Chapter 2

Tribus et mesures

2.1 Introduction... par les probabilités

2.1.1 Cas d'un problème "discret"

Pour introduire la série de définitions qui suivent, commençons par quelques exemples, tirés du calcul des probabilités. Le calcul des probabilités s'intéresse à mesurer la "chance" qu'un certain "événement", résultat d'une expérience, a de se produire. Considérons par exemple "l'expérience" qui consiste à lancer un dé. On appelle "éventualité" associée à cette expérience un des résultats possibles de cette expérience, et "univers des possibles" l'ensemble E de ces éventualités. Dans notre exemple, les éventualités peuvent être $1, 2, 3, 4, 5$ ou 6; on pourrait choisir aussi comme éventualités les résultats correspondant au "dé cassé". On peut donc tout de suite remarquer que l'ensemble E des univers du possible dépend de la modélisation, c'est à dire de la formalisation mathématique que l'on fait du problème. Notons qu'il est parfois difficile de définir l'ensemble E.

A partir des éventualités, qui sont, par définition, les éléments de l'univers des possibles E, on définit les "événements", qui forment un ensemble de parties de E. Dans notre exemple du lancer de dé, l'ensemble des événements est l'ensemble des parties de E, noté $\mathcal{P}(E)$. Dans l'exemple du dé, la partie $\{2, 4, 6\}$ de E est l'événement : "le résultat du lancer est pair". On appelle événement élémentaire un singleton, par exemple $\{6\}$ dans notre exemple du lancer de dé, événement certain l'ensemble E tout entier, et l'événement "vide" l'ensemble vide $\emptyset$ (qui a donc une "chance" nulle de se réaliser). Pour mesurer "la chance" qu'a un événement de se réaliser, on va définir une application p de l'ensemble des événements (donc de $\mathcal{P}(E)$ dans notre exemple du lancer de dé) dans $[0, 1]$ avec certaines propriétés (qui semblent naturelles...). La "chance" (ou probabilité) pour un événement $A \subset E$ de se réaliser sera donc le nombre $p(A)$, appartenant à $[0, 1]$.

L'exemple du lancer de dé, que nous venons de considérer, est un problème discret fini, au sens ou l'ensemble E est fini. On peut aussi envisager des problèmes discrets infinis, l'ensemble E est alors infini dénombrable (on rappelle qu'un ensemble I est "dénombrable" s'il existe une bijection de I dans $\mathbb{N}$, il est "au plus dénombrable" s'il existe une injection de I dans $\mathbb{N}$), ou des problèmes (parfois appelés "continus") où E est infini non dénombrable.

2.1.2 Exemple continu

Considérons maintenant “l'expérience” qui consiste à lancer une balle de ping-pong sur une table de ping-pong. Soit E l'ensemble des points de la table de ping-pong, on peut voir E comme un sous-ensemble de $\mathbb{R}^2$, un événement élémentaire est alors un point $(x, y) \in E$ (le point d'impact de la balle), et un événement semble être une partie quelconque A de $\mathcal{P}(E)$. On suppose qu'on a effectué le lancer “sans viser”, c'est à dire en supposant que “n'importe quel point de la table a une chance égale d'être atteint” (les événements élémentaires sont “équiprobables”), et que la balle tombe forcément sur la table (on est très optimiste...). On se rend compte facilement que la probabilité pour chacun des points de E d'être atteint doit être nulle, puisque le nombre des points est infini. On peut aussi facilement “intuiter” que la probabilité pour une partie A d'être atteinte (dans le modèle “équiprobable”) est le rapport entre la “surface de A” et la surface de E. La notion intuitive de “surface” correspond en fait à la notion mathématique de “mesure” que nous allons définir dans le prochain paragraphe. Malheureusement, comme on l'a dit dans le chapitre introductif, il ne nous sera pas mathématiquement possible de définir une application convenable, i.e. qui vérifie les propriétés (1.9)-(1.10) et qui “mesure” toutes les parties de $\mathbb{R}$, ou $\mathbb{R}^2$, ou même du sous-ensemble E de $\mathbb{R}^2$ (voir à ce sujet l'exercice 2.27). On va donc définir un sous-ensemble de $\mathcal{P}(E)$ (qu'on appelle “tribu”) sur lequel on pourra définir une telle application. Dans le cas d'un ensemble fini, la tribu sera, en général, $\mathcal{P}(E)$ tout entier. Mais, dans le cas de la balle de ping-pong que vous venons de décrire, l'ensemble des événements sera une tribu strictement incluse dans $\mathcal{P}(E)$.

2.2 Tribu ou $\sigma-$algèbre

Définition 2.1 (Tribu ou $\sigma-$algèbre) *Soient E un ensemble, T une famille de parties de E (i.e. $T \subset \mathcal{P}(E)$). La famille T est une tribu (on dit aussi une $\sigma-$algèbre) sur E si T vérifie :*

1. $\emptyset \in T, E \in T$,
2. *T est stable par union dénombrable, c'est-à-dire que pour toute famille dénombrable $(A_n)_{n\in\mathbb{N}} \subset T$, on a $\cup_{n\in\mathbb{N}} A_n \in T$.*
3. *T est stable par intersection dénombrable, c'est-à-dire que pour toute famille dénombrable $(A_n)_{n\in\mathbb{N}} \subset T$, on a $\cap_{n\in\mathbb{N}} A_n \in T$.*
4. *T est stable par passage au complémentaire, c'est-à-dire que pour tout $A \in T$, on a $A^c \in T$ (On rappelle que $A^c = E \setminus A$).*

Il est clair que, pour montrer qu'une partie T de $\mathcal{P}(E)$ est une tribu, il est inutile de vérifier les propriétés 1-4 de la proposition précédente. Il suffit de vérifier par exemple $\emptyset \in T$ (ou $E \in T$), 2 (ou 3) et 4.

Exemples de tribus sur E :

- $T = \{\emptyset, E\}$,
- $\mathcal{P}(E)$.

Définition 2.2 (Langage probabiliste) *Soient E un ensemble quelconque (“l'univers des possibles”) et T une tribu ; on appelle “éventualité” les éléments de E et “événements” les éléments de T. On appelle “événement élémentaire” un singleton de T.*
On dit que deux événements $A, B \in T$ sont incompatibles si $A \cap B = \emptyset$.

Proposition 2.1 (Stabilité par intersection des tribus) *Soient E et I deux ensembles. Pour tout $i \in I$, on se donne une tribu, T_i, sur E. Alors, la famille (de parties de E) $\cap_{i \in I} T_i = \{A \subset E;\ A \in T_i, \forall i \in I\}$ est encore une tribu sur E.*

Démonstration : La démonstration de cette proposition fait l'objet de la première question de l'exercice 2.2. ■

Cette proposition nous permet de définir ci-après la notion de tribu engendrée.

Définition 2.3 (Tribu engendrée) *Soient E un ensemble et $\mathcal{C} \subset \mathcal{P}(E)$. On appelle tribu engendrée par $\mathcal{C}$ la plus petite tribu contenant $\mathcal{C}$, c'est-à-dire la tribu $T(\mathcal{C})$ intersection de toutes les tribus sur E contenant $\mathcal{C}$ (cette intersection est non vide car $\mathcal{P}(E)$ est une tribu contenant $\mathcal{C}$).*

Il est parfois utile d'utiliser la notion d'algèbre, qui est indentique à celle de tribu en remplaçant "dénombrable" par "finie".

Définition 2.4 (Algèbre) *Soient E un ensemble, $\mathcal{A}$ une famille de parties de E (i.e. $\mathcal{A} \subset \mathcal{P}(E)$). La famille $\mathcal{A}$ est une algèbre sur E si $\mathcal{A}$ vérifie :*

1. *$\emptyset \in \mathcal{A}, E \in \mathcal{A}$,*
2. *$\mathcal{A}$ est stable par union finie, c'est-à-dire que pour tout $A, B \in \mathcal{A}$ on a $A \cup B \in \mathcal{A}$.*
3. *$\mathcal{A}$ est stable par intersection finie, c'est-à-dire que pour tout $A, B \in \mathcal{A}$ on a $A \cap B \in \mathcal{A}$.*
4. *$\mathcal{A}$ est stable par passage au complémentaire, c'est-à-dire que pour tout $A \in \mathcal{A}$, on a $A^c \in \mathcal{A}$.*

Remarque 2.1 *Soit E un ensemble et $\mathcal{C} \subset \mathcal{P}(\mathcal{E})$. Comme pour les tribus, on peut définir l'algèbre engendrée par $\mathcal{C}$. C'est la plus petite algèbre contenant $\mathcal{C}$, c'est-à-dire l'intersection de toutes les algèbres contenant $\mathcal{C}$ (voir l'exercice 2.9).*

Soit E un ensemble, $\mathcal{C} \subset \mathcal{P}(\mathcal{E})$ et $T(\mathcal{C})$ la tribu engendrée par $\mathcal{C}$ (voir la définition 2.3 et l'exercice 2.2). Il est important de remarquer que, contrairement à ce que l'on pourrait être tenté de croire, les éléments de la tribu engendrée par $\mathcal{C}$ ne sont pas tous obtenus, à partir des éléments de $\mathcal{C}$, en utilisant les opérations : "intersection dénombrable", "union dénombrable" et "passage au complémentaire". Plus précisément, on pose :

$$\begin{aligned}
\mathcal{R}^1(\mathcal{C}) &= \{A \subset E \text{ t.q. } A = \bigcup_{n \in \mathbb{N}} A_n, A_n \in \mathcal{C} \text{ ou } A_n^c \in \mathcal{C}\}, \\
\mathcal{R}^2(\mathcal{C}) &= \{A \subset E \text{ t.q. } A = \bigcap_{n \in \mathbb{N}} A_n, A_n \in \mathcal{C} \text{ ou } A_n^c \in \mathcal{C}\}, \\
\mathcal{R}(\mathcal{C}) &= \mathcal{R}^1(\mathcal{C}) \cup \mathcal{R}^2(\mathcal{C}).
\end{aligned}$$

Prenons $E = \mathbb{R}$ et $\mathcal{C}$ l'ensemble des ouverts de $\mathbb{R}$ (donc $T(\mathcal{C})$ est la tribu borélienne de $\mathbb{R}$, voir définition ci-après). Il est facile de voir que $\mathcal{R}(\mathcal{C}) \subset T(\mathcal{C})$, mais que, par contre (et cela est moins facile à voir), $\mathcal{R}(\mathcal{C})$ n'est pas une tribu. En posant : $\mathcal{S}_0 = \mathcal{C}$, et $\mathcal{S}_n = \mathcal{R}(\mathcal{S}_{n-1})$, pour $n \geq 1$, on peut aussi montrer que $\overline{\mathcal{S}} = \bigcup_{n \in \mathbb{N}} \mathcal{S}_n$ n'est pas une tribu (et que $\overline{\mathcal{S}} \subset T(\mathcal{C})$).

Remarque 2.2 *Soit E un ensemble et $\mathcal{C}_1 \subset \mathcal{C}_2 \subset \mathcal{P}(E)$. Il est alors facile de voir que $T(\mathcal{C}_1) \subset T(\mathcal{C}_2)$ (cf. Exercice 2.2).*

Soit E un ensemble. On rappelle qu'une "topologie" sur E est donnée par une famille de parties de E, appelées "ouverts de E", contenant $\emptyset$ et E, stable par union (quelconque) et stable par intersection finie. L'ensemble E, muni de cette famile de parties, est alors un "espace topologique".

Définition 2.5 (Tribu borélienne) *Soit E un ensemble muni d'une topologie (un espace métrique, par exemple). On appelle tribu borélienne (ou tribu de Borel) la tribu engendrée par l'ensemble des ouverts de E, cette tribu sera notée $\mathcal{B}(E)$. Dans le cas $E = \mathbb{R}$, cette tribu est donc notée $\mathcal{B}(\mathbb{R})$.*

Remarque 2.3

1. L'objectif de la section 2.5 est de construire une application $\lambda : \mathcal{B}(\mathbb{R}) \to \overline{\mathbb{R}}_+$ t.q. :

   (a) $\lambda(]\alpha, \beta[) = \beta - \alpha$, pour tout $\alpha, \beta \in \mathbb{R}$ $\alpha < \beta$,

   (b) $\lambda(\cup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} \lambda(A_n)$, pour toute suite $(A_n)_{n\in\mathbb{N}} \subset \mathcal{B}(\mathbb{R})$ t.q. $A_n \cap B_n = \emptyset$ si $n \neq m$. (Noter que $\cup_{n\in\mathbb{N}} A_n \in \mathcal{B}(\mathbb{R})$ grâce à la stabilité d'une tribu par union dénombrable.)

2. On peut se demander si $\mathcal{B}(\mathbb{R}) = \mathcal{P}(\mathbb{R})$. La réponse est non (voir les exercices 2.27 et 2.28). On peut même démontrer que $\text{card}(\mathcal{B}(\mathbb{R})) = \text{card}(\mathbb{R})$ (alors que $\text{card}(\mathcal{P}(\mathbb{R})) > \text{card}(\mathbb{R})$). On donne ci-après un rappel rapide sur les cardinaux (sans entrer dans les aspects difficiles de la théorie des ensembles, et donc de manière peut-être un peu imprécise).

3. Rappel sur les cardinaux. Soit A et B deux ensembles.

   (a) On dit que $\text{card}(A) = \text{card}(B)$ si il existe une application $\varphi : A \to B$, bijective.

   Pour montrer que deux ensembles ont même cardinaux, il est souvent très utile d'utiliser le théorème de Bernstein (voir l'exercice 1.12) qui montre que s'il existe une injection de A dans B et une injection de B dans A, alors il existe $\varphi : A \to B$, bijective (et donc $\text{card}(A) = \text{card}(B)$). Le théorème de Bernstein motive également la définition suivante.

   (b) On dit que $\text{card}(A) \leq \text{card}(B)$ s'il existe $\varphi : A \to B$, injective.

   (c) Un autre théorème intéressant, dû à Cantor, donne que, pour tout ensemble X, on a $\text{card}(X) < \text{card}(\mathcal{P}(X))$ (c'est-à-dire $\text{card}(X) \leq \text{card}(\mathcal{P}(X))$ et $\text{card}(X) \neq \text{card}(\mathcal{P}(X))$). On a donc, en particulier, $\text{card}(\mathcal{P}(\mathbb{R})) > \text{card}(\mathbb{R})$. La démonstration du théorème de Cantor est très simple. Soit $\varphi : X \to \mathcal{P}(X)$. On va montrer que φ ne peut pas être surjective. On pose $A = \{x \in X;\ x \notin \varphi(x)\}$ (A peut être l'ensemble vide). Supposons que $A \in \text{Im}(\varphi)$. Soit alors $a \in X$ t.q. $A = \varphi(a)$.

   Si $a \in A = \varphi(a)$, alors $a \notin A$ par définition de A.

   Si $a \notin A = \varphi(a)$, alors $a \in A$ par définition de A.

   On a donc montré que A ne peut pas avoir d'antécédent (par φ) et donc φ n'est pas surjective.

Proposition 2.2 *On note $\mathcal{C}_1$ l'ensemble des ouverts de $\mathbb{R}$, $\mathcal{C}_2 = \{]a, b[,\ a, b \in \mathbb{R},\ a < b\}$ et $\mathcal{C}_3 = \{]a, \infty[,\ a \in \mathbb{R}\}$. Alors $T(\mathcal{C}_1) = T(\mathcal{C}_2) = T(\mathcal{C}_3) = \mathcal{B}(\mathbb{R})$. (Noter que d'autres caractérisations de $\mathcal{B}(\mathbb{R})$, semblables, sont possibles.)*

Démonstration : On a, par définition de $\mathcal{B}(\mathbb{R})$, $T(\mathcal{C}_1) = \mathcal{B}(\mathbb{R})$. On va démontrer ci-après que $T(\mathcal{C}_1) = T(\mathcal{C}_2)$ (le fait que $T(\mathcal{C}_2) = T(\mathcal{C}_3)$ est laissé au lecteur).

Comme $\mathcal{C}_2 \subset \mathcal{C}_1$, on a $T(\mathcal{C}_2) \subset T(\mathcal{C}_1)$. Il suffit donc de démontrer l'inclusion inverse. On va montrer que $\mathcal{C}_1 \subset T(\mathcal{C}_2)$, on aura alors que $T(\mathcal{C}_1) \subset T(\mathcal{C}_2)$.

Soit O un ouvert de $\mathbb{R}$. On suppose $O \neq \emptyset$ (on sait déjà que $\emptyset \in T(\mathcal{C}_2)$). Le lemme 2.1 (plus simple que le lemme 2.4 page 35) ci-après nous donne l'existence d'une famille $(I_n)_{n\in A}$ d'intervalles ouverts t.q. $A \subset \mathbb{N}$ et $O = \cup_{n\in A} I_n$. Noter qu'on a aussi $O = \cup_{n\in\mathbb{N}} I_n$ en posant $I_n = \emptyset$ si $n \in \mathbb{N} \setminus A$. Comme $I_n \in \mathcal{C}_2 \subset T(\mathcal{C}_2)$ pour tout $n \in A$ et $\emptyset \in T(\mathcal{C}_2)$, on en déduit, par stabilité dénombrable d'une tribu, que $O \in T(\mathcal{C}_2)$. Donc, $\mathcal{C}_1 \subset T(\mathcal{C}_2)$ et donc $T(\mathcal{C}_1) \subset T(\mathcal{C}_2)$. On a bien montré que $T(\mathcal{C}_1) = T(\mathcal{C}_2)$.

■

Lemme 2.1 *Tout ouvert non vide de $\mathbb{R}$ est réunion au plus dénombrable d'intervalles ouverts bornés.*

Démonstration : Soit O un ouvert de $\mathbb{R}$, $O \neq \emptyset$. On pose $A = \{(\beta,\gamma) \in \mathbb{Q}^2;\ \beta < \gamma,\ ]\beta,\gamma[\subset O\}$. On a donc $\cup_{(\beta,\gamma)\in A}]\beta,\gamma[\subset O$. On va montrer que $O \subset \cup_{(\beta,\gamma)\in A}]\beta,\gamma[$ (et donc que $O = \cup_{(\beta,\gamma)\in A}]\beta,\gamma[$).

Soit $x \in O$, il existe $\alpha_x > 0$ t.q. $]x-\alpha_x, x+\alpha_x[\subset O$. En prenant $\beta_x \in \mathbb{Q} \cap]x-\alpha_x, x[$ et $\gamma_x \in \mathbb{Q} \cap]x, x+\alpha_x[$ (de tels β_x et γ_x existent) on a donc $x \in]\beta_x, \gamma_x[\subset O$ et donc $(\beta_x, \gamma_x) \in A$. D'où $x \in]\beta_x, \gamma_x[\subset \cup_{(\beta,\gamma)\in A}]\beta,\gamma[$. On a bien montré que $O \subset \cup_{(\beta,\gamma)\in A}]\beta,\gamma[$ et donc que $O = \cup_{(\beta,\gamma)\in A}]\beta,\gamma[$. Comme $\mathbb{Q}^2$ est dénombrable, A est au plus dénombrable et le lemme est démontré. ■

Définition 2.6 (Espace mesurable ou probabilisable, partie mesurable ou probabilisable) *Soient E un ensemble, et T une tribu sur E. Le couple (E,T) est appelé "espace mesurable" ou (en langage probabiliste !) "espace probabilisable". Les parties de E qui sont (resp. ne sont pas) des éléments de T sont dites mesurables ou probabilisables (resp. non mesurables, non probabilisables).*

2.3 Mesure, probabilité

Définition 2.7 (Mesure) *Soit (E,T) un espace mesurable. On appelle mesure une application $m : T \to \overline{\mathbb{R}}_+$ (avec $\overline{\mathbb{R}}_+ = \mathbb{R}_+ \cup (+\infty)$) vérifiant :*

1. *$m(\emptyset) = 0$*
2. *m est σ-additive, c'est-à-dire que pour toute famille $(A_n)_{n\in\mathbb{N}} \subset T$ de parties disjointes deux à deux, (i.e. t.q. $A_n \cap A_m = \emptyset$, si $n \neq m$), on a :*

$$m(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} m(A_n). \tag{2.1}$$

Remarque 2.4

1. Dans la définition précédente on a étendu à $\overline{\mathbb{R}}_+$ l'addition dans $\mathbb{R}_+$. On a simplement posé $x + \infty = \infty$, pour tout $x \in \overline{\mathbb{R}}_+$. Noter également que la somme de la série dans la définition précédente est à prendre dans $\overline{\mathbb{R}}_+$ et que, bien sûr, $a = \sum_{n\in\mathbb{N}} a_n$ signifie simplement que $\sum_{p=0}^{n} a_p \to a$ (dans $\overline{\mathbb{R}}_+$) quand $n \to \infty$.
2. Soient $x, y, z \in \overline{\mathbb{R}}_+$. Remarquer que $x + y = x + z$ implique $y = z$ si $x \neq \infty$.

3. Dans la définition précédente, la condition 1. peut être remplacée par la condition : $\exists\ A \in T, m(A) < \infty$. La vérification de cette affirmation est laissée au lecteur attentif.
4. Il est intéressant de remarquer que, pour une série à termes positifs, l'ordre de sommation est sans importance. Plus précisément, si $(a_n)_{n\in\mathbb{N}} \subset \overline{\mathbb{R}}_+$ et si φ est une bijection de $\mathbb{N}$ dans $\mathbb{N}$, on a $\sum_{n\in\mathbb{N}} a_n = \sum_{n\in\mathbb{N}} a_{\varphi(n)}$. C'est l'objet du lemme 2.2.
5. Une conséquence immédiate de la $\sigma-$additivité est l'additivité, c'est-à-dire que
$$m(\cup_{p=0}^n A_p) = \sum_{p=0}^n m(A_p)$$
pour toute famille finie $(A_p)_{p=0,\ldots,n}$ d'éléments de T, disjoints 2 à 2. L'additivité se démontre avec la $\sigma-$additivité en prenant $A_p = \emptyset$ pour $p > n$ dans (2.1).
6. Dans le cas $E = \mathbb{R}$ et $T = \mathcal{P}(\mathbb{R})$, il est facile de construire des mesures sur T, mais il n'existe pas de mesure sur T, notée m, telle que $m(]a, b[) = b - a$ pour tout a, $b \in \mathbb{R}$, $a < b$ (voir les exercices 2.28 et 2.27). Une telle mesure existe si on prend pour T la tribu borélenne de $\mathbb{R}$, c'est l'objet de la section 2.5.

Lemme 2.2 *Soit* $(a_n)_{n\in\mathbb{N}} \subset \overline{\mathbb{R}}_+$ *et soit* $\varphi\ :\ \mathbb{N} \to \mathbb{N}$ *bijective. Alors* $\sum_{n\in\mathbb{N}} a_n = \sum_{n\in\mathbb{N}} a_{\varphi(n)}$.

Démonstration :
On pose $A = \sum_{n\in\mathbb{N}} a_n (= \lim_{n\to\infty} \sum_{p=0}^n a_p) \in \overline{\mathbb{R}}_+$ et $B = \sum_{n\in\mathbb{N}} a_{\varphi(n)} (= \lim_{n\to\infty} \sum_{p=0}^n a_{\varphi(p)}) \in \overline{\mathbb{R}}_+$. On veut montrer que $A = B$.

On montre d'abord que $B \le A$. Soit $n \in \mathbb{N}$. On pose $N = \max\{\varphi(0), \ldots, \varphi(n)\}$. Comme $a_q \ge 0$ pour tout $q \in \mathbb{N}$, on a $\sum_{p=0}^n a_{\varphi(p)} \le \sum_{p=0}^N a_p \le A$. On en déduit, faisant tendre n vers ∞ que $B \le A$.

En raisonnant avec l'inverse de φ on a aussi $A \le B$ et finalement $A = B$.

■

Définition 2.8 (Mesure finie) *Soient E un ensemble et T une tribu sur E. On appelle mesure finie une mesure m sur T telle que $m(E) < \infty$.*

Définition 2.9 (Probabilité) *Soient E un ensemble et T une tribu sur E. On appelle probabilité une mesure p sur T t.q. $p(E) = 1$.*

Définition 2.10 (Espace mesuré, espace probabilisé) *Soient E un ensemble, T une tribu sur E et m une mesure (resp. une probabilité) sur T. Le triplet (E, T, m) est appelé "espace mesuré" (resp. "espace probabilisé").*

Définition 2.11 (Mesure σ-finie) *Soit (E, T, m) un espace mesuré, on dit que m est σ-finie (ou que (E, T, m) est σ-fini) si :*
$$\exists\, (A_n)_{n\in\mathbb{N}} \subset T,\quad m(A_n) < \infty,\quad \forall\, n \in \mathbb{N},\ \textit{et}\quad E = \bigcup_{n\in\mathbb{N}} A_n. \tag{2.2}$$

Remarque 2.5 (Langage probabiliste) En langage probabiliste, la propriété de σ-additivité (2.1) que l'on requiert dans la définition d'une mesure (et donc d'une probabilité) est souvent appelé "axiome complet des probabilités totales".

Exemple 2.1 (Mesure de Dirac) Soient E un ensemble, T une tribu sur E et $a \in E$. On définit sur T la mesure δ_a par (pour $A \in T$) :

$$\delta_a(A) = 0, \quad \text{si } a \notin A, \tag{2.3}$$

$$\delta_a(A) = 1, \quad \text{si } a \in A. \tag{2.4}$$

On peut remarquer que la mesure de Dirac est une probabilité.

Remarque 2.6 (Comment choisir la probabilité) Soit (E,T) un espace probabilisable, on peut évidemment définir plusieurs probabilités sur T. C'est tout l'art de la modélisation que de choisir une probabilité qui rende compte du phénomène aléatoire que l'on veut observer. On se base pour cela souvent sur la notion de fréquence, qui est une notion expérimentale à l'origine. Soit $A \in T$ un événement, dont on cherche à évaluer la probabilité $p(A)$. On effectue pour cela N fois l'expérience dont l'univers des possibles est E, et on note N_A le nombre de fois où l'événement A est réalisé. A N fixé, on définit alors la *fréquence* $f_A(N)$ de l'événement A par :

$$f_A(N) = \frac{N_A}{N}.$$

Expérimentalement, il s'avère que $f_N(A)$ admet une limite lorsque $N \to +\infty$. C'est ce qu'on appelle la "loi empirique des grands nombres". On peut donc définir "expérimentalement" $p(A) = \lim_{N\to+\infty} f_N(A)$. Cependant, on n'a pas ainsi démontré que p est une probabilité: il ne s'agit pour l'instant que d'une approche intuitive. On démontrera plus loin la loi forte des grands nombres, qui permettra de justifier mathématiquement la loi empirique. On peut remarquer que $f_N(E) = \frac{N}{N} = 1\ldots$

Exemple 2.2 (Le cas "équiprobable") Soit (E,T,p) un espace probabilisé .On suppose que tous les singletons de E appartiennent à la tribu et que les événements élémentaires sont équiprobables. . On a alors: $p(\{x\}) = \frac{1}{\operatorname{card} E}$ pour tout $x \in E$.

Définition 2.12 (mesure atomique) *Soit (E,T,m) un espace mesuré tel que : $\{x\} \in T$ pour tout x de E. On dit que m est portée par $S \in T$ si $m(S^c) = 0$. Soit $x \in E$, on dit que x est un atome ponctuel de m si $m(\{x\}) \neq 0$. On dit que m est purement atomique si elle est portée par la partie de E formée par l'ensemble de ses atomes ponctuels.*

Définition 2.13 (Mesure diffuse) *Soient (E,T) un espace mesurable et m une mesure sur T . On dit que m est diffuse si $\{x\} \in T$ et $m(\{x\}) = 0$ pour tout $x \in E$. (Cette définition est aussi valable pour une mesure signée sur T, définie dans la section 2.4.)*

Définition 2.14 (Partie négligeable) *Soient (E,T,m) un espace mesuré et $A \subset E$. On dit que A est négligeable s'il existe un ensemble $B \in T$ tel que $A \subset B$ et $m(B) = 0$.*

Définition 2.15 (Mesure complète) *Soit (E,T,m) un espace mesuré, on dit que m est complète (ou que (E,T,m) est complet) si toutes les parties négligeables sont mesurables, c'est-à-dire appartiennent à T.*

La proposition suivante donne les principales propriétés d'une mesure.

Proposition 2.3 (Propriétés des mesures) *Soit (E,T,m) un espace mesuré. La mesure m vérifie les quatre propriétés suivantes :*

1. *Monotonie : Soit A, $B \in T$, $A \subset B$, alors*

$$m(A) \leq m(B) \tag{2.5}$$

2. *σ-sous-additivité : Soit $(A_n)_{n\in\mathbb{N}} \subset T$, alors*

$$m(\bigcup_{n\in\mathbb{N}} A_n) \leq \sum_{n\in\mathbb{N}} m(A_n). \tag{2.6}$$

3. *Continuité croissante : Soit $(A_n)_{n\in\mathbb{N}} \subset T$, t.q. $A_n \subset A_{n+1}$, pour tout $n \in \mathbb{N}$, alors*

$$m(\bigcup_{n\in\mathbb{N}} A_n) = \lim_{n\to\infty}(m(A_n)) = \sup_{n\in\mathbb{N}}(m(A_n)) \tag{2.7}$$

4. *Continuité décroissante : Soit $(A_n)_{n\in\mathbb{N}} \subset T$, t.q. $A_{n+1} \subset A_n$, pour tout $n \in \mathbb{N}$, et t.q. il existe $n_0 \in \mathbb{N}$, $m(A_{n_0}) < \infty$, alors*

$$m(\bigcap_{n\in\mathbb{N}} A_n) = \lim_{n\to\infty}(m(A_n)) = \inf_{n\in\mathbb{N}}(m(A_n)) \tag{2.8}$$

Démonstration :
La démonstration de ces propriétés est facile: elles découlent toutes du caractère positif et du caractère σ-additif de la mesure. Attention: ces propriétés ne sont pas vérifiées par les mesures signées que nous verrons à la section 2.4.

1. *Monotonie.* Soit $A, B \in T$, $A \subset B$. On a $B = A \cup (B \setminus A)$ et $A \cap (B \setminus A) = \emptyset$. Comme $A \in T$ et $B \setminus A = B \cap A^c \in T$, l'additivité de m (voir la remarque 2.4) donne $m(B) = m(A) + m(B \setminus A) \geq m(A)$, car m prend ses valeurs dans $\overline{\mathbb{R}}_+$.

   Noter aussi que $m(B \setminus A) = m(B) - m(A)$ si $0 \leq m(A) \leq m(B) < \infty$ (mais cette relation n'a pas de sens si $m(A) = m(B) = \infty$).

2. *$\sigma-$sous additivité.* Soit $(A_n)_{n\in\mathbb{N}} \subset T$. On veut montrer que $m(\cup_{n\in\mathbb{N}} A_n) \leq \sum_{n\in\mathbb{N}} m(A_n)$. On pose $B_0 = A_0$ et, par récurrence sur n, $B_n = A_n \setminus (\cup_{i=0}^{n-1} B_i)$ pour $n \geq 1$. Par récurrence sur n on montre que $B_n \in T$ pour tout n en remarquant que, pour $n > 1$, $B_n = A_n \cap (\cap_{i=0}^{n-1} B_i^c)$. La construction des B_n assure que $B_n \cap B_m = \emptyset$ si $n \neq m$ et $\cup_{n\in\mathbb{N}} A_n = \cup_{n\in\mathbb{N}} B_n$. Pour vérifier cette dernière propriété, on remarque que $B_n \subset A_n$ donc $\cup_{n\in\mathbb{N}} B_n \subset \cup_{n\in\mathbb{N}} A_n$. Puis, si $x \in A_n$ et $x \notin \cup_{i=0}^{n-1} B_i$, on a alors $x \in A_n \cap (\cap_{i=0}^{n-1} B_i^c) = B_n$. Ceci prouve que $\cup_{n\in\mathbb{N}} A_n \subset \cup_{n\in\mathbb{N}} B_n$ et donc, finalement, $\cup_{n\in\mathbb{N}} A_n = \cup_{n\in\mathbb{N}} B_n$.

   On utilise maintenant la $\sigma-$additivité de m et la monotonie de m (car $B_n \subset A_n$) pour écrire $m(\cup_{n\in\mathbb{N}} A_n) = m(\cup_{n\in\mathbb{N}} B_n) = \sum_{n\in\mathbb{N}} m(B_n) \leq \sum_{n\in\mathbb{N}} m(A_n)$.

3. *Continuité croissante.* Soit $(A_n)_{n\in\mathbb{N}} \subset T$, t.q. $A_n \subset A_{n+1}$, pour tout $n \in \mathbb{N}$. Par monotonie de m, on a $m(A_{n+1}) \geq m(A_n)$, pour tout $n \in \mathbb{N}$, et donc $\lim_{n\to\infty} m(A_n) = \sup_{n\in\mathbb{N}} m(A_n) \in \overline{\mathbb{R}}_+$. On pose $A = \cup_{n\in\mathbb{N}} A_n$ et on définit la suite $(B_n)_{n\in\mathbb{N}}$ par $B_0 = A_0$ et $B_n = A_n \setminus A_{n-1}$ pour tout $n \geq 1$ (noter que $A_{n-1} \subset A_n$). On a $A = \cup_{n\in\mathbb{N}} A_n = \cup_{n\in\mathbb{N}} B_n$, $B_n \in T$ pour tout $n \in \mathbb{N}$ et $B_n \cap B_m = \emptyset$ si $n \neq m$.

La $\sigma-$additivité de m nous donne

$$m(A) = m(\cup_{n\in\mathbb{N}} B_n) = \sum_{n\in\mathbb{N}} m(B_n) = \lim_{n\to\infty} \sum_{p=0}^{n} m(B_p).$$

Puis, comme $A_n = \cup_{p=0}^{n} B_p$, l'additivité de m (qui se déduit de la $\sigma-$additivité) nous donne $\sum_{p=0}^{n} m(B_p) = m(A_n)$ et donc $m(A) = \lim_{n\to\infty} m(A_n)$.

4. *Continuité décroissante.* Soit $(A_n)_{n\in\mathbb{N}} \subset T$, t.q. $A_{n+1} \subset A_n$, pour tout $n \in \mathbb{N}$, et telle qu'il existe $n_0 \in \mathbb{N}$, $m(A_{n_0}) < \infty$.

   Par monotonie, on a $m(A_{n+1}) \le m(A_n)$ pour tout $n \in \mathbb{N}$ et donc $\lim_{n\to\infty} m(A_n) = \inf_{n\in\mathbb{N}} m(A_n) \in \overline{\mathbb{R}}_+$. On a aussi, par monotonie, $m(A) \le m(A_n)$, pour tout $n \in \mathbb{N}$, avec $A = \cap_{n\in\mathbb{N}} A_n$. Comme $m(A_{n_0}) < \infty$, on a aussi $m(A_n) < \infty$ pour tout $n \ge n_0$ et $m(A) < \infty$. On pose $B_n = A_{n_0} \setminus A_n = A_{n_0} \cap A_n^c \in T$, pour tout $n \ge n_0$. La suite $(B_n)_{n\ge n_0}$ est croissante ($B_n \subset B_{n+1}$ pour tout $n \ge n_0$) et $B = \cup_{n\ge 0} B_n = \cup_{n\ge n_0}(A_{n_0} \setminus A_n) = A_{n_0} \setminus \cap_{n\ge n_0} A_n = A_{n_0} \setminus A$.

   La continuité croissante donne

$$m(A_{n_0} \setminus A) = m(B) = \lim_{n\to\infty} m(B_n) = \lim_{n\to\infty} m(A_{n_0} \setminus A_n). \tag{2.9}$$

   Comme $A \subset A_{n_0}$, on a $m(A_{n_0} \setminus A) = m(A_{n_0}) - m(A)$ (car $m(A) \le m(A_{n_0}) < \infty$, on utilise ici la remarque à la fin de la preuve de la monotonie). De même, comme $A_n \subset A_{n_0}$ (pour $n \ge n_0$), on a $m(A_{n_0} \setminus A_n) = m(A_{n_0}) - m(A_n)$ (car $m(A_n) \le m(A_{n_0}) < \infty$). En utilisant une nouvelle fois que $m(A_{n_0}) < \infty$, on déduit de (2.9) que $m(A) = \lim_{n\to\infty} m(A_n)$.

■

Théorème 2.1 (Mesure complétée) *Soit (E, T, m) un espace mesuré, on note $\mathcal{N}_m$ l'ensemble des parties négligeables. On pose $\overline{T} = \{A \cup N,\ A \in T,\ N \in \mathcal{N}_m\}$. Alors $\overline{T}$ est une tribu, et il existe une et une seule mesure, notée $\overline{m}$, sur $\overline{T}$, égale à m sur T. De plus, une partie de E est négligeable pour $(E, \overline{T}, \overline{m})$ si et seulement si elle est négligeable pour (E, T, m). la mesure $\overline{m}$ est complète et l'espace mesuré $(E, \overline{T}, \overline{m})$ s'appelle le complété de (E, T, m). La mesure $\overline{m}$ s'appelle la mesure complétée de la mesure m.*

Démonstration : Cette démonstration est l'objet de l'exercice 2.32. ■

Définition 2.16 (Mesure absolument continue, mesure étrangère)
Soient (E, T) un espace mesurable, et m et μ des mesures (positives) sur T.

1. *On dit que la mesure μ est absolument continue par rapport à la mesure m (et on note $\mu << m$) si pour tout $A \in T$ tel que $m(A) = 0$, alors $\mu(A) = 0$.*

2. *On dit que la mesure μ est étrangère à la mesure m (et note $\mu \perp m$) s'il existe $A \in T$ tel que $m(A) = 0$ et $\mu(A^c) = 0$.*

Proposition 2.4 *Soient (E, T) un espace mesurable, et m et μ des mesures (positives) sur T; on suppose de plus que la mesure μ est σ-finie. Alors il existe une mesure μ_a absolument continue par rapport à m et une mesure μ_e étrangère à m (et à μ_a) t.q. $\mu = \mu_a + \mu_e$.*

Démonstration :
On suppose tout d'abord que μ est une mesure finie. On pose $\alpha = \sup\{\mu(A); A \in T, m(A) = 0\}$. Il existe donc une suite $(A_n)_{n\in\mathbb{N}} \subset T$ t.q. $m(A_n) = 0$, pour tout $n \in \mathbb{N}$, et $\mu(A_n) \to \alpha$, quand $n \to \infty$. On pose alors $C = \cup_{n\in\mathbb{N}} A_n$.

On a $C \in T$, $0 \le m(C) \le \sum_{n\in\mathbb{N}} m(A_n) = 0$ (par σ-sous additivité de m), $\mu(C) \ge \mu(A_n)$ pour tout $n \in \mathbb{N}$ (par monotonie de μ) et donc, en passant à la limite quand $n \to \infty$, $\mu(C) \ge \alpha$. Enfin, la définition de α donne alors $\mu(C) = \alpha$. On a donc trouvé $C \in T$ t.q. $m(C) = 0$ et $\mu(C) = \alpha$.

Pour $A \in T$, on pose $\mu_e(A) = \mu(A \cap C)$ et $\mu_a(A) = \mu(A \cap C^c)$.

Il est clair que μ_e et μ_a sont des mesures sur T et que $\mu = \mu_e + \mu_a$. Comme $\mu_e(C^c) = 0$ et $\mu_a(C) = 0$, les mesures μ_a et μ_e sont étrangères. Comme $m(C) = 0$ et $\mu_e(C^c) = 0$, les mesures μ_e et m sont aussi étrangères. Il reste à montrer que μ_a est absolument continue par rapport à m.

Soit $B \in T$ t.q. $m(B) = 0$. On veut montrer que $\mu_a(B) = 0$, c'est-à-dire que $\mu(B \cap C^c) = 0$. On pose $D = B \cap C^c$ et $F = C \cup D$. Comme $D \cap C = \emptyset$, On a $m(F) = m(C) + m(D) \le m(C) + m(B) = 0$ et $\mu(F) = \mu(C) + \mu(D) = \alpha + \mu(D)$. Comme $m(F) = 0$, la définition de α donne que $\mu(F) \le \alpha$. On a donc $\alpha + \mu(D) \le \alpha$, d'où l'on déduit, comme $\alpha \in \mathbb{R}$ (et c'est ici que l'on utilise le fait que μ est une mesure finie), que $\mu(D) = 0$, c'est-à-dire $\mu_a(B) = 0$. On a bien ainsi montré que μ_a est absolument continue par rapport à m.

On considère maintenant le cas général où μ est σ-finie. Il existe une suite $(E_n)_{n\in\mathbb{N}} \subset T$ t.q. $E = \cup_{n\in\mathbb{N}} E_n$, $\mu(E_n) < \infty$ pour tout $n \in \mathbb{N}$ et $E_n \cup E_m = \emptyset$ si $n \ne m$.

Pour $n \in \mathbb{N}$ et $A \in T$, on pose $\mu^{(n)}(A) = \mu(A \cap E_n)$. $\mu^{(n)}$ est donc une mesure finie sur T. Le raisonnement précédent donne donc l'existence de $\mu_a^{(n)}$ absolument continue par rapport à m et de $\mu_e^{(n)}$ étrangère à m (et à $\mu_a^{(n)}$) t.q. $\mu^{(n)} = \mu_a^{(n)} + \mu_e^{(n)}$. On pose alors, pour $A \in T$:

$$\mu_e(A) = \sum_{n\in\mathbb{N}} \mu_e^{(n)}(A);\ \mu_a(A) = \sum_{n\in\mathbb{N}} \mu_a^{(n)}(A).$$

μ_e et μ_a sont bien des mesures sur T (voir l'exercice 4.2) et il est clair que $\mu = \mu_e + \mu_a$, μ_a absolument continue par rapport à m et μ_e étrangère à m (et à μ_a). ■

Il est parfois utile (surtout en théorie des probabilités, mais une telle question apparaît aussi dans le section 2.5 et dans le chapitre 7) de montrer l'unicité d'une mesure ayant des propriétés données. La proposition suivante donne une méthode pour montrer une telle unicité (d'autres méthodes sont possibles, voir, par exemple, la proposition 5.4 dans le chapitre 5).

Proposition 2.5 *Soit (E, T) un espace mesurable et m, μ deux mesures sur T. On suppose qu'il existe $\mathcal{C} \subset T$ t.q.*

1. *$\mathcal{C}$ engendre T,*
2. *$\mathcal{C}$ est stable par intersection finie (c'est-à-dire $A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}$),*
3. *Il existe $(E_n)_{n\in\mathbb{N}} \subset \mathcal{C}$ t.q. $E_n \cap E_m = \emptyset$ si $n \ne m$, $m(E_n) < \infty$, pour tout $n \in \mathbb{N}$, et $E = \cup_{n\in\mathbb{N}} E_n$,*
4. *$m(A) = \mu(A)$ pour tout $A \in \mathcal{C}$.*

On a alors $m = \mu$ (c'est-à-dire $m(A) = \mu(A)$ pour tout $A \in T$).

Démonstration : Cette démonstration est l'objet de l'exercice 2.21. ■

2.4 mesure signée

Définition 2.17 (Mesure signée) *Soit (E, T) un espace mesurable. On appelle mesure signée (sur T) une application $m : T \to \mathbb{R}$ vérifiant la propriété de σ-additivité, c'est-à-dire t.q. pour toute famille $(A_n)_{n\in\mathbb{N}} \subset T$, t.q. $A_n \cap A_m = \emptyset$, si $n \neq m$,*

$$m(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} m(A_n). \tag{2.10}$$

Noter qu'une mesure signée prend ses valeurs dans $\mathbb{R}$. En prenant $A_n = \emptyset$ pour tout $n \in \mathbb{N}$ dans (2.10), on en déduit que $m(\emptyset) = 0$.

On peut aussi considérer des mesures à valeurs complexes (c'est-à-dire dans $\mathbb{C}$). Dans ce cas, les parties réelles et imaginaires de ces mesures à valeurs complexes sont des mesures signées.

Dans toute la suite du cours, les mesures considérées seront en général positives, c'est-à-dire (cf. définition 2.7) à valeurs dans $\overline{\mathbb{R}}_+$. Lorsque l'on s'intéressera à des mesures prenant leurs valeurs dans $\mathbb{R}$, on précisera qu'il s'agit de "mesures signées". Noter que les mesures signées ne vérifient pas, en général, les propriétés (2.5) et (2.6). Pour avoir un contre exemple, il suffit de considérer une mesure signée m (non nulle) t.q. $-m$ soit une mesure (positive).

Proposition 2.6 (Décomposition d'une mesure signée) *Soient (E, T) un espace mesurable et m une mesure signée sur T. Alors, il existe deux mesures (positives) finies, notées m^+ et m^-, t.q. :*

1. *$m(A) = m^+(A) - m^-(A)$, pour tout $A \in T$.*
2. *Les mesures m^+ et m^- sont étrangères, c'est-à-dire qu'il existe $C \in T$ tel que $m^+(C) = 0$, et $m^-(E \setminus C) = 0$.*

Une conséquence des propriétés ci-dessus est que $m^-(A) = -m(A\cap C)$ et $m^+(A) = m(A\cap C^c)$ pour tout $A \in T$.

De plus, la décompostion de m en différence de deux mesures (positives) finies étrangères est unique. Elle s'appelle "décomposition de Hahn" de m.

Démonstration :
La démonstration d'existence de m^+ et m^- est décomposée en trois étapes. Dans la première étape, on va montrer que, si $A \in T$, il existe $\tilde{A} \in T$ t.q. $\tilde{A} \subset A$, $m(\tilde{A}) \geq m(A)$ et :

$$B \in T,\ B \subset \tilde{A} \ \Rightarrow\ m(B) \geq 0.$$

Cette première étape nous permettra, dans l'étape 2, de montrer l'existence de $C \in T$ t.q. $m(C) = \sup\{m(A), A \in T\}$ (ceci montre, en particulier que $\sup\{m(A), A \in T\} < \infty$).

Enfin, dans l'étape 3, on pose $m^+(A) = m(A \cap C)$ et $m^-(A) = -m(A \cap C^c)$ (pour tout $A \in T$) et on remarque que m^+ et m^- sont des mesures finies, étrangères et t.q. $m = m^+ - m^-$.

Etape 1. Soit $A \in T$, on montre, dans cette étape, qu'il existe $\tilde{A} \in T$ t.q. $\tilde{A} \subset A$, $m(\tilde{A}) \geq m(A)$ et :

$$B \in T,\ B \subset \tilde{A} \ \Rightarrow\ m(B) \geq 0. \tag{2.11}$$

On commence par montrer, par récurrence sur n, l'existence d'une suite $(B_n)_{n\in\mathbb{N}}$ déléments de T t.q. :

1. $B_0 = A$,

2. $B_{n+1} \subset B_n$, pour tout $n \in \mathbb{N}$,

3. $m(B_n \setminus B_{n+1}) \le \beta_n = \max\{\frac{\alpha_n}{2}, -1\}$ où $\alpha_n = \inf\{m(C), C \in T, C \subset B_n\}$.

On prend $B_0 = A$. Soit maintenant $n \in \mathbb{N}$, on suppose B_p connu pour $p \le n$. On a $\alpha_n = \inf\{m(C), C \subset B_n\} \le 0$ (car $\emptyset \subset B_n$). Si $\alpha_n = -\infty$, il existe $C_n \in T$ t.q. $C_n \subset B_n$ et $m(C_n) \le \beta_n = -1$. Si $-\infty < \alpha_n < 0$, on a $\beta_n > \alpha_n$, il existe donc $C_n \in T$ t.q. $C_n \subset B_n$ et $m(C_n) \le \beta_n$. Si $\alpha_n = 0$, on prend $C_n = \emptyset$. Enfin, on prend $B_{n+1} = B_n \setminus C_n$ et on obtient bien les propriétés désirées en remarquant que $C_n = B_n \setminus B_{n+1}$.

La suite $(B_n)_{n\in\mathbb{N}}$ est décroissante (c'est-à-dire $B_{n+1} \subset B_n$ pour tout $n \in \mathbb{N}$). Pour $m > n$, on a donc $C_m \subset B_m \subset B_{n+1}$ et donc $C_m \cap C_n = \emptyset$ (car $B_{n+1} = B_n \setminus C_n$). Par $\sigma-$additivité de m, on en déduit $m(\cup_{n\in\mathbb{N}} C_n) = \sum_{n\in\mathbb{N}} m(C_n)$. Comme $m(\cup_{n\in\mathbb{N}} C_n) \in \mathbb{R}$, la série de terme général $m(C_n)$ est convergente. On a donc $m(C_n) \to 0$ quand $n \to \infty$ et donc $\beta_n \to 0$ quand $n \to \infty$ (car $m(C_n) \le \beta_n \le 0$) et, finalement, $\alpha_n \to 0$ quand $n \to \infty$.

On pose maintenant $\tilde{A} = A \setminus \cup_{n\in\mathbb{N}} C_n = \cap_{n\in\mathbb{N}} B_n$. On a, bien sûr, $\tilde{A} \in T$ et $\tilde{A} \subset A$. On montre maintenant que $\tilde{A}$ vérifie (2.11). Soit $C \in T$, $C \subset \tilde{A}$. On a, pour tout $n \in \mathbb{N}$, $C \subset B_n$ et donc $m(C) \ge \alpha_n$. Quand $n \to \infty$, on en déduit que $m(C) \ge 0$. ce qui donne bien (2.11).

Il reste à montrer que $m(\tilde{A}) \ge m(A)$. Comme $A = \tilde{A} \cup (\cup_{n\in\mathbb{N}} C_n)$ (et que cette union est "disjointe"), la $\sigma-$additivité de m donne que $m(A) = m(\tilde{A}) + \sum_{n\in\mathbb{N}} m(C_n) \le m(\tilde{A})$. Ce qui termine la première étape.

Etape 2. On pose $\alpha = \sup\{m(A), A \in T\}$ et on montre, dans cette étape, qu'il existe $C \in T$ t.q. $m(C) = \alpha$.

Par définition d'une borne supérieure, il existe une suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de T t.q. $m(A_n) \to \alpha$ quand $n \to \infty$. Grâce à l'étape 1, on peut supposer (quitte à remplacer A_n par $\tilde{A_n}$ construit comme dans l'étape 1) que A_n vérifie (2.11), c'est-à-dire que pour tout $n \in \mathbb{N}$:

$$B \in T,\ B \subset A_n \ \Rightarrow\ m(B) \ge 0. \tag{2.12}$$

On pose $C = \cup_{n\in\mathbb{N}} A_n$. On commence par montrer que $m(C) \ge m(A_m)$, pour tout $m \in \mathbb{N}$.

Soit $m \in \mathbb{N}$. On peut écrire C comme une union "disjointe" :

$$C = A_m \cup (\cup_{n \ne m} C_{n,m}),$$

avec $C_{n,m} \in T$ et $C_{n,m} \subset A_n$ pour tout $m \ne n$. En effet, il suffit pour cela de constuire par récurrence (sur n) la suite des $C_{n,m}$ en prenant pour $C_{n,m}$ l'intersection de C avec A_n à laquelle on retranche A_m et les $C_{n,m}$ précédemment construits.

Par $\sigma-$additivité de m, on a $m(C) = m(A_m) + \sum_{n\ne m} m(C_{n,m})$. puis, comme $C_{n,m} \subset A_n$, on a, par (2.12), $m(C_{n,m}) \ge 0$. On en déduit $m(C) \ge m(A_m)$.

En faisant tendre m vers ∞, on a alors $m(C) \ge \alpha$ et donc, finalement $m(C) = \alpha$.

Etape 3. Construction de m^+ et m^-.

Pour constuire m^+ et m^-, on utilise un élément C de T t.q. $m(C) = \alpha = \sup\{m(A), A \in T\}$ (l'existence de C a été montré à l'étape 2). Pour $A \in T$, on pose :

$$m^+(A) = m(A \cap C),\ m^-(A) = -m(A \cap C^c).$$

On a $m^+(\emptyset) = m^-(\emptyset) = 0$ (car $m(\emptyset) = 0$) et les applications m^+ et m^- sont des applications $\sigma-$additives de T dans $\mathbb{R}$ (car m est $\sigma-$additive). Pour montrer que m^+ et m^- sont des mesures finies, il suffit de montrer qu'elles prennent leurs valeurs dans $\mathbb{R}_+$, ce que l'on montre maintenant.

Soit $A \in T$, on a, par additivité de m et grâce à la définition de α, $\alpha = m(C) = m(A\cap C) + m(A^c \cap C) \leq m(A \cap C) + \alpha$. On en déduit $m(A \cap C) \geq 0$. ce qui prouve bien que $m^+(A) \in \mathbb{R}_+$. On a aussi, encore une fois par additivité de m et grâce à la définition de α, $\alpha \geq m(C) + m(A \cap C^c) = \alpha + m(A \cap C^c)$. On en déduit $m(A \cap C^c) \leq 0$ et donc $m^-(A) \in \mathbb{R}_+$.

Les applications m^+ et m^- sont des mesures finies (noter que $m^+(E) = m(E\cap C) < \infty$ et $m^-(E) = m(E \cap C^c) < \infty$). Elles sont étrangères car $m^+(C^c) = m(C^c \cap C) = m(\emptyset) = 0$ et $m^-(C) = -m(C \cap C^c) = 0$. Enfin, pour tout $A \in T$, on a, par $\sigma-$additivité de m :

$$m(A) = m(A \cap C) + m(A \cap C^c) = m^+(A) - m^-(A).$$

Ceci termine la démonstration de l'existence de m^+ et m^-.

Pour montrer l'unicité de cette décomposition de m, on suppose que μ et ν sont deux mesures finies étrangères t.q. $m = \mu - \nu$. Comme elle sont étrangères, il existe $D \in T$ t.q. $\mu(D^c) = \nu(D) = 0$. On montre alors que, pour tout $A \in T$, on a nécessairement :

$$\mu(A) = \sup\{m(B);\ B \in T,\ B \subset A\}. \tag{2.13}$$

En effet, si $A \in T$ et $B \in T$, $B \subset A$, on a $m(B) = \mu(B) - \nu(B) \leq \mu(B) \leq \mu(A)$ (par positivité de ν et monotonie de μ). Puis, en prenant $B = A \cap D$, on a $m(B) = m(A \cap D) = \mu(A \cap D) - \nu(A \cap D) = \mu(A \cap D) = \mu(A) - \mu(A \cap D^c) = \mu(A)$. Ceci prouve bien que (2.13) est vraie (et prouve que le sup est atteint pour $B = A \cap D$). L'égalité (2.13) donne donc de manière unique μ en fonction de m. L'unicité de ν découle alors du fait que $\nu = \mu - m$. ■

Remarque 2.7 Une conséquence de la proposition 2.6 est que la série $\sum_{n\in\mathbb{N}} m(A_n)$ apparaissant dans (2.10) est absolument convergente car (pour toute famille $(A_n)_{n\in\mathbb{N}} \subset T$ t.q. $A_n \cap A_m = \emptyset$, si $n \neq m$) on a $\sum_{p=0}^n |m(A_p)| \leq \sum_{p=0}^n m^+(A_p) + \sum_{p=0}^n m^-(A_p) \leq m^+(E) + m^-(E) < \infty$.

En fait, la définition 2.17 donne directement que la série $\sum_{n\in\mathbb{N}} m(A_n)$ apparaissant dans (2.10) est commutativement convergente (c'est-à-dire qu'elle est convergente, dans $\mathbb{R}$, quel que soit l'ordre dans lequel on prend les termes de la série et la somme de la série ne dépend pas de l'ordre dans lequel les termes ont été pris). Elle est donc absolument convergente (voir l'exercice 2.33). Nous verrons plus loin que cette équivalence entre les séries absolument convergentes et les séries commutativement convergentes est fausse pour des séries à valeurs dans un espace de Banach de dimension infinie.

2.5 La mesure de Lebesgue sur la tribu des boréliens

Il serait bien agréable, pour la suite du cours, de montrer l'existence d'une application λ, définie sur tout $\mathcal{P}(\mathbb{R})$ et à valeurs dans $\overline{\mathbb{R}}_+$, t.q. l'image par λ d'un intervalle de $\mathbb{R}$ soit la longueur de cet intervalle, et qui vérifie les propriétés (1.9) et (1.10). Malheureusement, on peut montrer qu'une telle application n'existe pas (voir les exercices 2.28 et 2.27). Le théorème suivant donne l'existence d'une telle application définie seulement sur la tribu des boréliens de $\mathbb{R}$, notée $\mathcal{B}(\mathbb{R})$ (l'exercice 2.28 donne alors que $\mathcal{B}(\mathbb{R}) \neq \mathcal{P}(\mathbb{R})$). Cette application s'appelle la mesure de Lebesgue.

Théorème 2.2 (Carathéodory) *Il existe une et une seule mesure sur $\mathcal{B}(\mathbb{R})$, notée λ et appelée mesure de Lebesgue sur les boréliens, t.q. $\lambda(]\alpha,\beta[) = \beta - \alpha$, pour tout $(\alpha,\beta) \in \mathbb{R}^2$ t.q. $-\infty < \alpha < \beta < +\infty$.*

Il y a plusieurs démonstrations possibles de ce théorème. Pour la partie "existence" de ce théorème, nous donnons dans cette section une démonstration due à Carathéodory. Soit $A \subset \mathbb{R}$. On définit $\lambda^\star(A)$ par :

$$\lambda^\star(A) = \inf_{(A_i)_{i\in\mathbb{N}} \in E_A} \sum_{i=1}^{n} \ell(A_i),$$

où E_A est l'ensemble des familles dénombrables d'intervalles ouverts dont l'union contient A, et $\ell(A_i)$ représente la longueur de l'intervalle A_i. On peut montrer (voir l'exercice 2.27) que l'application $\lambda^\star$ ainsi définie de $\mathcal{P}(\mathbb{R})$ dans $\overline{\mathbb{R}}_+$ n'est pas σ- additive (ce n'est donc pas une mesure).

On montre par contre dans cette section que la restriction de $\lambda^\star$ à $\mathcal{B}(\mathbb{R})$ est une mesure, qu'on note λ, mesure de Lebesgue. L'existence de la mesure de Lebesgue peut aussi être démontrée en utilisant un théorème plus général (de F. Riesz) que nous verrons dans un chapitre ultérieur.

Après la définition de $\lambda^\star$ et la démonstration de propriétés de $\lambda^\star$, on donne la démonstration de la partie "existence" du théorème de Carathéodory (voir page 32). La partie "unicité" du théorème de Carathéodory (voir page 35) peut être démontrée en utilisant le théorème de "régularité" sur les mesures sur $\mathcal{B}(\mathbb{R})$ (Théorème 2.3, très utille dans la suite du cours) et d'un lemme classique sur les ouverts de $\mathbb{R}$ (lemme 2.4). Cette partie "unicité" peut aussi être démontrée, plus directement, en utilisant la proposition 2.5.

Définition 2.18 (Définition de $\lambda^\star$) *Soit $A \in \mathcal{P}(\mathbb{R})$. On pose $\lambda^\star(A) = \inf\{\sum_{n\in\mathbb{N}} \ell(I_n);\ (I_n)_{n\in\mathbb{N}} \in E_A\}$, avec $E_A = \{(I_n)_{n\in\mathbb{N}};\ I_n =]a_n, b_n[,\ -\infty < a_n \le b_n < \infty,\ \forall n \in \mathbb{N},\ A \subset \cup_{n\in\mathbb{N}} I_n\}$ et $\ell(I) = b - a$ si $I =]a,b[,\ -\infty < a \le b < \infty$.*

Proposition 2.7 (Propriétés de $\lambda^\star$) *L'application $\lambda^\star : \mathcal{P}(\mathbb{R}) \to \overline{\mathbb{R}}_+$ (définie dans la définition 2.18) vérifie les propriétés suivantes :*

1. *$\lambda^\star(\emptyset) = 0$,*
2. *(Monotonie) $\lambda^\star(A) \le \lambda^\star(B)$, pour tout $A, B \in \mathcal{P}(\mathbb{R})$, $A \subset B$,*
3. *($\sigma-$sous additivité) Soit $(A_n)_{n\in\mathbb{N}} \subset \mathcal{P}(\mathbb{R})$ et $A = \cup_{n\in\mathbb{N}} A_n$, alors $\lambda^\star(A) \le \sum_{n\in\mathbb{N}} \lambda^\star(A_n)$,*
4. *$\lambda^\star(]a,b[) = b - a$ pour tout $(\alpha,\beta) \in \mathbb{R}^2$ t.q. $-\infty < \alpha < \beta < +\infty$.*

Démonstration :
On remarque tout d'abord que $\lambda^\star(A) \in \overline{\mathbb{R}}_+$ pour tout $A \in \mathcal{P}(\mathbb{R})$ (car $\lambda^\star(A)$ est la borne inférieure d'une partie de $\overline{\mathbb{R}}_+$).

Propriété 1. Pour montrer que $\lambda^\star(\emptyset) = 0$, il suffit de remarquer que $(I_n)_{n\in\mathbb{N}} \in E_\emptyset$ avec $I_n = \emptyset$ pour tout $n \in \mathbb{N}$, et donc $0 \le \lambda^\star(\emptyset) \le \sum_{n\in\mathbb{N}} \ell(I_n) = 0$.

Propriété 2. Soit $A, B \in \mathcal{P}(\mathbb{R})$ t.q. $A \subset B$. On a $E_B \subset E_A$ et donc $\lambda^\star(A) \le \lambda^\star(B)$.

Propriété 3. Soit $(A_n)_{n\in\mathbb{N}} \subset \mathcal{P}(\mathbb{R})$ et $A = \cup_{n\in\mathbb{N}} A_n$. Il suffit de considérer le cas où $\lambda^\star(A_n) < \infty$ pour tout $n \in \mathbb{N}$ (sinon, l'inégalité est immédiate).

Soit $\varepsilon > 0$. Pour tout $n \in \mathbb{N}$, il existe $(I_{n,m})_{m\in\mathbb{N}} \in E_{A_n}$ t.q. $\sum_{m\in\mathbb{N}} \ell(I_{n,m}) \le \lambda^\star(A_n) + \varepsilon/(2^n)$.

On remarque alors que $(I_{n,m})_{(n,m)\in\mathbb{N}^2}$ est un recouvrement de A par des intervalles ouverts et donc que :

$$\lambda^\star(A) \le \sum_{(n,m)\in\mathbb{N}^2} \ell(I_{n,m}).$$

Noter que $\sum_{(n,m)\in\mathbb{N}^2} \ell(I_{n,m}) = \sum_{n\in\mathbb{N}} \ell(I_{\varphi(n)})$, où φ est une bijection de $\mathbb{N}$ dans $\mathbb{N}^2$ (cette somme ne dépend pas de la bijection choisie, voir le lemme 2.2 page 21).Avec le lemme 2.3 ci dessous, on en déduit :

$$\lambda^\star(A) \le \sum_{n\in\mathbb{N}} (\sum_{m\in\mathbb{N}} \ell(I_{n,m})) \le \sum_{n\in\mathbb{N}} \lambda^\star(A_n) + 2\varepsilon,$$

ce qui donne bien, quand $\varepsilon \to 0$, $\lambda^\star(A) \le \sum_{n\in\mathbb{N}} \lambda^\star(A_n)$.

Propriété 4. Pour montrer la quatrième propriété. On commence par montrer :

$$\lambda^\star([a,b]) = b-a,\ \forall a,b \in \mathbb{R}, a<b. \tag{2.14}$$

Soit donc a, $b \in \mathbb{R}$, $a<b$.

Comme $[a,b] \subset]a-\varepsilon, b+\varepsilon[$, pour tout $\varepsilon > 0$, on a $\lambda^\star([a,b]) \le b-a+2\varepsilon$. On en déduit $\lambda^\star([a,b]) \le b-a$.

Pour démontrer l'inégalité inverse, soit $(I_n)_{n\in\mathbb{N}} \in E_{[a,b]}$. Par compacité de $[a,b]$, il existe $n \in \mathbb{N}$ t.q. $[a,b] \subset \cup_{p=0}^n I_p$. On peut alors construire (par récurrence) i_0, i_1, $\ldots$, $i_q \in \{0,\ldots,n\}$ t.q. $a_{i_0} < a$, $a_{i_{p+1}} < b_{i_p}$ pour tout $p \in \{0,\ldots,q-1\}$, $b < b_{i_q}$. On en déduit que $b-a < \sum_{p=0}^q b_{i_p} - a_{i_p} \le \sum_{n\in\mathbb{N}} \ell(I_n)$ et donc $b-a \le \lambda^\star([a,b])$. Ceci donne bien (2.14).

En remarquant que $[a+\varepsilon, b-\varepsilon] \subset]a,b[\subset [a,b]$ pour tout $a,b \in \mathbb{R}$, $a<b$, et $0<\varepsilon<(b-a)/2$, la monotonie de $\lambda^\star$ donne (avec (2.14)) que $\lambda^\star(]a,b[) = b-a$ pour tout $a,b \in \mathbb{R}, a<b$. La monotonie de $\lambda^\star$ donne alors aussi que $\lambda^\star([a,b[) = \lambda^\star(]a,b]) = \lambda^\star(]a,b[) = b-a$ pour tout $a,b\in\mathbb{R}, a<b$, et enfin que $\lambda^\star(]-\infty,a]) = \lambda^\star(]-\infty,a[) = \lambda^\star(]a,\infty]) = \lambda^\star([a,\infty]) = \infty$ pour tout $a \in \mathbb{R}$. ∎

Lemme 2.3 (Double série à termes positifs) *Soit $(a_{n,m})_{(n,m)\in\mathbb{N}^2} \subset \mathbb{R}_+$. Alors on a :*

$$\sum_{(n,m)\in\mathbb{N}^2} a_{n,m} = \sum_{n\in\mathbb{N}} (\sum_{m\in\mathbb{N}} a_{n,m}).$$

Démonstration : On pose $A = \sum_{(n,m)\in\mathbb{N}^2} a_{n,m}$ et $B = \sum_{n\in\mathbb{N}} (\sum_{m\in\mathbb{N}} a_{n,m})$. Soit φ une bijection de $\mathbb{N}$ dans $\mathbb{N}^2$. On rappelle que $\sum_{(n,m)\in\mathbb{N}^2} a_{n,m} = \sum_{p\in\mathbb{N}} a_{\varphi(p)}$.

Pour tout i, $j \in \mathbb{N}$, il existe $n\in\mathbb{N}$ t.q. $\{0,\ldots,i\}\times\{0,\ldots j\} \subset \{\varphi(0),\ldots,\varphi(n)\}$. Comme $a_{n,m} \ge 0$ pour tout (n,m), on en déduit que $A \ge \sum_{p=0}^n a_{\varphi(p)} \ge \sum_{n=0}^i (\sum_{m=0}^j a_{n,m})$ et donc, en faisant tendre j puis i vers ∞, que $A \ge B$. Un raisonnement similaire donne que $B \ge A$ et donc $A = B$. ∎

On introduit maintenant la tribu de Lebesgue, sur laquelle on montrera que $\lambda^\star$ est une mesure.

Définition 2.19 (Tribu de Lebesgue) *On pose $\mathcal{L} = \{E \in \mathcal{P}(\mathbb{R})$ t.q. $\lambda^\star(A) = \lambda^\star(A\cap E) + \lambda^\star(A \cap E^c)$ pour tout $A \in \mathcal{P}(\mathbb{R})\}$. On rappelle que $\lambda^\star$ est définie dans la définition 2.18 (et que $E^c = \mathbb{R}\setminus E$). Cet ensemble de parties de $\mathbb{R}$ noté $\mathcal{L}$ s'appelle "tribu de Lebesgue" (on montre dans la proposition 2.8 que $\mathcal{L}$ est bien une tribu).*

Remarque 2.8 On peut avoir une première idée de l'intérêt de la définition 2.19 en remarquant qu'elle donne immédiatement l'additivité de $\lambda^\star$ sur $\mathcal{L}$. En effet, soit $E_1, E_2 \subset \mathbb{R}$ t.q. $E_1 \cap E_2 = \emptyset$ et soit $A \subset \mathbb{R}$. On suppose que $E_1 \in \mathcal{L}$ et on utilise la définition de $\mathcal{L}$ avec $A \cap (E_1 \cup E_2)$, on obtient $\lambda^\star(A \cap (E_1 \cup E_2)) = \lambda^\star(A \cap (E_1 \cup E_2) \cap E_1) + \lambda^\star(A \cap (E_1 \cup E_2) \cap E_1^c) = \lambda^\star(A \cap E_1) + \lambda^\star(A \cap E_2)$ (car $E_1 \cap E_2 = \emptyset$).

Par récurrence sur n, on a donc aussi $\lambda^\star(A \cap (\cup_{i=1}^n E_i)) = \sum_{i=1}^n \lambda^\star(A \cap E_i)$, dès que $E_1, \ldots, E_{n-1} \in \mathcal{L}$, $A, E_n \subset \mathbb{R}$ et $E_i \cap E_j = \emptyset$ si $i \neq j$, $i, j \in \{1, \ldots, n\}$.

En particulier, en prenant $A = \mathbb{R}$, on obtient l'additivité de $\lambda^\star$ sur $\mathcal{L}$, c'est-à-dire

$$\lambda^\star(\cup_{i=1}^n E_i) = \sum_{i=1}^n \lambda^\star(E_i),$$

si $E_1, \ldots, E_{n-1} \in \mathcal{L}$ et $E_i \cap E_j = \emptyset$ si $i \neq j$, $i, j \in \{1, \ldots, n\}$.

Remarque 2.9 Pour tout E, $A \in \mathcal{P}(\mathbb{R})$, on a, par $\sigma-$sous additivité de $\lambda^\star$, $\lambda^\star(A) \leq \lambda^\star(A \cap E) + \lambda^\star(A \cap E^c)$. Pour montrer que $E \in \mathcal{L}$ (définie dans la définition 2.19), il suffit donc de montrer que $\lambda^\star(A) \geq \lambda^\star(A \cap E) + \lambda^\star(A \cap E^c)$, pour tout $A \in \mathcal{P}(\mathbb{R})$.

Proposition 2.8 (Propriétés de $\mathcal{L}$) *$\mathcal{L}$ est une tribu sur $\mathbb{R}$ et $\lambda^\star_{|\mathcal{L}}$ est une mesure. $\mathcal{L}$ et $\lambda^\star$ sont définies dans les définitions 2.18 et 2.19.*

Démonstration :
Il est immédiat que $\emptyset \in \mathcal{L}$ et que $\mathcal{L}$ est stable par "passage au complémentaire". On sait aussi que $\lambda^\star(\emptyset) = 0$. Il reste donc à démontrer que $\mathcal{L}$ est stable par union dénombrable et que la restriction de $\lambda^\star$ à $\mathcal{L}$ est une mesure. Ceci se fait en deux étapes décrites ci-après.

Etape 1. On montre, dans cette étape, que $\mathcal{L}$ est stable par union finie et que, si $n \geq 2$ et $(E_i)_{i=1,\ldots,n} \subset \mathcal{L}$ est t.q. $E_i \cap E_j = \emptyset$ si $i \neq j$, alors on a :

$$\lambda^\star(A \cap (\cup_{i=1}^n E_i)) = \sum_{i=1}^n \lambda^\star(A \cap E_i), \ \forall A \in \mathcal{P}(\mathbb{R}). \tag{2.15}$$

(Cette dernière propriété donne l'additivité de $\lambda^\star$ sur $\mathcal{L}$ en prenant $A = \mathbb{R}$, cette propriété d'additivité a déjà été signalée dans la remarque 2.8.)

Par une récurrence facile, il suffit de montrer que $E_1 \cup E_2 \in \mathcal{L}$ si E_1, $E_2 \in \mathcal{L}$ et de montrer la propriété (2.15) pour $n = 2$. Soit donc E_1, $E_2 \in \mathcal{L}$. On pose $E = E_1 \cup E_2$. Pour montrer que $E \in \mathcal{L}$, il suffit de montrer (voir la remarque 2.9) que $\lambda^\star(A) \geq \lambda^\star(A \cap E) + \lambda^\star(A \cap E^c)$, pour tout $A \in \mathcal{P}(\mathbb{R})$. Soit $A \in \mathcal{P}(\mathbb{R})$. Par $\sigma-$sous additivité de $\lambda^\star$ on a

$$\lambda^\star(A \cap (E_1 \cup E_2)) = \lambda^\star((A \cap E_1) \cup (A \cap E_1^c \cap E_2)) \leq \lambda^\star(A \cap E_1) + \lambda^\star(A \cap E_1^c \cap E_2),$$

et donc

$$\lambda^\star(A \cap (E_1 \cup E_2)) + \lambda^\star(A \cap (E_1 \cup E_2)^c) \leq \lambda^\star(A \cap E_1) + \lambda^\star(A \cap E_1^c \cap E_2) + \lambda^\star(A \cap E_1^c \cap E_2^c).$$

Comme $E_2 \in \mathcal{L}$, on a $\lambda^\star(A \cap E_1^c) = \lambda^\star(A \cap E_1^c \cap E_2) + \lambda^\star(A \cap E_1^c \cap E_2^c)$. Puis, comme $E_1 \in \mathcal{L}$, on a $\lambda^\star(A) = \lambda^\star(A \cap E_1) + \lambda^\star(A \cap E_1^c)$. On en déduit

$$\lambda^\star(A \cap (E_1 \cup E_2)) + \lambda^\star(A \cap (E_1 \cup E_2)^c) \leq \lambda^\star(A).$$

Ce qui prouve que $E \in \mathcal{L}$.

Pour montrer (2.15) avec $n = 2$ si E_1, $E_2 \in \mathcal{L}$ avec $E_1 \cap E_2 = \emptyset$, il suffit de remarquer que (pour tout $A \in \mathcal{P}(\mathbb{R})$) $\lambda^\star(A \cap (E_1 \cup E_2)) = \lambda^\star((A \cap E_1) \cup (A \cap E_2)) = \lambda^\star([(A \cap E_1) \cup (A \cap E_2)] \cap E_1) + \lambda^\star([(A \cap E_1) \cup (A \cap E_2)] \cap E_1^c) = \lambda^\star(A \cap E_1) + \lambda^\star(A \cap E_2)$. (On a utilisé le fait que $E_1 \in \mathcal{L}$.) Ceci termine l'étape 1.

Une conséquence de cette étape (et du fait que $\mathcal{L}$ est stable par passage au complémentaire) est que $\mathcal{L}$ est stable par intersection finie.

Etape 2. On montre, dans cette étape, que $\mathcal{L}$ est stable par union dénombrable et la restriction de $\lambda^\star$ à $\mathcal{L}$ est une mesure (ce qui termine la démonstration de la proposition 2.8).

Soit $(E_n)_{n\in\mathbb{N}} \subset \mathcal{L}$ et $E = \cup_{n\in\mathbb{N}} E_n$. On veut montrer que $E \in \mathcal{L}$. On commence par remarquer que $E = \cup_{n\in\mathbb{N}} F_n$ avec $F_0 = E_0$ et, par récurrence, pour $n \geq 1$, $F_n = E_n \setminus \cup_{p=0}^{n-1} F_p$. L'étape 1 nous donne que $(F_n)_{n\in\mathbb{N}} \subset \mathcal{L}$ et, comme $F_n \cap F_m = \emptyset$ si $n \neq m$, on peut utiliser (2.15). Pour tout $A \in \mathcal{P}(\mathbb{R})$, on a donc :

$$\lambda^\star(A) = \lambda^\star(A \cap (\cup_{p=0}^n F_p)) + \lambda^\star(A \cap (\cup_{p=0}^n F_p)^c) = \sum_{p=0}^n \lambda^\star(A \cap F_p) + \lambda^\star(A \cap (\cup_{p=0}^n F_p)^c). \tag{2.16}$$

En utilisant le fait que $E^c \subset (\cup_{p=0}^n F_p)^c$ et la monotonie de $\lambda^\star$, on a $\lambda^\star(A \cap (\cup_{p=0}^n F_p)^c) \geq \lambda^\star(A \cap E^c)$. En faisant tendre n vers ∞ dans (2.16) et en utilisant la $\sigma-$sous additivité de $\lambda^\star$, on en déduit alors que $\lambda^\star(A) \geq \lambda^\star(A \cap E) + \lambda^\star(A \cap E^c)$. Ce qui prouve que $E \in \mathcal{L}$ (voir remarque 2.9) et donc que $\mathcal{L}$ est une tribu.

Il reste à montrer que $\lambda^\star$ est une mesure sur $\mathcal{L}$. Soit $(E_n)_{n\in\mathbb{N}} \subset \mathcal{L}$ t.q. $E_i \cap E_j = \emptyset$ si $i \neq j$ et $E = \cup_{n\in\mathbb{N}} E_n$. Par monotonie de $\lambda^\star$ on a, pour tout $n \in \mathbb{N}$, $\lambda^\star(\cup_{p=0}^n E_p) \leq \lambda^\star(E)$ et donc, en utilisant l'additivité de $\lambda^\star$ sur $\mathcal{L}$ (démontrée à l'étape 1, voir (2.15) avec $A = E$), $\sum_{p=0}^n \lambda^\star(E_p) \leq \lambda^\star(E)$. Ce qui donne, passant à limite quand $n \to \infty$, $\sum_{p=0}^\infty \lambda^\star(E_p) \leq \lambda^\star(E)$. D'autre part, $\lambda^\star(E) \leq \sum_{p=0}^\infty \lambda^\star(E_p)$, par $\sigma-$sous additivité de $\lambda^\star$. On a donc $\lambda^\star(E) = \sum_{p=0}^\infty \lambda^\star(E_p)$. Ce qui prouve que $\lambda^\star_{|\mathcal{L}}$ est une mesure.

■

Démonstration de la partie "existence" du théorème 2.2 :

Pour montrer la partie "existence" du théorème 2.2, il suffit, grâce aux propositions 2.7 et 2.8, de montrer que $\mathcal{L}$ (définie dans la définition 2.19) contient $\mathcal{B}(\mathbb{R})$. Pour cela, il suffit de montrer que $]a,\infty[\subset \mathcal{L}$ pour tout $a \in \mathbb{R}$ (car $\{]a,\infty[,\ a \in \mathbb{R}\}$ engendre $\mathcal{B}(\mathbb{R})$). Soit donc $a \in \mathbb{R}$ et $E =]a,\infty[$. Pour tout $A \in \mathcal{P}(\mathbb{R})$, on veut montrer que $\lambda^\star(A) \geq \lambda^\star(A \cap E) + \lambda^\star(A \cap E^c)$. On peut supposer que $\lambda^\star(A) < \infty$ (sinon l'inégalité est immédiate).

Soit $\varepsilon > 0$. Par la définition de $\lambda^\star(A)$, il existe $(I_n)_{n\in\mathbb{N}} \in E_A$ t.q. $\lambda^\star(A) \geq \sum_{n\in\mathbb{N}} \ell(I_n) - \varepsilon$. Comme $A \cap E \subset (\cup_{n\in\mathbb{N}}(I_n \cap E))$ et $A \cap E^c \subset (\cup_{n\in\mathbb{N}}(I_n \cap E^c))$, la $\sigma-$sous additivité de $\lambda^\star$ donne

$$\lambda^\star(A \cap E) \leq \sum_{n\in\mathbb{N}} \lambda^\star(I_n \cap E) \quad \text{et} \quad \lambda^\star(A \cap E^c) \leq \sum_{n\in\mathbb{N}} \lambda^\star(I_n \cap E^c).$$

Comme $I_n \cap E$ et $I_n \cap E^c$ sont des intervalles, la fin de la démonstration de la proposition 2.7 donne $\lambda^\star(I_n \cap E) = \ell(I_n \cap E)$ et $\lambda^\star(I_n \cap E^c) = \ell(I_n \cap E^c)$. On en déduit $\lambda^\star(A \cap E) + \lambda^\star(A \cap E^c) \leq \sum_{n\in\mathbb{N}}(\ell(I_n \cap E) + \ell(I_n \cap E^c)) = \sum_{n\in\mathbb{N}} \ell(I_n)$ (car $\ell(I_n \cap E) + \ell(I_n \cap E^c) = \ell(I_n)$) et donc $\lambda^\star(A \cap E) + \lambda^\star(A \cap E^c) \leq \lambda^\star(A) + \varepsilon$. Quand $\varepsilon \to 0$ on trouve l'inégalité recherchée. On a bien montré que $E \in \mathcal{L}$. ■

On va maintenant démontrer un théorème important dont on peut déduire, en particulier, la partie "unicité" du théorème 2.2.

Théorème 2.3 (Régularité d'une mesure sur $\mathcal{B}(\mathbb{R})$, finie sur les compacts)

Soit m une mesure sur $\mathcal{B}(\mathbb{R})$. On suppose que m est finie sur les compacts, c'est à dire que $m(K) < \infty$ pour tout compact K de $\mathbb{R}$ (noter qu'un compact est nécessairement dans $\mathcal{B}(\mathbb{R})$). Alors, pour tout $A \in \mathcal{B}(\mathbb{R})$ et tout $\varepsilon > 0$, il existe un ouvert O et un fermé F t.q. $F \subset A \subset O$ et $m(O \setminus F) \leq \varepsilon$. En particulier, on a donc, pour tout $A \in \mathcal{B}(\mathbb{R})$, $m(A) = \inf\{m(O),\ O$ ouvert contenant $A\}$ et $m(A) = \sup\{m(K),\ K$ compact inclus dans $A\}$.

DÉMONSTRATION :
On appelle T l'ensemble des $A \in \mathcal{B}(\mathbb{R})$ t.q. pour tout $\varepsilon > 0$, il existe O ouvert et F fermé vérifiant $F \subset A \subset O$ et $m(O \setminus F) \leq \varepsilon$. On va montrer que T est une tribu contenant $\mathcal{C} = \{]a,b[,\ -\infty < a < b < \infty\}$. Comme $\mathcal{C}$ engendre $\mathcal{B}(\mathbb{R})$, ceci donnera $T = \mathcal{B}(\mathbb{R})$.

On montre tout d'abord que $\mathcal{C} \subset T$. Soit $-\infty < a < b < \infty$ et $A =]a,b[$.

Pour tout $n \geq n_0$ avec n_0 t.q. $(2/n_0) < b - a$ on a :

$$[a + (1/n), b - (1/n)] \subset A \subset]a,b[.$$

Pour $n \geq n_0$, on pose $B_n =]a, a + (1/n)[\cup]b - (1/n), b[)$. La suite $(B_n)_{n \geq n_0}$ est une suite décroissante et $\cap_{n \geq n_0} B_n = \emptyset$. Comme m est finie sur les compacts, on a $m(B_n) \leq m([a,b]) < \infty$. En utilisant la continuité décroissante de m (proposition 2.3), on a donc :

$$m(]a,b[\setminus [a + (1/n), b - (1/n)]) = m(]a, a + (1/n)[\cup]b - (1/n), b[) = m(B_n) \to 0 \text{ quand } n \to \infty.$$

Soit maintenant $\varepsilon > 0$ en prenant n assez grand on a $m(B_n) \leq \varepsilon$. En prenant $O = A$ et $F = [a + (1/n), b - (1/n)]$, on a bien O ouvert, F fermé, $F \subset A \subset O$ et $m(O \setminus F) \leq \varepsilon$. Ceci prouve que $]a,b[\in T$.

On montre maintenant que T est une tribu. On remarque tout d'abord que $\emptyset \in T$ (il suffit de prendre $F = O = \emptyset$) et que T est stable par passage au complémentaire (car, si $F \subset A \subset O$, on a $O^c \subset A^c \subset F^c$ et $F^c \setminus O^c = O \setminus F$). Il reste à montrer que T est stable par union dénombrable.

Soit $(A_n)_{n \in \mathbb{N}} \subset T$ et $A = \cup_{n \in \mathbb{N}} A_n$. On veut montrer que $A \in T$. On va commencer par traiter le cas (simple) où $m(A) < \infty$ puis le cas (plus difficile) où $m(A) = \infty$.

Premier cas. On suppose que $m(A) < \infty$.

Soit $\varepsilon > 0$. Pour tout $n \in \mathbb{N}$, il existe O_n ouvert et F_n fermé t.q. $F_n \subset A_n \subset O_n$ et $m(O_n \setminus F_n) \leq (\varepsilon/2^n)$. On pose $O = \cup_{n \in \mathbb{N}} O_n$ et $\tilde{F} = \cup_{n \in \mathbb{N}} F_n$. On a $\tilde{F} \subset A \subset O$, $m(O \setminus \tilde{F}) \leq 2\varepsilon$, car $(O \setminus \tilde{F}) \subset \cup_{n \in \mathbb{N}} (O_n \setminus F_n)$, et O ouvert mais $\tilde{F}$ n'est pas nécessairement fermé...

Cependant, puisque $m(A) < \infty$, on a aussi $m(\tilde{F}) < \infty$. Par continuité croissante de m (appliquée à la suite $(\cup_{p=0}^n F_p)_{n \in \mathbb{N}}$) on a $m(\cup_{p=0}^n F_p) \to m(\tilde{F})$, quand $n \to \infty$, d'où (puisque $m(\tilde{F}) < \infty$) $m(\tilde{F}) - m(\cup_{p=0}^n F_p) \to 0$. On prend alors $F = \cup_{p=0}^N F_p$ avec N assez grand pour que $m(\tilde{F} \setminus F) = m(\tilde{F}) - m(F) \leq \varepsilon$. On a bien $F \subset A \subset O$, O ouvert, F fermé et, comme $(O \setminus F) = (O \setminus \tilde{F}) \cup (\tilde{F} \setminus F)$, on a $m(O \setminus F) = m(O \setminus \tilde{F}) + m(\tilde{F} \setminus F) \leq 3\varepsilon$. Ceci prouve que $A \in T$.

Deuxième cas. On suppose maintenant que $m(A) = \infty$ (et le raisonnement précédent n'est plus correct si $m(\tilde{F}) = \infty$). On raisonne en 3 étapes :

1. Soit $p \in \mathbb{Z}$. On remarque d'abord que $A_n \cap [p, p+1[\in T$ pour tout $n \in \mathbb{N}$. En effet, soit $n \in \mathbb{N}$ et $\varepsilon > 0$. Il existe O ouvert et F fermé t.q. $F \subset A_n \subset O$ et $m(O \setminus F) \leq \varepsilon$. Pour $k \in \mathbb{N}^\star$, on a donc :

$$F_k = F \cap [p, p + 1 - \frac{1}{k}] \subset A_n \cap [p, p+1[\subset O_k = O \cap]p - \frac{1}{k}, p+1[.$$

On a F_k fermé, O_k ouvert et $(O_k \setminus F_k) \subset (O \setminus F) \cup]p - \frac{1}{k}, p[\cup]p + 1 - \frac{1}{k}, p + 1[$. On en déduit :

$$m(O_k \setminus F_k) \leq \varepsilon + m(]p - \frac{1}{k}, p[\cup]p + 1 - \frac{1}{k}, p + 1[).$$

Or, la continuité décroissante de m donne que $m((]p - \frac{1}{k}, p[\cup]p + 1 - \frac{1}{k}, p + 1[) \to 0$ quand $k \to \infty$ (on utilise ici le fait que $m([p - 1, p + 1]) < \infty$ car m est finie sur les compacts). Il existe donc $k \in \mathbb{N}^\star$ t.q. $m(O_k \setminus F_k) \leq 2\varepsilon$, ce qui donne bien que $A_n \cap [p, p + 1[\in T$.

2. Comme $m(A \cap [p, p + 1[) < \infty$, on peut maintenant utiliser le premier cas avec $A \cap [p, p + 1[= \cup_{n \in \mathbb{N}}(A_n \cap [p, p + 1[)$. Il donne que $A \cap [p, p + 1[\in T$ pour tout $p \in \mathbb{Z}$.

3. On montre enfin que $A \in T$. Soit $\varepsilon > 0$. Pour tout $p \in \mathbb{Z}$, il existe un ouvert O_p et un fermé G_p t.q. $G_p \subset A \cap [p, p + 1[\subset O_p$ et $m(O_p \setminus G_p) \leq \varepsilon/(2^{|p|})$. On prend $O = \cup_{p \in \mathbb{Z}} O_p$ et $F = \cup_{p \in \mathbb{Z}} G_p$. On obtient $F \subset A \subset O$, $m(O \setminus F) \leq 3\varepsilon$ et O est ouvert. Il reste à montrer que F est fermé.

   Soit $(x_n)_{n \in \mathbb{N}} \subset F$ t.q. $x_n \to x$ (dans $\mathbb{R}$) quand $n \to \infty$. On veut montrer que $x \in F$. Il existe $p \in \mathbb{Z}$ t.q. $x \in]p - 1, p + 1[$. Il existe donc $n_0 \in \mathbb{N}$ t.q. $x_n \in]p - 1, p + 1[$ pour tout $n \geq n_0$. Comme $x_n \in \cup_{q \in \mathbb{Z}} G_q$ et que $G_q \subset [q, q + 1[$ pour tout q, on a donc $x_n \in G_p \cup G_{p-1}$ pour tout $n \geq n_0$. Comme $G_p \cup G_{p-1}$ est fermé, on en déduit que $x \in G_p \cup G_{p-1} \subset F$ et donc que F est fermé.

   Ceci montre bien que $A \in T$ et termine la démonstration du fait que T est une tribu. Comme cela a déjà été dit, on en déduit que $T = \mathcal{B}(\mathbb{R})$.

On a donc bien montré que pour tout $A \in \mathcal{B}(\mathbb{R})$ et pour tout $\varepsilon > 0$, il existe O ouvert et F fermé vérifiant $F \subset A \subset O$ et $m(O \setminus F) \leq \varepsilon$.

On montre maintenant que $m(A) = \inf\{m(O),\ O$ ouvert contenant $A\}$ pour tout $A \in \mathcal{B}(\mathbb{R})$. Soit $A \in \mathcal{B}(\mathbb{R})$. On remarque d'abord que la monotonie d'une mesure donne $m(A) \leq \inf\{m(O),\ O$ ouvert contenant $A\}$. Puis, l'inégalité inverse est immédiate si $m(A) = \infty$. Enfin, si $m(A) < \infty$, pour tout $\varepsilon > 0$ il existe O ouvert et F fermé vérifiant $F \subset A \subset O$ et $m(O \setminus F) \leq \varepsilon$. On a donc $O \setminus A \subset O \setminus F$ et donc (par monotonie de m) $m(O \setminus A) \leq \varepsilon$ et $m(O) = m(A) + m(O \setminus A) \leq m(A) + \varepsilon$. On a donc trouvé un ouvert O contenant A t.q. $m(O) - \varepsilon \leq m(A)$. On en déduit que $\inf\{m(O),\ O$ ouvert contenant $A\} \leq m(A)$ et finalement que $m(A) = \inf\{m(O),\ O$ ouvert contenant $A\}$.

De manière semblable, on montre aussi que $m(A) = \sup\{m(K),\ K$ compact inclus dans $A\}$ pour tout $A \in \mathcal{B}(\mathbb{R})$. En effet, soit $A \in \mathcal{B}(\mathbb{R})$. Ici aussi, on commence par remarquer que la monotonie d'une mesure donne $m(A) \geq \sup\{m(K),\ K$ compact inclus dans $A\}$. On montre maintenant l'inégalité inverse. Soit $\varepsilon > 0$. Il existe F fermé t.q. $F \subset A$ et $m(A \setminus F) \leq \varepsilon$. Si $m(A) = \infty$, on en déduit que $m(F) = \infty$ et donc que $m(K_n) \uparrow \infty$ quand $n \to \infty$ (par continuité croissante de m) avec $K_n = F \cap [-n, n]$. Comme K_n est compact pour tout $n \in \mathbb{N}$, on a donc $\sup\{m(K),\ K$ compact inclus dans $A\} = \infty = m(A)$. Si $m(A) < \infty$, on a $m(A) \geq m(F) \geq m(A) - \varepsilon$ et donc, pour n assez grand, $m(K_n) \geq m(F) - \varepsilon \geq m(A) - 2\varepsilon$ (toujours par continuité croissante de m) avec $K_n = F \cap [-n, n]$. Comme K_n est compact pour tout $n \in \mathbb{N}$ et que ε est arbitraire, on en déduit que $\sup\{m(K),\ K$ compact inclus dans $A\} \geq m(A)$ et donc, finalement, $m(A) = \sup\{m(K),\ K$ compact inclus dans $A\}$. ■

Pour démontrer la partie "unicité" du théorème 2.2 avec le théorème 2.3 on aura aussi besoin du petit lemme suivant (plus précis que le lemme 2.1 car dans le lemme 2.1 il n'est pas demandé que les ouverts soient disjoints).

Lemme 2.4 (Ouverts de $\mathbb{R}$) *Soit O un ouvert de $\mathbb{R}$, alors O est une union dénombrable d'intervalles ouverts disjoints 2 à 2, c'est à dire qu'il existe $(I_n)_{n\in\mathbb{N}}$ t.q. I_n est un intervalle ouvert de $\mathbb{R}$ pour tout $n \in \mathbb{N}$, $I_n \cap I_m = \emptyset$ si $n \neq m$ et $O = \cup_{n\in\mathbb{N}} I_n$.*

DÉMONSTRATION :
Pour $x \in O$ on pose $O_x = \{y \in O;\ I(x,y) \subset O\}$, avec $I(x,y) = \{tx + (1-t)y,\ t \in [0,1]\}$ (on a donc $I(x,y) = [x,y]$ ou $[y,x]$). On remarque que $O = \cup_{x\in O} O_x$ et que O_x est, pour tout $x \in O$, un intervalle ouvert (c'est l'intervalle $]\inf O_x, \sup O_x[$, avec $\inf O_x$, $\sup O_x \in \overline{\mathbb{R}}$). Il est aussi facile de voir que, pour tout x, $y \in O$, $O_x \cap O_y \neq \emptyset$ implique que $O_x = O_y$. On peut trouver $A \subset O$ t.q. $O = \cup_{x\in A} O_x$ et $O_x \cap O_y = \emptyset$ si $x, y \in A$, $x \neq y$. Comme $O_x \neq \emptyset$ pour tout $x \in A$, on peut donc construire une application de A dans $\mathbb{Q}$ en choisissant pour chaque $x \in A$ un rationnel de O_x (ce qui est possible car tout ouvert non vide de $\mathbb{R}$ contient un rationnel). Cette application est injective car $O_x \cap O_y = \emptyset$ si $x, y \in A$, $x \neq y$. l'ensemble A est donc au plus dénombrable, ce qui termine la démonstration du lemme.

■

Remarque 2.10 Dans la démonstration du lemme 2.4, O_x est la "composante connexe" de x. Le lemme 2.4 consiste donc à remarquer qu'un ouvert est réunion de ses composantes connexes, que celles ci sont disjointes deux à deux et sont des ouverts connexes et donc des intervalles ouverts (car un connexe dans $\mathbb{R}$ est nécessairement un intervalle).

DÉMONSTRATION DE LA PARTIE "UNICITÉ" DU THÉORÈME 2.2 :

On a construit une mesure, notée λ, sur $\mathcal{B}(\mathbb{R})$ t.q. $\lambda(]a,b[) = b - a$ pour tout a, $b \in \mathbb{R}$, $a < b$. Supposons que m soit aussi une mesure sur $\mathcal{B}(\mathbb{R})$ t.q. $m(]a,b[) = b - a$ pour tout a, $b \in \mathbb{R}$, $a < b$. On veut montrer que $\lambda = m$ (sur tout $\mathcal{B}(\mathbb{R})$). Nous le montrons ici avec deux méthodes différentes, utilisant le théorème 2.3 ou la proposition 2.5.

Première méthode, avec le théorème 2.3. En utilisant le fait que tout ouvert est réunion dénombrable d'intervalles ouverts disjoints deux à deux (lemme 2.4) et les propriétés de $\sigma-$additivité de λ et de m, on montre que $\lambda(O) = m(O)$ pour tout ouvert O de $\mathbb{R}$. Puis, en utilisant la dernière assertion du théorème de régularité (qui s'applique pour m et pour λ, car m et λ sont des mesures sur $\mathcal{B}(\mathbb{R})$, finies sur les compacts), on obtient $\lambda(A) = m(A)$ pour tout $A \in \mathcal{B}(\mathbb{R})$, i.e. $m = \lambda$.

Deuxième méthode, avec la proposition 2.5. On utilise la proposition 2.5 avec $(E,T) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ et $\mathcal{C} = \{]a,b],\ a,b \in \mathbb{R}, a \leq b\}$. On sait que $\mathcal{C}$ engendre $\mathcal{B}(\mathbb{R})$, et il est clair que $\mathcal{C}$ est stable par intersection finie. On prend maintenant $F_n =]n, n+1]$ pour $n \in \mathbb{Z}$. La famille $(F_n)_{n\in\mathbb{Z}}$ est donc une famille dénombrable d'éléments de $\mathcal{C}$, disjoints deux à deux et t.q. $\mathbb{R} = \cup_{n\in\mathbb{Z}} F_n$. Pour $a, b \in \mathbb{R}$, $a \leq b$, on a, par continuité décroissante de m, $m(]a,b] = \lim_{p\to\infty} m(]a, b + \frac{1}{p}[= \lim_{p\to\infty} b - a + \frac{1}{p} = b - a = \lambda(]a,b])$. On a donc $m = \lambda$ sur $\mathcal{C}$ (et $m(F_n) < \infty$ pour tout $n \in \mathbb{Z}$). On peut donc appliquer la proposition 2.5. Elle donne $\lambda = m$ sur $\mathcal{B}(\mathbb{R})$. ■

Remarque 2.11 Nous avons vu que la mesure de Lebesgue, notée λ, était régulière. Ceci ne donne pas, pour $A \in \mathcal{B}(\mathbb{R})$, l'égalité de la mesure de A avec la mesure de son intérieur ou de son adhérence. Il suffit, pour s'en convaincre, de prendre, par exemple, $A = \mathbb{Q}$. On a alors $\lambda(A) = 0$ (voir la remarque 2.13) et $\lambda(\overline{A}) = \infty$.

Remarque 2.12 Nous avons donc, dans cette section, construit une application, notée $\lambda^\star$, de $\mathcal{P}(\mathbb{R})$ dans $\overline{\mathbb{R}}_+$. Cette application n'est pas une mesure mais nous avons montré que la restriction de $\lambda^\star$ à la tribu de Lebesgue, notée $\mathcal{L}$, était une mesure. Puis, nous avons démontré que $\mathcal{B}(\mathbb{R}) \subset \mathcal{L}$ et obtenu ainsi, en prenant la restriction de $\lambda^\star$ à $\mathcal{B}(\mathbb{R})$ la mesure que nous cherchions. On peut se demander toutefois quelle est la différence entre $\mathcal{L}$ et $\mathcal{B}(\mathbb{R})$. Du point de vue des cardinaux, cette différence est considérable car $\mathrm{card}(\mathcal{L}) = \mathrm{card}(\mathcal{P}(\mathbb{R}))$ alors que $\mathrm{card}(\mathcal{B}(\mathbb{R})) = \mathrm{card}(\mathbb{R})$ mais du point de vue de l'intégration, la différence est dérisoire, comme nous pourrons le voir avec l'exercice 4.18 (plus complet que l'exercice 2.32) car l'espace mesuré $(\mathbb{R}, \mathcal{L}, \lambda_{|\mathcal{L}})$ est simplement le complété de $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda_{|\mathcal{B}(\mathbb{R})})$.

On donne maintenant une propriété, spécifique à la mesure de Lebesgue, qui est à la base de toutes les formules de changement de variables pour l'intégrale de Lebesgue.

Proposition 2.9 (Invariance par translation "généralisée") *Soit $\alpha \in \mathbb{R}^\star$ et $\beta \in \mathbb{R}$. Pour $A \in \mathcal{P}(\mathbb{R})$, on note $\alpha A + \beta = \{\alpha x + \beta,\ x \in A\}$. On a alors :*

1. *$A \in \mathcal{B}(\mathbb{R})$ implique $\alpha A + \beta \in \mathcal{B}(\mathbb{R})$,*
2. *$\lambda(\alpha A + \beta) = |\alpha|\lambda(A)$ pour tout $A \in \mathcal{B}(\mathbb{R})$.*

Pour $\alpha = 1$, cette propriété s'appelle "invariance par translation de λ".

Démonstration :
Pour la première partie de la proposition, on pose $T = \{A \in \mathcal{B}(\mathbb{R}); \alpha A + \beta \in \mathcal{B}(\mathbb{R})\}$. On montre facilement que T est une tribu contenant les intervalles ouverts, on en déduit que $T = \mathcal{B}(\mathbb{R})$.

Pour la deuxième partie, on pose, pour tout $A \in \mathcal{B}(\mathbb{R})$, $m_1(A) = \lambda(\alpha A + \beta)$ et $m_2(A) = |\alpha|\lambda(A)$. Il est facile de voir que m_1 et m_2 sont des mesures sur $\mathcal{B}(\mathbb{R})$, finies sur les bornés, et qu'elles sont égales sur l'ensemble des intervalles ouverts. On raisonne alors comme dans la démonstration de la partie "unicité" du théorème 2.2, en utilisant le théorème 2.3 ou la proposition 2.5. Par exemple, en utilisant le lemme 2.4 et les propriétés de $\sigma-$additivité de m_1 et de m_2, on montre que $m_1(O) = m_2(O)$ pour tout ouvert O de $\mathbb{R}$. Puis, en utilisant la dernière assertion du théorème de régularité (qui s'applique pour m_1 et pour m_2), on obtient $m_1(A) = m_2(A)$ pour tout $A \in \mathcal{B}(\mathbb{R})$. On a donc $\lambda(\alpha A + \beta) = |\alpha|\lambda(A)$ pour tout $A \in \mathcal{B}(\mathbb{R})$.

■

Remarque 2.13 La mesure de Lebesgue est diffuse (c'est-à-dire que $\lambda(\{x\}) = 0$ pour tout $x \in \mathbb{R}$). Donc, si D est une partie dénombrable de $\mathbb{R}$, on a $\lambda(D) = 0$. Ainsi, $\lambda(\mathbb{N}) = \lambda(\mathbb{Z}) = \lambda(\mathbb{Q}) = 0$. La réciproque est fausse. On construit par exemple un ensemble (dit "ensemble de Cantor", K, qui est une partie compacte non dénombrable de [0,1], vérifiant $\lambda(K) = 0$, voir exercice 2.31).

Définition 2.20 (Mesure de Lebesgue sur un borélien de $\mathbb{R}$) *Soit I un intervalle de $\mathbb{R}$ (ou, plus généralement, $I \in \mathcal{B}(\mathbb{R})$) et $T = \{B \in \mathcal{B}(\mathbb{R});\ B \subset I\}$ (on peut montrer que $T = \mathcal{B}(I)$, où I est muni de la topologie induite par celle de $\mathbb{R}$, voir l'exercice 2.3 page 41). Il est facile de voir que T est une tribu sur I et que la restriction de λ (définie dans le théorème 2.2) à T est une mesure sur T, donc sur les boréliens de I (voir 2.16 page 45). On note toujours par λ cette mesure.*

2.6 Indépendance et probabilité conditionnelle

2.6.1 Probabilité conditionnelle

Commençons par expliquer la notion de probabilité conditionnelle sur l'exemple du lancer de dé. On se place dans le modèle équiprobable: soient $E = \{1, 2, 3, 4, 5, 6\}$, $T = \mathcal{P}(E)$ et p la probabilité définie par

$p(\{x\}) = \frac{1}{6}$, $\forall x \in E$. La probabilité de l'événement A "obtenir 6" est $\frac{1}{6}$. Supposons maintenant que l'on veuille évaluer la chance d'obtenir un 6, alors que l'on sait déjà que le résultat est pair (événement $B = \{2, 4, 6\}$). Intuitivement, on a envie de dire que la "chance" d'obtenir un 6 est alors $\frac{1}{\text{card}B} = \frac{1}{3}$.

Définition 2.21 (Probabilité conditionnelle) *Soient (E, T, p) un espace probabilisé et $A, B \in T$.*
Si $p(B) \neq 0$, la probabilité conditionnelle de A par rapport à B (on dit aussi probabilité de A par rapport à B), notée $p(A|B)$, est définie par $p(A|B) = \frac{p(A \cap B)}{p(B)}$.
Si $p(B) = 0$, la probabilité conditionnelle de A par rapport à B, notée $p(A|B)$, n'est pas définie. C'est un nombre arbitraire entre 0 et 1.

De cette définition on déduit la formule de Bayes: soient (E, T, p) un espace probabilisé et $A, B \in T$, alors:

$$p(B)p(A|B) = p(A \cap B) \tag{2.17}$$

Remarque 2.14 Soient (E, T, p) un espace probabilisé et A un événement tel que $p(A) \neq 0$. Alors l'application $p_A : T \to [0, 1]$ définie par:

$$p_A(B) = p(B|A) = \frac{p(A \cap B)}{p(A)}, \forall B \in T \tag{2.18}$$

est une probabilité sur T. On dit que "la masse de p_A est concentrée en A" : on a en effet : $p_A(B) = 0$, pour tout $B \in T$ t.q. $A \cap B = \emptyset$. On a aussi $p_A(A) = 1$.

Définition 2.22 *Soit (E, T, p) un espace probabilisé, on appelle système de constituants une famille $(C_n)_{n \in \mathbb{N}} \subset T$ d'ensembles disjoints deux à deux t.q. $\cup_{n \in \mathbb{N}} C_n = E$.*

On a comme corollaire immédiat de la relation 2.17:

Proposition 2.10 *Soient (E, T, p) un espace probabilisé, $(C_n)_{n \in \mathbb{N}} \subset T$ un système de constituants de probabilités non nulles et $A \in T$, alors:*

$$p(A) = \sum_{n \in \mathbb{N}} p(C_n)p(A|C_n). \tag{2.19}$$

Dans le cas où $p(B) = p(B|A)$, on a envie de dire que A n'influe en rien sur B ; on a dans ce cas: $p(A)p(B) = p(A \cap B)$.

2.6.2 Evènements indépendants, tribus indépendantes

Définition 2.23 (Indépendance de deux évènements) *Soient (E, T, p) on dit que deux événements A et B sont (stochastiquement) indépendants si $p(A)p(B) = p(A \cap B)$.*

Remarque 2.15 Attention: il est clair que, lors de la modélisation d'un phénomène aléatoire, si on a des évènements indépendants *a priori*, i.e. tels que la réalisation de l'un n'a aucune influence sur la réalisation de l'autre, on choisira, pour le modèle probabiliste, une probabilité qui respecte l'indépendance: on dit que l'indépendance *a priori* implique l'indépendance stochastique. Cependant, la notion d'indépendance est liée à la notion de probabilité; ainsi, pour une probabilité p donnée, deux évènements peuvent être indépendants alors qu'ils ne paraissent pas intuitivement indépendants.

Exemple 2.3 Prenons comme exemple le lancer simultané de deux dés: *a priori*, il parait raisonnable de supposer que les résultats obtenus pour chacun des deux dés n'influent pas l'un sur l'autre, et on va donc chercher une probabilité qui respecte cette indépendance. L'univers des possibles est ici $E = \{(i,j), 1 \le i \le 6, 1 \le j \le 6\}$. Les résultats de chaque lancer simultané des deux dés étant équiprobables, on a donc envie de définir, pour $A \in \mathcal{P}(E)$, $p(A) = \frac{\text{card}A}{36}$. Voyons maintenant si deux évènements *a priori* indépendants sont indépendants pour cette probabilité. Considérons par exemple l'évènement A: "obtenir un double 6"; on peut écrire: $A = B \cap C$, où B est l'évènement "obtenir un 6 sur le premier dé" et C l'évènement "obtenir un 6 sur le deuxième dé". On doit donc vérifier que: $p(A) = p(B)p(C)$. Or $B = \{(6,j), 1 \le j \le 6\}$ et $C = \{(i,6), 1 \le i \le 6\}$. On a donc $p(B) = p(C) = \frac{1}{6}$, et on a bien $p(A) = p(B)p(C)(= \frac{1}{36})$.

On généralise la notion d'indépendance de deux évènements en introduisant la notion d'indépendance de tribus.

Définition 2.24 (Indépendance des tribus) *Soit (E,T,p) un espace probabilisé et $(T_k)_{k\in\mathbb{N}^\star}$ une suite de tribus incluses dans T.*

1. *Soit $N > 1$. On dit que les N tribus T_k, $k = 1, \ldots, N$, sont indépendantes (on dit aussi que la suite $T_1, \ldots, T_N$ est indépendante) si pour toute famille $(A_1, \ldots, A_N)$ d'événements tels que $A_k \in T_k$ pour $k = 1, \ldots, N$ on a: $p(\cap_{k=1}^N A_k) = p(A_1)p(A_2)\ldots p(A_N)$.*
2. *On dit que la suite $(T_k)_{k\in\mathbb{N}^\star}$ est indépendante (ou que les tribus $T_1, \ldots, T_n, \ldots$ sont indépendantes) si„ pour tout $N \ge 1$, les N tribus T_k, $k = 1, \ldots, N$, sont indépendantes.*

On peut facilement remarquer que si A et B sont deux évènements d'un espace probabilisé (E,T,p), ils sont indépendants (au sens de la définition 2.23) si et seulement si les tribus $T_A = \{\emptyset, E, A, A^c\}$ et $T_B = \{\emptyset, E, B, B^c\}$ sont indépendantes (voir l'exercice 3.17). On en déduit la généralisation de la définition d'indépendance à plusieurs évènements:

Définition 2.25 (Evénements indépendants) *Soient (E,T,p) un espace probabilisé et $(A_k)_{k=1,\ldots,N}$ des événements, on dit que les N événements $(A_k)_{k=1,\ldots,N}$ sont indépendants si les N tribus engendrées par les événements A_k, $k = 1\ldots, N$ (c'est-à-dire les N tribus définies par $T_k = \{A_k, A_k^c, E, \emptyset\}$ pour $k = 1\ldots, N$) sont indépendantes.*

Sous les hypothèses de la définition précédente, on peut remarquer que les événements $A_1, \ldots, A_N$ sont indépendants (c'est-à-dire que les tribus engendrées par $A_1, \ldots, A_N$ sont indépendantes) si et seulement si $P(\cap_{i\in I} A_i) = \prod_{i\in I} P(A_i)$ pour tout $I \subset \{1, \ldots, N\}$), voir l'exercice 3.17. Nous terminons ce paragraphe par une proposition sur les tribus indépendantes :

Proposition 2.11 *Soit (E,T,p) un espace probabilisé.*

1. *Soit $N > 1$ et $(T_k)_{k\in\{0,\ldots N\}}$ une suite indépendante de tribus incluses dans T. La tribu T_0 est alors indépendante de la tribu engendrée par les tribus $T_1, \ldots, T_N$.*
2. *(Généralisation) Soit $N > 3$, $q > 1$, $n_0, \ldots, n_q$ t.q. $n_0 = 0$, $n_i < n_{i+1}$ (pour $i = 0, \ldots, q-1$), $n_q = N$ et $(T_k)_{k\in\{1,\ldots N\}}$ une suite indépendante de tribus incluses dans T. Pour $i = 1, \ldots, q$, on note τ_i la tribu engendrée par les tribus T_n pour $n = n_{i-1}, \ldots, n_i$. Alors, les tribus $\tau_1, \ldots, \tau_q$ sont indépendantes.*

DÉMONSTRATION : On montre tout d'abord le premier item de la proposition. On note S la tribu engendrée par les tribus $T_1, \ldots, T_N$. Comme S est la plus petite tribu contenant les tribus T_k ($k = 1, \ldots, N$), elle est incluse dans T. On veut montrer que T_0 et S sont indépendantes, c'est-à-dire que $p(A \cap B) = p(A)p(B)$ pour tout $A \in T_0$ et tout $B \in S$. Pour le montrer, on va utiliser la proposition 2.5 (donnant l'unicité d'une mesure). Soit $A \in T_0$, on définit les mesures m et μ sur T en posant :

$$m(B) = p(A \cap B),\ \mu(B) = p(A)p(B),\ \text{pour } B \in T,$$

et on pose :

$$\mathcal{C} = \{\cap_{k=1}^N A_k,\ A_k \in T_k \text{ pour } k = 1, \ldots, N\}.$$

Pour $B \in \mathcal{C}$, on a $B = \cap_{k=1}^N A_k$ avec $A_k \in T_k$ avec $k = 1, \ldots, N$. On a donc, en utilisant l'indépendance des tribus $T_0, T_1, \ldots, T_N$, $m(B) = p(A \cap B) = p(A)p(A_1)p(A_2) \ldots p(A_N) = p(A)p(B) = \mu(B)$. On a donc $m = \mu$ sur $\mathcal{C}$. Comme $\mathcal{C}$ est stable par intersection et que $E \in \mathcal{C}$, la proposition 2.5 nous donne $m = \mu$ sur la tribu engendrée par $\mathcal{C}$. Comme cette tribu contient toutes les tribus T_k ($k = 1, \ldots, N$), elle contient aussi S (en fait, elle est égale à S). On a donc bien montré que $p(A \cap B) = p(A)p(B)$ pour tout $B \in S$ et pour tout $A \in T_0$.

Pour montrer le deuxième item (qui est une généralisation du premier), il suffit de faire une récurrence finie de q étapes et d'utiliser la technique précédente. Par exemple, pour $q = 2$ la technique précédente donne :

$$p((\cap_{k=1}^{n_1} A_k) \cap B_2) = p(\cap_{k=1}^{n_1} A_k)p(B_2),$$

pour $A_k \in T_k$, $k = 1, \ldots, n_1$ et $B_2 \in \tau_2$. Puis en reprenant la technique précénete, on montre $p(B_1 \cap B_2) = p(B_1)p(B_2)$ pour $B_1 \in \tau_1$ et $B_2 \in \tau_2$. Ce qui donne bien l'indépendance de τ_1 et τ_2. ■

2.6.3 Probabilités sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$

Une probabilité est définie sur un espace probabilisable quelconque. Très souvent, on ne connait du problème aléatoire que l'on cherche à modéliser ni l'ensemble E ("univers des possibles") ni la tribu T (ensemble des évènements) ni la probabilité p. Par contre, on connait une "image" de la probabilité p par une application (dite mesurable, voir chapitre suivant) X de E dans $\mathbb{R}$. On travaille alors avec l'espace beaucoup plus sympathique (car mieux défini...) $(\mathbb{R}, \mathcal{B}(\mathbb{R}), p_X)$, ou p_X est une probabilité sur $\mathcal{B}(\mathbb{R})$, que les probabilistes appellent souvent "loi de probabilité" (elle dépend de p et de l'application X).
Nous donnons maintenant quelques notions propres aux lois de probabilités (ou probabilités définies sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$), ainsi que quelques exemples concrets utilisés dans la représentation de phénomènes aléatoires.

Définition 2.26 (Fonction de répartition) *Soit p une probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. On appelle fonction de répartition de la probabilité p la fonction F, définie de $\mathbb{R}$ dans $[0,1]$ par: $F(t) = p(]-\infty, t])$. Cette fonction est croissante et continue à droite.*

DÉMONSTRATION : Utiliser les propriétés de monotonie et de continuité croissante de la mesure.

■

Théorème 2.4 *Soit F une fonction croissante et continue à droite t.q.:*
$\lim_{t \to -\infty} F(t) = 0$ *et* $\lim_{t \to +\infty} F(t) = 1$.
Alors, il existe une unique probabilité p sur $\mathcal{B}(\mathbb{R})$ telle que F soit la fonction de répartition de p.

Plus généralement, on a le théorème suivant pour les mesures:

Théorème 2.5 (Lebesgue-Stieltjes)

1. *Pour toute mesure m sur $\mathcal{B}(\mathbb{R})$, finie sur les compacts (on dit "localement finie"), et pour tout $a \in \mathbb{R}$, la fonction F définie sur $\mathbb{R}$ par : $F(t) = m(]a,t])$ si $t \geq a$ et $F(t) = -m(]t,a])$ si $t \leq a$ est continue à droite et croissante.*

2. *Réciproquement, pour toute fonction F de $\mathbb{R}$ dans $\mathbb{R}$, croissante et continue à droite, il existe une unique mesure m sur $\mathcal{B}(\mathbb{R})$ t.q. pour tous $a, b \in \mathbb{R}$, t.q. $a \leq b$, on ait $m(]a,b]) = F(b) - F(a)$. Cette mesure s'appelle la mesure de Lebesgue-Stieltjes associée à F.*

Pour démontrer ce théorème, on introduit $p^\star$, application définie de l'ensemble des intervalles de $\mathbb{R}$ de la forme $]a,b]$ dans $\mathbb{R}$ par: $p^\star(]a,b]) = F(b) - F(a)$. La démonstration du fait qu'il existe un prolongement unique de cette application en une mesure sur $\mathcal{B}(\mathbb{R})$ est très voisine à celle du théorème de Carathéodory (théorème 2.2).

Donnons, pour clore ce chapitre, quelques exemples de lois de probabilités et leurs fonctions de répartition associées.

Définition 2.27 (Loi discrète) *Soit p une probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. On dit que p est discrète si elle est purement atomique (l'ensemble de ses atomes $\mathcal{A}$ est dénombrable, voir exercice 2.12). La probabilité p s'écrit alors $p = \sum_{a \in \mathcal{A}} p(\{a\})\delta_a$, où δ_a désigne la mesure de Dirac en a. La fonction de répartition de la probabilité p est définie par: $F(t) = \sum_{a \in \mathcal{A}, a \leq t} p(\{a\})$*

Exemple 2.4 (Exemples de lois discrètes) Donnons quelques exemples de probabilités discrètes, p, sur $\mathcal{B}(\mathbb{R})$ et de $\mathcal{A}$ l'ensemble (dénombrable) de leurs atomes.

- La loi uniforme : $N \in \mathbb{N}^\star$, $\mathcal{A} = \{a_1, \ldots, a_N\}$, $p(\{a_i\}) = \frac{1}{N}$
- La loi binômiale : $N \in \mathbb{N}^\star$, $\mathcal{A} = \{1, \ldots, N\}$, $P \in]0,1[$, $p(\{k\}) = C_N^k P^k (1-P)^{N-k}$
- La loi de Pascal : $\mathcal{A} = \mathbb{N}$, $P \in]0,1[$, $p(\{k\}) = P(1-P)^{k-1}$
- La loi de Poisson à paramètre λ : $\mathcal{A} = \mathbb{N}$, $\lambda > 0$, $p(\{k\}) = e^{-\lambda}\frac{\lambda^k}{k!}$

Définition 2.28 (Loi continue) *Soit p une probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. On dit que p est continue si sa fonction de répartition est continue.*

Exemple 2.5 (Exemple de loi continue) La plupart des exemples de probabilités continues provient de ce qu'on appelle "les mesures de densité" par rapport à la mesure de Lebesgue, pour lesquelles on a besoin de la notion d'intégrale de Lebesgue qu'on n'a pas encore introduite. On peut toutefois déjà citer l'exemple de la loi uniforme sur un intervalle $[a,b]$ de $\mathbb{R}$: Soient $-\infty < a < b < +\infty$; pour $A \in \mathcal{B}(\mathbb{R})$, on pose $p(A) = \frac{\lambda(A \cap [a,b])}{b-a}$. On vérifie facilement que p est une probabilité appelée probabilité uniforme sur $[a,b]$.

2.7 Exercices

2.7.1 Tribus

Exercice 2.1 (Caractérisation d'une tribu) *Corrigé 9 page 283*
Soit E un ensemble.

1. Soit T une partie de $\mathcal{P}(E)$ stable par union dénombrable, stable par passage au complémentaire et t.q. $\emptyset \in T$. Montrer que T est une tribu, c'est-à-dire qu'elle vérifie aussi $E \in T$ et qu'elle est stable par intersection dénombrable.

2. L'ensemble des parties finies de E est-il une tribu ?

Exercice 2.2 (Tribu engendrée) *Corrigé 10 page 283*
Soit E un ensemble.

1. Montrer qu'une intersection quelconque de tribus sur E est une tribu sur E.

2. Soit $\mathcal{A} \subset \mathcal{P}(E)$. On note $T_{\mathcal{A}}$ l'intersection de toutes les tribus sur E contenant $\mathcal{A}$ (une partie de E appartient donc à $T_{\mathcal{A}}$ si et seulement si elle appartient à toutes les tribus contenant $\mathcal{A}$, on remarquera qu'il y a toujours au moins une tribu contenant $\mathcal{A}$, c'est la tribu $\mathcal{P}(E)$). Montrer que $T_{\mathcal{A}}$ est la plus petite des tribus contenant $\mathcal{A}$ (c'est la tribu engendrée par $\mathcal{A}$).

3. Soient $\mathcal{A}$ et $\mathcal{B} \subset \mathcal{P}(E)$ et $T_{\mathcal{A}}$, $T_{\mathcal{B}}$ les tribus engendrées par $\mathcal{A}$ et $\mathcal{B}$. Montrer que si $\mathcal{A} \subset \mathcal{B}$ alors $T_{\mathcal{A}} \subset T_{\mathcal{B}}$.

Exercice 2.3 (Exemples de tribus) *Corrigé 11 page 284*

1. Tribu trace

   (a) Soit $\mathcal{T}$ une tribu sur un ensemble E et $F \subset E$. Montrer que $\mathcal{T}_F = \{A \cap F;\, A \in \mathcal{T}\}$ est une tribu sur F (tribu trace de $\mathcal{T}$ sur F).

   (b) Si E est un espace topologique et $\mathcal{T} = \mathcal{B}(E)$ ($\mathcal{B}(E)$ est la tribu borélienne de E), montrer que la tribu trace sur F, notée T_F, est la tribu engendrée par la topologie trace sur F (tribu borélienne de F, notée $\mathcal{B}(F)$). [Montrer que $\mathcal{B}(F) \subset T_F$. Pour montrer que $T_F \subset \mathcal{B}(F)$, considérer $\mathcal{C} = \{A \in \mathcal{P}(E);\, A \cap F \in \mathcal{B}(F)\}$ et montrer que $\mathcal{C}$ est une tribu (sur E) contenant les ouverts de E.] Si F est un borélien de E, montrer que T_F est égale à l'ensemble des boréliens de E contenus dans F.

2. Soit E un ensemble infini et $S = \{\{x\}, x \in E\}$. Déterminer la tribu engendrée par S (distinguer les cas E dénombrable et non dénombrable).

3. Soit E un ensemble et f une application de E dans lui-même. Montrer que l'ensemble des parties A de E t.q. $f^{-1}(f(A)) = A$ est une tribu sur E.

4. Soit E un ensemble et A un sous-ensemble de E. Trouver la tribu engendrée par $\mathcal{C} = \{B \subset E \mid A \subset B\}$. A quelle condition obtient-on $\mathcal{P}(E)$ ou la tribu grossière ?

5. Soit E un ensemble et f une bijection de E. Montrer que l'ensemble des parties A de E t.q. ($x \in A \iff f(x) \in A$ et $f^{-1}(x) \in A$) est une tribu sur E.

6. Dans $\mathbb{R}^2$, soit $\mathcal{C}$ l'ensemble des parties de $\mathbb{R}^2$ contenues dans une réunion dénombrable de droites. Décrire la tribu engendrée par $\mathcal{C}$. Est-ce une sous-tribu de $\mathcal{B}(\mathbb{R}^2)$?

Exercice 2.4 (Tribus images) *Corrigé 12 page 285*
Soient E et F des ensembles. Pour $\mathcal{A} \subset \mathcal{P}(E)$ (resp. $\mathcal{P}(F)$) on note $T(\mathcal{A})$ la tribu de E (resp. F) engendrée par $\mathcal{A}$.
Soit $f\ : E\ \to F$ une application.

1. Montrer que si $\mathcal{T}'$ est une tribu sur F, alors $f^{-1}(\mathcal{T}') = \{f^{-1}(B);\ B \in \mathcal{T}'\}$ est une tribu sur E (tribu image réciproque).

2. Montrer que si $\mathcal{T}$ est une tribu sur E, alors $\mathcal{T}' = \{B \subset F;\ f^{-1}(B) \in \mathcal{T}\}$ est une tribu sur F (tribu image directe).

3. Montrer que pour tout ensemble $\mathcal{C}$ de parties de F on a : $T(f^{-1}(\mathcal{C})) = f^{-1}(T(\mathcal{C}))$. [*Montrer que* $T(f^{-1}(\mathcal{C})) \subset f^{-1}(T(\mathcal{C}))$. *Puis, pour montrer que* $f^{-1}(T(\mathcal{C})) \subset T(f^{-1}(\mathcal{C}))$, *montrer que* $T = \{G \subset F;\ f^{-1}(G) \in T(f^{-1}(\mathcal{C}))\}$ *est une tribu contenant* $\mathcal{C}$.]

Exercice 2.5 (π-système, λ-système) *Corrigé 13 page 286*
Soit Ω un ensemble et $\mathcal{F} \subset \mathcal{P}(\Omega)$.

1. Montrer que $\mathcal{F}$ est une tribu si et seulement si $\mathcal{F}$ est un π-système (c'est-à-dire stable par intersection finie) et un λ-système (c'est-à-dire que $\mathcal{F}$ est stable par union dénombrable croissante, $\Omega \in \mathcal{F}$ et $A \setminus B \in \mathcal{F}$ si $A, B \in \mathcal{F}$ avec $B \subset A$).

2. On suppose que $\mathcal{F}$ est un λ-système. Soit $C \in \mathcal{F}$. On pose $\mathcal{G} = \{B \subset \Omega \text{ t.q. } C \cap B \in \mathcal{F}\}$. Montrer que $\mathcal{G}$ est un λ-système.

Exercice 2.6 (Tribu borélienne de $\mathbb{R}^2$) *Corrigé 14 page 287*
On note T la tribu (sur $\mathbb{R}^2$) engendrée par $\{A \times B;\ A,\ B \in \mathcal{B}(\mathbb{R})\}$. On va montrer ici que $T = \mathcal{B}(\mathbb{R}^2)$.

1. Montrer que tout ouvert de $\mathbb{R}^2$ est réunion au plus dénombrable de produits d'intervalles ouverts de $\mathbb{R}$. [S'inspirer d'une démonstration analogue faite pour $\mathbb{R}$ au lieu de $\mathbb{R}^2$.] En déduire que $\mathcal{B}(\mathbb{R}^2) \subset T$.

2. Soit A un ouvert de $\mathbb{R}$ et $T_1 = \{B \in \mathcal{B}(\mathbb{R});\ A \times B \in \mathcal{B}(\mathbb{R}^2)\}$. Montrer que T_1 est une tribu (sur $\mathbb{R}$) contenant les ouverts (de $\mathbb{R}$). En déduire que $T_1 = \mathcal{B}(\mathbb{R})$.

3. Soit $B \in \mathcal{B}(\mathbb{R})$ et $T_2 = \{A \in \mathcal{B}(\mathbb{R});\ A \times B \in \mathcal{B}(\mathbb{R}^2)\}$. Montrer que $T_2 = \mathcal{B}(\mathbb{R})$.

4. Montrer que $T \subset \mathcal{B}(\mathbb{R}^2)$ (et donc que $T = \mathcal{B}(\mathbb{R}^2)$).

Exercice 2.7 (Tribu borélienne sur $\mathbb{R}^N$) *Corrigé 15 page 288*

1. Montrer que la tribu borélienne de $\mathbb{R}^N$ est égale à celle engendrée par l'ensemble de toutes les boules ouvertes de $\mathbb{R}^N$. [On pourra montrer d'abord que tout ouvert de $\mathbb{R}^N$ est réunion dénombrable de boules ouvertes de $\mathbb{R}^N$.]

2. Montrer que la tribu borélienne de $\mathbb{R}^N$ est égale à celle engendrée par l'ensemble des produits d'intervalles ouverts à extrémités rationnelles.

3. Montrer que la tribu borélienne de $\mathbb{R}$ est engendrée par les intervalles $]a, b]$ où $a, b \in \mathbb{R}$, $a < b$.

4. Soit S un sous ensemble dense de $\mathbb{R}$. Montrer que $\mathcal{B}(\mathbb{R}^N)$ est engendrée par la classe des boules ouvertes (ou bien fermées) telles que les coordonnées du centre et le rayon appartiennent S.

Exercice 2.8 (Une tribu infinie est non dénombrable) ($\star\star$)
Montrer que toute tribu infinie T sur un ensemble (infini) E est non dénombrable. [Si T est dénombrable, on pourra introduire, pour tout élément $x \in E$, l'ensemble $A(x)$ intersection de tous les éléments de T contenant x. Puis, montrer à l'aide de ces ensembles qu'il existe une injection de $\mathcal{P}(\mathbb{N})$ dans T.]

Exercice 2.9 (Algèbre) *Corrigé 16 page 289*
Soit E un ensemble et $\mathcal{A} \subset \mathcal{P}(E)$.

1. Montrer que $\mathcal{A}$ est une algèbre (cf. définition 2.4) si et seulement si $\mathcal{A}$ vérifie les deux propriétés suivantes :

   (a) $E \in \mathcal{A}$,

   (b) $A, B \in \mathcal{A} \Rightarrow A \setminus B \in \mathcal{A}$.

2. Soit $(\mathcal{A}_i)_{i\in I}$ une famille d'algèbres (sur E). Montrer que $\cap_{i\in I}\mathcal{A}_i = \{A \in \mathcal{P}(E);\ A \in \mathcal{A}_i$ pour tout $i \in I\}$ est encore une algèbre.

Si $\mathcal{C} \subset \mathcal{P}(E)$, la deuxième question permet donc de définir l'algèbre engendrée par $\mathcal{C}$ comme l'intersection de toutes les algèbres sur E contenant $\mathcal{C}$.

Exercice 2.10 (Suite croissante de tribus)
Soit E un ensemble. Soit $(\mathcal{A}_n)_{n\in\mathbb{N}}$ une suite croissante de tribus de E. Montrer que $\mathcal{A} = \cup_{n\in\mathbb{N}}\mathcal{A}_n$ est une algèbre (cf. définition 2.4), mais n'est pas, en général, une tribu. Donner une suite d'algèbres finies de parties de $[0,1]$ dont la réunion engendre $\mathcal{B}([0,1])$.

Exercice 2.11 (Tribu engendrée par une partition)
Soit E un ensemble.

1. Soit $(A_i)_{i\in I}$ une partition dénombrable de E. Décrire la tribu engendrée par cette partition, c'est à dire par le sous ensemble de $\mathcal{P}(E)$ dont les éléments sont les A_i. Cette tribu est-elle dénombrable?

2. Montrer que toute tribu finie de parties de E est la tribu engendrée par une partition finie de E. Quel est le cardinal d'une telle tribu?

3. $(\star)$ Montrer que si E est dénombrable, toute tribu sur E est engendrée par une partition.

Exercice 2.12 *Corrigé 17 page 290*
Soit E un ensemble et $\mathcal{C}$ un ensemble de parties de E. On suppose que $\emptyset, E \in \mathcal{C}$, que $\mathcal{C}$ est stable par intersection finie et que le complémentaire de tout élément de $\mathcal{C}$ est une union finie disjointe d'éléments de $\mathcal{C}$, c'est-à-dire :

$$C \in \mathcal{C} \Rightarrow \text{ il existe } n \in \mathbb{N}^\star \text{ et } C_1, \ldots, C_n \in \mathcal{C} \text{ t.q. } C^c = \cup_{p=1}^n C_p \text{ et } C_p \cap C_q = \emptyset \text{ si } p \neq q.$$

On note $\mathcal{B}$ l'ensemble des réunions finies disjointes d'éléments de $\mathcal{C}$. Une partie de E est donc un élément de $\mathcal{B}$ si et seulement si il existe $n \in \mathbb{N}^\star$ et $(A_p)_{p=1,\ldots,n} \subset \mathcal{C}$ t.q. $A_p \cap A_q = \emptyset$ si $p \neq q$ et $A = \cup_{p=1}^n A_p$.

1. Montrer que $\mathcal{B}$ est stable par intersection finie et par passage au complémentaire.

2. Montrer que l'algèbre (cf. définition 2.4) engendrée par $\mathcal{C}$ est égale à $\mathcal{B}$.

Exercice 2.13 (Classes monotones) *Corrigé 18 page 291*
Soit E un ensemble. Pour $\Sigma \subset \mathcal{P}(E)$, on dit que Σ est une classe monotone (sur E) si Σ vérifie les deux propriétés suivantes (de stabilité par union croissante dénombrable et par intersection décroissante dénombrable) :

- $(A_n)_{n\in\mathbb{N}} \subset \Sigma$, $A_n \subset A_{n+1}$ pour tout $n \in \mathbb{N} \Rightarrow \cup_{n\in\mathbb{N}}A_n \in \Sigma$,
- $(A_n)_{n\in\mathbb{N}} \subset \Sigma$, $A_n \supset A_{n+1}$ pour tout $n \in \mathbb{N} \Rightarrow \cap_{n\in\mathbb{N}}A_n \in \Sigma$.

1. Soit $\Sigma \subset \mathcal{P}(E)$. Montrer que Σ est une tribu si et seulement si Σ est une classe monotone et une algèbre (cf. exercice 2.9).
2. Donner un exemple, avec $E = \mathbb{R}$, de classe monotone qui ne soit pas une tribu.
3. Soit $(\Sigma_i)_{i\in I}$ une famille de classes monotones (sur E). Montrer que $\cap_{i\in I}\Sigma_i = \{A \in \mathcal{P}(E);\ A \in \Sigma_i$ pour tout $i \in I\}$ est encore une classe monotone.

   Si $\mathcal{C} \subset \mathcal{P}(E)$, cette question permet donc de définir la classe monotone engendrée par $\mathcal{C}$ comme l'intersection de toutes les classes monotones sur E contenant $\mathcal{C}$.
4. (Lemme des classes monotones) Soit $\mathcal{A}$ une algèbre sur E. On note Σ la classe monotone engendrée par $\mathcal{A}$ et on note T la tribu engendrée par $\mathcal{A}$.
   (a) Montrer que $\Sigma \subset T$.
   (b) Soit $A \subset E$. On pose $\Sigma_A = \{B \subset E;\ A \setminus B \in \Sigma$ et $B \setminus A \in \Sigma\}$. Montrer que Σ_A est une classe monotone.
   (c) (Question plus difficile.) Montrer que Σ est une algèbre. [Utiliser la question (b) et la première question de l'exercice 2.9.] En déduire que $T = \Sigma$.

Exercice 2.14 (Caractérisation de la tribu engendrée) *Corrigé 19 page 293*
Soit E un ensemble et $\mathcal{A} \subset \mathcal{P}(E)$. On dit que $\mathcal{A}$ est stable par intersection finie si $A, B \in \mathcal{A} \Rightarrow A\cap B \in \mathcal{A}$. On dit que $\mathcal{A}$ est stable par différence si :

$$A, B \in \mathcal{A},\ B \subset A \Rightarrow A \setminus B = A \cap B^c \in \mathcal{A}.$$

On dit que $\mathcal{A}$ est stable par union dénombrable disjointe si :

$$(A_n)_{n\in\mathbb{N}} \subset \mathcal{A}, A_n \cap A_m = \emptyset \text{ pour } n \neq m \Rightarrow \cup_{n\in\mathbb{N}}A_n \in \mathcal{A}.$$

Soit $\mathcal{C} \subset \mathcal{P}(E)$.

1. On note $\mathcal{Z}$ l'ensemble des parties de $\mathcal{P}(E)$ stables par différence et stables par union dénombrable disjointe. Montrer qu'il existe $\mathcal{D} \in \mathcal{Z}$ t.q. $\mathcal{C} \subset \mathcal{D}$ et :

$$\mathcal{A} \in \mathcal{Z},\ \mathcal{C} \subset \mathcal{A} \Rightarrow \mathcal{D} \subset \mathcal{A}.$$

   Dans la suite, on note toujours $\mathcal{D}$ cette partie de $\mathcal{P}(E)$. On suppose maintenant que $\mathcal{C}$ est stable par intersection finie et que $E \in \mathcal{C}$.
2. Pour $A \in \mathcal{P}(E)$, on note $\mathcal{D}_A = \{D \in \mathcal{D}$ t.q. $A \cap D \in \mathcal{D}\}$.
   (a) Soit $A \in \mathcal{P}(E)$. Montrer que $\mathcal{D}_A$ est stable par union dénombrable disjointe et stable par différence.
   (b) Soit $A \in \mathcal{C}$. Montrer que $\mathcal{C} \subset \mathcal{D}_A$. En déduire que $\mathcal{D}_A = \mathcal{D}$.
   (c) Soit $A \in \mathcal{D}$. Montrer que $\mathcal{D}_A = \mathcal{D}$. En déduire que $\mathcal{D}$ est stable par intersection finie.
3. Montrer que $\mathcal{D}$ est une tribu. En déduire que $\mathcal{D}$ est la tribu engendrée par $\mathcal{C}$.

2.7.2 Mesures

Exercice 2.15 (Exemple de mesures) *Corrigé 20 page 295*
Soit E un ensemble infini non dénombrable. Pour toute partie A de E, on pose $m(A) = 0$ si A est au plus dénombrable, et $m(A) = +\infty$ sinon. L'application m est-elle une mesure sur $\mathcal{P}(E)$?

Exercice 2.16 (Mesure trace et restriction d'une mesure) *Corrigé 21 page 296*
Soit (E, T, m) un espace mesuré

1. Soit $F \in T$. Montrer que la tribu trace de T sur F, notée T_F, est incluse dans T (cette tribu est une tribu sur F). Montrer que la restriction de m à T_F est une mesure sur T_F. On l'appellera la *trace* de m sur F. Si $m(F) < \infty$, cette mesure est finie.

2. Soit $\mathcal{A}$ une tribu incluse dans T. La restriction de m à $\mathcal{A}$ est une mesure. Est-elle finie (resp. σ-finie) si m est finie (resp. σ-finie) ?

Exercice 2.17 *Corrigé 22 page 296*
Soit (E, T, m) un espace mesuré fini ("fini" signifie que $m(E) < \infty$) et $(A_n)_{n\in\mathbb{N}}$, $(B_n)_{n\in\mathbb{N}}$ des suites d'ensembles mesurables tels que $B_n \subset A_n$ pour tout $n \in \mathbb{N}$.

1. Montrer que $(\cup_{n\in\mathbb{N}} A_n) \setminus \cup_{n\in\mathbb{N}} B_n \subset \cup_{n\in\mathbb{N}} (A_n \setminus B_n)$.
2. Montrer que $m(\cup_{n\in\mathbb{N}} A_n) - m(\cup_{n\in\mathbb{N}} B_n) \le \sum_{n\in\mathbb{N}} (m(A_n) - m(B_n))$.

Exercice 2.18 *Corrigé 23 page 297*
Soit (E, T, m) un espace mesuré fini et $(A_n)_{n\in\mathbb{N}} \subset T$ t.q., pour tout $n \in \mathbb{N}$, $m(A_n) = m(E)$. Montrer que $m(\cap_{n\in\mathbb{N}} A_n) = m(E)$.

Exercice 2.19 (Sur la mesure d'une union...) *Corrigé 24 page 297*
Soit $(\Omega, \mathcal{A}, m)$ un espace mesuré et $n \in \mathbb{N}^\star$. Soit $A_1, \ldots, A_n \in \mathcal{A}$ et $B \in \mathcal{A}$. On suppose que $m(A_p) < \infty$ pour tout p. Montrer que $m(\cup_{p=1}^n (B \cap A_p)) = \sum_{k=1}^n (-1)^{k+1} \left(\sum_{1 \le i_1 < \ldots < i_k \le n} m(B \cap (\cap_{j=1}^k A_{i_j})) \right)$.

Exercice 2.20 (Contre exemples...) *Corrigé 25 page 297*

1. Soit λ la mesure de Lebesgue sur $\mathcal{B}(\mathbb{R})$ et $A \in \mathcal{B}(\mathbb{R})$ t.q. $\lambda(A) = 0$. A-t-on nécessairement A fermé ?

2. Soit (E, T) un espace mesurable et $\mathcal{C} \subset \mathcal{P}(E)$ qui engendre T. On considère m_1 et m_2 des mesures sur T. Montrer que $m_1(A) = m_2(A)$ pour tout $A \in \mathcal{C}$ n'implique pas que $m_1 = m_2$ sur T. [*On pourra trouver un exemple (facile) avec* $(E, T) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *et* m_1, m_2 *non finies. Un exemple avec* $(E, T) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *et* m_1, m_2 *finies est aussi possible mais plus difficile à trouver...*]

Exercice 2.21 (Résultat d'unicité) *Corrigé 26 page 298*
Soit (E, T) un espace mesurable et m, μ deux mesures sur T. Soit $\mathcal{C} \subset \mathcal{P}(E)$. On suppose que $\mathcal{C}$ engendre T et que $\mathcal{C}$ est stable par intersection finie.

On suppose que $m(A) = \mu(A)$ pour tout $A \in \mathcal{C}$.

1. On suppose que $E \in \mathcal{C}$ et que $m(E) < \infty$. Montrer que $m(A) = \mu(A)$ pour tout $A \in T$. [On pourra introduire $\mathcal{D} = \{A \in T,\ m(A) = \mu(A)\}$ et utiliser l'exercice 2.14.]

2. (Généralisation de la question précédente). On suppose qu'il existe une suite $(E_n)_{n\in\mathbb{N}} \subset \mathcal{C}$ t.q. $E_n \cap E_m = \emptyset$ si $n \neq m$, $m(E_n) < \infty$ pour tout $n \in \mathbb{N}$ et $E = \cup_{n\in\mathbb{N}} E_n$. Montrer que $m(A) = \mu(A)$ pour tout $A \in T$.

3. Avec $(E,T) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, donner un exemple pour lequel $E \in \mathcal{C}$ et $m \neq \mu$.

Exercice 2.22 (Existence d'une mesure, De l'algèbre à la σ-algèbre)
Soit Ω un ensemble, $\mathcal{F}_0$ une algèbre sur Ω et m une mesure sur $\mathcal{F}_0$ (c'est-à-dire que m est une application de $\mathcal{F}_0$ dans $\overline{\mathbb{R}}_+$, $m(\emptyset) = 0$ et $m(\cup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} m(A_n)$ pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{F}_0$ disjoints 2 à 2 et t.q. $\cup_{n\in\mathbb{N}} A_n \in \mathcal{F}_0$). On note $\mathcal{F} = \sigma(\mathcal{F}_0)$. Cette exercice montre qu'il est possible de prolonger m en une mesure sur $\mathcal{F}$.

Pour $A \subset \Omega$ on pose

$$m^\star(A) = \inf\{\sum_{n\in\mathbb{N}} m(A_n),\ (A_n)_{n\in\mathbb{N}} \subset \mathcal{F}_0, A \subset \cup_{n\in\mathbb{N}} A_n\}.$$

1. Montrer que $m^\star$ vérifie les 3 propriétés suivantes :
   - $m^\star(\emptyset) = 0$,
   - (monotonie de $m^\star$) pour tout $A, B \in \mathcal{P}(\Omega)$, $A \subset B \Rightarrow m^\star(A) \leq m^\star(B)$,
   - (σ-sous-additivité de $m^\star$) pour toute suite $(A_n)_{n\in\mathbb{N}} \subset \mathcal{P}(\Omega)$, $m^\star(\cup_{n\in\mathbb{N}} A_n) \leq \sum_{n\in\mathbb{N}} m^\star(A_n)$.

   N.B.: On dit que $m^\star$ est une mesure extérieure.

   Soit $A \in \mathcal{P}(\Omega)$.

   On dit que A est $m^\star$-mesurable si on a, pour tout $E \in \mathcal{P}(\Omega)$, $m^\star(E) = m^\star(E \cap A) + m^\star(E \cap A^c)$.

   On note $\mathcal{M}$ l'ensemble des parties de E $m^\star$-mesurables.

2. Soit $A \in \mathcal{P}(\Omega)$. Montrer que A est $m^\star$-mesurable si et seulement si on a, pour tout $E \in \mathcal{P}(\Omega)$ t.q. $m^\star(E) < \infty$, $m^\star(E) \geq m^\star(E \cap A) + m^\star(E \cap A^c)$.

3. Montrer que $\mathcal{M}$ est une algèbre. [On montrera que $\Omega \in \mathcal{M}$, puis que $A \cap B^c \in \mathcal{M}$ pour tout $A, B \in \mathcal{M}$.]

4. Montrer que $\mathcal{M}$ est une σ-algèbre. [On pourra montrer, par exemple, que $\mathcal{M}$ est stable par union dénombrable.]

5. Montrer que la restriction de $m^\star$ à $\mathcal{M}$ est une mesure.

6. Montrer que $\mathcal{F}_0 \subset \mathcal{M}$ et que $m^\star = m$ sur $\mathcal{F}_0$. En déduire que $\mathcal{F} \subset \mathcal{M}$ et que la restriction de $m^\star$ à $\mathcal{F}$ est une mesure sur $\mathcal{F}$ prolongeant m.

Exercice 2.23 (Un pas vers l'unicité d'une mesure)
Soit Ω un ensemble, $\mathcal{F}$ une tribu sur Ω et μ_1, μ_2 deux mesures sur $\mathcal{F}$. Soit $A \in \mathcal{F}$ t.q. $\mu_1(A) = \mu_2(A) < \infty$. On pose $\mathcal{L} = \{B \in \mathcal{F}$ t.q. $\mu_1(A \cap B) = \mu_2(A \cap B)\}$. Montrer que $\mathcal{L}$ est un λ-système (c'est-à-dire que $\mathcal{L}$ est stable par union dénombrable croissante, $\Omega \in \mathcal{F}$ et $B \setminus C \in \mathcal{F}$ si $B, C \in \mathcal{F}$ avec $C \subset B$).

Exercice 2.24 (Mesure atomique, mesure diffuse) *Corrigé 27 page 299*

Soit (E,T) un espace mesurable t.q. $\{x\} \in T$ pour tout $x \in E$. Une mesure m sur T est diffuse si $m(\{x\}) = 0$ pour tout $x \in E$. Une mesure m sur T est purement atomique si il existe $S \in T$ t.q. $m(S^c) = 0$ et $m(\{x\}) > 0$ si $x \in S$.

1. Montrer qu'une mesure purement atomique et diffuse est nulle. Donner, pour $(E,T) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ un exemple de mesure purement atomique et un exemple de mesure diffuse. [Montrer que la mesure de Lebesgue sur $\mathcal{B}(\mathbb{R})$ est diffuse.]
2. Soit m une mesure diffuse sur T. Montrer que tous les ensembles dénombrables sont de mesure nulle.
3. Soit m une mesure sur T. On suppose que m est σ-finie, c'est à dire qu'il existe $(E_n)_{n\in\mathbb{N}} \subset T$ t.q. $E = \cup_{n\in\mathbb{N}} E_n$ et $m(E_n) < +\infty$ pour tout $n \in \mathbb{N}$.
   (a) Montrer que l'ensemble des $x \in E$ t.q. $m(\{x\}) > 0$ (de tels x sont appelés "atomes" de m) est au plus dénombrable. [On pourra introduire l'ensemble $A_{n,k} = \{x \in E_n;\, m(\{x\}) \geq \frac{1}{k}\}$.]
   (b) Montrer qu'il existe une mesure diffuse m_d et une mesure purement atomique m_a sur T telles que $m = m_d + m_a$. Montrer que m_d et m_a sont étrangères, c'est à dire qu'il existe $A \in T$ t.q. $m_d(A) = 0$ et $m_a(A^c) = 0$.
   (c) Montrer que si m est finie il existe un singleton dont la mesure est supérieure ou égale à la mesure de tous les autres singletons. Montrer que ceci peut-être inexact si m n'est que σ-finie.
4. Pour $(E,T) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, donner un exemple de mesure purement atomique finie dont l'ensemble des atomes est infini.

Exercice 2.25 (limites sup et inf d'ensembles) *Corrigé 28 page 301*
Soit (E,T,m) un espace mesuré et $(A_n)_{n\in\mathbb{N}} \subset T$. On rappelle que $\limsup_{n\to\infty} A_n = \cap_{n\in\mathbb{N}} \cup_{p\geq n} A_p$ et $\liminf_{n\to\infty} A_n = \cup_{n\in\mathbb{N}} \cap_{p\geq n} A_p$.

1. On suppose qu'il existe $n_0 \in \mathbb{N}$ t.q. $m(\cup_{p\geq n_0} A_p) < \infty$. Montrer que $m(\liminf_{n\to\infty} A_n) \leq \liminf_{n\to\infty} m(A_n) \leq \limsup_{n\to\infty} m(A_n) \leq m(\limsup_{n\to\infty} A_n)$.
2. Donner un exemple (c'est-à-dire choisir (E,T,m) et $(A_n)_{n\in\mathbb{N}} \subset T$) pour lequel :
$$\limsup_{n\to\infty} m(A_n) > m(\limsup_{n\to\infty} A_n).$$
3. Donner un exemple avec m finie (c'est-à-dire $m(E) < \infty$) pour lequel
$m(\liminf_{n\to\infty} A_n) < \liminf_{n\to\infty} m(A_n) < \limsup_{n\to\infty} m(A_n) < m(\limsup_{n\to\infty} A_n)$.
4. $(\star)$ (Lemme de Borel-Cantelli) On suppose que $\sum_{n\in\mathbb{N}} m(A_n) < \infty$.
Montrer que $m(\limsup_{n\to\infty} A_n) = 0$.

Exercice 2.26 (Petit ouvert dense. . .) *Corrigé 29 page 303*
On considère ici l'espace mesuré $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$. Soit $\varepsilon > 0$, peut-on construire un ouvert dense dans $\mathbb{R}$ de mesure inférieure à ε ? [On rappelle qu'une partie A de R est dense dans $\mathbb{R}$ si $\overline{A} = \mathbb{R}$ ou encore si, pour tout $x \in \mathbb{R}$ et pour tout $\varepsilon > 0$, il existe $a \in A$ t.q. $|x - a| < \varepsilon$.]

Exercice 2.27 (Non existence d'une mesure sur $\mathcal{P}(\mathbb{R})$ exprimant la longueur) *Corrigé 30 page 303*

On définit la relation d'équivalence sur $[0,1[$: xRy si $x-y \in \mathbb{Q}$. En utilisant l'axiome du choix, on construit un ensemble $A \subset [0,1[$ tel que A contienne un élément et un seul de chaque classe d'équivalence. Pour $q \in \mathbb{Q} \cap [0,1[$, on définit $A_q = \{y \in [0,1[;\ y = x+q \text{ ou } y = x+q-1, x \in A\}$, c'est-à-dire $A_q = \{y \in [0,1[;\ y-q \in A \text{ ou } y-q+1 \in A\}$.

1. Montrer que $\bigcup_{q \in \mathbb{Q} \cap [0,1[} A_q = [0,1[$.

2. Montrer que si m est une application de $\mathcal{P}(\mathbb{R})$ dans $\overline{\mathbb{R}}_+$, invariante par translation et vérifiant $m([0,1[) = 1$, m ne peut pas être σ- additive. En déduire la non-existence d'une mesure m, sur $\mathcal{P}(\mathbb{R})$, invariante par translation et t.q. $m([a,b]) = b-a$ pour tout $a, b \in \mathbb{R}$, $a < b$. En particulier, montrer que l'application $\lambda^\star$, définie en cours, ne peut pas être une mesure sur $\mathcal{P}(\mathbb{R})$.

Exercice 2.28 (Non existence d'une mesure sur $\mathcal{P}(\mathbb{R})$ donnant la longueur des intervalles)
Cet exercice est plus général que le précédent car on veut montrer qu'il n'existe pas de mesure sur $\mathcal{P}(\mathbb{R})$ t.q. $m([a,b]) = b-a$ pour tout $a, b \in \mathbb{R}$, $a < b$, sans l'hypothèse d'invariance par translation de l'exercice précédent.

Soit E un ensemble non dénombrable, sur lequel on suppose qu'il existe un ordre total, noté $\leq$, tel que pour tout $x \in E$, l'ensemble $\{y \in E; y \leq x\}$ est dénombrable, c'est-à-dire qu'il existe une application f_x injective de cet ensemble dans $\mathbb{N}$. Si $E = \mathbb{R}$ ou $E = [0,1]$, on peut démontrer l'existence d'un tel ordre (ceci est une conséquence de l'axiome du continu). Soit m une mesure sur $\mathcal{P}(E)$; on suppose que m est finie, i.e. $m(E) < +\infty$, et diffuse. On se propose de montrer que m est nulle, i.e. $m(A) = 0$, pour tout $A \in \mathcal{P}(E)$. On pose, pour $x \in E$ et $n \in \mathbb{N}$, $A_{x,n} = \{y \in E; y \geq x \text{ et } f_y(x) = n\}$.

1. Montrer que pour tout $x, y \in E$ et $n \in \mathbb{N}$, $A_{x,n} \cap A_{y,n} = \emptyset$. En déduire que, pour tout $n \in \mathbb{N}$, $\{x \in E; m(A_{x,n}) \neq 0\}$ est au plus dénombrable (utiliser le fait que m est finie).

2. Montrer qu'il existe $x \in E$ t.q., pour tout $n \in \mathbb{N}$, $m(A_{x,n}) = 0$.

3. En déduire que m est nulle (montrer pour cela que $m(E) = 0$ en utilisant la question précédente et le fait que m est diffuse).

4. Montrer qu'il n'existe pas de mesure m sur $\mathcal{P}(\mathbb{R})$ t.q. $m(]a,b[) = b-a$ pour tout a, $b \in \mathbb{R}$, $a < b$.

Exercice 2.29 *Corrigé 31 page 304*
Soit m une mesure sur $\mathcal{B}(\mathbb{R})$ t.q. pour tout intervalle I et tout $x \in \mathbb{R}$ on ait $m(I) = m(I+x)$ (avec $I + x = \{a+x,\ a \in I\}$) et $m([0,1]) = 1$. Montrer que pour tout $x \in \mathbb{R}$, $m(\{x\}) = 0$ (i.e. m est diffuse). En déduire que m est la mesure de Lebesgue sur $\mathcal{B}(\mathbb{R})$. [On pourra découper $[0,1[$ en q intervalles de longueur $1/q$.]

Exercice 2.30 (Support d'une mesure sur les boréliens de $\mathbb{R}^d$) *Corrigé 32 page 305*
Soit m une mesure sur $\mathcal{B}(\mathbb{R}^d)$. Montrer qu'il existe un plus grand ouvert de mesure nulle pour m. L'ensemble fermé complémentaire de cet ouvert s'appelle *le support* de m. [On pourra, par exemple, considérer les pavés à extrémités rationnelles qui sont de mesure nulle pour m.]

Exercice 2.31 (Ensemble de Cantor) *Corrigé 33 page 305*
On considère l'espace mesuré $([0,1], \mathcal{B}([0,1]), \lambda)$.
On pose $C_0 = [0,1]$, $a_1^0 = 0$, $b_1^0 = 1$, et $\alpha_0 = 1$. Pour $n \geq 0$, on construit $C_{n+1} \subset [0,1]$ de la manière suivante : on suppose $C_n = \bigcup_{p=1}^{2^n} [a_p^n, b_p^n]$ connu, et on définit $C_{n+1} = \bigcup_{p=1}^{2^{n+1}} [a_p^{n+1}, b_p^{n+1}]$ où, pour $p = 1, \ldots, 2^n$, $a_{2p-1}^{n+1} = a_p^n$, $b_{2p-1}^{n+1} = a_p^n + \alpha_{n+1}$, $a_{2p}^{n+1} = b_p^n - \alpha_{n+1}$ et $b_{2p}^{n+1} = b_p^n$, avec $\alpha_{n+1} = \frac{\rho_n \alpha_n}{2}$, et $0 < \rho_n < 1$. On pose $C = \cap_{n \geq 0} C_n$ (C s'appelle "ensemble de Cantor", l'exemple le plus classique est obtenu avec $\rho_n = \frac{2}{3}$ pour tout $n \in \mathbb{N}$).

1. Montrer que $C_{n+1} \subset C_n$.

2. Montrer que C est compact et $\overset{\circ}{C} = \emptyset$.

3. (Question plus difficile.) Montrer que C est non dénombrable.

4. Montrer que si ρ_n ne dépend pas de n, alors $\lambda(C) = 0$. En déduire que si $A \in \mathcal{B}([0,1])$, $\lambda(A) = 0$ n'entraîne pas que A est dénombrable.

5. Soit $0 < \epsilon < 1$. Montrer qu'il existe une suite $(\rho_n)_{n \geq 0} \subset]0,1[$ t.q. $\lambda(C) = \epsilon$.

6. Soit f lipschitzienne de $\mathbb{R}$ dans $\mathbb{R}$. Montrer que si A est un compact de $[0,1]$ t.q. $\lambda(A) = 0$, alors $f(A)$ est un compact de $\mathbb{R}$ t.q. $\lambda(f(A)) = 0$.

7. Construire une fonction continue de $\mathbb{R}$ dans $\mathbb{R}$ t.q. si A est un compact de $[0,1]$ t.q. $\lambda(A) = 0$, on n'a pas forcément $\lambda(f(A)) = 0$ (mais $f(A)$ est un compact de $\mathbb{R}$). [Utiliser un ensemble de Cantor de mesure nulle (cf question 4) et un ensemble de Cantor de mesure $\epsilon > 0$ (cf question 5).]

Exercice 2.32 (Mesure complète) *Corrigé 34 page 308*
Soit (E, T, m) un espace mesuré. Une partie B de E est dite "négligeable" si elle est incluse dans un élément de T de mesure nulle. On note $\mathcal{N}_m$ l'ensemble des parties négligeables. On pose $\overline{T} = \{A \cup N;\ A \in T,\ N \in \mathcal{N}_m\}$.

1. Montrer que $\overline{T}$ est une tribu et que $T \cup \mathcal{N}_m \subset \overline{T}$.

2. Soit $A_1,\ A_2 \in T$ et $N_1,\ N_2 \in \mathcal{N}_m$ t.q. $A_1 \cup N_1 = A_2 \cup N_2$. Montrer que $m(A_1) = m(A_2)$.

   Pour $B \in \overline{T}$, soit $A \in T$ et $N \in \mathcal{N}_m$ t.q. $B = A \cup N$, on pose $\overline{m}(B) = m(A)$. (La question précédente montre que cette définition est cohérente.)

3. Montrer que $\overline{m}$ est une mesure sur $\overline{T}$ et $\overline{m}_{|T} = m$. Montrer que $\overline{m}$ est la seule mesure sur $\overline{T}$ égale à m sur T.

4. Montrer que $\mathcal{N}_{\overline{m}} = \mathcal{N}_m \subset \overline{T}$.

L'exercice 4.18 page 104 montre la différence "dérisoire", du point de vue de l'intégration, entre (E, T, m) et son complété $(E, \overline{T}, \overline{m})$.

Exercice 2.33 (Série commutativement convergente dans $\mathbb{R}$) *Corrigé 35 page 310*
Soit $(a_n)_{n \in \mathbb{N}} \subset \mathbb{R}$. Le but de l'exercice est de montrer que si la série $\sum_{n \in \mathbb{N}} a_{\varphi(n)}$ est convergente pour toute bijection $\varphi\ : \mathbb{N} \to \mathbb{N}$, alors la série $\sum_{n \in \mathbb{N}} a_n$ est absolument convergente.
Pour montrer ce résultat, on suppose, par exemple, que $\sum_{n \in \mathbb{N}} a_n^+ = \infty$. Montrer qu'il existe $\varphi\ :\ \mathbb{N} \to \mathbb{N}$, bijective, t.q. $\sum_{p=0}^{n} a_{\varphi(p)} \to \infty$ quand $n \to \infty$. Conclure.

Exercice 2.34 (Régularité d'une mesure sur $\mathcal{B}(\mathbb{R})$, finie sur les bornés)
Cet exercice redémontre le théorème de régularité (théorème 2.3).

I. Soit m une mesure finie sur $\mathcal{B}(\mathbb{R})$, tribu borélienne sur $\mathbb{R}$. On pose: $T = \{A \in \mathcal{B}(\mathbb{R})$ tel que $\forall \varepsilon > 0, \exists O$ ouvert de $\mathbb{R}$, et F fermé de $\mathbb{R}$, tels que $F \subset A \subset O$ et $m(O \setminus F) \leq \varepsilon\}$.

   1. Soient a et $b \in \mathbb{R}$, $a < b$. Montrer que $]a, b[\in T$.

2. Montrer que T est une tribu. En déduire que m est régulière.
3. En déduire que, $\forall A \in \mathcal{B}(\mathbb{R})$, $m(A) = \inf\{m(O), O \supset A, O$ ouvert de $\mathbb{R}\}$.

II. Soit m une mesure sur $\mathcal{B}(\mathbb{R})$, tribu borélienne sur $\mathbb{R}$. On suppose que pour toute partie B bornée de $\mathbb{R}$ t.q. $B \in \mathcal{B}(\mathbb{R})$, on a: $m(B) < +\infty$.

Pour $B \in \mathcal{B}(\mathbb{R})$, on pose: $\nu_B(A) = m(A \cap B)$, $\forall A \in \mathcal{B}(\mathbb{R})$.

1. Soit $B \in \mathcal{B}(\mathbb{R})$ une partie bornée de $\mathbb{R}$; montrer que ν_B est une mesure finie.
2. Soient $\varepsilon > 0$ et $A \in \mathcal{B}(\mathbb{R})$, montrer que, pour tout $n \in \mathbb{Z}$, il existe un ouvert O_n de $\mathbb{R}$ tel que $O_n \supset A \cap]n, n+1]$ et $m(O_n) \le m(A \cap]n, n+1]) + \dfrac{\varepsilon}{2^{|n|}}$. [*Appliquer* I.3 *avec* ν_{B_n}, B_n *ouvert borné contenant* $]n, n+1]$, *et* $A \cap]n, n+1]$.]
3. a) Soient $\varepsilon > 0$ et $A \in \mathcal{B}(\mathbb{R})$. Montrer que pour tout $n \in \mathbb{Z}$, il existe O_n ouvert de $\mathbb{R}$ et F_n fermé de $\mathbb{R}$ t.q. $F_n \subset A \cap]n, n+1] \subset O_n$ et $m(O_n \setminus F_n) \le \frac{\varepsilon}{2^{|n|}}$.
   b) Montrer que m est régulière. [*On pourra remarquer, en le démontrant, que si* F_n *est fermé et* $F_n \subset]n, n+1] \forall n \in \mathbb{Z}$, *alors* $\bigcup_{n \in \mathbb{Z}} F_n$ *est fermé.*]
   c) Montrer que $m(A) = \inf\{m(O), A \subset O, O$ ouvert de $\mathbb{R}\}$, $\forall A \in \mathcal{B}(\mathbb{R})$.

III. On note λ la mesure de Lebesgue sur $(\mathbb{R}, \mathcal{R})$. Soit $\alpha \in \mathbb{R}^\star$ et $\beta \in \mathbb{R}$. Soit $A \in \mathcal{R}$. On pose $\alpha A + \beta = \{\alpha x + \beta, \, x \in A\}$. Montrer que $\alpha A + \beta \in \mathcal{R}$ et que $\lambda(\alpha A + \beta) = |\alpha| \lambda(A)$. [*On pourra commencer par étudier le cas où* A *est un ouvert de* $\mathbb{R}$.] (Nous appellerons cette propriété : "invariance par translation généralisée pour la mesure de Lebesgue".)

Exercice 2.35 (Mesure sur S^1) *Corrigé 36 page 311*

On considère $S^1 = \{(x, y)^t \in \mathbb{R}^2, |x|^2 + |y|^2 = 1\}$ (S^1 est donc le cercle unité de $\mathbb{R}^2$). Pour $z = (x, y)^t \in S^1$, il existe un unique $\theta_z \in [0, 2\pi[$ t.q. $x = \cos(\theta_z)$ et $y = \sin(\theta_z)$. Pour $\alpha \in [0, 2\pi[$ et $z \in S^1$ on pose $R_\alpha(z) = (\cos(\theta_z + \alpha), \sin(\theta_z + \alpha))^t$. Noter que R_α est une bijection de S^1 sur S^1 (c'est la rotation d'angle α).

Définir une tribu T sur S^1, t.q. T contienne les parties de la forme $\{(\cos(\theta), \sin(\theta))^t, \ \theta \in]\alpha, \beta[\}$ avec $-\infty < \alpha < \beta < \infty$, et une mesure μ sur T de sorte que (S^1, T, μ) soit un espace mesuré avec $\mu(S^1) = 1$ et t.q. μ soit invariante par rotation (c'est à dire que, pour tout $A \in T$ et $\alpha \in [0, 2\pi[$, on ait $R_\alpha(A) = \{R_\alpha(z), z \in A\} \in T$ et $\mu(R_\alpha(A)) = \mu(A)$). [On pourra utiliser la tribu borélienne de $\mathbb{R}$, notée $\mathcal{B}(\mathbb{R})$, et la mesure de Lebesgue sur $\mathcal{B}(\mathbb{R})$.]

2.7.3 Probabilités

Exercice 2.36 (Exemple de probabilité)

Soit $E = \{x_k, k \in \mathbb{N}\}$ un ensemble infini dénombrable et $(p_k)_{k \in \mathbb{N}} \subset [0, 1]^{\mathbb{N}}$ t.q. $p_k \ge 0 \, \forall k \in \mathbb{N}$ et $\sum_{k \in \mathbb{N}} p_k = 1$.

1. Montrer que, pour tout $A \in \mathcal{P}(E)$, $A \neq \emptyset$, on peut définir $p(A) = \sum_{k; x_k \in A} p_k$. On pose $p(\emptyset) = 0$.
2. Montrer que p définie en 1. est une probabilité

Exercice 2.37 (Lemme de Borel-Cantelli) *Corrigé 37 page 312*

Soient (E, T, p) un espace probabilisé et $(A_n)_{n \in \mathbb{N}} \subset T$. On pose $B_n = \cup_{k \ge n} A_k$ et $A = \cap_{n \in \mathbb{N}} B_n$ (on rappelle que $A = \limsup_{n \to \infty} A_n$).

1. Montrer que si $\sum_{n\in\mathbb{N}} p(A_n) < +\infty$ alors $p(A) = 0$.
2. On suppose que, pour tout $n \in \mathbb{N}^\star$, les événements $A_1, \ldots, A_n$ sont indépendants. On suppose aussi que $\sum_{n\in\mathbb{N}} p(A_n) = \infty$. Montrer que $p(A) = 1$.

Exercice 2.38 Soient E une “population”, c'est-à-dire un ensemble fini,et $(C_n)_{n\in\mathbb{N}}$ une famille de ”sous-populations”, c'est-à-dire un système de constituants de l'espace probabilisable $(E, \mathcal{P}(E))$. Soit P_n la probabilité qu'un individu de la population appartienne à la sous-population C_n, c'est-à-dire $p(C_n)$, où p est une probabilité sur $(E, \mathcal{P}(E))$. Sachant que dans chaque sous-population la probabilité d'être gaucher est Q_n, trouver la probabilité qu'un gaucher appartienne à C_n.

Exercice 2.39 Soient S_1 (resp. S_2) un seau contenant n_1 cailloux noirs et b_1 cailloux blancs (resp. n_2 cailloux noirs et b_2 cailloux blancs). On tire au hasard, de manière équiprobable, un des deux seaux, et on tire ensuite au hasard, de manière équiprobable, un caillou dans ce seau. Sachant qu'on a tiré un caillou noir, quelle est la probabilité de l'avoir tiré du seau S_1 ?

Exercice 2.40 Soient p une probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ et F la fonction de répartition de p. Montrer que F est continue ssi $p(\{a\}) = 0$. En déduire que F est continue si F ne charge pas les points.

Chapter 3

Fonctions mesurables, variables aléatoires

3.1 Introduction, topologie sur $\overline{\mathbb{R}}_+$

Nous allons, dans ce chapitre, introduire différents outils nécessaires à la définition de l'intégrale de Lebesgue. De la même manière que les fonctions en escalier ont été introduites lors de la définition de l'intégrale des fonctions réglées, nous introduisons maintenant le concept de fonction étagée sur un espace mesurable (E, T). Nous introduirons ensuite les concepts de fonction mesurable et de variable aléatoire, ainsi que les premières notions de convergence de ces fonctions. La notion de variable aléatoire est fondamentale en calcul des probabilités: c'est en général par la connaissance de la variable aléatoire (et par sa "loi de probabilité") que se construit le modèle probabiliste, l'espace probabilisé (E, T, p) restant souvent mal connu.

Remarque 3.1

1. L'objectif est d'intégrer des fonctions de E (espace de départ) dans F (espace d'arrivée). Pour construire ainsi une notion d'intégrale, il faut un espace mesuré au départ et un espace topologique à l'arrivée, car nous aurons besoin dans l'espace d'arrivée d'une notion de convergence (pour les procédés de passage à la limite dans la définition de l'intégrale). Les espaces d'arrivée usuels sont (pour la théorie de l'intégration) $\mathbb{R}$, $\mathbb{C}$, $\mathbb{R}^N$ ou un espace de Banach. Le procédé de construction dû à Lebesgue donne un rôle fondamental aux fonctions à valeurs dans $\overline{\mathbb{R}}_+$ (et à la notion de convergence "croissante") et nous aurons besoin d'utiliser la topologie de $\overline{\mathbb{R}}_+$ (voir la définition 3.1).

2. On rappelle qu'un espace topologique est la donnée d'un ensemble F muni d'une famille de parties de F, appelées "ouverts de F", contenant $\emptyset$ et F, stable par union (quelconque) et stable par intersection finie. On rappelle aussi que, dans un espace topologique, $x_n \to x$, quand $n \to \infty$, signifie que, pour tout O ouvert contenant x, il existe n_0 t.q. $x_n \in O$ pour tout $n \geq n_0$.

3. Soit F un espace topologique et $G \subset F$. On appelle topologie trace sur G, ou topologie induite sur G, la topologie définie par l'ensemble des restrictions à G des ouverts de F. Si $O \subset G$, O est un ouvert de G si et seulement si il existe U ouvert de F t.q. $O = U \cap G$. Noter donc que O peut ne pas être un ouvert de F si G n'est pas un ouvert de F. Par contre, il est important de remarquer que si G est un borélien de F (c'est-à-dire $G \in \mathcal{B}(F)$, $\mathcal{B}(F)$ étant la tribu engendrée par les ouverts

de F), l'ensemble des boréliens de G est exactement l'ensemble des boréliens de F inclus dans G, c'est-à-dire $\mathcal{B}(G) = \{B \subset G;\ B \in \mathcal{B}(F)\}$, ceci est démontré dans l'exercice 2.3 page 41.

4. Un exemple fondamental de topologie sur l'ensemble F est celui de la topologie donnée par une distance sur F. Dans le cas de $F = \mathbb{R}$, nous considérerons toujours $\mathbb{R}$ muni de la topologie donnée par la structure métrique de $\mathbb{R}$, c'est-à-dire par l'application "distance" définie par $d(a, b) = |b - a|$.

Définition 3.1 (Topologie et tribu de Borel sur $\overline{\mathbb{R}}_+$) $\overline{\mathbb{R}}_+ = \mathbb{R}_+ \cup \{\infty\}$

1. *Soit $O \subset \overline{\mathbb{R}}_+$. O est un ouvert si pour tout $a \in O$ on a :*
   - *(a) Si $0 < a < \infty$, alors il existe $\varepsilon \in \mathbb{R}_+^\star$ t.q. $]a - \varepsilon, a + \varepsilon[\subset O$,*
   - *(b) si $a = 0$, alors il existe $\alpha \in \mathbb{R}_+^\star$ t.q. $[0, \alpha[\subset O$,*
   - *(c) si $a = \infty$, alors il existe $\alpha \in \mathbb{R}_+$ t.q. $]\alpha, \infty] \subset O$.*
2. *$\mathcal{B}(\overline{\mathbb{R}}_+)$ est la tribu (sur $\overline{\mathbb{R}}_+$) engendrée par les ouverts de $\overline{\mathbb{R}}_+$. Soit $B \subset \overline{\mathbb{R}}_+$, on peut montrer (voir la remarque 3.2 ci après) que $B \in \mathcal{B}(\overline{\mathbb{R}}_+)$ si et seulement si $B \cap \mathbb{R} \in \mathcal{B}(\mathbb{R})$ ($\mathcal{B}(\mathbb{R})$ est la tribu de Borel sur $\mathbb{R}$).*

Remarque 3.2

1. La topologie sur $\mathbb{R}_+$ est la topologie induite par celle de $\overline{\mathbb{R}}_+$, c'est aussi la topologie induite par celle de $\mathbb{R}$. L'ensemble des boréliens de $\mathbb{R}_+$ est donc égal à l'ensemble des boréliens de $\overline{\mathbb{R}}_+$ inclus dans $\mathbb{R}_+$ et c'est aussi l'ensemble des boréliens de $\mathbb{R}$ inclus dans $\mathbb{R}_+$ (voir la remarque 3.1). On remarque aussi que $\{\infty\} \in B(\overline{\mathbb{R}}_+)$ (car $\{\infty\}$ est, par exemple, une intersection dénombrable d'ouverts de $\overline{\mathbb{R}}_+$). On en déduit que, si $B \subset \overline{\mathbb{R}}_+$, on a $B \in \mathcal{B}(\overline{\mathbb{R}}_+)$ si et seulement si $B \cap \mathbb{R} \in \mathcal{B}(\mathbb{R})$ (noter que $B \cap \mathbb{R} = B \cap \mathbb{R}_+$).
2. Soit $A \subset \overline{\mathbb{R}}_+$, A est donc un borélien de $\overline{\mathbb{R}}_+$ si et seulement si $A \in \mathcal{B}(\mathbb{R})$ ou si $A = B \cup \{\infty\}$, avec $B \in \mathcal{B}(\mathbb{R})$.
3. La définition de la topologie sur $\overline{\mathbb{R}}_+$ donne bien que, pour $(x_n)_{n\in\mathbb{N}} \subset \overline{\mathbb{R}}_+$, on a $x_n \to \infty$ (dans $\overline{\mathbb{R}}_+$, quand $n \to \infty$) si et seulement si, pour tout $\alpha > 0$, il existe n_0 t.q. $x_n \in]\alpha, \infty]$ pour tout $n \geq n_0$ (ce qui est la définition usuelle de convergence vers ∞).
4. On peut aussi montrer (exercice…) que $\mathcal{B}(\overline{\mathbb{R}}_+)$ est la tribu engendrée par $\mathcal{C}_1 = \{]a, \infty];\ a \in \mathbb{R}_+\}$. C'est aussi la tribu engendrée par $\mathcal{C}_2 = \{]a, \infty[\cap \mathbb{R}_+;\ a \in \mathbb{R}\}$. Par contre, ce n'est pas la tribu engendrée (sur $\overline{\mathbb{R}}_+$) par $\mathcal{C}_3 = \{]a, \infty[;\ a \in \mathbb{R}_+\}$ (on a donc $T(\mathcal{C}_3) \subset \mathcal{B}(\overline{\mathbb{R}}_+)$ et $T(\mathcal{C}_3) \neq \mathcal{B}(\overline{\mathbb{R}}_+)$).

3.2 Fonctions étagées

Définition 3.2 (Fonction caractéristique) *Soient (E, T) un espace mesurable et soit $A \in T$. On appelle fonction caractéristique mesurable de l'ensemble A, et on note 1_A (ou χ_A) la fonction définie par : $1_A(x) = 1$ si $x \in A$ et $1_A(x) = 0$ si $x \in A^c$.*

Définition 3.3 (Fonction étagée) *Soient (E, T) un espace mesurable et f ; $E \to \mathbb{R}$.*

1. *On dit que f est étagée (ou T-étagée) si f est une combinaison linéaire (finie) de fonctions caractéristiques mesurables, c'est-à-dire s'il existe une famille finie $(A_i)_{i=1,\ldots,n} \subset T$ et n réels $a_1, ..., a_n$ tels que $f = \sum_{i=1}^{n} a_i 1_{A_i}$.*

2. *On dit que f est étagée positive si f est étagée et prend ses valeurs dans $\mathbb{R}_+$.*

On note $\mathcal{E}$ l'ensemble des fonctions étagées et $\mathcal{E}_+$ l'ensemble des fonctions étagées positives.

La notion de fonction étagée positive va nous permettre de définir l'intégrale à partir de la notion de mesure. On se limite pour l'instant aux fonctions positives afin de donner un sens à l'addition de mesures infinies. Notons que, dans la définition d'une fonction étagée, les ensembles A_i peuvent être d'intersection non vide. On aura besoin, pour introduire facilement la notion d'intégrale d'une fonction étagée positive, de considérer une décomposition de la fonction étagée sur des ensembles d'intersection vide. C'est l'objet du lemme suivant:

Lemme 3.1 (Décomposition canonique d'une fonction étagée positive)
Soit (E,T) un espace mesurable, et soit $f \in \mathcal{E}_+$ une fonction étagée positive, non identiquement nulle. Alors il existe une unique famille finie $(a_i, A_i)_{i=1,\ldots,n} \subset \mathbb{R}_+^\star \times T$ t.q. $0 < a_1 < \ldots < a_n$, $A_i \neq \emptyset$, pour tout i, $A_i \cap A_j = \emptyset$, si $i \neq j$, et $f = \sum_{i=1}^n a_i 1_{A_i}$.

Démonstration : Soient $(B_i)_{i=1,\ldots,p} \subset T$, $(b_i)_{i=1,\ldots,p} \subset \mathbb{R}$ et $f = \sum_{i=1}^p b_i 1_{B_i}$ une fonction étagée positive non nulle. L'ensemble $\mathrm{Im} f$ des valeurs prises par f est donc fini. Comme $\mathrm{Im} f \subset \mathbb{R}_+$, on a donc $\mathrm{Im} f \setminus \{0\} = \{a_1, \ldots, a_n\}$ avec $0 < a_1, \ldots < a_n$. En posant $A_i = \{x \in E; f(x) = a_i\}$, on a donc $f = \sum_{i=1}^n a_i 1_{A_i}$ avec $A_i \neq \emptyset$ et $A_i \cap A_j = \emptyset$. (Noter aussi que $\{x \in E; f(x) = 0\} = (\cup_{i=1}^n A_i)^c$.) Il reste à montrer que $A_i \in T$. Pour $i \in \{1, \ldots, n\}$, on pose $I_i = \{K \subset \{1, \ldots, p\}\ ;\ a_i = \sum_{k \in K} b_k\}$. On a alors, pour tout $i \in \{1, \ldots, n\}$, $A_i = \cup_{K \in I_i} C_K$, avec $C_K = \cap_{j=1}^p D_j$, $D_j = B_j$ si $j \in K$ et $D_j = B_j^c$ si $j \notin K$. Les propriétés de stabilité d'une tribu nous donnent alors que $A_i \in T$ pour tout $i \in \{1, \ldots, n\}$. On a donc trouvé la décomposition voulue de f. Le fait que cette décomposition est unique est immédiat car on a nécessairement $\{a_1, \ldots, a_n\} = \mathrm{Im} f \setminus \{0\}$ et $A_i = \{x \in E; f(x) = a_i\}$. ■

On aurait envie à partir de la notion de fonction étagée positive décomposée sous la forme précédente, de définir l'intégrale de f comme $\int f dm = \sum_{i=1}^n a_i m(A_i)$.
En fait, on pourra même (cf définition 4.1) définir l'intégrale d'une fonction étagée avec une décomposition plus générale (non unique) grâce au lemme suivant :

Lemme 3.2 *Soit (E,T,m) un espace mesuré et soit $f \in \mathcal{E}_+$ une fonction étagée positive non nulle, t.q. $f = \displaystyle\sum_{i=1}^n a_i 1_{A_i}$ et $f = \displaystyle\sum_{i=1}^p b_i 1_{B_i}$ où $a_1, ..., a_n$, $b_1, ..., b_p$ sont des réels strictement positifs, $(A_i)_{i=1,\ldots,n} \subset T$ et $(B_i)_{i=1,\ldots,p} \subset T$ sont des familles de parties disjointes deux à deux, i.e. telles que $A_i \cap A_j = \emptyset$ et $B_i \cap B_j = \emptyset$ si $i \neq j$. Alors :*

$$\sum_{i=1}^n a_i m(A_i) = \sum_{j=1}^p b_j m(B_j). \tag{3.1}$$

Démonstration : On pose, pour $i = 1, \ldots, n$ et $j = 1, \ldots, p$, $C_{ij} = A_i \cap B_j$. En remarquant que $\{x; f(x) > 0\} = \cup_{i=1}^n A_i = \cup_{j=1}^p B_j$, on écrit $A_i = \cup_{j=1}^p C_{ij}$ et $B_j = \cup_{i=1}^n C_{ij}$. On peut donc écrire

$$\sum_{i=1}^n a_i m(A_i) = \sum_{i=1}^n \sum_{j=1}^p a_i m(C_{ij}) \tag{3.2}$$

et

$$\sum_{j=1}^p b_j m(B_j) = \sum_{j=1}^p \sum_{i=1}^n b_j m(C_{ij}) \tag{3.3}$$

On remarque alors que $a_i = b_j$ dès que $C_{ij} \neq \emptyset$, d'où l'égalité 3.1. ■

Lemme 3.3 (Décomposition d'une fonction étagée avec une partition)
Soit (E, T) un espace mesurable, et soit $f \in \mathcal{E}$ une fonction étagée. Alors il existe une unique famille finie $(a_i, A_i)_{i=0,\ldots,n} \subset \mathbb{R} \times T$ t.q. $a_i \neq a_j$ si $i \neq j$, $A_i \neq \emptyset$, pour tout i, $A_i \cap A_j = \emptyset$, si $i \neq j$, $E = \cup_{i=0}^{n} A_i$ et $f = \sum_{i=0}^{n} a_i 1_{A_i}$.

Démonstration :
La démonstration est très voisine de celle donnée pour la décomposition d'une fonction étagée positive (lemme 3.1). L'ensemble $\{a_i,\ i \in \{0, \ldots, n\}\}$ est l'ensemble de toutes les valeurs prises par f (et pas seulement les valeurs non nulles) et $A_i = \{x \in E;\ f(x) = a_i\}$.
■

Enfin, on conclut ce paragraphe en remarquant que $\mathcal{E}$ est un espace vectoriel sur $\mathbb{R}$.

Proposition 3.1 (Structure vectorielle de $\mathcal{E}$) *Soit (E, T) un espace mesurable, l'ensemble des fonctions étagées, $\mathcal{E}$, est un espace vectoriel sur $\mathbb{R}$. De plus, si $f, g \in \mathcal{E}$, on a aussi $fg \in \mathcal{E}$.*

Démonstration :
Soit $f, g \in \mathcal{E}$ et soit $\alpha, \beta \in \mathbb{R}$. On utilise la décomposition de f et g donnée dans le lemme 3.3. Elle donne $f = \sum_{i=0}^{n} a_i 1_{A_i}$ et $g = \sum_{j=0}^{m} b_j 1_{B_i}$. Comme les familles $(A_i)_{\{i \in \{0,\ldots,n\}\}}$ et $(B_j)_{\{j \in \{0,\ldots,m\}\}}$ forment des partitions de E, on a:
$f = \sum_{i=0}^{n} \sum_{j=0}^{m} a_i 1_{A_i \cap B_j}$ et $g = \sum_{j=0}^{m} \sum_{i=0}^{n} b_j 1_{A_i \cap B_j}$,
de sorte que $\alpha f + \beta g = \sum_{i=0}^{n} \sum_{j=0}^{m} (\alpha a_i + \beta b_i) 1_{A_i \cap B_j}$, ce qui montre que $\alpha f + \beta g \in \mathcal{E}$, et donc que $\mathcal{E}$ est un espace vectoriel.
D'autre part, on remarque aussi que $fg = \sum_{i=0}^{n} \sum_{j=0}^{m} a_i b_i 1_{A_i \cap B_j}$, ce qui montre que $fg \in \mathcal{E}$.
■

On montrera aussi les propriétés de linéarité et de monotonie de l'intégrale des fonctions étagées (voir proposition 4.1).

3.3 Fonctions mesurables et variables aléatoires

Afin d'étendre le concept d'intégrale à une classe de fonctions plus générale que celle des fonctions étagées (positives), on introduit les fonctions mesurables (positives). On pourra ensuite utiliser une technique de "passage à la limite" pour définir l'intégrale de telles fonctions.

On va tout d'abord définir la notion de mesurabilité pour une fonction f de E dans F. L'espace de départ, E, est muni d'une tribu et l'espace d'arrivée, F, est, en général, muni d'une topologie (et donc de sa tribu de Borel, les exemples fondamentaux sont $F = \mathbb{R}$ ou $F = \overline{\mathbb{R}}_+$). On peut aussi considérer le cas où F est muni d'une tribu (non donnée par une topologie sur F).

Définition 3.4 (Fonction mesurable) *Soient (E, T) un espace mesurable et F un ensemble muni d'une topologie (par exemple : $F = \mathbb{R}$ ou $\overline{\mathbb{R}}_+$). Une fonction f, définie de E dans F, est une fonction T-mesurable si $f^{-1}(A) \in T$, pour tout $A \in \mathcal{B}(F)$. (Ce qui est équivalent à dire que la tribu $f^{-1}(\mathcal{B}(F)) = \{f^{-1}(B),\ B \in \mathcal{B}(F)\}$ est incluse dans T ou encore que la tribu $T_f = \{B \in \mathcal{P}(F);\ f^{-1}(B) \in T\}$ contient*

$\mathcal{B}(F)$, voir l'exercice 2.4 sur les tribus image directe et image réciproque.) En l'absence d'ambiguïté possible on dira "mesurable" au lieu de "T-mesurable".

Plus généralement, si F n'est pas muni d'une topologie (et donc de la tribu $\mathcal{B}(F)$) mais est muni directement d'une tribu $\mathcal{T}$ (on a alors deux espaces mesurables : (E,T) et $(F,\mathcal{T})$), une fonction f, définie de E dans F, est une fonction $(T,\mathcal{T})$-mesurable si $f^{-1}(A) \in T$, pour tout $A \in \mathcal{T}$. (Ce qui est équivalent à dire que la tribu $f^{-1}(\mathcal{T}) = \{f^{-1}(B),\, B \in \mathcal{T}\}$ est incluse dans T ou encore que la tribu $T_f = \{B \in \mathcal{P}(F);\ f^{-1}(B) \in T\}$ contient $\mathcal{T}$.) En l'absence d'ambiguïté possible on dira "mesurable" au lieu de "$(T,\mathcal{T})$-mesurable".

Remarque 3.3 Une fonction étagée est toujours mesurable. En effet, soit (E,T) un espace mesurable. Soit $f \in \mathcal{E}$ (donc f est une application de E dans $\mathbb{R}$). D'après la proposition 3.3, il existe $(A_0,\ldots,A_n)$, partition de E, et $a_0,\ldots,a_n \in \mathbb{R}$ t.q. $f = \sum_{i=0}^{n} a_i 1_{A_i}$ et $A_i \in T$ pour tout $i \in \{0,\ldots,n\}$. Pour tout $B \subset \mathbb{R}$, on a donc $f^{-1}(B) = \cup_{\{i;\, a_i \in B\}} A_i \in T$. Ce qui prouve que f est mesurable de E dans $\mathbb{R}$.

Noter que si $f \in \mathcal{E}_+$, on a donc aussi f mesurable de E dans $\overline{\mathbb{R}}_+$ (voir l'exercice 3.4).

La terminologie probabiliste utilise les termes "variable aléatoire" ou " élément aléatoire" (au lieu de "fonction mesurable" ou "application mesurable").

Définition 3.5 (Variable aléatoire, élément aléatoire)

1. *Soit (E,T) un espace probabilisable, on appelle variable aléatoire une fonction X définie de E dans $\mathbb{R}$ et T-mesurable, i.e. t.q. $X^{-1}(A) \in T$, pour tout $A \in \mathcal{B}(\mathbb{R})$.*
2. *Soit (E,T) et $(F,\mathcal{T})$ deux espaces probabilisables. Une fonction X, définie de E dans F, est un élément aléatoire si c'est une fonction $(T,\mathcal{T})$-mesurable (c'est-à-dire si $X^{-1}(A) \in T$, pour tout $A \in \mathcal{T}$). Lorsque F est un espace vectoriel, on dit que X est une variable aléatoire vectorielle ou un "vecteur aléatoire".*

Remarque 3.4 Comme cela a été dit dans la proposition 3.4, on dit, en l'absence d'ambiguïté, "mesurable" au lieu de "T-mesurable". On remarque d'ailleurs que le terme probabiliste "variable aléatoire" ne mentionne pas la dépendance par rapport à la tribu. Dans la définition 3.5, on a noté X la variable aléatoire plutôt que f car c'est l'usage dans la littérature probabiliste.

Définition 3.6 (Tribu engendrée par une fonction mesurable)
Soient (E,T) un espace mesurable (resp. probabilisable) et f (resp. X) une fonction mesurable de E dans $\mathbb{R}$ (resp. une variable aléatoire) alors l'ensemble $\{f^{-1}(A), A \in \mathcal{B}(\mathbb{R})\}$ (resp. $\{X^{-1}(A), A \in \mathcal{B}(\mathbb{R})\}$) est une tribu sur E qu'on appelle tribu engendrée par la fonction mesurable f (resp. la variable aléatoire X). Cette tribu est aussi la tribu image réciproque de $\mathcal{B}(\mathbb{R})$ par f (resp. X).

Définition 3.7 (espaces $\mathcal{M}$ et $\mathcal{M}_+$) *Soit (E,T) un espace mesurable, on note :*

- *$\mathcal{M}(E,T) = \{f : E \to \mathbb{R},$ mesurable $\}$,*
- *$\mathcal{M}_+(E,T) = \{f : E \to \overline{\mathbb{R}}_+,$ mesurable $\}$.*

En l'absence d'ambiguïté, on notera $\mathcal{M} = \mathcal{M}(E,T)$ et $\mathcal{M}_+ = \mathcal{M}_+(E,T)$.

Proposition 3.2 (Première caractérisation de la mesurabilité)
Soient $(E,\, T)$ un espace mesurable et $f : E \to F$, avec $F = \mathbb{R}$ ou $\overline{\mathbb{R}}_+$. Soit $\mathcal{C}$ une partie de $\mathcal{P}(F)$ engendrant la tribu borélienne de F. On a alors : f est mesurable si et seulement $f^{-1}(C) \in T$ pour tout $C \in \mathcal{C}$. En particulier, f est mesurable si et seulement si f vérifie l'une des deux propriétés suivantes :

1. $f^{-1}(]\alpha,\beta[) \in T$, *pour tout* $\alpha,\beta \in \mathbb{R}$, $\alpha < \beta$,

2. $f^{-1}(]\alpha,\infty[) \in T$, *pour tout* $\alpha \in \mathbb{R}$.

Dans cette caractérisation, l'ensemble $]\alpha,\beta[$ *(ou* $]\alpha,\infty[$*) désigne, bien sûr, l'ensemble des éléments de* F *appartenant à* $]\alpha,\beta[$ *(ou* $]\alpha,\infty[$*).*

La démonstration de cette proposition fait l'objet de l'exercice 3.1. Le lecteur pourra trouver lui-même d'autres caractérisations de la mesurabilité, en utilisant la proposition ci-dessus. Par exemple, soit f de E (muni de la tribu T) dans $\overline{\mathbb{R}}_+$, la fonction f est mesurable si et seulement si $f^{-1}(]\alpha,\infty]) \in T$ pour tout $\alpha > 0$ (par contre, $f^{-1}(]\alpha,\infty[) \in T$ pour tout $\alpha \geq 0$ n'implique pas que f est mesurable).

La proposition suivante nous permettra de définir l'intégrale des fonctions appartenant à $\mathcal{M}_+$ (comme limite d'intégrales de fonctions étagées, voir le chapitre suivant). Par contre, on ne pourra pas donner un sens, dans le cas général, à l'intégrale des fonctions appartenant à $\mathcal{M}$.

Proposition 3.3 (Mesurabilité positive) *Soient* (E,T) *un espace mesurable et* $f : E \to \overline{\mathbb{R}}_+$. *Alors* $f \in \mathcal{M}_+$ *si et seulement si il existe une suite* $(f_n)_{n\in\mathbb{N}} \subset \mathcal{E}_+$, *t.q. :*

1. *Pour tout* $x \in E$, $f_n(x) \to f(x)$, *quand* $n \to \infty$,

2. $f_{n+1}(x) \geq f_n(x)$, *pour tout* $x \in E$, *et tout* $n \in \mathbb{N}$.

les 2 conditions précédentes seront dénotées dans la suite sous la forme $f_n \uparrow f$ *quand* $n \to \infty$.

Démonstration : Soit $(f_n)_{n\in\mathbb{N}} \in \mathcal{E}_+$ t.q. $f_n \uparrow f$ quand $n \to \infty$. On remarque que, pour tout $\alpha \in \mathbb{R}$,

$$f^{-1}(]\alpha,+\infty]) = \bigcup_{n\in\mathbb{N}} f_n^{-1}(]\alpha,+\infty[). \tag{3.4}$$

Comme f_n est mesurable, pour tout $n \in \mathbb{N}$ (voir la remarque 3.3), on a $f_n^{-1}(]\alpha,+\infty[) \in T$ pour tout $n \in \mathbb{N}$ et donc, par stabilité de T par union dénombrable, $f^{-1}(]\alpha,+\infty]) \in T$. Ceci étant vrai pour tout $\alpha \in \mathbb{R}$, on en déduit, comme $\{]\alpha,+\infty],\ \alpha \geq 0\}$ engendre $\mathcal{B}(\overline{\mathbb{R}}_+)$, que f est mesurable de E dans $\overline{\mathbb{R}}_+$, c'est-à-dire $f \in \mathcal{M}_+$.

Réciproquement, on suppose que $f \in \mathcal{M}+$. On va construire $(f_n)_{n\in\mathbb{N}^\star} \subset \mathcal{E}_+$ t.q. $f_n \uparrow f$ quand $n \to \infty$.
Pour $n \in \mathbb{N}^\star$, on définit la fonction f_n par :

$$f_n(x) = \begin{cases} \dfrac{p}{2^n} \text{ si } f(x) \in [\dfrac{p}{2^n}, \dfrac{p+1}{2^n}[, \text{ avec } p \in \{0,\ldots,n2^n-1\}, \\ n \text{ si } f(x) \geq n, \end{cases} \tag{3.5}$$

de sorte que

$$f_n = n1_{\{x\in E;\, f(x)\geq n\}} + \sum_{p=0}^{n2^n-1} \frac{p}{2^n} 1_{\{x\in E;\, f(x)\in[\frac{p}{2^n},\frac{p+1}{2^n}[\}}.$$

Comme $f \in \mathcal{M}_+$, on a $\{x \in E;\, f(x) \geq n\} \in T$ et $\{x \in E;\, f(x) \in [\frac{p}{2^n}, \frac{p+1}{2^n}[\} \in T$ pour tout n et tout p, on a donc $(f_n)_{n\in\mathbb{N}^\star} \subset \mathcal{E}_+$.

On montre maintenant que, pour tout $x \in E$, on a $f_n(x) \to f(x)$, quand $n \to \infty$. Soit $x \in E$. On distingue deux cas :

Premier cas. On suppose $f(x) < \infty$. On a alors, pour $n \geq f(x)$, $|f(x) - f_n(x)| \leq \frac{1}{2^n}$. On a donc $f_n(x) \to f(x)$ quand $n \to \infty$.
Deuxième cas. On suppose $f(x) = \infty$. On a alors $f_n(x) = n$ pour tout $n \in \mathbb{N}^\star$ et donc $f_n(x) \to f(x)$ quand $n \to \infty$.

On montre enfin que, pour tout $x \in E$ et pour tout $n \in \mathbb{N}^\star$, on a $f_{n+1}(x) \geq f_n(x)$. Soit $x \in E$ et $n \in \mathbb{N}^\star$. On distingue trois cas :
Premier cas. On suppose $f(x) \geq n + 1$. On a alors $f_{n+1}(x) = n + 1 > n = f_n(x)$.
Deuxième cas. On suppose $n \leq f(x) < n + 1$. Il existe alors $i \in \{n2^{n+1}, \ldots, (n + 1)2^{n+1} - 1\}$ t.q. $f(x) \in [\frac{i}{2^{n+1}}, \frac{i+1}{2^{n+1}}]$. On a alors $f_n(x) = n \leq \frac{i}{2^{n+1}} = f_{n+1}(x)$.
Troisième cas. On suppose $f(x) < n$. Il existe alors $p \in \{0, \ldots, n2^n - 1\}$ t.q. $f(x) \in [\frac{p}{2^n}, \frac{p+1}{2^n}[= [\frac{2p}{2^{n+1}}, \frac{2(p+1)}{2^{n+1}}[$. Si $f(x) \in [\frac{2p}{2^{n+1}}, \frac{2p+1}{2^{n+1}}[$, on a $f_n(x) = \frac{p}{2^n} = \frac{2p}{2^{n+1}} = f_{n+1}(x)$. Si $f(x) \in [\frac{2p+1}{2^{n+1}}, \frac{2(p+1)}{2^{n+1}}[$, on a $f_n(x) = \frac{p}{2^n} < \frac{2p+1}{2^{n+1}} = f_{n+1}(x)$. On a toujours $f_n(x) \leq f_{n+1}(x)$.

On a bien ainsi construit $(f_n)_{n\in\mathbb{N}^\star} \subset \mathcal{E}_+$ t.q. $f_n \uparrow f$ quand $n \to \infty$.

■

Proposition 3.4 (Mesurabilité sans signe) *Soient (E, T) un espace mesurable, et $f : E \to \mathbb{R}$. On suppose que f est mesurable. Il existe alors une suite $(f_n)_{n\in\mathbb{N}} \subset \mathcal{E}$ t.q., pour tout $x \in E$, $f_n(x) \to f(x)$, quand $n \to \infty$.*

DÉMONSTRATION :
On définit la fonction $f^+ : E \to \mathbb{R}_+$ par $f^+(x) = \max(f(x), 0)$ pour tout $x \in E$. On remarque que $f^+ \in \mathcal{M}_+$ (et $f^+ \in \mathcal{M}$, voir l'exercice 3.4). En effet, f^+ prend ses valeurs dans $\mathbb{R}_+$ et $(f^+)^{-1}(]\alpha, \infty]) = f^{-1}(]\alpha, \infty[) \in T$ si $\alpha > 0$. On conclut en remarquant que $\{]\alpha, \infty], \alpha > 0\}$ engendre $\mathcal{B}(\overline{\mathbb{R}}_+)$. On définit également $f^- = (-f)^+$, de sorte que $f = f^+ - f^-$. On a donc aussi $f^- \in \mathcal{M}_+$. La proposition 3.3 donne l'existence de $(f_n)_{n\in\mathbb{N}} \subset \mathcal{E}_+$ et $(g_n)_{n\in\mathbb{N}} \subset \mathcal{E}_+$ t.q. $f_n \uparrow f^+$ et $g_n \uparrow f^-$ quand $n \to \infty$. On pose $h_n = f_n - g_n$, de sorte que $h_n(x) \to f(x)$, quand $n \to \infty$, pour tout $x \in E$. D'autre part, comme $\mathcal{E}$ est un espace vectoriel (voir la proposition 3.1 page 55), on a $(h_n)_{n\in\mathbb{N}} \in \mathcal{E}$.

■

La proposition précédente nous donnera, avec les propriétés de stabilité de $\mathcal{M}$ et $\mathcal{M}_+$ (voir la proposition 3.5) une deuxième caractérisation de la mesurabilité, voir la proposition 3.6.

L'ensemble des fonctions mesurables est un ensemble très "stable", c'est-à-dire que des opérations "usuelles" (comme "addition", "multiplication", "limite"...) sur des fonctions mesurables donnent encore des fonctions mesurables, ceci est précisé dans la proposition suivante. Dans le cas (fondamental) de $(E, T) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, il est "difficile" de trouver des fonctions non mesurables (comme il est "difficile" de trouver des parties non boréliennes, bien que le cardinal de $\mathcal{B}(\mathbb{R})$ soit égal au cardinal de $\mathbb{R}$ et donc strictement inférieur au cardinal de $\mathcal{P}(\mathbb{R})$). En pratique, on peut en gros supposer que les fonctions de $\mathbb{R}$ dans $\mathbb{R}$ sont *toutes* $\mathcal{B}(\mathbb{R})$-mesurables (bien qu'il y ait "beaucoup" de fonctions non mesurables...).

Proposition 3.5 (Stabilité de $\mathcal{M}$ et $\mathcal{M}_+$) *Soit (E, T) un espace mesurable.*

1. *Soit $I \subset \mathbb{N}$.*

   Soit $(f_n)_{n\in I} \subset \mathcal{M}_+$, alors $\sup_{n\in I} f_n \in \mathcal{M}_+$ et $\inf_{n\in I} f_n \in \mathcal{M}_+$.

   Soit $(f_n)_{n\in I} \subset \mathcal{M}$. Si $\sup_{n\in I} f_n$ prend ses valeurs dans $\mathbb{R}$, alors $\sup_{n\in I} f_n \in \mathcal{M}$. De même, si $\inf_{n\in I} f_n$ prend ses valeurs dans $\mathbb{R}$, alors $\inf_{n\in I} f_n \in \mathcal{M}$.

2. *Soit $(f_n)_{n\in\mathbb{N}} \subset \mathcal{M}_+$, alors $\limsup_{n\to\infty} f_n \in \mathcal{M}_+$ et $\liminf_{n\to\infty} f_n \in \mathcal{M}_+$.*

   Soit $(f_n)_{n\in\mathbb{N}} \subset \mathcal{M}$. Si $\limsup_{n\in\mathbb{N}} f_n$ prend ses valeurs dans $\mathbb{R}$, alors $\limsup_{n\in\mathbb{N}} f_n \in \mathcal{M}$. De même, si $\liminf_{n\in\mathbb{N}} f_n$ prend ses valeurs dans $\mathbb{R}$, alors $\liminf_{n\in\mathbb{N}} f_n \in \mathcal{M}$.

3. *Soit $(f_n)_{n\in\mathbb{N}} \subset \mathcal{M}_+$. On suppose que $f_n(x) \to f(x)$ dans $\overline{\mathbb{R}}_+$, pour tout $x \in E$. Alors $f \in \mathcal{M}_+$.*

   Soit $(f_n)_{n\in\mathbb{N}} \subset \mathcal{M}$. On suppose que $f_n(x) \to f(x)$ dans $\mathbb{R}$, pour tout $x \in E$. Alors $f \in \mathcal{M}$.

4. *$\mathcal{M}$ est un espace vectoriel sur $\mathbb{R}$ et si $f, g \in \mathcal{M}$, alors $fg \in \mathcal{M}$.*

Démonstration :

1. Soit $(f_n)_{n\in\mathbb{N}} \subset \mathcal{M}_+$. Il est clair que $f = \sup_{n\in I} f_n$ est bien définie et prend ses valeurs dans $\overline{\mathbb{R}}_+$. Puis, Pour tout $\alpha \in \mathbb{R}$, on a $f^{-1}(]\alpha,\infty]) = \cup_{n\in I} f_n^{-1}(]\alpha,\infty]) \in T$. Comme $\{]\alpha,\infty] \cap \overline{\mathbb{R}}_+,\ \alpha \in \mathbb{R}\}$ engendre $\mathcal{B}(\overline{\mathbb{R}}_+)$, on en déduit que f est mesurable de E dans $\overline{\mathbb{R}}_+$, c'est-à-dire $f \in \mathcal{M}_+$.

   De même la fonction $f = \inf_{n\in I} f_n$ est aussi bien définie et prend ses valeurs dans $\overline{\mathbb{R}}_+$ (elle prend même ses valeurs dans $\mathbb{R}_+$ si les f_n prennent leurs valeurs dans $\mathbb{R}_+$, ceci n'est pas vrai avec la fonction $\sup_{n\in I} f_n$). On remarque ensuite que $f^{-1}([-\infty,\alpha[) = \cup_{n\in I} f_n^{-1}([-\infty,\alpha[)$, pour tout $\alpha \in \mathbb{R}$. Comme $\{[-\infty,\alpha[\cap\overline{\mathbb{R}}_+,\ \alpha \in \mathbb{R}\}$ engendre $\mathcal{B}(\overline{\mathbb{R}}_+)$, on en déduit que f est mesurable de E dans $\mathcal{B}(\overline{\mathbb{R}}_+)$, c'est-à-dire $f \in \mathcal{M}_+$.

   Soit maintenant $(f_n)_{n\in\mathbb{N}} \subset \mathcal{M}$. La fonction $f = \sup_{n\in I} f_n$ est bien définie si on la considère comme étant à valeurs dans $\overline{\mathbb{R}}$ car elle peut prendre la valeur $+\infty$ en certains points. On la suppose maintenant à valeurs dans $\mathbb{R}$. On peut alors raisonner comme précédemment en remarquant que $f^{-1}(]\alpha,\infty[) = \cup_{n\in I} f_n^{-1}(]\alpha,\infty[)$ et que $\{]\alpha,\infty[,\ \alpha \in \mathbb{R}\}$ engendre $\mathcal{B}(\mathbb{R})$. De même, La fonction $f = \inf_{n\in I} f_n$ est bien définie si on la considère comme étant à valeurs dans $\overline{\mathbb{R}}$ car elle peut prendre la valeur $-\infty$ en certains points. On la suppose maintenant à valeurs dans $\mathbb{R}$. On peut alors raisonner comme précédemment en remarquant que $f^{-1}(]-\infty,\alpha[) = \cup_{n\in I} f_n^{-1}(]-\infty,\alpha[)$ et que $\{]-\infty,\alpha[,\ \alpha \in \mathbb{R}\}$ engendre $\mathcal{B}(\mathbb{R})$.

2. Soit $(f_n)_{n\in\mathbb{N}} \subset \mathcal{M}_+$. On pose $f = \limsup_{n\to\infty} f_n$, la fonction f est bien définie à valeurs dans $\overline{\mathbb{R}}_+$. Pour tout $x \in E$, on a $f(x) = \limsup_{n\to\infty} f_n(x) = \lim_{n\to\infty}(\sup_{p\geq n} f_p(x)) = \inf_{n\in\mathbb{N}}(\sup_{p\geq n} f_p(x))$, c'est-à-dire $f = \inf_{n\in\mathbb{N}}(\sup_{p\geq n} f_p)$. En utilisant les résultats précédents (avec sup puis inf), on a donc $f \in \mathcal{M}_+$. Un raisonnement similaire donne $\liminf_{n\to\infty} f_n = \sup_{n\in\mathbb{N}}(\inf_{p\geq n} f_p) \in \mathcal{M}_+$.

   Soit maintenant $(f_n)_{n\in\mathbb{N}} \subset \mathcal{M}$. On suppose que $f = \limsup_{n\to\infty} f_n = \inf_{n\in\mathbb{N}}(\sup_{p\geq n} f_p)$ (qui est bien définie dans $\mathbb{R} \cup \{\infty\}$) prend ses valeurs dans $\mathbb{R}$. Comme les f_n prennent leurs valeurs dans $\mathbb{R}$, on peut alors remarquer que la fonction $\sup_{p\geq n} f_p$ prend aussi ses valeurs dans $\mathbb{R}$, pour tout $n \in \mathbb{N}$. On a donc, avec la propriété démontrée en 1, $\sup_{p\geq n} f_p \in \mathcal{M}$ pour tout $n \in \mathbb{N}$. Puis, utilisant encore la propriété démontrée en 1, $f = \inf_{n\in\mathbb{N}}(\sup_{p\geq n} f_p) \in \mathcal{M}$. Un raisonnement analogue donne $\liminf_{n\to\infty} f_n = \sup_{n\in\mathbb{N}}(\inf_{p\geq n} f_p) \in \mathcal{M}$ dès que l'on suppose que $\liminf_{n\to\infty} f_n$ prend ses valeurs dans $\mathbb{R}$.

3. Cette question est immédiate grâce à la précédente. Il suffit de remarquer que dès que la limite de la suite $(f_n(x))_{n\in\mathbb{N}}$ existe, elle est nécessairement égale à la limite supérieure (ou la limite inférieure) de cette même suite $(f_n(x))_{n\in\mathbb{N}}$, c'est-à-dire que $\lim_{n\to\infty} f_n(x) = \limsup_{n\to\infty} f_n(x)$ (pour tout $x \in E$). Ici on remarque donc simplement que $f = \limsup_{n\to\infty} f_n$ et on applique la propriété 2.

4. Soit $f, g \in \mathcal{M}$ et soit $\alpha, \beta \in \mathbb{R}$ On pose $h = \alpha f + \beta g$. D'après la proposition 3.4, il existe $(f_n)_{n\in\mathbb{N}} \subset \mathcal{E}$ et $(g_n)_{n\in\mathbb{N}} \subset \mathcal{E}$ t.q. $f_n(x) \to f(x)$ et $g_n(x) \to g(x)$, quand $n \to \infty$, pour tout $x \in E$. On pose $h_n = \alpha f_n + \beta g_n$, de sorte que $h_n(x) \to h(x)$, quand $n \to \infty$, pour tout $x \in E$. La proposition 3.1 donne que $\mathcal{E}$ est un espace vectoriel sur $\mathbb{R}$, on a donc $(h_n)_{n\in\mathbb{N}} \subset \mathcal{E}$. Comme $\mathcal{E} \subset \mathcal{M}$ (voir la remarque 3.3), la propriété 3 ci dessus donne alors que $h \in \mathcal{M}$. L'ensemble $\mathcal{M}$ est donc un espace vectoriel (sur $\mathbb{R}$).

   Soit $f, g \in \mathcal{M}$. On pose $h = fg$. On raisonne comme ci dessus, il existe $(f_n)_{n\in\mathbb{N}} \subset \mathcal{E}$ et $(g_n)_{n\in\mathbb{N}} \subset \mathcal{E}$ t.q. $f_n(x) \to f(x)$ et $g_n(x) \to g(x)$, quand $n \to \infty$, pour tout $x \in E$. On pose $h_n = f_n g_n$, de sorte que $h_n(x) \to h(x)$, quand $n \to \infty$, pour tout $x \in E$. La proposition 3.1 donne aussi $(h_n)_{n\in\mathbb{N}} \subset \mathcal{E} \subset \mathcal{M}$. La propriété 3 ci dessus donne alors que $h \in \mathcal{M}$.

■

Proposition 3.6 (Deuxième caractérisation de la mesurabilité) *Soit (E, T) un espace mesurable et $f : E \to \mathbb{R}$. Alors, f est mesurable si et seulement si il existe une suite $(f_n)_{n\in\mathbb{N}} \subset \mathcal{E}$ t.q., pour tout $x \in E$, $f_n(x) \to f(x)$, quand $n \to \infty$.*

DÉMONSTRATION :
Cette caractérisation est donnée par la proposition 3.4 pour le sens "seulement si" et par la propriété 3 de la proposition 3.5 pour le sens "si".

■

On rappelle aussi qu'une fonction f de E (muni de la tribu T) dans $\overline{\mathbb{R}}_+$ est mesurable (c'est-à-dire appartient à $\mathcal{M}_+$) si et seulement si il existe $(f_n)_{n\in\mathbb{N}} \subset \mathcal{E}_+$ t.q. $f_n \uparrow f$ (voir la proposition 3.3).

Remarque 3.5 Il est intéressant de remarquer que la proposition 3.6 peut être fausse si on prend pour F un espace topologique quelconque (elle reste vraie, par exemple, si F est un espace vectoriel normé de dimension finie) avec une définition immédiate de $\mathcal{E}$ généralisant celle donnée pour les fonctions à valeurs dans $\mathbb{R}$ ou $\overline{\mathbb{R}}_+$.

Définition 3.8 *Soient E un ensemble et $f : E \to \mathbb{R}$. Pour tout $x \in E$, on pose :*

- $f^+(x) = \max(f(x), 0)$,
- $f^-(x) = -\min(f(x), 0) = (-f)^+(x)$,
- $|f|(x) = |f(x)|$.

Proposition 3.7 *Soient (E,T) un espace mesurable et $f \in \mathcal{M}$. On a alors $f = f^+ - f^-$, $|f| = f^+ + f^-$ et f^+, f^-, $|f| \in \mathcal{M}_+ \cap \mathcal{M}$.*

DÉMONSTRATION :
Le fait que $f = f^+ - f^-$ et $|f| = f^+ + f^-$ est immédiat. On a déjà vu, dans la démonstration de la proposition 3.4, que $f^+, f^- \in \mathcal{M}_+$ et donc que $f^+, f^- \in \mathcal{M}$ (voir l'exercice 3.4). La proposition 3.5 donne que $\mathcal{M}$ est un espace vectoriel sur $\mathbb{R}$. On a donc $|f| \in \mathcal{M}$ et donc aussi $|f| \in \mathcal{M}_+$ car $|f| \geq 0$.

■

3.4 Mesure image, loi d'une v.a., v.a. indépendantes

Soit (E, T) et $(F, \mathcal{T})$ deux espaces mesurables (l'exemple fondamental est $(F, \mathcal{T}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$) et f une fonction mesurable de E vers F. Si m est une mesure sur T, alors on peut définir, à partir de f et m, une mesure sur $\mathcal{T}$ de la manière suivante :

Proposition 3.8 (Mesure image) *Soient (E, T, m) un espace mesuré, $(F, \mathcal{T})$ un espace mesurable et f une fonction mesurable de E vers F (c'est-à-dire $(T, \mathcal{T})$-mesurable). Alors, l'application m_f définie de $\mathcal{T}$ dans $\mathbb{R}_+$ par : $m_f(A) = m(f^{-1}(A))$, pour tout $A \in \mathcal{T}$, est une mesure sur $\mathcal{T}$, appelée mesure image par f.*

DÉMONSTRATION : Il suffit de remarquer que m_f est bien définie, que $m_f(\emptyset) = 0$ et que m_f est $\sigma-$additive, ce qui découle naturellement des propriétés de m. ∎

Définition 3.9 (Loi de probabilité et fonction de répartition d'une variable aléatoire)
Soient (E, T, p) un espace probabilisé, X une variable aléatoire réelle (c'est-à-dire une fonction mesurable de E, muni de la tribu T, dans $\mathbb{R}$, muni de la tribu borélienne). On appelle loi de probabilité de la variable aléatoire X la probabilité p_X image de p par X (cette probabilité est donc définie sur $\mathcal{B}(\mathbb{R})$). On appelle fonction de répartition de la variable aléatoire X la fonction de répartition de la probabilité p_X.

Dans de nombreux cas, les modèles probabilistes seront déterminés par une loi de probabilité d'une variable aléatoire. Une conséquence immédiate du théorème 2.4 est que la loi de probabilité d'une variable aléatoire réelle est entièrement déterminée par sa fonction de répartition. Ceci est énoncé dans la proposition suivante.

Proposition 3.9 (Egalité de deux lois) *Soient (E, T, p) et (E', T', p') des espaces probabilisés, X une variable aléatoire réelle sur E (c'est-à-dire une fonction mesurable de E, muni de T, dans R muni de $\mathcal{B}(\mathbb{R})$) et X' une variable aléatoire réelle sur E'. On a alors $p_X = p_Y$ si et seulement si $p(\{X \le t\}) = p'(\{Y \le t\})$ pour tout $t \in \mathbb{R}$. On a aussi $p_X = p_Y$ si et seulement si $p(\{s \le X \le t\}) = p'(\{s \le Y \le t\})$ pour tout $s, t \in \mathbb{R}$, $s < t$. (Les inégalités strictes peuvent remplacées par des inégalités larges.)*

DÉMONSTRATION : Cette proposition est une conséquence des théorèmes 2.4 et 2.5. Il suffit de remarquer que $p(\{X \le t\}) = p_X(]-\infty, t])$ et $p(\{s \le X \le t\}) = p_X([s, t])$ (et les mêmes égalités avec Y au lieu de X). ∎

On rappelle que la notation $p(\{X \le t\})$ (si X est une v. a. réelle sur l'espace probabilisé (E, T, p)) signifie $p(\{\omega \in E;\ X(\omega) \le t\})$. Cette notation sera abrégée sous la forme $p(X \le t)$.

Définition 3.10 (Variables aléatoires équidistribuées)
Soient (E, T, p) et (E', T', p') des espaces probabilisés, X (resp. X') une variable aléatoire de (E, T, p) (resp. (E', T', p')) dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, on dit que les variables aléatoires X et X' sont équidistribuées si elles ont même loi de probabilité.

Définition 3.11 (Variable aléatoire discrète, entière, continue) *Soient (E, T, p) un espace probabilisé, X une variable aléatoire sur (E, T, p), p_X la loi de la variable aléatoire X et F_X sa fonction de répartition;*

1. *Si $X(E)$ est dénombrable, on dit que la variable aléatoire X est discrète.*
2. *Si $X(E) \subset \mathbb{N}$, on dit que la variable aléatoire X est entière.*

3. *Si la fonction de répartition F_X définie de $\mathbb{R}$ dans $[0,1]$ est continue, on dit que la variable aléatoire est continue.*

Définition 3.12 (Variables aléatoires indépendantes) *Soit $(E, T, p))$ un espace probabilisé.*

1. *Soit $N > 1$ et $X_1, \ldots, X_N$ une famille de variables aléatoires réelles. On dit que $X_1, \ldots, X_N$ sont indépendantes (ou que la famille ($X_1, \ldots, X_N$) est indépendante) si les tribus engendrées par $X_1, \ldots, X_N$ (on notera souvent $\tau(X)$ ou $\sigma(X)$ la tribu engendrée par la variable aléatoire X) sont indépendantes.*

2. *Soit $(X_n)_{n\in\mathbb{N}^\star}$ une suite de variables aléatoires réelles. On dit que la suite $(X_n)_{n\in\mathbb{N}^\star}$ est indépendante (ou que les v.a. $X_1, \ldots, X_n, \ldots$ sont indépendantes) si, pour tout $N > 1$, les v.a. $X_1, \ldots, X_N$ sont indépendantes.*

On appellera "suite de v.a.r.i.i.d. " une suite de variables aléatoires réelles indépendantes identiquement distribuées (ce dernier point signifiant que toutes les v.a. de la suite ont même loi).

Soit $(E, T, p))$ un espace probabilisé et X_1, X_2, X_3 trois v.a.r. (c'est-à-dire variables aléatoires réelles). Le fait que X_1 soit indépendante de X_2 et X_3 n'implique pas que X_1 soit indépendante de (par exemple) $X_2 + X_3$, même si X_2 et X_3 sont indépendantes. Mais, on a bien X_1 indépendante de $X_2 + X_3$ si la famille (X_1, X_2, X_3) est indépendante. Ceci est une conséquence da la proposition suivante.

Proposition 3.10 (Indépendance et composition)
Soit $(E, T, p))$ un espace probabilisé, $n \geq 1$, $m \geq 1$ et $X_1, \ldots, X_n, Y_1, \ldots, Y_m$ des v.a.r indépendantes. Soit φ une fonction borélienne de $\mathbb{R}^n$ dans $\mathbb{R}$ et ψ une fonction borélienne de $\mathbb{R}^m$ dans $\mathbb{R}$. Alors, les v.a.r. $\varphi(X_1, \ldots, X_n)$ et $\psi(Y_1, \ldots, Y_m)$ sont indépendantes. Nous avons ici décomposé la famille initiale de v.a.r. indépendantes en 2 groupes. la proposition peut se généraliser à une décomposition en un nombre quelconque de groupes.

Démonstration : La notation $\varphi(X_1, \ldots, X_n)$ est un peu incorrecte (mais toujours utilisée). Elle désigne (comme on le devine facilement) la composition de φ (qui va de $\mathbb{R}^n$ dans $\mathbb{R}$) avec l'application de E dans $\mathbb{R}^n$ donnée par les X_i, $i = 1, \ldots, n$.

La démonstration de cette proposition (et de sa généralisation à un nombre quelconque de groupes) est une conséquence simple de la proposition 2.11 dés que l'on remarque que la tribu engendrée par $\varphi(X_1, \ldots, X_n)$ est incluse dans la tribu engendrée par $X_1, \ldots, X_n$, ce que nous démontrons maintenant.

On note τ la tribu engendrée par $X_1, \ldots, X_n$ et X l'application de E dans $\mathbb{R}^n$ qui à $\omega \in E$ associe $(X_1(\omega,) \ldots, X_n(\omega))^t$. Il est facile de voir que $\{A \in \mathcal{B}(\mathbb{R}^n)$ t.q. $X^{-1}(A) \in \tau\}$ est une tribu (sur $\mathbb{R}^n$). Si $A = \prod_{i=1}^n A_i$ avec $A_i \in \mathcal{B}(\mathbb{R})$ pour tout $i = 1, \ldots, n$, on a $X^{-1}(A) = \cap_{i=1}^n X_i^{-1}(A_i) \in \tau$ (car $X_i^{-1}(A_i)$ appartient à $\tau(X_i)$ et donc à τ). Comme $\mathcal{B}(\mathbb{R}^n)$ est engendrée par l'ensemble des produits de boréliens de $\mathbb{R}$ (et même par l'ensemble des produits d'intervalles ouverts de $\mathbb{R}$, voir l'exercice 2.7), on en déduit que $\{A \in \mathcal{B}(\mathbb{R}^n)$ t.q. $X^{-1}(A) \in \tau\}$ contient $\mathcal{B}(\mathbb{R}^n)$. Pour tout $B \in \mathcal{B}(\mathbb{R})$, on a donc $(\varphi(X))^{-1}(B) = X^{-1}(\varphi^{-1}(B)) \in \tau$ car $\varphi^{-1}(B) \in \mathcal{B}(\mathbb{R}^n)$ (puisque φ est borélienne). ce qui prouve bien que la tribu engendrée par $\varphi(X_1, \ldots, X_n)$ est incluse dans la tribu engendrée par $X_1, \ldots, X_n$. ■

Nous verrons au chapitre 7 la conséquence principale de l'indépendance. Cette conséquence est que, si X, Y sont des v.a.r. indépendantes, la loi du couple (X, Y) est le produit des lois P_X et P_Y (c'est-à-dire , avec les notations du Chapitre 7, $P_{(X,Y)} = P_X \otimes P_Y$). Une propriété analogue est vraie pour une famille $(X_1, \ldots, X_n)$ de v.a.r. indépendantes.

Nous terminons ce paragraphe par un théorème très utile en probabilités sur la représentation d'une v.a. mesurable par rapport à une autre v.a..

Théorème 3.1 (V.a. mesurable par rapport à une autre v.a.) *Soient X et Y deux v.a. réelles définies sur un espace de probabilités $(\Omega, \mathcal{A}, P)$. Alors, la v.a. Y est mesurable par rapport à la tribu engendrée par X (notée $\tau(X)$) si et seulement si il existe une fonction borélienne f de $\mathbb{R}$ dans $\mathbb{R}$ telle que $Y = f(X)$.*

DÉMONSTRATION : La démonstration de ce résultat fait l'objet de l'exercice 3.15 (corrigé 48). Il est intéressant de remarquer que le démonstration de ce théorème faite dans le corrigé 48 donne les informations complémentaires suivantes :

- Y est $\tau(X)$-mesurable bornée si et seulement si il existe f borélienne bornée t.q. $Y = f(X)$,
- Y est $\tau(X)$-mesurable positive si et seulement si il existe f borélienne positive t.q. $Y = f(X)$.

La partie "si" de ces deux résultats est immédiate. Pour la partie "seulement si", il suffit de remarquer que la démonstration faite dans le corrigé 48 donne f t.q.

$$\mathrm{Im} f = \{f(t),\, t \in \mathbb{R}\} \subset \mathrm{Im}(Y) \cup \{0\}, \text{ avec } \mathrm{Im} Y = \{Y(\omega),\, \omega \in \Omega\}.$$

■

3.5 Convergence p.p., p.s., en mesure, en probabilité

On introduit ici plusieurs notions de convergence de fonctions définies sur un espace mesuré à valeurs dans $\mathbb{R}$ (ou $\overline{\mathbb{R}}_+$) et on donne des liens entre ces différentes convergences. On introduit les notions équivalentes pour les variables aléatoires en langage probabiliste.

Définition 3.13 (Egalité presque partout) *Soient (E, T, m) un espace mesuré, F un ensemble et f et g des fonctions définies de E dans F ($F = \mathbb{R}$ ou $F = \overline{\mathbb{R}}_+$, par exemple) ; on dit que $f = g$ m-presque partout (et on note $f = g$ m-p.p.) si l'ensemble $\{x \in E\,;\, f(x) \neq g(x)\}$ est négligeable, c'est à dire qu'il existe $A \in T$ t.q. $m(A) = 0$ et $f(x) = g(x)$ pour tout $x \in A^c$.*

On peut remarquer que si f et g sont des fonctions mesurables de E (muni de la tribu T et de la mesure m) dans $\mathbb{R}$ (ou $\overline{\mathbb{R}}_+$), l'ensemble $\{x \in E\,;\, f(x) \neq g(x)\}$ (noté aussi $\{f \neq g\}$) appartient à T. Le fait que $f = g\, m$ p.p. revient donc à dire que $m(\{f \neq g\}) = 0$. Dans la cas où f ou g n'est pas mesurable, l'ensemble $\{f \neq g\}$ peut être négligeable sans appartenir à T (il appartient nécessairement à T si la mesure est complète, voir la définition 2.15).

En l'absence de confusion possible, on remplace m-p.p. par p.p.. Cette définition se traduit en langage probabiliste par:

Définition 3.14 (Egalité presque sûre) *Soient (E, T, p) un espace probabilisé, X et Y des variables aléatoires X et Y de (E, T) dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$; on dit que $X = Y$ p-presque sûrement (et on note $X = Y$ p.s.), si l'ensemble $\{x \in E\,;\, X(x) \neq Y(x)\}$ est négligeable.*

Définition 3.15 (Convergence presque partout) *Soient (E, T, m) un espace mesuré, F un ensemble, $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de E dans F et f une fonction de E dans F ($F = \mathbb{R}$ ou $F = \overline{\mathbb{R}}_+$, par exemple) ; on dit que f_n converge presque partout vers f ($f_n \to f$ p.p.) si il existe une partie A de E, négligeable, t.q., pour tout élément x de A^c, la suite $(f_n(x))_{n\in\mathbb{N}}$ converge vers $f(x)$.*

Noter que la convergence simple entraîne la convergence presque partout.

La définition 3.15 se traduit en langage probabiliste par:

Définition 3.16 (Convergence presque sure) *Soient (E, T, p) un espace probabilisé, $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires et X une variable aléatoire ; on dit que X_n converge presque sûrement vers X ($X_n \to X$ p.s.) si il existe une partie A de E, négligeable, t.q., pour tout élément x de A^c, la suite $(X_n(x))_{n\in\mathbb{N}}$ converge vers $X(x)$.*

Définition 3.17 (Convergence presque uniforme) *Soient (E, T, m) un espace mesuré, $(f_n)_{n\in\mathbb{N}} \subset \mathcal{M}$ et $f \in \mathcal{M}$. On dit que f_n converge presque uniformément vers f ($f_n \to f$ p.unif.) si, pour tout $\varepsilon > 0$, il existe $A \in T$ t.q. $m(A) \le \varepsilon$ et f_n converge uniformément vers f sur A^c.*

La convergence presque uniforme entraîne la convergence presque partout (voir exercice 3.24).
Attention, la convergence "presque uniforme" ne donne pas la convergence "uniforme en dehors d'un ensemble de mesure nulle". La convergence "uniforme en dehors d'un ensemble de mesure nulle" est reliée à la convergence "essentiellement uniforme", c'est-à-dire la convergence pour le "sup essentiel", défini ci-après, ou pour la norme $\|\cdot\|_\infty$ que nous verrons dans la section 6.1.2.

Définition 3.18 (sup essentiel) *Soient (E, T, m) un espace mesuré et $f \in \mathcal{M}$. On dit que f est essentiellement bornée si il existe $C \in \mathbb{R}_+$ tel que $|f| \le C$ p.p.. On appelle alors "sup essentiel de $|f|$", et on le note $\|f\|_\infty$, l'infimum des valeurs C telles que $|f| \le C$ p.p.. Si f n'est pas essentiellement bornée, on pose $\|f\|_\infty = \infty$.*

Remarquons que dans le cas où $(E, T, m) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, le sup essentiel d'une fonction continue est la borne supérieure de sa valeur absolue (ceci fait l'objet de la proposition 6.4).

Définition 3.19 (Convergence essentiellement uniforme) *Soient (E, T, m) un espace mesuré, $f \in \mathcal{M}$ et $(f_n)_{n\in\mathbb{N}} \subset \mathcal{M}$; on dit que f_n converge essentiellement uniformément vers f ($f_n \to f$ ess. unif.) si $\|f_n - f\|_\infty \to 0$ lorsque $n \to +\infty$.*

Il est facile de voir que la convergence essentiellement uniforme entraîne la convergence presque uniforme, mais la réciproque est fausse (voir l'exercice 3.25). Le théorème suivant donne, dans le cas où la mesure est finie, un résultat très important qui fait le lien entre la convergence presque partout et la convergence presque uniforme.

Théorème 3.2 (Egorov) *Soient (E, T, m) un espace mesuré, tel que $m(E) < +\infty$, $(f_n)_{n\in\mathbb{N}} \subset \mathcal{M}$ et $f \in \mathcal{M}$. On suppose que $f_n \to f$ p.p.. Alors, pour tout $\varepsilon > 0$, il existe $A \in T$ tel que $m(A) \le \varepsilon$ et f_n converge uniformément vers f sur A^c. (Autrement dit, la suite $(f_n)_{n\in\mathbb{N}}$ converge presque uniformément vers f.)*

La démonstration de ce théorème fait l'objet de l'exercice 3.25. Attention, lorsque $m(E) = +\infty$, on peut trouver des suites de fonctions qui convergent presque partout et non presque uniformément.

Définition 3.20 (Convergence en mesure) *Soient (E, T, m) un espace mesuré, $(f_n)_{n\in\mathbb{N}} \subset \mathcal{M}$ et $f \in \mathcal{M}$. On dit que f_n converge en mesure vers f si :*

$$\forall\, \varepsilon > 0, \lim_{n\to+\infty} m(\{x \in E \;\; ; |f(x) - f_n(x)| \ge \varepsilon\}) = 0. \tag{3.6}$$

Cette définition se traduit en langage probabiliste par:

Définition 3.21 (Convergence stochastique ou en probabilité) *Soient* (E, T, p) *un espace probabilisé,* $(X_n)_{n\in\mathbb{N}}$ *une suite de variables aléatoires et* X *une variable aléatoire . On dit que* X_n *converge stochastiquement, ou en probabilité, vers* X *si :*

$$\forall \varepsilon > 0, \lim_{n\to+\infty} p(\{x \in X_n \;\; ; |X(x) - X_n(x)| \geq \varepsilon\}) = 0. \tag{3.7}$$

On peut montrer (cf exercice 3.23) que si $(f_n)_{n\in\mathbb{N}} \subset \mathcal{M}$ converge en mesure vers $f \in \mathcal{M}$ et $(f_n)_{n\in\mathbb{N}}$ converge en mesure vers $g \in \mathcal{M}$, alors $f = g$ p.p.. On montre aussi que si $(f_n)_{n\in\mathbb{N}} \subset \mathcal{M}$ converge en mesure vers $f \in \mathcal{M}$ et $(g_n)_{n\in\mathbb{N}}$ converge en mesure vers $g \in \mathcal{M}$, alors $(f_n + g_n)_{n\in\mathbb{N}} \subset \mathcal{M}$ converge en mesure vers $f + g \in \mathcal{M}$, et, si m est une mesure finie, $(f_n\, g_n)_{n\in\mathbb{N}} \subset \mathcal{M}$ converge en mesure vers $f\,g \in \mathcal{M}$. On montre à l'aide du théorème d'Egorov que si f_n converge vers f presque partout, et si $m(E) < +\infty$, alors f_n converge vers f en mesure. Réciproquement, si f_n converge vers f en mesure, alors il existe une sous-suite de $(f_n)_{n\in\mathbb{N}}$ qui converge vers f presque uniformément (et donc presque partout). Ce second résultat est vrai même si $m(E) = +\infty$ (voir exercice 3.26).

On donne maintenant un résumé des différents types de convergence connus jusqu'à présent avec les relations existantes entre eux; les relations entre convergence presque partout et convergence en mesure (resp. convergence presque sure et convergence stochastique) sont étudiées dans l'exercice 3.26. (On en introduira bientôt encore quelques-unes. . .)

Terminologie "analyste"	**Terminologie "probabiliste"**
convergence simple (cs)	
convergence uniforme (cu)	
convergence presque partout (cpp)	convergence presque sure (cps)
convergence presque uniforme (cpu)	
convergence en mesure (cm)	convergence stochastique (cst)

On a les implications suivantes :

Terminologie "analyste"	**Terminologie "probabiliste"**
(cu) $\Rightarrow$ (cs) $\Rightarrow$ (cpp)	
(cu) $\Rightarrow$ (cpu) $\Rightarrow$ (cpp)	
(cpp) $\Rightarrow$ (cpu) si la mesure est finie	
(cm) $\Rightarrow$ (cpu) (et donc (cpp)) pour une sous-suite	(cst) $\Rightarrow$ (cps) pour une sous-suite
(cpp) $\Rightarrow$ (cm) si la mesure est finie	(cps) $\Rightarrow$ (cst) si la mesure est finie
(cpu) $\Rightarrow$ (cm)	

3.6 Exercices

Exercice 3.1 (Caractérisation des fonctions mesurables) $(\star)$ *Corrigé 38 page 314*
Soient (E, T) un espace mesurable et f une application de E dans $\mathbb{R}$;

1. Montrer que $T_f = \{B \in \mathcal{P}(\mathbb{R})\, ;\, f^{-1}(B) \in T\}$ est une tribu.

2. Soit $\mathcal{C}$ un ensemble qui engendre $\mathcal{B}(\mathbb{R})$, montrer que les deux assertions suivantes sont équivalentes :

   (i) f est mesurable,

   (ii) $f^{-1}(C) \in T$, pour tout $C \in \mathcal{C}$.

Exercice 3.2 (Mesurabilité pour f à valeurs dans $\mathbb{R}$)
Soit Ω un ensemble et $\mathcal{F}$ une σ-algèbre sur Ω. Soit f une application de Ω dans $\mathbb{R}$. On munit $\mathbb{R}$ de la tribu borélienne.

1. Montrer que f est mesurable si et seulement si, pour tout $x \in \mathbb{R}$, $\{\omega \in \Omega,\ f(\omega) < x\} \in \mathcal{F}$.
2. Montrer que f est mesurable si et seulement si, pour tout $x \in \mathbb{R}$, $\{\omega \in \Omega,\ f(\omega) \leq x\} \in \mathcal{F}$.

Exercice 3.3 (Composition de fonctions mesurables) *Corrigé 39 page 314*
Soit (E,T) et (F,S) deux espaces mesurables. Soit $f : E \to F$ et $\varphi : F \to \mathbb{R}$ ($\mathbb{R}$ est muni, comme toujours, de la tribu borélienne). On suppose que f et φ sont mesurables. Montrer que $\varphi \circ f$ est mesurable (de E dans $\mathbb{R}$).

Exercice 3.4 ($\mathbb{R}$ ou $\overline{\mathbb{R}}_+ \ldots$) *Corrigé 40 page 314*
Soit $\varphi : \mathbb{R} \to \mathbb{R}$, $\varphi \geq 0$. On munit $\mathbb{R}$ (au départ et à l'arrivée) de la tribu borélienne. Montrer que φ est mesurable (on dit aussi borélienne) si et seulement si φ est mesurable quand on la considère comme une application de $\mathbb{R}$ dans $\overline{\mathbb{R}}_+$ ($\overline{\mathbb{R}}_+$ étant aussi muni de la tribu borélienne).

Exercice 3.5 (Stabilité de $\mathcal{M}$) *Corrigé 41 page 315*

1. Soient (E,T), (E',T'), (E'',T'') des espaces mesurables, f (resp. g) une application de E dans E' (resp. de E' dans E''). On suppose que f et g sont mesurables. Montrer que $g \circ f$ est une application mesurable de E dans E''.
2. Soit (E,T) un espace mesurable, on munit $\mathbb{R}$ de la tribu des boréliens $\mathcal{B}(\mathbb{R})$; soient f et g des fonctions mesurables de E dans $\mathbb{R}$.
   (a) Montrer que $f^+(= \sup(f,0))$, $f^-(= -\inf(f,0))$ sont des fonctions mesurables de E dans $\mathbb{R}$.
   (b) Montrer que $f+g$, fg et $|f|$ sont des fonctions mesurables de E dans $\mathbb{R}$.
3. Soient (E,T) un espace mesurable, $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables de E dans $\mathbb{R}$. On suppose que la suite $(f_n(x))_{n\in\mathbb{N}}$ converge (dans $\mathbb{R}$) pour tout $x \in E$. On pose $f(x) = lim_{n\to+\infty} f_n(x)$ (pour tout $x \in E$). Montrer que f est une fonction mesurable de E dans $\mathbb{R}$.
4. Soit (E,T) un espace mesurable, on suppose qu'il existe $A \in T$ dont les sous-ensembles ne soient pas tous mesurables. Il existe donc $B \subset A$ t.q. $B \notin T$. Montrer que $h = 1_B - 1_{A\setminus B}$ n'est pas mesurable (de E dans $\mathbb{R}$), alors que $|h|$ l'est.

Exercice 3.6 (Mesurabilité des fonctions continues) *Corrigé 42 page 316*
Soit f une application de $\mathbb{R}$ dans $\mathbb{R}$. On munit $\mathbb{R}$ (au départ et à l'arrivée) de la tribu borélienne

1. On suppose f continue. Montrer que f est mesurable (on dit aussi que f est borélienne).
2. On suppose f continue à droite (resp. gauche). Montrer que f est mesurable.
3. On suppose f croissante. Montrer que f est mesurable.

Exercice 3.7 (Mesurabilité de $1_{\mathbb{Q}}$)
On munit $\mathbb{R}$ de sa tribu borélienne. La fonction $1_{\mathbb{Q}}$ est-elle mesurable ?

Exercice 3.8

Soit (E, T, p) un espace probabilisé et X une variable aléatoire sur (E, T) ;

1. On prend pour X la variable aléatoire nulle, c'est-à-dire $X : E \to \mathbb{R}$, $X(x) = 0$, pour tout $x \in E$. Calculer la loi de probabilité p_X de X. En déduire que la connaissance de p_X ne permet pas en général de déterminer la probabilité p sur E.

2. Montrer que p_X détermine p de manière unique si la tribu engendrée par X (voir la définition 3.6), notée T_X, est égale à T. Cette condition est-elle nécessaire ?

Exercice 3.9
Soit $\mathcal{A}$ une tribu sur un ensemble E et $A \in \mathcal{A}$ tel que : $B \in \mathcal{A}$ et $B \subset A$ implique $B = \emptyset$ ou $B = A$. Montrer que toute fonction mesurable (de E dans $\mathbb{R}$) est constante sur A. En particulier si $\mathcal{A}$ est engendrée par une partition, une fonction mesurable est constante sur chaque élément de la partition. Donner une fonction constante sur tout élément d'une partition mais qui ne soit pas mesurable. [Prendre comme partition de $\mathbb{R}$ tous les singletons. . .]

Exercice 3.10 (Egalité presque partout) *Corrigé 43 page 316*

1. Soient f et g des fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$ et λ la mesure de Lebesgue ; montrer que $f = g$ λ p.p. si et seulement si $f = g$.

2. Soient f et g des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ et δ_0 la mesure de Dirac en 0 ; montrer que $f = g$ δ_0 p.p. si et seulement si $f(0) = g(0)$.

Exercice 3.11 *Corrigé 44 page 317*
Soit $f : \mathbb{R}^N \times \mathbb{R}$ dans $\mathbb{R}$. On munit $\mathbb{R}^p$ de sa tribu borélienne (pour tout $p \in \mathbb{N}^\star$). On suppose que f est mesurable par rapport à $x \in \mathbb{R}^N$, pour tout $y \in \mathbb{R}$, et que f est continue a gauche par rapport a $y \in \mathbb{R}$, pour tout $x \in \mathbb{R}^N$.
Pour $n > 1$ et $p \in \mathbb{Z}$, on pose : $a_p^n = \frac{p}{n}$, $p \in \mathbb{Z}$; on définit la fonction f_n, $n > 1$, de $\mathbb{R}^N \times \mathbb{R}$ dans $\mathbb{R}$ par :

$$f_n(x, y) = f(x, a_p^n), \text{ si } y \in [a_p^n, a_{p+1}^n[$$

1. Montrer que f_n converge simplement vers f lorsque $n \to +\infty$.

2. Montrer que f_n est mesurable. [On pourra utiliser, sans le démontrer, le fait que $A \times B \in \mathcal{B}(\mathbb{R}^{N+1})$ si $A \in \mathcal{B}(\mathbb{R}^N)$ et $B \in \mathcal{B}(\mathbb{R})$. Ceci est démontré dans l'exercice 2.6 page 42 pour $N = 1$ et dans l'exercice 7.1 pour $N > 1$.]

3. Montrer que f est mesurable.

[On pourra se contenter du cas $N = 1$. . .]

Exercice 3.12 (Tribu de Borel sur $\overline{\mathbb{R}}_+$) *Corrigé 45 page 318*

1. Montrer que $\{[0, \beta[, \beta \in \mathbb{R}_+^\star\}$ engendre $\mathcal{B}(\overline{\mathbb{R}}_+)$.

2. Montrer que $\{[0, \beta[, \beta \in \mathbb{Q} \cap \mathbb{R}_+^\star\}$ engendre $\mathcal{B}(\overline{\mathbb{R}}_+)$.

3. (Question plus difficile.) Montrer que $\{]0, \beta[, \beta \in \mathbb{R}_+^\star\}$ n'engendre pas $\mathcal{B}(\overline{\mathbb{R}}_+)$.

Exercice 3.13 *Corrigé 46 page 319*

Soit f une fonction mesurable de $\mathbb{R}$ dans $\mathbb{R}$ ($\mathbb{R}$ est muni de sa tribu borélienne, notée $\mathcal{B}(\mathbb{R})$). On se propose de montrer que le graphe de f est un borélien de $\mathbb{R}^2$. On admettra le résultat suivant, vu dans l'exercice 2.6:

$$A, B \in \mathcal{B}(\mathbb{R}) \Rightarrow A \times B \in \mathcal{B}(\mathbb{R}^2).$$

On munit aussi $\mathbb{R}^2$ de sa tribu borélienne. Pour $x, y \in \mathbb{R}$, on pose $F(x, y) = f(x)$ et $H(x, y) = y$.

1. Montrer que F et H sont mesurables de $\mathbb{R}^2$ dans $\mathbb{R}$.
2. On pose $G(f) = \{(x, y)^t \in \mathbb{R}^2;\ y = f(x)\}$ ($G(f)$ est donc le graphe de f). Montrer que $G(f) \in \mathcal{B}(\mathbb{R}^2)$.

Exercice 3.14 (mesurabilité au sens de Lusin) *Corrigé 47 page 319*
Soit m une mesure sur $\mathcal{B}(\mathbb{R}^N)$, finie sur les compacts de $\mathbb{R}^N$. On rappelle (cf. cours) que m est nécessairement régulière (c'est-à-dire que pour tout $A \in \mathcal{B}(\mathbb{R}^N)$ et pour tout $\varepsilon > 0$, il existe F fermé et O ouvert t.q. $F \subset A \subset O$ et $m(O \setminus F) < \varepsilon$).

Soit $f \in \mathbb{R}^N \to \mathbb{R}$. On dit que f est "mesurable au sens de Lusin" si pour tout compact K et pour tout $\varepsilon > 0$, il existe K_1 compact, $K_1 \subset K$, t.q. $m(K \setminus K_1) \leq \varepsilon$ et $f_{|K_1} \in C(K_1, \mathbb{R})$.

1. On suppose, dans cette question, que $f = 1_A$ avec $A \in \mathcal{B}(\mathbb{R}^N)$. Montrer que f est mesurable au sens de Lusin. [Construire K_1 avec K, F et O, où F et O sont donnés par la régularité de m appliquée à l'ensemble A.]
2. On suppose, dans cette question, que f est étagée (c'est-à-dire $f \in \mathcal{E}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$. Montrer que f est mesurable au sens de Lusin.
3. On suppose que f est mesurable (c'est-à-dire $f \in \mathcal{M}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$. Montrer que f est mesurable au sens de Lusin. [On rappelle qu'une fonction mesurable est limite simple de fonctions étagées. On pourra utiliser le théorème d'Egorov, Théorème 3.2, et la question précédente.]

Exercice 3.15 (V.a. mesurable par rapport à une autre v.a.) *Corrigé 48 page 321*
Dans cet exercice, on démontre le théorème 3.1. Soit X et Y deux variables aléatoires réelles définies sur un espace de probabilités $(\Omega, \mathcal{A}, P)$. On veut veut montrer que Y est mesurable par rapport à la tribu engendrée par X (notée $\tau(X)$ si et seulement si il existe une fonction borélienne f de $\mathbb{R}$ dans $\mathbb{R}$ telle que $Y = f(X)$ (c'est-à-dire , plus précisément, que $Y = f \circ X$).

1. Montrer que si Y est de la forme $Y = f(X)$ où f est une fonction borélienne de $\mathbb{R}$ dans $\mathbb{R}$, alors Y est $\tau(X)$-mesurable.

On suppose maintenant que Y est $\tau(X)$-mesurable.

2. On suppose, dans cette question, qu'il existe une suite de réels (a_j) tels que $a_j \neq a_k$ pour $j \neq k$ et une suite d'événements (A_j) disjoints deux à deux tels que
$$Y = \sum_j a_j 1_{A_j}.$$
On suppose aussi que $\cup_j A_j = \Omega$. Montrer que, pour tout j, $A_j \in \tau(X)$ et qu'il existe une fonction borélienne $f : \mathbb{R} \to \mathbb{R}$ telle que $Y = f(X)$.

3. Soit n un entier. On définit la fonction $\phi_n : \mathbb{R} \to \mathbb{R}$ par: $\phi_n(x) = \frac{1}{n}[nx]$ où $[\cdot]$ désigne la partie entière. ($[x]$ est le plus grand entier inférieur ou égal à x.)

   (a) Montrer que, pour tout $x \in \mathbb{R}$, $\phi_n(x)$ converge vers x, quand $n \to \infty$.

   (b) On pose $Y_n = \phi_n(Y)$. Montrer que Y_n est $\tau(X)$ mesurable.

4. Terminer la preuve du théorème.

Maintenant, on se demande dans quelle mesure la fonction f est unique. On note P_X la loi de X.

5. Soit f et g deux fonctions boréliennes t.q. $Y = f(X) = g(X)$. Montrer que

$$P_X(f = g) = 1.$$

Exercice 3.16 (Composition de v.a.) *Corrigé 49 page 323*
Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Soit N une variable aléatoire à valeurs dans $\mathbb{N}^\star$ et $(Y_n)_{n\in\mathbb{N}}$ une suite de variables alélatoires réelles. (c'est-à-dire à valeurs dans $\mathbb{R}$, muni de la tribu des boréliens). On définit Z par

$$\forall \omega \in \Omega\,, \quad Z(\omega) = Y_{N(\omega)}(\omega).$$

Montrer que Z est une variable aléatoire.

Exercice 3.17 (Evénements, tribus et v.a. indépendantes) *Corrigé 50 page 323*
Soit $(E, \mathcal{A}, P)$ un espace probabilisé.

1. (Indépendance de 2 évènements) Soit $A_1, A_2 \in \mathcal{A}$. Montrer que A_1 et A_2 sont indépendants (c'est-à-dire $P(A_1 \cap A_2) = P(A_1)P(A_2)$) si et seulement si les tribus $\tau(\{A_1\})$ et $\tau(\{A_2\})$ sont indépendantes (c'est-à-dire $P(B_1 \cap B_2) = P(B_1)P(B_2)$ pour tout $B_1 \in \tau(\{A_1\})$ et $B_2 \in \tau(\{A_2\})$).

2. (Indépendance de n évènements, $n \geq 2$) Soit $n \geq 2$, $A_1, \ldots, A_n \in \mathcal{A}$. Montrer que les événements $A_1, \ldots, A_n$ vérifient "$P(\cap_{i\in I} A_i) = \prod_{i\in I} P(A_i)$ pour tout $I \subset \{1, \ldots, n\}$" si et seulement si les tribus $\tau(\{A_1\}), \ldots, \tau(\{A_n\})$ sont indépendantes (c'est-à-dire $P(\cap_{i=1}^n B_i) = \prod_{i=1}^n P(B_i)$ pour tout $B_i \in \tau(\{A_i\})$, $i \in \{1, \ldots, n\}$).

3. En donnant un exemple (avec $n \geq 3$), montrer que l'on peut avoir n évévements, notés $A_1, \ldots, A_n$, indépendants deux à deux, sans que les événements $A_1, \ldots, A_n$ soient indépendants.

4. Soit $A \in \mathcal{A}$.

   (a) On suppose que $A \in \mathcal{A}_1$ et $A \in \mathcal{A}_2$ et que $\mathcal{A}_1$ et $\mathcal{A}_2$ sont deux tribus indépendantes (et contenues dans $\mathcal{A}$). Montrer que $P(A) \in \{0, 1\}$.

   (b) Montrer que $P(A) \in \{0, 1\}$ si et seulement si A est indépendant de tous les éléments de $\mathcal{A}$.

5. Soit $n \geq 1$ et $A_1, \ldots, A_n \in \mathcal{A}$. Montrer que les événements $A_1, \ldots, A_n$ sont indépendants si et seulement si les v.a. $1_{A_1}, \ldots, 1_{A_n}$ sont indépendantes.

Exercice 3.18 (Indépendance 3 par 3 et dépendance globale)
Trouver un espace de probabilités et 4 v.a.r. prenant leurs valeurs dans $\{-1, 1\}$ t.q. les 4 v.a. soient 3 par 3 indépendantes mais ne soient pas indépendantes. [On pourra considérer des produits de v.a. indépendantes.]

Exercice 3.19 (Transformation d'une v.a.r. de loi uniforme en une v.a.r. de loi donnée)
Soit $(E, \mathcal{A}, P)$ un espace de probabilités, X une v.a.r. et U une v.a.r. de loi $\mathcal{U}([0,1])$, Soit F la fonction de répartition de X (i. e. $F(x) = P(X \le x)$ pour $x \in \mathbb{R}$). On définit G de $\mathbb{R}$ dans $\mathbb{R}$ de la manière suivante :

$$G(u) = \inf\{x \in \mathbb{R};\ F(x) \ge u\},\ \text{si } u \in]0,1[,$$
$$G(u) = 0,\ \text{si } u \notin]0,1[.$$

Montrer que la v.a.r. $Y = G(U)$ a la même loi que X. [On pourra montrer que $P(G(U) \le x) = P(U \le F(x))$, pour tout $x \in \mathbb{R}$.]

Exercice 3.20 (Limite croissante d'une suite de v.a.r.)
Soit $(E, \mathcal{A}, P)$ un espace de probabilités, Y une v.a.r., $(X_n)_{n\in\mathbb{N}}$ une suite de v.a.r. et X une application de E dans $\mathbb{R}$. On suppose que $X_n \uparrow X$, quand $n \to \infty$, et que X_n et Y sont indépendantes, pour tout $n \in \mathbb{N}$. Montrer que X est une v.a.r. et que X et Y sont indépendantes. (N.B. La conclusion est encore vraie sans la croissance de la suite X_n.)

Exercice 3.21 (Construction de v.a. indépendantes de lois uniformes sur un intervalle)
Soit $(E, \mathcal{A}, P)$ un espace de probabilités.

1. Soit $(U_n)_{n\in\mathbb{N}^\star}$, une suite de v.a.r.i.i.d. avec $P(U_n = 0) = P(U_n = 1) = 1/2$. Montrer que V, définie par $V = \sum_{n\ge1} U_n 2^{-n}$ est une v.a. de loi $\mathcal{U}([0,1])$.
2. Soit $U_{n,k}$, $n,k \ge 1$, des v.a.r.i.i.d. avec $P(U_{n,k} = 0) = P(U_{n,k} = 1) = 1/2$. Montrer les v. a. V_n, $n \ge 1$ définies par $V_n = \sum_{k\ge1} U_{n,k} 2^{-k}$ sont des v.a.r.i.i.d. de loi $\mathcal{U}([0,1])$.

Exercice 3.22 (Loi du "produit de la loi exponentielle par ± 1")
Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et X, Y deux v.a.r. indépendantes. On suppose que X suit une loi exponentielle de paramètre λ (avec $\lambda > 0$), c'est-à-dire que P_X est une mesure de densité par rapport à la mesure de Lebesgue et cette densité est la fonction f définie par $f(x) = \lambda \exp(-\lambda x) 1_{]0,+\infty[}(x)$ pour $x \in \mathbb{R}$. On suppose que Y est t. q. $P(Y = 1) = P(Y = -1) = 1/2$. Donner la loi de XY.

Exercice 3.23 (Convergence en mesure) $(\star\star)$ *Corrigé 51 page 326*
Soient (E, T, m) un espace mesuré, $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables de E dans $\mathbb{R}$.

1. Montrer que si il existe f et g fonctions mesurables de E dans $\mathbb{R}$ telles que $(f_n)_{n\in\mathbb{N}}$ converge en mesure vers f et g, alors $f = g$ p.p..

   [*On pourra commencer par montrer que, pour tout* $\delta > 0$, $m(\{x \in E\,;\,|f(x) - g(x)| > \delta\}) = 0$].
2. Montrer que si $(f_n)_{n\in\mathbb{N}} \subset \mathcal{M}$ converge en mesure vers $f \in \mathcal{M}$ et $(g_n)_{n\in\mathbb{N}} \subset \mathcal{M}$ converge en mesure vers $g \in \mathcal{M}$, alors $(f_n + g_n)_{n\in\mathbb{N}} \subset \mathcal{M}$ converge en mesure vers $f + g \in \mathcal{M}$.
3. On suppose maintenant que m est une mesure finie. Montrer que si $(f_n)_{n\in\mathbb{N}} \subset \mathcal{M}$ converge en mesure vers $f \in \mathcal{M}$ et $(g_n)_{n\in\mathbb{N}} \subset \mathcal{M}$ converge en mesure vers g, alors $(f_n\, g_n)_{n\in\mathbb{N}} \subset \mathcal{M}$ converge en mesure vers $f\, g \in \mathcal{M}$.

   [*On pourra commencer par montrer que, si* $(f_n)_{n\in\mathbb{N}} \subset \mathcal{M}$ *converge en mesure vers* $f \in \mathcal{M}$, *alors, pour tout* $\varepsilon > 0$, *il existe* n_0 *et* $k_0 \in \mathbb{N}$ *tels que, si* $n \ge n_0$ *et* $k \ge k_0$, *on a* $m(\{x \in E\,;\,|f_n(x)| \ge k\}) \le \varepsilon$]. Donner un contre-exemple au résultat précédent lorsque $m(E) = \infty$.

Exercice 3.24 (Convergence presque uniforme et convergence p.p.) *Corrigé 52 page 328*

Soient (E, T, m) un espace mesuré, $(f_n)_{n\in\mathbb{N}} \subset \mathcal{M}$ (c'est-à-dire une suite de fonctions mesurables de E dans $\mathbb{R}$) et $f \in \mathcal{M}$. On suppose que $f_n \to f$ presque uniformément (c'est à dire que pour tout $\varepsilon > 0$ il existe $A \in T$ t.q. $m(A) \le \varepsilon$ et $f_n \to f$ uniformément sur A^c). Montrer que $f_n \to f$ p.p., quand $n \to \infty$.

Exercice 3.25 (Théorème d'Egorov) ($\star\star$) *Corrigé 53 page 328*
Soient (E, T, m) un espace mesuré fini, $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables de E dans $\mathbb{R}$, et f une fonction mesurable de E dans $\mathbb{R}$. On suppose que $f_n \to f$ p.p., lorsque $n \to +\infty$.
Pour $j \in \mathbb{N}^\star$ et $n \in \mathbb{N}$, on définit :

$$A_{n,j} = \{x\,;\, |f(x) - f_n(x)| \ge \frac{1}{j}\}, \text{ et } B_{n,j} = \bigcup_{p\ge n} A_{p,j} \tag{3.8}$$

1. Montrer que à j fixé, $\lim_{n\to+\infty} m(B_{n,j}) = 0$.
2. Montrer que, pour tout $\varepsilon > 0$, il existe A tel que $m(A) \le \varepsilon$ et $f_n \to f$ uniformément sur A^c lorsque $n \to +\infty$. En déduire le théorème d'Egorov (théorème 3.2).
   [*On cherchera A sous la forme :* $\bigcup_{j\in\mathbb{N}^\star} B_{n_j,j}$, *avec un choix judicieux de n_j.*]
3. Montrer, par un contre exemple, qu'on ne peut pas prendre $\varepsilon = 0$ dans la question précédente.
4. Montrer, par un contre exemple, que le résultat du théorème d'Egorov est faux lorsque $m(E) = +\infty$.

Exercice 3.26 (Convergence en mesure et convergence p.p.) *Corrigé 54 page 330*
Soient (E, T, m) un espace mesuré, $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables de E dans $\mathbb{R}$, et f une fonction mesurable de E dans $\mathbb{R}$. On rappelle que, par définition, la suite $(f_n)_{n\in\mathbb{N}}$ converge en mesure vers f si :

$$\forall\, \varepsilon > 0, \lim_{n\to+\infty} m(\{x \in E; |f_n(x) - f(x)| > \varepsilon\}) = 0. \tag{3.9}$$

1. On suppose ici que $m(E) < +\infty$.
   (a) Montrer que si $(f_n)_{n\in\mathbb{N}}$ tend vers f presque partout, alors $(f_n)_{n\in\mathbb{N}}$ tend vers f en mesure [*Utiliser le théorème d'Egorov.*]
   (b) (Question plus difficile.) Montrer par un contrexemple que la réciproque de la question précédente est fausse.
2. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables de E dans $\mathbb{R}$. On dit que $(f_n)_{n\in\mathbb{N}}$ est de Cauchy en mesure si :
   $$\forall\, \varepsilon > 0, \forall \delta > 0, \exists n \in \mathbb{N}; p, q \ge n \Rightarrow m(\{x \in E; |f_p(x) - f_q(x)| > \varepsilon\}) \le \delta. \tag{3.10}$$
   Montrer que si la suite $(f_n)_{n\in\mathbb{N}}$ converge en mesure vers f, alors $(f_n)_{n\in\mathbb{N}}$ est de Cauchy en mesure.
3. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de Cauchy en mesure ; montrer qu'il existe une fonction mesurable g et une sous-suite $(f_{n_k})_{k\in\mathbb{N}}$, telles que pour tout $\varepsilon > 0$, il existe $A \in T$ vérifiant $m(A) \le \varepsilon$ et tel que $(f_{n_k})_{k\in\mathbb{N}}$ converge uniformément vers g sur A^c. [*On pourra construire la sous-suite $(f_{n_k})_{k\in\mathbb{N}}$ de telle sorte que $m(A_k) \le 2^{-k}$, avec $A_k = \{x \in E; |f_{n_{k+1}}(x) - f_{n_k}(x)| > 2^{-k}\}$, et chercher A sous la forme $\bigcup_{k\ge p} A_k$, où p est convenablement choisi.*]

4. Montrer que si $(f_n)_{n\in\mathbb{N}}$ converge en mesure vers f, il existe une sous-suite qui converge vers f presque partout. [*On pourra commencer par montrer que la suite $(f_{n_k})_{k\in\mathbb{N}}$ construite à la question précédente converge presque partout et en mesure.*]

Exercice 3.27
Soient (E, T, m) un espace mesuré fini. On pose, pour f et g fonctions mesurables de E dans $\mathbb{R}$ (c'est-à-dire $f, g \in \mathcal{M}$) :

$$d(f, g) = \int \frac{|f - g|}{1 + |f - g|} dm.$$

1. Montrer que d est bien définie et prend ses valeurs dans $\mathbb{R}_+$ (c'est-à-dire que $\frac{|f-g|}{1+|f-g|} \in \mathcal{L}^1_{\mathbb{R}}(E, T, m)$ pour tout $f, g \in \mathcal{M}$) et que d est une semi-distance sur $\mathcal{M}$ (c'est-à-dire que $d(f, g) = d(g, f)$, pour tout $f, g \in \mathcal{M}$, et que $d(f, h) \leq d(f, g) + d(g, h)$, pour tout $f, g, h \in \mathcal{M}$).
2. Soient $(f_n)_{n\in\mathbb{N}} \subset \mathcal{M}$ et $f \subset \mathcal{M}$. Montrer que f_n converge en mesure vers f lorsque $n \to +\infty$ si et seulement si $\lim_{n\to+\infty} d(f_n, f) = 0$. [Il est probablement utile de considérer, pour $\varepsilon > 0$, les ensembles $A_n = \{x \in E; |f_n(x) - f(x)| > \varepsilon\}$.]

Exercice 3.28 (Mesurabilité d'une limite p.p.)
Soient (E, T, m) un espace mesuré, $(f_n)_{n\in\mathbb{N}} \subset \mathcal{M}(E, T)$ et $f : E \to \mathbb{R}$. On suppose que $f_n \to f$ p.p.. Montrer que $f \in \mathcal{M}(E, \overline{T})$, où $(E, \overline{T}, \overline{m})$ est le compété de (E, T, m) (voir le théorème 2.1). En donnant un exemple (c'est-à-dire en choisissant convenablement (E, T, m), $(f_n)_{n\in\mathbb{N}}$ et f), montrer qu'on peut avoir $f \notin \mathcal{M}(E, T)$.

Exercice 3.29 (Essentiellement uniforme versus presque uniforme) *Corrigé 55 page 331*
Soit (E, T, m) un espace mesuré. Pour $f \in \mathcal{M}$, on pose $A_f = \{C \in \mathbb{R}, |f| \leq C \text{ p.p.}\}$. Si $A_f \neq \emptyset$, on pose $\|f\|_\infty = \inf A_f$. Si $A_f = \emptyset$, on pose $\|f\|_\infty = \infty$.

1. Soit $f \in \mathcal{M}$ t.q. $A_f \neq \emptyset$. Montrer que $\|f\|_\infty \in A_f$.
2. Soient $(f_n)_{n\in\mathbb{N}} \subset \mathcal{M}$ et $f \in \mathcal{M}$.
   (a) On suppose, dans cette question, que $\|f_n - f\|_\infty \to 0$ quand $n \to \infty$ (on dit que $f_n \to f$ essentiellement uniformément). Montrer que $f_n \to f$ presque uniformément.
   (b) En donnant un exemple (c'est-à-dire en choisissant convenablement (E, T, m), $(f_n)_{n\in\mathbb{N}}$ et f), montrer qu'on peut avoir $f_n \to f$ presque uniformément, quand $n \to \infty$, et $\|f_n - f\|_\infty \not\to 0$.

Exercice 3.30
Soit $\mathcal{A}$ la classe des sous-ensembles de $\mathbb{Z}$ tels que

$$\text{pour } n > 0, \quad 2n \in A \Longleftrightarrow 2n + 1 \in A.$$

1. Montrer que $\mathcal{A}$ est une tribu.
2. Montrer que l'application $\varphi : n \mapsto n + 2$ est une bijection de $\mathbb{Z}$ dans $\mathbb{Z}$, mesurable quand $\mathbb{Z}$ est muni (au départ et à l'arrivée) de la tribu $\mathcal{A}$ (c'est à dire t.q. que $\varphi^{-1}(A) \in \mathcal{A}$ pour tout $A \in \mathcal{A}$). Montrer que l'inverse de φ n'est pas mesurable.

Exercice 3.31

On considère des applications f de E dans $\mathbb{R}$. On note $\sigma(f) = \{f^{-1}(A),\ A \in \mathcal{B}(\mathbb{R})\}$ ($\sigma(f)$ est la tribu image réciproque de $\mathcal{B}(\mathbb{R})$ par f).

1. Décrire $\sigma(f)$ dans chacun des cas suivants:

   (a) $E = \mathbb{R}$ et $f(x) = x$ ou $f(x) = x^2$ ou $f(x) = |x|$.

   (b) $E = \mathbb{R}^2$ et $f(x, y) = x + y$.

2. Montrer que les singletons sont tous dans $\sigma(f)$ si et seulement si f est injective.

3. Dans le cas général, montrer qu'une fonction g de E dans $\mathbb{R}$ est $\sigma(f)$-mesurable si et seulement si il existe une fonction borélienne φ de $\mathbb{R}$ dans $\mathbb{R}$ t.q. $g = \varphi \circ f$. [On pourra commencer par le cas où g est étagée, puis utiliser un argument de limite].

4. Montrer que l'on a toujours une fonction bornée g t.q. $\sigma(g) = \sigma(f)$.

Exercice 3.32 (Mesurabilité des troncatures) *Corrigé 56 page 332*
Soit $(X, \mathcal{T})$ un espace mesurable et f une fonction mesurable de X dans $\mathbb{R}$ ($\mathbb{R}$ est muni, comme toujours quand on ne le précise pas, de la tribu borélienne). Pour $a > 0$, on définit la fonction "tronquée" :

$$f_a(x) = \begin{cases} a & \text{si} \quad f(x) > a \\ f(x) & \text{si} \quad |f(x)| \le a \\ -a & \text{si} \quad f(x) < -a \end{cases}$$

Montrer que f_a est mesurable.

Exercice 3.33
Soit (E, T) un espace mesurable (on munit $\mathbb{R}$ de la tribu des boréliens $\mathcal{B}(\mathbb{R})$, comme toujours). Soient $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables de E dans $\mathbb{R}$, et A l'ensemble des points $x \in E$ tels que $(f_n(x))_{n\in\mathbb{N}}$ ne soit pas de Cauchy. Montrer que A est mesurable (i.e. $A \in T$).

Exercice 3.34
Soit $\varphi : \mathbb{R}^2 \to \mathbb{R}$ $\mathcal{B}(\mathbb{R}^2)$-mesurable. Pour $x \in \mathbb{R}$, on pose : $f_\varphi(x) = \varphi(x, x)$.

1. Montrer que f_φ est $\mathcal{B}(\mathbb{R})$-mesurable.

2. Soient φ et $\psi : \mathbb{R}^2 \to \mathbb{R}$ $\mathcal{B}(\mathbb{R}^2)$-mesurables. Montrer que $\varphi = \psi$ λ_2-p.p. $\not\Rightarrow f_\varphi = f_\psi$ λ-p.p.. (λ_2 est la mesure de Lebesgue sur $\mathcal{B}(\mathbb{R}^2)$ dont on suppose l'existence).

3. Soient φ et $\psi : \mathbb{R}^2 \to \mathbb{R}$ $\mathcal{B}(\mathbb{R}^2)$-mesurables t.q. :

   (a) $\varphi(x, .)$ et $\psi(x, .)$ sont continues p.p. en $x \in \mathbb{R}$

   (b) $\varphi(, .y)$ et $\psi(., y)$ sont mesurables pour tout $y \in \mathbb{R}$

   (Ces fonctions sont dites de "Carathéodory".)

   (a) Montrer que φ et ψ sont $\mathcal{B}(\mathbb{R}^2)$-mesurables.

   (b) Montrer que $\varphi = \psi$ λ_2-p.p. $\Rightarrow \varphi(x, .) = \psi(x, .)$ partout, p.p. en $x \in \mathbb{R}$. En déduire que si $\varphi = \psi$ λ_2-p.p, alors $f_\varphi = f_\psi$ λ-p.p..

Exercice 3.35 (Exemple de tribu engendrée) *Corrigé 57 page 332*

Dans cet exercice, on s'intéresse à la tribu $\tau(X)$ engendrée par la variable aléatoire X définie sur Ω, muni de la tribu $\mathcal{A}$, à valeurs dans $\mathbb{R}$, muni de la tribu borélienne.

1. (Cas d'un lancer de dé) Dans cette question, $\Omega = \{1, 2, 3, 4, 5, 6\}$, $\mathcal{A} = \mathcal{P}(\Omega))$ et X est la variable aléatoire définie par $X(\omega) = 1$ lorsque ω est pair, $X(\omega) = 0$ sinon. Montrer que $\tau(X)$ est formé de 4 éléments.

2. (Cas de n tirages à pile ou face) Soit $n \in \mathbb{N}^\star$, $\Omega = \{0, 1\}^n$, $\mathcal{A} = \mathcal{P}(\Omega))$ et $k \in \{1, \cdots, n\}$. La variable aléatoire X représente le k-ième tirage, X est donc l'application $\omega = (\omega_1, \cdots, \omega_n) \mapsto \omega_k$. Montrer que $\tau(X)$ est ici aussi formé de 4 éléments.

3. Dans cette question, on prend $\Omega = \mathbb{R}$, $\mathcal{A} = \mathcal{B}(\mathbb{R})$ et, pour tout $\omega \in \Omega$, $X(\omega) = \omega - [\omega]$, où $[\omega]$ désigne la partie entière de ω (c'est-à-dire $[\omega] = \max\{n \in \mathbb{Z}, \text{ t.q. } n \leq \omega\}$. Si C est un borélien inclus dans $[0, 1[$ (ce qui est équivalent à dire $C \in \mathcal{B}([0, 1[))$, on pose $\varphi(C) = \cup_{k \in \mathbb{Z}} C_k$, avec $C_k = \{x + k, x \in C\}$. Montrer que $\tau(X) = \{\varphi(C), C \in \mathcal{B}([0, 1[)\}$.

Exercice 3.36 (Tribu et partition)
Soit Ω un ensemble. On appelle partition de Ω une famille finie ou dénombrable de parties non vides de Ω et disjointes deux à deux et dont l'union est égale à Ω. Les éléments d'une partition sont appelés atomes.

1. Soit $a = \{A_i \,;\, i \in I\}$ une partition de Ω et soit $\tau(a)$ la tribu engendrée par a. Montrer que
$$\tau(a) = \{\bigcup_{j \in J} A_j \text{ où } J \subset I\}\,.$$
En déduire qu'une v.a. réelle est $\tau(a)$-mesurable si et seulement si elle est constante sur tous les atomes de a.

   Une partition a est dite plus fine qu'une partition b si tous les atomes de b s'écrivent comme union d'atomes de a.

2. Montrer que si a est plus fine que b et si b est plus fine que a alors a et b sont égales.

3. Montrer que si a et b sont deux partitions telles que $\tau(a) = \tau(b)$ alors a et b sont égales.

Exercice 3.37 (Fonctions constantes) *Corrigé 58 page 333*
Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilité et X une variable aléatoire (réelle). Pour $a \in \mathbb{R}$, on pose $\varphi(a) = P(X^{-1}(] - \infty, a]))$ (on note souvent $X^{-1}(] - \infty, a]) = \{X \leq a\}$). La fonction φ est donc la fonction de répartition de la probabilité P_X.

1. Montrer que φ est une fonction croissante de $\mathbb{R}$ dans $\mathbb{R}$ et que $\lim_{a \to +\infty} \varphi(a) = 1$, $\lim_{a \to -\infty} \varphi(a) = 0$.

On suppose maintenant que, pour tout $B \in \mathcal{B}(\mathbb{R})$, on a $P(X^{-1}(B)) = 0$ ou 1.

2. Montrer qu'il existe $\alpha \in \mathbb{R}$ t.q. $X = \alpha$ p.s..