. Par exemple $f :]0,1] \to \mathbb{R}$, $x \mapsto 1/x$. On a $|f(x)| < +\infty$ pour tout $x \in]0,1]$, donc f est bien finie en tout point de son domaine de définition. Par contre, $\sup_{x\in]0,1]} |f(x)| = +\infty$, donc f n'est pas bornée sur son domaine de définition.

Une mesure μ sur $(\Omega, \mathcal{F})$ est dite σ-finie si Ω est union dénombrable d'ensembles A_n membres de $\mathcal{F}$ et chacun de mesure finie.

1.3.2 Premiers exemples

Exemple 1.1 (masse de Dirac). Soit x_0 un élément fixé de Ω. On appelle masse de Dirac au point x_0, la mesure δ_{x_0} sur $(\Omega, \mathcal{P}(\Omega))$ définie par

$$\forall A \in \mathcal{P}(\Omega), \quad \delta_{x_0}(A) := \begin{cases} 1 & \text{si } x_0 \in A, \\ 0 & \text{si } x_0 \notin A. \end{cases}$$

Par restriction, δ_{x_0} est aussi une mesure sur tout espace mesurable $(\Omega, \mathcal{F})$ où $\mathcal{F}$ est une tribu sur Ω.

Vérification. Il est clair que $\delta_{x_0}(\emptyset) = 0$. Pour montrer la σ-additivité, soit $(A_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathcal{P}(\Omega)$, deux à deux disjoints.

Cas 1 : $x_0 \notin \bigcup_{n\in\mathbb{N}} A_n$. Alors $\delta_{x_0}\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) = 0$. D'autre part, x_0 ne peut appartenir à aucun des A_n, donc pour tout $n \in \mathbb{N}$, $\delta_{x_0}(A_n) = 0$ et $\sum_{n\in\mathbb{N}} \delta_{x_0}(A_n) = 0$.

Cas 2 : $x_0 \in \bigcup_{n\in\mathbb{N}} A_n$. Alors $\delta_{x_0}\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) = 1$. D'autre part, comme les A_n sont deux à deux disjoints, x_0 doit appartenir à *un seul* des A_n, disons A_{n_0}. Ainsi, $\delta_{x_0}(A_{n_0}) = 1$ et pour tout $n \neq n_0$, $\delta_{x_0}(A_n) = 0$, d'où $\sum_{n\in\mathbb{N}} \delta_{x_0}(A_n) = 1$.

Dans les deux cas, on a

$$\delta_{x_0}\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) = \sum_{n\in\mathbb{N}} \delta_{x_0}(A_n),$$

δ_{x_0} est donc bien σ-additive. C'est une mesure sur $(\Omega, \mathcal{P}(\Omega))$. $\square$

Exemple 1.2 (séries de mesures finies). Si les $(\mu_k)_{k\in\mathbb{N}}$ sont des mesures finies sur $(\Omega, \mathcal{F})$ et si $(a_k)_{k\in\mathbb{N}}$ est une suite de réels positifs, la fonction d'ensembles $\mu = \sum_{k\in\mathbb{N}} a_k\mu_k$ définie sur $\mathcal{F}$ par

$$\mu : \mathcal{F} \to \overline{\mathbb{R}}_+, \quad A \mapsto \mu(A) := \sum_{k\in\mathbb{N}} a_k\mu_k(A)$$

est une mesure sur $(\Omega, \mathcal{F})$.

Vérification. Pour tout k, $0 \leq a_k < +\infty$, donc $a_k\mu_k(\emptyset) = 0$ et $\mu(\emptyset) = \sum_{k\in\mathbb{N}} a_k\mu_k(\emptyset) = 0$. Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathcal{F}$, deux à deux disjoints. En utilisant la σ-additivité de chaque μ_k et la propriété de sommation par paquets dans $\overline{\mathbb{R}}_+$, on peut écrire :

$$\begin{aligned}
\mu\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) = \sum_{k\in\mathbb{N}} a_k\mu_k\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) &= \sum_{k\in\mathbb{N}} a_k \left(\sum_{n\in\mathbb{N}} \mu_k(A_n)\right) \\
&= \sum_{k\in\mathbb{N}} \left(\sum_{n\in\mathbb{N}} a_k\mu_k(A_n)\right) \\
&= \sum_{n\in\mathbb{N}} \left(\sum_{k\in\mathbb{N}} a_k\mu_k(A_n)\right) \\
&= \sum_{n\in\mathbb{N}} \mu(A_n),
\end{aligned}$$

ce qui établit la σ-additivité de μ. □

Exemple 1.3 (mesures ponctuelles). Soient $(x_k)_{k\in\mathbb{N}}$ une suite dans Ω et $(a_k)_{k\in\mathbb{N}}$ une suite de réels positifs. Alors $\mu := \sum_{k\in\mathbb{N}} a_k\delta_{x_k}$ est une mesure sur $(\Omega, \mathcal{P}(\Omega))$, donc aussi par restriction sur tout espace mesurable $(\Omega, \mathcal{F})$. C'est un cas particulier de l'exemple précédent. Il est clair que le résultat demeure si l'on remplace l'ensemble d'indexation $\mathbb{N}$ par n'importe quel ensemble dénombrable. Les mesures de ce type sont appelées *mesures ponctuelles*.

Si de plus $\sum_{k\in\mathbb{N}} a_k = 1$, μ est une probabilité, appelée « probabilité discrète ». C'est le cas notamment pour les *lois* des variables aléatoires discrètes, considérées comme mesures sur $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$. Si $X : \Omega \to \mathbb{R}$ est une variable aléatoire discrète, l'ensemble $X(\Omega)$ de toutes les valeurs possibles est dénombrable et peut s'écrire $X(\Omega) = \{x_k,\ k \in \mathbb{N}\}$ et la loi de X est la mesure $P_X = \sum_{k\in\mathbb{N}} \mathbf{P}(X = x_k)\delta_{x_k}$ (nous reviendrons sur cette question au chapitre 2).

Remarque 1.20. Si $\Omega = \{\omega_i;\ i \in I\}$ est fini ou dénombrable, toute mesure sur $(\Omega, \mathcal{P}(\Omega))$ qui est *finie sur les singletons* est une mesure ponctuelle. En particulier une probabilité P sur $(\Omega, \mathcal{P}(\Omega))$ est caractérisée par les relations

$$\forall i \in I, \quad P(\{\omega_i\}) = p_i,$$

où les p_i sont des réels positifs tels que $\displaystyle\sum_{i\in I} p_i = 1$.

Vérification. La tribu $\mathcal{P}(\Omega)$ possède les singletons (quand Ω est fini ou dénombrable, c'est même la seule tribu ayant cette propriété). Soit μ une mesure sur $(\Omega, \mathcal{P}(\Omega))$. On peut définir

$$\forall i \in I, \quad a_i := \mu(\{\omega_i\}),$$

en notant que puisque μ est finie sur les singletons, $0 \leq a_i < +\infty$. Considérons alors la mesure ponctuelle

$$\nu := \sum_{i \in I} a_i \delta_{\omega_i}.$$

Soit $A \in \mathcal{P}(\Omega)$, alors A est fini ou dénombrable et on peut l'écrire comme union disjointe finie ou dénombrable de singletons :

$$A = \bigcup_{\omega_i \in A} \{\omega_i\}.$$

Par σ-additivité (ou par additivité quand A est fini, cf. prop. 1.21 ci-dessous), on a donc

$$\mu(A) = \sum_{\omega_i \in A} \mu(\{\omega_i\}) = \sum_{\omega_i \in A} a_i = \sum_{i \in I} a_i \delta_{\omega_i}(A) = \nu(A).$$

Ainsi pour tout $A \in \mathcal{P}(\Omega)$, $\mu(A) = \nu(A)$, donc $\mu = \nu$ et μ est une mesure ponctuelle. □

Exemple 1.4 (trace et conditionnement). Soit μ une mesure sur $(\Omega, \mathcal{F})$ et $B \in \mathcal{F}$. La fonction d'ensembles $\nu = \mu(. \cap B)$ définie sur $\mathcal{F}$ par

$$\forall A \in \mathcal{F}, \quad \nu(A) := \mu(A \cap B)$$

est une mesure sur $(\Omega, \mathcal{F})$. Si $0 < \mu(B) < +\infty$, la fonction d'ensembles μ_B définie sur $\mathcal{F}$ par

$$\forall A \in \mathcal{F}, \quad \mu_B(A) := \frac{\mu(A \cap B)}{\mu(B)}$$

est une mesure sur $(\Omega, \mathcal{F})$, c'est même une probabilité. Quand $\mu = P$ est déjà une probabilité sur $(\Omega, \mathcal{F})$, la mesure P_B est appelée probabilité conditionnelle ; notation : $P_B(A) =: P(A \mid B)$.

Vérification. On se contente de vérifier que ν est σ-additive, tout le reste étant évident. Soit $(A_i)_{i \in \mathbb{N}}$ une suite d'éléments de $\mathcal{F}$, deux à deux disjoints. Pour $i \neq j$, $A_i \cap A_j = \emptyset$ et comme $(A_i \cap B) \cap (A_j \cap B) \subset A_i \cap A_j$, les $A_i \cap B$ sont aussi deux à deux disjoints. De plus on a

$$\Big(\bigcup_{i \in \mathbb{N}} A_i\Big) \cap B = \bigcup_{i \in \mathbb{N}} (A_i \cap B).$$

La σ-additivité de ν découle alors clairement de celle de μ. □

Exemple 1.5 (mesure de Lebesgue λ_d). Nous admettons provisoirement qu'il existe une unique mesure μ sur $(\mathbb{R}^d, \mathrm{Bor}(\mathbb{R}^d))$ telle que pour tout pavé $\prod_{i=1}^d]a_i, b_i]$ non vide[2],

$$\mu\left(\prod_{i=1}^d]a_i, b_i]\right) = \prod_{i=1}^d (b_i - a_i).$$

[2] Ce qui suppose implicitement que $a_i < b_i$ pour chaque $i = 1, \ldots, d$.

On l'appelle mesure de Lebesgue sur $\mathbb{R}^d$ et on la note λ_d. Pour $d = 1$, on étend ainsi la notion de longueur des intervalles à tous les ensembles membres de $\mathrm{Bor}(\mathbb{R})$, pour $d = 2$ on étend de même la notion d'aire des rectangles à tous les ensembles boréliens du plan $\mathbb{R}^2$, pour $d = 3$, on étend la notion de volume. Pour $d > 3$, on continuera à parler de volume ou d'hypervolume.

1.3.3 Propriétés générales d'une mesure

Proposition 1.21. *Toute mesure μ sur $(\Omega, \mathcal{F})$ vérifie les propriétés suivantes :*

1. *Additivité finie : si les A_i $(1 \leq i \leq n)$ sont deux à deux disjoints :*

$$\mu\Big(\bigcup_{i=1}^{n} A_i\Big) = \sum_{i=1}^{n} \mu(A_i).$$

2. *Croissance : $\forall A \in \mathcal{F}$, $\forall B \in \mathcal{F}$, $A \subset B \ \Rightarrow\ \mu(A) \leq \mu(B)$.*
3. *Continuité monotone séquentielle*
   (a) *Si $(B_n)_{n\geq 1}$ est une suite croissante dans $\mathcal{F}$, convergente*[3] *vers $B \in \mathcal{F}$, alors $\mu(B) = \lim_{n\to+\infty} \mu(B_n)$. Notation :*

$$B_n \uparrow B \Rightarrow \mu(B_n) \uparrow \mu(B) \quad (n \to +\infty).$$

   (b) *Si $(C_n)_{n\geq 1}$ est une suite décroissante dans $\mathcal{F}$, convergente*[4] *vers $C \in \mathcal{F}$* et si de plus *$\mu(C_1)$ est* fini, *alors $\mu(C) = \lim_{n\to+\infty} \mu(C_n)$. Notation :*

$$\mu(C_1) < +\infty \text{ et } C_n \downarrow C \Rightarrow \mu(C_n) \downarrow \mu(C) \quad (n \to +\infty).$$

4. (a) *Sous-additivité : $\forall A_1, \dots, A_n \in \mathcal{F}$,* $\quad \mu\Big(\bigcup_{i=1}^{n} A_i\Big) \leq \sum_{i=1}^{n} \mu(A_i).$

   (b) *Sous σ-additivité : $\forall A_1, \dots, A_n, \dots \in \mathcal{F}$,* $\quad \mu\Big(\bigcup_{i\in\mathbb{N}^*} A_i\Big) \leq \sum_{i=1}^{+\infty} \mu(A_i).$

Preuve de 1. Soient $A_1, \dots, A_n$, n éléments de $\mathcal{F}$ deux à deux disjoints. Pour $j > n$, posons $A_j = \emptyset$. On a ainsi une suite infinie $(A_j)_{j\geq 1}$ d'ensembles deux à deux disjoints, éléments de $\mathcal{F}$. En utilisant la σ-additivité, on obtient alors :

$$\mu\Big(\bigcup_{j=1}^{n} A_j\Big) = \mu\Big(\bigcup_{j\in\mathbb{N}^*} A_j\Big) = \sum_{j=1}^{n} \mu(A_j) + \sum_{j=n+1}^{+\infty} \mu(A_j).$$

Comme $\mu(\emptyset) = 0$, la somme pour $j \geq n + 1$ vaut 0. □

[3] Ce qui signifie : $\forall n \geq 1$, $B_n \subset B_{n+1}$ et $B = \bigcup_{n\geq 1} B_n$.

[4] Ce qui signifie : $\forall n \geq 1$, $C_{n+1} \subset C_n$ et $C = \bigcap_{n\geq 1} C_n$.

Remarque 1.22. Il résulte de l'additivité finie que si μ est une mesure *finie*,

$$\forall A, B \in \mathcal{F}, \text{ tels que } A \subset B, \quad \mu(B \setminus A) = \mu(B) - \mu(A). \tag{1.10}$$

Plus généralement (1.10) est vraie avec une mesure non finie à condition que $\mu(A)$ soit fini. Pour vérifier (1.10), on remarque simplement que B est la réunion disjointe de A et $B \setminus A$, d'où $\mu(B) = \mu(A) + \mu(B \setminus A)$. Quand $\mu(A) < +\infty$, on obtient une égalité équivalente en le retranchant des deux membres, d'où le résultat.

Preuve de 2. Si $A \subset B$, alors $B = A \cup (B \cap A^c)$ et cette réunion est disjointe. D'après *1* on a $\mu(B) = \mu(A) + \mu(B \cap A^c)$ et comme $\mu(B \cap A^c) \geq 0$, on en déduit $\mu(B) \geq \mu(A)$. □

Preuve de 3(a). On utilise la décomposition suivante de B en réunion dénombrable disjointe (cf. la preuve de la proposition 1.4 pour la justification) :

$$B = B_0 \cup \left(\bigcup_{i \in \mathbb{N}^*} (B_i \setminus B_{i-1}) \right).$$

Par σ-additivité de μ, ces deux décompositions nous donnent :

$$\mu(B_n) = \mu(B_0) + \sum_{i=1}^{n} \mu(B_i \setminus B_{i-1}), \quad \mu(B) = \mu(B_0) + \sum_{i=1}^{+\infty} \mu(B_i \setminus B_{i-1}).$$

Comme cette série converge, sa somme est la limite de la suite de ses sommes partielles de rang n, ce qui s'écrit :

$$\mu(B) = \lim_{n \to +\infty} \left\{ \mu(B_0) + \sum_{i=1}^{n} \mu(B_i \setminus B_{i-1}) \right\} = \lim_{n \to +\infty} \mu(B_n).$$

□

Preuve de 3(b). Posons $B_n = C_1 \setminus C_n$. Comme $C_n \subset C_1$, on a la réunion disjointe $C_1 = C_n \cup B_n$. Par additivité finie, $\mu(C_1) = \mu(C_n) + \mu(B_n)$. De plus comme $B_n \subset C_1$, $\mu(B_n) < +\infty$. On peut donc écrire

$$\mu(C_n) = \mu(C_1) - \mu(B_n).$$

La suite (B_n) est croissante de réunion $C_1 \setminus C$, car $B_n = C_1 \cap C_n^c$ d'où

$$\bigcup_{n \geq 1} B_n = C_1 \cap \left(\bigcup_{n \geq 1} C_n^c \right) = C_1 \cap \left(\bigcap_{n \geq 1} C_n \right)^c = C_1 \cap C^c = C_1 \setminus C.$$

Par 3(a) $\mu(B_n) \uparrow \mu(C_1 \setminus C)$, donc $\mu(C_n)$ converge vers $\mu(C_1) - \mu(C_1 \setminus C) = \mu(C)$. Noter que $\mu(C_1)$ étant fini et C étant inclus dans C_1, $\mu(C)$ et $\mu(C_1 \setminus C)$ sont tous deux finis. □

Il importe de noter l'utilisation de l'hypothèse $\mu(C_1) < +\infty$. Voici un contre-exemple sans elle : on prend $\Omega = \mathbb{N}$, $\mathcal{F} = \mathcal{P}(\Omega)$ et $\mu = \sum_{n \in \mathbb{N}} \delta_n$ mesure de comptage. La suite des $C_n = \mathbb{N} \cap [n, +\infty[$ est décroissante d'intersection vide, mais comme pour tout n, $\mu(C_n) = +\infty$, $\mu(C_n)$ ne peut converger vers 0.

Preuve de 4(a). On remarque que pour tout $n \geq 1$, on a :

$$\bigcup_{i=0}^{n} A_i = \bigcup_{i=0}^{n} B_i,$$

où les B_i sont des ensembles deux à deux disjoints définis comme suit :

$$\begin{aligned} B_0 = A_0, \quad B_1 &= A_1 \cap B_0^c, \quad B_2 = A_2 \cap (B_0 \cup B_1)^c, \dots \\ \dots B_n &= A_n \cap (B_0 \cup B_1 \cup \dots B_{n-1})^c, \dots \end{aligned}$$

Par additivité :

$$\mu\Big(\bigcup_{i=0}^{n} A_i\Big) = \mu\Big(\bigcup_{i=0}^{n} B_i\Big) = \sum_{i=0}^{n} \mu(B_i).$$

Par construction pour tout i, $B_i \subset A_i$, d'où $\mu(B_i) \leq \mu(A_i)$ et :

$$\mu\Big(\bigcup_{i=0}^{n} A_i\Big) = \sum_{i=0}^{n} \mu(B_i) \leq \sum_{i=0}^{n} \mu(A_i).$$

□

Preuve de 4(b). Posons pour tout $n \geq 1$,

$$D_n = \bigcup_{i=0}^{n} A_i, \qquad D = \bigcup_{n \geq 1} D_n = \bigcup_{i \in \mathbb{N}} A_i.$$

La suite $(D_n)_{n \geq 1}$ est *croissante* et a pour limite D. Donc d'après *3(a)*, $\mu(D_n) \uparrow \mu(D)$ $(n \to +\infty)$. D'après *4(a)* on a :

$$\forall n \geq 1, \quad \mu(D_n) \leq \sum_{i=0}^{n} \mu(A_i).$$

Les deux membres de cette inégalité étant les termes généraux de deux suites croissantes de réels positifs, on obtient en passant à la limite quand n tend vers $+\infty$:

$$\mu(\bigcup_{i \in \mathbb{N}} A_i) = \mu(D) \leq \sum_{i=0}^{+\infty} \mu(A_i).$$

Ce qui prouve *4(b)*. Remarquons que les sommes partielles de la série convergent dans $\mathbb{R}^+ \cup \{+\infty\}$. □

Proposition 1.23 (Formule de Poincaré).
Soit μ une mesure finie *sur $(\Omega, \mathcal{F})$. Pour tout entier $n \geq 2$ et tous $A_1, \dots, A_n \in \mathcal{F}$:*

$$\mu\Big(\bigcup_{i=1}^{n} A_i\Big) = \sum_{i=1}^{n} \mu(A_i) + \sum_{k=2}^{n} (-1)^{k+1} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \mu(A_{i_1} \cap \dots \cap A_{i_k}). \tag{1.11}$$

Démonstration. On raisonne par récurrence. Pour $n = 2$, la formule se réduit à

$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B). \tag{1.12}$$

On a les décompositions suivantes en unions disjointes :

$$\begin{aligned} A \cup B &= (A \cap B^c) \cup (A \cap B) \cup (A^c \cap B), \\ A &= (A \cap B^c) \cup (A \cap B), \\ B &= (A \cap B) \cup (A^c \cap B). \end{aligned}$$

En utilisant l'additivité on en déduit :

$$\begin{aligned} \mu(A \cup B) &= \mu(A \cap B^c) + \mu(A \cap B) + \mu(A^c \cap B) \\ &= \big[\mu(A \cap B^c) + \mu(A \cap B)\big] + \big[\mu(A \cap B) + \mu(A^c \cap B)\big] - \mu(A \cap B) \\ &= \mu(A) + \mu(B) - \mu(A \cap B). \end{aligned}$$

Supposons la formule de Poincaré vraie au rang n (plus précisément on suppose que pour *toute* suite de n éléments $A_1, \dots, A_n$ de $\mathcal{F}$, l'égalité (1.11) est vérifiée). Pour en déduire qu'elle est alors vraie au rang $n+1$, il nous faut calculer $\mu\Big(\overset{n+1}{\underset{i=1}{\cup}} A_i\Big)$. On commence par appliquer (1.12) avec $A = \overset{n}{\underset{i=1}{\cup}} A_i$ et $B = A_{n+1}$. On obtient ainsi :

$$\begin{aligned} \mu\Big(\overset{n+1}{\underset{i=1}{\cup}} A_i\Big) &= \mu\Big(\overset{n}{\underset{i=1}{\cup}} A_i\Big) + \mu(A_{n+1}) - \mu\Big(\Big[\overset{n}{\underset{i=1}{\cup}} A_i\Big] \cap A_{n+1}\Big) \\ &= \mu\Big(\overset{n}{\underset{i=1}{\cup}} A_i\Big) + \mu(A_{n+1}) - \mu\Big(\overset{n}{\underset{i=1}{\cup}} (A_i \cap A_{n+1})\Big). \end{aligned}$$

On applique maintenant l'hypothèse de récurrence (formule de Poincaré (1.11)) d'abord avec les n ensembles $A_1, \dots, A_n$ puis avec les n ensembles $A'_1, \dots, A'_n$, où l'on a posé $A'_i := A_i \cap A_{n+1}$. Il vient :

$$\begin{aligned} \mu\Big(\overset{n+1}{\underset{i=1}{\cup}} A_i\Big) = {} & \sum_{i=1}^n \mu(A_i) + \sum_{k=2}^n (-1)^{k+1} \sum_{1 \le i_1 < i_2 < \ldots < i_k \le n} \mu(A_{i_1} \cap \cdots \cap A_{i_k}) \\ & + \mu(A_{n+1}) \\ & - \sum_{i=1}^n \mu(A'_i) - \sum_{j=2}^n (-1)^{j+1} \sum_{1 \le i_1 < i_2 < \ldots < i_j \le n} \mu(A'_{i_1} \cap \cdots \cap A'_{i_j}) \\ = {} & \sum_{i=1}^{n+1} \mu(A_i) && (1.13) \\ & + \sum_{k=2}^n (-1)^{k+1} \sum_{1 \le i_1 < i_2 < \ldots < i_k \le n} \mu(A_{i_1} \cap \cdots \cap A_{i_k}) && (1.14) \\ & + (-1)^{2+1} \sum_{i=1}^n \mu(A_i \cap A_{n+1}) && (1.15) \\ & + \sum_{j=2}^n (-1)^{(j+1)+1} \sum_{1 \le i_1 < i_2 < \ldots < i_j \le n} \mu(A_{i_1} \cap \cdots \cap A_{i_j} \cap A_{n+1}). && (1.16) \end{aligned}$$

Comparons ce résultat avec ce que l'on espère trouver, c'est-à-dire avec

$$\sum_{i=1}^{n+1} \mu(A_i) + \underbrace{\sum_{k=2}^{n+1} (-1)^{k+1} \sum_{1 \leq i_1 < i_2 < \ldots < i_k \leq n+1} \mu(A_{i_1} \cap \cdots \cap A_{i_k})}_{=:T_{n+1}}.$$

Cela revient à vérifier que T_{n+1} est égal à la somme des lignes (1.14) à (1.16) ci-dessus. Partageons T_{n+1} en deux blocs comme suit. Le premier bloc regroupe tous les termes tels que $i_k < n+1$ (et donc $i_k \leq n$ et $k \leq n$). On le retrouve exactement à la ligne (1.14). Le deuxième bloc regroupe tous les termes pour lesquels $i_k = n+1$. Dans ce bloc, la somme des termes pour lesquels $k = 2$ se retrouve ligne (1.15). Il reste alors la somme des termes pour lesquels $3 \leq k \leq n+1$ et $i_k = n+1$ (donc $i_{k-1} \leq n$). Cette somme est exactement le contenu de la ligne (1.16), comme on peut le voir en faisant le changement d'indice $k = j + 1$ dans (1.16). Ceci achève la récurrence. □

1.3.4 Fonction de répartition

Cette sous-section est une illustration de la proposition 1.21 et un prélude à la caractérisation des mesures finies sur la tribu borélienne de $\mathbb{R}$.

Définition 1.24. *Soit μ une mesure* finie *sur $(\mathbb{R}, \mathrm{Bor}(\mathbb{R}))$. Sa fonction de répartition (f.d.r.) est l'application*

$$F : \mathbb{R} \to \mathbb{R}_+, \quad x \mapsto F(x) := \mu\big(]-\infty, x]\big).$$

Théorème 1.25 (propriétés des f.d.r.). *Soit F la fonction de répartition d'une mesure finie μ sur $(\mathbb{R}, \mathrm{Bor}(\mathbb{R}))$.*

i) F est croissante sur $\mathbb{R}$.

ii) F est continue à droite et pourvue d'une limite à gauche en tout point (fonction càdlàg*). De plus*

$$\lim_{x \to -\infty} F(x) = 0, \qquad \lim_{x \to +\infty} F(x) = \mu(\mathbb{R}).$$

iii) Pour tout $a \in \mathbb{R}$, $F(a) - F(a^-) = \mu(\{a\})$ et il n'y a qu'un ensemble au plus dénombrable de points a où $F(a) - F(a^-) > 0$.

iv) Si deux mesures finies μ et ν sur $(\mathbb{R}, \mathrm{Bor}(\mathbb{R}))$ ont même f.d.r., elles sont égales.

La propriété iv) est listée ici pour des raisons de commodité, mais nous n'avons pas encore les outils pour la prouver complètement. On la démontrera seulement pour des mesures ponctuelles au sens de l'exemple 1.3. Le cas général sera vu ultérieurement comme corollaire du théorème d'unicité (th. 1.34).

Preuve de i). Si $x < x'$, $]-\infty, x] \subset]-\infty, x']$ et par croissance de μ (prop. 1.21–2.), $\mu\big(]-\infty, x]\big) \leq \mu\big(]-\infty, x']\big)$, soit encore $F(x) \leq F(x')$. F est bien croissante. □

Preuve de ii). Comme fonction croissante, F a en tout point $a \in \mathbb{R}$, une limite à gauche $F(a^-) := \lim_{\varepsilon \to 0, \varepsilon > 0} F(a - \varepsilon)$ et une limite à droite $F(a^+) := \lim_{\varepsilon \to 0, \varepsilon > 0} F(a + \varepsilon)$. Pour établir la continuité à droite de F en a, il suffit donc de vérifier que $F(a^+) = F(a)$. L'*existence* de la limite à droite $F(a^+)$ étant acquise, il suffit finalement de vérifier que $F(a) = \lim_{n \to +\infty} F(a + 1/n)$. En posant $A_n :=]-\infty, a + 1/n]$ et $A =]-\infty, a]$, on a $F(a + 1/n) = \mu(A_n)$ et $F(a) = \mu(A)$. On voit immédiatement que $(A_n)_{n \geq 1}$ est une suite décroissante pour l'inclusion dans le tribu $\mathrm{Bor}(\mathbb{R})$ et que $\cap_{k \geq 1} A_k = A$. Par continuité décroissante séquentielle de μ (prop. 1.21–3.b)), on en déduit que $\mu(A) = \lim_{n \to +\infty} \mu(A_n)$.

La croissance de F sur $\mathbb{R}$ assure l'existence de limites en $-\infty$ et $+\infty$. Pour les identifier, il suffit donc de trouver $\lim_{n \to +\infty} F(-n)$ et $\lim_{n \to +\infty} F(n)$. En utilisant la continuité monotone séquentielle de μ (prop. 1.21–3.), on obtient quand n tend vers $+\infty$:

$$\begin{aligned} F(-n) = \mu\bigl(]-\infty, -n]\bigr) \ &\downarrow \ \mu\Bigl(\bigcap_{k \geq 1}]-\infty, -k]\Bigr) = \mu(\emptyset) = 0; \\ F(n) = \mu\bigl(]-\infty, n]\bigr) \ &\uparrow \ \mu\Bigl(\bigcup_{k \geq 1}]-\infty, k]\Bigr) = \mu(\mathbb{R}). \end{aligned}$$

□

Preuve de iii). Posons $C_n :=]a - 1/n, a]$. Parce que μ est une mesure finie, il résulte de l'additivité finie de μ (Rem. 1.22) que

$$\mu(C_n) = \mu\bigl(]-\infty, a] \setminus]-\infty, a-1/n]\bigr) = \mu\bigl(]-\infty, a]\bigr) - \mu\bigl(]-\infty, a-1/n]\bigr) = F(a) - F(a-1/n).$$

D'autre part $(C_n)_{n \geq 1}$ est une suite dans la tribu $\mathrm{Bor}(\mathbb{R})$, décroissante pour l'inclusion et son intersection est $\{a\}$. Donc par continuité décroissante séquentielle de μ, $\mu(C_n) = F(a) - F(a - 1/n)$ converge vers $\mu(\{a\})$, d'où $F(a) - F(a^-) = \mu(\{a\})$. Comme $F(a) = F(a^+)$, on voit ainsi que F est discontinue au point a si et seulement si $\mu(\{a\}) > 0$. Lorsque ceci arrive, $\mu(\{a\})$ est l'amplitude du « saut » de la fonction F en a et la discontinuité est de première espèce. On sait qu'une fonction monotone ne peut avoir de discontinuités que de première espèce (i.e. avec existence d'une limite à gauche et d'une limite à droite) et que l'ensemble de ces discontinuités est au plus dénombrable. □

Preuve de iv) dans le cas particulier des mesures finies ponctuelles. Soient μ et ν deux mesures finies ponctuelles au sens de l'exemple 1.3. On peut les représenter sous la forme

$$\mu = \sum_{k \in \mathbb{N}} a_k \delta_{x_k}, \quad \nu = \sum_{k \in \mathbb{N}} b_k \delta_{y_k}.$$

Quitte à modifier les coefficients a_k, b_k, on peut toujours se ramener au cas où les x_k sont tous distincts et de même pour les y_k. On a alors pour tout $x \in \mathbb{R}$,

$$\mu(\{x\}) = \sum_{k \in \mathbb{N}} a_k \delta_{x_k}(\{x\}) = \begin{cases} a_j & \text{si } x = x_j \text{ pour un } j \in \mathbb{N}, \\ 0 & \text{si } x \notin \{x_k;\ k \in \mathbb{N}\}. \end{cases}$$

De même,

$$\nu(\{x\}) = \sum_{k\in\mathbb{N}} b_k \delta_{y_k}(\{x\}) = \begin{cases} b_j & \text{si } x = y_j \text{ pour un } j \in \mathbb{N}, \\ 0 & \text{si } x \notin \{y_k;\ k \in \mathbb{N}\}. \end{cases}$$

Notant F et G les fonctions de répartitions respectives de μ et ν, on déduit de iii) et de l'hypothèse $F = G$ que $\mu(\{x\}) = \nu(\{x\})$ pour tout $x \in \mathbb{R}$. Le calcul de $\mu(\{x\})$ et $\nu(\{x\})$ ci-dessus nous montre alors que pour tout $k \in \mathbb{N}$, $x_k = y_k$ et $a_k = b_k$. L'égalité de μ et ν en découle. □

1.3.5 Mesures extérieures

Définition 1.26. *Soit* Ω *un ensemble. Une* mesure extérieure *sur* Ω *est une application* $\mu^* : \mathcal{P}(\Omega) \to \overline{\mathbb{R}}_+$, *telle que*

i) $\mu^*(\emptyset) = 0$ *;*

ii) μ^* *est croissante : si* $A \subset B \subset \Omega$, $\mu^*(A) \leq \mu^*(B)$.

iii) μ^* *est sous-*σ*-additive : pour toute suite* $(A_i)_{i\in\mathbb{N}}$ *de parties de* Ω,

$$\mu^*\Bigl(\bigcup_{i\in\mathbb{N}} A_i\Bigr) \leq \sum_{i\in\mathbb{N}} \mu^*(A_i). \tag{1.17}$$

Toute mesure sur $(\Omega, \mathcal{P}(\Omega))$ est une mesure extérieure (prop. 1.21–2. et 4.b)), mais bien sûr, l'intérêt de cette notion est qu'elle est plus générale que celle de mesure. Elle sert notamment à construire des mesures par restriction d'une mesure extérieure à une sous-tribu de $\mathcal{P}(\Omega)$. L'exemple fondamental suivant donne une construction très générale de mesures extérieures.

Exemple 1.6. Soit $\mathcal{E}$ une famille de parties de Ω, telle que $\emptyset \in \mathcal{E}$ et $\theta : \mathcal{E} \to \overline{\mathbb{R}}_+$ une fonction d'ensembles sur $\mathcal{E}$ telle que $\theta(\emptyset) = 0$. Pour tout $A \in \mathcal{P}(\Omega)$, on pose

$$\mu^*(A) := \inf \sum_{k\in\mathbb{N}} \theta(E_k), \tag{1.18}$$

où l'infimum est pris sur toutes les suites $(E_k)_{k\in\mathbb{N}}$ d'éléments de $\mathcal{E}$, dont la réunion $\cup_{k\in\mathbb{N}} E_k$ recouvre A. On fait la convention habituelle attribuant la valeur $+\infty$ à l'infimum d'un ensemble de réels lorsque ce dernier est vide (donc ici s'il n'y a aucun recouvrement de A par une suite dans $\mathcal{E}$). La fonction d'ensembles μ^* ainsi définie est une mesure extérieure. Remarquons la faiblesse des hypothèses faites sur $\mathcal{E}$ et θ. Pour obtenir des applications intéressantes, il nous faudra bien sûr enrichir la structure $(\mathcal{E}, \theta)$.

Vérification. D'abord il est clair que $\mu^*(A) \geq 0$ pour tout $A \in \mathcal{P}(\Omega)$ parce que la fonction θ est à valeurs dans $\overline{\mathbb{R}}_+$. Puisque $\emptyset$ est élément de $\mathcal{E}$, son recouvrement par la suite (E_k) dont tous les éléments sont $\emptyset$ donne $\mu^*(\emptyset) \leq 0$, donc $\mu^*(\emptyset) = 0$. Si $A \subset B \subset \Omega$, tout recouvrement de B est un recouvrement de A et on en déduit immédiatement que $\mu^*(A) \leq \mu^*(B)$. Pour montrer la sous-σ-additivité, soit $(A_i)_{i\in\mathbb{N}}$ une suite quelconque dans $\mathcal{P}(\Omega)$. Si l'un au moins des $\mu^*(A_i)$ vaut $+\infty$, (1.17) est automatiquement vérifiée.

Si tous les $\mu^*(A_i)$ sont finis, la définition de l'infimum (1.18) nous assure pour chaque i et chaque $\varepsilon > 0$, de l'existence d'un recouvrement $(E_{i,k},\ k \in \mathbb{N})$ de A_i tel que

$$\sum_{k\in\mathbb{N}} \theta(E_{i,k}) \leq \mu^*(A_i) + \frac{\varepsilon}{2^{i+1}}.$$

D'autre part, $A := \cup_{i\in\mathbb{N}} A_i$ a au moins un recouvrement par une union *dénombrable* d'éléments de $\mathcal{E}$, c'est $\cup_{(i,k)\in\mathbb{N}^2} E_{i,k}$, donc d'après la définition (1.18), on a

$$\mu^*(A) \leq \sum_{(i,k)\in\mathbb{N}^2} \theta(E_{i,k}) = \sum_{i\in\mathbb{N}}\sum_{k\in\mathbb{N}} \theta(E_{i,k}) \leq \sum_{i\in\mathbb{N}} \Bigl(\mu^*(A_i) + \frac{\varepsilon}{2^{i+1}}\Bigr) = \varepsilon + \sum_{i\in\mathbb{N}} \mu^*(A_i).$$

Ceci étant vrai pour tout $\varepsilon > 0$, on en déduit (1.17). □

Définition 1.27. *Soit μ^* une mesure extérieure sur Ω. On dit que $A \in \mathcal{P}(\Omega)$ est μ^*-mesurable si*

$$\forall E \in \mathcal{P}(\Omega), \quad \mu^*(E) = \mu^*(A \cap E) + \mu^*(A^c \cap E). \tag{1.19}$$

En raison de la sous-additivité de μ^*, (1.19) équivaut à

$$\forall E \in \mathcal{P}(\Omega), \quad \mu^*(E) \geq \mu^*(A \cap E) + \mu^*(A^c \cap E). \tag{1.20}$$

On note $\mathcal{M}(\mu^*)$ l'ensemble des parties μ^*-mesurables de Ω.

Nous prouverons dans un instant que $\mathcal{M}(\mu^*)$ est une tribu. Il n'est guère aisé de donner une interprétation convaincante de (1.19). Pour en saisir la pertinence, le plus simple est peut-être de regarder comment elle fonctionne dans la preuve du lemme suivant.

Lemme 1.28. *La trace de μ^* sur tout $E \in \mathcal{P}(\Omega)$ est additive sur $\mathcal{M}(\mu^*)$: pour tout $n \geq 1$ et tous $A_1, \dots, A_n \in \mathcal{M}(\mu^*)$ deux à deux disjoints,*

$$\mu^*\Bigl(E \cap \Bigl(\bigcup_{j=1}^{n} A_j\Bigr)\Bigr) = \sum_{j=1}^{n} \mu^*(E \cap A_j). \tag{1.21}$$

En particulier pour $E = \Omega$, on obtient l'additivité de μ^* sur $\mathcal{M}(\mu^*)$.

Démonstration. Remarquons qu'ici on ne peut pas se contenter de vérifier (1.21) pour $n = 2$ car on ne sait pas encore si $\mathcal{M}(\mu^*)$ est stable par union finie, ce qui empêche d'utiliser directement l'associativité de la réunion. L'égalité (1.21), vraie pour $n = 1$, se montre par récurrence. Supposons la vérifiée au rang n et soit $A_{n+1} \in \mathcal{M}(\mu^*)$, les A_j, $1 \leq j \leq n+1$ étant deux à deux disjoints. En utilisant (1.19) avec $E' := E \cap (\cup_{1\leq j\leq n+1} A_j)$ et $A := A_{n+1} \in \mathcal{M}(\mu^*)$, il vient

$$\mu^*(E') = \mu^*(E' \cap A_{n+1}) + \mu^*(E' \cap A_{n+1}^c). \tag{1.22}$$

Grâce à la disjonction des A_j, on a

$$E' \cap A_{n+1} = E \cap \Bigl(\bigcup_{j=1}^{n+1} A_j\Bigr) \cap A_{n+1} = \bigcup_{j=1}^{n+1} (E \cap A_j \cap A_{n+1}) = E \cap A_{n+1}$$

et comme les A_j pour $j \leq n$ sont tous inclus dans A_{n+1}^c, donc aussi leur réunion,

$$E' \cap A_{n+1}^c = \Bigl(E \cap \Bigl(\bigcup_{j=1}^{n} A_j\Bigr) \cap A_{n+1}^c\Bigr) \cup \bigl(E \cap A_{n+1} \cap A_{n+1}^c\bigr) = E \cap \Bigl(\bigcup_{j=1}^{n} A_j\Bigr).$$

En reportant ces évaluations dans (1.22), on obtient

$$\mu^*\Bigl(E \cap \Bigl(\bigcup_{j=1}^{n+1} A_j\Bigr)\Bigr) = \mu^*(E \cap A_{n+1}) + \mu^*\Bigl(E \cap \Bigl(\bigcup_{j=1}^{n} A_j\Bigr)\Bigr),$$

ce qui permet d'achever la récurrence. □

Théorème 1.29. *$\mathcal{M}(\mu^*)$ est une tribu et la restriction de μ^* à $\mathcal{M}(\mu^*)$ est une mesure.*

Démonstration. Pour montrer que $\mathcal{M}(\mu^*)$ est une tribu, nous utilisons la proposition 1.4. L'ensemble vide est clairement un élément de $\mathcal{M}(\mu^*)$, donc $\mathcal{M}(\mu^*)$ n'est pas vide (sic !). La stabilité par complémentaire étant évidente, il reste à prouver la stabilité pour l'intersection finie et pour la réunion dénombrable disjointe.
a) Stabilité pour l'intersection finie. Soient A et B dans $\mathcal{M}(\mu^*)$. On a pour tout E les trois égalités suivantes :

$$\begin{aligned} \mu^*(E) &= \mu^*(A \cap E) + \mu^*(A^c \cap E) \\ \mu^*(A \cap E) &= \mu^*(B \cap A \cap E) + \mu^*(B^c \cap A \cap E) \\ \mu^*(A^c \cap E) &= \mu^*(B \cap A^c \cap E) + \mu^*(B^c \cap A^c \cap E) \end{aligned}$$

Puis, remarquant que $(B \cap A^c \cap E) \cup (B^c \cap A \cap E) \cup (B^c \cap A^c \cap E) = (A^c \cup B^c) \cap E$ on obtient par sous σ-additivité de μ^*,

$$\mu^*(A^c \cup B^c \cap E) \leq \mu^*(B \cap A^c \cap E) + \mu^*(B^c \cap A \cap E) + \mu^*(B^c \cap A^c \cap E),$$

d'où découle finalement $\mu^*(E) \geq \mu^*(A^c \cup B^c \cap E) + \mu^*(B \cap A \cap E)$, qui prouve que $A \cap B$ est dans $\mathcal{M}(\mu^*)$.
b) Stabilité pour la réunion dénombrable disjointe. Soit $(A_n)_{n\geq 1}$ une suite d'éléments deux à deux disjoints de $\mathcal{M}(\mu^*)$ et A sa réunion. La stabilité par complémentaire et par intersection finie de $\mathcal{M}(\mu^*)$ entraînent sa stabilité par union finie. Ainsi pour tout $n \geq 1$, $B_n := \bigcup_{j=1}^{n} A_j$ est dans $\mathcal{M}(\mu^*)$ et l'on a pour tout $E \subset \Omega$,

$$\mu^*(E) = \mu^*(E \cap B_n) + \mu^*(E \cap B_n^c).$$

Comme $B_n \subset A$, $A^c \subset B_n^c$, et par croissance de la mesure extérieure μ^*, on en déduit

$$\mu^*(E) \geq \mu^*(E \cap B_n) + \mu^*(E \cap A^c). \tag{1.23}$$

Par le lemme 1.28,

$$\mu^*(E \cap B_n) = \sum_{j=1}^{n} \mu^*(E \cap A_j),$$

ce qui reporté dans (1.23) nous donne

$$\mu^*(E) \geq \sum_{j=1}^{n} \mu^*(E \cap A_j) + \mu^*\Big(E \cap A^c\Big).$$

Ceci étant vrai pour tout n, on en déduit

$$\mu^*(E) \geq \sum_{j=1}^{+\infty} \mu^*(E \cap A_j) + \mu^*\Big(E \cap A^c\Big). \tag{1.24}$$

Or $E \cap A = \cup_{j\in\mathbb{N}^*}(E \cap A_j)$, d'où par sous-$\sigma$-additivité de μ^*,

$$\mu^*(E) \geq \mu^*\Big(E \cap A\Big) + \mu^*\Big(E \cap A^c\Big),$$

ce qui prouve bien que A est dans $\mathcal{M}(\mu^*)$. Ainsi $\mathcal{M}(\mu^*)$ est stable par union dénombrable disjointe et est bien une tribu.

c) μ^* restreinte à $\mathcal{M}(\mu^*)$ est une mesure. C'est une retombée immédiate de la preuve du b). En effet, on sait déjà que $\mu^*(\emptyset) = 0$. Soit $(A_n)_{n\geq 1}$ une suite d'éléments deux à deux disjoints de $\mathcal{M}(\mu^*)$ et A sa réunion. Le cas particulier $E = A$ dans (1.24) s'écrit

$$\mu^*(A) \geq \sum_{j=1}^{+\infty} \mu^*(A_j). \tag{1.25}$$

La sous-σ-additivité de μ^* fournit l'inégalité inverse, de sorte qu'on a l'égalité dans (1.25), ce qui exprime la σ-additivité de μ^* restreinte à $\mathcal{M}(\mu^*)$. □

1.3.6 Théorème d'extension

Grosso modo, le problème que nous allons étudier maintenant peut se décrire ainsi. On dispose d'une fonction d'ensemble μ sur une famille $\mathcal{C}$ de parties de Ω. À titre d'exemple, on peut prendre pour $\mathcal{C}$ la classe $\mathcal{I}$ des intervalles $]a, b]$ de $\Omega = \mathbb{R}$ et y définir μ par $\mu(]a, b]) := b - a$. On cherche des conditions suffisantes assez générales permettant de prolonger μ à toute la tribu $\sigma(\mathcal{C})$ de sorte que ce prolongement soit une mesure sur $\mathcal{C}$. Avec l'exemple mentionné, comme $\sigma(\mathcal{I}) = \mathrm{Bor}(\mathbb{R})$, on aura ainsi construit la mesure de Lebesgue sur $\mathbb{R}$, qui étend la fonction « longueur » de $\mathcal{I}$ à toute la tribu borélienne de $\mathbb{R}$. Commençons par préciser la structure de $\mathcal{C}$.

Définition 1.30. *On appelle* semi-algèbre *de parties de Ω, toute famille $\mathcal{C}$ de parties de Ω telle que :*

a) $\emptyset \in \mathcal{C}$,

b) $\mathcal{C}$ est stable par intersection finie,

c) pour tous $A \in \mathcal{C}$, $B \in \mathcal{C}$ tels que $A \subset B$, il existe une suite finie $C_1, \dots, C_n$ d'éléments de $\mathcal{C}$, deux à deux disjoints tels que $B \setminus A = \bigcup_{k=1}^{n} C_k$.

Bien sûr le c) doit être compris comme englobant le cas particulier où $B \setminus A \in \mathcal{C}$, auquel cas on peut prendre $n = 1$, $C_1 = B \setminus A$ et la clause « deux à deux disjoints » devient sans objet. L'exemple fondamental suivant permet de motiver cette définition.

Exemple 1.7. La classe $\mathcal{I}$ des intervalles $]a, b]$ de $\mathbb{R}$ est une semi-algèbre de parties de $\mathbb{R}$. De même, la classe $\mathcal{C}$ des pavés de $\mathbb{R}^d$ de la forme $Q = \prod_{i=1}^d]a_i, b_i]$ est une semi-algèbre.

Vérification. L'intervalle $]a, b]$ est par définition $\{x \in \mathbb{R};\ a < x \leq b\}$. Ainsi pour $a = b$, $]a, a] = \emptyset \in \mathcal{I}$. Cette définition de l'intervalle nous montre immédiatement que $\cap_{1\leq k\leq n}]a_k, b_k] =]a, b]$ en posant $a := \max_{1\leq k\leq n} a_k$ et $b := \min_{1\leq k\leq n} b_k$ (et donc le résultat est $\emptyset$ quand $a \geq b$). Soit $]c, d] \subset]a, b]$ et supposons pour éviter le cas trivial que $c < d$. Alors nécessairement $a \leq c < d \leq b$ et $]a, b] \setminus]c, d] =]a, c] \cup]d, b]$. Ainsi $\mathcal{I}$ est bien une semi-algèbre. Notons au passage que notre vérification de la propriété c) n'est pas transposable aux cas des classes $\mathcal{I}_o$ d'intervalles ouverts $]a, b[$ ou $\mathcal{I}_f$ d'intervalles fermés $[a, b]$. Par contre la transposition à la classe $\mathcal{I}'$ des intervalles $[a, b[$ est immédiate.

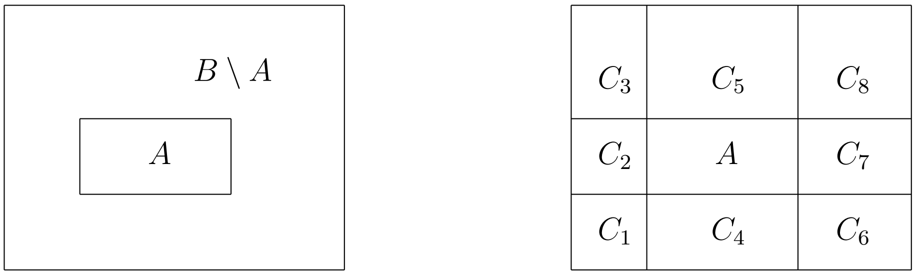
Passons à la classe $\mathcal{C}$ des pavés de $\mathbb{R}^d$ du type

$$Q = \prod_{i=1}^d]a_i, b_i] = \{x = (x_1, \dots, x_d) \in \mathbb{R}^d;\ \forall i = 1, \dots, d,\ a_i < x_i \leq b_i\}.$$

Sous cette dernière forme, on voit immédiatement que $\mathcal{C}$ contient $\emptyset$ (si $a_i \geq b_i$ pour au moins un i, $Q = \emptyset$) et est stable par intersection finie : en posant $a_i := \max_{1\leq k\leq n} a_{i,k}$ et $b_i := \min_{1\leq k\leq n} b_{i,k}$, on a

$$\bigcap_{k=1}^n \prod_{i=1}^d]a_{i,k}, b_{i,k}] = \prod_{i=1}^d]a_i, b_i].$$

Pour vérifier la condition c), soit $R := \prod_{i=1}^d]a'_i, b'_i]$ inclus dans $Q := \prod_{i=1}^d]a_i, b_i]$. Le cas $R = \emptyset$ est trivial. Si $R \neq \emptyset$, alors $Q \neq \emptyset$, d'où $a'_i < b'_i$ et $a_i < b_i$ pour tout $i = 1, \dots, d$. Le point $x = (b'_1, \dots, b'_d)$ appartient à R donc aussi à Q d'où $a_i < b'_i \leq b_i$ pour tout i. Si $m := \min_{1\leq i\leq d}(b'_i - a'_i)$, m est strictement positif et pour tout $0 < \varepsilon \leq m$, $y_\varepsilon := (a'_1 + \varepsilon, \dots, a'_d + \varepsilon)$ est dans R donc dans Q d'où $a_i < a'_i + \varepsilon \leq b'_i \leq b_i$ pour tout i. En faisant tendre ε vers 0, on en déduit que pour tout $i = 1, \dots, d$, $a_i \leq a'_i < b'_i \leq b_i$. On voit alors que le pavé Q peut être découpé en 3^d pavés $\prod_{i=1}^d I_{i,j(i)}$ où $j(i) \in \{1, 2, 3\}$ et $I_{i,1} :=]a_i, a'_i]$, $I_{i,2} :=]a'_i, b'_i]$, $I_{i,3} :=]b'_i, b_i]$. Ces 3^d pavés sont deux à deux disjoints (certains pouvant être vides). Le choix $j(i) = 2$ pour tout i donne le pavé R. On a donc trouvé $n = 3^d - 1$ pavés $C_1, \dots, C_n$ deux à deux disjoints tels que $Q \setminus R = \cup_{1\leq k\leq n} C_k$. La condition c) est vérifiée et $\mathcal{C}$ est une semi-algèbre. $\square$



La décomposition $B \setminus A = \overset{8}{\underset{i=1}{\cup}} C_i$, pour $d = 2$.

Théorème 1.31. *Soit $\mathcal{C}$ une semi-algèbre de parties de Ω et μ une application $\mathcal{C} \to \overline{\mathbb{R}}_+$ vérifiant :*

i) $\mu(\emptyset) = 0$ *;*

ii) *μ est additive sur $\mathcal{C}$: pour toute suite finie $C_0, \dots, C_n$ d'éléments deux à deux disjoints de $\mathcal{C}$ et telle que $\overset{n}{\underset{i=0}{\cup}} C_i$ appartienne à $\mathcal{C}$,*

$$\mu\Big(\overset{n}{\underset{i=0}{\cup}} C_i\Big) = \sum_{i=0}^{n} \mu(C_i);$$

iii) *μ est sous-σ-additive sur $\mathcal{C}$: pour toute suite $(A_i)_{i\in\mathbb{N}}$ d'éléments de $\mathcal{C}$ telle que $\underset{i\in\mathbb{N}}{\cup} A_i$ appartienne à $\mathcal{C}$,*

$$\mu\Big(\underset{i\in\mathbb{N}}{\cup} A_i\Big) \leq \sum_{i=0}^{+\infty} \mu(A_i).$$

Alors μ se prolonge en une mesure sur $\sigma(\mathcal{C})$.

Remarquons que la conclusion implique que μ devait déjà être σ-additive sur $\mathcal{C}$. Réciproquement, si μ est σ-additive sur $\mathcal{C}$ et $\mu(\emptyset) = 0$, alors elle vérifie ii) et iii). On n'aurait donc pas perdu de généralité en remplaçant dans l'énoncé ii) et iii) par l'hypothèse de σ-additivité de μ sur $\mathcal{C}$. Nous avons préféré la version énoncée ci-dessus pour des raisons pratiques.

Démonstration. La preuve s'appuie sur le théorème 1.29 : on associe à μ la mesure extérieure μ^* définie par

$$\forall A \in \mathcal{P}(\Omega), \quad \mu^*(A) := \inf \sum_{k\in\mathbb{N}} \mu(D_k), \tag{1.26}$$

où l'infimum est pris sur toutes les suites $(D_k)_{k\in\mathbb{N}}$ d'éléments de $\mathcal{C}$, dont la réunion $\cup_{k\in\mathbb{N}} D_k$ recouvre A. $\mathcal{M}(\mu^*)$ étant la tribu associée à μ^* (cf. déf. 1.27 et th. 1.29), on sait que la restriction de μ^* à $\mathcal{M}(\mu^*)$ est une mesure. Le théorème 1.31 sera donc prouvé si nous montrons que

a) $\sigma(\mathcal{C}) \subset \mathcal{M}(\mu^*)$;

b) μ^* coïncide avec μ sur $\mathcal{C}$: $\forall B \in \mathcal{C}$, $\mu^*(B) = \mu(B)$.

En préliminaire, vérifions d'abord que μ est croissante sur $\mathcal{C}$. Si $A, B \in \mathcal{C}$ et $A \subset B$, alors il existe une suite finie $C_1, \dots, C_n$ d'éléments deux à deux disjoints de $\mathcal{C}$ telle que $B \setminus A = \cup_{1\leq i\leq n} C_i$. Comme $B \setminus A$ est disjoint de A, chacun de ces C_i est disjoint de $A =: C_0$ et on peut décomposer B en réunion finie disjointe $B = \cup_{0\leq i\leq n} C_i$ d'éléments de $\mathcal{C}$. Par additivité et positivité de μ sur $\mathcal{C}$, on a alors :

$$\mu(B) = \sum_{i=0}^{n} \mu(C_i) = \mu(A) + \sum_{i=1}^{n} \mu(C_i) \geq \mu(A).$$

Ainsi $\mu(A) \leq \mu(B)$ et μ est croissante sur $\mathcal{C}$.

a) $\mathcal{M}(\mu^*)$ *contient* $\sigma(\mathcal{C})$. Par minimalité de $\sigma(\mathcal{C})$, il suffit de montrer que la tribu $\mathcal{M}(\mu^*)$ contient $\mathcal{C}$. Soit B quelconque dans $\mathcal{C}$, nous allons établir l'appartenance de B à $\mathcal{M}(\mu^*)$ en vérifiant (1.20) :

$$\forall E \in \mathcal{P}(\Omega), \quad \mu^*(E) \geq \mu^*(B \cap E) + \mu^*(B^c \cap E).$$

Si $\mu^*(E) = +\infty$, cette inégalité est automatiquement vérifiée. Si $\mu^*(E) < +\infty$, pour tout $\varepsilon > 0$, il existe une suite $(D_k)_{k\in\mathbb{N}}$ dans $\mathcal{C}$ telle que $E \subset \cup_{k\in\mathbb{N}} D_k$ et

$$\sum_{k\in\mathbb{N}} \mu(D_k) \leq \mu^*(E) + \varepsilon. \tag{1.27}$$

Comme $\mathcal{C}$ est stable par intersection finie, chaque $B \cap D_k$ est élément de $\mathcal{C}$. Ainsi l'inclusion $B \cap E \subset \cup_{k\in\mathbb{N}}(B \cap D_k)$ nous fournit un recouvrement de $B \cap E$ par une suite de $\mathcal{C}$ et la définition de μ^* implique

$$\mu^*(B \cap E) \leq \sum_{k\in\mathbb{N}} \mu(B \cap D_k). \tag{1.28}$$

Le même argument ne s'applique pas aux $B^c \cap D_k$, car $\mathcal{C}$ n'est pas nécessairement stable par complémentaire. Cependant en remarquant que $B^c \cap D_k = D_k \setminus (B \cap D_k)$, on dispose pour chaque k d'une suite finie $(C_{j,k})_{1\leq j\leq n_k}$ d'éléments deux à deux disjoints de $\mathcal{C}$ tels que

$$B^c \cap D_k = \bigcup_{j=1}^{n_k} C_{j,k}.$$

L'inclusion

$$B^c \cap E \subset \bigcup_{k\in\mathbb{N}} \bigcup_{j=1}^{n_k} C_{j,k}$$

nous fournit maintenant un recouvrement de $B^c \cap E$ par une union dénombrable d'éléments de $\mathcal{C}$, donc

$$\mu^*(B^c \cap E) \leq \sum_{k\in\mathbb{N}} \sum_{j=1}^{n_k} \mu(C_{j,k}). \tag{1.29}$$

En sommant (1.28) et (1.29), il vient

$$\mu^*(B\cap E)+\mu^*(B^c\cap E)\leq \sum_{k\in\mathbb{N}}\Big\{\mu(B\cap D_k)+\sum_{j=1}^{n_k}\mu(C_{j,k})\Big\}.$$

Grâce à l'additivité de μ sur $\mathcal{C}$ et à la décomposition

$$D_k=(B\cap D_k)\cup\bigcup_{j=1}^{n_k}C_{j,k}$$

en union finie disjointe d'éléments de $\mathcal{C}$, ceci peut se réécrire

$$\mu^*(B\cap E)+\mu^*(B^c\cap E)\leq \sum_{k\in\mathbb{N}}\mu(D_k).$$

En reportant dans (1.27), on obtient

$$\mu^*(E)\geq \mu^*(B\cap E)+\mu^*(B^c\cap E)-\varepsilon.$$

Cette inégalité étant vraie pour tout $\varepsilon>0$, on en déduit $\mu^*(E)\geq\mu^*(B\cap E)+\mu^*(B^c\cap E)$. Comme E était quelconque dans $\mathcal{P}(\Omega)$, on a bien vérifié (1.20) pour B quelconque dans $\mathcal{C}$ et ainsi la tribu $\mathcal{M}(\mu^*)$ contient $\mathcal{C}$, donc aussi $\sigma(\mathcal{C})$.

b) μ^* coïncide avec μ sur $\mathcal{C}$. Soit $B\in\mathcal{C}$. En considérant le recouvrement particulier de B par la suite $(D_k)_{k\in\mathbb{N}}$ où $D_0=B$ et $D_k=\emptyset$ pour tout $k\geq 1$, on obtient immédiatement

$$\mu^*(B)\leq\mu(B). \tag{1.30}$$

Soit maintenant $(D_k)_{k\in\mathbb{N}}$ un recouvrement dénombrable quelconque de B par des éléments de $\mathcal{C}$. On a alors $B=\cup_{k\in\mathbb{N}}(B\cap D_k)$, d'où par sous-$\sigma$-additivité de μ sur $\mathcal{C}$,

$$\mu(B)\leq\sum_{k\in\mathbb{N}}\mu(B\cap D_k).$$

Par croissance de μ sur $\mathcal{C}$, $\mu(B\cap D_k)\leq\mu(D_k)$, d'où

$$\mu(B)\leq\sum_{k\in\mathbb{N}}\mu(D_k).$$

Ceci étant vrai pour tout recouvrement dénombrable $(D_k)_{k\in\mathbb{N}}$ de B par des éléments de $\mathcal{C}$, on en déduit en prenant l'infimum sur tous ces recouvrements :

$$\mu(B)\leq\mu^*(B). \tag{1.31}$$

Cette inégalité jointe à (1.30) nous donne l'égalité $\mu(B)=\mu^*(B)$ pour tout $B\in\mathcal{C}$ et ceci termine la preuve du théorème. □

Nous terminons cette sous-section par deux lemmes élémentaires qui ont leur utilité pour la vérification pratique de la sous-σ-additivité sur $\mathcal{C}$.

Lemme 1.32. *Soient $\mathcal{C}$ une semi-algèbre de parties de Ω, $A \in \mathcal{C}$ et $A_1, A_2, \dots, A_n$ une suite finie dans $\mathcal{C}$, de réunion incluse dans A. Alors $A \setminus (A_1 \cup \cdots \cup A_n)$ est union finie disjointe d'éléments de $\mathcal{C}$.*

Démonstration. On le vérifie par récurrence sur n. Pour $n = 1$, la propriété énoncée se réduit au c) de la définition 1.30 d'une semi-algèbre. Supposons la propriété vraie au rang n. Soit $A_1, A_2, \dots, A_n, A_{n+1}$ une suite finie dans $\mathcal{C}$, de réunion incluse dans A. Par hypothèse de récurrence, on a une décomposition

$$A \setminus (A_1 \cup \cdots \cup A_n) = \bigcup_{j=1}^{m_n} C_j,$$

les C_j étant des éléments deux à deux disjoints de $\mathcal{C}$. On peut alors écrire

$$\begin{aligned} A \setminus (A_1 \cup \cdots \cup A_n \cup A_{n+1}) &= A \cap (A_1^c \cap \cdots \cap A_n^c) \cap A_{n+1}^c \\ &= \big(A \setminus (A_1 \cup \cdots \cup A_n)\big) \cap A_{n+1}^c \\ &= \bigcup_{j=1}^{m_n} (C_j \cap A_{n+1}^c) \\ &= \bigcup_{j=1}^{m_n} \big(C_j \setminus (C_j \cap A_{n+1})\big). \end{aligned}$$

Comme $\mathcal{C}$ est stable par intersection finie, $C_j \cap A_{n+1}$ appartient à $\mathcal{C}$ et en appliquant la propriété c) de la définition d'une semi-algèbre, on dispose pour chaque j d'une suite finie $(D_{j,k})_{1 \leq k \leq n_j}$ d'éléments deux à deux disjoints de $\mathcal{C}$ telle que

$$C_j \setminus (C_j \cap A_{n+1}) = \bigcup_{k=1}^{n_j} D_{j,k}.$$

La suite finie dans $\mathcal{C}$ obtenue en rassemblant tous les $D_{j,k}$ pour $1 \leq j \leq m_n$, $1 \leq k \leq n_j$ a ses éléments deux à deux disjoints, par construction pour les paires ayant le même indice j et pour les autres paires, parce que les $C_j \setminus (C_j \cap A_{n+1})$ sont inclus dans les C_j eux-mêmes deux à deux disjoints. On a donc obtenu la décomposition en réunion finie disjointe d'éléments de $\mathcal{C}$:

$$A \setminus (A_1 \cup \cdots \cup A_n \cup A_{n+1}) = \bigcup_{j=1}^{m_n} \bigcup_{k=1}^{n_j} D_{j,k},$$

ce qui achève la récurrence. □

Lemme 1.33. *Soient $\mathcal{C}$ une semi-algèbre de parties de Ω, μ une fonction d'ensembles définie sur $\mathcal{C}$, à valeurs dans $\overline{\mathbb{R}}_+$ et additive sur $\mathcal{C}$ et $A \in \mathcal{C}$.*

a) Si $A_1, \dots, A_n \in \mathcal{C}$ sont deux à deux disjoints et de réunion incluse dans A,

$$\sum_{j=1}^{n} \mu(A_j) \leq \mu(A).$$

b) Soient $B_1, \dots, B_n \in \mathcal{C}$. On suppose que $A \subset \bigcup_{j=1}^{n} B_j$. Alors

$$\mu(A) \leq \sum_{j=1}^{n} \mu(B_j).$$

Démonstration. Le a) est une conséquence immédiate du lemme 1.32 qui nous fournit la décomposition :

$$A = \Bigl(\bigcup_{j=1}^{n} A_j\Bigr) \cup \Bigl(\bigcup_{k=1}^{m_n} C_k\Bigr),$$

les C_k et les A_j étant tous dans $\mathcal{C}$ et deux à deux disjoints. Par additivité de μ sur $\mathcal{C}$, il en résulte

$$\mu(A) = \sum_{j=1}^{n} \mu(A_j) + \sum_{k=1}^{m_n} \mu(C_k) \geq \sum_{j=1}^{n} \mu(A_j).$$

Pour prouver le b), remarquons d'abord que

$$A = A \cap \Bigl(\bigcup_{j=1}^{n} B_j\Bigr) = \bigcup_{j=1}^{n} (A \cap B_j).$$

Ensuite, en posant $C_1 := A \cap B_1$ et pour $j > 1$, $C_j := (A \cap B_j) \setminus \bigcup_{i=1}^{j-1} (A \cap B_i)$, on a (cf. la preuve de la proposition 1.4)

$$A = \bigcup_{j=1}^{n} C_j \quad \text{(union disjointe)}.$$

Comme $\mathcal{C}$ est stable par intersection finie, on peut appliquer le lemme 1.32 à chaque C_j pour le décomposer en union finie disjointe d'éléments $D_{j,k}$ $(1 \leq k \leq m_j)$ de $\mathcal{C}$. On a alors l'union disjointe

$$A = \bigcup_{j=1}^{n} \bigcup_{k=1}^{m_j} D_{j,k} \quad \text{d'où} \quad \mu(A) = \sum_{j=1}^{n} \sum_{k=1}^{m_j} \mu(D_{j,k}).$$

Pour j fixé, les $D_{j,k}$ sont deux à deux disjoints et de réunion égale à C_j, donc incluse dans B_j. Par le a), on a alors

$$\sum_{k=1}^{m_j} \mu(D_{j,k}) \leq \mu(B_j).$$

En reportant dans la décomposition de $\mu(A)$ ci-dessus, on aboutit à

$$\mu(A) = \sum_{j=1}^{n} \sum_{k=1}^{m_j} \mu(D_{j,k}) \leq \sum_{j=1}^{n} \mu(B_j),$$

ce qui est bien la conclusion annoncée. □

1.3.7 Théorème d'unicité

Théorème 1.34. *Soient μ et ν deux mesures définies sur le même espace mesurable $(\Omega, \mathcal{F})$ et qui coïncident sur une π-classe $\mathcal{P}$ ($\mathcal{P} \subset \mathcal{F}$).*

a) Si $\mu(\Omega) = \nu(\Omega) < +\infty$, alors μ et ν coïncident sur $\sigma(\mathcal{P})$ (qui est une sous-tribu de $\mathcal{F}$).

b) Si $\mu(\Omega) = +\infty$ et s'il existe dans $\mathcal{P}$ une suite $(A_n)_{n\in\mathbb{N}^*}$ de réunion Ω et telle que pour tout $n \in \mathbb{N}^*$, $\mu(A_n) < +\infty$, alors μ et ν coïncident sur $\sigma(\mathcal{P})$.

Preuve de a). Soit $\mathcal{L} := \{A \in \mathcal{F};\ \mu(A) = \nu(A)\}$. Par hypothèse, $\mathcal{L}$ contient $\mathcal{P}$. Si on montre que $\mathcal{L}$ est une λ-classe, le théorème de Dynkin (th. 1.15) nous donnera l'inclusion $\sigma(\mathcal{P}) \subset \mathcal{L}$, qui signifie que $\mu(A) = \nu(A)$ pour tout $A \in \sigma(\mathcal{P})$, ce qui est bien la conclusion annoncée. Vérifions donc que $\mathcal{L}$ est une λ-classe. $\Omega \in \mathcal{L}$ puisque par hypothèse $\mu(\Omega) = \nu(\Omega)$. Si A et B sont dans $\mathcal{F}$ et $A \subset B$, comme μ et ν sont finies, on a (cf. remarque 1.22),

$$\mu(B \setminus A) = \mu(B) - \mu(A) = \nu(B) - \nu(A) = \nu(B \setminus A),$$

donc $B \setminus A$ est dans $\mathcal{L}$, ce qui établit la stabilité par différence propre. Soit enfin $(A_n)_{n\in\mathbb{N}}$ une suite dans $\mathcal{L}$, croissante pour l'inclusion et A sa réunion. Pour tout $n \in \mathbb{N}$, on a $\mu(A_n) = \nu(A_n)$ et par continuité croissante séquentielle des mesures, $\mu(A_n) \uparrow \mu(A)$ et $\nu(A_n) \uparrow \nu(A)$ quand n tend vers l'infini. On en déduit que $\mu(A) = \nu(A)$, d'où la stabilité de $\mathcal{L}$ par union dénombrable croissante. Ainsi $\mathcal{L}$ est bien une λ-classe et ceci achève la preuve du cas a). $\square$

Preuve de b). Soit $A \in \sigma(\mathcal{P})$. On montre que $\mu(A) = \nu(A)$ en utilisant un argument d'approximation par des mesures finies et le résultat du a). Pour celà, on note d'abord que

$$\mu(A) = \mu\Big(A \cap \Big(\bigcup_{i=1}^{+\infty} A_i\Big)\Big) = \mu\Big(\bigcup_{i=1}^{+\infty} (A \cap A_i)\Big) = \lim_{n\to+\infty} \mu\Big(\bigcup_{i=1}^{n} (A \cap A_i)\Big),$$

et que la même expression est valable en remplaçant μ par ν. Remarquons que sous cette forme et même sous l'hypothèse plus restrictive $A \in \mathcal{P}$ au lieu de $A \in \sigma(\mathcal{P})$, il n'est pas clair directement que $\mu\Big(\bigcup_{i=1}^{n} (A \cap A_i)\Big) = \nu\Big(\bigcup_{i=1}^{n} (A \cap A_i)\Big)$. En effet $\mathcal{P}$ n'est pas nécessairement stable par réunion finie. Pour se débarasser de cette réunion, on fait appel à la formule de Poincaré (prop. 1.23) qui s'écrit ici :

$$\mu\Big(\bigcup_{i=1}^{n} (A \cap A_i)\Big) = \sum_{i=1}^{n} \mu(A \cap A_i) + \sum_{k=2}^{n} (-1)^{k+1} \sum_{1\leq i_1 < i_2 < \ldots < i_k \leq n} \mu(A \cap A_{i_1} \cap \cdots \cap A_{i_k}).$$

En appliquant la même formule avec ν à la place de μ, on voit que le problème sera résolu si l'on prouve que pour tous choix de $1 \leq k \leq n$ et $1 \leq i_1 < i_2 < \ldots < i_k \leq n$,

$$\forall A \in \sigma(\mathcal{P}), \quad \mu(A \cap A_{i_1} \cap \cdots \cap A_{i_k}) = \nu(A \cap A_{i_1} \cap \cdots \cap A_{i_k}).$$

Remarquons maintenant que si l'on pose $B = A_{i_1} \cap \cdots \cap A_{i_k}$, B est dans $\mathcal{P}$ et $\nu(B) = \mu(B) \leq \max_{1\leq j\leq k} \mu(A_{i_j}) < +\infty$. L'ultime réduction du problème consiste ainsi à vérifier que :

$$\forall A \in \sigma(\mathcal{P}), \quad \forall B \in \mathcal{P} \text{ tel que } \mu(B) = \nu(B) < +\infty, \quad \mu(A \cap B) = \nu(A \cap B). \qquad (1.32)$$

Pour cela, on considère les mesures μ' et ν' définies par

$$\forall A \in \mathcal{F}, \quad \mu'(A) := \mu(A \cap B), \quad \nu'(A) := \nu(A \cap B).$$

Ce sont bien des mesures sur $\mathcal{F}$, d'après l'exemple 1.4. Elles sont finies puisque $\mu'(\Omega) = \mu(B) < +\infty$ (idem pour ν'). Elles coïncident sur $\mathcal{P}$ puisque si $A \in \mathcal{P}$, $A \cap B \in \mathcal{P}$ et $\mu = \nu$ sur $\mathcal{P}$. D'après le a), μ' et ν' coïncident sur $\sigma(\mathcal{P})$. Ceci étant vrai pour tout $B \in \mathcal{P}$ tel que $\mu(B) = \nu(B) < +\infty$, (1.32) est établie, ce qui achève la preuve du théorème d'unicité. □

À titre de première application du théorème d'unicité, nous pouvons maintenant compléter la preuve du théorème 1.25 iv).

Corollaire 1.35. *Si deux mesures finies μ et ν sur $(\mathbb{R}, \mathrm{Bor}(\mathbb{R}))$ ont même fonction de répartition, elles sont égales.*

Démonstration. La famille d'intervalles $\mathcal{P} := \{\,]-\infty, x];\ x \in \mathbb{R}\}$ est une π-classe qui engendre la tribu borélienne de $\mathbb{R}$. L'égalité des fonctions de répartition F de μ et G de ν n'est qu'une autre manière d'exprimer la coïncidence de μ et ν sur $\mathcal{P}$. De plus par continuité croissante séquentielle :

$$\mu(\mathbb{R}) = \lim_{n\to+\infty} \mu(]-\infty, n]) = \lim_{n\to+\infty} F(n) = \lim_{n\to+\infty} G(n) = \lim_{n\to+\infty} \nu(]-\infty, n]) = \nu(\mathbb{R}).$$

Donc par le théorème 1.34 a), μ et ν coïncident sur $\sigma(\mathcal{P}) = \mathrm{Bor}(\mathbb{R})$. □

1.3.8 Mesures de Stieltjes sur $\mathbb{R}$

Une application importante des théorèmes 1.31 et 1.34 est la construction de mesures sur $(\mathbb{R}, \mathrm{Bor}(\mathbb{R}))$ à partir d'une fonction croissante. En particulier on obtient ainsi la mesure de Lebesgue λ_1 sur $\mathbb{R}$.

Théorème 1.36. *Soit $F : \mathbb{R} \to \mathbb{R}$ une fonction croissante et continue à droite en tout point. Définissons sur $\mathcal{C} := \{\,]a, b];\ a, b \in \mathbb{R}\}$ la fonction d'ensembles μ par $\mu(\emptyset) = 0$ et $\mu(]a, b]) = F(b) - F(a)$ si $a < b$. Alors μ se prolonge de manière unique en une mesure sur la tribu borélienne de $\mathbb{R}$. Cette mesure est appelée* mesure de Stieltjes *associée à F. En particulier lorsque $F : x \mapsto x$ est l'identité sur $\mathbb{R}$, on obtient ainsi la mesure de Lebesgue λ (ou λ_1) qui prolonge à $\mathrm{Bor}(\mathbb{R})$ la fonction longueur des intervalles de $\mathcal{C}$ ($\lambda(]a, b]) = b - a$, si $a < b$).*

Démonstration. On sait que $\sigma(\mathcal{C}) = \mathrm{Bor}(\mathbb{R})$. On montre l'existence par le théorème 1.31 et l'unicité par le théorème 1.34.
Additivité de μ sur $\mathcal{C}$. Il s'agit de vérifier que pour toute décomposition en union finie disjointe $]a, b] = \bigcup_{k=1}^{n} I_k$, avec $I_k :=]a_k, b_k]$,

$$F(b) - F(a) = \sum_{k=1}^{n} \bigl(F(b_k) - F(a_k)\bigr). \tag{1.33}$$

On ne perd pas de généralité en supposant qu'aucun I_k n'est vide, donc que $a_k < b_k$ pour tout k. L'égalité (1.33) devient évidente par téléscopage des termes lorsque les extrémités des I_k forment une subdivision de $]a, b]$ de la forme

$$a = c_0 < c_1 < c_2 < \cdots < c_{n-1} < c_n = b, \quad a_k = c_{k-1},\ b_k = c_k, \quad 1 \leq k \leq n. \tag{1.34}$$

Il suffit donc de montrer que l'on peut se ramener à cette situation en réindexant les intervalles I_k. Réindexons les en intervalles I'_j $(1 \leq j \leq n)$ avec $I'_j =]a'_j, b'_j]$ de façon que la suite $(b'_j)_{1\leq j\leq n}$ soit croissante. S'il existait un j tel que $b'_j > a'_{j+1}$, on aurait $a'_{j+1} < b'_j \leq b'_{j+1}$, donc $b'_j \in I'_{j+1}$, d'où $I'_j \cap I'_{j+1} \neq \emptyset$, ce qui contredirait l'hypothèse que les intervalles sont deux à deux disjoints. Ainsi pour tout $j = 1, \ldots, n-1$, $b'_j \leq a'_{j+1}$. Supposons maintenant qu'il existe $1 \leq j_0 < n$ pour lequel cette inégalité soit stricte : $b'_{j_0} < a'_{j_0+1}$. Alors a'_{j_0+1} ne pourrait être dans $\cup_{j=1}^n I'_j$ puisque

$$a'_{j_0+1} > b_{j_0} \Rightarrow a'_{j_0+1} \notin \bigcup_{j=1}^{j_0} I'_j, \quad a'_{j_0+1} \notin]a'_{j_0+1}, b'_{j_0+1}], \quad a'_{j_0+1} < b'_{j_0+1} \Rightarrow a'_{j_0+1} \notin \bigcup_{j=j_0+2}^{n} I'_j.$$

Ceci est impossible car $a'_{j_0+1} \in]a, b]$ puisque d'une part $a'_{j_0+1} \geq b'_{j_0} \geq b'_1 > a$ et d'autre part $a'_{j_0+1} \leq b'_n \leq b$. On a donc pour tout $1 \leq j \leq n-1$, $b'_j = a'_{j+1}$. Tout ceci nous donne la configuration

$$a \leq a'_1 < b'_1 = a'_2 < b'_2 = a'_3 < \cdots < b'_{n-1} = a'_n < b'_n \leq b.$$

Si $a < a'_1$, alors $(a + a'_1)/2 \in]a, b]$ et $(a + a'_1)/2 \notin \cup_{1\leq j\leq n} I'_j$, donc nécessairement $a = a'_1$. De même $b'_n = b$. On a donc bien une configuration du type (1.34) avec $c_j = b'_j$, $c_{j-1} = a'_j$, $1 \leq j \leq n$. D'où

$$\sum_{k=1}^{n} \mu(I_k) = \sum_{j=1}^{n} \mu(I'_j) = \sum_{j=1}^{n} \big(F(c_j) - F(c_{j-1})\big) = F(c_n) - F(c_0) = F(b) - F(a) = \mu(]a, b]).$$

L'additivité de μ sur $\mathcal{C}$ est ainsi établie.

Sous-σ-additivité de μ sur $\mathcal{C}$. Il s'agit de montrer que

$$]a, b] = \bigcup_{k\in\mathbb{N}^*}]a_k, b_k] \quad \Rightarrow \quad F(b) - F(a) \leq \sum_{k=1}^{+\infty} \big(F(b_k) - F(a_k)\big). \tag{1.35}$$

Nous allons utiliser un argument de compacité, la continuité à droite de F permettant de remplacer le recouvrement de $]a, b]$ par les $]a_k, b_k]$ par un recouvrement d'ouverts avec une erreur aussi petite que l'on veut. Fixons pour commencer $\varepsilon > 0$. La croissance et la continuité à droite de F au point a nous assurent de l'existence d'un $\delta_0 \in]0, b-a[$ tel que

$$F(a) \leq F(a + \delta_0) \leq F(a) + \varepsilon. \tag{1.36}$$

On recouvre alors comme suit $[a + \delta_0, b]$:

$$[a + \delta_0, b] \subset]a, b] = \bigcup_{k\in\mathbb{N}^*}]a_k, b_k] \subset \bigcup_{k\in\mathbb{N}^*}]a_k, b_k + \delta_k[, \tag{1.37}$$

où les δ_k ($k \geq 1$) sont choisis grâce à la continuité à droite de F en b_k, de telle sorte que

$$\forall k \geq 1, \quad F(b_k) \leq F(b_k + \delta_k) \leq F(b_k) + \varepsilon 2^{-k}. \tag{1.38}$$

Du recouvrement du compact $[a + \delta_0, b]$ par la suite d'ouverts $\big(]a_k, b_k + \delta_k[\big)_{k\geq 1}$, on peut extraire un recouvrement *fini*. Il existe donc un entier n (dépendant de F, de ε et des suites (a_k), (b_k)) tel que

$$]a + \delta_0, b] \subset [a + \delta_0, b] \subset \bigcup_{k=1}^{n}]a_k, b_k + \delta_k[\subset \bigcup_{k=1}^{n}]a_k, b_k + \delta_k].$$

Ayant déjà établi l'additivité de μ sur la semi-algèbre $\mathcal{C}$, nous pouvons appliquer le lemme 1.33 b) à cette suite d'inclusions pour obtenir $\mu(\,]a+\delta_0, b]) \leq \sum_{k=1}^{n} \mu(\,]a_k, b_k+\delta_k])$, ce qui s'écrit aussi :

$$F(b) - F(a + \delta_0) \leq \sum_{k=1}^{n} \big(F(b_k + \delta_k) - F(a_k)\big). \tag{1.39}$$

Compte-tenu de (1.36) et (1.38) et de la positivité des $F(b_k) - F(a_k)$ par croissance de F, on en déduit :

$$F(b) - F(a) - \varepsilon \leq \sum_{k=1}^{n} \big(F(b_k) - F(a_k)\big) + \sum_{k=1}^{n} \varepsilon 2^{-k} \leq \sum_{k=1}^{+\infty} \big(F(b_k) - F(a_k)\big) + \varepsilon.$$

Finalement, nous avons montré que :

$$\forall \varepsilon > 0, \quad F(b) - F(a) \leq \sum_{k=1}^{+\infty} \big(F(b_k) - F(a_k)\big) + 2\varepsilon.$$

En faisant tendre ε vers 0, on en déduit (1.35). Ceci termine la vérification des hypothèses du théorème 1.31. L'existence d'une mesure prolongeant μ de $\mathcal{C}$ à $\mathrm{Bor}(\mathbb{R})$ est maintenant acquise.

Unicité du prolongement de μ à $\mathrm{Bor}(\mathbb{R})$. Si μ et ν sont deux mesures sur $\mathrm{Bor}(\mathbb{R})$, qui coïncident sur $\mathcal{C}$, elles sont σ-finies puisque $\mathbb{R} = \cup_{n\in\mathbb{N}^*}] - n, n]$ et

$$\mu(] - n, n]) = \nu(] - n, n]) = F(n) - F(-n) < +\infty,$$

en rappelant que F est à valeurs dans $\mathbb{R}$, donc $F(n) - F(-n)$ est un *nombre* réel, donc fini. Par continuité séquentielle croissante des mesures, comme $]-n, n] \uparrow \mathbb{R}$, on en déduit que $\mu(\mathbb{R}) = \nu(\mathbb{R}) \leq +\infty$. D'autre part, $\mathcal{C}$ est une π-classe qui engendre $\mathrm{Bor}(\mathbb{R})$. Par le théorème 1.34, cas a) si $\mu(\mathbb{R}) < +\infty$ et b) sinon, les deux mesures μ et ν coïncident sur $\sigma(\mathcal{C}) = \mathrm{Bor}(\mathbb{R})$. $\square$

Nous disposons donc maintenant d'une grande variété de mesures sur $(\mathbb{R}, \mathrm{Bor}(\mathbb{R}))$, puisque toute fonction croissante continue à droite F génère une mesure de Stieltjes $\mu = \mu_F$. Bien sûr cette représentation $\mu = \mu_F$ n'est pas unique puisque pour toute constante c, les fonctions F et $F + c$ génèrent la même mesure de Stieltjes. Malgré sa richesse, ce procédé de construction de mesures sur $(\mathbb{R}, \mathrm{Bor}(\mathbb{R}))$ ne les fournit pas toutes, comme on peut le voir par la caractérisation suivante des mesures de Stieltjes.

Proposition 1.37. *Une mesure μ sur $(\mathbb{R}, \mathrm{Bor}(\mathbb{R}))$ est de Stieltjes si et seulement si elle est finie sur les intervalles* bornés *de $\mathbb{R}$.*

Démonstration. Soit μ une mesure de Stieltjes, associée à la fonction croissante $F : \mathbb{R} \to \mathbb{R}$. Si I est un intervalle borné de $\mathbb{R}$, on peut l'inclure dans un $]a, b]$ avec a et b réels finis. Alors par croissance des mesures pour l'inclusion,

$$0 \leq \mu(I) \leq \mu(]a, b]) = F(b) - F(a) < +\infty,$$

puisque F est à valeurs dans $\mathbb{R}$. Ceci montre qu'une mesure de Stieltjes est finie sur tout intervalle borné[5].

Réciproquement, soit μ une mesure sur $(\mathbb{R}, \mathrm{Bor}(\mathbb{R}))$, telle que pour tout intervalle borné I de $\mathbb{R}$, $\mu(I) < +\infty$. Montrons que μ est de Stieltjes. Traitons d'abord le cas particulier où μ est une mesure finie ($\mu(\mathbb{R}) < +\infty$). On prend alors pour F la fonction de répartition de μ : $F(x) := \mu(]-\infty, x])$. On sait que F est croissante et continue à droite (th. 1.25). Notons μ_F la mesure de Stieltjes associée à F. Comme μ est finie, la remarque 1.22 nous permet d'écrire pour tout intervalle $]a, b]$,

$$\mu(]a, b]) = \mu(]-\infty, b] \setminus]-\infty, a]) = \mu(]-\infty, b]) - \mu(]-\infty, a]) = F(b) - F(a) = \mu_F(]a, b]).$$

Ainsi les mesures μ et μ_F coïncident sur la classe des intervalles $]a, b]$ donc sur $\mathrm{Bor}(\mathbb{R})$, par le théorème d'unicité. Donc $\mu = \mu_F$ est bien une mesure de Stieltjes.

Dans le cas général, associons à μ la fonction F définie par

$$F(x) := \begin{cases} \mu(]0, x]) & \text{si } x \geq 0, \\ -\mu(]x, 0]) & \text{si } x < 0. \end{cases}$$

μ étant finie sur les intervalles bornés, F est à valeurs dans $\mathbb{R}$. F est croissante car

- si $0 \leq x \leq y$, $]0, x] \subset]0, y]$ donc $F(x) = \mu(]0, x]) \leq \mu(]0, y]) = F(y)$;
- si $x < y < 0$, $]x, 0] \supset]y, 0]$, donc $\mu(]x, 0]) \geq \mu(]y, 0])$ et $-\mu(]x, 0]) \leq -\mu(]y, 0])$, d'où $F(x) \leq F(y)$;
- si $x < 0 \leq y$, $F(x) \leq 0$ et $F(y) \geq 0$, donc $F(x) \leq F(y)$.

Comme fonction monotone, F admet une limite à droite en tout point a de $\mathbb{R}$. Pour montrer que F est continue à droite au point a, il suffit donc de vérifier que $F(a + 1/n)$ tend vers $F(a)$ quand n tend vers $+\infty$. Ceci résulte immédiatement de la continuité monotone séquentielle de la mesure μ en notant que

- si $a \geq 0$, la suite d'intervalles $(]0, a + 1/n])_{n \geq 1}$ décroît (pour l'inclusion) et a pour intersection $]0, a]$;
- si $a < 0$, la suite $(]a + 1/n, 0])_{n \geq 1}$ croît (pour l'inclusion) et a pour réunion $]a, 0]$.

Ainsi la fonction croissante continue à droite F engendre une mesure de Stieltjes μ_F et pour voir que $\mu = \mu_F$, il ne reste plus qu'à vérifier que si $-\infty < a < b < +\infty$, $\mu(]a, b]) = F(b) - F(a)$. Par additivité finie de μ et la remarque 1.22, on a :

[5] Attention, ceci ne signifie pas qu'elle soit *bornée* sur la classe de tous les intervalles bornés. La situation n'est pas la même qu'à la sous-section 1.3.1 (μ finie sur $\mathcal{F}$ équivaut à μ bornée sur $\mathcal{F}$), car contrairement à la tribu $\mathcal{F}$, la classe des intervalles bornés n'a pas de plus grand élément pour l'inclusion.

– si $0 \leq a < b$, $\mu(]a,b]) = \mu(]0,b]\setminus]0,a]) = \mu(]0,b]) - \mu(]0,a]) = F(b) - F(a)$;
– si $a < 0 \leq b$, le découpage en union disjointe $]a,b] =]a,0] \cup]0,b]$ donne $\mu(]a,b]) = \mu(]a,0]) + \mu(]0,b]) = -F(a) + F(b)$;
– si $a < b < 0$, $\mu(]a,b]) = \mu(]a,0]\setminus]b,0]) = \mu(]a,0]) - \mu(]b,0]) = -F(a) - (-F(b)) = F(b) - F(a)$.

Le théorème d'unicité nous permet de conclure que $\mu = \mu_F$ est bien une mesure de Stieltjes. □

Exemple 1.8. Pour illustrer simplement la proposition 1.37, considérons les deux mesures ponctuelles

$$\mu := \sum_{k\in\mathbb{N}^*} \delta_k, \qquad \nu := \sum_{k\in\mathbb{N}^*} \delta_{1/k}.$$

μ est de Stieltjes car un intervalle borné ne contient qu'un ensemble fini d'entiers. Par contre ν n'est pas de Stieltjes puisque $\nu(]0,1]) = +\infty$.

1.3.9 Mesure de Lebesgue sur $\mathbb{R}^d$

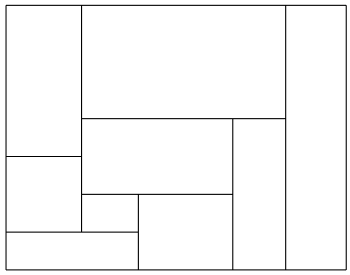
Théorème 1.38. *Il existe une unique mesure μ sur $(\mathbb{R}^d, \operatorname{Bor}(\mathbb{R}^d))$ telle que pour tout pavé $\prod_{i=1}^d]a_i, b_i]$ (avec $a_i < b_i$ pour tout $i = 1, \dots, d$),*

$$\mu\left(\prod_{i=1}^d]a_i, b_i]\right) = \prod_{i=1}^d (b_i - a_i). \tag{1.40}$$

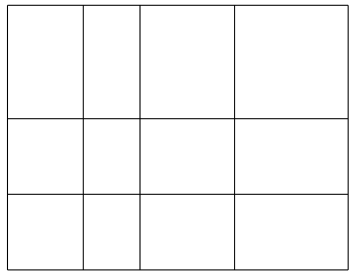
On l'appelle mesure de Lebesgue sur $\mathbb{R}^d$ *et on la note λ_d.*

Démonstration. On sait déjà que la classe $\mathcal{C}$ des pavés de la forme $\prod_{i=1}^d]a_i, b_i]$ est une semi-algèbre (exemple 1.7) qui engendre la tribu borélienne de $\mathbb{R}^d$ (prop. 1.19). Pour vérifier l'existence de la mesure de Lebesgue, on applique le théorème 1.31 à la fonction d'ensembles μ définie sur $\mathcal{C}$ par (1.40) lorsque le pavé $\prod_{i=1}^d]a_i, b_i]$ est non vide (c'est-à-dire lorsque $a_i < b_i$ pour chaque i) et par $\mu(\emptyset) = 0$ sinon. Il s'agit donc de vérifier que μ est additive sur $\mathcal{C}$ et sous-σ-additive sur $\mathcal{C}$.

Additivité de μ sur $\mathcal{C}$. Il nous faut vérifier que pour chaque décomposition d'un pavé $Q \in \mathcal{C}$ en réunion disjointe d'un nombre fini de pavés $C_j \in \mathcal{C}$, $1 \leq j \leq n$, $\mu(Q) = \sum_{j=1}^n \mu(C_j)$. Bien sûr, on suppose sans perdre de généralité qu'aucun des C_j n'est vide et dans cette égalité, $\mu(Q)$ et les $\mu(C_j)$ se calculent par la formule (1.40). Le cas facile est celui où les C_j forment ce que nous appelerons une « partition cartésienne » de Q, le cas d'une partition quelconque demandant un peu plus de travail. La figure ci-dessous illustre ces deux cas en dimension $d = 2$.



Partition quelconque



Partition cartésienne

Plus précisément, nous dirons que les $\{C_j;\ 1 \leq j \leq n\}$ forment une partition cartésienne de $Q = \prod_{i=1}^d]a_i, b_i]$ si pour chaque i, il existe une subdivision

$$a_i = c_{i,0} < \cdots < c_{i,k} < \cdots < c_{i,n_i} = b_i, \tag{1.41}$$

telle que $\{C_j;\ 1 \leq j \leq n\}$ soit l'ensemble de tous les pavés

$$Q_{k_1,\dots,k_d} := \prod_{i=1}^d]c_{i,k_i-1}, c_{i,k_i}], \quad 1 \leq k_i \leq n_i,\ 1 \leq i \leq d. \tag{1.42}$$

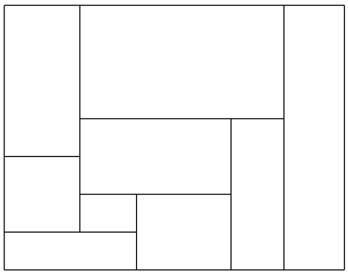
Notons $K := \{1, \dots, n_1\} \times \cdots \times \{1, \dots, n_d\}$ l'ensemble de tous les d-uples $(k_1, \dots, k_d)$, et remarquons que $n = \operatorname{card} K = n_1 \dots n_d$. Dans ce cas particulier, la formule traduisant l'additivité de μ se réduit au développement d'un produit fini de sommes finies, en écrivant $b_i - a_i = \sum_{k=1}^{n_i} (c_{i,k} - c_{i,k-1})$:

$$\begin{aligned}
\mu(Q) = \prod_{i=1}^d \left(\sum_{k=1}^{n_i} (c_{i,k} - c_{i,k-1}) \right) &= \sum_{(k_1,\dots,k_d)\in K} \prod_{i=1}^d (c_{i,k_i} - c_{i,k_i-1}) \\
&= \sum_{(k_1,\dots,k_d)\in K} \mu\big(Q_{k_1,\dots,k_d}\big) \\
&= \sum_{j=1}^n \mu(C_j).
\end{aligned}$$

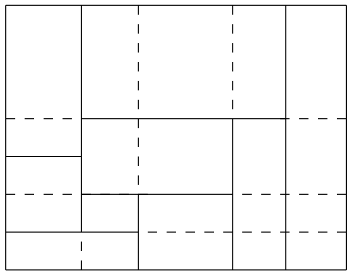
Passons au cas d'une partition quelconque de Q par les $C_j := \prod_{i=1}^d]a_{i,j}, b_{i,j}]$. L'idée est de se ramener à une partition cartésienne, en définissant pour tout $i = 1, \dots, d$, une subdivision (1.41) où les $c_{i,k}$ sont tels que

$$\big\{c_{i,k};\ 0 \leq k \leq n_i\big\} = \bigcup_{j=1}^n \{a_{i,j}, b_{i,j}\}.$$

Autrement dit, on prend toutes les extrémités des segments projections des C_j sur la i-ème coordonnée et on les range par ordre croissant après avoir effacé les répétitions.



Partition quelconque



Partition cartésienne induite

On induit ainsi une partition cartésienne de Q par les pavés $Q_{k_1,\dots,k_d}$ définis par (1.42), avec $n_i \leq n+2$. Avec K défini ci-dessus, on a donc

$$\mu(Q) = \sum_{(k_1,\dots,k_d)\in K} \mu\big(Q_{k_1,\dots,k_d}\big). \tag{1.43}$$

D'autre part, chacun des pavés C_j a lui aussi une partition cartésienne induite par les $Q_{k_1,\dots,k_d}$. En effet, pour tout i, $a_{i,j}$ et $b_{i,j}$ font partie par construction, de la subdivision $\{c_{i,k};\ 0 \leq k \leq n_i\}$. Il existe donc deux indices $l(i,j)$ et $m(i,j)$ tels que $a_{i,j} = c_{i,l(i,j)}$ et $b_{i,j} = c_{i,m(i,j)}$. Les $\{c_{i,k};\ l(i,j) \leq k \leq m(i,j)\}$ forment une subdivision de $]a_{i,j}, b_{i,j}]$. En notant $K_j := \{l(1,j),\dots,m(1,j)\} \times \cdots \times \{l(d,j),\dots,m(d,j)\}$, on a donc

$$\mu(C_j) = \sum_{(k_1,\dots,k_d)\in K_j} \mu\big(Q_{k_1,\dots,k_d}\big). \tag{1.44}$$

Pour conclure, il ne reste plus qu'à remarquer que les K_j (($1 \leq j \leq n$) forment une partition de l'ensemble de multi-indices K. Compte-tenu de (1.43) et (1.44), on en déduit

$$\sum_{j=1}^{n} \mu(C_j) = \sum_{j=1}^{n} \sum_{(k_1,\dots,k_d)\in K_j} \mu\big(Q_{k_1,\dots,k_d}\big) = \sum_{(k_1,\dots,k_d)\in K} \mu\big(Q_{k_1,\dots,k_d}\big) = \mu(Q).$$

Sous-σ-additivité de μ sur $\mathcal{C}$. La preuve est quasiment la même que celle déjà vue au théorème 1.36. On considère une décomposition en union dénombrable dans $\mathcal{C}$:

$$A = \bigcup_{j\in\mathbb{N}^*} A_j, \quad \text{où} \quad A := \prod_{i=1}^{d}]a_i, b_i], \quad A_j := \prod_{i=1}^{d}]a_{i,j}, b_{i,j}],\ j \in \mathbb{N}^*$$

et il s'agit de montrer que

$$\mu(A) \leq \sum_{j=1}^{+\infty} \mu(A_j). \tag{1.45}$$

Fixons $\varepsilon > 0$. Par continuité en zéro de la fonction polynomiale $t \mapsto \prod_{i=1}^{d}(b_i - a_i - t)$, il existe un $\delta > 0$ tel que

$$\mu(B) > \mu(A) - \varepsilon, \text{ où } B := \prod_{i=1}^{d}]a_i + \delta, b_i]. \tag{1.46}$$

Par un argument similaire, on a aussi

$$\forall j \geq 1,\ \exists \delta_j > 0 \text{ tel qu'avec } B_j := \prod_{i=1}^{d}]a_{i,j}, b_{i,j} + \delta_j], \quad \mu(B_j) < \mu(A_j) + \varepsilon 2^{-j}. \tag{1.47}$$

On a alors en notant $\overline{B} := \prod_{i=1}^{d} [a_i + \delta, b_i]$ et $\mathring{B}_j := \prod_{i=1}^{d}]a_{i,j}, b_{i,j} + \delta_j[$, les inclusions

$$\overline{B} \subset A = \bigcup_{j \in \mathbb{N}^*} A_j \subset \bigcup_{j \in \mathbb{N}^*} \mathring{B}_j.$$

Du recouvrement du compact $\overline{B}$ par les ouverts $\mathring{B}_j$, on peut extraire un recouvrement fini. Il existe donc un entier n tel que

$$B \subset \overline{B} \subset \bigcup_{j=1}^{n} \mathring{B}_j \subset \bigcup_{j=1}^{n} B_j.$$

L'additivité de μ sur $\mathcal{C}$ étant déjà prouvée, le lemme 1.33 b) nous dit alors que

$$\mu(B) \leq \sum_{j=1}^{n} \mu(B_j).$$

Compte-tenu de (1.46) et (1.47), on en déduit

$$\mu(A) - \varepsilon \leq \sum_{j=1}^{n} \mu(B_j) \leq \sum_{j=1}^{n} \mu(A_j) + \sum_{j=1}^{n} \varepsilon 2^{-j} \leq \sum_{j=1}^{+\infty} \mu(A_j) + \varepsilon.$$

Nous avons ainsi montré que

$$\forall \varepsilon > 0, \quad \mu(A) \leq \sum_{j=1}^{+\infty} \mu(A_j) + 2\varepsilon.$$

On en déduit (1.45), ce qui établit la sous-σ-additivité sur $\mathcal{C}$. Les hypothèses du théorème 1.31 sont maintenant vérifiées, ce qui nous assure de l'existence d'un prolongement de la fonction d'ensembles μ sur $\mathcal{C}$ en une mesure sur $\sigma(\mathcal{C}) = \mathrm{Bor}(\mathbb{R}^d)$.

Unicité de μ. Remarquons d'abord que toute mesure μ vérifiant (1.40) est σ-finie puisque :

$$\mathbb{R}^d = \bigcup_{n \in \mathbb{N}^*}]-n, n]^d \quad \text{et} \quad \forall n \in \mathbb{N}^*, \quad \mu\big(]-n, n]^d\big) = (2n)^d < +\infty.$$

Comme $\mathcal{C}$ est une π-classe qui engendre la tribu borélienne de $\mathbb{R}^d$, le théorème 1.34 b), appliqué avec $A_n =]-n, n]^d$ nous donne immédiatement l'unicité de μ. □

Proposition 1.39. *La mesure de Lebesgue λ_d sur $(\mathbb{R}^d, \mathrm{Bor}(\mathbb{R}^d))$ a les propriétés suivantes.*

i) λ_d est σ-finie.

ii) λ_d est invariante par translations.

iii) λ_d *est invariante par les « symétries » du type* s_ε *définies comme les applications linéaires transformant la base canonique* $(e_1, \dots, e_d)$ *en* $(\varepsilon_1 e_1, \dots, \varepsilon_d e_d)$*, où* $\varepsilon = (\varepsilon_1, \dots, \varepsilon_d) \in \{-1, 1\}^d$.

iv) Si h *est l'homothétie* $x \mapsto cx$ *dans* $\mathbb{R}^d$*, pour tout borélien* B*,* $\lambda_d\big(h(B)\big) = |c|^d \lambda_d(B)$.

v) λ_d *ne charge pas les points :* $\forall x \in \mathbb{R}^d$*,* $\lambda_d(\{x\}) = 0$*. Si* $A \subset \mathbb{R}^d$ *est fini ou dénombrable,* $\lambda_d(A) = 0$.

vi) $\lambda_d\left(\prod_{i=1}^{d}[a_i, b_i]\right) = \lambda_d\left(\prod_{i=1}^{d}]a_i, b_i[\right)$ *et cette égalité implique bien sûr, l'égalité des mesures des* 4^d *pavés obtenus en jouant sur l'ouverture ou la fermeture des extrémités* a_i*,* b_i *des intervalles.*

vii) Si E *est un sous-espace affine de* $\mathbb{R}^d$ *et* $E \neq \mathbb{R}^d$*,* $\lambda_d(E) = 0$.

La σ-finitude a déjà été vue ci-dessus (unicité de μ).

Preuve de ii)–iv). Notons encore $\mathcal{C}$ la classe des pavés $\prod_{i=1}^d]a_i, b_i]$. Soit g l'une des transformations translation, symétrie ou homothétie. En admettant provisoirement que $\lambda_d \circ g$ est une mesure sur $(\mathbb{R}^d, \mathrm{Bor}(\mathbb{R}^d))$, on voit immédiatement qu'elle coïncide sur $\mathcal{C}$ avec la mesure λ_d (cas ii) et iii)) ou $|c|^d \lambda_d$ (cas iv)). Le théorème d'unicité permet alors de conclure que ces mesures coïncident sur toute la tribu borélienne de $\mathbb{R}^d$ (détails laissés au lecteur). On en déduit que pour tout borélien B, $\lambda_d\big(g(B)\big) = \lambda_d(B)$ dans les cas ii) et iii) et $\lambda_d\big(g(B)\big) = |c|^d \lambda_d(B)$ dans le cas iv).

Revenons à la nature de la fonction d'ensembles $\lambda_d \circ g$. D'abord, est-elle bien définie sur $\mathrm{Bor}(\mathbb{R}^d)$? Il s'agit de vérifier que pour tout borélien B de $\mathbb{R}^d$, son image ensembliste directe $g(B)$ est encore un borélien de $\mathbb{R}^d$. L'application g étant bijective, soit $f := g^{-1}$ son inverse ponctuel. On a pour toute partie B de $\mathbb{R}^d$,

$$g(B) = \{g(y);\ y \in B\} = \{f^{-1}(y);\ y \in B\} = \{x \in \mathbb{R}^d;\ f(x) \in B\} = f^{-1}(B).$$

Par la proposition 1.11, $f^{-1}\big(\sigma(\mathcal{C})\big) = \sigma\big(f^{-1}(\mathcal{C})\big)$. On voit immédiatement que la classe des pavés $\mathcal{C}$ est globalement invariante par g : $f^{-1}(\mathcal{C}) = \mathcal{C}$. Par conséquent, $f^{-1}\big(\mathrm{Bor}(\mathbb{R}^d)\big) = \sigma(\mathcal{C}) = \mathrm{Bor}(\mathbb{R}^d)$. Ainsi pour tout borélien B, $f^{-1}(B)$ est lui-même un borélien et comme $f^{-1}(B) = g(B)$, ceci montre que la fonction d'ensembles $\lambda_d \circ g$ est bien définie sur $\mathrm{Bor}(\mathbb{R}^d)$.

Pour vérifier la σ-additivité de $\lambda_d \circ g$, soit $B = \cup_{i\in\mathbb{N}} B_i$ une réunion dénombrable disjointe de boréliens de $\mathbb{R}^d$. On a $g(\cup_{i\in\mathbb{N}} B_i) = \cup_{i\in\mathbb{N}} g(B_i)$ (c'est vrai pour une application quelconque) et les B_i étant deux à deux disjoints, les $g(B_i)$ le sont aussi, parce que g est *injective*. La σ-additivité de $\lambda_d \circ g$ se déduit alors de celle de λ_d en écrivant :

$$(\lambda_d \circ g)(B) = \lambda_d\big(g(B)\big) = \lambda_d\Big(\bigcup_{i\in\mathbb{N}} g(B_i)\Big) = \sum_{i\in\mathbb{N}} \lambda_d\big(g(B_i)\big) = \sum_{i\in\mathbb{N}} (\lambda_d \circ g)(B_i).$$

En outre, on a immédiatement $(\lambda_d \circ g)(\emptyset) = \lambda_d\big(g(\emptyset)\big) = \lambda_d(\emptyset) = 0$. Ainsi la fonction d'ensembles $\lambda_d \circ g$ est bien une mesure sur $\mathrm{Bor}(\mathbb{R}^d)$. □

Remarque 1.40. Dans ce qui précède, nous avons délibérément adopté le point de vue naïf selon lequel λ_d est invariante par une transformation f si $\lambda_d \circ f = \lambda_d$. Ceci a occasionné quelques contorsions pour vérifier que $\lambda_d \circ f$ est une mesure, la bijectivité de f jouant un rôle clé. Nous verrons au chapitre suivant que l'on peut définir la mesure image $\lambda_d \circ f^{-1}$ sous des conditions bien moins restrictives sur f (on demandera seulement que f soit mesurable). La « bonne définition » de l'invariance de λ_d par la transformation f s'écrira alors $\lambda_d \circ f^{-1} = \lambda_d$. Le lecteur pourra, à titre d'exercice, redémontrer ii) et iii) avec cette nouvelle définition.

Preuve de v) et vi). Ces deux propriétés se montrent par des arguments de continuité monotone séquentielle. Pour tout $x = (x_1, \dots, x_d) \in \mathbb{R}^d$, le singleton $\{x\}$ peut s'écrire comme l'intersection de la suite décroissante de pavés $C_n := \prod_{i=1}^d]x_i - 1/n, x_i]$. Par continuité décroissante de λ_d,

$$C_n \downarrow \{x\} \quad \Rightarrow \quad \lambda_d(\{x\}) = \lim_{n\to+\infty} \lambda_d(C_n) = \lim_{n\to+\infty} n^{-d} = 0.$$

Si A est fini ou dénombrable, par additivité finie ou σ-additivité de λ_d, on peut écrire

$$\lambda_d(A) = \lambda_d\Big(\bigcup_{x\in A}\{x\}\Big) = \sum_{x\in A} \lambda_d(\{x\}) = 0.$$

Noter que cet argument n'est plus valable si A n'est pas dénombrable : par exemple on a bien l'union disjointe $[0,1] = \bigcup_{x\in[0,1]}\{x\}$, mais $\lambda_1([0,1]) = 1 \neq 0$.

Pour prouver vi), on ne perd pas de généralité en supposant que $a_i \leq b_i$ pour tout i. Posons pour tout $n \geq 1$,

$$A_n := \prod_{i=1}^d \big]a_i - 1/n, b_i\big], \quad B_n := \prod_{i=1}^d \big]a_i, b_i - 1/n\big].$$

On voit facilement que le pavé fermé $\prod_{i=1}^d [a_i, b_i]$ est l'intersection de la suite décroissante $(A_n)_{n\geq1}$ et que le pavé ouvert $\prod_{i=1}^d]a_i, b_i[$ est la réunion de la suite croissante $(B_n)_{n\geq1}$. Par continuité monotone séquentielle de λ_d, on en déduit

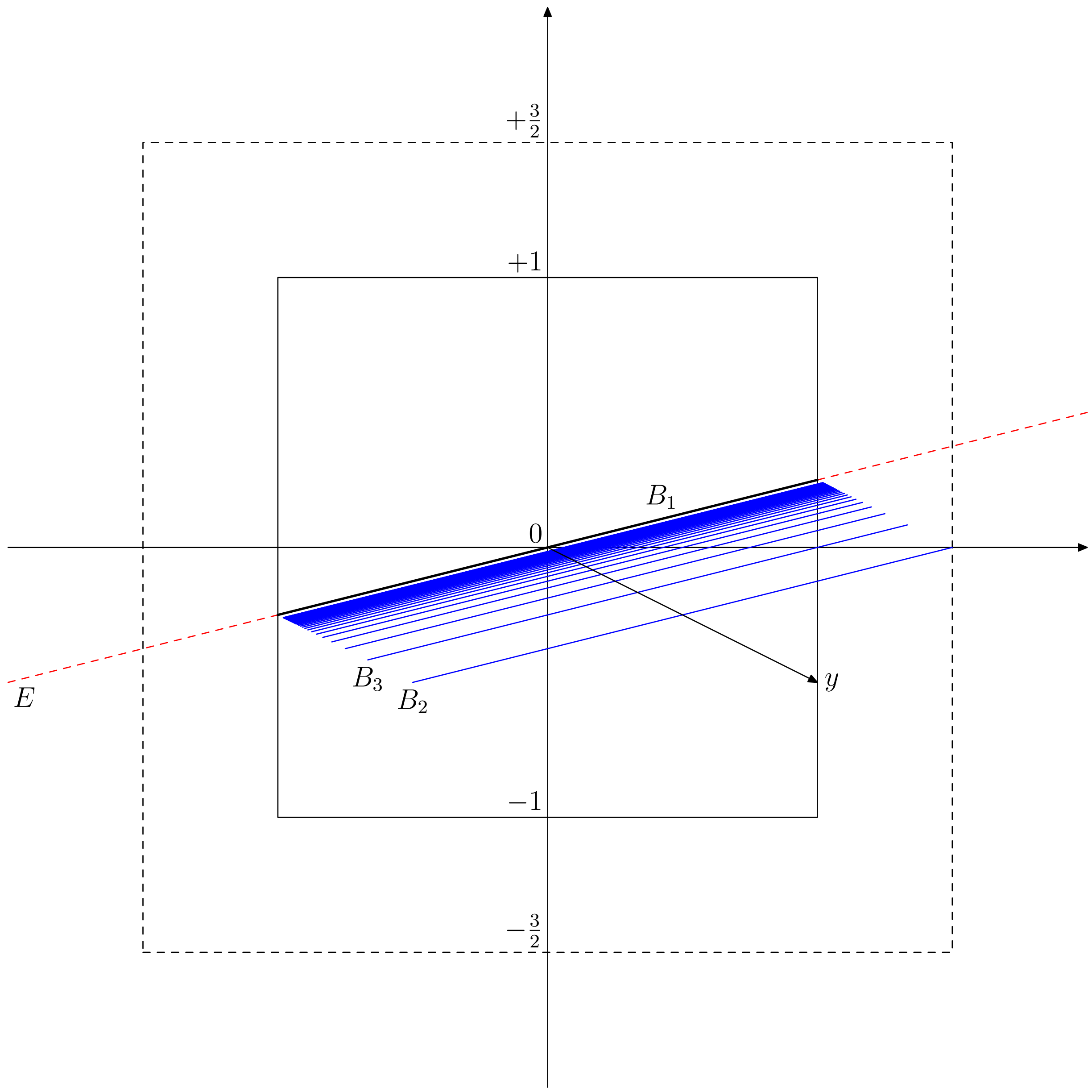
$$\lambda_d\left(\prod_{i=1}^d]a_i, b_i[\right) = \lim_{n\to+\infty} \lambda_d(A_n) = \lim_{n\to+\infty} \prod_{i=1}^d (b_i - a_i + 1/n) = \prod_{i=1}^d (b_i - a_i)$$

et

$$\lambda_d\left(\prod_{i=1}^d [a_i, b_i]\right) = \lim_{n\to+\infty} \lambda_d(B_n) = \lim_{n\to+\infty} \prod_{i=1}^d (b_i - a_i - 1/n)^+ = \prod_{i=1}^d (b_i - a_i),$$

d'où la conclusion. On a utilisé la notation $(b_i - a_i - 1/n)^+$ pour prendre en compte l'éventualité $a_i \geq b_i - 1/n$, auquel cas B_n est l'ensemble vide, de mesure nulle. □

Preuve de vii). L'idée est de regarder la trace B_1 de E sur le cube unité et d'en accumuler dans un volume *fini*, une infinité de translatés parallèles, donc disjoints. Comme tous ces translatés ont même mesure, celle-ci ne peut être que nulle et on en déduit $\lambda_d(E) = 0$ en utilisant les homothétiques de B_1. La figure suivante, dessinée dans le cas $d = 2$, illustre cette idée.



Voici les détails de la preuve en dimension d quelconque. En raison de l'invariance de λ_d par translation, on peut se ramener au cas où E est un sous-espace vectoriel de $\mathbb{R}^d$ (i.e. $0 \in E$). Par hypothèse, il existe $y \in \mathbb{R}^d$ tel que $y \notin E$ et comme $0 \in E$, y n'est pas nul. Comme E est un espace vectoriel, aucun vecteur cy, $c \in \mathbb{R}^*$ n'est dans E. Quitte à remplacer y par $\frac{1}{\|y\|_\infty} y$, on peut supposer que $\|y\|_\infty = 1$ (on note $\|x\|_\infty := \max_{1 \leq i \leq d} |x_i|$). Pour tout c réel, notons $E + cy$ le sous-espace affine, image de E par la translation de vecteur cy : $E + cy := \{x + cy ;\, x \in E\}$. Pour $c \neq c'$, les sous-espaces affines parallèles $E + cy$ et $E + c'y$ sont disjoints ; sinon ils auraient un élément commun z qui pourrait s'écrire $z = x + cy = x' + c'y$ pour des vecteurs x et x' de E et on en déduirait $y = \frac{1}{c - c'}(x - x')$, donc $y \in E$, ce qui est exclu. Définissons maintenant :

$$B_1 := E \cap [-1, 1]^d, \quad B_k := B_1 + \frac{1}{k}y, \quad (k \geq 2).$$

Comme $\|y\|_\infty = 1$, on a $\bigcup_{k\geq 1} B_k \subset [-3/2, 3/2]^d$, donc

$$\lambda_d\Bigl(\bigcup_{k\geq 1} B_k\Bigr) \leq 3^d < +\infty. \tag{1.48}$$

Les B_k sont deux à deux disjoints car chaque B_k est inclus dans $E + \frac{1}{k}y$. En utilisant la σ-additivité de λ_d et son invariance par translation, on voit que pour tout entier $n \geq 1$,

$$\lambda_d\Bigl(\bigcup_{k\geq 1} B_k\Bigr) = \sum_{k=1}^{+\infty} \lambda_d(B_k) \geq \sum_{k=1}^{n} \lambda_d(B_k) = n\lambda_d(B_1).$$

Compte-tenu de (1.48), on a ainsi $n\lambda_d(B_1) \leq 3^d$ pour tout $n \geq 1$ et ceci impose $\lambda_d(B_1) = 0$. Soit maintenant h_n l'homothétie $x \mapsto nx$. Par la propriété iv), $\lambda_d\bigl(h_n(B_1)\bigr) = n^d\lambda_d(B_1) = 0$. D'autre part il est clair que $h_n(B_1) = E \cap [-n, n]^d$. On en déduit que $\lambda_d(E) = 0$ par continuité croissante séquentielle puisque $E \cap [-n, n]^d \uparrow E$. $\square$