# PROCESSUS MARKOVIENS 1

Chaînes de Markov

Processus de Markov à temps continu et espace d'état discret

Nuages poissonniens

Thomas DUQUESNE

2023-2024

# Table des matières

# Chapitre I

# Chaînes de Markov.

## I.1 Introduction.

### I.1.a Premières définitions.

Dans tout ce chapitre, $E$ désigne un espace *dénombrable non-vide*, c'est-à-dire fini ou en bijection avec $\mathbb{N}$. On considère des suites de variables à valeurs dans un tel espace $E$ qui est appelé l'*espace d'états*. On le munit de la tribu de tous ses sous-ensembles, tribu notée $\mathscr{P}(E) = \{B \subset E\}$. Par conséquent, si $(E', \mathscr{E}')$ est un espace mesurable, toute fonction $g : E \to E'$ est $(\mathscr{P}(E), \mathscr{E}')$-mesurable. On note les points de $E$ de façon générique par $i, j, k, i', i_0, i_n, \ldots$ etc. Si $f : E \to [0, \infty]$, on pose

$$\sum_{i \in E} f(i) = \sup \left\{ \sum_{i \in F} f(i) \, ; \; F \subset E, F \text{ fini} \right\} .$$

Si $E$ est fini, $\sum_{i \in E} f(i)$ n'est qu'une somme finie. Si $E$ est infini, alors quelle que soit la bijection $\gamma : \mathbb{N} \to E$, $\sum_{i \in E} f(i)$ vaut la série $\sum_{n \geq 0} f(\gamma(n))$.

Une autre façon de voir cela est d'introduire la *mesure de comptage* $\#$ sur $E$ donnée pour tout $B \subset E$ par $\#(B) = n$ si $B$ compte $n$ éléments et $\#(B) = \infty$ si $B$ est un sous-ensemble infini. On vérifie facilement que $\# : \mathscr{P}(E) \to [0, \infty]$ est une mesure positive. Comme $E$ est dénombrable, on a $\# = \sum_{i \in E} \delta_i$, où $\delta_i$ désigne la masse de Dirac en $i$. On note que c'est une mesure sigma-finie. On vérifie facilement que $\sum_{i \in E} f(i) = \int_E f \, d\#$.

Si $f : E \to \mathbb{R}$ est telle que $\sum_{i \in E} |f(i)| < \infty$, alors, $\sum_{i \in E} (f(i))_+ < \infty$ et $\sum_{i \in E} (f(i))_- < \infty$ et on pose

$$\sum_{i \in E} f(i) = \sum_{i \in E} (f(i))^+ - \sum_{i \in E} (f(i))^- .$$

Si $E$ est fini, $\sum_{i \in E} f(i)$ n'est qu'une somme finie. Si $E$ est infini, on vérifie que quelle que soit la bijection $\gamma : \mathbb{N} \to E$, $\sum_{i \in E} f(i)$ vaut la série $\sum_{n \geq 0} f(\gamma(n))$ qui est absolument convergente. On vérifie également que $f$ est $\#$-intégrable et que

$$\sum_{i \in E} f(i) = \int_E f \, d\# .$$

Sans faire explicitement mention de la mesure de comptage $\#$, on lui appliquera, si nécessaire, les théorèmes de convergence monotone, dominée, le lemme de Fatou et les divers résultats d'interversion ainsi que Fubini.

Une *mesure positive* sur l'espace mesurable $(E, \mathscr{P}(E))$ est simplement donnée par une fonction $\mu : E \to [0, \infty]$. C'est une mesure de probabilité ssi $\sum_{i \in E} \mu(i) = 1$. On note $\mathscr{M}_1(E)$ l'ensemble des mesures de probabilité sur $(E, \mathscr{P}(E))$.

**Définition I.1.1** Un tableau de nombres réels indexés par $E \times E$, noté $Q = (p(i,j))_{i,j \in E}$ est une *matrice de transition* si

$$\forall i, j \in E \ , \quad p(i,j) \geq 0 \quad \text{et} \quad \sum_{j \in E} p(i,j) = 1 \ ,$$

Autrement dit, $Q$ est une matrice de transition si pour tout $i \in E$, $p(i, \cdot)$ est une mesure de probabilité sur $E$. □

Soit une fonction $f : E \to [0, \infty]$ et soit $Q = (p(i,j))_{i,j \in E}$, une matrice de transition. On définit la fonction $Q.f : E \to [0, \infty]$ par

$$\forall i \in E \ , \quad (Q.f)(i) = \sum_{j \in E} p(i,j) f(j) \ . \tag{I.1}$$

avec la convention $0 \times \infty = 0$. Si $f : E \to \mathbb{R}$, la fonction $Q.f$ est bien définie par (I.1) si

$$\forall i \in E \ , \quad \sum_{j \in E} p(i,j) |f(j)| < \infty \ .$$

Signalons que le produit $f \mapsto Q.f$ préserve les applications bornées. Plus précisément, pour tout $f : E \to \mathbb{R}$, on pose

$$\|f\|_\infty = \sup_{i \in E} |f(i)| \ .$$

Si $\|f\|_\infty < \infty$, alors $Q.f$ est bien définie, c'est aussi une fonction bornée et on a

$$\|Q.f\|_\infty \ \leq \|f\|_\infty \ . \tag{I.2}$$

En effet, $|Q.f(i)| \leq \sum_{j \in E} p(i,j)|f(j)| \leq \|f\|_\infty \sum_{j \in E} p(i,j) = \|f\|_\infty$, pour tout $i \in E$.

Dans tout ce chapitre $(\Omega, \mathscr{F})$ est l'espace mesurable sur lequel sont définies toutes les variables aléatoires, sauf si le contraire est explicitement mentionné. Ainsi, une variable aléatoire $Y$ à valeurs dans $E$ désignera une fonction $Y : \Omega \to E$, telle que $\{Y = i\} \in \mathscr{F}$, pour tout $i \in E$. La probabilité sur $(\Omega, \mathscr{F})$ sous laquelle on travaille sera en revanche toujours explicitement spécifiée.

**Définition I.1.2** **(1)** Soit $Q$ une matrice de transition sur $E$ et soit $\mu \in \mathscr{M}_1(E)$. Soit $(\mathscr{F}_n)_{n \geq 0}$ une filtration sur $(\Omega, \mathscr{F})$. Soit $\mathbf{X} = (X_n)_{n \geq 0}$ une suite de v.a. à valeurs dans $E$. Soit $\mathbf{P}$, une mesure de probabilité sur $(\Omega, \mathscr{F})$. On dit que sous $\mathbf{P}$, $\mathbf{X}$ est *une chaîne de Markov (homogène) relativement à* $(\mathscr{F}_n)_{n \geq 0}$*, de matrice de transition* $Q$ *et de loi d'entrée* $\mu$, si les conditions suivantes sont satisfaites.

$(a)$ $\mathbf{X}$ est $(\mathscr{F}_n)_{n \geq 0}$-adaptée.

$(b)$ $\forall i \in E$, $\mathbf{P}(X_0 = i) = \mu(i)$.

$(c)$ Pour tout $n \in \mathbb{N}$, et toute fonction $f : E \to \mathbb{R}$ positive ou bornée, on a

$$\mathbf{P}\text{-p.s.} \quad \mathbf{E}\big[f(X_{n+1}) \,\big|\, \mathscr{F}_n\big] = (Q.f)(X_n) \ . \tag{I.3}$$

Lorsqu'aucune filtration n'est précisée, il est implicitement supposé que la filtration à choisir dans (I.3) est

$$\mathscr{F}_n^{\mathbf{X}} := \sigma(X_0, \ldots, X_n) \ , \quad n \ \in \mathbb{N} \ ,$$

qui est la filtration naturelle de $\mathbf{X}$.

**(2)** On se donne une suite de matrices de transition $Q_n = (p^{(n)}(i,j))_{i,j \in E}$, $n \in \mathbb{N}$. On dit que sous $\mathbf{P}$, $\mathbf{X}$ est *une chaîne de Markov inhomogène relativement à* $(\mathscr{F}_n)_{n \geq 0}$*, de matrices de transition* $(Q_n)_{n \in \mathbb{N}}$ *et de loi d'entrée* $\mu$, si elle satisfait $(a)$, $(b)$ et

$(c')$ Pour tout $n \in \mathbb{N}$, et toute fonction $f : E \to \mathbb{R}$ positive ou bornée, on a

$$\mathbf{P}\text{-p.s.} \quad \mathbf{E}\big[f(X_{n+1}) \,\big|\, \mathscr{F}_n\,\big] = (Q_n.f)(X_n)\ . \tag{I.4}$$

À part dans quelques exercices et dans la section sur le recuit simulé, nous ne considérerons que des chaînes homogènes. □

**Remarque I.1.3** On suppose que $\mathbf{X} = (X_n)_{n\geq 0}$ est, sous $\mathbf{P}$, une chaîne de Markov relativement à une filtration $(\mathscr{F}_n)_{n\geq 0}$, de matrice de transition $Q$ et de loi d'entrée $\mu$. Comme la suite de v.a. $\mathbf{X}$ est adaptée par rapport à $(\mathscr{F}_n)_{n\geq 0}$, on a $\mathscr{F}_n^{\mathbf{X}} \subset \mathscr{F}_n$, $n \geq 0$. Pour toute fonction $f : E \to \mathbb{R}$ positive ou bornée, (I.3) implique donc

$$\forall B \in \mathscr{F}_n^{\mathbf{X}}\ , \quad \mathbf{E}[f(X_{n+1})\mathbf{1}_B] = \mathbf{E}\big[(Q.f)(X_n)\mathbf{1}_B\big]\ , \tag{I.5}$$

car $B \in \mathscr{F}_n^{\mathbf{X}} \subset \mathscr{F}_n$. On remarque que $(Q.f)(X_n)$ est $\mathscr{F}_n^{\mathbf{X}}$-mesurable. Donc (I.5) implique que

$$\mathbf{P}\text{-p.s.} \quad \mathbf{E}\big[f(X_{n+1}) \,\big|\, \mathscr{F}_n^{\mathbf{X}}\,\big] = (Q.f)(X_n)\ ,$$

et $\mathbf{X} = (X_n)_{n\geq 0}$ *est donc, sous* $\mathbf{P}$*, une chaîne de Markov relativement à la filtration* $(\mathscr{F}_n^{\mathbf{X}})_{n\geq 0}$*, de matrice de transition* $Q$ *et de loi d'entrée* $\mu$*.*

Par ailleurs, cela montre que $\mathbf{E}\big[f(X_{n+1}) \,\big|\, X_0, \ldots, X_n\,\big] = \mathbf{E}\big[f(X_{n+1}) \,\big|\, X_n\,\big]$. Autrement dit, la loi de $X_{n+1}$, conditionnellement à son passé donné par $\sigma(X_0, \ldots, X_n)$, ne dépend en fait que de $X_n$. Les variables d'une chaîne de Markov ne sont pas, en général, indépendantes, mais leur est en quelque sorte "faible". □

Le lemme suivant peut se voir comme une définition alternative des chaînes de Markov.

**Lemme I.1.4** *Soit* $Q = (p(i,j))_{i,j\in E}$*, une matrice de transition sur* $E$ *et soit* $\mu \in \mathscr{M}_1(E)$*. Soit* $\mathbf{X} = (X_n)_{n\geq 0}$*, une suite de v.a. à valeurs dans* $E$*. Soit* $\mathbf{P}$*, une mesure de probabilité sur* $(\Omega, \mathscr{F})$*. Les deux assertions suivantes sont équivalentes*

*(i) Sous* $\mathbf{P}$*,* $\mathbf{X}$ *est une chaîne de Markov de matrice de transition* $Q$ *et de loi d'entrée* $\mu$*.*

*(ii) Pour tout* $n \in \mathbb{N}$*, et pour tous* $i_0, \ldots, i_n \in E$*, on a*

$$\mathbf{P}(X_0 = i_0; \ldots; X_n = i_n) = \mu(i_0)p(i_0, i_1)p(i_1, i_2)\ldots p(i_{n-1}, i_n)\ . \tag{I.6}$$

**Preuve :** on suppose d'abord $(i)$ et on montre (I.6) par récurrence sur $n$. Par définition de la loi d'entrée, $\mathbf{P}(X_0 = i_0) = \mu(i_0)$, pour tout $i_0 \in E$. Supposons que pour tous $i_0, \ldots, i_n \in E$, (I.6) ait lieu. On fixe $i_0, \ldots, i_n, j \in E$. On rappelle que $\mathscr{F}_n^{\mathbf{X}} = \sigma(X_0, \ldots, X_n)$. On applique la définition I.1.2 à $f = \mathbf{1}_{\{j\}}$. Pour cela on remarque (facilement) au préalable que $(Q.f)(i) = p(i,j)$, pour tout $i \in E$. On a donc

$$\begin{aligned}
\mathbf{E}[\mathbf{1}_{\{X_0=i_0;\ldots;X_n=i_n;X_{n+1}=j\}}|\mathscr{F}_n^{\mathbf{X}}] &= \mathbf{1}_{\{X_0=i_0;\ldots;X_n=i_n\}}\mathbf{E}[\mathbf{1}_{\{j\}}(X_{n+1})|\mathscr{F}_n^{\mathbf{X}}] \\
&= \mathbf{1}_{\{X_0=i_0;\ldots;X_n=i_n\}}(Q.f)(X_n) \\
&= \mathbf{1}_{\{X_0=i_0;\ldots;X_n=i_n\}}p(X_n, j) = \mathbf{1}_{\{X_0=i_0;\ldots;X_n=i_n\}}p(i_n, j).
\end{aligned}$$

En intégrant et en utilisant l'hypothèse de récurrence, on obtient (I.6) au rang $n+1$, ce qui implique le point $(ii)$ par récurrence.

Réciproquement, on suppose $(ii)$. On fixe $j \in E$. On définit $\phi_j : E^{n+1} \to [0,1]$, en posant $\phi_j(i_0, \ldots, i_n) = p(i_n, j)$, pour tous $i_0, \ldots, i_n \in E$. Par (I.6), pour tous $i_0, \ldots, i_n \in E$, on a

$$\mathbf{E}[\mathbf{1}_{\{X_0=i_0;\ldots;X_n=i_n\}}\mathbf{1}_{\{X_{n+1}=j\}}] = \mu(i_0)p(i_0, i_1)p(i_1, i_2)\ldots p(i_{n-1}, i_n)p(i_n j)$$

$$
\begin{aligned}
&= \mathbf{E}[\mathbf{1}_{\{X_0=i_0;\dots;X_n=i_n\}} p(i_n, j)] \\
&= \mathbf{E}[\mathbf{1}_{\{X_0=i_0;\dots;X_n=i_n\}} \phi_j(X_0, \dots, X_n)] \qquad \text{(I.7)}
\end{aligned}
$$

Il est clair que les événements $\{X_0 = i_0; \dots; X_n = i_n\}$, $i_0, \dots, i_n \in E$, forment une partition dénombrable de $\Omega$ qui engendre $\mathscr{F}_n^{\mathbf{X}}$, qui est donc une tribu discrète. On rappelle que tout événement $B$ de $\mathscr{F}_n^{\mathbf{X}}$ est de la forme $\bigcup_{(i_0,\dots i_n)\in J}\{X_0 = i_0; \dots; X_n = i_n\}$, pour un certain sous-ensemble $J \subset E^{n+1}$. On a donc par (I.7) et par interversion série / espérance positive,

$$
\begin{aligned}
\mathbf{E}[\mathbf{1}_B \mathbf{1}_{\{X_{n+1}=j\}}] &= \mathbf{E}\Big[\sum_{(i_0,\dots i_n)\in J} \mathbf{1}_{\{X_0=i_0;\dots;X_n=i_n\}} \mathbf{1}_{\{X_{n+1}=j\}}\Big] \\
&= \sum_{(i_0,\dots i_n)\in J} \mathbf{E}[\mathbf{1}_{\{X_0=i_0;\dots;X_n=i_n\}} \mathbf{1}_{\{X_{n+1}=j\}}] \\
&= \sum_{(i_0,\dots i_n)\in J} \mathbf{E}[\mathbf{1}_{\{X_0=i_0;\dots;X_n=i_n\}} \phi_j(X_0, \dots, X_n)] \\
&= \mathbf{E}\Big[\sum_{(i_0,\dots i_n)\in J} \mathbf{1}_{\{X_0=i_0;\dots;X_n=i_n\}} \phi_j(X_0, \dots, X_n)\Big] \\
&= \mathbf{E}[\mathbf{1}_B \phi_j(X_0, \dots, X_n)].
\end{aligned}
$$

Comme $\phi_j(X_0, \dots, X_n)$ est $\mathscr{F}_n^{\mathbf{X}}$-mesurable cela implique que

$$\forall j \in E\,,\ \text{p.s.} \quad \mathbf{E}[\mathbf{1}_{\{X_{n+1}=j\}} | \mathscr{F}_n^{\mathbf{X}}] = \phi_j(X_0, \dots, X_n) = p(X_n, j)\ .$$

Soit $f : E \to \mathbb{R}$, bornée. Par le théorème d'interversion série / espérance conditionnelle, on a p.s.

$$\mathbf{E}[f(X_{n+1}) | \mathscr{F}_n^{\mathbf{X}}] = \sum_{j\in E} f(j) \mathbf{E}[\mathbf{1}_{\{X_{n+1}=j\}} | \mathscr{F}_n^{\mathbf{X}}] = \sum_{j\in E} f(j) p(X_n, j) = (Q.f)(X_n),$$

ce qui montre $(i)$. ■

**Exemple I.1.5** (*Marches aléatoires sur* $\mathbb{Z}^d$) Dans cet exemple l'espace d'états est $\mathbb{Z}^d$. On fixe une loi de probabilité $\pi$ sur $\mathbb{Z}^d$, que l'on appelle la *loi de saut*. On suppose qu'est définie sur un espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$ une suite de variables aléatoires $\xi_n : \Omega \to \mathbb{Z}^d$, $n \geq 1$, $\mathscr{F}$-mesurables et i.i.d. de loi commune $\pi$ :

$$\forall n \geq 1\,, \forall \mathbf{i} \in \mathbb{Z}^d\,, \quad \mathbf{P}(\xi_n = \mathbf{i}) = \pi(\mathbf{i})\ .$$

On se donne également $X_0 : \Omega \to \mathbb{Z}^d$, une variable $\mathscr{F}$-mesurable qui est supposée indépendante de la suite $(\xi_n)_{n\geq 1}$. On pose

$$\forall n \geq 1\,, \quad X_n = X_0 + \xi_1 + \dots + \xi_n\ .$$

Alors $(X_n)_{n\geq 0}$ est une chaîne de Markov. En effet pour toute fonction $f : \mathbb{Z}^d \to \mathbb{R}$ bornée on a clairement

$$\mathbf{E}\big[f(X_{n+1}) | X_0, \dots, X_n\big] = \sum_{\mathbf{i}\in\mathbb{Z}^d} f(X_n + \mathbf{i}) \pi(\mathbf{i})\ ,$$

comme le membre de droite ne dépend que de $X_n$, la suite $(X_n)_{n\geq 0}$ satisfait bien la définition I.1.2 des chaîne de Markov. On note ensuite $\mu$ la loi de $X_0$ : $\mu(\mathbf{i}) = \mathbf{P}(X_0 = \mathbf{i})$, $\mathbf{i} \in \mathbb{Z}^d$. On fixe $\mathbf{i}_0, \dots, \mathbf{i}_n \in \mathbb{Z}^d$ et on remarque alors que

$$\mathbf{P}(X_0 = \mathbf{i}_0; \dots; X_n = \mathbf{i}_n) \;=\; \mathbf{P}(X_0 = \mathbf{i}_0; \xi_1 = \mathbf{i}_1 - \mathbf{i}_0; \dots; \xi_n = \mathbf{i}_n - \mathbf{i}_{n-1})$$

$$\begin{aligned} &= \mathbf{P}(X_0 = \mathbf{i}_0)\mathbf{P}(\xi_1 = \mathbf{i}_1 - \mathbf{i}_0)\dots\mathbf{P}(\xi_n = \mathbf{i}_n - \mathbf{i}_{n-1}) \\ &= \mu(\mathbf{i}_0)\pi(\mathbf{i}_1 - \mathbf{i}_0)\dots\pi(\mathbf{i}_n - \mathbf{i}_{n-1}) \ . \end{aligned}$$

On définit $Q = (p(\mathbf{i}, \mathbf{j}))_{\mathbf{i},\mathbf{j}\in\mathbb{Z}^d}$ en posant

$$\forall \mathbf{i}, \mathbf{j} \in \mathbb{Z}^d \,, \quad p(\mathbf{i}, \mathbf{j}) = \pi(\mathbf{j} - \mathbf{i}) \ .$$

On vérifie facilement que $Q$ est une matrice de transition sur $\mathbb{Z}^d$ : $(X_n)_{n\geq 0}$ est une chaîne de Markov à valeurs dans $\mathbb{Z}^d$ de loi initiale $\mu$ et de matrice de transition $Q$. Cette chaîne de Markov s'appelle la *marche aléatoire de loi de saut* $\pi$.

On note $(\mathbf{e}_1, \dots, \mathbf{e}_d)$, la base canonique de $\mathbb{R}^d$. Si

$$\pi(\mathbf{e}_1) = \pi(-\mathbf{e}_1) = \ldots = \pi(\mathbf{e}_d) = \pi(-\mathbf{e}_d) = \frac{1}{2d} \,,$$

alors la marche aléatoire de loi de saut $\pi$ est appelée *marche aléatoire simple sur* $\mathbb{Z}^d$. Voir les figures I.1.5 page 6 et I.1.5 page 6.

**Exemple I.1.6** (*Processus de naissance et de mort*). Une chaîne de Markov à valeurs dans $\mathbb{N}$ et dont la matrice de transition $Q = (p(i, j))_{i,j\in\mathbb{N}}$ est telle que $p(i, j) = 0$ dès que $|i - j| \geq 2$, est appelée *processus de naissance et de mort*. ■

**Exemple I.1.7** (*Marches aléatoires sur les graphes pondérés*) Un graphe est la donnée d'un ensemble de sommets noté $\mathbf{S}$, que l'on supposera toujours dénombrable, et d'un ensemble d'arêtes reliant éventuellement certains sommets. L'ensemble des arêtes est noté $\mathbf{A}$. Nous nous restreindrons ici à l'idée la plus simple que l'on puisse s'en faire. Une arête reliant deux sommets $s$ et $s'$ est simplement formalisée par le couple non-ordonné $\{s, s'\}$. Le fait que l'on choisisse un couple non-ordonné signifie que nous n'orientons pas l'arête. Par ailleurs, nous ne considérons pas les graphes à arêtes multiples, c'est-à-dire les graphes dont deux sommets peuvent être reliés par plusieurs arêtes. On parlera dans ce cas de *graphes simples*. Si $\{s, s\} = \{s\}$ est une arête, alors on l'appelle *boucle*. Il nous arrivera parfois de considérer des graphes sans boucles (dans ce cas, si $\{s, s'\}$ est une arête du graphe, alors $s \neq s'$). L'ensemble des arêtes est simplement un sous-ensemble des paires non-ordonnées de sommets :

$$\mathbf{A} \subset \big\{ \{s, s'\} \,; \quad s, s' \in \mathbf{S} \big\} \ .$$

Le graphe est formellement donné par $\mathbf{G} = (\mathbf{S}, \mathbf{A})$. Les sommets seront notés en général par les lettres $s, s', s'', s_1, s_k \dots$ et les arêtes par $a, a', a'', a_1, a_k \dots$

On dit que deux sommets $s, s' \in \mathbf{S}$ sont *voisins*, ce que l'on note $s \sim s'$, s'ils sont reliés par une arête : $\{s, s'\} \in \mathbf{A}$. Le *degré* d'un sommet $s$ est le nombre de ses voisins. Ce nombre, qui peut être nul ou infini en général, est noté $\deg(s)$ :

$$\deg(s) = \#\{s' \in \mathbf{S} \ : \ s \sim s'\} \in \mathbb{N} \cup \{\infty\} \ .$$

Un *système de poids sur les arêtes* est la donnée d'une famille de réels *strictement positifs* indexés par les arêtes : $C_a \in ]0, \infty[$, $a \in \mathbf{A}$. Le graphe $\mathbf{G} = (\mathbf{S}, \mathbf{A})$ muni du du système de poids $\mathbf{C} = (C_a ; a \in \mathbf{A})$ est appelé *graphe pondéré*. Le *poids d'un sommet* $s$ est la somme des poids des arêtes reliant $s$ à ses voisins. On note cette quantité (qui peut être nulle ou infinie) par $\pi(s)$ :

$$\pi(s) = \sum_{s'\sim s} C_{\{s,s'\}} \ \in \ [0, \infty] \ .$$

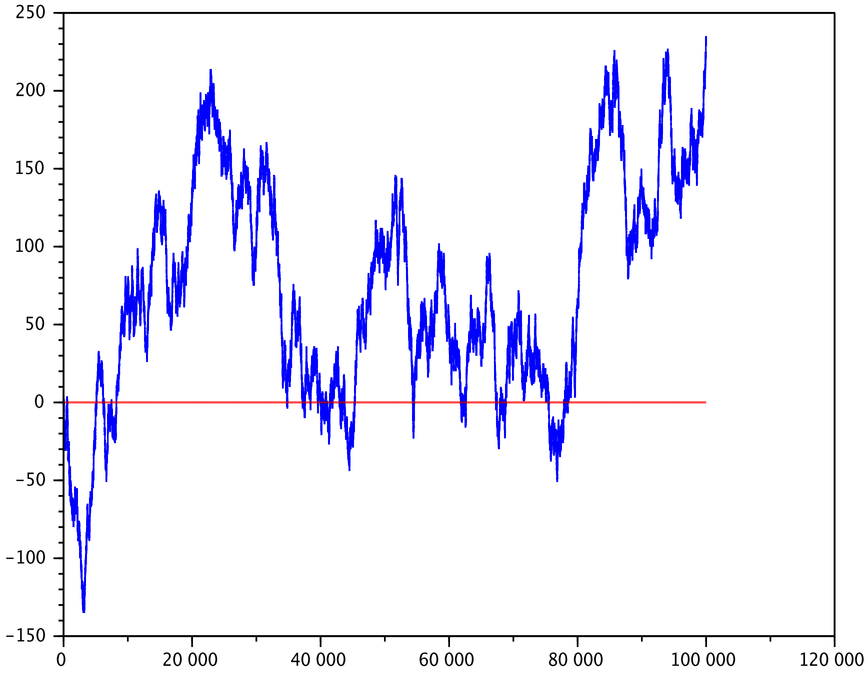


FIGURE I.1 – Le graphe $\{(n, X_n); 0 \leq n \leq 10^5\}$ d'une marche aléatoire simple sur $\mathbb{Z}$

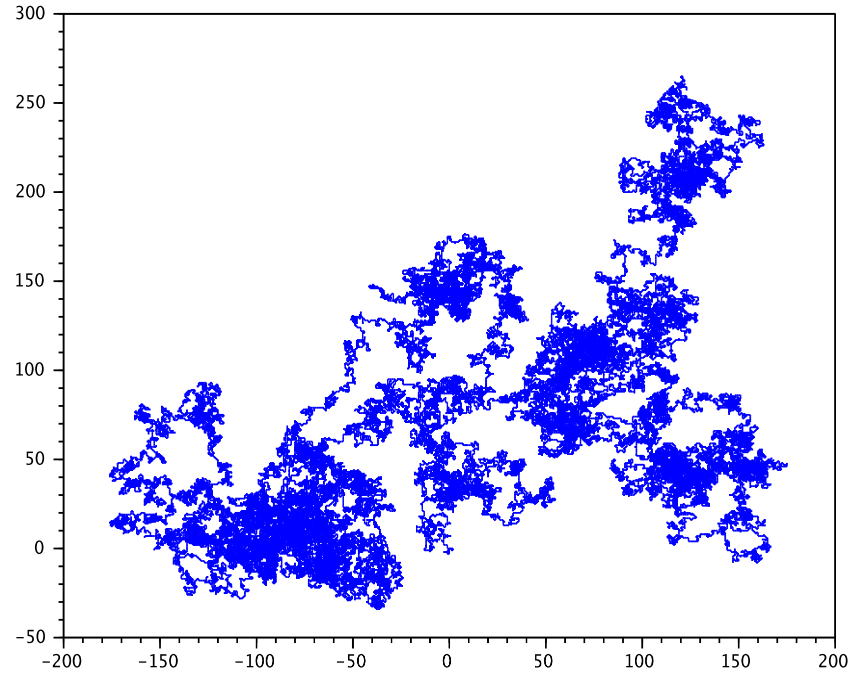


FIGURE I.2 – La trace $\{X_n; 0 \leq n \leq 10^5\} \subset \mathbb{Z}^2$ d'une marche aléatoire simple sur $\mathbb{Z}^2$

On fait l'hypothèse suivante :

$$\forall s \in \mathbf{S}\,, \quad 0 < \pi(s) < \infty\ . \tag{I.8}$$

On définit ensuite $Q = (p(s,s'))_{s,s' \in \mathbf{S}}$ par

$$p(s,s') = \begin{cases} C_{\{s,s'\}}/\pi(s) & \text{si } \{s,s'\} \in \mathbf{A} \\ 0 & \text{sinon.} \end{cases}$$

On voit que $Q$ est une matrice de transition sur $\mathbf{S}$. Une chaîne de Markov ayant $Q$ pour matrice de transition est appelée *marche aléatoire sur le graphe pondéré* $\mathbf{G} = (\mathbf{S}, \mathbf{A}, \mathbf{C})$ . Informellement une telle marche correspond au déplacement aléatoire d'une particule qui saute de sommets voisins en sommets voisins : à chaque étape la particule choisit d'emprunter une arête avec une probabilité proportionnelle au poids de cette arête.

Lorsque les poids $C_a$ sont constants à un réel strictement positif $c$, alors $\pi(s) = c\deg(s)$ et l'hypothèse (I.8) revient à supposer que tout sommet possède un nombre fini de voisins au moins égal à un. Dans ce cas, on a $p(s,s') = 1/\deg(s)$ et on parle de *marche simple sur le graphe* $\mathbf{G} = (\mathbf{S}, \mathbf{A})$. □

## EXERCICES.

**Exercice I.1** On suppose que sur $(\Omega, \mathscr{F}, \mathbf{P})$ est définie une suite $(\xi_n)_{n\geq 0}$, de variables de Bernoulli indépendantes de paramètre $p \in [0,1]$ : $\mathbf{P}(\xi_n = 1) = p$ et $\mathbf{P}(\xi_n = 0) = 1-p$, pour tout $n \in \mathbb{N}$. Pour tout $n \in \mathbb{N}$, on pose $X_n = \mathbf{1}_{\{\xi_n = \xi_{n+1}\}}$. Pour quel(s) $p \in [0,1]$, la suite $(X_n)_{n\geq 0}$ est-elle une chaîne de Markov ? □

**Exercice I.2** Toutes les variables sont définies sur l'espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$. On fixe un entier $N \geq 2$ et on se donne une suite $\mathbf{p} = (p_1, \ldots, p_N)$ telle que $p_k > 0$, pour tout $1 \leq k \leq N$, et $p_1 + \ldots + p_N = 1$. Autrement $\mathbf{p}$ est une loi de probabilité sur $\{1, \ldots, N\}$. Soient $(\xi_n)_{n\geq 1}$ une suite de variables à valeurs dans $\{1, \ldots, N\}$, indépendantes et de même loi $\mathbf{p}$ :

$$\forall k \in \{1, \ldots, N\}\,,\ \forall n \geq 1\,, \quad \mathbf{P}(\xi_n = k) = p_k\ .$$

On définit une suite $(X_n)_{n\geq 0}$ de variables à valeurs dans $\{0, \ldots, N\}$ en posant

$$X_0 = 0 \quad \text{et} \quad \forall n \geq 1\,, \quad X_n = \#\{\xi_1, \ldots, \xi_n\}\ .$$

1. Pour quelle(s) loi(s) $\mathbf{p}$, $(X_n)_{n\geq 0}$ est une chaîne de Markov (inhomogène ici) ?
2. On suppose que $\mathbf{p} = (\frac{1}{N}, \ldots, \frac{1}{N})$ la loi uniforme sur $\{1, \ldots, N\}$. Pour tout $1 \leq k \leq N$, on pose
$$T_k = \inf\{n \geq 0\ :\ X_n = k\}$$
Montrer que les variables $(T_{k+1} - T_k)_{0\leq k<N}$ sont indépendantes. Elles n'ont pas la même loi : calculer la loi de $T_{k+1} - T_k$.
3. On suppose toujours que $\mathbf{p}$ est uniforme. Montrer que
$$\lim_{N\to\infty} \frac{\mathbf{E}[T_N]}{N \log N} = 1\ .$$
4. On suppose toujours que $\mathbf{p}$ est uniforme. Montrer que pour tout $x \in \mathbb{R}$, on a
$$\lim_{N\to\infty} \mathbf{P}\big(T_N - N\log N \leq Nx\big) = \exp\left(-e^{-x}\right)\ .$$
5. Dans un village de $6 \times 365 = 2190$ personnes, quelle est approximativement la probabilité que chacun des 365 jours de l'année soit la date anniversaire d'au moins un habitant du village ? (*Réponse :* $40\%$). Montrer que dans un village de $5 \times 365 = 1825$ personne, cette probabilité tombe à environ $8\%$ (*Indication :* $365 \log 365$ *vaut environ* $2153$). □

**Exercice I.3** Toutes les variables sont définies sur l'espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$. Soit $(\xi_n)_{n\geq 0}$ une suite de variables indépendantes à valeurs dans un ensemble dénombrable.

1. Montrer que $(\xi_n)_{n\geq 0}$ est une chaîne de Markov (a priori inhomogène) dont on calculera la matrice de transition. Quelle hypothèse doit-on faire sur les lois des $\xi_n$ pour cette chaîne soit homogène ?
2. On suppose que la suite $(\xi_n)_{n\geq 0}$ est i.i.d. de loi $\pi$. On fixe $k \geq 2$ et on pose

$$X_n = (\xi_n, \dots, \xi_{n+k}) \ , \quad n \geq 0 \ .$$

3. Montrer que $(X_n)_{n\geq 0}$ est une chaîne de Markov homogène dont on précisera la matrice de transition. □

**Exercice I.4** (*Serpent*) Toutes les variables sont définies sur l'espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$. Soit $(X_n)_{n\geq 0}$, une chaîne de Markov à valeurs dans l'espace d'état $E$, homogène et de matrice de transition $Q = (p(i,j))_{i,j\in E}$. On fixe $k \geq 1$ et on pose

$$X_n^{(k)} = (X_n, \dots, X_{n+k}) \ , \quad n \geq 0 \ .$$

On appelle la suite $(X_n^{(k)})_{n\geq 0}$ le *k-serpent* de la chaîne $(X_n)_{n\geq 0}$. On introduit l'espace

$$E^{(k)} = \{\mathbf{i} = (\mathbf{i}_0, \dots \mathbf{i}_k) \in E^{k+1} \ : \ p(\mathbf{i}_0, \mathbf{i}_1) \dots p(\mathbf{i}_{k-1}, \mathbf{i}_k) > 0\} \ .$$

Montrer que $(X_n^{(k)})_{n\geq 0}$ est une chaîne de Markov homogène à valeurs dans $E^{(k)}$. Calculer sa matrice de transition. □

**Exercice I.5** Toutes les variables sont définies sur l'espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$. Soit une filtration sur $(\Omega, \mathscr{F})$ notée$(\mathscr{F}_n)_{n\geq 0}$ et soit $(X_n)_{n\geq 0}$, une suite de variables à valeurs dans un espace dénombrable $E$ qui est $(\mathscr{F}_n)_{n\geq 0}$-adaptée. On fixe $k \geq 1$. On suppose que pour toute fonction $f : E^k \to \mathbb{R}$, bornée, il existe une fonction $g : E^k \to \mathbb{R}$, qui ne dépend pas de $n$ (uniquement de $f$) telle qu'on ait pour tout $n \in \mathbb{N}$,

$$\mathbf{P}\text{- p.s.} \quad \mathbf{E}\big[f(X_{n+1}, \dots, X_{n+1+k}) | \mathscr{F}_{n+k}\big] = g(X_n, \dots, X_{n+k}).$$

On pose

$$X_n^{(k)} = (X_n, \dots, X_{n+k}) \ , \quad n \geq 0 \ .$$

1. Montrer que $(X_n^{(k)})_{n\geq 0}$ est une chaîne de Markov homogène à valeurs dans $E^k$.
2. Montrer que $(X_n)_{n\geq 0}$ est une chaîne de Markov homogène à valeurs dans $E$
3. Déduire la matrice de transition de $(X_n^{(k)})_{n\geq 0}$ en fonction de celle de $(X_n)_{n\geq 0}$, en utilisant l'exercice précédent I.4. □

**Exercice I.6** On suppose que sur $(\Omega, \mathscr{F}, \mathbf{P})$ est définie une suite $(\xi_n)_{n\geq 1}$, de variables à valeurs dans $\{-1, 1\}$, dont la loi est donnée par $\mathbf{P}(\xi_n = 1) = \mathbf{P}(\xi_n = -1) = \frac{1}{2}$, pour tout $n \geq 1$. On pose ensuite

$$S_0 = 0 \quad \text{et} \quad S_n = \xi_1 + \ldots + \xi_n \ , \quad n \geq 1 \ .$$

On pose ensuite

$$X_n = \max_{0\leq m\leq n} S_m \ , \quad n \geq 0.$$

Montrer que $(X_n)_{n\geq 0}$ *n'est pas* une chaîne de Markov. □

**Exercice I.7** Toutes les variables sont définies sur l'espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$. Soient $U, U_1, \dots, U_n, \dots$ des variables indépendantes uniformes sur $]0,1[$. On pose $X_0 = 0$ et

$$X_n = (2.\mathbf{1}_{\{U_1<U\}} - 1) + \ldots + (2.\mathbf{1}_{\{U_n<U\}} - 1) \ , \quad n \geq 1.$$

Montrer que $(X_n)_{n\geq 0}$ est une chaîne de Markov inhomogène. □

**Exercice I.8** Soit un jeu de $N$ cartes distinctes numérotées de 1 à $N$. On bat les cartes selon la procédure suivante : on prend la carte qui est au dessus du paquet et on l'insère uniformément au hasard parmi les $N-1$ du dessous (on insiste sur le fait que l'on peut choisir de mettre cette carte tout en bas du paquet ou de la laisser à sa place : il y a donc $N$ possibilités d'insertion). Modéliser les battages successifs du jeu de cartes par une suite de permutations aléatoires de $\{1, \dots, N\}$ qui est une chaîne de Markov dont dont précisera la matrice de transition. □

**Exercice I.9** On considère $N$ individus distincts qui portent portent deux sortes d'allèle d'un même gène : l'allèle $A$ et l'allèle $a$. La population évolue en restant de taille constante $N$ et les gènes se transmettent d'une génération à l'autre de la manière suivante (très simplifiée) : chaque individu de la génération $n$ choisit indépendamment au hasard un "père" à la génération précédente ; il hérite de la forme ($A$ ou $a$) de son "père". On note $X_n$ le nombre d'individus porteur de la forme $A$ à la génération $n$. Montrer que $(X_n)_{n\geq 0}$ est une chaîne de Markov dont on précisera l'espace d'états et la matrice de transition. □

**Exercice I.10** Soit $(X_n)_{n\geq 0}$ une chaîne de Markov à valeurs dans $E$. A priori elle est inhomogène et ses matrices de transition sont notées $Q_n = (p_n(i,j))_{i,j\in E}$. On pose $Y_n = (n, X_n)$, pour tout $n \in \mathbb{N}$. Montrer que $(Y_n)_{n\geq 0}$ est une chaîne de Markov à valeurs dans $\mathbb{N} \times E$, qui est homogène et on précisera sa matrice de transition. □

**Exercice I.11** Toutes les variables sont définies sur l'espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$. On se donne une filtration $(\mathscr{F}_n)_{n\geq 0}$ sur $(\Omega, \mathscr{F})$ et soit $(X_n)_{n\geq 0}$, une $(\mathscr{F}_n)_{n\geq 0}$-chaîne de Markov à valeurs dans $E$, homogène et de matrice de transition $Q = (p(i,j))_{i,j\in E}$. On fixe une fonction $\psi : E \to F$, où $F$ est un ensemble dénombrable.

1. On suppose $\psi$ bijective. Montrer que $(\psi(X_n))_{n\geq 0}$ est une $(\mathscr{F}_n)_{n\geq 0}$-chaîne de Markov à valeurs dans $F$, homogène, dont on calculera la matrice de transition.
2. Montrer, en trouvant un contre-exemple, que si on ne suppose pas $\psi$ bijective le résultat précédent peut devenir faux.
3. On ne suppose pas que $\psi$ est nécessairement bijective. Mais on suppose que $\psi$ est surjective et que pour tout $i \in E$ et tout $\ell \in F$, on pose

$$\forall i, j \in E \text{ tels que } \psi(i) = \psi(j)\,,\ \forall a \in F\,, \qquad \sum_{k\in E:\psi(k)=a} p(i,k) = \sum_{k\in E:\psi(k)=a} p(j,k)\ .$$

   Montrer que sous cette hypothèse, $(\psi(X_n))_{n\geq 0}$ est une $(\mathscr{F}_n)_{n\geq 0}$-chaîne de Markov à valeurs dans $F$, homogène, dont on calculera la matrice de transition. □

### I.1.b Exemples : processus de branchement, arbres aléatoires.

**Définitions et premiers résultats.** Dans cette section nous détaillons un modèle décrivant l'évolution de la taille d'une population se reproduisant selon le mécanisme suivant. On fixe une loi de probabilité sur $\mathbb{N}$, notée $\xi = (\xi(k))_{k\in\mathbb{N}}$, que l'on voit comme une *loi de reproduction* :

$$\forall k \geq 0\,, \quad \xi(k) \geq 0 \quad \text{et} \quad \sum_{k\geq 0} \xi(k) = 1\ .$$

On suppose qu'à la génération 0, il y a un seul ancêtre. Cet ancêtre a un nombre aléatoire d'enfants distribué selon la loi $\xi$ ; ses enfants sont à la génération 1. Ensuite chacun des enfants de la génération 1 se reproduit indépendamment selon la même loi $\xi$ : leurs enfants sont alors situés à la génération 2. A leur tour, les individus de la générations 2 se reproduisent indépendamment selon la loi $\xi$ et leurs enfants sont à la génération 3 ...etc. On note $Z_n$ le nombre total d'individus à la génération $n$ et on s'intéresse à la suite de variables aléatoires $(Z_n)_{n\geq 0}$ qui est appelée processus de branchement de loi de reproduction $\xi$, dont une définition plus formelle est la suivante.

**Définition I.1.8** Soit $\xi = (\xi(k))_{k\in\mathbb{N}}$, une mesure de probabilité sur $\mathbb{N}$. Soit $(X_{n,j})_{n\geq 0, j\geq 1}$ une famille de variables indépendantes à valeurs dans $\mathbb{N}$ et de loi $\xi$. Soit $Z_0 : \Omega \to \mathbb{N}$, indépendante des $(X_{n,j})_{n\geq 0, j\geq 1}$. On définit récursivement la suite de variables aléatoires $(Z_n)_{n\geq 0}$ pour tout $n \geq 0$ par :

$$Z_{n+1} = \sum_{1\leq j\leq Z_n} X_{n,j} \quad \text{si} \quad Z_n \geq 1 \qquad \text{et} \qquad Z_{n+1} = 0 \quad \text{si} \quad Z_n = 0\ .$$

La suite $(Z_n)_{n\geq 0}$ est appelée *processus de branchement de loi de reproduction* $\xi$ (ou encore processus de Galton-Watson ou processus de Bienaymé). □

Si chaque individu n'a qu'un enfant au plus, c'est-à-dire si $\xi(0) + \xi(1) = 1$, alors le modèle n'est pas très intéressant. C'est pour cela que nous supposons toujours que la loi de reproduction n'est pas triviale, ce qui signifie ici que

$$\xi(0) + \xi(1) < 1\ . \tag{I.9}$$

On introduit ensuite la moyenne $m_\xi$ et la fonction génératrice $\varphi$ de $\xi$, qui jouent un grand rôle dans les résultats.

$$m_\xi = \sum_{k\geq 0} k\xi(k) \in ]0,\infty] \quad \text{et} \quad \varphi(r) = \sum_{k\geq 0} \xi(k) r^k \,, \quad r \in [0,1] \,.$$

On note qu'a priori $m_\xi$ peut-être infinie et on observe que $\varphi$ est positive, croissante et que $\varphi(1) = 1$. Donc $\varphi([0,1]) \subset [0,1]$. Cela permet de définir les itérées successives de la fonction $\varphi$, qui sont notées $\varphi_n$, $n \geq 0$. Formellement, pour tout $r \in [0,1]$,

$$\varphi_0(r) = r \quad \text{et} \quad \varphi_{n+1}(r) = \varphi(\varphi_n(r)) = \varphi_n(\varphi(r)) \,, \quad n \geq 0.$$

**Lemme I.1.9** *Pour tout $n \in \mathbb{N}$ et pour tout $r \in [0,1]$, on a presque sûrement*

$$\mathbf{E}\big[r^{Z_{n+1}} \,|\, Z_n\big] = \mathbf{E}\big[r^{Z_{n+1}} \,|\, \sigma(Z_0,\dots,Z_n)\big] = \varphi(r)^{Z_n} \,. \tag{I.10}$$

*Par conséquent $\mathbf{E}\big[r^{Z_n}|Z_0\big] = \varphi_n(r)^{Z_0}$, $r \in [0,1]$, $n \geq 0$.*

**Preuve :** on pose $\mathscr{F}_0 = \sigma(Z_0)$ et pour tout $n \geq 1$, on pose $\mathscr{F}_n = \sigma(Z_0, X_{p,j} \,;\, j \in \mathbb{N}^*\,,\, 0 \leq p \leq n-1)$, la tribu incluant l'information généalogique concernant les $n$ premières générations. Clairement $(Z_n)_{n\geq 0}$ est adaptée à la filtration $(\mathscr{F}_n)_{n\geq 0}$ et on a donc

$$\sigma(Z_n) \subset \sigma(Z_0,\dots,Z_n) \subset \mathscr{F}_n \,. \tag{I.11}$$

Par interversion série / espérance conditionnelle positive, p.s. on a

$$\begin{aligned}
\mathbf{E}\big[r^{Z_{n+1}} \,|\, \mathscr{F}_n\big] &= \mathbf{E}\Big[\mathbf{1}_{\{Z_n=0\}} + \sum_{k\geq 1} \mathbf{1}_{\{Z_n=k\}} \prod_{1\leq j\leq k} r^{X_{n,j}} \,\Big|\, \mathscr{F}_n\Big] \\
&= \mathbf{1}_{\{Z_n=0\}} + \sum_{k\geq 1} \mathbf{1}_{\{Z_n=k\}} \mathbf{E}\Big[\prod_{1\leq j\leq k} r^{X_{n,j}} \,\Big|\, \mathscr{F}_n\Big].
\end{aligned} \tag{I.12}$$

Clairement $\sigma(X_{n,j}, j \in \mathbb{N}^*)$ est indépendante de de $\mathscr{F}_n$. Cela montre que pour tout $k \geq 1$, le vecteur $(r^{X_{n,1}},\dots,r^{X_{n,k}})$ est indépendant de $\mathscr{F}_n$ et donc pour tout $k \geq 1$, p.s. on a

$$\mathbf{E}\Big[\prod_{1\leq j\leq k} r^{X_{n,j}} \,\Big|\, \mathscr{F}_n\Big] = \prod_{1\leq j\leq k} \mathbf{E}\big[r^{X_{n,j}}\big].$$

On remarque ensuite que $\varphi(r) = \mathbf{E}[r^{X_{n,j}}]$, pour tout $n$ et tout $j$. Ce qui précède montre donc que pour tout $k \geq 1$, p.s. on a

$$\mathbf{E}\Big[\prod_{1\leq j\leq k} r^{X_{n,j}} \,\Big|\, \mathscr{F}_n\Big] = \varphi(r)^k \,,$$

et par (I.12), p.s. on a

$$\mathbf{E}\big[r^{Z_{n+1}} \,|\, \mathscr{F}_n\big] = \mathbf{1}_{\{Z_n=0\}} + \sum_{k\geq 1} \mathbf{1}_{\{Z_n=k\}} \varphi(r)^k = \varphi(r)^{Z_n} \,,$$

ce qui permet de conclure grâce à (I.11). ■

Ce résultat montre que les itérées de la fonction caractéristique jouent un rôle important dans l'étude des processus de branchement. Le lemme suivant donne un résultat analytique concernant le comportement de la suite $(\varphi_n(0))_{n\geq 0}$.

**Lemme I.1.10** *Sous l'hypothèse (I.9) de non-trivialité de $\xi$, la fonction $\varphi$ est strictement convexe, c'est-à-dire que $\varphi'$ est strictement croissante. On en déduit les assertions suivantes.*

*(i) Si $m_\xi \leq 1$, l'équation $\varphi(r) = r$ n'a qu'une seule solution dans $[0, 1]$ qui est $r = 1$.*

*(ii) Si $m_\xi > 1$, l'équation $\varphi(r) = r$ a deux solutions distinctes dans $[0, 1]$, dont $r = 1$.*

*On note $q_\xi$ la plus petite solution de l'équation $\varphi(r) = r$ dans $[0, 1]$. Ce qui précède montre donc que $q_\xi = 1$ si $m_\xi \leq 1$ et que $q_\xi < 1$ si $m_\xi > 1$. De plus, on a*

$$\varphi(r) > r \text{ si } r \in [0, q_\xi[ \quad \text{et} \quad \varphi(r) < r \text{ si } r \in ]q_\xi, 1[ . \tag{I.13}$$

*Par conséquent, $(\varphi_n(0))_{n\geq 0}$ est une suite croissante qui converge vers $q_\xi$.*

**Preuve :** exercice. ■

On introduit le *temps d'extinction* :

$$T = \inf \big\{ n \in \mathbb{N} \, : \, Z_n = 0 \big\} \ ,$$

avec la convention que $\inf \emptyset = \infty$, c'est-à-dire que $\{T = \infty\} = \{\forall n \in \mathbb{N}\, , \, Z_n \geq 1\}$. On voit donc que $\{T < \infty\}$ est l'événement de *l'extinction de la population* et que $\{T = \infty\}$ est l'événement de *la survie de la population*. Par définition, si $Z_n = 0$, alors $Z_p = 0$, pour tout $p \geq n$. On a donc

$$\{Z_n = 0\} \subset \{Z_{n+1} = 0\} \quad \text{et} \quad \{T < \infty\} = \bigcup_{n\in\mathbb{N}} \{Z_n = 0\} \ .$$

Par conséquent $\mathbf{P}(T < \infty) = \lim_{n\to\infty} \uparrow \mathbf{P}(Z_n = 0)$. On rappelle ensuite que $\mathbf{E}[r^{Z_n}] = \varphi_n(r)$. Or $\mathbf{E}[r^{Z_n}] = \mathbf{P}(Z_n = 0) + \mathbf{E}[r^{Z_n} \mathbf{1}_{\{Z_n\geq 1\}}]$, donc $\varphi_n(0) = \mathbf{P}(Z_n = 0)$. Cela implique que

$$\mathbf{P}(T < \infty) = q_\xi \quad \text{ et } \quad \mathbf{P}(T = \infty) = \mathbf{P}(\forall n \geq 0\, , \ Z_n \geq 1) = 1 - q_\xi. \tag{I.14}$$

Au vu du lemme I.1.10 et de ce qui précède, on utilise la terminologie suivante

- Si $m_\xi \leq 1$, la population s'éteint avec probabilité 1. Si $m_\xi < 1$ on dit qu'on est dans le cas *sous-critique* et si $m_\xi = 1$ on dit qu'on est dans le cas *critique*.
- Si $m_\xi > 1$, on a $q_\xi < 1$ et la population a une probabilité $1 - q$ de jamais s'éteindre. On dit qu'on est dans le cas *sur-critique*. On observe que $m_\xi > 1$ implique (I.9), c'est-à-dire que $\xi$ est non-triviale nécessairement.

**Processus de branchement vus comme chaînes de Markov.** Nous allons montrer que $(Z_n)_{n\geq 0}$ est une chaîne de Markov en précisant sa matrice de transition. Pour cela, il est commode d'introduire, la notion de *produit de convolution* dans le cadre très simple des mesures de probabilité sur $\mathbb{N}$ : soient $\mu$ et $\nu$, deux mesures de probabilité sur $\mathbb{N}$. On définit la mesure $\mu * \nu$ par

$$(\mu * \nu)(i) = \sum_{0\leq j\leq i} \mu(j)\nu(i - j) \ , \quad i \in \mathbb{N}.$$

On vérifie facilement que $\mu * \nu$ est une mesure de probabilité sur $\mathbb{N}$, que $\mu * \nu = \nu * \mu$ et que $(\mu * \nu) * \pi = \mu*(\nu*\pi)$, c'est-à-dire que $*$ est une opération associative. Cette opération est appelée *le produit de convolution*.

Il s'interprête de façon probabiliste comme suit : soient $X, Y : \Omega \to \mathbb{N}$, deux v.a. indépendantes telle que $X$ a pour loi $\mu$ et $Y$ a pour loi $\nu$. On vérifie facilement que $X + Y$ a pour loi $\mu * \nu$. Pour toute mesure de probabilité $\mu$ sur $\mathbb{N}$, on note $\varphi_\mu$ sa fonction génératrice :

$$\forall r \in [0,1] , \quad \varphi_\mu(r) = \sum_{i \in \mathbb{N}} r^i \mu(i) ,$$

qui caractérise $\mu$. On vérifie facilement que

$$\pi = \mu * \nu \iff \varphi_\pi(r) = \varphi_\mu(r)\varphi_\nu(r) , \quad r \in [0,1] . \tag{I.15}$$

Pour tout $i \in \mathbb{N}$, on introduit récursivement la $i^{\text{ième}}$ *convolée de* $\xi$, notée $\xi^{*i}$, en posant $\xi^{*0} = \delta_0$, la masse de Dirac en 0, qui est l'élément neutre pour le produit $*$, $\xi^{*1} = \xi$, $\xi^{*2} = \xi * \xi$, et plus généralement $\xi^{*(i+1)} = \xi^{*i} * \xi = \xi * \xi^{*i}$. L'équivalence (I.15) montre que $\varphi(r)^i = \varphi_{\xi^{*i}}(r)$, $r \in [0,1]$. On pose alors

$$Q = (p(i,j))_{i,j \in \mathbb{N}} \text{ où } \quad \forall i, j \in \mathbb{N} , \quad p(i,j) = \xi^{*i}(j) . \tag{I.16}$$

Il est facile de voir que $Q$ est une matrice de transition. Pour tout $n \in \mathbb{N}$, et pour tout $r \in [0,1]$, on a donc p.s.

$$\varphi(r)^{Z_n} = \sum_{j \in \mathbb{N}} p(Z_n, j) r^j . \tag{I.17}$$

Pour tout $n \geq 0$, on pose $\mathscr{F}_n = \sigma(Z_0, \ldots, Z_n)$. Par interversion série / espérance conditionnelle, implique que pour tout $r \in [0,1]$, p.s. on a

$$\mathbf{E}\big[r^{Z_{n+1}} \mid \mathscr{F}_n\big] = \sum_{j \in \mathbb{N}} r^j \mathbf{E}[\mathbf{1}_{\{Z_{n+1}=j\}} | \mathscr{F}_n] . \tag{I.18}$$

Par (I.10) ainsi qu'un argument simple sur l'identification des coefficients aléatoires, on en déduit que :

$$\forall n, j \in \mathbb{N} , \text{ p.s.} \quad \mathbf{E}[\mathbf{1}_{\{Z_{n+1}=j\}} | \mathscr{F}_n] = p(Z_n, j) ,$$

ce qui montre facilement que

- $(Z_n)_{n \geq 0}$ *est une chaîne de Markov de matrice de transition* $Q$ *donnée par (I.16).*

La proposition suivante donne une caractérisation intrinsèque des processus de branchement.

**Proposition I.1.11** *Soit* $Q = (p(i,j))_{i,j \in \mathbb{N}}$ *une matrice de transition sur* $\mathbb{N}$. *On suppose qu'elle satisfait la* ***propriété de branchement*** *:*

$$\forall i_1, i_2 \in \mathbb{N}, \quad p(i_1, \cdot) * p(i_2, \cdot) = p(i_1 + i_2, \cdot) .$$

*Alors* $Q$ *est la matrice de transition d'un processus de branchement de loi de reproduction* $\xi(k) = p(1,k)$, $k \in \mathbb{N}$.

**Preuve :** exercice. ■

**Résultats complémentaires sur les processus de branchement.** Le lemme suivant détaille dans le cas surcritique le comportement qualitatif de $(Z_n)_{n \geq 0}$ lorsque la population survit.

**Lemme I.1.12** *On se place dans le cas sur-critique où* $m_\xi > 1$. *Si la population ne s'éteint pas, alors elle tend vers* $\infty$ *:*

$$\text{p.s.} \quad \mathbf{1}_{\{T=\infty\}} = \mathbf{1}_{\{\forall n \in \mathbb{N}\,,\, Z_n \geq 1\}} = \mathbf{1}_{\{\lim_{n\to\infty} Z_n = \infty\}} \; .$$

**Preuve :** si $\xi(0) = 0$, alors $q_\xi = 0$. Chaque individu a au moins un enfant. Il est facile de montrer par récurrence à partir des définitions que $(Z_n)_{n\geq 0}$ est une suite croissante. On pose $Z_\infty = \sup_{n\geq 0} Z_n$. On a donc $\lim_{n\to\infty} \uparrow Z_n = Z_\infty$. Pour tout $r \in [0,1[$, la convergence dominée implique que $\mathbf{E}[r^{Z_\infty}] = \lim_{n\to\infty} \mathbf{E}[r^{Z_n}]$. Or $\mathbf{E}[r^{Z_n}] = \varphi_n(r)$, et il est facile de montrer grâce au lemme I.1.10 que pour tout $r \in [0,1[$, la suite $(\varphi_n(r))_{n\geq 0}$ décroît vers $q_\xi = 0$. On en déduit que $\mathbf{E}[r^{Z_\infty}] = 0$, pour tout $r \in [0,1[$, ce qui implique que $Z_\infty = \infty$, p.s., qui est bien le résultat désiré dans le cas où $\xi(0) = q_\xi = 0$.

On suppose ensuite que $\xi(0) > 0$, ce qui implique que $q_\xi > 0$. Comme $m_\xi > 1$, on a également $q_\xi < 1$. Pour tout $n \in \mathbb{N}$, on pose

$$\mathscr{F}_n = \sigma(Z_0, \ldots, Z_n) \quad \text{et} \quad M_n = q_\xi^{Z_n} \in ]0,1] \; .$$

Le lemme I.1.9 montre que $\mathbf{E}[M_{n+1}|\mathscr{F}_n] = \mathbf{E}\big[q_\xi^{Z_{n+1}}|\mathscr{F}_n\big] = \varphi(q_\xi)^{Z_n} = q_\xi^{Z_n} = M_n$. Autrement dit, $(M_n)_{n\geq 0}$ est une $(\mathscr{F}_n)_{n\geq 0}$-martingale positive qui est constituée de variables bornées par 1. C'est une martingale régulière qui converge donc p.s. et dans $\mathrm{L}^1$ vers $M_\infty \in [0,1]$. La convergence $\mathrm{L}^1$ implique le passage à la limite sous l'espérance suivant :

$$\begin{aligned} q_\xi &= \mathbf{E}[M_0] = \mathbf{E}[M_n] = \lim_{n\to\infty} \mathbf{E}[M_n] \\ &= \mathbf{E}[M_\infty] = \mathbf{E}[M_\infty \mathbf{1}_{\{T<\infty\}}] + \mathbf{E}[M_\infty \mathbf{1}_{\{T=\infty\}}]. \end{aligned}$$

Or $\mathbf{1}_{\{T<\infty\}} M_\infty = \mathbf{1}_{\{T<\infty\}}$ et comme $\mathbf{P}(T<\infty) = q_\xi$, les égalités précédentes impliquent que $M_\infty \mathbf{1}_{\{T=\infty\}} = 0$, p.s., c'est-à-dire que $\lim_{n\to\infty} q_\xi^{Z_n} \mathbf{1}_{\{T=\infty\}} = 0$, presque sûrement. En utilisant la notation $q_\xi^\infty = 0$, et en posant $Y = \liminf_{n\to\infty} Z_n$, ce qui précède montre donc que $q_\xi^Y \mathbf{1}_{\{T=\infty\}} = 0$ p.s. Cela implique que $\mathbf{1}_{\{T=\infty\}} \leq \mathbf{1}_{\{Y=\infty\}}$ p.s., c'est-à-dire que $\mathbf{1}_{\{T=\infty\}} \leq \mathbf{1}_{\{\lim_{n\to\infty} Z_n = \infty\}}$ p.s. Par ailleurs, il est évident que $\mathbf{1}_{\{\lim_{n\to\infty} Z_n=\infty\}} \leq \mathbf{1}_{\{T=\infty\}}$, ce qui permet de conclure. ■

Affinons l'étude de $(Z_n)_{n\geq 0}$ en montrant que si la population survit, alors elle croît *exponentiellement vite*.

**Théorème I.1.13** *On suppose que* $1 < m_\xi < \infty$ *et que* $\sum_{k\geq 0} k^2 \xi(k) < \infty$. *Pour tout* $n \in \mathbb{N}$, *on pose* $\mathscr{F}_n = \sigma(Z_0, \ldots, Z_n)$ *et* $W_n = m_\xi^{-n} Z_n$. *Alors, les assertions suivantes sont vérifiées.*

*(i)* $(W_n)_{n\geq 0}$ *est une* $(\mathscr{F}_n)_{n\geq 0}$*-martingale positive et bornée en norme* $\mathrm{L}^2$. *Elle converge p.s. et dans* $\mathrm{L}^2$ *vers une v.a. notée* $W_\infty$.

*(ii) Presque sûrement,* $\mathbf{1}_{\{T=\infty\}} = \mathbf{1}_{\{W_\infty>0\}}$. *Cela montre que* $Z_n \sim_{n\to\infty} m_\xi^n W_\infty$ *est le rythme de croissance exact de la suite* $(Z_n)_{n\geq 0}$.

**Preuve :** comme tous les termes sont positifs, on a l'interversion série/espérance conditionnelle suivante $\mathbf{E}[r^{Z_{n+1}}|\mathscr{F}_n] = \sum_{k\geq 0} r^k \mathbf{E}[\mathbf{1}_{\{Z_{n+1}=k\}}|\mathscr{F}_n]$. Puisque c'est une série entière à coefficients aléatoires mais bornés positifs, on a p.s. pour tout $r \in [0,1[$,

$$\frac{d}{dr}\mathbf{E}[r^{Z_{n+1}}|\mathscr{F}_n] = \sum_{k\geq 1} k r^{k-1} \mathbf{E}[\mathbf{1}_{\{Z_{n+1}=k\}}|\mathscr{F}_n] = \mathbf{E}[Z_{n+1} r^{Z_{n+1}-1}|\mathscr{F}_n] \; .$$

Or le lemme I.1.9 implique que $\frac{d}{dr}\mathbf{E}[r^{Z_{n+1}}|\mathscr{F}_n] = Z_n \varphi'(r)\varphi(r)^{Z_n-1}$. Donc p.s., pour tout $r \in [0,1[$, on a

$$\mathbf{E}[Z_{n+1} r^{Z_{n+1}-1}|\mathscr{F}_n] = Z_n \varphi'(r)\varphi(r)^{Z_n-1} \; . \tag{I.19}$$

De même, en redérivant, on montre que

$$\mathbf{E}[Z_{n+1}(Z_{n+1}-1)r^{Z_{n+1}-2}|\mathscr{F}_n] = Z_n\varphi''(r)\varphi(r)^{Z_n-1} + Z_n(Z_n-1)\varphi'(r)^2\varphi(r)^{Z_n-2} \; . \tag{I.20}$$

On fait tendre $r$ vers $1^-$ dans (I.19), par convergence monotone, on obtient $\mathbf{E}[Z_{n+1}\,|\,\mathscr{F}_n] = m_\xi Z_n$. Cela implique que $(W_n)_{n\geq 0}$ est une $(\mathscr{F}_n)_{n\geq 0}$-martingale et que $\mathbf{E}[Z_n] = m_\xi^n$, $n \geq 0$. On fait tendre $r$ vers $1^-$ dans (I.20) et, par convergence monotone, obtient

$$\mathbf{E}[Z_{n+1}(Z_{n+1}-1)\,|\,\mathscr{F}_n] = \alpha Z_n + m_\xi^2 Z_n(Z_n-1) \; . \tag{I.21}$$

où on a posé $\alpha = \lim_{r\to 1^-}\varphi''(r) = \sum_{k\geq 2} k(k-1)\xi(k)$, qui est une quantité positive finie. On pose $\sigma^2 = \sum_{k\geq 0} k^2\xi(k) - \mu^2$, qui est la variance de $\xi$. On a $\alpha = \sigma^2 + \mu^2 - \mu$. Un calcul simple déduit de (I.21) implique que $\mathbf{E}[Z_{n+1}^2|\mathscr{F}_n] = (\mu Z_n)^2 + \sigma^2 Z_n$. Cela implique

$$\forall n \in \mathbb{N}\, , \quad \mathbf{E}[W_{n+1}^2|\mathscr{F}_n] = W_n^2 + \tfrac{\sigma^2}{m_\xi^2} m_\xi^{-n} W_n \; .$$

On pose $u_n = \mathbf{E}[W_n^2]$ (a priori dans $[0,\infty]$). En intégrant l'égalité précédente et en utilisant le fait que $\mathbf{E}[W_n] = 1$, on obtient l'équation $u_{n+1} = u_n + \frac{\sigma^2}{m_\xi^2} m_\xi^{-n}$ et $u_0 = 1$. Cela entraîne facilement que $u_n = 1 + \frac{\sigma^2}{m_\xi^2}\frac{1-m_\xi^{-n}}{1-m_\xi^{-1}}$ et donc que $\lim_{n\to\infty}\mathbf{E}[W_n^2] = 1 + \frac{\sigma^2}{m_\xi(m_\xi-1)}$, ce qui montre $(i)$.

Pour montrer le point $(ii)$, il suffit de prouver que $\mathbf{P}(W_\infty = 0) = q_\xi$. En effet, cela implique que $\mathbf{P}(W_\infty > 0) = 1 - q_\xi = \mathbf{P}(T = \infty)$, c'est-à-dire que $\mathbf{E}[\mathbf{1}_{\{T=\infty\}} - \mathbf{1}_{\{W_\infty>0\}}] = 0$ ; or clairement, on a $\mathbf{1}_{\{W_\infty>0\}} \leq \mathbf{1}_{\{T=\infty\}}$, ce qui entraîne bien que $\mathbf{1}_{\{W_\infty>0\}} = \mathbf{1}_{\{T=\infty\}}$ presque sûrement, qui est le résultat désiré.

Montrons donc $\mathbf{P}(W_\infty = 0) = q_\xi$. Pour tout $\lambda > 0$, on pose $L(\lambda) = \mathbf{E}[\exp(-\lambda W_\infty)]$. Le théorème de convergence dominée implique que $L(\lambda) = \lim_{n\to\infty}\mathbf{E}[\exp(-\lambda W_n)]$. Or lemme I.1.9 implique que

$$\varphi\big(\mathbf{E}\big[e^{-\lambda W_n}\big]\big) = \varphi\big(\varphi_n(e^{-\lambda\mu^{-n}})\big) = \varphi_{n+1}(e^{-m_\xi\lambda m_\xi^{-n-1}}) = \mathbf{E}\big[e^{-m_\xi\lambda W_{n+1}}\big],$$

ce qui entraîne

$$\forall \lambda \in \mathbb{R}_+\, , \quad \varphi(L(\lambda)) = L(m_\xi\lambda) \; . \tag{I.22}$$

On remarque ensuite que

$$L(\lambda) = \mathbf{P}(W_\infty = 0) + \mathbf{E}[\exp(-\lambda W_\infty)\mathbf{1}_{\{W_\infty>0\}}] \; .$$

Or, par convergence dominée, $\lim_{\lambda\to\infty}\mathbf{E}[\exp(-\lambda W_\infty)\mathbf{1}_{\{W_\infty>0\}}] = 0$, on a donc

$$\mathbf{P}(W_\infty = 0) = \lim_{\lambda\to\infty} L(\lambda) \; .$$

L'égalité fonctionnelle (I.22) implique alors que $\varphi(\mathbf{P}(W_\infty = 0)) = \mathbf{P}(W_\infty = 0)$. D'après le lemme I.1.10, $\mathbf{P}(W_\infty = 0) \in \{q_\xi, 1\}$. Comme $(W_n)_{n\geq 0}$ converge dans $\mathrm{L}^2$, elle converge également dans $\mathrm{L}^1$ et on doit avoir $\mathbf{E}[W_\infty] = \mathbf{E}[W_0] = 1$. Donc on ne peut pas avoir $\mathbf{P}(W_\infty = 0) = 1$. On a donc $\mathbf{P}(W_\infty = 0) = q_\xi$, ce qui permet de conclure. ■

**Remarque I.1.14** Lorsque $\mu > 1$, la loi de $W_\infty$ est en général très difficile à estimer explicitement, bien que sa transformée de Laplace satisfasse l'équation fonctionnelle (I.22). On ne la connaît explicitement que pour de rares exemples. □

L'hypothèse de moment d'ordre deux n'est, bien entendu, pas la meilleure possible. Citons le théorème de Kesten-Stigum qui répond complètement à cette question (sa preuve n'est pas simple).

**Théorème I.1.15** **(Kesten-Stigum)** *Il y a équivalence entre les trois conditions suivantes.*

**(a)** *Presque sûrement* $\mathbf{1}_{\{W_\infty > 0\}} = \mathbf{1}_{\{T=\infty\}}$.

**(b)** *La martingale* $(W_n)_{n\geq 0}$ *converge presque sûrement et dans* $\mathrm{L}^1$ *vers* $W_\infty$.

**(c)** $\sum_{k\geq 2} k\log(k).\xi(k) < \infty$.

**Arbres aléatoires.** Dans cette section on formalise la construction de l'arbre généalogique des processus de branchement introduits précédemment. Un arbre désigne en général un graphe sans cycle dont on connaît les sommets, autrement dit, un *arbre étiqueté*. Si on distingue un point que l'on voit comme la racine, on parle alors d'arbre étiqueté *enraciné*. Dans le modèle que l'on considère, il s'agit d'un arbre généalogique où les individus au sein d'une même fratrie (c'est-à-dire les enfants d'un même individu) sont ordonnés par leur ordre de naissance. D'un point de vue combinatoire l'arbre obtenu est un *arbre ordonné enraciné*. Pour cela on introduit le *formalisme de S. Ulam* des les arbres ordonnés enracinés, formalisme qui constitue une base claire à diverses constructions combinatoires.

On étiquette les individus de la population étudiée par des *mots écrits avec des entiers* de la manière suivante : l'ancêtre est étiqueté par le *mot vide* noté $\varnothing$, le premier enfant de l'ancêtre est étiqueté par le mot d'une lettre $(1)$, le deuxième enfant de l'ancêtre par le mot d'une lettre $(2)$, le troisième par le mot $(3)$, etc. Le premier enfant de l'individu étiqueté par le mot $(1)$ est étiqueté par le mot $(1,1)$, son deuxième enfant par le mot $(1,2)$, etc, et par exemple, le quatrième enfant de $(3)$ est étiqueté pat $(3,4)$, ainsi de suite. On identifie ensuite les individus/sommets de l'arbre généalogique avec leur étiquette. Un arbre enraciné ordonné se voit alors comme un certain sous-ensemble de mots écrits avec des entiers.

Plus précisément, on note $\mathbb{U}$ l'ensemble des mots finis écrits dans l'aphabet $\mathbb{N}^*$ qui se définit formellement par

$$\mathbb{U} = \bigcup_{n\in\mathbb{N}} (\mathbb{N}^*)^n$$

avec la convention que $(\mathbb{N}^*)^0 = \{\varnothing\}$. Avant de définir ce qu'est un arbre dans ce contexte, on introduit quelques notations sur les mots de $\mathbb{U}$. Soit $u = (a_1, \dots, a_n) \in \mathbb{U}$.

- On note $|u| = n$ la *longueur du mot* $u$ (ou sa *hauteur* ou encore sa *génération*) avec la convention que $|\varnothing| = 0$.
- Pour tout entier $0 \leq k \leq n = |u|$, on note $u_{|k} = (a_1, \dots, a_k)$ qui est l'ancêtre de $u$ à la génération $k$. On adopte la convention que $u_{|0} = \varnothing$ et on observe que $u_{|\,|u|} = u$. On note $[\![\varnothing, u]\!] = \{u_{|k}; 0 \leq k \leq |u|\}$ qui est *la lignée ancestrale de* $u$, avec les notations annexes $[\![\varnothing, u[\![ = [\![\varnothing, u]\!] \backslash \{u\}$, $]\!]\varnothing, u]\!] = [\![\varnothing, u]\!] \backslash \{\varnothing\}$ et $]\!]\varnothing, u[\![ = [\![\varnothing, u]\!] \backslash \{\varnothing, u\}$.
- On note $\overleftarrow{u} = u_{|n-1} = (a_1, \dots, a_{n-1})$, le *parent* de $u$, avec la convention $\overleftarrow{\varnothing} = \varnothing$.
- Soit $v = (b_1, \dots, b_m) \in \mathbb{U}$. On note $u * v = (a_1, \dots, a_n, b_1, \dots, b_m)$ la *concaténation de* $u$ *avec* $v$. On adopte la convention que $u * \varnothing = \varnothing * u = u$. On observe que $|u * v| = |u| + |v|$.
- Soient $u, v \in \mathbb{U}$. On note $u \wedge v \in \mathbb{U}$ le mot qui est *le plus grand préfixe commun à* $u$ *et* $v$, c'est-à-dire que $u \wedge v = u_{|m} = v_{|m}$ où $m = \max\big\{k \in \{0, \dots, \min(|u|, |v|)\} : u_{|k} = v_{|k}\big\}$. Dans l'interprétation des arbres en termes de généalogies, $u \wedge v$ est *le plus récent ancêtre commun* à $u$ et $v$.

**Définition I.1.16** Un *arbre ordonné et enraciné* est un sous-ensemble $t \subset \mathbb{U}$ satisfaisant les conditions suivantes.

$(a)$ $\varnothing \in t$ et si $u \in t$, alors $\overleftarrow{u} \in t$.

$(b)$ Pour tout $u \in t$, il existe $k_u(t) \in \mathbb{N}$ tel que $u * (j) \in t$ si et seulement si $1 \leq j \leq k_u(t)$ (ici le mot $u * (j)$ est la concaténation du mot $u$ avec le mot d'une lettre $(j)$).

Si $k_u(t) = 0$, la condition $(b)$ est vide et si on voit $t$ comme un arbre généalogique, $k_u(t)$ est appelé le *nombre d'enfants de* $u$ dans $t$. La racine de $t$ est le mot vide $\varnothing$.

On note $\mathbb{T}$ l'ensemble des arbres ordonnés enracinés, $\mathbb{T}^f$ l'ensemble des arbres enracinés et ordonnés *finis* et pour tout entier $n \geq 1$, on pose $\mathbb{T}_n = \{t \in \mathbb{T} : \#t = n\}$. $\square$

Soit $t \in \mathbb{U}$ et $u \in t$. Si on voit $t$ comme l'arbre généalogique d'une population issue de l'ancêtre $\varnothing$, on note $\theta_u t$ *l'arbre des descendants de* $u$ dans $t$. On voit $\theta_u t$ comme un arbre enraciné et ordonné ce qui conduit à la définition formelle suivante

$$\theta_u t = \{v \in \mathbb{U} : u * v \in t\} . \tag{I.23}$$

On observe que $\theta_\varnothing t = t$.

Pour tout $t \in \mathbb{T}$ et tout $n \in \mathbb{N}$, on introduit l'arbre $t_{|n}$ des $n$ premières générations :

$$t_{|n} = \{u \in t : |u| \leq n\} .$$

On observe que $t_{|n} = \{u_{|n}; u \in t\}$, que $t_{|n} \in \mathbb{T}^f$ et que $t_{|0} = \{\varnothing\}$.

On cherche a définir des arbres aléatoires. Pour cela il est nécessaire d'introduire une tribu sur $\mathbb{T}$ qui ne soit ni trop grosse (car il n'existerait alors que très peu d'arbres aléatoires) ni trop petite (car certaines opérations élémentaires ne seraient pas mesurables). Il semble nécessaire de pouvoir évaluer la probabilité qu'un sommet $u \in \mathbb{U}$ donné appartienne à un arbre aléatoire donc il est naturel d'introduire la tribu suivante.

**Définition I.1.17** On note $\mathscr{F}(\mathbb{T})$ la tribu engendrée sur $\mathbb{T}$ par les sous-ensembles $A_u = \{t \in \mathbb{T} : u \in t\}$.

$$\mathscr{F}(\mathbb{T}) = \sigma\big(A_u; u \in \mathbb{U}\big) .$$

Soit $(\Omega, \mathscr{F})$, un espace mesurable. Une application $\tau : \Omega \to \mathbb{T}$ est un *arbre aléatoire $\mathscr{F}$-mesurable* si c'est une application $(\mathscr{F}, \mathscr{F}(\mathbb{T}))$-mesurable, autrement dit, si et seulement si l'événément $\{u \in \tau\} \in \mathscr{F}$ pour tout $u \in \mathbb{U}$. $\square$

**Lemme I.1.18** *Pour tout* $t \in \mathbb{T}$ *et tout* $n \in \mathbb{N}$*, on pose* $B(t, n) = \{t' \in \mathbb{T} : t'_{|n} = t_{|n}\}$*. On pose ensuite*

$$\mathscr{C} = \{\emptyset\} \cup \big\{B(t, n);\, t \in \mathbb{T}, n \in \mathbb{N}\big\} .$$

*Alors les assertions suivantes sont vérifiées.*

$(i)$ *$\mathscr{C}$ est un pi-système qui engendre $\mathscr{F}(\mathbb{T})$ : $\sigma(\mathscr{C}) = \mathscr{F}(\mathbb{T})$.*

$(ii)$ *Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité et $\tau : \Omega \to \mathbb{T}$, un arbre aléatoire $\mathscr{F}$-mesurable. Sa loi est caractérisée par les probabilités $\mathbf{P}(\tau_{|n} = t_{|n})$, pour tous $t \in \mathbb{T}$ et $n \in \mathbb{N}$.*

**Preuve :** $B(t, n) = \bigcap_{u \in t_{|n}} A_u$ (avec la notation $A_u$ de la définition I.1.17). Comme $t_{|n}$ est fini et que par définition les $A_u \in \mathscr{F}(\mathbb{T})$, on en déduit que $B(t, n) \in \mathscr{F}(\mathbb{T})$ et donc que $\mathscr{C} \subset \mathscr{F}(\mathbb{T})$, ce qui implique que $\sigma(\mathscr{C}) \subset \mathscr{F}(\mathbb{T})$.

On observe ensuite que $A_u = \bigcup_{t \in \mathbb{T}^f \cap A_u} B(t, |u|)$. Comme $\mathbb{T}^f$ est dénombrable, on en déduit que $A_u \in \sigma(\mathscr{C})$. Or $\mathscr{F}(\mathbb{T})$ est par définition la plus petite tribu contenant les $A_u$, $u \in \mathbb{U}$ ; cela entraîne donc que $\mathscr{F}(\mathbb{T}) \subset \sigma(\mathscr{C})$ et donc que $\sigma(\mathscr{C}) = \mathscr{F}(\mathbb{T})$.

Soient $t_1, t_2 \in \mathbb{T}$ et soient $n_1, n_2 \in \mathbb{N}$ tels que $n_1 \leq n_2$. $B(t_1, n_1) \cap B(t_2, n_2) = B(t_2, n_2)$ si $t_{2\,|n_1} = t_{1\,|n_1}$ et $B(t_1, n_1) \cap B(t_2, n_2) = \emptyset$ sinon. Cela montre que $\mathscr{C}$ est stable par intersection simple. De plus $\mathbb{T} = B(t, 0)$, donc $\mathscr{C}$ est un pi-système, ce qui termine la preuve du $(i)$.

Montrons $(ii)$. On note $\mathtt{m}$ la loi de $\tau$ sous $\mathbf{P}$. On observe que pour tout $t \in \mathbb{T}$ et tout $n \in \mathbb{N}$, $\{\tau_{|n} = t_{|n}\} = \{\tau \in B(t, n)\}$ et donc $\mathbf{P}(\tau_{|n} = t_{|n}) = \mathtt{m}(B(t, n))$. Soient $(\Omega', \mathscr{F}', \mathbf{P}')$, un espace de probabilité et $\tau' : \Omega' \to \mathbb{T}$, un arbre aléatoire $\mathscr{F}'$-mesurable. On note $\mathtt{m}'$ la loi de $\tau'$ sous $\mathbf{P}'$. On suppose que $\mathbf{P}(\tau_{|n} = t_{|n}) = \mathbf{P}'(\tau'_{|n} = t_{|n})$. Alors $\mathtt{m}(B(t, n)) = \mathtt{m}'(B(t, n))$, c'est-à-dire que les mesures de probabilité $\mathtt{m}$ et $\mathtt{m}'$ sur l'espace mesurable $(\mathbb{T}, \mathscr{F}(\mathbb{T}))$ coïncident sur le pi-système $\mathscr{C}$. Comme $\mathscr{C}$ engendre $\mathscr{F}(\mathbb{T})$, le théorème d'unicité du prolongement des mesures implique que $\mathtt{m} = \mathtt{m}'$, ce qui est le résultat voulu. ■

**Définition I.1.19** Soit $\xi = (\xi(k))_{k \in \mathbb{N}}$, une loi de reproduction, c'est-à-dire une mesure de probabilité sur $\mathbb{N}$. Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité et $\tau : \Omega \to \mathbb{T}$, un arbre aléatoire $\mathscr{F}$-mesurable. C'est un *arbre de Galton-Watson de loi de reproduction* $\xi$ s'il satisfait les deux conditions suivantes.

$(a)$ Pour tout $k \in \mathbb{N}$, $\mathbf{P}(k_\varnothing(\tau) = k) = \xi(k)$.

$(b)$ Soit $k \geq 1$ tel que $\xi(k) > 0$. Alors sous $\mathbf{P}(\,\cdot\,|k_\varnothing(\tau) = k)$ les arbres des descendants des $k$ enfants de l'ancêtre $\varnothing$, arbres notés $\theta_{(1)}\tau, \ldots, \theta_{(k)}\tau$ sont indépendants et de même loi que $\tau$. □

**Une construction.** On introduit tout d'abord la notation suivante pour tout $u = (a_1, \ldots, a_n) \in \mathbb{U}$ :

$$\mathbf{a}(u) = a_n \quad \text{avec la convention que} \quad \mathbf{a}(\varnothing) = 0.$$

On observe ici que $a_k = \mathbf{a}(u_{|k})$, $1 \leq k \leq |u| = n$. On rappelle ensuite que $\mathbb{U}$ est dénombrable.

Soit $\xi = (\xi(k))_{k \in \mathbb{N}}$, une loi de reproduction. Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité qui est supposé suffisamment grand pour qu'y puisse être définie une suite i.i.d. de variables aléatoires réelles de loi arbitraire. On peut donc y définir la famille de variable i.i.d. $X(u) : \Omega \to \mathbb{N}^*$, $u \in \mathbb{U}$, de loi $\xi$. On pose alors

$$\tau = \{\varnothing\} \cup \big\{ u \in \mathbb{U} \backslash \{\varnothing\} : \forall k \in \{1, \ldots, |u|\},\ \mathbf{a}(u_{|k}) \leq X(u_{|k-1}) \big\} . \tag{I.24}$$

Montrons que $\tau$ est bien un arbre ordonné et enraciné comme dans la définition I.1.16. Par définition, $\varnothing \in \tau$ et on vérifie immédiatement que si $u \in \tau$ alors $\overleftarrow{u} \in \tau$. On voit ensuite que si $u \in \tau$, alors $u * (j) \in \tau$ si et seulement si $1 \leq j \leq X(u)$, c'est-à-dire que $\tau$ satisfait la condition $(b)$ de la définition I.1.16 avec $k_u(\tau) = X(u)$. Donc $\tau \in \mathbb{T}$.

Montrons ensuite la mesurabilité de $\tau$ : soit $u \in \mathbb{U}$. On observe que

$$\{u \in \tau\} = \bigcap_{0 \leq k \leq |u|} \{\mathbf{a}(u_{|k}) \leq X(u_{|k-1})\} \in \mathscr{F}$$

car les $X(u)$ sont $\mathscr{F}$-mesurables. Cela montre donc que $\tau$ est un arbre aléatoire $\mathscr{F}$-mesurable comme à la définition I.1.17.

Montrons ensuite que $\tau$ est bien un arbre de Galton-Watson de loi de reproduction $\xi$ comme introduit à la définition I.1.19. Comme remarqué précédemment, $k_\varnothing(\tau) = X(\varnothing)$ a pour loi $\xi$. Montrons que $\tau$ satisfait le point $(b)$ de la définition I.1.19 : pour tout $j \in \mathbb{N}^*$, on pose $X_j(u) = X((j) * u)$, $u \in \mathbb{U}$. On observe que les familles $(X_j(u))_{u \in \mathbb{U}}$, $j \in \mathbb{N}^*$ sont d'une part mutuellement indépendantes, d'autre part toutes indépendantes de $X(\varnothing)$ et enfin que pour tout $j$ fixé, $(X_j(u))_{u \in \mathbb{U}}$ a même loi que $(X(u))_{u \in \mathbb{U}}$. Pour tout $j \in \mathbb{N}^*$, on pose alors

$$\tau_j = \{\varnothing\} \cup \big\{ u \in \mathbb{U} \backslash \{\varnothing\} : \forall k \in \{1, \ldots, |u|\},\ \mathbf{a}(u_{|k}) \leq X_j(u_{|k-1}) \big\} .$$

On en déduit que les $\tau_j$, $j \in \mathbb{N}^*$, sont d'une part mutuellement indépendants, d'autre part tous indépendants de $X(\varnothing)$ et enfin qu'ils ont même loi que $\tau$. On remarque alors que $\theta_{(j)}\tau = \tau_j$ pour tout $1 \leq j \leq k_\varnothing(\tau) = X(\varnothing)$, ce qui entraîne facilement la condition $(b)$ de la définition I.1.19.

Cela montre que *l'arbre $\tau$ construit par (I.24) à partir de la famille de variable i.i.d. $(X(u))_{u\in\mathbb{U}}$ de loi commune $\xi$ est bien un arbre de Galton-Watson de loi de reproduction $\xi$.* ■

**Lemme I.1.20** *Soient $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité et $\tau : \Omega \to \mathbb{T}$, un arbre de Galton-Watson $\mathscr{F}$-mesurable de loi de reproduction $\xi = (\xi(k))_{k\in\mathbb{N}}$. On suppose que $\xi(0) + \xi(1) < 1$. On note rappelle que $m_\xi$ est la moyenne de $\xi$, que $\varphi_\xi$ est la fonction génératrice de $\xi$ et que $q_\xi$ est la plus petite solution sur $[0, 1]$ de l'équation $\varphi_\xi(r) = r$. Les assertions suivantes sont vérifiées.*

*$(i)$ Pour tout $t \in \mathbb{T}$ et tout $n \in \mathbb{N}^*$, $\mathbf{P}(\tau_{|n} = t_{|n}) = \prod_{u\in t:|u|<n} \xi(k_u(t))$ et la loi de $\tau$ est entièrement caractérisée par $\xi$. De plus*
$$\forall t \in \mathbb{T}^f, \quad \mathbf{P}(\tau = t) = \prod_{u\in t} \xi(k_u(t)) \ . \tag{I.25}$$

*$(ii)$ Pour tout $r \in [0, 1]$, on pose $g(r) = \mathbf{E}[r^{\#\tau}\mathbf{1}_{\{\#\tau<\infty\}}]$. La fonction $g$ vérifie l'équation fonctionnelle $g(r) = r\varphi_\xi(g(r))$, $r \in [0, 1]$.*

*$(iii)$ On a $\mathbf{P}(\#\tau < \infty) = q_\xi$. Donc si $m_\xi \leq 1$, alors $\mathbf{P}$-p.s. $\tau$ est un arbre fini et si $m_\xi > 1$, alors $\mathbf{P}(\#\tau = \infty) = 1 - q_\xi > 0$.*

**Preuve :** on fixe $t \in \mathbb{T}$ et on montre le premier point de $(i)$ par récurrence sur $n \geq 1$. On observe d'abord que $\{\tau_{|1} = t_{|1}\} = \{k_\varnothing(\tau) = k_\varnothing(t)\}$ et donc $\mathbf{P}(\tau_{|1} = t_{|1}) = \xi(k_\varnothing(t))$.

Supposons $(i)$ vrai jusqu'au rang $n \geq 1$. Pour simplifier les notations on pose $k = k_\varnothing(t)$. On observe alors que
$$\{\tau_{|n+1} = t_{|n+1}\} = \{k_\varnothing(\tau) = k\} \bigcap_{1\leq j\leq k} \{(\theta_{(j)}\tau)_{|n} = (\theta_{(j)}t)_{|n}\} \ .$$

Par la propriété $(b)$ de la définition I.1.19, on en déduit que
$$\mathbf{P}(\tau_{|n+1} = t_{|n+1}) = \mathbf{P}(k_\varnothing(\tau) = k) \prod_{1\leq j\leq k} \mathbf{P}(\tau_{|n} = (\theta_{(j)}t)_{|n})$$

et l'hypothèse de récurrence s'applique au rang $n$ pour montrer que $\mathbf{P}(\tau_{|n} = (\theta_{(j)}t)_{|n}) = \prod_{u\in U_j} \xi(k_u(t))$ où $U_j = \{(j) * u \in \mathbb{U} : |u| < n \text{ et } (j)*u \in t\}$. Comme les $U_1, \ldots, U_k$, forment une partition de $\{u \in t : 1 \leq |u| < n + 1\}$, on a bien
$$\mathbf{P}(\tau_{|n+1} = t_{|n+1}) = \xi(k) \prod_{1\leq j\leq k} \prod_{u\in U_j} \xi(k_u(t)) = \prod_{u\in t:|u|<n+1} \xi(k_u(t)) \ ,$$

ce qui montre bien le premier point de $(i)$ par récurrence. Ce premier point combiné au lemme I.1.18 $(ii)$ implique que $\xi$ caractérise la loi de $\tau$.

Enfin on se donne $t \in \mathbb{T}^f$. Il existe $n$ tel que $|u| < n$ pour tout $u \in t$ et on observe alors que $t_{|n} = t$ et que $\{\tau_{|n} = t_{|n}\} = \{\tau = t\}$ et le premier point de $(i)$ implique alors (I.25).

Montrons $(ii)$. On observe que si $k_\varnothing(\tau) = k \geq 1$, $\#\tau$ est fini si et seulement si les sous-arbres $\theta_{(1)}\tau, \ldots, \theta_{(k)}\tau$ le sont aussi et dans ce cas $\#\tau = 1 + \sum_{1\leq j\leq k} \#(\theta_{(j)}\tau)$. Par interversion série/ espérance positive, on a donc
$$g(r) = r\mathbf{P}(k_\varnothing(\tau) = 0) + \sum_{k\geq 1} \mathbf{E}\Big[\mathbf{1}_{\{k_\varnothing(\tau)=k\}} r \prod_{1\leq j\leq} r^{\#\theta_{(j)}\tau}\mathbf{1}_{\{\#\theta_{(j)}\tau<\infty\}}\Big] \ .$$

Par les conditions $(a)$ et $(b)$ de la définition I.1.19, on a bien pour tout $k \geq 1$,

$$\mathbf{E}\Big[\mathbf{1}_{\{k_\varnothing(\tau)=k\}} r \prod_{1\leq j\leq} r^{\#\theta_{(j)}\tau} \mathbf{1}_{\{\#\theta_{(j)}\tau<\infty\}}\Big] = r\xi(k)g(r)^k,$$

ce qui implique que $g(r) = r\varphi(g(r))$.

Montrons $(iii)$ à l'aide de $(ii)$. Si $\xi(0) = 0$, alors clairement $\#\tau = \infty$ et donc $\mathbf{P}(\#\tau < \infty) = 0 = q_\xi$. Supposons que $\xi(0) > 0$ ce qui implique que $q_\xi > 0$. Pour tout $r \in [0, 1[$ , on a $g(r) = r\varphi_\xi(g(r)) < \varphi_\xi(g(r))$. Par (I.13), cela implique que $g(r) < q_\xi$. Or $g(1^-) = \lim_{r\to 1^-} g(r) = \mathbf{P}(\#\tau < \infty)$ satisfait $g(1^-) = \varphi(g(1^-))$ : c'est donc un point fixe de $\varphi_\xi$ et on en déduit nécessairement que $g(1^-) = q_\xi$. Le reste du point $(iii)$ est une conséquence du lemme I.1.10. ■

La proposition suivante montre que le cas particulier des arbres de Galton-Watson de loi de reproduction géométrique permet de réaliser la loi d'un arbre enraciné ordonné uniforme à $n$ sommets. Voir la figure I.1.b pour une simulation d'un arbre enraciné ordonné uniforme à $n = 5000$ sommets.

**Proposition I.1.21** *Soit $\rho \in ]0, 1/2]$. On pose $\xi(k) = (1-\rho)\rho^k$, $k \in \mathbb{N}$. Soient $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité et $\tau : \Omega \to \mathbb{T}$, un arbre de Galton-Watson $\mathscr{F}$-mesurable de loi de reproduction $\xi = (\xi(k))_{k\in\mathbb{N}}$. Alors $\tau$ sous $\mathbf{P}(\,\cdot\,|\#\tau = n)$ est de loi uniforme sous $\mathbb{T}_n$, l'ensemble des arbres enracinés ordonnés à $n$ sommets. On a de plus*

$$\#\mathbb{T}_n = \frac{1}{4n-2}\binom{2n}{n} \sim_{n\to\infty} \frac{1}{2\sqrt{\pi}} n^{-3/2} 4^n \ . \tag{I.26}$$

**Preuve :** soit $t \in \mathbb{T}^f$. On observe que $\sum_{u\in t} k_u(t) = \#t - 1$. Par (I.25), on a

$$\mathbf{P}(\tau = t) = \prod_{u\in t} \xi(k_u(t)) = (1-\rho)^{\#t}\rho^{\#t-1}$$

qui ne dépend que de $\rho$ et de $\#t$. Donc

$$\mathbf{P}(\#\tau = n) = \sum_{t\in\mathbb{T}_n} \mathbf{P}(\tau = t) = (1-\rho)^n\rho^{n-1}\#\mathbb{T}_n \ ,$$

ce qui montre que $\mathbf{P}(\tau = t|\#\tau = n) = 1/(\#\mathbb{T}_n)$, pour tout $t \in \mathbb{T}_n$, qui est le premier point de la proposition.

On observe ensuite que $m_\xi = \rho/(1-\rho) \leq 1$, $\mathbf{P}$-p.s. $\#\tau < \infty$ et $g(r) = \mathbf{E}[r^{\#\tau}]$ satisfait l'équation $g(r) = r\varphi_\xi(g(r))$, $r \in [0, 1]$. Or ici $\varphi_\xi(r) = (1-\rho)/(1-\rho r)$ donc $g$ satisfait $g(r)(1-\rho g(r)) = r(1-\rho)$ c'est-à-dire que

$$g(r) = \frac{1 - \sqrt{1 - 4\rho(1-\rho)r}}{2\rho} \ , \quad r \in [0, 1) \ .$$

Pour simplifier on choisit $\rho = 1/2$ et on a le développement en série entière

$$g(r) = 1 - \sqrt{1-r} = \sum_{n\geq 1} (-1)^{n+1}\binom{1/2}{n} r^n \ , r \in [0, 1] \ .$$

où $n!\binom{1/2}{n} = 1/2(1/2-1)(1/2-2)\dots(1/2-n+1)$. Or $g(r) = \sum_{n\geq 0} r^n\mathbf{P}(\#\tau = n)$ par définition. Donc par identification des coefficients, on a

$$\mathbf{P}(\#\tau = n) = (-1)^{n+1}\binom{1/2}{n} = \frac{(2n-3).(2n-5)\dots 3.1}{n!2^n} = \frac{1}{2^{2n}(2n-1)}\binom{2n}{n} \ .$$

Or, dans le cas où $\rho = 1/2$, on a montré que $\mathbf{P}(\#\tau = n) = 2.4^{-n}\#\mathbb{T}_n$, ce qui entraîne (I.26). ■

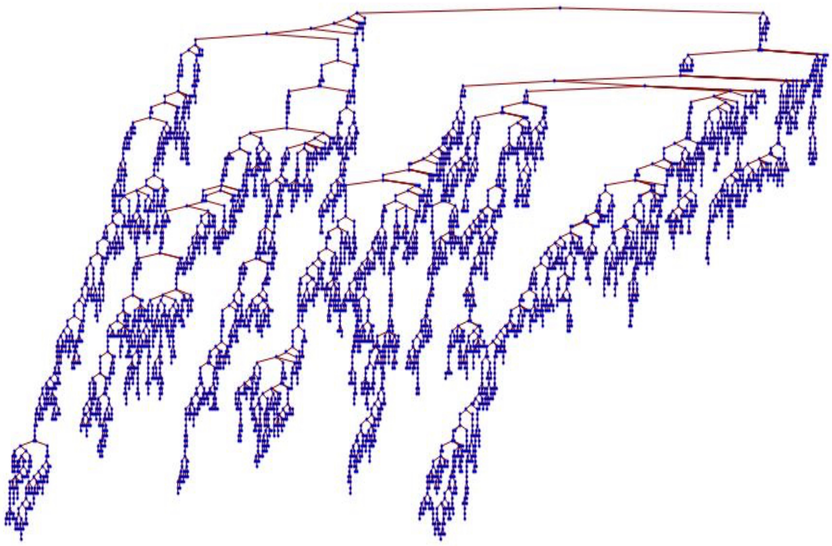

FIGURE I.3 – Une représentation d'un arbre enraciné ordonné uniforme à $n$ sommets avec $n = 5000$.

**Lien avec les processus de branchement.** Pour tout $t \in \mathbb{T}$ et tout $n \in \mathbb{N}$, on pose

$$\mathcal{L}_n(t) = \{u \in t : |u| = n\} \quad \text{et} \quad Z_n(t) = \#\mathcal{L}_n(t) \tag{I.27}$$

qui sont respectivement l'ensemble des sommets de hauteur $n$ dans $t$ et leur nombre. La fonction $Z_n : \mathbb{T} \to \mathbb{N}$ est $\mathscr{F}(\mathbb{T})$-mesurable car elle s'écrit $Z_n = \sum_{u \in \mathbb{U}:|u|=n} \mathbf{1}_{A_u}$, c'est-à-dire comme une somme dénombrable de fonctions indicatrices d'événements appartenant à $\mathscr{F}(\mathbb{T})$.

Soit $\xi = (\xi(k))_{k\in\mathbb{N}}$, une loi de reproduction. On rappelle que pour tout $i \in \mathbb{N}$, $\xi^{*i} \in \mathscr{M}_1(\mathbb{N})$ est la loi $\xi$ convolée $i$ fois avec elle-même, en adoptant la convention $\xi^{*0} = \delta_0$. On rappelle qu'un processus de branchement de loi de reproduction $\xi$ est une chaîne de Markov de matrice de transition $Q = (p(i,j))_{i,j\in\mathbb{N}}$ donnée par $p(i,j) = (\xi^{*i})(j)$. La proposition suivante montre que si $\tau$ est un arbre de Galton-Watson, $(Z_n(\tau))_{n\in\mathbb{N}}$ est un processus de branchement de même loi de reproduction que $\tau$.

**Lemme I.1.22** *Soient $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité et $\tau : \Omega \to \mathbb{T}$, un arbre de Galton-Watson $\mathscr{F}$-mesurable de loi de reproduction $\xi = (\xi(k))_{k\in\mathbb{N}}$. Pour tout $n \in \mathbb{N}$, on note $\mathscr{F}_n$ la tribu engendrée par $\tau_{|n}$. Les assertions suivantes sont vérifiées.*

$(i)$ *Soit $n \in \mathbb{N}$. Alors $\mathcal{L}_n(\tau)$ est $\mathscr{F}_n$-mesurable et conditionnellement à $\mathscr{F}_n$ les sous-arbres $\theta_u\tau$, $u \in \mathcal{L}_n(\tau)$, sont mutuellement indépendants de même loi que $\tau$ sous $\mathbf{P}$.*

$(ii)$ *$(Z_n(\tau))_{n\geq 0}$ est un $(\mathscr{F}_n)_{n\geq 0}$ processus de branchement de loi de reproduction $\xi$ de valeur initiale $Z_0 = 1$.*

**Preuve :** on montre $(i)$ par récurrence sur $n$, la cas $n = 1$ étant une reformulation de la condition $(b)$ de la définition I.1.19. On suppose $(i)$ vérifiée au rang $n$. On fixe $t \in \mathbb{U}$ tel que $t_{|n+1} = t$ et pour tout $u \in \mathcal{L}_{n+1}(t)$, on fixe des fonctions $F_u : \mathbb{T} \to \mathbb{R}$ mesurable bornées. On observe que

$$\mathbf{1}_{\{\tau_{|n+1}=t\}} \prod_{u\in\mathcal{L}_{n+1}(t)} F_u(\theta_u\tau) = \mathbf{1}_{\{k_\varnothing(\tau)=k_\varnothing(t)\}} \prod_{1\leq j\leq k_\varnothing(t)} \Big(\mathbf{1}_{\{(\theta_{(j)}\tau)_{|n}=\theta_{(j)}t\}} \prod_{v\in\mathcal{L}_n(\theta_{(j)}t)} F_{(j)*v}(\theta_v(\theta_{(j)}\tau))\Big)\,.$$

En appliquant la condition $(b)$ de la définition I.1.19, on obtient

$$\mathbf{E}\Big[\mathbf{1}_{\{\tau_{|n+1}=t\}}\prod_{u\in\mathcal{L}_{n+1}(t)}F_u(\theta_u\tau)\Big]=\xi(k_\varnothing(t))\prod_{1\leq j\leq k_\varnothing(t)}\mathbf{E}\Big[\mathbf{1}_{\{\tau_{|n}=\theta_{(j)}t\}}\prod_{v\in\mathcal{L}_n(\theta_{(j)}t)}F_{(j)*v}(\theta_v(\tau))\Big]\ .$$

L'hypothèse de récurrence implique pour tout $j\in\{1,\ldots,k_\varnothing(t)\}$ que

$$\mathbf{E}\Big[\mathbf{1}_{\{\tau_{|n}=\theta_{(j)}t\}}\prod_{v\in\mathcal{L}_n(\theta_{(j)}t)}F_{(j)*v}(\theta_v(\tau))\Big]=\mathbf{P}(\tau_{|n}=\theta_{(j)}t)\prod_{v\in\mathcal{L}_n(\theta_{(j)}t)}\mathbf{E}[F_{(j)*v}(\tau)]\ .$$

Donc

$$\begin{aligned}\mathbf{E}\Big[\mathbf{1}_{\{\tau_{|n+1}=t\}}\prod_{u\in\mathcal{L}_{n+1}(t)}F_u(\theta_u\tau)\Big] &= \xi(k_\varnothing(t))\prod_{1\leq j\leq k_\varnothing(t)}\Big(\mathbf{P}(\tau_{|n}=\theta_{(j)}t)\prod_{v\in\mathcal{L}_n(\theta_{(j)}t)}\mathbf{E}[F_{(j)*v}(\tau)]\Big)\\ &= \xi(k_\varnothing(t))\Big(\prod_{1\leq j\leq k_\varnothing(t)}\mathbf{P}(\tau_{|n}=\theta_{(j)}t)\Big)\prod_{u\in\mathcal{L}_{n+1}(t)}\mathbf{E}\big[F_u(\tau)\big]\end{aligned}$$

car les ensembles $\{(j)*v\,;\,v\in\mathcal{L}_n(\theta_{(j)}t)\}$, $1\leq j\leq k_\varnothing(t)$, forment une partition de $\mathcal{L}_{n+1}(t)$. En prenant les fonctions $F_u$ égales à 1 dans ce qui précède, on obtient $\mathbf{P}(\tau_{|n+1}=t)=\xi(k_\varnothing(t))\prod_{1\leq j\leq k_\varnothing(t)}\mathbf{P}(\tau_{|n}=\theta_{(j)}t)$ et donc

$$\mathbf{E}\Big[\mathbf{1}_{\{\tau_{|n+1}=t\}}\prod_{u\in\mathcal{L}_{n+1}(t)}F_u(\theta_u\tau)\Big]=\mathbf{P}(\tau_{|n+1}=t)\prod_{u\in\mathcal{L}_{n+1}(t)}\mathbf{E}\big[F_u(\tau)\big]\ .$$

ce qui entraîne le résultat pour $n+1$ et ce qui complète la preuve de $(i)$.

Montrons $(ii)$ : pour cela on observe que

$$Z_{n+1}(\tau)=\sum_{u\in\mathcal{L}_n(\tau)}k_u(\tau)=\sum_{u\in\mathcal{L}_n(\tau)}k_\varnothing(\theta_u\tau)\ .$$

Donc par $(i)$, conditionnellement à $\mathscr{F}_n$, $Z_{n+1}(\tau)$ est la somme de $Z_n(\tau)$ variables (conditionnellement) indépendantes de loi $\xi$, c'est-à-dire que conditionnelement à $\mathscr{F}_n$, $Z_{n+1}(\tau)$ a pour loi $\xi^{*Z_n(\tau)}$. Pour toute fonction $f:\mathbb{N}\to[0,\infty[$, on a par conséquent

$$\mathbf{E}[F(Z_{n+1}(\tau))|\mathscr{F}_n]=\sum_{j\in\mathbb{N}}f(j)\xi^{*Z_n(\tau)}(j)=(Q.f)(Z_n(\tau)),$$

ce qui prouve $(ii)$. ■

## EXERCICES.

**Exercice I.12** Montrer la proposition I.2.4. □

### I.1.c Construction comme système dynamique aléatoire.

On fournit ici une construction des chaînes de Markov comme des *systèmes dynamiques aléatoires*. Cette construction peut aussi se voir comme une méthode de simulation mais elle est très naïve et n'est en général pas praticable quand l'espace d'états est grand.

**Définition I.1.23** (*Fonctions d'échantillonnage*) Soit $\mathbf{P}$, une loi de probabilité sur $(\Omega, \mathscr{F})$. Soit $U : \Omega \to [0,1[$, une v.a. de loi uniforme.

(a) (*Echantillonnage d'une loi*) Soit $\mu \in \mathscr{M}_1(E)$. Une fonction d'échantillonnage de $\mu$ est une fonction $\phi_\mu : [0,1[ \to E$, mesurable telle que

$$\forall i \in E\ , \quad \mu(i) = \mathbf{P}\big(\, \phi_\mu(U) = i\,\big) = \int_0^1 \mathbf{1}_{\{\phi_\mu(x)=i\}} dx\ .$$

(b) (*Echantillonnage d'une matrice de transition*) Soit $Q = (p(i,j))_{i,j\in E}$, une matrice de transition. Une fonction d'échantillonnage de $Q$ est une fonction $\Phi_Q : [0,1[ \times E \to E$, mesurable telle que

$$\forall i,j \in E\ , \quad p(i,j) = \mathbf{P}\big(\, \Phi_Q(U,i) = j\,\big) = \int_0^1 \mathbf{1}_{\{\Phi_Q(x,i)=j\}} dx\ .$$

Autrement dit $\Phi_Q(\cdot, i)$ est une fonction d'échantillonnage de la probabilité $p(i,\cdot)$. □

**Lemme I.1.24** *Toute mesure de probabilité $\mu$ sur $E$, ainsi que toute matrice de transition $Q = (p(i,j))_{i,j\in E}$ sur $E$ admettent des fonctions d'échantillonnage.*

**Preuve :** on pose $N = \#E$. Si $E$ est fini, $N$ est un entier et on indexe les éléments de $E$ d'une façon quelconque par $i_0, \ldots, i_{N-1}$. Si $E$ est infini, alors $N = \infty$ et $E$ est en bijection avec les entiers $\mathbb{N}$ : on indexe les éléments de $E$ en une suite d'éléments distincts, suite notée $(i_n)_{n\in\mathbb{N}}$. Dans les deux cas, on se donne une indexation des éléments de $E$ par les entiers, qui est donnée par $E = \big\{i_n\,;\, 0 \leq n < N\big\}$ où $N = \#E$.

On définit une première fonction $\phi_\mu : [0,1[ \to E$ de la manière suivante :

— $\forall x \in [0, \mu(i_0)[$, on pose $\phi_\mu(x) = i_0$ ;

— $\forall\ 0 \leq n < N, \forall x \in [0,1[$ tel que $\sum_{0\leq p\leq n} \mu(i_p) \leq x < \sum_{0\leq p\leq n+1} \mu(i_p)$, on pose $\phi_\mu(x) = i_{n+1}$.

On fixe $i \in E$. Il existe un unique entier $0 \leq n < N$ tel que $i = i_n$.

$$\text{Si } n \geq 1, \quad \mathbf{P}(\phi_\mu(U) = i) = \mathbf{P}\Big( \sum_{0\leq p\leq n-1} \mu(i_p) \leq U < \sum_{0\leq p\leq n} \mu(i_p)\Big) = \mu(i_n) = \mu(i)\ .$$

Si $n = 0$, alors $i = i_0$ et $\mathbf{P}(\phi_\mu(U) = i) = \mathbf{P}(U \leq \mu(i_0)) = \mu(i_0) = \mu(i)$. Cela montre que $\phi_\mu$ est une fonction d'échantillonnage de $\mu$.

On construit ensuite une fonction $\Phi_Q : [0,1[ \times E \to E$ de la manière suivante : on fixe $i \in E$ et

— $\forall x \in [0, p(i,i_0)[$, on pose $\Phi_Q(x,i) = i_0$ ;

— $\forall\ 0 \leq n < N$ et $\forall x \in [0,1[$ tel que $\sum_{0\leq p\leq n} p(i,i_p) \leq x < \sum_{0\leq p\leq n+1} p(i,i_p)$, on pose $\Phi_Q(x,i) = i_{n+1}$.

En raisonnant de même, on voit que $\Phi_Q$ est une fonction d'échantillonnage de $Q$. ■

La proposition suivante montre que les chaînes de Markov sont, en quelque sorte, une généralisation des suites définies par récurrence, c'est-à-dire des systèmes dynamiques aléatoires.

**Proposition I.1.25** *Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité sur lequel est définie une suite $U_n : \Omega \to [0,1[$, $n \in \mathbb{N}$, de v.a. indépendantes de loi uniforme. Soit $\mu$, une mesure de probabilité et soit $Q$, une matrice de transition. Soient $\phi_\mu$ et $\Phi_Q$, des fonctions d'échantillonnage de $\mu$ et $Q$. On définit récursivement une suite $\mathbf{X} = (X_n)_{n\geq 0}$ de v.a. à valeurs dans $E$ en posant*

$$X_0 = \phi_\mu(U_0) \quad \text{et} \quad \forall n \in \mathbb{N}, \quad X_{n+1} = \Phi_Q(U_{n+1}, X_n) \ .$$

*Alors, sous $\mathbf{P}$, $\mathbf{X}$ est une chaîne de Markov de loi d'entrée $\mu$ et de matrice de transition $Q$.*

**Preuve :** on remarque tout d'abord que $\sigma(X_0, \ldots, X_n) \subset \sigma(U_0, \ldots, U_n)$ par conséquent $U_{n+1}$ est indépendante de $X_0, \ldots, X_n$. On en déduit que pour tout $j \in E$,

$$\mathbf{E}\big[\mathbf{1}_{\{X_{n+1}=j\}} |\, X_0, \ldots, X_n\big] = \int_0^1 \mathbf{1}_{\{\Phi_Q(x,X_n)=j\}} dx = p(X_n, j) \ .$$

Cela implique que pour tout $n \in \mathbb{N}$ et toute fonction $f : E \to \mathbb{R}_+$,

$$\begin{aligned}
\mathbf{E}[f(X_{n+1}) |\, X_0, \ldots, X_n] &= \mathbf{E}\Big[\sum_{j\in E} f(j)\mathbf{1}_{\{X_{n+1}=j\}} \Big|\, X_0, \ldots, X_n\Big] \\
&= \sum_{j\in E} f(j)\mathbf{E}\big[\mathbf{1}_{\{X_{n+1}=j\}} |\, X_0, \ldots, X_n\big] \\
&= \sum_{j\in E} p(X_n, j) f(j) = (Q.f)(X_n),
\end{aligned}$$

ce qui montre le résultat voulu. ■

La proposition précédente suppose l'existence d'un espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$ sur lequel peut être définie une suite $U_n : \Omega \to [0,1[$, $n \in \mathbb{N}$, de variables $\mathscr{F}$-mesurables indépendantes de loi uniforme. Un tel espace de probabilité existe et peut se construire simplement à l'aide de la mesure de Lebesgue (voir l'exercice I.13 23). La proposition précédente entraîne donc le résultat d'existence suivant.

**Proposition I.1.26** *Soit $Q$, une matrice de transition et soit $\mu$, une mesure de probabilité sur $E$. Alors il existe un espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$ et une suite $X_n : \Omega \to E$, $n \in \mathbb{N}$, de variables $\mathscr{F}$-mesurables telle que sous $\mathbf{P}$, $(X_n)_{n\geq 0}$ soit une chaîne de Markov de matrice de transition $Q$ et de loi d'entrée $\mu$.*

## EXERCICES.

**Exercice I.13** Pour tout $x \in [0,1[$ et pour tout $n \geq 1$, on pose

$$\varepsilon_n(x) = \lfloor 2^{n+1}x \rfloor - 2\lfloor 2^n x \rfloor \ .$$

1. Montrer que $\varepsilon_n(x) \in \{0, 1\}$, pour tout $x \in [0, 1[$ et tout $n \geq 1$. Montrer que pour tout $x \in [0, 1[$, la suite $(\varepsilon_n(x))_{n\geq 1}$ n'est pas constante à 1 à partir d'un certain rang. Montrer de plus que
$$x = \sum_{n\geq 1} 2^{-n} \varepsilon_n(x) \ .$$
Réciproquement, on suppose qu'il existe une suite $\eta_n \in \{0, 1\}$, $n \geq 1$ qui ne stationne pas à la valeur 1 à partir d'un certain rang et telle que
$$x = \sum_{n\geq 1} 2^{-n} \eta_n \ .$$
Montrer que $\eta_n = \varepsilon_n(x)$, pour tout $n \geq 1$. La suite $(\varepsilon_n(x))_{n\geq 1}$ est le *développement dyadique* (ou *en base* 2) de $x \in [0, 1[$.
2. On pose $\Omega = [0, 1[$, $\mathscr{F} = \mathscr{B}([0, 1[)$, la tribu des Boréliens et $\mathbf{P} = \ell$, la mesure de Lebesgue sur $[0, 1[$. Montrer que sous $\mathbf{P}$, $(\varepsilon_n)_{n\geq 1}$ est une suite de variables de Bernoulli $\mathscr{F}$-mesurable, indépendantes de paramètre $1/2$ : $\mathbf{P}(\varepsilon_n = 0) = \mathbf{P}(\varepsilon_n = 1) = 1/2$.
3. Montrer qu'il existe une *partition* de $\mathbb{N}^*$ en une suite de sous-ensembles *infinis* notés $I_k$, $k \geq 0$. Pour chaque $k \geq 0$, on indexe de façon croissante les éléments de $I_k = \{n(k, 1) < n(k, 2) < \ldots\}$. On pose ensuite
$$\forall k \geq 0 \, , \quad U_k = \sum_{m\geq 1} 2^{-m} \varepsilon_{n(k,m)} \ .$$
Montrer que $U_k : \Omega \to [0, 1)$ est $\mathscr{F}$-mesurable et que la suite $(U_k)_{k\geq 0}$ est une suite de variables indépendantes et de loi uniforme sur $[0, 1[$. I.1.26. □

### I.1.d Trajectoire d'une chaîne de Markov.

Il est assez naturel de voir une chaîne de Markov comme une suite aléatoire, c'est-à-dire une trajectoire aléatoire, ce qui demande quelques définitions.

**Définition I.1.27** L'ensemble des suites à valeurs dans $E$ est noté
$$E^{\mathbb{N}} = \{\mathbf{x} = (x_n)_{n\in\mathbb{N}} : x_n \in E, n \in \mathbb{N}\} \ .$$

(a) Les sous-ensembles de $E^{\mathbb{N}}$ de la forme suivante :
$$C_{i_0,\ldots,i_n} = \{\mathbf{x} = (x_k)_{k\in\mathbb{N}} \in E^{\mathbb{N}} \ : \ x_0 = i_0 \ ; \ldots ; x_n = i_n\} \, , \ n \in \mathbb{N} \, , \ i_0, \ldots, i_n \in E, \qquad \text{(I.28)}$$
sont appelés des *cylindres élémentaires*. On note $\mathscr{C}$ l'ensemble des cylindres élémentaires auquel on adjoint $\emptyset$ et $E^{\mathbb{N}}$. C'est clairement un pi-système.

(b) On note $\mathscr{P}(E)^{\otimes\mathbb{N}} := \sigma(\mathscr{C})$, la tribu sur $E^{\mathbb{N}}$ engendrée par les cylindres élémentaires. On l'appelle la *tribu produit sur* $E^{\mathbb{N}}$ (un ensemble de la tribu produit est donc un ensemble de suites à valeurs dans $E$).

(c) Soit $(\Omega, \mathscr{F})$ un espace mesurable et $\mathbf{X} = (X_n)_{n\geq 0} : \Omega \to E^{\mathbb{N}}$. On dit que c'est une *suite aléatoire* $\mathscr{F}$*-mesurable* (ou un *processus* $\mathscr{F}$*-mesurable*) si $\mathbf{X}$ est $(\mathscr{F}, \mathscr{P}(E)^{\otimes\mathbb{N}})$-mesurable. □

**Lemme I.1.28** *Soit $(\Omega, \mathscr{F})$ un espace mesurable. Soit $X_n : \Omega \to E$, $n \in \mathbb{N}$, une suite de fonctions, sur lesquelles on ne fait pas d'hypothèse de mesurabilité a priori. On pose $\mathbf{X} = (X_n)_{n\geq 0} : \Omega \to E^{\mathbb{N}}$. Alors il y a équivalence entre les assertions suivantes.*

*(i) $\mathbf{X}$ est une suite aléatoire $\mathscr{F}$-mesurable (c-à-d que la fonction $\mathbf{X}$ est $(\mathscr{F}, \mathscr{P}(E)^{\otimes\mathbb{N}})$-mesurable).*

*(ii) Pour tout $n \in \mathbb{N}$, $X_n$ est $\mathscr{F}$-mesurable (c-à-d que pour tout $n \in \mathbb{N}$ et tout $i \in E$, $\{X_n = i\} \in \mathscr{F}$).*

**Preuve :** on suppose que les v.a. $X_n$ sont $\mathscr{F}$-mesurables. On observe que

$$\{\mathbf{X} \in C_{i_0,\dots,i_n}\} = \bigcap_{0 \le p \le n} \{X_p = i_p\} \in \mathscr{F} ,$$

pour tout cylindre élémentaire $C_{i_0,\dots,i_n} \in \mathscr{C}$. Comme $\sigma(\mathscr{C}) = \mathscr{P}(E)^{\otimes \mathbb{N}}$, cela montre que $\mathbf{X}$ est $(\mathscr{F}, \mathscr{P}(E)^{\otimes \mathbb{N}})$-mesurable.

Réciproquement, on suppose que $\mathbf{X}$ est $(\mathscr{F}, \mathscr{P}(E)^{\otimes \mathbb{N}})$-mesurable. On fixe $n \in \mathbb{N}$ et $j \in E$. On vérifie que $\{X_n = j\} = \bigcup_{i_0,\dots,i_{n-1} \in E} \{\mathbf{X} \in C_{i_0,\dots,i_{n-1},j}\} \in \mathscr{F}$, ce qui entraîne que $X_n$ est $\mathscr{F}$-mesurable, et ce pour tout $n \in \mathbb{N}$. ■

On fixe une probabilité $\mathbf{P}$ sur $(\Omega, \mathscr{F})$ et on se donne une suite aléatoire $\mathbf{X} = (X_n)_{n \ge 0}$ à valeurs dans $E$. On rappelle que sa loi sous $\mathbf{P}$, notée $\mu_{\mathbf{X}}$, est une mesure de probabilité sur l'espace mesurable $(E^{\mathbb{N}}, \mathscr{P}(E)^{\otimes \mathbb{N}})$ donnée par

$$\forall B \in \mathscr{P}(E)^{\otimes \mathbb{N}} \qquad \mu_{\mathbf{X}}(B) = \mathbf{P}(\mathbf{X} \in B) .$$

On a notamment

$$\forall n \in \mathbb{N} , \; \forall i_0, \dots, i_n \in E , \quad \mu_{\mathbf{X}}(C_{i_0,\dots,i_n}) = \mathbf{P}(X_0 = i_0; \dots ; X_n = i_n) . \tag{I.29}$$

Le théorème de transfert montre que pour toute fonction $F : E^{\mathbb{N}} \to \mathbb{R}$, $\mathscr{P}(E)^{\otimes \mathbb{N}}$-mesurable bornée ou positive, on a

$$\mathbf{E}\big[F(\mathbf{X})\big] = \int_{E^{\mathbb{N}}} F \, d\mu_{\mathbf{X}} .$$

Le lemme suivant montre que la loi d'une chaîne de Markov est entièrement caractérisée par sa matrice de transition et sa loi d'entrée.

**Lemme I.1.29** *Soient $(\Omega, \mathscr{F}, \mathbf{P})$ et $(\Omega', \mathscr{F}', \mathbf{P}')$, deux espaces de probabilité sur lesquels sont définies resp. $\mathbf{X} = (X_n)_{n \ge 0}$ et $\mathbf{X}' = (X'_n)_{n \ge 0}$, deux suites aléatoires à valeurs dans E. Alors, les assertions suivantes sont équivalentes.*

*(i)* $\mu_{\mathbf{X}} = \mu_{\mathbf{X}'}$.

*(ii)* $\mathbf{P}(\mathbf{X} \in B) = \mathbf{P}'(\mathbf{X}' \in B)$, $\forall B \in \mathscr{P}(E)^{\otimes \mathbb{N}}$.

*(iii)* $\mathbf{E}\big[F(\mathbf{X})\big] = \mathbf{E}'\big[F(\mathbf{X}')\big]$, $\forall F : E^{\mathbb{N}} \to \mathbb{R}$, $\mathscr{P}(E)^{\otimes \mathbb{N}}$*-mesurable bornée (ou positive).*

*(iv)* $\mathbf{P}(X_0 = i_0; \dots ; X_n = i_n) = \mathbf{P}'(X'_0 = i_0; \dots ; X'_n = i_n)$, $\forall n \in \mathbb{N}$, $\forall i_0, \dots, i_n \in E$.

*Par conséquent, si $\mathbf{X}$ sous $\mathbf{P}$ et $\mathbf{X}'$ sous $\mathbf{P}'$, sont deux chaînes de Markov de même loi d'entrée et de même matrice de transition, alors elles ont même loi.*

**Preuve :** par définition $(i) \iff (ii) \iff (iii)$. Trivialement, $(ii) \Longrightarrow (iv)$. Supposons $(iv)$ : (I.29) implique que $\mu_{\mathbf{X}}$ et $\mu_{\mathbf{X}'}$ coïncident sur le pi-système $\mathscr{C}$ des cylindres élementaires. Le théorème d'unicité du prolongement des mesures A.1.11, en appendice page 271, implique que $\mu_{\mathbf{X}}$ et $\mu_{\mathbf{X}'}$ coïncident sur $\sigma(\mathscr{C}) = \mathscr{P}(E)^{\otimes \mathbb{N}}$, c'est-à-dire $\mu_{\mathbf{X}} = \mu_{\mathbf{X}'}$. Cela montre $(iv) \Longrightarrow (i)$.

Supposons que $\mathbf{X}$ sous $\mathbf{P}$ et $\mathbf{X}'$ sous $\mathbf{P}'$, soient deux chaînes de Markov de même loi d'entrée et de même matrice de transition. Le lemme I.1.4 implique que $\mathbf{X}$ et $\mathbf{X}'$ satisfont $(iv)$, qui implique $(i)$, ce qu'il fallait démontrer. ■

**Le processus canonique.** On fixe $Q = (p(i,j))_{i,j \in E}$, une matrice de transition et $\mu \in \mathscr{M}_1(E)$. La proposition I.1.26 montre l'existence d'un espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$ et une suite aléatoire $\mathbf{X} = (X_n)_{n \ge 0}$, définie sur cet espace telle que, sous $\mathbf{P}$, $\mathbf{X}$ soit une chaîne de Markov de matrice de transition $Q$ et de loi d'entrée $\mu$. Le lemme I.1.29 précédent montre que la loi $\mu_{\mathbf{X}}$ de $\mathbf{X}$ sous $\mathbf{P}$, qui est une mesure de probabilité sur $(E^{\mathbb{N}}, \mathscr{P}(E)^{\otimes \mathbb{N}})$, est entièrement déterminée par $Q$ et $\mu$. On note $\mu_{\mathbf{X}}$ par $\mathbb{P}_\mu^Q$. Le lemme I.1.29 précédent et le lemme I.1.4 montrent que c'est *l'unique mesure de probabilité sur* $(E^{\mathbb{N}}, \mathscr{P}(E)^{\otimes})$ *telle que*

$$\mathbb{P}_\mu^Q(C_{i_0,\dots,i_n}) = \mu(i_0)p(i_0,i_1)\dots p(i_{n-1},i_n) \ , \quad n \in \mathbb{N} \, , \; i_0,\dots,i_n \in E \ . \tag{I.30}$$

On appelle $\mathbb{P}_\mu^Q$ *la loi canonique d'une chaîne de matrice de transition* $Q$ *et de loi d'entrée* $\mu$. On pose ensuite

$$\Omega^o = E^{\mathbb{N}} \quad \text{et} \quad \mathscr{F}^o = \mathscr{P}(E)^{\otimes \mathbb{N}} \ .$$

Pour tout $n \in \mathbb{N}$, et pour toute suite $\mathbf{x} = (x_p)_{p \in \mathbb{N}} \in E^{\mathbb{N}}$, on pose $X_n^o(\mathbf{x}) = x_n$. Cela définit bien une fonction $X_n^o : \Omega^o \to E$ : c'est simplement *la* $n$*-ième projection canonique*. Même si c'est un peu ridicule, on pose $\mathbf{X}^o = (X_n^o)_{n \in \mathbb{N}}$, qui est une fonction de $\Omega^o$ dans $E^{\mathbb{N}}$. On vérifie immédiatement que pour toute suite $\mathbf{x} \in \Omega^o = E^{\mathbb{N}}$, $\mathbf{X}^o(\mathbf{x}) = \mathbf{x}$. La fonction $\mathbf{X}^o$ est donc l'identité sur $\Omega^o = E^{\mathbb{N}}$. Elle est donc clairement $(\mathscr{F}^o, \mathscr{P}(E)^{\otimes \mathbb{N}})$-mesurable. Le lemme I.1.28 implique donc que les projections canoniques $X_n^o : \Omega^o \to E$ sont bien $\mathscr{F}^o$-mesurables. *On appelle* $\mathbf{X}^o = (X_n^o)_{n \geq 0}$ *le processus canonique.*

On vérifie alors que $C_{i_0,\dots,i_n} = \{X_0^o = i_0; \dots, X_n^o = i_n\}$. Donc (I.30) et le lemme I.1.4 impliquent que *sous* $\mathbb{P}_\mu^Q$*,* $\mathbf{X}^o$ *est une chaîne de Markov de matrice de transition* $Q$ *et de loi d'entrée* $\mu$. On introduit la terminologie suivante.

- $(\Omega^o, \mathscr{F}^o, \mathbf{X}^o, \mathbb{P}_\mu^Q)$ *est appelée la chaîne de Markov canonique de matrice de transition* $Q$ *et de loi d'entrée* $\mu$.

Cette réalisation particulière sert de référence : soit $(\Omega', \mathscr{F}', \mathbf{P}')$, un espace de probabilité et soit $\mathbf{X}' = (X_n')_{n \geq 0}$, une suite aléatoire définie sur cet espace telle que, sous $\mathbf{P}'$, $\mathbf{X}'$ soit une chaîne de Markov de matrice de transition $Q$ et de loi d'entrée $\mu$. Alors,

$$\forall B \in \mathscr{P}(E)^{\otimes \mathbb{N}} , \quad \mathbf{P}'(\mathbf{X}' \in B) = \mathbb{P}_\mu^Q(\mathbf{X}^o \in B) = \mathbb{P}_\mu^Q(B) \ .$$

La chaîne canonique n'interviendra que dans certains résultats d'existence. □

**Définition « globale » des chaînes.** Il arrivera très fréquemment que l'on change la loi d'entrée d'une chaîne de Markov. Pour ne pas affronter des problèmes de notation insurmontables, on ne veut pas changer de suite de variables aléatoires. On préfère fixer l'espace mesurable et la suite des variables qui y sont définies, et changer les probabilités sur l'espace mesurable sur lequel on travaille. C'est pour cela que l'on introduit la définition, définitive pour ce cours, des chaînes de Markov, qui portera temporairement le nom de « définition globale des chaînes de Markov ». Le terme « global » n'est pas standard et nous l'abandonnerons rapidement.

**Définition I.1.30** Une chaîne de Markov *globale*, de matrice de transition $Q$, est la donnée des objets mathématiques suivants :

$$\big(\Omega \, ; \, \mathscr{F} \, ; \, (\mathscr{F}_n)_{n \geq 0} \, ; \, \mathbf{X} = (X_n)_{n \geq 0} \, ; \, Q \, ; \, \mathbf{P}_\mu \, , \, \mu \in \mathscr{M}_1(E) \, \big).$$

où on précise que :

- $(\Omega, \mathscr{F})$ est un espace mesurable, $(\mathscr{F}_n)_{n \geq 0}$ est une filtration sur $(\Omega, \mathscr{F})$, $\mathbf{X} = (X_n)_{n \geq 0}$ est une suite de variables $(\mathscr{F}_n)_{n \geq 0}$-adaptée à valeurs dans $E$.
- Pour toute mesure de probabilité $\mu \in \mathscr{M}_1(E)$, $\mathbf{P}_\mu$ est une mesure de probabilité sur $(\Omega, \mathscr{F})$ telle que, sous $\mathbf{P}_\mu$, $\mathbf{X}$ est une chaîne de Markov relativement à $(\mathscr{F}_n)_{n \geq 0}$, de loi d'entrée $\mu$ et de matrice de transition $Q$, selon la définition I.1.2. □

Pour tout $n \in \mathbb{N}$, on note $\mathscr{F}_n^o = \sigma(X_0^o, \ldots, X_n^o)$ : $(\mathscr{F}_n^o)_{n\geq 0}$ est la filtration canonique. On appelle

$$\big(\Omega^o\,;\, \mathscr{F}^o\,;\, (\mathscr{F}_n^o)_{n\geq 0}\,;\, \mathbf{X}^o = (X_n^o)_{n\geq 0}\,;\, Q\,;\, \mathbb{P}_\mu^Q\,,\, \mu \in \mathscr{M}_1(E)\,\big)$$

la chaîne de Markov (globale) canonique de matrice de transition $Q$. Cela montre immédiatement le théorème d'existence suivant.

**Théorème I.1.31** *Pour toute matrice de transition Q, il existe une chaîne de Markov globale ayant Q pour matrice de transition.*

### I.1.e Propriétés de Markov.

Dans la suite $\big(\Omega; \mathscr{F}; (\mathscr{F}_n)_{n\geq 0}; \mathbf{X} = (X_n)_{n\geq 0}; Q = (p(i,j))_{i,j\in E}; \mathbf{P}_\mu, \mu \in \mathscr{M}_1(E)\big)$ désigne une chaîne de Markov (globale). La propriété de Markov s'énonce informellement de la manière suivante : soit $n_0 \in \mathbb{N}$; la suite $(X_{n_0+n})_{n\geq 0}$, qui représente la trajectoire de la chaîne de Markov après l'instant $n_0$, ne dépend du passé $\mathscr{F}_{n_0}$ qu'à travers sa position (ou son état) $X_{n_0}$ à l'instant $n_0$ et elle se comporte comme une chaîne de Markov issue de l'état $X_{n_0}$ et de matrice de transition $Q$. Pour pouvoir énoncer formellement cette propriété, on introduit l'opérateur de décalage sur les suites à valeurs dans $E$.

**Définition I.1.32** On fixe $n_0 \in \mathbb{N}$. On définit l'application $\theta_{n_0} : E^{\mathbb{N}} \to E^{\mathbb{N}}$ par $\theta_{n_0}\mathbf{x} = (x_{n_0+n})_{n\in\mathbb{N}}$, pour toute suite $\mathbf{x} = (x_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$. L'application $\theta_{n_0}$ est appelée l'*opérateur de décalage au temps* $n_0$. Il est facile de vérifier que $\theta_{n_0} \circ \theta_{n_1} = \theta_{n_0+n_1}$, pour tous $n_0, n_1 \in \mathbb{N}$. □

**Lemme I.1.33** $\theta_{n_0} : E^{\mathbb{N}} \to E^{\mathbb{N}}$ *est* $(\mathscr{P}(E)^{\otimes\mathbb{N}}, \mathscr{P}(E)^{\otimes\mathbb{N}})$*-mesurable.*

**Preuve :** on remarque que

$$\begin{aligned} \theta_{n_0}^{-1}(C_{i_0,\ldots,i_n}) &= \big\{\mathbf{x} = (x_k)_{k\in\mathbb{N}} \in E^{\mathbb{N}} \;:\; x_{n_0} = i_0\,;\,\ldots\,;\, x_{n_0+n} = i_n\big\} \\ &= \bigcup_{j_0,\ldots,j_{n_0-1}\in E} C_{j_0,\ldots,j_{n_0-1},i_0,\ldots,i_n} \quad \in \mathscr{P}(E)^{\otimes\mathbb{N}}\,. \end{aligned}$$

Comme $\sigma(\mathscr{C}) = \mathscr{P}(E)^{\otimes\mathbb{N}}$, cela permet de conclure. ■

**Conventions et notations importantes.** Avant d'énoncer la propriété de Markov simple, un peu de "gymnastique notationnelle" s'avère utile.

• On note $\mathbf{E}_\mu$ l'espérance associée à $\mathbf{P}_\mu$. Par définition la loi de $\mathbf{X}$ sous $\mathbf{P}_\mu$ est $\mathbb{P}_\mu^Q$, la loi "canonique". On note de même $\mathbb{E}_\mu^Q$ l'espérance associée à $\mathbb{P}_\mu^Q$. Soit $F : E^{\mathbb{N}} \to \mathbb{R}$, une fonction $\mathscr{P}(E)^{\otimes\mathbb{N}}$-mesurable bornée ou positive. On notera l'intégrale $\int_{E^{\mathbb{N}}} F\, d\mathbb{P}_\mu^Q$ plutôt par $\mathbb{E}_\mu^Q[F]$ que par $\mathbb{E}_\mu^Q[F(\mathbf{X}^o)]$. Avec ces conventions, le théorème de transfert s'exprime donc $\mathbf{E}_\mu[F(\mathbf{X})] = \mathbb{E}_\mu^Q[F]$. On a donc cinq manières d'écrire la même quantité avec une préférence de notation pour le premier et le dernier membre de ces égalités :

$$\mathbf{E}_\mu[F(\mathbf{X})] = \int_\Omega F(\mathbf{X}(\omega))d\mathbf{P}(\omega) = \int_{E^{\mathbb{N}}} F\, d\mathbb{P}_\mu^Q = \mathbb{E}_\mu^Q[F(\mathbf{X}^o)] = \mathbb{E}_\mu^Q[F]\,.$$

• On fixe $i \in E$. On note $\delta_i$ *la masse de Dirac en* $i$, qui est une mesure de probabilité sur $E$ : $\delta_i \in \mathscr{M}_1(E)$. On rappelle que $\delta_i = (\delta_i(j))_{j\in E}$ est donnée par $\delta_i(j) = 0$ si $i \neq j$ et $\delta_i(i) = 1$. Pour simplifier les notations, on pose

$$\forall i \in E\,, \quad \mathbf{P}_i := \mathbf{P}_{\delta_i}\,.$$

On a donc $\mathbf{P}_i(X_0 = i) = 1$, c'est-à-dire que $\mathbf{P}_i$-p.s. $X_0 = i$. On note $\mathbf{E}_i$ l'espérance associée à $\mathbf{P}_i$. De même on préfère noter $\mathbb{P}^Q_{\delta_i}$ par $\mathbb{P}^Q_i$ et $\mathbb{E}^Q_{\delta_i}$ par $\mathbb{E}^Q_i$.

• Soient $Z : \Omega \to E$, une v.a. $\mathscr{F}$-mesurable et $F : E^{\mathbb{N}} \to \mathbb{R}$, une fonction $\mathscr{P}(E)^{\otimes\mathbb{N}}$-mesurable bornée ou positive. On définit la fonction $Y : \Omega \to \mathbb{R}$ en posant $Y(\omega) = \mathbf{E}_{Z(\omega)}[F(\mathbf{X})]$, pour tout $\omega \in \Omega$. On note $Y$ simplement par $\mathbf{E}_Z[F(\mathbf{X})]$. On a donc $Y(\omega) = \mathbf{E}_i[F(\mathbf{X})]$ si $Z(\omega) = i$, ou encore

$$\mathbf{E}_Z[F(\mathbf{X})] = \sum_{i\in E} \mathbf{1}_{\{Z=i\}}\mathbf{E}_i[F(\mathbf{X})] = \sum_{i\in E} \mathbf{1}_{\{Z=i\}}\, \mathbb{E}^Q_i[F] \ .$$

Si $\mathscr{G}$ une sous-tribu de $\mathscr{F}$ et si $Z$ est $\mathscr{G}$-mesurable, alors $\mathbf{E}_Z[F(\mathbf{X})]$ est $\mathscr{G}$-mesurable également.

**Théorème I.1.34** **(Propriété de Markov simple)** *Pour tout $n_0 \in \mathbb{N}$, pour toute loi d'entrée $\mu \in \mathscr{M}_1(E)$ et pour toute application $F : E^{\mathbb{N}} \to \mathbb{R}$ qui est $\mathscr{P}(E)^{\otimes\mathbb{N}}$-mesurable bornée, on a*

$$\mathbf{P}_\mu\text{-}p.s. \quad \mathbf{E}_\mu\left[F(\theta_{n_0}\mathbf{X}) \,|\, \mathscr{F}_{n_0}\right] = \mathbf{E}_{X_{n_0}}\left[F(\mathbf{X})\right]. \tag{I.31}$$

**Preuve :** on fixe $\mu \in \mathscr{M}_1(E)$. Pour tout $m \in \mathbb{N}$ et tout $j \in E$, on a

$$\mathbf{P}_\mu\text{-p.s.} \quad \mathbf{E}_\mu[\mathbf{1}_{\{X_{m+1}=j\}}|\mathscr{F}_m] = p(X_m, j) \ . \tag{I.32}$$

Ceci résulte de la définition I.1.2 avec $\mathbf{P} = \mathbf{P}_\mu$, $n = m$ et $f = \mathbf{1}_{\{j\}}$, et donc $(Q.f)(i) = p(i,j)$, pour tout $i \in E$.

On fixe ensuite $i_0, \dots, i_n \in E$ et on montre d'abord la propriété de Markov (I.31) pour $F = \mathbf{1}_{C_{i_0,\dots,i_n}}$. On remarque que $F(\theta_{n_0}\mathbf{X}) = \mathbf{1}_{\{X_{n_0}=i_0;X_{n_0+1}=i_1;\dots;X_{n_0+n-1}=i_{n-1}\}}\mathbf{1}_{\{X_{n_0+n}=i_n\}}$. Comme $\mathbf{X}$ est $(\mathscr{F}_n)_{n\geq 0}$-adaptée, on a $\{X_{n_0} = i_0; X_{n_0+1} = i_1; \dots; X_{n_0+n-1} = i_{n-1}\} \in \mathscr{F}_{n_0+n-1}$ et en appliquant (I.32), avec $m = n_0+n-1$ et $j = i_n$, on voit que

$$\begin{aligned}
\mathbf{E}_\mu\big[F(\theta_{n_0}\mathbf{X})|\mathscr{F}_{n_0+n-1}\big] &= \mathbf{1}_{\{X_{n_0}=i_0;X_{n_0+1}=i_1;\dots;X_{n_0+n-1}=i_{n-1}\}}\mathbf{E}\big[\mathbf{1}_{\{X_{n_0+n}=i_n\}} \,|\, \mathscr{F}_{n_0+n-1}\big] \\
&= \mathbf{1}_{\{X_{n_0}=i_0;X_{n_0+1}=i_1;\dots;X_{n_0+n-1}=i_{n-1}\}}p(X_{n_0+n-1}, i_n) \\
&= \mathbf{1}_{\{X_{n_0}=i_0;X_{n_0+1}=i_1;\dots;X_{n_0+n-1}=i_{n-1}\}}p(i_{n-1}, i_n).
\end{aligned}$$

En répétant l'argument, on voit que

$$\mathbf{E}_\mu\big[F(\theta_{n_0}\mathbf{X})|\mathscr{F}_{n_0+n-2}\big] = \mathbf{1}_{\{X_{n_0}=i_0;X_{n_0+1}=i_1;\dots;X_{n_0+n-2}=i_{n-2}\}}p(i_{n-2}, i_{n-1})p(i_{n-1}, i_n),$$

et en recommençant, on finit par obtenir

$$\mathbf{E}_\mu\big[F(\theta_{n_0}\mathbf{X})|\mathscr{F}_{n_0}\big] = \mathbf{1}_{\{X_{n_0}=i_0\}}p(i_0, i_1)\dots p(i_{n-1}, i_n) \ . \tag{I.33}$$

On observe ensuite que pour tout $i \in E$

$$\mathbf{E}_i[F(\mathbf{X})] = \mathbf{P}_i(\mathbf{X} \in C_{i_0,\dots,i_n}) = \mathbf{P}_i(X_0 = i_0; \dots, X_n = i_n) = \delta_i(i_0)p(i_0, i_1)\dots p(i_{n-1}, i_n)$$

et par (I.33) $\mathbf{E}_\mu\big[F(\theta_{n_0}\mathbf{X})|\mathscr{F}_{n_0}\big] = \mathbf{P}_{X_{n_0}}(\mathbf{X} \in C_{i_0,\dots,i_n})$, $\mathbf{P}_\mu$-p.s., ce qui montre la propriété de Markov (I.31) pour $F = \mathbf{1}_C$, où $C$ est un cylindre élémentaire quelconque.

On note ensuite $\mathscr{H}$ l'ensemble de toutes les fonctions $F : E^{\mathbb{N}} \to \mathbb{R}$ qui sont $\mathscr{P}(E)^{\otimes\mathbb{N}}$-mesurables bornées et qui satisfont (I.31). Par linéarité de l'espérance et de l'espérance conditionnelle, on voit que $\mathscr{H}$ est un $\mathbb{R}$-espace vectoriel. On vient de montrer que $\mathbf{1}_C \in \mathscr{H}$, pour tout $C \in \mathscr{C}$. Supposons que l'on ait $F_p \in \mathscr{H}$, $p \in \mathbb{N}$, une suite de fonctions telles qu'il existe une constante $c > 0$, satisfaisant $0 \leq F_p(\mathbf{x}) \leq F_{p+1}(\mathbf{x}) \leq c$, pour tout

$\mathbf{x} \in E^{\mathbb{N}}$ et tout $p \in \mathbb{N}$. On pose $F = \sup_p F_p$. C'est clairement une fonction $\mathscr{P}(E)^{\otimes \mathbb{N}}$-mesurable positive et majorée par $c$, donc bornée. Par convergence monotone conditionnelle, on a $\mathbf{P}_\mu$-p.s.

$$\mathbf{E}_\mu \left[ F(\theta_{n_0} \mathbf{X}) \,|\, \mathscr{F}_{n_0} \right] = \lim_{p \to \infty} \mathbf{E}_\mu \left[ F_p(\theta_{n_0} \mathbf{X}) \,|\, \mathscr{F}_{n_0} \right] = \lim_{p \to \infty} \mathbf{E}_{X_{n_0}} \left[ F_p(\mathbf{X}) \right] = \mathbf{E}_{X_{n_0}} \left[ F(\mathbf{X}) \right] \ ,$$

ce qui montre que $F \in \mathscr{H}$. La version fonctionnelle du théorème A.1.20 de la classe monotone (prouvé en appendice page 273) implique donc que $\mathscr{H}$ contient toutes les fonctions $\mathscr{P}(E)^{\otimes \mathbb{N}}$-mesurables bornées, ce qui termine la preuve. ∎

En appliquant la propriété de Markov (I.31) au temps $n_0 = 0$, et en intégrant l'expression obtenue, on obtient le corollaire suivant.

**Corollaire I.1.35** *Pour toute fonction $F : E^{\mathbb{N}} \to \mathbb{R}$ qui est $\mathscr{P}(E)^{\otimes \mathbb{N}}$-mesurable bornée, et pour toute loi d'entrée $\mu \in \mathscr{M}_1(E)$, on a $\mathbf{E}_\mu[F(\mathbf{X})] = \sum_{i \in E} \mu(i) \mathbf{E}_i[F(\mathbf{X})]$.*

**Propriété de Markov forte.** Nous généralisons la propriété de Markov à des temps d'arrêt. Pour cela, on a besoin de préciser les notations suivantes. On fixe $i^* \in E$, qui ne joue aucun rôle spécifique. Soit $T : \Omega \to \mathbb{N} \cup \{\infty\}$, un temps aléatoire. On pose alors

$$X_T = \begin{cases} X_n & \text{si } T = n \\ i^* & \text{si } T = \infty \end{cases} \qquad \text{et} \qquad \theta_T \mathbf{X} = \begin{cases} \theta_n \mathbf{X} = (X_{m+n})_{m \geq 0} & \text{si } T = n \\ (i^*, i^*, i^*, i^*, \ldots) & \text{si } T = \infty. \end{cases}$$

**Théorème I.1.36** **(Propriété de Markov forte)** *Soit $T$, un $(\mathscr{F}_n)_{n \geq 0}$-temps d'arrêt. Alors les assertions suivantes sont vérifiées.*

*(i)* *$X_T : \Omega \to E$ est $\mathscr{F}_T$-mesurable.*

*(ii)* *$\theta_T \mathbf{X} : \Omega \to E^{\mathbb{N}}$ est $(\mathscr{F}, \mathscr{P}(E)^{\otimes \mathbb{N}})$-mesurable.*

*(iii)* *Pour toute $\mu \in \mathscr{M}_1(E)$, et toute $F : E^{\mathbb{N}} \to \mathbb{R}$, $\mathscr{P}(E)^{\otimes \mathbb{N}}$-mesurable bornée, on a*

$$\mathbf{P}_\mu\text{-}p.s. \quad \mathbf{E}_\mu \left[ \mathbf{1}_{\{T < \infty\}} F(\theta_T \mathbf{X}) \,|\, \mathscr{F}_T \right] = \mathbf{1}_{\{T < \infty\}} \mathbf{E}_{X_T} \left[ F(\mathbf{X}) \right]. \tag{I.34}$$

**Preuve :** on fixe $i \in E$ et on observe que pour tout $n \in \mathbb{N}$, on a $\{X_T = i\} \cap \{T = n\} = \{X_n = i\} \cap \{T = n\} \in \mathscr{F}_n$, car $X_n$ est $\mathscr{F}_n$-mesurable. Cela montre que $\{X_T = i\} \in \mathscr{F}_T$, pour tout $i \in E$, donc que $X_T$ est $\mathscr{F}_T$-mesurable, ce qui est le point $(i)$.

On note $[i^*]$, la suite constante à $i^*$. On fixe $C \in \mathscr{C}$ et on pose $A^* = \emptyset$ si $[i^*] \notin C$ et $A^* = \Omega$ si $[i^*] \in C$. On remarque que

$$\{\theta_T \mathbf{X} \in C\} = \big( A^* \cap \{T = \infty\} \big) \cup \bigcup_{k \in \mathbb{N}} \{\theta_k \mathbf{X} \in C\} \cap \{T = k\} \ .$$

Comme les $\theta_k$ sont mesurables, $\{\theta_k \mathbf{X} \in C\} \cap \{T = k\} \in \mathscr{F}$, pour tout $k \in \mathbb{N}$. Cela implique que $\{\theta_T \mathbf{X} \in C\} \in \mathscr{F}$ pour tout $C \in \mathscr{C}$. Comme $\sigma(\mathscr{C}) = \mathscr{P}(E)^{\otimes \mathbb{N}}$, cela montre que $\theta_T \mathbf{X}$ est $(\mathscr{F}, \mathscr{P}(E)^{\otimes \mathbb{N}})$-mesurable, ce qui est exactement le point $(ii)$.

Montrons la propriété forte de Markov qui est le point $(iii)$ : on fixe $\mu$ et $F$ comme dans l'énoncé. On fixe $A \in \mathscr{F}_T$. Comme $F$ est bornée, on en déduit l'interversion série / espérance suivante :

$$\mathbf{E}_\mu \left[ \mathbf{1}_{\{T < \infty\} \cap A} F(\theta_T \mathbf{X}) \right] = \sum_{n \in \mathbb{N}} \mathbf{E}_\mu \left[ \mathbf{1}_{\{T = n\} \cap A} F(\theta_n \mathbf{X}) \right] \ .$$

Par définition de $\mathscr{F}_T$, $\{T = n\} \cap A \in \mathscr{F}_n$, pour tout $n \in \mathbb{N}$. La propriété de Markov simple au temps $n$ implique alors que $\mathbf{E}_\mu\big[\mathbf{1}_{\{T=n\}\cap A} F(\theta_n \mathbf{X})\big] = \mathbf{E}_\mu\big[\mathbf{1}_{\{T=n\}\cap A} \mathbf{E}_{X_n}\big[F(\mathbf{X})\big]\,\big]$. On a donc

$$\mathbf{E}_\mu\left[\mathbf{1}_{\{T<\infty\}\cap A} F(\theta_T \mathbf{X})\right] = \sum_{n\in\mathbb{N}} \mathbf{E}_\mu\big[\mathbf{1}_{\{T=n\}\cap A} \mathbf{E}_{X_n}\big[F(\mathbf{X})\big]\,\big] = \mathbf{E}_\mu\left[\mathbf{1}_{\{T<\infty\}\cap A} \mathbf{E}_{X_T}\big[F(\mathbf{X})\big]\,\right].$$

Comme cette égalité est vraie pour tout $A \in \mathscr{F}_T$, et comme $\mathbf{E}_{X_T}\big[F(\mathbf{X})\big]$ est $\mathscr{F}_T$-mesurable, cela entraîne (I.34). ■

**Exemple I.1.37** On se place dans le cas où la chaîne de Markov est une marche aléatoire sur $\mathbb{Z}^d$. Il est facile de voir que la propriété de Markov forte implique le résultat suivant : soit $T$, un $(\mathscr{F}_n)_{n\geq 0}$-temps d'arrêt et soit $\mu \in \mathscr{M}_1(\mathbb{Z}^d)$. On suppose que $\mathbf{P}_\mu(T < \infty) = 1$. Alors, sous $\mathbf{P}_\mu$, la suite de v.a. $(X_{n+T} - X_T)_{n\geq 0}$ est indépendante de la suite $(X_{n\wedge T})_{n\geq 0}$ et elle est distribuée comme $\mathbf{X}$ sous $\mathbf{P}_0$ (les détails de cette vérification sont laissés au lecteur). □

## EXERCICES.

**Exercice I.14** Soit $\big(\Omega\,;\, \mathscr{F}\,;\, (\mathscr{F}_n)_{n\geq 0}\,;\, \mathbf{X} = (X_n)_{n\geq 0}\,;\, Q = (p(i,j))_{i,j\in E}\,;\, \mathbf{P}_\mu\,,\, \mu \in \mathscr{M}_1(E)\,\big)$, une chaîne de Markov. On fixe $F \subset E$, un sous-ensemble non-vide de $E$. On pose

$$T_F = \inf\{n \in \mathbb{N} \,:\, X_n \in F\}$$

avec la convention que $T_F = \infty$ ssi pour tout $n \in \mathbb{N}$, on a $X_n \notin F$. $T_F$ est le temps d'atteinte de $F$.

1. Montrer que $T_F$ est un $(\mathscr{F}_n)_{n\geq 0}$-temps d'arrêt.
2. Pour tout $n \geq 0$, on pose $Y_n = X_{n\wedge T_F}$. Montrer que

$$\big(\Omega\,;\, \mathscr{F}\,;\, (\mathscr{F}_n)_{n\geq 0}\,;\, \mathbf{Y} = (Y_n)_{n\geq 0}\,;\, Q_* = (p_*(i,j))_{i,j\in E}\,;\, \mathbf{P}_\mu\,,\, \mu \in \mathscr{M}_1(E)\,\big)$$

est une chaîne de Markov dont on précisera la matrice de transition $Q_*$. Cette chaîne s'appelle *la chaîne absorbée en $F$* associée à $\mathbf{X}$. □

**Exercice I.15** Soit $\big(\Omega\,;\, \mathscr{F}\,;\, (\mathscr{F}_n)_{n\geq 0}\,;\, \mathbf{X} = (X_n)_{n\geq 0}\,;\, Q = (p(i,j))_{i,j\in E}\,;\, \mathbf{P}_\mu\,,\, \mu \in \mathscr{M}_1(E)\,\big)$, une chaîne de Markov. On fixe $F \subset E$, un sous-ensemble non-vide de $E$. On pose

$$T_F = \inf\{n \in \mathbb{N} \,:\, X_n \in F\}$$

avec la convention que $T_F = \infty$ ssi pour tout $n \in \mathbb{N}$, on a $X_n \notin F$. $T_F$ est le temps d'atteinte de $F$, qui est un $(\mathscr{F}_n)_{n\geq 0}$-temps d'arrêt. On se donne un point cimetière $\partial \notin E$ et on définit pour tout $n \in \mathbb{N}$

$$Y_n = \begin{cases} X_n & \text{si } n < T_F, \\ \partial & \text{si } n \geq T_F. \end{cases}$$

1. On pose $E^* = (E\backslash F) \cup \{\partial\}$. Montrer que

$$\big(\Omega\,;\, \mathscr{F}\,;\, (\mathscr{F}_n)_{n\geq 0}\,;\, \mathbf{Y} = (Y_n)_{n\geq 0}\,;\, Q_* = (p_*(i,j))_{i,j\in E^*}\,;\, \mathbf{P}_\mu\,,\, \mu \in \mathscr{M}_1(E^*)\,\big)$$

est une chaîne de Markov dont on calculera la matrice de transition $Q_*$. Cette chaîne s'appelle *la chaîne tuée en $F$* associée à $\mathbf{X}$.
2. Quel est le lien avec la chaîne absorbée en $F$ ?
3. On pose $\zeta = \inf\{n \in \mathbb{N} :\, Y_n = \partial\}$, avec les convention habituelles. Montrer que $T_F = \zeta$. □

**Exercice I.16** Soit $\big(\Omega\,;\,\mathscr{F}\,;\,(\mathscr{F}_n)_{n\geq 0}\,;\,\mathbf{X}=(X_n)_{n\geq 0}\,;\,Q=(p(i,j))_{i,j\in E}\,;\,\mathbf{P}_\mu\,,\,\mu\in\mathscr{M}_1(E)\,\big)$, une chaîne de Markov. On fixe $F\subset E$, un sous-ensemble non-vide de $E$. Pour tout $k\in\mathbb{N}$, on définit

$$T_F^{(0)}=\inf\{n\geq 0 : X_n\in F\} \quad \text{et} \quad \forall k\geq 0 \quad T_F^{(k+1)}=\inf\{n>T_F^{(k)}\ :\ X_n\in F\}\ ,$$

avec la convention habituelle $\inf\emptyset=\infty$.

1. Montrer que les temps $T_F^{(k)}$, $k\geq 0$ sont des $(\mathscr{F}_n)_{n\geq 0}$-temps d'arrêt. Ce sont les temps de retour successifs de la chaîne en $F$.
2. On se donne un point cimetière $\partial\notin E$ et on définit pour tout $k\in\mathbb{N}$

$$Y_k=\begin{cases} X_{T_F^{(k)}} & \text{si } T_F^{(k)}<\infty,\\ \partial & \text{si } T_F^{(k)}=\infty.\end{cases}$$

On pose $F^*=F\cup\{\partial\}$. Montrer que

$$\big(\Omega\,;\,\mathscr{F}\,;\,(\mathscr{F}_{T_F^{(k)}})_{k\geq 0}\,;\,\mathbf{Y}=(Y_k)_{k\geq 0}\,;\,Q_*=(p_*(i,j))_{i,j\in F^*}\,;\,\mathbf{P}_\mu\,,\,\mu\in\mathscr{M}_1(F^*)\,\big)$$

est une chaîne de Markov dont on calculera la matrice de transition $Q_*$. □

**Exercice I.17** Soit $\mathbf{G}=(\mathbf{S},\mathbf{A})$ un graphe simple, *fini*. On rappelle qu'une suite finie $\gamma=(s_0,s_1,\ldots,s_{n+1})$ est un chemin dans le graphe ssi $\{s_k,s_{k+1}\}\in\mathbf{A}$, pour tout $0\leq k\leq n$. On dit que le chemin $\gamma$ *relie* le sommet $s_0$ au sommet $s_n$. On dit que le graphe $\mathbf{G}$ est *connexe* ssi toute paire de sommets distincts sont reliés par un chemin. Le chemin $\gamma$ est un *cycle* ssi (1) :$s_0=s_{n+1}$, (2) : $n\geq 2$ *et* $(3)$ : tous les sommets $s_0,s_1,\ldots,s_n$ sont distincts. Un *arbre est un graphe connexe et sans cycle*.

On suppose que $\mathbf{T}=(\mathbf{S},\mathbf{A})$ est un arbre. On se donne un sous-arbre $\mathbf{T}'=(\mathbf{S}',\mathbf{A}')$, c'est-à-dire un arbre tel que $\mathbf{S}'\subset\mathbf{S}$ et $\mathbf{A}'\subset\mathbf{A}$. Soit

$$\big(\Omega\,;\,\mathscr{F}\,;\,(\mathscr{F}_n)_{n\geq 0}\,;\,\mathbf{X}=(X_n)_{n\geq 0}\,;\,Q=(p(i,j))_{i,j\in\mathbf{S}}\,;\,\mathbf{P}_\mu\,,\,\mu\in\mathscr{M}_1(\mathbf{S})\,\big)\ ,$$

une marche aléatoire simple sur l'arbre $\mathbf{T}=(\mathbf{S},\mathbf{A})$. On pose

$$T_{\mathbf{S}'}^{(0)}=\inf\{n\geq 0 : X_n\in\mathbf{S}'\} \quad \text{et} \quad \forall k\geq 0 \quad T_{\mathbf{S}'}^{(k+1)}=\inf\{n>T_{\mathbf{S}'}^{(k)}\ :\ X_n\in\mathbf{S}'\}\ ,$$

avec la convention habituelle $\inf\emptyset=\infty$. On suppose que

$$\forall\mu\in\mathscr{M}_1(\mathbf{S})\,,\quad \mathbf{P}_\mu(T_{\mathbf{S}'}^{(0)}<\infty)=1\ .$$

1. Montrer que pour toute loi d'entrée $\mu\in\mathscr{M}_1(\mathbf{S})$ et pour tout $k\geq 0$, $\mathbf{P}_\mu(T_{\mathbf{S}'}^{(k)}<\infty)=1$ et donc que

$$\mathbf{P}_\mu(\forall k\geq 0\,,\quad T_{\mathbf{S}'}^{(k)}<\infty)=1.$$

Que signifie ce résultat ?

2. Montrer que la question précédente implique que pour toute loi d'entrée $\mu\in\mathscr{M}_1(\mathbf{S})$, il a un sens à définir $\mathbf{P}_\mu$-presque sûrement la suite $(Y_k)_{k\geq 0}$ par

$$Y_k=X_{T_{\mathbf{S}'}^{(k)}}\ ,\quad k\geq 1\ .$$

Montrer que

$$\big(\Omega\,;\,\mathscr{F}\,;\,(\mathscr{F}_{T_{\mathbf{S}'}^{(k)}})_{k\geq 0}\,;\,\mathbf{Y}=(Y_k)_{k\geq 0}\,;\,Q_*=(p_*(i,j))_{i,j\in\mathbf{S}'}\,;\,\mathbf{P}_\mu\,,\,\mu\in\mathscr{M}_1(\mathbf{S}')\,\big)$$

est une marche aléatoire simple sur le sous-arbre $T'$. Où la propriété d'arbre est-elle utilisée de façon essentielle ? Ce résultat peut-il rester vrai pour des marches aléatoires simples des graphes qui ne sont pas sur des arbres ? □

### I.1.f Exemple d'application : le problème de Dirichlet sur les graphes.

**Problème de Dirichlet dans $\mathbb{R}^d$.** Soit $D\subset\mathbb{R}^d$ un *domaine*, à savoir un ensemble ouvert connexe. Nous désignons par $\overline{D}$ son adhérence et par $\partial D=\overline{D}\backslash D$ sa frontière qui est un ensemble fermé. Soit $f:\partial D\to\mathbb{R}_+$ une fonction bornée et continue.

Imaginons que $\overline{D}$ soit un objet métallique : nous imposons la température à la frontière de $D$, la valeur de la température à $x\in\partial D$ étant notée $f(x)$. On attend qu'un régime stationnaire s'établisse.

*Quelle est la température à l'intérieur de D ?*

Si on note par $u(x)$ la température en $x \in \overline{D}$. Alors $u = f$ sur $\partial D$. De plus, la physique implique que $u$ doit être continue sur $\overline{D}$ (et bornée). Enfin la température dans l'objet métallique est le résultat d'un flux de chaleur en régime stationnaire et $u$ est la moyenne de ses propres valeurs dans un voisinage.

Plus précisément, on désigne par $B(x,r)$ la boule ouverte de centre $x \in D$ et de rayon $r \in ]0,\infty[$. On note $\ell_d$ la mesure de Lebesgue. Alors, $u$ satisfait à la *propriété de la valeur moyenne dans $D$* si elle satisfait les conditions suivantes

$$\forall x \in D, \forall r \in ]0,\infty[ \text{ tel que } B(x,r) \subset \overline{D}, \quad u(x) = \frac{1}{\ell_d(B(0,1))} \int_{B(0,1)} u\big(x + \tfrac{r}{\|y\|} y\big)\, \ell_d(dy). \tag{I.35}$$

Le membre de droite correspond à intégrer uniformément sur la sphère $\partial B(x,r) = \{y \in \mathbb{R}^d : \|y - x\| = r\}$. Nous disons, de façon synonyme, que $u$ est *harmonique dans $D$* (Fourier a développé l'analyse harmonique pour résoudre ce problème). Mentionnons que la propriété de la valeur moyenne peut être reformulée (localement) de la manière suivante.

- *Soit $v : D \to \mathbb{R}$, une fonction mesurable et localement bornée. Alors les assertions suivantes sont équivalentes.*
  (a) *$v$ satisfait la propriété de la valeur moyenne dans $D$.*
  (b) *$v$ est $C^2$ dans $D$ et*

$$\forall x \in D\ , \quad \Delta v(x) := \sum_{k=0}^{d} \frac{\partial^2 v}{\partial x_k^2}(x) = 0\ . \tag{I.36}$$

Ici l'opérateur différentiel $\Delta$ est appelé le Laplacien. Le problème que nous avons décrit est appelé *le $(D,f)$-problème de Dirichlet* et il peut être reformulé de la manière suivante.

- Soit $D \subset \mathbb{R}^d$ un domaine et soit $f : \partial D \to \mathbb{R}$ une fonction continue. Alors $u : \overline{D} \to \mathbb{R}$ est une solution du problème de $(D,f)$-*Dirichlet* si

$$u \in C^2(D) \cap C^0(\overline{D}), \quad u = f \text{ sur } \partial D \quad \text{et} \quad \Delta u = 0 \text{ sur } D. \tag{I.37}$$

**Sphères de Poincaré.** H. Poincaré a conçu une solution probabiliste au problème de Dirichlet, connue sous le nom de « sphères de Poincaré », qui se décrit comme suit. Pour simplifier, nous supposons que $D$ est un domaine borné. On se donne $(U_n)_{n\geq 1}$ une suite de vecteurs aléatoires indépendants de loi uniforme sur la sphère unité $\partial B(0,1)$, c'est-à-dire que pour toute fonction $g : \mathbb{R}^d \to \mathbb{R}$ mesurable bornée, on a

$$\mathbf{E}[g(U_n)] = \tfrac{1}{\ell_d(B(0,1))} \int_{B(0,1)} g\big(\tfrac{1}{\|y\|} y\big)\, \ell_d(dy).$$

À partir de $x \in D$ et de la suite $(U_n)_{n\geq 1}$, nous définissons une suite de v.a. $(X_n(x))_{n\in\mathbb{N}}$ à valeurs dans $D$ de la manière récursive suivante.

$(a)$ On pose $X_0(x) = x$.

$(b)$ Supposons que $X_n(x)$ ait déjà été construit. On définit alors $R_n = \sup\{r \in ]0,\infty[ : B(X_n(x), r) \subset \overline{D}\}$. On pose alors

$$X_{n+1}(x) = X_n(x) + R_n U_{n+1}\ .$$

Autrement dit $X_{n+1}(x)$ est choisi uniformément sur la sphère $\partial B(X_n(x), R_n)$

On remarque que $(X_n(x))_{n\in\mathbb{N}}$ est une chaîne de Markov à temps discret mais à espace d'état continu. Modulo, des hypothèses de régularité sur $\partial D$, Poincaré affirme ce qui suit.

$(i)$ Presque sûrement, $\lim_{n\to\infty} X_n(x) = X_\infty(x)$ existe et p.s. $X_\infty(x) \in \partial D$.

$(ii)$ Pour tout $x \in \overline{D}$, on définit $u(x) = \mathbf{E}[f(X_\infty(x))]$. Alors, si $\partial D$ est suffisamment régulier, $u$ est une solution au $(D, f)$-problème de Dirichlet.

Cette procédure remarquable explique parfaitement comment la propriété de la valeur moyenne s'étend à la frontière. Le point $(i)$ est relativement facile à prouver à l'aide du théorème de convergence des martingales (et lorsque $D$ est borné). Cependant $(ii)$ est plus difficile à prouver directement mais si le problème de Dirichlet admet une solution (et si $D$ est borné) alors il n'est pas difficile de montrer qu'elle elle est nécessairement donnée par $(ii)$ : voir l'exercice I.18, page 36 en fin de section.

**Le problème de Dirichlet sur les graphes.** La méthode des sphères de Poincaré ne permet pas une preuve facile de l'existence d'une solution au problème de Dirichlet. Néanmoins, une version discrétisée de ce problème sur les graphes est prouvable, ce qui est le but de cette partie. Cela fournit également une méthode d'approximation possible des solutions au problème de Dirichlet dans $\mathbb{R}^d$.

Soit $\mathbf{G} = (\mathbf{S}, \mathbf{A})$, un graphe simple (non-orienté sans boucle et sans arête multiple). On suppose $\mathbf{S}$ dénombrable et *connexe*, c'est-à-dire que pour tout $s, s' \in \mathbf{S}$ il existe $s_0, s_1, \ldots, s_n \in \mathbf{S}$ tels que $s_0 = s$, $s_n = s'$ et $s_0 \sim s_1 \sim \ldots \sim s_n$. On le munit d'un système de poids $\mathbf{C} = (C_a \,;\, a \in \mathbf{A})$. On rappelle que $0 < C_a < \infty$, pour toute arête $a \in \mathbf{A}$ et que

$$\forall s \in \mathbf{S}\,, \quad \pi(s) = \sum_{s' \in \mathbf{S}: s' \sim s} C_{\{s,s'\}} \ \in ]0, \infty[\,.$$

On considère la marche aléatoire sur $\mathbf{G}$ associée au système de poids $\mathbf{C}$ :

$$\big(\Omega\,;\, \mathscr{F}\,;\, (\mathscr{F}_n)_{n \geq 0}\,;\, (X_n)_{n \geq 0}\,;\, Q = (p(s,s'))_{s,s'}\,;\, \mathbf{P}_\mu\,,\, \mu \in \mathscr{M}_1(\mathbf{S})\ \big)$$

On rappelle que

$$\forall s, s' \in \mathbf{S}\,, \quad p(s,s') = C_{\{s,s'\}}/\pi(s) \quad \text{si } s \sim s' \quad \text{et} \quad p(s,s') = 0 \quad \text{sinon.}$$

**Définition I.1.38** Soit $D \subset \mathbf{S}$.

(a) On dit que $D$ est un *domaine* du graphe $\mathbf{G}$ si et seulement si $D$ est *intrinsèquement connexe*, c'est-à-dire que pour toute paire de sommets distincts $s$ et $s'$ de $D$, il existe un chemin $s_0 = s \sim s_1 \sim \ldots \sim s_n = s'$ tel que $s_1, \ldots, s_n \in D$.

(b) Le bord d'un domaine $D$ de $\mathbf{G}$ est donné par

$$\partial D := \{ s \in \mathbf{S} \backslash D \ :\ \exists s' \in D\ ,\ s \sim s' \}\ .$$

(c) Soit $D \subset \mathbf{S}$, un domaine. Soit $u : D \cup \partial D \to \mathbb{R}$. On dit que $u$ is *harmonique sur* $D$ si

$$\forall s \in D\ , \quad u(s) = \frac{1}{\pi(s)} \sum_{s' \in \mathbf{S}: s \sim s'} C_{s,s'} u(s')\ . \tag{I.38}$$

**Remarque I.1.39** Soit $D$, un domaine du graphe $\mathbf{G}$ set soit $u : D \cup \partial D \to \mathbb{R}$. Alors pour toute mesure intiale $\mu \in \mathscr{M}_1(\mathbf{S})$, $\mathbf{P}_\mu$-p.s. si $X_0 \in D$, alors $X_1 \in D \cup \partial D$. On observe également que $u$ est harmonique sur $D$ si et seulement si pour tout $s \in D$, $u(s) = \mathbf{E}_s[u(X_1)]$. □

**Proposition I.1.40 (Principe du maximum discret)** *Soit $D$, un domaine de* $\mathbf{G}$. *Soit $u : D \cup \partial D \to \mathbb{R}$, une fonction harmonique sur $D$. Supposons que*

$$\exists s_0 \in D \ : \quad u(s_0) = \max_{s \in D \cup \partial D} u(s) \ .$$

*Alors, $u$ est constante sur $D \cup \partial D$.*

**Preuve :** on définit $K = \{ s \in D : u(s) = \max_{s' \in D \cup \partial D} u(s') \}$. Alors, $s_0 \in K$. Soit $s \in K$. Puisque $u$ est harmonique sur $D$,

$$\sum_{s' \in \mathbf{S} : s \sim s'} C_{\{s,s'\}}(u(s) - u(s')) = 0 \ .$$

Puisque, par définition $C_{\{s,s'\}} > 0$ et puisque $s \in K$, $u(s) - u(s') \geq 0$, pour tout $s'$ tel que $s' \sim s$. Par conséquent, $u(s') = u(s)$ et donc $s' \in K$ pour tout $s'$ tel que $s \sim s'$. Puisque $D$ est connexe, on obtient alors que $K = D$, ce qui implique facilement la proposition. ■

**Définition I.1.41** Soit $D \subset \mathbf{S}$, un domaine tel que $\partial D \neq \emptyset$. Soit $f : \partial D \to \mathbb{R}_+$. Alors $u : D \cup \partial D \to \mathbb{R}_+$ est une solution au $(D, f)$-*problème de Dirichlet discret* si cette fonction satisfait les conditions suivantes.

$(a)$ $u$ est harmonique dans $D$,

$(b)$ $u(i) = f(i)$, pour tout $i \in \partial D$. □

Le principe du maximum implique le résultat d'unicité suivant pour le problème de Dirichlet.

**Proposition I.1.42 (Principe d'unicité)** *Soit $D$ un domaine dans* $\mathbf{S}$. *Nous supposons que $D$ est fini. Soit $f : \partial D \to \mathbb{R}_+$. Supposons que $u_1$ et $u_2$ soient deux solutions au $(D, f)$-problème de Dirichlet. Alors, $u_1 = u_2$.*

**Preuve :** on pose $u = u_1 - u_2$. On observe tout d'abord que $u$ est une solution au $(D, 0)$-problème de Dirichlet. Supposons que $u_1 \neq u_2$. Puisque $u_1$ et $u_2$ coïncident sur $\partial D$, il existe un sommet $s \in D$. tel que $u_1(s) \neq u_2(s)$. Puisque $u_1$ et $u_2$ jouent le même rôle, on peut supposer sans perte de généralité que $u_1(s) > u_2(s)$. Puisque $D$ est fini,

$$\exists s_0 \in D \ : \quad u(s_0) = \max_{s' \in D \cup \partial D} u(s') \ > \ 0 \ .$$

Par conséquent $u$ est constant sur $D \cup \partial D$ par le principe du maximum discret. Puisque, $u$ est nul sur $\partial D$, $u$ doit être nul sur $D$, ce qui contredit $u_1(s) \neq u_2(s)$. Cela montre par l'absurde que $u_1 = u_2$. ■

Énonçons maintenant le résultat principal concernant le problème de Dirichlet discret.

**Théorème I.1.43 (Solution probabiliste au problème de Dirichlet discret)** *Soit $D$ un domaine de* $\mathbf{G}$ *tel que $\partial D \neq \emptyset$. Soit $f : \partial D \to \mathbb{R}_+$. On définit $T_{D^c} = \inf\{ n \in \mathbb{N} : X_n \notin D \}$, avec la convention que $\inf \emptyset = \infty$. On définit ensuite la fonction $u : D \cup \partial D \to \mathbb{R}_+$ par*

$$\forall s \in D \cup \partial D \ , \quad u(s) = \mathbf{E}_s \left[ \mathbf{1}_{\{T_{D^c} < \infty\}} f(X_{T_{D^c}}) \right] \ .$$

*Alors $u$ est une solution au $(D, f)$-problème de Dirichlet discret. De plus, si $D$ est fini, c'est la seule solution.*

**Preuve :** pour simplifier les notations, on définit $F(\mathbf{X}) = \mathbf{1}_{\{T_{D^c} < \infty\}} f(X_{T_{D^c}})$. On observe que si $X_0 \in \partial D$, $T_{D^c} = 0$ et $F(\mathbf{X}) = f(X_0)$. Ainsi, si $s \in \partial D$, alors $u(s) = f(s)$. Remarquons ensuite que si $X_0 \in D$, alors $T_{D^c} \geq 1$ et $F(\mathbf{X}) = F(\theta_1 \mathbf{X})$. Ainsi, si $s \in D$, la propriété de Markov simple au temps 1 implique ce qui suit :

$$u(s) \quad = \quad \mathbf{E}_s \left[ F(\mathbf{X}) \right] = \mathbf{E}_s \left[ F(\theta_1 \mathbf{X}) \right]$$

$$= \mathbf{E}_s\left[\mathbf{E}_{X_1}\left[F(\mathbf{X})\right]\right]$$
$$= \mathbf{E}_s\left[u(X_1)\right] \ ,$$

ce qui implique que $u$ est harmonique sur $D$ par la remarque I.1.39 (page 33). Nous avons donc prouvé que $u$ est solution au $(D, f)$-problème de Dirichlet discret. Le principe d'unicité complète la preuve du théorème. ■

Lorsque $D$ n'est pas fini, le problème de Dirichlet n'a pas nécessairement de solution unique comme montré dans l'exercices I.19, en fin de section page 36. Nous renvoyons également à l'exercice I.36 (page 73) pour plus de détails lorsque le domaine $D$ n'est pas fini, situation dans lesquelles des hypothèses probabilistes sont à préciser.

Pour conclure expliquons comment un problème de Dirichlet discret sur $\mathbb{Z}^d$ (utilisant la marche aléatoire symétrique) peut être utilisé pour approcher un problème de Dirichlet sur $\mathbb{R}^d$. Ici cette présentation n'est pas mathématiquement justifiée dans les détails. Voici l'esquisse de la méthode : on fixe un domaine borné $D$ dans $\mathbb{R}^d$ ; soit $f : \partial D \to \mathbb{R}$, une fonction continue. Nous supposons que le $(D, f)$-problème de Dirichlet (dans $\mathbb{R}^d$) a une solution unique (ce qui est vrai si $D$ a une frontière suffisamment régulière).

— Nous fixons un pas $\varepsilon > 0$ qui est petit. Nous approchons $\mathbb{R}^d$ par $\varepsilon\mathbb{Z}^d = \{\varepsilon x; x \in \mathbb{Z}^d\}$, que l'on voit comme un graphe. Nous définissons ensuite

$$D_\varepsilon = \{x \in \mathbb{Z}^d : \varepsilon x \in D\} \ ; .$$

Si $D$ est suffisamment régulier et $\varepsilon$ suffisamment petit, $D_\varepsilon$ est aussi un domaine de $\mathbb{Z}^d$ (au sens de la définition I.1.38, page 33). On note $\partial D_\varepsilon$ le bord de $D_\varepsilon$, au sens des graphes (définition I.1.38, page 33). Dans un certain sens $\partial D$ est proche de $\{\varepsilon x; x \in \partial D_\varepsilon\}$. Supposons que soit définie une fonction $f_\varepsilon : \partial D_\varepsilon \to \mathbb{R}$ qui soit également « proche » de $f$, c'est-à-dire que $f_\varepsilon(x) \simeq f(x_\varepsilon)$ où $x_\varepsilon$ est point de $\partial D$ le plus proche de $\varepsilon x$. Soit $u_\varepsilon$ une solution du problème $(D_\varepsilon, f_\varepsilon)$ dans $\varepsilon\mathbb{Z}^d$. Nous affirmons alors que

$$\forall x \in D_\varepsilon, \quad u_\varepsilon(x) \simeq u(\varepsilon x) \ ,$$

avec un contrôle uniforme sur $|u_\varepsilon(x) - u(\varepsilon x)|$ en termes de $\varepsilon$.

— Nous obtenons $u_\varepsilon$ grâce au théorème I.1.43 et à la loi des grands nombres. Plus précisément, on fixe $p \in \mathbb{N}$ (grand) et pour n'importe quel $x \in D_\varepsilon$, on génère $p$ marches aléatoires symétriques simples sur $\mathbb{Z}^d$ indépendantes

$$(X_n(1))_{n\in\mathbb{N}}, \ldots, (X_n(p))_{n\in\mathbb{N}},$$

toutes issues du vecteur $x$. Ensuite, on définit

$$\forall k \in \{1, \ldots, p\}, \quad T_{D_\varepsilon}(k) = \inf\{n \in \mathbb{N} : X_n(k) \notin D_\varepsilon\} \quad \text{et} \quad Z_k = X_{T_{D^c_\varepsilon(k)}}(k) \ .$$

Autrement dit, $Z_k$ est la position de la marche aléatoire $(X_n(k))_{n\in\mathbb{N}}$ quand elle sort pour la première fois sort de $D_\varepsilon$. Alors, par la loi des grands nombres

$$u_\varepsilon(x) = \mathbf{E}[f(X_{T_{D_\varepsilon}})] \simeq \frac{1}{p}\big(f_\varepsilon(Z_1) + \ldots + f_\varepsilon(Z_p)\big),$$

l'erreur étant (moralement) de l'ordre de $O(p^{-1/2})$ par le théorème de la limite centrale.

La figure I.1.f page 36 montre le résultat de cette méthode dans le cas suivant : la dimension est $d = 2$ et le domaine est le carré $D := [-1, 1]^2$. La fonction $f : \partial D \to \mathbb{R}_+$ est donnée par

$$\forall x = (x_1, x_2) \in \partial D, \quad f(x) = 4\cos^2\left(\tfrac{x_1+x_2}{6}\right) \ ; .$$

Le pas du maillage est $\varepsilon = 1/30$ et le nombre $p$ de marches aléatoire est $p = 400$.

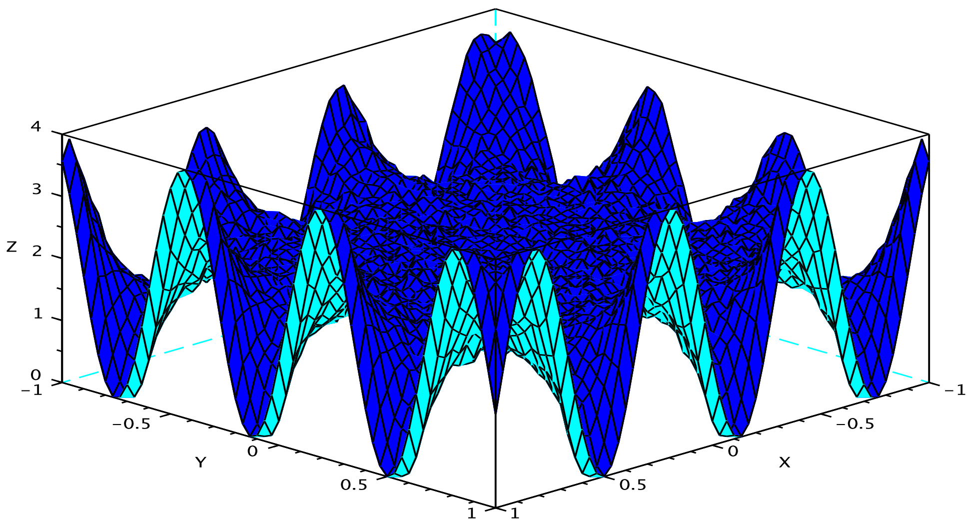


FIGURE I.4 – Approximation probabiliste d'une solution de Dirichlet. Ici $z = u(x,y)$.

## EXERCICES.

**Exercice I.18** Soit $D$ un domaine borné de $\mathbb{R}^d$. Soit $x \in D$. Soient $X_n(x)$ et $R_n$, $n \in \mathbb{N}$, les variables construites par l'algorithme des *sphères de Poincaré* à la page 32.

1. Soit $v : D \to \mathbb{R}$ une fonction bornée et $C^2$ telle que $\Delta v = 0$. En d'autres termes, $v$ satisfait à la propriété de la valeur moyenne. On pose alors $Y_n = v(X_n(x))$. Prouver que $(Y_n)_{n\in\mathbb{N}}$ est une martingale bornée. Elle est donc convergente.
2. Pour tout $k \in \{1,\ldots,d\}$, définir $p_k : x = (x_1,\ldots,x_d) \in D \mapsto x_k \in \mathbb{R}$. Prouvez que $\Delta p_k = 0$ sur $D$. Prouvez que presque sûrement, $\lim_{n\to\infty} X_n(x) = X_\infty(x)$ existe.
3. Pour tout $x \in \overline{D}$ on pose $\mathrm{dist}(x,\partial D) = \inf\{\|x-y\|; y \in \partial D\}$, où $\|\cdot\|$ représente la norme euclidienne sur $\mathbb{R}^d$. Prouver que
$$\forall x_1, x_2 \in \overline{D}, \quad |\mathrm{dist}(x_1,\partial D) - \mathrm{dist}(x_2,\partial D)| \le \|x_1 - x_2\|.$$
4. Prouver que pour tout $x \in \overline{D}$, $\mathrm{dist}(x,\partial D) = 0$ si $x \in \partial D$.
5. Prouver que tout $n \in \mathbb{N}$, on a $\mathrm{dist}(X_n(x),\partial D) \le 2R_n = 2\|X_n(x) - X_{n+1}(x)\|$.
6. Prouver que presque sûrement $X_\infty(x) \in \partial D$.
7. Rappelons que $D$ est borné. Soit $f : \partial D \to \mathbb{R}$, une fonction continue. Supposons que $u$ soit une solution au $(D,f)$-problème de Dirichlet. Prouvez que pour tout $x \in \overline{D}$, $u(x) = \mathbf{E}[f(X_\infty(x))]$. □

**Exercice I.19** Nous considérons $\mathbb{Z}$ comme un graphe $(\mathbf{S},\mathbf{A})$ : ici $\mathbf{S} = \mathbb{Z}$ et $\mathbf{A} = \big\{\{i,i+1\}; i \in \mathbb{Z}\big\}$, c'est-à-dire le graphe linéaire habituel. Nous fixons $\lambda \in ]0,\infty[$ et nous définissons l'ensemble de poids suivant : $C_{\{i,i+1\}} = \lambda^i$, $i \in \mathbb{Z}$. On se donne comme dans la définition I.1.30 une marche aléatoire associée à l'ensemble de poids $C$. Nous choisissons ici $\lambda > 1$. Cela implique que facilement que ue pour tout $\mu \in \mathscr{M}_1(\mathbb{Z})$), $\mathbf{P}_\mu$-p.s. $\lim_{n\to\infty} X_n = \infty$.

On définit $D = \{i \in \mathbb{Z} : i \ge 1\}$ qui est un domaine de $\mathbb{Z}$ tel que $\partial D = \{0\}$. Soit $u$ la fonction harmonique dans $D$ et telle que $u(0) = 1$ qui est donnée par le théorème I.1.43. Prouver que $\lim_{i\to\infty} u(i) = 0$. Trouvez une autre solution au même problème de Dirichlet discret. □

## I.2 Algèbre linéaire pour les matrices de transition.

Dans la suite $\big(\Omega; \mathscr{F}; (\mathscr{F}_n)_{n\geq 0}; \mathbf{X} = (X_n)_{n\geq 0}; Q = (p(i,j))_{i,j\in E}; \mathbf{P}_\mu, \mu \in \mathscr{M}_1(E)\big)$ désigne une chaîne de Markov.

### I.2.a Notation et calcul matriciel.

**Produit de matrices de transition.** On peut définir un produit pour les matrices de transition comme suit : soient $Q = (p(i,j))_{i,j\in E}$ et $Q' = (p'(i,j))_{i,j\in E}$, deux matrices de transition ; on définit la matrice produit $QQ' = Q'' = (p''(i,j))_{i,j\in E}$ par

$$\forall i,j \in E\ , \quad p''(i,j) = \sum_{k\in E} p(i,k)p'(k,j)\ ,$$

qui a toujours un sens dans $[0,\infty]$, a priori. On observe que $Q''$ est elle-même une matrice de transition : pour tout $i \in E$, on a en effet,

$$\sum_{j\in E} p''(i,j) = \sum_{j\in E}\sum_{k\in E} p(i,k)p'(k,j) = \sum_{k\in E} p(i,k)\sum_{j\in E} p'(k,j) = \sum_{k\in E} p(i,k).1 = 1\ .$$

On définit récursivement $Q^n$ par $Q^{n+1} = QQ^n = Q^nQ$ et $Q^0 = \mathrm{Id}_E$, l'identité sur $E$. Il est parfois utile de poser la notation suivante $[Q](i,j) = p(i,j)$ où $p(i,j)$ est *l'entrée* $(i,j)$ *de* $Q$. On vérifie facilement par récurrence que

$$\forall n \geq 2\ ,\ \forall i,j \in E\ , \quad [Q^n](i,j) = \sum_{i_1,\ldots,i_{n-1}\in E} p(i,i_1)p(i_1,i_2)\ldots p(i_{n-1},j)\ . \tag{I.39}$$

**Action sur les fonctions.** Comme déjà expliqué au début de ce cours, une matrice de transition $Q = (p(i,j))_{i,j\in E}$ agit sur les fonctions de $E$ dans $\mathbb{R}$ : à une telle fonction $f$, on associe la fonction $Q.f$ donnée par $(Q.f)(i) = \sum_{j\in E} p(i,j)f(j)$ qui a bien un sens dès que

$$\forall i \in E\ , \quad \sum_{j\in E} p(i,j)|f(j)| < \infty\ .$$

Cette action préserve les fonctions bornées. Plus précisément on rappelle que $\|Q.f\|_\infty \leq \|f\|_\infty$ dès que $\|f\|_\infty < \infty$.

**Action sur les mesures.** La matrice de transition $Q$ agit également sur les mesures positives sur $E$ de la façon suivante : si $\mu$ est une mesure positive, on définit la mesure $\mu.Q$ en posant

$$\forall j \in E\ , \quad (\mu.Q)(j) = \sum_{i\in E} \mu(i)p(i,j)\ \in [0,\infty]\ . \tag{I.40}$$

Il faut bien remarquer que $(\mu.Q)(j)$ peut être une quantité infinie. On note $\langle\mu\rangle = \sum_{j\in E}\mu(j)$ la *masse totale* de $\mu$. Notamment $\mu \in \mathscr{M}_1(E)$ ssi $\langle\mu\rangle = 1$. On note

$$\mathscr{M}_f(E) = \big\{\mu : E \to [0,\infty] : \langle\mu\rangle < \infty\big\}\ ,$$

l'ensemble des *mesures de masse finie sur* $E$. L'action de $Q$ sur les mesures de masse finie préserve leur masse, c'est-à-dire que

$$\forall \mu \in \mathscr{M}_f(E), \quad \langle \mu \rangle = \langle \mu.Q \rangle$$

car $\langle \mu.Q \rangle = \sum_{j\in E} \mu.Q(j) = \sum_{i,j\in E} \mu(i)p(i,j) = \sum_{i\in E} \mu(i) \sum_{j\in E} p(i,j) = \sum_{j\in E} \mu(i) = \langle \mu \rangle$.

En particulier, on voit que

$$\mu \in \mathscr{M}_1(E) \implies \mu.Q \in \mathscr{M}_1(E) .$$

Pour toute fonction $f : E \to \mathbb{R}$ et toute mesure positive $\mu$, on définit

$$\langle \mu, f \rangle = \sum_{i\in E} \mu(i)f(i)$$

qui a bien un sens si $\sum_{i\in E} \mu(i)|f(i)| < \infty$. Si $f$ est bornée et $\mu \in \mathscr{M}_f(E)$, alors $|\langle \mu, f\rangle| \leq \langle \mu \rangle \|f\|_\infty$.

**Lemme I.2.1** *Soit* $\mu \in \mathscr{M}_1(E)$. *La loi de* $X_n$ *sous* $\mathbf{P}_\mu$ *est* $\mu.Q^n$. *Donc, pour toute fonction* $f : E \to \mathbb{R}$ *bornée,*

$$\forall n \in \mathbb{N}, \quad \mathbf{E}_\mu[f(X_n)] = \langle \mu.Q^n \, , \, f \rangle = \langle \mu \, , \, Q^n.f \rangle.$$

**Preuve :** on remarque que (I.39) implique que pour tout $j \in E$, on a

$$\begin{aligned}
\mathbf{P}_\mu(X_n = j) &= \sum_{i,i_1,\dots,i_{n-1}\in E} \mathbf{P}_\mu(X_0 = i; \dots ; X_n = j) = \sum_{i,i_1,\dots,i_{n-1}\in E} \mu(i)p(i,i_1)p(i_1,i_2)\dots p(i_{n-1},j) \\
&= \sum_{i\in E} \mu(i) \Big( \sum_{i_1,\dots,i_{n-1}\in E} p(i,i_1)p(i_1,i_2)\dots p(i_{n-1},j) \Big) = \sum_{i\in E} \mu(i)[Q^n](i,j) \\
&= \mu.Q^n(j) .
\end{aligned}$$

Cela montre bien que $\mu.Q^n$ est la loi de $X_n$ sous $\mathbf{P}_\mu$. ■

Un des buts généraux de ce chapitre est de trouver des hypothèses satisfaisantes sous lesquelles $\lim_n [Q^n](i,j)$ existe.

**Remarque I.2.2** Lorsque l'espace $E$ est fini, ce qui précède se traduit par du "vrai" calcul matriciel. On peut toujours se ramener à $E = \{1, \dots, N\}$, où $N = \#E$. Une fonction est un vecteur, c'est-à-dire une matrice $N \times 1$, une mesure est un vecteur du dual (une forme linéaire), c'est-à-dire une matrice $1 \times N$. ■

**Exemple I.2.3** On considère une marche aléatoire sur $\mathbb{Z}^d$ de loi de saut $\pi$. La matrice de transition de la marche $Q = (p(\mathbf{i},\mathbf{j}))_{\mathbf{i},\mathbf{j}\in\mathbb{Z}^d}$ est donnée par $\pi(\mathbf{j}-\mathbf{i}) = p(i,j)$ (voir l'exemple I.1.5, page 4). On étend la notion de produit de convolution $*$ introduite à l'exemple I.1.b, page 9, aux mesures de probabilité sur $\mathbb{Z}^d$ et on vérifie facilement que

$$\forall n \in \mathbb{N}, \, \forall \mathbf{i}, \mathbf{j} \in \mathbb{Z}^d, \quad [Q^n](\mathbf{i},\mathbf{j}) = \pi^{*n}(\mathbf{j} - \mathbf{i}) .$$

Les détails sont laissés au lecteur. □

**Exemple I.2.4** Soit $Q = (p(i,j))_{i,j\in\mathbb{N}}$, une matrice de transition sur $\mathbb{N}$. On dit qu'elle satisfait la *propriété de branchement* si la condition suivante est satisfaite :

$$\forall i, i' \in \mathbb{N}, \quad p(i,\cdot) * p(i',\cdot) = p(i+i',\cdot) , \tag{I.41}$$

où $*$ désigne le produit de convolution introduit à l'exemple I.1.b, page 9. On pose $\xi = p(1,\cdot)$. Il est alors facile de voir que $p(i,\cdot) = \xi^{*i}$ et donc que $Q$ est la matrice d'un processus de branchement de loi de reproduction $\xi$.

### I.2.b Irréductibilité, période.

Dans cette section nous étudions les propriétés *algèbriques* d'une matrice de transition $Q = (p(i,j))_{i,j \in E}$.

**Définition I.2.5** Soit $F \subset E$, un sous-ensemble de $E$. On dit qu'il est *Q-fermé* si $p(i,j) = 0$, pour tout $i \in F$ et tout $j \notin F$. □

**Lemme I.2.6** *Soit $F \subset E$, un sous-ensemble de $E$. On note $Q_{|F}$ la matrice $Q$, restreinte à $F$, c'est-à-dire $Q_{|F} = (p(i,j))_{i,j \in F}$. Alors, $F$ est un sous-ensemble Q-fermé ssi $Q_{|F}$ est une matrice de transition sur $F$.*

**Preuve :** pour tout $i \in F$, on a $1 = p(i,F) + p(i, E\backslash F) := \sum_{j \in F} p(i,j) + \sum_{j \notin F} p(i,j)$. Par conséquent, $\sum_{j \notin F} p(i,j) = 0$ ssi $\sum_{j \in F} p(i,j) = 1$, ce qui permet de conclure. ■

**Exemple I.2.7** On considère la marche aléatoire sur $\mathbb{Z}$ de loi de saut $\pi$, donnée par

$$\pi(-2) = \pi(2) = \tfrac{1}{2} \ .$$

On note $Q$ la matrice de transition d'une telle marche. On voit que l'ensemble $2\mathbb{Z}$ des entiers pairs est fermé ainsi que l'ensemble $2\mathbb{Z}+1$ des entiers impairs. On voit d'ailleurs que l'étude d'une telle marche se ramène à étudier la marche aléatoire simple sur $\mathbb{Z}$. □

Le lemme suivant montre que les parties $Q$-fermées constituent des pièges pour les chaînes de matrice de transition $Q$.

**Lemme I.2.8** *Si $F$ est Q-fermé, alors pour toute loi d'entrée $\mu \in \mathscr{M}_1(E)$,*

$$\mathbf{P}_\mu\text{-}p.s. \quad \forall n \in \mathbb{N}\,, \quad \big(X_n \in F\big) \Longrightarrow \big(\forall m \in \mathbb{N}\,,\ X_{n+m} \in F\big) \ . \tag{I.42}$$

*Supposons que $\mu(F) = 1$, alors sous $\mathbf{P}_\mu$, $\mathbf{X}$ est une chaîne de Markov de loi d'entrée $\mu$ et de matrice de transition $Q_{|F}$.*

**Preuve :** il est facile de montrer que pour tout $i \in F$, tout $j \in E$ et tout $\ell \in \mathbb{N}$, on a

$$\mathbf{P}_i(X_\ell = j) = 0 \text{ si } j \notin F \quad \text{et} \quad \mathbf{P}_i(X_\ell = j) = \big[(Q_{|F})^\ell\big](i,j) \text{ si } j \in F. \tag{I.43}$$

Pour montrer (I.42), il suffit de montrer que

$$\mathbf{P}_\mu\big(\exists n, m \in \mathbb{N} : X_n \in F \text{ et } X_{n+m} \notin F\big) = 0 \ . \tag{I.44}$$

Or on remarque que

$$\big\{\exists n, m \in \mathbb{N} : X_n \in F \text{ et } X_{n+m} \notin F\big\} = \bigcup_{n,m \in \mathbb{N}} \bigcup_{i \in F} \bigcup_{j \notin F} \{X_n = i ; X_{n+m} = j\} \ ,$$

Il suffit donc montrer que pour tout $n \in \mathbb{N}$, tout $m \in \mathbb{N}^*$, tout $i \in F$ et tout $j \notin F$ que $\mathbf{P}_\mu(X_n = i; X_{n+m} = j) = 0$. La propriété de Markov au temps $n$ sous $\mathbf{P}_\mu$ implique que $\mathbf{P}_\mu(X_n = i; X_{n+m} = j) = \mathbf{P}_\mu(X_n = i)\mathbf{P}_i(X_m = j)$ et (I.43) montre que $\mathbf{P}_i(X_m = j) = 0$, ce qui prouve donc (I.42). Le dernier point du lemme est une conséquence simple du lemme I.2.6, dont la preuve est laissée au lecteur. ■

**Définition I.2.9** Rappelons que $Q = (p(i,j))_{i,j \in E}$ est une matrice de transition sur $E$.

**(Accessibilité)** Soient $i, j \in E$. On dit que $j$ est *accessible* depuis $i$, ce qui est noté $i \to j$, si $i = j$ ou si $i \neq j$ et s'il existe $n \geq 1$ et $i_1, \dots i_{n-1} \in E$ tels que

$$p(i, i_1)p(i_1, i_2) \dots p(i_{n-1}, j) > 0 .$$

On observe que si $p(i, j) > 0$, alors $i \to j$. D'autre part (I.39) montre que

$$i \to j \iff \exists n \geq 0 , \quad [Q^n](i, j) > 0 .$$

**(Communication)** Soient $i, j \in E$. On dit que $i$ et $j$ *communiquent* (ce qui est noté $i \leftrightarrow j$) ssi $i \to j$ et $j \to i$. On remarque que $\leftrightarrow$ est une *relation d'équivalence sur* $E$.

**(Irréductibilité)** On dit que $Q$ est *irréductible* ssi tous ses états communiquent.

**(Etat absorbant)** Un état $i \in E$ est dit *absorbant* ssi $p(i, i) = 1$. On remarque que si une matrice de transition sur un espace d'états admettant au moins deux éléments distincts, possède un état absorbant, alors elle n'est pas irréductible.

**(Période)** Soit $i \in E$. On pose $D_i = \{n \geq 1 : [Q^n](i, i) > 0\}$. Si $D_i$ est vide, alors on définit *la période de* $i$ par $d_i := \infty$. Si $D_i$ n'est pas vide, on définit la période de $i$ comme le plus grand commun diviseur de l'ensemble $D_i$ :

$$d_i = \mathrm{PGCD}\{n \geq 1 : [Q^n](i, i) > 0\} .$$

Si $d_i = 1$, on dit que $i$ est *apériodique*. □

**Remarque I.2.10** Il est nécessaire de préciser ce que l'on entend par le *PGCD d'un ensemble infini d'entiers*. Soit $D$, un ensemble infini d'entiers. Pour tout $q \in \mathbb{N}$, on pose $d(q) = \mathrm{PGCD}(D \cap \{1, \dots, q\})$. Il est facile de voir que $q \mapsto d(q) \geq 1$ décroît : cette suite d'entiers est donc stationnaire à sa limite notée $d$, que l'on définit comme $\mathrm{PGCD}(D)$. □

**Exemple I.2.11** On considère le graphe circulaire à $d$ sommets : $\mathbf{S} = \{0, \dots, d-1\}$ et

$$\mathbf{A} = \big\{ \{0, 1\} , \dots , \{s, s+1\} , \dots , \{d-2, d-1\} , \{d-1, 0\} \big\} .$$

On note $Q = (p(s, s'))_{s,s' \in \mathbf{S}}$ la matrice donnée par $p(s, \overline{s+1}) = 1$ où $\overline{s+1}$ désigne le seul entier $\ell \in \{0, \dots, d-1\}$ tel que $d$ divise $s + 1 - \ell$. On voit que $Q$ est irréductible et $d$ périodique. Cette matrice de transition correspond à une chaîne de Markov déterministe qui tourne sur le graphe circulaire à $d$ sommets. □

**Exemple I.2.12** On considère la marche aléatoire simple sur $\mathbb{Z}^d$. Il est assez facile de voir qu'elle est irréductible et 2 périodique. □

**Exemple I.2.13** Pour les processus de branchement de loi de reproduction, l'état 0 est absorbant : cela découle de la définition. □

**Exemple I.2.14** Soit $\mathbf{G} = (\mathbf{S}, \mathbf{A})$, un graphe non-orienté, simple (pas d'arêtes multiples), et bien sûr, dénombrable (c'est-à-dire que $\mathbf{S}$ est dénombrable). On munit $\mathbf{G}$ d'un système de poids $(C_a ; a \in \mathbf{A})$ et on suppose toujours que

$$\forall s \in \mathbf{S} , \quad \pi(s) = \sum_{s' \sim s} C_{\{s,s'\}} \ \in \mathbb{R}_+^* .$$

On note $Q = (p(s, s'))_{s,s' \in \mathbf{S}}$ la matrice de transition de la marche sur $\mathbf{G}$ associée aux poids $\mathbf{C} = (C_a; a \in \mathbf{A})$, comme défini à l'exemple I.1.7, page 5.

Soit $n \in \mathbb{N}^*$ et soient $n + 1$ sommets $s_0, s_1, \ldots, s_n \in \mathbf{S}$. Alors on introduit la notion de *chemin* comme suit :

$$\gamma = (s_0, s_1, \ldots, s_n) \text{ est un } \textit{chemin} \textbf{ ssi } \text{pour tout } 0 < k \leq n, \text{ on a } \{s_{k-1}, s_k\} \in \mathbf{A}. \tag{I.45}$$

On dit alors que $\gamma$ *relie* $s_0$ *à* $s_n$ *dans* $\mathbf{G}$. On définit ensuite le *poids d'un chemin* $\gamma = (s_0, s_1, \ldots, s_n)$ comme le produit des poids des arêtes qui le composent :

$$\pi_{\mathbf{C}}(\gamma) = C_{\{s_0,s_1\}} C_{\{s_1,s_2\}} \ldots C_{\{s_{n-1},s_n\}} \ .$$

On voit facilement que

$$\pi_{\mathbf{C}}(\gamma) := \big(\pi(s_0)\pi(s_1)\ldots\pi(s_{n-1})\big) \cdot \big(p(s_0, s_1)p(s_1, s_2)\ldots p(s_{n-1}, s_n)\big) \ .$$

Cela montre que $Q$ est irréductible **ssi** pour tous sommets $s, s'$, il existe un chemin les reliant qui est de poids strictement positif. On observe aussi que pour une marche sur un graphe, $s \to s'$ implique $s' \to s$ et donc que $s \leftrightarrow s'$, ce qui est une propriété remarquable sur laquelle nous reviendrons plus loin.

Lorsque l'on considère une marche aléatoire simple sur $\mathbf{G}$, alors les poids sont constants et on rappelle que $p(s, s') = 1/\deg(s)$ si $s \sim s'$ et $p(s, s') = 0$, sinon. On voit alors qu'*une marche aléatoire simple sur un graphe est irréductible* **ssi** *pour tous sommets $s, s' \in \mathbf{S}$, il existe un chemin les reliant* et on dit dans ce cas que le graphe est *connexe*. □

**Exemple I.2.15** Soit $Q = (p(i, j))_{i,j \in \mathbb{N}}$, la matrice de transition d'un processus de naissance et de mort, comme défini à l'exemple I.1.6, c'est-à-dire que $p(i, j) = 0$ dès que $|i - j| \geq 2$. On voit que si $p(n, n+1) = 0$, alors $F = \{0, \ldots, n\}$ est un ensemble $Q$-fermé. De même si $p(n, n - 1) = 0$, alors $\{n, n + 1, \ldots\}$ est un ensemble $Q$-fermé. On suppose que $p(n, n+1)p(n+1, n) > 0$, pour tout $n \in \mathbb{N}$. Alors, $Q$ est irréductible. La réciproque est vraie (nous laissons cela en exercice). □

On s'intéressera surtout aux chaînes irréductibles et la plupart des exemples de chaînes non-irréductibles que nous traiterons seront des chaînes possédant des éléments absorbants. Examinons ici la notion de période bien qu'elle n'intervienne qu'à la fin de ce cours.

**Théorème I.2.16** *Soient $i, j \in E$ deux états qui communiquent : $i \leftrightarrow j$. On note $d_i$ la période de $i$ et $d_j$, celle de $j$. Alors, $d_i = d_j := d$. De plus, il existe un entier $n(i, j) \geq 1$, qui dépend de $i$ et de $j$, tel que*

$$\forall n \geq 0 \ , \quad [Q^{n(i,j)+nd}](i, j) > 0 \ .$$

*En particulier, si $d = 1$ (c-à-d si $i$ et $j$ sont apériodiques) alors $[Q^n](i, j) > 0$, pour tout entier $n$ assez grand.*

**Preuve :** on pose $D_i = \{n \geq 1 \ : \ [Q^n](i, i) > 0\}$ et $D_j = \{n \geq 1 \ : \ [Q^n](j, j) > 0\}$. Puisque $i \leftrightarrow j$, il existe $m_0, n_0 \geq 1$ tels que $[Q^{m_0}(i, j)] > 0$ et $[Q^{n_0}](j, i) > 0$. Soit $n \in D_j$. Pour tout $p \geq 0$, on observe que

$$\begin{aligned}
[Q^{m_0+n_0+pn}](i, i) \quad &= \quad \sum_{i_0,\ldots,i_p} [Q^{m_0}](i, i_0) \cdot [Q^n](i_0, i_1) \ldots [Q^n](i_{p-1}, i_p) \cdot [Q^{n_0}](i_p, i) \\
&\geq \quad [Q^{m_0}](i, j) \cdot ([Q^n](j, j))^p \cdot [Q^{n_0}](j, i) > 0 \ .
\end{aligned}$$

On voit que pour tout $n \in D_j$ et tout $p \in \mathbb{N}$, on a $m_0 + n_0 + pn \in D_i$. Cela implique notamment que $m_0 + n_0 \in D_i$ et que $m_0 + n_0 + n \in D_i$. Par conséquent $d_i | n$, pour tout $n \in D_j$, ce qui entraîne que $d_i | d_j$. De même, on montre que $d_j | d_i$, ce qui montre que $d_i = d_j = d$.

Soient $m, n \in D_j$. On observe que

$$[Q^{m+n}](j,j) = \sum_{k \in E} [Q^m](j,k) \cdot [Q^n](k,j) \geq [Q^m](j,j)[Q^n](j,j) > 0 \,,$$

ce qui montre que $m + n \in D_j$. Par conséquent, $D_j$ *est stable par addition*. Il est facile ensuite de voir qu'il existe $n_1, \ldots, n_\ell \in D_j$ tels que $d = d_j = \mathrm{PGCD}\{n_1, \ldots, n_\ell\}$. Par le théorème de Bezout, il existe $k_1, \ldots, k_\ell \in \mathbb{Z}$ tels que $k_1 n_1 + \ldots + k_\ell n_\ell = d$. On pose alors

$$a = \sum_{\substack{1 \leq p \leq \ell \\ k_p > 0}} k_p n_p \quad \text{et} \quad b = \sum_{\substack{1 \leq p \leq \ell \\ k_p < 0}} |k_p| n_p \,.$$

Puisque $D_j$ est stable par addition, $a, b \in D_j$. De plus, on a $d = a - b$, par définition. On considère deux cas.
*Cas 1 :* $b = 0$. Alors $d = a \in D_j$. Puisque $D_j$ est stable par addition, on a $nd \in D_j$, pour tout $n \geq 1$. Cela entraîne facilement que

$$\forall n \geq 1 \,, \quad [Q^{m_0 + nd}](i,j) \geq [Q^{m_0}](i,j) \cdot [Q^{nd}](j,j) > 0 \,,$$

et on a prouvé le théorème avec $n(i,j) = m_0$.
*Cas 2 :* $b \geq 1$. Soit $n \geq b(b-1)$. On note $q$ le quotient de la division Euclidienne de $n$ par $b$ : $n = qb + r$, où $r \in \{0, \ldots, b-1\}$. On a alors $nd = dqb + dr$. Rappelons que $d = a - b$. On a donc $nd = dqb + (a-b)r = ar + b(dq - r)$. Puisque $n \geq b(b-1)$, $q \geq b - 1$ et donc $qd - r \geq d(b-1) - r \geq 0$. Comme $a, b \in D_j$ et que $D_j$ est stable par addition, cela implique que $nd \in D_j$, pour tout $n \geq b(b-1)$. Cela entraîne facilement que

$$\forall n \geq b(b-1) \,, \quad [Q^{m_0 + nd}](i,j) \geq [Q^{m_0}](i,j) \cdot [Q^{nd}](j,j) > 0 \,,$$

et on a prouvé le théorème avec $n(i,j) = m_0 + db(b-1)$. ■

**Corollaire I.2.17** *Si $Q$ est irréductible, tous les états ont la même période $d$, qui est appelée période de $Q$ (ou de la chaîne de Markov associée). Si $d = 1$, on parle de matrice de transition (ou de chaîne) apériodique.*

**Exemple I.2.18** Soit $Q = (p(i,j))_{i,j \in \mathbb{N}}$ la matrice de transition d'un processus de naissance et de mort, comme défini à l'exemple I.1.6, c'est-à-dire que $p(i,j) = 0$ dès que $|i - j| \geq 2$. On a vu à l'exemple I.2.15 que $Q$ est irréductible ssi

$$\forall n \geq 0 \,, \quad p(n, n+1) p(n+1, n) > 0 \,.$$

*On suppose que $Q$ est irréductible et on détermine la période de $Q$* : si $p(n,n) > 0$, alors il est clair que $n$ est apériodique. Il est facile de voir que si $p(n,n) = 0$, pour tout $n \in \mathbb{N}$, alors la période est 2 : en effet, pour revenir en $n$, il faut monter autant de fois que l'on est descendu et par ailleurs $[Q^2](n,n) \geq p(n,n+1)p(n+1,n) > 0$. Le résultat précédent implique donc que *$Q$ est apériodique ssi il existe $n_0 \in \mathbb{N}$ tel que $p(n_0, n_0) > 0$ et que si $p(n,n) = 0$, pour tout $n \in \mathbb{N}$, alors la période de $Q$ est* 2. □

La proposition suivante montre que l'étude d'une chaîne $d$-périodique se ramène à l'étude de $d$ chaînes apériodiques.

**Proposition I.2.19** *On suppose que $Q$ est* **irréductible de période plus grande ou égale à 2**. *Alors, la période de $Q$ est $d$* **ssi** *il existe une partition de $E$ en $d$ sous-ensembles $E_0, \ldots, E_{d-1}$ tels que*

$$\forall k \in \{0, \ldots, d-1\} \,, \ \forall j \in E \,, \quad \Big( \exists i \in E_k \,, \quad p(i,j) > 0 \Big) \iff \Big( j \in E_{\overline{k+1}} \Big) \,, \tag{I.46}$$

*où $\overline{k+1}$ est l'unique entier $\ell \in \{0, \dots, d-1\}$ tel que $d$ divise $k+1-\ell$. De plus, si les ensembles $E_0, \dots, E_{d-1}$ satisfont la condition précédente, alors pour tout $0 \leq k \leq d-1$, la matrice*

$$Q^d_{|E_k} = ([Q^d](i,j))_{i,j \in E_k}$$

*est une matrice de transition irréductible et aperiodique.*

**Preuve :** on suppose d'abord que $Q$ est $d$-périodique, avec $d \geq 2$. On fixe $i_0 \in E$ et pour tout $0 \leq k \leq d-1$, on pose

$$E_k = \{ i \in E \; : \; \exists n \geq 0 \; , \; [Q^{k+nd}](i_0, i) > 0 \} \; .$$

Puisque $Q$ est irréductible, il est facile de vérifier que $\bigcup_{k=0}^{d-1} E_k = E$. On fixe ensuite $0 \leq k < \ell \leq d-1$ et on raisonne par l'absurde en supposant que $i \in E_k \cap E_\ell$. Il existe donc $n_1, n_2 \geq 0$, tels que $[Q^{k+n_1 d}](i_0, i) > 0$ et $[Q^{\ell + n_2 d}](i_0, i) > 0$. D'autre part, puisque $Q$ est irréductible, il existe $n_3 \geq 1$,tel que $[Q^{n_3}(i, i_0)] > 0$. On voit facilement que cela implique que

$$[Q^{k+n_1 d + n_3}](i_0, i_0) \geq [Q^{k+n_1 d}](i_0, i) \cdot [Q^{n_3}](i, i_0) > 0 \; .$$

Cela entraîne que $d$ divise $k + n_1 d + n_3$. Par un argument similaire on montre que $d$ divise $\ell + dn_2 + n_3$. Par conséquent, $d$ divise $\ell - k \in \{1, \dots, d-1\}$, ce qui entraîne une contradiction. On a donc montré par l'absurde que $E_k \cap E_\ell = \emptyset$ dès que $k < \ell$.

Soit $k \in \{0, \dots, d-1\}$ et soit $j \in E$. On remarque que par définition de $E_k$ et puisque $Q^{k+1+nd} = Q^{k+nd}.Q$, on a

$$\forall n \in \mathbb{N} \, , \; \forall i \in E_k \, , \quad [Q^{k+1+nd}](i_0, j) = \sum_{i' \in E_k} [Q^{k+nd}](i_0, i') \cdot p(i', j) \; \geq \; [Q^{k+nd}](i_0, i) \cdot p(i, j) \; . \quad \text{(I.47)}$$

Si $i \in E_k$, alors, par définition il existe $n \geq 0$, tel que $[Q^{k+nd}](i_0, i) > 0$ et si $p(i,j) > 0$, (I.47) implique que $[Q^{k+1+nd}](i_0, j) > 0$, c'est-à-dire $j \in E_{\overline{k+1}}$. Si $j \in E_{\overline{k+1}}$, alors par définition, il existe $n \geq 0$ tel que $[Q^{k+1+nd}](i_0, j) > 0$ et (I.47) entraîne qu'il existe $i' \in E_k$ tel que $p(i', j) > 0$. On a donc montré que si $Q$ est irréductible $d$-périodique, il existe une partition de $E$ en $d$ sous-ensembles $E_0, \dots, E_{d-1}$ qui satisfont (I.46).

Montrons la réciproque ainsi que le dernier point du théorème. Supposons que $E$ est partitionné en $d-1$ ensembles $E_0, \dots, E_{d-1}$ qui satisfont (I.46). On fixe $i \in E_0$. Il est facile de montrer par récurrence que si $[Q^n](i,j) > 0$, alors $j \in E_{\overline{n}}$. Nous laissons les détails au lecteur. On en déduit que la période de $i$, notée $d_i$, divise $d$. Comme $Q$ est irréductible, on a montré que tous les états ont même période, notée $d'$. On a donc montré que $d' | d$.

Montrons ensuite que $Q^d_{|E_0}$ est une matrice de transition qui est irréductible : on se donne $i \in E_0$ et $j \in E$. On a montré que si $[Q^d](i,j) > 0$, alors $j \in E_0$. On en déduit que

$$\forall i \in E_0 \; , \quad \sum_{j \in E_0} [Q^d](i,j) = \sum_{j \in E} [Q^d](i,j) = 1 \; . \quad \text{(I.48)}$$

Cela montre que $Q^d_{|E_0}$ est une matrice de transition.

On fixe $i, j \in E_0$. Comme $Q$ est irréductible, il existe $n \geq 0$, tel que $[Q^n](i,j) > 0$. Soit $q$ le quotient de $n$ par $d$ : $n = qd + r$, $0 \leq r \leq d-1$. On observe que

$$[Q^n](i,j) = \sum_{i_1, \dots, i_q} [Q^d](i, i_1) \dots [Q^d](i_{q-1}, i_q) [Q^r](i_q, j) \; .$$

Puisque $[Q^n](i,j) > 0$, il existe $i_1, \dots, i_q \in E$ tels que $[Q^d](i,i_1) > 0$, ..., $[Q^d](i_{q-1}, i_q) > 0$ et $[Q^r](i_q, j) > 0$. Par conséquent $i_1, \dots, i_q \in E_0$ et puisque $j \in E_0$, on a $r = 0$. On a donc prouvé que

$$\forall i,j \in E_0 \, , \forall n \geq 1 \; , \quad [Q^n](i,j) > 0 \Longrightarrow d|n \;\text{ et }\; \exists q \geq 1 \; , \; [(Q^d)^q](i,j) > 0 \; , \tag{I.49}$$

ce qui implique bien que $Q^d_{|E_0}$ est irréductible. Rappelons que $d'$ désigne la période de $Q$ et que $d'|d$. On fixe $i \in E_0$. Il existe $n_1$, ..., $n_\ell$ tels que

$$[Q^{n_1}](i,i), \dots, [Q^{n_\ell}](i,i) > 0 \;\text{ et }\; \mathrm{PGCD}\{n_1 \dots, n_\ell\} = d' \; .$$

Par (I.49), on voit que $d$ divise $n_1$, ..., $n_\ell$. Par conséquent $d|d'$, et donc $d = d'$. On a prouvé que $Q$ est $d$-périodique. Ce qui implique que $Q^d_{|E_0}$ est apériodique. On prouve un résultat similaire pour $Q^d_{|E_k}$, en réindexant les ensembles. Cela termine la preuve du théorème. ■

Si on suppose que $Q$ est irréductible de période $d$ et si $E_0, \dots, E_{d-1}$ sont comme dans le théorème précédent, alors il est facile de montrer que pour tout $0 \leq k \leq d-1$ et pour tout $i \in E_k$, on a

$$\mathbf{P}_i \text{ - p.s.} \quad X_n \in E_{\overline{k+n}} \, , \quad n \geq 0 \; .$$

### I.2.c Fonctions harmoniques.

**Définition I.2.20** Soit $Q$ une matrice de transition sur $E$. Soit $f : E \to \mathbb{R}_+$. Soit $F \subset E$, un sous ensemble non-vide de $E$.

— On dit que $f$ est *Q-harmonique* <u>sur $F$</u> si $(Q.f)(i) = f(i)$, pour tout $i \in F$.

— On dit que $f$ est *Q-sur-harmonique* <u>sur $F$</u> si $(Q.f)(i) \leq f(i)$, pour tout $i \in F$.

— On dit que $f$ *est Q-sous-harmonique* <u>sur $F$</u> si $(Q.f)(i) \geq f(i)$, pour tout $i \in F$.

— Enfin, on dit que $f$ est (sur/sous)-harmonique, si elle est (sur/sous)-harmonique sur $E$ tout entier. □

**Proposition I.2.21** *Soit $f : E \to \mathbb{R}_+$. Soit $F \subset E$, non-vide. On introduit le temps aléatoire $T = \inf\{n \in \mathbb{N} : X_n \notin F\}$, avec la convention que $T = \infty$ ssi pour tout $n \in \mathbb{N}$, on a $X_n \in F$. Alors, $T$ est un $(\mathscr{F}_n)_{n\geq 0}$-temps d'arrêt et il y a équivalence entre les deux assertions suivantes.*

*(i) Pour toute $\mu \in \mathscr{M}_1(E)$, sous $\mathbf{P}_\mu$, $(f(X_{n\wedge T}))_{n\geq 0}$ est une (sur/sous)-martingale positive relativement à $(\mathscr{F}_n)_{n\geq 0}$.*

*(ii) $f$ est (sur/sous)-harmonique sur $F$.*

**Preuve :** $T$ est un $(\mathscr{F}_n)_{n\geq 0}$-temps d'arrêt car c'est le premier temps d'atteinte de $E\backslash F$, et car la suite $(X_n)_{n\geq 0}$ est $(\mathscr{F}_n)_{n\geq 0}$-adaptée. On traite le cas harmonique, les autres cas se prouvant de la même manière. On suppose $(i)$. On fixe $i \in F$. On remarque que $\mathbf{P}_i$-p.s. $T \geq 1$ et donc $X_{1\wedge T} = X_1$. Donc

$$f(i) = \mathbf{E}_i[f(X_0)] = \mathbf{E}_i[f(X_{0\wedge T})] = \mathbf{E}_i[f(X_{1\wedge T})] = \mathbf{E}_i[f(X_1)] = (Q.f)(i) \, ,$$

qui entraîne bien $(ii)$.

Réciproquement, on suppose $(ii)$ et on remarque que

$$f(X_{(n+1)\wedge T}) = \mathbf{1}_{\{T\leq n\}} f(X_T) + \mathbf{1}_{\{T>n\}} f(X_{n+1}) \; .$$

Or $\mathbf{1}_{\{T\leq n\}} f(X_T) = \sum_{0\leq k\leq n} \mathbf{1}_{\{T=k\}} f(X_k)$, qui est clairement $\mathscr{F}_n$-mesurable. De plus $\{T > n\} \in \mathscr{F}_n$ car $T$ est un temps d'arrêt. La propriété de Markov simple au temps $n$ implique alors

$$\begin{aligned} \mathbf{P}_\mu\text{-p.s.} \quad \mathbf{E}_\mu\big[f(X_{(n+1)\wedge T})|\mathscr{F}_n\big] &= \mathbf{1}_{\{T\leq n\}} f(X_T) + \mathbf{1}_{\{T>n\}} \mathbf{E}_{X_n}\big[f(X_1)\big] \\ &= \mathbf{1}_{\{T\leq n\}} f(X_T) + \mathbf{1}_{\{T>n\}} (Q.f)(X_n). \end{aligned}$$

Or si $T > n$, $X_n \in F$ et on a $(Q.f)(X_n) = f(X_n)$. Donc

$$\mathbf{P}_\mu\text{-p.s.} \quad \mathbf{E}_\mu\big[f(X_{(n+1)\wedge T})|\mathscr{F}_n\big] = \mathbf{1}_{\{T\leq n\}} f(X_T) + \mathbf{1}_{\{T>n\}} f(X_n) = f(X_{n\wedge T}) \ ,$$

ce qui montre $(i)$. ■

**Exemple I.2.22** Soit $Q = (p(i,j))_{i,j\in\mathbb{N}}$, la matrice de transition d'un processus de naissance et de mort, comme défini à l'exemple I.1.6, page 5, c'est-à-dire que $p(i,j) = 0$, dès que $|i-j| \geq 2$. On a vu à l'exemple I.2.15, page 41, que $Q$ est irréductible ssi

$$\forall n \geq 0\,, \quad p(n+1,n)p(n,n+1) > 0 \ .$$

*On suppose que $Q$ est irréductible et on se propose de trouver les fonctions $Q$-harmoniques sur* $\mathbb{N}^* = \{1, 2, \ldots\}$. Supposons que $f : \mathbb{N} \to \mathbb{R}_+$ soit une telle fonction, c'est-à-dire que $(Q.f)(n) = f(n)$, pour tout $n \in \mathbb{N}^*$. Cela est équivalent à dire

$$f(n) = p(n,n-1)f(n-1) + p(n,n)f(n) + p(n,n+1)f(n+1)\,, \quad n \geq 1 \ .$$

Comme $Q$ est une matrice de transition, on a $p(n,n) = 1 - p(n,n-1) - p(n,n+1)$. Le système d'équations précédent est donc équivalent à

$$f(n+1) - f(n) = \tfrac{p(n,n-1)}{p(n,n+1)}\big(f(n) - f(n-1)\big)\ , \quad n \geq 1\ ,$$

ce qui est équivalent à

$$f(n+1) - f(n) = \big(f(1) - f(0)\big) \prod_{1\leq k\leq n} \tfrac{p(k,k-1)}{p(k,k+1)}\ ,$$

ce qui est encore équivalent à

$$f(n+1) = f(0) + \big(f(1) - f(0)\big)\Big[1 + \sum_{1\leq \ell\leq n} \prod_{1\leq k\leq \ell} \tfrac{p(k,k-1)}{p(k,k+1)}\Big]\ , \quad n \geq 1\ .$$

On voit que $f(0)$ et $f(1)$ peuvent être fixés arbitrairement. On remarque, en raisonnant de la même façon, que les fonctions $Q$-harmoniques sur $\mathbb{N}$ sont nécessairement constantes. □

Donnons deux exemples importants. Soit $F \subset E$, non-vide. On adopte la convention $\inf \emptyset = \infty$ et on pose

$$T_F = \inf\{n \in \mathbb{N} \ : \ X_n \in F\,\} \quad \text{et} \quad T_F^{(1)} = \inf\{n \geq 1 \ : \ X_n \in F\,\}\ ,$$

On a donc $T_F = \infty$ (resp. $T_F^{(1)} = \infty$) ssi $X_n \notin F$ pour tout $n \in \mathbb{N}$ (resp. pour tout $n \in \mathbb{N}^*$). Ce sont des $(\mathscr{F}_n)_{n\geq 0}$-temps d'arrêt : $T_F$ est appelé le premier temps d'atteinte de la partie $F$ et $T_F^{(1)}$ est appelé premier temps de retour en $F$. On observe que si $X_0 \notin F$, alors $T_F = T_F^{(1)}$.

**Lemme I.2.23** *Avec les notations précédentes, on pose*

$$\forall i \in E\,, \quad f(i) = \mathbf{P}_i(T_F < \infty) \quad \text{et} \quad f^*(i) = \mathbf{P}_i\big(T_F^{(1)} < \infty\big)\ .$$

*Alors $f$ et $f^*$ sont des fonctions sur-harmoniques : $f \geq Q.f$ et $f^* \geq Q.f^*$. De plus, ces fonctions prennent leurs valeurs dans $[0,1]$ et on a $f(i) = f^*(i)$, pour tout $i \notin F$.*

**Preuve :** si $i \in F$, alors $f(i) = 1$. Comme $f(j) \leq 1$, pour tout $j \in E$, on a bien

$$\forall i \in F\,, \quad (Q.f)(i) = \sum_{j \in E} p(i,j) f(j) \leq \sum_{j \in E} p(i,j) = 1 = f(i)\ .$$

On suppose que $i \notin F$. On utilise la notation $T_F(\theta_1 \mathbf{X}) = \inf\{n \in \mathbb{N} : X_{n+1} \in F\}$ et on vérifie que $\{X_0 \notin F; T_F < \infty\} = \{X_0 \notin F;\ T_F(\theta_1 \mathbf{X}) < \infty\}$. La propriété de Markov et le fait que $i \notin F$ entraînent que

$$\mathbf{P}_i\text{-p.s.} \quad \mathbf{E}_i\big[\mathbf{1}_{\{T_F < \infty\}} | \mathscr{F}_1\big] = \mathbf{P}_{X_1}(T_F < \infty) = f(X_1)$$

et en intégrant, $f(i) = \mathbf{P}_i(T_F < \infty) = \mathbf{E}_i[f(X_1)] = (Q.f)(i)$, pour tout $i \notin F$, ce qui implique le résultat voulu pour $f$.

On introduit la notation $T_F^{(1)}(\theta_1 \mathbf{X}) = \inf\{n \geq 1 : X_{n+1} \in F\}$, et on vérifie que que $\mathbf{1}_{\{T_F^{(1)} < \infty\}} = \mathbf{1}_{\{X_1 \in F\}} + \mathbf{1}_{\{X_1 \notin F\}} \mathbf{1}_{\{T_F^{(1)}(\theta_1 \mathbf{X}) < \infty\}}$. La propriété de Markov au temps 1, implique que $\mathbf{P}_i$-p.s.

$$\begin{aligned} \mathbf{E}_i\big[\mathbf{1}_{\{T_F^{(1)} < \infty\}} | \mathscr{F}_1\big] &= \mathbf{1}_{\{X_1 \in F\}} + \mathbf{1}_{\{X_1 \notin F\}} \mathbf{P}_{X_1}\big(T_F^{(1)} < \infty\big) \\ &= \mathbf{1}_{\{X_1 \in F\}} + \mathbf{1}_{\{X_1 \notin F\}} f^*(X_1) \\ &\geq f^*(X_1)\,, \end{aligned}$$

car $0 \leq f^* \leq 1$. En intégrant, on a $f^*(i) = \mathbf{P}_i(T_F^{(1)} < \infty) \geq \mathbf{E}_i[f^*(X_1)] = (Q.f^*)(i)$, pour tout $i \in E$, ce qui termine la preuve. ■

### I.2.d Mesure invariantes, chaînes réversibles.

**Définition I.2.24** Soit $Q = (p(i,j))_{i,j \in E}$, une matrice de transition et soit $\mu$, une mesure positive sur $E$. C'est une *$Q$-mesure invariante* si elle satisfait les conditions suivantes.

(a) $\mu$ n'est pas identiquement nulle, c'est-à-dire qu'il existe $i \in E$ tel que $\mu(i) \neq 0$.

(b) Pour tout $i \in E$, on a $\mu(i) < \infty$.

(c) On a $\mu = \mu.Q$, c'est-à-dire $\mu(j) = \sum_{i \in E} \mu(i) p(i,j)$, pour tout $j \in E$.

La mesure $\mu$ est dite *$Q$-excessive* si elle satisfait (a), (b) et ($\gamma$), où la condition ($\gamma$) est donnée comme suit.

($\gamma$) On a $\mu \geq \mu.Q$, c'est-à-dire $\mu(j) \geq \sum_{i \in E} \mu(i) p(i,j)$, pour tout $j \in E$.

Lorsque la mesure $\mu$ est $Q$-invariante et de masse totale 1, on parle plutôt de *loi $Q$-invariante* ou de *loi $Q$-stationnaire* ou encore de *loi d'équilibre*. (ce sont des synonymes). □

Justifions le terme de loi invariante : on suppose que $\pi$ est une loi $Q$-invariante. Le lemme I.2.1 montre que la loi de $X_n$ sous $\mathbf{P}_\pi$ est $\pi.Q^n$. Or $\pi.Q^n = \pi$. Donc, sous $\mathbf{P}_\pi$, toutes les variables $X_n$ ont même loi $\pi$.

**Remarque I.2.25** Si $\mu$ est $Q$-invariante, il en est de même pour $c.\mu$, où $c$ est n'importe quel nombre réel strictement positif. Par conséquent, si $Q$ admet un mesure invariante $\mu$, de masse finie la loi $\pi$ donnée par $\pi(i) = \mu(i)/\langle\mu\rangle$, pour tout $i \in E$, est une loi invariante. □

**Exemple I.2.26** Soit $Q = (p(i,j))_{i,j\in\mathbb{N}}$, la matrice de transition d'un processus de naissance et de mort, comme défini à l'exemple I.1.6, c'est-à-dire que $p(i,j) = 0$ dès que $|i-j| \geq 2$. On a vu à l'exemple I.2.15 (page 41) que $Q$ est irréductible ssi

$$\forall n \geq 0\,, \quad p(n+1,n)p(n,n+1) > 0\ .$$

On suppose que $Q$ est irréductible et *on se propose de trouver les mesures $Q$-invariantes*. Une mesure positive est $Q$-invariante ssi $\mu.Q = \mu$, c'est-à-dire ssi pour tout $n \geq 1$, on a

$$\begin{cases} \mu(0) &=& \mu(1)p(1,0) + \mu(0)p(0,0) \\ \mu(n) &=& \mu(n+1)p(n+1,n) + \mu(n)p(n,n) + \mu(n-1)p(n-1,n) \end{cases}$$

On voit que si $\mu(0)$ est fixé, il n'y a qu'une seule solution à ce système d'équations : en effet, on vérifie que

$$\mu(n) = \mu(0) \prod_{1\leq k\leq n} \tfrac{p(k-1,k)}{p(k,k-1)}\,, \quad n \geq 1\,,$$

ce qui est équivalent à dire que $\mu(n+1)p(n+1,n) = \mu(n)p(n,n+1)$, pour tout $n \in \mathbb{N}$.

Un processus de vie et de mort n'admet donc qu'une seule mesure $Q$-invariante, à une constante multiplicative près. Cela implique le résultat suivant : on pose

$$S = 1 + \sum_{n\geq 1} \prod_{1\leq k\leq n} \tfrac{p(k-1,k)}{p(k,k-1)} \ \in\, ]1,\infty]\ .$$

— Si $S < \infty$, alors il existe une *unique mesure de probabilité $Q$-invariante* donnée par

$$\mu(0) = \tfrac{1}{S} \quad \text{et} \quad \mu(n) = \tfrac{1}{S} \prod_{1\leq k\leq n} \tfrac{p(k-1,k)}{p(k,k-1)}\,, \quad n \geq 1\ .$$

— Si $S = \infty$, alors toutes les mesures $Q$-invariantes sont proportionnelles entre elles et donc de masse totale infinie. Il n'y a pas de probabilité $Q$-invariante. □

Mentionnons que certaines matrices de transition (qui peuvent être irréductibles) n'admettent pas de mesure invariante. Le problème consistant à trouver une mesure invariante n'est, en général, pas simple. Il y a cependant des chaînes dont la mesure invariante est simple à trouver : ce sont les chaînes *réversibles*.

**Définition I.2.27** Soit $Q = (p(i,j))_{i,j\in E}$, une matrice de transition. On dit qu'elle est *réversible* s'il existe une mesure positive $\mu$ sur $E$ telle que

$$\forall i,j \in E\,, \quad 0 < \mu(i) < \infty \quad \text{et} \quad \mu(i)p(i,j) = \mu(j)p(j,i)\ .$$

On appelle une telle mesure $\mu$, une *mesure de réversibilité de $Q$*. □

**Lemme I.2.28** *Soit $Q$ une matrice de transition réversible et on note $\mu$ une mesure de réversibilité de $Q$. Alors $\mu$ est $Q$-invariante.*

**Preuve :** pour tout $j \in E$, on a $\sum_{i \in E} \mu(i)p(i,j) = \sum_{i \in E} \mu(j)p(j,i) = \mu(j)\sum_{i \in E} p(j,i) = \mu(j)$, ce qui montre le lemme. ■

La proposition suivante explique le terme "réversible".

**Proposition I.2.29** *Soit $Q = (p(i,j))_{i,j \in E}$ une matrice de transition qui admet une mesure invariante notée $\mu$. On suppose que $\mu(i) > 0$, pour tout $i \in E$. On définit $Q_* = (p^*(i,j))_{i,j \in E}$ par*

$$\forall i, j \in E\,, \quad p^*(i,j) = p(j,i)\mu(j)/\mu(i)\ .$$

*Alors, les assertions suivantes sont vérifiées.*

*(i) $Q_*$ est une matrice de transition sur $E$ (on l'appelle matrice de transition duale de $Q$). Elle admet $\mu$ comme mesure invariante. De plus, $Q_* = Q$ ssi $Q$ est réversible de mesure de réversibilité $\mu$.*

*(ii) On suppose $\mu$ est une probabilité et que, sous $\mathbf{P}$, $(X_n)_{n \geq 0}$ et $(X^*_n)_{n \geq 0}$ sont deux chaînes de Markov de loi d'entrée $\mu$ et de matrices de transition respectives $Q$ et $Q_*$. Alors pour tout $n \in \mathbb{N}$,*

$$\big(X^*_0, X^*_1, \ldots, X^*_{n-1}, X^*_n\big) \overset{\text{(loi)}}{=} (X_n, X_{n-1}, \ldots, X_1, X_0)\ . \tag{I.50}$$

*Par conséquent, $Q$ est réversible de mesure de réversibilité $\mu$, ssi, lorsque $X_0$ est de loi $\mu$, on a*

$$\big(X_0, X_1, \ldots, X_{n-1}, X_n\big) \overset{\text{(loi)}}{=} (X_n, X_{n-1}, \ldots, X_1, X_0)\ . \tag{I.51}$$

**Preuve :** puisque $\mu$ est une loi $Q$-invariante, on voit que pour tout $i \in E$, on a

$$\sum_{j \in E} p^*(i,j) = \tfrac{1}{\mu(i)} \sum_{j \in E} \mu(j)p(j,i) = \mu(i)/\mu(i) = 1\ ,$$

ce qui montre que $Q_*$ est une matrice de transition. On observe ensuite que pour tout $j \in E$, on a

$$\sum_{i \in E} \mu(i)p^*(i,j) = \mu(j)\sum_{i \in E} p(j,i) = \mu(j)\ ,$$

ce qui montre que $\mu$ est une mesure $Q_*$-invariante. Il est par ailleurs immédiat de vérifier que $Q$ est $\mu$-réversible ssi $Q_* = Q$. Cela prouve $(i)$.

On fixe ensuite $n \in \mathbb{N}$ et $i_0, \ldots, i_n \in E$. On a

$$\begin{aligned}
\mathbf{P}(X^*_0 = i_0, \ldots; X^*_n = i_n) &= \mu(i_0) \prod_{0 \leq k < n} p^*(i_k, i_{k+1}) = \mu(i_0) \prod_{0 \leq k < n} \tfrac{\mu(i_{k+1})}{\mu(i_k)} p(i_{k+1}, i_k) \\
&= \mu(i_n) \prod_{0 \leq k < n} p(i_{n-k}, i_{n-k-1}) \\
&= \mathbf{P}(X_0 = i_n; X_1 = i_{n-1}; \ldots; X_n = i_0)\,,
\end{aligned}$$

ce qui implique (I.50). On voit donc que (I.51) a lieu ssi $Q = Q_*$, et on conclut. ■

**Exemple I.2.30** (*Les marches sur les graphes pondérés sont exactement les chaînes réversibles*) On se donne $\mathbf{G} = (\mathbf{S}, \mathbf{A})$, un graphe non-orienté, simple (pas d'arêtes multiples) et bien sûr, dénombrable (c'est-à-dire que

$\mathbf{S}$ est dénombrable). On le munit d'un système de poids $\mathbf{C} = (C_a; a \in \mathbf{A})$ (on rappelle que $0 < C_a < \infty$, pour toute arête $a \in \mathbf{A}$). On suppose toujours que

$$\forall s \in \mathbf{S}\,, \quad \pi(s) = \sum_{s' \sim s} C_{\{s,s'\}} \ \in \ \mathbb{R}_+^* \ .$$

On note $Q = (p(s, s'))_{s,s' \in \mathbf{S}}$, la matrice de transition de la marche sur $\mathbf{G}$ associée aux poids $\mathbf{C}$, comme défini à l'exemple I.1.7, page 5 : on rappelle que $p(s, s') = C_{\{s,s'\}}/\pi(s)$ si $s \sim s'$ et $p(s, s') = 0$ sinon. On voit alors que

$$\forall s, s' \in \mathbf{S} \text{ tels que } s \sim s', \quad \pi(s)p(s, s') = C_{\{s,s'\}} = C_{\{s',s\}} = \pi(s')p(s', s) \ . \tag{I.52}$$

Si $s$ et $s'$ ne sont pas voisins alors $p(s, s') = p(s', s) = 0$. Cela implique donc dans tous les cas que

$$\forall s, s' \in \mathbf{S}\,, \quad \pi(s)p(s, s') = \pi(s')p(s', s) \ .$$

On voit donc que $\pi$ est une mesure de réversibilité pour $Q$ et donc que la matrice de transition $Q$ est réversible. Cela implique par la même occasion que $\pi$ est une mesure $Q$-invariante. Si on a une marche simple sur le graphe $\mathbf{G} = (\mathbf{S}, \mathbf{A})$, alors $\pi(s) = \deg(s)$ est une mesure $Q$-invariante.

*Réciproquement.* Soit $Q$ une matrice de transition sur $E$ qui est rendue réversible par la mesure $\mu$, c'est-à-dire que $\mu(i)p(i, j) = \mu(j)p(j, i)$, pour tous $i, j \in E$. On rappelle que la mesure rendant $Q$ réversible doit satisfaire $0 < \mu(i) < \infty$, pour tout $i \in E$. On voit donc que si $p(i, j) > 0$, alors $p(j, i) > 0$. Bien qu'on ait le choix pour définir le graphe pondéré correspondant à la matrice réversible $Q$, on prend le plus *économique* qui est donné par

$$\mathbf{S} = E \ , \ \mathbf{A} = \big\{ \{i, j\}\,;\, i, j \in E \ : \ p(i, j) > 0 \,\big\} \quad \text{et} \quad C_{\{i,j\}} = \mu(i)p(i, j)\,, \ \{i, j\} \in \mathbf{A} \ .$$

On voit facilement que la matrice de transition de la marche aléatoire sur les graphe pondéré ainsi défini, est *exactement* la matrice réversible $Q$. □

Examinons maintenant le lien entre loi invariante et période.

**Proposition I.2.31** *Soit $Q = (p(i, j))_{i,j \in E}$, une matrice de transition que l'on suppose* **irréductible et $d$-périodique avec** $d \geq 2$. *Soit $E_0$, ... , $E_{d-1}$ une partition de $E$ en $d$ sous-ensembles non-vides qui satisfont la condition (I.46) de la proposition I.2.19. On suppose que $Q$ admet une loi invariante notée $\pi$. Alors les assertions suivantes sont vérifiées.*

*(i) On a $\pi(E_0) = \ldots = \pi(E_{d-1}) = 1/d$.*

*(ii) Pour tout $0 \leq k \leq d-1$, on pose $\pi_k = (d\pi(i); i \in E_k)$. Alors $\pi_k$ est une loi invariante de la matrice de transition $Q^d_{|E_k}$.*

*Réciproquement, si on se donne $\pi_0 = (\pi_0(i), i \in E_0)$, une loi $Q^d_{|E_0}$-invariante, on définit, $\pi_k = (\pi_k(j), j \in E_k)$, pour tout $1 \leq k \leq d-1$, par*

$$\forall j \in E_k \ , \quad \pi_k(j) := \sum_{i \in E_0} \pi_0(i)\, [Q^k](i, j) \ . \tag{I.53}$$

*Alors, $\pi_k$ est une loi $Q^d_{|E_k}$-invariante. De plus $\pi = d^{-1}\pi_0 + \ldots + d^{-1}\pi_{d-1}$ est une loi $Q$-invariante.*

**Preuve :** nous laissons la preuve de résultat en exercice. ■

### I.2.e Constructions de chaînes stationnaires dont le temps est indexé par $\mathbb{Z}$.

Dans cette section on étend le résultat de la proposition I.2.29 $(ii)$ page 48 à tous les temps $n$ variant dans $\mathbb{Z}$. Soit $Q = (p(i,j))_{i,j\in E}$, une matrice de transition sur un espace d'états $E$ dénombrable discret. Soit $\mu$ une mesure de probabilité telle que.

$$\forall i \in E, \quad \mu(i) > 0 \quad \text{et } \mu \text{ est } Q\text{-invariante : } \quad \mu.Q = \mu. \tag{I.54}$$

On introduit la matrice duale $Q_* = (p^*(i,j))_{i,j\in E}$ comme à la proposition I.2.29 :

$$\forall i,j \in E\,, \quad p^*(i,j) = p(j,i)\mu(j)/\mu(i)\ . \tag{I.55}$$

La proposition I.2.29 montre que c'est une matrice de transition dont $\mu$ est une mesure invariante.

Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité que l'on suppose suffisamment riche pour qu'y soit définie une v.a. $U : \Omega \to [0,1]$ $\mathscr{F}$-mesurable de loi uniforme.

On considère des suites de v.a. indexées par $\mathbb{Z}$ et à valeurs dans $E$. Pour cela on note $E^{\mathbb{Z}}$ l'espace des suites à valeurs dans $E$ et indexées par $\mathbb{Z}$, espace que l'on munit de la *tribu produit*, notée $\mathscr{P}(E)^{\otimes\mathbb{Z}}$ et définie comme la tribu engendrée par les *cylindres élémentaires* de $E^{\mathbb{Z}}$ qui sont les ensembles de la forme

$$C = \big\{ \mathbf{x} = (x_n)_{n\in\mathbb{Z}} \in E^{\mathbb{Z}} : \forall k \in [\![m,n]\!] : x_k = i_k \big\}$$

où $m,n \in \mathbb{Z}$ sont tels que $m \leq n$, où $[\![m,n]\!]$ désigne $\mathbb{Z} \cap [m,n]$ et où $i_k \in E$, $k \in [\![m,n]\!]$. On note $\mathscr{C}$ l'ensemble constitué des cylindres élémentaires de $E^{\mathbb{Z}}$, de l'ensemble $E^{\mathbb{Z}}$ lui-même et de l'ensemble vide : on remarque que $\mathscr{C}$ est un pi-système. On a donc

$$\mathscr{P}(E)^{\otimes\mathbb{Z}} := \sigma(\mathscr{C})\ .$$

On raisonne comme dans les énoncés de la section I.1.d (page 24) pour montrer que *si pour tout $n \in \mathbb{Z}$, on se donne des fonctions $X_n : \Omega \to E$, alors il y a équivalence entre les deux assertions suivantes.*

- *Pour tout $n \in \mathbb{Z}$, $X_n$ est $\mathscr{F}$-mesurable.*
- *$\mathbf{X} = (X_n)_{n\in\mathbb{Z}} : \Omega \to E^{\mathbb{Z}}$ est $(\mathscr{F}, \mathscr{P}(E)^{\otimes\mathbb{Z}})$-mesurable.*

De même : soit $Y_n$, $n \in \mathbb{Z}$, une autre suite de v.a. à valeurs dans $E$ et $\mathscr{F}$-mesurables ; alors *il y a équivalence entre les deux assertions suivantes.*

- *Sous $\mathbf{P}$, $(X_n)_{n\in\mathbb{Z}}$ et $(Y_n)_{n\in\mathbb{Z}}$ ont même loi.*
- *Pour tous $m,n \in \mathbb{Z}$ tels que $m \leq n$ et tous $i_k \in E$, $k \in [\![m,n]\!]$,*

$$\mathbf{P}\big(\forall k \in [\![m,n]\!] : X_k = i_k\big) = \mathbf{P}\big(\forall k \in [\![m,n]\!] : Y_k = i_k\big)$$

Nous laissons la preuve de ces faits en exercice.

**Proposition I.2.32** *On reprend les notations et les hypothèses ci-dessus. Alors, il existe une suite de v.a. $X_n : \Omega \to E$, $n \in \mathbb{Z}$ qui satisfait les propriétés suivantes pour tout $n_0 \in \mathbb{Z}$.*

$(i)$ *$(X_{n_0+n})_{n\in\mathbb{N}}$ est une chaîne de Markov de loi initiale $\mu$ et de matrice de transition $Q$.*

$(ii)$ *$(X_{n_0-n})_{n\in\mathbb{N}}$ est une chaîne de Markov de loi initiale $\mu$ et de matrice de transition $Q_*$.*

$(iii)$ *Pour tout $i \in E$, sous $\mathbf{P}(\,\cdot\,|X_{n_0} = i)$, $(X_{n_0+n})_{n\in\mathbb{N}}$ et $(X_{n_0-n})_{n\in\mathbb{N}}$ sont indépendantes.*

$(iv)$ *$(X_{n_0+n})_{n\in\mathbb{Z}}$ et $(X_n)_{n\in\mathbb{Z}}$ ont même loi sous $\mathbf{P}$.*

*En particulier, pour tout $n \in \mathbb{Z}$, $X_n$ a pour loi $\mu$, c'est pourquoi $(X_n)_{n\in\mathbb{Z}}$ est appelée* chaîne stationnaire.

**Remarque I.2.33** Si $\mu$ est une mesure de réversibilité pour $Q$, alors $Q = Q_*$ et $(X_n)_{n\in\mathbb{Z}}$ et $(X_{-n})_{n\in\mathbb{Z}}$ ont même loi. $\square$

**Preuve de la proposition I.2.32.** On a supposé qu'était définie sur $(\Omega, \mathscr{F}, \mathbf{P})$ une v.a. $U$ de loi uniforme. À partir de $U$, on obtient facilement deux suites indépendantes de v.a. uniformes sur $[0,1]$, notées $U_n$ et $U_n^*$, $n \in \mathbb{N}$, comme cela est expliqué dans l'exrcice I.13 23. On fixe les fonctions suivantes.

– Une fonction d'échantillonnage $\phi_\mu : [0,1] \to E$ de la loi $\mu$, c'est-à-dire que $\mathbf{P}(\phi_\mu(U_0) = i) = \mu(i)$, $i \in E$.

– Deux fonctions $\Phi_Q$ et $\Phi_{Q_*}$ : $[0,1] \times E \to E$, d'échantillonnage des matrices de transition rspectives $Q$ et $Q_*$, c'est-à-dire que $\mathbf{P}(\Phi_Q(U_0, i) = j) = p(i,j)$ et $\mathbf{P}(\Phi_{Q_*}(U_0, i) = j) = p^*(i,j)$ pour tous $i, j \in E$.

On définit alors par récurrence deux suites de v.a. $X_n$ et $X_n^* : \Omega \to E$, $n \in \mathbb{N}$, en posant :

$$X_0 = X_0^* = \phi_\mu(U_0) \quad \text{et} \quad \forall n \in \mathbb{N}, \quad X_{n+1} = \Phi_Q(U_{n+1}, X_n) \quad \text{et} \quad X_{n+1}^* = \Phi_{Q*}(U_{n+1}^*, X_n). \tag{I.56}$$

On observe d'abord que $U_0^*$ n'est pas utilisée dans la construction. La proposition I.1.25 (page 23) permet ensuite d'affirmer que $(X_n)_{n\geq 0}$ et $(X_n^*)_{n\geq 0}$ sont deux chaînes de Markov de matrices de transition respectives $Q$ et $Q_*$. On pose alors

$$\forall n \in \mathbb{N}, \quad X_{-n} := X_n^* \, .$$

On vérifie ensuite que la suite de v.a. $(X_n)_{n\in\mathbb{Z}}$ que l'on a contruite vérifie bien les propriétés $(i)$-$(iv)$ de l'énoncé.

On voit que la construction implique que les propriétés $(i)$ et $(ii)$ sont vraies pour $n_0 = 0$. Ensuite sous $\mathbf{P}(\,\cdot\,|X_0 = i)$, on observe que $(X_n)_{n\in\mathbb{N}}$ est une fonction déterministe de $(U_n)_{n\geq 1}$ et que $(X_n^*)_{n\in\mathbb{N}}$ est une fonction déterministe de $(U_n^*)_{n\geq 1}$. Comme $(U_n)_{n\geq 1}$ et $(U_n^*)_{n\geq 1}$ sont indépendantes, cela montre la propriété $(iii)$ lorsque $n_0 = 0$.

On fixe $n_0 \in \mathbb{Z}$ et on montre ensuite que la propriété $(iv)$ pour $n_0$ implique $(i)$, $(ii)$ et $(iii)$ pour $n_0$. Pour simplifier, on pose $X'_n = X_{n_0+n}$, $n \in \mathbb{Z}$. La propriété $(iv)$ signifie donc que $(X'_n)_{n\in\mathbb{Z}}$ et $(X_n)_{n\in\mathbb{Z}}$ ont même loi sous $\mathbf{P}$. En particulier $(X'_n)_{n\in\mathbb{N}}$ et $(X_n)_{n\in\mathbb{N}}$ ont même loi sous $\mathbf{P}$, c'est-à-dire que ce sont deux chaînes de Markov de matrice de transition $Q$ et de loi intiale $\mu$, ce qui montre $(i)$ pour $n_0$. De même $(X'_{-n})_{n\in\mathbb{N}}$ et $(X_{-n})_{n\in\mathbb{N}}$ ont même loi sous $\mathbf{P}$, c'est-à-dire que ce sont des chaînes de Markov de matrice de transition $Q_*$ et de loi intiale $\mu$, ce qui montre $(ii)$ pour $n_0$. Enfin, pour tout $i \in E$ et pour toutes fonctions $F, G : E^{\mathbb{N}} \to \mathbb{R}$ qui sont $\mathscr{P}(E)^{\otimes\mathbb{N}}$-mesurables bornées

$$\begin{aligned}
&\mathbf{E}\big[F\big((X_{n_0+n})_{n\in\mathbb{N}}\big)G\big((X_{n_0-n})_{n\in\mathbb{N}}\big)\big|X_{n_0} = i\big] \\
&\qquad = \mathbf{E}\big[F\big((X_{n_0+n})_{n\in\mathbb{N}}\big)G\big((X_{n_0-n})_{n\in\mathbb{N}}\big)\mathbf{1}_{\{X_{n_0}=i\}}\big]/\mathbf{P}(X_{n_0} = i) \\
&\qquad = \mathbf{E}\big[F\big((X'_n)_{n\in\mathbb{N}}\big)G\big((X'_{-n})_{n\in\mathbb{N}}\big)\mathbf{1}_{\{X'_0=i\}}\big]/\mathbf{P}(X'_0 = i) \\
&\qquad = \mathbf{E}\big[F\big((X_n)_{n\in\mathbb{N}}\big)G\big((X_{-n})_{n\in\mathbb{N}}\big)\mathbf{1}_{\{X_0=i\}}\big]/\mathbf{P}(X_0 = i) \\
&\qquad = \mathbf{E}\big[F\big((X_n)_{n\in\mathbb{N}}\big)G\big((X_{-n})_{n\in\mathbb{N}}\big)\big|X_0 = i\big] \\
&\qquad = \mathbf{E}\big[F\big((X_n)_{n\in\mathbb{N}}\big)\big|X_0 = i\big]\mathbf{E}\big[G\big((X_{-n})_{n\in\mathbb{N}}\big)\big|X_0 = i\big]
\end{aligned}$$

la dernière égalité venant de l'indépendance (prouvée) de $(X_n)_{n\in\mathbb{N}}$ et de $(X_{-n})_{n\in\mathbb{N}}$ sous $\mathbf{P}(\,\cdot\,|\,X_0 = i)$. Cela implique facilement que

$$\mathbf{E}\big[F\big((X_{n_0+n})_{n\in\mathbb{N}}\big)G\big((X_{n_0-n})_{n\in\mathbb{N}}\big)\big|X_{n_0} = i\big] = \mathbf{E}\big[F\big((X_{n_0+n})_{n\in\mathbb{N}}\big)\big|X_{n_0} = i\big]\mathbf{E}\big[G\big((X_{n_0-n})_{n\in\mathbb{N}}\big)\big|X_{n_0} = i\big]$$

et donc la propriété $(iii)$ pour $n_0$.

Il reste donc à prouver $(iv)$. Montrons qu'il suffit de prouver que $(X_n)_{n\in\mathbb{Z}}$ et $(X_{n+1})_{n\in\mathbb{Z}}$ ont même loi sous $\mathbf{P}$. En effet, pour tout $n_0$, on définit l'opérateur de décalage $\theta_{n_0} : E^{\mathbb{Z}} \to E^{\mathbb{Z}}$, en posant $\theta_{n_0}\mathbf{x} = (x_{n_0+n})_{n\in\mathbb{Z}}$

pour toute suite $\mathbf{x}=(x_n)_{n\in\mathbb{Z}}\in E^{\mathbb{Z}}$. On vérifie facilement que l'image réciproque d'un cylindre élémentaire par $\theta_{n_0}$ est un cylindre élémentaire et donc que $\theta_{n_0}$ est $(\mathscr{P}(E)^{\otimes\mathbb{Z}}, \mathscr{P}(E)^{\otimes\mathbb{Z}})$-mesurable. On vérifie immédiatemnt que $\theta_{n_0}\circ\theta_{n_1}=\theta_{n_0+n_1}$, pour tous $n_0, n_1\in\mathbb{Z}$. Si on montre que $\theta_1\mathbf{X}$ a même loi que $\mathbf{X}$ alors $\theta_2\mathbf{X}=\theta_1(\theta_1\mathbf{X}))$ a même loi que $\theta_1\mathbf{X}$, qui a même loi que $\mathbf{X}$ et on montre ainsi facilement par récurrence que pour tout $n_0\in\mathbb{N}$, $\theta_{n_0}\mathbf{X}$ a même loi que $\mathbf{X}$. Pour les décalage de temps négatifs l'argument est le même $\mathbf{X}=\theta_{-1}(\theta_1\mathbf{X}))$ a même loi que $\theta_{-1}\mathbf{X}$. Donc $\theta_{-2}\mathbf{X}=\theta_{-1}(\theta_{-1}\mathbf{X}))$ a même loi que $\theta_{-1}\mathbf{X}$ qui a même loi que $\mathbf{X}$, et on montre ainsi facilement par récurrence que pour tout $n_0\in\mathbb{N}$, $\theta_{-n_0}\mathbf{X}$ a même loi que $\mathbf{X}$.

Il reste donc à prouver que $(X_n)_{n\in\mathbb{Z}}$ et $(X_{n+1})_{n\in\mathbb{Z}}$ ont même loi sous $\mathbf{P}$. Pour cela on fixe $(i_k)_{k\in\mathbb{Z}}$ et pour tout $N\in\mathbb{N}^*$, on calcule tout d'abord $a_N=\mathbf{P}\big(\forall k\in[\![-N,N]\!] : X_{k+1}=i_k\big)$ explicitement pour montrer que cela vaut $\mathbf{P}\big(\forall k\in[\![-N,N]\!] : X_k=i_k\big)$. Par $(iii)$ pour $n_0=0$ (qui est prouvée et est une conséquence de la construction), on a

$$\begin{aligned}
a_N &= \mathbf{P}(X_0=i_{-1})\mathbf{P}\big((X_{-k})_{0\le k\le N-1}=(i_{-k-1})_{0\le k\le N-1}\,;\,(X_k)_{0\le k`N+1}=(i_{k-1})_{0\le k\le N}\,\big|\,X_0=i_{-1}\big)\\
&= \mu(i_{-1})\,\mathbf{P}\big((X^*_k)_{0\le k\le N-1}=(i_{-k-1})_{0\le k\le N-1}\,\big|\,X_0=i_{-1}\big)\,\mathbf{P}\big((X_k)_{0\le k\le N+1}=(i_{k-1})_{0\le k\le N+1}\,\big|\,X_0=i_{-1}\big)\\
&= \mu(i_{-1})\,p^*(i_{-1},i_{-2})\dots p^*(i_{-(N-1)},i_{-N})\,p(i_{-1},i_0)p(i_0,i_1)\dots p(i_{N-1},i_N)\\
&= \mu(i_{-1})\frac{p(i_{-2},i_{-1})\mu(i_{-2})}{\mu(i_{-1})}\dots\frac{p(i_{-N},i_{-(N-1)})\mu(i_{-N})}{\mu(i_{-(N-1)})}p(i_{-1},i_0)p(i_0,i_1)\dots p(i_{N-1},i_N)\\
&= \mu(i_0)\frac{p(i_{-1},i_0)\mu(i_{-1})}{\mu(i_0)}\frac{p(i_{-2},i_{-1})\mu(i_{-2})}{\mu(i_{-1})}\dots\frac{p(i_{-N},i_{-(N-1)})\mu(i_{-N})}{\mu(i_{-(N-1)})}\,p(i_0,i_1)\dots p(i_{N-1},i_N)\\
&= \mu(i_0)\,p^*(i_0,i_{-1})p^*(i_{-1},i_{-2})\dots p^*(i_{-(N-1)},i_{-N})\,p(i_0,i_1)\dots p(i_{N-1},i_N)\\
&= \mu(i_0)\,\mathbf{P}\big((X^*_k)_{0\le k\le N-1}=(i_{-k})_{0\le k\le N-1}\,\big|\,X_0=i_0\big)\,\mathbf{P}\big((X_k)_{0\le k\le N+1}=(i_k)_{0\le k\le N+1}\,\big|\,X_0=i_0\big)\\
&= \mathbf{P}(X_0=i_0)\mathbf{P}\big((X_{-k})_{0\le k\le N-1}=(i_{-k})_{0\le k\le N-1}\,;\,(X_k)_{0\le k`N+1}=(i_k)_{0\le k\le N}\,\big|\,X_0=i_0\big)\\
&= \mathbf{P}\big(\forall k\in[\![-N,N]\!] : X_k=i_k\big).
\end{aligned}$$

Cela prouve que $\mathbf{P}(\forall k\in[\![-N,N]\!]:X_{k+1}=i_k)=\mathbf{P}\big(\forall k\in[\![-N,N]\!]:X_k=i_k\big)$, pour toute suite $(i_k)_{k\in\mathbb{Z}}\in E^{\mathbb{Z}}$ et tout $N\in\mathbb{N}^*$. Cela entraîne facilement que pour tout cylindre élémentaire $C$ de $E^{\mathbb{Z}}$, $\mathbf{P}(\mathbf{X}\in C)=\mathbf{P}(\theta_1\mathbf{X}\in C)$ et donc que $\mathbf{X}$ et $\theta_1\mathbf{X}$ ont même loi sous $\mathbf{P}$, ce qui prouve $(iv)$ comme démontré précédemment. Cela clôt la preuve de la proposition. ■

## EXERCICES.

**Exercice I.20** Prouver la proposition I.2.31. □

**Exercice I.21** (*Critère de Kolmogorov*) Soit $Q=(p(i,j))_{i,j\in E}$, une matrice de transition sur un ensemble dénombrable $E$. On suppose que $Q$ est *irréductible*. Montrer que $Q$ est réversible **ssi** les deux conditions suivantes sont vérifiées :

**(a)** Pour tous $i,j\in E$, on a $p(i,j)>0$ ssi $p(j,i)>0$.

**(b)** Pour toute suite finie d'états $(i_0,i_1,\dots,i_n=i_0)$ telle que $p(i_0,i_1)\dots p(i_{n-1},i_n)>0$, on a

$$\prod_{0\le k<n}\frac{p(i_k,i_{k+1})}{p(i_{k+1},i_k)}=1\ .$$

□

**Exercice I.22** (*Serpent*) Soit $\big(\Omega\,;\,\mathscr{F}\,;\,(\mathscr{F}_n)_{n\geq 0}\,;\,(X_n)_{n\geq 0}\,;\,Q\,;\,\mathbf{P}_\mu\,,\,\mu\in\mathscr{M}_1(E)\;\big)$, une chaîne de Markov. On fixe $k\geq 1$ et on pose

$$X_n^{(k)}=(X_n,\ldots,X_{n+k})\;,\quad n\geq 0\;.$$

On pose également

$$E^{(k)}=\{\mathbf{i}=(\mathbf{i}_0,\ldots\mathbf{i}_k)\in E^{k+1}\;:\;p(\mathbf{i}_0,\mathbf{i}_1)\ldots p(\mathbf{i}_{k-1},\mathbf{i}_k)>0\}\;.$$

On définit aussi la matrice $Q_{(k)}=(p^{(k)}(\mathbf{i},\mathbf{j}))_{\mathbf{i},\mathbf{j}\in E^{(k)}}$, une matrice sur $E^{(k)}$ en posant

$$p^{(k)}(\mathbf{i},\mathbf{j})=\mathbf{1}_{\{\mathbf{j}_0=\mathbf{i}_1;\ldots;\mathbf{j}_{k-1}=\mathbf{i}_k\}}p(\mathbf{i}_k,\mathbf{j}_k)\;,\quad \mathbf{i}=(\mathbf{i}_0,\ldots,\mathbf{i}_k),\mathbf{j}=(\mathbf{j}_0,\ldots,\mathbf{j}_k)\in E^{(k)}\;.$$

1. Montrer que $Q_{(k)}=(p^{(k)}(\mathbf{i},\mathbf{j}))_{\mathbf{i},\mathbf{j}\in E^{(k)}}$, est une matrice de transition. Montrer que

   $$\big(\Omega\,;\,\mathscr{F}\,;\,(\mathscr{F}_{n+k})_{n\geq 0}\,;\,(X_n^{(k)})_{n\geq 0}\,;\,Q_{(k)}\,;\,\mathbf{P}_\mu\,,\,\mu\in\mathscr{M}_1(E^{(k)})\;\big)\;,$$

   est une chaîne de Markov. C'est le $k$-serpent associé à $(X_n)_{n\geq 0}$.
2. Montrer que si $Q$ est irréductible, il en est de même pour $Q_{(k)}$.
3. Montrer que si $Q$ est irréductible et $d$-périodique, il en est de même pour $Q_{(k)}$.
4. Soit $\mu$ une mesure $Q_{(k)}$-invariante. On définit $\mu_k$ sur $E^{(k)}$ par

   $$\mu_k(\mathbf{i})=\mu(\mathbf{i}_0)p(\mathbf{i}_0,\mathbf{i}_1)\ldots p(\mathbf{i}_{k-1},\mathbf{i}_k)\;,\quad \mathbf{i}=(\mathbf{i}_0,\ldots,\mathbf{i}_k)\in E^{(k)}\;.$$

   Montrer que $\mu_k$ est $Q_{(k)}$-invariante. □

**Exercice I.23** (*Certaines chaînes irréductibles n'ont pas de mesure invariante*) L'espace d'états est $E=\mathbb{N}$. Soit $(a_i,i\geq 0)$ une suite de réels strictement positifs tels que $a_0=1$ et $a_i\in\,]0,1[$, pour tout $i\geq 1$. On définit $Q=(p(i,j))_{i,j\in\mathbb{N}}$ en posant

$$p(i,j)=0\;\text{ si }\;j\notin\{0,i+1\}\quad\text{et}\quad p(i,0)=1-p(i,i+1)=a_i\;.$$

1. Montrer que pour tous $i,j\in\mathbb{N}$, où $j\geq 1$, on a

   $$[Q^{j+1}](i,j)\geq(1-a_i)a_0\ldots a_{j-1}>0\;.$$

   En déduire que $Q$ est irréductible apériodique.
2. On suppose que

   $$\sum_{i\geq 0}(1-a_i)\;<\infty\;. \tag{I.57}$$

   On peut prendre par exemple $a_i=1-\frac{1}{(i+1)^2}$, $i\geq 1$ et $a_0=1$ (comme requis). On suppose que $Q$ admet une mesure $\mu$ qui est invariante, c'est-à-dire que $(1)$ : $\mu$ n'est pas identiquement nulle, $(2)$ : on a $0\leq\mu(i)<\infty$, pour tout $i\in\mathbb{N}$ et $(3)$ : on a $\mu.Q=\mu$. Montrer que

   $$\forall i\geq 1\;,\quad \mu(i)=a_0a_1\ldots a_{i-1}\mu(0)\;.$$

3. On conserve les hypothèses de la question précédente. Montrer que

   $$\mu(0)=\mu(0)\big(1-\lim_n\downarrow a_1\ldots a_n\big)\;.$$

   Montrer que (I.57) implique que

   $$\lim_n\downarrow a_1\ldots a_n\leq\exp\Big(\sum_{i\geq 1}(1-a_i)\Big)<1\;.$$

   En déduire que $\mu(0)=0$.
4. Conclure. □

**Exercice I.24** On reprend la matrice de transition ci-dessus l'exercice I.23 ci-dessus

1. Sous quelle hypothèse nécessaire et suffisante sur les $(a_i)_{i\geq 0}$, la matrice $Q$ qui y est associée admet-elle une mesure invariante ? Montrer que sous cette hypothèse toutes les mesure $Q$-invariantes sont proportionnelles et telles que $\mu(i)>0$, $i\in\mathbb{N}$.
2. Sous quelle hypothèse nécesaire et suffisante sur les $(a_i)_{i\geq 0}$ admet-elle une probabilité invariante ? Montrer que sous cette hypothèse, cette probabilité invariante est unique.

3. On fixe $\beta \in ]0, \beta[$ et on pose $a_0 = 1$ et $a_i = 1 - \frac{1}{(i+1)^\beta}$, pour tout $i \geq 1$. Pour quels $\beta$, la matrice $Q$ admet-elle une loi de probabilité invariante ? Pour quels $\beta$ (s'il y en a) $Q$ admet-elle une mesure invariante de masse infinie ? Pour quels $\beta$ (s'il y en a) $Q$ n'admet-elle aucune mesure invariante ?
4. On fixe $c \in ]0, 2[$ et on pose $a_0 = 1$ et $a_i = 1 - \frac{c}{i+1}$, pour tout $i \geq 1$. Pour quels $c$, $Q$ admet-elle une loi de probabilité $Q$-invariante ? Pour quels $c$, $Q$ admet-elle une mesure invariante de masse infinie ? $Q$ admet-elle toujours une mesure invariante ? □

**Exercice I.25** Quelles sont les périodes possibles d'une matrice réversible irréductible ? Donner des exemples. □

**Exercice I.26** (*$h$-transformée de Doob*) Soit

$$\big(\Omega\,;\,\mathscr{F}\,;\,(\mathscr{F}_n)_{n\geq 0}\,;\,(X_n)_{n\geq 0}\,;\,Q = (p(i,j))_{i,j\in E}\,;\,\mathbf{P}_\mu\,,\,\mu \in \mathscr{M}_1(E)\,\big)\;,$$

une chaîne de Markov. On suppose qu'il existe une fonction $h : E \to \mathbb{R}_+$ qui n'est pas identiquement nulle et qui est $Q$-harmonique : $Q.h = h$. On pose

$$E_+ = \{i \in E \,:\, h(i) > 0\} \quad \text{et} \quad E_0 = E\backslash E_+ = \{i \in E \,:\, h(i) = 0\}\;.$$

1. On remarque que $Q^n.h = h$, pour tout $n \geq 0$. En déduire que si $i \in E_0$ et si $i \to j$, alors $j \in E_0$. Montrer que $E_0$ est $Q$-fermée.
2. Pour tout $i, j \in E_+$, on pose $p_h(i,j) = p(i,j)h(j)/h(i)$. Montrer que $Q_h = (p_h(i,j))_{i,j\in E_+}$ est une matrice de transition. On l'appelle la *$h$-transformée de Doob de $Q$*.
3. On fixe une loi initiale $\mu$ telle que $\mu(E_0) = 0$, c'est-à-dire que $\mu \in \mathscr{M}_1(E_+)$. Pour tout $n \in \mathbb{N}$ et pour tout $A \in \mathscr{F}_n$, on pose

$$\mathbf{P}^{n,h}_\mu(A) = \mathbf{E}_\mu\big[\mathbf{1}_A \frac{h(X_n)}{h(X_0)}\big]\;.$$

Montrer que $\mathbf{P}^{n,h}_\mu$ est une loi de probabilité sur l'espace mesurable $(\Omega, \mathscr{F}_n)$. On note $\mathbf{E}^{n,h}_\mu$ l'espérance (l'intégrale) qui lui est associée. Soient $m, n \in \mathbb{N}$ tels que $m < n$. Montrer que pour toute variable $Z : \Omega \to \mathbb{R}$, $\mathscr{F}_m$-mesurable bornée, on a

$$\mathbf{E}^{n,h}_\mu[Z] = \mathbf{E}^{m,h}_\mu[Z]\;,$$

4. Pour tous $m, n \in \mathbb{N}$ tels que $m < n$, montrer que la suite finie $(X_0, \ldots, X_m)$ sous $\mathbf{P}^{n,h}_\mu$ est une chaîne de Markov à valeurs dans $E_+$ de loi initiale $\mu \in \mathscr{M}_1(E_+)$ et de matrice de transition $Q_h$ la $h$-transformée de Doob de $Q$. □

*Les deux exercices suivants donnent une interprétation de la $h$-transformée de Doob sur deux exemples.*

**Exercice I.27** Soit $\big(\Omega\,;\,\mathscr{F}\,;\,(\mathscr{F}_n)_{n\geq 0}\,;\,(Z_n)_{n\geq 0}\,;\,Q = (p(i,j))_{i,j\in\mathbb{N}}\,;\,\mathbf{P}_\mu\,,\,\mu \in \mathscr{M}_1(\mathbb{N})\,\big)$, un processus de branchement de loi de reproduction $\xi$. On suppose que

$$m_1 = \sum_{i\in\mathbb{N}} i\xi(i) = 1\;.$$

On rappelle que $T = \inf\{n \geq 0 \,:\, Z_n = 0\}$ est le temps d'extinction. Comme le nombre moyen d'enfant est égal à 1, pour toute loi initiale, la population s'éteint presque sûrement : $\mathbf{P}_\mu(T < \infty) = 1$, c'est-à-dire $\mathbf{P}_\mu(\lim_n Z_n = 0) = 1$. On note $\varphi_\xi$ la fonction génératrice de $\xi$. On rappelle que

$$\mathbf{E}_\mu\big[r^{Z_{n+1}}|\mathscr{F}_n\big] = \varphi_\xi(r)^{Z_n}\;, \quad r \in [0,1]\,,\; n \in \mathbb{N}.$$

On note $\varphi_\xi^{\circ n}$ la $n$-ième itérée de $\varphi_\xi$. On déduit de ce qui précède que

$$\mathbf{E}_\mu\big[r^{Z_n}\big] = \mathbf{E}_\mu\big[\varphi_\xi^{\circ n}(r)^{Z_0}\big]\;,$$

et donc que

$$\forall n \in \mathbb{N}\,, \quad \mathbf{P}_\mu(T > n) = 1 - \mathbf{E}_\mu\big[\varphi_\xi^{\circ n}(0)^{Z_0}\big]\;.$$

On rappelle que $Q$ est caractérisée par

$$\forall i \in \mathbb{N}\,, \quad \varphi_\xi(r)^i = \sum_{j\in\mathbb{N}} p(i,j)r^j\;, \quad r \in [0,1]\;.$$

1. En remarquant que $(Z_n)_{n\geq 0}$ est une $(\mathscr{F}_n)_{n\geq 0}$-martingale positive sous $\mathbf{P}_\mu$ pour toute loi intiale $\mu$, en déduire que $h : \mathbb{N} \to \mathbb{R}_+$ donnée par $h(i) = i$, $i \in \mathbb{N}$, est $Q$-harmonique. On a $\mathbb{N}_+ = \mathbb{N}^* = \{1, 2, \dots\}$ et $\mathbb{N}_0 = \{0\}$.
2. On note $Q_h = (p_h(i,j))_{i,j\in\mathbb{N}^*}$, la $h$-transformée de Doob de $Q$. Montrer qu'elle est caractérisée par
$$\forall i \in \mathbb{N}^* , \quad \varphi'_\xi(r)\varphi_\xi(r)^{i-1} = \sum_{j\in\mathbb{N}^*} p_h(i,j)r^j \ , \quad r \in [0,1] \ .$$
3. Montrer que pour tout entier $b \in \mathbb{N}$, on a les inégalités
$$\forall r \in [0,1] , \quad 0 \leq (1-r)^b - 1 + br \leq \tfrac{1}{2} b(b-1)r^2 \quad \text{et} \quad 0 \leq (1-r)^b - 1 + br \leq br$$
En déduire que
$$\forall r \in [0,1] , \quad 0 \leq (1-r)^b - 1 + br \leq br \cdot \big(1 \wedge \tfrac{(b-1)r}{2}\big) \ .$$
4. On fixe $m \geq 1$. Soit $F : \mathbb{N}^m \to \mathbb{R}$, une fonction bornée. On fixe $i \in \mathbb{N}^*$. Pour tout $k \geq 1$, on pose $r_k = \varphi_\xi^{\circ k}(0)$. On rappelle que $\lim_k \uparrow r_k = 1$. Montrer que
$$\forall n \geq 1 , \quad \mathbf{E}_i\big[F(Z_0, \dots, Z_m)\mathbf{1}_{\{T>m+n\}}\big] = \mathbf{E}_i\big[F(Z_0, \dots, Z_m)\big(1 - (1-r_n)^{Z_m}\big)\big] .$$
On pose $\varepsilon_n = \mathbf{E}_i\big[Z_m\big(1 \wedge \frac{(Z_m-1)r_n}{2}\big)\big]$. Montrer que $\lim_n \varepsilon_n = 0$. Montrer ensuite que
$$\begin{aligned} &\Big| \tfrac{r_n}{1-(1-r_n)^i}\mathbf{E}_i\big[F(Z_0, \dots, Z_m)Z_m\big] - \mathbf{E}_i\big[F(Z_0, \dots, Z_m) \,\big|\, T > m+n\big] \Big| \\ &\quad \leq \quad \|F\|_\infty \cdot \tfrac{r_n}{1-(1-r_n)^i} \cdot \varepsilon_n \ . \end{aligned}$$
5. En déduire que pour tout $m \in \mathbb{N}$ et tout $i \in \mathbb{N}^*$,
$$(Z_0, \dots, Z_m) \quad \text{sous} \quad \mathbf{P}_i(\,\cdot\,|\, T > n+m) \xrightarrow[n\to\infty]{\text{(loi)}} (Z_0, \dots, Z_m) \quad \text{sous} \quad \mathbf{P}_i^{m,h} \ .$$
C'est-à-dire que la chaîne dont la matrice de transition est la $h$-transformée de celle du processus de branchement critique s'interprête comme la loi du processus critique conditionné à ne jamais s'éteindre. Ici, il faut remarquer que l'événement "ne jamais s'éteindre" est de probabilité nulle pour le processus de branchement critique : il faut comprendre ce conditionnement exactement comme la limite en loi donnée précédemment. □

**Exercice I.28** Soit $\big(\Omega\,;\, \mathscr{F}\,;\, (\mathscr{F}_n)_{n\geq 0}\,;\, (X_n)_{n\geq 0}\,;\, Q = (p(i,j))_{i,j\in\mathbb{N}}\,;\, \mathbf{P}_\mu \,,\, \mu \in \mathscr{M}_1(\mathbb{N})\ \big)$, une marche aléatoire simple sur $\mathbb{N}$, considéré comme un graphe, l'ensemble des sommets étant $\mathbf{S} = \mathbb{N}$ et l'ensemble des arêtes étant $\mathbf{A} = \{\,\{i, i+1\}; i \in \mathbb{N}\}$.

1. Montrer que cette marche est un processus de naissance et de mort dont on précisera la matrice de transition $Q$.
2. Si $(Y_n)_{n\geq 0}$ est une marche simple sur $\mathbb{Z}$, montrer que $(|Y_n|)_{n\geq 0}$ est une marche aléatoire sur $\mathbb{N}$, considéré comme un graphe, c'est-à-dire une chaîne de Markov de matrice de transition $Q$.
3. Montrer que la fonction $h : \mathbb{N} \to \mathbb{R}_+$, donnée par $h(i) = i$, $i \in \mathbb{N}$, est $Q$-harmonique. On a donc $\mathbb{N}_+ = \mathbb{N}^* = \{1, 2, \dots\}$ et $\mathbb{N}_0 = \{0\}$.
4. On pose $T_0 = \inf\{n \geq 0 : X_n = 0\}$ et pour tout $M$, $T_M = \inf\{n \geq 0 : X_n = M\}$. On fixe $1 \leq i < M$. D'après la question 2), $\mathbf{P}_i(T_M < T_0)$ est égale à la probabilité qu'une marche simple issue de $i$ atteigne $M$ avant 0. Montrer que $\mathbf{P}_i(T_M < T_0) = \frac{i}{i+M}$.
On fixe $m \geq 1$. Soit $F : \mathbb{N}^m \to \mathbb{R}$, une fonction bornée. On fixe $i \in \mathbb{N}^*$. Montrer que pour tout $M > 1$
$$\lim_{M\to\infty} \mathbf{E}_i\big[F(X_0, \dots, X_m) \,\big|\, T_M < T_0\big] = \mathbf{E}_i^{m,h}\big[F(X_0, \dots, X_m)\big] \ .$$
C'est-à-dire que la chaîne dont la matrice de transition est la $h$-transformée de celle de la marche simple sur $\mathbb{N}$ s'interprête comme la loi de la marche simple sur $\mathbb{N}$ conditionnée à ne jamais revenir en 0. Ici, il faut remarquer que l'événement "ne jamais revenir en 0" est de probabilité nulle pour la marche simple sur $\mathbb{N}$ : il faut comprendre ce conditionnement exactement comme la limite en loi donnée précédemment. □

**Exercice I.29** On explique ici le mot "duale" dans l'expression "matrice de transition duale". Soit $Q = (p(i,j))_{i,j\in E}$ qui admet une mesure invariante $\mu$ telle que $\mu(i) > 0$, pour tout $i \in E$. On rappelle la définition de la matrice duale $Q_* = ((p^*(i,j))_{i,j\in E}$ donnée par $p^*(i,j) = \mu(j)p(j,i)/\mu(i)$, pour tous $i,j \in E$. On équipe $E$ de la tribu $\mathscr{P}(E)$ de tous ses sous-ensembles : $\mathscr{P}(E) = \{B \subset E\}$. On note $\mathrm{L}^2(E, \mathscr{P}(E), \mu)$ l'espace des fonctions $f : E \to \mathbb{R}$ dont le carré est $\mu$-intégrable :

$$\mathrm{L}^2(E, \mathscr{P}(E), \mu) = \Big\{ f : E \to \mathbb{R} \;:\; \langle \mu, f^2 \rangle = \sum_{i \in E} f^2(i)\mu(i) \;< \infty \Big\}.$$

Pour toutes fonctions $f, g \in \mathrm{L}^2(E, \mathscr{P}(E), \mu)$, on pose

$$(f,g)_\mu = \sum_{i\in E} f(i)g(i)\mu(i) \quad \text{et} \quad \|f\|_\mu^2 = (f,f)_\mu = \langle \mu, f^2\rangle = \sum_{i\in E} f^2(i)\mu(i) < \infty$$

1. Vérifie que $(\,\cdot\,,\,\cdot\,)_\mu$ est un produit scalaire bien défini sur $\mathrm{L}^2(E, \mathscr{P}(E), \mu)$. On note que cela n'est plus vrai s'il existe $i_0 \in E$ tel que $\mu(i_0) = 0$. Expliquer (ou redémontrer) pourquoi $\big(\mathrm{L}^2(E, \mathscr{P}(E), \mu)\,;\,(\,\cdot\,,\,\cdot\,)_\mu\big)$ est un espace de Hilbert, c'est-à-dire qu'il est complet.
2. Soient $f, g \in \mathrm{L}^2(E, \mathscr{P}(E), \mu)$. En utilisant Cauchy-Schwarz, montrer que

$$\big|\,(Q.f, g)_\mu\,\big| \leq \sum_{j\in E} |f(j)|\sqrt{\mu(j)} \cdot \sqrt{\sum_{i\in E} \mu(i)p(i,j)g^2(i)},$$

   l'inégalité ayant lieu a priori dans $[0,\infty]$. En utilisant encore Cauchy-Schwarz, montrer que

$$\big|\,(Q.f, g)_\mu\,\big| \leq \sqrt{\sum_{j\in E} \mu(j)f^2(j)} \cdot \sqrt{\sum_{i,j\in E} \mu(i)p(i,j)g^2(i)} = \|f\|_\mu \cdot \|g\|_\mu \,.$$

3. On rappelle le théorème de représentation de Riesz : si $(H, [\cdot,\cdot])$ est un espace de Hilbert et si $\phi : H \to \mathbb{R}$ est une forme linéaire continue, alors il existe un unique élément $x_\phi \in H$ tel que $\phi(y) = [x_\phi, y]$, pour tout $y \in H$. Pour tout $g \in \mathrm{L}^2(E, \mathscr{P}(E), \mu)$, on définit la forme linéaire $\phi_g : \mathrm{L}^2(E, \mathscr{P}(E), \mu) \to \mathbb{R}$ par

$$\forall f \in \mathrm{L}^2(E, \mathscr{P}(E), \mu)\,, \quad \phi_g(f) = (Q.f, g)_\pi \,.$$

   La fonction $\phi_g$ est bien définie grâce à la question précédente. En utilisant cette question, constater que $\phi_g$ est une forme linéaire continue. En utilisant le théorème de Riesz, montrer qu'il existe un unique élément, noté $Q^\bullet g$ de $\mathrm{L}^2(E, \mathscr{P}(E), \mu)$ tel que $(Q.f, g)_\pi = (f, Q^\bullet g)_\pi$. En utilisant la linéarité de $Q$ et l'unicité dans le théorème de Riesz, montrer que $g \mapsto Q^\bullet g$ est une application linéaire. En utilisant la question précédente, montrer que

$$\forall g \in \mathrm{L}^2(E, \mathscr{P}(E), \mu)\,, \quad \|Q^\bullet.g\|_\mu \leq \|g\|_\mu \,.$$

   C'est donc un opérateur continu (borné en norme) sur $\mathrm{L}^2(E, \mathscr{P}(E), \mu)$.
4. Montrer que $Q^\bullet = Q_*$.
5. Puisque $(Q_*)_* = Q$, en appliquant les résultats précendents à $Q_*$, montrer que $Q$ est tel que $\|Q.f\|_\mu \leq \|f\|_\mu$ et donc que $Q : \mathrm{L}^2(E, \mathscr{P}(E), \mu) \to \mathrm{L}^2(E, \mathscr{P}(E), \mu)$ (ce qui n'a rien d'évident a priori).

   *En résumé : on a montré que si $f \in \mathrm{L}^2(E, \mathscr{P}(E), \mu)$, alors $Q.f$ et $Q_*.f \in \mathrm{L}^2(E, \mathscr{P}(E), \mu)$ et que*

$$\forall f, g \in \mathrm{L}^2(E, \mathscr{P}(E), \mu)\,, \quad (Q.f, g)_\mu = (f, Q_*.g)_\mu$$

   *ce qui justifie que l'on appelle $Q_*$ la matrice de transition duale de $Q$. De plus on a montré les inégalités suivantes :*

$$\forall f, g \in \mathrm{L}^2(E, \mathscr{P}(E), \mu)\,, \quad \big|\,(Q.f, g)_\mu\,\big| \leq \|f\|_\mu \cdot \|g\|_\mu \,,\; \|Q.f\|_\mu \text{ et } \|Q_*.f\|_\mu \;\leq \|f\|_\mu \,,$$

   *$Q$ et $Q_*$ sont des opérateurs linéaires continus (bornés en norme) pour la norme $\|\cdot\|_\mu$* □

## I.3 Asymptotique des chaînes : résultats qualitatifs.

Dans la suite $\big(\Omega; \mathscr{F}; (\mathscr{F}_n)_{n\geq 0}; \mathbf{X} = (X_n)_{n\geq 0}; Q = (p(i,j))_{i,j\in E}; \mathbf{P}_\mu, \mu \in \mathscr{M}_1(E)\big)$ désigne une chaîne de Markov.

### I.3.a Excursions, états récurrents et transitoires.

Commençons par quelques notations déterministes : soit $\mathbf{x} = (x_n)_{n\geq 0} \in E^{\mathbb{N}}$, une suite à valeurs dans $E$ et soit $i \in E$, un état fixé. Les *temps de retour successifs* de $\mathbf{x}$ en $i$ sont récursivement définis par

$$T_i^{(0)}(\mathbf{x}) = 0 \quad \text{et} \quad T_i^{(p+1)}(\mathbf{x}) = \inf\{n > T_i^{(p)}(\mathbf{x}) \, : \, x_n = i\} \ ,$$

avec la convention habituelle $\inf \emptyset = \infty$. On pose ensuite

$$N_i(\mathbf{x}) = \sup\{p \in \mathbb{N} \, : \, T_i^{(p)}(\mathbf{x}) < \infty \ \} \ \in \mathbb{N} \cup \{\infty\} \ .$$

$N_i(\mathbf{x})$ est égal au *nombre de visites de la suite* $\mathbf{x}$ *en* $\{i\}$ *à partir du temps* 1. En effet il est facile de vérifier que

$$N_i(\mathbf{x}) = \sum_{n\geq 1} \mathbf{1}_{\{x_n = i\}} \ .$$

Pour alléger les notations, on notera simplement

$$T_i^{(p)} := T_i^{(p)}(\mathbf{X}) \quad \text{et} \quad N_i := N_i(\mathbf{X}) \ .$$

Il est facile de montrer par récurrence que pour tout $p \geq 0$, $T_i^{(p)}$ est un $(\mathscr{F}_n)_{n\geq 0}$-temps d'arrêt. Nous donnons plus loin un contenu plus parlant à la définition suivante.

**Définition I.3.1** Avec les notations introduites ci-dessus, on définit les notions suivantes.

**(Etat récurrent)** $i \in E$ est *récurrent* si $\mathbf{P}_i(T_i^{(1)} < \infty) = 1$.

**(Etat transient)** $i \in E$ est *transient* (ou *transitoire*) ssi $\mathbf{P}_i(T_i^{(1)} < \infty) < 1$. □

**Excursions.** On définit ensuite les *excursions de la suite* $\mathbf{x}$ *en dehors de l'état* $\{i\}$. Pour cela on choisit un point $\partial$ qui n'est pas dans l'espace d'état $E$ : $\partial \notin E$. On voit $\partial$ comme un *point cimetière* (ou bien un point à l'infini). On pose $E^* = E \cup \{\partial\}$, qui reste un ensemble dénombrable. On note $(\partial)$, la suite à valeurs dans $E^*$ qui est constante à $\partial$. Soit $\mathbf{x} = (x_n)_{n\geq 0} \in E^{\mathbb{N}}$. On définit la première excursion de $\mathbf{x}$ hors de $\{i\}$ comme la suite $\mathscr{E}^{(1)}(\mathbf{x}) = (\mathscr{E}_n^{(1)}(\mathbf{x}))_{n\geq 0}$, à valeurs dans $E^*$ donnée par

$$\mathscr{E}_n^{(1)}(\mathbf{x}) = \begin{cases} x_n & \text{si } n < T_i^{(1)}(\mathbf{x}) \\ \partial & \text{si } n \geq T_i^{(1)}(\mathbf{x}). \end{cases}$$

Si $T_i^{(1)}(\mathbf{x}) = \infty$, alors $\mathscr{E}^{(1)}(\mathbf{x}) = \mathbf{x}$. On définit ensuite les excursions successives $\mathscr{E}^{(p)}(\mathbf{x})$, $p \geq 1$, en posant

$$\mathscr{E}^{(p+1)}(\mathbf{x}) = \begin{cases} \mathscr{E}^{(1)}\big(\theta_{T_i^{(p)}(\mathbf{x})}\mathbf{x}\big) & \text{si } T_i^{(p)}(\mathbf{x}) < \infty \\ (\partial) & \text{si } T_i^{(p)}(\mathbf{x}) = \infty. \end{cases}$$

Il est facile de voir que pour tout $p \geq 1$,

$$T_i^{(p)}(\mathbf{x}) < \infty \implies \mathscr{E}^{(p+1)}(\mathbf{x}) = \mathscr{E}^{(p)}\big(\theta_{T_i^{(1)}(\mathbf{x})}\mathbf{x}\big) \ . \tag{I.58}$$

On voit aussi que si $T_i^{(p+1)}(\mathbf{x}) < \infty$, alors

$$\mathscr{E}^{(p+1)}(\mathbf{x}) = (x_{T_i^{(p)}(\mathbf{x})}, x_{T_i^{(p)}(\mathbf{x})+1}, \ldots, x_{T_i^{(p+1)}(\mathbf{x})-1}, \partial, \partial, \ldots)$$

et si $T_i^{(p)}(\mathbf{x}) < \infty$ mais $T_i^{(p+1)}(\mathbf{x}) = \infty$, alors

$$\mathscr{E}^{(p+1)}(\mathbf{x}) = (x_{T_i^{(p)}(\mathbf{x})}, x_{T_i^{(p)}(\mathbf{x})+1}, \dots, x_{T_i^{(p)}(\mathbf{x})+n}, \dots) .$$

On s'épargne la peine de montrer formellement que pour tout $p \geq 1$, la fonction

$$\mathbf{x} \in E^{\mathbb{N}} \longmapsto \mathscr{E}^{(p)}(\mathbf{x}) \in (E^*)^{\mathbb{N}}$$

est bien $(\mathscr{P}(E)^{\otimes \mathbb{N}}, \mathscr{P}(E^*)^{\otimes \mathbb{N}})$-mesurable (ce résultat, très simple à montrer mais fastidieux à écrire, est laissé au lecteur). Pour simplifier les notations, on note pour tout $p \geq 1$,

$$\mathscr{E}^{(p)} = \big(\mathscr{E}_n^{(p)}\big)_{n \geq 0} := \mathscr{E}^{(p)}(\mathbf{X}) .$$

$\mathscr{E}^{(p)} : \Omega \to E^{*\mathbb{N}}$ est $(\mathscr{F}, \mathscr{P}(E^*)^{\otimes \mathbb{N}})$-mesurable.

**Théorème I.3.2** *On note $\mathscr{E}^{(p)}$, $p \geq 1$, la suite des excursions de $\mathbf{X}$ en dehors de l'état $\{i\}$, comme définies précédemment. Alors, on a l'alternative suivante.*

**(Cas 1)** *Si $i$ est récurrent, c-à-d si $\mathbf{P}_i(T_i^{(1)} < \infty) = 1$, alors*

$$\mathbf{P}_i\text{-}p.s. \quad \forall p \geq 1 \ , \ T_i^{(p)} < \infty \quad \text{et} \quad N_i = \infty .$$

*De plus, sous $\mathbf{P}_i$, les excursions $\mathscr{E}^{(p)}$, $p \geq 1$, sont indépendantes et de même loi.*

**(Cas 2)** *Si $i$ est transient, c-à-d si $\mathbf{P}_i(T_i^{(1)} < \infty) < 1$, alors sous $\mathbf{P}_i$, la variable aléatoire $N_i$ a une loi géométrique. Plus précisément, si on pose $\rho_i := \mathbf{P}_i(T_i^{(1)} = \infty) > 0$, on a*

$$\forall p \geq 0 \ , \quad \mathbf{P}_i(N_i = p) = \rho_i(1-\rho_i)^p .$$

*De plus, sous la probabilité conditionnelle $\mathbf{P}_i(\,\cdot \mid N_i = p)$, les assertions suivantes sont vérifiées.*

- *Les excursions $\mathscr{E}^{(1)}, \dots, \mathscr{E}^{(p+1)}$ sont indépendantes.*
- *$\mathscr{E}^{(1)}, \dots, \mathscr{E}^{(p)}$ ont même loi que $\mathscr{E}^{(1)}$ sous $\mathbf{P}_i(\,\cdot\, | T_i^{(1)} < \infty)$.*
- *$\mathscr{E}^{(p+1)}$ a même loi que $\mathscr{E}^{(1)}$ sous $\mathbf{P}_i(\,\cdot\, | T_i^{(1)} = \infty)$.*

**Preuve :** on fixe $p \geq 1$. Soient $G_1, \dots, G_{p+1}$, $p+1$ fonctions de $(E^*)^{\mathbb{N}}$ dans $\mathbb{R}$ qui sont $\mathscr{P}(E^*)^{\otimes \mathbb{N}}$-mesurables bornées. Pour tout $1 \leq \ell \leq p+1$, on pose

$$F_\ell(\mathbf{X}) = \mathbf{1}_{\{T_i^{(\ell-1)} < \infty\}} G_\ell(\mathscr{E}^{(1)}) \dots G_1(\mathscr{E}^{(\ell)}) .$$

On observe alors que (I.58) implique

$$F_{p+1}(\mathbf{X}) = \mathbf{1}_{\{T_i^{(1)} < \infty\}} G_{p+1}(\mathscr{E}^{(1)}) F_p(\theta_{T_i^{(1)}} \mathbf{X}). \tag{I.59}$$

On remarque que $X_{T_i^{(1)}} = i$ sur l'événement $\{T_i^{(1)} < \infty\}$. Par conséquent la propriété de Markov forte au temps $T_i^{(1)}$ implique que $\mathbf{P}_i$-p.s. on a

$$\mathbf{E}_i\big[F_{p+1}(\mathbf{X}) | \mathscr{F}_{T_i^{(1)}}\big] \quad = \quad \mathbf{1}_{\{T_i^{(1)} < \infty\}} G_{p+1}(\mathscr{E}^{(1)}) \mathbf{E}_i\big[F_p(\theta_{T_i^{(1)}} \mathbf{X}) | \mathscr{F}_{T_i^{(1)}}\big]$$

$$= \mathbf{1}_{\{T_i^{(1)}<\infty\}} G_{p+1}(\mathscr{E}^{(1)}) \mathbf{E}_i \left[ F_p(\mathbf{X}) \right] ,$$

car $G_{p+1}(\mathscr{E}^{(1)})$ est clairement $\mathscr{F}_{T_i^{(1)}}$-mesurable. On a donc

$$\mathbf{E}_i \left[ F_{p+1}(\mathbf{X}) \right] = \mathbf{E}_i \left[ \mathbf{1}_{\{T_i^{(1)}<\infty\}} G_{p+1}(\mathscr{E}^{(1)}) \right] \mathbf{E}_i \left[ F_p(\mathbf{X}) \right] . \tag{I.60}$$

Une simple récurrence entraîne alors que

$$\begin{aligned} \mathbf{E}_i \Big[ \mathbf{1}_{\{T_i^{(p)}<\infty\}} & G_{p+1}(\mathscr{E}^{(1)}) \dots G_2(\mathscr{E}^{(p)}) G_1(\mathscr{E}^{(p+1)}) \Big] \\ &= \mathbf{E}_i[G_1(\mathscr{E}^{(1)})] \prod_{1\le \ell \le p} \mathbf{E}_i[\mathbf{1}_{\{T_i^{(1)}<\infty\}} G_{\ell+1}(\mathscr{E}^{(1)})]. \end{aligned} \tag{I.61}$$

Comme, $\{N_i \ge p\} = \{T_i^{(p)} < \infty\}$, en prenant $G_1 = \dots = G_{p+1} = \mathbf{1}$ dans (I.61), on obtient

$$\forall p \ge 1 , \quad \mathbf{P}_i(N_i \ge p) = \mathbf{P}_i\big(T_i^{(1)} < \infty\big)^p . \tag{I.62}$$

Si on se place dans le premier cas où $\mathbf{P}(T_i^{(1)} < \infty) = 1$, on voit que (I.61) et (I.62) impliquent immédiatement les résultats désirés.

On se place dans le second cas où $\rho_i = \mathbf{P}(T_i^{(1)} = \infty) > 0$. (I.62) implique immédiatement que $N_i$ sous $\mathbf{P}_i$ suit une loi géométrique de paramètre $\rho_i$. On déduit ensuite facilement de (I.61) que

$$\begin{aligned} \mathbf{E}_i \Big[ \mathbf{1}_{\{N_i=p\}} & G_{p+1}(\mathscr{E}^{(1)}) \dots G_2(\mathscr{E}^{(p)}) G_1(\mathscr{E}^{(p+1)}) \Big] \\ &= \mathbf{E}_i[\mathbf{1}_{\{T_i^{(1)}=\infty\}} G_1(\mathscr{E}^{(1)})] \prod_{1\le \ell \le p} \mathbf{E}_i[\mathbf{1}_{\{T_i^{(1)}<\infty\}} G_{\ell+1}(\mathscr{E}^{(1)})] . \end{aligned}$$

Par conséquent

$$\begin{aligned} \mathbf{E}_i \Big[ & G_{p+1}(\mathscr{E}^{(1)}) \dots G_2(\mathscr{E}^{(p)}) G_1(\mathscr{E}^{(p+1)}) \,|\, N_i = p \Big] \\ &= \mathbf{E}_i[G_1(\mathscr{E}^{(1)}) \,|\, T_i^{(1)} = \infty] \prod_{1\le \ell \le p} \mathbf{E}_i[G_{\ell+1}(\mathscr{E}^{(1)}) \,|\, T_i^{(1)} < \infty \,] , \end{aligned}$$

ce qui entraîne le résultat voulu. ■

Bien que les variables constituant une chaîne de Markov soient dépendantes en général, le résultat précédent montre qu'il existe de larges portions de trajectoire qui sont indépendantes. Ce phénomène de *renouvellement* est la clef d'une partie des résultats des prochaines sections. Le théorème précédent montre également la loi du 0-1 suivante : ou bien $\mathbf{P}_i(N_i = \infty) = 1$ ou bien $\mathbf{P}_i(N_i = \infty) = 0$, ce qui est une propriété remarquable des chaînes de Markov. Le corollaire suivant donne un contenu plus concret aux notions de récurrence et de transience des états.

**Corollaire I.3.3** *Il y a équivalence entre les assertions suivantes.*

*(i) L'état $i \in E$ est récurrent, c'est-à-dire,* $\mathbf{P}_i(T_i^{(1)} < \infty) = 1$.

*(ii)* $\mathbf{P}_i(N_i = \infty) = 1$.

*(iii)* $\mathbf{E}_i[N_i] = \sum_{n\ge 1}[Q^n](i,i) = \infty$.

*De même, il y a équivalence entre les assertions suivantes.*

*(i*)* *L'état $i \in E$ est transient, c'est-à-dire,* $\mathbf{P}_i(T_i^{(1)} < \infty) < 1$.

*(ii*)* $\mathbf{P}_i(N_i < \infty) = 1$.

*(iii*)* $\mathbf{E}_i[N_i] = \sum_{n \geq 1} [Q^n](i,i) < \infty$.

*Par ailleurs si (iii*) est satisfaite, on a* $\mathbf{E}_i[N_i] = \mathbf{P}_i(T_i^{(1)} < \infty)/\mathbf{P}_i(T_i^{(1)} = \infty)$.

**Preuve :** on remarque d'abord que $(i) \iff (ii)$ et que $(i^*) \iff (ii^*)$, d'après le théorème I.3.2. Ce même théorème implique que ou bien $N_i = \infty$, $\mathbf{P}_i$-presque sûrement, ou bien $N_i$ suit une loi géométrique sous $\mathbf{P}_i$. Cela implique que $\mathbf{E}_i[N_i] = \infty$ ssi $\mathbf{P}_i(N_i = \infty) = 1$ et aussi que $\mathbf{E}_i[N_i] < \infty$ ssi $\mathbf{P}_i(N_i < \infty) = 1$, ce qui n'a rien d'évident a priori. On conclut par Fubini positif qui implique que

$$\mathbf{E}_i[N_i] = \mathbf{E}_i\Big[\sum_{n \geq 1} \mathbf{1}_{\{X_n = i\}}\Big] = \sum_{n \geq 1} \mathbf{P}_i(X_n = i) = \sum_{n \geq 1} [Q^n](i,i) \ .$$

Le calcul $\mathbf{E}_i[N_i] = \mathbf{P}_i(T_i^{(1)} < \infty)/\mathbf{P}_i(T_i^{(1)} = \infty)$ est élémentaire. ■

## I.3.b Classification des états.

**Lemme I.3.4** *Soient $i, j \in E$, deux états qui communiquent : $i \leftrightarrow j$. Alors, seules les deux situations suivantes sont possibles : ou bien $i$ et $j$ sont tous les deux récurrents, ou bien $i$ et $j$ sont tous les deux transients.*

**Preuve :** comme $i$ et $j$ communiquent, il existe $k, \ell \in \mathbb{N}$ tels que $[Q^k](i,j) > 0$ et $[Q^\ell](j,i) > 0$. Pour tout $n \in \mathbb{N}$, on voit facilement que

$$[Q^{k+\ell+n}](i,i) \geq [Q^k](i,j)[Q^n](j,j)[Q^\ell](j,i) \text{ et } [Q^{k+\ell+n}](j,j) \geq [Q^\ell](j,i)[Q^n](i,i)[Q^k](i,j) \ .$$

Cela entraîne que les séries $\sum_{n \geq 1}[Q^n](i,i)$ et $\sum_{n \geq 1}[Q^n](j,j)$ sont de même nature, ce qui permet de conclure grâce au corollaire I.3.3. ■

Le lemme suivant montre que, en un certain sens, la récurrence se propage.

**Lemme I.3.5** *On suppose que $j$ est accessible depuis $i$ : $i \to j$. Si $i$ est récurrent, alors $j \to i$ et $j$ est récurrent. De plus on a $\mathbf{P}_i(N_j = \infty) = 1$, et $\mathbf{P}_j(N_i = \infty) = 1$ également.*

**Preuve :** raisonnons par l'absurde en supposant que $i$ n'est pas accessible depuis $j$. Cela signifie que pour tout $n \in \mathbb{N}$, on a $\mathbf{P}_j(X_n = i) = [Q^n](j,i) = 0$, ce qui implique que $\mathbf{P}_j(T_i^{(1)} = \infty) = 1$. Comme $i \to j$, il existe $k \in \mathbb{N}$ tel que $\mathbf{P}_i(X_k = j) = [Q^k](i,j) > 0$. En appliquant la propriété de Markov au temps $k$, on voit que

$$\mathbf{P}_i\text{-p.s.} \quad \mathbf{E}_i\big[\mathbf{1}_{\{X_k = j\} \cap \{T_i^{(1)}(\theta_k \mathbf{X}) = \infty\}} \,|\, \mathscr{F}_k\big] = \mathbf{1}_{\{X_k = j\}} \mathbf{P}_{X_k}\big(T_i^{(1)} = \infty\big) = \mathbf{1}_{\{X_k = j\}} \ .$$

En intégrant, on obtient $\mathbf{P}_i\big(X_k = j; T_i^{(1)}(\theta_k \mathbf{X}) = \infty\big) = \mathbf{P}_i(X_k = j) > 0$, Or $\{X_k = j; T_i^{(1)}(\theta_k \mathbf{X}) = \infty\} \subset \{N_i < \infty\}$ et on aurait $\mathbf{P}_i(N_i < \infty) > 0$, ce qui contredit le fait que $i$ est récurrent (on utilise ici le corollaire I.3.3). Cela montre par l'absurde que $j \to i$. Comme on a supposé $i \to j$, cela implique que $i \leftrightarrow j$. Le lemme (I.3.4) entraîne alors que $j$ est récurrent.

Il reste à montrer que $\mathbf{P}_i(N_j = \infty) = 1$ : le théorème I.3.2 montre que sous $\mathbf{P}_i$, les excursions $\mathscr{E}^{(p)}$, $p \geq 1$, de la chaîne en dehors de l'état $\{i\}$ sont indépendantes et de même loi. Pour tout $p \geq 1$, on pose

$B_p = \{\exists n \in \mathbb{N} : \mathscr{E}_n^{(p)} = j\}$. Il est clair que les variables de Bernoulli $\mathbf{1}_{B_p}$, $p \geq 1$, sont i.i.d. sous $\mathbf{P}_i$. On pose $\alpha = \mathbf{P}(B_1) = \mathbf{E}_i[\mathbf{1}_{B_p}]$. La loi des grands nombres sous $\mathbf{P}_i$, implique que

$$\mathbf{P}_i\text{-p.s.} \quad \frac{1}{p} \sum_{1 \leq q \leq p} \mathbf{1}_{B_q} \xrightarrow[p \to \infty]{} \alpha .$$

Si $\alpha > 0$, alors cela entraîne que $\mathbf{P}_i$-p.s. $\sum_{p \geq 1} \mathbf{1}_{B_p} = \infty$. Par ailleurs, il est clair que

$$\mathbf{P}_i\text{-p.s.} \quad N_j \geq \sum_{p \geq 1} \mathbf{1}_{B_p} .$$

On a donc montré que si $\alpha > 0$, alors $\mathbf{P}_i(N_j = \infty) = 1$. Il suffit donc de prouver que

$$\alpha = \mathbf{P}(\exists n \in \mathbb{N} : \mathscr{E}_n^{(1)} = j) > 0 . \tag{I.63}$$

Comme $i \to j$, il existe $i_1, \ldots, i_\ell$ tel que $p(i, i_1) \ldots p(i_\ell, j) > 0$. On note $r = \max\{s \in \{1, \ldots, \ell\} : i_s = i\}$. On a donc $i_{r+1}, \ldots, i_\ell \neq i$ et $p(i, i_{r+1}) \ldots p(i_\ell, j) > 0$. Cela montre qu'il existe $i_1^*, \ldots, i_m^* \in E \backslash \{i\}$ tel que $p(i, i_1^*) \ldots p(i_m^*, j) > 0$. Or on observe que

$$\mathbf{P}_i\text{-p.s.} \quad \mathbf{1}_{\{\exists n \in \mathbb{N} : \mathscr{E}_n^{(1)} = j\}} \geq \mathbf{1}_{\{X_0 = i; X_1 = i_1^*; \ldots; X_m = i_m^*; X_{m+1} = j\}} .$$

En prenant l'espérance de cette inégalité sous $\mathbf{P}_i$, on a donc $\alpha \geq p(i, i_1^*) \ldots p(i_m^*, j) > 0$, ce qui entraîne (I.63) et donc que $\mathbf{P}_i(N_j = \infty) = 1$. On a également, $\mathbf{P}_j(N_i = \infty) = 1$, en échangeant les rôles de $i$ et $j$. ■

**Définition I.3.6** On note $\mathbf{R}$ l'ensemble des états récurrents et on note $\mathbf{T}$ l'ensemble des états transients. Par définition $\mathbf{R} \cup \mathbf{T} = E$. Si $\mathbf{R} \neq \emptyset$, on note $\mathbf{R}_1, \mathbf{R}_2 \ldots$ les classes d'équivalence de $\mathbf{R}$ pour la relation de communication $\leftrightarrow$. Les classes $\mathbf{R}_1, \mathbf{R}_2 \ldots$ sont appelées les **classes de récurrences** et on a la partition suivante de l'espace des états :

$$E = \mathbf{T} \cup \mathbf{R} = \mathbf{T} \cup \mathbf{R}_1 \cup \mathbf{R}_2 \cup \ldots .$$

On note que $\mathbf{T}$ ou $\mathbf{R}$ peuvent être vides mais que les deux ne peuvent être vides à la fois car $E$ est non-vide. □

Supposons que $\mathbf{R}$ ne soit pas vide et soit $\mathbf{R}_\ell$, une classe de récurrence. Si $i \in \mathbf{R}_\ell$ et si $i \to j$, alors $j$ est récurrent et $i \leftrightarrow j$, par le lemme de propagation de la récurrence I.3.5. Cela implique donc que $j \in \mathbf{R}_\ell$. On voit donc que c'est une partie $Q$-fermée. Autrement dit, si la chaîne visite à un certain temps un état récurrent, elle reste piégée par la suite dans la classe de récurrence de cet état (ce d'après le lemme I.2.8) et elle évolue dans cette classe $\mathbf{R}_\ell$ comme une chaîne de matrice de transition $Q_{|\mathbf{R}_\ell}$. Le lemme de propagation de la récurrence I.3.5 indique qu'elle visite alors tous les états de $\mathbf{R}_\ell$ une infinité de fois.

La classe des états transients n'est pas nécessairement fermée : une chaîne peut commencer par ne visiter que des états transients puis à un certain instant, elle visite pour la première fois un état récurrent ; à partir de cet instant la chaîne reste piégée dans la classe de récurrence de cet état récurrent et visite une infinité de fois tous les états de cette classe.

Il se peut également que la chaîne ne visite aucun état récurrent. Dans ce cas, elle ne visite qu'un nombre fini de fois chaque état car si elle visitait une infinité de fois un état, la propriété de Markov au premier temps d'atteinte de cet état impliquerait que cet état soit récurrent.

Nous énonçons sous forme de théorème cette analyse du comportement d'une chaîne : ce théorème est appelé théorème de classification des états. Pour l'énoncer, on introduit pour tout sous-ensemble $F \subset E$ le temps d'arrêt suivant :

$$T_F = \inf \{n \in \mathbb{N} : \ X_n \in F\} ,$$

avec la convention $\inf \emptyset = \infty$, c'est-à-dire que $\{T_F < \infty\} = \{\exists n \in \mathbb{N} \ : \ X_n \in F\}$.

**Théorème I.3.7** **(Classification des états)** *Pour toute loi d'entrée $\mu \in \mathscr{M}_1(E)$, les assertions suivantes sont vérifiées $\mathbf{P}_\mu$-presque sûrement.*

*(i) Si $\mathbf{R}_\ell$ est une classe de récurrence telle que $T_{\mathbf{R}_\ell} < \infty$, alors :*

*(a) pour tout $n \geq T_{\mathbf{R}_\ell}$, $X_n \in \mathbf{R}_\ell$ ;*

*(b) pour tout $0 \leq n < T_{\mathbf{R}_\ell}$, $X_n \in \mathbf{T}$ ;*

*(c) pour tout $j \in \mathbf{R}_\ell$, on a $N_j = \infty$.*

*(ii) Si $T_{\mathbf{R}} = \infty$, c'est-à-dire si la chaîne n'atteint aucun état récurrent, alors pour tout $j \in E$, on a $N_j < \infty$.*

**Preuve :** on suppose qu'il y a au moins un état récurrent. Alors, il existe une ou plusieurs classes de récurrence comme définies dans l'énoncé. Soit $\mathbf{R}_\ell$, une telle classe. Supposons que $i \in \mathbf{R}_\ell$ et que $j \in E$ soit tel que $p(i,j) > 0$. Alors $i \to j$. Le lemme I.3.5 entraîne que $i \leftrightarrow j$ et que $j$ est récurrent. Par conséquent $j \in \mathbf{R}_\ell$. Cela montre que $\mathbf{R}_\ell$ est une classe fermée.

On fixe ensuite $\mu \in \mathscr{M}_1(E)$. On suppose qu'il y a au moins une classe de récurrence $\mathbf{R}_\ell$. Comme elle est fermée, pour tout $i \in \mathbf{R}_\ell$, et pour tout $n \geq 0$, on a $\mathbf{P}_i(X_n \in \mathbf{R}_\ell) = 1$ (voir (I.42). Par conséquent, $\mathbf{P}_i(\forall n \in \mathbb{N},\ X_n \in \mathbf{R}_\ell) = 1$. La propriété de Markov forte implique que $\mathbf{P}_\mu$-p.s.

$$\mathbf{E}_\mu\big[\mathbf{1}_{\{T_{\mathbf{R}_\ell}<\infty\}}\mathbf{1}_{\{\exists n\in\mathbb{N}:X_{n+T_{\mathbf{R}_\ell}}\notin\mathbf{R}_\ell\}}|\mathscr{F}_{T_{\mathbf{R}_\ell}}\big] = \mathbf{1}_{\{T_{\mathbf{R}_\ell}<\infty\}}\mathbf{P}_{X_{T_{\mathbf{R}_\ell}}}\big(\exists\, n \in \mathbb{N} : X_n \notin \mathbf{R}_\ell\big) = 0.$$

En intégrant, on a $\mathbf{P}_\mu(T_{\mathbf{R}_\ell} < \infty\,;\, \exists n \geq T_{\mathbf{R}_\ell} : X_n \notin \mathbf{R}_\ell) = 0$, ce qui montre (a).

Soient $\mathbf{R}_k$ et $\mathbf{R}_\ell$ deux classes de récurrences distinctes. Elles sont fermées et le point (a) entraîne facilement que $\mathbf{P}_\mu(T_{\mathbf{R}_k} < \infty\,;\, T_{\mathbf{R}_\ell} < \infty) = 0$. De plus, par définition, on a $X_n \notin \mathbf{R}_\ell$, pour tout $n < T_{\mathbf{R}_\ell}$, et donc $\mathbf{P}_\mu(T_{\mathbf{R}_\ell} < \infty\,;\, \exists n < T_{\mathbf{R}_\ell} \,:\, X_n \in \mathbf{R}) = 0$, ce qui prouve (b).

On remarque tout d'abord que le lemme I.3.5 implique que pour tout $i, j \in \mathbf{R}_\ell$, on ait $\mathbf{P}_i(N_j = \infty) = 1$. Par conséquent, pour tout $i \in \mathbf{R}_\ell$, on a $\mathbf{P}_i(\forall j \in \mathbf{R}_\ell\,,\, N_j = \infty) = 1$. On remarque ensuite que

$$\{T_{\mathbf{R}_\ell} < \infty\} \cap \{\exists j \in \mathbf{R}_\ell \,:\, N_j < \infty\} = \{T_{\mathbf{R}_\ell} < \infty\} \cap \{\exists j \in \mathbf{R}_\ell \,:\, N_j(\theta_{T_{\mathbf{R}_\ell}}\mathbf{X}) < \infty\}. \tag{I.64}$$

Or la propriété de Markov, combinée avec ce qui précède, implique que $\mathbf{P}_\mu$-p.s.

$$\mathbf{E}_\mu\big[\mathbf{1}_{\{T_{\mathbf{R}_\ell}<\infty\}}\mathbf{1}_{\{\exists j\in\mathbf{R}_\ell:N_j(\theta_{T_{\mathbf{R}_\ell}}\mathbf{X})<\infty\}}|\mathscr{F}_{T_{\mathbf{R}_\ell}}\big] = \mathbf{1}_{\{T_{\mathbf{R}_\ell}<\infty\}}\mathbf{P}_{X_{T_{\mathbf{R}_\ell}}}\big(\exists j \in \mathbf{R}_\ell \,:\, N_j < \infty\big) = 0.$$

En intégrant et en utilisant (I.64), on obtient que $\mathbf{P}_\mu(T_{\mathbf{R}_\ell} < \infty\,;\, \exists j \in \mathbf{R}_\ell \,:\, N_j < \infty) = 0$, ce qui prouve (c).

Il reste à montrer $(ii)$ : on suppose que $i \in E$ est tel que $\mathbf{P}_\mu(N_i = \infty) > 0$. On rappelle que $T_{\{i\}}$ est le premier temps d'atteinte de l'état $i$ et on observe que $\{N_i = \infty\} \subset \{T_{\{i\}} < \infty\}$. Cela montre que $\mathbf{P}_\mu(T_{\{i\}} < \infty) > 0$. La propriété de Markov entraîne ensuite que

$$\begin{aligned} \mathbf{P}_\mu\text{ - p.s.}\quad \mathbf{E}_\mu\big[\mathbf{1}_{\{N_i=\infty\}}|\mathscr{F}_{T_{\{i\}}}\big] &= \mathbf{E}_\mu\big[\mathbf{1}_{\{T_{\{i\}}<\infty\}}\mathbf{1}_{\{N_i(\theta_{T_{\{i\}}}\mathbf{X})=\infty\}}|\mathscr{F}_{T_{\{i\}}}\big] \\ &= \mathbf{1}_{\{T_{\{i\}}<\infty\}}\mathbf{P}_i(N_i = \infty)\,. \end{aligned}$$

En intégrant, on a donc $0 < \mathbf{P}_\mu(N_i = \infty) = \mathbf{P}_\mu(T_{\{i\}} < \infty)\mathbf{P}_i(N_i = \infty)$. Par conséquent $\mathbf{P}_i(N_i = \infty) > 0$, et le corollaire I.3.3 entraîne donc que $\mathbf{P}_i(N_i = \infty) = 1$ et que $i$ est récurrent.

On a donc montré que si $\mathbf{P}_\mu(N_i = \infty) > 0$, alors $i$ est récurrent. Autrement dit, si $i$ est transient, alors $\mathbf{P}_\mu(N_i = \infty) = 0$. Cela implique que $\mathbf{P}_\mu(\exists i \in \mathbf{T} \,:\, N_i = \infty) = 0$. Ensuite on constate facilement que si $T_{\mathbf{R}} = \infty$, alors $N_j = 0$, pour tout $j \in \mathbf{R}$. Par conséquent

$$\mathbf{P}_\mu(T_{\mathbf{R}} = \infty\,;\, \exists i \in E \,:\, N_i = \infty) \leq \mathbf{P}_\mu(\exists i \in \mathbf{T} \,:\, N_i = \infty) = 0,$$

ce qui montre $(ii)$. ■

On peut rapidement réexprimer le théorème de classification comme suit. Une chaîne a deux types de comportement asymptotique :

— *ou bien*, avec une certaine probabilité, elle a un comportement récurrent, c'est-à-dire qu'elle visite un nombre fini d'états transients avant d'atteindre une classe de récurrence où elle reste indéfiniment en y visitant tous les états une infinité de fois ;

— *ou bien*, avec une certaine probabilité, elle a un comportement transient, c'est-à-dire qu'elle ne visite que des états transients et dans ce cas elle ne visite qu'un nombre fini de fois chaque état.

En général la probabilité qu'elle ait un comportement récurrent et la probabilité qu'elle ait un comportement transient sont des probabilités qui peuvent être strictement positives et strictement inférieures à 1 et ces quantités ne sont pas toujours faciles à calculer. Mais lorsque la chaîne est irréductible, la situation est beaucoup plus simple, comme le montre le corollaire suivant.

**Corollaire I.3.8** *On suppose que $Q$ est irréductible. Alors $E = \mathbf{T}$ ou $E = \mathbf{R} = \mathbf{R}_1$, c'est-à-dire que l'on a l'alternative suivante.*

*(i) Ou bien tous les états sont récurrents et pour toute $\mu \in \mathscr{M}_1(E)$, on a*

$$\mathbf{P}_\mu\text{-}p.s. \quad \forall\, i \in E\,, \quad N_i = \infty\ .$$

*(ii) Ou bien tous les états sont transients et pour toute $\mu \in \mathscr{M}_1(E)$, on a*

$$\mathbf{P}_\mu\text{-}p.s. \quad \forall\, i \in E\,, \quad N_i < \infty\ .$$

*Par conséquent, si tous les états sont transients, $E$ est nécessairement infini et toute chaîne de Markov irréductible à valeurs dans un espace d'états fini est récurrente.*

**Preuve :** si $Q$ est irréductible et s'il existe un état récurrent, alors le lemme I.3.5 implique que tous les états sont récurrents. Comme tous les états communiquent, il n'existe qu'une seule classe de récurrence qui est tout l'espace d'état et le théorème I.3.7 de classification implique $(i)$. Si tous les états sont transients, on est clairement dans le cas $(ii)$ du théorème I.3.7 de classification qui implique immédiatement le point $(ii)$ du corollaire.

Supposons ensuite que $E$ soit un ensemble fini. On constate donc que $\{\forall i \in E\,,\, N_i < \infty\} = \{\sum_{i\in E} N_i < \infty\}$. Or $\{\sum_{i\in E} N_i < \infty\} = \emptyset$ car la chaîne reste indéfiniment dans $E$. On a donc $\mathbf{P}_\mu(\forall i \in E\,,\, N_i < \infty) = 0$. Or si tous les états sont transients, on doit avoir $\mathbf{P}_\mu(\forall i \in E\,,\, N_i < \infty) = 1$, ce qui termine la preuve. ■

**Exemple I.3.9** Soit $(Z_n)_{n\geq 0}$ un processus de branchement de loi de reproduction $\xi$. On a montré à l'exemple I.1.b que $(Z_n)_{n\geq 0}$ est une chaîne de Markov dont la matrice de transition $Q = (p(i,j))_{i,j\in\mathbb{N}}$ est donnée par $p(i,j) = \xi^{*i}(j)$ (voir l'exemple I.1.b, page 9). On rappelle que $\varphi$ désigne la fonction génératrice de $\xi$.

On s'intéresse aux classes de récurrence et de transience du processus de branchement en se plaçant dans le cas le plus intéressant où

$$\xi(0) > 0 \quad \text{et} \quad \xi(0) + \xi(1) < 1\ .$$

On vérifie d'abord que $\{0\}$, qui est un élément absorbant est l'unique classe de récurrence : on voit en effet que pour tout $i \in \mathbb{N}$, $p(i,0) = \xi(0)^i > 0$ ; on a donc $i \to 0$ ; comme 0 est absorbant, si $i \neq 0$, on n'a pas $0 \to i$. Cela montre que $\mathbb{N}^*$ est l'ensemble des états transients :

$$\mathbf{R} = \mathbf{R}_1 = \{0\} \quad \text{et} \quad \mathbf{T} = \mathbb{N}^* .$$

Le théorème de classification des états permet alors d'affirmer que pour toute $\mu \in \mathscr{M}_1(\mathbb{N})$, $\mathbf{P}_\mu$-p.s. : *ou bien* il existe un temps $n$ (aléatoire) où $Z_n = 0$ et la chaîne reste absorbée en 0 (la population s'éteint), *ou bien* $\lim_n Z_n = \infty$, c'est-à-dire

$$\mathbf{P}_\mu(\exists n \in \mathbb{N} \,:\, \forall m \in \mathbb{N},\, Z_{n+m} = 0\,) + \mathbf{P}_\mu\big(\lim_{n\to\infty} Z_n = \infty\big) = 1 .$$

Cela démontre de nouveau le lemme I.1.12, page 13, qui a été prouvé à l'aide de technique de martingales. On rappelle que $q$ désigne la plus petite racine de $\varphi(r) = r$ (voir le lemme I.1.10, page 11). Il est facile de déduire que

$$\mathbf{P}_\mu(\exists n \in \mathbb{N} \,:\, \forall m \in \mathbb{N},\, Z_{n+m} = 0\,) = \mathbf{E}_\mu\big[q^{Z_0}\big] = \sum_{i\in E} q^i \mu(i) .$$

Cette probabilité (de tomber dans une classe de récurrence) peut, selon les lois de reproduction $\xi$, prendre toutes les valeurs réelles entres 0 et 1. □

**Exemple I.3.10** (*Le problème de Polyà*) On s'intéresse à la récurrence ou à la transience des marches aléatoires simples symétriques dans $\mathbb{Z}^d$. On note $(\mathbf{e}_1, \ldots, \mathbf{e}_d)$, la base canonique de $\mathbb{R}^d$. On définit la mesure de probabilité $\pi$ sur $\mathbb{Z}^d$ par $\pi(\mathbf{e}_1) = \pi(-\mathbf{e}_1) = \ldots = \pi(\mathbf{e}_d) = \pi(-\mathbf{e}_d) = \frac{1}{2d}$. Pour simplifier, on se place sur $(\Omega, \mathscr{F}, \mathbf{P})$, sur lequel on suppose définie une suite i.i.d. de v.a. notée $(\xi_n)_{n\geq 1}$ de loi commune $\pi$ et on pose $X_0 = 0$ et $X_n = \xi_1 + \ldots + \xi_n$, $n \geq 1$. Il est facile de montrer que la matrice de transition d'une telle marche est irréductible. Le corollaire I.3.8 montre l'alternative suivante.

- La marche est récurrente ssi $\mathbf{E}[N_0] = \sum_{n\geq 1} \mathbf{P}(X_n = 0) = \infty$.
- La marche est transiente ssi $\mathbf{E}[N_0] = \sum_{n\geq 1} \mathbf{P}(X_n = 0) < \infty$.

*Cas de la dimension* $d = 1$. Un argument combinatoire simple montre que $\mathbf{P}(X_{2n+1} = 0) = 0$ et que

$$\mathbf{P}(X_{2n} = 0) = 2^{-2n} \binom{2n}{n} = 2^{-2n} \frac{(2n)!}{(n!)^2} .$$

On rappelle la formule de Stirling $n! \sim \sqrt{2\pi n}(n/e)^n$, ce qui implique que

$$\mathbf{P}(X_{2n} = 0) \sim \frac{1}{\sqrt{\pi n}} .$$

On a donc $\mathbf{E}[N_0] = \sum_{n\geq 1} \mathbf{P}(X_n = 0) = \infty$, et la marche est récurrente.

*Cas de la dimension* $d = 2$. Là aussi, un argument combinatoire simple montre que $\mathbf{P}(X_{2n+1} = 0) = 0$. On trie les trajectoires de longueur $2n$ qui partent de l'origine 0 et qui y reviennent, selon le nombre $\ell \in \{0, \ldots, n\}$ de fois où elles sont montées vers le Nord : elles ont donc été $\ell$ fois vers le Sud, $n - \ell$ vers l'Ouest ainsi que vers l'Est. On a donc

$$\mathbf{P}(X_{2n} = 0) \quad = \quad 4^{-2n} \sum_{\ell=0}^{n} \frac{(2n)!}{(2\ell)!(2n-2\ell)!} \frac{(2\ell)!}{(\ell!)^2} \frac{(2n-2\ell)!}{((n-\ell)!)^2}$$

$$
\begin{aligned}
&= 4^{-2n}\frac{(2n)!}{(n!)^2}\sum_{\ell=0}^{n}\frac{n!}{\ell!(n-\ell)!}\frac{n!}{(n-\ell)!\ell!} \\
&= 4^{-2n}\frac{(2n)!}{(n!)^2}\cdot\frac{(2n)!}{(n!)^2} \ ,
\end{aligned}
$$

ce qui implique que

$$
\mathbf{P}(X_{2n}=0) \underset{n\to\infty}{\sim} \frac{1}{\pi n} \ .
$$

On a donc $\mathbf{E}[N_0]=\sum_{n\geq 1}\mathbf{P}(X_n=0)=\infty$, et la marche est récurrente.

*Cas de la dimension* $d=3$. Un argument combinatoire simple montre que $\mathbf{P}(X_{2n+1}=0)=0$. En triant les trajectoires de la même façon on a

$$
\mathbf{P}(X_{2n}=0)=6^{-2n}\sum_{k+\ell+m=n}\frac{(2n)!}{(k!\ell!m!)^2} \ = \ 2^{-2n}\frac{(2n)!}{(n!)^2}\cdot\sum_{k+\ell+m=n}\Big(\frac{3^{-n}n!}{k!\ell!m!}\Big)^2 \ .
$$

La loi multinomiale implique que $3^n=(1+1+1)^n=\sum_{k+\ell+m=n}\frac{n!}{k!\ell!m!}$. On a donc

$$
\mathbf{P}(X_{2n}=0)\leq 2^{-2n}\frac{(2n)!}{(n!)^2}\cdot M_n \quad \text{où} \quad M_n=\max_{k+\ell+m=n}3^{-n}\frac{n!}{k!\ell!m!}.
$$

Supposons que $k+\ell+m=n$ et que $0\leq k\leq\ell\leq m\leq n$ avec $k<m$. On a $\frac{k!m!}{(k+1)!(m-1)!}=\frac{n}{k+1}\leq 1$, ce qui entraîne que $\frac{n!}{k!\ell!m!}\leq\frac{n!}{(k+1)!\ell!(m-1)!}$. Cela implique par symétrie, que si $k_n+\ell_n+m_n=n$, alors

$$
\frac{n!}{k_n!\ell_n!m_n!}=M_n \implies \Big|k_n-\frac{n}{3}\Big| \ , \ \Big|\ell_n-\frac{n}{3}\Big| \ , \ \Big|m_n-\frac{n}{3}\Big| \leq 1 \ .
$$

Par la formule de Stirling, il existe une constante $C$, telle que $M_n\leq C/n$. Il existe donc une constante $C^*$ telle que

$$
\mathbf{P}(X_{2n}=0)\leq\frac{C^*}{n^{3/2}} \ .
$$

On a donc $\mathbf{E}[N_0]=\sum_{n\geq 1}\mathbf{P}(X_n=0)<\infty$, et la marche est transiente.

*Cas des dimensions* $d\geq 4$. On introduit la notation suivante

$$
X_n=(X_n^1,\ldots,X_n^d) \quad \text{et} \quad Y_n=(X_n^1,X_n^2,X_n^3)
$$

où $X_n^1,\ldots,X_n^d$ sont les coordonnées de $X_n$ dans la base canonique $(\mathbf{e}_1,\ldots,\mathbf{e}_d)$. On introduit également les temps aléatoires $\tau_k$, $k\geq 0$, définis par

$$
\tau_0=0 \quad \text{et} \quad \forall k\geq 0 \ , \quad \tau_{k+1}=\inf\{n>\tau_k \ : \ Y_n\neq Y_{\tau_k}\} \ .
$$

Il est facile de voir que les variables $(\tau_{k+1}-\tau_k)_{k\geq 0}$ sont indépendantes et de loi géométrique :

$$
\mathbf{P}(\tau_{k+1}-\tau_k=m)=\tfrac{3}{d}\big(1-\tfrac{3}{d}\big)^{m-1} \ , \quad m\geq 1 \ .
$$

On voit également que $(Y_{\tau_k})_{k\geq 0}$ est une marche aléatoire simple symétrique sur $\mathbb{Z}^3$. Elle est donc transiente : elle ne visite qu'un nombre fini de fois chaque site, ce qui implique que $\lim_k\|Y_{\tau_k}\|=\infty$, $\mathbf{P}$-presque sûrement.

Ici $\|\cdot\|$ désigne la norme euclidienne. Comme $Y_n$ est constante pour tout $\tau_k \le n < \tau_{k-1}$, on a donc $\lim_n \|Y_n\| = \infty$, $\mathbf{P}$-presque sûrement. Enfin on remarque que $\|Y_n\|^2 \le \|X_n\|^2$, ce qui implique que

$$\mathbf{P}-\text{p.s.} \quad \lim_n \|X_n\| = \infty \ .$$

La marche aléatoire est donc transiente en dimension $d \ge 4$.

**En résumé**, *on a montré qu'en dimensions $d = 1, 2$, la marche aléatoire simple est irréductible récurrente : quelle que soit sa loi initiale, presque sûrement elle visite tous les points de $\mathbb{Z}^d$ une infinité de fois. En dimension $d \ge 3$, la marche aléatoire simple est irréductible transiente : quelle que soit sa loi initiale, presque sûrement elle ne visite les points de $\mathbb{Z}^d$ qu'un nombre fini de fois (et elle tend en norme vers l'infini).*

Ce résultat est connu sous le nom de *théorème de Polyà*. À la fin des années 20, Polyà a posé la question de la récurrence ou de la transience des marches simples sur $\mathbb{Z}^d$ comme un exercice dans une revue allemande. Il en a rédigé la solution une ou deux années plus tard. Ce théorème est le premier d'une longue série concernant la récurrence ou la transience des marches aléatoires, problème qui requiert en général des techniques plus sophistiquées que celles que nous avons présentées ici (voir les exercices pour quelques idées). $\square$

**Remarque I.3.11** La figure I.3.10 page 67 montre le graphe de $n \mapsto \|X_n\|$, où $n \in \{1, \ldots, 50000\}$. En dimension 1 la marche aléatoire visite assez souvent 0 (à peu près $\sqrt{50000} \asymp 220$ fois). Notez cependant qu'il y a des excursions assez longues en dehors de 0 (comme la dernière qui dure plus de 20000 étapes). En dimension 2 la marche aléatoire visite 0 moins souvent (à peu près $\log(50000) \asymp 10$ fois) et les excursions en dehors de 0 sont assez longues.

La figure I.3.10 page 67 montre le graphe de $\|X_n\|$ en dimensions $d = 3$ et $d = 20$ : la marche semble ne jamais revenir à 0, ce qui est cohérent avec le fait qu'en dimension $d \ge 3$, la marche est transitoire.

La figure I.7 page 68 montre le graphe de la marche aléatoire sur $\mathbb{Z}^{500}$. et compare $\|X_n\|$ avec $\sqrt{n}$. En fait, d'après le théorème de la limite centrale, en toute dimension $d$ (y compris les dimensions 1 et 2) la v.a. $\|X_n\|/\sqrt{n}$ tend en distribution à la loi de la norme d'une v.r. gaussienne standard en dimension $d$ qui est la racine carrée de $d$ v.a. indépendantes de Rayleigh. Ainsi, si $d = 500$, dans l'intervalle $n \in \{25, 20000\}$, la loi des grands nombres s'applique grossièrement à la distribution limite et le graphe semble assez proche de celui de $n \mapsto \sqrt{n}$, mais l'image n'est plus valable pour $n >> 50000$. $\square$

Pour conclure, nous donnons un critère de récurrence des chaînes irréductibles utilisant la théorie des martingales (ce critère servira par la suite).

**Proposition I.3.12** *On suppose $Q$ irréductible. Alors la chaîne est récurrente ssi toute fonction positive $Q$-sur-harmonique est constante.*

**Preuve :** on suppose d'abord que la chaîne est récurrente. Soit $f : E \to \mathbb{R}_+$ telle que $f \ge Q.f$ ($f$ est sur-harmonique). On pose $Y_n = f(X_n)$, $n \ge 0$. La proposition I.2.21 montre que pour tout $i \in E$, sous $\mathbf{P}_i$, $(Y_n)_{n\ge 0}$ est une sur-martingale positive relativement à la filtration $(\mathscr{F}_n)_{n\ge 0}$. Le théorème de convergence des sur-martingales positives implique qu'il existe $Y_\infty : \Omega \to [0, \infty]$, telle que $\mathbf{P}_i$-p.s. on ait $\lim_n Y_n = Y_\infty$. Comme les sur-martingales décroissent en espérance, on a de plus $\mathbf{E}_i[Y_n] \le \mathbf{E}_i[Y_0] = f(i)$. Par Fatou, on en déduit que $\mathbf{E}_i[Y_\infty] \le f(i) < \infty$. Cela implique donc que $\mathbf{P}_i(Y_\infty < \infty) = 1$. On a donc

$$\mathbf{P}_i - \text{p.s.} \quad \lim_n f(X_n) = Y_\infty < \infty \ .$$

Comme la chaîne est supposée récurrente, pour tout $j \in E$, $\mathbf{P}_i$-p.s. l'ensemble des temps $\{n \in \mathbb{N} : X_n = j\}$ est infini. Donc, pour tout $j \in E$, on a $\mathbf{P}_i(Y_\infty = f(j)) = 1$, ce qui entraîne que $\mathbf{P}_i(\forall j \in E\,, Y_\infty = f(i) = f(j)) = 1$, et donc que $f(i) = f(j)$, pour tout $j \in E$. La fonction sur-harmonique $f$ est donc constante.

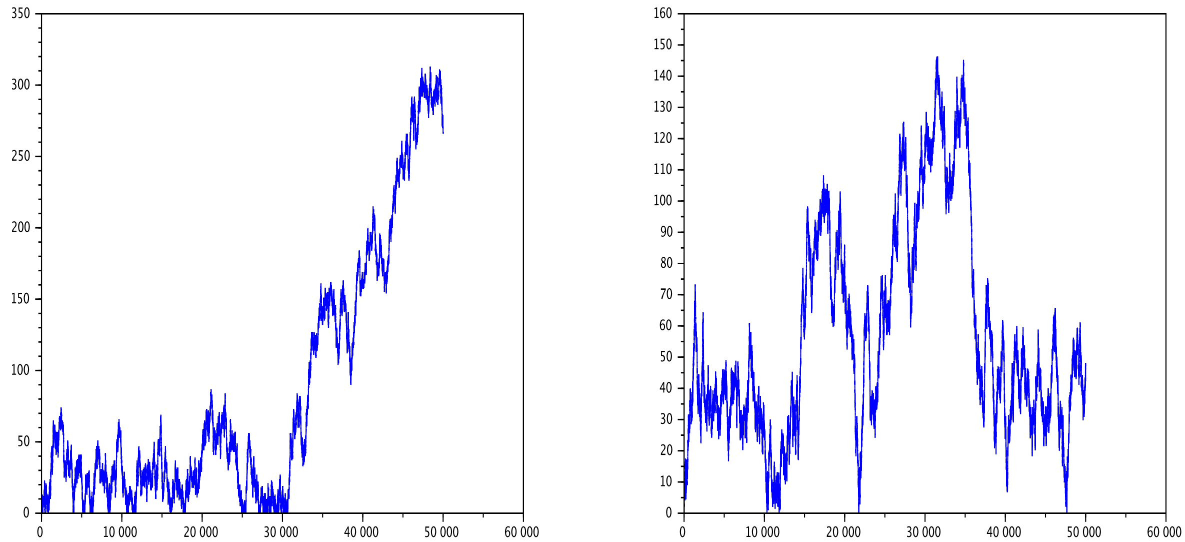


FIGURE I.5 – Marches simples symétriques sur $\mathbb{Z}^d$, graphe de $n \mapsto \|X_n\|$, $0 \leq n \leq 5.10^4$. A gauche : $d = 1$, à droite : $d = 2$.

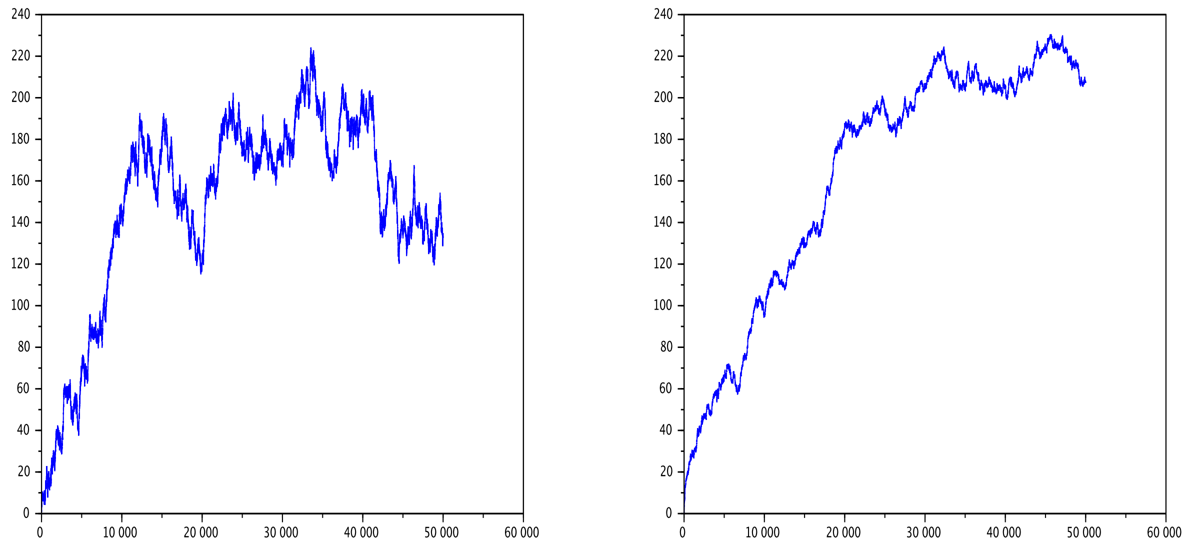


FIGURE I.6 – Marches simples symétriques sur $\mathbb{Z}^d$, graphe de $n \mapsto \|X_n\|$, $0 \leq n \leq 5.10^4$. A gauche : $d = 3$, à droite : $d = 20$.

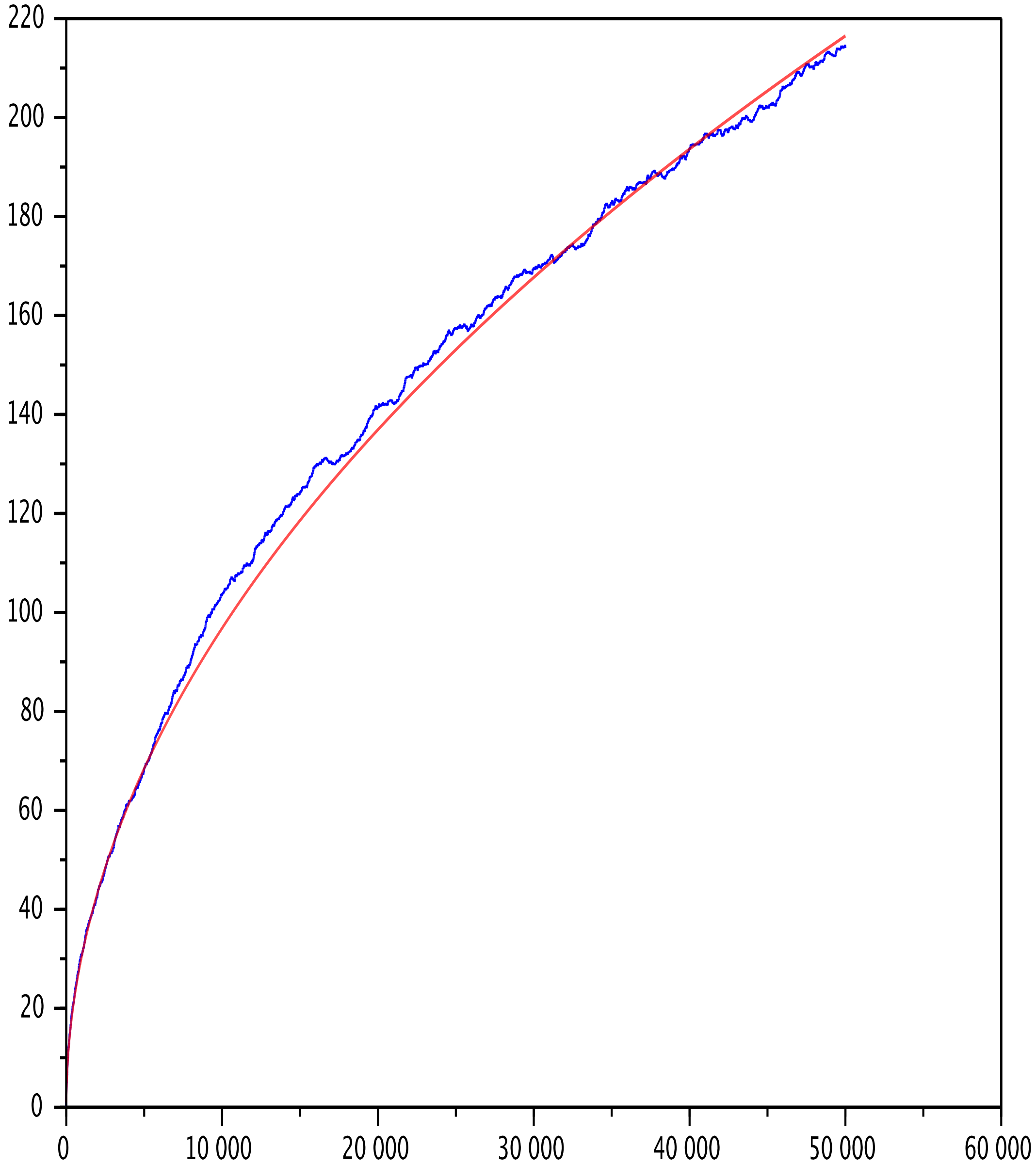


FIGURE I.7 – Marche simple symétrique sur $\mathbb{Z}^{500}$, graphe de $n \mapsto \|X_n\|$ et de $n \mapsto \sqrt{n}$ pour $0 \leq n \leq 5.10^4$.

Montrons la réciproque. Pour tout $i_0 \in E$, et pour tout $i \in E$, on pose $f(i) = \mathbf{P}_i(T_{i_0} < \infty)$ et $f^*(i) = \mathbf{P}_i(T_{i_0}^{(1)} < \infty)$, où on rappelle que $T_{i_0}$ est le premier temps d'atteinte de $i_0$ et $T_{i_0}^{(1)}$ est le premier temps de retour en $i_0$. Le lemme I.2.23 implique que $f$ et $f^*$ sont sur-harmoniques. Elles sont donc constantes, selon notre hypothèse. Il existe alors deux constantes $c, c^* \geq 0$, telles que $f \equiv c$ et $f^* \equiv c^*$. Or $f(i_0) = 1$, par définition. Donc $c = 1$. De plus $f(i) = f^*(i)$, pour tout $i \neq i_0$, donc $c = c^* = 1$. Cela montre que $f^*(i_0) = \mathbf{P}_{i_0}(T_{i_0}^{(1)} < \infty) = 1$ et donc $i_0$ est récurrent. Comme la chaîne est supposée irréductible, tous les états sont récurrents. ■

**Exemple I.3.13** On considère une chaîne de naissance et de mort, c'est-à-dire que sa matrice de transition $Q = (p(i,j))_{i,j \in \mathbb{N}}$ est telle que $p(i,j) = 0$ dès que $|i-j| \geq 2$. On la suppose irréductible, ce qui est équivalent à supposer que $p(i,i+1)p(i+1,i) > 0$, pour tout $i \in \mathbb{N}$. Pour tout $i \in \mathbb{N}$, on pose $T_i = \inf\{n \geq 0 \, : \, X_n = i\}$ avec la convention que $T_i = \infty$ ssi pour tout $n \in \mathbb{N}$, $X_n \neq i$. $T_i$ est un $(\mathscr{F}_n)_{n \geq 0}$-temps d'arrêt : c'est le premier temps d'atteinte de l'état $i$.

Soient $i, a \in \mathbb{N}$ tels que $i \leq a$. Comme la chaîne est irréductible, elle est soit transiente soit récurrente mais dans les deux cas $\mathbf{P}_i(T_a < \infty) = 1$. Soit $f : \mathbb{N} \to \mathbb{R}_+$ une fonction $Q$-harmonique sur $\mathbb{N}^*$ *qui n'est pas constante*, comme on en a construit à l'exemple I.2.22, page 45. La proposition I.2.21 page 44 montre que pour toute loi d'entrée $\mu \in \mathscr{M}_1(\mathbb{N})$, sous $\mathbf{P}_\mu$, $(Y_n)_{n \geq 0} = (f(X_{n \wedge T_0}))_{n \geq 0}$ est une $(\mathscr{F}_n)_{n \geq 0}$-martingale. La proposition **??**, page **??**, implique que sous $\mathbf{P}_\mu$, $(Y_{n \wedge T_a})_{n \geq 0}$ est une $(\mathscr{F}_n)_{n \geq 0}$-martingale également. Or on a

$$\forall n \in \mathbb{N}, \quad 0 \leq Y_{n \wedge T_a} \leq \max(f(0), \ldots, f(a)) \, .$$

C'est donc une martingale bornée. De plus,

$$\mathbf{P}_i\text{-p.s.} \quad \lim_{n \to \infty} Y_{n \wedge T_a} = f(a) \mathbf{1}_{\{T_a < T_0\}} + f(0) \mathbf{1}_{\{T_0 < T_a\}} \, .$$

Or $Y_0 = f(i)$, $\mathbf{P}_i$-p.s. Le théorème de convergence p.s. des martingales combiné avec de la convergence dominée, implique que

$$f(i) = \mathbf{E}_i[Y_0] = \mathbf{E}[Y_{n \wedge T_a}] = \mathbf{E}\big[\lim_{n \to \infty} Y_{n \wedge T_a}\big] = f(a) \mathbf{P}_i(T_a < T_0) + f(0) \mathbf{P}_i(T_0 < T_a).$$

ce qui implique que

$$\mathbf{P}_i(T_0 < T_a) = \frac{f(a) - f(i)}{f(a) - f(0)} \, , \quad 0 < i < a \, .$$

On observe ensuite que $\mathbf{P}_i$-p.s., $T_a \geq a - i$, et donc $\lim_{a \to \infty} T_a = \infty$. Par conséquent $\mathbf{P}_i(T_0 < \infty) = \lim_{a \to \infty} \uparrow \mathbf{P}_i(T_0 < T_a)$. On choisit $f(0) = 0$ et $f(1) = 1$. Alors, on pose

$$L = \lim_{a \to \infty} f(a) = \sum_{n \geq 1} \prod_{1 \leq k \leq n} \frac{p(k, k-1)}{p(k, k+1)} \; \in \; [0, \infty] \, .$$

Si $L = \infty$, alors $\mathbf{P}_i(T_0 < \infty) = 1$, pour tout $i \in \mathbb{N}^*$. En appliquant Markov fort, cela montre que

$$\begin{aligned} \mathbf{P}_i(T_i^{(1)} = \infty) &= \mathbf{P}_i(T_0 < \infty; T_i^{(1)} = \infty) = \mathbf{P}_i(T_0 < \infty; T_i^{(1)}(\theta_{T_0} \mathbf{X}) = \infty) \\ &= \mathbf{P}_i(T_0 < \infty) \mathbf{P}_0(T_i = \infty) = 0. \end{aligned}$$

Donc $\mathbf{P}_i(T_i^{(1)} < \infty) = 1$ et la chaîne est récurrente.

Si $L < \infty$, alors $\mathbf{P}_i(T_0 < \infty) = \frac{L - f(i)}{L - f(0)} < 1$. On a donc $\mathbf{P}_i(N_0 = 0) > 0$, ce qui montre que la chaîne est transiente. Par conséquent

— $Q$ est récurrente si $\sum_{n\geq 1}\prod_{1\leq k\leq n}\frac{p(k,k-1)}{p(k,k+1)}=\infty$.

— $Q$ est transiente si $\sum_{n\geq 1}\prod_{1\leq k\leq n}\frac{p(k,k-1)}{p(k,k+1)}<\infty$. □

## EXERCICES.

**Exercice I.30** Soit $\big(\Omega\,;\,\mathscr{F}\,;\,(\mathscr{F}_n)_{n\geq 0}\,;\,(X_n)_{n\geq 0}\,;\,Q\,;\,\mathbf{P}_\mu\,,\,\mu\in\mathscr{M}_1(E)\;\big)$, une chaîne de Markov. On suppose que la chaîne est **irréductible** (c'est-à-dire que $Q$ l'est). Soit $f:E\to\mathbb{R}_+$, une fonction. On utilise les notations suivantes : on dit que "$f$ tend vers $\infty$ en l'infini", ce qui est noté $\lim_\infty f=\infty$ ssi pour tout réel $c>0$, l'ensemble $\{i\in E : f(i)\leq c\}$ est fini. De même on dit que "$f$ tend vers 0 en l'infini" ssi pour tout réel $c>0$, l'ensemble $\{i\in E : c<f(i)\}$ est fini. Soit $F\subset E$ un sous-ensemble **fini** d'états. On suppose que $E$ est **infini**. On suppose que

$$\forall i\in E\backslash F\,,\quad Q.f(i)\leq f(i)\;.$$

On veut montrer les critères de récurrence et de transience suivants :

**(i)** Si $\lim_\infty f=\infty$, alors la chaîne est récurrente.

**(ii)** On suppose que $f(i)>0$, pour tout $i\in F$ et que $\lim_\infty f=0$. Alors la chaîne est transiente.

1. Montrer (et c'est un fait général) que puisque $Q$ est irréductible, pour toute loi d'entrée $\mu$, sous $\mathbf{P}_\mu$, la chaîne visite une infinité de site, c'est-à-dire que
$$\mathbf{P}_\mu\text{-p.s.}\quad \#\{X_n\,;\,n\in\mathbb{N}\}=\infty\;.$$
2. On pose $T_F^{(1)}=\inf\{n\geq 1 : X_n\in F\}$ avec la convention que $T_F^{(1)}=\infty$ ssi pour tout $n\geq 1$, on a $X_n\notin F$. Il est facile de voir que $T_F^{(1)}$ est un $(\mathscr{F}_n)_{n\geq 0}$-temps d'arrêt : c'est le premier temps de *retour* de la chaîne en $F$. On pose $Y_n:=f(X_{n\wedge T_F^{(1)}})$, $n\geq 0$. Montrer que la proposition I.2.21 page 44 montre que pour $i\in E\backslash F$, sous $\mathbf{P}_i$, $(Y_n)_{n\geq 0}$ est une $(\mathscr{F}_n)_{n\geq 0}$-sur-martingale positive. Elle converge presque sûrement vers $Y_\infty$ et montrer que
$$\forall i\in E\backslash F\,,\quad \mathbf{E}_i[Y_\infty]\leq f(i)\;.$$
En déduire que si $\lim_\infty f=\infty$, alors
$$\forall i\in E\,,\ \mathbf{P}_i\text{-p.s.}\quad \#\{X_{n\wedge T_F^{(1)}}\,;\,n\in\mathbb{N}\}\;<\;\infty\;.$$
3. Montrer que si $\lim_\infty f=\infty$, alors pour tout $i\in E$, on a $\mathbf{P}_i(T_F^{(1)}<\infty)=1$. En utilisant la propriété de Markov, montrer que pour tout $i\in E$, on a
$$\mathbf{P}_i\text{-p.s.}\quad \#\{n\in\mathbb{N}\ :\ X_n\in F\}=\infty\;.$$
Comme $F$ est fini, montrer que
$$\sum_{j\in F}N_j=\#\{n\in\mathbb{N}\ :\ X_n\in F\}$$
en déduire que la chaîne est récurrente, ce qui montre (**i**).
4. En raisonnant de façon proche, montrer (**ii**).
5. En appliquant ce critère aux processus de naissance et de mort. □

**Exercice I.31** Soit $\big(\Omega\,;\,\mathscr{F}\,;\,(\mathscr{F}_n)_{n\geq 0}\,;\,(X_n)_{n\geq 0}\,;\,Q\,;\,\mathbf{P}_\mu\,,\,\mu\in\mathscr{M}_1(E)\;\big)$, une chaîne de Markov. Soit $F\subset E$, un sous-ensemble. On note $T_F=\inf\{n\geq 0\ :\ X_n\in F\}$ avec la convention habituelle $T_F=\infty$ ssi pour tout $n\in\mathbb{N}$, on a $X_n\notin F$. On suppose qu'il existe $\eta>0$ et $m\in\mathbb{N}$ tels que

$$\forall i\in E\backslash F\,,\quad \mathbf{P}_i(X_m\in F)>\eta\;.$$

1. Montrer que
$$\forall k\in\mathbb{N}\,,\ \forall i\in E\,,\quad \mathbf{P}_i(T_F>km)\leq(1-\eta)^k\;.$$
2. En déduire que $\mathbf{E}_i[T_F]\leq m/\eta$ et donc que $\mathbf{P}_i(T_F<\infty)=1$, pour tout $i\in E$.

3. Montrer que pour tout $0 \le u - \frac{1}{m}\log(1-\eta)$, pour tout $i \in E$ on a $\mathbf{E}_i[e^{uT_F}] < \infty$. □

**Exercice I.32** Soit $\big(\Omega\,;\,\mathscr{F}\,;\,(\mathscr{F}_n)_{n\ge 0}\,;\,(X_n)_{n\ge 0}\,;\,Q\,;\,\mathbf{P}_\mu\,,\,\mu\in\mathscr{M}_1(E)\,\big)$, une chaîne de Markov.

1. Montrer qu'une classe de récurrence est réduite à un singleton $\{i\}$ ssi $i$ est absorbant.
2. On suppose que toutes les classes de la chaîne sont réduites à des points (donc des états absorbants). Décrire le comportment possible de la chaîne.
3. On suppose que $E$ est fini, que la chaîne possède $q$ états absorbants, notés $a_1,\dots,a_q$. On suppose également que
$$\forall i \in E\,,\quad \exists\, 1 \le k \le q\ :\ i \to a_k\ .$$
Montrer que la chaîne admet exactement $q$ classes de récurrence $R_1 = \{a_1\},\dots,R_q = \{a_q\}$ et que la classe des états transients est $T = E\backslash\{a_1,\dots,a_q\}$. Montrer que quelle que soit la loi d'entrée $\mu \in \mathscr{M}_1(E)$, sous $\mathbf{P}_\mu$ la chaîne va être absorbée :
$$\mathbf{P}_\mu(\exists n \in \mathbb{N}\ :\ X_n \in \{a_1,\dots,a_q\}\,) = 1\ .$$
4. On suppose que $E$ est fini, que la chaîne possède deux états absorbants notés $a$ et $b$. On suppose que
$$\forall i \in E\,,\quad \text{soit } i \to a, \text{ soit } i \to b.$$
On pose $A = \{\exists n \in \mathbb{N} :\ X_n = a\}$ et $B = \{\exists n \in \mathbb{N} :\ X_n = b\}$. Montrer que pour toute loi d'entréee $\mu \in \mathscr{M}_1(E)$, on a
$$\mathbf{P}_\mu(A \cap B) = 0 \quad \text{et} \quad \mathbf{P}_\mu(A \cup B) = \mathbf{P}_\mu(A) + \mathbf{P}_\mu(B) = 1\ .$$
On se donne $f : \mathbf{E} \to \mathbb{R}_+$ qui est $Q$-harmonique sur $E\backslash\{a,b\}$. Montrer que
$$\forall \mu \in \mathscr{M}_1(E)\ \langle \mu, f\rangle = f(a)\mathbf{P}_\mu(A) + f(b)\mathbf{P}_\mu(B)\ .$$
Quelle est l'hypothèse à faire sur $f$ et $\mu$ pour en déduire un calcul $\mathbf{P}_\mu(A)$ ?
5. On rappelle l'exercice I.9 page 8 où $Q$ est donnée de la manière suivante : on considère $N$ individus distincts qui portent deux sortes d'allèles d'un même gène : l'allèle $A$ et l'allèle $a$. La population évolue en restant de taille constante $N$ et les gènes se transmettent d'une génération à l'autre de la manière suivante (très simplifiée) : chaque individu de la génération $n$ choisit indépendamment uniformément un "père" à la génération précédente et il hérite de la forme ($A$ ou $a$) de son "père". On note $X_n$ le nombre d'individus porteur de la forme $A$ à la génération $n$. Montrer que $(X_n)_{n\ge 0}$ est une chaîne de Markov à valeurs dans $E = \{0,\dots,N\}$. Montrer que 0 et $N$ sont les seuls éléments absorbants. Montrer que $(X_n)_{n\ge 0}$ est une martingale. En déduire que si la population au temps 0 compte $i$ individus porteur de l'allèle $A$, la probabilité pour que tous les individus finissent par porter l'allèle $A$ est $i/N$ et la probabilité pour que tous les individus finissent par porter l'allèle $a$ est $(N-i)/N$. □

**Exercice I.33** (*Décomposition de Riesz*) Soit
$$\big(\Omega\,;\,\mathscr{F}\,;\,(\mathscr{F}_n)_{n\ge 0}\,;\,(X_n)_{n\ge 0}\,;\,Q\,;\,\mathbf{P}_\mu\,,\,\mu\in\mathscr{M}_1(E)\,\big)$$
une chaîne de Markov. Pour tous $i,j \in E$, on pose
$$U(i,j) = \delta_{i,j} + \sum_{n\ge 1}[Q^n](i,j) = \mathbf{E}_i\Big[\sum_{n\in\mathbb{N}} \mathbf{1}_{\{X_n = j\}}\Big] \ \in [0,\infty]\ ,$$
où $\delta_{i,j} = 1$ si $i = j$ et $\delta_{i,j} = 0$ si $i \ne j$. La matrice $(U(i,j))_{i,j\in E}$ est appelée la *matrice potentielle de* $Q$.

1. Soit $h : E \to \mathbb{R}_+$. On définit $U.h(i) = \sum_{j\in E} U(i,j)h(j) \in [0,\infty]$ avec la convention que si $U(i,j) = \infty$ et $h(j) = 0$, alors $U(i,j)h(j) = 0$. On a donc $U.h(i) = \mathbf{E}_i\Big[\sum_{n\in\mathbb{N}} h(X_n)\Big]$, pour tout $i \in E$. Montrer que $U.h : E \to [0,\infty]$ est $Q$-sur-harmonique. La fonction $U.h$ est appelée le *potentiel de* $h$.
2. Soit $f : E \to \mathbb{R}_+$ une fonction $Q$-sur-harmonique. Pour tout $i \in E$, on pose $h(i) = f(i) - Q.f(i)$. Montrer que $0 \le h(i) < \infty$, pour tout $i \in E$. Montrer que
$$\forall n \ge 1\,,\qquad \sum_{0\le k<n} Q^k.h = f - Q^n.f\ .$$
En déduire que pour tout $i \in E$, la suite $(Q^n.f(i))_{n\ge 0}$ est décroissante et qu'il existe une fonction $Q$-harmonique positive $f' : E \to \mathbb{R}_+$ telle que
$$f' = \lim_n \downarrow Q^n.f = \inf_{n\ge 1} Q^n.f\ .$$
Montrer également que $0 \le U.h(i) < \infty$, pour tout $i \in E$ et que $f = U.h + f'$.

3. Montrer que pour toute fonction $f : E \to \mathbb{R}_+$ qui est $Q$-sur-harmonique, il existe un *unique* couple de fonction $h, f' : E \to \mathbb{R}_+$ telles que les conditions suivantes soient satisfaites.

   — La fonction $f'$ est $Q$-harmonique.

   — On a $f = U.h + f'$.

   Cette décomposition est appelée *décomposition de Riesz de* $f$.

4. Si les fonctions $h, f'$ sont issues de la décomposition de Riesz de la fonction $Q$-sur-harmonique $f : E \to \mathbb{R}_+$, montrer que pour tout $i \in E$, on a
$$\mathbf{P}_i\text{-p.s.} \quad \lim_n f(X_n) \text{ existe dans } \mathbb{R}_+ \text{ et } \quad \lim_n f(X_n) = \lim_n f'(X_n) \ ,$$
   (*Indication : on montrera que* $(U.h(X_n))_{n\geq 0}$ *converge en norme* $\mathrm{L}^1$ *vers* 0.) □

**Exercice I.34** *Cet exercice est une application de la décomposition de Riesz obtenue à l'exercice précédent I.33.* Soit
$$\big(\Omega\,;\,\mathscr{F}\,;\,(\mathscr{F}_n)_{n\geq 0}\,;\,(X_n)_{n\geq 0}\,;\,Q\,;\,\mathbf{P}_\mu\ ,\ \mu\in\mathscr{M}_1(E)\ \big)$$
une chaîne de Markov. Soit $F \subset E$. On pose $T_F = \inf\{n \geq 0 \ :\ X_n \in F\}$ avec la convention habituelle $T_F = \infty$ ssi pour tout $n \in \mathbb{N}$, on a $X_n \notin F$. On pose également
$$N_F = \sum_{n\in\mathbb{N}} \mathbf{1}_{\{X_n\in F\}} \quad \text{et} \quad \forall i \in E\ , \quad f(i) = \mathbf{P}_i(T_F < \infty)\ .$$

1. Montrer que $f : E \to [0,1]$ est $Q$-sur-harmonique. D'après l'exercice précédent I.33, il existe donc un unique couple de fonctions $h, f' : E \to \mathbb{R}_+$ tels que $f'$ est $Q$-harmonique et $f = U.h + f'$.
2. Montrer que $f'(i) = \mathbf{P}_i(N_F = \infty)$, pour tout $i \in E$.
3. Montrer que
$$\forall i \in E\ , \quad h(i) = \mathbf{P}_i\big(\forall\, n \geq 1\ ,\ X_n \notin F\ \big)\ .$$
4. Montrer que
$$\forall i \in E\ ,\ \mathbf{P}_i\text{-p.s.} \quad \lim_n f(X_n) = \lim_n f'(X_n) = \mathbf{1}_{\{N_F=\infty\}}\ .$$

□

**Exercice I.35** *En application des deux exercices précédents I.33 et I.34, on redémontre des résultats du cours.* Soit
$$\big(\Omega\,;\,\mathscr{F}\,;\,(\mathscr{F}_n)_{n\geq 0}\,;\,(X_n)_{n\geq 0}\,;\,Q\,;\,\mathbf{P}_\mu\ ,\ \mu\in\mathscr{M}_1(E)\ \big)$$
une chaîne de Markov. Pour tout $i \in E$. On pose $T_i = \inf\{n \geq 0 \ :\ X_n = i\}$ avec la convention habituelle $T_i = \infty$ ssi pour tout $n \in \mathbb{N}$, on a $X_n \neq i$. On pose également
$$N_i = \sum_{n\in\mathbb{N}} \mathbf{1}_{\{X_n=i\}}\ .$$
et $\phi_i(j) = \mathbf{P}_i(T_j < \infty)$. On a démontré à l'exercice précédent I.34 que $\phi_i : E \to [0,1]$ est $Q$-surharmonique. On note $h_i : E \to \mathbb{R}_+$ et $\phi'_i : E \to \mathbb{R}_+$, les fonctions issues de la décomposition de Riesz de $\phi_i$ : on a $\phi_i = U.h_i + \phi'_i$ avec $\phi'_i$ qui est $Q$-harmonique.

1. Montrer que

   — soit $\phi'_i \equiv 0$ et $\mathbf{P}_i(N_i = \infty) = 0$ : $i$ est donc récurrent,

   — soit $U.h_i \equiv 0$ et $\mathbf{P}_i(N_i = \infty) = 1$ : $i$ est donc transient.

2. Montrer que si $i$ est récurrent alors pour tout $j \in E$, ou bien $\mathbf{P}_i(N_j = \infty) = 0$, ou bien $\mathbf{P}_i(N_j = \infty) = 1$.
3. Pour tout état $i \in E$ récurrent, on pose $R_i = \{j \in E \ :\ \mathbf{P}_i(N_j = \infty) > 0\ .\}$. Montrer que si $j \in R_i$ alors $j$ est récurrent et $C_i = C_j$. □

**Exercice I.36** (*Problème de Dirichlet sur les graphes.*) Soit $\mathbf{G} = (\mathbf{S}, \mathbf{A})$, un graphe simple, non-orienté, dénombrable, connexe et sans boucle. On le munit d'un système de poids $\mathbf{C} = (C_a\ ;\ a \in \mathbf{A})$. On rappelle que $0 < C_e a < \infty$, pour toute arête $a \in \mathbf{A}$ et que

$$\forall s \in \mathbf{S}\,, \quad \pi(s) = \sum_{s' \in \mathbf{S} : s' \sim s} C_{\{s,s'\}} \ \in ]0, \infty[\,.$$

On considère une marche aléatoire sur $\mathbf{G}$ associée au système de poids $\mathbf{C}$ :

$$\big(\Omega\ ;\ \mathscr{F}\ ;\ (\mathscr{F}_n)_{n\geq 0}\ ;\ (X_n)_{n\geq 0}\ ;\ Q = (p(s,s'))_{s,s'}\ ;\ \mathbf{P}_\mu\ ,\ \mu \in \mathscr{M}_1(\mathbf{S})\ \big)$$

On rappelle que

$$\forall s, s' \in \mathbf{S}\,, \quad p(s,s') = C_{\{s,s'\}}/\pi(s) \quad \text{si } s \sim s' \quad \text{et} \quad p(s,s') = 0 \quad \text{sinon.}$$

Puisqu'on a supposé $\mathbf{G}$ connexe et que $0 < C_a < \infty$, pour toute arête $a \in \mathbf{A}$, on voit facilement que $Q$ est **irréductible**. De plus, on rappelle que la matrice de transition est réversible et $\pi$ en est une mesure de réversibilité.

On fixe un sous-ensemble $D$ de sommets. On dit que $D$ est un *domaine* du graphe $\mathbf{G}$ ssi $D$ est connexe, c'est-à-dire que pour toute paire de sommets distincts $s$ et $s'$ de $D$, il existe un chemin $\gamma = (s_0 = s, s_1, \ldots, s_n = s')$ tels que $s_1, \ldots, s_n \in D$. On fixe un domaine $D \subset \mathbf{S}$ et pour éviter les trivialités, on suppose que $D$ est non-vide et distinct de $\mathbf{S}$. On fixe $g : \mathbf{S}\backslash D \to \mathbb{R}$, une fonction **bornée**. Le problème de Dirichlet associé à $D$ et $g$ consiste à trouver un fonction $f : \mathbf{S} \to \mathbb{R}$ qui est $Q$-harmonique sur $D$ et telle que $f(i) = g(i)$, pour tout $i \notin D$.

$$\text{Pb Dirichlet}(D,g) : \exists ? f : E \to \mathbb{R}\ :\ \begin{cases} Q.f(i) = f(i) & \text{pour tout } i \in D, \\ f(i) = g(i) & \text{pour tout } i \in \mathbf{S}\backslash D. \end{cases}$$

On introduit $T_{D^c} = \inf\{n \geq 0\ :\ X_n \in \mathbf{S}\backslash D\}$ avec la convention que $T_{D^c} = \infty$ ssi pour tout $n \geq 0$, $X_n \in D$.

1. Pour tout $s \in \mathbf{S}$, on pose $f(s) = \mathbf{E}_s[g(X_{T_{D^c}})\mathbf{1}_{\{T_{D^c}<\infty\}}]$. Montrer que $f : E \to \mathbb{R}$ est une fonction bien définie bornée. Montrer qu'elle est une solution au problème de Dirichlet$(D,g)$.
2. On suppose que $Q$ est récurrente. Montrer que pour tout $s \in \mathbf{S}$, on a $\mathbf{P}_s(T_{D^c} < \infty) = 1$. Montrer que le problème de Dirichlet$(D,g)$ a une unique solution qui soit bornée et elle est donnée par la formule de la question précédente. (*Indication : si $f'$ est une solution bornée, $(f'(X_{n\wedge T_{D^c}}))_{n\geq 0}$ est une martingale bornée sous $\mathbf{P}_s$.*)
3. Que dit le résultat précédent si le graphe est fini ou si $D$ est fini ?
4. On suppose que $D$ est infini. On suppose que $g$ est positive. Soit $f$ une solution problème de Dirichlet $(D,g)$ qui est positive également. Montrer que
$$\forall s \in \mathbf{S}\,, \quad \mathbf{E}_s[g(X_{T_{D^c}})\mathbf{1}_{\{T_{D^c}<\infty\}}] \leq f(s)\ .$$
5. On suppose que $D$ est infini. Soient $f_1$ et $f_2$ deux solutions du problème de Dirichlet$(D,g)$. On suppose que $\lim_\infty f_1 = \lim_\infty f_2 = 0$ dans $D$ : c'est-à-dire que
$$\forall \eta > 0\,, \quad \#\{s \in D\ :\ |f_1(s)| > \eta\} < \infty \quad \text{et} \quad \#\{s \in D\ :\ |f_2(s)| > \eta\} < \infty\ .$$
On suppose qu'il existe $s_0 \in D$ tel que $f_1(s_0) < f_2(s_0)$. Montrer que $f_2 - f_1$ est solution de du problème de Dirichlet$(D,0)$. On pose
$$C = \{s \in \mathbf{S}\ :\ f_2(s) - f_1(s) = \sup_{s' \in \mathbf{S}} f_2(s') - f_1(s')\ \}\ .$$
Montrer que $C \neq \emptyset$. Montrer que $C \subset D$. Montrer que si $s \in C$ et $s' \sim s$, alors $s' \in C$. En déduire que $C = D$ et que $f_2 - f_1$ est constante et strictement positive. Montrer que cela n'est pas possible. Montrer que le problème de Dirichlet $(D,g)$ a au plus une solution $f$ telle que $\lim_\infty f = 0$ dans $D$.
6. Montrer que dans $\mathbb{Z}^3$, si on prend $D : \mathbb{Z}^3\backslash\{0\}$ et $g(0) = 1$, le problème de Dirichlet admet au moins deux solutions bornées distinctes. Est-ce le cas dans $\mathbb{Z}^2$ ? □

**Exercice I.37** (*Marches nostalgiques sur les arbres symétriques*) Soit $\mathbf{T} = (\mathbf{S}, \mathbf{A})$ un graphe simple, non-orienté, sans boucle. On rappelle qu'une suite finie $\gamma = (s_0, s_1, \ldots, s_{n+1})$ est un chemin dans le graphe ssi $\{s_k, s_{k+1}\} \in \mathbf{A}$, pour tout $0 \leq k \leq n$. On dit que le chemin $\gamma$ *relie* le sommet $s_0$ au sommet $s_n$. On rappelle que le graphe $\mathbf{T}$ est dit *connexe* ssi toute paire de sommets distincts sont reliés par un chemin. La longueur du chemin $\gamma = (s_0, s_1, \ldots, s_n)$ est $n$ et elle notée $|\gamma|$. On définit la *distance du graphe* $\mathbf{T}$ qui est une une fonction $d : \mathbf{S} \times \mathbf{S} \to \mathbb{R}_+$ donnée par

$$\forall s, s' \in \mathbf{S}\,, \quad d(s,s') = \min\{|\gamma|\ ;\ \gamma, \text{chemin joignant } s \text{ à } s'.\} \quad \text{si } s \neq s' \quad \text{et} \quad d(s,s) = 0\ .$$

Un chemin $\gamma = (s_0, s_1, \ldots, s_n)$ est un *cycle* ssi (1) :$s_0 = s_n$, (2) : $n \geq 3$ *et* (3) : tous les sommets $s_0, s_1, \ldots, s_n$ sont distincts. Un *arbre est un graphe connexe et sans cycle*.

On suppose que $\mathbf{T} = (\mathbf{S}, \mathbf{A})$ est un **arbre infini**. On distingue un sommet $r \in \mathbf{S}$ que l'on appellera la *racine*. On suppose que l'arbre $\mathbb{T}$ est *à symétrie sphérique*, c'est-à-dire qu'il existe une suite $(a_n)_{n\geq 1}$ d'entiers *strictement positifs* tel que

$$\forall n \geq 1 \; , \; \forall s \in \mathbf{S} \;\text{ tel que } d(r,s) = n, \text{ on a } \deg(s) = a_n + 1 \; .$$

On fixe $\lambda \in ]0, \infty[$ et on considère la marche aléatoires

$$\big(\Omega \; ; \; \mathscr{F} \; ; \; (\mathscr{F}_n)_{n\geq 0} \; ; \; (X_n)_{n\geq 0} \; ; \; Q_\lambda = (p_\lambda(s,s'))_{s,s'} \; ; \; \mathbf{P}_\mu \; , \; \mu \in \mathscr{M}_1(\mathbf{S}) \; \big)$$

associée au système de poids $\mathbf{C}^\lambda = (C_a^\lambda \; ; \; a \in \mathbf{A})$ donné par $C_{\{r,s'\}} = 1/a_0$, pour tout $s' \sim r$ et

$$\forall s, s' \in \mathbf{S} \text{ tels que } s \sim s' \text{ et } s \neq r, \quad C_{\{s,s'\}} = \begin{cases} \lambda^{d(r,s)} & \text{si } d(r,s') = d(r,s) + 1, \\ \lambda^{d(r,s)-1} & \text{si } d(r,s') = d(r,s) - 1. \end{cases}$$

On voit que si $\lambda < 1$ la marche va avoir tendance à de rapprocher de la racine. Comme c'est surtout ce cas qui est intéressant, on donne à ce modèle le nom de marche nostalgique.

1. Montrer que si $s \sim r$, $p(r,s) = 1/a_0$ et
$$\forall s, s' \in \mathbf{S} \text{ tels que } s \sim s' \text{ et } s \neq r, \quad p(s,s') = \begin{cases} \frac{\lambda}{\lambda a_n + 1} & \text{si } d(r,s') = d(r,s) + 1, \\ \frac{1}{\lambda a_n + 1} & \text{si } d(r,s') = d(r,s) - 1. \end{cases}$$
2. On pose $Y_n = d(r, X_n)$, $n \geq 0$. Montrer que sous $\mathbf{P}_r$, $(Y_n)_{n\geq 0}$ est distribué comme un processus de naissance et de mort issu de 0. Calculer la matrice de transition de $(Y_n)_{n\geq 0}$.
3. Montrer que la marche nostalgique est récurrenet **ssi** $(Y_n)_{n\geq 0}$ est récurrent.
4. Pour tout $n \geq 0$, on pose $\mathcal{L}_n = \#\{s \in \mathbf{S} \; : \; d(r,s) = n\}$. Montrer qu'on a le résultat suivant.

— Si $\lambda \leq \liminf_{n\to\infty} (\mathcal{L}_n)^{1/n}$, alors la marche nostalgique est récurrente.

— Si $\lambda > \liminf_{n\to\infty} (\mathcal{L}_n)^{1/n}$, alors la marche nostalgique est transiente. □

— ——— —

*les exercice suivant donnent des critères analytiques de transience et de récurrence pour es marches aléatoires sur $\mathbb{Z}^d$.*

— ——— —

**Notations.** Dans tous les exercices qui suivent, nous adoptons les mêmes notations : $(\Omega, \mathscr{F}, \mathbf{P})$ est un espace de probabilité sur lequel est définie une suite de variables $(\xi_n)_{n\geq 1}$ qui sont à valeurs dans $\mathbb{Z}^d$, indépendantes et de loi commune notée $\pi$ :

$$\forall n \geq 1 \; , \; \forall \mathbf{k} \in \mathbb{Z}^d \; , \quad \mathbf{P}(\xi_n = \mathbf{k}) = \pi(\mathbf{k}) \; .$$

On pose $X_0 = 0$ et pour tout $n \geq 1$, $X_n = \xi_1 + \ldots + \xi_n$. La suite $(X_n)_{n\geq 0}$ est donc une marche aléatoire issue de 0 et de loi de saut $\pi$. On s'intéresse à la récurrence de 0. On note $N_0$ le nombre de visite en 0 de la marche

$$N_0 = \sum_{n\geq 1} \mathbf{1}_{\{X_n = 0\}} \; \in \mathbb{N} \cup \{\infty\} \; .$$

Les critères font intervenir la fonction caractéristique des variables $\xi_n$ : on note $\langle \cdot \, , \cdot \rangle$ le produit scalaire Euclidien. On note $\|\cdot\|$ la norme Euclidienne et on pose

$$\forall \mathbf{u} = (u_1, \ldots, u_d) \in \mathbb{R}^d \; , \quad \psi(\mathbf{u}) = \mathbf{E}\big[e^{i\langle \mathbf{u}, \xi_1\rangle}\big] = \sum_{\mathbf{k}\in\mathbb{Z}^d} e^{i\langle \mathbf{u}, \mathbf{k}\rangle} \pi(k) \; .$$

On note $\ell$ la mesure de Lebesgue sur $\mathbb{R}^d$ et pour tout nombre complexe $z \in \mathbb{C}$, on note $\mathrm{Re}(z)$ la partie réelle de $z$.

**Exercice I.38** On note $\xi_1 = (\xi_1^{(1)}, \xi_1^{(2)}, \ldots, \xi_1^{(d)})$, les coordonnées de $\xi_1$ dans la base canonique. Montrer que si le vecteur moyen est non-nul :

$$\mathbf{E}[\xi_1] = \big( \mathbf{E}[\xi_1^{(1)}], \mathbf{E}[\xi_1^{(2)}], \ldots, \mathbf{E}[\xi_1^{(d)}] \,\big) \; \neq \; 0 \; ,$$

alors $\mathbf{E}[N_0] < \infty$ et 0 est transient. □

**Exercice I.39** (*Critère de Chung*) Le but de l'exercice est de montrer le critère de récurrence suivant : *soit* $\eta\in]0,1[$ *; l'état* 0 *est récurrent pour la marche* $(X_n)_{n\geq 0}$ **ssi**

$$\sup_{r\in]0,1[}\int_{]-\eta,\eta[^d}\mathrm{Re}\Big(\frac{1}{1-r\psi(\mathbf{u})}\Big)\,\ell(d\mathbf{u})=\infty\;. \tag{I.65}$$

Ce critère s'appelle le critère de Chung.

1. Montrer que $1-\cos y\geq y^2/4$, pour tout $y\in[-\pi/3,\pi/3]$.
2. Montrer que pour tout $y\in\mathbb{R}$, on a
$$\frac{1-\cos(\eta y)}{(\eta y)^2}=\frac{1}{2}\int_{-\eta}^{\eta}e^{ity}\frac{\eta-|t|}{\eta^2}\,dt\quad\textbf{et}\quad\int_{-\infty}^{\infty}e^{ity}\frac{1-\cos(t/\eta)}{\pi t^2/\eta}\,dt=(1-|\eta y|)^+\leq\mathbf{1}_{\{|y|<1/\eta\}}.$$
3. Pour tout $\mathbf{u}=(u_1,\ldots,u_d)\in\mathbb{R}^d$, on pose
$$g(\mathbf{u})=\prod_{1\leq k\leq d}\frac{(\eta-|u_k|)^+}{\eta^2}\;.$$
En utilisant les questions précédentes, montrer que
$$\forall\mathbf{x}\in\mathbb{R}^d\,,\quad\mathbf{1}_{\{\|\mathbf{x}\|<\pi/(3\eta)\}}\leq 2^d\int_{]-\eta,\eta[^d}e^{i\langle\mathbf{u},\mathbf{x}\rangle}g(\mathbf{u})\,\ell(d\mathbf{u})\;.$$
4. On remarque que pour tout $\mathbf{x}=(x_1,\ldots,x_d)\in\mathbb{R}^d$, on a
$$\mathbf{1}_{\{\|\mathbf{x}\|<\sqrt{d}/\eta\}}\geq\prod_{1\leq k\leq d}\mathbf{1}_{\{|\eta x_k|<1\}}\geq\prod_{1\leq k\leq d}(1-|\eta x_k|)^+\;.$$
Pour tout $\mathbf{u}=(u_1,\ldots,u_d)\in\mathbb{R}^d$, on pose
$$f(\mathbf{u})=\prod_{1\leq k\leq d}\frac{1-\cos(u_k/\eta)}{\pi u_k^2/\eta}\;\geq 0\;\;.$$
Montrer que
$$\forall\mathbf{x}\in\mathbb{R}^d\,,\quad\mathbf{1}_{\{\|\mathbf{x}\|<\sqrt{d}/\eta\}}\geq\int_{\mathbb{R}^d}e^{i\langle\mathbf{u},\mathbf{x}\rangle}f(\mathbf{u})\,\ell(d\mathbf{u})\;.$$
5. Déduire des deux questions précédentes que pour tout $r\in]0,1[$ on a
$$\sum_{n\geq 0}r^n\mathbf{P}\big(\,\|X_n\|<\pi/(3\eta)\big)\leq 2^d\int_{]-\eta,\eta[^d}\mathrm{Re}\Big(\frac{1}{1-r\psi(\mathbf{u})}\Big)\,g(\mathbf{u})\,\ell(d\mathbf{u})$$
et
$$\sum_{n\geq 0}r^n\mathbf{P}\big(\,\|X_n\|<\sqrt{d}/\eta\big)\geq\int_{\mathbb{R}^d}\mathrm{Re}\Big(\frac{1}{1-r\psi(\mathbf{u})}\Big)\,f(\mathbf{u})\,\ell(d\mathbf{u}).$$
6. Montrer que $\mathrm{Re}\Big(\frac{1}{1-r\psi(\mathbf{u})}\Big)\geq 0$.
7. Montrer que $g(\mathbf{u})\leq 1/\eta^d$. En déduire que
$$\sum_{n\geq 0}\mathbf{P}\big(\,\|X_n\|<\pi/(3\eta)\big)\leq(2/\eta)^d\sup_{r\in]0,1[}\int_{]-\eta,\eta[^d}\mathrm{Re}\Big(\frac{1}{1-r\psi(\mathbf{u})}\Big)\,\ell(d\mathbf{u})\;.$$
8. Puisque $\pi/3>1$, montrer que si $\mathbf{u}\in]-1/\eta,1/\eta[^d$, on a $f(\mathbf{u})\geq(4\eta)^{-d}$. Montrer alors que
$$\sum_{n\geq 0}\mathbf{P}\big(\,\|X_n\|<\sqrt{d}/\eta\big)\geq(4\eta)^{-d}\sup_{r\in]0,1[}\int_{]-\frac{1}{\eta},\frac{1}{\eta}[^d}\mathrm{Re}\Big(\frac{1}{1-r\psi(\mathbf{u})}\Big)\,\ell(d\mathbf{u}).$$
9. Conclure. □

**Exercice I.40** (*Critère pour les marches symétriques*) On suppose que la loi de saut est symétrique, c'est-à-dire que

$$\xi_1\overset{\text{(loi)}}{=}-\xi_1\;.$$

Le but de l'exercise est de montrer le critère suivant : *soit* $\eta\in]0,1[$ *; l'état* 0 *est récurrent pour la marche* $(X_n)_{n\geq 0}$ **ssi**

$$\int_{]-\eta,\eta[^d}\frac{\ell(d\mathbf{u})}{1-\psi(\mathbf{u})}=\infty\;. \tag{I.66}$$

1. Montrer que $\psi$ est à valeurs réelles.
2. Montrer que
$$\forall \mathbf{u} \in \mathbb{R}^d\,, \quad \mathbf{E}\big[e^{\langle \mathbf{u}, \xi_1+\xi_2\rangle}\big] = \psi(\mathbf{u})^2\ .$$
3. Montrer que
$$\sup_{r\in ]0,1[} \int_{]-\eta,\eta[^d} \frac{\ell(d\mathbf{u})}{1-r\psi(\mathbf{u})^2} = \int_{]-\eta,\eta[^d} \frac{\ell(d\mathbf{u})}{1-\psi(\mathbf{u})^2}\ .$$
4. En utilisant le critère de Chung, montrer que l'état 0 est récurrent pour la marche $(X_{2n})_{n\geq 0}$ ssi
$$\int_{]-\eta,\eta[^d} \frac{\ell(d\mathbf{u})}{1-\psi(\mathbf{u})^2} = \infty\ .$$
5. Montrer que l'état 0 est récurrent pour la marche $(X_{2n})_{n\geq 0}$ ssi il est récurrent pour la marche $(X_n)_{n\geq 0}$.
6. Montrer que $1-\psi(\mathbf{u})^2 \sim_{\mathbf{u}\to 0} 2(1-\psi(\mathbf{u})$.
7. Conclure. □

**Exercice I.41** On se place en dimension $d=1$. On fixe $\alpha \in ]0,2[$ et on choisit une loi de saut $\pi$ sur $\mathbb{Z}$ donnée par
$$\pi(0)=0 \quad \text{et} \quad \forall k \in \mathbb{Z}\backslash\{0\}\,, \quad \pi(k) = \frac{C_\alpha}{|k|^{1+\alpha}}\ ,$$
où $1/C_\alpha = 2\sum_{k\geq 1} k^{-1-\alpha}$. La loi $\pi$ est donc symétrique. On remarque que la marche est **irréductible**.

1. On rappelle que $1-\cos y \geq y^2/4$, pour tout $y\in[-\pi/3,\pi/3]$. Montrer que pour tout $u\neq 0$, on a
$$1-\psi(u) \geq \mathbf{E}\big[(1-\cos(u\xi_1)\,)\mathbf{1}_{\{|\xi_1|\leq \pi/(3|u|)\}}\big] \geq \tfrac{1}{4}u^2 \sum_{1\leq k\leq \pi/(3|u|)} k^{1-\alpha} \sim_{u\to 0} C'|u|^\alpha\ ,$$
où $C'$ est une constante strictement positive que l'on s'amusera à préciser.
2. On suppose maintenant que $1\leq\alpha<2$. En observant que la fonction $y\mapsto y^{-2}(1-\cos y)$ est bornée sur $[-1,1]$, montrer qu'il existe une constante $C''>0$, pour tout $u\in\mathbb{R}$,
$$\mathbf{E}\big[(1-\cos(u\xi_1)\,)\mathbf{1}_{\{|\xi_1|\leq 1/|u|)\}}\big] \leq C''|u|^\alpha\ .$$
3. On suppose toujours que $1\leq\alpha<2$. Montrer que
$$\mathbf{E}\big[(1-\cos(u\xi_1)\,)\mathbf{1}_{\{|\xi_1|> 1/|u|)\}}\big] \leq 2C_\alpha \sum_{k>1/|u|} k^{-1-\alpha} \sim_{u\to 0} C^*|u|^\alpha$$
où $C^*$ est une constante strictement positive que l'on s'amusera à préciser.
4. Déduire des deux questions précédentes que si $1\leq\alpha<2$, alors il existe une constante $K>0$ telle que
$$\forall u\in[-1,1]\,, \quad 1-\psi(u)\leq K\,|u|^\alpha\ .$$
5. En utilisant le critère précédent montrer que si $1\leq\alpha<2$, alors la marche est récurrente.
6. En utilisant le critère précédent montrer que si $0<\alpha<1$, alors la marche est transiente. Montrer également que dans ce cas
$$\mathbf{P}\text{-p.s.} \quad \limsup_{n\to\infty} X_n = +\infty \quad \text{et} \quad \liminf_{n\to\infty} X_n = -\infty\ .$$
Est-ce surprenant ? Décrivez qualitativement le comportement de la marche. □

**Exercice I.42** Le but de cet exercice est de montrer que les marches aléatoires dans $\mathbb{Z}^d$ qui sont "vraiment plus que tri-dimensionnelles" sont transientes. Plus précisément : *si l'espace vectoriel engendré par les* $\mathbf{k}\in\mathbb{Z}^d$ *tels que* $\pi(\mathbf{k})>0$ *est de dimension supérieure ou égale à* 3*, alors l'état* 0 *est transient.*

1. On montre d'abord le résultat en dimension $d=3$. L'hypothèse signifie donc que l'espace vectoriel engendré par les $\mathbf{k}\in\mathbb{Z}^d$ tels que $\pi(\mathbf{k})>0$ est $\mathbb{R}^3$. Montrer que
$$\forall r\in[0,1[\,,\ \forall\mathbf{u}\in\mathbb{R}^3\backslash\{0\}\,, \quad \frac{1}{1-\mathrm{Re}(\psi(\mathbf{u}))} \geq \mathrm{Re}\Big(\frac{1}{1-\psi(\mathbf{u})}\Big) \geq 0\ .$$

2. Pour tout $\mathbf{u} \in \mathbb{R}^3 \backslash \{0\}$, on pose $\phi(\mathbf{u}) = \frac{1}{\|\mathbf{u}\|}.\mathbf{u}$. On rappelle que $1 - \cos y \geq y^2/4$, pour tout $y \in [-\pi/3, \pi/3]$. Montrer que pour tout $\mathbf{u} \in \mathbb{R}^3 \backslash \{0\}$, on a
$$\begin{aligned} 1 - \mathrm{Re}(\psi(\mathbf{u})) &\geq \mathbf{E}\big[(1 - \cos\langle \mathbf{u}, \xi_1 \rangle)\mathbf{1}_{\{|\langle \mathbf{u}, \xi_1\rangle| \leq \pi/3\}}\big] \\ &\geq \tfrac{1}{4}\|\mathbf{u}\|^2 \mathbf{E}\big[\langle \phi(\mathbf{u}), \xi_1\rangle^2 \mathbf{1}_{\{|\langle \phi(\mathbf{u}), \xi_1\rangle| \leq \pi/(3\|\mathbf{u}\|)\}}\big] \; . \end{aligned}$$

3. Supposons qu'il existe un suite $\mathbf{u}_n \in [-1,1]^3 \backslash \{0\}$, $n \geq 0$, telle que $\lim_n \mathbf{u}_n = 0$ et telle que
$$\lim_n \mathbf{E}\big[\langle \phi(\mathbf{u}_n), \xi_1\rangle^2 \mathbf{1}_{\{|\langle \phi(\mathbf{u}_n), \xi_1\rangle| \leq \pi/(3\|\mathbf{u}_n\|)\}}\big] = 0$$
Montrer qu'on peut trouver une suite strictement croissante d'indices $(n_k)_{k \geq 0}$ telle que
$$\lim_k \phi(\mathbf{u}_{n_k}) = \mathbf{v} \; .$$
On a donc $\|\mathbf{v}\| = 1$. Fatou et l'hypothèse précédente impliquent que
$$\mathbf{E}[\big[\langle \mathbf{v}, \xi_1\rangle^2\big] = 0 \; .$$
Montrer que cela contredit le fait que l'espace vectoriel engendré par les $\mathbf{k} \in \mathbb{Z}^3$ tels que $\pi(\mathbf{k}) > 0$ est $\mathbb{R}^3$. En déduire qu'il existe une constante $C > 0$ et $r_0 > 0$ tel que
$$\forall \mathbf{u} \in \mathbb{R}^3 \backslash \{0\} \text{ tel que } \|\mathbf{u}\| \leq r_0, \quad \mathbf{E}\big[\langle \phi(\mathbf{u}), \xi_1\rangle^2 \mathbf{1}_{\{|\langle \phi(\mathbf{u}), \xi_1\rangle| \leq \pi/(3\|\mathbf{u}\|)\}}\big] \geq C > 0 \; .$$

4. Deduire des questions précédentes que
$$\forall \mathbf{u} \in \mathbb{R}^3 \text{ tel que } \|\mathbf{u}\| \leq r_0, \quad 1 - \mathrm{Re}(\psi(\mathbf{u})) \geq \tfrac{C}{4}\|\mathbf{u}\|^2 \; .$$

5. Par un changement de coordonnées radiales, montrer que
$$\int_{\mathbb{R}^d} \mathbf{1}_{[0, r_0]}(\|\mathbf{u}\|) \frac{\ell(d\mathbf{u})}{\|\mathbf{u}\|^2} < \infty \; .$$

6. En utilisant le critère de Chung et les exercices précédents, montrer que 0 est un état transient pour la marche $(X_n)_{n \geq 0}$.

7. On suppose que $d \geq 4$. Utiliser le résultat en dimension 3 et raisonner comme pour la preuve du théorème de Polyà dans l'exemple I.3.10 page 64 (attention : l'idée est la même mais ne s'adapte pas sans un peu de travail) □

# I.4 Asymptotique des chaînes : résultats quantitatifs.

Dans la section précédente nous avons étudié d'un point de vue qualitatif le comportement asymptotique des chaînes de Markov, c'est-à-dire que nous avons expliqué quels étaient les états qui étaient visités une infinité de fois (états récurrents) et ceux qui ne l'étaient pas (états transients). Une étude plus précise de la transience des chaînes devrait en théorie permettre de dire de façon satisfaisante comment la chaîne s'échappe vers "l'infini" mais cela dépasse les buts modestes de ce cours et cette question ne sera donc pas abordée ici. Nous nous concentrerons plutôt sur une étude plus fine de la récurrence. Le théorème de classification montre qu'une chaîne restreinte à chaque classe de récurrence se comporte comme une chaîne irréductible récurrente. En théorie, l'étude de la récurrence des chaînes se ramène donc à étudier en général les chaînes *irréductibles récurrentes* et c'est l'objet de cette section dont le but est de mesurer quels sont les états qui sont visités plus fréquemment que d'autres, étant entendu que tous le sont une infinité de fois. Cette étude est facilitée par le fait que dans le cas récurrent, les excursions de la chaîne sont indépendantes et de même loi : c'est essentiellement cette propriété de renouvellement qui est à l'oeuvre derrière la plupart des résultats énoncés dans cette section.

Dans la suite $\big(\Omega; \mathscr{F}; (\mathscr{F}_n)_{n \geq 0}; \mathbf{X} = (X_n)_{n \geq 0}; Q = (p(i,j))_{i,j \in E}; \mathbf{P}_\mu, \mu \in \mathscr{M}_1(E)\big)$ désigne comme d'habitude la chaîne de Markov que l'on considère. Elle sera en général irréductible récurrente mais cette hypothèse sera spécifiée explicitement à chaque fois que cela sera nécessaire.

### I.4.a Existence de mesures invariantes.

Nous avons introduit dans la section I.2.d la notion de mesure invariante, qui joue un rôle essentiel dans l'étude précise de la récurrence que nous nous proposons de mener. Nous avons déjà mentionné que certaines chaînes irréductibles ne possèdent pas de mesure invariante. Le but de cette section est de montrer que toute chaîne irréductible récurrente possède des mesures invariantes. Nous allons même les construire explicitement et montrer qu'elles sont uniques à une constante multiplicative près. Pour cela, on rappelle les notations introduites à la section I.3.a (page 57) sur les temps de retour successifs d'une suite à valeurs dans $E$ : soit $\mathbf{x} = (x_n)_{n\geq 0} \in E^{\mathbb{N}}$, soit $i \in E$, un état fixé ; les temps de retour successifs de la suite $\mathbf{x}$ en $i$ sont définis par $T_i^{(0)}(\mathbf{x}) = 0$ et $T_i^{(p+1)}(\mathbf{x}) = \inf\{n > T_i^{(p)}(\mathbf{x}) \ : \ x_n = i\}$ avec la convention habituelle : $\inf \emptyset = \infty$. On introduit également la notation suivante

$$\forall i,j \in E\,, \quad N_i^{(j)}(\mathbf{x}) = \sum_{0\leq n<T_j^{(1)}(\mathbf{x})} \mathbf{1}_{\{x_n=i\}} \quad \in \mathbb{N}\cup\{\infty\}\,.$$

$N_i^{(j)}(\mathbf{x})$ est donc le nombre de visites en $i$ de la suite $\mathbf{x}$ *strictement avant son premier retour en* $j$. On rappelle également que $\mathscr{E}^{(p)}(\mathbf{x})$, $p \geq 1$, désigne les *excusions successives de* $\mathbf{x}$ *en dehors de l'état* $\{i\}$. Pour simplifier les notations, on pose $T_i^{(p)} = T_i^{(p)}(\mathbf{X})$, $\mathscr{E}^{(p)} = \mathscr{E}^{(p)}(\mathbf{X})$ et

$$\forall i,j \in E\,, \quad N_i^{(j)} := N_i^{(j)}(\mathbf{X}) = \sum_{0\leq n<T_j^{(1)}} \mathbf{1}_{\{X_n=i\}}\,.$$

On rappelle également que si $i$ est récurrent, le théorème I.3.2 implique que $\mathbf{P}_i(\forall p\geq 1\,,\, T_i^{(p)} < \infty) = 1$ et que sous $\mathbf{P}_i$, les excursions hors de $\{i\}$ sont indépendantes et de même loi. On pose également,

$$\forall i,j \in E\,, \quad \rho_{i,j} := \mathbf{P}_i\big(T_j^{(1)} < T_i^{(1)}\big)\,. \tag{I.67}$$

**Proposition I.4.1** *Soient $i,j \in E$, deux états distincts. On suppose que $i \leftrightarrow j$ et que $i$ (et donc $j$) est récurrent. Alors, les assertions suivantes sont vérifiées.*

*(i)* $\rho_{i,j}\rho_{j,i} > 0$.

*(ii)* $\mathbf{P}_i(N_i^{(j)} = p) = \rho_{i,j}(1-\rho_{i,j})^{p-1}$*, pour tout* $p \geq 1$.

*(iii)* $0 < \mathbf{E}_j[N_i^{(j)}] < \infty$*. Plus précisément, on a* $\mathbf{E}_j[N_i^{(j)}] = \rho_{j,i}/\rho_{i,j}$.

**Preuve :** comme $i \to j$, il existe $i_0 = i, i_1, \ldots, i_{n-1}, i_n = j \in E$ tels que $p(i,i_1)\ldots p(i_{n-1},j) > 0$. On pose $r_* = \max\{0 \leq r \leq n : i_r = i\}$ et $r^* = \inf\{r_* \leq r \leq n : i_r = j\,\}$. On remarque premièrement que $r_* < r^*$, deuxièmement que $i_{r_*} = i$, $i_{r^*} = j$, troisièmement que s'il existe $r_* < r < r^*$, alors $i_r \in E\backslash\{i,j\}$ et que $p(i, i_{r_*+1})\ldots p(i_{r^*-1},j) > 0$. Tout cela permet d'affirmer que

$$\exists i_1^*, \ldots, i_m^* \in E\backslash\{i,j\}\ : \quad p(i,i_1^*)p(i_1^*,i_2^*)\ldots p(i_m^*,j) > 0\,.$$

On observe ensuite que

$$\mathbf{P}_i\text{-p.s.} \quad \mathbf{1}_{\{T_j^{(1)}<T_i^{(1)}\}} \geq \mathbf{1}_{\{X_0=i;X_1=i_1^*;\ldots;X_m=i_m^*;X_{m+1}=j\}}\,.$$

En intégrant cette inégalité sous $\mathbf{P}_i$, on obtient donc

$$\rho_{i,j} \quad = \quad \mathbf{P}_i(T_j^{(1)} < T_i^{(1)})$$

$$\begin{aligned} &\geq \mathbf{P}_i(X_0 = i; X_1 = i_1^*; \ldots ; X_m = i_m^*; X_{m+1} = j) \\ &= p(i, i_1^*)p(i_1^*, i_2^*) \ldots p(i_m^*, j) > 0 \ . \end{aligned}$$

Comme $i$ et $j$ jouent un rôle symétrique, on prouve de même que $\rho_{j,i} > 0$, ce qui montre $(i)$.

Pour tout $p \geq 1$, on pose $\xi_p = \mathbf{1}_{B_p}$ où $B_p = \{\forall n \geq 1 : \mathscr{E}_n^{(p)} \neq j\}$. Comme $i$ est récurrent, sous $\mathbf{P}_i$, les variables $(\xi_p)_{p\geq 1}$ sont des variables de Bernoulli i.i.d. Or on observe que $B_1 = \{T_i^{(1)} < T_j^{(1)}\}$, donc $\mathbf{P}_i(\xi_p = 1) = \mathbf{P}_i(B_p) = 1 - \rho_{i,j}$. On voit ensuite que pour tout $p \geq 1$, on a $\mathbf{P}_i(N_i^{(j)} = p) = \mathbf{P}_i(\xi_1 = 1; \ldots ; \xi_{p-1} = 1; \xi_p = 0) = (1 - \rho_{i,j})^{p-1}\rho_{i,j}$, ce qui montre $(ii)$. On a donc $\mathbf{E}_i[N_i^{(j)}] = 1/\rho_{i,j}$, qui est un nombre strictement positif et fini.

Pour montrer le point $(iii)$, on remarque d'abord que

$$\mathbf{P}_j\text{-p.s.} \quad N_i^{(j)} = \mathbf{1}_{\{T_i^{(1)} < T_j^{(1)}\}} N_i^{(j)}\big(\theta_{T_i^{(1)}} \mathbf{X}\big)$$

et ensuite que pour tout $n \in \mathbb{N}$, $\{T_i^{(1)} < T_j^{(1)}\} \cap \{T_i^{(1)} = n\} = \{n < T_j^{(1)}\} \cap \{T_i^{(1)} = n\}$. Or $\{n < T_j^{(1)}\} \in \mathscr{F}_n$ car $T_j^{(1)}$ est un $(\mathscr{F}_n)_{n\geq 0}$-temps d'arrêt. Donc $\{T_i^{(1)} < T_j^{(1)}\} \in \mathscr{F}_{T_i^{(1)}}$. Comme $X_{T_i^{(1)}} = i$, la propriété de Markov entraîne alors que

$$\mathbf{P}_j\text{-p.s.} \quad \mathbf{E}_j\big[N_i^{(j)} | \mathscr{F}_{T_i^{(1)}}\big] = \mathbf{1}_{\{T_i^{(1)} < T_j^{(1)}\}} \mathbf{E}_i\big[N_i^{(j)}\big] \ ,$$

et en intégrant sous $\mathbf{P}_j$ cette égalité on obtient que $\mathbf{E}_j\big[N_i^{(j)}\big] = \mathbf{P}_j\big(T_i^{(1)} < T_j^{(1)}\big)\mathbf{E}_i\big[N_i^{(j)}\big] = \rho_{j,i}/\rho_{i,j}$, ce qui est le résultat désiré. ■

**Théorème I.4.2** **(Existence des mesures invariantes)** *On suppose la chaîne irréductible et récurrente. On fixe $j \in E$ et on définit une mesure positive $\nu_j$ en posant $\nu_j(i) = \rho_{j,i}/\rho_{i,j}$, pour tout $i \in E$. Alors, on a*

$$\forall i \in E \, , \quad \nu_j(i) = \mathbf{E}_j[N_i^{(j)}] = \mathbf{E}_j\Big[ \sum_{0 \leq n < T_j^{(1)}} \mathbf{1}_{\{X_n = i\}} \Big] = \sum_{n\geq 0} \mathbf{P}_j\big( X_n = i \, ; \, T_j^{(1)} > n \big) \ . \tag{I.68}$$

*De plus, $\nu_j(j) = 1$ et $0 < \nu_j(i) < \infty$, pour tout $i \in E$. Enfin, $\nu_j$ est une mesure Q-invariante.*

**Preuve :** les deux premières égalités (I.68) sont une conséquence de la proposition I.4.1 et de la définition de $N_i^{(j)}$. On remarque ensuite que

$$\begin{aligned} \nu_j(i) &= \mathbf{E}_j\Big[ \sum_{0 \leq n < T_j^{(1)}} \mathbf{1}_{\{X_n = i\}} \Big] \\ &= \mathbf{E}_j\Big[ \sum_{n\geq 0} \mathbf{1}_{\{X_n = i \, ; \, T_j^{(1)} > n\}} \Big] \\ &= \sum_{n\geq 0} \mathbf{P}_j\big( X_n = i \, ; \, T_j^{(1)} > n \big) \ , \end{aligned}$$

par interversion série / espérance positive. Le fait que $0 < \nu_j(i) < \infty$, est exactement le point $(iii)$ de la proposition I.4.1. Comme $\mathbf{P}_j(N_j^{(j)} = 1) = 1$, on a bien $\nu_j(j) = 1$. Il reste simplement à prouver que pour tout $i^* \in E$, on a

$$\sum_{i\in E} \nu_j(i)p(i, i^*) = \nu_j(i^*) \ , \tag{I.69}$$

**(Cas 1)** : $j \neq i^*$. Comme $\{T_j^{(1)} > n\} \in \mathscr{F}_n$, Markov au temps $n$ implique $\mathbf{P}_j$-p.s. que

$$\mathbf{E}_j\big[\mathbf{1}_{\{X_n=i;X_{n+1}=i^*;T_j^{(1)}>n\}}|\mathscr{F}_n\big] = \mathbf{1}_{\{X_n=i;T_j^{(1)}>n\}}\mathbf{P}_i(X_1 = i^*) = \mathbf{1}_{\{X_n=i;T_j^{(1)}>n\}}p(i, i^*).$$

En intégrant cette égalité sous $\mathbf{P}_j$, on obtient

$$\mathbf{P}_j(X_n = i; X_{n+1} = i^*; T_j^{(1)} > n) = \mathbf{P}_j(X_n = i; T_j^{(1)} > n)p(i, i^*) .$$

Puisque $j \neq i^*$, si $X_{n+1} = i^*$ et $T_j^{(1)} > n$, alors $T_j^{(1)} > n + 1$, par conséquent

$$\mathbf{P}_j(X_n = i; X_{n+1} = i^*; T_j^{(1)} > n) = \mathbf{P}_j(X_n = i; X_{n+1} = i^*; T_j^{(1)} > n + 1)$$

On obtient alors

$$\begin{aligned}
\sum_{i\in E} \nu_j(i)p(i, i^*) &= \sum_{n\geq 0}\sum_{i\in E} \mathbf{P}_j(X_n = i; T_j^{(1)} > n)p(i, i^*) \\
&= \sum_{n\geq 0}\sum_{i\in E} \mathbf{P}_j(X_n = i; X_{n+1} = i^*; T_j^{(1)} > n + 1) \\
&= \sum_{n\geq 0} \mathbf{P}_j(X_{n+1} = i^*; T_j^{(1)} > n + 1) = \mathbf{E}_j\Big[\sum_{1\leq n\leq T_j^{(1)}} \mathbf{1}_{\{X_n=i^*\}}\Big] \\
&= \mathbf{E}_j\Big[\sum_{0\leq n<T_j^{(1)}} \mathbf{1}_{\{X_n=i^*\}}\Big] = \nu_j(i^*) ,
\end{aligned}$$

car $\mathbf{P}_j$-p.s. on a $X_0 = X_{T_j^{(1)}} = j$.

**(Cas 2)** : $j = i^*$. Comme $\nu_j(j) = 1$, il suffit de montrer que $\sum_{i\in E} \nu_j(i)p(i, i^*) = 1$. En raisonnant comme précédemment, la propriété de Markov montre que

$$\mathbf{P}_j(X_n = i; T_j^{(1)} > n)p(i, j) = \mathbf{P}_j(X_n = i; X_{n+1} = j; T_j^{(1)} > n) = \mathbf{P}_j(X_n = i; T_j^{(1)} = n + 1).$$

On en déduit

$$\begin{aligned}
\sum_{i\in E} \nu_j(i)p(i, j) =& \sum_{n\geq 0}\sum_{i\in E} \mathbf{P}_j(X_n = i; T_j^{(1)} > n)p(i, j) = \sum_{n\geq 0}\sum_{i\in E} \mathbf{P}_j(X_n = i; T_j^{(1)} = n + 1) \\
=& \sum_{n\geq 0} \mathbf{P}_j(T_j^{(1)} = n + 1) = \mathbf{P}_j(T_j^{(1)} < \infty) = 1,
\end{aligned}$$

ce qui termine la preuve du théorème. ■

Le théorème suivant montre l'unicité, à une constante multiplicative près, des mesures invariantes pour les chaînes irréductibles récurrentes (il montre même un peu plus).

**Théorème I.4.3** *On suppose la chaîne irréductible et récurrente. Soit $\nu$, une mesure Q-invariante et soit $\mu$ une mesure Q-excessive, c-à-d que $\mu \geq \mu.Q$. Alors les assertions suivantes sont vraies.*

*(i) La mesure $\mu$ est Q-invariante et il existe $0 < c < \infty$, telle que $\mu(i) = c\nu(i)$, pour tout $i \in E$.*

*(ii) Si $\nu^*$ et $\nu$ sont deux mesures Q-invariantes, alors il existe $0 < c^* < \infty$ telle que $\nu^*(i) = c^*\nu(i)$, pour tout $i \in E$.*

*(iii)* $0 < \nu(i) < \infty$, *pour tout* $i \in E$.

*(iv)* $\nu(i) = \nu(j)\nu_j(i)$, *pour tous* $i, j \in E$ *(ici,* $\nu_j$ *désigne la mesure définie au théorème d'existence I.4.2).*

*(v)* $Q$ *admet une probabilité invariante ssi elle admet une mesure invariante de masse finie. Dans ce cas, il n'existe qu'une seule probabilité invariante.*

**Preuve :** on note $Q = (p(i,j))_{i,j\in E}$, on fixe $j_0 \in E$ et on pose $\kappa = \nu_{j_0}$, la mesure $Q$-invariante définie par le théorème I.4.2. On a donc $\kappa.Q = \kappa$ et aussi : $0 < \kappa(i) < \infty$, pour tout $i \in E$, ce qui permet de poser $Q_* = (p_*(i,j))_{i,j\in E}$, où

$$\forall i, j \in E\,, \quad p_*(i,j) = \kappa(j)p(j,i)/\kappa(i)\ .$$

Puisque $\kappa.Q = \kappa$, on en déduit par un calcul simple que $Q_*$ est une matrice de transition. On montre facilement par récurrence que

$$\forall n \geq 1\,,\ \forall i, j \in E\,, \quad [Q_*^n](i,j) = \tfrac{\kappa(j)}{\kappa(i)}\,[Q^n](j,i)\ . \tag{I.70}$$

*En effet*, c'est vrai au rang $n = 1$, par définition de $Q_*$ ; on le suppose vrai au rang $n$ et on observe que

$$\begin{aligned} [Q_*^{n+1}](i,j) &= \sum_{k\in E} p_*(i,k)\cdot[Q_*^n](k,j) = \sum_{k\in E} \tfrac{\kappa(k)p(k,i)}{\kappa(i)}\tfrac{\kappa(j)\,[Q^n](j,k)}{\kappa(k)} \\ &= \tfrac{\kappa(j)}{\kappa(i)}\sum_{k\in E}[Q^n](j,k)\cdot p(k,i) = \tfrac{\kappa(j)}{\kappa(i)}[Q^{n+1}](j,i)\ , \end{aligned}$$

ce qui montre (I.70) par récurrence.

Comme $Q$ est irréductible, pour tous $i, j \in E$, il existe $n \geq 1$, tel que $[Q^n](j,i) > 0$. Comme $\kappa(j) > 0$, (I.70) implique que $[Q_*^n](i,j) > 0$, ce qui montre que $Q_*$ est irréductible. De plus (I.70) implique aussi que $[Q_*^n](i,i) = [Q^n](i,i)$. Donc $\sum_{n\geq 1}[Q_*^n](i,i) = \sum_{n\geq 1}[Q^n](i,i) = \infty$, par le corollaire I.3.3 puisque $Q$ est récurrente. Cela montre que $Q_*$ est récurrente par le même corollaire I.3.3 appliqué à $Q_*$.

On pose ensuite $f(i) = \mu(i)/\kappa(i)$, pour tout $i \in E$. Comme $\mu$ est non-identiquement nulle, $f$ est également non-identiquement nulle. On observe ensuite que pour tout $i \in E$, on a

$$\begin{aligned} (Q_*.f)(i) &= \sum_{j\in E} p_*(i,j)f(j) = \sum_{j\in E}\tfrac{\kappa(j)p(j,i)}{\kappa(i)}\tfrac{\mu(j)}{\kappa(j)} \\ &= \tfrac{1}{\kappa(i)}\sum_{j\in E}\mu(j)p(j,i) \ = \tfrac{1}{\kappa(i)}(\mu.Q)(i) \ \leq \mu(i)/\kappa(i) = f(i)\ . \end{aligned}$$

Par conséquent $f$ est une fonction $Q_*$-sur-harmonique positive. Comme $Q_*$ est irréductible récurrente, la proposition I.3.12 implique que $f$ est constante. Comme $f$ n'est pas identiquement nulle, il existe une constante $0 < c < \infty$ telle que $f \equiv c$, et on a bien montré que

$$\forall i \in E\ , \quad \mu(i) = c\,\kappa(i) = c\nu_{j_0}(i)\ .$$

Soient $\nu$ et $\nu^*$ sont deux mesures $Q$-invariantes : ce sont des cas particuliers de mesures $Q$-excessives. Il existe donc deux réels strictement positifs $a$ et $b$ tels que $\nu = a\nu_{j_0}$ et $\nu^* = b\nu_{j_0}$. Si on pose $c^* = b/a$, on a donc $\nu^* = c^*\nu$, ce qui montre le point $(ii)$. Comme on a montré que $0 < \nu_{j_0}(i) < \infty$, pour tout $i \in E$, on en déduit que $0 < \nu(i) < \infty$, pour tout $i \in E$, ce qui montre $(iii)$.

On voit que si $\nu$ est $Q$-invariante, pour tout $j \in E$, il existe un réel strictement positif $c_j$ tel que $\nu = c_j\nu_j$. On rappelle que, par définition, on a $\nu_j(j) = 1$. Donc $\nu(j) = c_j$, ce qui implique $(iv)$.

Si $Q$ admet une probabilité invariante $\pi$, alors pour toute mesure $Q$-invariante, il existe un réel strictement positif $c$ tel que $\nu = c\pi$ et on a $\langle\nu\rangle = c < \infty$. Réciproquement, si $Q$ admet une mesure invariante $\nu$ de masse totale finie, on voit que $\pi = \frac{1}{\langle \nu \rangle} \nu$ est une mesure de probabilité invariante. Pour une chaîne irréductible récurrente, lorsqu'une probabilité invariante existe, elle est donc unique. ■

Nous allons considérer de plus près l'existence ou non d'une probabilité invariante. Pour cela, on montre la proposition suivante.

**Proposition I.4.4** *On suppose que la chaîne est irréductible (seulement). Si $Q$ admet une mesure de probabilité invariante, alors la chaîne est récurrente.*

**Preuve :** on note $\pi$ une probabilité $Q$-invariante. Il existe nécessairement $i \in E$ tel que $\pi(i) > 0$. On rappelle que pour toute suite $\mathbf{x} = (x_n)_{n \in \mathbb{N}}$ à valeurs dans $E$, on note $N_i(\mathbf{x}) = \sum_{n \geq 1} \mathbf{1}_{\{x_n = i\}}$ le nombre (peut-être infini) de visites de la suite à l'état $i$. Pour simplifier, on pose $N_i := N_i(\mathbf{X})$. On vérifie immédiatement que

$$N_i = \mathbf{1}_{\{T_i^{(1)} < \infty\}} + \mathbf{1}_{\{T_i^{(1)} < \infty\}} N_i\big(\theta_{T_i^{(1)}} \mathbf{X}\big) \; .$$

La propriété de Markov sous $\mathbf{P}_\pi$ et le fait que $X_{T_i^{(1)}} = i$ si $T_i^{(1)} < \infty$, impliquent que

$$\mathbf{P}_\pi\text{ - p.s.} \quad \mathbf{E}_\pi\big[N_i | \mathscr{F}_{T_i^{(1)}}\big] = \mathbf{1}_{\{T_i^{(1)} < \infty\}} + \mathbf{1}_{\{T_i^{(1)} < \infty\}} \mathbf{E}_i[N_i]$$

En intégrant sous $\mathbf{P}_\pi$, on obtient

$$\mathbf{E}_\pi\big[N_i\big] = \mathbf{P}_\pi\big(T_i^{(1)} < \infty\big) \cdot \big(1 + \mathbf{E}_i[N_i]\big) \leq 1 + \mathbf{E}_i[N_i], \tag{I.71}$$

ces inégalités ayant un sens dans $[0, \infty]$. Par interversion série / espérance,

$$\mathbf{E}_\pi\big[N_i\big] = \sum_{n \geq 1} \mathbf{P}_\pi(X_n = i) = \sum_{n \geq 1} \pi(i) = \infty \; ,$$

car $\pi$ est invariante et car on a choisi $i$ tel que $\pi(i) > 0$. Or (I.71) implique que $\mathbf{E}_i[N_i] = \infty$ et le corollaire I.3.3 implique que $i$ est récurrent. Comme la chaîne est irréductible tous les états sont récurrents. ■

**Remarque I.4.5** On prendra garde qu'il existe des chaînes irréductibles *transientes* qui admettent des mesures invariantes. La proposition précédente entraîne simplement que dans ce cas, elles sont de masse infinie. On se rappelera également qu'il existe des chaînes irréductibles qui n'admettent pas de mesure invariante. Le théorème d'existence I.4.2 entraîne que dans ce cas ces chaînes sont transientes. □

**Définition I.4.6** On introduit les notions suivantes.

**(Récurrence positive)** $i \in E$ est récurrent positif si $\mathbf{E}_i\big[T_i^{(1)}\big] < \infty$.

**(Récurrence nulle)** $i \in E$ est récurrent nul si $i$ est récurrent mais $\mathbf{E}_i\big[T_i^{(1)}\big] = \infty$.

Rappelons que l'on dit toujours que $i$ est transient si $\mathbf{P}_i\big(T_i^{(1)} < \infty\big) < 1$. □

On voit qu'être récurrent nul est une qualité intermédiaire entre la récurrence positive (où la chaîne revient vite) et la transience (où la chaîne peut, avec une probabilité non-nulle, ne pas revenir). Les termes "récurrent positif" et "récurrent nul" seront expliqués par les théorèmes ergodiques qui affirment qu'une chaîne passe asymptotiquement une fraction strictement positive de son temps en un état récurrent positif et passe une fraction asymptotiquement nulle de son temps en un état récurrent nul (bien qu'elle y passe infiniment souvent). Le théorème suivant donne une première utilisation de ces nouvelles notions.

**Théorème I.4.7** *On suppose que la chaîne est irréductible (seulement). Alors les assertions suivantes sont équivalentes.*

*(i) La chaîne admet une probabilité invariante.*

*(ii) Tous les états sont récurrents positifs.*

*(iii) Il existe un état récurrent positif.*

*Si l'une de ces trois conditions équivalentes est vérifiée, on dit que la chaîne est* **récurrente positive**. *De plus la probabilité invariante, notée $\pi$, est unique et on a*

$$\forall i \in E\,, \quad \pi(i) = \frac{1}{\mathbf{E}_i[T_i^{(1)}]}\ . \tag{I.72}$$

**Preuve :** on suppose $(i)$. La proposition I.4.4 implique que la chaîne est récurrente. On fixe $j \in E$ et on rappelle la définition de la mesure $Q$-invariante $\nu_j$ donnée au théorème I.4.2. En intervertissant série et espérance, on a

$$\langle \nu_j \rangle = \sum_{i\in E} \nu_j(i) = \mathbf{E}_j\Big[ \sum_{0\leq n<T_j^{(1)}} \sum_{i\in E} \mathbf{1}_{\{X_n=i\}} \Big] = \mathbf{E}_j\big[T_j^{(1)}\big]. \tag{I.73}$$

Or, si $Q$ admet une probabilité invariante le théorème I.4.3 $(iv)$ implique que toutes les mesures invariantes sont de masse finie et (I.73) entraîne que $j$ est récurrent positif. Comme on a raisonné avec $j$ quelconque, tous les états sont récurrent positifs. On a donc montré que $(i) \Rightarrow (ii)$. L'implication $(ii) \Rightarrow (iii)$ est triviale. Montrons $(iii) \Rightarrow (i)$ : on suppose que $j$ est récurrent positif. Il est donc récurrent. Comme la chaîne est irréductible, tous les états sont récurrents. Cela permet d'appliquer le théorème d'existence I.4.2. De plus (I.73) implique que la mesure $Q$-invariante $\nu_j$ donnée au théorème I.4.2 est de masse finie et le théorème I.4.3 $(iv)$ implique que $Q$ admet une probabilité invariante. On a donc montré l'équivalence des trois points.

On suppose ensuite que $Q$ admet une probabilité invariante notée $\pi$. Le théorème I.4.3 implique qu'il existe un réel strictement positif $c$ tel que $\nu_j = c\pi$ et (I.73) implique que $\mathbf{E}_j\big[T_j^{(1)}\big] = \langle \nu_j \rangle = c\pi(E) = c$. Or on rappelle que, par définition, on a $\nu_j(j) = 1$. Donc, $1 = \nu_j(j) = c\pi(j) = \mathbf{E}_j\big[T_j^{(1)}\big]\,\pi(j)$, ce qui entraîne donc la formule (I.72). ∎

Bien que moins utile, nous énonçons le corollaire suivant qui est une immédiate conséquence du théorème précédent.

**Corollaire I.4.8** *On suppose la chaîne irréductible et récurrente. Les assertions suivantes sont équivalentes.*

*(i) La chaîne admet une mesure invariante de masse totale infinie.*

*(ii) Toutes les mesures invariantes sont de masse totale infinie.*

*(iii) Tous les états sont récurrents nuls.*

*(iv) Il existe un état récurrent nul.*

**Exemple I.4.9** On sait que la marche aléatoire simple symétrique sur $\mathbb{Z}$ est irréductible récurrente. Il est par ailleurs facile de vérifier que la mesure $\nu = \sum_{i\in\mathbb{Z}} \delta_i$ est une mesure invariante. Comme elle est de masse infinie, on en déduit notamment que

$$\mathbf{E}_0\big[T_0^{(1)}\big] = \infty\ ,$$

ce qui n'est pas si simple à démontrer "à la main". De plus comme $\nu_0(0) = 1$, on voit que $\nu = \nu_0$ et donc

$$\forall i \in \mathbb{Z}\,, \quad \mathbf{E}_0\Big[\sum_{n=0}^{T_0^{(1)}-1} \mathbf{1}_{\{X_n=i\}}\Big] = 1\ ,$$

ce qui peut sembler surprenant si on prend $i = 10^9$ par exemple. □

**Exemple I.4.10** On considère une chaîne de naissance et de mort, c'est-à-dire que sa matrice de transition $Q = (p(i,j))_{i,j\in\mathbb{N}}$ est telle que $p(i,j) = 0$ dès que $|i-j| \geq 2$. On la suppose irréductible, ce qui est équivalent à supposer que $p(i,i+1)p(i+1,i) > 0$, pour tout $i \in \mathbb{N}$. En résumé des calculs faits aux exemples I.3.13, page 69, et I.2.26, page 47, on a les critères suivants.

— $Q$ est récurrente positive si $\sum_{n\geq 1}\prod_{1\leq k\leq n}\frac{p(k,k-1)}{p(k,k+1)} = \infty$ et $\sum_{n\geq 1}\prod_{1\leq k\leq n}\frac{p(k-1,k)}{p(k,k-1)} < \infty$

— $Q$ est récurrente nulle si $\sum_{n\geq 1}\prod_{1\leq k\leq n}\frac{p(k,k-1)}{p(k,k+1)} = \infty$ et $\sum_{n\geq 1}\prod_{1\leq k\leq n}\frac{p(k-1,k)}{p(k,k-1)} = \infty$.

— $Q$ est transiente si $\sum_{n\geq 1}\prod_{1\leq k\leq n}\frac{p(k,k-1)}{p(k,k+1)} < \infty$, mais elle admet une mesure invariante non-triviale. □

**Exemple I.4.11** Il existe une matrice de transition irréductible (apériodique) qui n'admet pas de mesure invariante. Le théorème I.4.2 implique alors qu'elle est nécessairement transiente. On peut construite une telle matrice $Q = (p(i,j))_{i,j\in\mathbb{N}}$, de la façon suivante. On suppose que

- $p(i,j) = 0$ si $j \notin \{0, i+1\}$, $i,j \in \mathbb{N}$.
- $p(i,i+1) = 1 - p(i,0) := a_i \in ]0,1[$.

On suppose que

$$\sum_{i\in\mathbb{N}}(1-a_i) < \infty\ . \tag{I.74}$$

On vérifie facilement que pour tout $i,j \in \mathbb{N}$, avec $j \geq 1$, on a

$$[Q^{j+1}](i,j) \geq (1-a_i)a_0 \dots a_{j-1} > 0\,,$$

ce qui implique que $Q$ est irréductible (et apériodique). Supposons que $Q$ admette une mesure invariante, c'est-à-dire une mesure $\mu$, *non-nulle*, telle que $0 \leq \mu(i) < \infty$, pour tout $i \in \mathbb{N}$ et $\mu.Q = \mu$. Si $j \geq 1$, l'équation $(\mu.Q)(j) = \mu(j)$, donne $\mu(j-1)p(j-1,j) = \mu(j)$, donc

$$\forall j \in \mathbb{N}^*\,, \quad \mu(j) = \mu(0)a_0a_1\dots a_{j-1}\ . \tag{I.75}$$

L'équation $(\mu.Q)(0) = \mu(0)$, donne $\sum_{i\in\mathbb{N}}\mu(i)p(i,0) = \mu(0)$. Or pour tout $n \geq 1$, on a

$$\begin{aligned}
\sum_{0\leq i\leq n}\mu(i)p(i,0) &= \mu(0)(1-a_0) + \sum_{1\leq i\leq n}\mu(i)p(i,0) \\
&= \mu(0)(1-a_0) + \sum_{1\leq i\leq n}\mu(0)a_0a_1\dots a_{i-1}(1-a_i) \\
&= \mu(0)(1-a_0) + \sum_{1\leq i\leq n}\mu(0)a_0a_1\dots a_{i-1} - \sum_{1\leq i\leq n}\mu(0)a_0a_1\dots a_i \\
&= \mu(0)(1-a_0) + \mu(0)a_0 - \mu(0)a_0a_1\dots a_n = \mu(0)(1-a_0a_1\dots a_n).
\end{aligned}$$

Clairement, $n \mapsto a_0 \dots a_n$, est une suite strictement décroissante positive et strictement inférieure à 1. On note $L$ sa limite. L'hypothèse (I.74) implique que $\lim_{i\to\infty} 1 - a_i = 0$ et donc que $\log a_i \sim -(1-a_i)$. On en déduit l'existence de $c \in ]0, \infty[$ telle que $-c = \sum_{i\in\mathbb{N}} \log a_i$. On voit donc que $L = e^{-c} > 0$. Or les calculs précédents impliquent que

$$\mu(0) = \sum_{i\in\mathbb{N}} \mu(i)p(i,0) = \lim_{n\to\infty} \sum_{0\le i\le n} \mu(i)p(i,0) = \mu(0)(1-L) \ .$$

Comme $L > 0$, cela implique que $\mu(0) = 0$ mais (I.75) implique que $\mu(j) = 0$, pour tout $j \in \mathbb{N}$, ce qui est une contradiction. La chaîne n'admet donc pas de mesure invariante. □

## EXERCICES.

**Exercice I.43** (*Récurrence positive et processus de naissance et de mort*) Soit

$$\big(\Omega\,;\,\mathscr{F}\,;\,(\mathscr{F}_n)_{n\ge 0}\,;\,(X_n)_{n\ge 0}\,;\,Q = (p(i,j))_{i,j\in\mathbb{N}}\,;\,\mathbf{P}_\mu\,,\,\mu\in\mathscr{M}_1(\mathbb{N})\,\big)$$

un processus de naissance et de mort que l'on suppose irréductible, c'est-à-dire que $p(i+1,i)p(i,i+1) > 0$, pour tout $i \in \mathbb{N}$. On a montré à l'exemple I.2.26 page 47 que $Q$ admet des mesure invariantes $\mu$ qui sont toutes de la forme

$$\mu(n) = \mu(0) \prod_{1\le k\le n} \tfrac{p(k-1,k)}{p(k,k-1)} \ , \quad n \ge 0 \ .$$

On rappelle l'exercice **??** page **??** où on a montré que $Q$ est récurrente ssi

$$\sum_{n\ge 1} \prod_{1\le k\le n} \tfrac{p(k,k-1)}{p(k,k+1)} \ = \ \infty \ .$$

1. Sous quelle hypothèse $Q$ est récurrente positive ?
2. Sous quelle hypothèse $Q$ est récurrente nulle ?
3. On suppose que $p(i,i) = 0$ et que $p(i,i+1) = \frac{1}{2} + \eta_i$, où $\eta_i \sim Ci^{-\alpha}$ avec $\alpha, C \in ]0,\infty[$. Pour quels $\alpha$ et $C$ la chaîne est transiente et récurrente ?
4. On suppose que $p(i,i) = 0$ et que $p(i,i+1) = \frac{1}{2} - \eta_i$, où $\eta_i \sim Ci^{-\alpha}$ avec $\alpha, C \in ]0,\infty[$. Pour quels $\alpha$ et $C$ la chaîne est positive récurrente ? □

**Exercice I.44** Soit $\mathbf{G} = (\mathbf{S}, \mathbf{A})$ un graphe simple, non-orienté dénombrable, connexe et sans boucle. La marche aléatoire simple est donc irréductible.

1. On suppose que la marche aléatoire simple sur $\mathbf{G}$ est récurrente. Montrer qu'elle est récurrente nulle si $\mathbf{G}$ est infini et récurrente positive si $\mathbf{G}$ est fini.
2. On suppose que $\mathbf{G}$ est fini. Trouver la loi invariante de la marche simple.
3. On considère un cavalier sur l'échiquier qui se déplace uniformément au hasard. Calculer le temps moyen de retour à son point de départ (cela dépend du point de départ). (On peut prendre un échiquier $4 \times 4$, si la motivation n'est pas au rendez-vous.) □

**Exercice I.45** (*Urne d'Ehrenfest*) La chaîne que nous présentons est un modèle rudimentaire de diffusion d'un gaz introduit par Ehrenfest : $N$ boules numérotées de 1 à $N$ sont réparties sur deux urnes, l'urne A et l'urne B. On note $X_0$ le nombre de boules dans l'urne A au temps 0. Au temps 1 on choisit uniformment au hasard un entier compris entre 1 et $N$, on prend la boule portant ce numéro et on la change d'urne ; on note $X_1$ le nombre de boules dans l'urne A. On recommence la même chose au temps 2 ...etc. On note $X_n$ le nombre de boules dans l'urne A à l'étape $n$.

1. Montrer que $(X_n)_{n\ge 0}$ est une chaîne de Markov dont l'espace d'états est $E = \{0,\dots,N\}$ et dont on calculera la matrice de transition.
2. Montrer que la chaîne est irréductible, ce qui implique, puisque l'espace d'états est fini, qu'elle est récurrente positive. Calculer la période de la chaîne.

3. On note $\pi$ sa loi invariante et on note $\varphi$ sa fonction génératrice.

$$\varphi(r) = \sum_{0 \leq i \leq N} r^i \pi(i) \; , \quad r \in [0,1] \; .$$

Trouver l'équation fonctionnelle vérifiée par $\varphi$ et en déduire $\pi$.

4. On suppose que $X_0 = N$. Comme la chaîne est récurrente $\mathbf{P}(\exists n \in \mathbb{N} \; : \; X_n = 0) = 1$. Faut-il nécessairement rejeter le modèle lorsque $N$ est très grand ? □

**Exercice I.46** (*Modèle de diffusion de Bernoulli-Laplace*) On considère deux urnes A et B qui contiennent chacune $N$ boules : il y a au total $2N$ boules. Les boules ont deux couleurs : rouge et verte. On note $X_0$ le nombre de boules rouges que contient l'urne A. On mélange les boules successivement en procédant comme suit : à chaque étape, on tire uniformément au hasard une boule dans l'urne A et indépendamment, une boule dans l'urne B et on échange d'urne les deux boules tirées. On note $X_n$ le nombre de boules rouges dans l'urne A à l'étape $n$. On observe que le nombre de boules dans chaque urne reste constant égal à $N$.

1. Montrer que $(X_n)_{n \geq 0}$ est une chaîne de Markov dont l'espace d'états est $E = \{0, \ldots, N\}$ et dont on calculera la matrice de transition.
2. Montrer que la chaîne est irréductible, ce qui implique, puisque l'espace d'états est fini, qu'elle est récurrente positive. Calculer la période de la chaîne.
3. On note $\pi$, la loi invariante. On remarque que la chaîne est un processus de vie et de mort sur l'espace d'états $\{0, \ldots, N\}$. Ecrire les équations déterminant $\pi$ et calculer $\pi$. (*Indication : on pourra utiliser le fait que le coefficient d'ordre $n$ du polynôme $(1-x^2)^n$ est le coefficient d'ordre $n$ du produit de polynômes $(1-x)^n(1+x)^n$.*) □

**Exercice I.47** (*Chaîne de renouvellement*) Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité sur lequel est définie une suite de variable $(\xi_k)_{k \geq 1}$, à valeurs dans $\mathbb{N}$, indépendantes et de loi $\nu : \nu(i) = \mathbf{P}(\xi_k = i)$, $i \in \mathbb{N}$. On suppose également que sur $(\Omega, \mathscr{F}, \mathbf{P})$ est également définie une variable $S_0 : \Omega :\to \mathbb{N}$, qui est indépendante de $(\xi_n)_{n \geq 1}$. On pose $S_k = S_0 + \xi_1 + \ldots + \xi_k$, pour tout $k \geq 1$. Pour tout $n \geq 0$, on pose

$$X_n = \inf \left\{ \ell - n \; : \; \ell \in \{n, n+1, \ldots\} \cap \{S_0, S_1, \ldots\} \right\} \, .$$

1. Montrer que $X_0 = S_0$ et montrer que $(X_n)_{n \geq 0}$ est une chaîne de Markov à valeurs dans $\mathbb{N}$ dont on précisera la matrice de transition.
2. On pose $K = \sup\{i \in \mathbb{N} \; : \; \nu(i) > 0\}$. Montrer que $\{i \in \mathbb{N} : i < K\}$ est une classe fermée de récurrence et $\mathbf{T} = \{i \in \mathbb{N} : i \geq K\}$ (qui est vide si $K = \infty$) est la classe des états transient.
3. Montrer que 0 est récurrent.
4. On suppose que $K = \infty$ et donc que la chaîne de renouvellement est irréductible récurrente. Montrer qu'elle est récurrente positive ssi $\sum_{i \in \mathbb{N}} i\nu(i) < \infty$. □

**Exercice I.48** (*File d'attente $M/G/1$*) Toutes les variables sont définies sur un même espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$. On suppose qu'à un guichet des clients arrivent : le temps que met chaque client à être servi est 1 (puis il quitte la file d'attente). Entre les temps $n$ et $n+1$, $\xi_{n+1}$ clients arrivent dans la file d'attente. On suppose que les variables $(\xi_n)_{n \geq 1}$ sont indépendantes et de même loi $\nu$. On note $X_0$ le nombre de clients présents dans la file au temps 0 et on note $X_n$ le nombre de clients dans la file à l'instant $n$

1. Montrer que $X_{n+1} = (X_n + \xi_{n+1} - 1)^+$, $n \geq 0$. Montrer que $(X_n)_{n \geq 0}$ est une chaîne de Markov dont on calculera la matrice de transition.
2. On suppose que $\nu(i) > 0$, pour tout $i \in \mathbb{N}$. Montrer que la chaîne est irréductible.
3. On suppose que $\nu(i) > 0$, pour tout $i \in \mathbb{N}$ et on pose $a = \sum_{i \in \mathbb{N}} i\nu(i)$. En utilisant la loi des grands nombres pour $(\xi_n - 1)_{n \geq 1}$, montrer que la marche aléatoire dont les sauts sont les variables $(\xi_n - 1)_{n \geq 1}$ est transiente si $a > 1$. En déduire que $(X_n)_{n \geq 0}$ est transiente si $a > 1$.
4. On suppose que $\nu(i) > 0$, pour tout $i \in \mathbb{N}$ et $a \leq 1$. On suppose également que $\mathbf{P}(X_0 = i) = 1$, pour un certain $i \geq 1$ et on pose $T_0 = \inf\{n \geq 0 \; : \; X_n = 0\}$, avec les conventions habituelles : $T_0 = \infty$ ssi pour tout $n \geq 0$, on a $X_n \geq 1$. Montrer que $(X_{n \wedge T_0})_{n \geq 0}$ est une sur-martingale positive. Pour tout $\ell \in \mathbb{N}$, on pose $T^*_\ell = \inf\{n \geq 0 \; : X_n > \ell\}$ ( avec les convention habituelles). En appliquant (et justifiant )un théorème d'arrêt pour la sur-martingale $(X_{n \wedge T_0})_{n \geq 0}$ au temps $T^*_\ell$, montrer que

$$\mathbf{P}(T^*_\ell < T_0) \leq i/\ell \; .$$

En déduire que $\mathbf{P}(T_0 = \infty) = 0$. En déduire que la chaîne est récurrente.

5. On suppose toujours que $\nu(i) > 0$, pour tout $i \in \mathbb{N}$. On suppose que $a < 1$ et que $\mathbf{P}(X_0 = i) = 1$, pour un certain $i \geq 1$. On pose $Y_n = X_{n\wedge T_0} - (a-1)n \wedge T_0$, pour tout $n \geq 0$. Montrer que $(Y_n)_{n\geq 0}$ est une martingale. En déduire que
$$\forall n \in \mathbb{N}\,, \quad \mathbf{E}[T_0] \leq \frac{i}{1-a}\;.$$
6. On se donne $\big(\Omega\,;\,\mathscr{F}\,;\,(\mathscr{F}_n)_{n\geq 0}\,;\,(X_n)_{n\geq 0}\,;\,Q\,;\,\mathbf{P}_\mu\,,\,\mu \in \mathscr{M}_1(\mathbb{N})\,\big)$ la chaîne globale dont la matrice de transition correspond à la file d'attente que l'on considère. Pour tout $i \in E$, on note $T_i = \inf\{n \geq 0 \,:\, X_n = i\}$, avec la convention habituelle que $T_i = \infty$ ssi pour tout $n \in \mathbb{N}$, on a $X_n \neq i$. On suppose que $\nu(i) > 0$, pour tout $i \in \mathbb{N}$, que $a \leq 1$. On a montré à la question précédente que si $a < 1$, alors $\mathbf{E}_i[T_0] \leq i/(1-a)$. En remarquant que la suite $X_n$ ne peut décroître au pire de 1, montrer que $\mathbf{E}_i[T_{i-1}] = \mathbf{E}_1[T_0]$, pour tout $i \geq 1$. En appliquant la propriété de Markov forte, montrer que pour tout $i \geq 1$, on a $\mathbf{E}_i[T_0] = C.i$ , où $C = \mathbf{E}_1[T_0]$, qui appartient a priori à $[0,\infty]$. En appliquant la propriété de Markov au temps 1, montrer que
$$C := \mathbf{E}_1[T_0] = \nu(0) + \sum_{i\geq 0} \nu(i+1)\mathbf{E}_i[T_0] = 1 + Ca\;.$$
On pose $T_0^{(1)} = \inf\{n \geq 1\,;\, X_n = 0\}$, le premier temps de *retour en* 0. Déduire de ce qui précède que si $a < 1$, alors $\mathbf{E}_0[T_0^{(1)}] = \nu(0)/(1-a)$. Cela montre que $(X_n)_{n\geq 0}$ est récurrente positive si $a < 1$.
7. On suppose que $\nu(i) > 0$, pour tout $i \in \mathbb{N}$, que $a \leq 1$. On suppose que la chaîne est positive récurrente c'est-à-dire que $\mathbf{E}_0[T_0^{(1)}] < \infty$. En appliquant la propriété de Markov au temps 1, montrer que
$$\mathbf{E}_0[T_0^{(1)}] = \nu(0) + \nu(1) + \sum_{i\geq 1} \nu(i+1)\mathbf{E}_i[T_0]\;.$$
Montrer que les question précédentes entraînent que $a < 1$.

*Pour résumer : la chaîne est irréductible dès que* $\nu(i) > 0$*, pour tout* $i \in \mathbb{N}$*; elle est récurrente positive ssi* $a := \sum_{i\in E} i\nu(i) < 1$*, récurrente nulle ssi* $a = 1$ *et transiente si* $a > 1$.
8. Interpréter le résultat précédent. Est-ce surprenant ? Dans le cas récurrent positif écrire les équations caractérisant la probabilité invariante. Cela a-t-il l'air calculable explicitement ? □

**Exercice I.49** (*Critère de récurrence positive Forster*) Soit
$$\big(\Omega\,;\,\mathscr{F}\,;\,(\mathscr{F}_n)_{n\geq 0}\,;\,(X_n)_{n\geq 0}\,;\,Q\,;\,\mathbf{P}_\mu\,,\,\mu \in \mathscr{M}_1(E)\,\big)\,,$$
une chaîne de Markov que l'on suppose **irréductible**. On suppose qu'il existe un reel strictement positif $\eta > 0$, une partie **finie** de $E$ notée $F \subset E$ et une fonction $h : E \to \mathbb{R}_+$, tels que
$$\forall i \in E\,, \quad Q.h(i) < \infty \quad \text{et} \quad \forall i \in E\backslash F\,, \quad Q.h(i) \leq h(i) - \eta\;.$$
Le but de l'exercice est de montrer que *la chaîne est nécessairement positive récurrente*. Pour cela on pose $T_F = \inf\{n \geq 0 \,:\, X_n \in F\}$, $T_F^{(1)} = \inf\{n \geq 1 \,:\, X_n \in F\}$ et plus généralement $T_F^{(k+1)} = \inf\{n > T_F^{(k)} \,:\, X_n \in F\}$ qui sont les temps de retour successifs de la chaîne dans $F$, avec la convention habituelle que $\inf \emptyset = \infty$.

1. On pose $Y_n = h(X_{n\wedge T_F}) + n\eta$, $n \geq 0$. Montrer que pour tout $i \in E\backslash F$, sous $\mathbf{P}_i$, $(Y_n)_{n\geq 0}$ est une $(\mathscr{F}_n)_{n\geq 0}$-sur-martingale positive. En déduire que
$$\forall i \in E\backslash F\,, \quad \mathbf{E}_i[T_F] \leq h(i)/\eta < \infty\;.$$
2. En déduire que pour tout $i \in E$, $\mathbf{E}_i[T_F^{(1)}] < \infty$ et plus généralement que $\mathbf{E}_i[T_F^{(k)}] < \infty$, pour tout $k \geq 1$.
3. On fixe $i \in F$ et on pose $M = \max_{i\in F} \mathbf{E}_i[T_F^{(1)}]$, qui est une quantité finie car $F$ est fini et d'après la question précédente. On pose $r = \mathbf{P}_i(X_{T_F^{(1)}} = i)$. Expliquer pourquoi $r > 0$ et justifier que
$$\mathbf{E}_i[T_i^{(1)}] \leq \sum_{k\geq 1} (1-r)^{k-1} rkM < \infty\,,$$
ce qui entraîne le résultat voulu. □

### I.4.b Théorèmes ergodiques.

On suppose ici que la chaîne de Markov sous-jacente est **irréductible récurrente**. Pour tout $n \in \mathbb{N}$ et tout $i \in E$, on pose

$$N_i(n) = \sum_{m=1}^{n} \mathbf{1}_{\{X_m = i\}} \;,$$

qui est le nombre de visites entre les temps 1 et $n$ de la chaîne en l'état $i$. On voit facilement que $N_i(n) \leq n$. On rappelle également que $T_i^{(p)}$ est le $p$-ième temps de retour de la chaîne en $i$ (avec la convention que $T_i^{(0)} = 0$). Comme la chaîne est irréductible récurrente, pour toute $\mu \in \mathscr{M}_1(E)$, et pour tout $p \geq 0$, on a $\mathbf{P}_\mu(T_i^{(p)} < \infty) = 1$. On voit alors que

$$\forall n \geq 0 \,, \quad T_i^{(N_i(n))} \leq n < T_i^{(N_i(n)+1)} \;. \tag{I.76}$$

On rappelle que $\mathscr{E}^{(p)}$, $p \geq 1$, sont les excursions de la chaîne en dehors de l'état $\{i\}$ :

$$\mathscr{E}^{(p)} = \big( X_{T_i^{(p-1)}}, X_{T_i^{(p-1)}+1}, X_{T_i^{(p-1)}+2}, \dots, X_{T_i^{(p)}-1}, \partial, \partial, \partial, \dots \big) \;.$$

D'après le théorème I.3.2, page 58, sous $\mathbf{P}_i$, ces excursions sont i.i.d. On pose ensuite

$$D_p = T_i^{(p)} - T_i^{(p-1)} \quad \text{et} \quad V_p(f) = \sum_{n=0}^{D_p - 1} f(X_{T_i^{(p-1)}+n})$$

La variable $D_p$ est la *durée de l'excursion* $\mathscr{E}^{(p)}$. Sous $\mathbf{P}_i$, la suite $(D_p)_{p\geq 1}$ est i.i.d. ainsi que la suite $(V_p(f))_{p\geq 1}$. On rappelle la notation $\nu_i$ pour la mesure $Q$-invariante donnée par le théorème I.4.2. Si $\langle \nu_i, |f| \rangle < \infty$, alors il est facile de vérifier que $V_1(f)$ est une variable $\mathbf{P}_i$-intégrable et que

$$\forall p \geq 1 \,, \quad \mathbf{E}_i[V_p(f)] = \mathbf{E}_i[V_1(f)] = \sum_{i' \in E} \nu_i(i') f(i') = \langle \nu_i, f \rangle \;.$$

Si $\nu$ est une mesure $Q$-invariante, le théorème I.4.3 $(iv)$ implique que $\nu = \nu(i)\nu_i$. Par conséquent, la condition $\langle \nu, |f| \rangle < \infty$ est équivalente à la condition $\langle \nu_i, |f| \rangle < \infty$, et on a

$$\forall p \geq 1 \,, \quad \mathbf{E}_i[V_p(f)] = \mathbf{E}_i[V_1(f)] = \sum_{i' \in E} \nu_i(i') f(i') = \tfrac{1}{\nu(i)} \langle \nu, f \rangle \;. \tag{I.77}$$

Lorsque la loi d'entrée de la chaîne n'est plus $\delta_i$, mais une loi quelconque, les propriétés d'indépendance des excursions sont encore préservées, comme le montre le lemme suivant.

**Lemme I.4.12** *On suppose la chaîne irréductible récurrente. Soient $\mu \in \mathscr{M}_1(E)$ et $i \in E$. Alors, les assertions suivantes sont vérifiées.*

*(i) Sous $\mathbf{P}_\mu$, les excursions de $\mathbf{X}$ hors de $i$, notées $\mathscr{E}^{(p)}$, $p \geq 1$, sont indépendantes et la suite $\mathscr{E}^{(p)}$, $p \geq 2$, est i.i.d. de même loi que $\mathscr{E}^{(1)}$ sous $\mathbf{P}_i$.*

*(ii) Soit $\nu$, une mesure $Q$-invariante et soit $f : E \to \mathbb{R}$ telle que $\langle \nu, |f| \rangle < \infty$. Alors*

$$\mathbf{P}_\mu\text{-}p.s. \quad \frac{1}{p}\big( V_1(f) + V_2(f) + \dots + V_p(f) \big) \xrightarrow[p\to\infty]{} \tfrac{1}{\nu(i)} \langle \nu, f \rangle \;.$$

**Preuve :** soit $\ell \geq 2$ et $\Phi_1, \ldots, \Phi_\ell : (E^*)^{\mathbb{N}} \to \mathbb{R}$, $\ell$ fonctions $\mathscr{P}(E^*)^{\otimes \mathbb{N}}$-mesurables bornées. Puisque la chaîne est récurrente, on a $\mathbf{P}_\mu(T_i^{(1)} < \infty) = 1$ et on remarque que $\mathbf{P}_\mu$-p.s.

$$\Phi_1(\mathscr{E}^{(1)})\Phi_2(\mathscr{E}^{(2)}) \ldots \Phi_\ell(\mathscr{E}^{(\ell)}) = \Phi_1(\mathscr{E}^{(1)})\Phi_2\big(\mathscr{E}^{(1)}(\theta_{T_i^{(1)}}\mathbf{X})\big) \ldots \Phi_\ell\big(\mathscr{E}^{(\ell-1)}(\theta_{T_i^{(1)}}\mathbf{X})\big).$$

Par Markov au temps $T_i^{(1)}$ et le théorème I.3.2, on obtient $\mathbf{P}_\mu$-p.s. que

$$\begin{aligned} \mathbf{E}_\mu\Big[\prod_{1\leq k\leq \ell} \Phi_k(\mathscr{E}^{(k)}) \,\Big|\, \mathscr{F}_{T_i^{(1)}}\Big] &= \Phi_1(\mathscr{E}^{(1)})\mathbf{E}_i\Big[\prod_{1\leq k\leq \ell-1} \Phi_{k+1}(\mathscr{E}^{(k)})\Big] \\ &= \Phi_1(\mathscr{E}^{(1)}) \prod_{1\leq k\leq \ell-1} \mathbf{E}_i\big[\Phi_{k+1}(\mathscr{E}^{(1)})\big] \end{aligned}$$

et le point $(i)$ s'obtient en intégrant sous $\mathbf{P}_\mu$. Par la loi des grands nombres, on a $\mathbf{P}_\mu$-p.s.

$$\lim_p \frac{1}{p}\big(V_1(f) + V_2(f) + \ldots + V_p(f)\big) = \lim_p \frac{1}{p-1}\big(V_2(f) + \ldots + V_p(f)\big) = \mathbf{E}_i[V_1(f)]$$

et (I.77) permet de conclure. ■

**Théorème I.4.13** **(Théorème ergodique quotient)** *On suppose la chaîne irréductible récurrente. Soit $\nu$, une mesure $Q$-invariante et soient $f, g : E \to \mathbb{R}$, deux fonctions telles que*

$$\langle \nu, |f|\rangle < \infty\ , \quad \langle \nu, |g|\rangle < \infty \quad \text{et} \quad \langle \nu, g\rangle \neq 0\ .$$

*Alors, pour toute $\mu \in \mathscr{M}_1(E)$, on a*

$$\mathbf{P}_\mu\text{-p.s.} \qquad \frac{1}{N_i(n)}\big(f(X_0) + f(X_1) + \ldots + f(X_n)\big) \xrightarrow[n\to\infty]{} \frac{1}{\nu(i)}\langle \nu, f\rangle\ .$$

*et*

$$\mathbf{P}_\mu\text{-p.s.} \qquad \frac{f(X_0) + f(X_1) + \ldots + f(X_n)}{g(X_0) + g(X_1) + \ldots + g(X_n)} \xrightarrow[n\to\infty]{} \frac{\langle \nu, f\rangle}{\langle \nu, g\rangle}\ .$$

**Preuve :** on pose $S_p(f) = V_2(f) + \ldots + V_p(f)$ si $p \geq 2$, et $S_0(f) = S_1(f) = 0$. On pose également

$$U_n = \sum_{0\leq k< n\wedge T_i^{(1)}} f(X_k) \quad \text{et} \quad R_n(f) = \sum_{T_i^{(N_i(n))}\leq k\leq n} f(X_k)\,,$$

si bien que

$$f(X_0) + \ldots + f(X_n) = U_n(f) + S_{N_i(n)}(f) + R_n(f)\ .$$

Comme la chaîne est irréductible récurrente, $\mathbf{P}_\mu$-p.s. $\lim_n N_i(n) = \infty$. La loi des grands nombres établie au lemme I.4.12 $(ii)$ implique que

$$\mathbf{P}_\mu\text{-p.s.} \quad \lim_{n\to\infty} S_{N_i(n)}(f)/N_i(n) = \frac{1}{\nu(i)}\langle \nu, f\rangle\ .$$

Il est maintenant facile de voir que $\mathbf{P}_\mu$-p.s. $\lim_n U_n(f) = V_1(f)$, et donc $\lim_n U_n(f)/N_i(n) = 0$. Enfin, on voit que $\mathbf{P}_\mu$-p.s.

$$\Big|\frac{1}{N_i(n)}R_n(f)\Big| \leq \frac{1}{N_i(n)}R_n(|f|) = \frac{1}{N_i(n)}\big(S_{N_i(n)+1}(|f|) - S_{N_i(n)}(|f|)\big) \xrightarrow[n\to\infty]{} 0\,,$$

ce qui, combiné avec les autres limites, implique la première convergence du théorème. La seconde est une conséquence immédiate de la première. ■

**Corollaire I.4.14** *On suppose la chaîne irréductible récurrente. Alors, les assertions suivantes sont vérifiées.*

*(i) Si la chaîne est récurrente nulle, alors pour toute $\mu \in \mathscr{M}_1(E)$, on a*

$$\mathbf{P}_\mu\text{-}p.s. \quad \forall i \in E\,, \qquad \frac{N_i(n)}{n} \xrightarrow[n\to\infty]{} 0\,.$$

*(ii) Si la chaîne est récurrente positive, alors on note $\pi$ son unique loi invariante, et pour toute $\mu \in \mathscr{M}_1(E)$, on a*

$$\mathbf{P}_\mu\text{-}p.s. \quad \forall i \in E\,, \qquad \frac{N_i(n)}{n} \xrightarrow[n\to\infty]{} \pi(i) = \frac{1}{\mathbf{E}_i[T_i^{(1)}]}\,.$$

**Preuve :** on suppose la chaîne récurrente nulle et on note $\nu$, une mesure invariante. Le corollaire I.4.8 implique que $\nu$ est de masse totale infinie : $\langle \nu \rangle = \infty$. Cela implique clairement que l'espace d'états est infini. On se donne une suite de sous-ensembles *finis* $F_\ell \subset E$, $\ell \geq 0$, tel que $F_\ell \subset F_{\ell+1}$ et $\bigcup_\ell F_\ell = E$. On a donc $\lim_\ell \nu(F_\ell) = \infty$. On fixe $i \in E$ et on applique le théorème ergodique quotient à la fonction $f_\ell = \mathbf{1}_{F_\ell}$. Il est clair que $\langle \nu, f \rangle = \nu(F_\ell)$. De plus, on a clairement

$$\forall n \in \mathbb{N}\,, \quad \frac{n}{N_i(n)} \geq \frac{1}{N_i(n)}\big(\mathbf{1}_{F_\ell}(X_0) + \ldots + \mathbf{1}_{F_\ell}(X_n)\big)\,.$$

Pour toute loi d'entrée $\mu$, on a donc $\mathbf{P}_\mu$-p.s. $\liminf_n n/N_i(n) \geq \nu(F_\ell)/\nu(i)$ et en faisant tendre $\ell$ vers l'infini, on a donc $\mathbf{P}_\mu(\liminf_n n/N_i(n) = \infty) = 1$, ce qui implique $(i)$.

On suppose ensuite que la chaîne récurrente positive. On applique le théorème I.4.13 ergodique quotient avec $\nu = \pi$ et à la fonction constante à 1 qui est telle que $\langle \pi, \mathbf{1} \rangle = 1$. On a alors pour toute loi d'entrée $\mu$, $\mathbf{P}_\mu$-p.s. $\lim_n n/N_i(n) = 1/\pi(i)$, ce qui implique $(ii)$. ■

**Remarque I.4.15** On voit que $N_i(n)/n$ représente la fraction de temps passé par la chaîne en l'état $i$ entre les instants 1 et $n$. Un état est récurrent nul ssi cette fraction tend vers 0 et il est récurrent positif ssi cette fraction tend vers un nombre strictement positif (la chaîne passe une fraction strictement positive de son temps en l'état $i$). Si $i$ est récurrent nul, il n'est en général pas très facile de trouver un équivalent à $N_i(n)$ (c'est-à-dire qu'il n'y a pas de résultat général donnant un equivalent de $N_i(n)$ à l'aide des fonctions usuelles telles que $n/\log n$, $\sqrt{n}$ ... etc). □

Donnons un second corollaire, qui est aussi important que le théorème ergodique quotient : nous l'appellons théorème ergodique et nous l'énonçons comme un théorème bien qu'il soit une conséquence immédiate des deux résultats précédents.

**Théorème I.4.16** **(Théorème ergodique)** *On suppose la chaîne irréductible et récurrente positive. On note $\pi$ son unique probabilité invariante. Soit $f : E \to \mathbb{R}$, telle que $\langle \pi, |f| \rangle < \infty$. Alors pour toute $\mu \in \mathscr{M}_1(E)$, on a*

$$\mathbf{P}_\mu\text{-}p.s. \quad \tfrac{1}{n}\big(f(X_0) + \ldots + f(X_n)\big) \xrightarrow[n\to\infty]{} \langle \pi, f \rangle\,.$$

**Preuve :** on fixe $i \in E$ et on remarque que

$$\frac{1}{n}\big(f(X_0) + \ldots + f(X_n)\big) = \frac{N_i(n)}{n} \cdot \frac{1}{N_i(n)}\big(f(X_0) + \ldots + f(X_n)\big)\,.$$

On conclut alors par le théorème ergodique quotient et le corollaire I.4.14 $(ii)$. ■

## EXERCICES.

**Exercice I.50** Soit $(X_n)_{n\geq 0}$ une marche aléatoire simple sur $\mathbb{Z}$ qui est définie sur un espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$. Montrer qu'il existe $\ell \in \mathbb{R}_+$ que l'on calculera explictement tel que $\mathbf{P}$-p.s. on ait

$$\lim_{n\to\infty} \Big( \sum_{1\leq k\leq n} \frac{1}{(1+|X_n|)(2+|X_n|)} \Big) \Big/ \Big( \sum_{1\leq k\leq n} 2^{-|X_n|} \Big) \; = \ell \; .$$

□

**Exercice I.51** (*Théorèmes ergodiques pour les serpents*) On rappelle l'exercice I.22 page 53 sur les serpents associés à une chaîne. Soit

$$\big(\Omega\, ;\, \mathscr{F}\, ;\, (\mathscr{F}_n)_{n\geq 0}\, ;\, (X_n)_{n\geq 0}\, ;\, Q = (p(i,j))_{i,j\in E}\, ;\, \mathbf{P}_\mu\, ,\; \mu \in \mathscr{M}_1(E)\; \big)$$

une chaîne de Markov que l'on suppose **irréductible récurrente**.

1. On note $\nu$ une mesure $Q$-invariante. On fixe $k \geq 1$ et on se donne deux fonctions $f, g : E^{k+1} \to \mathbb{R}$ telles que

$$\sum_{i_0,\ldots,i_k\in E} \nu(i_0)p(i_0,i_1)\ldots p(i_{k-1},i_k)\,\big|\, f(i_0,i_1,\ldots i_k)\,\big| \; < \infty$$

   et la même chose en remplaçant $g$ par $f$. Montrer que pour toute loi d'entrée $\mu \in \mathscr{M}_1(E)$, on a

$$\mathbf{P}_\mu \text{ - p.s.} \quad \frac{f(X_0,\ldots,X_k) + f(X_1,\ldots,X_{k+1}) + \ldots + f(X_n,\ldots,X_{n+k})}{g(X_0,\ldots,X_k) + g(X_1,\ldots,X_{k+1}) + \ldots + g(X_n,\ldots,X_{n+k})} \quad \xrightarrow[n\to\infty]{} \quad \frac{\sum_{i_0,\ldots,i_k\in E} \nu(i_0)p(i_0,i_1)\ldots p(i_{k-1},i_k) f(i_0,i_1,\ldots i_k)}{\sum_{i_0,\ldots,i_k\in E} \nu(i_0)p(i_0,i_1)\ldots p(i_{k-1},i_k) g(i_0,i_1,\ldots i_k)}$$

2. On suppose que $Q$ est positive récurrente. On note $\pi$ la probabilité invariante. On fixe $k \geq 1$, et on se donne une fonction $f : E^{k+1} \to \mathbb{R}$ telles que

$$\sum_{i_0,\ldots,i_k\in E} \pi(i_0)p(i_0,i_1)\ldots p(i_{k-1},i_k)\,\big|\, f(i_0,i_1,\ldots i_k)\,\big| \; < \infty \; .$$

   Montrer que pour toute loi d'entrée $\mu \in \mathscr{M}_1(E)$, on a

$$\mathbf{P}_\mu \text{ - p.s.,} \quad \tfrac{1}{n}\big( f(X_0,\ldots X_k) + f(X_1,\ldots,X_{k+1}) + \ldots + f(X_n,\ldots,X_{n+k}) \big) \quad \xrightarrow[n\to\infty]{} \quad \sum_{i_0,\ldots,i_k\in E} \pi(i_0)p(i_0,i_1)\ldots p(i_{k-1},i_k) f(i_0,i_1,\ldots i_k) \; .$$

3. On suppose qu'un ivrogne tape sur un clavier totalement au hasard en produisant une suite de lettres. Quelle est la fréquence asymptotique de l'apparition de la phrase "in vino veritas" ? (Pour simplifier, sur le clavier, il n'y a pas d'accent, ni de majuscule, ni de ponctuation : l'espace est compté comme un caractère : au total, il y a donc 27 touches sur le clavier.) □

**Exercice I.52** (*Gestion de l'argent liquide dans une banque*) On note chaque jour la quantité d'argent liquide présente dans les coffres d'une banque à son ouverture. Cette quantité évolue de la manière suivante. Durant une journée, entre l'ouverture et la fermeture de la banque, *ou bien* 1 kilo-euro a été déposé par les clients avec probabilité $1/2$, *ou bien* 1 kilo-euro a été retiré avec probabilité $1/2$. Si à la fin de la journée, il n'y a plus d'argent liquide dans les coffres, le banquier appelle une compagnie de transport de fonds qui, durant la nuit, apporte $s$ kilo-euros (ici, $s \geq 2$). Comme immobiliser de l'argent sous forme de liquide coûte de l'argent, si à la fin de la journée il y a $S$ kilo-euros (ici, $S > s$), le banquier appelle une compagnie de transport de fonds qui, durant la nuit, emporte $S - s$ kilo-euros et les apporte à la maison mère qui place cet argent : la banque ouvre le lendemain avec $s$ kilo-euros d'argent liquide dans ses coffres. On note $X_n$ la quantité d'argent liquide à l'ouverture de la banque le $n$-ième jour.

1. Montrer que $(X_n)_{n\geq 0}$ est une chaîne de Markov dont l'espace d'états est $\{1,\ldots,S-1\}$ et dont la matrice de transition est donnée par $p(i,i+1) = p(i,i-1) = 1/2$, pour tout $2 \leq i \leq S-2$ et

$$p(1,2) = p(1,s) = p(S-1,S-2) = p(S-1,s) = 1/2 \; .$$

2. Montrer que cette chaîne est irréductible et calculer sa période en fonction de la parité de $s$ et de $S$.

3. Calculer sa probabilité invariante, qui est notée $\pi$.
4. Chaque kilo-euro immobilisé en liquide pendant une journée à la banque coûte $r$ euros par jour. Par ailleurs la compagnie de transport de fonds facture $c$ euros par kilo-euro transporté. On note $C_n$ le coût pour la banque de sa gestion de l'argent liquide du matin du jour 0 au matin du jour $n+1$. Montrer qu'il existe une fonction $C(s, S, r, c)$ telle que

$$\mathbf{P}\text{-p.s.} \quad \lim_n C_n/n = C(s, S, r, c) \ .$$

5. Les constantes $r$ et $c$ étant fixées par le marché, trouver les valeurs de $s$ et $S$ qui minimisent le coût journalier moyen pour la banque. □

**Exercice I.53** (*Théorème Central-Limite pour les chaînes*) Soit

$$\left(\Omega \,;\, \mathscr{F} \,;\, (\mathscr{F}_n)_{n\geq 0} \,;\, (X_n)_{n\geq 0} \,;\, Q = (p(i,j))_{i,j\in E} \,;\, \mathbf{P}_\mu \,,\, \mu \in \mathscr{M}_1(E) \right)$$

une chaîne de Markov que l'on suppose **irréductible, récurrente positive**. On note $\pi$ sa loi invariante. Soit $f : E \to \mathbb{R}$ telle que $\langle \pi, f^2 \rangle < \infty$. Jensen implique que $\langle \pi, |f| \rangle^2 < \langle \pi, f^2 \rangle < \infty$. On note son écart-type par

$$\sigma(f) = \sqrt{\langle \pi, f^2 \rangle - \langle \pi, f \rangle^2} \ .$$

On pose $g = f - \langle \pi, f \rangle < \infty$, si bien que $\langle \pi, g^2 \rangle = \sigma^2(f)$. On rappelle la notation $V_k(f)$.

1. Pour toute loi d'entrée $\mu \in \mathscr{M}_1(E)$, et pour tout $\eta > 0$, montrer que

$$\lim_{n\to\infty} \mathbf{P}_\mu \Big( n^{-1/2} \max_{1\leq k\leq n} V_k(|g|) \geq \eta \Big) = 0 \ .$$

2. On note $Z$ une variable Gaussienne standard réelle (de moyenne 0 et de variance 1). Pour toute loi d'entrée $\mu \in \mathscr{M}_1(E)$, montrer que

$$\frac{1}{\sigma(f)\sqrt{n}} \Big( \sum_{1\leq k\leq n} f(X_k) - \, n\langle \pi, f \rangle \Big) \quad \text{sous } \mathbf{P}_\mu \quad \xrightarrow[n\to\infty]{(\text{loi})} Z \ .$$

□

### I.4.c Convergence vers la loi stationnaire.

On suppose la chaîne sous-jacente *irréductible et récurrente positive*. On note $\pi$ son unique probabilité invariante. On s'intéresse au problème de la convergence en loi de la chaîne. C'est-à-dire que pour toute loi d'entrée $\mu$ et tout état $i$, on se demande si

$$\mathbf{P}_\mu(X_n = i) \xrightarrow[n\to\infty]{} \ ? \tag{I.78}$$

Supposons que la limite dans (I.78) existe. On remarque que $0 \leq N_i(n)/n \leq 1$. Par le corollaire I.4.14 et le théorème de convergence dominée, on voit que $\lim_n \mathbf{E}_\mu[N_i(n)/n] = \pi(i)$ et donc $\mathbf{E}_\mu[N_i(n)/n] = \frac{1}{n}\big(\mathbf{P}_\mu(X_0 = i) + \ldots + \mathbf{P}_\mu(X_n = i)\big) \to \pi(i)$. Si la limite dans (I.78) existe, par Cesaro, cette limite vaut nécessairement $\pi(i)$. On peut donc affiner la question de la convergence en loi en se demandant, pour toute loi d'entrée $\mu$ et tout état $i$, si

$$\mathbf{P}_\mu(X_n = i) \xrightarrow[n\to\infty]{} \pi(i) \quad ? \tag{I.79}$$

Supposons maintenant que la chaîne soit $d$-périodique avec $d \geq 2$. On note $E_0, \ldots, E_{d-1}$, la partition de l'espace d'états $E$ donnée par la proposition I.2.19. On fixe par exemple $i \in E_0$, alors on observe que $\mathbf{P}_i$-p.s. $X_n \in E_{\overline{n}}$, ce qui implique que $\mathbf{P}_i(X_n = i) = 0$, si $n$ n'est pas divisible par $d$. Cela entraîne que la suite de réels $\mathbf{P}_i(X_n = i)$, $n \geq 0$, *ne converge pas vers* $\pi(i)$ et donc que l'on a pas (I.79).

Cette brève discussion, montre que la convergence en loi de la chaîne (I.79) n'a pas lieu lorsque la chaîne a une période plus grande que 2. Nous allons cependant montrer que (I.79) a lieu pour toutes les chaînes irréductibles, récurrentes positives et apériodiques en utilisant un argument probabiliste basé sur la notion de *couplage*.

**Théorème I.4.17** **(Théorème de convergence vers la loi stationnaire)** *On suppose la chaîne irréductible, récurrente positive et apériodique. On note $\pi$ son unique loi invariante. Alors, pour toute $\mu \in \mathscr{M}_1(E)$, on a*

$$\sum_{j\in E} |\mathbf{P}_\mu(X_n = j) - \pi(j)| \quad \xrightarrow[n\to\infty]{} 0 \; .$$

**Preuve :** la preuve que nous présentons est une preuve par *couplage*, dont l'idée générale est la suivante : sur un même espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$ on définit deux chaînes *indépendantes* $(X_n)_{n\in\mathbb{N}}$ et que $(X'_n)_{n\in\mathbb{N}}$ de même matrice de transition $Q$ et on suppose que $X_0$ a pour loi $\mu$ et $X'_0$ a pour loi $\pi$. On veut montrer que ces chaîne se rencontrent , c'est-à-dire que $T = \inf\{n \in \mathbb{N} : X_n = X'_n\}$ est fini $\mathbf{P}$-presque sûrement (ce qui est fait en utilisant la récurrence positive de $Q$ ainsi que son apériodicité). Il est assez naturel de penser ensuite que si on pose $X''_n = X'_n$ si $n \leq T$ et $X''_n = X_n$ si $n > T$, alors $(X''_n)_{n\in\mathbb{N}}$ est une chaîne de Markov de matrice de transition $Q$ (ce n'est pas aussi simple) et comme $X''_0 = X'_0$ suit la loi $\pi$, la loi de $X''_n$ serait $\pi$. Avec, cette idée en tête, de façon informelle, si $n > T$, $X_n$ aurait pour loi $\pi$ et « donc » $|\mathbf{P}(X_n = i) - \pi(i)| \leq \mathbf{P}(n \leq T)$ qui tend vers 0 lorsque $n$ tend vers $\infty$. Tout cela doit être précisé : voir l'exercice I.58 page 97 pour dissiper quelques idées naïves sur cette approche. Pour rendre rigoureuse cette approche, on commence par voir le couple $(X_n, X'_n)$ comme une chaîne de Markov à valeurs dans $E^2$ dont on calcule la matrice de transition comme suit.

On définit une matrice indexée par $E^2 \times E^2$, notée

$$Q_* = \big(p_*(\,(i,i'),(j,j')\,)\big)_{(i,i')\,,\,(j,j')\in E^2}$$

en posant $p_*(\,(i,i'),(j,j')\,) = p(i,j)p(i',j')$. On veut montrer que $Q_*$ est une matrice de transition irréductible et positive récurente. On remarque tout d'abord que

$$\sum_{(j,j')\in E^2} p_*(\,(i,i'),(j,j')\,) = \Big(\sum_{j\in E} p(i,j)\Big)\Big(\sum_{j'\in E} p(i',j')\Big) = 1\cdot 1 = 1 \; ,$$

ce qui montre que $Q_*$ est une matrice de transition sur $E^2$. On vérifie facilement par récurrence que pour tout $n \geq 1$, et tous $i, i', j, j' \in E$, on a

$$[Q_*^n](\,(i,i'),(j,j')\,) = [Q^n](i,j)\cdot[Q^n](i',j') \; .$$

On fixe $(i,i'),(j,j') \in E^2$. Puisque que $Q$ est irréductible et apériodique, le théorème I.2.16 $(ii)$ implique l'existence de $n_0 \geq 1$ (qui dépend de $i, i', j$ et $j'$), tel que

$$\forall n \geq n_0 \; , \quad [Q^n](i,j)\cdot[Q^n](i',j') > 0 \; .$$

Cela montre que $Q_*$ est irréductible. Signalons que *c'est le seul endroit dans la preuve où on utilise le fait que $Q$ soit apériodique*. Pour tout $(i,i') \in E^2$, on pose $\pi_*((i,i')) = \pi(i)\pi(i')$. On voit que $\pi_*$ est la loi produit $\pi\otimes\pi$ sur $E^2$. C'est donc une probabilité sur $E^2$. De plus, pour tous $j, j' \in E$, on a

$$\begin{aligned}\sum_{(i,i')\in E^2} \pi_*(\,(i,i')\,)\,p_*(\,(i,i'),(j,j')\,) &= \Big(\sum_{i\in E}\pi(i)p(i,j)\Big)\Big(\sum_{i'\in E}\pi(i')p(i',j')\Big)\\ &= \pi(j)\,\pi(j') = \pi_*((j,j'))\; ,\end{aligned}$$

ce qui montre que $\pi_*$ est une loi $Q_*$-invariante. La matrice de transition $Q_*$ satisfait donc la condition $(i)$ du théorème I.4.7, qui implique que $Q_*$ est récurrente positive. Mentionnons que *nous aurons besoin uniquement du fait que $Q_*$ soit irréductible récurrente*.

Pour simplifier les notations, on suppose qu'il est possible de définir une loi de probabilité $\mathbf{P}$ sur $(\Omega, \mathscr{F})$, et deux suites de variables $X_n, X'_n : \Omega \to E$, $n \geq 0$, ($\mathscr{F}$-mesurables, bien sûr) qui remplissent les conditions suivantes.

- Sous $\mathbf{P}$, $(X_n)_{n\geq 0}$ est une chaîne de Markov de matrice de transition $Q$ et de loi d'entrée $\mu \in \mathscr{M}_1(E)$.
- Sous $\mathbf{P}$, $(X'_n)_{n\geq 0}$ est une chaîne de Markov de matrice de transition $Q$ et de loi d'entrée $\pi$, l'unique probabilité invariante de $Q$.
- Sous $\mathbf{P}$, les deux suites $(X_n)_{n\geq 0}$ et $(X'_n)_{n\geq 0}$ sont *indépendantes*.

**Etape 1** : *on veut montrer que, sous* $\mathbf{P}$, *la suite* $\big( (X_n, X'_n) \big)_{n\geq 0}$ *est une chaîne de Markov à valeurs dans* $E^2$ *de matrice de transition* $Q_*$ *et de loi d'entrée* $\mu \otimes \pi \in \mathscr{M}_1(E^2)$.

**Preuve :** pour tout $n \in \mathbb{N}$ et tous $(i_0, i'_0), \ldots, (i_n, i'_n) \in E^2$ , on a

$$\begin{aligned} &\mathbf{P}\big((X_0, X'_0) = (i_0, i'_0); \ldots ; (X_n, X'_n) = (i_n, i'_n)\big) \\ &\quad = \mathbf{P}\big(X_0 = i_0; \ldots ; X_n = i_n\big)\mathbf{P}\big(X'_0 = i'_0; \ldots ; X'_n = i'_n\big) \\ &\quad = \mu(i_0)p(i_0, i_1) \ldots p(i_{n-1}, i_n) \cdot \pi(i'_0)p(i'_0, i'_1) \ldots p(i'_{n-1}, i'_n) \\ &\quad = (\mu \otimes \pi)\big( (i_0, i'_0) \big) \cdot p_*\big( (i_0, i'_0), (i_1, i'_1)\big) \ \ldots \ p_*\big( (i_{n-1}, i'_{n-1}), (i_n, i'_n)\big) \ , \end{aligned}$$

ce qui montre bien le résultat voulu par le lemme I.1.4. □

**Etape 2** : *on pose ensuite* $T = \inf\{n \geq 0 \ : \ X_n = X'_n\}$, *avec la convention que* $T = \infty$ *ssi* $X_n \neq X'_n$, *pour tout* $n \in \mathbb{N}$. *On veut montrer qu'alors :* $\mathbf{P}(T < \infty) = 1$.

**Preuve :** on fixe $i \in E$ et on note $T_{(i,i)} = \inf\{n \geq 0 \ : \ (X_n, X'_n) = (i, i)\}$, avec la convention que $T_{(i,i)} = \infty$ ssi $(X_n, X'_n) \neq (i, i)$, pour tout $n \in \mathbb{N}$. D'après l'étape (1), et puisque $Q_*$ est irréductible récurrente, la chaîne de Markov $\big( (X_n, X'_n) \big)_{n\geq 0}$ visite une infinité de fois chaque point de $E^2$ et notamment $(i, i)$, ce qui montre que $\mathbf{P}(T_{(i,i)} < \infty) = 1$. Or il est clair que $T \leq T_{(i,i)}$, ce qui permet de conclure. □

**Etape 3** : *on veut montrer que* $\mathbf{P}(X_n = j; T \leq n) = \mathbf{P}(X'_n = j; T \leq n)$, $n \in \mathbb{N}$, $j \in E$.

**Preuve :** on a les égalité suivantes

$$\begin{aligned} \mathbf{P}(X_n = j; T \leq n) &= \sum_{0\leq m\leq n} \sum_{i\in E} \mathbf{P}(X_n = j; T = m; X_m = i) \\ &= \sum_{0\leq m\leq n} \sum_{i\in E} \mathbf{P}(T = m; X_m = i)[Q^{n-m}](i, j) \\ &= \sum_{0\leq m\leq n} \sum_{i\in E} \mathbf{P}(T = m; X'_m = i)[Q^{n-m}](i, j) \\ &= \sum_{0\leq m\leq n} \sum_{i\in E} \mathbf{P}(X'_n = j; T = m; X'_m = i) \\ &= \mathbf{P}(X'_n = j; T \leq n), \end{aligned}$$

ce qui montre le résultat voulu. □

**Fin de la preuve :** on déduit de l'étape (3) que pour tout $n \in \mathbb{N}$ et tout $j \in E$,

$$\mathbf{P}(X_n = j) \quad = \quad \mathbf{P}(X_n = j; T \leq n) + \mathbf{P}(X_n = j; T > n)$$

$$\begin{aligned} &= \mathbf{P}(X'_n = j; T \leq n) + \mathbf{P}(X_n = j; T > n) \\ &= \mathbf{P}(X'_n = j) - \mathbf{P}(X'_n = j; T > n) + \mathbf{P}(X_n = j; T > n), \end{aligned}$$

ce qui implique

$$\big| \mathbf{P}(X_n = j) - \mathbf{P}(X'_n = j) \big| \leq \mathbf{P}(X_n = j; T > n) + \mathbf{P}(X'_n = j; T > n) \ .$$

Comme $\sum_{j \in E} \mathbf{P}(X_n = j; T > n) = \sum_{j \in E} \mathbf{P}(X'_n = j; T > n) = \mathbf{P}(T > n)$, on a donc

$$\forall n \in \mathbb{N}, \quad \sum_{j \in E} \big| \mathbf{P}(X_n = j) - \mathbf{P}(X'_n = j) \big| \leq 2\mathbf{P}(T > n).$$

On observe maintenant que $\mathbf{P}(X'_n = j) = \pi(j)$. Ce qui précède, combiné à l'étape (2), entraîne

$$\sum_{j \in E} \big| \mathbf{P}(X_n = j) - \pi(j) \big| \leq 2\mathbf{P}(T > n) \quad \xrightarrow[n \to \infty]{} 0 \ ,$$

ce qui termine la preuve du théorème. ■

## EXERCICES.

**Exercice I.54** Soit $E$ un ensemble dénombrable. Soient $\mu, \nu \in \mathscr{M}_1(E)$, deux mesures de probabilité sur $E$. Leur *distance en variation* est donnée par

$$d(\mu, \nu) = \frac{1}{2} \sum_{i \in E} |\mu(i) - \nu(i)| \ .$$

1. Montrer que $(\mathscr{M}_1(E), d)$ est un espace métrique complet.
2. Soient $\mu, \nu \in \mathscr{M}_1(E)$, deux mesures de probabilité sur $E$. Montrer que

$$d(\mu, \nu) = \sup_{A \subset E} |\mu(A) - \nu(A)| \ = \ \sup_{A \subset E} \big(\mu(A) - \nu(A)\big) \ .$$

3. Soient $\mu, \nu \in \mathscr{M}_1(E)$, deux mesures de probabilité sur $E$. Soient $(\Omega, \mathscr{F}, \mathbf{P})$ un espace de probabilité et $Y, Z : \Omega \to E$, deux variables aléatoires $\mathscr{F}$-mesurables telles que $Y$ ait pour loi $\mu$ sous $\mathbf{P}$ et $Z$ ait pour loi $\nu$ sous $\mathbf{P}$. Alors on a

$$\mathbf{P}(Y = Z) \leq 1 - d(\mu, \nu) \ .$$

4. Soient $\mu, \nu \in \mathscr{M}_1(E)$, deux mesures de probabilité sur $E$. On suppose que $\mu \neq \nu$. Toutes les variables de cette question sont définies sur $(\Omega, \mathscr{F}, \mathbf{P})$. Montrer que

$$d(\mu, \nu) = \sum_{i \in E} (\mu(i) - \nu(i))_+ = \sum_{i \in E} (\nu(i) - \mu(i))_+ = 1 - \sum_{i \in E} \min(\mu(i), \nu(i)) \ .$$

Soit $U : \Omega \to \{0, 1\}$ telle que $\mathbf{P}(U = 1) = 1 - d(\mu, \nu)$. Soient $R, S, T : \Omega \to E$, trois variables telles que

$$\mathbf{P}(R = i) = \frac{\min(\mu(i), \nu(i))}{1 - d(\mu, \nu)} \ , \quad \mathbf{P}(S = i) = \frac{(\mu(i) - \nu(i))_+}{d(\mu, \nu)}$$

et

$$\mathbf{P}(T = i) = \frac{(\nu(i) - \mu(i))_+}{d(\mu, \nu)} \ .$$

On suppose de plus que $R, S, T$ et $U$ sont indépendantes. On pose

$$X = UR + (1 - U)S \quad \text{et} \quad Y = (1 - U)R + UT \ .$$

Montrer que la loi de $X$ est $\mu$, celle de $Y$ est $\nu$ et que $1 - d(\mu, \nu) = \mathbf{P}(X = Y)$.

5. Soit $(\Omega, \mathscr{F}, \mathbf{P})$ un espace de probabilité. Soient $Y_n, Z_n : \Omega \to E$, $n \geq 0$, deux suites de variables $\mathscr{F}$-mesurables. On adopte les notations et les hypothèses suivantes.
   **(a)** Pour tout $n \in \mathbb{N}$, on note $\mu_n$ la loi de $Y_n$ sous $\mathbf{P}$ :
   $$\forall i \in E\,, \quad \mathbf{P}(Y_n = i) = \mu_n(i)\ .$$
   **(b)** On suppose que les variables $Z_n$, $n \geq 0$, ont même loi sous $\mathbf{P}$. Cette loi commune est notée $\nu$ :
   $$\forall n \in \mathbb{N}\,,\ \forall i \in E\,, \quad \mathbf{P}(Z_n = i) = \nu(i)\ .$$
   **(c)** On suppose qu'il existe $T : \Omega \to \mathbb{N}$ tel que $\mathbf{P}$-p.s. $Y_n = Z_n$, pour tout $n \geq T$. Le temps aléatoire $T$ est appelé **temps de couplage**.

   Montrer que
   $$d(\mu_n, \nu) \ \leq\ \mathbf{P}(T > n) \quad \xrightarrow[n\to\infty]{} \ 0\ .$$
   C'est *un principe général de couplage*. □

**Exercice I.55** (*Claudine et la Kabbale*) Prenez un exemplaire du livre de Colette *Claudine à l'école* ; choisissez un mot parmi ceux du premier paragraphe : s'il a $n_0$ lettres, allez $n_0$ mots plus loin dans le texte ; vous tombez sur un mot de $n_1$ lettres ; allez $n_1$ mots plus loin : vous tombez sur un mot de $n_2$ lettres ... etc.

Nous affirmons que quel que soit le mot que vous avez choisi au départ, le dernier mot que vous obtenez avant de sortir du chapitre I est *toujours le même*. Faut-il considérer, pour cette raison, que l'auteur Colette a caché dans son livre un des grands mystères de l'univers ? □

**Exercice I.56** On considère un jeu de $N$ cartes numérotées de 1 à $N$ que l'on mélange successivement comme suit : on prend la carte du dessus du paquet, on l'insère uniformément au hasard parmi les $n-1$ cartes restantes (il y a donc $n$ possibilités, car on n'exclut pas de mettre la carte en bas du paquet ou de la laisser en haut du paquet). Toutes les variables sont définies sur un espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$. On modélise ce battage de cartes par une suite $(\Sigma_n)_{n\geq 0}$ de permutations aléatoires de $\{1, \ldots, N\}$ : $\Sigma_n(k)$ donne le rang à l'étape $n$ dans le paquet en partant du bas de la carte portant le numéro $k$. Pour tout $k \in \{1, \ldots, N\}$, on note $\gamma_k$ la permutation donnée par $\gamma_k(N) = k$, $\gamma_k(\ell) = \ell + 1$, pour tout $k \leq \ell < N$ et $\gamma_k(\ell) = \ell$ pour tout $1 \leq \ell < k$. On voit alors que
$$\forall 1 \leq k \leq N\,, \quad \mathbf{P}(\Sigma_{n+1} = \gamma_k \circ \Sigma_n \,|\Sigma_n) = \frac{1}{N}\ ,$$
ce qui implique que $(\Sigma_n)_{n\geq 0}$ est une chaîne de Markov.

1. Montrer que cette chaîne est irréductible apériodique et qu'elle admet la loi uniforme sur toutes les permutations comme loi stationnaire.
2. Le but du battage du jeu de cartes est, à partir d'un jeu trié au temps 0, de bien mélanger le jeu, c'est-à-dire d'atteindre la loi uniforme : la question est donc : "combien de fois faut-il battre le jeu pour qu'il soit bien mélangé ?". On note $T_N$ le temps où la carte qui est au bas du paquet au temps 0 se retrouve pour la première fois en haut du paquet : le numéro de la carte au bas du paquet au temps 0 est $\Sigma_0^{-1}(1)$, on a donc
   $$T_N = \inf\{n \geq 0\ , \Sigma_n(\Sigma_0^{-1}(1)) = N\}\ .$$
   Montrer que $\Sigma_{T_N+1}$ est de loi uniforme sur l'ensemble des permutations de $\{1, \ldots, N\}$, c'est-à-dire que pour toute permutation $\sigma$ de $\{1, \ldots, N\}$, on
   $$\mathbf{P}(\Sigma_{T_N+1} = \sigma) = \frac{1}{N!}\ .$$
3. Montrer que pour tout $n \geq 1$, $\Sigma_{T_N+n}$ est de loi uniforme sur l'ensemble des permutations de $\{1, \ldots, N\}$. Un tel temps $T_N + 1$ est appelé *Temps de mélange*.
4. En utilisant l'exercice I.2 page 7, montrer que $\lim_N T_N/(N \log N) = 1$ en probabilité. Combien de fois faut-il (à peu prés) battre (de cette façon qui n'est pas très subtile) un jeu de 52 cartes ? □

**Exercice I.57** On considère une chaîne à deux états (distincts) : $E = \{a, b\}$. On définit la matrice de transition $Q = (p(i,j))_{i,j\in E}$ en posant
$$p(a, b) = p(b, a) = 1 \quad \text{et} \quad p(a, a) = p(b, b) = 0\ .$$

1. Montrer que $Q$ est irréductible. Calculer sa période et sa probabilité invariante, notée $\pi$.
2. Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité sur lequel sont définies deux suites de variables $X_n, X'_n : \Omega \to \mathbf{E}$, $n \geq 0$, ($\mathscr{F}$-mesurables) qui remplissent les conditions suivantes.

   **(a)** Sous $\mathbf{P}$, $(X_n)_{n\geq 0}$ est une chaîne de Markov de matrice de transition $Q$ telle que $\mathbf{P}(X_0 = a) = 1$.

   **(b)** Sous $\mathbf{P}$, $(X'_n)_{n\geq 0}$ est une chaîne de Markov de matrice de transition $Q$ et de loi d'entrée $\pi$, l'unique probabilité invariante de $Q$.

   **(c)** Sous $\mathbf{P}$, les deux suites $(X_n)_{n\geq 0}$ et $(X'_n)_{n\geq 0}$ sont *indépendantes*.

   On rappelle la notation $T = \inf\{n \geq 0 \ : \ X_n = X'_n\}$ avec la convention que $T = \infty$ ssi pour tout $n \in \mathbb{N}$, on a $X_n \neq X'_n$. Montrer que $\mathbf{P}(T = \infty) = 1/2$. Mettre cela en relation avec la preuve du théorème I.4.17 de convergence vers la loi stationnaire. □

**Exercice I.58** On considère une chaîne à deux états (distincts) : $E = \{a, b\}$. On définit la matrice de transition $Q = (p(i,j))_{i,j\in E}$ en posant

$$p(a,b) = p(b,a) = p(a,a) = p(b,b) = 1/2 \ .$$

1. Montrer que $Q$ est irréductible (donc positive récurrente) apériodique et calculer sa probabilité invariante, notée $\pi$.
2. Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité sur lequel sont définies deux suites de variables $X_n, X'_n : \Omega \to E$, $n \geq 0$, ($\mathscr{F}$-mesurables) qui remplissent les conditions suivantes.

   **(a)** Sous $\mathbf{P}$, $(X_n)_{n\geq 0}$ est une chaîne de Markov de matrice de transition $Q$ telle que $\mathbf{P}(X_0 = a) = 1$.

   **(b)** Sous $\mathbf{P}$, $(X'_n)_{n\geq 0}$ est une chaîne de Markov de matrice de transition $Q$ et de loi d'entrée $\pi$, l'unique probabilité invariante de $Q$.

   **(c)** Sous $\mathbf{P}$, les deux suites $(X_n)_{n\geq 0}$ et $(X'_n)_{n\geq 0}$ sont *indépendantes*.

   On rappelle la notation $T = \inf\{n \geq 0 \ : \ X_n = X'_n\}$ avec la convention que $T = \infty$ ssi pour tout $n \in \mathbb{N}$, on a $X_n \neq X'_n$. Redémontrer que $\mathbf{P}(T < \infty) = 1$. On pose $Z = X_T = X'_T$. Montrer que $\mathbf{P}(Z = a) \geq 5/8$. En déduire que la loi de $Z$ *n'est pas la loi stationnaire*. □

**Exercice I.59** On fixe un entier $N \geq 2$ et on note $E = \{0,1\}^N$. On dit que $s = (s_1, \dots, s_N)$ et $s' = (s'_1, \dots, s'_N) \in E$ sont *voisins*, ce que l'on note $s \sim s'$ ssi il existe $1 \leq k \leq N$ tel que

$$\forall \ell \in \{1, \dots, N\} \backslash \{k\} \, , \quad s_\ell = s'_\ell \quad \text{et} \quad s_k \neq s'_k \ .$$

Un état $s \in E$ a donc $N$ voisins. On définit la matrice de transition $Q = (p(s,s'))_{s,s'\in E}$ sur $E$ en posant

$$p(s,s') = \begin{cases} 1/(2N) & \text{si } s \sim s', \\ 1/2 & \text{si } s = s', \\ 0 & \text{dans les autres cas.} \end{cases}$$

1. Montrer que $Q$ est irréductible (donc positive récurrente) et apériodique. Calculer sa loi invariante, notée $\pi$.
2. Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité sur lequel est définie une suite de variables $U_n : \Omega \to \{1, \dots, 2N\}$, $n \geq 0$, $\mathscr{F}$-mesurables, indépendantes et de loi uniforme. On fixe une loi de probabilité $\mu \in \mathscr{M}_1(E)$ et suppose en plus qu'il existe $X_0 : \Omega \to E$, $\mathscr{F}$-mesurable, de loi $\mu$, et qui est indépendante de la suite $(U_n)_{n\geq 0}$. On définit de manière récursive deux suites de variables à valeurs dans $E$ : $X_n = (X_n^1, \dots, X_n^N) \in E$ et $Y_n = (Y_n^1, \dots, Y_n^N) \in E$, $n \geq 0$ :

   **(a)** Pour tout $1 \leq k \leq N$, si $U_0 = 2k$ ou $2k - 1$, on pose $Y_0 = k$.

   **(b)** On suppose définies $X_0, Y_0, \dots, X_n, Y_n$. Pour tout $1 \leq k \leq N$, si $U_{n+1} = 2k - 1$, la $k$-ième coordonnée de $X_{n+1}$ et de $Y_{n+1}$ sont choisies égales à 1, les autres coordonées de $X_{n+1}$ (resp. de $Y_{n+1}$) sont égales à celles de $X_n$ (resp. à celles de $Y_n$). Pour tout $1 \leq k \leq N$, si $U_{n+1} = 2k$, la $k$-ième coordonnée de $X_{n+1}$ et de $Y_{n+1}$ sont choisies égales à 0, les autres coordonées de $X_{n+1}$ (resp. de $Y_{n+1}$) sont égales à celles de $X_n$ (resp. à celles de $Y_n$).

   Montrer que $(X_n)_{n\geq 0}$ est une chaîne de Markov de matrice de transition $Q$ et de loi d'entrée $\mu$ et montrer que $(Y_n)_{n\geq 0}$ est une chaîne de Markov de matrice de transition $Q$ et de loi d'entrée $\pi$.
3. On pose $T_N = \inf\{n \geq 1 \ : \ \{U_1, \dots, U_n\} = \{1, \dots, N\} \ \}$. Montrer que $X_n = Y_n$, pour tout $n \geq T_N$. En déduire que

$$\sum_{s\in E} |\mu.Q^n(s) - \pi(s)| \leq 2\mathbf{P}(T_N > n) \ .$$

4. En utilisant l'exercice I.2 page 7, montrer que $\lim_N T_N/(N \log N) = 1$ en probabilité. □

### I.4.d Vitesse de convergence pour les chaînes à espace d'états fini.

Dans cette section, nous considérons les chaînes de Markov à espace d'états fini du point de vue de l'algèbre linéaire. On ne perd pas en généralité en supposant que

$$E = \{1, \dots, N\}$$

où $N$ est un entier supérieur ou égal à 2. On note $Q = (p(i,j))_{1 \leq i,j \leq N}$ une matrice de transition : c'est une "vraie matrice" $N \times N$. Nous allons fournir des résultats asymptotiques pour $Q^n$, lorsque $n$ tend vers l'infini. On utilisera la notation $\mathbf{1}_N$ pour le vecteur colonne ne comportant que des 1 :

$$\mathbf{1}_N = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

Le premier résultat dans cette direction est le théorème de Perron-Frobenius (dans une version spécifique).

**Théorème I.4.18** **((un cas du) théorème de Perron-Frobenius)** *On suppose que $Q = (p(i,j))_{1 \leq i,j \leq N}$ est une matrice de transition* **irréductible et apériodique**. *Alors d'une part, $1$ est valeur propre de $Q$ de multiplicité algébrique $1$ (en particulier, tout vecteur propre $v$ associé à la valeur propre $1$ est colinéaire à $\mathbf{1}_N$). D'autre part, si $\lambda$ est une valeur propre de $Q$ distincte de $1$, alors $|\lambda| < 1$ nécessairement.*

**Preuve :** montrons tout d'abord que les valeurs propres de $Q$ sont de module inférieur ou égal à 1. Pour cela on rappelle que $\|v\|_\infty = \max_{1 \leq i \leq N} |v(i)|$ désigne la norme du max sur $\mathbb{C}^N$. On rappelle ensuite la propriété élémentaire des matrices de transition (I.2) : $\|Q.v\|_\infty \leq \|v\|_\infty$. Soit $v \in \mathbb{C}^N$, un vecteur propre de $Q$ associé à la valeur propre $\lambda \in \mathbb{C}$. On a donc $|\lambda| \|v\|_\infty = \|Q.v\|_\infty \leq \|v\|_\infty$. Donc $|\lambda| \leq 1$.

Le fait que $Q$ soit une matrice de transition implique immédiatement que $Q.\mathbf{1}_N = \mathbf{1}_N$, ce qui signifie que 1 est valeur propre de $Q$ associée au vecteur propre $\mathbf{1}_N$.

Soit $v \in \mathbb{C}^N$, un vecteur propre de $Q$ associé à la valeur propre $\lambda \in \mathbb{C}$. Nous montrons ensuite que si $|\lambda| = 1$ alors d'une part $\lambda = 1$ et d'autre part $v$ est colinéaire à $\mathbf{1}_N$. En effet supposons qu'il existe $x \in [0, 1[$ tel que $e^{2i\pi x}$ soit également valeur propre de $Q$. Il existe donc un vecteur propre $v$ qui lui est associé ($v$ n'est pas le vecteur nul, par définition des vecteurs propres). Pour tout $n \in \mathbb{N}$, on donc $Q^n.v = e^{2in\pi x} v$. Puisque $Q$ est supposée irréductible apériodique et puisque l'espace d'état est fini, le théorème I.2.16 implique qu'il existe $n_0 \in \mathbb{N}$ tel que

$$\forall n \geq n_0 \, , \; \forall 1 \leq i,j \leq N \, , \quad [Q^n](i,j) > 0 \, . \tag{I.80}$$

Soit $i_0 \in \{1, \dots N\}$ tel que $|v(i_0)| = \max_{1 \leq j \leq N} |v(j)|$. Puisque $v$, est non-nul, on a $|v(i_0)| > 0$. Pour tout $n \geq n_0$, on a

$$\sum_{j=1}^{N} [Q^n](i_0, j) e^{-2ni\pi x} \frac{v(j)}{v(i_0)} = 1 = \sum_{j=1}^{N} [Q^n](i_0, j) \; ,$$

ce qui, en considérant les parties réelles , implique que

$$\sum_{j=1}^{N} [Q^n](i_0, j) \; \mathrm{Re}\big(1 - e^{-2in\pi x} \tfrac{v(j)}{v(i_0)}\big) = 0 \; .$$

On remarque ensuite que si $z$ est un complexe tel que $|z| \leq 1$, on a $\mathrm{Re}(1 - z) \geq 0$ avec égalité ssi $z = 1$ (faire un dessin dans le plan complexe pour s'en convaincre). En appliquant ce résultat à $z = e^{-2in\pi x} \frac{v(j)}{v(i_0)}$, pour tout $1 \leq j \leq N$, et en observant que $[Q^n](i_0, j) > 0$, pour tout $1 \leq j \leq N$, on en déduit que

$$\forall n \geq n_0 \, , \; \forall 1 \leq j \leq N \, , \quad v(j) = e^{2in\pi x} v(i_0) \; ,$$

ce qui implique que $x = 0$ et que $v$ est colinéaire à $\mathbf{1}_N$.

On a donc également montré que si $\lambda$ est une valeur propre de $Q$ distincte de 1, alors $|\lambda| < 1$

On note $m$ la mutiplicité algébrique de la valeur propre 1. Il reste à montrer que $m = 1$. Pour cela on observe qu'il existe un polynôme $R$ de degré $N-m$ tel que $R(1) \neq 0$ et tel que $\det(zId-Q) = (z-1)^m R(z)$. Par Cayley-Hamilton et le lemme des noyaux on a $\mathbb{C}^N = E \oplus F$ où $E = \mathrm{Ker}(Id - Q)^m$ et $F = \mathrm{Ker} R(Q)$. Ici $\dim E = m$ et $E$ est laissé stable par $Q$ ; on abuse des notations en continuant de noter $Q$ l'endomorphisme de $E$ induit par $Q$ et on pose $M = Id_E - Q$ qui est un endomorphisme nilpotent puisque $M^m = 0$. Soit $d \in \{1, \ldots, m\}$ tel que $M^{d-1} \neq 0$ et $M^d = 0$. Supposons que $d \geq 2$, alors $Q^n = \binom{n}{d-1} M^{d-1} + \sum_{0 \leq k < d-1} \binom{n}{k} M^k$. Comme $M^{d-1}$ n'est pas la matrice nulle, il existe $v \in E$ tel que $M^{d-1} v \neq 0$ et on en déduit

$$\begin{aligned} \|M^{d-1}v\|_\infty &\leq \binom{n}{d-1}^{-1} \|Q^n v\|_\infty + + \sum_{0 \leq k < d-1} \binom{n}{k} \binom{n}{d-1}^{-1} \|M^k v\|_\infty \\ &\leq \binom{n}{d-1}^{-1} \|v\|_\infty + \sum_{0 \leq k < d-1} \binom{n}{k} \binom{n}{d-1}^{-1} \|M^k v\|_\infty \xrightarrow[n \to \infty]{} 0 \end{aligned}$$

car $\lim_{n \to \infty} \binom{n}{k} \binom{n}{d-1}^{-1} = 0$, pour tout $0 \leq k < d-1$. Cela implique une contradiction. Donc $d = 1$ et donc $M$ est la matrice nulle. Autrement dit $Q$ est l'identité sur $E$ : tout vecteur de $E$ est donc un vecteur propre de $Q$ associé à la valeur propre 1. Or on a montré qu'ils étaient tous colinéaires à $\mathbf{1}_N$. Cela montre que la dimension $m$ de $E$ vaut 1. Cela termine la preuve du théorème. ■

**Notation.** Soit $Q$ une matrice de transition $N \times N$ irréductible apériodique. On note $\lambda_1 = 1$, $\lambda_2$, ... , $\lambda_r$ les différentes valeurs propres distinctes de $Q$ et on note $m_1 = 1$, $m_2$, ... , $m_r$ leurs multiplicité algébriques respectives. On choisit de les indexer de sorte qu'on ait d'une part

$$\lambda_1 = 1 > |\lambda_2| \geq |\lambda_3| \geq \ldots \geq |\lambda_r|$$

et d'autre part, si $|\lambda_2| = |\lambda_s|$ pour un certain $s \in \{3, \ldots, r\}$, alors $m_2 \geq m_s$. Cela est toujours possible d'après le théorème de Perron-Frobenius. □

On note $\mathscr{M}_N(\mathbb{C}))$ l'algèbre des matrices de taille $N \times N$ à coefficients complexes. On munit $\mathscr{M}_N(\mathbb{C})$ d'une norme de matrice notée $\|\cdot\|$, c'est-à-dire que $\|\cdot\|$ est une norme sur $\mathscr{M}_N(\mathbb{C})$, vu comme espace vectoriel complexe de dimension finie $N^2$, et $\|A.B\| \leq \|A\| \cdot \|B\|$, pour toutes matrices $A, B \in \mathscr{M}_n(\mathbb{C})$. Si $M = (m(i,j))_{1 \leq i,j \leq N}$, on peut par exemple choisir de poser $\|M\| = N \max_{1 \leq i,j \leq N} |m(i,j)|$. On observe que puisque $\mathscr{M}_n(\mathbb{C})$ est de dimension finie, $(\mathscr{M}_n(\mathbb{C}), \|\cdot\|)$ est un espace vectoriel normé complet. Par ailleurs, on utilise la notation $[M](i,j) = m(i,j)$, qui est l'entrée $(i,j)$ de la matrice $M$. On remarque que pour tout $1 \leq i,j \leq N$, la fonction $M \in \mathscr{M}_n(\mathbb{C}) \mapsto [M](i,j) \in \mathbb{C}$ est continue.

Soit $r_0$ un réel strictement positif. On note $D(0, r_0) = \{z \in \mathbb{C} : |z| < r_0\}$, le disque ouvert du plan complexe de centre 0 et de rayon $r_0$. Soit $f(z) = \sum_{n \geq 0} a_n z^n$, une série entière dont le rayon de convergence est plus grand ou égal à $r_0$. Alors pour tout $0 \leq r < r_0$, on a $\sum_{n \geq 0} |a_n| r^n < \infty$. Pour tout $M \in \mathscr{M}_n(\mathbb{C})$, et tout $n \geq 0$, on pose $f_n(M) = \sum_{0 \leq k \leq n} a_k M^k$, qui est bien définie comme matrice de $\mathscr{M}_N(\mathbb{C})$. On suppose que $\|M\| < r_0$, alors on a

$$\sup_{n,m \geq q} \|f_n(M) - f_m(M)\| \leq \sum_{k \geq q} |a_k| \cdot \|M\|^k \xrightarrow[q \to \infty]{} 0 \ ,$$

ce qui entraîne que la suite de matrices $(f_n(M))_{n \geq 0}$ est de Cauchy pour $\|\cdot\|$. Cette suite converge et on note symboliquement cette limite par

$$\|\cdot\|\text{-} \lim_n f_n(M) := f(M) = \sum_{n \geq 0} a_n M^n \ . \tag{I.81}$$

Comme pour tout $1 \leq i, j \leq N$, la fonction $M \in \mathscr{M}_n(\mathbb{C}) \mapsto [M](i,j) \in \mathbb{C}$ est continue, cela est équivalent à la convergence dans $\mathbb{C}$ des séries suivantes :

$$\forall 1 \leq i, j \leq N \, , \quad [f(M)](i,j) = \sum_{n \geq 0} a_n \, [M^n](i,j) \; . \tag{I.82}$$

Ces résultats sont utilisés spour prouver le théorème suivant.

**Théorème I.4.19** **(Convergence vers la loi stationnaire : cas fini)** *On pose $E = \{1, \ldots, N\}$ avec $N \geq 2$. Soit*

$$\big(\Omega \, ; \, \mathscr{F} \, ; \, (\mathscr{F}_n)_{n \geq 0} \, ; \, (X_n)_{n \geq 0} \, ; \, Q \, ; \, \mathbf{P}_\mu \, , \, \mu \in \mathscr{M}_1(E) \, \big) \; ,$$

*une chaîne de Markov. On suppose que $Q$ est* **irréductible et apériodique**. *On adopte la convention sur l'indexation des valeurs propres de $Q$ comme indiqué précédemment. Alors les assertions suivantes sont vérifiées.*

**(i)** *$Q$ admet une unique loi invariante $\pi$ qui est telle que $\pi(i) > 0$, pour tout $1 \leq i \leq N$.*

**(ii)** *Il existe une constante positive $C_Q$, qui ne dépend que de $Q$ telle que*

$$\forall \mu \in \mathscr{M}_1(E) \, , \; \forall j \in E \, , \; \forall n \geq 0 \, , \quad | \, \mathbf{P}_\mu(X_n = j) - \pi(j) \, | \leq \; C_Q n^{m_2 - 1} |\lambda_2|^n \; \xrightarrow[n \to \infty]{} 0. \tag{I.83}$$

**Preuve :** on montre facilement que pour tout $z \in \mathbb{C}$, on a

$$P(z) := \det(\mathrm{Id} - zQ) = (1 - z) \prod_{s=2}^{r} (1 - z\lambda_s)^{m_s} \; .$$

On voit donc qu'il existe $r_0 > 0$ tel que $P(z) \neq 0$, pour tout $z \in D(0, r_0)$. On note $R(z) = (R_{i,j}(z))_{1 \leq i,j \leq N}$ la matrice inverse de $\mathrm{Id} - zQ$, qui est bien définie pour tout $z \in D(0, r_0)$. On rappelle que pour toute matrice $A \in \mathscr{M}_N(\mathbb{C})$, qui est inversible, on a

$$A^{-1} = \frac{1}{\det A} {}^{\mathrm{tr}}\mathrm{Com}(A),$$

où $\mathrm{Com}(A)$ est la co-matrice de $A$ et ${}^{\mathrm{t}}\mathrm{Com}(A)$ est la transposée de la co-matrice : on rappelle également que $[\mathrm{Com}(A)](i,j) = (-1)^{i+j} D_{i,j}$, où $D_{i,j}$ est le déterminant de la matrice $(N-1) \times (N-1)$ obtenue en retirant de $A$ la $i$-ème ligne et la $j$-ième colonne. Par conséquent,

$$R(z) = \frac{1}{P(z)} {}^{\mathrm{tr}}\mathrm{Com}(\mathrm{Id} - zQ).$$

Cela prouve que $R_{i,j}(z)$ est une fraction rationnelle dont le degré est compris entre $-N$ et $N-1$. La décomposition des fractions rationnelles en élements simples implique que

$$R_{i,j}(z) = P_{i,j}(z) + \frac{a_{i,j}}{1 - z} + \sum_{s=2}^{r} \sum_{k=1}^{m_s} \frac{c_{i,j}^{(s)}(k)}{(1 - z\lambda_s)^k} \; ,$$

où $P_{i,j}$ est une fonction polynomiale de degré au plus $N-1$ et où les $a_{i,j}$, $c_{i,j}^{(s)}(k)$, $1 \leq k \leq m_s$, $2 \leq s \leq r$, sont des nombres complexes. On rappelle que pour tout complexe $z_0$ tel que $|z_0| < 1$ et pour tout $k \geq 1$, on a le développement

$$\frac{1}{(1 - z_0)^k} = \sum_{n \geq 0} \binom{n + k - 1}{k - 1} z_0^n \; .$$

On note $r_1$ le minimum de $r_0$ et de $1/(1+|\lambda_2|)$. Pour tout $z \in D(0, r_1)$, on a

$$R_{i,j}(z) = P_{i,j}(z) + \sum_{n\geq 0} z^n \left( a_{i,j} + \sum_{s=2}^{r} \lambda_s^n \sum_{k=1}^{m_s} \binom{n+k-1}{k-1} c_{i,j}^{(s)}(k) \right) . \tag{I.84}$$

On calcule ensuite $R(z)$ d'une façon différente. On pose $M_n(z) = \sum_{0\leq k\leq n} z^k M^k$, pour tout $z \in \mathbb{C}$ et tout $n \in \mathbb{N}$. Si $|z| < 1/\|Q\|$, alors

$$\sup_{m,n\geq q} \|M_n(z) - M_m(z)\| \leq \sum_{k\geq q} \big(|z| \cdot \|Q\|\big)^k \xrightarrow[q\to\infty]{} 0 \, .$$

De plus $M_n(z)(\mathrm{Id} - zQ) = (\mathrm{Id} - zQ)M_n(z) = \mathrm{Id} - z^{n+1}Q^{n+1}$. On voit donc que

$$\forall z \in D\big(0, \tfrac{1}{\|Q\|}\big) , \quad \|\cdot\|\text{-} \lim_n M_n(z) = \sum_{n\geq 0} z^n Q^n = R(z) \, ,$$

et (I.82) permet de dire que pour tous $1 \leq i, j \leq N$ et tout $z$ tel que $|z| < 1/\|Q\|$, on a

$$R_{i,j}(z) = \sum_{n\geq 0} z^n \, [Q^n](i,j) \, .$$

En identifiant les coefficients des série entières, (I.84) entraîne que

$$\forall n \geq N \, , \; \forall 1 \leq i, j \leq N \, , \quad [Q^n](i,j) = a_{i,j} + \sum_{s=2}^{r} \lambda_s^n \sum_{k_s=1}^{m_s} \binom{n+k-1}{k-1} c_{i,j}^{(s)}(k) \, . \tag{I.85}$$

On remarque que $\binom{n+k-1}{k-1} \leq N.n^{k-1}$, pour tout $1 \leq k \leq N$ et tout $n \geq 1$. On voit facilement qu'il existe $C_Q$ qui ne dépend que de $Q$ telle que

$$\forall n \geq 1 \, , \; \forall 1 \leq i, j \leq N \, , \quad |\, [Q^n](i,j) - a_{i,j} \,| \leq \; C_Q.n^{m_2 - 1}|\lambda_2|^n \, . \tag{I.86}$$

On a donc $\lim_{n\to\infty}[Q^n](i,j) = a_{i,j}$, ce qui implique que les $a_{i,j} \in [0,1]$. Comme $Q^n.Q = Q.Q^n = Q^{n+1}$, on obtient à la limite que

$$\forall i, j \in \{1, \ldots, N\} \, , \quad \sum_{i'\in E} p(i,i')a_{i',i} = a_{i,j} \quad \text{et} \quad \sum_{j'\in E} a_{i,j'}p(j',j) = a_{i,j}. \tag{I.87}$$

On fixe $j_0 \in \{1, \ldots, N\}$ et on pose

$$v_{j_0} = \begin{pmatrix} a_{1,j_0} \\ \vdots \\ a_{N,j_0} \end{pmatrix} .$$

La première égalité dans (I.87) implique que $Q.v_{j_0} = v_{j_0}$. Par le théorème de Perron Frobenius, il existe $\pi(j_0) \in \mathbb{C}$ tel que $v_{j_0} = \pi(j_0)\mathbf{1}_N$. Comme $v_{j_0}$ a des coordonnées à valeurs dans $[0,1]$, $\pi(j_0) \in [0,1]$. On a donc montré que $a_{i,j_0} = \pi(j_0)$, pour tout $1 \leq i \leq j_0$. On note $\pi = (\pi(j); 1 \leq j \leq N)$. La seconde égalité dans (I.87), implique que $\pi.Q = \pi$, c'est-à-dire que $\pi$ est invariante. Comme $Q^n$ est une matrice de transition on a $\sum_{1\leq j\leq N}[Q^n](i,j) = 1$ et en passant à la limite on voit que $\pi$ est une mesure de probabilité. On a donc prouvé que $Q$ admet une loi invariante.

Supposons que $\pi(j) = 0$. Comme $Q$ est irréductible apériodique sur un espace d'états fini, il existe $n$ tel que $[Q^n](i,j) > 0$, pour tous $1 \leq i, j \leq n$. Comme $\pi.Q^n = \pi$, on devrait avoir

$$\sum_{i=1}^{N} \pi(i)[Q^n](i,j) = \pi(j) = 0 \ ,$$

ce qui entraînerait $\pi(1) = \ldots = \pi(N) = 0$, or $\pi$ est de masse totale 1, d'où une contradiction. On en déduit que $\pi(j) > 0$, pour tout $1 \leq j \leq N$. Montrons l'unicité : supposons que $\pi^*$ soit une autre loi invariante pour $Q$. On a donc $\pi^*.Q^n = \pi^*$, pour tout $n \in \geq 0$, c'est-à-dire que

$$\forall n \geq 1 \ , \ \forall 1 \leq j \leq N \ , \quad \sum_{1 \leq i \leq N} \pi^*(i)[Q^n](i,j) = \pi^*(j) \ .$$

En passant à la limite on obtient alors pour tout $1 \leq j \leq N$,

$$\pi^*(j) = \sum_{1 \leq i \leq N} \pi^*(i)\pi(j) = \pi(j) \ ,$$

ce qui termine la preuve du théorème. ■

Un des buts de la théorie consiste à généraliser cette convergence vers la loi stationnaire (lorsqu'elle existe) aux chaînes de Markov à valeurs dans un espace d'état infini. Ce résultat ne se laisse pas démontrer facilement par des méthodes analytiques ou reposant uniquement sur de l'algèbre linéaire en dimension infinie. Pour la plupart des applications concrètes des chaînes de Markov, notamment à la simulation de variables aléatoires, on a seulement besoin du théorème de convergence pour les chaîne finies, que l'on vient de démontrer. Ce théorème indique que la convergence vers la loi stationnaire est très rapide (géométrique). On voit que le contrôle de la vitesse de convergence, crucial pour les applications numériques, se fait grâce à la seconde plus grande (en module) valeur propre $\lambda_2$, qu'il n'est en général pas possible de calculer facilement. De nombreuses techniques ont pour but d'estimer de façon effective $|\lambda_2|$. Le cas le plus fréquent est celui des chaînes réversibles pour lesquelles il existe une technique plus adaptée (et un peu plus probabiliste) d'estimer la vitesse de convergence.

Ces « détails » font l'objet d'un certains nombre d'exercices dont l'exercice I.61 page 103 qui traite le cas de chaînes réversibles, plus simple et qui joue un rôle central dans les applications.

## EXERCICES.

**Exercice I.60** Soit $Q = (p(i,j))_{1 \leq i,j \leq N}$ une matrice de transition **irréductible**.

1. Montrer qu'une matrice et sa transposée ont même déterminant. En déduire que $Q$ et sa transposée ont le même polynôme caractéristique. Montrer qu'il existe un vecteur ligne $\pi = (\pi(1), \ldots, \pi(N))$ non-nul tel que

$$\forall j \in \{1, \ldots, N\} \ , \quad \sum_{1 \leq i \leq N} \pi(i)p(i,j) = \pi(j) \ .$$

2. On note $I_+ = \{1 \leq i \leq N \ : \ \pi(i) > 0\}$ et $I_- = \{1 \leq i \leq N \ : \ \pi(i) < 0\}$. On raisonne par l'absurde et on suppose que $I_+ \neq \emptyset$ et $I_- \neq \emptyset$. Montrer alors que

$$\sum_{j \in I_+} \pi(j) = \sum_{i \in I_+} \pi(i)p(i, I_+) - \sum_{i \in I_-} \pi(i)p(i, I_+)$$

En déduire que pour tout $i \in I_+$, $p(i, I_+) = 1$ et pour tout $i \in I_-$, $p(i, I_+) = 0$. De même, montrer que pour tout $i \in I_-$, on a $p(i, I_-) = 1$ et pour tout $i \in I_+$, on a $p(i, I_-) = 0$. Montrer que cela contredit l'irréductibilitélité de $Q$. On en déduit que soit $I_+$ est vide, soit $I_-$ est vide.

3. Déduire de la question précédente que $Q$ admet une probabilité invariante que l'on notera $\pi$. Montrer que $\pi(i) > 0$, pour tout $1 \leq i \leq N$.
4. Montrer que si $Q$ est $d$-périodique, alors les complexes $\exp(2i\pi k/d)$, $0 \leq k \leq d-1$, sont valeurs propres de multiplicité 1 et toutes les autres valeurs propres sont de module strictement inférieures à 1 (*Indication : penser que $Q^d$ est irréductible apériodique.*). Montrer également que
$$\forall 1 \leq i,j \leq N \, , \quad \frac{1}{n}\big([Q](i,j) + [Q^2](i,j) + \ldots + [Q^n](i,j)\big) \xrightarrow[n\to\infty]{} 0 \, .$$
5. Montrer que toute matrice de transition irréductible sur un espace d'état fini admet une unique mesure de probabilité invariante.
6. Montrer qu'il existe une matrice de transition sur un espace d'états à deux éléments qui admet une infinité de probabilité invariante (cette matrice de transition n'est donc par irréductible). □

**Exercice I.61** Soit $E$ un espace d'états **fini**. On suppose que $Q = (p(i,j))_{i,j\in E}$ est une matrice de transition qui est réversible : il existe donc un mesure de **probabilité** $\pi$ telle que $\pi(i) > 0$, pour tout $i \in E$ et telle que $\pi(i)p(i,j) = \pi(j)p(j,i)$, pour tous $i,j \in E$. On introduit le produit scalaire $(\,\cdot\,,\,\cdot\,)_\pi$ :
$$\forall f,g : E \to \mathbb{R} \, , \quad (f,g)_\pi = \sum_{i\in E} f(i)g(i)\pi(i) \, .$$
C'est clairement un produit scalaire sur l'espace vectoriel de dimension finie des fonctions de $E$ sur $\mathbb{R}$.

1. Montrer que $(Q.f,g)_\pi = (f,Q.g)_\pi$, c'est à dire que $Q$ est auto-adjointe par rapport au produit scalaire $(\,\cdot\,,\,\cdot\,)_\pi$. En déduire que $Q$ est une matrice diagonalisable dans $\mathbb{R}$ et que ses valeurs propres sont réelles. Montrer de plus que les valeurs sont comprises entre $-1$ et 1. Montrer que 1 est valeur propre et qu'un vecteur propre possible est la fonction constante égale à 1. On les indexe de façon décroissante.
$$1 = \beta_1 \geq \beta_2 \geq \ldots \geq \beta_{\#E} \geq -1 \, .$$
2. Montrer que si $Q$ est irréductible alors 1 est valeur propre simple
$$1 = \beta_1 > \beta_2 \geq \ldots \geq \beta_{\#E} \geq -1 \, .$$
3. Montrer que si $Q$ est irréductible et apériodique, $-1$ n'est pas valeur propre :
$$1 = \beta_1 > \beta_2 \geq \ldots \geq \beta_{\#E} > -1$$
4. On suppose que $Q$ est irréductible et apériodique. Comme $Q$ est auto-adjointe par rapport au produit scalaire $(\,\cdot\,,\,\cdot\,)_\pi$, il existe une base orthonormée de diagonalisation $f_1 = \mathbf{1}_E, f_2, \ldots, f_{\#E} : E \to \mathbb{R}$ :
$$Q.\mathbf{1}_E = \mathbf{1}_E \, , \; Q.f_2 = \beta_2 f_2, \ldots, Q.f_{\#E} = \beta_{\#E} f_{\#E} \quad \text{et} \quad (f_k, f_\ell)_\pi = \delta_{k,\ell} \, .$$
Montrer que pour toute fonction $f : E \to \mathbb{R}$, et tout $n \geq 1$, on a
$$Q^n.f - \langle \pi, f\rangle = \sum_{2\leq k\leq \#E} \beta_k^n (f,f_k)_\pi f_k \, .$$
5. On suppose que $Q$ est irréductible et apériodique. Pour toute fonction $f : E \to \mathbb{R}$, on pose $\sigma_\pi^2(f) = (f,f)_\pi - \langle \pi, f\rangle^2$, qui est la variance de $f$ relativement à la probabilité $\pi$. Montrer que pour toute fonction $f : E \to \mathbb{R}$, et tout $n \geq 1$, on a
$$\forall i \in E \, , \quad \big| \, Q^n.f(i) - \langle \pi, f\rangle \, \big| \leq \sigma_\pi^2(f) \frac{Q^2(i,i)}{2\pi(i)} \cdot \rho^{2n-2} \, ,$$
où $\rho = \max\{|\beta_k| \, ; \, 2 \leq k \leq \#E\} < 1$. Cela donne une nouvelle preuve de la convergence vers la loi stationnaire dans le cas des matrices réversibles sur un espace d'état fini. □

**Exercice I.62** Soit $E$ un espace d'états **fini**. On suppose que $Q = (p(i,j))_{i,j\in E}$ est une matrice de transition qui est réversible : il existe donc un mesure de **probabilité** $\pi$ telle que $\pi(i) > 0$, pour tout $i \in E$ et telle que $\pi(i)p(i,j) = \pi(j)p(j,i)$, pour tous $i,j \in E$. On introduit le produit scalaire $(\,\cdot\,,\,\cdot\,)_\pi$ :
$$\forall f,g : E \to \mathbb{R} \, , \quad (f,g)_\pi = \sum_{i\in E} f(i)g(i)\pi(i) \, .$$

C'est clairement un produit scalaire sur l'espace vectoriel de dimension finie des fonctions de $E$ sur $\mathbb{R}$. Pour toute fonction $f : E \to \mathbb{R}$, on pose

$$\mathscr{E}(f,f) = (f,f)_\pi - (Q.f,f)_\pi \quad \text{et} \quad \sigma^2_\pi(f) = (f,f)_\pi - \langle \pi, f\rangle^2 \ .$$

La forme quadratique $\mathscr{E}$ est appelée la forme de Dirichlet de $Q$. On suppose que $Q$ est **irréductible apériodique**. Comme elle est auto-adjointe par rapport au produit scalaire $(\cdot\,,\,\cdot)_\pi$ , on peut indexer ses valeurs propres de la sorte :

$$1 = \beta_1 > \beta_2 \geq \ldots \geq \beta_{\#E} > -1$$

On pose

$$\lambda(Q) = \inf\left\{ \frac{\mathscr{E}(f,f)}{\sigma^2(f)} \ ; \ f : E \to \mathbb{R} \ : \ \sigma^2(f) \neq 0 \right\} .$$

La quantité $\lambda(Q)$ est appelée le trou spectral.

1. Montrer que $\mathscr{E}(f,f) \geq 0$, pour toute fonction $f : E \to \mathbb{R}$.
2. Montrer que $\lambda(Q) = 1 - \beta_2$.
3. On suppose que
$$\forall i \in E\,, \quad p(i,i) \geq \tfrac{1}{2} \ . \tag{I.88}$$
Montrer que les valeurs propres de $Q$ sont toutes positives, ce qui implique que $\beta_2 \geq |\beta_k| = \beta_k$, $3 \leq k \leq \#E$. On rappelle la notation $\rho = \max\{|\beta_k| \,;\, 2 \leq k \leq \#E\} < 1$ de l'exercice précédent I.61. On a alors $\beta_2 = \rho$ et donc $1 - \lambda(Q) = \rho$. En déduire que pour toute fonction $f : E \to \mathbb{R}$, et tout $n \geq 1$, on a
$$\forall i \in E\ , \quad \left| Q^n.f(i) - \langle \pi, f\rangle \right| \leq \sigma^2_\pi(f) \frac{Q^2(i,i)}{2\pi(i)} \cdot (1-\lambda(Q))^{2n-2} \ . \tag{I.89}$$
4. On suppose ne suppose plus (I.88) mais on suppose que
$$\forall\, f : E \to \mathbb{R}\,, \quad (Q.f,f)_\pi \geq 0 \ . \tag{I.90}$$
Montrer que les valeurs propres de $Q$ sont toutes positives et en déduire que (I.89) a bien lieu.
5. On ne suppose ni (I.88), ni (I.90). Montrer que $(Q^2.f,f) \geq 0$, pour tout $i \in E$. Montrer que
$$\forall i \in E\ , \quad \left| Q^n.f(i) - \langle \pi, f\rangle \right| \leq \sigma^2_\pi(f) \frac{Q^2(i,i)}{2\pi(i)} \cdot (1-\lambda(Q^2))^{n-1} \ .$$

□

— ——— —

*Les exercices qui suivent sont des résultats préliminaires au dernier exercice I.65*

— ——— —

**Exercice I.63** Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité sur lequel est définie une suite de variables $(\xi_n, n \geq 1)$ indépendantes dont la loi est donnée par $\mathbf{P}(\xi_n = -1) = \mathbf{P}(\xi_n = 1) = 1/2$. On pose $S_0 = 0$ et pour tout $n \geq 1$, on pose $S_n = \xi_1 + \ldots + \xi_n$. Autrement dit $(S_n)_{n\geq 0}$ est une marche aléatoire simple sur $\mathbb{Z}$ issue de 0. Pour tout $n \in \mathbb{N}$ et tout $k \in \mathbb{Z}$, on pose

$$s_n(k) = \mathbf{P}(S_n = k) \ .$$

1. Montrer que pour tout $d \in \mathbb{N}$, on
$$\sum_{k\in\mathbb{Z}:|k|\geq n} s_n(k) \leq 2\exp(-\tfrac{d^2}{2n}) \ .$$
2. Montrer que pour tout entier $\ell \in \mathbb{Z}$, il existe un unique polynôme $T_\ell$ de degré $|\ell|$, à coefficient réels tel que
$$\forall \alpha \in \mathbb{R}\,, \quad T_\ell(\cos(\alpha)) = \cos(\ell\alpha) \ .$$
$T_\ell$ est appelé le $\ell$-ième polynôme de Tchebychev.
3. Montrer que pour tout $\ell \in \mathbb{Z}$, on a
$$\max_{x\in[-1,1]} |T_\ell(x)| \leq 1 \ .$$

4. Montrer que
$$\forall n \geq 1 \,,\ \forall z \in \mathbb{C} \,, \quad z^n = \sum_{\ell \in \mathbb{Z}} s_n(\ell) T_\ell(z) \ .$$

□

**Exercice I.64** Soit $(H, [\,\cdot\,,\,\cdot\,])$ un espace de Hilbert. On note $\|\cdot\|$ la norme correspondant au produit scalaire. Soit $\Phi : H \to H$ un opérateur (c'est-à-dire une fonction linéaire) continu (donc bornée en norme) : il existe donc $C > 0$ telle que
$$\forall x \in H \,, \quad \|\Phi(x)\| \leq C\|x\| \ .$$
On pose
$$|\!|\!|\Phi|\!|\!| = \sup \Big\{ \frac{[\Phi(x), x]}{\|x\|} \ ; \ x \in H \,,\ x \neq 0 \Big\} \ .$$

1. Montrer que $|\!|\!| \cdot |\!|\!|$ est bien définie.
2. Soient $\Phi, \Psi : H \to H$, deux opérateurs continus. Montrer que pour tous réels $a, b \in \mathbb{R}$, on a
$$|\!|\!|a\Phi + b\Psi|\!|\!| \leq |a| \cdot |\!|\!|\Phi|\!|\!| + |b| \cdot |\!|\!|\Psi|\!|\!| \quad \text{et} \quad |\!|\!|\Phi \circ \Psi|\!|\!| \leq |\!|\!|\Phi|\!|\!| \cdot |\!|\!|\Psi|\!|\!| \ .$$
Autrement dit $|\!|\!| \cdot |\!|\!|$ est une norme d'algèbre.
3. Soit $R(X) = \sum_{0 \leq q \leq d} a_q X^q$, un polynôme à coefficient réels : $a_q \in \mathbb{R}$, $0 \leq q \leq d$. On note $R(\Phi)$ l'opérateur borné $\sum_{0 \leq q \leq d} a_q \Phi^q$, où $\Phi$ est le $q$-ième itéré de l'opérateur continu $\Phi : H \to H$, avec la convention que $\Phi^0$ est l'identité sur $H$. Montrer que
$$|\!|\!|R(\Phi)|\!|\!| \leq \max_{x \in [\,-|\!|\!|\Phi|\!|\!|\,,\,|\!|\!|\Phi|\!|\!|\,]} |\,R(x)\,| \ .$$

□

**Exercice I.65** (*Estimée de Carne-Varopoulos*) *Il est nécessaire d'avoir traité les exercices précédents I.63 et I.64 ainsi que l'exercice I.29 page 55.* Soit $G = (\mathbf{S}, \mathbf{A})$ un graphe dénombrable (c'est-à-dire que $\mathbf{S}$ est dénombrable) simple, non-orienté et connexe. On rappelle qu'un chemin joignant les sommets $s$ et $s'$ est une suite finie $\gamma = (s_0 = s, s_1, \ldots, s_{n-1}, s_n = s')$ telle que $\{s, s_0\}, \{s_1, s_2\}, \ldots, \{s_{n-1}, s'\} \in \mathbf{A}$. La longueur du chemin $\gamma$ est $n$ et elle notée $|\gamma|$. On définit la *distance du graphe* $\mathbf{G}$ qui est une une fonction $d : \mathbf{S} \times \mathbf{S} \to \mathbb{R}_+$ donnée par
$$\forall s, s' \in \mathbf{S} \,, \quad d(s, s') = \min\{|\gamma| \ ; \ \gamma : \text{chemin joignant } s \text{ à } s'.\} \quad \text{si } s \neq s' \quad \text{et} \quad d(s, s) = 0 \ .$$
On munit $\mathbf{G}$ d'un système de poids $\mathbf{C} = (C_a \ ; \ a \in \mathbf{A})$. On rappelle la notation
$$\forall s \in \mathbf{S} \,,\ \pi(s) = \sum_{s' \in \mathbf{S} : s' \sim s} C_{\{s,s'\}} \ .$$
On fait l'hypothèse habituelle
$$\forall s \in \mathbf{S} \,, \quad 0 < \pi(s) < \infty \ .$$
On note $Q = (p(s, s'))_{s,s' \in \mathbf{S}}$, la matrice de transition de la mache aléatoire sur le graphe pondéré $(\mathbf{G}, \mathbf{C})$ :
$$\forall s, s' \in \mathbf{S} \,, \quad p(s, s') = \frac{C_{\{s,s'\}}}{\pi(s)} \ \text{ si } s \sim s' \quad \text{et} \quad p(s, s') = 0 \ \text{ sinon.}$$
On rappelle qu'on a vu à l'exemple I.2.30 page 48, que $Q$ est réversible et que
$$\forall s, s' \in \mathbf{S} \,, \quad \pi(s)p(s, s') = \pi(s')p(s', s) \ .$$
$\pi$ est donc $Q$-invariante et $Q$ est égale à sa matrice duale. On équipe $\mathbf{S}$ de la tribu de tous ses sous-ensembles, tribu notée $\mathscr{P}(\mathbf{S}) = \{B \subset \mathbf{S}\}$. On introduit l'espace $\mathrm{L}^2(\mathbf{S}, \mathscr{P}(\mathbf{S})\, \pi)$ des fonctions de $\mathbf{S}$ dont le carré est intégrable contre $\pi$ :
$$\mathrm{L}^2(\mathbf{S}, \mathscr{P}(\mathbf{S})\, \pi) = \Big\{ f : \mathbf{S} \to \mathbb{R} \ : \ \sum_{s \in \mathbf{S}} f^2(s)\pi(s) < \infty \Big\} \,,$$
On équipe $\mathrm{L}^2(\mathbf{S}, \mathscr{P}(\mathbf{S})\, \pi)$ du produit scalaire $(\,\cdot\,,\,\cdot\,)_\pi$ :
$$\forall f, g \in \mathrm{L}^2(\mathbf{S}, \mathscr{P}(\mathbf{S})\, \pi) \ , \quad (f, g)_\pi = \sum_{s \in \mathbf{S}} f(s)g(s)\pi(s) \ ,$$

si bien que $\left(\mathrm{L}^2(\mathbf{S},\mathscr{P}(\mathbf{S})\,\pi)\,;\,(\,\cdot\,,\,\cdot\,)_\pi\right)$ est un espace de Hilbert et on note $\|\cdot\|_\pi$ la norme correspondante. On a montré à l'exercice I.29 que $Q:\mathrm{L}^2(\mathbf{S},\mathscr{P}(\mathbf{S})\,\pi)\to\mathrm{L}^2(\mathbf{S},\mathscr{P}(\mathbf{S})\,\pi)$ et que

$$\forall f\in\mathrm{L}^2(\mathbf{S},\mathscr{P}(\mathbf{S})\,\pi)\ ,\quad (Q.f,f)_\pi\le\|f\|_\pi\ .$$

On définit alors comme à l'exercice I.64 la norme

$$\|\!|Q|\!\|=\sup\Big\{\,\frac{(Q.f,f)_\pi}{\|f\|_\pi}\ ;\ f\in\mathrm{L}^2(\mathbf{S},\mathscr{P}(\mathbf{S})\,\pi)\,,\ f\neq 0\Big\}\ .$$

On a donc $\|\!|Q^n|\!\|\le\|\!|Q|\!\|^n$. Le but de l'exercice est de montrer l'estimée de Carne-Varopoulos qui s'énonce comme suit :

$$\forall\,n\ge 1\ ,\ \forall s,s'\in\mathbf{S}\,,\quad [Q^n](s,s')\le 2\sqrt{\tfrac{\pi(s')}{\pi(s)}}\cdot\|\!|Q|\!\|^n\cdot\exp\left(-\tfrac{d(s,s')^2}{2n}\right)\ . \tag{I.91}$$

1. On rappelle que pour tout $\ell\in\mathbb{Z}$, $T_\ell$ est le $\ell$-ième polynôme de Tchebychev. Montrer que
$$Q^n=\|\!|Q|\!\|^n\sum_{\ell\in\mathbb{Z}}s_n(\ell)T_\ell\left(\tfrac{1}{\|\!|Q|\!\|}Q\right)\ .$$
Observer que cette série est finie car $s_n(\ell)=0$, pour tout $\ell>n$.
2. Pour tout $s\in\mathbf{S}$, on pose $g_s(s'')=(\pi(s))^{-1/2}\mathbf{1}_{\{s\}}(s'')$. Montrer que
$$[Q^n](s,s')=\sqrt{\pi(s')/\pi(s)}\cdot(Q^n.g_s,g_{s'})_\pi$$
Remarquer que $[Q^m](s,s')=0$, dès que $m<d(s,s')$. En déduire que
$$\left(T_\ell(\tfrac{1}{\|\!|Q|\!\|}Q).g_s,g_{s'}\right)_\pi=0\ ,\quad |\ell|<d(s,s').$$
3. En utilisant l'exercice I.64, montrer que pour tout $\ell\in\mathbb{Z}$ tel que $|\ell|\ge d(s,s')$, on a
$$\left|\left(T_\ell(\tfrac{1}{\|\!|Q|\!\|}Q).g_s,g_{s'}\right)_\pi\right|\le\|\!|T_\ell(\tfrac{1}{\|\!|Q|\!\|}Q)|\!\|\cdot\|g_s\|_\pi\cdot\|g_{s'}\|_\pi\le 1.$$
4. Conclure. □
5. Montrer que l'inégalité I.91 est optimale.
6. Que donne cette inégalité dans le cas des marches aléatoires simples sur $\mathbb{Z}^d$ ? □

# I.5 Chaînes de Markov et simulation de variables.

## I.5.a Problème général, mesures de Gibbs.

Une des principales applications des chaînes de Markov concerne la simulation de variables aléatoires. Le problème général s'énonce comme suit.

> *Soit un ensemble fini $E$, en général très gros. Soit $\pi$, une mesure de probabilité, qui n'est pas nécessairement connue de façon très explicite. On veut simuler (souvent à partir de variables indépendantes et uniformes sur $[0,1]$) une variable aléatoire $X$ ayant $\pi$ comme loi.*

Une solution naïve consiste à indexer les points de $E$ en une suite et à construire une fonction d'échantillonnage $\phi_\pi:[0,1[\to E$ comme expliqué à la section I.1.c, dans la preuve du lemme I.1.24 et il suffit alors de poser $\phi_\pi(U)=X$, où $U$ est une variable uniforme sur $[0,1]$. Dans la pratique cette approche ne fonctionne que lorsque $E$ est petit; si $E$ est très gros, l'ordinateur ne pourra pas stocker les valeurs de la fonctions $\phi_\pi$. De plus, il se peut que $\pi$ ne soit pas connue explicitement, c'est-à-dire numériquement, auquel cas, cette approche est complétement inopérante. Donnons quelques exemples de telles situations.

**Exemple I.5.1** (*Arbres couvrants)* Donnons d'abord définitions (certaines ont déjà été introduites).

**Définition I.5.2** Soit $\mathbf{G} = (\mathbf{S}, \mathbf{A})$ un graphe *simple* (les arêtes sont non-orientées, pas d'arêtes multiples, pas de boucle).

$(a)$ Un *chemin dans* $\mathbf{G}$ est une suite (finie ou infinie) $w = (s_n)$ de sommets qui sont successivement adjacents $s_n \sim s_{n+1}$, c'est-à-dire $\{s_n, s_{n+1}\} \in \mathbf{A}$, pour tout $n$.

$(b)$ Un chemin $w = (s_n)$ est dit *auto-évitant* si les sommets $s_n$ qui le composent sont distincts : $s_n \neq s_m$ dès que $n \neq m$.

$(c)$ Un chemin $\gamma = (s_0, s_1, \ldots, s_{n-1}, s_n)$ est un *cycle* si sa longueur $n$ est supérieures à 3, si $s_0 = s_n$ et si $(s_0, \ldots, s_{n-1})$ est auto-évitant.

$(d)$ Un graphe est dit *connexe* si toute paire de sommets distincts $s, s'$ est reliée au moins par un chemin $s = s_0 \sim s_1 \sim \ldots \sim s_n = s'$.

$(e)$ Un graphe simple sans cycle qui est connexe est appelé un *arbre*.

$(f)$ Un graphe $\mathbf{G}' = (\mathbf{S}', \mathbf{A}')$ est un *sous-graphe de* $\mathbf{G}$ si $\mathbf{S}' \subset \mathbf{S}$ et $\mathbf{A}' \subset \mathbf{A}$.

$(g)$ Un *arbre couvrant* de $\mathbf{G}$ est un sous-graphe de $\mathbf{G}$ qui est un arbre et dont l'ensemble de sommets est celui de $\mathbf{G}$.

<u>Convention et notations</u>. Comme les sommets d'un arbre couvrant $T$ de $\mathbf{G}$ sont ceux de $\mathbf{G}$, cet arbre couvrant est simplement la donnée de ses arêtes. Il est commode de confondre $T$ avec l'ensemble de ses arêtes : c'est-à-dire que le graphe $(\mathbf{S}, T)$ *est* l'arbre couvrant $T$. On note $\mathbb{T}(\mathbf{G})$ l'ensemble des arbres couvrants de $\mathbf{G}$. □

Donnons un premier procédé pour obtenir un arbre couvrant. Supposons que $\mathbf{G}$ soit connexe. Il existe alors un chemin du graphe $\mathbf{x} = (x(0), x(1), \ldots, x(n))$ passant au moins une fois par tous les sommets : c'est-à-dire que $\{x(0), \ldots, x(n)\} = \mathbf{S}$. Pour tout sommet $s \in \mathbf{S}$ on peut alors définir le *dernier instant de passage en* $s$ le long du chemin $\mathbf{x}$ par

$$\ell_s = \max\left\{k \in \{0, \ldots, n\} : x(k) = s\right\} .$$

On vérifie alors que le sous graphe dont les sommet sont $\mathbf{S}$ et les arêtes sont données par

$$T = \left\{\{s, x(\ell_s + 1)\}; s \in \mathbf{S}\backslash\{x(n)\}\right\} \tag{I.92}$$

est un arbre couvrant de $\mathbf{G}$. On appelle $T$ *l'arbre de dernier passage le long du chemin* $\mathbf{x}$. La preuve de ce résultat est laissé en exercice.

De même, on peut définir également *l'arbre des premiers passages le long du chemin* $\mathbf{x}$ par

$$T' = \left\{\{s, x(\sigma_s - 1)\}; s \in \mathbf{S}\backslash\{x(0)\}\right\} \tag{I.93}$$

où $\sigma(s)$ est le premier temps de passage en $x$ le long du chemin $\mathbf{x}$ :

$$\sigma_s = \min\left\{k \in \{0, \ldots, n\} : x(k) = s\right\} .$$

On déduit de ce qui précède que $T'$ est également un arbre couvrant de $\mathbf{G}$. En effet, notons $\widehat{\mathbf{x}} = (x(n), \ldots, x(0))$, le chemin $\mathbf{x}$ parcouru dans le sens inverse. Alors $T'$ est l'arbre de dernier passage de $\widehat{\mathbf{x}}$. Autrement dit, *l'arbre de premier passage le long d'un chemin est l'arbre de dernier passage le long du chemin inverse.*

Ces construction précédentes montrent par ailleurs qu'*un graphe admet un arbre couvrant dès qu'il est connexe* (il y a bien sûr une façon plus simple de montrer ce résultat).

En informatique, il est important d'être capable de tirer uniformément au hasard un arbre couvrant d'un graphe connexe $\mathbf{G}$, c'est-à-dire de simuler un arbre couvrant de loi $\pi$ donnée par

$$\forall T \in \mathbb{T}(\mathbf{G}) , \quad \pi(T) = \frac{1}{\#\mathbb{T}(\mathbf{G})} .$$

Or si $\mathbf{G}$ est n'est même que modérément grand, l'ensemble $\mathbb{T}(\mathbf{G})$ peut être énorme. Par exemple si $\mathbf{G}_n$ est une grille $n \times n$ de $\mathbb{Z}^2$, le graphe compte $n^2$ sommets mais on peut montrer que

$$\lim_{n\to\infty} \frac{1}{n^2} \log \#\mathbb{T}(\mathbf{G}_n) = \frac{4}{\pi} \sum_{k\geq 0} \frac{(-1)^k}{(2k+1)^2} = 1,66...$$

En général, $\mathbb{T}(\mathbf{G})$ n'est pas élémentaire à énumérer. □

**Exemple I.5.3** (*$q$-coloriages de graphes*) Soit $\mathbf{G} = (\mathbf{S}, \mathbf{A})$ un graphe simple fini et connexe. On dispose de $q$ couleurs distinctes, numérotées de 1 à $q$. Un *coloriage* du graphe est un coloriage de chaque sommet avec l'une des $q$ couleurs qui respecte la contrainte que deux sommets adjacents ont des couleurs distinctes. Un coloriage est donc une fonction $\eta : \mathbf{S} \to \{1, \ldots, q\}$ telle que $\eta(s) \neq \eta(s')$ si $s \sim s'$.

On note $\mathrm{Col}_q(\mathbf{G})$ l'ensemble des $q$-coloriages de $\mathbf{G}$. Il est clair que si $q < \min_{s\in\mathbf{S}} \deg(s)$, alors aucun $q$-coloriage n'existe. Il est également clair que si $q \geq \#S$, un $q$-coloriage existe. Mais c'est un problème combinatoire en général difficile de montrer pour un graphe donné qu'il existe ou pas un $q$-coloriage : le fameux *théorème des quatre couleurs* affirme par exemple que si le graphe $\mathbf{G}$ est planaire (c'est-à-dire, pour faire court, s'il peut être dessiné dans un plan sans croiser ses arêtes), alors il existe toujours un $4$-coloriage.

On suppose ici que $q$ est assez grand (mais tout de même plus petit que $\#S$) pour qu'il existe au moins un $q$-coloriage. On veut chercher à générer un $q$-coloriage uniforme $\pi$ :

$$\forall \eta \in \mathrm{Col}_q(\mathbf{G})\,, \quad \pi(\eta) = \frac{1}{\#\mathrm{Col}_q(\mathbf{G})}\,.$$

On peut montrer que $\#\mathrm{Col}_q(\mathbf{G}) = P_{\mathbf{G}}(q)$ où $P_{\mathbf{G}}$ est un polynôme à coefficients entiers ne dépendant que du graphe : c'est le *polynôme chromatique* du graphe. En général $\#\mathrm{Col}_q(\mathbf{G})$ n'est pas connu précisément et ce nombre peut être énorme, si bien que simuler un $q$-coloriage uniforme ne peut pas être fait par une méthode naïve. □

**Entropie, principes de thermodynamique, mesure de Gibbs.** On imagine un système *fermé* de $N$ particules (c'est-à-dire sans interaction avec l'extérieur). On s'intéresse à un *aspect spécifique* de ce système physique que l'on décrit par un ensemble d'états $E$ : par exemples toutes les vitesses et masse de molécules d'un gaz enfermé dans une boîte hermétique ou encore les spins des atomes d'un réseau métallique, etc. Les autres phénomènes physiques de ce système ne nous sont pas connus ou ne sont pas directement accessibles (ou sont volontairement ignorés dans la description des états du système). Notre description n'est donc que *partielle* : par exemple, dans le cas d'un système de molécules de gaz dans une boîte hermétiques, nous pouvons ignorer le champ électromagnetique rayonné par ces molécules à divers moments de leurs interactions.

On suppose que les états du système correspondent à un type de phénomène physique assez bien identifié et que tout état $i \in E$ décrit dans notre système possède un niveau d'énergie $V(i)$ où $V : E \to \mathbb{R}_+$ (il est naturel de choisir ici $V$ positive). Par exemple, l'énergie cinétique des particules dont on connait la masse et la vitesse. *Pour simplifier, on suppose que chaque état $i$ a un niveau distinct des autres états* : autrement dit $E$ se confond avec les niveaux d'énergie de l'aspect étudié du système physique.

**La question que l'on se pose** est de déterminer *la fraction $\pi(i)$ de particules dans l'état $i$ lorsque le système est à l'équilibre* (on suppose qu'un tel équilibre se produit, c'est-à-dire qu'après un certain temps les fractions $\pi(i)$ de particules d'état $i$ d'énergie $V(i)$ n'évoluent plus de façon décelable). On obtient à cette question une réponse en appliquant les *deux principes de la thermodynamique*, le premier étant la *conservation de l'énergie*, le second la *maximisation de l'entropie*.

**Conséquences de la conservation de l'énergie.** On note $\mathscr{E}_{\mathtt{syst}}$ l'énergie totale du système qui est l'énergie initiale du système, divisée par le nombre $N$ de particules. Le système va évoluer jusqu'à un équilibre. Une partie de l'énergie "identifiée", c'est-à-dire celle qui correspond aux états décrits par l'ensemble $E$ va se convertir en une énergie concernant les aspects non-pris en compte dans la description et cela *de façon irréversible*. À l'équilibre, les aspects étudiés du système correpondent à une énergie $N\mathscr{E}$ et l'énergie du reste du système physique est notée $N\mathscr{E}_{\mathtt{déso}}$, pour énergie de « désordre » car l'énergie transformée est vue comme une énergie irrécupérable ou inemployable pour l'aspect physique décrit (rayonnement de chaleur ou autre), ce qui peut se voir comme un plus grand désordre du système, cette notion étant ici assez vague et en tout cas relative à l'aspect physique que l'on a choisi de considérer. Autrement dit, l'énergie à l'équilibre tend à se répandre dans toutes les configurations du vrai système étudié qui est plus complexe que celui décrit par $E$. Le *premier principe de la thermodynamique* affirme que *l'énergie est conservée*, c'est-à-dire que $N\mathscr{E}_{\mathtt{syst}} = N\mathscr{E} + N\mathscr{E}_{\mathtt{déso}}$. Comme $\mathscr{E} = \sum_{i\in E} V(i)\pi(i)$, et comme le système est fermé et conserve le nombre de particules et l'énergie, on a les deux équations de conservation suivantes que doit nécessairement satisfaire $\pi$ :

$$\sum_{i\in E} \pi(i) = 1 \quad \text{et} \quad \mathscr{E}_{\mathtt{déso}} + \sum_{i\in E} V(i)\pi(i) = \mathscr{E}_{\mathtt{syst}} \ . \tag{I.94}$$

**Conséquences du second principe de la thermodynamique.** Le second principe peut de formuler qualitativement par l'affirmation qu'à l'équilibre, l'énergie du système tend à se répandre dans *toutes* les configurations du système physique. Il faut donc avoir une idée quantitative du nombre de configurations d'un système et suivant Boltzmann, il est pratique d'adopter une échelle logarithmique. Autrement dit on va considérer le logarithme des configurations possibles du système et on va postuler l'existence d'une fonction *extensive* $H$ des systèmes, c'est-à-dire telle que si un système physique $\mathtt{Syst}$ se décompose en deux systèmes indépendants, c'est-à-dire si $\mathtt{Syst} = \mathtt{Syst}_1 \times \mathtt{Syst}_2$, alors $H(\mathtt{Syst}) = H(\mathtt{Syst}_1) + H(\mathtt{Syst}_2)$. On suppose que $H$ est positive.

Ici, on considère que le système est constitué des $N$ particules qui peuvent être dans les états $i \in E$ et du reste du système physique qui est indépendant car l'énergie y a été convertie de façon irréversible et que l'on est à l'équilibre. Donc, $H(\mathtt{Syst}) = H(\mathtt{Syst}_1) + H(\mathtt{Syst}_2)$ où $\mathtt{Syst}_1$ représente l'ensemble des configurations de $N$ particules telles que $N\pi(i)$ d'entre elles sont dans l'état $i \in E$ et $H(\mathtt{Syst}_2)$ est un quantité qui ne dépend pas des $\pi(i)$, $i \in E$. Ici $H(\mathtt{Syst}_1)$ est une fonction dépendant de $N$ et de $\pi = (\pi(i))_{i\in E}$. L'espace des configurations de particules est donné en fait par le nombre de particules dans un état $i$ c'est donc $N!/\prod_{i\in E}(N_i)!$, où pour simplifier on a posé $N_i = N\pi(i)$. Donc par la formule de Stirling et puisque $N = \sum_{i\in E} N\pi(i)$,

$$\begin{aligned} H(\mathtt{Syst}_1) &= \log N! - \sum_{i\in E} \log(N_i!) \\ &\approx N\log N - N - \tfrac{1}{2}\log(2\pi N) - \sum_{i\in E} \big(N_i \log N_i - N_i - \tfrac{1}{2}\log(2\pi N_i)\big) \\ &\approx N\log N - \sum_{i\in E} (N_i \log N_i) = -N\sum_{i\in E} \pi(i)\log \pi(i). \end{aligned}$$

On « voit » qu'il est raisonnable de prendre comme définition de l'entropie par particule d'un système ayant une fraction $\mu(i)$ de particules dans l'état $i \in E$

$$H(\mu) = -\sum_{i\in E} \mu(i)\log\mu(i)\,, \quad \mu \in \mathscr{M}_1(E)\ .$$

On peut obtenir ce résultat à partir de postulats mathématiques plus clairs, c'est-à-dire une formulation plus nette de la nature additive de l'entropie. Le second principe de la thermodynamique implique alors que *la* *mesure d'équilibre $\pi$ maximise $H$ parmi les mesures satisfaisant les contraintes (I.94)* c'est-à-dire

$$H(\pi) = \max\left\{H(\mu)\,;\ \mu \in \mathscr{M}_1(E) : \langle \mu, \mathbf{1}_E\rangle = 1 \text{ et } \langle \mu, V\rangle = \mathscr{E}_{\mathtt{syst}} - \mathscr{E}_{\mathtt{déso}}\right\} .$$

Il s'agit d'un problème d'extrema liés qui prescrit la forme générale de $\pi$. Rappelons brièvement le raisonnement : soit $t \in \mathbb{R} \to \mu_t \in \mathscr{M}_1(E)$ un arc $C^1$ tel que $\langle \mu_t, \mathbf{1}_E\rangle = 1$ et $\langle \mu_t, V\rangle = \mathscr{E}_{\mathtt{syst}} - \mathscr{E}_{\mathtt{déso}}$ ; on suppose que $\mu_0 = \pi$ ; alors le vecteur dérivé $\dot\mu_0$ est tel que $\langle \dot\mu_0, \mathbf{1}_E\rangle = 0$ et $\langle \dot\mu_0, V\rangle = 0$ et

$$0 = \frac{d}{dt}_{|t=0} H(\mu_t) = -\sum_{i\in E} \dot\mu_0(i)(1 + \log \pi(i)) = -\langle \dot\mu_0, R\rangle \ ,$$

où $R$ est le vecteur/fonction $R(i) = 1 + \log \pi(i)$, $i \in E$ ; comme les valeurs de $\dot\mu_0$ peuvent varier et engendrer l'espace vectoriel $\langle \cdot\,, \cdot\rangle$-orthogonal à $\mathbf{1}_E$ et $V$, cela implique que $R$ appartient à l'espace vectoriel engendré par $\mathbf{1}_E$ et $V$ : c'est-à-dire qu'il existe $a, b \in \mathbb{R}$ tels que $R = a\mathbf{1}_E + bV$. Donc $\pi(i) = e^{a-1}e^{bV(i)}$, $i \in E$. Autrement dit, $\pi$ est de la forme

$$\pi(i) = e^{bV(i)}/z(b) \text{ où } b \in \mathbb{R} \text{ (a priori) et où } z(b) = \sum_{i\in E} \exp(bV(i)) \ .$$

Étudions ensuite la fonction $z(b)$. Par l'inégalité de Hölder on observe tout d'abord que pour tout $\theta \in ]0, 1[$ et $b_1, b_2 \in \mathbb{R}$, on a $z(\theta b_1 + (1-\theta)b_2) \leq z(b_1)^\theta z(b_2)^{1-\theta}$. Pour éviter les trivialités, on ne considère que des systèmes ayant au moins deux niveaux d'énergie distincts, ce qui se traduit par le fait que $V$ ne soit pas constante. L'inégalité de Hölder précédente est alors stricte dès que $b_1 \neq b_2$. Cela montre que la fonction $b \in \mathbb{R} \mapsto \log z(b)$ est strictement convexe et donc que la dérivée de cette fonction est strictement croissante. Or $\dot z(b)/z(b) = \sum_{i\in E} V(i)e^{bV(i)}/z(b)$. On a par conséquent montré le fait suivant.

*Il existe un unique $b_0 \in \mathbb{R}$ tel que $\mathscr{E}_{syst} - \mathscr{E}_{déso} = \sum_{i\in E} \pi(i)V(i) = \sum_{i\in E} V(i)e^{b_0V(i)}/z(b_0)$.*

Considérons ensuite deux cas extrémaux : le premier est celui où il n'y a aucun désordre ; dans ce cas toutes les particules ont depuis le début été dans le niveau le plus stable d'énergie potentielle, c'est-à-dire le niveau minimal (on suppose qu'il existe) et donc $\min_{i\in E} V(i) = \mathscr{E}_{\mathtt{syst}}$. Cela correspond au cas-limite $b_0 \to -\infty$. Si $\mathscr{E}_{\mathtt{syst}} > \min_{i\in E} V(i)$, un trop plein d'énergie a été introduit dès le départ dans le système fermé, ce qui se traduit par du désordre et la conversion à terme d'une partie de l'énergie du système en énergie de désordre. Considérons ensuite l'autre cas extrême où le désordre est maximal c'est-à-dire tel que les niveaux d'énergie potentielle n'ont plus d'impact sur la répartition des particules selon ces niveaux d'énergie : dans ce cas $\pi$ est la mesure uniforme sur $E$ (la probabilité uniforme sur $E$ maximise clairement l'entropie) : le niveau d'énergie par particule introduit au départ est alors $\mathscr{E}_{\mathtt{syst}} = \mathscr{E}_{\mathtt{déso}} + \sum_{i\in E} V(i)/\#E$, ce qui correspond à $b_0 = 0$. La situation réelle doit se situer entre ces deux cas, c'est-à-dire que $\min_{i\in E} V(i) < \mathscr{E}_{\mathtt{syst}} - \mathscr{E}_{\mathtt{déso}} < \sum_{i\in E} V(i)/\#E$. Cela montre donc plus précisément que $b_0 \in ]-\infty, 0[$. On a donc justifié le fait suivant.

> *Soit un système* fermé *de particules dont une partie est décrite par un ensemble fini d'états $E$. L'état $i \in E$ a l'énergie $V(i)$ où $V : E \to \mathbb{R}_+$. On note $\mathscr{E}_{syst}$ l'énergie totale du système par particule ; on note $\mathscr{E}_{déso}$ l'énergie de désordre divisé par le nombre de particules, c'est-à-dire l'énergie qui est transférée irréversiblement au reste du système physique durant son évolution jusqu'à son équilibre. On note $\pi$ la probabilité sur $E$ telle que $\pi(i)$ est fraction de particules l'atat $i \in E$. Alors, il existe un unique $\beta \in ]0, \infty[$ tel que*
>
> $$\forall i \in E\,, \quad \pi(i) = \frac{1}{Z_\beta} \exp\big(-\beta V(i)\big)\,,$$

*où $Z_\beta$ est la fonction de partition associée à $V$ qui vaut nécessairement*

$$Z_\beta = \sum_{j \in E} \exp\big(-\beta V(j)\big) \ .$$

*Le nombre $\beta$ est déterminé par $V$, $\mathscr{E}_{syst}$ et $\mathscr{E}_{déso}$ par l'équation $\mathscr{E}_{syst} - \mathscr{E}_{déso} = \sum_{i \in E} V(i) e^{-\beta V(i)} / Z_\beta$. Le nombre $\beta$ est* l'inverse de la température $T$ du système *: $T = 1/\beta$.*

Observons que si, en général $V(i)$ peut être calculé pour tout $i \in E$, $Z_\beta$ est plus difficilement explicitable lorsque $E$ est gros.

**Exemple I.5.4** (*Modèle d'Ising*) Soit $\mathbf{G} = (\mathbf{S}, \mathbf{A})$, un graphe simple, fini et connexe. En chaque sommet $s \in \mathbf{S}$, on place un *spin* $\eta(s)$ qui vaut $+1$ ou $-1$. Une configuration de spin est donc la donnée de $\eta = (\eta(s))_{s \in \mathbf{S}}$. L'ensemble des configurations de spin est donc $E = \{+1, -1\}^{\mathbf{S}}$ qui est simplement l'ensemble des fonctions de $\mathbf{S}$ à valeurs dans $\{+1, -1\}$. On a donc $\# E = 2^{\# S}$, qui peut être très grand. On se donne une famille de nombres réels indexés par les arêtes $(J_a ; a \in \mathbf{A})$, appelée énergie interne, et une autre famille de nombres réels indéxés par les sommets $(B_s ; s \in \mathbf{S})$, appelée champ magnétique extérieur. On définit *l'énergie d'une configuration de spin* $\eta$ par

$$\forall \eta \in \{+1, -1\}^{\mathbf{S}} \, , \quad V(\eta) = -\tfrac{1}{2} \sum_{\{s, s'\} \in \mathbf{A}} J_{\{s,s'\}} \eta(s) \eta(s') \ + \sum_{s \in \mathbf{S}} B_s \eta(s) \ .$$

On interprête $\beta > 0$, comme l'inverse d'une température. On définit la mesure d'Ising $\pi_\beta$ sur l'espace $\{+1, -1\}^{\mathbf{S}}$ des configurations de spin par

$$\forall \eta \in \{+1, -1\}^{\mathbf{S}} \, , \quad \pi_\beta(\eta) = \tfrac{1}{Z_\beta} \exp\big(-\beta V(\eta)\big) \quad \text{où} \quad Z_\beta = \sum_{\eta' \in \{+1,-1\}^{\mathbf{S}}} \exp\big(-\beta V(\eta')\big).$$

Autrement dit, la mesure d'Ising $\pi_\beta$ est la mesure de Gibbs sur l'espace $\{+1, -1\}^{\mathbf{S}}$ des configurations de spins associée au potentiel $V$.

Le modèle d'Ising a été introduit dans la première moitié du XX-ième siècle pour étudier la magnétisation spontanée à basse température de certain métaux. Il a fait, et fait toujours, l'objet d'intenses recherches en mathématiques : la question centrale consistant à savoir quels types de configurations la mesure $\pi_\beta$ favorise et quelle est la géométrie d'une configuration de loi $\pi_\beta$.

Simplifions la situation en choisissant $J_a = 1$, pour toute arête $a \in \mathbf{A}$ et $B_s = 0$, pour tout sommet $s \in \mathbf{S}$. Pour toute configuration $\eta$, on note $N(\eta)$ le nombre d'arêtes $\{s, s'\}$ telle que $\eta(s) = \eta(s')$. On voit donc que $V(\eta) = -N(\eta) + \frac{1}{2} \# \mathbf{A}$ et on peut réécrire la mesure d'Ising par

$$\forall \eta \in \{+1, -1\}^{\mathbf{S}} \, , \quad \pi_\beta(\eta) = \frac{1}{Z'_\beta} \exp\big(\beta N(\eta)\big) \quad \text{où} \quad Z'_\beta = \sum_{\eta' \in \{+1,-1\}^{\mathbf{S}}} \exp\big(\beta N(\eta')\big).$$

On observe que si $N(\eta) < N(\eta')$, alors $\pi_\beta(\eta) < \pi_\beta(\eta')$. Autrement dit, dans ce cas simplifié la mesure d'Ising favorise les configurations de spin $\eta$ pour lesquelles $N(\eta)$ est grand. Lorsque la température tend vers 0, c'est-à-dire lorsque $\beta$ tend vers l'infini, on observe que

$$\forall \eta \in \{+1, -1\}^{\mathbf{S}} \, , \quad \lim_{\beta \to \infty} \pi_\beta(\eta) = \tfrac{1}{2} \mathbf{1}_{\{\eta = \eta_*\}} + \tfrac{1}{2} \mathbf{1}_{\{\eta = \eta^*\}} \quad \text{où} \quad \forall s \in \mathbf{S} \, , \quad \eta_*(s) = -1 \text{ et } \eta^*(s) = +1.$$

Autrement dit, la mesure se concentre avec probabilité égale sur les deux configurations où tous les spins sont touts alignés en $+1$, ou tous en $-1$. Physiquement, cela correspond à une magnétisation complète du métal. Lorsque la température tend vers l'infini, c'est-à-dire lorsque $\beta$ tend vers 0, on voit en revanche que

$$\forall \eta \in \{+1, -1\}^{\mathbf{S}}, \quad \lim_{\beta \to 0} \pi_\beta(\eta) = \pi_0(\eta) = 2^{-\#\mathbf{S}} .$$

Autrement dit sous $\pi_0$, les spins sont *indépendants et de même loi* $\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_{+1}$.

On peut montrer qu'il existe un paramètre critique $\beta_c > 0$ tel que si $\beta < \beta_c$ (les températures sont hautes) alors les spins sont assez largement indépendants sous $\pi_\beta$ et en moyenne ne créent pas de champ magnétique global observable, et si $\beta > \beta_c$ (les températures sont basses), alors sous $\pi_\beta$, les spins vont avoir une forte tendance à s'aligner et créer un champ magnétique macroscopique : ce phénomène est appelé *magnétisation spontané de Curie-Weiss* et la température critique $T_c = 1/\beta_c$ est appelée température de Curie-Weiss. Lorsque le graphe est une grille $n \times n$ de $\mathbb{Z}^2$ et lorsque $n$ tend vers l'infini on peut démontrer que $\beta_c = \frac{1}{2}\log(1+\sqrt{2}) = 0,441...$. C'est l'un des rares cas où $\beta_c$ est connu explicitement (ici elle s'appelle la constante de Onsager).

Signalons qu'en plus d'être un modèle simple rendant compte de la magnétisation spontanée, une variante du modèle d'Ising est utilisée en imagerie, pour préciser pour restaurer le contour d'images endommagées (dans ce cas, ici les spin sont des pixels blancs ou noirs). $\square$

L'idée générale pour résoudre le problème de simulation est la suivante.

*Idée/ méthode générale* Soit un ensemble fini $E$, en général très gros. Soit $\pi$, une mesure de probabilité, qui n'est pas nécessairement connue de façon très explicite (une mesure de Gibbs par exemple). On trouve une "bonne" matrice de transition $Q = (p(i,j))_{i,j \in E}$ qui est irréductible et apériodique et qui admette $\pi$ comme loi invariante. On simule ensuite une chaîne de Markov $(X_n)_{n \geq 0}$ de matrice de transition $Q$. Pour tout $n \geq 0$ suffisament grand, la loi de $X_n$ est très proche de celle de $\pi$ : on peut donc dire que $X_n$ est une variable suivant la loi $\pi$ avec une marge d'erreur faible.

Cette méthode générale soulève quelques questions.

1) Par une "bonne" matrice de transition $Q = (p(i,j))_{i,j \in E}$, il faut comprendre, une matrice dont la chaîne de Markov correspondante soit facilement simulable. Autrement dit pour chaque états $i \in E$ le nombre d'états $j$ tels que $p(i,j) > 0$ ne doit pas être trop grand et une fonction d'échantillonnage $\Phi_Q$ (ou quelque chose d'équivalent) doit pouvoir être définie. Une solution à ce problème (l'algorithme de Metropolis-Hasting) est donnée dans la section qui suit.

2) Cette méthode fournit seulement, au bout d'un certain temps $n$, une variable dont la loi approche $\pi$. Si on veut contrôler l'erreur que l'on commet, il est nécessaire d'avoir des estimations de la vitesse à laquelle la chaîne converge vers sa loi stationnaire : nous avons vu une borne générale en fonction de la seconde plus grande valeur propre de la matrice de transition mais signalons qu'obtenir un critère pratique d'estimation de la vitesse de convergence est une des difficultés de cette approche. Dans la suite, nous revenons sur ce problème en donnant un algorithme de simulation exacte dû à Propp et Wilson en 1996.

### I.5.b L'algorithme de Metropolis-Hasting.

On fixe $E$, un espace d'états fini et $\pi$ une mesure de probabilité sur $E$. Dans cette section, nous donnons une méthode générale permettant de donner une matrice de transition $Q$ irréductible, apériodique et réversible qui admet $\pi$ comme loi invariante. Cette méthode est due à Metropolis et Hasting. Notons que pour que cela soit possible, il est nécessaire que

$$\forall i \in E, \quad \pi(i) > 0 \ .$$

**Matrice de contrôle.** On suppose qu'il existe une *matrice de contrôle* $A = (a(i,j))_{i,j\in E}$ qui satisfait les hypothèses suivantes.

$(a)$ $A$ est une matrice de transition.

$(b)$ Elle est irréductible.

$(c)$ Elle a la propriété de symétrie suivante :

$$\forall i,j \in E \ , \quad a(i,j) > 0 \text{ si et seulement si } a(j,i) > 0 \ . \tag{I.95}$$

Le rôle de la matrice de contrôle est de prescrire les transitions possibles de la chaîne à simuler. Dans les situations concrètes, cette matrice s'impose naturellement. La condition de faisabilité doit impliquer qu'une fonction d'échantillonnage (ou un mécanisme équivalent) puisse être constructible, ce qui impose que d'un état donné on ne puisse pas effectuer trop de transitions. Cette matrice de contrôle doit donc, en plus des conditions $(a)$, $(b)$ et $(c)$, satisfaire la condition de faisabilité (assez vague) suivante :

$$\forall i \in E \ , \quad \#\{j \in E : a(i,j) > 0\} \quad \text{n'est pas trop gros.}$$

**La fonction de troncature.** On utilise une fonction $h : ]0,\infty[ \to ]0,1]$ telle que

$$\forall u \in (0,\infty) \ , \quad uh(1/u) = h(u) \ . \tag{I.96}$$

Par exemple, on peut prendre

$$h(u) = \min(u,1) \quad \text{ou} \quad h(u) = \frac{u}{1+u} \ .$$

**Matrices de Metropolis-Hasting.** Soit $\pi$, une mesure de probabilité sur $E$ qui satisfait $\pi(i) > 0$ pour tout $i \in E$. La matrice de Metropolis-Hasting associée $A$, $h$ et $\pi$ est donnée par

$$\forall i,j \in E \ , \quad p(i,j) = \begin{cases} 0 & \text{si } i \neq j \text{ et } a(i,j) = 0 \\ a(i,j)h\left(\frac{\pi(j)a(j,i)}{\pi(i)a(i,j)}\right) & \text{si } i \neq j \text{ et si } a(i,j) > 0 \\ 1 - \sum_{k\in E\backslash\{i\}} p(i,k) & \text{si } i = j. \end{cases} \tag{I.97}$$

**Théorème I.5.5** *Soit $E$, un espace d'état fini. Soit $\pi$, une mesure de probabilité sur $E$. Soit $A = (a(i,j))_{i,j\in E}$, une matrice de contrôle sur $E$ (c'est-à-dire une matrice de transition irréductible satisfaisant (I.95)). Soit $h$ : $(0,\infty) \to [0,1]$, une fonction de troncature (c'est-à-dire satisfaisant (I.96)). Soit $Q$, la matrice de Metropolis-Hasting associée à $A$, $h$ et $\pi$ (c'est-à-dire donnée par (I.97)). Alors $Q$ est irréductible, réversible et de loi invariante $\pi$. De plus, si $h(u) < 1$, pour tout $u \in ]0,\infty[$, alors $Q$ est apériodique.*

**Preuve :** l'irréductibilité de $Q$ se déduit de celle de $A$ facilement. De plus, la définition même de $Q$ assure que c'est une matrice de transition. Soient $i,j \in E$, tels que $i \neq j$. On pose $u = \frac{\pi(i)a(i,j)}{\pi(j)a(j,i)}$ et on remarque que

$$\pi(i)p(i,j) = \pi(i)a(i,j)h\left(\frac{\pi(j)a(j,i)}{\pi(i)a(i,j)}\right) \quad = \quad \pi(j)a(j,i)uh(1/u) = \pi(j)a(j,i)h(u)$$

$$= \quad \pi(j)a(j,i)h\left(\frac{\pi(i)a(i,j}{\pi(j)a(j,i)}\right) = \pi(j)p(j,i) \ .$$

Cela implique que $Q$ est réversible et que $\pi$ est $Q$-invariante. Si $h(u) < 1$ pour tout $u > 0$, on voit que

$$\forall i \in E\,, \quad 1 - p(i,i) = \sum_{k \in E\setminus\{i\}} p(i,k) < \sum_{k \in E\setminus\{i\}} a(i,k) \leq 1\,,$$

ce qui montre que $p(i,i) > 0$ et donc que $i$ est apériodique ainsi que $Q$. ■

On en déduit un algorithme de simulation. On suppose qu'un logiciel fournit un espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$ ainsi que deux suites $U_n, U'_n : \Omega \to [0,1]$, $n \geq 0$, de v.a. $\mathscr{F}$-mesurable indépendantes de loi uniforme sur $[0,1]$. On suppose construite (et implémentée) une fonction d'échantillonnage $\Phi_A : E \times [0,1] \to E$ de la matrice de transition $A$, c'est-à-dire que pour toute variable $U$ de loi uniforme sur $[0,1]$ sous $\mathbf{P}$ et pour tous états $i, j \in E$, on a $\mathbf{P}(\Phi_A(i,U) = j) = a(i,j)$. On suppose la matrice de Metropolis-Hasting associée à $A$, $h$ et $\pi$ apériodique.

**Algorithme de Metropolis-Hasting.**

- $\underline{n = 0}$ : on choisit $X_0 \in E$, à l'aide de $U_0$. On laisse un large choix possible pour cette étape d'initialisation de l'algorithme. En particulier on peut choisir $X_0$ égal, presque sûrement, à un certain état de référence $i_0$.
- $\underline{n \to n+1}$ : on pose on note $X'_{n+1} = \Phi_A(X_n, U'_{n+1})$.
  - $*$ **Si**
  $$U_{n+1} < h\left(\frac{\pi(X'_{n+1})a(X'_{n+1}, X_n)}{\pi(X_n)a(X_n, X'_{n+1})}\right)$$
  on pose $X_{n+1} = X'_{n+1}$.
  - $*$ **Sinon**, on pose $X_{n+1} = X_n$.

On vérifie que $(X_n)_{n\geq 0}$ est une chaîne de Markov de matrice de transition $Q$. comme dans l'énoncé du théorème précédent. Lorsque $n$ est jugé suffisamment grand, on pose $X_* = X_n$ et la loi de $X_*$ est proche de $\pi$. □

On voit que l'implémentation de cet algorithme nécessite de pouvoir calculer les quotients $\pi(i)/\pi(j)$. Cette méthode est particulièrement bien adaptée à la simulation de mesures de Gibbs comme le montre l'application suivante.

**Simulation de mesures de Gibbs sur les graphes.**

Soit $\mathbf{G} = (\mathbf{S}, \mathbf{A})$, un graphe *connexe fini*. Ici $E = \mathbf{S}$ et la construction du graphe correspond à un choix raisonnable de transitions possibles. La matrice de contrôle $A = (a(s,s'))_{s,s' \in \mathbf{S}}$ que l'on choisit est celle de la marche simple sur le graphe :

$$\forall s, s' \in \mathbf{S}\ , \quad a(s,s') = \mathbf{1}_{\{s \sim s'\}} \frac{1}{\deg(s)}\ ,$$

où on rappelle que $\deg(s)$ est le degré du sommet $s$, c'est-à-dire le nombre de sommets qui lui sont voisins. Soit $V : \mathbf{S} \to \mathbb{R}$ et $\beta > 0$. On rappelle que la mesure de Gibbs associée est donnée par

$$\forall s \in \mathbf{S}\,, \quad \pi_\beta(s) = \frac{1}{Z_\beta} \exp\big(-\beta V(s)\big) \quad \text{où} \quad Z_\beta = \sum_{s' \in \mathbf{S}} \exp\big(-\beta V(s')\big)\ .$$

On suppose qu'un logiciel fournit un espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$ ainsi que deux suites $U_n, U'_n : \Omega \to [0,1]$, $n \geq 0$, de v.a. $\mathscr{F}$-mesurable indépendantes de loi uniforme sur $[0,1]$. On suppose construite (et implémentée) une fonction d'échantillonnage $\Phi_A : \mathbf{S} \times [0,1] \to \mathbf{S}$ de la matrice de transition $A$ qui est celle de la marche simple uniforme sur $\mathbf{G}$.

**Algorithme de Metropolis-Hasting pour une mesure de Gibbs sur un graphe.**

- $\underline{n = 0}$ : on choisit $X_0 \in \mathbf{S}$, éventuellement à l'aide de $U_0$.
- $\underline{n \to n+1}$ : on pose on note $X'_{n+1} = \Phi_A(X_n, U'_{n+1})$. Autrement dit, $X'_{n+1}$ est un sommet voisin de $X_n$ choisi uniformément au hasard.

  $*$ **Si**
  $$U_{n+1} < h\left(\frac{\deg(X_n)}{\deg(X'_{n+1})} e^{\beta\left(V(X_n) - V(X'_{n+1})\right)}\right)$$
  on pose $X_{n+1} = X'_{n+1}$.

  $*$ **Sinon**, on pose $X_{n+1} = X_n$.

On vérifie que $(X_n)_{n \geq 0}$ est une chaîne de Markov de matrice de transition $Q$. Lorsque $n$ est jugé suffisamment grand, on pose $X_* = X_n$ et la loi de $X_*$ est proche de $\pi_\beta$. □

La proposition suivante donne une estimée de la seconde plus grande valeur propre en module dans le cas de la matrice qui vient d'être définie.

**Proposition I.5.6** *Soit $\mathbf{G} = (\mathbf{S}, \mathbf{A})$, un graphe simple connexe, fini et sans boucle. Soit $V : \mathbf{S} \to \mathbb{R}$ et $\beta > 0$. On note $Q_\beta$ la matrice de transition de Metropolis-Hasting subordonnée à la matrice de transition de la marche simple sur $\mathbf{G}$ et à la fonction $h(u) = \min(1, u)$, $u > 0$, dont la mesure invariante est la mesure de Gibbs de potentiel $V$ à température $1/\beta$. On note $\lambda_2(\beta)$ la seconde plus grande valeur propre en module de $Q_\beta$. On suppose que $V$ n'est pas constant et on note $d_* = \max_{s \in \mathbf{S}} \deg(s)$. Il existe alors une constante $C(V) > 0$, qui ne dépend que de $V$ telle que*

$$\forall \beta > 0\,, \quad |\lambda_2(\beta)| \;\leq\; 1 \;-\; \frac{\exp\left(-\beta C(V)\right)}{d_*^4 (\#\mathbf{S})^3}\;.$$

**Preuve :** résultat admis. ■

Bien qu'il existe des formules exprimant $C(V)$, cette constante n'est pas calculable en pratique. Ce résultat n'a donc qu'une importance théorique.

## I.5.c Le recuit simulé.

En application des résultats précédents, nous introduisons brièvement la méthode du *recuit simulé* qui permet de donner une solution stochastique approchée au problème de minimisation suivant.

> *Problème de minimisation.* Soit $E$, un espace d'état fini de très grande taille. Soit $V : E \to \mathbb{R}$, une fonction de coût. On cherche à calculer $\min_E V$ et surtout à trouver un élément de $E_{\min} = \{i \in E : V(i) = \min_E V\}$ (ou bien de trouver un état $i$ tel que $V(i)$ est proche de $\min_E V$). □

Le principe est le suivant : on fixe $\beta > 0$ et on introduit la mesure de Gibbs de potentiel $V$ et de température $1/\beta$ :

$$\forall i \in E\,, \quad \pi_\beta(i) = \tfrac{1}{Z_\beta} \exp\big(-\beta V(i)\big) \quad \text{où} \quad Z_\beta = \sum_{j \in E} \exp\big(-\beta V(j)\big)\ .$$

Lorsque la température tend vers 0, c'est-à-dire lorsque $\beta$ tend vers $\infty$, on vérifie facilement que

$$\forall i \in E\,, \quad \lim_{\beta \to \infty} \pi_\beta(i) := \pi_\infty(i) = \frac{1}{\# E_{\min}} \mathbf{1}_{\{E_{\min}\}}(i)\ ,$$

c'est-à-dire que $\pi_\beta$ converge en loi vers la loi uniforme sur $E_{\min}$ lorsque $\beta$ tend vers l'infini. L'idée consiste alors à choisir $\beta$ suffisamment grand, à simuler une variable $X_*$ de loi $\pi_\beta$ ; comme $\pi_\beta$ se concentre sur les états $i$ tels que $V(i)$ est minimal ou proche du minimum, on doit avoir une forte probabilité que $V(X_*)$ soit proche de $\min_E V$. La simulation de $X_*$ peut se faire grâce à l'algorithme de Metropolis-Hasting.

**Exemple I.5.7** (*Le problème du voyageur de commerce*) Un voyageur de commerce en partant d'une ville 0, doit visiter une fois exactement $N$ villes numérotées de 1 à $N$ et revenir à 0. Le but est d'effectuer la tournée la plus courte possible. Les distances entre les villes sont données par des nombres positifs $(d_{k,\ell}, 0 \leq k, \ell \leq N)$ tels que $d(k,\ell) = d(\ell,k)$. Une tournée du voyageur se modélise par une permutation $\sigma$ de $\{1, \ldots, N\}$. La distance totale parcourue par le voyageur de commerce lors de sa tournée donnée par la permutation $\sigma$ est donc

$$V(\sigma) = d_{0,\sigma(1)} + d_{\sigma(1),\sigma(2)} + \ldots d_{\sigma(N-1),\sigma(N)} + d_{\sigma(N),0}\ .$$

Le voyageur de commerce veut trouver un tour $\sigma$ qui minimise $V(\sigma)$, ou plus modestement de trouver $\sigma$ tel que $V(\sigma)$ soit proche de $\min V$.

Une méthode naïve consisterait à calculer $V(\sigma)$ pour les $N!$ permutations possibles de $\{1, \ldots, N\}$ et ensuite de choisir la ou les permutations minimales. Comme $28! \geq 10^{28}$, cette méthode est irréaliste dès que $N$ est même modérément grand. Observons aussi que même une méthode qui permettrait de trouver un parcours relativement économique sans être forcément optimale constituerait un outil appréciable. □

Tout le problème de la méthode du recuit simulé est de trouver une température $1/\beta$ qui soit suffisamment petite pour qu'avec une grande probabilité, pour que la simulation d'une variable $X_*$ approchant la mesure de Gibbs $\pi_\beta$ ait une forte chance de se trouver dans $E_{\min}$, ou dans un état $i$ tel que $V(i)$ soit proche de $\min V$.

> *Idée générale du recuit simulé*. Elle consiste à trouver une « bonne suite » de températures $1/\beta_n$ décroissant vers 0 et à simuler une chaîne de Markov $(X_n)_{n\geq 0}$ inhomogène : la matrice de transition du temps $n$ au temps $n+1$ étant $Q_{\beta_n}$, une matrice de type Metropolis-Hasting dont la loi invariante est $\pi_{\beta_n}$. On espère qu'asymptotiquement, la loi de la chaîne va tendre vers $\pi_\infty$ et que $(X_n)_{n\geq 0}$ va finir par être stationnaire à un état pour lequel $V$ est minimale. □

Nous donnons l'algorithme de recuit correspondant à une fonction de coût définie sur une graphe bien choisi $\mathbf{G} = (\mathbf{S}, \mathbf{A})$ qui est simple, fini, connexe et sans boucle. Ici $E = \mathbf{S}$ et la construction du graphe correspond à un choix raisonnables des transitions possibles. La matrice de contrôle $A = (a(s,s'))_{s,s' \in \mathbf{S}}$ que l'on choisit est cette de la marche simple sur le graphe :

$$\forall s, s' \in \mathbf{S}\,, \quad a(s,s') = \mathbf{1}_{\{s \sim s'\}} \frac{1}{\deg(s)}\ ,$$

où on rappelle que $\deg(s)$ est le degré du sommet $s$, c'est-à-dire le nombre de sommets qui lui sont voisins. Soit $V : \mathbf{S} \to \mathbb{R}$. Pour tout $\beta > 0$, on pose

$$\forall s \in \mathbf{S}\,, \quad \pi_\beta(s) = \frac{1}{Z_\beta} \exp\big(-\beta V(s)\big) \quad \text{où} \quad Z_\beta = \sum_{s' \in \mathbf{S}} \exp\big(-\beta V(s')\big)\,.$$

On choisit une suite croissante $(\beta_n)_{n \geq 0}$ telle que $\lim_{n\to\infty} \beta_n = \infty$, qui est un *schéma de refroidissement*.

On suppose qu'un logiciel fournit un espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$ ainsi que deux suites $U_n, U'_n : \Omega \to [0,1]$, $n \geq 0$, de v.a. $\mathscr{F}$-mesurable indépendantes de loi uniforme sur $[0,1]$. On suppose construite (et implémentée) une fonction d'échantillonnage $\Phi_A : \mathbf{S} \times [0,1] \to \mathbf{S}$ de la matrice de transition $A$ qui est celle de la marche simple uniforme sur $\mathbf{G}$.

**Algorithme de recuit simulé**

- $\underline{n = 0}$ : on choisit $X_0 \in \mathbf{S}$, éventuellement à l'aide de $U_0$.
- $\underline{n \to n+1}$ : on pose on note $X'_{n+1} = \Phi_A(X_n, U'_{n+1})$. Autrement dit $X'_{n+1}$ est un sommet voisin de $X_n$ choisi uniformément au hasard.

  $*$ **Si**
  $$U_{n+1} < h\left(\frac{\deg(X_n)}{\deg(X'_{n+1})} e^{\beta_n \left(V(X_n) - V(X'_{n+1})\right)}\right)$$

  $*$ **Sinon**, on pose $X_{n+1} = X_n$.

Si le schéma de refroidissement est bien choisi, alors (on espère que) $X_n \in E_{\min}$ pour $n$ assez grand. □

Informellement, l'algorithme de recuit évolue d'une manière qui résulte de deux phénomènes concurrents : si le schéma de refroidissement est bien choisi, pour $n$ assez grand, la loi de $X_n$ est proche de $\pi_{\beta_n}$ qui se concentre plutôt en des états pour lesquels $V$ est proche de son minimum. Mais il se peut que $V$ en $X_n$ ne réalise qu'un minimum local ($X_n$ est dans un mauvais puits de potentiel). On peut se représenter la température comme une agitation de molécules qui viennent entrechoquer la particule $X_n$ et la faire changer d'état : si cette température est trop basse, $X_n$ n'a pas l'énergie nécessaire pour sortir du puits de potentiel dont le fond n'est pas le minimum global : la chaîne va alors rester coincée dans ce mauvais puits de potentiel pendant un temps déraisonnable (pas observable dans la réalité). Si la température est suffisamment élevée, il va pouvoir quitter ce mauvais puits de potentiel et peut-être ensuite finir par tomber dans le puit de potentiel correspondant au minimum global mais le danger est qu'une agitation trop forte le lui fasse quitter trop rapidement.

Plus précisément, supposons qu'à l'étape $n$ de l'algorithme de recuit ci-dessus, la chaîne $X_n$ soit en un sommet $s \in \mathbf{S}$ tel que pour tout sommet voisin $s' \sim s$ on ait $V(s') > V(s) + c$, où $c$ est une constante positive. On suppose par ailleurs que $s$ n'est pas un minimum de $V$. On voit que $s$ est dans un mauvais puits de potentiel et on a

$$h\left(\frac{\deg(X_n)}{\deg(X'_{n+1})} e^{\beta_n \left(V(X_n) - V(X'_{n+1})\right)}\right) \leq h\left(\frac{\deg(X_n)}{\deg(X'_{n+1})} e^{-c\beta_n}\right) \xrightarrow[\beta_n \to \infty]{} 0$$

Lorsque $\beta_n$ est très grand, on voit donc que $\lim_{\beta_n \to \infty} \mathbf{P}(X_{n+1} = X_n) = 1$ et la chaîne va vraisemblablement passer un temps très (trop) long, voire, si $\beta_n$ croît trop vite, définitivement rester en $X_n = s$, qui n'est pourtant pas le minimum de $V$. L'exemple élémentaire suivant illustre ce phénomène.

**Exemple I.5.8** Soit $\mathbf{G} = (\mathbf{S}, \mathbf{A})$, un "carré" :

$$\mathbf{S} = \{1, 2, 3, 4\} \quad \text{et} \quad \mathbf{A}\big\{ \{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 1\} \big\}$$

On définit la fonction de coût suivante

$$V(3) = 0 < V(1) = 1 < V(2) = V(4) = 2 .$$

On note $(X_n)_{n \geq 0}$ la chaîne donnée par l'algorithme précédent, associée au schéma de refroidissement $(\beta_n)_{n \geq 0}$. Lors de l'étape d'initialisation, on a choisi $\mathbf{P}(X_0 = 1) = 1$. Il est facile de vérifier que

$$\forall n \geq 1 , \quad \mathbf{P}(X_0 = X_1 = \ldots = X_n = 1) = \prod_{0 \leq k < n} \left(1 - e^{-\beta_k}\right) .$$

On a donc

$$\mathbf{P}\big( \forall n \geq 0 , \; X_n = 1 \big) = \prod_{n \geq 1} \left(1 - e^{-\beta_n}\right) .$$

On voit que cette probabilité est strictement positive si et seulement si $\sum_{n \in \mathbb{N}} \exp(-\beta_n) < \infty$. Par exemple si on pose $\beta_n = c \log(n + 1)$, on voit que

$$\mathbf{P}\big( \forall n \geq 0 , \; X_n = 1 \big) = \begin{cases} > 0 & \text{si } c > 1, \\ = 0 & \text{si } 0 < c \leq 1. \end{cases}$$

Si $c > 1$, le refroidissement est trop rapide et l'algorithme de recuit a une probabilité non-nulle de ne pas choisir un minimum de $V$. Si $0 < c < 1$, il resterait encore à montrer que l'algorithme converge bien vers l'état minimum 3. On voit sur cet exemple qu'il peut être assez délicat de trouver le bon schéma de refroidissement. □

Nous énonçons une théorème qui permet de dire qu'un algorithme de convergence est possible.

**Théorème I.5.9** *Soit $\mathbf{G} = (\mathbf{S}, \mathbf{A})$ qui est simple, fini, connexe et sans boucle. Soit $V : \mathbf{S} \to \mathbb{R}$. qui n'est pas constante. On se donne une schéma de refroidissement $(\beta_n)_{n \geq 0}$, c'est-à-dire une suite positive croissante tendant vers l'infini. On note $(X_n)_{n \geq 0}$, la chaîne de Markov inhomogène donnée par l'algorithme de recuit ci-dessus.*

*Alors, il existe une constante $C(V) > 0$ qui ne dépend que de $V$ telle que*

$$\textit{si} \quad \sum_{n \geq 0} \exp\big(-\beta_n C(V)\big) = \infty \quad \textit{alors} \quad \lim_n \mathbf{P}(X_n \in E_{\min}) = 1 .$$

*On choisit $C > C(V)$. On obtient un algorithme de recuit concluant si l'on choisit par exemple l'un des deux schémas de refroidissement suivants.*

**(a)** *Pour tout $n \geq 0$, on pose $\beta_n = C \log(n + 1)$.*

**(b)** *(Refroidissement par paliers) pour tout $k \in \mathbb{N}$ et pour tout $e^{kC} \leq n < e^{(k+1)C}$, on pose $\beta_n = k$.*

**Preuve :** admis. Peut-être sous forme d'exercice. ■

Il est important de mentionner que ce résultat théorique ne fournit en général pas de borne ou d'estimation de $C(V)$ qui soit utilisable dans la pratique.

**Illustration informatiqe.** Les figures suivantes illustrent l'algorithme de recuit simulé le voyageur de commerce (exemple I.5.7 page 116) pour les conditions suivantes.

— $N = 20$ villes. Les villes sont prises comme des points aléatoires dans le carré $[0,1] \times [0,1]$ avec des coordonnées indépendantes uniformément distribuées. A savoir, $C_k = (U_k, V_k)$, $1 \leq k \leq N$, et les $U_1, V_1, U_2, V_2, \ldots, U_N, V_N$ sont des coordonnées i.i.d. uniformes sur $[0,1]$. La distance entre deux villes est la norme euclidienne :

$$\forall k, \ell \in \{1, \ldots, N\}, \quad d_{k,\ell} = \sqrt{(U_k - U_\ell)^2 + (V_k - V_\ell)^2}\,;.$$

— Le schéma de refroidissement est le suivant.

  — Pour $n \in \{1, \ldots, 10\}$, $\beta_n = 40$ ;

  — Pour $n \in \{11, \ldots, 110\}$, $\beta_n = 80$ ;

  — Pour $n \in \{111, \ldots, 110\}$, $\beta_n = 120$ ;

  — Pour $n \in \{1111, \ldots, 11110\}$, $\beta_n = 160$ ;

  — Pour $n \in \{11111, \ldots, 111110\}$, $\beta_n = 200$ ;

— La chaîne de Markov inhomogène $(X_n)_{n \in \mathbb{N}}$ est générée grâce à l'algorithme de Metropolis-Hasting suivant le schéma de refroidissement $(\beta_n)_{n \in \mathbb{N}}$ spécifié ci-dessus. L'état initial est la permutation identique $\sigma_0 = (1, \ldots, N)$.

Les graphiques de gauche représentent les graphes de $V(X_n)$. Observez que dans le dernier graphique, pour les temps $n \in \{11111, \ldots, 111110\}$ la température est encore suffisamment élevée pour faire bouger un peu la chaîne de Markov, mais suffisamment basse pour la faire revenir dans le minimum (peut-être local) de $V$.

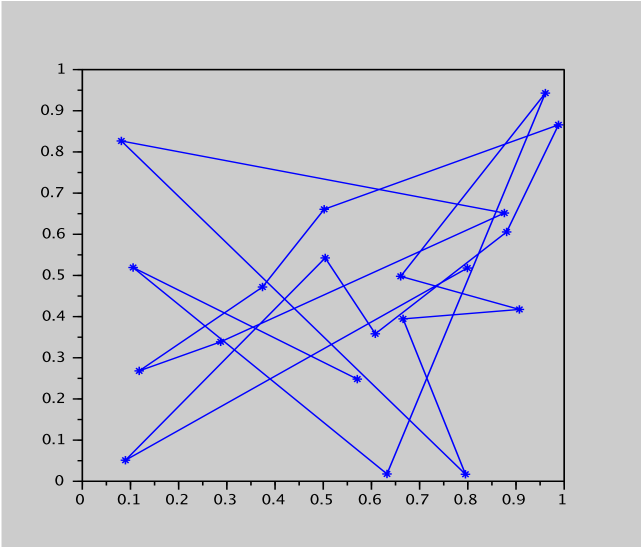


FIGURE I.8 – 20 villes de coordonnées uniformes indépendantes dans $[0, 1]^2$. Chemin initial de longueur= **9.864**.

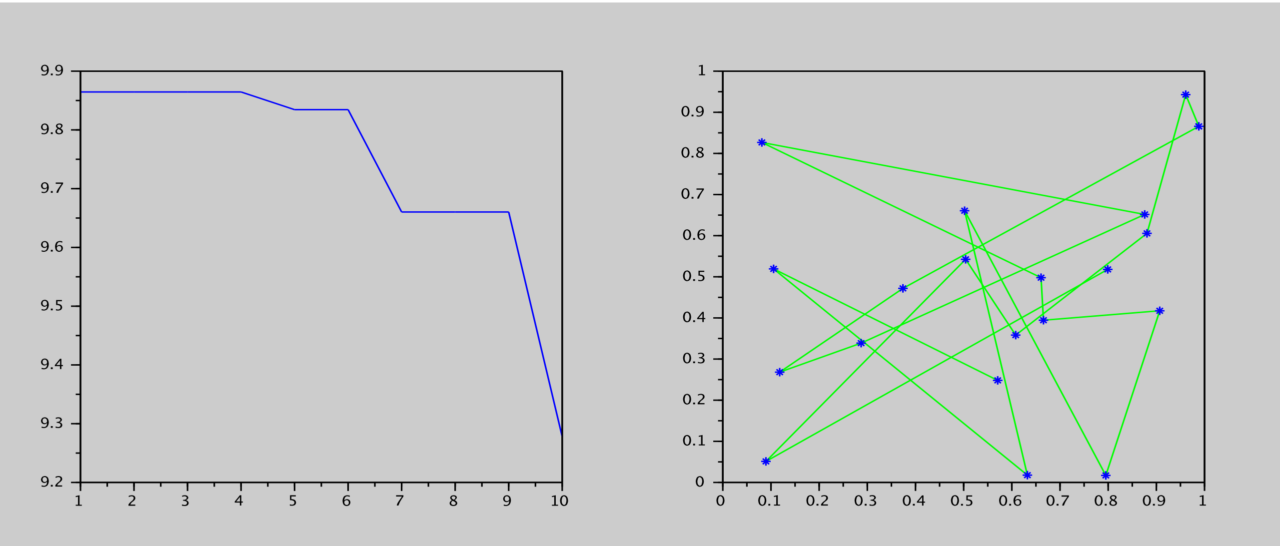


FIGURE I.9 – iterations $1$ à $10$, $\beta = 40$. Gauche : $V(X_1)$ à $V(X_1 0) =$ **9.278**. Droite : le chemin au temps 10

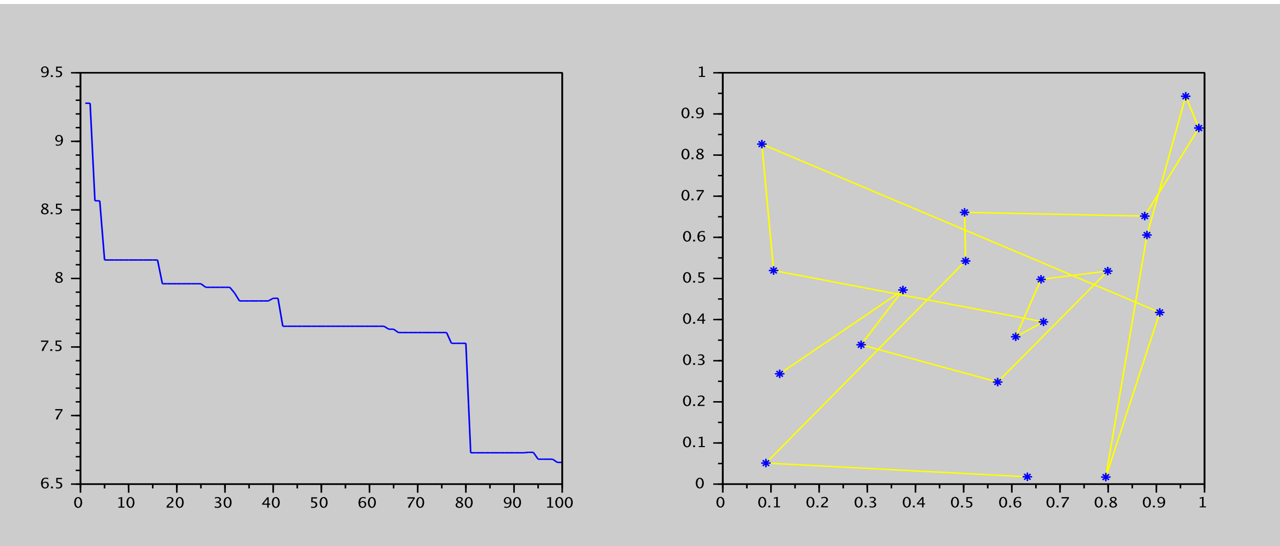


FIGURE I.10 – iterations $11$ à $110$, $\beta = 80$. Gauche : $V(X_{11})$ à $V(X_{110}) =$ **6.658**. Droite : le chemin au temps 110

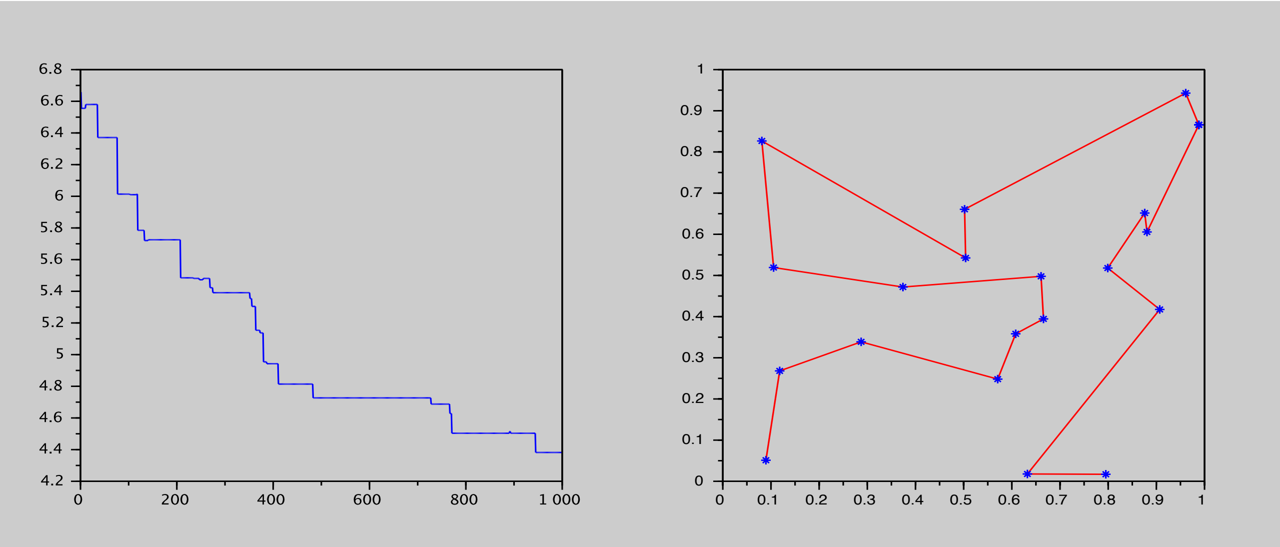


FIGURE I.11 – iterations 111 à 1110, $\beta = 120$. Gauche : $V(X_{111})$ à $V(X_{1110}) = \mathbf{4.382}$. Droite : le chemin au temps 1110

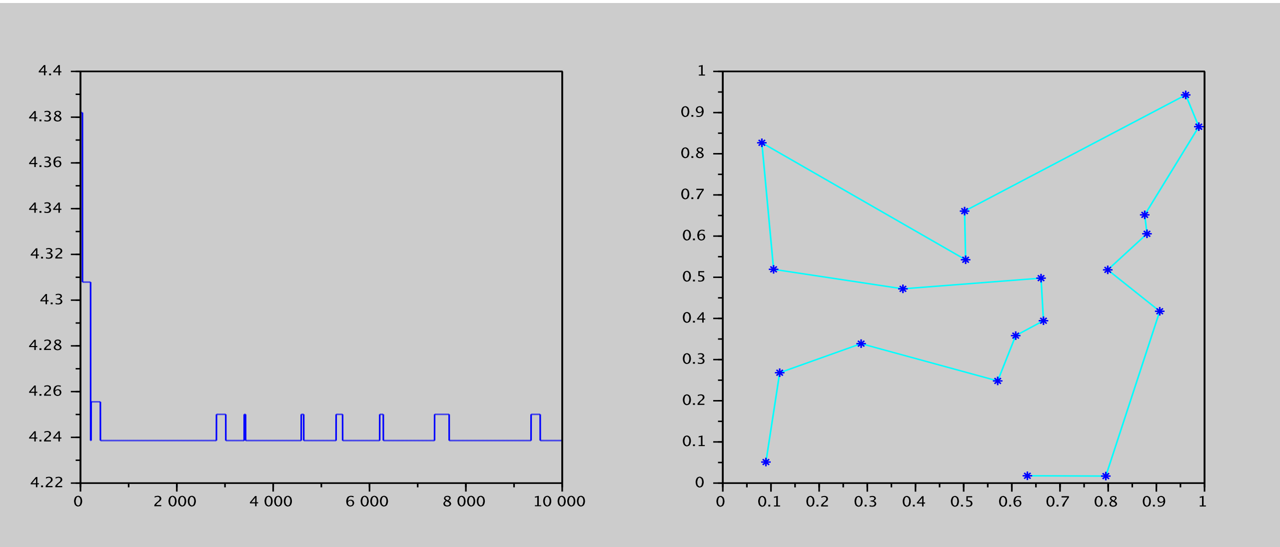


FIGURE I.12 – iterations 1111 à 11110, $\beta = 160$. Gauche : $V(X_{1111})$ à $V(X_{11110}) = \mathbf{4.238}$. Droite : le chemin au temps 11110

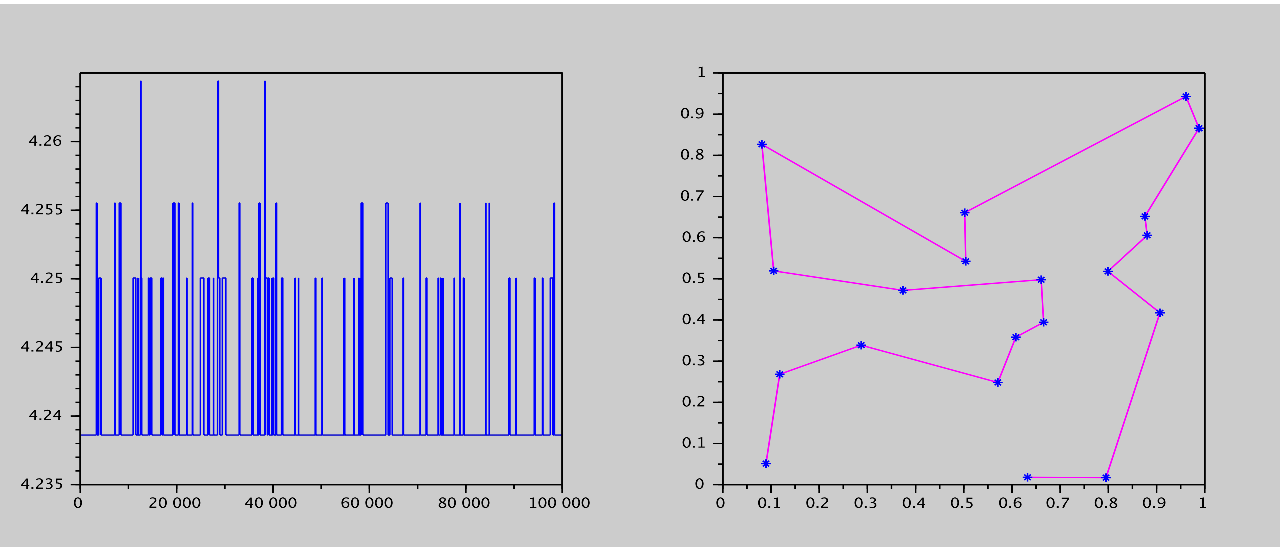


FIGURE I.13 – iterations 11111 à 111110, $\beta = 200$. Gauche : $V(X_{11111})$ à $V(X_{111110}) = \mathbf{4.238}$. Droite : le chemin au temps 111110

### I.5.d Simulation exacte : l'algorithme de Propp-Wilson.

Pour fixer les idées voici les objets que l'on se donne.

– On fixe $\pi$, une loi sur $E$, telle que $\pi(i) > 0$ pour tout $i \in E$.

– On suppose que l'on dispose d'une matrice de transition $Q = (p(i,j))_{i,j \in E}$ qui est *irréductible et apériodique* et dont la loi invariante est $\pi$ (on a vu des méthodes permettant d'obtenir une telle matrice).

Lorsque l'on applique l'algorithme de Metropolis-Hasting, l'un des problèmes importants consiste à déterminer le temps $n$ pour lequel la loi de $X_n$ est suffisamment proche de la loi invariante $\pi$ que l'on souhaite simuler. Le contrôle de la vitesse de convergence nécessite une bonne connaissance des valeurs propres de la matrice de transition ou (d'autres quantités plus probabilistes), ce qui n'est pas toujours simple. Pour contourner cette difficulté, on souhaite disposer d'un algorithme de simulation exacte, c'est-à-dire retournant une variable ayant exactement la loi désirée. Une idée ayant été efficace dans la preuve du théorème de convergence vers la loi stationnaire consiste à utiliser le *couplage* de deux chaînes.

**Couplage direct.** Expliquons briévement une version simple de l'idée de couplage et une difficulté importante dans son utilisation pour construire un algorithme de simulation exacte. On suppose que sur un espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$, on dispose de deux suites de v.a. à valeurs dans $E$, $(X_n)_{n\in\mathbb{N}}$ et $(X'_n)_{n\in\mathbb{N}}$. Elles ne sont pas nécessairement indépendantes mais :

– sous $\mathbf{P}$, $(X_n)_{n\in\mathbb{N}}$ et $(X'_n)_{n\in\mathbb{N}}$ sont deux chaînes de Markov de matrice de transition $Q$ (ce sont deux chaînes *couplées*).

Comme dans la preuve du théorème de la convergence des chaînes de Markov vers leur loi stationnaire, on peut d'une part supposer que $X'_0$ a pour loi $\pi$, stationnaire pour $Q$ ; par conséquent, pour tout $n \in \mathbb{N}$, la v.a. $X'_n$ a pour loi $\pi$ également. On peut ensuite introduire le *temps de couplage* :

$$T = \inf \left\{ n \in \mathbb{N} : X_n = X'_n \right\}$$

avec la convention a priori que $\inf \emptyset = \infty$. On fait l'hypothèse que $\mathbf{P}(T < \infty) = 1$ (rappelons que montrer cette condition occupe l'essentiel de la preuve du théorème de convergence vers la loi stationnaire). Il est ensuite naturel de penser que $Y = X_T = X'_T$ a pour loi $\pi$ *mais ce n'est en général pas le cas*. Une façon simple de voir cela est de considérer une situation où la chaîne possède une *zone d'étranglement*, c'est-à-dire

$$\exists i_0, j_0 \in E : \{i_0\} = \{i \in E : p(i, j_0) > 0\} \ . \tag{I.98}$$

(C'est par exemple le cas d'un processus de naissance et de mort irréductible apériodique dont l'espace d'état est $\{0, \dots, N\}$ : les états 0 et $N$ sont des zones d'étranglement dès que $p(0,0) = p(N,N) = 0$.) On observe que puisque par définition on a $X_{T-1} \neq X'_{T-1}$ et $X_T = X'_T$, alors nécessairement $Y \neq j_0$, c'est-à-dire que $\mathbf{P}(Y = j_0) = 0$. Or $\pi(j_0) > 0$ (c'est toujours le cas de loi stationnaire d'une chaîne irréductible à espace d'états finis). Donc $Y$ ne peut pas avoir $\pi$ pour loi.

Cet exemple peut sembler caricatural et on peut penser qu'il ne constitue pas un problème sérieux (on pourrait par exemple modifier les transitions de $j_0$ vers les autres états de façon à ce que $p(j_0, j_0) > 0$ et éliminer ainsi tous les étranglements de type (I.98)). Il met cependant en lumière le phénomène général suivant : *deux chaînes couplées vont avoir tendendance à se croiser dans une zone de l'espace d'état $E$ qui favorise la coalescence : cette contrainte géométrique (locale) des transitions introduit un* biais *fort dans la loi de $Y$, biais qui est sans doute plus fort que ce que « ressent » la mesure invariante $\pi$ de ces contraintes géométriques qu'elle moyenne sur des temps longs.* Baser un algorithme sur une idée de couplage simple à partir de la variable $Y$ semble donc difficile. Une idée naturelle consiste à continuer l'évolution de la chaîne après le temps de couplage, c'est-à-dire à considérer la variable $X_{T+n}$ pour un $n$ suffisamment grand afin de diluer le biais. Cette idée a été poursuivie sans succès pratique jusqu'à ce que Propp et Wilson en 1996 modifient l'approche par couplage de la manière expliquée dans le paragraphe suivant.

**Couplage rétrograde.** Une méthode de simulation exacte et implémentable dans la pratique a été proposée par Propp et Wilson en 1996. Il existe depuis d'autres algorithmes (Fill 1998) mais nous en tiendrons à une brève introduction de la méthode de Propp et Wilson. Comme mentionné dans le paragraphe précédent, l'idée repose sur un couplage similaire à celui utilisé dans la preuve du théorème de la convergence des chaînes vers leur loi stationnaire mais au lieu d'établir ce couplage en progressant dans le temps, on le fait de façon rétrograde (le nom en anglais de cet algorithme est *the coupling from the past*).

En plus de $Q$ et $\pi$, dans ce qui suit, on suppose qu'il est possible de construire une fonction d'échantillonnage de $Q$ qui soit facilement implémentable. Mentionnons ici *qu'une bonne partie de l'efficacité de l'algorithme de Propp-Wilson provient d'un choix judicieux de la fonction d'échantillonnage* et dans la pratique il est parfois intéressant d'en avoir plusieurs qui varient au cours de la simulation. Donc, bien que cela ne change rien sur le plan théorique, on suppose plus précisément que l'on dispose (c'est-à-dire qu'il est possible d'implémenter) d'*une suite* $(\Phi_Q^{(-n)})_{n\geq 1}$ de fonctions d'échantillonnage de la matrice de transition $Q$, c'est-à-dire que pour tout entier $n \geq 1$, $\Phi_Q^{(-n)} : [0,1] \times E \to E$ est telle que

$$\mathbf{P}\big(\Phi_Q^{(-n)}(U,i) = j\big) = p(i,j)$$

où $U$ est une variable uniforme définie sur un espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$. Ici l'indexation par des entiers négatifs se justifie par la nature rétrograde du couplage expliqué ci-desouss

Construction du couplage rétrograde. On suppose que sur $(\Omega, \mathscr{F}, \mathbf{P})$, il existe une suite $U_{-n} : \Omega \to [0,1]$, $n \geq 1$, de variables indépendantes et de loi uniforme (l'indexation par des entiers négatifs reflète la nature rétrograde du couplage). Pour tout $n \in \mathbb{N}^*$, et pour tout état $i \in E$, on définit une suite finie de variables $(X_k^{(n)}[\,i\,])_{0\leq k\leq n}$ de la manière récursive suivante.

- Pour tout $i \in E$, on pose $X_0^{(1)}[\,i\,] = i$ et $X_1^{(1)}[\,i\,] = \Phi_Q^{(-1)}(U_{-1}, i)$.
- Supposons que pour tout $i \in E$, la suite $(X_k^{(n)}[\,i\,])_{0\leq k\leq n}$ soit définie. On pose alors

$$X_0^{(n+1)}[\,i\,] = i\ ,\ X_1^{(n+1)}[\,i\,] = \Phi_Q^{(-n-1)}(U_{-n-1}, i) \ \text{ et } \quad \forall\, 2 \leq k \leq n+1\,, \quad X_k^{(n+1)}[\,i\,] = X_{k-1}^{(n)}[\,X_1^{(n+1)}[\,i\,]\,]\,.$$

□

Autrement dit, pour tout $n \geq 1$, tout $1 \leq k \leq n$ et tout $i \in E$, on a

$$X_n^{(n)}[\,i\,] = \Phi_Q^{(-1)}\Big(U_{(-1)}\,,\, \Phi_Q^{(-2)}\Big(U_{-2}\,,\,\ldots\,,\,\Phi_Q^{(-(n-1))}\big(U_{-(n-1)}\,,\,\Phi_Q^{(-n)}(U_{-n}, i)\big)\ \ldots\ \Big)\Big).$$

Faisons deux remarques cruciales.

**1)** *Les mêmes variables uniformes* $(U_{-n}, \ldots, U_{-1})$ sont utilisées pour simuler à $n$ fixé les transitions de toutes les suites $(X_k^{(n)}[\,i\,]\,;\, 0 \leq k \leq n)\,,\, i \in E$. □

**2)** Les suites *coalescent* :

$$X_{k_0}^{(n)}[\,i\,] = X_{k_0}^{(n)}[\,j\,] \implies \forall k \geq k_0\,, \quad X_k^{(n)}[\,i\,] = X_k^{(n)}[\,j\,].$$

Notamment, si $X_{k_0}^{(n)}[\,i\,] = X_{k_0}^{(n)}[\,j\,]$ alors $X_n^{(n)}[\,i\,] = X_n^{(n)}[\,j\,]$. □

On pose alors

$$\Delta_n = \left\{ X_n^{(n)}[\,i\,] \; ; \; i \in E \right\} \quad \text{et} \quad D_n = \# \Delta_n \; .$$

$D_n$ est donc le nombre de points distincts dans l'ensemble $\Delta_n$. La remarque précédente montre que $\Delta_{n+1} \subset \Delta_n$ et donc $D_{n+1} \leq D_n$. On introduit ensuite le *temps de coalescence*

$$T = \inf \left\{ n \geq 1 \; : \; D_n = 1 \right\}$$

avec la convention que $T = \infty$ si pour tout $n \geq 1$, on a $D_n \geq 2$. On voit que si $T < \infty$, alors pour tout $n \geq T$, on a $D_n = 1$, $\Delta_n$ est réduit à un singleton $\{X_*\}$, c'est-à-dire que :

$$\text{sur } \{T < \infty\}, \quad \forall n \geq T \, , \; \forall i \in E \, , \quad X_n^{(n)}[\,i\,] = X_* \; .$$

Sur $\{T = \infty\}$, on pose $X_* = i_{\texttt{toc}}$, où $i_{\texttt{toc}} \in E$, est un état que l'on fixé à l'avance qui ne joue aucun rôle spécifique. Le théorème suivant montre que la simulation rétrograde donne une simulation exacte d'une variable de loi $\pi$.

**Théorème I.5.10** *Avec les notations précédentes, si* $\mathbf{P}(T < \infty) = 1$, *alors la loi de* $X_*$ *est* $\pi$.

**Preuve :** on fixe $i \in E$. On observe que $(X_k^{(n)}[\,i\,]\,;\, 0 \leq k \leq n)$ a même loi qu'une chaîne de Markov issue de $i$ presque sûrement et de matrice de transition $Q$. Donc, pour tout $j \in E$.

$$\mathbf{P}\big( X_k^{(n)}[\,i\,] = j \big) = [Q^n](i,j) \; .$$

On observe ensuite que

$$\begin{aligned} \mathbf{P}\big( X_k^{(n)}[\,i\,] = j \big) &= \mathbf{P}(X_* = j\,;\, T \leq n) + \mathbf{P}(X_k^{(n)}[\,i\,] = j\,;\, T > n) \\ &= \mathbf{P}(X_* = j) + \mathbf{P}(X_k^{(n)}[\,i\,] = j\,;\, T > n) - \mathbf{P}(X_* = j\,;\, T > n). \end{aligned}$$

Cela implique que pour tout $j \in E$,

$$\big| \, [Q^n](i,j) - \mathbf{P}(X_* = j) \, \big| \leq \mathbf{P}(T > n) \; .$$

Comme $Q$ est irréductible apériodique et que l'espace d'état est fini, on a $\lim_{n \to \infty} [Q^n](i,j) = \pi(j)$ ; par le théorème I.4.17 de la convergence vers la loi stationnaire. De plus, comme on a supposé $\mathbf{P}(T < \infty)$, on a $\lim_{n \to \infty} \mathbf{P}(T > n) = 0$ et donc on a bien $\mathbf{P}(X_* = j) = \pi(j)$, pour tout $j \in E$. ■

À la lumière des définitions précédentes, l'algorithme de Propp-Wilson peut se formuler comme suit.

**Algorithme de Propp-Wilson.** On suppose que sur $(\Omega, \mathscr{F}, \mathbf{P})$, il existe une suite $U_{-n} : \Omega \to [0,1]$, $n \geq 1$, de variables indépendantes et de loi uniforme sur $[0,1]$.

$(a)$ (*Initialisation*) La variable qui compte les boucles est $n$ ; elle est initialisée à $n = 1$. Le stock de v.a. uniformes est noté `Stock` ; il est est initialisé à `Stock` $= \emptyset$.

$(b)$ (*Simulation des chaînes couplées*) On a conservé en mémoire les variables

$$\texttt{Stock} = (U_{-(n-1)}, U_{-(n-2)}, \ldots, U_{-1})$$

générées lors des $n-1$ étapes précédentes. On tire $U_{-n}$ et pour chaque $i \in E$, on construit les suites $(X_k^{(n)}[i])_{0 \leq k \leq n}$ en posant récursivement

$$X_0^{(n)}[\,i\,] = i \quad \text{et} \quad \forall\, 0 \leq k < n \, , \quad X_{k+1}^{(n)}[\,i\,] = \Phi_Q^{-(n-k)} \big( U_{-(n-k)}, X_k^{(n)}[\,i\,] \big) \; .$$

On passe alors à l'instruction $(c)$.

$(c)$ (*Arrêt ou poursuite*)

— Si $\#\big\{X_n^{(n)}[\,i\,]\,;\, i \in E\,\big\} = 1$, toutes les variables $X_n^{(n)}[\,i\,]$, $i \in E$, sont égales à une même variable $X_*$. Le temps de coalescence est $T = n$ et *l'algorithme s'arrête et retourne* $X_*$.

— Si $\#\big\{X_n^{(n)}[\,i\,]\,;\, i \in E\,\big\} \geq 2$ (il est possible de s'en apercevoir au cours de $(b)$), alors on passe à l'instruction $(d)$ ci-dessous.

$(d)$ On met $U_{-n}$ dans `Stock` : `Stock` $\to (U_{-n}, U_{-(n-1)}, U_{-(n-2)}, \ldots, U_{-1})$. On augmente $n$ de 1 : $n \to n+1$, et on passe à l'instruction $(b)$, après ces mises à jour.

Si $\mathbf{P}(T < \infty)$, l'algorithme s'arrête avec probabilité 1 et $X_*$ suit la loi $\pi$. □

Faisons tout d'abord quelques remarques « théoriques ». Signalons ici que l'algorithme dépend fortement du choix des fonctions d'échantillonnage $\Phi_Q^{(-n)}$ et qu'un mauvais choix peut entraîner qu'avec probabilité strictement positive, l'algorithme de se termine jamais (c'est-à-dire que les chaîne simulées ne coalescent jamais) comme le montre l'exemple suivant.

**Exemple I.5.11** . On considère $E = \{i_1, i_2\}$, $i_1 \neq i_2$ et la matrice de transition $Q = (p(i,j))_{i,j\in E}$ donnée par

$$p(i_1, i_1) = p(i_1, i_2) = p(i_2, i_2) = p(i_2, i_1) = \tfrac{1}{2} \ .$$

Elle est clairement irréductible et sa probabilité invariante est $\pi(i_1) = \pi(i_2) = \frac{1}{2}$. On définit

$$\Phi_Q(i_1, x) = i_2 \mathbf{1}_{[0,\frac{1}{2}[}(x) + i_1 \mathbf{1}_{[\frac{1}{2},1[}(x) \quad \text{et} \quad \Phi_Q(i_2, x) = i_1 \mathbf{1}_{[0,\frac{1}{2}[}(x) + i_2 \mathbf{1}_{[\frac{1}{2},1[}(x).$$

On voit que pour tout $x \in [0,1[$, $\Phi_Q(i_1,x) \neq \Phi_Q(i_2,x)$. Par conséquent, il ne peut jamais y avoir coalescence. En revanche si on choisit $\Phi_Q(i_1,x) = \Phi_Q(i_2,x) = i_1 \mathbf{1}_{[0,\frac{1}{2}[}(x) + i_2 \mathbf{1}_{[\frac{1}{2},1[}(x)$, alors il y a coalescence dès la première étape de l'algorithme : $\mathbf{P}(T = 1) = 1$. □

Le lemme suivant donne une condition théorique sous laquelle l'algorithme de Propp-Wilson converge rapidement (en temps géométrique) : la condition est que les $\Phi_Q^{(-n)}$ doivent favoriser la coalescence (cela n'est pas surprenant).

**Lemme I.5.12** *Soit $Q = (p(i,j))_{i,j\in E}$, une matrice de transition sur $E$ irréductible apériodique de loi invariante $\pi$. Pour tout entier $n \geq 1$, soit $\Phi_Q^{(-n)} : [0,1] \times E \to E$, une fonction d'échantillonnage. Pour tout sous ensemble $F \subset E$ et tout $x \in [0,1]$, on note $\Phi_Q^{(-n)}(x,F)$ l'ensemble image $\{\Phi_Q^{(-n)}(x,i)\,;\, i \in F\}$. Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité sur lequel sont définies les variables $X_k^{(n)}[i]$ avec couplage rétrograde comme expliqué précédemment. Soit $U : \Omega \to [0,1]$ est une v.a. $\mathscr{F}$-mesurable de loi uniforme sur $[0,1]$. On suppose qu'il existe $0 < \rho < 1$ tel que pour tout $n \geq 1$ on ait*

$$\forall F \subset E \quad \textit{tel que} \quad \#F \geq 2\,, \quad \mathbf{P}\big(\#\Phi_Q^{(-n)}(U,F) \ \leq \ \#F\ -1\big) \geq \rho\,. \tag{I.99}$$

*Si $T$ est le temps de coalescence associé à l'algorithme de Propp-Wilson, alors pour tout entier $q \geq 1$, on a*

$$\mathbf{P}\big(\,T\ \geq q\#E\,\big) \leq \big(1 - \rho^{\#E-1}\big)^q \ \xrightarrow[q\to\infty]{} \ 0\ .$$

Avant de donner une preuve de ce résultat, faisons deux remarques.

**Remarque I.5.13** Supposons que la chaîne possède un état d'étranglement comme en (I.98). Soit $i$, un état distinct de $i_0$. Si on pose $F = \{i, i_0\}$, alors $\#\Phi_Q(U,F)) = 2$ et l'hypothèse (I.99) du lemme ne peut être vérifiée. □

**Remarque I.5.14** Ce théorème montre que la queue de $T$ est exponentiellement petite mais l'estimée qu'il contient implique que $\mathbf{E}[T] \leq \rho^{-\#E}\#E$. Dans la pratique, pour que l'algorithme tourne en un temps raisonnable, $T$ doit être beaucoup plus petit que $\rho^{-\#E}\#E$ qui est une durée potentiellement astronomique. La borne pour $T$ donnée dans l'énoncé est d'ailleurs en général (extraordinairement) mauvaise et n'a qu'un interêt purement théorique. $\square$

**Preuve :** on pose $\Delta_k^{(n)} = \big\{X_k^{(n)}[\,i\,]\,;\, i \in E\big\}$. On rappelle que $\Delta_{k+1}^{(n)} \subset \Delta_k^{(n)}$. On remarque que

$$\mathbf{E}\big[\mathbf{1}_{\{\#\Delta_{k+1}^{(n)} \leq \#\Delta_k^{(n)}-1\}} |\, U_{-n}, \ldots, U_{-(n-k+1)}\big] \geq \rho > 0 \,.$$

Pour simplier on pose $a = \#E - 1$. En itérant ce qui précède, on voit que pour tout $n \geq a$, et tout $m \leq n-a$,

$$\mathbf{E}\big[\mathbf{1}_{\{\forall\, 0\leq k<a\,,\, \#\Delta_{m+k+1}^{(n)} < \#\Delta_{m+k}^{(n)}\}} |\, U_{-n}, \ldots, U_{-(n-m+1)}\big] \geq \rho^a > 0 \,.$$

Cela implique immédiatement que pour tout $n \geq a$, et tout $m \leq n-a$,

$$\mathbf{E}\big[\mathbf{1}_{\{\#\Delta_{m+a}^{(n)}=1\}} |\, U_{-n}, \ldots, U_{-(n-m+1)}\big] \geq \rho^{a-1} > 0 \,.$$

On pose $n = qa$ avec $q \geq 1$ et pour tout $0 \leq k < q$, on pose $B_k = \{\#\Delta_{(k+1)a}^{(n)} = 1\}$. On remarque que

$$\mathbf{1}_{\{T>n\}} \leq \big(1-\mathbf{1}_{B_0}\big)\big(1-\mathbf{1}_{B_1}\big)\ldots\big(1-\mathbf{1}_{B_{q-1}}\big).$$

Ce qui précède montre que

$$\mathbf{E}\big[\mathbf{1}_{\{T>qa\}}|U_{-qa}, \ldots, U_{-(q+1)}\big] \leq \big(1-\mathbf{1}_{B_0}\big)\big(1-\mathbf{1}_{B_1}\big)\ldots\big(1-\mathbf{1}_{B_{q-2}}\big)(1-\rho^a).$$

En itérant, on obtient

$$\mathbf{E}\big[\mathbf{1}_{\{T>qa\}}|U_{-qa}, \ldots, U_{-(2q+1)}\big] \leq \big(1-\mathbf{1}_{B_0}\big)\big(1-\mathbf{1}_{B_1}\big)\ldots\big(1-\mathbf{1}_{B_{q-3}}\big)(1-\rho^a)^2$$

et ainsi de suite ... etc. On obtient $\mathbf{P}(T > qa) \leq (1-\rho^a)^q$, ce qui est le résultat désiré. $\blacksquare$

**Deux modifications pratiques de l'algorithme de Propp-Wilson.** On observe que l'algorithme de Propp-Wilson dans la version donnée précédemment fait appel aux fonctions d'échantillonnage $\Phi_Q^{-(n)}$ au pire $\#E(1+2+\ldots+T) = \frac{1}{2}T(T+1)\#E$ fois. Cela n'est pas optimal. Il est possible de modifier dans certaines situations l'algorithme de façon à ce qu'il dure un temps de l'ordre de quelques multiples de $T$, comme expliqué ci-dessous.

*Modification 1* : **le « sandwiching »**. Il est nécessaire à chaque étape $n$ de simuler $\#E$ chaînes et lorsque $E$ est gros, c'est extrêmement long. Le but de cette section est de donner une idée générale (le « sandwiching ») permettant de résoudre ce problème dans certaines situations.

Supposons que $E$ puisse être muni d'un ordre partiel $\preceq$. Supposons aussi qu'il existe un état $\preceq$-minimal noté $i_*$ et un état $\preceq$-maximal noté $i^*$ :

$$\forall i \in E\,, \quad i_* \preceq i \preceq i^* \,.$$

Nous supposons enfin qu'il soit possible de contruire des fonctions d'échantillonnage $\Phi_Q^{-(n)}$ qui soient $\preceq$-monotones dans le sens suivant : pour tout entier $n \geq 1$,

$$i_1 \preceq i_2 \implies \forall x \in [0,1]\,, \quad \Phi_Q^{-(n)}(x, i_1) \preceq \Phi_Q^{-(n)}(x, i_2) \,.$$

Cela implique la propriété suivante : à l'étape $n$ de l'algorithme de Propp-Wilson, si $i_1 \preceq i_2$, alors $X_k^{(n)}[\, i_1\,] \preceq X_k^{(n)}[\, i_2\,]$, pour tout $0 \leq k \leq n$. Cela entraîne que

$$T = n \quad \text{si et seulement si} \quad X_k^{(n)}[\, i_*\,] = X_k^{(n)}[\, i^*\,] \; .$$

À chaque étape, on n'a plus besoin que de simuler deux chaînes : l'une issue de $i_*$, l'autre issue de $i^*$, ce qui est beaucoup mieux que de simuler $\# E$ chaînes. Ce principe s'appelle le *sandwiching*. L'algorithme se reformule alors comme suit.

**Propp-Wilson avec sandwiching.** On suppose que sur $(\Omega, \mathscr{F}, \mathbf{P})$, il existe une suite $U_{-n} : \Omega \to [0,1]$, $n \geq 1$, de variables indépendantes et de loi uniforme sur $[0,1]$.

$(a)$ (*Initialisation*) La variable qui compte les boucles est $n$ ; elle est initialisée à $n = 1$. Le stock de v.a. uniformes est noté `Stock` ; il est est initialisé à `Stock` $= \emptyset$.

$(b)$ (*Simulation des chaînes couplées*) On a conservé en mémoire les variables

$$\texttt{Stock} = (U_{-(n-1)}, U_{-(n-2)}, \ldots, U_{-1})$$

générées lors des $n-1$ étapes précédentes. On tire $U_{-n}$ et on construit récursivement deux suites $(Y_k^{(n)})_{0 \leq k \leq n}$ et $(Z_k^{(n)})_{0 \leq k \leq n}$ en posant $Y_0^{(n)} = i_*$, $Z_0^{(n)} = i^*$ et

$$\forall\, 0 \leq k < n\,, \quad Y_{k+1}^{(n)} = \Phi_Q^{(-(n-k))}\big(\, U_{-(n-k)}, Y_k^{(n)}\,\big) \text{ et } Z_{k+1}^{(n)} = \Phi_Q^{(-(n-k))}\big(\, U_{-(n-k)}, Z_k^{(n)}\,\big).$$

On passe alors à l'instruction $(c)$.

$(c)$ (*Arrêt ou poursuite*)

— Si $Y_n^{(n)} = Z_n^{(n)}$, alors $T = n$, *l'algorithme s'arrête* et retourne $X_* = Y_n^{(n)} = Z_n^{(n)}$.

— Si $Y_n^{(n)} \neq Z_n^{(n)}$, (il est possible de s'en apercevoir au cours de $(b)$), alors on passe à l'instruction $(d)$ ci-dessous.

$(d)$ On met $U_{-n}$ dans `Stock` : `Stock` $\to (U_{-n}, U_{-(n-1)}, U_{-(n-2)}, \ldots, U_{-1})$. On augmente $n$ de $1$ : $n \to n+1$, et on passe à l'instruction $(b)$, après ces mises à jour.

Si $\mathbf{P}(T < \infty)$, l'algorithme s'arrête avec probabilité 1 et $X_*$ suit la loi $\pi$. □

On observe que l'algorithme de Propp-Wilson dans la version donnée précédemment fait appel à la fonction d'échantillonnage $\Phi_Q^{-(n)}$ au pire $2(1+2+\ldots+T) = T(T+1)$ fois, ce qui n'est toujours pas optimal.

Modification 2 : **simulation par blocs de temps**. L'idée (combinée avec le sandwiching) est qu'au lieu de simuler deux chaînes de durée 1, puis deux de durée 2, . . ., puis deux chaînes de durée $n$ etc, juqu'au temps de coalescence $T$, on effectue des simulations des deux chaînes pour des blocs de temps $N_1 < N_2 < \ldots N_m < \ldots$. Dans des variantes de l'algorithme, les $N_m$ peuvent être aléatoires mais dans ce qui suit nous nous en tiendrons aux cas où la suite de temps $N_m$ est déterminée à l'avance et où $N_{m+1} \geq RN_m$ avec un ratio $R > 1$.

**Propp-Wilson par blocs temporels et sandwiching.** On suppose que sur $(\Omega, \mathscr{F}, \mathbf{P})$, il existe une suite $U_{-n} : \Omega \to [0,1]$, $n \geq 1$, de variables indépendantes et de loi uniforme sur $[0,1]$. On suppose que $(N_m)_{m \geq 1}$ une suite strictement croissante de temps fixés.

$(a)$ (*Initialisation*) La variable qui compte les boucles est $m$ ; elle est initialisée à $m = 1$. Le stock de v.a. uniformes est noté `Stock` ; il est est initialisé à `Stock` $= \emptyset$.

$(b)$ (*Simulation des chaînes couplées*) On a conservé en mémoire les variables

$$\texttt{Stock} = (U_{-N_{m-1}}, U_{-(N_{m-1}-1)}, U_{-(N_{m-1}-2)}, \ldots, U_{-2}, U_{-1})$$

générées lors des $m-1$ étapes précédentes. On tire les variables $U_{-N_m}$, $U_{-(N_m-1)}$, …, $U_{-(N_{m-1}+2)}$, $U_{-(N_{m-1}+1)}$ et on construit récursivement deux suites $(Y_k^{(N_m)})_{0\le k\le N_m}$ et $(Z_k^{(N_m)})_{0\le k\le N_m}$ en posant $Y_0^{(N_m)} = i_*$, $Z_0^{(N_m)} = i^*$ et

$$\forall k\in\{0,\ldots,N_m\},\ Y_{k+1}^{(N_m)}=\Phi_Q^{(-(N_m-k))}\big(U_{-(N_m-k)}, Y_k^{(N_m)}\big) \text{ et } Z_{k+1}^{(N_m)}=\Phi_Q^{(-(N_m-k))}\big(U_{-(N_m-k)}, Z_k^{(N_m)}\big).$$

On passe alors à l'instruction $(c)$.

$(c)$ (*Arrêt ou poursuite*)

— Si $Y_{N_m}^{(N_m)} = Z_{N_m}^{(N_m)}$, alors le temps de coalescence est tel que $N_{m-1} < T \le N_m$ et *l'algorithme s'arrête* et retourne $X_* = Y_{N_m}^{(N_m)} = Z_{N_m}^{(N_m)}$.

— Si $Y_{N_m}^{(N_m)} \neq Z_{N_m}^{(N_m)}$, (il est possible de s'en apercevoir au cours de $(b)$), alors on passe à l'instruction $(d)$ ci-dessous.

$(d)$ On met les v.a. $U_{-N_m}$, $U_{-(N_m-1)}$, …, $U_{-(N_{m-1}+2)}$, $U_{-(N_{m-1}+1)}$ dans `Stock` :

$$\texttt{Stock} \to (U_{-N_m}, U_{-(N_m-1)}, U_{-(N_m-2)}, \ldots, U_{-2}, U_{-1})\ .$$

On augmente $m$ de $1$ : $m \to m+1$, et on passe à l'instruction $(b)$, après ces mises à jour.

Si $\mathbf{P}(T < \infty)$, l'algorithme s'arrête avec probabilité $1$ et $X_*$ suit la loi $\pi$. $\square$

On voit que si on prend

$$\forall m \in \mathbb{N}, \quad N_m = 2^m$$

l'algorithme de Propp-Wilson ainsi modifié (sandwiching et simulation par blocs temporels) fait appel aux fonctions d'échantillonnage $\Phi_Q^{(-n)}$ au pire $2(N_1 + \ldots + N_{m^*})$ fois, où $m^*$ est l'indice (aléatoire) tel que $N_{m^*-1} < T \le N_{m^*}$. Or si les $N_m$ sont des puissantes de deux $N_{m^*} \le 2T$ et $2(N_1 + \ldots + N_{m^*}) = 4(N_{m^*} - 1) \le 8T$.

*L'algorithme fait appel au pire $8T$ fois aux fonctions d'échantillonnage $\Phi_Q^{(-n)}$,*

ce qui est bien mieux que les $\frac{1}{2}T(T+1)\#E$ fois de la première version.

**Exemple I.5.15** On reprend l'exemple du modèle d'Ising. Soit $\mathbf{G} = (\mathbf{S}, \mathbf{A})$ un graphe simple, fini et connexe. L'ensemble des configurations de spin est donc $E = \{+1,-1\}^{\mathbf{S}}$. On se donne $(J_a; a \in \mathbf{A})$, l'énergie interne. Pour simplifier on prend un champ magnétique nul. L'énergie des configurations de spins est donc donnée par

$$\forall \eta \in \{+1,-1\}^{\mathbf{S}}, \quad V(\eta) = -\tfrac{1}{2} \sum_{\{s,s'\}\in\mathbf{A}} J_{\{s,s'\}}\eta(s)\eta(s')\ .$$

La mesure d'Ising à température $1/\beta$ est donc

$$\forall \eta \in \{+1,-1\}^{\mathbf{S}}, \quad \pi_\beta(\eta) = \tfrac{1}{Z_\beta}\exp\big(-\beta V(\eta)\big) \quad \text{où} \quad Z_\beta = \sum_{\eta'\in\{+1,-1\}^{\mathbf{S}}} \exp\big(-\beta V(\eta')\big).$$

On définit une fonction d'échantillonnage $\Phi$ correspondant informellement au méchanisme suivant : on choisit un sommet au hasard $s$, on aligne le spin en $s$ en position $+1$ avec une certaine probabilité à préciser et en position $-1$ avec la probabilité complémentaire, on laisse inchangés les autres spins de la configuration. Pour éviter les complications inutiles $\Phi$ prend comme source aléatoires *deux variables* uniformes indépendantes $U$ et $U'$ : l'une sert à choisir un sommet uniformément au hasard, l'autre sert à changer le spin au sommet choisi. Pour toute configuration de spins $\eta \in \{+1,-1\}^{\mathbf{S}}$ et tout sommet $s \in \mathbf{S}$, on pose

$$\forall s \in \mathbf{S}\,, \quad \sigma(\eta, s) = \sum_{s' : s \sim s'} J_{\{s,s'\}} \eta(s')$$

et on définit deux nouvelles configurations $a_s^+\eta$ et $a_s^-\eta$ par

$$a_s^+\eta(s') = a_s^+\eta(s') = \eta(s')\,,\ s' \in \mathbf{S}\backslash\{s\} \quad \text{et} \quad a_s^+\eta(s) = +1 \text{ et } a_s^-\eta(s) = -1\ .$$

Autrement dit $a_s^+\eta$ (resp. $a_s^-\eta$) coïncide avec $\eta$ sauf en $s$ où le spin vaut $+1$ (resp. $-1$). On suppose donnée, $\phi : [0,1[\to \mathbf{S}$, une fonction d'échantillonnage de la loi uniforme sur $\mathbf{S}$ :

$$\forall\, s \in \mathbf{S}\,, \quad \mathbf{P}(\phi(U) = s) = \frac{1}{\#\mathbf{S}}\ .$$

On fixe $x, y \in [0,1]$ ; on pose $s_0 = \phi(x)$, et on pose $\Phi(x, y; \eta) = \eta'$, où $\eta'(s) = \eta(s)$, pour tout $s \in \mathbf{S}\backslash\{s_0\}$ et

$$\eta'(s_0) = \begin{cases} +1 & \text{si } y < \frac{\exp(-\beta\sigma(\eta,s_0))}{1+\exp(-\beta\sigma(\eta,s_0))}\,, \\ -1 & \text{sinon.} \end{cases}$$

Autrement dit, pour tous $x, y \in [0,1]$ et toute configuration de spins $\eta \in \{+1,-1\}^{\mathbf{S}}$.

$$\Phi(x, y; \eta) = a^+_{\phi(x)}\eta \quad \text{si} \quad y < \frac{\exp(-\beta\sigma(\eta,\phi(x)))}{1+\exp(-\beta\sigma(\eta,\phi(x)))} \quad \text{et} \quad \Phi(x, y; \eta) = a^-_{\phi(x)}\eta \quad \text{sinon.}$$

On définit une matrice de transition $Q = (p(\eta, \eta'))_{\eta,\eta' \in \{+1,-1\}^{\mathbf{S}}}$ sur l'espace des configurations de spin en posant

$$\mathbf{P}\big(\, \Phi(U, U'; \eta) = \eta'\,\big) = p(\eta, \eta')\,, \quad \eta, \eta' \in \{+1,-1\}^{\mathbf{S}}\ .$$

Par définition $\Phi$ est une fonction d'échantillonnage de $Q$. Il est facile de vérifier que $Q$ est irréductible et comme $p(\eta, \eta) > 0$, elle est apériodique. Montrer que $\pi_\beta$ est sa probabilité invariante.

**En effet**, on fixe $x \in [0,1]$ et on note $s_0 = \phi(x)$. On note $h(u) = \frac{u}{1+u}$. On voit que

$$\mathbf{P}\big(\, \Phi(x, U'; \eta) = \eta'\,\big) = \begin{cases} 0 & \text{si } \eta' \notin \{a^+_{s_0}\eta, a^+_{s_0}\eta\} \\ h\big(e^{-\beta\sigma(\eta,s_0)}\big) & \text{si } \eta' = a^+_{s_0}\eta\,, \\ h\big(e^{\beta\sigma(\eta,s_0)}\big) & \text{si } \eta' = a^-_{s_0}\eta\,, \end{cases}$$

On observe ensuite que

$$\mathbf{P}\big(\, \Phi(x, U'; a^+_{s_0}\eta) = \eta\,\big) = \mathbf{P}\big(\, \Phi(x, U'; a^-_{s_0}\eta) = \eta\,\big) = h\big(e^{-\eta(s_0)\beta\sigma(\eta,s_0)}\big).$$

Un calcul élémentaire montre que

$$\frac{\mathbf{P}\big(\, \Phi(x, U'; \eta) =\ a^+_{s_0}\eta\big)}{\mathbf{P}\big(\, \Phi(x, U'; a^+_{s_0}\eta) =\ \eta\big)} = \frac{\pi_\beta(a^+_{s_0}\eta)}{\pi_\beta(\eta)} \quad \text{et} \quad \frac{\mathbf{P}\big(\, \Phi(x, U'; \eta) =\ a^-_{s_0}\eta\big)}{\mathbf{P}\big(\, \Phi(x, U'; a^-_{s_0}\eta) =\ \eta\big)} = \frac{\pi_\beta(a^-_{s_0}\eta)}{\pi_\beta(\eta)}.$$

On en déduit alors que

$$\forall x \in [0,1]\,, \forall \eta' \in \{-1,1\}^{\mathbf{S}}\,, \quad \pi_\beta(\eta)\mathbf{P}\big(\, \Phi(x, U'; \eta) = \eta'\,\big) = \pi_\beta(\eta')\mathbf{P}\big(\, \Phi(x, U'; \eta') = \eta\,\big).$$

En intégrant en $x$, on obtient que $\pi_\beta(\eta)p(\eta,\eta') = \pi_\beta(\eta')p(\eta',\eta)$, pour tous $\eta,\eta' \in \{+1,-1\}^{\mathbf{S}}$, ce qui montre que $\pi_\beta$ est une mesure de réversibilité de $Q$, ce qui permet de conclure. ■

On introduit l'ordre partiel $\preceq$ sur l'espace des configurations de spin : soient $\eta_1,\eta_2 \in \{+1,-1\}^{\mathbf{S}}$, alors

$$\eta_1 \preceq \eta_2 \iff \forall s \in \mathbf{S}, \quad \eta_1(s) \le \eta_2(s)\ .$$

Cet ordre admet la configuration $\preceq$-maximale $\eta^*$ dont tous les spins valent $+1$ et la configuration $\preceq$-minimale $\eta_*$ dont tous les spins valent $-1$. Montrons que $\Phi$ est $\preceq$-croissante.

$$\eta_1 \preceq \eta_2 \implies \forall x,y \in [0,1]\,, \quad \Phi(x,y;\eta_1) \preceq \Phi(x,y;\eta_2)\ .$$

**En effet**, supposons $\eta_1 \preceq \eta_2$ et fixons $x,y \in [0,1]$. On pose $s_0 = \phi(x)$. Comme $J_{s_0,s_0'} > 0$, pour tout $s_0' \sim s_0$, on voit facilement que $\sigma(\eta_1,s_0) \le \sigma(\eta_2,s_0)$. On voit donc que

$$\frac{\exp(-\beta\sigma(\eta_1,s_0)\,)}{1+\exp(-\beta\sigma(\eta_1,s_0)\,)} \le \frac{\exp(-\beta\sigma(\eta_2,s_0)\,)}{1+\exp(-\beta\sigma(\eta_2,s_0)\,)}.$$

On en déduit que si $\Phi(x,y;\eta_1) = a^+_{s_0}\eta_1$ alors $\Phi(x,y;\eta_2) = a^+_{s_0}\eta_2$ et on remarque facilement que $a^+_{s_0}\eta_1 \preceq a^+_{s_0}\eta_2$. On observe ensuite que $a^-_{s_0}\eta_1 \preceq a^-_{s_0}\eta_2 \preceq a^+_{s_0}\eta_2$. Dans tous les cas on a $\Phi(x,y;\eta_1) \preceq \Phi(x,y;\eta_2)$. ■

On a donc montré que $\Phi$ est la fonction d'échantillonnage d'une matrice de transition sur l'espace des configurations qui est irréductible, apériodique, reversible et dont la mesure d'Ising à la température $1/\beta$ est la loi invariante. De plus nous avons montré que $\Phi$ est monotone pour l'ordre partiel $\preceq$, ordre qui admet un élement maximal $\eta^*$ et un élément minimal $\eta_*$. L'algorithme de Propp-Wilson avec sandwiching peut donc s'appliquer.

Nous laissons au lecteur intéressé le soin d'implémenter cet algorithme sur le graphe constitué de la grille toroïdale de taille $n$, c'est-à-dire $(\mathbb{Z}/n\mathbb{Z})^2$ pour lequel la température de Curie-Weiss est connue. Mais avant de se lancer dans cette tâche amusante mais longue, voici quelques avertissements.

a) La méthode exposée ci-dessus peut marcher telle quelle sur une grille qui n'est pas trop grande ($n = 300$) et ce, à des températures plus hautes que la température critique, (pour des températures hautes les spins sont désordonnés et donc relativement indépendants). Pour des températures hautes la chaîne de Markov converge en un temps polynomial (de petit degré). Néanmoins, si la taille de la grille augmente le temps d'exécution s'allonge et le stockage des variables uniformes peut excéder la mémoire de l'ordinateur. Il s'agit d'un inconvénient général de l'algorithme de Propp-Wilson. Mentionnons que Wilson en 2000 a apporté des solutions pour éviter de stocker trop de variables (Propp-Wilson with *with read-once randomness*).

b) Le choix d'une chaîne qui choisit au hasard *successivement* les sites des spins à mettre à jour n'est, dans le cas d'une grille, pas optimal. Comme la quantité $\sigma(s,\eta)$ ne dépend que des valeurs des spins sur les sites voisins de $s$, sur la grille toroïdale $(\mathbb{Z}/n\mathbb{Z})^2$ il est possible de changer *d'un coup* tous les sommets de coordonnées paires puis d'un coup tous les sommets de coordonnées impaires en parallélisant ce calcul (il s'agit toujours d'une chaîne de loi invariante $\pi_\beta$). Cette parallélisation permet évidemment de gagner énormément de temps.

c) Lorsque que l'on s'approche de la température critique $T_c = 1/\beta_c$ où $\beta_c = 0,441...$ est la constante de Onsager, l'algorithme dure de plus en plus longtemps. Lorsque la température est en dessous de $T_c$ (c'est-à-dire $\beta > \beta_c$), le temps de convergence de la chaîne de loi invariante $\pi_\beta$ est exponentiellement grand et l'algorithme ne se termine pas dans des délais raisonnables. Pour simuler le modèle d'Ising dans ces cas, il faut utiliser une propriété qui le lie à un autre modèle, le *cluster aléatoire de Fortuin-Kasteleyn* : on applique l'algorithme de Propp-Wilson à cet autre modèle qui est immunisé contre le ralentissement à la phase critique puis on reconvertit ce résultat pour le modèle d'Ising.

□

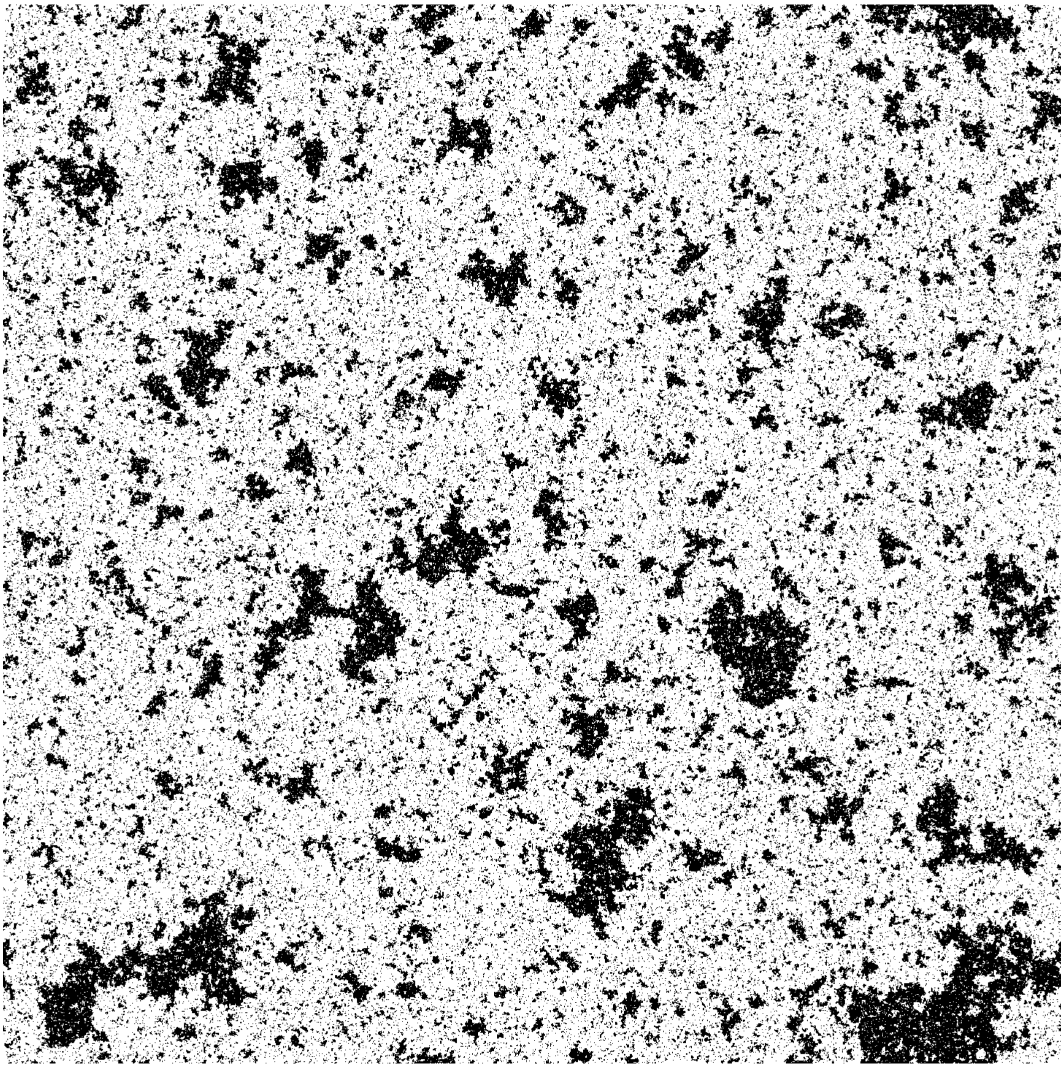

FIGURE I.14 – Le modèle d'Ising à la température critique sur une grille toroïdale $4200 \times 4200$. Image obtenue par D.B Wilson.

## I.6 Génération d'arbres couvrants aléatoires.

Dans cette section $(\Omega, \mathscr{F}, \mathbf{P})$ désigne un espace de probabilité sur lequel sont définies les variables mentionnées. Soit $\mathbf{G} = (\mathbf{S}, \mathbf{A})$ un graphe simple (les arêtes ne sont pas orientées, pas d'arêtes multiples, pas de boucles) muni d'un système de poids $\mathbf{C} = (C_a)_{a \in \mathbf{A}}$ *strictement positifs* : $C_a > 0$, $a \in \mathbf{A}$. On rappelle les définitions I.5.2, page 107, notamment celle d'arbre couvrant (et on rappelle qu'un arbre couvrant est donné par l'ensemble de ses arêtes). On suppose le graphe $\mathbf{G}$ *connexe*. On a vu à l'exemple I.5.1 (page 106) que dans ce cas, le graphe admet au moins un arbre couvrant. On rappelle également la notation $\mathbb{T}(\mathbf{G})$ pour l'ensemble des arbres couvrants de $\mathbf{G}$. On introduit le *poids d'un arbre couvrant* comme la quantité suivante.

$$\forall T \in \mathbb{T}(\mathbf{G}), \quad \mathtt{poids}(T) = \prod_{a \in T} C_a \ . \tag{I.100}$$

Le but de cette section est, en autre, de donner une méthode pour générer un arbre couvrant aléatoire $\mathcal{T} : \Omega \to \mathbb{T}(\mathbf{G})$ dont la loi est proportionnelle aux poids introduits en (I.100) :

$$\forall T \in \mathbb{T}(\mathbf{G}), \quad \mathbf{P}(\mathcal{T} = T) = \frac{\mathtt{poids}(T)}{Z(\mathbf{G}, \mathbf{C})} \quad \text{où} \quad Z(\mathbf{G}, \mathbf{C}) = \sum_{T' \in \mathbb{T}(\mathbf{G})} \mathtt{poids}(T') \ . \tag{I.101}$$

En particulier, lorsque tous les poids des arêtes sont égaux à une même constante $C = C_a$, $a \in \mathbf{A}$, alors $\mathtt{poids}(T) = C^{\#S-1}$ donc $Z(\mathbf{G}, \mathbf{C}) = C^{\#S-1} \# \mathbb{T}(\mathbf{G})$ et $\mathcal{T}$ est de *loi uniforme sur* $\mathbb{T}(\mathbf{G})$.

Dans la sous-section qui suit, on prouve le théorème de Kirchhoff (appelé parfois le *Tree-Matrix-Theorem*) qui interprête $Z(\mathbf{G}, \mathbf{C})$ en fonction du spectre de la matrice Laplacienne du graphe pondéré $(\mathbf{G}, \mathbf{C})$. Dans la sous-section suivante, on montre un premier algorithme permettant d'obtenir un arbre couvrant aléatoire dont la loi est donnée par (I.101) : cet algorithme a été obtenu indépendamment par Aldous et Broder en 1989-90. Une modification (due à Aldous) de cet algorithme dans le cas du graphe complet permet de construire (et simuler efficacement) un arbre uniforme à $N$ sommets dont une simulation avec $N = 50\,000$ est illustrée à la figure I.6.b, page 145.

### I.6.a Le théorème de Kirchhoff comptant le nombre d'arbres couvrants.

On considère un graphe simple $\mathbf{G} = (\mathbf{S}, \mathbf{A})$ connexe, possédant un nombre fini $N \geq 2$ de sommets. On munit $\mathbf{G}$ de poids $\mathbf{C} = (C_a)_{a \in \mathbf{A}}$ strictement positifs. Pour simplifier les notations dans ce qui suit, on *fixe une énumération de* $\mathbf{S}$ *puis on identifie* $\mathbf{S}$ *avec l'ensemble des entiers* $\{1, \ldots, N\}$, que l'on continue de noter génériquement $s, s'$, etc. L'ensemble des arêtes $\mathbf{A}$ de $\mathbf{G}$ est donc un sous ensemble de $\big\{\{s, s'\}; 1 \leq s < s' \leq N\big\}$.

Pour tout sommet $s \in \mathbf{S}$, on pose $\pi(s) = \sum_{s' \in \mathbf{S}} \mathbf{1}_{\{s' \sim s\}} C_{\{s, s'\}}$ et on définit la *matrice laplacienne* $\Delta$ du graphe pondéré $(\mathbf{G}, \mathbf{C})$ comme la matrice $N \times N$ dont les entrées réelles sont données pour tous $s, s' \in \{1, \ldots, N\}$ par :

$$[\Delta](s, s') = \begin{cases} \pi(s) & \text{si } s = s', \\ -\mathbf{1}_{\{s \sim s'\}} C_{s,s'} & \text{si } s \neq s'. \end{cases} \tag{I.102}$$

Il s'agit d'une matrice réelle symétrique : ses valeurs propres sont réelles. On les notes $\lambda_0, \ldots, \lambda_{N-1}$ où on a choisi une indexation telle que $|\lambda_0| \leq |\lambda_1| \leq \ldots \leq |\lambda_{N-1}|$.

Puisque $\pi(s) - \sum_{s' \in \mathbf{S}} \mathbf{1}_{\{s' \sim s\}} C_{\{s, s'\}} = 0$ pour tout $s \in \mathbf{S}$, on a $\Delta . \mathbf{1}_{\mathbf{S}} = 0$ et donc $\mathbf{1}_{\mathbf{S}}$ (qui est le vecteur colonne dont toutes les entrées sont égales à 1) est dans le noyau de $\Delta$. Cela entraîne $\lambda_0 = 0$. Le théorème de Kirchhoff ci-dessous montre que $0$ est de multiplicité $1$ (c'est-à-dire que le noyau de $\Delta$ est de dimension 1) et que le produit $\lambda_1 \ldots \lambda_{N-1}$ et lié à la fonction de partition $Z(\mathbf{G}, \mathbf{C})$ introduite en (I.101) et donc à $\#\mathbb{T}(\mathbf{G})$ dans le cas de poids constants.

**Théorème I.6.1** **(Théorème de Kirchhoff ou Tree Matrix Theorem)** *Soit $\mathbf{G} = (\mathbf{S}, \mathbf{A})$, un graphe simple et connexe. On suppose que $\mathbf{S} = \{1, \ldots, N\}$ où $N \geq 2$. On munit $\mathbf{G}$ de poids $\mathbf{C} = (C_a)_{a \in \mathbf{A}}$ strictement positifs. On rappelle la définition (I.100) du poids d'un arbre couvrant et la définition (I.101) de $Z(\mathbf{G}, \mathbf{C})$. On rappelle la définition (I.102) de $\Delta$, la matrice laplacienne du graphe pondéré, dont les valeurs propres sont notées $\lambda_0 = 0, \lambda_1, \ldots, \lambda_{N-1}$. Pour tous $s, s' \in \{1, \ldots, N\}$, on note $\Delta^{(s,s')}$ la matrice $(N-1) \times (N-1)$ obtenue en retirant à $\Delta$ sa ligne numero $s$ et sa colonne numero $s'$. Alors pour tous $s, s' \in \{1, \ldots, N\}$, on a*

$$Z(\mathbf{G}, \mathbf{C}) = (-1)^{s+s'} \mathtt{det}\big(\Delta^{(s,s')}\big) = \frac{(-1)^{N-1}}{N} \cdot \frac{d}{dx}_{|x=0} \mathtt{det}\big(x\mathtt{Id} - \Delta\big) = \frac{1}{N} \lambda_1 \ldots \lambda_{N-1} \, . \tag{I.103}$$

*Cette égalité permet notamment de calculer $\#\mathbb{T}(\mathbf{G})$ lorsque les poids $\mathbf{C}$ sont tous égaux à $1$.*

**Preuve :** supposons que l'on ait montré que $Z(\mathbf{G}, \mathbf{C}) = \mathtt{det}\big(\Delta^{(N,N)}\big)$. Comme $Z(\mathbf{G}, \mathbf{C})$ ne dépend pas d'une indexation des sommets de $\mathbf{G}$, on doit avoir $Z(\mathbf{G}, \mathbf{C}) = \mathtt{det}\big(\Delta^{(s,s)}\big)$ pour tout $s \in \mathbf{S}$. Par ailleurs, comme $Z(\mathbf{G}, \mathbf{C}) > 0$, cela implique donc que le noyau est de dimension 1 (et donc que $\lambda_1 \ldots \lambda_{N-1} \neq 0$). On fixe $s_0 \in \{1, \ldots, N\}$ et on pose $f(s) = (-1)^{s+s_0} \mathtt{det}\big(\Delta^{(s,s_0)}\big)$. En développant le déterminant selon la colonne numéro $s_0$, on observe que

$$\begin{aligned} \mathtt{det}(\Delta) &= \sum_{1 \leq s \leq N} [\Delta](s, s_0)(-1)^{s+s_0} \mathtt{det}\big(\Delta^{(s,s_0)}\big) \\ &= \pi(s_0) f(s_0) - \sum_{s \in \mathbf{S}} \mathbf{1}_{\{s \sim s_0\}} f(s) C_{s,s_0} \, . \end{aligned}$$

Comme $\mathtt{det}(\Delta) = 0$, on a $\Delta f = 0$ et puisque que le noyau est de dimension 1, $f$ est colinéaire à $\mathbf{1_S}$, c'est-à-dire que $f$ et constante et donc que $(-1)^{s+s_0} \mathtt{det}\big(\Delta^{(s,s_0)}\big) = (-1)^{s_0+s_0} \mathtt{det}\big(\Delta^{(s_0,s_0)}\big) = Z(\mathbf{G}, \mathbf{C})$, pour tout $s \in \mathbf{S}$. Cela entraîne la première égalité de (I.103).

Le fait que le déterminant soit une forme linéaire alternée des colonnes implique ensuite que

$$\frac{d}{dx}_{|x=0} \mathtt{det}\big(x\mathtt{Id} - \Delta\big) = \sum_{1 \leq s \leq N} \mathtt{det}\big(-\Delta^{(s,s)}\big) = (-1)^{N-1} N \mathtt{det}\big(\Delta^{(N,N)}\big),$$

ce qui entraîne la deuxième égalité de (I.103). On remarque ensuite que le polynôme caractéristique s'écrit $\mathtt{det}\big(x\mathtt{Id} - \Delta\big) = x(x - \lambda_1) \ldots (x - \lambda_{N-1})$ et sa dérivée en zéro vaut donc $(-1)^{N-1} \lambda_1 \ldots \lambda_{N-1}$, ce qui entraîne la troisième et dernière égalité de (I.103).

Pour montrer le théorème il suffit donc de prouver que

$$Z(\mathbf{G}, \mathbf{C}) = \mathtt{det}\big(\Delta^{(N,N)}\big) \, . \tag{I.104}$$

Pour cela on introduit tout d'abord la notion de *champ discret sur* $\mathbf{G}$ : un tel champ est une application $\overrightarrow{h} : \mathbf{S} \backslash \{N\} \to \mathbf{S}$ qui en chaque sommet du graphe distinct de $N$ ($N$ qui est un sommet *mort*) indique un sommet voisin, c'est-à-dire $s \sim \overrightarrow{h}(s)$ pour tout $s \in \mathbf{S} \backslash \{N\}$. On note $\Gamma$ l'ensemble des champs discrets sur $\mathbf{G}$. Pour tout $s \in \mathbf{S} \backslash \{N\}$, la *ligne de champ issue de* $s$ est un chemin $(s_n)$ qui a une durée maximale (possiblement infinie) et qui est tel que $s_0 = s$ et $s_{n+1} = \overrightarrow{h}(s_n)$, pour tout $n \in \mathbb{N}$. Deux cas sont possibles.

- Ou bien la ligne de champ issue de $s$ a une durée finie notée $m$ et elle se termine nécessairement au sommet $N$ : $s = s_0 \sim s_1 \sim \ldots \sim s_m = N$. Ce chemin $(s_0, \ldots, s_m)$ est nécessairement auto-évitant.
- Ou bien la durée du chemin $(s_n)$ est infinie. Dans ce cas on observe tout d'abord que le chemin *ne passe pas par* $N$. On voit ensuite que le chemin finit par être périodique, c'est-à-dire qu'il existe des entiers $k_0 \geq 0$ et $\ell \geq 2$ tels que

$(i)$ les sommets $s_0, \ldots, s_{k_0}$ sont distincts (c'est-à-dire que $(s_0, \ldots, s_{k_0})$ est auto-évitant),
$(ii)$ $\gamma = (s_{k_0}, s_{k_0+1} \ldots, s_{k_0+l})$ est un cycle,
$(iii)$ $s_{k_0+k} = s_{k_0+k+\ell}$ pour tout $k \in \mathbb{N}$.

En suivant les lignes de champ de $\overrightarrow{h}$ issues de chaque sommet $s \in \mathbf{S}\backslash\{N\}$, on obtient son *portrait de phase* qui se décompose en ls objets suivants.

$(a)$ Un ensemble (possiblement vide) de cycles $\mathtt{Cycles}(\overrightarrow{h}) = \{\gamma_1, \ldots, \gamma_p\}$. Ces cycles ont des supports disjoints deux-à-deux. Aucun cycle ne passe par $N$.

$(b)$ Un ensemble de lignes de champ coalescentes qui se terminent en $N$ : On note $S_N(\overrightarrow{h})$ l'ensemble des sommets $s \in \mathbf{S}\backslash\{N\}$ dont la ligne de champ est de durée finie et se termine donc en $N$ ; on pose aussi $T_N(\overrightarrow{h}) = \big\{\{s, \overrightarrow{h}(s)\}; s \in S_N(\overrightarrow{h})\big\}$. Alors on vérifie que $\big(S_N(\overrightarrow{h}) \cup \{N\}, T_N(\overrightarrow{h})\big)$ est un sous-graphe de $\mathbf{G}$ qui est connexe et sans cycle : c'est donc un sous-arbre de $\mathbf{G}$. On remarque ensuite que $S_N(\overrightarrow{h}) = \{1, \ldots, N-1\}$ si et seulement si $\overrightarrow{h}$ n'a pas de cycle. Autrement dit, $T_N(\overrightarrow{h})$ est un arbre couvrant de $\mathbf{G}$ si et seulement si $\mathtt{Cycles}(\overrightarrow{h}) = \emptyset$. On voit donc que

$$\overrightarrow{h} \mapsto T_N(\overrightarrow{h}) \text{ est une bijection entre les champs } \overrightarrow{h} \text{ tels que } \mathtt{Cycles}(\overrightarrow{h}) = \emptyset \text{ et } \mathbb{T}(\mathbf{G}). \qquad \text{(I.105)}$$

$(c)$ En chaque cycle $\gamma_j \in \mathtt{Cycles}(\overrightarrow{h})$ et en chaque sommet $s$ du support du cycle $\gamma_j$ on note $\mathbf{S}_s(\overrightarrow{h})$ l'ensemble (possiblement vide) des sommets qui n'appartiennent par au support du cycle $\gamma_j$ mais dont la ligne de champ passe par $s$. On note $T_s(\overrightarrow{h}) = \big\{\{s', \overrightarrow{h}(s')\}; s' \in S_s(\overrightarrow{h})\big\}$. Alors on vérifie que $\big(S_s(\overrightarrow{h}) \cup \{s\}, T_s(\overrightarrow{h})\big)$ est un sous-graphe de $\mathbf{G}$ qui est connexe et sans cycle : c'est donc un sous-arbre de $\mathbf{G}$ qui est le *bassin d'attraction de* $s$. On note qu'il peut être réduit au point $s$ et que dans ce cas $T_s(\overrightarrow{h})$ est vide.

On définit ensuite un poids pour chacun des types d'objets introduits précédemment :
– le *poids du champ* $\overrightarrow{h}$ par $\mathtt{poids}(\overrightarrow{h}) = \prod_{s \in \mathbf{S}\backslash\{N\}} C_{\{s, \overrightarrow{h(s)}\}}$/
– À un cycle $\gamma = (s_0, \ldots, s_\ell)$ de longueur $\ell$ on associe également le poids $\mathtt{poids}(\gamma) = \prod_{0 \leq k < \ell} C_{\{s_k, s_{k+1}\}}$, qui est le produit des poids des arêtes parcourues par le cycle jusqu'au retour au point de départ.
– Plus généralement si $A \subset \mathbf{A}$, on pose $\mathtt{poids}(A) = \prod_{a \in A} C_a$, avec la convention que $\mathtt{poids}(\emptyset) = 1$.

On observe alors que

$$\mathtt{poids}(\overrightarrow{h}) = \mathtt{poids}\big(T_N(\overrightarrow{h})\big) \prod_{\gamma \in \mathtt{Cycles}(\overrightarrow{h})} \Big(\mathtt{poids}(\gamma) \prod_{s \in \gamma} \mathtt{poids}\big(T_s(\overrightarrow{h})\big)\Big) . \qquad \text{(I.106)}$$

On développe ensuite le déterminant de $\Delta^{(N,N)}$ de la manière à faire apparaître des poids de champs. Plus précisément, on note $\mathfrak{S}_{N-1}$ l'ensemble des permutations de $\{1, \ldots, N-1\} = \mathbf{S}\backslash\{N\}$. On note la signature d'une permutation $\sigma \in \mathfrak{S}_{N-1}$ par $\varepsilon(\sigma)$ (qui est la parité du nombre de transpositions dont la composition fait $\sigma$). On a donc

$$\mathtt{det}(\Delta^{(N,N)}) = \sum_{\sigma \in \mathfrak{S}_{N-1}} \alpha(\sigma) \,, \text{ où on a posé } \alpha(\sigma) = \varepsilon(\sigma) \prod_{1 \leq s \leq N-1} [\Delta](s, \sigma(s)), \text{ pour toute } \sigma \in \mathfrak{S}_{N-1}.$$

Une permutation $\sigma \in \mathfrak{S}_{N-1}$ est ensuite dite $\mathbf{G}$-*compatible* si pour tout $s \in \{1, \ldots, N-1\}$ tel que $s \neq \sigma(s)$, on a $s \sim \sigma(s)$ (c'est-à-dire $\{s, \sigma(s)\} \in \mathbf{A}$ dès que $s \neq \sigma(s)$). On note l'ensemble des permutation $\mathbf{G}$-compatibles par $\mathfrak{S}_{\mathbf{G}}$ et on observe que si $\sigma$ n'est pas $\mathbf{G}$-compatible, alors $\alpha(\sigma) = 0$. Donc on a

$$\mathtt{det}(\Delta^{(N,N)}) = \sum_{\sigma \in \mathfrak{S}_{\mathbf{G}}} \alpha(\sigma) \;. \qquad \text{(I.107)}$$

On donne ensuite une description des permutation $\mathbf{G}$-compatible à l'aide de champs discrets de la manière suivante. On note $\mathscr{C}$ l'ensemble des ensembles des cycles de $\mathbf{G}$ notés génériquement $C = \{\gamma_1, \dots, \gamma_p\}$ tel que les cycles $\gamma_1, \dots, \gamma_p$ ont des supports qui ne *contiennent pas $N$ et qui sont deux-à-deux disjoints*. Attention, on veut inclure le cas où $C$ est vide. On note $\mathtt{Fix}(C)$ les sommets de $\{1, \dots, N-1\}$ qui ne sont dans aucun des supports des cycles $\gamma_1, \dots, \gamma_p$, c'est-à-dire $\mathtt{Fix}(C) = \{1, \dots, N-1\} \backslash (\mathtt{supp}(\gamma_1) \cup \dots \cup \mathtt{supp}(\gamma_p))$ ; notons que cet ensemble est possiblement vide. Notons également que si $C = \emptyset$, alors $\mathtt{Fix}(\emptyset) = \{1, \dots, N-1\}$.

À $C = \{\gamma_1, \dots, \gamma_p\} \in \mathscr{C}$, on associe une permutation $\mathbf{G}$-compatible $\sigma_C$ dont l'ensemble des points fixes est $\mathtt{Fix}(C)$ et dont les cycles de longueur $\geq 2$ sont les $\gamma_1, \dots, \gamma_p$, c'est-à-dire : $\sigma_C = \gamma_1 \dots \gamma_p \prod_{s \in \mathtt{Fix}(C)} (s)$, avec l'écriture habituelles en produit (pour la composition) de la décomposition des permutations en cycles à supports disjoints. On voit facilement tout d'abord que

$$C \in \mathscr{C} \longmapsto \sigma_C \in \mathfrak{S}_{\mathbf{G}} \text{ est une bijection.} \tag{I.108}$$

On fixe ensuite $C = \{\gamma_1, \dots, \gamma_p\} \in \mathscr{C}$. Si on note $\ell_m$ la longueur du cycle $\gamma_m$, alors la signature de $\sigma_C$ est $\varepsilon(\sigma_C) = (-1)^{\ell_1 - 1} \dots (-1)^{\ell_p - 1}$. On ensuite utilise les deux conventions suivantes : un produit sur un ensemble vide d'incides est pris égal à 1 et une somme sur un ensemble vide d'indices est prise égale à 0, et on remarque alors que

$$\begin{aligned}
\alpha(\sigma_C) &= \varepsilon(\sigma_C)\Big(\prod_{s \in \mathtt{Fix}(C)} [\Delta](s,s)\Big)\Big(\prod_{1 \leq m \leq p} \prod_{s \in \mathtt{supp}(\gamma_m)} [\Delta](s, \gamma_m(s))\Big) \\
&= \varepsilon(\sigma_C)\Big(\prod_{s \in \mathtt{Fix}(C)} \pi(s)\Big)\Big(\prod_{1 \leq m \leq p} (-1)^{\ell_m} \mathtt{poids}(\gamma_m)\Big) \\
&= (-1)^p \Big(\prod_{s \in \mathtt{Fix}(C)} \pi(s)\Big)\Big(\prod_{1 \leq m \leq p} \mathtt{poids}(\gamma_m)\Big) \\
&= (-1)^{\# C} \Big(\prod_{s \in \mathtt{Fix}(C)} \pi(s)\Big)\Big(\prod_{\gamma \in C} \mathtt{poids}(\gamma)\Big) \\
&= (-1)^{\# C} \Big(\prod_{s \in \mathtt{Fix}(C)} \Big(\sum_{1 \leq s' \leq N-1} \mathbf{1}_{\{s \sim s'\}} C_{\{s,s'\}}\Big)\Big)\Big(\prod_{\gamma \in C} \mathtt{poids}(\gamma)\Big) \\
&= (-1)^{\# C} \sum_{\overrightarrow{g}} \Big(\Big(\prod_{s \in \mathtt{Fix}(C)} C_{\{s, \overrightarrow{g}(s)\}}\Big)\Big(\prod_{\gamma \in C} \mathtt{poids}(\gamma)\Big)\Big),
\end{aligned}$$

où la sommation s'effectue sur *tous les champs partiels* qui sont les fonctions $\overrightarrow{g} : \mathtt{Fix}(C) \to \mathbf{S}$ telles que $s \sim \overrightarrow{g}(s)$ (c'est-à-dire que $\overrightarrow{g}$ est la restriction d'un champ discret à l'ensemble de sommets $\mathtt{Fix}(C)$). On observe ensuite que la donnée d'un champ partiel $\overrightarrow{g} : \mathtt{Fix}(C) \to \mathbf{S}$ et des cycles $\gamma \in C$ permet *d'étendre de manière unique* $\overrightarrow{g}$ en un champ discret $\overrightarrow{h} : \mathbf{S} \backslash \{N\} \to \mathbf{S}$ tel que $C \subset \mathtt{Cycles}(\overrightarrow{h})$. Réciproquement, si $\overrightarrow{h}$ est un champ sur $\mathbf{S}$ tel que $C \subset \mathtt{Cycles}(\overrightarrow{h})$ et si on note $\overrightarrow{g}$ sa restriction à $\mathtt{Fix}(C)$ alors $\mathtt{poids}(\overrightarrow{h}) = (\prod_{s \in \mathtt{Fix}(C)} C_{\{s, \overrightarrow{g}(s)\}})(\prod_{\gamma \in C} \mathtt{poids}(\gamma))$ on a donc

$$\alpha(\sigma_C) = \sum_{\overrightarrow{h} \in \Gamma} \mathbf{1}_{\{C \subset \mathtt{Cycles}(\overrightarrow{h})\}} (-1)^{\# C} \mathtt{poids}(\overrightarrow{h}) ,$$

où on rappelle que $\Gamma$ est l'ensemble des champs discrets sur $\mathbf{G}$. On observe que cette formule est vérifiée pour $C = \emptyset$. On a donc par (I.107) et (I.108)

$$\mathtt{det}(\Delta^{(N,N)}) = \sum_{C \in \mathscr{C}} \sum_{\overrightarrow{h} \in \Gamma} \mathbf{1}_{\{C \subset \mathtt{Cycles}(\overrightarrow{h})\}} (-1)^{\# C} \mathtt{poids}(\overrightarrow{h})$$

$$
\begin{aligned}
&= \sum_{\overrightarrow{h}\in\Gamma} \mathtt{poids}(\overrightarrow{h})\Big(\sum_{C\in\mathscr{C}} \mathbf{1}_{\{C\subset \mathtt{Cycles}(\overrightarrow{h})\}}(-1)^{\#C}\Big)\\
&= \sum_{\overrightarrow{h}\in\Gamma} \mathtt{poids}(\overrightarrow{h})\mathbf{1}_{\{\mathtt{Cycles}(\overrightarrow{h})=\emptyset\}} + \sum_{\overrightarrow{h}\in\Gamma} \mathtt{poids}(\overrightarrow{h})\mathbf{1}_{\{\mathtt{Cycles}(\overrightarrow{h})\neq\emptyset\}}\Big(\sum_{C\subset \mathtt{Cycles}(\overrightarrow{h})} (-1)^{\#C}\Big). \qquad \text{(I.109)}
\end{aligned}
$$

On rappelle ensuite le fait général élémentaire suivant : si $E$ est un ensemble non-vide, alors

$$
\begin{aligned}
\sum_{J\subset E}(-1)^{\#J} &= \sum_{0\leq k\leq \#E}\ \sum_{J\subset E:\#J=k}(-1)^k\\
&= \sum_{0\leq k\leq \#E}\binom{\#E}{k}(-1)^k = (1-1)^{\#E} = 0
\end{aligned}
$$

Cela implique donc que

$$
\begin{aligned}
\mathtt{det}(\Delta^{(N,N)}) &\overset{\text{par (I.109)}}{=} \sum_{\overrightarrow{h}\in\Gamma} \mathtt{poids}(\overrightarrow{h})\mathbf{1}_{\{\mathtt{Cycles}(\overrightarrow{h})=\emptyset\}}\\
&\overset{\text{par (I.106)}}{=} \sum_{\overrightarrow{h}\in\Gamma} \mathtt{poids}(T_N(\overrightarrow{h}))\mathbf{1}_{\{\mathtt{Cycles}(\overrightarrow{h})=\emptyset\}}\\
&\overset{\text{par (I.105)}}{=} \sum_{T\in\mathbb{T}(\mathbf{G})} \mathtt{poids}(T),
\end{aligned}
$$

ce qui prouve (I.104) et donc le théorème. ■

**Exemple I.6.2** (*arbres à $n$ sommets étiquetés (arbres de Cayley)*) Soit un entier $N \geq 2$. On considère l'ensemble $\mathbb{T}_N^{\mathsf{Cay}}$ des arbres à $N$ sommets, c'est-à-dire des graphes sans cycle, connexes et dont les sommets forment l'ensemble $\mathbf{S} = \{1,\dots,N\}$ (il n'y a pas de racine). Ces arbres peuvent se voir comme l'ensemble des arbres couvrants du *graphe complet d'ordre* $N$, noté $\mathbf{K}_N$ qui a $N$ sommets et qui a toutes les arêtes possibles (non-orientées, simples, sans boucle). C'est-à-dire que $\mathbf{K}_N = (\mathbf{S},\mathbf{A})$ où

$$
\mathbf{S} = \{1,\dots,N\}, \quad \text{et} \quad \mathbf{A} = \big\{\{s,s'\}; 1\leq s < s' \leq N\big\}\ .
$$

Il est clair que

$$
\mathbb{T}_N^{\mathsf{Cay}} = \mathbb{T}(\mathbf{K}_N)\ . \qquad \text{(I.110)}
$$

Cette interprétation permet d'appliquer le théorème de Kirchhoff pour (re)démontrer un résultat de Cayley affirmant que

$$
\#\mathbb{T}_N^{\mathsf{Cay}} = N^{N-2}\ . \qquad \text{(I.111)}
$$

*En effet,* pour cela on détermine le spectre de la matrice laplacienne $\Delta$ du graphe complet $\mathbf{K}_N$ où toutes les arêtes ont un même poids pris égal à 1. Il est facile de voir que $\Delta = N\mathtt{Id} - J$ où $J$ est la matrice dont toutes les entrées valent 1. Soit $v = (v_1,\dots,v_N) \in \mathbb{R}^N$ tel que $v_1 + \ldots + v_n = 0$ ; alors $J.v = 0$ et donc $\Delta v = Nv$. L'espace des vecteurs dont la somme des coordonnées est nulle est de dimension $N-1$ (c'est un hyperplan). Par conséquent, en plus de 0, l'autre valeur propre de $\Delta$ est $N$ et elle a une multiplicité de $N-1$. Dans la formule (I.103) du théorème I.6.1 de Kirchhoff, page 133, on a donc $\#\mathbb{T}(\mathbf{K}_N) = \#Z(\mathbf{K}_N,\mathbf{1}) = \frac{1}{N}\lambda_1\ldots\lambda_{N-1} = N^{N-2}$, qui est le résultat voulu. □

### I.6.b Algorithme d'Aldous-Broder.

Dans la section suivante, on détaille un algorithme permettant d'obtenir un arbre couvrant aléatoire dont la loi est donnée par (I.101), algorithme obtenu indépendamment par Aldous et Broder en 1989-90. Une des idées sur lesquelles repose cet algorithme est l'étude dynamique de l'*arbre de dernier passage* introduit dans l'exemple I.5.1, page 106.

Plus précisément, fixons un graphe $\mathbf{G}(\mathbf{S}, \mathbf{A})$ supposé simple connexe et fini et que l'on munit de poids $\mathbf{C} = (C_a)_{a \in \mathbf{A}}$ strictement positifs. On présente tout d'abord quelques propriétés déterministes de l'arbre de dernier passage. Pour cela, cela on fixe un chemin $\mathbf{x} = (x(n))_{n \in \mathbb{Z}}$ dans le graphe (on remarquera que son indexation est bi-infinie) et on suppose que

$$\forall n \in \mathbb{Z}, \quad \{x(k); k \leq n\} = \{x(k); k \geq n\} = \mathbf{S}\,,$$

c'est-à-dire que *le chemin a déjà visité tous les sommets de* $\mathbf{G}$ *à l'instant* $n$ *et visitera tous les sommets de* $\mathbf{G}$ *après l'instant* $n$. On peut donc définir pour tout sommet $s \in \mathbf{S}$ son *dernier instant de passage en* $s$ *avant l'instant* $n$ (l'instant $n$inclus) et son *premier temps de passage après l'instant* $n$ par

$$\ell_s(n) = \max\big\{k \leq n : x(k) = s\big\} \quad \text{et} \quad \sigma_s(n) = \inf\big\{k \geq n : x(k) = s\big\} \tag{I.112}$$

et définir l'*arbre de dernier passage au temps* $n$ *associé à* $\mathbf{x}$ ainsi que *son arbre de premier temps de passage au temps* $n$ par

$$T_n(\mathbf{x}) = \big\{\{s, x(\ell_s(n){+}1)\}\,;\ s \in \mathbf{S}\backslash\{x(n)\}\big\} \quad \text{et} \quad \widehat{T}_n(\mathbf{x}) = \big\{\{s, x(\sigma_s(n){-}1)\}\,;\ s \in \mathbf{S}\backslash\{x(n)\}\big\} \tag{I.113}$$

On remarque que

$$T_0(\widehat{\mathbf{x}}) = \widehat{T}_0(\mathbf{x}) \quad \text{et} \quad \widehat{T}_0(\widehat{\mathbf{x}}) = T_0(\mathbf{x}) \quad \text{où} \quad \widehat{\mathbf{x}} = (x(-n))_{n \in \mathbb{Z}}. \tag{I.114}$$

De plus, $T_0(\mathbf{x})$ *et* $\widehat{T}_0(\mathbf{x})$ *sont deux arbres couvrants de* $\mathbf{G}$.

**Preuve.** Par (I.114) il suffit de montrer que $T_n(\mathbf{x})$ est un arbre couvrant. Raisonnons par l'absurde en supposant que le graphe dont les sommets sont $\mathbf{S}$ et les arêtes $T_n(\mathbf{x})$. On constate évidemment que c'est un sous-graphe de $\mathbf{G}$. Supposons que $(s_0, \ldots, s_m)$ soit un cycle de ce graphe (c'est-à-dire que $m \geq 3$, que $\{s_k, s_{k+1}\} \in T_n(\mathbf{x})$ pour tous $0 \leq k < m$ et que les sommets $s_0, \ldots, s_{m-1}$ soient distincts). Sans perte de généralité, on peut supposer que l'indexation du cycle est telle que $\ell_{s_0}(n) = \min\{\ell_{s_k}(n); 0 \leq k < m\}$. Par définition $\{s_0, s_1\} \in T_n(\mathbf{x})$ et l'un des deux sommets $s_0$ ou $s_1$ est distinct de $x(n)$.

– Si $s_1 \neq x(n)$ alors $s_0 = x(\ell_{s_1}(n){+}1)$ et $\mathbf{x}$ visite $s_0$ après le temps $\ell_{s_1}(n)$. Cela implique $\ell_{s_0}(n) > \ell_{s_1}(n)$ ce qui est contradictoire.
– Si $s_0 \neq x(n)$ alors $s_1 = x(\ell_{s_0}(n) + 1)$. Donc $s_{m-1} = x(\ell_{s_0}(n) + 1)$, ce qui implique $s_1 = s_{m-1}$ or cela contredit $m \geq 3$.

Dans les deux cas, on obtient une contradiction. Le graphe de sommets $\mathbf{S}$ et d'arêtes $T_n(\mathbf{x})$ est donc sans cycle : c'est donc un arbre couvrant. ■

**Transitions de $T_n(\mathbf{x})$ à $T_{n+1}(\mathbf{x})$.** Comme $\mathbf{x}$ est un chemin, $\{x(n), x(n+1)\}$ est une arête du graphe. On observe tout d'abord que $\ell_{x(n)}(n+1) = n$ et donc $\{x(n), x(n+1)\} \in T_{n+1}(\mathbf{x})$. Pour tout $s \in \mathbf{S}\backslash\{x(n+1)\}$, on a $\ell_s(n) = \ell_s(n+1) < n+1$ et donc $\{s, x(\ell_s(n+1)+1)\} = \{s, x(\ell_s(n)+1)\}$. On pose $s' = x(n+1)$ pour simplifier les notations et on observe que $\{s', x(\ell_{s'}(n){+}1)\}$ est une arête présente dans $T_n(\mathbf{x})$ mais qu'elle a été retirée de $T_{n+1}(\mathbf{x})$. Cela suggère le mécanisme d'évolution suivant.

**Définition I.6.3** Soient $s, s' \in \mathbf{S}$ deux sommets adjacents : $\{s, s'\} \in \mathbf{A}$. Soit $T \in \mathbb{T}(\mathbf{G})$. À partir de $(s, T)$ et $s'$ on associe deux nouveaux arbres couvrants $T'$ et $\widehat{T}'$ de la manière suivante.

– Si $\{s, s'\} \in T$ alors $T = T' = \widehat{T}'$.
– Si $\{s, s'\} \notin T$, alors dans le sous-graphe de sommets $\mathbf{S}$ et d'arêtes $T \cup \big\{\{s, s'\}\big\}$, il y a la présence d'un seul cycle $s_0 = s \sim s_1 = s' \sim s_2 \sim \ldots \sim s_{m-1} \sim s_m = s$. Comme $m \geq 3$, $\{s', s_2\} \in T$ et $\{s_{m-1}, s\} \in T$ nécessairement on pose alors

$$T' = \big\{\{s, s'\}\big\} \cup \big(T \backslash \big\{\{s', s_2\}\big\}\big) \quad \text{et} \quad \widehat{T}' = \big\{\{s, s'\}\big\} \cup \big(T \backslash \big\{\{s_{m-1}, s\}\big\}\big)$$

On observe que $T'$ et $\widehat{T}'$ sont des arbres couvrants de $\mathbf{G}$. On utilise les notations

$$\mathtt{Last}\big(s'; (s, T)\big) = (s', T') \quad \text{et} \quad \mathtt{First}\big(s'; (s, T)\big) = (s', \widehat{T}')$$

qui sont appelés les *mécanismes d'évolutions de respectivement dernier et premier passage* des arbres couvrants enracinés. □

En gardant les notations $T_n(\mathbf{x})$ et $\widehat{T}_n(\mathbf{x})$ pour les arbres de dernier passage et de premier passage du chemin $\mathbf{x} = (x(k))_{k \in \mathbb{Z}}$ au temps $n$, on voit que pour tout $n \in \mathbb{Z}$,

$$\begin{aligned}
\mathtt{Last}\Big(x(n+1); \big(x(n), T_n(\mathbf{x})\big)\Big) &= \big(x(n+1), T_{n+1}(\mathbf{x})\big) \quad \text{et} \\
\mathtt{First}\Big(x(n+1); \big(x(n), \widehat{T}_n(\mathbf{x})\big)\Big) &= \big(x(n+1), \widehat{T}_{n+1}(\mathbf{x})\big)
\end{aligned} \tag{I.115}$$

Lorsque la trajectoire $\mathbf{x}$ est une marche aléatoire simple sur le graphe pondéré, cela induit une matrice de transition sur l'espace des arbres couvrants enracinés, c'est-à-dire des paires $(s, T) \in \mathbf{S} \times \mathbb{T}(\mathbf{S})$ où le sommet $s$ est vu comme une racine de $T$. Rappelons que la matrice de transition $Q = (p(s, s'))_{s, s' \in \mathbf{S}}$ de la marche sur le graphe pondéré $(\mathbf{G}, \mathbf{C})$ est donnée par

$$p(s, s') = \mathbf{1}_{\{s \sim s'\}} \frac{C_{\{s, s'\}}}{\pi(s)}$$

où $\pi(s) = \sum_{s' \in \mathbf{S} : s' \sim s} C_{\{s, s'\}}$. On rappelle que $\pi$ est une mesure de réversibilité : comme le graphe est connexe et fini, la marche est irreductible récurrente positive et sa loi stationnaire est donnée par

$$\forall s \in \mathbf{S}, \quad \varpi(s) = \frac{\pi(s)}{\langle \pi \rangle} \text{ où } \langle \pi \rangle = \sum_{s' \in \mathbf{S}} \pi(s') \ . \tag{I.116}$$

On définit alors une matrice $\overline{Q} = (\overline{p}((s, T), (s', T')))_{(s,T),(s',T') \in \mathbf{S} \times \mathbb{T}(\mathbf{G})}$ en posant

$$\overline{p}\big((s, T), (s', T')\big) = \begin{cases} p(s, s') & \text{si } s \sim s' \text{ et si } \mathtt{Last}\big(s', (s, T)\big) = (s', T'), \\ 0 & \text{sinon } s \neq s'. \end{cases} \tag{I.117}$$

**Lemme I.6.4** *On reprend les notations et les hypothèses précédentes. Alors $\overline{Q}$ est une matrice de transition sur $\mathbf{S} \times \mathbb{T}(\mathbf{G})$ qui est irréductible et de loi stationnaire donnée par*

$$\overline{\pi}(s, T) = \varpi(s) \frac{\mathtt{poids}(T)}{Z(\mathbf{G}, \mathbf{C})} \ , \tag{I.118}$$

*où on rappelle la définition (I.116) de $\varpi$ et la définition (I.100) de $\mathtt{poids}(T)$ et la définition (I.101) de $Z(\mathbf{G}, \mathbf{C})$.*

**Preuve :** montrons d'abord l'irréductibilité. On fixe $(s_0, T_0)$ et $(s, T)$ dans $\mathbf{S} \times \mathbb{T}(\mathbf{G})$. On note $y_1, \ldots, y_p$ les *feuilles* de $T$, c'est-à-dire les sommets distincts de $s$ qui ont un degré égal à 1 dans l'arbre couvrant $T$. Comme $T$ est un sous-graphe connexe, il existe un chemin $(x_n)_{0 \leq n \leq N}$ qui n'emprunte que des arêtes de $T$ et qui effectue un tour commençant en $x_0 = s_0$ et qui passe successivement par $s$ puis $y_1$ puis par $s$ puis par $y_2$, ..., puis par $s$ puis par $y_{p-1}$ puis par $s$ puis par $y_p$ puis par $s$ où il s'arrête. On définit ensuite une suite d'arbres couvrants $(T_n)_{0 \leq n \leq N}$ par récurrence en posant $(x_{n+1}, T_{n+1}) = \mathtt{Last}(x_{n+1}, (x_n, T_n))$, $0 \leq n < N$. On voit tout d'abord que

$$\prod_{0 \leq n < N} \overline{p}\big((x_n, T_n), (x_{n+1}, T_{n+1})\big) = \prod_{0 \leq n < N} p(x_n, x_{n+1}) > 0 \, .$$

Il est clair que $\{x_n\, ;\ 0 \leq n < N\} = \mathbf{S}$ et par conséquent $T_N$ est l'arbre de dernier passage du chemin $(x_n)_{0 \leq n < N}$ : comme ce chemin n'emprunte que des arêtes de $T$, un sommet distinct de $s$ lorsqu'il est visité une dernière fois est quitté par l'arête qui pointe vers la racine $s$. Cela implique que $T_N = T$ et cela prouve que $\overline{Q}$ est irréductible. .

Montrons ensuite que $\overline{\pi}$ est $\overline{Q}$-invariante. On voit que $\overline{Q}$ *n'est pas réversible*. Il faut donc décrire toutes les transitions possibles arrivant en un arbre couvrant enraciné donné $(s', T') \in \mathbf{S} \times \mathbb{T}(\mathbf{G})$. On remarque qu'il existe une bijection $\phi$ entre l'ensemble $B_1$ des sommets $s'' \in \mathbf{S}$ voisins de $s'$ et l'ensemble $B_2$ des $(s, T) \in \mathbf{S} \times \mathbb{T}(\mathbf{G})$ tels que $\mathtt{Last}(s', (s, T)) = (s', T')$. Cette bijection se construit comme suit. Soit $s''$ voisin de $s'$ dans $\mathbf{G}$ ; si $\{s'', s'\} \in T'$ alors on pose $\phi(s'') = (s'', T')$ et on vérifie bien dans ce cas que $\mathtt{Last}(s', (s'', T')) = (s', T')$. Si $\{s'', s'\} \notin T'$, alors l'ajout de l'arête $\{s', s''\}$ à $T'$ créée un cycle dans ce nouveau graphe dont les arêtes sont $T' \cup \{\{s', s''\}\}$, cycle qui est noté $s_0 = s' \sim s_1 = s'' \sim \ldots s_{m-1} \sim s'$ avec $m \geq 3$. On a donc $\{s_{m-1}, s'\} \in T'$ (car $\{s_{m-1}, s'\}$ doit être une arête de $T' \cup \{\{s', s''\}\}$ et $s_{m-1} \neq s''$ puisque $m \geq 3$). On pose alors $\phi(s'') = (s, T)$ où $s = s_{m-1}$ et $T = (T' \backslash \{\{s_{m-1}, s'\}\}) \cup \{\{s', s''\}\}$. On vérifie bien que $\mathtt{Last}(s', (s, T)) = (s', T')$. Donc on a construit une fonction de $B_1$ dans $B_2$. On vérifie facilement que c'est une bijection en construisant de façon similaire explicitement son inverse $\phi^{-1}$. On remarque que pour tout $(s, T) \in B_2$, si on pose $s'' = \phi^{-1}(s, T)$ alors

$$\mathtt{poids}(T) = \frac{C_{\{s, s''\}}}{C_{\{s', s\}}} \mathtt{poids}(T') = \frac{C_{\{s, \phi^{-1}(s,T)\}}}{C_{\{s', s\}}} \mathtt{poids}(T')$$

On a donc pour tout $(s', T') \in \mathbf{S} \times \mathbb{T}(\mathbf{G})$,

$$\begin{aligned}
\langle \pi \rangle Z(\mathbf{G}, \mathbf{C}) \sum_{(s,T) \in B_2} \overline{\pi}(s, T) \overline{p}\big((s, T), (s', T')\big) &= \sum_{(s,T) \in B_2} \mathtt{poids}(T) \pi(s) p(s, s') \\
&= \sum_{(s,T) \in B_2} \frac{C_{\{s, \phi^{-1}(s,T)\}}}{C_{\{s', s\}}} \mathtt{poids}(T') \pi(s) p(s, s') = \sum_{(s,T) \in B_2} \frac{C_{\{s, \phi^{-1}(s,T)\}}}{C_{\{s', s\}}} \mathtt{poids}(T') C_{\{s', s\}} \\
&= \mathtt{poids}(T') \sum_{(s,T) \in B_2} C_{\{s, \phi^{-1}(s,T)\}} = \mathtt{poids}(T') \sum_{s'' \in \mathbf{S} : s'' \sim s'} C_{\{s, s''\}} = \langle \pi \rangle Z(\mathbf{G}, \mathbf{C}) \overline{\pi}(s', T'),
\end{aligned}$$

ce qui montre que $\overline{\pi}$ est $\overline{Q}$ invariante. Une matrice de transition irréductible sur un espace fini ne peut avoir qu'une seule loi invariante. Cela termine la preuve du lemme. ■

On utilise ensuite la proposition I.2.32 page 50 qui donne une construction d'une version stationnaire de la marche aléatoire simple sur le graphe $\mathbf{G}$ muni des poids $C_a$, $a \in \mathbf{A}$ : c'est-à-dire, une famille $(X_n)_{n \in \mathbb{Z}}$ de variables telles que pour tout $n_0 \in \mathbb{Z}$,

$(a)$ $X_{n_0}$ a pour loi $\varpi$, la loi invariante de la marche.

$(b)$ $(X_{n_0+n})_{n \in \mathbb{Z}}$ a même loi que $(X_{n_0-n})_{n \in \mathbb{Z}}$.

$(c)$ sous $\mathbf{P}(\,\cdot\,|\,X_{n_0}=i)$, $(X_{n_0+n})_{n\in\mathbb{N}}$ et $(X_{n_0-n})_{n\in\mathbb{N}}$ sont deux chaînes de Markov qui sont indépendantes et de même loi.

Il est clair également que $\mathbf{P}$-p.s. $\{X_k\,;\,k\leq n_0\}=\{X_k\,;\,k\geq n_0\}=\mathbf{S}$

**Théorème I.6.5 (Aldous-Broder)** *Soit un espace de probabilité $(\Omega,\mathscr{F},\mathbf{P})$ sur lequel est définie $\mathbf{X}=(X_n)_{n\in\mathbb{Z}}$ une marche stationnaire sur le graphe pondéré $(\mathbf{G},\mathbf{C})$ comme ci-dessus. Pour tout $s\in\mathbf{S}$, on pose*

$$\ell_s=\max\{k\leq 0: X_k=s\}\quad \textit{et}\quad \boldsymbol{\sigma}(s)=\min\{k\geq 0: X_k=s\}$$

*qui sont respectivement le dernier temps de passage en $s$ avant $0$ et le premier temps de passage après $0$ de la marche stationnaire. On définit également les deux arbres couvrants suivants.*

$$\mathcal{T}=\big\{\{s,X_{\ell_s+1}\};s\in\mathbf{S}\backslash\{X_0\}\big\}\quad \textit{et}\quad \widehat{\mathcal{T}}=\big\{\{s,X_{\boldsymbol{\sigma}(s)-1}\}\,;\,s\in\mathbf{S}\backslash\{X_0\}\big\}.$$

*Alors $\mathcal{T}$ et $\widehat{\mathcal{T}}$ ont la même loi donnée pour tout $(s,T)\in\mathbf{S}\times\mathbb{T}(\mathbf{G})$ par*

$$\mathbf{P}(X_0=s;\widehat{\mathcal{T}}=T)=\varpi(s)\frac{\texttt{poids}(T)}{Z(\mathbf{G},\mathbf{C})}\quad \textit{et donc}\quad \mathbf{P}(\widehat{\mathcal{T}}=T|X_0=s)=\frac{\texttt{poids}(T)}{Z(\mathbf{G},\mathbf{C})}.$$

**Preuve :** pour tout $n\in\mathbb{Z}$, on note $T_n(\mathbf{X})$ l'arbre de dernier passage avant $n$ comme défini aux (I.113) et (I.112). On voit donc que $\mathcal{T}=T_0(\mathbf{X})$. Par (I.115), on voit que $(X_{n+1},T_{n+1}(\mathbf{X}))=\texttt{Last}\big(X_{n+1},(X_n,T_n(\mathbf{X}))\big)$. Si on note $\mathscr{F}_n$ la filtration engendrée par les $X_k$, $k\in\mathbb{Z}$ tels que $k\leq n$, alors

$$\forall(s,T)\in\mathbf{S}\times\mathbb{T}(\mathbf{G}),\quad \mathbf{E}\big[\mathbf{1}_{\{(X_{n+1},T_{n+1}(\mathbf{X}))=(s,T)\}}\,\big|\,\mathscr{F}_n\big]=\overline{p}\big((X_n,T_n(\mathbf{X})),(s,T)\big)\ ,$$

Donc $((X_n,T_n(\mathbf{X})))_{n\in\mathbb{N}}$ est une chaîne de Markov sur $\mathbf{S}\times\mathbb{T}(\mathbf{G})$ ayant pour matrice de transition $\overline{Q}$ qui est définie par (I.117). On pose ensuite $\mathbf{X}'=(X_{n+1})_{n\in\mathbb{Z}}$. Il est clair que $\mathbf{X}'$ et $\mathbf{X}$ ont même loi or on voit que $(X_{n+1},T_{n+1}(\mathbf{X}))=(X'_n,T_n(\mathbf{X}'))$ : on en déduit que la loi de $(X_n,T_n(\mathbf{X}))$ ne dépend pas de $n\in\mathbb{Z}$ (la suite est stationnaire en loi). Puisque la mesure $\overline{\pi}$ sur $\mathbf{S}\times\mathbb{T}(\mathbf{G})$ définie par (I.118) est l'unique loi stationnaire de la matrice de transition irréductible $\overline{Q}$, la loi commune à tous les $(X_n,T_n(\mathbf{X}))$ est $\overline{\pi}$. On observe ensuite que $\widehat{\mathcal{T}}=\widehat{T}_0(\mathbf{X})$, qui est l'arbre de premier passage de la marche $\mathbf{X}=(X_n)_{n\in\mathbb{Z}}$. Par la remarque (I.114), on voit que $\widehat{T}_0(\mathbf{X})=T_0(\widehat{\mathbf{X}})$ où $\widehat{\mathbf{X}}=(X_{-n})_{n\in\mathbb{Z}}$. Comme $\widehat{\mathbf{X}}$ et $\mathbf{X}$ ont même loi, on en déduit que $(X_0,\mathcal{T})$ et $(X_0,\widehat{\mathcal{T}})$ ont même loi, ce qui implique aisément le reste du théorème. ■

Dans l'énoncé du théorème d'Aldous-Broder, les temps de premier passage en chaque sommet peuvent être pris séquentiellement dans l'*ordre de découverte de nouveaux sommets* par la marche $X$. Cela conduit à l'algorithme suivant.

**Algorithme d'Aldous-Broder.** On suppose que sur $(\Omega,\mathscr{F},\mathbf{P})$, on puisse définir (c'est-à-dire que l'on puisse simuler jusqu'à des temps suffisamment grands) une marche aléatoire simple $(X_n)_{n\in\mathbb{N}}$ sur le graphe pondéré $(\mathbf{G},\mathbf{C})$. Alors on construit récursivement une suite d'ensemble de sommets $(S_n)$ et d'arêtes $(T_n)$ de la manière suivante.

- À $n=0$, on pose $S_0=\{X_0\}$ et $T_0=\emptyset$.
- Si $X_{n+1}\in S_n$, alors $S_{n+1}=S_n$ et $T_{n+1}=T_n$
- Si $X_{n+1}\notin S_n$, alors on pose $S_{n+1}=S_n\cup\{X_{n+1}\}$ et $T_{n+1}=T_n\cup\big\{\{X_n,X_{n+1}\}\big\}$
- L'algorithme s'arrête à l'instant aléatoire $\mathtt{t_{couv}}=\min\{n\in\mathbb{N}: S_n=\mathbf{S}\}$ et il retourne $\mathcal{T}:=T_{\mathtt{t_{couv}}}$. On a

$$\forall T\in\mathbb{T}(\mathbf{G}),\quad \mathbf{P}(\mathcal{T}=T)=\frac{\texttt{poids}(T)}{Z(\mathbf{G},\mathbf{C})}\ .$$

Le temps $\mathtt{t_{couv}}$ appelé *temps de couverture* du graphe pondéré $(\mathbf{G},\mathbf{C})$. □

**Simulation de l'arbre uniforme à $n$ sommets**

Nous expliquons ci-dessous une modification (due à Aldous) de l'algorithme d'Aldous-Broder dans le cas du graphe complet qui permet de construire (et de simuler efficacement) un arbre uniforme à $N$ sommets.

Soit un entier $N \geq 2$. On rappelle l'interprétation (I.110) (dans l'exemple I.6.b, page 141) d'un arbre à $N$ sommets étiquetés (un arbre dit de *Cayley*) comme un arbre couvrant du graphe complet $\mathbf{K}_N$ : l'ensemble des sommets du le graphe complet est $\mathbf{S} = \{1, \ldots, N\}$ et l'ensemble de ses arêtes sont toutes les arêtes possibles (boucles exceptées), c'est-à-dire que l'ensemble des arêtes du graphe complet est $\mathbf{A} = \big\{\{s, s'\}; 1 \leq s < s' \leq N\big\}$. On munit $\mathbf{K}_N$ de poids tous égaux à 1. Dans ce cas, la marche simple uniforme sur $\mathbf{K}_N$ est particulièrement simple : $p(s, s') = \frac{1}{N-1}$ pour tous $s, s' \in \{1, \ldots, N\}$ distincts : c'est-à-dire que si la marche se trouve en $s$ à l'instant $n$, on obtient la position suivante de la marche en tirant uniformément au hasard un entier de $\{1, \ldots, N\}$ distinct de $s$.

On peut générer la marche simple $(X_n)_{n\in\mathbb{N}}$ sur $\mathbf{K}_N$ à partir d'une suite $(U_p)_{p\in\mathbb{N}}$ i.i.d. de variables aléatoires uniformes sur $\{1, \ldots, N\}$. Plus précisément, on définit récursivement les $(X_n)$ et des indices aléatoires $(\mathbf{p}_n)_{n\geq 0}$ de la manière suivante.

$-$ Initialement on dispose de $X_0 = U_0$ (par exemple).

$-$ On pose $\mathbf{p}_1 = \min\{p \geq 1 : U_p \neq X_0\}$ et $X_1 := U_{\mathbf{p}_1}$.

$-$ Supposons construites les variables $X_0, \ldots, X_n$ et $\mathbf{p}_1, \ldots, \mathbf{p}_n$. On pose alors

$$\mathbf{p}_{n+1} = \min\{p > \mathbf{p}_n : U_p \neq X_n\} \quad \text{et} \quad X_{n+1} := U_{\mathbf{p}_{n+1}} .$$

Dans l'algorithme d'Aldous-Broder à partir de la marche $(X_n)$ ainsi construite, on voit que $S_n = \{X_0\} \cup \{U_1, U_2, \ldots, U_{\mathbf{p}_n}\}$ et qu'une arête nouvelle est ajoutée si et seulement si $U_{\mathbf{p}_{n+1}} \notin S_n$ : dans ce cas l'arête ajoutée est en fait $\{U_{\mathbf{p}_{n+1}-1}, U_{\mathbf{p}_{n+1}}\}$ car pour tout $k$ tel que $\mathbf{p}_n \leq k < \mathbf{p}_{n+1}$, $U_k = U_{\mathbf{p}_n}$, par définition de $\mathbf{p}_{n+1}$. Dans l'algorithme d'Aldous-Broder, un sommet nouveau exploré par la marche correspond donc exactement à un sommet nouveau découvert par la suite $(U_p)$ de v.a. uniformes sur $\{1, \ldots, N\}$. Par conséquent, l'algorithme peut se reformuler comme suit.

**Aldous-Broder pour l'arbre uniforme à $N$ sommets (1ère modification).** On suppose que l'on dispose d'une suite $(U_p)_{p\in\mathbb{N}}$ i.i.d. de variables aléatoires uniformes sur $\{1, \ldots, N\}$.

- À $n = 0$, on pose $S_0 = \{U_0\}$ et $T_0 = \emptyset$.
- Si $U_{p+1} \in S_p$, alors $S_{p+1} = S_p$ et $T_{p+1} = T_p$.
- Si $U_{p+1} \notin S_p$, alors on pose $S_{p+1} = S_p \cup \{U_{p+1}\}$ et $T_{p+1} = T_p \cup \big\{\{U_p, U_{p+1}\}\big\}$.
- L'algorithme s'arrête à l'instant aléatoire $\mathbf{t}_{\mathtt{couv}} = \min\{p \in \mathbb{N} : S_p = \mathbf{S}\}$ et il retourne $\mathcal{T} := T_{\mathbf{t}_{\mathtt{couv}}}$. On a

$$\forall T \in \mathbb{T}_N^{\mathtt{Cay}}, \quad \mathbf{P}(\mathcal{T} = T) = \frac{1}{\#\mathbb{T}_N^{\mathtt{Cay}}} = N^{-(N-2)} .$$

On voit que $\mathbf{t}_{\mathtt{couv}}$ (qui dépend bien entendu de $N$ même si cela n'apparaît pas dans la notation) correspond au problème du *collectionneur de coupons* (voir l'exercice I.2, page 7) pour lequel on montre facilement que

$$\mathbf{E}[\mathbf{t}_{\mathtt{couv}}] \sim_{N\to\infty} N \log N \quad \text{et} \quad \sqrt{\mathbf{var}(\mathbf{t}_{\mathtt{couv}})} \sim_{N\to\infty} \tfrac{\pi}{\sqrt{6}} N .$$

ce qui implique $\mathbf{t}_{\mathtt{couv}}/(N \log N) \to 1$ en probabilité. Pour simuler un arbre uniforme à 20 000 (et les 20 000 étiquettes !) uniforme il faudra appeler de l'ordre de 200 000 uniformes.

Une partie de l'algorithme est utilisée pour bien mélanger les étiquettes : si on ne s'interesse qu'à la *forme* de l'arbre $\mathcal{T}$ et non au numero porté par chaque sommet, les symétries de l'algorithme suggèrent des simplifications. Précisons tout d'abord ce que l'on entend par la « forme » de $\mathcal{T}$ : entre chaque sommet $s, s' \in \mathcal{T}$ distincts, il existe un chemin utilisant un nombre minimal d'arêtes les reliant ; on note $d_{\mathtt{gr}}(s, s')$ ce nombre d'arêtes, c'est-à-dire la longueur du chemin le plus court joignant dans $\mathcal{T}$ le sommet $s$ au sommet $s'$. On vérifie que $d_{\mathtt{gr}}$ est une distance sur $\mathcal{T}$. Par « forme » de $\mathcal{T}$, on veut dire qu'on envisage l'espace métrique $(\mathcal{T}, d_{\mathtt{gr}})$ à isométrie près et en particulier à *ré-étiquetage* des sommets près.

Précisons ce que nous entendons par *ré-étiquetage* : soit $T \in \mathbb{T}_N^{\mathtt{Cay}}$ un arbre dont les sommets sont $\{1, \ldots, N\}$ ; soit $\sigma \in \mathfrak{S}_N$, une permutation de $\{1, \ldots, N\}$. L'arbre $T$ *ré-tiqueté par* $\sigma$ est donné par

$$T^{(\sigma)} = \big\{ \{\sigma(s), \sigma(s')\} \, ; \; s, s' \in \{1, \ldots, N\} : \{s, s'\} \in T \big\} \, .$$

On voit que $\sigma : T \to T^{(\sigma)}$ est une isométrie de graphe, c'est-à-dire une isométrie des espaces métriques $(T, d_{\mathtt{gr}})$ et $(T^{(\sigma)}, d_{\mathtt{gr}})$, c'est-à-dire que pour toute paire de sommets $s, s' \in \{1, \ldots, N\}$, $d_{\mathtt{gr}}(s, s') = d_{\mathtt{gr}}(\sigma(s), \sigma(s'))$. Si on ne s'intéresse qu'à la forme d'un arbre uniforme, l'algorithme précédent se simplifie et conduit à la proposition suivante due à Aldous.

**Proposition I.6.6 (Algorithme d'Aldous pour les arbres de Cayley)** *Soit un espace de probabilité* $(\Omega, \mathscr{F}, \mathbf{P})$ *sur lequel est définie une suite* $(V_k)_{k \geq 1}$ *de v.a. indépendantes de loi uniforme sur* $\{1, \ldots, N\}$. *Soit* $\boldsymbol{\sigma} : \Omega \to \mathfrak{S}_N$*, une permutation aléatoire de* $\{1, \ldots, N\}$ *de loi uniforme, indépendante de* $(V_k)_{k \geq 1}$. *On définit l'arbre de sommets* $\{1, \ldots, N\}$ *et d'arêtes*

$$\mathcal{T} = \big\{ \{k, \min(V_k, k-1)\} \, ; \; 2 \leq k \leq N \big\} \, .$$

*Alors l'arbre ré-étiqueté* $\mathcal{T}^{(\boldsymbol{\sigma})}$ *est de loi uniforme sur l'ensemble* $\mathbb{T}_N^{\mathtt{Cay}}$ *des arbres de Cayley à* $N$ *sommets.*

**Preuve.** Soient $(U_p)_{p \in \mathbb{N}}$ des variables i.i.d. uniformes sur $\{1, \ldots, N\}$. Pour tout entier $k \in \{1, \ldots, N\}$, on note

$$r_k = \min\{p \in \mathbb{N} : \#\{U_0, U_1, \ldots, U_p\} = k\}.$$

On observe que $r_1 = 0$. Les temps $r_1, \ldots, r_N$ sont les temps où la suite $(U_p)$ « découvre » un nouveau nombre de $\{1, \ldots, N\}$. L'algorithme d'Aldous Broder modifié montre que le graphe dont les sommets sont $\{1, \ldots, N\}$ et les arêtes

$$\mathcal{T}_* := \big\{ \{U_{r_k - 1}, U_{r_k}\} ; 2 \leq k \leq N \big\}$$

est un arbre et que sa loi est uniforme sur l'ensemble $\mathbb{T}_N^{\mathtt{Cay}}$ des arbres de Cayley à $N$ sommets.

On suppose que l'on dispose également d'une suite de v.a. auxiliaires notées $(J_k)_{1 \leq k \leq N}$ qui sont d'une part *indépendantes de* $(U_p)$ mais aussi *indépendantes entre elles* et telles que $J_k$ *soit de loi uniforme sur l'ensemble des entiers successifs* $\{k, k+1, \ldots, N\}$.

À l'aide des $(U_p)$ et des $(J_k)_{1 \leq k \leq N}$ on construit les v.a. $(V_k)_{1 \leq k \leq N}$ et la permutation aléatoire $\boldsymbol{\sigma}$ de l'énoncé de la manière suivante. On observe tout d'abord que les nombres « découverts » par $(U_p)$, à savoir $U_{r_1}, \ldots, U_{r_N}$, sont par définition *distincts* et si on pose

$$\forall k \in \{1, \ldots, N\}, \quad \boldsymbol{\sigma}(k) := U_{r_k} \, ,$$

on définit bien une *permutation* de $\{1, \ldots, N\}$ ; nous montrons plus loin dans la preuve qu'elle est de loi uniforme sur $\mathfrak{S}_N$. On observe également que

$$\forall k \in \{1, \ldots, N\}, \quad \{\boldsymbol{\sigma}(1), \boldsymbol{\sigma}(2), \ldots, \boldsymbol{\sigma}(k)\} = \{U_0, U_1, \ldots, U_{r_k}\}.$$

On pose ensuite $V_1 = J_1$ (elle ne joue aucun rôle) et pour $k \in \{2, 3, \ldots, N\}$, nous définissons la variable $V_k$ de la façon suivante.

– Si $r_k - r_{k-1} = 1$, c'est-à-dire si $U_{r_{k-1}+1} \notin \{\boldsymbol{\sigma}(1), \ldots, \boldsymbol{\sigma}(k-1)\}$ (et donc $\boldsymbol{\sigma}(k) = U_{r_{k-1}+1}$), alors on pose $V_k = J_k$. Dans ce cas on observe que

$$\min(V_k, k-1) = k-1 \quad \text{et} \quad \{\boldsymbol{\sigma}(k-1), \boldsymbol{\sigma}(k)\} \in \mathcal{T}_* \,. \tag{I.119}$$

– Si $r_k - r_{k-1} \geq 2$, alors $U_{r_k - 1} \in \{\boldsymbol{\sigma}(1), \ldots, \boldsymbol{\sigma}(k-1)\}$ et il existe un unique $V_k \in \{1, \ldots, k-1\}$ tel que $U_{r_k - 1} = \boldsymbol{\sigma}(V_k)$. Dans ce cas, on observe que

$$\min(V_k, k-1) = V_k \quad \text{et} \quad \{\boldsymbol{\sigma}(V_k), \boldsymbol{\sigma}(k)\} \in \mathcal{T}_* \,. \tag{I.120}$$

Nous pouvons alors établir le constat suivant : on note $\boldsymbol{\sigma}^{<-1>}$ la permutation inverse de $\boldsymbol{\sigma}$ et on note $\mathcal{T} = \mathcal{T}_*^{(\boldsymbol{\sigma}^{<-1>})}$ l'arbre $\mathcal{T}_*$ ré-étiqueté par la permutation $\boldsymbol{\sigma}^{<-1>}$ ; alors (I.119) et (I.120) impliquent que

$$\mathcal{T} = \big\{ \{k, \min(V_k, k-1)\} \, ; \ 2 \leq k \leq N \big\} \quad \text{et} \quad \mathcal{T}^{(\boldsymbol{\sigma})} = \mathcal{T}_*$$

et donc $\mathcal{T}^{(\boldsymbol{\sigma})}$ *est de loi uniforme sur l'ensemble* $\mathbb{T}_N^{\mathsf{Cay}}$ *des arbres de Cayley à* $N$ *sommets.* Pour compléter la preuve du théorème, il suffit donc de montrer que $(V_k)_{2\leq k\leq N}$ et $\boldsymbol{\sigma}$ sont indépendantes, que les $(V_k)_{2\leq k \leq N}$ sont indépendantes et uniformes sur $\{1, \ldots, N\}$ et que $\boldsymbol{\sigma}$ est de loi uniforme sur $\mathfrak{S}_N$.

Pour cela, on fixe une permutation $\gamma \in \mathfrak{S}_N$ et $N$ nombres $j_1, \ldots, j_N \in \{1, \ldots, N\}$ ; pour tout $k \in \{1, \ldots, N\}$, on introduit l'événement

$$A_k = \bigcap_{1 \leq \ell \leq k} \{\boldsymbol{\sigma}(\ell) = \gamma(\ell); V_\ell = j_\ell\} \,.$$

On note $\mathscr{G}_{k-1}$ la tribu engendrée par les v.a. $U_0, \ldots, U_{r_{k-1}}$ et par $J_1, \ldots, J_{k-1}$. Pour simplifier on pose $U'_n = U_{r_{k-1}+n}$ pour tout entier $n \geq 1$. On vérifie que les v.a. $(U'_n)_{n\geq 1}$ sont i.i.d., de loi uniforme sur $\{1, \ldots, N\}$ et également indépendantes de $\mathscr{G}_{k-1}$ ; de même $J_k$ est indépendante de $\mathscr{G}_{k-1}$. Si $j_k \geq k$, on a donc

$$\begin{aligned} \mathbf{E}\big[\mathbf{1}_{A_k} \mathbf{1}_{\{r_k = r_{k-1}+1\}} \,\big|\, \mathscr{G}_{k-1}\big] &= \mathbf{E}\big[\mathbf{1}_{A_{k-1}} \mathbf{1}_{\{U'_1 = \gamma(k); J_k = j_k\}} \,\big|\, \mathscr{G}_{k-1}\big] \\ &= \mathbf{1}_{A_{k-1}} \mathbf{P}\big(U'_1 = \gamma(k)\big) \mathbf{P}(J_k = j_k) \\ &= \mathbf{1}_{A_{k-1}} \tfrac{1}{N} \tfrac{1}{N-k+1}. \end{aligned}$$

Bien sûr, si $j_k < k$, par définition de $V_k$ on a $\mathbf{E}\big[\mathbf{1}_{A_k} \mathbf{1}_{\{r_k = r_{k-1}+1\}} \,\big|\, \mathscr{G}_{k-1}\big] = 0$.

Soit un entier $m \geq 1$. Pour simplifier les notations, on pose $S_{k-1} = \{\gamma(1), \ldots, \gamma(k-1)\}$. On suppose que $j_k < k$ ; on observe alors que

$$\begin{aligned} \mathbf{E}\big[\mathbf{1}_{A_k} \mathbf{1}_{\{r_k = r_{k-1}+m+1\}} \,\big|\, \mathscr{G}_{k-1}\big] &= \mathbf{E}\big[\mathbf{1}_{A_{k-1}} \mathbf{1}_{\{U'_1, \ldots, U'_{m-1} \in S_{k-1}\}} \mathbf{1}_{\{U'_m = \gamma(j_k)\}} \mathbf{1}_{\{U'_{m+1} = \gamma(k)\}} \,\big|\, \mathscr{G}_{k-1}\big] \\ &= \mathbf{1}_{A_{k-1}} \mathbf{P}\big(U'_1 \in S_k\big) \ldots \mathbf{P}\big(U'_{m-1} \in S_k\big) \mathbf{P}(U'_m = \gamma(j_k)) \mathbf{P}(U'_{m+1} = \gamma(k)) \\ &= \mathbf{1}_{A_{k-1}} \big(\tfrac{k-1}{N}\big)^{m-1} \tfrac{1}{N} \tfrac{1}{N}. \end{aligned}$$

et en sommant sur $m \geq 1$, on en déduit

$$\mathbf{E}\big[\mathbf{1}_{A_k} \mathbf{1}_{\{r_k > r_{k-1}+1\}} \,\big|\, \mathscr{G}_{k-1}\big] = \mathbf{1}_{A_{k-1}} \tfrac{1}{N} \tfrac{1}{N-k+1}.$$

Bien sûr, si $j_k \geq k$, par définition de $V_k$ on a $\mathbf{E}\big[\mathbf{1}_{A_k} \mathbf{1}_{\{r_k > r_{k-1}+1\}} \,\big|\, \mathscr{G}_{k-1}\big] = 0$. Dans tous les cas on a montré que

$$\mathbf{E}\big[\mathbf{1}_{A_k} \,\big|\, \mathscr{G}_{k-1}\big] = \mathbf{1}_{A_{k-1}} \tfrac{1}{N(N-k+1)}$$

et en prenant l'espérance de cette expression, on obtient, $\mathbf{P}(A_k) = \mathbf{P}(A_{k-1})\frac{1}{N(N-k+1)}$. Une récurrence immédiate implique que $\mathbf{P}(A_N) = \mathbf{P}(A_1)N^{-(N-1)}((N-1)!)^{-1}$. Or $\boldsymbol{\sigma}(1) = U_0$ et $V_1 = J_1$ qui sont deux variables indépendantes de loi uniforme sur $\{1, \ldots, N\}$. On e déduit donc que

$$\mathbf{P}\big(\boldsymbol{\sigma}=\gamma \,;\, (V_1, \ldots, V_N)=(j_1, \ldots, j_N)\big) = \mathbf{P}(A_N) = \frac{1}{N^N N!},$$

ce qui montre bien que les v.a. $(V_k)_{1\leq k\leq N}$ sont i.i.d. de loi uniforme sur $\{1, \ldots, N\}$, que $\boldsymbol{\sigma}$ est de loi uniforme sur $\mathfrak{S}_N$ et que $(V_k)_{1\leq k\leq N}$ et $\boldsymbol{\sigma}$ sont indépendantes. Comme montré plus haut, cela suffit pour conclure. ■

La proposition permet de simuler très rapidement un arbre uniforme à $N$ sommets si on ne tient pas compte de l'étiquetage des sommets. Une simulation faite sous *mathematica* avec 50 000 sommets est donnée à la figure I.6.b (page 145).

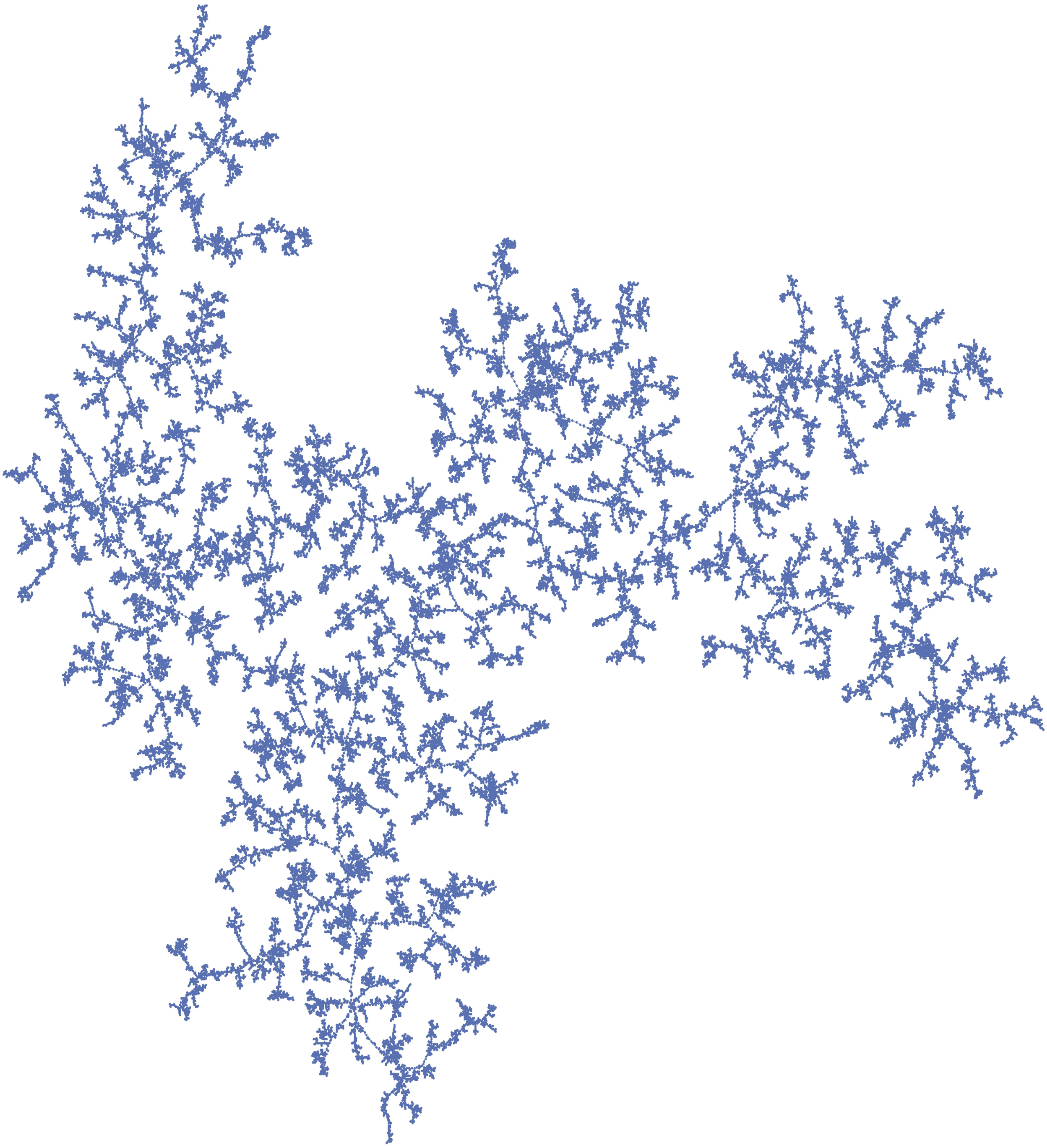

FIGURE I.15 – Un arbre uniforme à $N = 50\,000$ sommets.

# Chapitre II

# Processus markoviens à espace d'états discret

## II.1 Définitions et premières propriétés.

### II.1.a Préliminaires sur les processus à valeurs dans un espace dénombrable discret.

Dans tout le chapitre $E$ désigne un ensemble dénombrable de plus de deux éléments. On munit $E$ la topologie discrète, ce qui correspond à la distance $d : E \times E \to \mathbb{R}_+$ par $d(i,j) = 0$ si $i = j$ et $d(i,j) = 1$ si $i \neq j$. Ainsi une suite d'états $i_n \in E$, $n \geq 0$, qui converge vers $i$ relativement à $d$ doit être stationnaire à $i$ (c'est-à-dire qu'il existe $n_0 \in \mathbb{N}$ tel que $i_n = i$, pour tout $n \geq n_0$). On munit $E$ de la tribu de tous ses sous-ensembles notée $\mathscr{P}(E)$. Toute fonction définie sur $E$ discret et à valeurs dans un espace mesuré est donc mesurable.

On se donne un espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$. Une variable $Z : \Omega \to E$ est dite $\mathscr{F}$-mesurable ssi pour tout état $i \in E$, $\{Z = i\} \in \mathscr{F}$. On s'intéresse tout d'abord aux collections de variables aléatoires indexées par $\mathbb{R}_+$ et à valeurs dans $E$ : précisément on se donne $(X_t)_{t \in \mathbb{R}_+}$ une collection d'applications $X_t : \Omega \to E$ qui sont $\mathscr{F}$-mesurables. On veut voir cette collection de variables comme une fonction aléatoire, c'est-à-dire une application qui à $\omega \in \Omega$ associe la fonction $t \in \mathbb{R}_+ \mapsto X_t(\omega)$. L'ensemble des fonctions de $\mathbb{R}_+$ dans $E$ est noté $E^{\mathbb{R}_+}$ :

$$E^{\mathbb{R}_+} = \{\mathbf{x} = (x(t))_{t \in \mathbb{R}_+} \ : \ x(t) \in E \ , \ t \in \mathbb{R}_+\} \ .$$

Pour voir $X : \Omega \to E^{\mathbb{R}_+}$ comme une variable aléatoire, il est nécessaire de définir une tribu sur $E^{\mathbb{R}_+}$, ce qui est l'objet de la définition suivante.

**Définition II.1.1** (*Cylindres et tribu produit.*) Soit $E$ un ensemble dénombrable discret.

**(a)** (*Ensemble des cylindres élémentaires*) On dit qu'un ensemble $C \subset E^{\mathbb{R}_+}$ de fonctions de $\mathbb{R}_+$ dans $E$ est un *cylindre élémentaire* ssi il existe $n \in \mathbb{N}$, $0 \leq t_0 < t_1 < \ldots < t_n$ et $i_0, \ldots, i_n \in E$ tels que

$$C = \{\mathbf{x} = (x(t))_{t \in \mathbb{R}_+} \in E^{\mathbb{R}_+} \ : \ x(t_0) = i_0 \ ; \ldots ; x(t_n) = i_n\} \ .$$

On note $\mathscr{C}$ l'ensemble des cylindres élémentaires auxquels on adjoint $\emptyset$ et $E^{\mathbb{R}_+}$. C'est clairement un pi-système.

**(b)** (*Tribu produit*) On note $\mathscr{P}(E)^{\otimes \mathbb{R}_+}$ la tribu engendrée par $\mathscr{C}$. qui est la *tribu produit sur* $E^{\mathbb{R}_+}$. □

**Lemme II.1.2** *Soit $X_t : \Omega \to E$, $t \in \mathbb{R}_+$, une collection de variables aléatoires $\mathscr{F}$-mesurables. Alors l'application*

$$\omega \in \Omega \longmapsto \mathbf{X}(\omega) = (X_t(\omega))_{t \in \mathbb{R}_+} \ \in E^{\mathbb{R}_+}$$

*est* $(\mathscr{F}, \mathscr{P}(E)^{\otimes\mathbb{R}_+})$*-mesurable. Une telle variable aléatoire à valeurs dans* $E^{\mathbb{R}_+}$ *est appelé un* **processus à valeurs dans** $E$.

**Preuve :** comme $\mathscr{C}$ génère $\mathscr{B}(E^{\mathbb{R}_+})$, il suffit de vérifier que $\mathbf{X}^{-1}(C) \in \mathscr{F}$, pour tout $C \in \mathscr{C}$. Soit $C = \{\mathbf{x} = (x(t))_{t\in\mathbb{R}_+} \in E^{\mathbb{R}_+} \ : \ x(t_0) = i_0 \, ; \ldots ; x(t_n) = i_n\}$. Alors

$$\mathbf{X}^{-1}(C) = \{\omega \in \Omega : (X_t(\omega))_{t\in\mathbb{R}_+} \in C\} = \{\omega \in \Omega \ : \ X_{t_0}(\omega) = i_0 \, ; \ldots \, ; X_{t_n}(\omega) = i_n\} = \bigcap_{0\le p\le n} \{X_{t_p} = i_p\}.$$

Puisque $X_{t_p}$ est $\mathscr{F}$-mesurable, $\{X_{t_p} = i_p\} \in \mathscr{F}$, ce qui termine la preuve. ■

Nous introduisons ensuite l'*opérateur de décalage* sur les trajectoires.

**Définition II.1.3** (*Opérateur de décalage.*) On fixe $t_0 \in \mathbb{R}_+$. On définit l'application $\theta_{t_0} : E^{\mathbb{R}_+} \to E^{\mathbb{R}_+}$ par

$$\theta_{t_0}\mathbf{x} = (x(t_0 + t))_{t\in\mathbb{R}_+} \ ,$$

pour tout $\mathbf{x} = (x(t))_{t\in\mathbb{R}_+}$. L'application $\theta_{t_0}$ est appelée l'*opérateur de décalage au temps* $t_0$ des trajectoires. Il est facile de vérifier que $\theta_{t_0} \circ \theta_{t_1} = \theta_{t_0+t_1}$, pour tous $t_0, t_1 \in \mathbb{R}_+$. □

**Lemme II.1.4** *On fixe* $t \in \mathbb{R}_+$*. L' application* $\theta_t : E^{\mathbb{R}_+} \to E^{\mathbb{R}_+}$ *est* $(\mathscr{P}(E)^{\otimes\mathbb{R}_+}, \mathscr{P}(E)^{\otimes\mathbb{R}_+})$*-mesurable.*

**Preuve :** comme $\mathscr{C}$ engendre $\mathscr{P}(E)^{\otimes\mathbb{R}_+}$, il suffit de vérifier que $\theta_t^{-1}(C) \in \mathscr{P}(E)^{\otimes\mathbb{R}_+}$, pour tout $C \in \mathscr{C}$. On se donne $C = \{\mathbf{x} = (x(t))_{t\in\mathbb{R}_+} \in E^{\mathbb{R}_+} : x(t_0) = i_0; \ldots ; x(t_n) = i_n\}$. Alors

$$\theta_t^{-1}(C) = \{\mathbf{x} = (x(s))_{s\in\mathbb{R}_+} \in E^{\mathbb{R}_+} \ : \ x(t_0 + t) = i_0 \, ; \ldots ; x(t + t_n) = i_n\},$$

qui est un cylindre élémentaire, ce qui achève la preuve du lemme. ■

Nous introduisons la notion de filtration en temps continu.

**Définition II.1.5** (*Filtration en temps continu*) On rappelle que $(\Omega, \mathscr{F})$ est un espace mesurable.

**(a)** Une *filtration* sur $(\Omega, \mathscr{F})$ est une collection $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$ de sous-tribus de $\mathscr{F}$ qui est croissante pour l'inclusion $\mathscr{F}_s \subset \mathscr{F}_t \subset \mathscr{F}$, pour tous $t \ge s$.

**(b)** (*Filtration continue à droite*) La filtration continue à droite, notée $(\mathscr{F}_{t+})_{t\ge 0}$ est définie par $\mathscr{F}_{t+} = \bigcap_{h>0} \mathscr{F}_{t+h}$, pour tout $t \in \mathbb{R}_+$. Il est facile de vérifier que $\mathscr{F}_{t+} = \bigcap_{h>0} \mathscr{F}_{(t+h)+}$, $t \in \mathbb{R}_+$ d'où le nom de filtration continue à droite.

**(c)** (*Processus adapté* ) On se donne un processus $\mathbf{X} = (X_t)_{t\in\mathbb{R}_+}$ défini sur $(\Omega, \mathscr{F})$, à valeurs dans $E$, qui est $(\mathscr{F}, \mathscr{P}(E)^{\otimes\mathbb{R}_+})$-mesurable. On se donne une filtration $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$. On dit que $\mathbf{X}$ est $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$-adapté ssi pour tout $t \in \mathbb{R}_+$, $X_t$ est $\mathscr{F}_t$-mesurable.

**(d)** (*Filtration naturelle* ) On se donne un processus $\mathbf{X} = (X_t)_{t\in\mathbb{R}_+}$ défini sur $(\Omega, \mathscr{F})$, à valeurs dans $E$, qui est $(\mathscr{F}, \mathscr{P}(E)^{\otimes\mathbb{R}_+})$-mesurable. Pour tout $t \in \mathbb{R}_+$, on note $\mathscr{F}_t^{\mathbf{X}}$ la tribu engendrée par les variables aléatoires $X_s$, $s \in [0, t]$. Autrement dit

$$\mathscr{F}_t^{\mathbf{X}} = \sigma\left(\{X_s = i\} \ ; \ i \in E \ , \ s \in [0, t]\right) \ .$$

Il est facile de voir que $(\mathscr{F}_t^{\mathbf{X}})_{t\in\mathbb{R}_+}$ est une filtration sur $(\Omega, \mathscr{F})$. On l'appelle *la filtration naturelle associée à* $\mathbf{X}$. Il est facile de voir également que c'est la plus petite filtration (pour l'inclusion) par rapport à laquelle $\mathbf{X}$ est adapté. □

### II.1.b Semi-groupes.

Dans le cas des processus à temps continu, ce qui joue le rôle des matrices de transition est le *semi-groupe* qui est défini comme suit.

**Définition II.1.6** (*Semi-groupes markoviens et sous-markoviens*) On rappelle que $E$ est un espace dénombrable discret. Pour tout $t \in \mathbb{R}_+$, on se donne une matrice sur $E$ notée $P_t = (p_t(i,j))_{i,j\in E}$. La famille $(P_t)_{t\in\mathbb{R}_+}$ forme un *semi-groupe sous-markovien* si elle satisfait les propriétés suivantes.

$(a)$ Pour tous $t \in \mathbb{R}_+$ et tous $i,j \in E$, on a $p_t(i,j) \geq 0$.

$(b)$ Pour tous $t,s \in \mathbb{R}_+$ et tous $i,j \in E$, on a $p_{s+t}(i,j) = \sum_{k\in E} p_s(i,k)p_t(k,j)$, ce que l'on note « matriciellement » par

$$\forall s,t \in \mathbb{R}_+ \ , \quad P_sP_t = P_{t+s} \ ,$$

C'est la propriété de *semi-groupe*. On requiert également que $P_0 = \mathtt{Id}_E$, c'est-à-dire que $p_0(i,j) = 1$ si $i = j$ et $p_0(i,j) = 0$ si $i \neq j$.

$(c)$ Pour tout $t \in \mathbb{R}_+$ et tous $i,j \in E$, $\sum_{j\in E} p_t(i,j) \leq 1$.

La famille $(P_t)_{t\in\mathbb{R}_+}$ forme un *semi-groupe markovien* si elle satisfait $(a)$, et $(b)$ comme ci-dessus et $(c')$ :

$(c')$ Pour tout $t \in \mathbb{R}_+$ et tous $i,j \in E$, $\sum_{j\in E} p_t(i,j) = 1$. □

**Exemple II.1.7** On prend $\mathbb{N} = E$ et on fixe $c > 0$. Pour tout $t \in \mathbb{R}_+$, on pose

$$p_t(i,j) = \mathbf{1}_{\{i\leq j\}} c^{j-i} e^{-ct}/(j-i)! \ , \quad i,j \in \mathbb{N}.$$

On vérifie que la famille $(P_t)_{t\in\mathbb{R}_+}$ est un semi-groupe markovien. On verra plus loin que c'est le semi-groupe du processus de Poisson linéaire de paramètre $c$. □

**Exemple II.1.8** On prend $E = \{0,1\}$ et pour tout $t > 0$ on pose $p_t(i,j) = 1/2$, $i,j \in \{0,1\}$ et on pose $p_0(1,1) = p_0(0,0) = 1$, $p_0(1,0) = p_0(0,1) = 0$. On vérifie que la famille $(P_t)_{t\in\mathbb{R}_+}$ est un semi-groupe markovien. □

**Remarque II.1.9** (*Rendre markovien un semi-groupe sous-markovien*) Considérons un semi-groupe sous-markovien $(P_t)_{t\in\mathbb{R}_+}$. Quitte à considérer un espace d'état plus grand, il est possible de le « rendre » markovien par la procédure suivante. On se donne un point *cimetière*, c'est-à-dire un point $\partial$ qui n'est pas dans $E$. On pose $E^* = E \cup \{\partial\}$. Pour tout $t \in \mathbb{R}_+$, on définit la matrice $P_t^* = (p_t^*(i,j))_{i,j\in E^*}$, comme suit.

$$p_t^*(i,j) = p_t(i,j) \ , \ i,j \in E \ , \quad p_t^*(\partial,\partial) = 1 \ , \quad p_t^*(\partial,j) = 0 \ , \ j \in E$$

et pour tout $i \in E$, on pose

$$p_t^*(i,\partial) = 1 - p_t(i,E) = 1 - \sum_{j\in E} p_t(i,j) \ .$$

On vérifie alors que $(P_t^*)_{t\in\mathbb{R}_+}$ est un semi-groupe markovien sur $E^*$. □

**Semi-groupes (sous) markoviens sur un espace d'état fini.**

Lorsque $E$ est fini, pour simplifier les notations, on identifie $E$ à $\{1, \ldots, N\}$, $N$ étant le nombre d'éléments de $E$. On note $\mathscr{M}_N(\mathbb{C})$ l'ensemble des matrices $N \times N$ à coefficients complexes. $\mathscr{M}_N(\mathbb{C})$ est un espace vectoriel réel de dimension $N^2$ et il peut être muni d'une norme d'algèbre $\|\cdot\|$, c'est-à-dire qu'en plus d'être une norme d'espace vectoriel $\|\cdot\|$ satisfait $\|AB\| \leq \|A\| \, \|B\|$, pour tout $A, B \in \mathscr{M}_N(\mathbb{R})$. Si $A = (a_{i,j})_{1 \leq i,j \leq N} \in \mathscr{M}_N(\mathbb{R})$, on utilisera la notation $[A](i,j) = a_{i,j}$ pour l'entrée $(i,j)$ de la matrice $A$. On voit que pour $(i,j)$ fixe $A \mapsto [A](i,j)$ est une application continue.

**Rappels sur l'exponentielle de matrices.** Soit $A \in \mathscr{M}_N(\mathbb{C})$. Pour tout $n \in \mathbb{N}$, on pose $f_n(A) = \sum_{0 \leq k \leq n} \frac{1}{k!} A^k$, avec la convention $A^0 = \mathrm{Id}$. On remarque que

$$\sup_{n_1, n_2 \geq n} \|f_{n_1}(A) - f_{n_2}(A)\| \leq \sum_{m \geq n} \frac{1}{m!} \|A\|^m \xrightarrow[n \to \infty]{} 0.$$

La suite de matrices $(f_n(A), n \geq 0)$ est de Cauchy dans $(\mathscr{M}_N(\mathbb{C}), \|\cdot\|)$, qui est complet car de dimension finie. Cette suite de matrices a donc une limite qui est notée symboliquement

$$\exp(A) := \|\cdot\|\text{-}\lim_n f_n(A) = \sum_{n \geq 0} \frac{1}{n!} A^n \; .$$

Cela montre que pour tout $1 \leq i, j \leq N$, la série numérique suivante est absolument convergente

$$[\exp(A)](i,j) = \sum_{n \geq 0} \frac{1}{n!} [A^n](i,j) \; .$$

On suppose que $A, B \in \mathscr{M}_N(\mathbb{C})$ commutent : $A.B = B.A$. On montre facilement que

$$f_n(A).f_n(B) - f_n(A+B) = \sum_{n+1 \leq q \leq 2n} \frac{1}{q!} \sum_{q-n \leq p \leq n} \binom{q}{p} A^p.B^{q-p} \; .$$

Par conséquent

$$\|f_n(A).f_n(B) - f_n(A+B)\| \leq \sum_{q > n} \frac{1}{q!} \big(\|A\| + \|B\|\big)^q \xrightarrow[n \to \infty]{} 0.$$

Cela entraîne la propriété suivante

$$\forall \, A, B \in \mathscr{M}_N(\mathbb{C}) \, , \quad A.B = B.A \implies \exp(A+B) = \exp(A).\exp(B) \; . \tag{II.1}$$

L'exponentielle de matrice permet de définir des semi-groupes de matrices. Plus précisément, soit $G \in \mathscr{M}_N(\mathbb{C})$. Pour tout $t \in \mathbb{R}_+$ on pose $P_t = \exp(tG)$. Comme $tG$ et $sG$ commutent et que $tG + sG = (t+s)G$, $(P_t)_{t \in \mathbb{R}_+}$ est un semi-groupe pour le produit matriciel c'est-à-dire que $P_0 = \mathtt{Id}$ et $P_{t+s} = P_t P_s$. Le théorème ci-dessous montre une réciproque à ce fait sous une hypothèse de dérivabilité du semi-groupe.

**Théorème II.1.10** *Pour tout $t \in \mathbb{R}_+$, on se donne une matrice $P_t = (p_t(i,j))_{1 \leq i,j \leq N} \in \mathscr{M}_N(\mathbb{C})$. une matrice On fait les hypothèses suivantes.*

$(a)$ *On suppose que $P_t P_s = P_{t+s}$, pour tous $s, t \in \mathbb{R}_+$.*
$(b)$ *On suppose que $P_0 = \mathtt{Id}$.*

$(c)$ *On suppose que* $t \in \mathbb{R}_+ \mapsto P_t$ *est continument dérivable à droite en* $0^+$*, c'est-à-dire qu'il existe* $G = (q_{i,j})_{1\leq i,j\leq N} \in \mathscr{M}_N(\mathbb{C})$ *telle que*

$$\forall i,j \in \{1,\ldots,N\} \quad \lim_{t\to 0^+} \tfrac{1}{t}(p_t(i,j) - \delta_{i,j}) = q_{i,j} \; .$$

*Alors, les assertions suivantes sont vraies.*

$(i)$ *Pour tout* $t \in \mathbb{R}_+$*, on a* $P_t = \exp(tG)$*. La fonction* $t \in \mathbb{R}_+ \mapsto P_t \in \mathscr{M}_N(\mathbb{C})$ *est dérivable et sa dérivée* $\frac{d}{dt}P_t$ *satisfait l'équation différentielle ordinaire*

$$\forall t \in \mathbb{R}_+, \quad \frac{d}{dt}P_t = GP_t = P_tG \quad et \quad P_0 = \mathtt{Id} \; .$$

$(ii)$ $(P_t)_{t\in\mathbb{R}}$ *est* markovien *(resp.* sous-markovien*) si et seulement si les coefficients de* $G$ *sont des* réels *satisfaisant*

$$\forall\, 1 \leq i,j \leq N\,, \quad q_{i,j} \geq 0 \; si \; i \neq j \quad et \quad \sum_{1\leq j\leq N} q_{i,j} = -q_{i,i} \quad (resp. \; \leq -q_{i,i}\;). \tag{II.2}$$

*On observe que* $-q_{i,i} \geq 0$ *pour tout* $1 \leq i \leq N$.

**Remarque II.1.11** Le théorème II.1.10 reste vrai sous les hypothèses $(a)$, $(b)$ et l'hypothèse plus faible $(c')$ que $t \in \mathbb{R}_+ \mapsto P_t$ est continu à droite en $0^+$. Voir l'exercice II.3 pour une preuve. □

**Preuve du théorème.** On commence par montrer $(i)$ sous les hypothèses $(a)$ $(b)$ et $(c)$. Par $(c)$, il existe un réel $\eta > 0$ tel que la fonction $h \in [-\eta,\eta] \mapsto P_{h+\eta}$ soit continue. Soit $t_0 \geq 2\eta$. Par la propriété de semi-groupe, $P_{t_0+h} = P_{t_0-\eta}P_{h+\eta}$, ce qui implique que la fonction $h \in [-\eta,\eta] \mapsto P_{t_0+h}$ est continue et donc que $t \mapsto P_t$ l'est également sur tout $\mathbb{R}_+$.

Puisqu'on est en dimension finie et que la convergence pour toute norme sur $\mathscr{M}_N(\mathbb{C})$ est équivalente à la convergence des entrées des matrices, $(a)$ se reformule de manière équivalente par :

$$\lim_{h\to 0^+} \|\tfrac{1}{h}(P_h - \mathtt{Id}) - G\| = 0 \; .$$

On observe ensuite que pour tout $t > 0$, on a $\|\frac{1}{h}(P_{t+h} - P_t) - GP_t\| \leq \|P_t\|\, \|\frac{1}{h}(P_h - \mathtt{Id}) - G\|$. Par conséquent $t \mapsto P_t$ est dérivable à droite sur tout $\mathbb{R}_+$ de dérivée $GP_t$. De même on montre que la dérivée à droite est aussi égale à $P_tG$. Donc $G$ commute avec $P_t$, pour tout $t \in \mathbb{R}_+$.

Soit un réel $t > 0$ et soit $h \in ]0,t[$ . Comme on a montré que le semi-groupe est continu, $c = \sup_{s\in[0,t]}\|P_s\|$ est une quantité finie et on observe que

$$\begin{aligned}
\|\tfrac{1}{h}(P_{t-h} - P_t) + GP_t\| &\leq \|\tfrac{1}{h}(P_{t-h} - P_t) + GP_{t-h}\| + \|GP_t - GP_{t-h}\|, \\
&\leq \|P_{t-h}(\tfrac{1}{h}(\mathtt{Id} - P_h) + G)\| + \|G\|\,\|P_{t-h}\|\,\|P_h - \mathtt{Id}\|, \\
&\leq \|P_{t-h}\|\,\|\tfrac{1}{h}(\mathtt{Id} - P_h) + G\| + c\|G\|\,\|P_h - \mathtt{Id}\|, \\
&\leq c\|\tfrac{1}{h}(\mathtt{Id} - P_h) + G\| + c\|G\|\,\|P_h - \mathtt{Id}\|
\end{aligned}$$

On en déduit que $\lim_{h\to 0^+} \frac{1}{h}\|\frac{1}{h}(P_{t-h} - P_t) + GP_t\| = 0$. Par conséquent $t \mapsto P_t$ est dérivable à gauche et sa dérivée coïncide avec la dérivée à droite. En définitive, $t \mapsto P_t$ est dérivable sur $\mathbb{R}_+$ et satisfait l'équation différentielle : $\frac{d}{dt}P_t = G\cdot P_t$ avec condition initiale $P_0 = \mathtt{Id}$. Cette équation différentielle linéaire d'ordre 1 n'a qu'une seule solution de donnée initiale $\mathtt{Id}$. Or on vérifie que $t \mapsto \exp(tG)$ est solution de la même équation. On a donc $P_t = \exp(tG)$, ce qui montre le premier point.

Montrons ensuite $(ii)$. On suppose d'abord que $(P_t)_{t\in\mathbb{R}_+}$ est markovien. C'est donc une matrice réelle et donc $\frac{d}{dt}P_{0+} = G$ est une matrice réelle également. Puisque par $(i)$ $P_t = \exp(tG)$, on peut donc écrire pour tout $t \in [0,1]$,

$$P_t = \mathtt{Id} + tG + t^2 B_t \ ,$$

avec $t \in [0,1] \mapsto B_t \in \mathscr{M}_N(\mathbb{R})$ continue bornée sur $[0,1]$, ses entrées notées $b_t(i,j)$ sont donc continues bornées. Pour tous $1 \leq i,j \leq N$, on a donc

$$p_t(i,j) = tq_{i,j} + t^2 b_t(i,j) \text{ si } i \neq j \quad \text{et} \quad p_t(i,i) = 1 + tq_{i,i} + t^2 b_t(i,i). \tag{II.3}$$

Supposons d'abord que $i \neq j$. Comme $p_t(i,j) \geq 0$, on en déduit $q_{i,j} \geq -tb_t(i,j)$, pour tout $t \in ]0,1]$ et donc $q_{i,j} \geq 0$ en faisant tendre $t$ vers $0^+$. Supposons que le semi-groupe soit sous-markovien. On remarque ensuite

$$1 \geq \sum_{1\leq j\leq N} p_t(i,j) = 1 + t \sum_{1\leq j\leq N} q_{i,j} \ + t^2 \sum_{1\leq j\leq N} b_t(i,j) \tag{II.4}$$

Donc $0 \geq \sum_{1\leq j\leq N} q_{i,j} + t\sum_{1\leq j\leq N} b_t(i,j)$. En faisant tendre $t$ vers $0^+$, on obtient $\sum_{1\leq j\leq N} q_{i,j} \leq 0$, ce qui implique (II.2) dans le cas sous-markovien. Le cas markovien se traite de façon similiaire, l'inégalité dans (II.4) étant une égalité. en faisant tendre $t$ vers $0^+$. On a donc prouvé une implication de $(ii)$.

Montrons l'implication contraire : on suppose que $(P_t)_{t\in\mathbb{R}}$ satisfait $(a)$, $(b)$ et $(c)$ et (II.2). On montre d'abord que pour tout $t \in \mathbb{R}_+$, la matrice $P_t$ est à coefficients positifs. Pour cela, on fait d'abord l'hypothèse que $q_{i,j} \neq 0$, pour tous $1 \leq i,j \leq N$. On remarque que puisque $P_t = \exp(tG)$, on peut écrire $P_t = \mathtt{Id} + tG + t^2 B_t$ avec $B_t$ bornée, ce qui implique (II.3). Il existe donc un réel $\varepsilon > 0$ tel que $p_t(i,j) > 0$, pour tout $t \in ]0,\varepsilon]$ et pour tous $i,j$. Soit un réel $t_0 > 0$. Il existe un entier $n \geq 1$ tel que $0 < t_0/n < \varepsilon$. Ce qui précède implique alors que $P_{t_0/n}$ est à coefficients positifs. Il en est donc même pour $P^n_{t_0/n}$ qui vaut $P_{t_0}$, par la propriété de semi-groupe. Par conséquent, pour tout $t \in \mathbb{R}_+$, $P_t$ est à coefficient positifs.

Montrons ensuite que pour tout $t \in \mathbb{R}_+$, la matrice $P_t$ est à coefficients positifs, sans plus nécessairement supposer que les $q_{i,j}$ soient non-nuls. On procède par perturbation pour se ramener au cas de coefficients non-nuls. Soit $\eta \in ]0,1[$ . On introduit $G_\eta = (q^\eta_{i,j})_{i,j\in E}$, où $q^\eta_{i,j} = q_{i,j} + \eta > 0$ pour tous $i \neq j$ et $q^\eta_{i,i} = q_{i,i} - N\eta < 0$. Autrement dit $G^\eta = G + \eta M$, où $M$ est une certaine matrice carrée. On pose ensuite $P^\eta_t = \exp(tG_\eta) = (p^\eta_t(i,j))_{i,j\in E}$. Par ce qui précède $p^\eta_t(i,j) \geq 0$, pour tous $i,j \in E$, $t \in \mathbb{R}_+$ et tout $\eta \in ]0,1[$ . Or la fonction exponentielle $\exp : \mathscr{M}_N(\mathbb{R}) \to \mathscr{M}_N(\mathbb{R})$ est continue donc pour tout $t \in \mathbb{R}_+$, $\|\cdot\| - \lim_{\eta\to 0^+} \exp(tG_\eta) = \exp(tG)$. Par conséquent, $p_t(i,j) = \lim_{\eta\to 0^+} p^\eta_t(i,j)$ et donc $p_t(i,j) \geq 0$, ce pour tout $t \in \mathbb{R}_+$ et pour tous $1 \leq i,j \leq N$.

On note ensuite $v$, le vecteur colonne dont les $N$ composantes valent 1. On a donc $Gv = 0$ (resp. les coordonnées du vecteur $Gv$ sont négatives). On remarque que la $i$-éme composante de $P_t.\mathbf{1}_N$ est $\sum_{1\leq j\leq N} p_t(i,j)$. On remarque ensuite que

$$\frac{d}{dt}(P_t v) = (\frac{d}{dt}P_t)v = (P_t G)v = P_t(Gv) = \Big(\sum_{1\leq j\leq N} p_t(i,j)(Gv)(j)\Big)_{1\leq i\leq N}.$$

Si $Gv = 0$ (si resp. les coordonnées du vecteur $Gv$ sont négatives) alors pour tout $i \in \{1,\ldots,N\}$, $t \mapsto P_t v(i) = \sum_{1\leq j\leq N} p_t(i,j)$ est une fonction constante (resp. décroissante). Or $P_0 v = v$, qui est le vecteur dont les compsantes sont toutes égales à 1. Donc $\sum_{1\leq j\leq N} p_t(i,j) = 1$ (resp. $\leq 1$). Cela termine la preuve de la seconde implication de $(ii)$ et du théorème. ■

EXERCICES.

**Exercice II.1** Montrer que le semi-groupe donné dans l'exemple II.1.8 (b) n'est pas de la forme exponentielle du théorème précédent. □

**Exercice II.2** Il s'agit d'un exercice préparatoire à l'exercice II.3. On munit $\mathscr{M}_N(\mathbb{C})$ d'une norme matricielle $\|\cdot\|$. Soit $t \in \mathbb{R}_+ \mapsto R_t \in \mathscr{M}_N(\mathbb{C})$, une fonction continue. Soient $a, b \in \mathbb{R}_+$ tels que $a \leq b$. On note $\int_a^b R_t \, dt$ la matrice $M$ telle que $[M](i,j) = \int_a^b [R_t](i,j) \, dt$.

1. Montrer que $\|\cdot\| - \lim_{n\to\infty} \frac{b-a}{n} \sum_{0\leq k<n} R_{a+(b-a)k/n} = \int_a^b R_t \, dt$.
2. Montrer que $\|\int_a^b R_t \, dt\| \leq \int_a^b \|R_t\| \, dt$.
3. Soit $t \in \mathbb{R}_+ \mapsto Q_t \in \mathscr{M}_N(\mathbb{C})$, une fonction continue. Soient $P, M \in \mathscr{M}_N(\mathbb{C})$. Montrer que $M \int_a^b R_t \, dt = \int_a^b M R_t \, dt$ et $(\int_a^b R_t \, dt) M = \int_a^b R_t M \, dt$. Montrer que $\int_a^b R_t \, dt + \int_a^b Q_t \, dt = \int_a^b (R_t + Q_t) \, dt$.
4. Montrer soit $c \in \mathbb{R}_+$ tel que $c \geq b$. Montrer que $\int_a^c R_t \, dt = \int_a^b R_t \, dt + \int_b^c R_t \, dt$.
5. Montrer que $t \mapsto \int_0^t R_s \, ds$ a des entrées $C^1$ et montrer que
$$\frac{d}{dt} \int_0^t R_s \, ds = R_t \, .$$
6. On suppose désormais que $\int_0^\infty \|R_t\| \, dt < \infty$. Montrer que $\|\cdot\| - \lim_{b\to\infty} \int_0^b R_t \, dt$ existe dans $\mathscr{M}_N(\mathbb{C})$. On note simplement $\int_0^\infty R_t \, dt$ cette limite.
Montrer que $\|\int_0^\infty R_t \, dt\| \leq \int_0^\infty \|R_t\| \, dt$.
Montrer que $M \int_0^\infty R_t \, dt = \int_0^\infty M R_t \, dt$ et $(\int_0^\infty R_t \, dt) M = \int_0^\infty R_t M \, dt$.
On suppose également que $\int_0^\infty \|Q_t\| \, dt < \infty$. Montrer que $\int_0^\infty R_t \, dt + \int_0^\infty Q_t \, dt = \int_0^\infty (R_t + Q_t) \, dt$.
7. Soit $v \in \mathbb{C}^N$. On suppose que $\|\cdot\|$ est la norme matricielle subordonnée à une norme $|\cdot|$ sur $\mathbb{C}^N$ : $\|M\| = \max_{v\in\mathbb{C}^N : |v|\leq 1} |Mv|$.
Soit $v \in \mathbb{C}^N$. Montrer que $\int_a^b R_t v \, dt = (\int_a^b R_t \, dt) v$.
Montrer que $|\cdot| - \lim_{b\to\infty} \int_0^b R_t v \, dt$ existe et vaut $(\int_0^\infty R_t \, dt) v$.
Soit $\phi : \mathbb{C}^N \to \mathbb{C}$ une forme linéaire. Montrer que $t \in \mathbb{R}_+ \mapsto \phi(R_t v)$ est intégrable sur $\mathbb{R}_+$ et que $\phi((\int_0^\infty R_t \, dt) v) = \phi(\int_0^\infty R_t v \, dt) = \int_0^\infty \phi(R_t v) \, dt$. □

**Exercice II.3** Il est recommandé de faire d'abord l'exercice II.2. Soit $(P_t)_{t\in\mathbb{R}_+}$, une famille de matrice de $\mathscr{M}_N(\mathbb{C})$. On suppose que la fonction $t \mapsto P_t$ est continue à droite en 0. On suppose également qu'elle satisfait les hypothèses $(a)$ et $(b)$ du théorème II.1.10, page 150. On suppose que $\|\cdot\|$ est la norme matricielle subordonnée à une norme $|\cdot|$ sur $\mathbb{C}^N$

1. Montrer que la fonction $t \mapsto P_t$ est continue.
2. On note $D = \{v \in \mathbb{C}^N \ : \ |\cdot| - \lim_{h\to 0^+} \frac{1}{h}(P_h v - v) \text{ existe}\}$. Montrer que $D$ est un $\mathbb{C}$-sous-espace vectoriel de $\mathbb{C}^N$.
3. Soit $\lambda_0 \in \mathbb{R}$ tel que $\max_{t\in[0,1]} \|P_t\| \leq e^{\lambda_0}$. Montrer que $\|P_t\| \leq e^{\lambda_0} e^{\lambda_0 t}$ pour tout $t \in \mathbb{R}_+$. Montrer que pour tout réel $\lambda > \lambda_0$, $\int_0^\infty e^{-\lambda t} \|P_t\| \, dt < \infty$ et en utilisant l'exercice II.2, page 153, en déduire que $R_\lambda := \int_0^\infty e^{-\lambda t} P_t \, dt$ est une matrice de $\mathscr{M}_N(\mathbb{C})$ bien définie.
4. Soit $h \in \mathbb{R}_+$, justifier (soigneusement) que $R_\lambda P_h = P_h R_\lambda = e^{\lambda h} R_\lambda - e^{\lambda h} \int_0^h e^{-\lambda s} P_s \, ds$.
5. Déduire de la question précédente et de l'exercice II.2, page 153 que pour tout $\lambda > \lambda_0$ et tout $v \in \mathbb{C}^N$, $R_\lambda v \in D$.
6. Pour tout $v \in \mathbb{C}^N$, montrer que $\lim_{\lambda\to\infty} |\lambda R_\lambda v - v| = 0$.
7. Déduire des questions précédentes que $D = \mathbb{C}^N$, c'est-à-dire que $t \mapsto P_t$ est dérivable à droite en 0.
8. Conclure. □

## II.1.c Définition des processus markoviens, propriété de Markov.

On rappelle que $(\Omega, \mathscr{F}, \mathbf{P})$ désigne un espace de probabilité sur lequel sont définies, sauf mention explicite du contraire, les variables aléatoires que l'on considère. On commence par une première définition des processus de Markov.

**Définition II.1.12** (*Première définition des processus markoviens à espace d'états discret.*) Soit $E$ un espace dénombrable discret. Un processus $\mathbf{X} = (X_t)_{t\in\mathbb{R}_+}$ défini sur $(\Omega, \mathscr{F})$ et à valeurs dans $E$ est un processus markovien sous $\mathbf{P}$ ssi il existe un semi-groupe markovien sur $E$ noté $P_t = (p_t(i,j))_{i,j\in E}$, $t \in \mathbb{R}_+$ tel que pour tous $t_0 = 0 < t_1 < \ldots < t_n$ et pour tout tous $i_0, \ldots, i_n \in E$, on ait

$$\mathbf{P}\left(X_0 = i_0; X_{t_1} = i_1; \ldots; X_{t_n} = i_n\right) = \mathbf{P}(X_0 = i_0) p_{t_1}(i_0, i_1) p_{t_2 - t_1}(i_1, i_2) \ldots p_{t_n - t_{n-1}}(i_{n-1}, i_n) \ .$$

La loi de $X_0$ sous $\mathbf{P}$ est appelée loi initiale ou loi d'entrée de $\mathbf{X}$. On la notera souvent $\mu$. Pour résumer, on dit que $\mathbf{X}$ est un processus de Markov à valeurs dans $E$ de semi-groupe $(P_t)_{t\in\mathbb{R}_+}$ et de loi d'entrée $\mu$. □

Comme pour les chaînes de Markov, il est commode d'introduire une filtration dans la définition d'un processus de Markov. Cela donne lieu à la définition suivante.

**Définition II.1.13** Soit $E$ un espace dénombrable discret. Soit $\mathbf{X} = (X_t)_{t\in\mathbb{R}_+}$, un processus de Markov à valeurs dans $E$ qui est défini sur l'espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$. Soit $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$, une filtration définie sur $(\Omega, \mathscr{F})$. On dit que $\mathbf{X}$ est un processus de Markov relativement à la filtration $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$ ssi pour tous $s, t \in \mathbb{R}_+$ et pour toute fonction $f : E \to \mathbb{R}_+$ bornée, il existe une fonction $g : E \to \mathbb{R}$ telle que

$$\mathbf{P}\text{-p.s.} \quad \mathbf{E}\big[f(X_{t+s})\,|\mathscr{F}_t\big] = g(X_t) \ .$$

On note que $g$ dépend a priori de $f$ et de $t$ et $s$. □

**Lemme II.1.14** *Soit $E$ un espace dénombrable discret. Soit $(\Omega, \mathscr{F}, \mathbf{P})$ un espace de probabilité muni d'une filtration $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$. Soit $\mathbf{X} = (X_t)_{t\in\mathbb{R}_+}$, un processus de de Markov à valeurs dans $E$ relativement à la filtration $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$. On note $P_t = (p_t(i,j))_{i,j\in E}$, $t \in \mathbb{R}_+$, son semi-groupe. Alors pour tous $s, t \in \mathbb{R}_+$ et tout $j \in E$, on a*

$$\mathbf{P}\text{-}p.s. \quad \mathbf{E}\big[\mathbf{1}_{\{X_{t+s}=j\}}\,|\mathscr{F}_t\big] = p_s(X_t, j) \ .$$

**Preuve :** par définition il existe $g : E \to \mathbb{R}$ telle que $\mathbf{P}$-p.s. $\mathbf{E}\big[\mathbf{1}_{\{X_{t+s}=j\}}\,|\mathscr{F}_t\big] = g(X_t)$. Pour tout $i, i_0 \in E$, on a donc

$$\begin{aligned}
\mathbf{P}(X_0 = i_0) p_t(i_0, i) p_s(i, j) &= \mathbf{P}(X_0 = i_0\,;\, X_t = i\,;\, X_{t+s} = j) \\
&= \mathbf{E}\big[\mathbf{1}_{\{X_0 = i_0\,;\, X_t = i\}} g(X_t)\big] \\
&= \mathbf{P}(X_0 = i_0; X_t = i) g(i) = \mathbf{P}(X_0 = i_0) p_t(i_0, i) g(i) \ .
\end{aligned}$$

On en déduit que si $\mathbf{P}(X_t = i) > 0$, alors $g(i) = p_s(i,j)$, ce qui implique le résultat voulu. ■

Comme pour les chaînes de Markov, nous introduisons la notion de *processus de Markov global*. Le terme « global » n'étant pas standard et il ne sera plus employé dans la suite.

**Définition II.1.15** (*Processus de Markov « global ».*) On se donne un espace dénombrable discret $E$. Un processus markovien (global) à valeurs dans $E$ est la donnée des objets suivants.

$$\big(\,\Omega\,;\,\mathscr{F}\,;\,(\mathscr{F}_t)_{t\in\mathbb{R}_+}\,;\,\mathbf{X} = (X_t)_{t\in\mathbb{R}_+}\,;\,(P_t)_{t\in\mathbb{R}_+}\,;\,\mathbf{P}_\mu\,,\,\mu \in \mathscr{M}_1(E)\,\big) \ .$$

($a$) $(\Omega, \mathscr{F})$ est un espace mesurable et $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$ est une filtration sur $(\Omega, \mathscr{F})$.

($b$) Pour tout $t \in \mathbb{R}_+$, $X_t : \Omega \to E$ est $\mathscr{F}_t$-mesurable. Autrement dit $\mathbf{X} = (X_t)_{t\in\mathbb{R}_+}$ est un processus à valeurs dans $E$ qui est $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$-adapté.

($c$) $(P_t)_{t\in\mathbb{R}_+}$ est un semi-groupe markovien. Pour tout $t \in \mathbb{R}_+$, on adopte la notation $P_t = (p_t(i,j))_{i,j\in E}$.

$(d)$ Pour toute $\mu \in \mathscr{M}_1(E)$, $\mathbf{P}_\mu$ est une mesure de probabilité sur $(\Omega, \mathscr{F})$ telle que

$$\forall\, s,t \in \mathbb{R}_+ \,,\ \forall\, j \in E \,, \quad \mathbf{P}_\mu\text{-p.s.} \quad \mathbf{E}_\mu\big[\mathbf{1}_{\{X_{t+s}=j\}} \,\big|\, \mathscr{F}_t\,\big] = p_s(X_t, j) \;. \tag{II.5}$$

ce qui est equivalent à ce que pour toute fonction $f : E \to \mathbb{R}$ bornée ou positive on ait

$$\mathbf{P}_\mu\text{-p.s.} \quad \mathbf{E}_\mu\big[f(X_{t+s}) \,\big|\, \mathscr{F}_t\,\big] = (P_s f)(X_t) = \sum_{j\in E} p_s(X_t, j) f(j) \;.$$

**Notations :** on note $\delta_i$ la masse de Dirac en $i \in E$, c'est-à-dire la mesure de probabilité sur $E$ telle que $\delta_i(j) = 1$ ssi $i = j$ (et par conséquent $\delta_i(j) = 0$ si $i \neq j$). On notera $\mathbf{P}_{\delta_i}$ simplement par $\mathbf{P}_i$. L'espérance associée à $\mathbf{P}_i$ sera notée $\mathbf{E}_i$ au lieu de $\mathbf{E}_{\delta_i}$. □

**Théorème II.1.16** **(Propriété de Markov simple)** *On se donne $E$ un espace dénombrable discret et un processus markovien, noté*

$$\big(\,\Omega\,;\, \mathscr{F}\,;\, (\mathscr{F}_t)_{t\in\mathbb{R}_+}\,;\, \mathbf{X} = (X_t)_{t\in\mathbb{R}_+}\,;\, (P_t)_{t\in\mathbb{R}_+}\,;\, \mathbf{P}_\mu \,,\, \mu \in \mathscr{M}_1(E)\,\big) \;.$$

*Pour tout $\mu \in \mathscr{M}_1(E)$, pour tout $t \in \mathbb{R}_+$ et pour toute fonction $F : E^{\mathbb{R}_+} \to \mathbb{R}$, $\mathscr{P}(E)^{\otimes}$-mesurable bornée, on a l'égalité suivante, appelée* propriété de Markov au temps *$t$ :*

$$\mathbf{P}_\mu\text{-}p.s. \quad \mathbf{E}_\mu\big[F(\theta_t \mathbf{X}\,) \,\big|\, \mathscr{F}_t\,\big] = \mathbf{E}_{X_t}[F(\mathbf{X})] \;. \tag{II.6}$$

**Preuve :** remarquons tout d'abord que $\omega \in \Omega \mapsto F(\theta_t \mathbf{X}(\omega)) \in \mathbb{R}$ est une application $\mathscr{F}$-mesurable d'après le lemme II.1.4. D'autre part, si pour tout $i \in E$, on pose $\phi(i) = \mathbf{E}_i[F(\mathbf{X})]$, alors $\phi(X_t) = \mathbf{E}_{X_t}[F(\mathbf{X})]$ est $\mathscr{F}_t$-mesurable. On fixe $\mu \in \mathscr{M}_1(E)$ et $t \in \mathbb{R}_+$. On se donne aussi $C \in \mathscr{C}$ : il existe donc $0 \leq t_0 < \ldots < t_n$ et $i_0, \ldots, i_n \in E$ tel que $C = \{\mathbf{x} = (x(t))_{t\in\mathbb{R}_+} \in E^{\mathbb{R}_+} \;:\; x(t_0) = i_0; \ldots ; x(t_n) = i_n\;\}$. On montre d'abord la propriété de Markov simple pour $F = \mathbf{1}_C$.

Pour cela on observe que $F(\theta_t \mathbf{X}) = \mathbf{1}_{\{X_{t+t_0}=i_0;\ldots;X_{t+t_n}=i_n\}}$. Comme $\mathbf{X}$ est $(\mathscr{F}_s)_{s\in\mathbb{R}_+}$-adapté, $\{X_{t+t_0} = i_0; \ldots ; X_{t+t_{n-1}} = i_{n-1}\} \in \mathscr{F}_{t+t_{n-1}}$ et on a $\mathbf{P}_\mu$-p.s.

$$\begin{aligned}
\mathbf{E}_\mu\big[F(\theta_t \mathbf{X}) \,\big|\, \mathscr{F}_{t+t_{n-1}}\,\big] &= \mathbf{1}_{\{X_{t+t_0}=i_0;\ldots;X_{t+t_{n-1}}=i_{n-1}\}} p_{t_n - t_{n-1}}(X_{t+t_{n-1}}, i_n) \\
&= \mathbf{1}_{\{X_{t+t_0}=i_0;\ldots;X_{t+t_{n-1}}=i_{n-1}\}} p_{t_n - t_{n-1}}(i_{n-1}, i_n) \;.
\end{aligned}$$

En itérant ce raisonnement, on obtient $\mathbf{P}_\mu$-p.s.

$$\mathbf{E}_\mu\big[F(\theta_t \mathbf{X}) \,\big|\, \mathscr{F}_{t+t_{n-2}}\,\big] = \mathbf{1}_{\{X_{t+t_0}=i_0;\ldots;X_{t+t_{n-2}}=i_{n-2}\}} p_{t_{n-2}-t_{n-1}}(i_{n-2}, i_{n-1}) p_{t_n - t_{n-1}}(i_{n-1}, i_n)$$

et en poursuivant, on aboutit à

$$\mathbf{P}_\mu\text{-p.s.} \quad \mathbf{E}_\mu\big[F(\theta_t \mathbf{X}) \,\big|\, \mathscr{F}_t\,\big] = p_{t_0}(X_t, i_0) p_{t_1 - t_0}(i_0, i_1) \ldots p_{t_n - t_{n-1}}(i_{n-1}, i_n) \;. \tag{II.7}$$

On fixe $i \in E$, on choisit $\mu = \delta_i$ et $t = 0$ dans l'expression précédente et on intégre l'expression obtenue, en notant bien que, sous $\mathbf{P}_i(X_0 = i) = 1$. On obtient alors

$$\mathbf{P}_i(\mathbf{X} \in \mathbf{C}) = \mathbf{E}_i[F(\mathbf{X})] = p_{t_0}(i, i_0) p_{t_1 - t_0}(i_0, i_1) \ldots p_{t_n - t_{n-1}}(i_{n-1}, i_n) \;.$$

Ceci combiné à (II.7) avec $\mu$ et $t$ quelconques, montre que

$$\mathbf{P}_\mu\text{-p.s.} \quad \mathbf{E}_\mu\big[F(\theta_t \mathbf{X}) \,\big|\, \mathscr{F}_t\,\big] = \mathbf{E}_{X_t}[F(\mathbf{X})] \;,$$

ce qui est bien la propriété de Markov pour $F = \mathbf{1}_C$.

Le cas général s'obtient par l'utilisation de la *version fonctionnelle du théorème de la classe monotone* : on fixe $\mu \in \mathscr{M}_1(E)$ et $t \in \mathbb{R}_+$ et on note $\mathscr{H}$ l'ensemble des fonctions $F : E^{\mathbb{R}_+} \to \mathbb{R}$, $\mathscr{P}(E)^{\otimes \mathbb{R}_+}$-mesurables bornées qui satisfont la propriété de Markov simple (II.6). On a montré que $\mathbf{1}_C \in \mathscr{H}$, pour tout cylindre élémentaire $C \in \mathscr{C}$. Il est facile à l'aide du théorème de convergence dominée conditionnelle de montrer que $\mathscr{H}$ satisfait les hypothèses de la version fonctionnelle du théorème de la classe monotone (théorème A.1.20, page 273). En vertu de ce théorème on conclut que $\mathscr{H}$ contient toute fonction $F : E^{\mathbb{R}_+} \to \mathbb{R}$ qui est $\sigma(\mathscr{C})$-mesurable bornée. Or, par définition, $\sigma(\mathscr{C}) = \mathscr{P}(E)^{\otimes \mathbb{R}_+}$, ce qui termine la preuve du théorème. ■

### II.1.d Trajectoires régulières, décomposition jusqu'au temps d'explosion éventuelle.

Il est nécessaire d'aborder une difficulté théorique : les diverses définitions que nous avons introduites précédemment n'éliminent pas un certain nombre de processus pathologiques. Examinons l'exemple suivant : on pose $E = \{0, 1\}$, et on fixe $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité. On suppose que pour tout $t \in \mathbb{R}_+$, il existe une variable $X_t : \Omega \to \{0, 1\}$, qui est $\mathscr{F}$-mesurable et on fait l'hypothèse que $\mathbf{P}(X_0 = 0) = 1$, que les variables $X_t$, $t \in \mathbb{R}_+$, sont mutuellement indépendantes et que pour tout $t > 0$, on a $\mathbf{P}(X_t = 0) = \mathbf{P}(X_t = 1) = 1/2$. Le lemme II.1.2 assure que $\mathbf{X} = (X_t)_{t \in \mathbb{R}_+} : \Omega \to \{0, 1\}^{\mathbb{R}_+}$ est bien $(\mathscr{F}, \mathscr{P}(\{0, 1\})^{\otimes \mathbb{R}_+})$-mesurable. De plus, il est facile de vérifier que $\mathbf{X} = (X_t)_{t \in \mathbb{R}_+}$ est un processus markovien de semi-groupe $P_t = (p_t(i, j))_{i,j \in \{0,1\}}$, $t \in \mathbb{R}_+$, donné par $p_t(0, 1) = p_t(1, 0) = p_t(0, 0) = p_t(1, 1) = 1/2$. (voir l'exemple II.1.8 page 149). Ce processus de Markov très simple cadre avec les définitions que nous avons données mais il est extrémement irrégulier. Son existence même suppose que l'espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$ soit de. cardinal extrèmement très gros. La trajectoire de $\mathbf{X}$ est très irrégulière et nous laissons le lecteur s'interroger sur le sous-ensemble aléatoire $S \subset \mathbb{R}_+$ des zéros de $\mathbf{X}$ : $S = \{t \in \mathbb{R}_+ : \ X_t = 0\}$.

Pour éviter, ce genre de processus très irréguliers, il est nécessaire de se restreindre à une classe de processus dont la trajectoire est *régulière*, Plus précisément nous allons nous restreindre à des processus à valeurs dans $E$ qui sont *continus à droite*.

**Définition II.1.17** (*Trajectoires continues à droites*) Soit $E$, un ensemble dénombrable non-vide. On le voit comme un espace topologique discret : on le munit de la distance $d$ définie pour tous $i, j \in E$, par $d(i, j) = 1$ si $i \neq j$ et $d(i, i) = 0$ ; la topologie de $(E, d)$ est alors la topologie discrète. Comme déjà mentionné la tribu des Boréliens de $(E, d)$ est la tribu de tous les sous-ensembles de $E$, notée $\mathscr{P}(E) := \{B \subset E\}$.

Soit $\mathbf{x} = (x(t))_{t \in \mathbb{R}_+} \in E^{\mathbb{R}_+}$, une fonction de $\mathbb{R}_+$ dans $E$. On dit qu'elle est continue à droite (*càd* en abrégé) si et seulement si pour tout $t \in \mathbb{R}_+$, on a $\lim_{h>0\, ,\, h \to 0} d(x(t), x(t + h)) = 0$. On voit facilement que cela est équivalent à la définition suivante :

$$\mathbf{x} = (x(t))_{t \in \mathbb{R}_+} \in E^{\mathbb{R}_+} \text{ est càd ssi } \quad \forall\, t \in \mathbb{R}_+ \ , \ \exists h_t > 0 \ : \ x(t) = x(t + u) \ , \ u \in [0, h_t] \ . \qquad \text{(II.8)}$$

Autrement dit, une fonction de $\mathbb{R}_+$ à valeurs dans $E$ dénombrable discret est continue à droite si et seulement si pour tout $t \in \mathbb{R}_+$ elle reste constante après $t$ pendant un intervalle de temps strictement positif. □

L'exemple suivant montre que les fonctions continues à droites à valeurs dans un espace dénombrable discret peuvent être néanmoins compliquées.

**Exemple II.1.18** On choit $E = \mathbb{Z}$, discret. On définit une fonction $\mathbf{x} = (x(t))_{t \in \mathbb{R}_+}$, une fonction 1-périodique en posant pour tout $t \in [0, 1[$

$$x(t) = (-1)^n n \quad \text{si } t \in [1 - 2^{-n}, 1 - 2^{-n-1}[\, , n \in \mathbb{N}\ .$$

On voit que $\mathbf{x}$ est bien contiue à droite. En revanche, elle n'admet aucune limite à gauche aux temps $t$ qui sont des entiers strictement positifs :

$$\forall p \in \mathbb{N}^*, \quad \liminf_{\substack{h \to 0 \\ h>0}} x(p-h) = -\infty \quad \text{et} \quad \limsup_{\substack{h \to 0 \\ h>0}} x(p-h) = +\infty .$$

Une fonction continue à droite n'admet pas forcément des limites à gauche et même si elle progresse par palliers successifs, les temps où elle saute d'une valeur à l'autre peuvent s'accumuler : on appelle ce phénomène *l'explosion.* □

**Définition II.1.19** (*Processus continus à droite*) Soit $E$, un espace dénombrable discret. Soit $(\Omega, \mathscr{F})$, un espace mesurable. Pour tout $t \in \mathbb{R}_+$, soit $X_t : \Omega \to E$, une variable $\mathscr{F}$-mesurable. On dit que le processus $\mathbf{X} = (X_t)_{t \in \mathbb{R}_+}$ est un *processus continu à droite* (partout sur $\Omega$) si et seulement si

$$\forall\, \omega \in \Omega, \quad t \in \mathbb{R}_+ \longmapsto X_t(\omega) \in E \ \text{ est une fonction continue à droite.}$$

□

On fixe $\mathbf{x} = (x(t))_{t \in \mathbb{R}_+} \in E^{\mathbb{R}_+}$, une fonction continue à droite et on s'intéresse ensuite à la trajectoire de $\mathbf{x}$ jusqu'à son premier temps d'explosion éventuelle. Pour décomposer cette trajectoire, nous introduisons les notations suivantes qui seront utilisées *dans tout le reste du chapitre*.

$(a)$ (*Temps de saut*) On pose $J_0(\mathbf{x}) = 0$ et on définit récursivement la suite $J_n(\mathbf{x}) \in [0, \infty]$, $n \in \mathbb{N}$ par :

$$J_{n+1}(\mathbf{x}) = \inf\{t > J_n(\mathbf{x}) \ : \ x(t) \neq x(J_n(\mathbf{x}))\} ,$$

avec la convention que $\inf \emptyset = \infty$. La suite $(J_n(\mathbf{x}))_{n \geq 1}$ est la *suite des temps de sauts de* $\mathbf{x}$ *avant la première explosion*. On observe que si $J_n(\mathbf{x}) < \infty$ mais $J_{n+1}(\mathbf{x}) = \infty$, alors $x(t) = x(J_n(\mathbf{x}))$, pour tout temps $t \geq J_n(\mathbf{x})$. On dit que la trajectoire est *absorbée* en $x(J_n(\mathbf{x}))$, et on a alors $J_p(\mathbf{x}) = \infty$, pour tout $p \geq n+1$.

$(b)$ (*Temps d'explosion*) On pose

$$\zeta(\mathbf{x}) = \sup_{n \geq 0} J_n(\mathbf{x}) \ \in [0, \infty] .$$

C'est le premier *temps d'explosion de* $\mathbf{x}$. Si $\zeta(\mathbf{x}) = \infty$, on dit que $\mathbf{x}$ n'explose pas.

$(c)$ On note $N(\mathbf{x}) = \sup\{n \geq 0 \ : \ J_n(\mathbf{x}) < \infty\}$ qui est *le nombre de sauts avant absorption éventuelle*. Donc $N(\mathbf{x}) < \infty$ si et seulement si la trajectoire $\mathbf{x}$ est absorbée en un état.

$(d)$ (*Squelette*) On note les valeurs successives de $\mathbf{x}$ avant l'explosion par

$$Y_n(\mathbf{x}) = x(J_n(\mathbf{x})) \quad \text{si } n \leq N(\mathbf{x}) \quad \text{et} \quad Y_n(\mathbf{x}) = x(J_{N(\mathbf{x})}(\mathbf{x})) \quad \text{si } n > N(\mathbf{x}) ,$$

la seconde égalité étant une convention qui s'avère commode en cas d'absorption. La suite

$$\mathbf{Y}(\mathbf{x}) = (Y_n(\mathbf{x}))_{n \geq 0}$$

est appelée le *squelette de* $\mathbf{x}$ (ou la chaîne (ou suite) des sauts). On remarque que si $N(\mathbf{x}) < \infty$ (c'est-à-dire si la trajectoire est aborbée), la suite $(Y_n(\mathbf{x}))_{n \geq 0}$ est constante à partir d'un certain rang.

$(e)$ (*Durées d'attente*) On définit la suite $D_n(\mathbf{x}) \in [0, \infty]$, $n \geq 1$, des durées d'attente entre les sauts de $\mathbf{x}$ de la manière suivante :

$$D_{n+1}(\mathbf{x}) = J_{n+1}(\mathbf{x}) - J_n(\mathbf{x}) \quad \text{si } 0 \leq n \leq N(\mathbf{x}) \quad \text{et} \quad D_{n+1}(\mathbf{x}) = \infty \quad \text{si } n > N(\mathbf{x}) .$$

□

**Il est important de noter** que la donnée de $(D_n(\mathbf{x}))_{n\geq 0}$ et de $(Y_n(\mathbf{x}))_{n\geq 0}$ permet de reconstituer entièrement $\mathbf{x}$ sur l'intervalle de temps $[0, \zeta(\mathbf{x})[$ et que l'on a

$$\zeta(\mathbf{x}) = \sum_{n\geq 1} D_n(\mathbf{x}) \; .$$

En revanche ces deux suites ne disent rien sur la trajectoire après le temps $\zeta(\mathbf{x})$. Signalons dès maintenant que nous ne nous préoccuperons pas de l'étude des processus markoviens après leur premier temps d'explosion.

### II.1.e Propriété de Markov forte.

On restreint notre étude des processus de Markov à valeurs dans un espace dénombrable discret aux processus qui sont continus à droite et on va prouver qu'ils satisfont la propriété de Markov forte. Pour cela il est nécessaire de définir la notion de temps d'arrêt à valeurs dans $[0, \infty]$.

#### Temps d'arrêts.

Bien que très proche des temps d'arrêt discrets, la notion de temps d'arrêt à valeurs dans $[0, \infty]$ comportent quelques complications.

**Définition II.1.20** On fixe un espace mesurable $(\Omega, \mathscr{F})$ et une filtration $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$ sur cet espace. On pose également $\mathscr{F}_\infty = \sigma(\mathscr{F}_t, t \in \mathbb{R}_+)$, donc $\mathscr{F}_\infty \subset \mathscr{F}$. On rappelle la notation $\mathscr{F}_{t+} = \bigcap_{h>0} \mathscr{F}_{t+h}$ (qui en général est distincte de $\mathscr{F}_t$).

$(a)$ (*Temps d'arrêt*) Soit $T : \Omega \to [0, \infty]$. $T$ est un temps d'arrêt relativement à la filtration $(\mathscr{F}_t)_{t\geq 0}$ ssi

$$\forall t \in \mathbb{R}_+ \; , \quad \{\omega \in \Omega : \, T(\omega) \leq t\} \in \mathscr{F}_t \; .$$

$(b)$ Soit $T$ un temps d'arrêt relativement à $(\mathscr{F}_t, t \geq 0)$. On pose

$$\mathscr{F}_T = \big\{ \, A \in \mathscr{F}_\infty : \, \forall \, t \in \mathbb{R}_+ \, , \; A \cap \{T \leq t\} \in \mathscr{F}_t \, \big\}$$

et

$$\mathscr{F}_{T+} = \big\{ \, A \in \mathscr{F}_\infty : \, \forall \, t \in \mathbb{R}_+ \, , \; A \cap \{T \leq t\} \in \mathscr{F}_{t+} \, \big\}.$$

Il est facile de vérifier que $\mathscr{F}_T$ et $\mathscr{F}_{T+}$ sont des sous-tribus de $\mathscr{F}_\infty$ et que $\mathscr{F}_T \subset \mathscr{F}_{T+}$. La tribu $\mathscr{F}_T$ est la tribu des évévement *antérieurs à* $T$. □

**Proposition II.1.21** *Soient $(\Omega, \mathscr{F})$, un espace mesurable et $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$, une filtration sur cet espace. On pose $\mathscr{F}_\infty = \sigma(\mathscr{F}_t, t \geq 0)$. On rappelle la notation $\mathscr{F}_{t+} = \bigcap_{h>0} \mathscr{F}_{t+h}$, $t \in \mathbb{R}_+$. On rappelle que $(\mathscr{F}_{t+})_{t\in\mathbb{R}_+}$ est une filtration sur $(\Omega, \mathscr{F})$ qui est appelée filtration continue à droite associée à $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$. Soient $S, T : \Omega \to [0, \infty]$, deux fonctions. Alors les assertions suivantes sont vraies.*

$(i)$ *La variable $T$ est un $(\mathscr{F}_{t+})_{t\geq 0}$-temps d'arrêt si et seulement si $\{T < t\} \in \mathscr{F}_t$, pour tout $t \in \mathbb{R}_+$.*

$(ii)$ *On suppose que $T$ est un $(\mathscr{F}_t)_{t\geq 0}$-temps d'arrêt. Alors,*

$$\mathscr{F}_{T+} = \big\{ \, A \in \mathscr{F}_\infty : \, \forall \, t \in \mathbb{R}_+ \, , \; A \cap \{T < t\} \in \mathscr{F}_t \, \big\} \; .$$

$(iii)$ *On suppose que $S$ et $T$ sont deux $(\mathscr{F}_t)_{t\geq 0}$-temps d'arrêt. Alors, les variables $S \wedge T$ et $S \vee T$ sont deux $(\mathscr{F}_t)_{t\geq 0}$-temps d'arrêt et $\mathscr{F}_{T\wedge S} = \mathscr{F}_S \cap \mathscr{F}_T$. Par conséquent, si $S(\omega) \leq T(\omega)$, pour tout $\omega \in \Omega$, alors $\mathscr{F}_S \subset \mathscr{F}_T$.*

*Soit $(S_n)_{n\geq 1}$, une suite de temps d'arrêt relativement à $(\mathscr{F}_t)_{t\geq 0}$.*

$(iv)$ $S^* = \sup_{n\geq 1} S_n$ *est un $(\mathscr{F}_t)_{t\geq 0}$-temps d'arrêt.*

$(v)$ $S_* = \inf_{n\geq 1} S_n$ *est un $(\mathscr{F}_{t+})_{t\geq 0}$-temps d'arrêt et $\mathscr{F}_{S_*+} = \bigcap_{n\geq 1} \mathscr{F}_{S_n+}$.*

**Preuve :** on suppose tout d'abord que $T$ est un $(\mathscr{F}_{t+})_{t\geq 0}$-temps d'arrêt, c'est-à-dire que pour tout $s \in \mathbb{R}_+$, $\{T \leq s\} \in \mathscr{F}_{s+}$. On fixe $t > 0$ et on fixe une suite strictement croissante $t_n < t$, $n \in \mathbb{N}$ telle que $\lim_n t_n = t$. On a $\{T \leq t_n\} \in \mathscr{F}_{t_n+}$. Mais on observe facilement que $\mathscr{F}_{t_n+} \subset \mathscr{F}_t$. Donc

$$\{T < t\} = \bigcup_{n\geq 0} \{T \leq t_n\} \ \in \mathscr{F}_t \ .$$

Par ailleurs, on observe que $\{T < 0\} = \emptyset \in \mathscr{F}_0$. On a donc montré que si $T$ est un $(\mathscr{F}_{t+})_{t\geq 0}$-temps d'arrêt, alors $\{T < t\} \in \mathscr{F}_t$, pour tout $t \in \mathbb{R}_+$.

Montrons la réciproque et pour cela supposons que $\{T < t\} \in \mathscr{F}_t$, pour tout $t \in \mathbb{R}_+$. On fixe $t \in \mathbb{R}_+$ et une suite $t_n > t$, $n \in \mathbb{N}$, qui décroît strictement vers $t$. On fixe $h > 0$. Pour tout $n \in \mathbb{N}$ tel que $t_n \leq t + h$, on a donc $\{T < t_n\} \in \mathscr{F}_{t_n} \subset \mathscr{F}_{t+h}$. Cela implique donc que

$$\forall\, h > 0\, , \quad \{T \leq t\} = \bigcap_{n\geq 0} \{T < t_n\} \ \in \mathscr{F}_{t+h} \ .$$

Par conséquent, pour tout $t \in \mathbb{R}_+$, on a $\{T \leq t\} \in \bigcap_{h>0} \mathscr{F}_{t+h} = \mathscr{F}_{t+}$, ce qui termine la preuve de $(i)$.

Prouvons le second point. Soit $A \in \mathscr{F}_{T+}$. On fixe $t > 0$, et une suite strictement croissante $t_n < t$, $n \in \mathbb{N}$, telle que $\lim_n t_n = t$. Alors $A \cap \{T \leq t_n\} \in \mathscr{F}_{t_n+} \subset \mathscr{F}_t$ et donc

$$A \cap \{T < t\} = \bigcup_{n\geq 1} A \cap \{T \leq t_n\} \ \in \mathscr{F}_t \ .$$

Réciproquement, supposons que $A \in \mathscr{F}_\infty$ soit tel que pour tout $s \in \mathbb{R}_+$, on ait $A \cap \{T < s\} \in \mathscr{F}_s$. On fixe $t \in \mathbb{R}_+$ et une suite $t_n > t$, $n \in \mathbb{N}$, qui décroît strictement vers $t$. On fixe $h > 0$. Pour tout $n \in \mathbb{N}$ tel que $t_n \leq t + h$, on a $A \cap \{T < t_n\} \in \mathscr{F}_{t+h}$ et donc

$$\forall\, h > 0\, , \quad A \cap \{T \leq t\} = \bigcap_{n\geq 0} A \cap \{T < t_n\} \ \in \mathscr{F}_{t+h} \ .$$

Par conséquent, pour tout $t \in \mathbb{R}_+$, on a $A \cap \{T \leq t\} \in \bigcap_{h>0} \mathscr{F}_{t+h} = \mathscr{F}_{t+}$, ce qui termine la preuve de $(ii)$.

Montrons le point $(iii)$. Il est clair que pour tout $t \in \mathbb{R}_+$, $\{S \wedge T > t\} = \{T > t\} \cap \{S > t\}$ qui est bien dans $\mathscr{F}_t$ car le complémentaire de $\{T > t\}$ et de $\{S > t\}$ y sont, par définition. Par passage au complémentaire cela implique que $\{S \wedge T \leq t\} \in \mathscr{F}_t$, pour tout $t \in \mathbb{R}_+$ et donc que $S \wedge T$ est un $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$-temps d'arrêt. Le cas de $S \vee T$ se traite de façon similaire.

Soit $A \in \mathscr{F}_{S\wedge T}$. Alors $A \cap \{S \wedge T \leq t\} \in \mathscr{F}_t$. Puisque $\{T \leq t\} \in \mathscr{F}_t$, on remarque que $A \cap \{T \leq t\} = A \cap \{S \wedge T \leq t\} \cap \{T \leq t\} \in \mathscr{F}_t$. Cela implique que $A \in \mathscr{F}_T$. On montre de même que $A \in \mathscr{F}_S$ et donc que $A \in \mathscr{F}_S \cap \mathscr{F}_T$, c'est-à-dire que $\mathscr{F}_{S\wedge T} \subset \mathscr{F}_S \cap \mathscr{F}_T$. Montrons l'inclusion inverse et pour cela supposons que $A \in \mathscr{F}_S \cap \mathscr{F}_S$. Cela implique que pour tout $t \in \mathbb{R}_+$, on a $A \cap \{T \leq t\}$ et $A \cap \{S \leq t\}$ sont dans $\mathscr{F}_t$. Par conséquent $A \cap \{S \wedge T \leq t\} = (A \cap \{S \leq t\}) \cup (A \cap \{T \leq t\} \in \mathscr{F}_t$, ce qui implique que $A \in \mathscr{F}_{S\wedge T}$ et donc $\mathscr{F}_S \cap \mathscr{F}_T \subset \mathscr{F}_{S\wedge T}$, ce qui termine la preuve de $(iii)$.

Le point $(iv)$ est un conséquente directe du fait que pour tout $t \in \mathbb{R}_+$, $\{S^* \leq t\} = \bigcap_{n\geq 0}\{S_n \leq t\}$. Le point $(v)$ est un peu plus compliqué. On observe que pour tout $t \in \mathbb{R}_+$ et pour tout $n \in \mathbb{N}$, on a $\{S_n < t\} \subset \{S_n \leq t\} \in \mathscr{F}_t$. Par conséquent

$$\{S_* < t\} = \bigcup_{n\geq 0} \{S_n < t\} \ \in \mathscr{F}_t \ ,$$

ce qui montre que $S_*$ est un $(\mathscr{F}_{t+})_{t\geq 0}$-temps d'arrêt par $(i)$. On se donne ensuite $A \in \mathscr{F}_{S_*+}$. Par $(ii)$, cela est équivalent à ce que pour tout $t \in \mathbb{R}_+$, on ait $A \cap \{S_* < t\} \in \mathscr{F}_t$. On remarque ensuite que

$$\forall\, n \in \mathbb{N}\,, \quad A \cap \{S_n < t\} = (A \cap \{S_* < t\}) \cap \{S_n < t\} \in \mathscr{F}_t\ .$$

Cela montre que si $A \in \mathscr{F}_{S_*+}$, alors $A \in \mathscr{F}_{S_n+}$, pour tout $n \in \mathbb{N}$. Par conséquent $\mathscr{F}_{S_*+} \subset \bigcap_{n\geq 0} \mathscr{F}_{S_n+}$.

Montrons l'inclusion contraire et supposons que $A \in \bigcap_{n\geq 0} \mathscr{F}_{S_n+}$. Par $(ii)$, pour tout $n \in \mathbb{N}$, et pour tout $t \in \mathbb{R}_+$, on a $A \cap \{S_n < t\} \in \mathscr{F}_t$. Par conséquent

$$A \cap \{S_* < t\} = \bigcup_{n\geq 1} A \cap \{S_n < t\} \ \in \mathscr{F}_t\ ,$$

ce qui montre, en utilisant encore $(ii)$, que $A \in \mathscr{F}_{S_*+}$ : cela implique donc que $\bigcap_{n\geq 0} \mathscr{F}_{S_n+} \subset \mathscr{F}_{S_*+}$. ■

La proposition suivante fournit des exemples de temps d'arrêts utiles.

**Proposition II.1.22** *Soit $E$, un espace dénombrable discret. Soient $(\Omega, \mathscr{F})$, un espace mesurable et $(\mathscr{F}_t)_{t\geq 0}$, une filtration sur cet espace. Pour tout $t \in \mathbb{R}_+$, on se donne une variable $X_t : \Omega \to E$ qui est $\mathscr{F}_t$-mesurable. Autrement dit le processus $\mathbf{X} = (X_t)_{t\in\mathbb{R}_+}$ est $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$-adapté. On suppose de plus que $\mathbf{X}$ est* continu à droite *partout sur $\Omega$. Alors les assertions suivantes sont vérifiées.*

$(i)$ *Le premier temps de saut de $\mathbf{X}$, noté $J_1(\mathbf{X})$, est un $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$-temps d'arrêt. Plus généralement, pour tout $n \in \mathbb{N}$, le $n$-ième temps de saut $J_n(\mathbf{X})$ est aussi un $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$-temps d'arrêt, ainsi que le temps d'explosion $\zeta(\mathbf{X})$.*

$(ii)$ *On fixe $i \in E$. On pose*

$$T_i(\mathbf{X}) = \inf\{t \in \mathbb{R}_+\,,\, X_t = i\} \quad \text{et} \quad T_i^+(\mathbf{X}) = \inf\{t \geq J_1(\mathbf{X})\ :\ X_t = i\}\ ,$$

*avec la convention que* $\inf \emptyset = \infty$. *Alors $T_i(\mathbf{X})$ et $T_i^+(\mathbf{X})$ sont deux $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$-temps d'arrêt : $T_i(\mathbf{X})$ est le* premier temps d'atteinte de $i$ *et $T_i^+(\mathbf{X})$ est le* premier temps de retour en $i$.

$(iii)$ *Soit $K \subset E$. On pose $T_K(\mathbf{X}) = \inf\{t \in \mathbb{R}_+\,,\, X_t \in K\}$, avec la convention que* $\inf \emptyset = \infty$. *Alors, $T_K(\mathbf{X})$ est un $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$-temps d'arrêt : $T_K(\mathbf{X})$ est le* premier temps d'atteinte de l'ensemble $K$.

**Preuve :** comme $\mathbf{X}$ est continu à droite, on a pour tout $t \in \mathbb{R}_+$,

$$\{J_1(\mathbf{X}) > t\} = \{\forall s \in [0,t]\ ,\ X_s = X_0\} = \{\forall r \in ([0,t] \cap \mathbb{Q}) \cup \{t\}\ ,\ X_r = X_0\} = \bigcap_{r\in([0,t]\cap\mathbb{Q})\cup\{t\}} \{X_r = X_0\}.$$

Or pour tout $r \in ([0,t] \cap \mathbb{Q}) \cup \{t\}$, $\{X_r = X_0\} \in \mathscr{F}_r \subset \mathscr{F}_t$, car $\mathbf{X}$ est $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$-adapté. Donc $\{J_1(\mathbf{X}) > t\} \in \mathscr{F}_t$, ainsi que son complémentaire. Cela montre que $J_1(\mathbf{X})$ est un $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$-temps d'arrêt. Plus généralement, on montre que pour tout $n \in \mathbb{N}$ et pour tout $t \in \mathbb{R}_+$, on a

$$\begin{aligned} \{J_n(\mathbf{X}) \leq t\} &= \big\{\exists\, 0 < s_1 < \ldots < s_n \leq t\ :\ X_0 \neq X_{s_1}; \ldots ; X_{s_{n-1}} \neq X_{s_n}\big\} \\ &= \bigcap_{\substack{r_1,\ldots,r_n\in([0,t]\cap\mathbb{Q})\cup\{t\} \\ 0<r_1<\ldots<r_n}} \big\{\, X_0 \neq X_{r_1}; \ldots ; X_{r_{n-1}} \neq X_{r_n}\,\big\} \end{aligned}$$

ce qui entraîne que $J_n(\mathbf{X})$ de $X$ est un $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$-temps d'arrêt (on a également utilisé la continuité à droite pour passer à la seconde égalité). La proposition II.1.21 $(iv)$ enfin implique que $\zeta(\mathbf{X}) = \sup_{n\geq 0} J_n(\mathbf{X})$ est également un $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$-temps d'arrêt. Cela montre le point $(i)$.

Montrons une partie du point $(ii)$ et le point $(iii)$ : pour tout $t \in \mathbb{R}_+$, on remarque que

$$\{T_K(\mathbf{X}) \leq t\} = \bigcup_{i \in K} \bigcup_{r \in (]0,t] \cap \mathbb{Q}) \cup \{t\}} \{X_r = i\} \; .$$

Or pour tout $r \in ([0,t] \cap \mathbb{Q}) \cup \{t\}$, $\{X_r = i\} \in \mathscr{F}_r \subset \mathscr{F}_t$, car $\mathbf{X}$ est $(\mathscr{F}_t)_{t \in \mathbb{R}_+}$-adapté. Donc $\{T_K(\mathbf{X}) \leq t\} \in \mathscr{F}_t$, pour tout $t \in \mathbb{R}_+$. Cela montre que $T_K(\mathbf{X})$ est un $(\mathscr{F}_t)_{t \in \mathbb{R}_+}$-temps d'arrêt. Si on prend $K = \{i\}$, cela implique que $T_i(\mathbf{X})$ est un $(\mathscr{F}_t)_{t \in \mathbb{R}_+}$-temps d'arrêt.

On observe ensuite que $\{T_i^+(\mathbf{X}) \leq 0\} = \emptyset \in \mathscr{F}_0$ et que pour tout $t > 0$,

$$\{T_i^+(\mathbf{X}) \leq t\} = \bigcup_{\substack{r,r' \in (]0,t] \cap \mathbb{Q}) \cup \{t\} \\ r' < r}} \{X_{r'} \neq X_r ; X_r = i\} \; ,$$

ce qui entraîne également que $T_i^+(\mathbf{X})$ est un $(\mathscr{F}_t)_{t \in \mathbb{R}_+}$-temps d'arrêt. ■

**Propriété de Markov forte.**

**Notations.** On fixe $(\Omega, \mathscr{F})$, un espace mesurable et $(\mathscr{F}_t)_{t \in \mathbb{R}_+}$, une filtration sur cet espace. Pour tout $t \in \mathbb{R}_+$, on se donne une variable $X_t : \Omega \to E$ qui est $\mathscr{F}$-mesurable. On fixe $i^* \in E$, qui ne joue aucun rôle spécifique dans la suite. On fixe également une fonction $T : \Omega \to [0, \infty]$. On introduit alors les notations suivantes. Pour tout $\omega \in \Omega$,

$$X_T(\omega) = \begin{cases} X_{T(\omega)}(\omega) & \text{si } T(\omega) < \infty, \\ i^* & \text{si } T(\omega) = \infty \end{cases}$$

On note $\mathbf{1}_{i^*} : \mathbb{R}_+ \to E$, la fonction constante à $i^*$ et pour tout $\omega \in \Omega$, on pose ensuite

$$\theta_T \mathbf{X}(\omega) = \begin{cases} (X_{T(\omega)+t}(\omega))_{t \in \mathbb{R}_+} & \text{si } T(\omega) < \infty, \\ \mathbf{1}_{i^*} & \text{si } T(\omega) = \infty \end{cases}$$

On énonce la propriété de Markov forte comme suit.

**Théorème II.1.23** **(Propriété de Markov forte)** *On se donne $E$ un espace dénombrable discret et un processus markovien, noté*

$$\big( \Omega \, ; \, \mathscr{F} \, ; \, (\mathscr{F}_t)_{t \in \mathbb{R}_+} \, ; \, \mathbf{X} = (X_t)_{t \in \mathbb{R}_+} \, ; \, (P_t)_{t \in \mathbb{R}_+} \, ; \, \mathbf{P}_\mu \, , \, \mu \in \mathscr{M}_1(E) \big) \; .$$

*On suppose que $\mathbf{X}$ est* continu à droite. *Soit $T$ un $(\mathscr{F}_t)_{t \in \mathbb{R}_+}$-temps d'arrêt. Alors les assertions suivantes sont vraies.*

$(i)$ *La variable aléatoire $X_T : \Omega \to E$ est $\mathscr{F}_T$-mesurable.*

$(ii)$ *La trajectoire aléatoire $\theta_T \mathbf{X} : \Omega \to E^{\mathbb{R}_+}$ est $(\mathscr{F}, \mathscr{P}(E)^{\otimes \mathbb{R}_+})$-mesurable.*

$(iii)$ *Pour toute loi initiale $\mu \in \mathscr{M}_1(E)$ et toute fonction $F : E^{\mathbb{R}_+} \to \mathbb{R}$ qui est $\mathscr{P}(E)^{\otimes \mathbb{R}_+}$-mesurable bornée, on a*

$$\mathbf{P}_\mu\text{-}p.s. \quad \mathbf{E}_\mu \left[ \mathbf{1}_{\{T < \infty\}} F(\theta_T \mathbf{X}) \mid \mathscr{F}_{T+} \right] = \mathbf{1}_{\{T < \infty\}} \mathbf{E}_{X_T} \left[ F(\mathbf{X}) \right] \; . \tag{II.9}$$

*Comme $\mathscr{F}_T \subset \mathscr{F}_{T+}$ et comme $X_T$ est $\mathscr{F}_T$-mesurable, un résultat analogue a lieu en conditionnant par rapport à $\mathscr{F}_T$.*

**Preuve :** pour prouver $(i)$, on utilise l'approximation suivante du temps d'arrêt $T$ : pour tout réel positif $a$, on note $\lceil a \rceil$ le plus petit entier qui soit strictement plus grand que $a$ ($\lceil a \rceil = n$ si $n-1 \leq a < n$) c'est-à-dire que $\lceil a \rceil$ est la partie entière de $a$ augmentée de $1$ ; pour tout $n \in \mathbb{N}$, on pose $T_n = 2^{-n}\lceil 2^n T_n \rceil$, si bien que $T_n = (k+1)2^{-n}$ si et seulement si $k2^{-n} \leq T < (k+1)2^{-n}$ ; on observe alors que $T_n = \infty$ ssi $T = \infty$ et que $(T_n)_{n\geq 0}$ est une suite décroissante de temps qui converge vers $T$. La continuité à droite de $X$ implique alors que pour tout $\omega \in \Omega$, $X_{T_n}(\omega)$ converge vers $X_T(\omega)$ lorsque $n$ tend vers l'infini. Comme $E$ est discret cela implique que la suite $(X_{T_n}(\omega))_{n\geq 0}$ stationne à $X_T(\omega)$ à partir d'un certain rang. On fixe $t \in \mathbb{R}_+$ et on observe que

$$\{T_n < t\} \cap \{X_{T_n} = i\} = \bigcup_{1\leq k < 2^n t} \{X_{k2^{-n}} = i\} \cap \{(k-1)2^{-n} \leq T < k2^{-n}\} \, .$$

Or $\{(k-1)2^{-n} \leq T < k2^{-n}\} \in \mathscr{F}_{k2^{-n}}$ et si $k < 2^n t$, alors $\mathscr{F}_{k2^{-n}} \subset \mathscr{F}_t$. Par conséquent $\{T_n < t\} \cap \{X_{T_n} = i\} \in \mathscr{F}_t$. On observe ensuite que comme $\mathbf{X}$ est continu à droite, on a

$$\{T < t\} \cap \{X_T = i\} = \bigcap_{p\geq 0} \bigcup_{n\geq p} \{T_n < t\} \cap \{X_{T_n} = i\}$$

et ce qui précède montre que $\{T < t\} \cap \{X_T = i\} \in \mathscr{F}_t$. On note ensuite que $\{T = t\} \in \mathscr{F}_t$ et donc que $\{T = t\} \cap \{X_t = i\} \in \mathscr{F}_t$. Cela montre finalement que

$$\forall i \in E\,, \quad \forall t \in \mathbb{R}_+\,, \quad \{T \leq t\} \cap \{X_T = i\} = (\{T < t\} \cap \{X_T = i\}) \cup (\{T = t\} \cap \{X_t = i\}) \in \mathscr{F}_t \, .$$

Cela entraîne que $\{X_T = i\} \in \mathscr{F}_T$, pour tout $i \in E$ et donc que $X_T$ est $\mathscr{F}_T$-mesurable, ce qui prouve $(i)$. Le point $(ii)$ ne présente aucune difficulté : il est laissé en exercice au lecteur.

Montrons le point $(iii)$. On fixe $\mu \in \mathscr{M}_1(E)$ et on prouve la propriété de Markov forte pour une fonction de la forme $F = \mathbf{1}_C$ où $C \in \mathscr{C}$ est un cylindre élémentaire. Il existe donc $n \in \mathbb{N}$, $0 \leq t_1 < \ldots < t_n$ dans $\mathbb{R}_+$ et $i_1, \ldots, i_n \in E$, tels que $C = \{\mathbf{x} \in E^{\mathbb{R}_+} : x(t_1) = i_1; \ldots; x(t_n) = i_n\}$. On rappelle que la suite de temps $(T_n)_{n\geq 0}$ approche $T$ en décroissant. Le fait que $\mathbf{X}$ soit continu à droite pour la topologie discrète sur $E$ implique facilement que

$$\forall \omega \in \Omega\,, \lim_{n\to\infty} \mathbf{1}_{\{T_n(\omega)<\infty\}} \mathbf{1}_C(\theta_{T_n}\mathbf{X}(\omega)) = \mathbf{1}_{\{T(\omega)<\infty\}} \mathbf{1}_C(\theta_T \mathbf{X}(\omega)) \tag{II.10}$$

et que

$$\forall \omega \in \Omega\,, \lim_{n\to\infty} \mathbf{1}_{\{T_n(\omega)<\infty\}} \mathbf{P}_{X_{T_n}(\omega)}(\mathbf{X} \in C) = \mathbf{1}_{\{T(\omega)<\infty\}} \mathbf{P}_{X_T(\omega)}(\mathbf{X} \in C) \, . \tag{II.11}$$

Ensuite on fixe $A \in \mathscr{F}_{T+}$. Alors (II.10) et (II.11) combinés avec le théorème de convergence dominée impliquent facilement

$$\begin{aligned} &\lim_{n\to\infty} \mathbf{E}_\mu[\mathbf{1}_A \mathbf{1}_{\{T_n<\infty\}} \mathbf{1}_C(\theta_{T_n} X)] = \mathbf{E}_\mu[\mathbf{1}_A \mathbf{1}_{\{T<\infty\}} \mathbf{1}_C(\theta_T X)] \quad \text{et} \\ &\qquad \lim_{n\to\infty} \mathbf{E}_\mu[\mathbf{1}_A \mathbf{1}_{\{T_n<\infty\}} \mathbf{P}_{X_{T_n}}(\mathbf{X} \in C)] = \mathbf{E}_\mu[\mathbf{1}_A \mathbf{1}_{\{T<\infty\}} \mathbf{P}_{X_T}(\mathbf{X} \in C)]. \end{aligned} \tag{II.12}$$

On observe alors que pour tout $n \in \mathbb{N}$,

$$\mathbf{E}_\mu[\mathbf{1}_A \mathbf{1}_{\{T_n<\infty\}} \mathbf{1}_C(\theta_{T_n} X)] = \sum_{k\geq 1} \mathbf{E}_\mu[\mathbf{1}_{A\cap\{(k-1)2^{-n}\leq T<k2^{-n}\}} \mathbf{1}_C(\theta_{k2^{-n}} X)] \, .$$

Puisque $A \in \mathscr{F}_{T+}$, l'événement $A \cap \{(k-1)2^{-n} \leq T < k2^{-n}\}$ appartient à la tribu $\mathscr{F}_{k2^{-n}}$. La propriété de Markov simple au temps $k2^{-n}$ implique alors

$$\mathbf{E}_\mu[\mathbf{1}_{A\cap\{(k-1)2^{-n}\leq T<k2^{-n}\}} \mathbf{1}_C(\theta_{k2^{-n}} X)] = \mathbf{E}_\mu[\mathbf{1}_{A\cap\{(k-1)2^{-n}\leq T<k2^{-n}\}} \mathbf{P}_{X_{k2^{-n}}}(\mathbf{X} \in C)].$$

Cela entraîne les égalités suivantes :

$$\begin{aligned}\mathbf{E}_\mu[\mathbf{1}_A\mathbf{1}_{\{T_n<\infty\}}\mathbf{1}_C(\theta_{T_n}X)] &= \sum_{k\geq 1}\mathbf{E}_\mu[\mathbf{1}_{A\cap\{(k-1)2^{-n}\leq T<k2^{-n}\}}\mathbf{P}_{X_{k2^{-n}}}(\mathbf{X}\in C)] \\ &= \mathbf{E}_\mu[\mathbf{1}_A\mathbf{1}_{\{T_n<\infty\}}\mathbf{P}_{X_{T_n}}(\mathbf{X}\in C)]\end{aligned}$$

et (II.12) implique alors que pour tout $A\in\mathscr{F}_{T+}$,

$$\mathbf{E}_\mu[\mathbf{1}_A\mathbf{1}_{\{T<\infty\}}\mathbf{1}_C(\theta_T X)] = \mathbf{E}_\mu[\mathbf{1}_A\mathbf{1}_{\{T<\infty\}}\mathbf{P}_{X_T}(\mathbf{X}\in C)].$$

On observe que $\mathbf{P}_{X_T}(\mathbf{X}\in C) = \mathbf{E}_{X_T}[\mathbf{1}_C(\mathbf{X})]$ est $\mathscr{F}_T$-mesurable, donc $\mathscr{F}_{T+}$-mesurable. Comme l'égalité précédente est vraie pour tout $A\in\mathscr{F}_{T+}$, on obtient

$$\mathbf{P}_\mu\text{–a.s.}\quad \mathbf{E}_\mu\left[\mathbf{1}_{\{T<\infty\}}\mathbf{1}_C(\theta_T\mathbf{X})\,|\,\mathscr{F}_{T+}\right] = \mathbf{1}_{\{T<\infty\}}\mathbf{E}_{X_T}[\mathbf{1}_C(\mathbf{X})],$$

ce qui prouve la propriété de Markov forte pour toute fonction de la forme $F=\mathbf{1}_C$, $C\in\mathscr{C}$.

Le résultat général s'obtient en raisonnant par classe monotone, comme dans la preuve de la propriété de Markov simple : On note $\mathscr{H}$ l'ensemble des fonctions $F: E^{\mathbb{R}_+}\to\mathbb{R}$ qui sont $\mathscr{P}(E)^{\otimes\mathbb{R}_+}$-mesurables bornées et qui satisfont (II.9). On vient de montrer que $\mathbf{1}_C\subset\mathscr{H}$, pour tout cylindre élémentaire $C\in\mathscr{C}$. Il est facile à l'aide du théorème de convergence dominée conditionnelle de montrer que $\mathscr{H}$ satisfait les hypothèses de la version fonctionnelle du théorème de la classe monotone (théorème A.1.20, page 273). En vertu de ce théorème on conclut que $\mathscr{H}$ contient toute fonction $F: E^{\mathbb{R}_+}\to\mathbb{R}$ qui est $\sigma(\mathscr{C})$-mesurable bornée. Or, par définition, $\sigma(\mathscr{C}) = \mathscr{P}(E)^{\otimes\mathbb{R}_+}$, ce qui termine la preuve du théorème. ■

### II.1.f Décomposition d'un processus markovien, générateur infinitésimal.

On rappelle les notations $J_n(\mathbf{x})$, $\zeta(\mathbf{x})$, $N(\mathbf{x})$, $Y_n(\mathbf{x})$ et $D_n(\mathbf{x})$, définies pour une trajectoire $\mathbf{x}$ continue à droite à valeurs dans un espace discret. On adopte les conventions suivantes sur les lois exponentielles : on dit qu'une variable nulle presque surement est une variable exponentielle de paramètre infini et qu'une variable infinie presque surement est une variable exponentielle de paramètre nul. On parle alors de variable exponentielle *étendue* pour signifier que son paramètre varie dans $[0,\infty]$. On utilise le fait suivant laissé en exercice : soit $(\Omega,\mathscr{F},\mathbf{P})$, un espace de probabilité sur lequel est définie $\mathbf{e}:\Omega\to[0,\infty]$, une variables aléatoire $\mathscr{F}$-mesurable.

$$\text{La v.a. } \mathbf{e} \text{ suit une loi exponentielle étendue} \iff \forall s,t\in\mathbb{R}_+,\ \mathbf{P}(\mathbf{e}>s+t)=\mathbf{P}(\mathbf{e}>s)\mathbf{P}(\mathbf{e}>t). \quad \text{(II.13)}$$

On commence par démontrer la proposition suivante.

**Proposition II.1.24** *On se donne $E$ un espace dénombrable discret et un processus markovien, noté*

$$\big(\,\Omega\,;\,\mathscr{F}\,;\,(\mathscr{F}_t)_{t\in\mathbb{R}_+}\,;\,\mathbf{X}=(X_t)_{t\in\mathbb{R}_+}\,;\,(P_t)_{t\in\mathbb{R}_+}\,;\,\mathbf{P}_\mu\,,\,\mu\in\mathscr{M}_1(E)\,\big)\;.$$

*On suppose que $\mathbf{X}$ est* continu à droite. *Pour tout $i\in E$, sous $\mathbf{P}_i$, le premier temps de saut $J_1(\mathbf{X})$ suit une loi exponentielle (étendue) de paramètre noté $q(i)\in\mathbb{R}_+$. Par ailleurs, sous $\mathbf{P}_i$, $J_1(\mathbf{X})$ est indépendante de la positition $Y_1(\mathbf{X})$ au premier saut.*

**Preuve :** pour simplifier les notations on pose $J_1:=J_1(\mathbf{X})$. On fixe $i\in E$. pour tous $t,s\geq 0$, on a

$$\mathbf{P}_i(J_1>s+t) \;=\; \mathbf{E}_i\left[\mathbf{1}_{\{J_1>t\}}\mathbf{1}_{\{\forall u\in[0,s]\ ,\ X_{u+t}=X_t\}}\right]$$

$$= \quad \mathbf{E}_i \left[ \mathbf{1}_{\{J_1 > t\}} \mathbf{1}_{\{J_1(\theta_t \mathbf{X}) > s\}} \right] .$$

Puisque, d'après la proposition II.1.22, $J_1$ est un $(\mathscr{F}_t)_{t\geq 0}$-temps d'arrêt, l'événement $\{J_1 > t\}$ appartient à $\mathscr{F}_t$. La propriété de Markov au temps $t$ entraîne que $\mathbf{P}_i$-p.s.

$$\mathbf{E}_i \left[ \mathbf{1}_{\{J_1 > t\}} \mathbf{1}_{\{J_1(\theta_t \mathbf{X}) > s\}} \,|\, \mathscr{F}_t \right] = \mathbf{1}_{\{J_1 > t\}} \mathbf{P}_{X_t}(J_1 > s) \ .$$

On observe ensuite que $\mathbf{P}_i$-p.s. $X_t = i$ si $J_1 > t$. Donc

$$\mathbf{E}_i \left[ \mathbf{1}_{\{J_1 > t\}} \mathbf{1}_{\{J_1(\theta_t \mathbf{X}) > s\}} \,|\, \mathscr{F}_t \right] = \mathbf{1}_{\{J_1 > t\}} \mathbf{P}_i(J_1 > s) \ .$$

En intégrant cette égalité sous $\mathbf{P}_i$, on obtient $\mathbf{P}_i(J_1 > t + s) = \mathbf{P}_i(J_1 > t)\mathbf{P}_i(J_1 > s)$ et (II.13) permet d'affirmer que sous $\mathbf{P}_i$, $J_1$ suit une loi exponentielle étendue. On note son paramètre par $q(i)$. Puisque $X$ est continu à droite, on a $\mathbf{P}_i(J_1 > 0) = 1$, ce qui entraîne que $q(i) < \infty$ et cela complète la preuve du premier point de la proposition. .

Pour simplifier les notations, on pose $Y_1 = Y_1(\mathbf{X})$. Montrons l'indépendance de $J_1$ et de $Y_1$ : on fixe $t \in \mathbb{R}_+$ et $g : E \to \mathbb{R}$, une application bornée. La propriété de Markov au temps $t$ implique que

$$\mathbf{E}_i[\mathbf{1}_{\{J_1 > t\}} g(Y_1)] = \mathbf{E}_i \left[ \mathbf{1}_{\{J_1 > t\}} g\left(Y_1(\theta_t \mathbf{X})\right) \right] = \mathbf{E}_i \left[ \mathbf{1}_{\{J_1 > t\}} \mathbf{E}_{X_t}[g(Y_1)] \right] = \mathbf{P}_i(J_1 > t) \mathbf{E}_i[g(Y_1)].$$

cela qui implique facilement le résultat voulu (voir la proposition B.2.8 $(i) \Leftrightarrow (iii)$, page 308). ■

**Définition II.1.25** (*Générateur infinitésimal associé à un Markov càd*) Soit

$$\big(\Omega; \mathscr{F}, ; (\mathscr{F}_t)_{t\in\mathbb{R}_+}; \mathbf{X} = (X_t)_{t\in\mathbb{R}_+}; (P_t)_{t\in\mathbb{R}_+}; \mathbf{P}_\mu, \ \mu \in \mathscr{M}_1(E)\big),$$

un processus markovien continu à droite partout sur $\Omega$. La proposition précédente II.1.24 montre que sous $\mathbf{P}_i$, $J_1$ est une variable exponentielle de paramètre noté $q(i) \in \mathbb{R}_+$. Pour tous $i, j \in E$, on pose $\pi_{i,j} = \mathbf{P}_i(Y_1 = j)$. On remarque d'abord que $Q = (\pi_{i,j})_{i,j\in E}$ est une matrice de transition sur $E$. On observe plus précisément que :

- si $q(i) > 0$, alors $\pi_{i,i} = 0$ ;
- si $q(i) = 0$, alors $\pi_{i,i} = 1$ et donc $\pi_{i,j} = 0$ si $i \neq j$.

Dans les deux cas $q(i)\pi_{i,i} = 0$. On définit le générateur infinitésimal du processus markovien comme la matrice $G = (q_{i,j})_{i,j\in E}$ définie par

$$\forall\, i, j \in E\,, \quad q_{i,j} = \begin{cases} q(i)\pi_{i,j} \in \mathbb{R}_+ & \text{si } i \neq j, \\ -q(i) \in \mathbb{R}_- & \text{si } i = j. \end{cases}$$

On remarque que $\sum_{j\in\setminus\{i\}} q_{i,j} = -q_{i,i} \geq 0$ pour tout $i \in E$. □

Le théorème suivant donne explicitement la loi des temps de saut et du squelette d'un processus markovien (continu à droite) jusqu'au temps d'explosion.

**Théorème II.1.26** *Soit $E$ un espace dénombrable discret. Soit*

$$\big(\,\Omega\,;\, \mathscr{F}\,;\, (\mathscr{F}_t)_{t\in\mathbb{R}_+}\,;\, \mathbf{X} = (X_t)_{t\in\mathbb{R}_+}\,;\, (P_t)_{t\in\mathbb{R}_+}\,;\, \mathbf{P}_\mu\,,\, \mu \in \mathscr{M}_1(E)\,\big)$$

*un processus markovien que l'on suppose* continu à droite*. Pour simplifier les notations, on pose $J_n := J_n(\mathbf{X})$, $Y_n := Y_n(\mathbf{X})$, $D_n := D_n(\mathbf{X})$, $n \in \mathbb{N}$.*

$(i)$ *Le processus à temps discret*

$$\big(\,\Omega\,;\,\mathscr{F};(\mathscr{F}_{J_n})_{n\in\mathbb{N}}\,;\,\mathbf{Y}=(Y_n)_{n\in\mathbb{N}}\,;\,Q=(\pi_{i,j})_{i,j\in E}\,;\,\mathbf{P}_\mu,\mu\in\mathscr{M}_1(E)\,\big)$$

*est une chaîne de Markov.*

$(ii)$ *Pour toute $\mu\in\mathscr{M}_1(E)$, sous $\mathbf{P}_\mu$, conditionnellement à $(Y_n)_{n\geq 0}$, les temps d'attentes $(D_n)_{n\geq 1}$ sont mutuellement indépendants et la loi conditionnelle de $D_n$ est la loi exponentielle de paramètre $q(Y_{n-1})$.*

**Preuve :** pour simplifier les notations, on suppose que $(\Omega,\mathscr{F})$ est suffisamment riche pour qu'y puisse être définie une suite de v.a. $(\mathbf{e}_n)_{n\in\mathbb{N}}$, à valeurs dans $\mathbb{R}_+$ et telle que pour toute loi d'entrée $\mu\in\mathscr{M}_1(E)$, sous $\mathbf{P}_\mu$, les $(\mathbf{e}_n)_{n\in\mathbb{N}}$ sont des exponentielles i.i.d. de paramètre $1$ indépendantes de $\mathbf{X}$. On fixe $i\in E$ et on suppose d'abord que $q(i)>0$, si bien que $\mathbf{P}_i$-p.s. $J_1<\infty$. Soit $F:E^{\mathbb{R}_+}\to\mathbb{R}$ une fonction $\mathscr{P}(E)^{\otimes\mathbb{R}_+}$-mesurable bornée. La propriété de Markov forte au temps $J_1$ implique que $\mathbf{P}_i$-a.s.

$$\mathbf{E}_i\left[F(\theta_{J_1}\mathbf{X})\,|\,\mathscr{F}_{J_1}\right]=\mathbf{E}_{Y_1}[F(X)]\ .$$

Par conséquent, pour toute fonctions $f:[0,\infty]\to\mathbb{R}$ et $g:E\to\mathbb{R}$ mesurables et bornées, on obtient

$$\mathbf{E}_i[f(J_1)g(Y_1)F(\theta_{J_1}\mathbf{X})]=\mathbf{E}_i[f(J_1)g(Y_1)\,\mathbf{E}_{Y_1}[F(\mathbf{X})]]\ ,$$

La proposition II.1.24 permet d'affirmer ensuite que

$$\mathbf{E}_i[f(J_1)g(Y_1)F(\theta_{J_1}\mathbf{X})]=\mathbf{E}_i[f(J_1)]\,\mathbf{E}_i\big[\,g(Y_1)\,\mathbf{E}_{Y_1}[F(\mathbf{X})]\,\big]\ . \tag{II.14}$$

Si $q(i)=0$, on voit que $\mathbf{E}_i[f(J_1)g(Y_1)]=f(\infty)g(i)$. On rappelle que $N:=N(\mathbf{X})=\sup\{n\in\mathbb{N}:J_n<\infty\}$ est le nombre de sauts avant absorption. Dans tous les cas, on a

$$\mathbf{E}_i[f(J_1)g(Y_1)F(\theta_{J_{1\wedge N}}\mathbf{X})]=\mathbf{E}_i[f(\mathbf{e}_1/q(i))]\,\mathbf{E}_i\big[\,g(Y_1)\,\mathbf{E}_{Y_1}[F(\mathbf{X})]\big]\ , \tag{II.15}$$

car $\mathbf{e}_1/q(i)$ suit une loi exponentielle de paramètre $q(i)$ (on utilise la convention $\mathbf{e}_1/q(i)=\infty$ si $q(i)=0$). En appliquant récursivement cette identité, on voit que pour tous $i_1,\dots,i_n\in E$ et pour toutes fonctions $f_1$, ... , $f_n$ mesurables bornées de $[0,\infty]$ vers $\mathbb{R}$,

$$\mathbf{E}_i\left[f_1(D_1)\dots f_n(D_n)\mathbf{1}_{\{Y_1=i_1;\dots;Y_n=i_n\}}\right]=$$
$$\mathbf{E}_i\left[f_1(\mathbf{e}_1/q(i))\dots f_n(\mathbf{e}_n/q(i_{n-1}))\right]\pi_{i,i_1}\dots\pi_{i_{n-1},i_n}\ .$$

La propriété de Markov au temps $0$, combinée avec l'égalité précèdente, entraîne ensuite que pour toute loi d'entrée $\mu\in\mathscr{M}_1(X)$, on ait

$$\mathbf{E}_\mu\left[f_1(D_1)\dots f_n(D_n)\mathbf{1}_{\{Y_0=i_0;\dots;Y_n=i_n\}}\right]=$$
$$\mathbf{E}_\mu\left[f_1(\mathbf{e}_1/q(i_0))\dots f_n(\mathbf{e}_n/q(i_{n-1}))\right]\mu(i_0)\pi_{i_0,i_1}\dots\pi_{i_{n-1},i_n}\ ,$$

ce qui complète la preuve du théorème. ■

### II.1.g Les équations de Kolmogorov-Chapman.

Cette section est consacrée au théorème suivant qui étend aux espace d'état infinis le théorème II.1.10, page 150.

**Théorème II.1.27** **(Equations de Kolmogorov-Chapman)** *Soit $E$ un espace dénombrable discret. Soit*

$$\big(\Omega\,;\,\mathscr{F}\,;\,(\mathscr{F}_t)_{t\in\mathbb{R}_+}\,;\,\mathbf{X}=(X_t)_{t\in\mathbb{R}_+}\,;\,(P_t)_{t\in\mathbb{R}_+}\,;\,\mathbf{P}_\mu\,,\,\mu\in\mathscr{M}_1(E)\,\big)$$

*un processus markovien que l'on suppose* continu à droite. *On note $G=(q_{i,j})_{i,j\in E}$, son générateur infinitésimal (voir la définition II.1.25)). On utilise la notation $P_t=(p_t(i,j))_{i,j\in E}$, $t\in\mathbb{R}_+$.*

$(i)$ *Les applications $t\in\mathbb{R}_+\mapsto p_t(i,j)$, $i,j\in E$, sont continument dérivables et sont solutions du système d'équations différentielles linéaires suivant :*

$$\forall i,j\in E\ ,\ \forall t\in\mathbb{R}_+\ ,\quad p'_t(i,j)=\sum_{k\in E}q_{i,k}p_t(k,j)\ .$$

*Si on note $\frac{d}{dt}P_t=(p'_t(i,j))_{i,j\in E}$, $t\in\mathbb{R}_+$, alors on réécrit symboliquement ce système par*

$$\forall t\in\mathbb{R}_+\ ,\quad \tfrac{d}{dt}P_t=GP_t\quad\text{et}\quad P_0=\mathtt{Id}_E\ ,$$

*qui est appelée* équation de Kolmogorov-Chapman directe.

$(ii)$ *On rappelle que $\zeta:=\zeta(\mathbf{X})$ est le premier temps d'explosion de $\mathbf{X}$. On pose*

$$\forall\,i,j\in E\,,\ \forall\,t\in\mathbb{R}_+\,,\quad p_t^{\min}(i,j)=\mathbf{P}_i(X_t=i\,;\,\zeta>t)$$

*et on note $P_t^{\min}=(p_t^{\min}(i,j))_{i,j\in E}$. Alors $(P_t^{\min})_{t\in\mathbb{R}_+}$ un semi-groupe sous-markovien qui est solution de l'équation de Kolmogorov-Chapman directe :*

$$\forall t\in\mathbb{R}_+,\quad \tfrac{d}{dt}P_t^{\min}=GP_t^{\min}\quad\textit{et}\quad P_0^{\min}=\mathtt{Id}_E\ .$$

C'en est même la solution minimale positive *: si $\widetilde{P}_t=(\widetilde{p}_t(i,j))_{i,j\in E}$, $t\in\mathbb{R}_+$, est une famille de matrices à coefficients positifs continument dérivables qui satisfont $\frac{d}{dt}\widetilde{P}_t=G\widetilde{P}_t$, $t\in\mathbb{R}_+$, avec $\widetilde{P}_0=\mathrm{Id}_E$, alors*

$$\forall\,i,j\in E\,,\ \forall\,t\in\mathbb{R}_+\,,\quad p_t^{\min}(i,j)\le\widetilde{p}_t(i,j)\ .$$

$(iii)$ $(P_t^{\min})_{t\in\mathbb{R}_+}$ *est aussi solution de l'*équation de Kolmogorov-Chapman rétrograde *:*

$$\forall i,j\in E\ ,\ \forall t\in\mathbb{R}_+\ ,\quad \tfrac{d}{dt}p_t^{\min}(i,j)=\sum_{k\in E}p_t^{\min}(i,k)q_{k,j}$$

*que l'on réécrit symboliquement*

$$\forall t\in\mathbb{R}_+,\quad \tfrac{d}{dt}P_t^{\min}=P_t^{\min}G\quad\textit{et}\quad P_0^{\min}=\mathtt{Id}_E\ .$$

C'en est même la solution minimale positive*, c'est-à-dire que si $\widetilde{P}_t=(\widetilde{p}_t(i,j))_{i,j\in E}$, $t\in\mathbb{R}_+$, est une famille de matrices à coefficients positifs continument dérivables qui satisfont $\frac{d}{dt}\widetilde{P}_t=\widetilde{P}_tG$, $t\in\mathbb{R}_+$, avec $\widetilde{P}_0=\mathrm{Id}_E$, alors $p_t^{\min}(i,j)\le\widetilde{p}_t(i,j)$ pour tous $i,j\in E$ et tout $t\in\mathbb{R}_+$.*

**Preuve :** puisque $\mathbf{X}$, le théorème de convergence dominée, implique que $t\mapsto p_t(i,j)=\mathbf{P}_i(X_t=j)$ est continue à droite, pour tous $i,j\in E$. On fixe ensuite $i,j\in E$ et $t>0$. On utilise le symbole de Kronecker $\delta_{i,j}$ qui vaut 1 si $i=j$ et 0 sinon. La propriété de Markov forte au temps $J_1$ ou le théorème II.1.26 entraînent

$$p_t(i,j)\ =\ \mathbf{E}_i[\mathbf{1}_{\{J_1>t\}}]\delta_{i,j}+\mathbf{E}_i[\mathbf{1}_{\{J_1\le t\}}\mathbf{1}_{\{(\theta_{J_1}\mathbf{X})_{t-J_1}=j\}}]$$

$$\begin{aligned} &= \delta_{i,j} e^{-q(i)t} + \int_0^t \mathbf{E}_i[\mathbf{P}_{Y_1}(X_{t-s}=j)] q(i) e^{-q(i)s} \, ds \\ &= \delta_{i,j} e^{-q(i)t} + \int_0^t \mathbf{E}_i[p_{t-s}(Y_1,j)] q(i) e^{-q(i)s} \, ds \\ &= \delta_{i,j} e^{-q(i)t} + e^{-q(i)t} \int_0^t \mathbf{E}_i[p_s(Y_1,j)] q(i) e^{q(i)s} \, ds \end{aligned}$$

où on a effectué le changement de variable $s \mapsto t-s$ pour obtenir la dernière égalité. En multipliant cette égalité par $e^{q(i)t}$ on obtient :

$$e^{q(i)t} p_t(i,j) = \delta_{i,j} + \int_0^t q(i) e^{q(i)s} \mathbf{E}_i[p_s(Y_1,j)] \, ds \ . \tag{II.16}$$

Par convergence dominée $s \mapsto \mathbf{E}_i[p_s(Y_1,j)]$ est continue à droite. Par conséquent, l'égalité précédente implique que $t \mapsto e^{q(i)t} p_t(i,j)(t)$ est dérivable, de dérivée continue à droite. Il est facile alors de voir que c'est également le cas de $t \mapsto p_t(i,j)$. En continuant ainsi on montre que $t \mapsto p_t(i,j)$ est $C^\infty$. Par ailleurs si on dérive (II.16) on obtient après simplificattion par $\exp(q(i)t)$,

$$p'_t(i,j) = -q(i)p_t(i,j) + q(i)\mathbf{E}_i[p_t(Y_1,j)] = \sum_{k\in E} q(i)\pi_{i,k} p_t(k,j) = \sum_{k \in E} q_{i,k} p_t(k,j) \ ,$$

ce qui prouve $(i)$.

Montrons $(ii)$ : pour tous $i,j \in E$ et tous $t,s \in \mathbb{R}_+$, on observe que

$$p_t^{\min}(i,j) = \mathbf{P}_i(X_{t+s}=j; \zeta > t+s) = \sum_{k\in E} \mathbf{P}_i(X_t = k; X_{s+t}=j; \zeta > t+s) \ . \tag{II.17}$$

Par ailleurs, on rappelle que $\{\zeta > t\} \in \mathscr{F}_t$ car $\zeta$ est un temps d'arrêt. En appliquant la propriété de Markov au temps $t$, on a

$$\begin{aligned} \mathbf{P}_i(X_t = k; X_{s+t}=j; \zeta > t+s) &= \mathbf{P}_i(X_t = k; \zeta > t; (\theta_t \mathbf{X})_s = j; \zeta(\theta_t\mathbf{X}) > s) \\ &= \mathbf{P}_i(X_t = k; \zeta(X) > t) \mathbf{P}_k(X_s = j; \zeta(X) > s) \\ &= p_t^{\min}(i,k) p_s^{\min}(k,j) \end{aligned}$$

ce qui combiné avec (II.17), montre que $(P_t^{\min})_{\mathbb{R}_+}$ est un semi-groupe sous-markovien. En appliquant la propriété de Markov au temps $J_1$ sous $\mathbf{P}_i$ et en raisonnant comme précédemment on obtient pour tous $t \in \mathbb{R}_+$, tout $n \in \mathbb{N}$ et tous $i,j \in E$ :

$$\mathbf{P}_i(X_t = j; t < J_{n+1}) = e^{-q(i)t}\delta_{i,j} + \sum_{k \in E\backslash\{i\}} \int_0^t ds \, q_{i,k} e^{-q(i)s} \mathbf{P}_k(X_{t-s} = j; t-s < J_n) \ . \tag{II.18}$$

Comme la suite de v.a. $(J_n)_{n\in\mathbb{N}}$ croît vers $\zeta$, on a

$$\lim_n \uparrow \mathbf{P}_k(X_{t-s} = j; t-s < J_n) = \mathbf{P}_k(X_{t-s} = j; t-s < \zeta) = p_{t-s}^{\min}(k,j) \ .$$

De même $\mathbf{P}_i(X_t = j; t < J_{n+1})$ tend en croissant vers $\mathbf{P}_i(X_t = j; t < \zeta)$. Le théorème de la convergence monotone implique alors

$$p_t^{\min}(i,j) \ = \ \mathbf{P}_i(X_t = j; t < \zeta) = e^{-q(i)t}\delta_{i,j} + \sum_{k\in E\backslash\{i\}} \int_0^t ds \, q_{i,k} e^{-q(i)s} \mathbf{P}_k(X_{t-s} = j; t-s < \zeta)$$

$$= \quad e^{-q(i)t}\delta_{i,j} + \sum_{k\in E\backslash\{i\}} \int_0^t ds\, q_{i,k} e^{-q(i)s} p_{t-s}^{\min}(k,j)$$

ce qui implique facilement que $(P_t^{\min})_{\mathbb{R}_+}$ satisfait l'équation de Kolmogorov-Chapman rétrograde, par des arguments analogues à ceux utilisés dans la preuve du point $(i)$.

Montrons la minimalité. Pour cela, on se donne une solution $\widetilde{P}_t = (\widetilde{p}_t(i,j))_{i,j\in E}$, $t \in \mathbb{R}_+$, à l'équation de Kolmogorov-Chapman rétrograge. Un calcul simple d'intégration par parties implique tout d'abord que

$$\widetilde{p}_t(i,j) = e^{-q(i)t}\delta_{i,j} + \sum_{k\in E\backslash\{i\}} \int_0^t ds\, q_{i,k} e^{-q(i)s} \widetilde{p}_{t-s}(k,j)\ . \tag{II.19}$$

On remarque ensuite que $\mathbf{P}_i(X_t = j; t < J_0) = 0 \leq \widetilde{p}_t(i,j)$. Supposons alors que

$$\forall t \in \mathbb{R}_+\ ,\ \forall i,j \in E\ , \quad \mathbf{P}_i(X_t = j; t < J_n) \leq \widetilde{p}_t(i,j)\ . \tag{II.20}$$

Par (II.18) et (II.19), l'inégalité précédente (II.20) au rang $n$ implique l'inégalité au rang $n+1$. Comme elle est vraie au rang 0, on a démontré par récurrence la validité de (II.20) pour tout $n \in \mathbb{N}$.

En passant à la limite en $n$ dans (II.20), on obtient que $p_t^{\min}(i,j) = \mathbf{P}_i(X_t = j; t < \zeta) \leq \widetilde{p}_t(i,j)$, pour tous $i; j \in E$ et tout $t \in \mathbb{R}_+$, ce qui montre la minimalité. Ceci termine la preuve du point $(ii)$.

La preuve de $(iii)$ est plus délicate. On rappelle que $J_{n+1} = J_n + D_{n+1} = J_{n-1} + D_n + D_{n+1}$. On pose $\alpha = \mathbf{E}_i\big[\mathbf{1}_{\{J_n\leq t<J_{n+1};Y_{n-1}=k;Y_n=j\}}|Y_{n-1}, J_{n-1}\big]$. En utilisant le théorème de décomposition on en déduit que

$$\begin{aligned}
\alpha &= \pi_{k,j}q(k)\int_0^\infty ds\, e^{-q(k)s}\int_0^\infty du\, q(j)e^{-q(j)u}\mathbf{1}_{\{J_{n-1}+s\leq t<J_{n-1}+s+u;Y_{n-1}=k\}} \\
&= \pi_{k,j}q(k)\int_0^{t-J_{n-1}} ds\, e^{-q(k)s}e^{-q(j)(t-J_{n-1}-s)}\mathbf{1}_{\{J_{n-1}+s\leq t;Y_{n-1}=k\}} \\
&= \pi_{k,j}q(k)\int_0^{t-J_{n-1}} ds\, e^{-q(k)(t-J_{n-1}-s)}e^{-q(j)s}\mathbf{1}_{\{J_{n-1}\leq t;Y_{n-1}=k\}} \\
&= \pi_{k,j}q(k)\int_0^{t-J_{n-1}} ds\, e^{-q(j)s}\mathbf{E}_i\big[\mathbf{1}_{\{J_{n-1}\leq t;D_n>t-J_{n-1}-s;Y_{n-1}=k\}}|J_{n-1}, Y_{n-1}\big] \\
&= \pi_{k,j}q(k)\int_0^{t} ds\, e^{-q(j)s}\mathbf{E}_i\big[\mathbf{1}_{\{J_{n-1}\leq t-s<J_n;Y_{n-1}=k\}}|J_{n-1}, Y_{n-1}\big].
\end{aligned}$$

En prenant l'espérance de cette expression et en intervertissant espérance/intégrale, on obtient

$$\begin{aligned}
&\mathbf{P}_i\big(J_n \leq t < J_{n+1}; Y_{n-1} = k; Y_n = j\big) \\
&\quad = \mathbf{1}_{\{k\neq j\}}q_{k,j}\int_0^t ds e^{-q(j)s}\mathbf{P}_i\big(J_{n-1} \leq t-s < J_n; Y_{n-1} = k\big)\ ,
\end{aligned}$$

car $\pi_{k,j}q(k) = 0$ si $k = j$ et $\pi_{k,j}q(k) = q_{k,j}$ si $k \neq j$. On observe alors que

$$\begin{aligned}
\mathbf{P}_i(X_t = j; t < J_{n+1}) &= \delta_{i,j}\mathbf{P}_i(J_1 > t) + \sum_{m=1}^n \sum_{k\in E\backslash\{j\}} \mathbf{P}_i(J_m \leq t < J_{m+1}; Y_{m-1} = k, Y_m = j) \\
&= \delta_{i,j}e^{-q(j)t} + \sum_{k\in E\backslash\{j\}} q_{k,j}\int_0^t ds\, e^{-q(j)s}\sum_{m=1}^n \mathbf{P}_i\left(Y_{m-1} = k; J_{m-1} \leq t-s \leq J_m\right)
\end{aligned}$$

$$\delta_{i,j}e^{-q(j)t} \quad + \sum_{k \in E\backslash\{j\}} q_{k,j} \int_0^t ds\, e^{-q(j)s} \mathbf{P}_i(X_{t-s} = k; t - s < J_n) \; . \tag{II.21}$$

En faisant tendre $n$ vers l'infini, le théorème de convergence monotone implique facilement que

$$p_t^{\min}(i,j) = \delta_{i,j}e^{-q(j)t} + \sum_{k \in E\backslash\{j\}} q_{k,j} \int_0^t ds\, e^{-q(j)s} p_{t-s}^{\min}(i,k) \; .$$

En effectuant le changement de variables $s \mapsto t - s$, puis en multipliant par $e^{q(i)t}$ cela donne

$$e^{q(j)t} p_t^{\min}(i,j) = \delta_{i,j} + \sum_{k \in E\backslash\{j\}} q_{k,j} \int_0^t ds\, e^{q(j)s} p_s^{\min}(i,k) \; .$$

Par Fubini positif on a

$$e^{q(j)t} p_t^{\min}(i,j) = \delta_{i,j} + \int_0^t ds \sum_{k \in E\backslash\{j\}} q_{k,j} e^{q(j)s} p_s^{\min}(i,k) \; . \tag{II.22}$$

Par ailleurs, cette équation (ou bien l'équation de Kolmogorov-Chapman directe) entraînent que $s \mapsto e^{q(j)s} p_t^{\min}(i,k)$ est croissante. Donc, par Dini, $s \mapsto \sum_{k \in E\backslash\{j\}} q_{k,j} e^{q(j)s} p_s^{\min}(i,k)$ est continue croissante. On peut dériver (II.22) et on obtient alors

$$e^{q(j)t} \left( \tfrac{d}{dt} p_t^{\min}(i,j) + q(j) p_t^{\min}(i,j) \right) = e^{q(j)t} \sum_{k \in E\backslash\{j\}} p_s^{\min}(i,k) q_{k,j} \; ,$$

ce qui entraîne que

$$\forall i,j \in E \; , \; \forall t \in \mathbb{R}_+ \; , \quad \tfrac{d}{dt} p_t^{\min}(i,j) = \sum_{k \in E} p_t^{\min}(i,k) q_{k,j} \; .$$

Le semi-groupe $(P_t^{\min})_{t \in \mathbb{R}_+}$ satisfait donc l'équation de Kolmogorov-Chapman rétrograde.

Montrons que c'en est la solution minimale positive : on se donne une solution positive $\widetilde{P}_t = (\widetilde{p}_t(i,j), i,j \in E)$, $t \in \mathbb{R}_+$, de Kolmogorov-Chapman retrodgrade. Une intégration par parties implique que

$$\widetilde{p}_t(i,j) = e^{-q(i)t} \delta_{i,j} + \sum_{k \in E\backslash\{j\}} \int_0^t ds\, e^{-q(i)s} \widetilde{p}_{t-s}(i,k) q_{k,j} \; . \tag{II.23}$$

En raisonnant par récurrence à partir de (II.21) et de (II.23), comme pour le cas de l'équation directe, on montre que

$$\forall n \in \mathbb{N} \; , \; \forall t \in \mathbb{R}_+ \; , \; \forall i,j \in E \; , \quad \mathbf{P}_i(X_t = j; t < J_n) \leq \widetilde{p}_t(i,j),$$

ce qui implique le résultat désiré en faisant tendre $n$ vers l'infini. ■

**Remarque II.1.28** Soit $E$ un espace discret dénombrable. Soit $(P_t)_{t \in \mathbb{R}_+}$ un semi-groupe markovien sur $E$. Pour tout $t \in \mathbb{R}_+$, on écrit $P_t = (p_t(i,j))_{i,j \in E}$. On se donne un processus markovien continu à droite :

$$\big( \Omega \, ; \, \mathscr{F} \, ; \, (\mathscr{F}_t)_{t \in \mathbb{R}_+} \, ; \, \mathbf{X} = (X_t)_{t \in \mathbb{R}_+} \, ; \, (P_t)_{t \in \mathbb{R}_+} \, ; \, \mathbf{P}_\mu \, , \, \mu \in \mathscr{M}_1(E) \big)$$

et on note $G = (q_{i,j})_{i,j \in E}$, son générateur infinitésimal. Soit $(\widetilde{P}_t)_{t \in \mathbb{R}_+}$, une solution de l'équation de Kolmogorov-Chapman directe : $\frac{d}{dt}\widetilde{P}_t = GP_t$ pour tout $t \in \mathbb{R}_+$ et $\widetilde{P}_0 = \mathtt{Id}_E$. Si le processus $\mathbf{X}$ explose, c'est-à-dire s'il existe $i \in E$ tel que $\mathbf{P}_i(\zeta < \infty) > 0$, alors cette équation n'admet pas qu'une seule soluttion

En effet, cela implique facilement qu'il existe $j \in E$ et $t > 0$ tels que $\mathbf{P}_i(X_t = j; \zeta \leq t) > 0$ et on voit que

$$\begin{aligned} p_t(i,j) = P_i(X_t = j) &= \mathbf{P}_i(X_t = j; \zeta > t) + \mathbf{P}_i(X_t = j; \zeta \leq t) \\ &> \mathbf{P}_i(X_t = j; \zeta(X) > t) = p_t^{\min}(i,j) \ . \end{aligned}$$

Donc $P_t \neq P_t^{\min}$ or ces deux semi-groupes sous-markoviens sont deux solutions positives de l'équation de Kolmogorov-Chapman directe. Par ailleurs, on remarquera que l'on ne mentionne pas que $(P_t)_{t \in \mathbb{R}_+}$ est solution de l'équation de Kolmogorov-Chapman rétrograde. □

## II.2 Processus minimal associé à un générateur infinitésimal.

### II.2.a Définition.

Comme nous l'avons mentionné plus haut, notre étude des processus markoviens à valeurs dans un espace d'état discret dénombrable ne porte que sur ce qui se passe avant une explosion éventuelle. Par les résultats précédents, cette partie de la trajectoire est entièrement caractérisée par générateur infinitésimal et on souhaite la voir comme un processus de Markov auquel on peut appliquer les résultats déjà obtenus (notamment la propriété de Markov). Pour cela, une idée consiste à choisir un point cimetière et à y absorber définitivement le processus après le temps d'explosion (s'il survient). On appelle de tels processus des *processus minimaux*. Ils forment une classe de processus markoviens à espace d'état dénombrable discret suffisamment riche pour la plupart des applications (files d'attente, modèles issus de la biologie, mécanique statistique, etc). Commençons par se donner l'objet analytique caractérisant les lois de ces processus minimaux : leur générateur infinitésimal.

**Définition II.2.1** (*Générateur infinitésimal* ) On se donne un espace $E$ discret dénombrable. Une matrice $G = (q_{i,j})_{i,j \in E}$ est un générateur infinitésimal si et seulement si

$$\forall i,j \in E\ ,\ q_{i,j} \in \mathbb{R}_+ \text{ si } i \neq j \text{ et } q_{i,i} \in \mathbb{R}_- \quad \text{et} \quad \forall i \in E\ , \quad -q_{i,i} = \sum_{j \in E \backslash \{i\}} q_{i,j} \ .$$

À un générateur infinitésimal $G = (q_{i,j})_{i,j \in E}$, on associe deux objets.

$(a)$ Une famille $(q(i))_{i \in E}$ de réels positifs donnés par $q(i) = -q_{i,i}$, $i \in E$, et appelée *famille des paramètres des temps d'attente*.

$(b)$ Une matrice $Q = (\pi_{i,j})_{i,j \in E}$ donnée par

$$\forall\, i,j \in E\ , \quad \pi_{i,j} = \begin{cases} -q_{i,j}/q_{i,i} & \text{si } i \neq j \text{ et } q_{i,i} < 0, \\ 0 & \text{si } i = j \text{ et } q_{i,i} < 0, \\ \delta_{i,j} & \text{si } q_{i,i} = 0, \end{cases}$$

où $\delta_{i,j}$ désigne le symbole de Kronecker qui vaut 1 si $i = j$ et 0 sinon. On remarque facilement que $Q$ est une matrice de transition ; elle est appelée *matrice de transition du squelette*. On note que si $q_{i,i} = 0$ alors $i$ est absorbant pour $Q$ et si $q_{i,i} < 0$ alors $\pi_{i,i} = 0$. □

Nous donnons ensuite la définition d'un processus minimal en y incluant quelques commentaires expliquant le terme de « minimal » ainsi que des notations utilisées dans la suite du chapitre.

**Définition II.2.2** (*Processus minimal*) Soit $E$ un espace dénombrable discret. Soit $G = (q_{i,j})_{i,j\in E}$ un générateur infinitésimal sur $E$. On se donne un point cimetière $\partial \notin E$. On pose $E^* = E \cup \{\partial\}$, que l'on voit comme un espace dénombrable *discret*. Un processus de Markov à valeurs dans $E^*$, noté

$$\big( \Omega\,;\, \mathscr{F}\,;\, (\mathscr{F}_t)_{t\in\mathbb{R}_+}\,;\, \mathbf{X} = (X_t)_{t\in\mathbb{R}_+}\,;\, P^*_t = (p^*_t(i,j))_{i,j\in E^*},\, t \in \mathbb{R}_+\,;\, \mathbf{P}_\mu\,,\, \mu \in \mathscr{M}_1(E^*)\,\big)$$

continu à droite partout sur $\Omega$ est un *processus minimal* associé au générateur $G$ si et seulement s'il satisfait les conditions suivantes.

$(a)$ $\mathbf{X}$ est continu à droite dans $E^*$ (qui est on le rappelle supposé discret).

$(b)$ Le point cimetière est absorbant : pour tout $t \in \mathbb{R}_+$, $p^*_t(\partial, i) = 0$ pour tout $i \in E$ et $p^*_t(\partial, \partial) = 1$.

$(c)$ Le processus explose si et seulement s'il atteint $\partial$ (état où il est définitivement absorbé), c'est-à-dire que

$$\forall \omega \in \Omega\,, \quad \zeta = \inf\{t \in \mathbb{R}_+ \;:\; X_t = \partial\,\}\,,$$

(avec la convention que $\inf \emptyset = \infty$) et pour tout $t \geq \zeta$, alors $X_t = \partial$. Donc $X_t \in E$, pour tout $t < \zeta$.

$(d)$ On note $G^* = (q^*_{i,j})_{i,j\in E^*}$ le générateur infinitésimal du processus, comme à la définition II.1.25 (page 164). Alors

$$\forall i, j \in E\,, \quad q_{i,j} = q^*_{i,j}\;.$$

Les conditions $(a)$ à $(d)$ combinées avec les résultats des sections précédentes ont les conséquences immédiates suivantes.

$(i)$ Comme $\partial$ est absorbant, on vérifie que $q^*_{\partial,\partial} = q^*_{\partial,i} = q^*_{i,\partial} = 0$, pour tout $i \in E$. Comme le temps d'explosion est exactement le temps d'atteinte de $\partial$. Il ne peut donc atteindre $\partial$ depuis un état $i \in E$ sans faire une infinité de sauts et on ne peut donc sauter d'un état $i$ sur $\partial$ directement.

$(ii)$ Pour tout $t \in \mathbb{R}_+$ et pour tous $i, j \in E$, on pose $p_t(i,j) = p^*_t(i,j)$ et $P_t = (p_t(i,j))_{i,j\in E}$. Puisque le processus, après l'explosion stationne à $\partial$, on voit que

$$\forall t \in \mathbb{R}_+\,,\ \forall\, i, j \in E\,, \quad p_t(i,j) = p^*_t(i,j) = \mathbf{P}_i(X_t = j) = \mathbf{P}_i(X_t = j; \zeta > t)\;.$$

Cela implique que $(P_t)_{t\in\mathbb{R}_+}$ est un semi-groupe sous-markovien qui est *la solution minimale* positive des équations de Kolmogorov-Chapman (directe et rétrograde) associées à $G$ :

$$\tfrac{d}{dt} P_t = G P_t = P_t G\,, \quad t \in \mathbb{R}_+ \quad \text{et} \quad P_0 = \mathtt{Id}_E\;.$$

$(iii)$ Ce qui précède montre que le semi-groupe sous-markovien $(P_t)_{t\in\mathbb{R}_+}$ est entièrement déterminé par $G$. Par ailleurs, il détermine entièrement le semi-groupe markovien $(P^*_t)_{t\in\mathbb{R}_+}$ à travers les équations

$$\forall t \in \mathbb{R}_+\,,\ \forall\, i \in E\,, \quad p^*_t(i, \partial) = 1 - \sum_{j\in E} p_t(i,j)$$

et bien sûr $p^*_t(\partial, i) = 0$ et $p^*_t(i,j) = p_t(i,j)$, $j \in E$.

**Notation.** Comme le processus $\big(\Omega; \mathscr{F}; (\mathscr{F}_t)_{t\in\mathbb{R}_+}; \mathbf{X} = (X_t)_{t\in\mathbb{R}_+}; (P^*_t)_{t\in\mathbb{R}_+}; \mathbf{P}_\mu,\, \mu \in \mathscr{M}_1(E^*)\big)$ est entièrement déterminé par $G$ et comme $\partial$ est absorbant, on préfère se donner le processus minimal associé à $G$ dans faire référence au point cimetière en le notant simplement

$$\big(\,\Omega\,;\, \mathscr{F}\,;\, (\mathscr{F}_t)_{t\in\mathbb{R}_+}\,;\, \mathbf{X} = (X_t)_{t\in\mathbb{R}_+}\,;\, G\,;\, \mathbf{P}_\mu\,,\, \mu \in \mathscr{M}_1(E)\,\big)$$

Il faut simplement avoir conscience que le processus markovien qui correspond à cette notation est un processus markovien à valeurs dans l'espace dénombrable discret $E^* = E \cup \{\partial\}$. □

Dans la section suivante, on construit les processus minimaux associés à un générateur (il n'y a pas de restriction). Ensuite, nous étudierons les processus minimaux comme nous l'avons fait pour les chaînes de Markov (récurrence, transience, loi stationnaires, théorèmes ergodique et convergence vers la loi stationnaire) en insistant sur les processus irréductibles récurrents. Enfin, nous donnerons quelques outils de calculs utilisant le générateur infinitésimal et nous introduirons quelques exemples.

### II.2.b Construction du processus minimal.

Dans cette section, on construit un processus minimal associé à un générateur quelconque (comme déjà mentionné, il n'y a pas de restriction). Cette construction est instructive car elle montre comment un générateur infinitésimal code une dynamique markovienne, ce qui sera utile dans les applications où les modèles sont décrits à travers leur générateur.

Soit $E$ un espace dénombrable discret. Soit $G = (q_{i,j})_{i,j\in E}$ un générateur infinitésimal sur $E$, selon la définition II.2.1, page 170. On note $(q(i))_{i\in E}$ la famille des paramètres des temps d'attente et $Q = (\pi_{i,j})_{i,j\in E}$, la matrice de transition du squelette, objets qui sont associés à $G$, comme expliqué dans la définition II.2.1.

On se donne un point cimetière $\partial \notin E$. On pose $E^* = E\cup\{\partial\}$, que l'on voit comme un espace dénombrable *discret*. On étend $Q$ à $E^*$ en considérant que $\partial$ est un état absorbant en posant :

$$\forall i \in E\,, \quad \pi_{\partial,i} = \pi_{i,\partial} = 0 \quad \text{et} \quad \pi_{\partial,\partial} = 1.$$

On continue de noter $Q$ la nouvelle matrice de transition $(\pi_{i,j})_{i,j\in E^*}$. De même on pose

$$q(\partial) = 0\ .$$

La construction que nous donnons ci-dessous présuppose l'existence des objets suivants. (une construction a été donné dans le chapitre sur les chaînes).

$(a)$ Une chaîne de Markov à valeurs dans $E^*$

$$(\,\Omega; \mathscr{F}; \mathbf{Y} = (Y_n)_{n\in\mathbb{N}}; Q; \mathbf{P}_\mu, \mu \in \mathscr{M}_1(E^*)\,)\ .$$

La filtration discrète est la filtration naturelle de $\mathbf{Y}$ : $\mathscr{F}_n^{\mathbf{Y}} = \sigma(Y_0, \ldots, Y_n)$, $n \in \mathbb{N}$.

$(b)$ Une suite $\mathbf{e}_n : \Omega \to ]0, \infty[$ , $n \in \mathbb{N}$, de v.a. $\mathscr{F}$-mesurable. telles que pour toute loi d'entrée $\mu \in \mathscr{M}_1(E^*)$, sous $\mathbf{P}_\mu$, les $(\mathbf{e}_n)_{n\in\mathbb{N}}$ sont des exponentielles i.i.d. de paramètre $1$ *indépendantes de la chaîne* $\mathbf{Y}$.

On utilise la convention $a/0 = \infty$ pour tout $a \in ]0, \infty[$ , si bien que $\mathbf{e}_n(\omega)/0 = \infty$, pour tout $n \in \mathbb{N}$ et pour tout $\omega \in \Omega$. Cela permet de définir la suite de variables $\mathscr{F}$-mesurables $D_n : \Omega \to ]0, \infty]$, $n \geq 1$, par

$$D_n = \frac{\mathbf{e}_n}{q(Y_{n-1})} \ \in ]0, \infty]\ .$$

On pose ensuite $J_0 = 0$ et pour tout $n \geq 1$,

$$J_n = D_1 + \ldots + D_n \ \in ]0, \infty]\ ,$$

avec la convention que $a + \infty = \infty$, pour tout $a \in [0, \infty]$. On pose enfin

$$N = \sup\{n \in \mathbb{N} :\ J_n < \infty\} \in \mathbb{N} \cup \{\infty\} \quad \text{et} \quad \zeta = \sup_{n\in\mathbb{N}} J_n \in [0, \infty]\ .$$

Puisque les $(J_n)$ croissent, l'intervalle $[0, \zeta[$ peut être partitionné de la manière suivante.

$$[0, \zeta[= \bigcup_{n\in\mathbb{N}} [J_n, J_{n+1}[\ . \tag{II.24}$$

On construit ensuite un processus $\mathbf{X} = (X_t)_{t\in\mathbb{R}_+}$ de la manière suivante.

- Si $t \in [0, \zeta[$, alors, par (II.24), il existe $n \in \mathbb{N}$ tel que $t \in [J_n, J_{n+1}[$ . On pose alors $X_t = Y_n$.
- Si $t \in [\zeta, \infty[$ , on pose $X_t = \partial$.

On voit donc que pour tout $t \in \mathbb{R}_+$,

$$\forall\, i \in E\,,\ \{X_t = i\} = \bigcup_{n \geq 0} \{J_n \leq t < J_{n+1}\} \cap \{Y_n = i\}\ \in \mathscr{F} \quad \text{et} \quad \{X_t = \partial\} = \{\zeta \leq t\} \in \mathscr{F}\ .$$

Cela montre que pour tout $t \in \mathbb{R}_+$ , $X_t : \Omega \to E^*$ est $\mathscr{F}$-mesurable. Par le lemme II.1.2, on peut affirmer que $X : \Omega \to E^{*\mathbb{R}_+}$ est une application $(\mathscr{F}, \mathscr{P}(E^*)^{\otimes \mathbb{R}_+})$-mesurable. De plus, la construction implique que $t \mapsto X_t$ est continu à droite dans $E^*$, vu comme un espace d'états dénombrable discret et

$$J_n(\mathbf{X}) = J_n\,,\ D_n(\mathbf{X}) = D_n\,,\ N(\mathbf{X}) = N\,,\ \zeta(\mathbf{X}) = \zeta = \inf\{t \geq 0\ :\ X_t = \partial\}.$$

Pour tout $t \in \mathbb{R}_+$, on pose alors $P^*_t = (p^*_t(i,j))_{i,j \in E^*}$ où $p^*_t(i,j) = \mathbf{P}_i(X_t = j)$, pour tous $i, j \in E^*$. On pose $\mathscr{F}^{\mathbf{X}}_t = \sigma(X_s, s \in [0,t])$, $t \in \mathbb{R}_+$, la filtration naturelle de $\mathbf{X}$.

**Proposition II.2.3** *Avec les notations et les définitions introduites ci-dessus,*

$$\big(\, \Omega\,;\ \mathscr{F}\,;\ (\mathscr{F}^{\mathbf{X}}_t)_{t \in \mathbb{R}_+}\,;\ \mathbf{X}\,;\ (P^*_t)_{t \in \mathbb{R}_+}\,;\ \mathbf{P}_\mu, \mu \in \mathscr{M}_1(E^*)\,\big)$$

*est un processus markovien minimal associé à $G$ introduit à la définition II.2.2.*

**Preuve :** il suffit de montrer que pour toute $\mu \in \mathscr{M}_1(E^*)$, sous $\mathbf{P}_\mu$, $\mathbf{X}$ est un processus de Markov : *en effet*, par construction son semi-groupe est donné par $G$ (étendu à $E^*$) ; de plus il est continu à droite, le temps d'atteinte de $\partial$ est égal au premier temps d'explosion ; c'est un processus minimal associé à $G$ selon la définition II.2.2 page 171.

Fixons donc $\mu \in \mathscr{M}_1(E^*)$ et montrons que, sous $\mathbf{P}_\mu$, $\mathbf{X}$ est un processus de Markov. Pour cela, on prouve d'abord que pour tout $m \in \mathbb{N}$, tous $i, j \in E^*$, tous $t, s \geq 0$, tous $i_0, \ldots, i_{m-1}, i_m = i \in E$ et tous $t_0 = 0 \leq t_1 \leq \ldots \leq t_m = t$, on a

$$\begin{aligned} &\mathbf{P}_\mu(X_0 = i_0; \ldots; X_{t_{m-1}} = i_{m-1}; X_t = i; X_{t+s} = j) \\ &\quad = \ \mathbf{P}_\mu(X_0 = i_0; \ldots; X_{t_{m-1}} = i_{m-1}; X_t = i) p^*_s(i,j)\ . \end{aligned} \tag{II.25}$$

**Preuve de (II.27).** La construction donnée ci-dessus montre que $\mathbf{X}$ s'écrit comme une certaine fonction $\Phi$ des deux suites $(D_n)_{n \geq 0}$ et $(Y_n)_{n \geq 0}$ :

$$\mathbf{X} = \Phi\,(\,(D_n)_{n \geq 0}; (Y_n)_{n \geq 0}\,)\ .$$

On admet la mesurabilité de $\Phi$ : ce point simple est laissé en exercice. Soit $F : (E^*)\mathbb{R}_+ \to \mathbb{R}$ une fonction $\mathscr{P}(E^*)^{\otimes \mathbb{R}_+}$-mesurable bornée. On fixe d'abord $i \in E$ et on remarque que

$$\begin{aligned} &\mathbf{1}_{\{J_n \leq t < J_{n+1}; X_t = i\}} F(\theta_t \mathbf{X}) \\ &\quad = \ \mathbf{1}_{\{J_n \leq t; Y_n = i\}} \mathbf{1}_{\{D_{n+1} > t - J_n\}} F\big(\,\Phi(D_{n+1} - (t - J_n), D_{n+2}, \ldots; Y_n, Y_{n+1} \ldots\,\big) \end{aligned}$$

Pour tout $u \in \mathbb{R}_+$, on pose

$$G(u) = \mathbf{E}_i\big[\mathbf{1}_{\{D_1 > u\}} F\big(\,\Phi(D_1 - u, D_2, \ldots; Y_0, Y_1 \ldots\,)\big]\ .$$

L'indépendante des exponentielles $\mathbf{e}_p$ et la propriété de Markov au temps $n$ pour $\mathbf{Y}$, impliquent que

$$\mathbf{E}_\mu\big[\mathbf{1}_{\{J_n \leq t < J_{n+1}; X_t = i\}} F(\theta_t \mathbf{X}) | D_1, \ldots, D_n, Y_0, \ldots Y_n\big] = \mathbf{1}_{\{J_n \leq t; Y_n = i\}} G(t - J_n)\ .$$

La propriété d'oubli des loi exponentielles implique ensuite que

$$\mathbf{E}_i\big[\, F\big(\,\Phi(D_1-u, D_2,\ldots;Y_0,Y_1\ldots\,)\,\big|D_1>u\big] = \mathbf{E}_i\big[\, F\big(\,\Phi(D_1,D_2,\ldots;Y_0,Y_1\ldots\,)\,\big] = \mathbf{E}_i[F(\mathbf{X})]\ .$$

Par conséquent, $G(u) = \mathbf{P}_i(D_1>u)\mathbf{E}_i[\,F(\mathbf{X})] = e^{-q(i)u}\mathbf{E}_i[\,F(\mathbf{X})]$, ce qui combiné avec les égalités précédentes implique que

$$\mathbf{E}_\mu\big[\mathbf{1}_{\{J_n\le t<J_{n+1};X_t=i\}}F(\theta_t\mathbf{X})|D_1,\ldots,D_n,Y_0,\ldots Y_n\big] = \mathbf{1}_{\{J_n\le t;Y_n=i\}}e^{-q(i)(t-J_n)}\mathbf{E}_i[\,F(\mathbf{X})]\ . \tag{II.26}$$

En prenant $F\equiv 1$, on obtient donc que

$$\mathbf{E}_\mu\big[\mathbf{1}_{\{J_n\le t<J_{n+1};X_t=i\}}|D_1,\ldots,D_n,Y_0,\ldots Y_n\big] = \mathbf{1}_{\{J_n\le t;Y_n=i\}}e^{-q(i)(t-J_n)}$$

et ré-injectant dans (II.26), on obtient

$$\begin{aligned}&\mathbf{E}_\mu\big[\mathbf{1}_{\{J_n\le t<J_{n+1};X_t=i\}}F(\theta_t\mathbf{X})|D_1,\ldots,D_n,Y_0,\ldots Y_n\big]\\ &\quad = \mathbf{E}_\mu\big[\mathbf{1}_{\{J_n\le t<J_{n+1};X_t=i\}}|D_1,\ldots,D_n,Y_0,\ldots Y_n\big]\,\mathbf{E}_i[\,F(\mathbf{X})].\end{aligned}$$

Or sur $\{J_n\le t<J_{n+1}\}$, les variables $X_0,X_{t_1},\ldots,X_{t_{m-1}}$ ne dépendent que de $D_1,\ldots,D_n,Y_0,\ldots Y_n$. Cela montre donc que

$$\begin{aligned}&\mathbf{E}_\mu\big[\mathbf{1}_{\{X_0=i_0;X_{t_1}=i_1;\ldots;X_{t_{m-1}}=i_{m-1};J_n\le t<J_{n+1};X_t=i\}}F(\theta_t\mathbf{X})\,|\,D_1,\ldots,D_n,Y_0,\ldots Y_n\big]\\ &\quad = \mathbf{E}_\mu\big[\mathbf{1}_{\{X_0=i_0;X_{t_1}=i_1;\ldots;X_{t_{m-1}}=i_{m-1};J_n\le t<J_{n+1};X_t=i\}}\,|\,D_1,\ldots,D_n,Y_0,\ldots Y_n\big]\\ &\qquad \times\quad \mathbf{E}_i[\,F(\mathbf{X})]\ .\end{aligned} \tag{II.27}$$

En prenant l'espérance sous $\mathbf{P}_\mu$ de cette égalité et en sommant sur $n$ on obtient donc que

$$\begin{aligned}&\mathbf{E}_\mu\big[\mathbf{1}_{\{X_0=i_0;X_{t_1}=i_1;\ldots;X_{t_{m-1}}=i_{m-1};\zeta>t;X_t=i\}}F(\theta_t\mathbf{X})\,\big] =\\ &\qquad \mathbf{P}_\mu\big(X_0=i_0;X_{t_1}=i_1;\ldots;X_{t_{m-1}}=i_{m-1};\zeta>t;X_t=i\big)\mathbf{E}_i[\,F(\mathbf{X})]\ .\end{aligned} \tag{II.28}$$

Comme $i\in E$, $\{X_t=i\}=\{\zeta>t;X_t=i\}$. En prenant $F(\mathbf{X})=\mathbf{1}_{\{X_s=j\}}$ dans l'égalité précédente, on obtient (II.27). Par ailleurs (II.27) est trivialement vérifiée si $i=\partial$. On a donc montré (II.27). ■

**Fin de la preuve de la proposition II.2.3.** En appliquant récursivement (II.27), on montre que pour toute $\mu\in\mathscr{M}_1(E_\partial)$, tous $t_0=0<t_1<\ldots<t_n$ et tous $i_0,\ldots,i_n\in E^*$, on obtient

$$\mathbf{P}_\mu(X_0=i_0;X_{t_1}=i_1;\ldots;X_{t_n}=i_n) = \mu(i_0)\prod_{1\le m\le n}p^*_{t_m-t_{m-1}}(i_{m-1},i_m)\ .$$

Cela entraîne d'une part que $(P^*_t)_{t\in\mathbb{R}_+}$ est un semi-groupe markovien. D'autre part cela implique que le processus $\mathbf{X}$ est un processus markovien. On a vu que la construction entraînait la minimalité de $\mathbf{X}$, ce qui permet de conclure. ■

Cette construction combinée au théorème II.1.27, page 166, implique immédiatement le résultat suivant d'existence de solutions positives aux équations de Kolmogorov-Chapman directe et rétrograde.

**Corollaire II.2.4** *Pour tout générateur G comme à la définition II.2.1, page 170, il existe une solution $(P_t)_{t\in\mathbb{R}_+}$ aux équations de Kolmogorov-Chapman directe et rétrograde dont les coefficients sont positifs :*

$$\forall t\in\mathbb{R}_+,\quad \tfrac{d}{dt}P_t = GP_t = P_tG \quad \textit{et} \quad P_0=\mathtt{Id}_E\ .$$

*Plus précisément, ces équations ont une même solution minimale à coefficients positifs qui s'avère être un semi-groupe sous-Markovien.*

### II.2.c Classification des états.

Dans toute cette section on se donne un espace d'état $E$, dénombrable discret et $G = (q_{i,j})_{i,j\in E}$ un générateur infinitésimal sur $E$. On se donne également un point cimetière $\partial \notin E$ et on utilise la notation $E^* = E\cup\{\partial\}$, que l'on voit comme un espace dénombrable discret. On se donne un processus minimal

$$\big( \Omega\,;\, \mathscr{F}\,;\, (\mathscr{F}_t)_{t\in\mathbb{R}_+}\,;\, \mathbf{X} = (X_t)_{t\in\mathbb{R}_+}\,;\, G\,;\, \mathbf{P}_\mu, \mu \in \mathscr{M}_1(E)\ \big)$$

continu à droite partout sur $\Omega$. On note $(P_t)_{t\in\mathbb{R}_+}$ le semi-groupe minimal associé à $G$. On utilisera la notation $P_t = (p_t(i,j))_{i,j\in E}$. On rappelle les notations $(q(i))_{i\in E}$ pour les paramètres des temps d'attente et $Q = (\pi_{i,j})_{i,j\in E}$ pour la matrice de transition du squelette. Pour simplifier les notations, on posera

$$D_n = D_n(\mathbf{X})\ ,\ J_n = J_n(\mathbf{X})\ ,\ \zeta := \zeta(\mathbf{X})\ ,\ Y_n := Y_n(\mathbf{X})\ .$$

**Définition II.2.5** (*Irréductibilité.*) Soient $i, j \in E$ deux états distincts.

$(a)$ On dit que $j$ est accessible depuis $i$ ssi $\mathbf{P}_i(\exists t \geq 0\ :\ X_t = j) > 0$, ce que l'on note $i \to j$.

$(b)$ On dit que $i$ et $j$ communiquent ssi $i \to j$ et $j \to i$, ce que l'on note $i \leftrightarrow j$.

$(c)$ On dit que $G$ (ou que le processus $\mathbf{X}$) est irréductible ssi tous ses états communiquent. □

La proposition suivante donne des formulations équivalentes pour qu'un état soit accessible depuis un autre.

**Proposition II.2.6** *Soient $i, j \in E$ deux états distincts. Il y a équivalence entre les assertions suivantes.*

$(i)$ $i \to j$

$(ii)$ *$i$ est accessible depuis $j$ pour la chaîne du squelette $\mathbf{Y}$, c'est-à-dire qu'il existe $n \geq 1$, et $i_1, \ldots, i_n \in E$ tels que $\pi_{i,i_1}\pi_{i_1,i_2}\ldots\pi_{i_n,j} > 0$.*

$(iii)$ *Il existe $n \geq 1$, et $i_1, \ldots, i_n \in E$, tels que $i \neq i_1$, $i_1 \neq i_2$, ..., $i_n \neq j$ et que $q_{i,i_1}q_{i_1,i_2}\ldots q_{i_n,j} > 0$.*

$(iv)$ *Pour tout $t > 0$, $p_t(i,j) > 0$.*

$(v)$ *Il existe $t_0 > 0$, tel que $p_{t_0}(i,j) > 0$.*

**Preuve :** on a clairement $(iv) \Rightarrow (v) \Rightarrow (i) \Rightarrow (ii)$. On suppose $(ii)$. Quitte à raccourcir le chemin $i_1, \ldots, i_n$ à la fin, on peut supposer qu'aucun état $i, \ldots, i_n$ n'est absorbant, c'est -à-dire que $q(i), \ldots, q(i_n) > 0$. Alors $(ii)$ implique $(iii)$ car

$$q_{i,i_1}q_{i_1,i_2}\ldots q_{i_n,j} = q(i)\ldots q(i_n)\pi_{i,i_1}\pi_{i_1,i_2}\ldots\pi_{i_n,j} > 0\ .$$

Ensuite, on suppose $(iii)$. On remarque que si $k \neq \ell$ dans $E$ et si $q_{k,\ell} > 0$, alors on a d'une part $\pi_{k,\ell} > 0$ et $q(k) > 0$ car $q_{k,\ell} = q(k)\pi_{k,\ell}$, et d'autre part

$$\begin{aligned} p_s(k,\ell) &\geq \mathbf{P}_k(J_1 \leq s < J_2; Y_1 = \ell) \\ &= \int_0^s du\, q(k)e^{-q(k)u}\pi_{k,\ell}\,\mathbf{P}_\ell(J_1 > s-u) \\ &= \int_0^s du\, q(k)e^{-q(k)u}\pi_{k,\ell}\, e^{-q(\ell)(s-u)}\ >\ 0\ . \end{aligned}$$

Comme $P_t = (P_{t/n})^n$ on en déduit que $p_t(i,j) \geq p_{t/n}(i,i_1)p_{t/n}(i_1,i_2)\ldots p_{t/n}(i_n,j) > 0$, ce implique $(iv)$ et ce qui achève la preuve de la proposition. ■

**Récurrence et transience.**

**Définition II.2.7** (*Etats récurrents et transients*) On fixe $i \in E$.

$(a)$ On dit que $i$ est *récurrent* si $\mathbf{P}_i$-p.s. l'ensemble (aléatoire) de temps $\{t \geq 0 \;:\; X_t = i\}$ n'est pas borné.

$(b)$ On dit que $i$ est *transient* si $\mathbf{P}_i(\{t \geq 0 \;:\; X_t = i\}$ est borné$) = 1$. □

**Théorème II.2.8** *Les assertions suivantes sont vérifiées.*

$(i)$ *Si $i$ est récurrent pour la chaîne du squelette $\mathbf{Y}$, alors $i$ est récurrent pour le processus markovien $\mathbf{X}$.*

$(ii)$ *Si $i$ est transient pour la chaîne du squelette $\mathbf{Y}$ alors $i$ est transient pour le processus markovien $\mathbf{X}$.*

$(iii)$ *Chaque état de $E$ est soit transient soit récurrent.*

$(iv)$ *Si $i \leftrightarrow j$, alors $i$ et $j$ sont tous les deux transients ou bien $i$ et $j$ sont tous les deux récurrents.*

$(v)$ *Si $G$ est irréductible, alors ou bien les états sont tous récurrents ou bien ils sont tous transients.*

**Preuve :** on suppose que $i$ est récurrent pour $\mathbf{Y}$ ; on introduit les temps de retours successifs $T_i^{(m)} := T_i^{(m)}(\mathbf{Y})$, $m \in \mathbb{N}$, de $\mathbf{Y}$ en $i$ : $T_i^{(0)} = 0$ et $T_i^{(m+1)} = \inf\{n > T_i^{(m)} \;:\; Y_n = i\}$. Comme $i$ est récurrent pour $\mathbf{Y}$, $\mathbf{P}_i$-p.s. les $T_i^{(m)}$ sont finis. On remarque ensuite que

$$\{\, t \in [0, \zeta[ \,:\; X_t = i \,\} = \bigcup_{m \geq 0} [\, J_{T_i^{(m)}} \,,\, J_{T_i^{(m)}+1} \,[$$

Or le théorème II.1.26 implique que sous $\mathbf{P}_i$, les variables $J_{T_i^{(m)}+1} - J_{T_i^{(m)}} = D_{T_i^{(m)}+1}$, $m \geq 1$, sont indépendantes et suivent une loi exponentielle de paramètre $q(i)$. Par conséquent

$$\infty = \sum_{m \geq 0} D_{T_i^{(m)}+1} \leq \zeta \;.$$

Cela montre que $\mathbf{P}_i(\zeta = \infty) = 1$. On en déduit donc que $\mathbf{P}_i$-p.s.

$$X_{J_{T_i^{(m)}}} = Y_{T_i^{(m)}} = i \quad \text{et} \quad \lim_{m \to \infty} J_{T_i^{(m)}} = \infty \;.$$

Par conséquent, $\mathbf{P}_i$-p.s. $\{t \geq 0 \;:\; X_t = i\}$ n'est pas borné et $i$ est récurrent pour $\mathbf{X}$.

Supposons ensuite que $i$ soit transient pour $\mathbf{Y}$. Il est clair que cela implique que $X$ n'est pas aborbé dans $i$ et donc que $q(i) > 0$. Par ailleurs il existe $M \in \mathbb{N}$ (aléatoire) tel que

$$T_i^{(M)} < \infty \quad \text{et} \quad T_i^{(M+1)} = \infty \;.$$

On a donc

$$\{\, t \in [0, \zeta[ \,:\; X_t = i \,\} = \bigcup_{m=0}^{M} [\, J_{T_i^{(m)}} \,,\, J_{T_i^{(m)}+1} \,[$$

Or le théorème II.1.26 implique que sous $\mathbf{P}_i$ la variable $J_{T_i^{(M)}+1} - J_{T_i^{(M)}} = D_{T_i^{(M)}+1}$ suit une loi exponentielle de paramètre $q(i) > 0$. Par conséquent $J_{T_i^{(M)}+1} < \infty$ et $\mathbf{P}_i$-p.s. $\{t \geq 0 \;:\; X_t = i\}$ est borné, ce qui implique que $i$ est transient. Les points $(iii)$, $(iv)$ et $(v)$ sont des conséquences de $(i)$ et $(ii)$ qui permettent de se ramener aux chaînes de Markov pour lesquelles ces résultats sont vérifiés. ■

On introduit *le premier temps de retour en $i$* du processus $\mathbf{X}$ :

$$T_i^+ = \inf\{t > J_1 \;:\; X_t = i\}$$

(avec la convention habituelle que $\inf \emptyset = \infty$). On peut alors reformuler la notion de récurrence et de transience de $i$ à l'aide de $T_i^+$ (et aussi uniquement en termes du semi-groupe) comme suit.

**Proposition II.2.9** *Soit $i \in E$. On a l'alternative suivante.*

**(I)** *Si $q(i) = 0$ ou si $\mathbf{P}_i(T_i^+ < \infty) = 1$, alors $i$ est récurrent et on a*

$$\mathbf{E}_i \left[ \int_0^\infty \mathbf{1}_{\{X_t = i\}} \, dt \right] = \int_0^\infty p_t(i,i) \, dt = \infty \; .$$

**(II)** *Si $q(i) > 0$* **et** *si $\mathbf{P}_i(T_i^+ < \infty) < 1$, alors $i$ est transient et on a*

$$\mathbf{E}_i \left[ \int_0^\infty \mathbf{1}_{\{X_t = i\}} \, dt \right] = \int_0^\infty p_t(i,i) \, dt < \infty \; .$$

**Preuve :** si $q(i) = 0$ alors $p_t(i,i) = 1$ ($i$ est absorbant) et donc $\mathbf{P}_i(\forall t \in \mathbb{R}_+ \, , \; X_t = i) = 1$. On suppose ensuite que $\mathbf{P}_i(T_i^+ < \infty) = 1$, ce qui implique que $q(i) > 0$. On note alors $T_i^{(1)}(\mathbf{Y})$ le premier temps de retour en $i$ pour la chaîne du squelette $\mathbf{Y}$ : $T_i^{(1)}(\mathbf{Y}) = \inf\{n \geq 1 : Y_n = i\}$. On a clairement $\mathbf{P}_i$-p.s. $T_i^{(1)}(\mathbf{Y}) < \infty$ ssi $T_i^+ < \infty$, et donc $\mathbf{P}_i(T_i^{(1)}(\mathbf{Y}) < \infty) = 1$, ce qui implique que $i$ est récurrent pour $\mathbf{Y}$. Le point $(i)$ du théorème II.2.8, page 176, implique que $i$ est récurrent pour $\mathbf{X}$. Par ailleurs, on a

$$\int_0^\infty \mathbf{1}_{\{X_t = i\}} \, dt = \sum_{n \geq 0} D_{n+1} \mathbf{1}_{\{Y_n = i\}} \; .$$

En prenant l'esperance sous $\mathbf{P}_i$ de cette expression et en remarquant que d'après le théorème II.1.26, conditionnellement à $\{Y_n = i\}$, $D_{n+1}$ suit une loi exponentielle de paramètre $q(i)$, par Fubini, on obtient

$$\begin{aligned} \int_0^\infty p_t(i,i) \, dt &= \mathbf{E}_i \left[ \int_0^\infty \mathbf{1}_{\{X_t = i\}} dt \right] \\ &= \sum_{n \geq 0} \mathbf{E}_i \left[ D_{n+1} \mathbf{1}_{\{Y_n = i\}} \right] \\ &= q(i)^{-1} \sum_{n \geq 0} \mathbf{P}_i(Y_n = i) \\ &= q(i)^{-1} \sum_{n \geq 0} [Q^n](i,i) \; . \end{aligned}$$

Cela permet de conclure, d'après le corollaire I.3.3 (page 59) qui porte sur les chaînes de Markov. ■

## II.2.d Explosion.

Dans toute cette section on se donne un espace d'état $E$, dénombrable discret et $G = (q_{i,j})_{i,j \in E}$ un générateur infinitésimal sur $E$. On se donne également un point cimetière $\partial \notin E$ et on utilise la notation $E^* = E \cup \{\partial\}$, que l'on voit comme un espace dénombrable discret. On se donne un processus minimal

$$\big( \Omega \, ; \, \mathscr{F} \, ; \, (\mathscr{F}_t)_{t \in \mathbb{R}_+} \, ; \, \mathbf{X} = (X_t)_{t \in \mathbb{R}_+} \, ; \, G \, ; \, \mathbf{P}_\mu, \mu \in \mathscr{M}_1(E) \, \big)$$

continu à droite partout sur $\Omega$. On note $(P_t)_{t \in \mathbb{R}_+}$ le semi-groupe minimal associé à $G$. On utilisera la notation $P_t = (p_t(i,j))_{i,j \in E}$. On rappelle les notations $(q(i))_{i \in E}$ pour les paramètres des temps d'attente et $Q = (\pi_{i,j})_{i,j \in E}$ pour la matrice de transition du squelette. Pour simplifier les notations, on posera

$$D_n = D_n(\mathbf{X}) \; , \; J_n = J(\mathbf{X}) \; , \; \zeta := \zeta(\mathbf{X}) \; , \; Y_n := Y_n(\mathbf{X}) \; .$$

Le résultat suivant donne des conditions simples sous lequelles un processus markovien n'explose pas.

**Proposition II.2.10** *Soit $\mu \in \mathscr{M}_1(E)$ (en particulier $\mu(\partial) = 0$). Si l'une des conditions suivante est vérifiée alors $\mathbf{P}_\mu(\zeta < \infty) = 0$, c'est-à-dire que le processus $\mathbf{X}$ de loi d'entrée $\mu$* n'explose pas.

$(a)$ *$E$ est fini.*

$(b)$ *La quantité $c := \sup_{i \in E} q(i)$ est finie.*

$(c)$ *$\mu = \delta_i$ et $i$ est récurrent pour la chaîne squelette $\mathbf{Y}$.*

**Preuve :** Il est clair que $(a)$ implique $(b)$. On suppose que $C := \sup_{i \in E} q(i) < \infty$. D'après le théorème II.1.26, on peut trouver une suite de variables aléatoires $(\mathbf{e}_n)_{n \geq 1}$ à valeurs dans $\mathbb{R}_+$, qui, sous $\mathbf{P}_\mu$ sont indépendantes entre elles, et indépendantes également de $\mathbf{Y}$, de loi exponentielle de paramètre 1, et qui sont telles que $D_n = \mathbf{e}_n / q(Y_{n-1})$, $n \geq 1$, avec la convention $\mathbf{e}_n / 0 = \infty$. Alors,

$$\mathbf{P}_\mu\text{-p.s.} \quad \zeta = \sum_{n \geq 1} \frac{\mathbf{e}_n}{q(Y_{n-1})} \geq \frac{1}{c} \sum_{n \geq 1} \mathbf{e}_n = \infty \ .$$

La preuve de $(c)$ se fait en réutilisant les arguments de la preuve du point $(i)$ du théorème II.2.8 : nous laissons les détails au lecteur. ■

Nous donnons ensuite une condition nécessaire et suffisante pour qu'un procesuss markovien explose.

**Théorème II.2.11** **(Critère de Reuter)** *Le générateur est non-explosif si et seulement si pour tout réel $c > 0$, la seule solution positive bornée de l'équation*

$$\sum_{j \in E} q_{i,j} f(j) = c f(i) \ , \quad i \in E \tag{II.29}$$

*est la fonction nulle.*

**Preuve :** pour tout réel $c > 0$ et tout $i \in E$, on pose $g_c(i) = \mathbf{E}_i \left[ \exp(-c\zeta) \right]$, avec la convention $e^{-\infty} = 0$. On voit que $0 \leq g_c(i) \leq 1$ pour tout $i \in E$. On remarque ensuite que

$$\begin{aligned} g_c(i) &= \mathbf{E}_i \left[ \exp(-cJ_1) \exp\left( \zeta(\theta_{J_1} \mathbf{X}) \right) \right] = \mathbf{E}_i \left[ \exp(-cJ_1) \right] \sum_{j \in E} \pi_{i,j} \mathbf{E}_j \left[ \exp(-c\zeta) \right] \\ &= \mathbf{E}_i \left[ \exp(-cJ_1) \right] \sum_{j \in E} \pi_{i,j} g_c(j) = \frac{q(i)}{q(i) + c} \sum_{j \in E} \pi_{i,j} g_c(j). \end{aligned}$$

Par conséquent $(q(i) + c) g_c(i) = \sum_{j \in E \setminus \{i\}} q(i) \pi_{i,j} g_c(j)$. Si $q(i) = 0$ alors $\mathbf{P}_i(J_1 = \infty) = 1$ et $g_c(i) = 0$. Si $q(i) > 0$, alors $\pi_{i,i} = 0$ et l'égalité précédente montre que $(q(i) + c) g_c(i) = \sum_{j \in E \setminus \{i\}} q_{i,j} g_c(j)$. Dans les deux cas on a bien

$$\sum_{j \in E} q_{i,j} g_c(j) = c g_c(i) \ , \quad i \in E \ .$$

Si $G$ explose alors il existe $i \in E$ tel que $\mathbf{P}_i(\zeta < \infty) > 0$. Cela implique que $g_c(i) > 0$. La fonction $g_c$ est bien une solution positive bornée non-nulle de (II.29).

Supposons maintenant que (II.29) admette une solution positive non-nulle bornée notée $f$. Cela entraîne facilement que

$$f(i) = \frac{q(i)}{q(i) + c} \sum_{j \in E} \pi_{i,j} f(j) = \mathbf{E}_i \left[ \exp(-cJ_1) f(Y_1) \right] \ ,$$

En itérant ce raisonnement, on montre que pour tout $n \geq 1$, on a

$$f(i) = \mathbf{E}_i \left[\exp(-cJ_n) f(Y_n)\right] \leq \|f\|_\infty \mathbf{E}_i \left[\exp(-cJ_n)\right] \ , \quad i \in E \ .$$

En passant à la limite, on obtient donc par convergence dominée que $f(i) \leq \|f\|_\infty g_c(i)$, pour tout $i \in E$. Puisque $f$ est non-nulle positive, il existe $i_0 \in E$ tel que $f(i_0) > 0$, ce qui implique que $0 < g_c(i_0) = \mathbf{E}_{i_0}\left[\exp(-c\zeta)\right]$. Comme $c > 0$, on en déduit que $\mathbf{P}_{i_0}(\zeta < \infty) > 0$ et donc que $G$ est explosif, ce qui achève la preuve du théorème. ■

**Exemple II.2.12** En exploitant le critère de Reuter, nous allons donner une condition nécessaire et suffisante de non-explosion des processus de naissance et de mort : on considère le générateur $G = (q_{i,j})_{i,j\in\mathbb{N}}$ d'un processus de naissance et de mort ( $\Omega; \mathscr{F}; (\mathscr{F}_t)_{t\in\mathbb{R}_+}; \mathbf{X} = (X_t)_{t\in\mathbb{R}_+}; G; \mathbf{P}_\mu, \mu \in \mathscr{M}_1(\mathbb{N})$ ). On suppose que $G$ est irréductible c'est-à-dire que

$$\forall i \in \mathbb{N} \ , \quad q_{i,i+1} q_{i+1,i} > 0 \ .$$

On fixe $c > 0$ et on se donne une fonction $f : \mathbb{N} \to \mathbb{R}_+$, que l'on suppose bornée et telle que $G \cdot f = cf$. On voit que cela implique que

$$cf(0) = q(0)(f(1) - f(0)) \ , \quad cf(i) = q_{i,i+1} f(i+1) - (q_{i,i+1} + q_{i,i-1}) f(i) + q_{i,i-1} f(i-1) \ , \quad i \in \mathbb{N}. \tag{II.30}$$

Commençons par quelque remarques : on voit que si $f(0) = 0$ alors $f(i) = 0$, $i \geq 1$. S'il existe une solution non-nulle alors nécessairement $f(0) > 0$. Par ailleurs, si $f$ est solution du système alors $(af(i), i \in \mathbb{N})$ l'est aussi pour tout $a \geq 0$. Cela implique que s'il y a une solution non-nulle, il en existe une telle que $f(0) = 1$.

Pour tout $i \in \mathbb{N}$, on pose $g(i) = f(i+1) - f(i)$ ; si $f$ est solution du système $G \cdot f = cf$, alors on a

$$\forall i \geq 1 \ , \quad g(i) = \frac{c}{q_{i,i+1}} f(i) + \frac{q_{i,i-1}}{q_{i,i+1}} g(i-1) \ . \tag{II.31}$$

Cela permet de montrer par récurrence sur $i$ que

$$\begin{aligned} g(i) \quad &= \quad \frac{c}{q_{i,i+1}} f(i) + \frac{c \cdot q_{i,i-1}}{q_{i,i+1} q_{i-1,i}} f(i-1) + \frac{c \cdot q_{i,i-1} q_{i-1,i_2}}{q_{i,i+1} q_{i-1,i} q_{i-2,i+1}} f(i-2) + \\ &\quad + \ \ldots \ + \frac{c \cdot q_{i,i-1} \cdots q_{2,1}}{q_{i,i+1} q_{i-1,i} \cdots q_{1,2}} f(1) + \frac{q_{i,i-1} \cdots q_{1,0}}{q_{i,i+1} \cdots q_{0,1}} \cdot \frac{c}{q(0)} f(0) \ . \end{aligned} \tag{II.32}$$

Cette formule implique tout d'abord que $g(i) \geq 0$, $i \in \mathbb{N}$. Une fonction positive solution de $G \cdot f = cf$ est donc croissante. Par conséquent $f(i) \geq f(0)$ pour tout $i \in \mathbb{N}$ et alors (II.32) implique que

$$g(i) \geq cf(0) \left( \frac{1}{q_{i,i+1}} + \ldots + \frac{q_{i,i-1} \cdots q_{2,1}}{q_{i,i+1} q_{i-1,i} \cdots q_{1,2}} + \frac{q_{i,i-1} \cdots q_{1,0}}{q_{i,i+1} \cdots q_{0,1} q(0)} \right) . \tag{II.33}$$

On pose

$$S := \sum_{i=1}^{\infty} \left( \frac{1}{q_{i,i+1}} + \frac{q_{i,i-1}}{q_{i,i+1} q_{i-1,i}} \ldots + \frac{q_{i,i-1} \cdots q_{2,1}}{q_{i,i+1} q_{i-1,i} \cdots q_{1,2}} + \frac{q_{i,i-1} \cdots q_{1,0}}{q_{i,i+1} \cdots q_{0,1} q(0)} \right) \in [0, \infty] \ . \tag{II.34}$$

Supposons que $S = \infty$. S'il existe une solution bornée positive non-nulle alors, les remarques précédentes impliquent qu'il en existe une, notée $f$ telle que $f(0) = 1$. On observe ensuite que si $f$ est bornée, alors $g$ l'est aussi, ce qui contredit (II.33). Cela montre que si $S = \infty$, le processus n'explose pas par le critère de Reuter (théorème II.2.11).

Montrons la réciproque : on se donne une fonction $f : \mathbb{N} \to \mathbb{R}_+$ qui satisfait (II.30) et telle que $f(0) = 1$. Une telle fonction peut toujours être définie. La question est de savoir si elle peut-être bornée. La remarque que $f$ est croissante combinée avec (II.32) implique que

$$f(i+1) - f(i) = g(i) \leq f(i) \left( \frac{c}{q_{i,i+1}} + \ldots + \frac{c \cdot q_{i,i-1} \cdots q_{2,1}}{q_{i,i+1} q_{i-1,i} \cdots q_{1,2}} + \frac{c \cdot q_{i,i-1} \cdots q_{1,0}}{q_{i,i+1} \cdots q_{0,1} q(0)} \right) . \tag{II.35}$$

Ce qui implique que

$$\begin{aligned} f(i+1) \quad &\leq \quad f(i) \left( 1 + \frac{c}{q_{i,i+1}} + \ldots + \frac{c \cdot q_{i,i-1} \cdots q_{2,1}}{q_{i,i+1} q_{i-1,i} \cdots q_{1,2}} + \frac{c \cdot q_{i,i-1} \cdots q_{1,0}}{q_{i,i+1} \cdots q_{0,1} q(0)} \right) \\ &\leq \quad f(i) \exp \left( \frac{c}{q_{i,i+1}} + \ldots + \frac{c \cdot q_{i,i-1} \cdots q_{2,1}}{q_{i,i+1} q_{i-1,i} \cdots q_{1,2}} + \frac{c \cdots q_{i,i-1} \cdots q_{1,0}}{q_{i,i+1} \cdots q_{0,1} q(0)} \right) \end{aligned}$$

d'après l'inégalité de convexité élémentaire $1+x \leq e^x$. Cela implique que $f(i) \leq \exp(cS)$, pour tout $i \in E$, où $S$ est défini par (II.34). En particulier si $S$ est une quantité finie, $f$ est une solution de $G \cdot f = cf$ qui est non-nulle positive bornée. Par le critère de Reuter, $G$ explose.

*En résumé,* nous avons montré que $G$ n'explose pas si et seulement si

$$\sum_{i=1}^{\infty} \left( \frac{1}{q_{i,i+1}} + \frac{q_{i,i-1}}{q_{i,i+1}q_{i-1,i}} \dots + \frac{q_{i,i-1}\cdots q_{2,1}}{q_{i,i+1}q_{i-1,i}\cdots q_{1,2}} + \frac{q_{i,i-1}\cdots q_{1,0}}{q_{i,i+1}\cdots q_{0,1}q(0)} \right) = \infty . \tag{II.36}$$

□

## II.2.e Mesures invariantes.

Dans toute cette section on garde les notations $X_t$, $\mathscr{F}_t$, $G$, etc, pour un processus minimal.

**Définition II.2.13** (*Mesure invariante*) Soit $G = (q_{i,j})_{i,j\in E}$ un générateur infinitésimal. Une mesure $m = (m(i))_{i\in E}$ est dit $G$-invariante si elle n'est pas la mesure nulle et si $mG = 0$, c'est-à-dire

$$\forall j \in E \; , \quad \sum_{i\in E} m(i) q_{i,j} = 0 \; .$$

(Ici $m.G = 0$ a bien un sens, car c'est équivalent à $q(j)m(j) = \sum_{i\in E\setminus\{j\}} m(i)q_{i,j}$, où on ne somme que des termes positifs.) □

**Lemme II.2.14** *Soit $G = (q_{i,j})_{i,j\in E}$, un générateur infinitésimal que l'on suppose* irréductible. *Cela implique que $q(i) > 0$, pour tout $i \in E$. Soit $\mu \in \mathscr{M}_1(E)$ qui est liée à $m$ par les équations*

$$q(i)m(i) = \mu(i) \; , \quad i \in E \; . \tag{II.37}$$

*Alors $m$ est $G$-invariante si et seulement si $\mu$ est $Q$-invariante (c'est-à-dire $\mu.Q = \mu$).*

**Preuve :** le fait que $G$ soit irréductible implique qu'il n'y ait pas d'état absorbant et donc que $q(i) > 0$, pour tout $i \in E$. On a donc $\pi_{i,i} = 0$ et $q(i)\pi_{i,j} = q_{i,j}$ dès que $i \neq j$. Soient deux mesures reliées par (II.37). Alors $m \cdot G = 0$ implique que

$$q(j)m(j) = \sum_{i\in E\setminus\{j\}} m(i)q_{i,j} = \sum_{i\in E} m(i)q(i)\pi_{i,j} \; ,$$

ce qui termine la preuve du lemme. ■

La proposition suivante, qui établit l'existence de mesures invariantes, se déduit facilement du théorème analogue pour les chaînes de Markov grâce au lemme précédent.

**Proposition II.2.15** *Soit $G$ un générateur* irréductible et récurrent. *Alors, il existe une mesure $G$-invariante $m$. Par ailleurs, si $m'$ est une autre mesure $G$-invariante, alors il existe un réel $c > 0$ tel que*

$$\forall i \in E \; , \quad m'(i) = cm(i) \; .$$

*Autrement dit, il existe une unique mesure invariante à un coefficient positif de proportionnalité près. Par ailleurs, pour tout $i \in E$, on a : $0 < m(i) < \infty$ et pour tout $i_0 \in E$,*

$$\forall i \in E \; , \quad m(i) = m(i_0)\, q(i_0)\, \mathbf{E}_{i_0}\left[ \int_0^{T_{i_0}^+} \mathbf{1}_{\{X_t = i\}}\, dt \right] \; , \tag{II.38}$$

*où on rappelle que $T_{i_0}^+ = \inf\{t \geq J_1 : X_t = i_0\}$.*

**Preuve :** puisque $G$ est irréductible et récurrent, il en est de même pour la matrice de transition du squelette $Q$. Le théorème d'existence I.4.2 (page 79) et le théorème d'unicité I.4.3 (page 80) sur les chaînes de Markov implique qu'il existe une mesure $\mu$ qui est $Q$-invariante et unique à coefficient multiplicatif près. Le lemme II.2.14 permet alors de conclure la preuve du premier point de la proposition, à savoir l'existence et l'unicité à coefficient multiplicatif $> 0$ près d'une mesure $G$-invariante.

On pose ensuite $T_{i_0}^{(1)}(\mathbf{Y}) = \inf\{n \geq 1 : Y_n = i_0\}$, le premier temps de retour en $i_0$ pour la chaîne du squelette. Montrons (II.38) : on remarque que puisque $G$ est irréductible récurrent, $G$ n'explose pas et d'autre part $\mathbf{P}_{i_0}(T_{i_0}^{(1)}(\mathbf{Y}) < \infty) = \mathbf{P}_{i_0}(T_{i_0}^{+} < \infty) = 1$. On remarque alors que

$$\int_0^{T_{i_0}^+} \mathbf{1}_{\{X_t=i\}}\, dt = \sum_{n=0}^{T_{i_0}^{(1)}(\mathbf{Y})-1} D_{n+1}\mathbf{1}_{\{Y_n=i\}} = \sum_{n\geq 0} D_{n+1}\mathbf{1}_{\{Y_n=i\,;\, n<T_{i_0}^{(1)}(\mathbf{Y})\}} .$$

Par Fubini on a donc

$$\mathbf{E}_{i_0}\left[\int_0^{T_{i_0}^+} \mathbf{1}_{\{X_t=i\}}\, dt\right] = \sum_{n\geq 0} \mathbf{E}_{i_0}\left[D_{n+1}\mathbf{1}_{\{Y_n=i\,;\, n<T_{i_0}^{(1)}(\mathbf{Y})\}}\right] .$$

Or le théorème II.1.26 implique que conditionnellement à $\mathbf{Y}$, les variables $D_n$, $n \geq 1$ sont indépendantes de paramètres respectifs $q(Y_{n-1})$, $n \geq 1$. Donc, on a

$$\mathbf{E}_{i_0}\left[\int_0^{T_{i_0}^+} \mathbf{1}_{\{X_t=i\}}\, dt\right] = \frac{1}{q(i)}\sum_{n\geq 0} \mathbf{E}_{i_0}\left[\mathbf{1}_{\{Y_n=i\,;\, n<T_{i_0}^{(1)}(\mathbf{Y})\}}\right] = \frac{1}{q(i)}\mathbf{E}_{i_0}\Big[\sum_{n=0}^{T_{i_0}^{(1)}(\mathbf{Y})-1} \mathbf{1}_{\{Y_n(X)=i\}}\Big] .$$

On reconnaît dans le membre de droite la mesure $\nu_{i_0}$ définie au théorème d'existence I.4.2 (page 79) qui est l'unique mesure $Q$-invariante telle que $\nu_{i_0}(i_0) = 1$. On conclut alors grâce au lemme II.2.14. ■

**Définition II.2.16** (*Réversibilité*) Un générateur infinitésimal $G = (q_{i,j})_{i,j\in E}$ est dit réversible s'il existe une mesure $m$ non-identiquement nulle telle que

$$\forall i, j \in E\ , \quad m(i)q_{i,j} = m(j)q_{j,i}\ .$$

On dit que la mesure $m$ rend réversible $G$. □

**Lemme II.2.17** *Soit $G = (q_{i,j})_{i,j\in E}$, un générateur infinitésimal rendu réversible par la mesure $m$. Alors $m.G = 0$, c'est-à-dire que $m$ est $G$-invariante.*

**Preuve :** pour tout $j \in E$, $\sum_{i\in E} m(i)q_{i,j} = \sum_{i\in E} m(j)q_{j,i} = m(j)\sum_{i\in E} q_{j,i} = 0$. ■

Ce résultat permet de trouver facilement des mesures invariantes pour les générateurs qui sont réversibles comme le montre l'exemple suivant qui traite du cas des processus de naissance et de mort.

**Exemple II.2.18** On se donne $G = (q_{i,j})_{i,j\in\mathbb{N}}$, le générateur d'un processus de naissance et de mort : les paramètres de naissance sont notés $(a_i)_{i\geq 0}$ et les paramètres de mort, $(b_i)_{i\geq 1}$. On le suppose irréductible ce qui est équivalent à supposer que

$$\forall i \in \mathbb{N}\ , \quad a_i b_{i+1} > 0\ .$$

Supposons qu'il soit rendu réversible par une mesure non-identiquement nulle $m$, ce qui est équivalent à supposer

$$\forall i \in \mathbb{N}\ , \quad m(i)q_{i,i+1} = m(i)a_i = m(i+1)q_{i+1,i} = m(i+1)b_i\ .$$

Ce système d'équation est équivalent à

$$\forall i \geq 1 \ , \quad m(i) = m(0) \prod_{\ell=1}^{i} \frac{q_{\ell-1,\ell}}{q_{\ell,\ell-1}} = m(0) \prod_{\ell=1}^{i} \frac{a_{\ell-1}}{b_\ell} \ . \tag{II.39}$$

□

Le thèorème suivant justifie le terme de mesure invariante dans la définition II.2.13.

**Théorème II.2.19** *Soit $G$ un générateur infinitésimal* irréductible récurrent. *On note $(P_t)_{t\in\mathbb{R}_+}$ le semi-groupe minimal associé. Soit $m$ une mesure $G$-invariante. Alors, pour tout $t \in \mathbb{R}_+$, $m$ est $P_t$-invariante, c'est-à-dire que $m.P_t = m$ :*

$$\forall j \in E \ , \quad m(j) = \sum_{i\in E} m(i) p_t(i,j) \ .$$

**Preuve :** on rappelle que $T_i^+ = \inf\{t \geq J_1 : X_t = i\}$. D'après la proposition II.2.15 on a pour tout $i \in E$

$$\forall j \in E \ , \quad m(j) = m(i)q(i)\mathbf{E}_i \left[ \int_0^{T_i^+} \mathbf{1}_{\{X_t=j\}} \, dt \right] \ .$$

On fixe $t > 0$ et $i$. La propriété de Markov forte au temps d'arrêt $T_i^+$ implique facilement que

$$\mathbf{E}_i \left[ \int_{T_i^+(X)}^{t+T_i^+(X)} \mathbf{1}_{\{X_s=j\}} \, ds \right] = \mathbf{E}_i \left[ \int_0^t \mathbf{1}_{\{X_s=j\}} \, ds \right] \ .$$

Par conséquent on a

$$\begin{aligned}
m(i)q(i)\mathbf{E}_i \left[ \int_t^{t+T_i^+} \mathbf{1}_{\{X_s=j\}} \, ds \right] &= m(i)q(i)\mathbf{E}_i \left[ \int_0^{t+T_i^+} \mathbf{1}_{\{X_s=j\}} \, ds \right] - m(i)q(i)\mathbf{E}_i \left[ \int_0^t \mathbf{1}_{\{X_s=j\}} \, ds \right] \\
&= m(i)q(i)\mathbf{E}_i \left[ \int_0^{t+T_i^+} \mathbf{1}_{\{X_s=j\}} \, ds \right] - m(i)q(i)\mathbf{E}_i \left[ \int_{T_i^+}^{t+T_i^+} \mathbf{1}_{\{X_s=j\}} \, ds \right] \\
&= m(i)q(i)\mathbf{E}_i \left[ \int_0^{T_i^+} \mathbf{1}_{\{X_s=j\}} \, ds \right] = m(j) \ .
\end{aligned}$$

Par ailleurs, un changement de variable simple combiné avec Fubini et la propriété de Markov impliquent que

$$\begin{aligned}
m(j) &= m(i)q(i)\mathbf{E}_i \left[ \int_t^{t+T_i^+} \mathbf{1}_{\{X_s=j\}} \, ds \right] = m(i)q(i) \int_0^\infty ds \, \mathbf{P}_i(X_{t+s} = j \, ; \, s < T_i^+) \\
&= m(i)q(i) \int_0^\infty ds \sum_{k\in E} \mathbf{P}_i(X_s = k \, ; \, s < T_i^+) p_t(k,j) \\
&= \sum_{k\in E} \left( m(i)q(i)\mathbf{E}_i \left[ \int_0^{T_i^+} \mathbf{1}_{\{X_s=k\}} \, ds \right] \right) p_t(k,j) \\
&= \sum_{k\in E} m(k) p_t(k,j) \ ,
\end{aligned}$$

qui montre bien que $m$ est $P_t$-invariante. ■

On s'intéresse maintenant à l'existence d'une *probabilité invariante* (loi stationnaire).

**Définition II.2.20** Comme pour les chaîne de Markov, on introduit les notions suivantes.

**(a)** (*Etat récurrent positif*) On dit qu'un état est récurrent positif ssi $\mathbf{E}_i[T_i^+] < \infty$. (On voit que $i$ est récurrent selon la définition II.2.7, page 176, d'après le théorème II.2.8 **(i)**, page 176)

**(b)** (*Etat récurrent nul*) On dit qu'un état est récurrent nul ssi $\mathbf{P}_i(T_i^+ < \infty) = 1$ mais $\mathbf{E}_i[T_i^+] = \infty$. (On voit que $i$ est récurrent selon la définition II.2.7, page 176, d'après le théorème II.2.8 **(i)**, page 176.)

**(c)** Si $\mathbf{P}_i(T_i^+ < \infty) < 1$, alors l'état est transient selon la définition II.2.7, page 176, d'après le théorème II.2.8 **(ii)**, page 176. □

Le lien entre la récurrence positive et l'exitence d'une probabilité invariante est presque le même que dans le cas des chaînes de Markov

**Théorème II.2.21** *Soit $G$ un générateur infinitésimal* **irréductible** *(on insiste sur le fait que l'on ne suppose pas $G$ récurrent a priori). Les assertions suivantes sont équivalentes.*

**(i)** *Il existe un état récurrent positif.*

**(ii)** *Tous les états sont récurrents positifs.*

**(iii)** *Le générateur infinitésimal $G$ est non-explosif et il admet une mesure de probabilité $G$-invariante.*

*Si l'une de ces trois conditions équivalentes est vérifiée, alors l'unique mesure de probabilité $G$-invariante est donnée par*

$$\forall i \in E\,, \quad m(i) = \frac{1}{q(i)\mathbf{E}_i\left[T_i^+\right]}\,.$$

**Preuve :** on rappelle que $i$ est absorbant ssi $q(i) = 0$. Si $G$ est supposé irréductible, il n'y a aucun état absorbant et donc

$$\forall i \in E\,, \quad q(i) > 0\,. \tag{II.40}$$

On montre tout d'abord que **(i)** implique **(ii)**. Soit $i_0$ récurrent positif. Il est donc récurrent. Puisque $G$ est supposé irréductible, tous les états sont récurrents, d'après le théorème II.2.8. Pour tout $i \in E$, on définit la mesure $m_i = (m_i(j), j \in E)$, en posant

$$\forall j \in E\,, \quad m_i(j) = q(i)\mathbf{E}_i\Big[\int_0^{T_i^+} \mathbf{1}_{\{X_t = j\}}\,dt\Big]\,. \tag{II.41}$$

On fixe $t \in [0, T_i^+[$ et $j \in E$. On remarque que $\{X_t = j\} = \{X_t = j\} \cap \{t < \zeta\}$, car après l'explosion eventuelle, le processus minimal reste au point cimetière. On a donc $\sum_{j\in E} \mathbf{1}_{\{X_t = j\,;\, t < T_i^+\,;\, t<\zeta\}} = \mathbf{1}_{\{t < \zeta \wedge T_i^+\}}$ et en sommant sur $j$ dans (II.41), on obtient

$$\forall i \in E\,, \quad \sum_{j\in E} m_i(j) = q(i)\mathbf{E}_i\big[\zeta \wedge T_i^+\big]\,. \tag{II.42}$$

Comme $G$ est supposé récurrent, la proposition II.2.10 **(iii)** (page 178) implique que $\mathbf{P}_i(\zeta = \infty) = 1$. Par conséquent, on a

$$\forall i \in E\,, \quad \sum_{j\in E} m_i(j) = q(i)\mathbf{E}_i\big[T_i^+\big]\,. \tag{II.43}$$

D'après la proposition II.2.15, $m_i$ est l'unique mesure $G$-invariante telle que $m_i(i) = 1$ et (II.38) page 180 permet d'affirmer que

$$\forall i, j \in E \ , \quad m_i(j) = m_i(i_0) m_{i_0}(j) \ . \tag{II.44}$$

Par (II.43) appliqué à $i = i_0$ donne

$$\sum_{j \in E} m_{i_0}(j) = q(i_0) \mathbf{E}_{i_0} [T^+_{i_0}] \ < \infty \ , \tag{II.45}$$

car $i_0$ est supposé récurrent positif. Donc $m_{i_0}$ est de masse totale finie. Par (II.44), il en est de même pour toutes les mesures $m_i$ et (II.43) implique que $\mathbf{E}_i[T_i^+] < \infty$, pour tout $i \in E$, ce qui implique **(ii)**.

On suppose toujours (**i**) : on a montré que $G$ est irréductible récurrent et qu'il existe une mesure invariante $m_{i_0}$ de masse finie. On a montré qu'il en est de même avec toutes les mesures $m_i$. Par le théorème II.2.15, toutes les mesures $G$-invariantes sont proportionnelles entre elles, elles ont toute une masse totale finie. Il existe donc une unique mesure de probabilité $G$-invariante que l'on note $m$. Pour tout $i \in E$, il existe une constante $c(i) \in ]0, \infty[$ telle que $m(j) = c(i) m_i(j)$, pour tout $j \in E$. Comme $m_i(i) = 1$, on a $m(i) = c(i)$. D'autre part on a

$$1 = \sum_{j \in E} m(j) = c(i) \sum_{j \in E} m_i(j) = c(i) q(i) \mathbf{E}_i \left[ T_i^+ \right] \ ,$$

d'après (II.43). Par conséquent

$$\forall i \in E \ , \quad m(i) = \frac{1}{q(i) \mathbf{E}_i \left[ T_i^+ \right]} \ .$$

Montrons que **(ii)** implique **(iii)**. Puisque $G$ est irréductible récurrent, la proposition II.2.10 **(iii)** implique que pour tout $i \in E$, $\mathbf{P}_i(\zeta < \infty) = 0$. Markov simple au temps 0 implique ensuite que pour tout $\mu \in \mathscr{M}_1(E)$ on a $\mathbf{P}_\mu(\zeta < \infty) = 0$ et $G$ n'explose pas. D'autre part, si on pose

$$\forall j \in E \ , \quad m(j) = \frac{1}{q(i_0) \mathbf{E}_{i_0} [T^+_{i_0}]} m_{i_0}(j) \ ,$$

on voit facilement que $m$ est $G$-invariante, puisque $m_{i_0}$ l'est et qu'elle est de masse totale 1, par (II.45), ce qui prouve **(iii)**.

Montrons que **(iii)** implique **(i)** : on suppose que $G$ n'explose pas et on suppose qu'il existe une probabilité $G$-invariante notée $m$. On fixe $i_0 \in E$ tel que $m(i_0) > 0$ (il en existe forcément un). On définit la mesure $\nu$ par

$$\forall i \in E \ , \quad \nu(i) = \frac{q(i) m(i)}{q(i_0) m(i_0)} \ ,$$

Le lemme II.2.14 implique que $\nu$ est $Q$-invariante. On fixe $j \in E$. On observe ce qui suit.

$$\begin{aligned} \nu(j) &= \sum_{i \in E} \nu(i) \pi_{i,j} \\ &= \pi_{i_0,j} + \sum_{i \in E \backslash \{i_0\}} \nu(i) \pi_{i,j} \ . \end{aligned}$$

On montre facilement par récurrence que

$$\nu(j) = \pi_{i_0,j} + \sum_{i_1 \in E \backslash \{i_0\}} \pi_{i_0,i_1} \pi_{i_1,j} \ + \ \ldots \ + \sum_{i_1, \ldots, i_n \in E \backslash \{i_0\}} \pi_{i_0,i_1} \ldots \pi_{i_n,j}$$

$$+ \sum_{i_1,\ldots,i_{n+1} \in E\setminus\{i_0\}} \nu(i_{n+1})\pi_{i_{n+1},i_n} \ldots \pi_{i_0,j}. \tag{II.46}$$

Cela implique que

$$\begin{aligned}
\nu(j) &\geq \pi_{i_0,j} + \sum_{i_1 \in E\setminus\{i_0\}} \pi_{i_0,i_1}\pi_{i_1,j} \;+\; \ldots \;+ \sum_{i_1,\ldots,i_n \in E\setminus\{i_0\}} \pi_{i_0,i_1} \ldots \pi_{i_n,j} \\
&= \mathbf{P}_{i_0}(Y_1 = j\,;\, T_{i_0}^{(1)}(\mathbf{Y}) > 1) + \ldots + \mathbf{P}_{i_0}(Y_n = j\,;\, T_{i_0}^{(1)}(\mathbf{Y}) > n)\;,
\end{aligned}$$

où $T_{i_0}^{(1)}(\mathbf{Y}) = \inf\{n \geq 1 \,:\, Y_n = i_0\}$ (avec la convention $\inf \emptyset = \infty$) est le premier temps de retour de la chaîne du squelette $\mathbf{Y}$ en $i_0$. En faisant tendre $n$ vers l'infini, on obtient

$$\nu(j) \geq \mathbf{E}_{i_0}\Big[\sum_{n=0}^{T_{i_0}^{(1)}(\mathbf{Y})-1} \mathbf{1}_{\{Y_n=j\}}\Big] = \sum_{n\geq 0} \mathbf{P}_{i_0}\big(Y_n = j\,;\, n < T_{i_0}^{(1)}(\mathbf{Y})\big) \tag{II.47}$$

On remarque ensuite que consitionnellement à $Y_n = j$, $D_{n+1}$ suit une loi exponentielle de durée $q(j)$, qui est d'espérance $1/q(j)$. On a alors

$$\begin{aligned}
\frac{m(j)}{q(i_0)m(i_0)} = \frac{\nu(j)}{q(j)} &\geq \sum_{n\geq 0} \frac{1}{q(j)} \mathbf{P}_{i_0}\big(Y_n = j\,;\, n < T_{i_0}^{(1)}(\mathbf{Y})\big) \\
&= \sum_{n\geq 0} \mathbf{E}_{i_0}\big[D_{n+1}\mathbf{1}_{\{Y_n=j\,;\, n<T_{i_0}^{(1)}(\mathbf{Y})\}}\big] \\
&= \mathbf{E}_{i_0}\Big[\int_0^{T_{i_0}^+} \mathbf{1}_{\{X_t=j\}}\, dt\Big].
\end{aligned}$$

On a donc montré que

$$\forall j \in E\,, \quad \frac{m(j)}{q(i_0)m(i_0)} = \frac{\nu(j)}{q(j)} \geq \mathbf{E}_{i_0}\Big[\int_0^{T_{i_0}^+} \mathbf{1}_{\{X_t=j\}}\, dt\Big]\;. \tag{II.48}$$

On fixe $t \in [0, T_{i_0}^+[$ et $j \in E$. On remarque que $\{X_t = j\} = \{X_t = j\} \cap \{t < \zeta\}$, car après l'explosion eventuelle, le processus minimal reste au point cimetière. On a donc $\sum_{j\in E} \mathbf{1}_{\{X_t=j\,;\, t<T_{i_0}^+\,;\, t<\zeta\}} = \mathbf{1}_{\{t<\zeta\wedge T_{i_0}^+\}}$ et en sommant sur $j$ dans (II.48), on obtient

$$\frac{1}{q(i_0)m(i_0)} \geq \mathbf{E}_{i_0}\big[\zeta \wedge T_{i_0}^+\big]\;.$$

Comme on a supposé que le générateur $G$ n'est *pas explosif*, on a $\mathbf{P}_{i_0}(\zeta = \infty) = 1$. Donc $\mathbf{E}_{i_0}\big[\zeta \wedge T_{i_0}^+\big] = \mathbf{E}_{i_0}\big[T_{i_0}^+\big]$ et donc que

$$\mathbf{E}_{i_0}\big[T_{i_0}^+\big] \leq \frac{1}{q(i_0)m(i_0)} < \infty\;.$$

On en déduit donc que $i_0$ est récurrent positif, ce qui implique (**i**). Cela termine la preuve du théorème. ■

**Remarque II.2.22** Contrairement aux chaînes de Markov, il existe des processus de Markov dont le générateurs est irréductible, qui ne sont pas récurrents mais qui admettent quand même une probabilité invariante (ils explosent nécessairement, d'après le théorème précédent). On renvoie aux exercices pour un exemple d'un tel processus. □

### II.2.f Théorèmes ergodiques et convergence vers la loi stationnaire.

Sur un espace $E$ dénombrable discret on rappelle le lemme suivant dont la preuve est laissée en exercice.

**Lemme II.2.23** *Soient $\mu, \mu_n \in \mathscr{M}_1(E)$, $n \in \mathbb{N}$. Les assertion suivantes ont équivalentes.*

$(a)$ *$\mu_n \to \mu$ en loi sur $E$.*

$(b)$ *Pour tout $i \in E$, $\mu_n(i) \to \mu(i)$.*

$(c)$ $\sum_{i\in E} |\mu_n(i) - \mu(i)| \to 0$.

**Preuve.** Exercice. ■

**Théorème II.2.24** **(Convergence vers la loi stationnaire)** *Soit $G$ un générateur infinitésimal que l'on suppose* **irréductible et récurrent positif**. *On note $(X_t)_{t\in\mathbb{R}_+}$ le processus minimal associé. Il admet une unique probabilité invariante notée $m$ qui est donnée par*

$$\forall i \in E \;, \quad m(i) = \frac{1}{q(i)\mathbf{E}_i[T_i^+]} \;.$$

*Pour toute mesure d'entrée $\mu \in \mathscr{M}_1(E)$, on a*

$$\lim_{t\to\infty} \sum_{j\in E} |\, \mathbf{P}_\mu(X_t = j) - m(j)\,| \;= 0 \;.$$

**Preuve :** on remarque tout d'abord que, puisque $G$ est irréductible récurrent, la proposition II.2.10 **(iii)** (page 178) implique que $G$ n'explose pas. Donc le semi-groupe minimal $(P_t)_{t\in\mathbb{R}_+}$ est un *vrai* semi-groupe markovien. On fixe ensuite $h > 0$ et on note $Z_n^h = X_{hn}$, pour tout $n \in \mathbb{N}$. Il est facile de voir que pour tout $n \in \mathbb{N}$, tous $i_0, \dots, i_n \in E$ et tout $\mu \in \mathscr{M}_1(E)$ on a

$$\mathbf{P}_\mu(Z_0^h = i_0; \dots ; Z_n^h = i_n) = \mu(i_0) p_h(i_0, i_1) \dots p_h(i_{n-1}, i_n) \;.$$

Par conséquent, $(Z_n^h)_{n\in\mathbb{N}}$ est une chaîne de Markov de matrice de transition $P_h$. La proposition II.2.6 **(iv)** (page 175) permet d'affirmer que $p_h(i,j) > 0$, pour tous $i, j \in E$, et $P_h$ est irréductible apériodique. Par le théorème II.2.19 (page 182), $m$ est une probabilité invariante pour $P_h$. Le théorème I.4.7, page 83 pour les chaînes de Markov, implique que la matrice $P_h$ est récurrente positive. Le théorème I.4.17 de la convergence des chaînes de Markov vers leur loi stationnaire, page 93, implique que

$$\forall h > 0 \,,\; \forall \mu \in \mathscr{M}_1(E) \,, \quad \lim_{n\to\infty} \sum_{j\in E} \Big| \mathbf{P}_\mu(Z_n^h = j) - m(j) \Big| \;= 0 \;. \tag{II.49}$$

On fixe ensuite $t$. Il existe $n \in \mathbb{N}$ tel que $nh \le t < (n+1)h$, c'est-à-dire que $n = \lfloor t/h \rfloor$, où $\lfloor \cdot \rfloor$ désigne la fonction partie entière. Pour tous $i, j \in E$, on a

$$\begin{aligned} |p_t(i,j) - p_{nh}(i,j)| &= \Big| \sum_{k\in E} p_{t-nh}(i,k) p_{nh}(k,j) - p_{nh}(i,j) \Big| \\ &\le (1 - p_{t-nh}(i,i)\,)\, p_{nh}(i,j) + \sum_{k\in E\backslash\{i\}} p_{t-nh}(i,k) p_{nh}(k,j) \\ &\le 1 - p_{t-nh}(i,i) + \sum_{k\in E\backslash\{i\}} p_{t-nh}(i,k) \end{aligned}$$

$$= \quad 2(1 - p_{t-nh}(i,i)\,).$$

Or $p_{t-nh}(i,i) = \mathbf{P}_i(X_{t-nh} = i) \geq \mathbf{P}_i(J_1 > t - nh) \geq \mathbf{P}_i(J_1 > h) = e^{-q(i)h}$. On a donc

$$\forall i, j \in E\,, \forall t, h > 0\,, \quad |p_t(i,j) - p_{\lfloor t/h \rfloor h}(i,j)| \leq 2\big(1 - e^{-q(i)h}\big)$$

On fixe $\mu \in \mathscr{M}_1(E)$, et on observe ensuite que

$$\begin{aligned} \Big|\mathbf{P}_\mu(X_t = j) - \mathbf{P}_\mu(Z^h_{\lfloor t/h \rfloor h} = j)\Big| &= \Big|\sum_{i \in E} \mu(i)\big(p_t(i,j) - p_{\lfloor t/h \rfloor h}(i,j)\big)\Big| \\ &\leq 2\sum_{i \in E} \mu(i)\big(1 - e^{-q(i)h}\big)\ . \end{aligned}$$

Par conséquent pour tout $t, h > 0$, toute loi d'entrée $\mu \in \mathscr{M}_1(E)$ et tout $j \in E$, on a

$$|\mathbf{P}_\mu(X_t = j) - m(j)| \leq \Big|\mathbf{P}_\mu(Z^h_{\lfloor t/h \rfloor h} = j) - m(j)\Big| + 2\sum_{i \in E} \mu(i)\big(1 - e^{-q(i)h}\big)\ . \qquad \text{(II.50)}$$

En utilisant (II.49), on voit que cela implique que

$$\limsup_{t \to \infty} |\mathbf{P}_\mu(X_t = j) - m(j)| \ \leq\ 2\sum_{i \in E} \mu(i)\big(1 - e^{-q(i)h}\big)\ . \qquad \text{(II.51)}$$

On observe que pour tout $i \in E$, $0 \leq \mu(i)(1 - e^{-q(i)h}) \leq \mu(i)$, que $\sum_{i \in E} \mu(i) = 1$ et que $\lim_{h \to 0} \mu(i)(1 - e^{-q(i)h}) = 0$. Le théorème de convergence dominée s'applique, et il entraîne

$$\lim_{h \downarrow 0} \downarrow \sum_{i \in E} \mu(i)\big(1 - e^{-q(i)h}\big) = 0\ .$$

Comme le membre de gauche dans (II.51) ne dépend pas de $h$, on en déduit que

$$\forall j \in E\,,\ \forall \mu \in \mathscr{M}_1(E)\,, \quad \lim_{t \to \infty} |\mathbf{P}_\mu(X_t = j) - m(j)| = 0\ . \qquad \text{(II.52)}$$

On conclut facilement par le lemme II.2.23. ■

Nous concluons cette étude en énonçant les théorèmes ergodiques. Le principe de leur preuve est le même pour le cas des chaînes de Markov : on peut introduire les excursions en dehors d'un état qui sont indépendantes et utiliser la forme des mesures invariantes donnée à la proposition II.2.15 (page 180). Nous laissons en exercice la preuve de l'énoncé suivant.

**Théorème II.2.25** **(Théorème ergodique)** *Soit $G$ un générateur infinitésimal que l'on suppose* **irréductible et récurrente**. *On note $(X_t)_{t \in \mathbb{R}_+}$ le processus minimal associé. On note $m$ une mesure $G$-invariante. On se donne également $f, g : E \to \mathbb{R}$, deux fonctions telles que*

$$\langle m, |f|\rangle\ < \infty\ , \quad \langle m, |g|\rangle\ < \infty \quad \text{et} \quad \langle m, g\rangle \neq 0\ .$$

**(i)** **(Théorème ergodique quotient)** *Pout toute loi d'entrée $\mu \in \mathscr{M}_1(E)$, on a*

$$\mathbf{P}_\mu\text{ - p.s.} \quad \frac{\int_0^t f(X_s)\,ds}{\int_0^t g(X_s)\,ds} \ \xrightarrow[t \to \infty]{}\ \frac{\langle m, f\rangle}{\langle m, g\rangle}\ .$$

**(ii)** **(Théorème ergodique)** *On suppose que $G$ est* **irréductible, et récurrente positive**. *On suppose que $m$ est l'unique mesure de* **probabilité $G$-invariante**. *Alors, pour toute mesure d'entrée $\mu \in \mathscr{M}_1(E)$, on a*

$$\mathbf{P}_\mu\text{-}p.s. \quad \frac{1}{t}\int_0^t f(X_s)\,ds \ \xrightarrow[t \to \infty]{}\ \langle m, f\rangle\ .$$

## II.3 Quelques calculs sur le générateur.

Dans cette section on fixe

$$\big(\Omega\,;\,\mathscr{F}\,;\,(\mathscr{F}_t)_{t\in\mathbb{R}_+}\,;\,\mathbf{X}=(X_t)_{t\in\mathbb{R}_+}\,;\,G\,;\,\mathbf{P}_\mu\,,\,\mu\in\mathscr{M}_1(E)\big)\,,$$

un processus markovien minimal. On se donne $G=(q_{i,j})_{i,j\in E}$. On rappelle que le processus markovien qui correspond à cette notation est un processus markovien à valeurs dans l'espace dénombrable discret $E^*=E\cup\{\partial\}$ qui est absorbé au point cimetière $\partial$. Cette section est consacrée à quelques opérations sur les processus markoviens qui préservent le caractère markovien (mais pas forcément l'espace d'état).

### II.3.a Processus stoppés et changés de temps.

**Processus stoppé.** On fixe $K\subset E$, *qui n'est pas tout* $E$ *et qui n'est pas vide*. On pose

$$T_K=\inf\{t\in\mathbb{R}_+ : X_t\in K\}\,,$$

avec la convention habituelle $\inf\emptyset=\infty$. On pose

$$X'_t=X_{t\wedge T_K}\,,\quad t\in\mathbb{R}_+\,.$$

On voit que $\mathbf{X}'=(X'_t)_{t\in\mathbb{R}_+}$ est la trajectoire $\mathbf{X}$ *stoppée* (ou *absorbée*) dès qu'elle atteint $K$. Il et assez facile de voir que $\mathbf{X}'$ est un processus markovien minimal, continu à droite, qui possède les mêmes transitions que $\mathbf{X}$ hors de $K$ mais où tous les états de $K$ sont absorbants. Son générateur infinitésimal $G'=(q'_{i,j})_{i,j\in E}$ est donné par

$$\forall i,j\in E\,,\quad q'_{i,j}=q_{i,j}\ \text{ si } i\notin K\quad \text{et}\quad q'_{i,j}=0\ \text{ si } i\in K\,.$$

**Preuve :** exercice. ■

**Changement de temps.** On fixe une fonction $v: E\to]0,\infty[$ que l'on prolonge en $\partial$ en posant $v(\partial)=0$. Pour tout $t\in\mathbb{R}_+$, on pose alors

$$A_t(\mathbf{X})=\int_0^t v(X_s)\,ds\quad\text{et}\quad \tau_t(\mathbf{X})=\inf\big\{s\in\mathbb{R}_+\,:\,A_s(\mathbf{X})>t\,\big\}\,,$$

avec la convention habituelle $\inf\emptyset=\infty$. Pour simplifier les notations, on pose $A_t:=A_t(\mathbf{X})$ et $\tau_t:=\tau_t(\mathbf{X})$. On définit un nouveau processus $\mathbf{X}'=(X'_t)_{t\in\mathbb{R}_+}$, appelé processus changé de temps en posant

$$X'_t=\begin{cases} X_{\tau_t} & \text{si } \tau_t<\infty,\\ \partial & \text{si } \tau_t=\infty.\end{cases}$$

On remarque que $t\mapsto A_t$ est continu strictement croissant sur $[0,\zeta[$ et que $A_t:\Omega\to\mathbb{R}_+$ est $\mathscr{F}_t$-mesurable (on peut voir cela en montrant par exemple que $A_t=\lim_{n\to\infty}\frac{1}{n}\sum_{0\le k<n}v(X_{kt/n})$, partout sur $\Omega$). Autrement dit $t\in[0,\zeta[\mapsto A_t\in[0,A_\infty[$ est une bijection continue strictement croissante de $[0,\zeta[$ sur $[0,A_\infty[$. On observe ensuite que $A_\zeta=A_\infty$, c'est-à-dire que $A_t$ est constant sur $[\zeta,\infty[$.

On voit également que $\tau_t<\infty$ ssi $t<A_\infty$. De plus, $t\mapsto\tau_t$ est continu strictement croissant sur $[0,A_\infty[$. On observe que $\{\tau_t\le s\}=\{A_s\ge t\}$, ce qui implique que $\tau_t$ est un $(\mathscr{F}_s)_{s\in\mathbb{R}_+}$-temps d'arrêt. Par ailleurs, on voit facilement que $\mathbf{1}_{\{\tau_t<\infty\}}A_{s+\tau_t}=\mathbf{1}_{\{\tau_t<\infty\}}\big(t+A_s(\theta_{\tau_t}\mathbf{X})\big)$. On en déduit que

$$\tau_{t+s}=\tau_t+\mathbf{1}_{\{\tau_t<\infty\}}\tau_s(\theta_{\tau_t}\mathbf{X})\,. \tag{II.53}$$

On observe que puisque $t \mapsto \tau_t$ est continu strictement croissant sur $[0, A_\infty[$ , $\mathbf{X}'$ est simplement la même trajectoire que $\mathbf{X}$ parcourue à une vitesse différente : c'est donc un processus continu à droite, $\zeta(\mathbf{X}') = A_\zeta = A_\infty$, et $X'_t = \partial$, pour tout $t \in [\zeta(\mathbf{X}'), \infty[$ . De plus $X'_t \in E$, pour tout $t < \zeta(\mathbf{X}')$. On fixe $t, s \in \mathbb{R}_+$, $j \in E$ et $\mu \in \mathscr{M}_1(E)$. Sous $\mathbf{P}_\mu$, si $X'_{t+s} = j \in E$, alors $t + s < A_\zeta$ et $\tau_t < \infty$. On applique la propriété de Markov au temps $\tau_t$ et (II.53) implique que

$$\mathbf{P}_\mu\text{-p.s.} \quad \mathbf{E}_\mu\big[\mathbf{1}_{\{X'_{t+s}=j\}} \,\big|\, \mathscr{F}_{\tau_t}\big] = \mathbf{1}_{\{\tau_t<\infty\}}\mathbf{P}_{X_{\tau_t}}(X_{\tau_s} = j) = \mathbf{1}_{\{\tau_t<\infty\}}\mathbf{P}_{X'_t}(X'_s = j) \ .$$

Cela montre que $\mathbf{X}'$ est un processus de Markov minimal relativement à la filtration $(\mathscr{F}_{\tau_t})_{t\in\mathbb{R}_+}$. Calculons son générateur infinitésimal : comme la chaîne du squelette de $\mathbf{X}'$ est la même que celle de $\mathbf{X}$, il suffit de calculer ses paramètres des temps d'attente. On fixe $i \in E$ et on observe que

$$\mathbf{P}_i\text{-p.s.} \quad J_1(\mathbf{X}') = A_{J_1(\mathbf{X})} = v(i)J_1(\mathbf{X}) \ .$$

Par conséquent

$$\mathbf{P}_i(J_1(\mathbf{X}') > t) = \exp\big(-\tfrac{q(i)}{v(i)}t\big) \ .$$

On a donc montré la proposition suivante.

**Proposition II.3.1** *Le processus $\mathbf{X}'$ est un processus de Markov minimal relativement à la filtration $(\mathscr{F}_{\tau_t})_{t\in\mathbb{R}_+}$ et dont le générateur infinitésimal $G' = (q'_{i,j})_{i,j\in E}$ est donné par*

$$\forall i, j \in E\,, \quad q'_{i,j} = q_{i,j}/v(i) \ .$$

**Remarque II.3.2** On peut autoriser $v$ à prendre des valeurs infinies sans modifier le résultat précédent : la seule chose à remarquer est que les états $i \in E$ pour lesquels $v(i) = \infty$ sont des états absorbants pour le processus changé de temps $\mathbf{X}'$. Il est possible d'autoriser $v$ à prendre des valeurs nulles mais le résultat précédent doit être modifié : les états $i \in E$ tels que $v(i) = 0$ sont instantanément "sautés" par le processus changé de temps ; pour rester dans le cadre des processus standard, on est alors conduit à restreindre l'espace d'états pour $\mathbf{X}'$ a $E' = \{i \in E : v(i) > 0\}$. De plus, on ne peut plus exclure que $\mathbf{X}'$ passe d'un éléments de $E'$ à $\partial$ sans effecteur une infinité de sauts ; il n'est donc plus un processus minimal à strictement parler. Tout ceci ne constitue pas un obstacle fondamental mais nous n'irons pas plus loin dans cette direction. $\square$

## II.3.b La formule de Dynkin.

Les équations de Kolmogorov-Chapman (ou bien un argument simple basé sur la propriété de Markov au temps $J_1$) montrent que pour tous $i, j \in E$, il existe une fonction continue $\varepsilon_{i,j} : \mathbb{R}_+ \to \mathbb{R}$ telle que

$$\forall i, j \in E \quad \lim_{t\to 0} \varepsilon_{i,j}(t) = 0 \quad \text{et} \quad p_t(i,j) = \begin{cases} tq_{i,j} + t\varepsilon_{i,j}(t) & \text{si } i \neq j, \\ 1 - tq(i) + t\varepsilon_{i,i}(t) & \text{si } i \neq j. \end{cases}$$

La difficulté principale est que *les fonctions $\varepsilon_{i,j}$ peuvent ne pas tendre vers* 0 *uniformément en i et j, c'est-à-dire que* $\sup_{i,j} \varepsilon_{i,j}$ *peut ne pas tendre vers* 0, ce qui occasionne de nombreuses difficultés techniques.

**Heuristique.** Malgré cela écrivons quelques calculs heuristiques. Ces calculs sont faux en général bien que les idées formelles qui sont à leur origine soient fondées et dans certains cas rigoureusement justifiables.

Soit $f : E \to \mathbb{R}$, une fonction. On étend $f$ à $E^*$ en posant $f(\partial) = 0$. La fonction $P_t.f : E \to \mathbb{R}$ s'interprète de manière probabiliste de la façon suivante :

$$\forall i \in E\,, \quad P_t.f(i) = \sum_{j\in E} p_t(i,j)f(j) = \sum_{j\in E} f(j)\mathbf{P}_i(X_t = j\,;\zeta > t) = \mathbf{E}_i\big[f(X_t)\mathbf{1}_{\{\zeta>t\}}\big].$$

Mais comme $f(\partial) = 0$ et $X_t \in E$ ssi $t < \zeta$, on a

$$\forall i \in E\,, \quad P_t f(i) = \mathbf{E}_i\big[f(X_t)\big]\ .$$

On rappelle que $(P_t)_{t\in\mathbb{R}_+}$ est la solution (minimale) de l'équation de Kolmogorov-Chapman directe, ce qui « entraîne formellement » (c'est le point délicat à justifier quand cela est possible, et ce n'est pas toujours vrai) que

$$P_t.f(i) - P_0.f(i) = \int_0^t ds\, \tfrac{d}{ds}P_s.f(i) = \int_0^t ds\,(P_s.(G.f))(i). \tag{II.54}$$

On rappelle que le processus minimal prend ses valeurs dans $E^*$, et que $\partial$ est absorbant. Par conséquent, on a $q_{\partial,\partial} = q_{\partial,i} = 0$ et donc $G.f(\partial) = 0$. On a donc

$$(P_s.(G.f))(i) = \mathbf{E}_i\big[G.f(X_s)\big]\ .$$

Comme $P_0 = \mathrm{Id}_E$, on a $P_0.f = f$. On en déduit

$$\forall i \in E\,, \quad \mathbf{E}_i\big[f(X_t)\big] = f(i) + \mathbf{E}_i\Big[\int_0^t G.f(X_s)\,ds\Big]\ . \tag{II.55}$$

On pose alors

$$\forall t \in \mathbb{R}_+\,, \quad M_t^f = f(X_t) - f(X_0) - \int_0^t Gf(X_s)\,ds\ .$$

et par (II.55), $(M_t^f)_{t\in\mathbb{R}_+}$ est une $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$-martingale. En appliquant un théorème d'arrêt, on obtient alors $\mathbf{E}_i[M_T^f] = \mathbf{E}_i[M_0^f] = f(i)$ c'est-à-dire

$$\forall i \in E\,, \quad \mathbf{E}_i\big[f(X_T)\big] = f(i) + \mathbf{E}_i\Big[\int_0^T Gf(X_s)\,ds\Big]\ . \tag{II.56}$$

qui est la *formule de Dynkin.*

L'étape qu'il semble difficile à justifier directement est (II.54). La seconde étape est le théorème d'arrêt : une hypothèse d'intégrabilité de $T$ sous $\mathbf{P}_i$ est nécessaire.

**Un énoncé rigoureux.** On introduit l'espace $C_c(E)$ des fonctions dont le support est fini.

$$C_c(E) = \big\{f : E \to \mathbb{R}\ :\ \#\{i \in E : f(i) \neq 0\} < \infty\,\big\}\ .$$

Il est clair que c'est un espace vectoriel dont une base est fournie par les fonctions $\mathbf{1}_{\{j\}}$, $j \in E$.

**Proposition II.3.3** *Soit $f \in C_c(E)$. Les assertions suivantes sont vérifiées.*

**(i)** *Il existe $a, b \in\, ]0,\infty[$, qui ne dépendent que de $f$ et de $G$, telles que $\sup_{i\in E}\int_0^t \mathbf{E}_i[\,|Gf|(X_s)]\,ds\ \le a{+}b\,t$. De plus, la formule de Dynkin (II.55) est vérifiée. pour tout $i \in E$ et tout $t \in \mathbb{R}_+$.*

**(ii)** *Pour toute fonction $f \in C_c(E)$, on pose*

$$\forall t \in \mathbb{R}_+ , \quad M_t^f = f(X_t) - f(X_0) - \int_0^t Gf(X_s)\, ds \, .$$

*Il existe $c, d \in ]0, \infty[$ qui ne dépendent que de $f$ et de $G$, telles que* $\sup_{\mu \in \mathscr{M}_1(E)} \mathbf{E}_\mu\big[\, |M_t^f|\,\big] \leq c + d\,t$. *De plus pour tout $t, s \in \mathbb{R}_+$, on a*

$$\mathbf{P}_\mu\text{-}p.s. \quad \mathbf{E}_\mu\big[M_{t+s}^f \,|\mathscr{F}_t\,\big] = M_t^f \ ,$$

*ce qui implique que* $\mathbf{E}_\mu[f(X_t)] = \langle \mu, f\rangle + \mathbf{E}_\mu\Big[\int_0^t Gf(X_s)\, ds\Big]$.

**(iii)** *Soit $\mu \in \mathscr{M}_1(E)$ et $T$, un $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$-temps d'arrêt $T$ tel que $\mathbf{E}_\mu[T] < \infty$. Alors $\mathbf{E}_\mu[\int_0^T |Gf|(X_s)ds] < \infty$ et*

$$\mathbf{E}_\mu\big[f(X_T)\big] = \langle \mu, f\rangle + \mathbf{E}_\mu\Big[\int_0^T Gf(X_s)\, ds\Big] \, . \tag{II.57}$$

**Preuve :** on fixe $j_0 \in E$ et on commence par prouver les choses pour $f = \mathbf{1}_{\{j_0\}}$. L'équation de Kolmogorov-Chapman directe implique pour tout $s \in \mathbb{R}_+$ que

$$p_s'(i, j_0) = -q(j_0)p_s(i, j_0) + \sum_{k\in E\backslash\{j_0\}} p_s(i,k)q_{k,j_0} \ .$$

Il est clair que $G.\mathbf{1}_{\{j_0\}}(k) = q_{k,j_0}$. On en déduit dont que

$$p_s'(i, j_0) = p_s(i, j_0)G.\mathbf{1}_{\{j_0\}}(j_0) + \sum_{k\in E\backslash\{j_0\}} p_s(i,k)G.\mathbf{1}_{\{j_0\}}(k) \ .$$

On observe que $\int_0^t p_s'(i, j_0)ds$ a un sens et vaut $p_t(i, j_0) - p_0(i, j_0) = p_t(i, j_0) - \mathbf{1}_{\{j_0)}(i)$. D'autre part,

$$\int_0^t p_s(i, j_0)\big|G.\mathbf{1}_{\{j_0\}}(j_0)\big|\, ds = q(j_0)\int_0^t p_s(i, j_0) ds \leq q(j_0)t \ .$$

On en déduit que $\int_0^t ds\, \sum_{k\in E\backslash\{j_0\}} p_s(i,k)G.\mathbf{1}_{\{j_0\}}(k)$ est bien défini et que

$$\begin{aligned}
\int_0^t ds \sum_{k\in E\backslash\{j_0\}} p_s(i,k)\big|G.\mathbf{1}_{\{j_0\}}(k)\big| \ . \quad &= \quad \int_0^t ds \sum_{k\in E\backslash\{j_0\}} p_s(i,k)G.\mathbf{1}_{\{j_0\}}(k) \\
&= \quad p_t(i, j_0) - \mathbf{1}_{\{j_0)}(i) + q(j_0)\int_0^t p_s(i, j_0)\, ds \\
&\leq \quad 1 + q(j_0)t \ .
\end{aligned}$$

Par conséquent,

$$p_t(i, j_0) - \mathbf{1}_{\{j_0)}(i) = \int_0^t ds \sum_{k\in E} p_s(i,k)G.\mathbf{1}_{\{j_0\}}(k) = \int_0^t ds\, \mathbf{E}_i[G.\mathbf{1}_{\{j_0\}}(X_s)] \ ,$$

et que

$$\int_0^t ds\, \mathbf{E}_i\big[\,\big|G.\mathbf{1}_{\{j_0\}}(X_s)\big|\,\big] \quad \leq \quad \int_0^t ds \sum_{k\in E\backslash\{j_0\}} p_s(i,k)\big|G.\mathbf{1}_{\{j_0\}}(k)\big| + q(j_0)\int_0^t p_s(i, j_0) ds$$

$$\leq \ 1 + 2q(j_0)t \ < \ \infty \ . \tag{II.58}$$

Par Fubini, on en déduit

$$p_t(i, j_0) = \mathbf{E}_i\big[\mathbf{1}_{\{j_0\}}(X_t)\big] = \mathbf{1}_{\{j_0\}}(i) + \mathbf{E}_i\Big[\int_0^t G.\mathbf{1}_{\{j_0\}}(X_s)\, ds\Big] \ .$$

Soit $f \in C_c(E)$. On observe que $f = \sum_{j_0 \in E} f(j_0)\mathbf{1}_{\{j_0\}}$. Comme $G$ est linéaire, on déduit de (II.58) que

$$\sup_{i \in E} \int_0^t \mathbf{E}_i[\,|G.f|(X_s)]\, ds \leq a + t\, b \tag{II.59}$$

où on a posé

$$a = \sum_{j_0 \in E} |f(j_0)| \quad \text{et} \quad b = 2\sum_{j_0 \in E} q(j_0)|f(j_0)| \ .$$

qui sont deux quantités finies car $f \in C_c(E)$. De plus $f$ vérifie la formule de Dynkin (II.55). Cela montre le premier point de la proposition.

On observe tout d'abord que pour toute loi d'entrée $\mu \in \mathscr{M}_1(E)$, la propriété de Markov au temps $0$ implique que

$$\mathbf{P}_\mu\text{-p.s.} \quad \mathbf{E}_\mu\Big[\int_0^t |G.f|(X_s)\, ds \,\big|\, \mathscr{F}_0\Big] = \mathbf{E}_{X_0}\Big[\int_0^t |G.f|(X_s)\, ds\Big] \leq a + b\,.t$$

En intégrant on a donc $\mathbf{E}_\mu[\int_0^t |G.f|(X_s)\, ds\,] \leq a + b\,t$. On observe ensuite que

$$\big|M_t^f\big| \leq \big|f(X_t)\big| + \big|f(X_0)\big| + \int_0^t |G.f|(X_s)\, ds \ .$$

On pose $\|f\|_\infty = \sup_{i \in E} |f(i)|$ qui est manifestement une quantité finie car $f \in C_c(E)$. Ce qui précède implique que pour toute loi d'entrée $\mu \in \mathscr{M}_1(E)$, on a

$$\mathbf{E}_\mu\big[\,|M_t^f|\,\big] \leq 2\|f\|_\infty + a + b\,t \ ,$$

ce qui implique le premier résultat du **(ii)** en posant $c = 2\|f\|_\infty + a$ et $d = b$. On remarque ensuite que

$$M_{t+s}^f - M_t^f = f((\theta_t \mathbf{X})_s) - f((\theta_t \mathbf{X})_0) + \int_0^s G.f((\theta_t \mathbf{X})_r)\, dr \ .$$

La propriété de Markov implique et la formule de Dynkin (II.55) au temps $s$ pour tout $i \in E$, qui est démontrée au **(i)**, implique que $\mathbf{E}_\mu\big[M_{t+s}^f - M_t^f \,\big|\, \mathscr{F}_t\big] = 0$. Or on remarque que $M_t^f$ est $\mathscr{F}_t$-mesurable (ce détail est laissé au lecteur). On a donc prouvé le point **(ii)**.

On note $\lceil \cdot \rceil$ la fonction $\lfloor \cdot \rfloor + 1$, où $\lfloor \cdot \rfloor$, désigne la partie entière. On observe que

$$\int_0^T |G.f|(X_s)ds \leq \int_0^{\lceil T \rceil} |G.f|(X_s)ds = \sum_{n \geq 1} \mathbf{1}_{\{T \geq n\}} \int_n^{n+1} |G.f|(X_s)ds \ .$$

La propriété de Markov au temps $n$, combinée avec la borne du point **(i)**, implique que

$$\mathbf{P}_\mu\text{-p.s.} \quad \mathbf{E}_\mu\Big[\mathbf{1}_{\{T \geq n\}} \int_n^{n+1} |G.f|(X_s)ds \,\big|\, \mathscr{F}_n\Big] \ = \ \mathbf{1}_{\{T \geq n\}} \mathbf{E}_\mu\Big[\int_n^{n+1} |G.f|(X_s)ds \,\big|\, \mathscr{F}_n\Big]$$

$$
\begin{aligned}
&= \mathbf{1}_{\{T\geq n\}}\mathbf{E}_{X_n}\Big[\int_0^1 |G.f|(X_s)\,ds\Big] \\
&\leq (a+b)\mathbf{1}_{\{T\geq n\}} .
\end{aligned}
$$

On en déduit donc

$$
\begin{aligned}
\mathbf{E}_\mu\Big[\int_0^{\lceil T\rceil} |G.f|(X_s)ds\Big] &\leq (a+b)\sum_{n\geq 0}\mathbf{P}_\mu(T\geq n) \leq (a+b)\mathbf{E}_\mu\big[\lceil T\rceil\big] \\
&\leq (a+b)(1+\mathbf{E}_\mu[T]) \;<\; \infty . \qquad \text{(II.60)}
\end{aligned}
$$

Cela montre le premier résultat du point **(iii)**. On en déduit d'une part que

$$
|M^f_T| \leq U := 2\|f\|_\infty + \int_0^{\lceil T\rceil} |G.f|(X_s)ds ,
$$

et $U$ est intégrable sous $\mathbf{P}_\mu$. On pose ensuite $T_n = 2^{-n}\lceil 2^n T\rceil$ et on observe que

$$
|M^f_{T_n}| \leq U \quad \text{et} \quad M^f_{T_n} = \sum_{k\geq 0}\mathbf{1}_{\{T\geq k2^{-n}\}}\big(M^f_{(k+1)2^{-n}} - M^f_{k2^{-n}}\big) .
$$

En utilisant le point **(ii)** on voit que $\mathbf{P}_\mu$-p.s.

$$
\mathbf{E}_\mu\big[\mathbf{1}_{\{T\geq k2^{-n}\}}\big(M^f_{(k+1)2^{-n}} - M^f_{k2^{-n}}\big)\,\big|\,\mathscr{F}_{k2^{-n}}\big] = \mathbf{1}_{\{T\geq k2^{-n}\}}\mathbf{E}_\mu\big[\big(M^f_{(k+1)2^{-n}} - M^f_{k2^{-n}}\big)\,\big|\,\mathscr{F}_{k2^{-n}}\big] = 0 .
$$

On en déduit que $\mathbf{E}_\mu[M^f_{T_n}] = 0$. Par continuité à droite de $\mathbf{X}$, on voit que $\lim_n M^f_{T_n} = M^f_T$, $\mathbf{P}_\mu$-presque sûrement. Comme $|M^f_{T_n}| \leq U$, pour tout $n\geq 0$, et puisque $U$ est intégrable sous $\mathbf{P}_\mu$, le théorème de convergence dominée implique que $\mathbf{E}_\mu[M^f_T] = 0$, ce qui implique (II.57). ■

### II.3.c Compléments : le cas des générateurs uniformes.

En faisant une hypothèse (forte) sur le générateur nous allons étendre les résultats de la proposition II.3.3 aux fonctions bornées : pour toute fonction $f : E\to\mathbb{R}$, on pose

$$
\|f\|_\infty = \sup_{i\in E}|f(i)| \quad \text{et} \quad C_b(E) = \big\{ f : E\to\mathbb{R} \,:\, \|f\|_\infty < \infty \big\} .
$$

Il est facile de vérifier que $(C_b(E), \|\cdot\|_\infty)$ est un espace vectoriel normé complet (un espace de Banach). Notre premier but est d'étendre les résultats obtenus à la section II.1.b pour les espace d'états finis. Pour cela on introduit l'ensemble $M_E(\mathbb{R})$ des « matrices » $B = (b(i,j))_{i,j\in E}$ sur $E$ à coefficients réels telles que

$$
|B| := \sup_{i\in E}\sum_{j\in E}|b(i,j)| \;<\; \infty .
$$

Si $A = (a(i,j))_{i,j\in E}$ et $B = (b(i,j))_{i,j\in E}$ sont des éléments de $M_E(\mathbb{R})$, pour tout réel $\lambda\in\mathbb{R}$, on pose $C = A+\lambda.B = (c(i,j))_{i,j\in E}$ où $c(i,j) = a(i,j)+\lambda b(i,j)$. Il est facile de montrer que

$$
|\lambda.B| = |\lambda|\,|B| \quad \text{et} \quad |A+\lambda.B| \leq |A| + |\lambda|\,|B| .
$$

De plus on observe que

$$
\forall i\in E , \quad \sum_{j\in E}\sum_{k\in E}|a(i,k)|\,|b(k,j)| \leq |B|\sum_{k\in E}|a(i,k)| \leq |A|\,|B| .
$$

Cela permet de définir $D := A.B = (d(i,j))_{i,j\in E}$ avec

$$\forall i,j \in E\,,\quad d(i,j) = \sum_{k\in E} a(i,k)b(k,j)\ .$$

On a

$$\forall A,B \in \mathbb{M}_E(\mathbb{R})\,,\quad A.B \in M_E(\mathbb{R}) \quad \text{et} \quad |A.B| \le |A|\,|B|\ .$$

On vérifie ensuite que le produit matriciel est associatif et que $(M_E(\mathbb{R})\,, +\,, \cdot\,)$ est une algèbre (non-commutative). De plus, on constate que $|\cdot|$ est une norme d'algèbre sur $M_E(\mathbb{R})$ et il facile de montrer que $(M_E(\mathbb{R})\,, |\cdot|\,)$ est complet. Autrement dit $M_E(\mathbb{R})$ munie de cette norme est une algèbre de Banach : toute suite de Cauchy y est convergente. On utilise la notation $[B](i,j) = b(i,j)$, pour l'entrée $(i,j)$ de la matrice $B$ : il est évident que pour tous $(i,j) \in E$ la fonction $B \in M_E(\mathbb{R}) \mapsto [B](i,j) \in \mathbb{R}$ est linéaire continue : c'est une forme linéaire continue. Soit $f \in C_b(E)$. On définit $B.f : E \to \mathbb{R}$ par

$$\forall i \in E\,,\quad B.f(i) = \sum_{j\in E} [B](i,j)f(j)\ .$$

Il est facile de vérifier que

$$\forall f \in C_b(E)\,,\quad \|B.f\|_\infty \le |B|\,\|f\,\|_\infty\ .$$

De plus $f \in C_b(E) \mapsto B.f \in C_b(E)$ est linéaire. $B$ peut donc se voir comme une application linéaire de $C_b(E)$ dans lui-même qui est continu (borné en norme). De plus

$$\forall A,B \in M_E(\mathbb{R})\,,\ \forall f,g \in C_b(E)\,,\quad \|A.f - B.g\|_\infty \le |A|\,\|f-g\|_\infty + |A-B|\,\|g\|_\infty\ .$$

**Exponentielle.** Pour tout $B \in M_E(\mathbb{R})$, et pour tout $n \in \mathbb{N}$, on pose $S_n = \sum_{0\le p\le n} \frac{1}{p!}B^p$, avec la convention que $B^0 = \mathrm{Id}_E$, qui est la matrice identité sur $E$. On observe alors que

$$\sup_{n_1,n_2\ge n} \big|S_{n_1} - S_{n_2}\big| \le \sum_{p\ge n} \tfrac{1}{p!}|B|^p \xrightarrow[n\to\infty]{} 0\ .$$

Cela montre que la suite $(S_n)_{n\ge 0}$ est de Cauchy dans $(M_E(\mathbb{R}), |\cdot|\,)$ qui est complet. Elle converge donc vers une limite qui est notée $\exp(B)$ et on écrit

$$\exp(B) = \sum_{n\ge 0} \tfrac{1}{n!}B^n\ .$$

Il est facile de vérifier (nous n'en aurons pas vraiment besoin) que

$$\exp(A+B) = \exp(A).\exp(B) \quad \text{si} \quad A.B = B.A\ .$$

Soit $t \in \mathbb{R}+ \mapsto P_t \in M_E(\mathbb{R})$, une fonction. On dit que cette fonction est dérivable en $t$ de dérivée $P'_t$ ssi

$$\lim_{h\to 0} \big|\tfrac{1}{h}(P_{t+h} - P_t) - P'_t\big| = 0\ .$$

**Lemme II.3.4** *Soient $A, B \in M_E(\mathbb{R})$, deux matrices. On considère l'équation différentielle suivante :*

$$\forall t \in \mathbb{R}_+\,,\quad P'_t = B.P_t \quad \text{et} \quad P_0 = A\,, \tag{II.61}$$

*(où la fonction $t \mapsto P_t$ est supposée a priori continue et dérivable partout). Cette équation a une unique solution donnée par $P_t = \exp(t.B).A$, $t \in \mathbb{R}_+$.*

*De même, l'équation différentielle*

$$\forall t \in \mathbb{R}_+\,,\quad P'_t = P_t.B \quad \text{et} \quad P_0 = A\,, \tag{II.62}$$

*a une unique solution donnée par $P_t = A.\exp(t.B)$, $t \in \mathbb{R}_+$.*

**Preuve :** on fixe $f \in C_b(E)$ et $i \in E$. On suppose de plus que $t \mapsto P_t$ est continue, dérivable partout et satisfait (II.61). On observe que

$$\left|\tfrac{1}{h}(P_{t+h}.f(i) - P_t.f(i)) - B.P_t.f(i)\right| \leq \left|\tfrac{1}{h}(P_{t+h} - P_t) - P'_t\right| \|f\|_\infty \xrightarrow[h\to 0]{} 0.$$

Cela montre que $t \mapsto P_t.f(i)$ est une fonction réelle, bornée continue et dérivable de dérivée $t \mapsto B.P_t.f(i)$, elle-même bornée continue. On a donc

$$P_t.f(i) - A.f(i) = \int_0^t B.P_s.f(i)\, ds\ .$$

En itérant cette propriété on montre facilement par récurrence que

$$P_t.f(i) = A.f(i) + t.B.A.f(i) + \ldots + \frac{t^n}{n!}.B^n.A.f(i) + R_n(i,t)$$

où on a posé

$$R_n(i,t) = \int_{\{0\leq s_n \leq s_{n-1}\leq \ldots \leq s_0 \leq t\}} ds_0 \ldots ds_n \left(B^{n+1}.P_{s_n}.f(i) - B^n.A.f(i)\right)\ .$$

On pose $c_t = \sup_{s\in[0,t]} |P_s|$. On montre alors que

$$|R_n(i,t)| \leq \big(c_t|B|+1\big).|B|^n|A|.\|f\|_\infty \int_{\{0\leq s_n \leq s_{n-1}\leq \ldots \leq s_0 \leq t\}} ds_0 \ldots ds_n = \frac{t^{n+1}}{n!}\big(c_t|B|+1\big)|B|^n.|A|.\|f\|_\infty.$$

On a donc $\lim_n R_n(i,t) = 0$ et on en déduit la convergence suivante :

$$P_t.f(i) = \sum_{n\geq 0} \frac{t^n}{n!}.B^n.A.f(i) = \exp(t.B).A.f(i)$$

Cela entraîne que $P_t = \exp(t.B).A$, $t \in \mathbb{R}_+$. Cela montre l'unicité de la solution de (II.61). Il est ensuite facile de vérifier que $t \mapsto \exp(t.B).A$ est une solution, ce qui montre le premier point du lemme. On raisonne de la même façon pour montrer le second point. ■

**Définition II.3.5** On dit qu'un générateur infinitésimal $G$ est *uniforme* ssi $\sup_{i\in E}\, q(i) := C < \infty$. □

Le petit travail préparatoire effectué au lemme II.3.4 porte immédiatement ses fruits.

**Théorème II.3.6** *Soit*

$$\big(\Omega\,;\, \mathscr{F}\,;\, (\mathscr{F}_t)_{t\in\mathbb{R}_+}\,;\, \mathbf{X} = (X_t)_{t\in\mathbb{R}_+}\,;\, G\,;\, \mathbf{P}_\mu\,,\, \mu \in \mathscr{M}_1(E)\big)\,,$$

*un processus markovien minimal. On suppose que G un générateur infinitésimal uniforme. Alors les assertions suivantes sont vérifiées*

**(i)** $G \in M_E(\mathbb{R})$ *:* $|G| \leq 2C$*, où* $C = \sup_{i\in E} q(i)$*. Le semi-groupe* $(P_t)_{t\in\mathbb{R}_+}$ *associé à G est l'unique solution des équations de Kolmogorov-Chapman directe et rétrograde. On a de plus*

$$\lim_{t\to 0} \left|P_t - \mathrm{Id}_E\right| = 0 \quad \text{et} \quad \lim_{t\to 0}\left|\tfrac{1}{t}(P_t - \mathrm{Id}_E) - G\right| = 0\ .$$

*Plus généralement, on a* $P_t = \exp(tG)$*, c'est-à-dire*

$$\forall t \in \mathbb{R}_+ \quad \lim_n \Big|P_t - \sum_{n\geq 0} \tfrac{t^n}{n!}.G^n\Big| = 0\ .$$

*Cela entraîne donc que*

$$\forall f \in C_b(E)\,,\ \forall \mu \in \mathscr{M}_1(E)\,,\ \forall t \in \mathbb{R}_+\,,\quad \mathbf{E}_\mu[f(X_t)] = \langle \mu, f\rangle + \mathbf{E}_\mu\Big[\int_0^t G.f(X_s)\, ds\Big]\ .$$

**(ii)** *Il n'y a pas d'explosion : pour toute loi d'entrée $\mu \in \mathscr{M}_1(E)$, $\mathbf{P}_\mu(\zeta = \infty) = 1$.*

**(iii)** *Pour toute loi d'entrée $\mu \in \mathscr{M}_1(E)$, pour tout $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$-temps d'arrêt tel que $\mathbf{E}_\mu[T] < \infty$ et pour toute fonction bornée $f \in C_b(E)$,*

$$\mathbf{E}_\mu\Big[\int_0^T |G.f|(X_s)\,ds\Big] \leq 2C\|f\|_\infty \mathbf{E}_\mu[T] \quad \text{et} \quad \mathbf{E}_\mu[f(X_T)] = \langle \mu, f\rangle + \mathbf{E}_\mu\Big[\int_0^T G.f(X_s)\,ds\Big] .$$

**Preuve :** si on pose $G = (q_{i,j})_{i,j\in E}$, on observe que $\sum_{j\in E} |q_{i,j}| = 2q(i)$, ce qui implique que $|G| = 2C < \infty$ et donc $G \in M_E(\mathbb{R}_+)$. Le reste du premier point est une conséquence du lemme II.3.4 et du calcul formel effectué au début de la sous-section, qui est ici parfaitement justifié. Le point **(ii)** découle de la formule de Dynkin : on fixe $\mu \in \mathscr{M}_1(E)$ et on observe que

$$\mathbf{P}_\mu(X_t \in E) = \langle \mu, \mathbf{1}_E\rangle + \mathbf{E}_\mu\Big[\int_0^t G.\mathbf{1}_E(X_s)\,ds\Big] .$$

Or on remarque que $G.\mathbf{1}_E \equiv 0$. Donc $\mathbf{P}_\mu(X_t \in E) = 1$, pour tout $t \in \mathbb{R}_+$. On en déduit qu'il n'y a pas d'explosion. Le dernier point se déduit de la formule de Dynkin à temps fixé $t$ par les mêmes arguments que ceux utilisés dans la preuve de la proposition II.3.3. ■

Beaucoup de générateurs uniformes sont de la forme suivante : on se donne une matrice de transition $Q = (\pi_{i,j})_{i,j\in E}$ qui peut être la matrice de transition de la chaîne du squelette d'un processus markovien ; cela est équivalent à supposer que

$$\forall i \in E\,, \quad \text{ou bien } \pi_{i,i} = 0, \text{ ou bien } \pi_{i,i} = 1. \tag{II.63}$$

On fixe $c \in ]0, \infty[$ et on pose

$$G = c.\big(Q - \mathrm{Id}_E\big) . \tag{II.64}$$

On voit que $G$ est un générateur infinitésimal uniforme : tous les paramètres d'attente sont égaux à $c$ et la chaîne du squelette associée à pour matrice de transition $Q$. Par ailleurs, il est facile de voir que $\exp(-ct.\mathrm{Id}_E) = e^{ct}.\mathrm{Id}_E$, par conséquent, le semi-groupe associé à $G$ de la forme (II.64) est donné par

$$\forall t \in \mathbb{R}_+\,, \quad P_t = e^{-ct}.\exp(ct.Q) = \sum_{n\geq 0} e^{-ct}\frac{(ct)^n}{n!}.Q^n . \tag{II.65}$$

On voit que les générateurs uniformes sont une extension simple du cas des processus de Markov à espace d'états fini. Cependant, l'hypothèse d'uniformité est très forte et de nombreux processus markoviens intéressants et naturels *ne sont pas uniformes*. Il n'y a malheureusement pas d'hypothèse élémentaire sur le générateur qui englobe tous les cas intéressant ou utiles et qui permette d'utiliser simplement le générateur pour effectuer des calculs.

## II.4 Exemples.

### II.4.a Processus de Poisson.

On imagine la situation suivante : un compteur compte combien d'électrons frappent une plaque de métal au cours du temps. On commence le comptage au temps $t = 0$ où $N_0$ électrons ont déjà été enregistrés. On note $N_t$ le nombre d'électrons détecté au temps $t$. On remarque que $t \mapsto N_t$ est continu à droite. De plus on suppose qu'il est quasiment impossible que deux électrons arrivent en même temps sur la plaque. Par conséquent, les sauts du processus de comptage sont égaux à 1 : si on note $N_{t-}$ la limite à gauche du processus de comptage à l'instant $t$, alors $N_t - N_{t-} \in \{0, 1\}$, pour tout $t \in \mathbb{R}_+$. On note $T_n$ le temps d'arrivée du $n$-ième électron (à partir du temps 0). On a donc

$$N_t = N_0 + \sum_{n\geq 1} \mathbf{1}_{[0,t]}(T_n) .$$

On considère que le phénomène à l'origine de l'arrivée des électrons est aléatoire mais ne change pas statistiquement et que les électrons se comportent de façon aléatoire. Si on change l'origine des temps en $t$, le nombre d'électrons supplémentaires ayant frappé la plaque entre les instants $t$ et $t+s$, qui est égal à $N_{t+s}-N_t$ doit donc avoir même loi que $N_s - N_0$. De plus le nombre d'électrons supplémentaires $N_{t+s}-N_t$ ne dépend pas des d'électrons qui sont arrivés dans le passé avant $t$. Cela implique que $\mathbf{N}=(N_t)_{t\in\mathbb{R}_+}$ est un processus markovien. On appelle un tel processus de comptage un *processus de Poisson*. Il et assez naturel de formaliser ce processus de comptage de la manière suivante.

**Définition II.4.1** Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité. Soit $\mathbf{N}=(N_t)_{t\in\mathbb{R}_+}$, un processus *càd croissant* partout sur $\Omega$ à valeurs dans $\mathbb{N}$. C'est un *processus de Poisson* s'il satisfait les hypothèses suivantes.

$(a)$ *Les accroissement de* $\mathbf{N}$ *sont indépendants*, c'est-à-dire que pour tous $0<t_1<\ldots<t_n$, les variables

$$N_0\, ,\, N_{t_1}-N_0\, ,\, N_{t_2}-N_{t_1}\, ,\, \ldots\, ,\, N_{t_n}-N_{t_{n-1}}\, ,$$

sont indépendantes.

$(b)$ *Les accroissements de* $\mathbf{N}$ *sont homogènes en loi* : c'est-à-dire que pour tous $s,t\in\mathbb{R}_+$, $N_{t+s}-N_s$ a même loi que $N_t-N_0$.

$(c)$ *C'est un processus de comptage*, c'est-à-dire que ses sauts valent au plus $1$ : $N_t(\omega)-N_{t-}(\omega)\in\{0,1\}$, pour tout $\omega\in\Omega$ et tout réel $t>0$. □

Démontrons qu'un processus de Poisson $\mathbf{N}=(N_t)_{t\in\mathbb{R}_+}$ comme dans la définition précédente est processus de Markov. Pour tous $i,j\in\mathbb{N}$, et tout $t\in\mathbb{R}_+$, on pose

$$p_t(i,j)=\mathbf{P}(N_t-N_0=j-i)\ .$$

On observe que $p_t(i,j)=0$ si $j<i$ et que $p_t(i,j)=p_t(0,j-i)$. Pour tous $t_0=0<t_1<\ldots<t_n$ et pour tout tous $i_0,\ldots,i_n\in\mathbb{N}$, on a

$$\begin{aligned}\mathbf{P}\left(N_0=i_0;N_{t_1}=i_1;\ldots;N_{t_n}=i_n\right) &= \mathbf{P}(N_0=i_0;N_{t_1}-N_0=i_1-i_0;\ldots;N_{t_n}-N_{t_{n-1}}=i_n-i_{n-1})\\ &= \mathbf{P}(N_0=i_0)\mathbf{P}(N_{t_1}-N_0=i_1-i_0)\mathbf{P}(N_{t_n}-N_{t_{n-1}}=i_n-i_{n-1})\\ &= \mathbf{P}(N_0=i_0)\mathbf{P}(N_{t_1}-N_0=i_1-i_0)\mathbf{P}(N_{t_n-t_{n-1}}-N_0=i_n-i_{n-1})\\ &= \mathbf{P}(N_0=i_0)p_{t_1}(i_0,i_1)p_{t_2-t_1}(i_1,i_2)\ldots p_{t_n-t_{n-1}}(i_{n-1},i_n)\end{aligned}$$

Cela implique d'une part que les $P_t=(p_t(i,j))_{i,j\in\mathbb{N}}$, $t\in\mathbb{R}_+$, forment un semi-groupe et d'autre part que $(N_t)_{t\in\mathbb{R}_+}$ est un processus markovien pour ce semi-groupe (au sens de la définition élémentaire II.1.12, à la page 153).

Calculons le générateur infinitésimal du processus de Poisson : on pose $\theta=q(0)=-q_{0,0}$, le paramètre d'attente en 0. Comme les sauts du processus sont exactement égaux 1, la chaîne du squelette est particulièrement simple : c'est un processus déterministe qui augmente de 1 à chaque étape. On en déduit notamment que $q_{i,j}=0$ si $j\notin\{i+1,i\}$ et donc $q_{i,i+1}=-q_{i,i}=q(i)$. Donc

$$1-q(i)t+o(t)=p_t(i,i)=p_t(0,0)=1-\theta t+o(t)\ .$$

On en déduit donc que $q(i)=q(0)=\theta$. Le semi-groupe, qui est très simple, est donc donné par

$$q_{i,j}=\begin{cases}\theta & \text{si } j=i+1,\\ -\theta & \text{si } j=i,\\ \text{et donc } 0 & \text{si } j\notin\{i,i+1\}\end{cases}$$

La décomposition du processus en un squelette discret et des durée d'attentes exponentielle donnée au théorème II.1.26 (page 164) montre ici que si $J_n$ est le $n$-ième temps de saut de $\mathbf{N}=(N_t)_{t\in\mathbb{N}}$, alors

$(a)$ $N_{J_n} - N_0 = n, n \in \mathbb{N}$.

$(b)$ les v.a. $D_n = J_n - J_{n-1}$, $n \geq 1$, sont i.i.d. de loi exponentielle de paramètre $\theta$.

On a donc

$$\forall t \in \mathbb{R}_+, \quad N_t = N_0 + \sum_{n \geq 1} \mathbf{1}_{[0,t]}(J_n) \ . \tag{II.66}$$

Il est possible de résoudre les équations de Kolmogorov-Chapman dans ce cas pour trouver explicitement les probabilités de transition $p_t(i,j)$, mais un calcul intégral direct s'appuyant sur $(a)$ et $(b)$ permet ce calcul :

$$\begin{aligned}
p_t(0,j) &= \mathbf{P}(N_t - N_0 = j) = \mathbf{P}(J_j \leq t < J_j + D_{j+1}) \\
&= \mathbf{E}\big[\mathbf{1}_{\{0 \leq t - J_j\}} e^{-\theta(t-J_j)}\big] \\
&= \int_0^\infty \cdots \int_0^\infty ds_1 \ldots ds_j \theta^j e^{-\theta(s_1+\ldots+s_j)} \mathbf{1}_{[0,t]}(s_1 + \ldots + s_j) e^{-\theta(t-(s_1+\ldots+s_j))} \\
&= \theta^j e^{-\theta t} \int_0^\infty \cdots \int_0^\infty ds_1 \ldots ds_j \mathbf{1}_{[0,t]}(s_1 + \ldots + s_j) \\
&= \theta^j e^{-\theta t} \int_0^\infty \cdots \int_0^\infty dt_1 \ldots dt_j \mathbf{1}_{\{0 \leq t_1 \leq \ldots \leq t_j \leq t\}} \\
&= \theta^j e^{-\theta t} t^j / j! \ .
\end{aligned}$$

Donc

$$\forall t \in \mathbb{R}_+, \forall i,j \in \mathbb{N}, \quad p_t(i,j) = \mathbf{1}_{\{i \leq j\}} e^{-\theta t} \frac{(\theta t)^{j-i}}{(j-i)!} \ .$$

La définition II.4.1 des processus de Poisson peut se reformuler comme suit à l'aide de la filtration naturelle $\mathscr{F}_t^{\mathbf{N}} = \sigma(N_s; s \in [0,t])$, $t \in \mathbb{R}_+$ : l'indépendance des accroissements $(a)$ est équivalente à ce que

$(a')$ *Pour tous* $t, s \in \mathbb{R}_+$, $N_{t+s} - N_t$ *est indépendant de* $\mathscr{F}_t^{\mathbf{N}}$.

Cela justifie l'introduction de la définition suivante d'un processus de Poisson *relativement à une filtration*.

**Définition II.4.2** (*Processus de Poisson relativement à une filtration*) Soit $\mathbf{N} = (N_t)_{t \in \mathbb{R}_+}$ un processus à valeurs dans $\mathbb{N}$ qui est défini sur un espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$. On suppose qu'il est càd don les sauts sont égaux à 1, partout sur $\Omega$. Soit $(\mathscr{F}_t)_{t \in \mathbb{R}_+}$, une filtration sur $(\Omega, \mathscr{F})$. On dit que $\mathbf{N}$ est un processus de Poisson *relativement à la filtration* $(\mathscr{F}_t)_{t \in \mathbb{R}_+}$, s'il satisfait les trois conditions suivantes.

$(a)$ Pour tout $t \in \mathbb{R}_+$, $N_t$ est $\mathscr{F}_t$-mesurable, c'est-à-dire que $\mathbf{N}$ est $(\mathscr{F}_t)_{t \in \mathbb{R}_+}$-adapté.

$(b)$ Pour tous $s, t \in \mathbb{N}$, $N_{t+s} - N_t$ est une variable indépendante de la tribu $\mathscr{F}_t$.

$(c)$ Pour tous $t, s \in \mathbb{R}_+$, $N_{t+s} - N_t$ a même loi que $N_s - N_0$. □

Il est facile de montrer qu'un processus de Poisson au sens de la définition II.4.1 , page 197, est un processus de Poisson relativement à sa filtration naturelle $(\mathscr{F}_t^{\mathbf{N}})_{t \in \mathbb{R}_+}$, comme à définition précédente.

Il peut paraître un peu artificiel d'introduire une filtration dans la définition des processus de Poisson mais comme on le verra plus loin c'est une idée utile lorsque l'on considère plusieurs processus de Poisson (ou plusieurs processus, dont des processus de Poisson). Pour illustrer ce fait, nous énonçons ci-dessous un résultat (dont la preuve est donnée dans les exercices) où la filtration joue un rôle à la fois crucial, qui n'est pas élémentaire.

**Proposition II.4.3** *Soit* $(\Omega, \mathscr{F}, \mathbf{P})$ *un espace de probabilité. Soit* $(\mathscr{F}_t)_{t \in \mathbb{R}_+}$, *une filtration sur* $(\Omega, \mathscr{F})$. *Soit un entier* $d \geq 2$. *Soient*

$$\mathbf{N}^1 = (N_t^1)_{t \in \mathbb{R}_+} \ , \ \ldots \ , \ \mathbf{N}^d = (N_t^d)_{t \in \mathbb{R}_+} \ ,$$

*des processus de Poisson* relativement à la même filtration $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$ *qui sont tous issus de* $0$ *:* $\mathbf{P}(N_0^1 = \ldots = N_0^d = 0) = 1$. *Pour tout* $k \in \{1, \ldots, d\}$, *on note* $(T_n^k)_{n\geq 1}$, *la suite des temps de sauts de* $\mathbf{N}^k$, *c'est-à-dire que*

$$N_t^k = \sum_{n\geq 1} \mathbf{1}_{[0,t]}(T_n^k) \ , \quad t \in \mathbb{R}_+ .$$

*On suppose que les processus de Poisson* $(\mathbf{N}^k)_{1\leq k\leq d}$ ne sautent pas en même temps, *c'est-à-dire que*

$$\mathbf{P}\text{-}p.s. \quad \forall\, 1 \leq k < \ell \leq d\,, \quad \{T_1^k < \ldots < T_n^k < \ldots\} \cap \{T_1^\ell < \ldots < T_n^\ell < \ldots\} = \emptyset\ . \tag{II.67}$$

*Alors les processus de Poisson* $\mathbf{N}^1, \ldots, \mathbf{N}^d$ *sont indépendants, ce qui est équivalent à dire que les suites* $(T_n^1)_{n\geq 1}$, ... , $(T_n^d)_{n\geq 1}$, *sont indépendantes.*

**Preuve :** voir l'exercice II.11 page 204. ■

### II.4.b Marches aléatoires.

**Définition II.4.4** Soit $\mathbf{X} = (X_t)_{t\in\mathbb{R}_+}$ un processus à valeurs dans $\mathbb{Z}^d$ qui est défini sur l'espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$. On le suppose non-constant et continu à droite. On dit que c'est une marche aléatoire ssi il satisfait les deux conditions suivantes.

$(a)$ (*Homogénéité en loi*) Pour tous $s, t \in \mathbb{R}_+$, $X_{t+s} - X_t$ a même loi que $X_s - X_0$.

$(b)$ (*Indépendance des accroissements*) Pour tous $t_1 \leq \ldots \leq t_n$, les variables

$$X_0\ ,\ X_{t_1} - X_0\ ,\ X_{t_2} - X_{t_1}\ ,\ \ldots\ ,\ X_{t_n} - X_{t_{n-1}}\ ,$$

sont indépendantes. □

La proposition suivante justifie le terme de marche aléatoire.

**Proposition II.4.5** *Soit* $(X_t)_{t\in\mathbb{R}_+}$, *une marche aléatoire sur* $\mathbb{Z}^d$ *homogène en loi définie sur un espace de probabilité* $(\Omega, \mathscr{F}, \mathbf{P})$. *Alors, il existe un paramètre* $\theta \in ]0, \infty[$ *et une loi de probabilité* $\nu$ *sur* $\mathbb{Z}^d$ *telle que* $\nu(0) = 0$ *qui décrivent la loi de* $\mathbf{X}$ *de la façon suivante.*

$(i)$ $\mathbf{X}$ *est un processus de Markov à valeurs dans* $\mathbb{Z}^d$ *dont le générateur* $G = (q_{i,j})_{i,j\in\mathbb{Z}^d}$ *est donné par*

$$\forall i, j \in \mathbb{Z}^d\,, \quad q_{i,j} = \begin{cases} -\theta & si\ i = j, \\ \theta\nu(j - i) & si\ i \neq j. \end{cases}$$

$(ii)$ *Il existe une suite de variables* $(\mathscr{E}_n)_{n\geq 1}$, *indépendantes et de même loi exponentielle de paramètre* $\theta$, *et il existe une suite* $(\xi_n)_{n\geq 1}$ *de variables indépendantes, à valeurs dans* $\mathbb{Z}^d$, *de même loi* $\nu$, *suites qui satisfont les conditions suivantes.*

*(ii-a)* $X_0$, $(\mathscr{E}_n)_{n\geq 1}$ *et* $(\xi_n)_{n\geq 1}$ *sont indépendantes.*

*(ii-b) On a*

$$X_t = X_0 + \sum_{1\leq k\leq N_t} \xi_k \ = \sum_{n\geq 1} \mathbf{1}_{[0,t]}(T_n)\, \xi_n\ ,$$

*où on a posé*

$$\forall t \in \mathbb{R}_+\,, \quad N_t = \sum_{n\geq 1} \mathbf{1}_{[0,t]}(T_n) \quad avec \quad T_n = \mathscr{E}_1 + \ldots + \mathscr{E}_n\ ,$$

*qui est un processus de Poisson de paramètre* $\theta$ *indépendant de la suite des sauts* $(\xi_n)_{n\geq 1}$.

*La loi* $\nu$ *est appelé la loi des sauts de la marche* $\mathbf{X}$.

**Preuve :** voir l'exercice II.4. ■

### II.4.c Contruction des processus minimaux à l'aide des processus de Poisson.

On se propose de donner une construction du processus minimal associé à un générateur infinitésimal $G = (q_{i,j})_{i,j \in E}$ sur $E$ en utilisant *un processus de Poisson* et *un changement de temps*. L'approche n'est pas fondamentalement différente de celle de la section II.2.b page 172. On note $q(i) = -q_{i,i}$, $i \in E$ la famille des paramètres de temps d'attente associée à $G$ et on note $Q = (\pi_{i,j})_{i,j \in E}$, la matrice de transition du squelette qui lui est également associée. On se donne les deux processus suivants.

$(a)$ Une chaîne de Markov à valeurs dans $E$ :

$$\big( \Omega \, ; \, \mathscr{F} \, ; \, \mathbf{Y} = (Y_n)_{n \in \mathbb{N}} \, ; \, Q \, ; \, \mathbf{P}_\mu , \; \mu \in \mathscr{M}_1(E^*) \big) \; .$$

La filtration choisie est la filtration naturelle de $\mathbf{Y}$ : $\mathscr{F}_n^{\mathbf{Y}} = \sigma(Y_0, \ldots, Y_n)$, $n \in \mathbb{N}$. Elle ne jouera aucun rôle.

$(b)$ Une suite de variables $\mathscr{E}_n : \Omega \to \mathbb{R}_+$, $n \geq 1$, $\mathscr{F}$-mesurables telle que pour toute loi d'entrée $\mu \in \mathscr{M}_1(E)$, sous $\mathbf{P}_\mu$, la suite $(\mathscr{E}_n)_{n \geq 1}$ est i.i.d. de loi exponentielle de paramètre 1 indépendante de la chaîne $\mathbf{Y}$.

On pose alors $T_n = \mathscr{E}_1 + \ldots + \mathscr{E}_n$, $n \geq 1$ et

$$\forall t \in \mathbb{R}_+ \, , \quad N_t = \sum_{n \geq 1} \mathbf{1}_{[0,t]}(T_n) \; ,$$

si bien que pour toute loi d'entrée $\mu \in \mathscr{M}_1(E)$, sous $\mathbf{P}_\mu$, $\mathbf{N} = (N_t)_{t \in \mathbb{R}_+}$ est un processus de Poisson de paramètre 1 indépendant de la chaîne $\mathbf{Y}$. On pose également

$$\widetilde{G} = Q - \mathrm{Id}_E \quad \text{et} \quad \forall t \in \mathbb{R}_+ \, , \quad \widetilde{X}_t = Y_{N_t} \; .$$

On voit que la matrice $\widetilde{G}$ est un générateur infinitésimal sur $E$. On voit également que $\widetilde{\mathbf{X}} = (\widetilde{X}_t)_{t \in \mathbb{R}_+}$ est un processus à valeurs dans $E$ qui est continu à droite. On pose

$$\forall t \in \mathbb{R}_+ \, , \; \forall i, j \in E \, , \quad \widetilde{p}_t(i,j) = \mathbf{P}_i(\widetilde{X}_t = j) \; .$$

D'après les propriétés d'indépendances on a

$$\mathbf{P}_i\text{-p.s.} \quad \mathbf{E}_i\big[ \mathbf{1}_{\{\widetilde{X}_t = i\}} \, | \mathbf{N} \big] = [Q^{N_t}](i,j) \; .$$

Par conséquent,

$$\widetilde{p}_t(i,j) = \mathbf{E}\big[ \, [Q^{N_t}](i,j) \, \big] = \sum_{n \geq 0} e^{-t} \frac{t^n}{n!} [Q^n](i,j) \; .$$

On pose $\widetilde{P}_t = (\widetilde{p}_t(i,j))_{i,j \in E}$. Ce qui précède montre que

$$\widetilde{P}_t = \sum_{n \geq 0} e^{-t} \frac{t^n}{n!} Q^n = \exp\big( t.(Q - \mathrm{Id}_E) \big) = \exp(t.\widetilde{G}) \; ,$$

la convergence de la série de matrices ayant lieu dans $M_E(\mathbb{R})$ muni de la norme $|\cdot|$, comme expliqué au (II.65) page 196, à la section II.3.

Donc $(\widetilde{P}_t)_{t \in \mathbb{R}_+}$ est un semi-groupe markovien sur $E$ associé au générateur uniforme $\widetilde{G}$. On peut voir cela plus directement comme suit. On fixe les temps $t_1 < \ldots < t_n$ et les états $i_0, \ldots, i_n \in E$. On remarque pour toute loi d'entrée $\mu \in \mathscr{M}_1(E)$, on a $\mathbf{P}_\mu$-presque sûrement,

$$\mathbf{E}_\mu\big[ \mathbf{1}_{\{\widetilde{X}_0 = i_0 ; \widetilde{X}_{t_1} = i_1 ; \ldots ; \widetilde{X}_{t_n} = i_n\}} | \, \mathbf{N} \, \big] = \mu(i_0) [Q^{N_{t_1}}](i_0, i_1) \cdot [Q^{N_{t_2} - N_{t_1}}](i_1, i_2) \ldots [Q^{N_{t_n} - N_{t_{n-1}}}](i_{n-1}, i_n).$$

Comme $N_{t_1}, N_{t_2} - N_{t_1}, \ldots, N_{t_n} - N_{t_{n-1}}$ sont des variables de Poisson indépendantes de paramètres $t_1, t_2 - t_1, \ldots, t_n - t_{n-1}$, on en déduit que

$$\mathbf{P}_\mu\big(\widetilde{X}_0 = i_0; \widetilde{X}_{t_1} = i_1; \ldots; \widetilde{X}_{t_n} = i_n\big) = \mu(i_0) p_{t_1}(i_0, i_1) \ldots p_{t_n - t_{n-1}}(i_{n-1}, i_n) ,$$

ce qui montre que $\widetilde{\mathbf{X}}$ est un processus markovien de semi-groupe $(\widetilde{P}_t)_{t\in\mathbb{R}_+}$ : ce n'est pas un processus explosif car $\widetilde{q}(i) = 1$ pour tout $i \in E$ et on utilise le critère d'après le critère $(b)$ de la proposition II.2.10, page 178. C'est donc le processus minimal associé au générateur uniforme $\widetilde{G}$.

Pour tout $i \in E$, on pose ensuite

$$v(i) = 1/q(i) \text{ si } q(i) > 0 \quad \text{et} \quad v(i) = 1 \text{ si } q(i) = 0.$$

On voit que $v : E \to ]0, \infty[$ . On effectue le changement de temps suivant en posant pour tout $t \in \mathbb{R}_+$,

$$A_t = \int_0^s v(\widetilde{X}_s)\, ds \quad \text{et} \quad \tau_t = \inf\{s \in \mathbb{R}_+ \ : \ A_s > t\} ,$$

avec la convention $\inf \emptyset = \infty$. On pose ensuite

$$X_t = \widetilde{X}_{\tau_t} \quad \text{si } \tau_t < \infty \quad \text{et} \quad X_t = \partial \quad \text{si } \tau_t = \infty,$$

où $\partial \notin E$ est un point cimetière.

La proposition II.3.1 page 189 implique $(X_t)_{t\in\mathbb{R}_+}$ est un processus minimal associé au générateur dont les entrées sont données par $\widetilde{q}_{i,j}/v(i)$, $i, j \in E$, où $\widetilde{G} = Q - \mathrm{Id}_E = (\widetilde{q}_{i,j})_{i,j\in E}$. Or si $q(i) > 0$ et $i \neq j$, alors $\widetilde{q}_{i,j} = \pi_{i,j}$ et $\widetilde{q}_{i,j}/v(i) = q(i)\pi_{i,j} = q_{i,j}$ ; de plus $\widetilde{q}_{i,i} = -1$ donc $\widetilde{q}_{i,i}/v(i) = -q(i) = q_{i,i}$. Si $q(i) = 0$ et $i \neq j$, alors $\widetilde{q}_{i,j} = \pi_{i,j} = 0$ et $\widetilde{q}_{i,j}/v(i) = 0 = q_{i,j}$ ; de plus $\pi_{i,i} = 1$, donc $\widetilde{q}_{i,i} = 0$, et on a $\widetilde{q}_{i,i}/v(i) = 0 = q_{i,i}$. Cela montre que $(X_t)_{t\in\mathbb{R}_+}$ est un *processus markovien minimal de générateur $G$*, ce qui achève la construction.

## EXERCICES.

**Exercice II.4** Cet exercice détaille une preuve de la proposition II.4.5. On suppose que $\mathbf{X} = (X_t)_{t\in\mathbb{R}_+}$ est une marche aléatoire dans $\mathbb{Z}^d$ définie sur l'espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$.

1. Pour tout $i, j \in \mathbb{Z}^d$ et tout $t \in \mathbb{R}_+$, on pose $p_t(i,j) = \mathbf{P}(X_t - X_0 = i - j)$. En raisonnant comme dans la preuve du lemme **??**, montrer que $P_t = (p_t(i,j))_{i,j\in\mathbb{Z}^d}$, $t \in \mathbb{R}_+$, est un semi-groupe markovien sur $\mathbb{Z}^d$.
2. Montrer que $\mathbf{X}$ est un processus de Markov de semi-groupe $(P_t)_{t\in\mathbb{R}_+}$.
3. On note $G = (q_{i,j})_{i,j\in\mathbb{Z}^d}$ le générateur infinitésimal associé à $(P_t)_{t\in\mathbb{R}_+}$. Montrer que puisque $p_t(i,j) = p_t(0, j-i)$, on a

$$\forall i, j \in \mathbb{Z}^d , \quad q_{i,j} = q_{0,j-i} .$$

   On pose $\theta = -q_{0,0}$. Déduire de ce qui précède que $q_{i,i} = -\theta$, pour tout $i \in \mathbb{Z}^d$. Expliquer pourquoi le fait que $\mathbf{X}$ est supposé non-constant implique que $\theta > 0$.
4. On pose $\nu(k) = q_{0,k}/\theta$, pour tout $k \in \mathbb{Z}^d \backslash \{0\}$ et $\nu(0) = 0$. Montrer que $\nu$ est une mesure de probabilité et montrer que $G$ est bien de la forme donné par le point **(i)** de la proposition II.4.5.
5. On note $\mathscr{E}_n = D_n(\mathbf{X})$, $n \geq 1$, les durées d'attentes de $\mathbf{X}$. Montrer que la suite $(\mathscr{E}_n)_{n\geq 1}$ est i.i.d. de loi exponentielle de paramètre $\theta$ et montrer qu'elle est indépendante de la chaîne du squelette $\mathbf{Y}(\mathbf{X}) = (Y_n(\mathbf{X}))_{n\geq 0}$, qui est associée à $\mathbf{X}$.
6. Montrer que la chaîne du squelette est une marche aléatoire sur $\mathbb{Z}^d$ de loi de saut $\nu$ et conclure. □

**Exercice II.5** Soit $\mathbf{X} = (X_t)_{t\in\mathbb{R}_+}$ un processus de Markov minimal à valeurs dans $\mathbb{Z}^d$ de générateur $G = (q_{i,j})_{i,j\in\mathbb{Z}^d}$, non-nul et tel que

$$\forall i,j,k \in \mathbb{Z}^d\,, \quad q_{i,j} = q_{i+k,j+k}\ .$$

Montrer que $\mathbf{X}$ est une marche aléatoire. $\square$

**Exercice II.6** On considère une marche aléatoire homogène en loi de loi de saut $\nu$, qui est une probabilité sur $\mathbb{Z}^d$ et de paramètre des durées d'attente $\theta$. On suppose que toutes les variables sont définies sur l'espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$. On note la marche $\mathbf{X} = (X_t)_{t\in\mathbb{R}_+}$ et on suppose que $\mathbf{P}(X_0 = 0) = 1$. Pour tout $\mathbf{u} = (u_1, \ldots, u_d) \in \mathbb{R}^d$, on pose

$$\psi(\mathbf{u}) = \sum_{k\in\mathbb{Z}^d} e^{i\langle \mathbf{u},k\rangle}\nu(k)\ .$$

La fonction $\psi : \mathbb{R}^d \to \mathbb{C}$ est donc la fonction caractéristique de $\nu$. Calculer $\mathbf{E}\big[e^{i\langle \mathbf{u},X_t\rangle}\big]$ en fonction de $t$, $\mathbf{u}$, $\psi$ et $\theta$. $\square$

**Exercice II.7** On fixe

$$\big(\,\Omega\,;\,\mathscr{F}\,;\,(\mathscr{F}_t)_{t\in\mathbb{N}}\,;\,\mathbf{N} = (N_t)_{t\in\mathbb{R}_+}\,;\,G\,;\,\mathbf{P}_\mu\,,\ \mu\in\mathscr{M}_1(\mathbb{N})\,\big)$$

un processus de Poisson de paramètre $\theta$. On fixe $\mathbf{V} = (V_t)_{t\in\mathbb{R}_+}$, une famille de variables aléatoires telle que

— pour tout $t \in \mathbb{R}_+$, $V_t : \Omega \to \mathbb{C}$ est $\mathscr{F}_t$-mesurable ;

— pour tout $\omega \in \Omega$, $t \mapsto V_t(\omega)$ est continue à gauche ;

— pour tout $t \in \mathbb{R}_+$, il existe une constante $c_t$ telle que

$$\forall \omega\in\Omega\,,\quad \sup_{s\in[0,t]} |V_s(\omega)| \le c_t\ .$$

Pour tout $t \in \mathbb{R}_+$, on définit le symbole suivant

$$\int_0^t V_s dN_s = \sum_{n\ge 1} \mathbf{1}_{[0,t]}(T_n)\, V_{T_n}\ ,$$

où la suite $T_n := J_n(\mathbf{N})$, $n \ge 1$, est la suite des temps de sauts de $\mathbf{N}$.

1. On note $\lfloor\cdot\rfloor$ l'application partie entière. Montrer que

$$\forall\omega\in\Omega\,,\quad \int_0^t V_s(\omega)ds = \lim_n \int_0^t V_{\lfloor ns\rfloor/n}(\omega)ds.$$

et que

$$\int_0^t V_{\lfloor ns\rfloor/n}ds = \tfrac{nt-\lfloor nt\rfloor}{n}\, V_{\lfloor nt\rfloor/n} + \sum_{0\le\lfloor nt\rfloor} \tfrac{1}{n}V_{k/n}\ .$$

Montrer que $\int_0^t V_s\, ds : \Omega \to \mathbb{C}$ est $\mathscr{F}_t$-mesurable intégrable.

2. Pour tout $n \ge 1$, on pose ensuite

$$U_n = V_{\lfloor nt\rfloor/n}\big(\,N_t - N_{\lfloor nt\rfloor/n}\,\big) + \sum_{0\le k\le\lfloor nt\rfloor} V_{k/n}\,\big(\,N_{(k+1)/n} - N_{k/n}\,\big)\ .$$

Montrer que, partout sur $\Omega$, on a

$$\lim_n U_n = \int_0^t V_s dN_s \quad \text{et} \quad |U_n| \le c_t\, N_t\ .$$

En déduire que $\int_0^t V_s dN_s$ est $\mathscr{F}_t$-mesurable et intégrable.

3. On fixe $0 \le k \le \lfloor nt\rfloor$. Montrer que

$$\mathbf{E}\big[V_{k/n}\,\big(\,N_{(k+1)/n} - N_{k/n}\,\big)\,\big|\,\mathscr{F}_{k/n}\big] = \theta\cdot\tfrac{1}{n}V_{k/n}\ .$$

En déduire que $\mathbf{E}[U_n] = \theta\cdot\mathbf{E}[\int_0^t V_{\lfloor ns\rfloor/n}ds]$.

4. Montrer que
$$\mathbf{E}\Big[\int_0^t V_s dN_s\Big] = \theta \cdot \int_0^t \mathbf{E}[V_s]\, ds\ .$$

$\square$

**Exercice II.8** Toutes les variables sont définies sur un même espace de probabilité. Soit $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$, une filtration sur $(\Omega, \mathscr{F})$. Pour tout $1 \leq k \leq d$, on se donne un processus de Poisson $\mathbf{N}^k = (N^k_t)_{t\in\mathbb{R}_+}$ de paramètre $\theta_k$ et on suppose que c'est un processus de Markov relativement à la filtration $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$ issu de 0, c'est-à-dire que

— $\mathbf{P}(N^k_0 = 0) = 1$ ;

— pour tous $s, t \in \mathbb{R}_+$, $N^k_{t+s} - N^k_t$ est indépendante de $\mathscr{F}_t$ et suit une loi de poisson de paramètre $s.\theta_k$.

On suppose de plus que les $d$-processus de Poisson $\mathbf{N}^1, \ldots \mathbf{N}^d$ sont *indépendants*. On pose
$$N_t = (N^1_t, \ldots, N^d_t) \in \mathbb{N}^d\ , \quad t \in \mathbb{R}_+\ .$$

1. Montrer que $\mathbf{N} = (N_t)_{t\in\mathbb{R}_+}$ est un processus de Markov à valeurs dans $\mathbb{N}^d$ relativement à la filtration $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$.
2. Montrer que $\mathbf{N}$ est une marche aléatoire à valeurs dans $\mathbb{N}^d$ de paramètre de durée d'attence $\theta = \theta_1 + \ldots + \theta_d$ et de loi de saut $\nu$ donnée par $\nu(e_k) = \theta_k/\theta$, $1 \leq k \leq d$, où $e_1, \ldots, e_d$ désignent les vecteurs de la base canonique de $\mathbb{R}^d$.
3. On note $T_n = J_n(\mathbf{N})$, $n \geq 1$, la suite des temps de sauts de $\mathbf{N}$. Pour tout $n \geq 1$, on pose $\chi_n = k$ ssi le $n$-ième saut de $\mathbf{N}$ est dû à $\mathbf{N}^k$, c'est-à-dire qu'il s'effectue dans la direction $k$ portée par $e_k$ :
$$\chi_n = k \iff N_{T_n} - N_{T_n-} = e_k\ .$$
On pose également
$$N^*_t = N^1_t + \ldots + N^d_t\ , \quad t \in \mathbb{R}_+\ .$$
Montrer que
$$N^*_t = \sum_{n\geq 1} \mathbf{1}_{[0,t]}(T_n)\ , \quad t \in \mathbb{R}_+\ .$$
En déduire que $\mathbf{N}^* = (N^*_t)_{t\in\mathbb{R}_+}$ est un processus de Poisson de paramètre $\theta$.
4. Montrer que $(\chi_n)_{n\geq 1}$ est une suite i.i.d. à valeurs dans $\{1, \ldots, d\}$ de loi donnée par
$$\mathbf{P}(\chi_n = k) = \theta_k/\theta\ , \quad 1 \leq k \leq d\ .$$
Montrer que $\mathbf{N}^*$ et $(\chi_n)_{n\geq 1}$ sont indépendants (ce qui est équivalent, on le notera à l'indépendance de $(T_n)_{n\geq 1}$ et de $(\chi_n)_{n\geq 1}$).
5. Donner un énoncé sans mentionner de filtration.
6. On se donne maintenant une suite de variables indépendantes $(\mathscr{E}_n)_{n\geq 1}$ de même loi exponentielle de paramètre $\theta$. On se donne également une suite de variables $(\kappa_n)_{n\geq 1}$, à, valeurs dans $\{1, \ldots, d\}$, indépendantes de même loi donnée par
$$\mathbf{P}(\kappa_n = k) = p_k\ , \quad 1 \leq k \leq d\ .$$
On suppose que $0 < p_k < 1$, pour tout $1 \leq k \leq d$. On suppose de plus que $(\mathscr{E}_n)_{n\geq 1}$ et $(\kappa_n)_{n\geq 1}$ sont indépendantes. On pose $T^\bullet_n = \mathscr{E}_1 + \ldots + \mathscr{E}_n$. Pour tout $1 \leq k \leq d$, on pose également
$$M^k_t = \sum_{n\geq 1} \mathbf{1}_{[0,t]}(T^\bullet_n)\mathbf{1}_{\{\kappa_n = k\}}\ , \quad t \in \mathbb{R}_+\ .$$
Montrer que $(M^1)_{t\in\mathbb{R}_+}, \ldots, (M^d_t)_{t\in\mathbb{R}_+}$ sont $d$ processus de Poisson indépendants de paramètres respectifs $p_1\theta, \ldots, p_d\theta$. $\square$

**Exercice II.9** Toutes les variables sont définies sur un même espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$, muni d'une filtration $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$. On se donne un $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$-processus markovien à valeurs dans $\mathbb{N}^d$. Il est noté $\mathbf{N} = (N_t)_{t\in\ \mathbb{R}_+}$. On pose
$$N_t = (N^1_t, \ldots, N^d_t)\ , \quad t \in \mathbb{R}_+\ ,$$
$N^k_t$ étant la $k$-ième coordonnée de $N_t$ dans la base canonique de $\mathbb{R}^d$. On fait les deux hypothèses suivantes.

— Pour tout $1 \leq k \leq d$, $\mathbf{N}^k = (N^k_t)_{t\in\mathbb{R}_+}$ est un processus de Poisson relativement à la filtration $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$, issu de 0 : on note $\theta_k$ son paramètre.

— On note $(T^k_n)_{n\geq 1}$ la suite des temps de sauts de $\mathbf{N}^k$. On suppose que

$$\mathbf{P}\text{-p.s.} \quad \forall 1 \leq k < \ell \leq d\,, \quad \{T^k_n\,;\, n\geq 1\} \cap \{T^\ell_n\,;\, n\geq 1\} = \emptyset\,.$$

Autrement dit les processus de Poisson $\mathbf{N}^k$ ne sautent pas en même temps.

1. Calculer le générateur infinitésimal de $\mathbf{N}$. Constater que c'est le même que celui du processus de l'exercice II.8.
2. En déduire que $\mathbf{N}^1, \ldots, \mathbf{N}^d$ sont indépendants. □

**Exercice II.10** Toutes les variables sont définies sur un même espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$. Cet espace est muni d'une filtration notée $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$. On considère un processus de Markov à valeurs dans $\mathbb{Z}^d$, relativement à la filtration $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$. On le note $\mathbf{X} = (X_t)_{t\in\mathbb{R}_+}$ et on suppose que c'est une marche aléatoire de loi de saut $\nu$ et de paramètre des durée d'attente noté $\theta$. On suppose deux choses sur $\nu$.

**(a)** L'espace vectoriel engendré par les vecteurs $j \in \mathbb{Z}^d$ tels que $\nu(j)$ est $\mathbb{Z}^d$.

**(b)** Si $j \in \mathbb{Z}^d$ est tel que $\nu(j) > 0$, alors toutes les coordonées de $j$ dans la base canonique sont nulles sauf une.

On pose $X_t = (X^1_t, \ldots, X^d_t)$, $t \in \mathbb{R}_+$, qui sont les coordonnées de $X_t$ dans la base canonique.

1. En s'aidant de l'exercice II.11 (ou de la proposition II.4.3) montrer que $\mathbf{X}^1 = (X^1_t)_{t\in\mathbb{R}_+}, \ldots, \mathbf{X}^d = (X^d_t)_{t\in\mathbb{R}_+}$ sont $d$ marches aléatoires relativement à $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$ qui sont **indépendantes** à valeurs dans $\mathbb{Z}$ dont on précisera les lois de saut et les paramètres de durées d'attentes respectifs.
2. Réciproquement, montrer que si $\mathbf{X}^* = (X^*_t)_{t\in\mathbb{R}_+}$ est une marche aléatoire homogène en loi sur $\mathbb{Z}^d$ dont les coordonnées sont des processus de Markov indépendants, alors sa loi de saut satisfait l'hypothèse **(b)**. □

**Exercice II.11** *Le but de cet exercice est de détailler une preuve de la proposition II.4.3. Il est nécessaire d'avoir traité l'exercice II.7. On adopte les hypothèses et les notations de la proposition II.4.3.* On note $\theta_k$ le paramètre de $\mathbf{N}^k$, $1 \leq k \leq d$. On se donne $f_1, \ldots, f_d : \mathbb{R}_+ \to \mathbb{R}$, $d$ fonctions *continues à gauches bornées* et on pose pour tout $t \in \mathbb{R}_+$

$$Y_t = \exp\Big(i \sum_{1\leq k\leq d} \int_0^t f_k(s)dN^k_s\Big)\,.$$

1. Montrer que $Y_t$ est $\mathscr{F}_t$-mesurable, que $|Y_t| = 1$ et que $t \mapsto Y_t$ admet une limite à gauche en tout temps $t > 0$.
2. Pour tout $1 \leq k \leq d$, on pose alors
$$\forall s \in \mathbb{R}_+\,, \quad V^k_s = (e^{i\, f_k(s)} - 1)Y_{s-}\,.$$
Montrer que que $s \mapsto V^k_s$ est continue à gauche, et enfin que $\mathbf{V}^k = (V^k_t)_{t\in\mathbb{R}_+}$ est $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$-adapté. Montrer que le fait que les $\mathbf{N}^k$ ne sautent pas en même temps implique que
$$Y_t - 1 = \sum_{1\leq k\leq d} \int_0^t V^k_s dN^k_s\,.$$
3. Pour tout $t \in \mathbb{R}_+$, on pose $\phi(t) = \mathbf{E}[Y_t]$. Déduire de la question précédente et de l'exercice II.7 que
$$\phi(t) = 1 + \sum_{1\leq k\leq d} \theta_k \int_0^t (e^{i\, f_k(s)} - 1)\phi(s)\, ds\,.$$
4. En déduire que
$$\phi(t) = \mathbf{E}\Big[\prod_{1\leq k\leq d} \exp\Big(i\int_0^t f_k(s)dN^k_s\Big)\Big] = \prod_{1\leq k\leq q} \exp\Big(\theta_k \int_0^t (e^{i\, f_k(s)} - 1)\, ds\Big)\,.$$
5. On fixe ensuite $p$ réels positifs $0 \leq t_1 < \ldots t_p \leq t$; pour tout $1 \leq k \leq d$, on fixe aussi $d$ vecteurs $\mathbf{u}^k = (u^k_1, \ldots, u^k_p) \in \mathbb{R}^p$, et on pose
$$\forall s \in \mathbb{R}_+\,, \quad f_k(s) = \sum_{1\leq j\leq p} u^k_j \mathbf{1}_{[0,t_j]}(s)\,.$$
Vérifier que $f_k$ est continue à gauche, bornée. Montrer que
$$\int_0^t f_k(s)dN^k_s = \sum_{1\leq j\leq p} u^k_j N^k_{t_j}\,.$$
Montrer que
$$\mathbf{E}\Big[\prod_{1\leq k\leq d} \exp\big(i\, u^k_1 N^k_{t_1} + \ldots + iu^k_p N^k_{t_p}\big)\Big] = \prod_{1\leq k\leq d} \mathbf{E}\big[\exp\big(i\, u^k_1 N^k_{t_1} + \ldots + iu^k_p N^k_{t_p}\big)\big]\,.$$
6. Conclure. □

### II.4.d Processus de naissance et de mort

Dans cette section nous choisissons comme espace d'états $E = \mathbb{N}$ et nous considérons les générateurs infinitésimaux qui sont "tridiagonaux". Plus précisément, on se donne $G = (q_{i,j})_{i,j\in\mathbb{N}}$ un générateur infinitésimal tels que

$$\forall i, j \in \mathbb{N} \ , \quad q_{i,j} = 0 \quad \text{si} \quad |i - j| \geq 2 \ . \tag{II.68}$$

Autrement dit les seules entrées non-nulles de la matrice $G$ sont parmi les nombres $q_{0,0}$, $q_{0,1}$ et $q_{i,i-1}$, $q_{i,i}$ et $q_{i,i+1}$, pour $i \geq 1$. Puisque $G$ est un générateur infinitésimal on a :

**(a)** $q_{0,1}, q_{i,i-1}, q_{i,i+1} \geq 0$, pour tout $i \geq 1$ ;

**(b)** $a_0 := q_{0,1} = -q_{0,0}$ ;

**(c)** $q_{i,i-1} + q_{i,i+1} = -q_{i,i}$, et on pose $a_i = q_{i,i+}$ et $b_i = q_{i,i-1}$ pour tout $i \geq 1$.

On voit donc que $G$ est entièrement caractérisée par la suite $(a_i)_{i\geq 0}$ des *taux de naissance* et la suite $(b_i)_{i\geq 1}$ des *taux de décès*. On dit que le processus de générateur $G$ est un processus de *naissance pure* si $b_i = 0$, si pour tout $i \geq 1$. C'est un processus de *mort pure* ssi $a_i = 0$ pour tout $i \geq 0$. La matrice de transition de la chaîne du squelette $Q = (\pi_{i,j}, i, j \in \mathbb{N})$ est également tridiagonale car on a $\pi_{i,j} = 0$ dès que $|i - j| \geq 2$. Le squelette est une chaîne de naissance et de mort.

L'évolution d'un processus minimal de générateur $G$ peut se décrire grossièrement de la manière suivante : soit le processus est aborbé dans un état $i$ tel que $q(i) = -q_{i,i} = 0$, soit il attend un temps exponentiel de paramètre $q(i) = -q_{i,i}$ et il augmente de 1 avec probabilité $\pi_{i,i+1}$ ou baisse de 1 (s'il n'est pas en $i = 0$) avec probabilité $\pi_{i,i-1}$. Ce modèle peut se voir comme l'évolution d'une population au cours du temps ou un seul individu à la fois meurt ou bien naît, avec une probabilité et un taux qui ne dépend que de la taille de la population. C'est pourquoi de tels processus markoviens portent le nom de processus de naissance et de mort. Cependant de tels modèles recouvrent de nombreuses situations qui ne sont pas nécessairement liées à la biologie (voir les files d'attentes dans la section suivante).

**Remarque II.4.6** Si pour un état $K \geq 1$, $q_{K,K+1} = 0$, on voit qu'un processus issu de $i \in \{0, \dots, K\}$, ne visitera jamais les états $\{K + 1, K + 2, \dots\}$. Il peut se voir alors comme un processus de naissance et de mort dont l'espace d'états est $\{0, \dots, K\}$, et de générateur infinitésimal $(q_{i,j})_{0\leq i,j\leq K}$. □

Comme déjà mentionné, la chaîne du squelette associée à un processus de naissance et de mort est une chaîne de Markov de naissance et de mort qui ont été traitée de façon très complète au chapitre sur les chaînes de Markov. La seule question générale propre aux processus markoviens en temps continu, qui reste à régler est l'explosion éventuelle des processus de naissance et de mort.

EXERCICES.

**Exercice II.12** Soient $(\mathbf{e}_n, n \geq 0)$ une suite de v.a. exponentielles indépendantes de paramètres respectifs $(q_n, n \geq 0)$, définies sur un espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$. Montrer que l'on a l'alternative suivante.

**(i)** $\mathbf{P}(\sum_{n\geq 0} \mathbf{e}_n = \infty) = 1$ si $\sum_{n\geq 0} \frac{1}{q_n} = \infty$.

**(ii)** $\mathbf{P}(\sum_{n\geq 0} \mathbf{e}_n < \infty) = 1$ si $\sum_{n\geq 0} \frac{1}{q_n} < \infty$.

*(On notera que* $\mathbf{E}[\sum_{n\geq 0}\mathbf{e}_n]=\sum_{n\geq 0}\frac{1}{q_n}$). □

**Exercice II.13** On considère un processus de naissance pure associé aux paramètres $a_i>0$, $i\geq 0$. On suppose que toutes les variables sont définies sur un même espace de probabilité $(\Omega,\mathscr{F},\mathbf{P})$. On note $\mathbf{X}=(X_t)_{t\in\mathbb{R}_+}$ un tel processus. On suppose que $\mathbf{P}(X_0=0)=1$.

1. Montrer à l'aide de l'exercice II.12 précédent que le processus explose ssi la série de terme général $1/a_i$ est convergente et que l'on a
$$\mathbf{P}(\zeta<\infty)=1 \quad \text{si} \quad \sum_{i\geq 0}\frac{1}{a_i}<\infty \quad \text{et} \quad \mathbf{P}(\zeta=\infty)=1 \quad \text{si} \quad \sum_{i\geq 0}\frac{1}{a_i}=\infty.$$
2. On suppose que si $i\neq j$ alors $a_i\neq a_j$. Montrer que
$$\mathbf{P}(X_t=j)=\frac{1}{a_j}\sum_{0\leq i\leq j}(a_i)^n e^{-a_it}\prod_{\substack{0\leq k\leq j\\ k\neq i}}\left(1-\frac{a_k}{a_i}\right)^{-1}.$$
3. Calculer $\mathbf{P}(X_t=j)$ lorsque $a_i:=c.i$, pour une certaine constante $c$.
4. Pour tout $i\in\mathbb{N}$, on pose $v(i)=a_i$. Pour tout $t\in\mathbb{R}_+$, on pose également $A_t=\int_0^t v(X_s)\,ds$ et $\tau_t=\inf\{s\in\mathbb{R}_+ : A_s>t\}$ avec la condition habituelle $\inf\emptyset=\infty$. Montrer que $(X_{\tau_t})_{t\in\mathbb{R}_+}$ est un processus de Poisson de paramètre 1 issu de 0. □

**Exercice II.14** On considère un processus de naissance et de mort associés aux paramètres $a_i=i.a$, $i\geq 0$ et $b_i=i.b$, $i\geq 1$, où $a,b$ sont deux réels strictement positif. On suppose que toutes les variables sont définies sur un même espace de probabilité $(\Omega,\mathscr{F},\mathbf{P})$. On note $\mathbf{X}=(X_t)_{t\in\mathbb{R}_+}$ un tel processus. On suppose que $\mathbf{P}(X_0=i_0)=1$. On pose $v(r,t)=\mathbf{E}[r^{X_t}]$, $r\in[0,1]$ et $t\in\mathbb{R}_+$.

1. On suppose que $i_0=1$. En utilisant les équationde Kolmogorov rétrogrades et directe montrer que $t\mapsto v(r,t)$ satisfait une équation différentielle que l'on explicitera et que l'on résoudra.
2. Trouver une forme explicite à $v(r,t)$ (il faut distinguer les cas $a=b$ et $a\neq b$). □

### II.4.e Quelques modèles de files d'attentes.

Les principaux sytèmes de files d'attentes ont été décrits et classifiés par l'ingénieur des télécommunication danois K. Erlang et le mathématicien anglais D. Kendall de la manière suivante : on suppose que des clients entrent les uns après les autres dans une salle d'attente ; les temps entre les arrivées successives des clients sont appelés *temps d'interarrivée* ; dans cette salle d'attente se trouvent un certain nombre $K$ de guichets derrière lesquels des agents servent les clients (un seul client à la fois par guichet) ; on suppose que la salle d'attente a une contenance maximale de $C$ clients possibles ; si la salle d'attente est pleine et qu'un nouveau client veut rentrer, celui-ci est tout simplement refoulé (et on n'en tient plus compte) ; chaque client est appelé à un guichet dans un ordre fixé à l'avance qui est la *règle de priorité* ; le temps mis pour servir un client est appelé *temps de service* ; lorsqu'un client est servi, on considère qu'il quitte le guichet où il était servi et que l'agent appelle instantanément un nouveau client ; par ailleurs on suppose qu'en dehors de l'attente due aux services des autres clients, chaque client se comporte indépendamment des autres ; cela conduit à supposer que les temps d'interarrivée sont des variables indépendantes, ainsi que les temps de service. La loi d'une file d'attente de ce type est donc complètement décrite par cinq symboles :

Symbole 1 / Symbole 2 / K / C / Règle de priorité ,

qui signifient les choses suivantes.

— Le premier symbole décrit la loi des temps d'interarrivée. On suppose que ces temps forment un suite de variables aléatoires indépendantes de même loi. Cela n'est pas tout à fait réaliste car la statistiques du nombre de clients qui arrivent peut changer au cours du temps (on peut penser aux heures de pointe dans une poste). Mais il est raisonnable de penser que cette fluctuation de la statistiques du nombre de clients se présentant dans la salle d'attente varie assez lentement. On supposera ici que sur le laps de temps où la file d'attente est étudiée, la statistique du nombre d'arrivées de clients est constante. Comme les clients se comportent de façon indépendante les uns des autres, il est naturel de penser qu'un bon modèle pour les temps d'arrivée est un processus de Poisson ; dans ce cas, la loi d'interarrivée est une loi exponentielle. On signale cette hypothèse en prenant pour "Symbole 1" la lettre M, comme "markovien". Si la loi des temps d'interarrivée n'est pas exponentielle, on le signale par les lettres $GI$, comme "Généralisé et Indépendant" . On note $(A_n, n \geq 1)$ la suite des temps d'interarrivée. Les temps d'arrivée sont donc donnés par la suite $(T_n, n \geq 1)$ définie par

$$T_n = A_1 + \ldots + A_n \;, \quad n \geq 1.$$

— Le second symbole décrit la loi des temps de service de chaque client. On suppose que ces temps de service sont indépendants et de même loi : cela signifie que tous les clients sont statistiquement du même type. Cela n'est pas toujours réaliste, mais nous nous contentons de décrire les modèles de files d'attente les plus simples. Lorsque la loi des temps de service est une loi exponentielle, alors on signale cette hypothèse en prenant pour "Symbole 2" la lettre M, comme "markovien". En effet nous verrons que dans ce cas, la file d'attente est complètement décrite par un processus markovien. Il faut cependant noter que cette hypothèse n'est pas très réaliste, c'est pour cela que sont souvent considérées des lois plus générales : on signale cela en prenant pour s "Symbole 2" les lettres $GI$. Dans ce cas, la file d'attente est plus délicate à étudier mais elle peut cependant être analysée par des méthodes markoviennes.

— Comme déjà mentionné, le nombre $K$ représente le nombre de guichets. Mais signalons qu'il peut être pris égal à $\infty$, éventuellement.

— Comme déjà mentionné le nombre $C$ représente la contenance maximale de la salle. Signalons qu'il peut être pris égal à $\infty$ ou à $0$.

— Il existe de nombreuse règles de priorité ; nous n'en mentionnerons que deux. La règle FIFO, pour "First In First Out". Pour cette règle, les clients sont servis dans l'ordre de leur arrivée c'est-à-dire que dès qu'un guichet se libère, le client dans la salle d'attente qui a attendu le plus longtemps s'y dirige (instantanément) et est servi. On peut aussi considèrer la règle LIFO, pour "Last In First Out" ; selon cette règle, dès qu'un nouveau client rentre, il supplante un client en train d'être servi, le guichetier auquel il se présente se met à le servir ; il faut préciser ce qui arrive au client supplanté : on impose que le client supplanté revienne dans la salle d'attente mais il y figure avec la priorité sur les autres clients attendant, qui conservent leur ordre de priorité entre eux, par ailleurs. Cette règle est appelée "LIFO preemptive resume". Lorsqu'il y a plusieurs guichets et que l'on adopte la règle LIFO preemptive resume, il faut préciser comment le nouveau client choisit son nouveau guichet. La règle LIFO preemptive resume peut sembler totalement injuste au premier abord, mais le nouveau client qui est passé devant tout le monde sera peut être supplanté si son temps de service est trop long ; cette règle a tendance à pénaliser les clients au temps de service long ; elle est parfois globalement plus efficace que la règle FIFO.

Il existe une littérature abondante sur les files d'attente, qui donne lieu à des développements mathématiques compliqués. Nous nous restreignons aux cas suivants :

(M/M/1/$\infty$/FIFO), (M/M/K/0/-), (M/M/K/$\infty$/FIFO) et (M/GI/1/$\infty$/FIFO).

Par ailleurs, il n'est pas particulièrement facile de *prouver* que les files d'attentes markoviennes sont effectivement markoviennes. Nous nous contenterons ici de nous en convaincre, sans beaucoup de détails, et d'en calculer le générateur infinitésimal. Pour ce calcul nous allons utiliser les deux résultats suivants sur les variables exponentielles.

**Lemme II.4.7** **(Lemme des réveils)** *Soient $(\mathbf{e}_n, n \geq 0)$ une suite de v.a. exponentielles indépendantes de paramètres respectifs $(q_n, n \geq 0)$, définies sur un espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$. Alors, $\inf_{n\geq 0} \mathbf{e}_n$ suit une loi exponentielle de paramètre $\sum_{n\geq 0} q_n$. De plus, si*

$$0 < \sum_{n\geq 0} q_n < \infty ,$$

*alors, $\mathbf{P}$-p.s. il existe un unique indice aléatoire $N \in \mathbb{N}$ tel que $\mathbf{e}_N = \inf_{n\geq 0} \mathbf{e}_n$. Par ailleurs $N$ est indépendante de $\inf_{n\geq 0} \mathbf{e}_n$ et on a*

$$\forall k \geq 0 , \quad \mathbf{P}(N = k) = \frac{q_k}{\sum_{n\geq 0} q_n} .$$

**Preuve :** nous laissons la preuve de ce lemme en exercice. ■

• *File d'attente (M/M/1/∞/FIFO).* Il y a un unique guichet et une salle d'attente infinie. On note $a$ le paramètre des temps d'interarrivée et $b$ le paramètre des temps de service. Les client sont numérotés par leur ordre d'arrivée. On note $X_t$ le nombres de clients dans la salle d'attente au temps $t$ et on note $(Y_n)_{n\geq 0}$ les états successifs, c'est-à-dire le squelette. On note $J_1$ le premier temps de changement d'état.

— Si $X_0 = 0$, le premier saut d'effectue lorsque le premier client arrive. Alors on a $Y_1 = 1$ et et $J_1$ est une variable exponentielle de paramètre $a$. On doit donc avoir $q(0) = a = -q_{0,0}$. Comme $\pi_{0,1} = \mathbf{P}_0(Y_1 = 1)$, on en déduite que $\pi_{0,i} = q_{0,i} = 0$ dès que $i \geq 2$. On a donc

$$q_{0,0} = -a \ , \ q_{0,1} = a \quad \text{et donc } q_{0,i} = 0 \text{ dès que } i \geq 2.$$

— Supposons que $X_0 = i \geq 1$. Alors le processus change d'état dès qu'un client entre ou dès qu'un client de la salle d'attente est servi : $J_1$ est donc le minimum d'un temps d'arrivée et d'un temps de service, qui sont des exponentielles indépendantes de paramètres respectifs $a$ et $b$. $J_1$ est donc une loi exponentielle de paramètre $a + b = -q_{i,i}$. La probabilité qu'un client arrive avant qu'un client de la salle d'attente soit appelé pour être servi est, d'après le lemme des réveil, $\mathbf{P}(Y_1 = i + 1) = a/(a + b)$ et la probabilité contraire est donc $\mathbf{P}(Y_1 = i - 1) = b/(a + b)$, de plus la durée d'attente $J_1$ et $Y_1$ sont indépendants. On en déduit facilement que $\pi_{i,j} = 0$, dès que $|i - j| \geq 0$ et $\pi_{i,i+1} = 1 - \pi_{i,i-1} = a/(a + b)$. on a donc

$$q_{i,i} = a + b \ , \ q_{i,i+1} = a \ , \ q_{i,i-1} = b \ \text{ et donc } q_{i,j} = 0 \text{ dès que } |i - j| \geq 2.$$

On a complétement calculé le générateur infinitésimal de la taille de la file d'attente, *si celle-ci suit une dynamique markovienne*. Pour ce convaincre de cela, il faut penser à la propriété "d'oubli" des lois exponentielles : à chaque modification de la file, par arrivée ou fin de service, les exponentielles modèlisant le temps d'attente jusqu'à l'arrivée du prochain client ou le temps de service restant au client actuellement servi, sont rafraîchies et statistiquement "comme neuves". Ce n'est ici pas une situation compliquée et on pourrait tout à fait rigoureusement justifier ce fait.

On constate que cette file d'attente est un cas particulier de processus de naissance et de mort. Les questions que l'on peut se poser sur une telle file d'attente sont les suivantes.
∗ Est-ce-que la file d'attente s'engorge, c'est-à-dire : est-ce que $X_t$ tend vers l'infini ?
∗ Lorsque la file ne s'engorge pas, a-t-elle un régime stationnaire et si c'est le cas, lequel ?
∗ Lorsque l'on est en régime stationnaire, quel est le temps d'attente moyen d'un client avant d'être servi ?

• *File d'attente (M/M/K/0/-).* Ici, il n'y a pas de salle d'attente, il n'est donc pas nécessaire d'avoir une règle de priorité. Il vaut mieux voir cette file d'attente en termes de serveur téléphonique : on considère que ce serveur téléphonique possède $K \geq 1$ canaux ; dès que quelqu'un appelle le serveur, il est connecté instantanément à son correspondant via un des canaux libres, qui est choisi uniformément au hasard ; les temps de service représentent alors les temps de conversation des clients avec leur interlocuteur ; lorsque les $K$ canaux sont occupés et que quelqu'un appelle le serveur, il est refoulé. Cette situation est celle, par exemple d'une borne relai de téléphones portables. On suppose que les temps d'interarrivées sont des variables exponentielles indépendantes de paramètre $a$ et que les temps de service/conversation sont des variables exponentielles indépendantes de paramètre $b$.

On note $X_t$ le nombre de canaux occupés au temps $t$. On admet qu'il s'agit d'un processus markovien dont l'espace d'état est alors $\{0, 1, \ldots, K\}$. On veut calculer son générateur infinitésimal $G = (q_{i,j})_{0 \leq i,j \leq K}$. Pour cela on se donne $i \in \{0, 1, \ldots, K\}$ et on suppose que $X_0 = i$.

— Si $i = 0$, le premier saut de $X$ à lieu au temps $A_1$ qui est le temps d'arrivée du premier appel. Donc $J_1 = A_1$ suit une loi exponentielle de paramètre $a$ et on a $Y_1 = X_{J_1} = 1$. Cela donne donc

$$-q_{0,0} = a = q_{0,1} \quad \text{et} \quad q_{0,j} = 0 \quad \text{si} \quad j \geq 2 \ .$$

— On suppose que $1 \leq i \leq K - 1$. On note $B_1, \ldots, B_i$ les temps de service/conversation des $i$ clients connectés au temps $t = 0$. Ce sont des exponentielles indépendantes de paramètre $b$, indépendantes des temps d'interarrivée $(A_{n+i}, n \geq 1)$. Le premier temps de saut de $X$ est donc

$$J_1 = \min(A_{1+i}, B_1, \ldots B_i) \ .$$

Le lemme des réveils implique que $J_1$ est une exponentielle de paramètre $a + ib$ qui est indépendante des deux événements $\{A_{1+i} = J_1\}$ et $\{A_{1+i} > J_1\}$ qui sont tels que

$$\mathbf{P}(A_{1+i} = J_1) = \frac{a}{a+ib} = 1 - \mathbf{P}(A_{1+i} > J_1) \ .$$

On pose $Y_1 = X_{J_1} = i + 1$ si $A_{1+i} = J_1$ et $Y_1 = X_{J_1} = i - 1$ si $A_{1+i} > J_1$. On voit donc que $Y_1$ est indépendant de $J_1$ et que

$$\pi_{i,i+1} = \mathbf{P}(Y_1 = i + 1) = \frac{a}{a+ib} = 1 - \pi_{i,i-1} = 1 - \mathbf{P}(Y_1 = i - 1) \ ,$$

ce qui implique que $\pi_{i,j} = 0$ si $|i - j| \neq 1$. On a donc

$$q_{i,i+1} = a \ , \ q_{i,i-1} = ib \ , \ q_{i,i} = -(a + ib) \quad \text{et} \quad q_{i,j} = 0 \quad \text{si} \quad |i - j| \geq 2 \ .$$

— On suppose que $i = K$, alors le premier saut de $X$ à lieu au temps $J_1 = \min(B_1, \ldots, B_K)$ qui suit une loi exponentielle de paramètre $Kb$. De plus on a $Y_1 = X_{J_1} = K - 1$. Cela implique que

$$-q_{K,K} = bK = q_{K,K-1} \quad \text{et} \quad q_{K,j} = 0 \quad \text{si} \quad j \geq 2 \ .$$

On voit que $X$ est un processus de naissance et de mort dont l'espace d'état est fini. Les questions que l'on peut se poser sont les suivantes.

$*$ Le processus $X$ a-t-il un régime stationnaire ? Si c'est le cas, lequel ?

$*$ Quelle est la proportion moyenne de temps durant lequel le serveur est saturé ?

$*$ En régime stationnaire, combien de clients sont refoulés en moyenne par unité de temps ?

• *File d'attente (M/M/K/$\infty$/FIFO).* Il s'agit du même sytème que le précédent, excepté qu'il y a une capacité d'attente infinie, si bien qu'aucun appel n'est refoulé. On note toujours $X_t$ le nombre de clients en service ou présents dans la salle d'attente. On admet que c'est un processus markovien. L'espace d'états est $\mathbb{N}$. On suppose que les temps d'interarrivées sont des variables exponentielles indépendantes de paramètre $a$ et que les temps de service/conversation sont des variables exponentielles indépendantes de paramètre $b$. On remarque que lorsque $X_t \geq K$, les $K$ canaux sont occupés et il y a $X_t - K$ appels/clients en attente. En revanche, lorsque $X_t < K$, il n'y a aucun appel/client en attente. On veut calculer le générateur infinitésimal $G = (q_{i,j})_{i,j \in \mathbb{N}}$ de $\mathbf{X}$ Pour cela on se donne $i \in \mathbb{N}$ et on suppose que $X_0 = i$.

— Si $i = 0$, le premier saut de $\mathbf{X}$ à lieu au temps $A_1$ qui est le temps d'arrivée du premier appel. Donc $J_1 = A_1$ suit une loi exponentielle de paramètre $a$ et on a $Y_1 = X_{J_1} = 1$. Cela donne donc

$$-q_{0,0} = a = q_{0,1} \quad \text{et} \quad q_{0,j} = 0 \quad \text{si} \quad j \geq 2 \ .$$

— On suppose que $1 \leq i \leq K - 1$. On note $B_1, \ldots, B_i$ les temps de service/conversation des $i$ clients connectés au temps $t = 0$. Ce sont des exponentielles indépendantes de paramètre $b$, indépendantes des temps d'interarrivée $(A_{n+i}, n \geq 1)$. Le premier temps de saut de $X$ est donc

$$J_1 = \min(A_{1+i}, B_1, \ldots B_i) \ .$$

Le lemme des réveils implique que $J_1$ est une exponentielle de paramètre $a + ib$ qui est indépendante des deux événements $\{A_{1+i} = J_1\}$ et $\{A_{1+i} > J_1\}$, qui sont tels que

$$\mathbf{P}(A_{1+i} = J_1) = \frac{a}{a + ib} = 1 - \mathbf{P}(A_{1+i} > J_1) \ .$$

On pose $Y_1 = X_{J_1} = i + 1$ si $A_{1+i} = J_1$ et $Y_1 = X_{J_1} = i - 1$ si $A_{1+i} > J_1$. On voit donc que $Y_1$ est indépendant de $J_1$ et que

$$\pi_{i,i+1} = \mathbf{P}(Y_1 = i + 1) = \frac{a}{a + ib} = 1 - \pi_{i,i-1} = 1 - \mathbf{P}(Y_1 = i - 1) \ ,$$

ce qui implique que $\pi_{i,j} = 0$ si $|i - j| \neq 1$. On a donc

$$q_{i,i+1} = a \ , \ q_{i,i-1} = ib \ , \ q_{i,i} = -(a + ib) \quad \text{et} \quad q_{i,j} = 0 \quad \text{si} \quad |i - j| \geq 2 \ .$$

— On suppose que $i \geq K$, alors le premier saut de $X$ à lieu au temps $J_1 = \min(A_{1+i}, B_1, \ldots, B_K)$ qui suit une loi exponentielle de paramètre $a + Kb$, qui est indépendante des deux événements $\{A_{1+i} = J_1\}$ et $\{A_{1+i} > J_1\}$, qui sont tels que

$$\mathbf{P}(A_{1+i} = J_1) = \frac{a}{a + Kb} = 1 - \mathbf{P}(A_{1+i} > J_1) \ .$$

On pose $Y_1 = X_{J_1} = i + 1$ si $A_{1+i} = J_1$ et $Y_1 = X_{J_1} = i - 1$ si $A_{1+i} > J_1$. On voit donc que $Y_1$ est indépendant de $J_1$ et que

$$\pi_{i,i+1} = \mathbf{P}(Y_1 = i + 1) = \frac{a}{a + Kb} = 1 - \pi_{i,i-1} = 1 - \mathbf{P}(Y_1 = i - 1) \ ,$$

ce qui implique que $\pi_{i,j} = 0$ si $|i - j| \neq 1$. On a donc

$$q_{i,i+1} = a \ , \ q_{i,i-1} = Kb \ , \ q_{i,i} = -(a + Kb) \quad \text{et} \quad q_{i,j} = 0 \quad \text{si} \quad |i - j| \geq 2 \ .$$

## EXERCICES.

**Exercice II.15** On considère une file d'attente $(M/M/1/\infty/FIFO)$. Les temps d'interarrivées des clients sont des variables exponentielles indépendantes de paramètre noté $a$ et les temps de services sont des exponentielles notées $b$. On suppose que $ab > 0$. On note $X_t$ le nombre de clients dans la file à l'instant $t$.

1. Montrer que $\mathbf{X}$ est irréductible et réversible.
2. A quelles conditions sur $a$ et $b$ est-il transient ? Que cela signifie-t-il ?
3. A quelles conditions sur $a$ et $b$ $\mathbf{X}$ est-il récurrent positif ? Quelle est sa probabilité invariante, notée $m$.
4. On suppose que $a$ et $b$ sont tels que la file est récurrente positive. On suppose que l'on est en régime stationnaire, c'est-à-dire que $X_0$ suive la loi $m$. On marque un client lorsqu'il arrive dans la file et on note $L$ sa durée d'attente totale avant d'être complètement servi. Trouver la loi de $L$ et calculer sa moyenne. □

**Exercice II.16** On considère une file d'attente $(M/M/K/0/-)$. Les temps d'interarrivées des clients sont des variables exponentielles indépendantes de paramètre noté $a$ et les temps de services sont des exponentielles notées $b$. On suppose que $ab > 0$. On note $X_t$ le nombre de guichets occupés à l'instant $t$.

1. Calculer la probabilité invariante, notée $m$.
2. On note $A_t = \int_0^t \mathbf{1}_{X_s=K} ds$ qui est le temps pendant lequel l'agence est saturée entre les instants 0 et $t$. Montrer que
$$\lim_{t\to\infty} \frac{1}{t} A_t = \lim_{t\to\infty} \frac{1}{t} \mathbf{E}[A_t] \ .$$
3. On note $R_t$ le nombre de clients refoulés par l'agence entre les instants 0 et $t$. Calculer la loi conditionnelle de $R_t$ sachant $A_t$.
4. Calculer $\lim_{t\to\infty} \frac{1}{t}\mathbf{E}[R_t]$ en fonction de $a$, $b$ et $N$. On sup On suppose que $b = 1$ et que les temps d'interarrivées des clients sont $K$ fois plus courts en moyenne que les temps de services. On pose $c(K) = \lim_{t\to\infty} \frac{1}{t}\mathbf{E}[R_t]$. Montrer que
$$c(K) \ \sim \ \sqrt{\frac{K}{8\pi}} \ .$$

□

**Exercice II.17** On considère une file d'attente $(M/M/K/\infty/FIFO)$. Les temps d'interarrivées des clients sont des variables exponentielles indépendantes de paramètre noté $a$ et les temps de services sont des exponentielles notées $b$. On suppose que $ab > 0$. On note $X_t$ le nombre de clients en attente ou en train d'être servis.

1. Déterminer en fonction de $a, b$ et $K$ si le processus est transient, récurrent nul ou récurrent positif : on calculera ses loi invariantes. Dans le cas récurrent positif on notera $m$ son unique probabilité invariante, que l'on calculera.
2. On note $N_t$ le nombre de guichets occupés à l'instant $t$. On se place dans le cas récurrent positif : montrer que $\mathbf{P}$-presque sûrement
$$\lim_{t\to\infty} \frac{1}{t} \int_0^t N_s ds = c \ ,$$
où $c$ est une constante que l'on exprimera en fonction de $m$ : interpréter cette quantité. □

**Exercice II.18** On considère une file d'attente $(M/M/K/C/FIFO)$, avec $0 < K, C < \infty$. Les temps d'interarrivées des clients sont des variables exponentielles indépendantes de paramètre noté $a$ et les temps de services sont des exponentielles notées $b$. On suppose que $ab > 0$. On note $X_t$ le nombre de clients en attente ou en train d'être servis.

1. Calculer la mesure de probabilité invariante.
2. Calculer combien de clients sont refoulés en moyenne par unité de temps. □

### II.4.f Processus de branchement en temps continu.

Nous allons examiner le cas d'une population qui évolue de la façon suivante.

- Chaque individu a une durée de vie indépendante et les durées de vies suivent la même loi. Un choix naturel est une loi exponentielle, dont on notera $c \in ]0, \infty[$ le paramètre. Ce choix est naturel car la loi exponentielle modélise assez correctement la durée de vie d'individus (ou de machines) en fonctionnement normal (pas de maladie infantile ou de période de rodage et pas de vieillissement).

- Chaque individu, en mourant donne naissance à un nombre aléatoire d'enfants qui suit une certaine loi de reproduction notée $\xi = (\xi(i))_{i\in\mathbb{N}}$, qui est donc une probabilité sur $\mathbb{N}$ : $\xi(i) \geq 0$, pour tout $i \in \mathbb{N}$ et $\sum_{i\in\mathbb{N}} \xi(i) = 1$. De plus chaque individu se reproduit indépendamment des autres.

On suppose qu'au temps $t = 0$, il y a $Z_0$ ancêtres dans la population et on note $Z_t$ le nombre d'individus vivants à l'instant $t$. On suppose que $(Z_t)_{t\in\mathbb{R}_+}$ est un processus de Markov, ce qui peut se justifier heuristiquement par la propriété d'oubli des durées de vie exponentielles. Signalons que si les durées de vies ne suivent pas une loi exponentielle, $(Z_t)_{t\in\mathbb{R}_+}$ n'est en général pas markovien. *On admet* qu'un modèle markovien $\mathbf{Z} = (Z_t)_{t\in\mathbb{R}_+}$ correspond à la situation que nous avons décrite et nous allons caractériser sa loi. Pour cela, on peut faire les observations suivantes.

— Si la population est éteinte à l'instant $t$, c'est-à-dire si $Z_t = 0$, aucun individu ne va naître spontanément sans parent : $Z_{s+t} = 0$, pour tout $s \in \mathbb{R}_+$. Donc $0$ est un élément absorbant.

— On ne peut pas exclure que la population explose en un temps fini $t$ (aléatoire), c'est-à-dire $\lim_{s\uparrow t} Z_s = \infty$. On convient alors que la population reste infinie après ce temps d'explosion et qu'avant cette explosion elle est de taille finie. Autrement dit $\infty$ est un point cimetière naturel qui est absorbant. L'espace d'états de $\mathbf{Z}$ est donc $\mathbb{N}_\infty = \mathbb{N} \cup \{\infty\}$, avec $0$ et $\infty$ comme éléments absorbants, $\infty$ étant le point cimetière, c'est-à-dire le seul état que l'on ne puisse atteindre sans effectuer une infinité de sauts. Autrement dit $(Z_t)_{t\in\mathbb{R}_+}$ est un processus minimal dont le générateur infinitésimal, noté $G$ devrait être caractérisé par $c$ et $\xi$.

— (*Propriété de branchement*) On considère deux populations *indépendantes* se reproduisant selon le même mécanisme (fixé par le paramètre des temps de vie $c$ et la loi de reproduction $\xi$). La taille de la première est donnée par le processus $\mathbf{Z} = (Z_t)_{t\in\mathbb{R}_+}$ et la taille de la seconde est donnée par $\mathbf{Z}' = (Z'_t)_{t\in\mathbb{R}_+}$. La réunion des deux populations forme une nouvelle population qui se reproduit selon le même mécanisme, c'est-à-dire que $\mathbf{Z}'' = (Z_t + Z'_t)_{t\in\mathbb{R}_+}$ est encore un processus markovien minimal dont le générateur infinitésimal est $G$.

On note $P_t = (p_t(i,j))_{i,j\in\mathbb{N}_\infty}$, $t \in \mathbb{R}_+$, le semi-groupe du processus de branchement de paramètre de temps de vie $c$ et de loi de reproduction $\xi$ (c'est-à-dire le semi-groupe associé au générateur $G$). On reformule en termes du semi-groupe la propriété de branchement : soient $\mu$ et $\nu$ deux probabilités sur $\mathbb{N}_\infty$. On définit la mesure convolée $\mu * \nu$ en posant

$$\forall i \in \mathbb{N}_\infty\,, \quad \mu * \nu(i) = \sum_{\substack{k,\ell\in\mathbb{N}_\infty\\ k+\ell=i}} \mu(k)\nu(\ell)$$

avec la convention que $n + \infty = \infty$. On voit $\mu * \nu$ est une mesure de probabilité sur $\mathbb{N}_\infty$ et on vérifie facilement que

$$\mu * \nu(\infty) = 1 - (1 - \mu(\infty))(1 - \nu(\infty))\ .$$

On vérifie également que si $Y$ et $Z$ sont deux variables à valeurs dans $\mathbb{N}_\infty$, indépendantes, $Y$ ayant pour loi $\mu$ et $Z$ ayant pour loi $\nu$, alors $Y+Z$ a pour loi $\mu * \nu$. On voit donc que la propriété de branchement implique que $p_t(i,\cdot) * p_t(i',\cdot) = p_t(i+i',\cdot)$, pour tout $t \in \mathbb{R}_+$ et pour tous états $i, i'$. *Nous allons prendre cette propriété comme définition des processus de branchement en temps continu.*

**Définition II.4.8** On note $\mathbb{N}_\infty = \mathbb{N} \cup \{\infty\}$, où $\infty$ est vu comme un point cimetière. On considère un processus de Markov minimal sur $\mathbb{N}$ de générateur $G$. On note $P_t = (p_t(i,j))_{i,j\in\mathbb{N}}$, $t \in \mathbb{R}_+$ le semi-groupe minimal généré par $G$ et on l'étend à $\mathbb{N}_\infty$ en posant $p_t(\infty, i) = 0$, si $i \in \mathbb{N}$, $p_t(\infty,\infty) = 1$ et $p_t(i,\infty) = 1 - p_t(i,\mathbb{N})$. On dit que $G$ est le générateur d'un processus de branchement **ssi** le semi-groupe satisfait la *propriété de branchement*

$$\forall i, i' \in \mathbb{N}_\infty\,,\ \forall t \in \mathbb{R}_+\,, \quad p_t(i,\cdot) * p_t(i',\cdot) = p_t(i+i',\cdot). \tag{II.69}$$

On observe que cette définition ne fait aucune mention de la loi de reproduction ni du paramètre des temps de vie. □

**Proposition II.4.9** *Soit* $\big(\Omega\,;\,\mathscr{F}\,;\,(\mathscr{F}_t)_{t\in\mathbb{R}_+}\,;\,\mathbf{Z} = (Z_t)_{t\in\mathbb{R}_+}\,;\,G = (q_{i,j})_{i,j\in\mathbb{N}}\,;\,\mathbf{P}_\mu\,,\,\mu \in \mathscr{M}_1(\mathbb{N})\big)$, *un* processus de branchement. *Alors les assertions suivantes sont vérifiées.*

$(i)$ *Il existe une mesure de probabilité sur* $\mathbb{N}$, *notée* $\xi = (\xi(i))_{i\in\mathbb{N}}$ *telle que* $\xi(1) = 0$ *et il existe un réel* $c \in ]0,\infty[$, *qui satisfont*

$$\forall i,j \in \mathbb{N}\,, \quad q_{i,j} = c\cdot i\,\xi(j-i+1)\mathbf{1}_{\{j-i+1\geq 0\}} \text{ si } i \neq j \quad \text{et} \quad q_{i,i} = -c\cdot i\,.$$

*On observe que* $q_{0,0} = 0$, *ce qui implique que* $0$ *est absorbant. On appelle* $c$ *le* paramètre des temps de vie *et* $\xi$ *la* loi de reproduction.

$(ii)$ *On adopte les conventions suivantes :* $r^0 = 1$ *et* $r^\infty = 0$, *pour tout* $r \in [0,1]$. *On pose ensuite*

$$\forall t \in \mathbb{R}_+\,,\ \forall r \in [0,1]\,, \quad v(r,t) = \mathbf{E}_1\big[r^{Z_t}\big]\,.$$

*Alors*

$$\forall i \in \mathbb{N}\,,\ \forall t \in \mathbb{R}_+\,, \quad \forall r \in [0,1]\,, \quad \mathbf{E}_i\big[r^{Z_t}\big] = v(r,t)^i\,. \tag{II.70}$$

*De plus pour tout* $r \in [0,1]$, *la fonction* $t \in \mathbb{R}_+ \mapsto v(r,t)$ *est continument dérivable et elle est solution de l'équation différentielle suivante*

$$\frac{\partial}{\partial t} v(r,t) = c\big(\varphi_\xi(v(r,t)) - v(r,t)\big)\,,\ t \in \mathbb{R}_+ \quad \text{et} \quad v(r,0) = r\,, \tag{II.71}$$

*où* $\varphi_\xi$ *est la fonction génératrice de* $\xi$, *c'est-à-dire que* $\varphi_\xi(r) = \sum_{i\in\mathbb{N}} \xi(i) r^i$.

$(iii)$ *Pour toute loi d'entrée* $\mu$, *pour tout* $r \in [0,1]$ *et pour tous* $s,t \in \mathbb{R}_+$, *on a*

$$\mathbf{P}_\mu\text{-}p.s. \quad \mathbf{E}_\mu\big[r^{Z_{s+t}}\,\big|\,\mathscr{F}_s\big] = v(r,t)^{Z_s}\,.$$

**Preuve :** pour toute loi $\pi$ sur $\mathbb{N}_\infty$, on note $\varphi_\pi$ sa fonction génératrice : $r \in [0,1] \mapsto \sum_{i\in\mathbb{N}} \pi(i) r^i$ (on rappelle la convention que $r^\infty = 0$). Soient $\mu$ et $\nu$, deux lois sur $\mathbb{N}_\infty$. On vérifie facilement que pour tout $r \in [0,1]$, on a $\varphi_{\mu*\nu}(r) = \varphi_\mu(r).\varphi_\nu(r)$, $r \in [0,1]$. La propriété de branchement (II.69) implique que

$$\forall i,j \in \mathbb{N}\,,\ \forall t \in \mathbb{R}_+\,, \quad \forall r \in [0,1]\,, \quad \mathbf{E}_i\big[r^{Z_t}\big]\cdot\mathbf{E}_j\big[r^{Z_t}\big] = \mathbf{E}_{i+j}\big[r^{Z_t}\big]\,,$$

ce qui entraîne (II.70). Par ailleurs, il est clair que $(iii)$ est une conséquence du $(ii)$ et de la propriété de Markov. En décomposant ce qui se passe au premier temps de saut $J_1$, comme dans la preuve de l'équation de

Kolmogorov-Chapman directe (voir la preuve du théorème II.1.27 et la formule (II.16) page 167), on voit que pour tout $t \in \mathbb{R}_+$ et tous $i, j \in \mathbb{N}$, on a

$$e^{q(i)t} p_{i,j}(t) = \delta_{i,j} + \int_0^t q(i) e^{q(i)s} \mathbf{E}_i[p_s(Y_1, j)] \, ds \, . \tag{II.72}$$

En utilisant Fubini positif, on en déduit que

$$\begin{aligned} e^{q(i)t} v(r,t)^i &= e^{q(i)t} \mathbf{E}_i[r^{Z_t}] = e^{q(i)t} \sum_{j \in \mathbb{N}} p_t(i,j) r^j = r^i + \int_0^t ds \, q(i) e^{q(i)s} \mathbf{E}_i \Big[ \sum_{j \in \mathbb{N}} p_s(Y_1, j) r^j \Big] \\ &= r^i + \int_0^t ds \, q(i) e^{q(i)s} \mathbf{E}_i \big[ v(r,s)^{Y_1} \big] \, . \end{aligned} \tag{II.73}$$

On note $\xi$ la loi de $Y_1$ sous $\mathbf{P}_1$ : $\mathbf{P}_1(Y_1 = i) = \xi(i)$. Il est clair que $\xi(1) = 0$ ; on pose également $c = -q_{1,1} = q(1)$. L'équation précédente avec $i = 1$, implique que

$$e^{ct} v(r,t) = r + \int_0^t ds \, c e^{cs} \varphi_\xi(v(r,s)) \, .$$

On voit que $v(r, 0) = r$ et que pour tout $r \in [0, 1]$, $t \mapsto v(r, t)$ est continument dérivable. En dérivant cette équation on obtient facilement que

$$\frac{\partial}{\partial t} v(r,t) = c\big( \varphi_\xi(v(r,t)) - v(r,t) \big) \, , \ \ t \in \mathbb{R}_+ \, , \quad \text{et} \quad v(r,0) = r \, .$$

On en déduit que $\mathbf{E}_i[r^{Z_t}] = v(r,t)^i$ est dérivable en $t$ et notamment en $0+$. On cherche à calculer cette dérivée en justifiant le passage à la limite sous la somme suivante :

$$\tfrac{1}{t} \big( \mathbf{E}_i[r^{Z_t}] - r^i \big) = \tfrac{1}{t} (p_t(i,i) - 1) r^i + \sum_{j \in \mathbb{N} \backslash \{i\}} \tfrac{1}{t} p_t(i,j) r^j \, .$$

On suppose que $i \neq j$, alors on déduit de (II.72)

$$\tfrac{1}{t} p_t(i,j) \leq \tfrac{1}{t} (1 - e^{-q(i)t}) \leq q(i) \, .$$

On fixe $r \in [0, 1[$. On a donc $\frac{1}{t} p_t(i,j) \leq q(i) r^j$, pour tout $t > 0$ et $\sum_{j \in \mathbb{N}} q(i) r^j = q(i)/(1-r) < \infty$. Par convergence dominée, on a

$$\lim_{t \downarrow 0} \tfrac{1}{t} \sum_{j \in \mathbb{N} \backslash \{i\}} p_t(i,j) r^j = \sum_{j \in \mathbb{N} \backslash \{i\}} \lim_{t \downarrow 0} \tfrac{1}{t} p_t(i,j) r^j = \sum_{j \in \mathbb{N} \backslash \{i\}} q_{i,j} r^j \leq \frac{q(i)}{1-r} < \infty \, .$$

On en déduit donc que pour tout $r \in [0, 1[$, $t \mapsto \sum_{j \in \mathbb{N}} p_t(i,j) r^j$ est dérivable à droite en 0 et on a

$$\forall i \in \mathbb{N} \, , \ \forall r \in [0,1[ \, , \quad \lim_{t \downarrow 0} \tfrac{1}{t} \big( \mathbf{E}_i[r^{Z_t}] - r^i \big) = \sum_{j \in \mathbb{N}} q_{i,j} r^j. \tag{II.74}$$

Mais comme $\mathbf{E}_i[r^{Z_t}] = v(r,t)^i$ et que $t \mapsto v(r,t)$ est dérivable en $0+$, on a

$$\forall i \in \mathbb{N} \, , \ \forall r \in [0,1[ \, , \quad \lim_{t \downarrow 0} \tfrac{1}{t} \big( \mathbf{E}_i[r^{Z_t}] - r^i \big) = i \tfrac{\partial}{\partial t} v(r, 0+) v(r,0)^{i-1} = i \tfrac{\partial}{\partial t} v(r, 0+) r^{i-1} \, .$$

Or (II.74) appliqué à $i = 1$, donne

$$\frac{\partial}{\partial t} v(r, 0+) = \sum_{j \in \mathbb{N}} q_{1,j} r^j \, .$$

On obtient donc

$$\forall i \in \mathbb{N}\, , \; \forall r \in [0, 1[\, , \quad \sum_{j \in \mathbb{N}} q_{i,j} r^j = i \sum_{j \in \mathbb{N}} q_{1,j} r^{j+i-1} \, .$$

En identifiant les coefficients des deux séries entières, on obtient le point $(i)$, ce qui termine la preuve. ■

Les processus de branchement en temps continu sont très semblables aux processus de branchement à temps discret (qui sont des chaînes de Markov). On se pose les mêmes questions à leur propos : on introduit

$$T_{\mathrm{ex}} = \inf\{t \in \mathbb{R}_+ \; : \; Z_t = 0\} \, ,$$

avec la convention que $T_{\mathrm{ex}} = \infty$ si pour tout temps $t \in \mathbb{R}_+$, $Z_t \geq 1$. La variable $T_{\mathrm{ex}}$ est le *temps d'extinction*. Il est naturel de se poser les questions suivantes.

$*$ Quelle est la probabilité d'extinction $\mathbf{P}_\mu(T_{\mathrm{ex}} < \infty)$ ?

$*$ Sur l'événement de la non-extinction $\{T_{\mathrm{ex}} = \infty\}$, comment se comporte le processus ?

$*$ Quelles sont les lois de reproduction pour lesquelles le processus de branchement explose : $\mathbf{P}_\mu(\zeta < \infty) > 0$ ?

Le résultat suivant montre que tout processus de branchement à temps continu est une marche aléatoire changée de temps : le changement de temps explicité ci-dessous est appelé *changement de temps de Lamperti*.

**Proposition II.4.10** **(Transformée de Lamperti)** *On suppose que sur* $(\Omega, \mathscr{F}, \mathbf{P})$ *est défini un processus de branchement en temps continu de loi de reproduction* $\xi$ *et de paramètre des temps de vie* $c$*. On note ce processus par* $\mathbf{Z} = (Z_t)_{t \in \mathbb{R}_+}$*. Son générateur est donné par la proposition II.4.9* $(i)$ *et il est noté* $G = (q_{i,j})_{i,j \in \mathbb{N}}$*.*

*On définit la loi de probabilité* $\nu$ *sur* $\mathbb{Z}$ *en posant* $\nu(k) = \xi(k+1)$, $k \in \mathbb{N} \cup \{-1\}$ *et on suppose que* $\mathbf{X} = (X_t)_{t \in \mathbb{R}_+}$ *est une marche aléatoire sur* $\mathbb{Z}$ *de paramètre des durée d'attente* $c$ *et de loi de saut* $\nu$*. On suppose que* $\mathbf{P}(X_0 = Z_0 \geq 1) = 1$*. On pose* $T = \inf\{t \in \mathbb{R}_+ : X_t = 0\}$*, avec la convention habituelle :* $\inf \emptyset = \infty$*. On considère la marche* $(X_{t \wedge T})_{t \in \mathbb{R}_+}$ *stoppée en* $0$*, qui est un processus markovien (minimal) dont on note le générateur* $G' = (q'_{i,j})_{i,j \in \mathbb{N}}$*. Alors les assertions suivantes sont vérifiées.*

$(i)$ *Le générateur* $G'$ *de* $(X_{t \wedge T})_{t \in \mathbb{R}_+}$ *est donné par*

$$\forall i, j \in \mathbb{N}\, , \quad q'_{i,j} = \begin{cases} 0 & \textit{si } i = 0, \\ -c & \textit{si } i = j \geq 1, \\ c\nu(j-i)\mathbf{1}_{\{j > i\}} & \textit{si } i \geq 1 \textit{ et } i \neq j \end{cases}$$

*(ici on a restreint l'espace d'états à* $\mathbb{N}$ *car on a supposé* $X_0 \geq 1$*, et donc* $(X_{t \wedge T})_{t \in \mathbb{R}_+}$ *ne peut pas visiter les entiers strictement négatifs).*

$(ii)$ *On définit les changements de temps suivants :*

$$\forall t \in \mathbb{R}_+ \quad \begin{cases} A_t = \int_0^t \max(1, Z_s)\, ds & \textit{et} \quad \tau_t = \inf\{s \in \mathbb{R}_+ : A_s > t\} \\ A'_t = \int_0^t \frac{ds}{\max(1, X_{s \wedge T})} & \textit{et} \quad \tau'_t = \inf\{s \in \mathbb{R}_+ : A'_s > t\} \, , \end{cases}$$

*avec la convention habituelle* $\inf \emptyset = \infty$*, et on pose*

$$X'_t = \begin{cases} Z_{\tau_t} & \textit{si } \tau_t < \infty, \\ \partial & \textit{si } \tau_t = \infty. \end{cases} \qquad \textit{et} \qquad Z'_t = \begin{cases} X_{T \wedge \tau'_t} & \textit{si } \tau'_t < \infty, \\ \partial & \textit{si } \tau'_t = \infty. \end{cases}$$

*Alors on a*

$$(Z_t)_{t\in\mathbb{R}_+} \stackrel{\text{(loi)}}{=} (Z'_t)_{t\in\mathbb{R}_+} \quad \text{et} \quad (X_t)_{t\in\mathbb{R}_+} \stackrel{\text{(loi)}}{=} (X'_t)_{t\in\mathbb{R}_+} \, .$$

*Comme* $\mathbf{X}$ *est de générateur uniforme, il n'explose pas et donc* $\tau_t < \infty$ *pour tout* $t \in \mathbb{R}_+$.

**Preuve :** le premier point est une conséquence immédiate de la forme du générateur d'une marche aléatoire et de la discussion au début de la section II.3. La proposition II.3.1 implique que $\mathbf{X}' = (X'_t)_{t\in\mathbb{R}_+}$ est un processus minimal dont le générateur est donc $q_{i,j}/(1 \vee i)$, pour tous $i, j \in \mathbb{N}$. On constate alors que c'est le générateur $G'$ de la marche stoppée $(X_{t\wedge T})_{t\in\mathbb{R}_+}$. On peut prendre le changement de temps inverse et, à partir de la marche stoppée, obtenir un processus de branchement. ■

On voit donc qu'un processus de branchement $\mathbf{Z}$ et la marche stoppée qui lui est associée comme dans la proposition précédente, ont la même chaîne du squelette, qui est une marche aléatoire $\mathbf{Y} = (Y_n)_{n\geq 0}$ sur $\mathbb{Z}$ de loi de saut $\nu$ et stoppée en $0$.

Si $\nu(-1) = \xi(0) = 0$, comme $\nu(0) = \xi(1) = 0$, alors $Y_{n+1} - Y_n \geq 2$, pour tout $n \geq 0$ et on a $\lim_n Y_n = \infty$ presque sûrement, ce qui implique que le processus de branchement $\mathbf{Z}$ tend vers l'infini : $\mathbf{P}(\lim_{t\to\zeta} Z_t = \infty) = 1$.

Si $\nu(-1) = \xi(0) = 1$, alors la marche est absorbée et le processus de branchement tend vers $0$.

On suppose alors

$$\nu(-1) = \xi(0) \ \in ]0, 1[ \, , \tag{II.75}$$

Sous cette hypothèse, une marche aléatoire à temps discret de loi de saut $\nu$ est une chaîne de Markov irréductible : si on la stoppe en $0$, la seule classe d'irréductibilité de la marche stoppée est $\{0\}$ : le théorème de classification des états des chaînes de Markov montre que si $\mathbf{P}(Y_0 \geq 0) = 1$, alors $\mathbf{P}(\lim_n Y_n = 0) + \mathbf{P}(\lim_n Y_n = \infty) = 1$. Il n'y a donc que deux comportements asymptotiques pour le processus de branchement à temps continu :

$$\mathbf{P}\big(\exists t \in \mathbb{R}_+ \, , \ Z_t = 0\big) + \mathbf{P}\big(\lim_{t\uparrow\zeta} Z_t = \infty\big) = 1 \, , \tag{II.76}$$

Nous calculons ensuite $\mathbf{P}\big(\exists t \in \mathbb{R}_+ \, , \ Z_t = 0\big) = \mathbf{P}(T_{\mathtt{ex}} < \infty)$ qui est la probabilité d'extinction. Pour cela on fixe

$$\big(\Omega \, ; \, \mathscr{F} \, ; \, (\mathscr{F}_t)_{t\in\mathbb{R}_+} \, ; \, \mathbf{Z} = (Z_t)_{t\in\mathbb{R}_+} \, ; \, G = (q_{i,j})_{i,j\in\mathbb{N}} \, ; \, \mathbf{P}_\mu \, , \, \mu \in \mathscr{M}_1(\mathbb{N})\big) \, .$$

un processus de branchement de loi de reproduction $\xi$ et de paramètre des temps de vie $c$. Pour tout $j \in \mathbb{N}$, on pose $T_j = \inf\{n \in \mathbb{N} \, , \ Y_n = j\}$, avec la convention que $\inf \emptyset = \infty$. La propriété de Markov pour $\mathbf{Y}$ implique que

$$\forall r \in [0, 1[ \, , \ \forall i \geq 1 \, , \quad \mathbf{E}_i\big[r^{T_0}\big] = \mathbf{E}_i\big[r^{T_{i-1}}\big]\mathbf{E}_{i-1}\big[r^{T_0}\big] = \mathbf{E}_1\big[r^{T_0}\big]\mathbf{E}_{i-1}\big[r^{T_0}\big].$$

On pose $g(r) = \mathbf{E}_1[r^{T_0}]$, $r \in [0, 1[$. Ce qui précèdse montre que $\mathbf{E}_i\big[r^{T_0}\big] = g(r)^i$. La propriété de Markov au temps $1$ pour $\mathbf{Y}$ implique que

$$g(r) = \nu(-1)r + \sum_{i\geq 2} r\nu(i)\mathbf{E}_{i+1}\big[r^{T_0}\big] = \nu(-1)r + \sum_{k\geq 2} r\nu(i)g(r)^{i+1}.$$

On note $\varphi_\xi(r) = \sum_{k \, \geq 0} r^k \xi(k)$, la fonction génératrice de $\xi$. On a donc montré que

$$g(r) = r\varphi_\xi(g(r)) \, , \quad r \in [0, 1[ \, . \tag{II.77}$$

Comme $r^\infty = 0$, si $0 < r < 1$, on en déduit que $\lim_{r\uparrow 1} g(r) = \mathbf{P}_1(T_0 < \infty)$. On en déduit que $\mathbf{P}_1(T_0 < \infty)$ est une racine de $\varphi_\xi(r) = r$. On rappelle que cette équation a une seule solution $q_\xi = 1$ si $m_\xi = \sum_{k\geq 0} k\xi(k) \leq$

1 et qu'elle a deux solutions distinctes $q_\xi$ et 1 si $m_\xi > 1$. Dans les deux cas, on a noté $q_\xi$ la plus petite solution sur $[0,1]$ de $\varphi_\xi(r) = r$. On remarque aussi que la convexité stricte de $\varphi_\xi$ (qui est une conséquence de (II.75)) implique que $\varphi_\xi(r) > r$ ssi $r < q_\xi$. Or (II.77) implique que $g(r) > \varphi_\xi(g(r))$ et donc $g(r) < q_\xi$ pour tout $r \in [0,1[$. Donc $\mathbf{P}_1(T_0 < \infty) = q_\xi$. Cela montre

$$\forall i \in \mathbb{N}, \quad \mathbf{P}_i\big(\exists t \in \mathbb{R}_+ ,\ Z_t = 0\big) = (q_\xi)^i . \tag{II.78}$$

En particulier si $m_\xi \leq 1$, $\mathbf{P}_i\big(\exists t \in \mathbb{R}_+ ,\ Z_t = 0\big) = 1$. Si $m_\xi > 1$ alors cette probabilité d'extinction est strictement plus petite que 1 et il y a une probabilité non-nulle que la population tende vers l'infini. Le processus de branchement a deux façons de tendre vers l'infini : ou bien il explose en temps fini ou bien il n'explose pas mais sa taille augmente et tend vers l'infini en un temps infini :

$$\big\{ \lim_{t\uparrow\zeta} Z_t = \infty \big\} = \{\zeta < \infty\} \cup \big\{ \zeta = \infty \quad \text{et} \quad \lim_{t\uparrow\infty} Z_t = \infty \big\} .$$

**Étude de l'explosion.** On suppose que $m_\xi > 1$, ce qui implique que $q_\xi < 1$. On rappelle la notation $\mathbf{E}_1[r^{Z_t}] = v(r,t)$, qui est continûment dérivable en temps et satisfait l'équation (II.71) (page 213, c'est-à-dire l'équation différentielle suivante

$$y'(t) = F(y(t)) , \quad y(0) = r \in [0,1] \quad \text{où} \quad F(z) = c(\varphi_\xi(z) - z) ,\ z \in [0,1] . \tag{II.79}$$

L'étude de cette équation différentielle requiert quelques précautions car il n'y a pas toujours unicité des solutions (précisément dans le cas où il y a explosion). Le théorème de Cauchy-Lipschitz assure l'existence et l'unicité locale des solutions si $F$ est Lipschitzienne : plus précisément supposons que $z \in ]a,b[ \subset [0,1]$ : on a

$$|F'(z)| \leq c\varphi'_\xi(b) + c < \infty .$$

Donc la restriction de $F$ à $]a,b[$ est Lipschitzienne et Cauchy-Lipschitz implique pour tout $r \in ]a,b[$ il existe $t_r > 0$ tel que $t \in [0,t_r[ \mapsto v(r,t) \in ]a,b[$ soit l'unique solution de (II.79) sur l'intervalle de temps $[0,t_r[$. Il y a plusieurs cas à considérer

*Cas où* $r = q_\xi$. On observe que la fonction constante $y(t) = q_\xi$ est solution et un argument simple montre qu'il n'y a globalement qu'une seule solution à (II.79). On a donc

$$\forall t \in \mathbb{R}_+ , \quad v(q_\xi, t) = q_\xi .$$

*Cas où* $r \in [0,q_\xi[$. On se donne une solution locale $y$ telle que $y(0) = r$ définie sur l'intervalle de temps $[0,t_r[$. On constate que si $y(t) \in [0,q_\xi[$, $\varphi_\xi(y(t)) > y(t)$ et donc $y'(t) > 0$. De plus il est impossible qu'il existe un temps $t \in [0,t_r[$ tel que $y(t) = q_\xi$ à cause de l'unicité locale des solutions. Donc $y(t) < q_\xi$, $t \in [0,t_r[$. On en déduit facilement qu'il existe une unique solution issue de $r \in [0,q_\xi[$ définie pour tout temps $t \in \mathbb{R}_+$ :

$$t \in \mathbb{R}_+ \mapsto v(r,t) \in [0,q_\xi[ \text{ est sol. unique de (II.79) ; elle croît et } \lim_{t\uparrow\infty} \uparrow v(r,t) = q_\xi .$$

*Cas où* $r \in ]q_\xi, 1[$. On se donne une solution locale $y$ telle que $y(0) = r$ définie sur l'intervalle de temps $[0,t_r[$. On constante que si $y(t) \in [0,q_\xi[$, $\varphi_\xi(y(t)) < y(t)$ et donc $y'(t) < 0$. De plus il est impossible qu'il existe un temps $t \in [0,t_r[$ tel que $y(t) = q_\xi$ à cause de l'unicité locale des solutions. Donc $y(t) \in ]q_\xi, 1[$, $t \in [0,t_r[$. On en déduit facilement qu'il existe une unique solution issue de $r \in ]q_\xi, 1[$ définie pour tout temps $t \in \mathbb{R}_+$ :

$$t \in \mathbb{R}_+ \mapsto v(r,t) \in ]q_\xi, 1[ \text{ est sol. unique de (II.79) ; elle décroît et } \lim_{t\uparrow\infty} \downarrow v(r,t) = q_\xi .$$

Nous laissons pour l'instant de côté les solutions de l'équation issues de $r = 1$. On observe que si $r \neq q_\xi$, alors $t \mapsto v(r,t)$, qui est unique solution de (II.79), est strictement monotone. On fixe donc $r$ et on voit que

$$\forall t \in ]0,\infty[\ , \quad \int_0^t \frac{\frac{\partial}{\partial s} v(r,s)}{\varphi_\xi(v(r,s)) - v(r,s)} ds = c\, t\ .$$

On effectue le changement de variable $u = v(r,s)$ et on obtient donc

$$\forall r \in [0,1]\backslash\{q_\xi, 1\}\ ,\ \forall t \in ]0,\infty[\ , \quad \int_{v(r,t)}^r \frac{du}{u - \varphi_\xi(u)} = c\, t\ . \tag{II.80}$$

Comme $Z_t = \infty$ si $t \geq \zeta$, on a $v(r,t) = \mathbf{E}_1[r^{Z_t} \mathbf{1}_{\{t<\zeta\}}]$, pour tout $r < 1$. On a donc

$$\mathbf{P}_1(t < \zeta) = v(1-, t) = \lim_{r\uparrow 1} \uparrow v(r,t)\ .$$

On rappelle que l'on a supposé $m_\xi > 1$, ce qui implique que $q_\xi < 1$. La stricte convexité de $\varphi_\xi$, garantie par (II.75), implique que $u - \varphi_\xi(u) > 0$, pour tout $u \in ]q_\xi, 1[$ . Il y a deux cas à considérer

*Cas où* $\int^{1-} \frac{du}{u-\varphi_\xi(u)} < \infty$, c'est-à-dire qu'il existe $\eta > 0$ tel que $\int_{1-\eta}^1 \frac{du}{u-\varphi_\xi(u)} < \infty$. Alors en passant à la limite dans (II.80) on obtient que $q_\xi < v(1-,t) < 1$ et

$$\forall t \in ]0,\infty[, \quad \int_{v(1-,t)}^1 \frac{du}{u - \varphi_\xi(u)} = ct\ . \tag{II.81}$$

De plus

$$t \in \mathbb{R}_+^* \mapsto v(1-,t) \in ]q_\xi, 1[ \text{ est sol. de (II.79) avec } r = 1 \text{ ; elle décroît et } \lim_{t\uparrow\infty} \downarrow v(1-,t) = q_\xi\ .$$

(**Attention**, elle n'est pas solution unique : en effet la solution constante $y(t) = 1$, $t \in \mathbb{R}_+$ est aussi solution de (II.79) avec condition initiale $y(0) = r = 1$.) Il y a explosion et on a

$$\mathbf{P}_i(t < \zeta) = (v(1-,t))^i \xrightarrow[t\to\infty]{} \mathbf{P}_i(\zeta = \infty) = \ q_\xi^i = \mathbf{P}_i(\exists t \in \mathbb{R}_+\ ,\ Z_t = 0)\ .$$

Par conséquent on a

$$\mathbf{P}_i(\zeta = \infty\ ;\ \lim_{t\to\infty} Z_t = \infty) = 0 \quad \text{et} \quad \mathbf{P}_i(\exists t \in \mathbb{R}_+\ ,\ Z_t = 0) + \mathbf{P}_i(\zeta < \infty) = 1.$$

C'est-à-dire que le processus de branchement s'éteint en temps fini ou bien explose en temps fini. On observe que

$$\int^{1-} \frac{du}{u - \varphi_\xi(u)} < \infty \quad \Longrightarrow \quad m_\xi = \infty\ .$$

*Cas où* $\int^{1-} \frac{du}{u-\varphi_\xi(u)} = \infty$. En passant à la limite dans (II.80), on voit que $\lim_{r\to 1} v(r,t) = 1$, nécessairement. Donc $\mathbf{P}(t < \zeta) = v(1-,t) = 1$, pour tout $t \in \mathbb{R}_+$. La fonction constante égale à 1 et la seule solution de l'équation (II.79) de condition initiale égale à 1. De plus, on a $\mathbf{P}_1(\zeta = \infty) = 1$. Il n'y a pas d'explosion et on a

$$\mathbf{P}_i(\zeta < \infty) = 0 \quad \text{et} \quad \mathbf{P}_i(\exists t \in \mathbb{R}_+\ ,\ Z_t = 0) + \mathbf{P}_i\big(\zeta = \infty\ ;\ \lim_{t\to\infty} Z_t = \infty\big) = 1\ .$$

Autrement dit, le processus n'explose jamais : ou bien il s'éteint en temps fini ou bien il tend vers l'infini en un temps infini.

## EXERCICES.

**Exercice II.19** On considère un processus de branchement de loi de reproduction $\xi$ et de paramètre des temps de vie $c$. On utilise les notations introduites précédemment. On suppose que $m_\xi > 1$ et on se place dans le cas explosif :

$$\int^{1-} \frac{du}{u-\varphi_\xi(u)} < \infty \; .$$

1. Montrer que $m_\xi = \infty$.
2. On considère l'équation différentielle

$$y \in C^1(\mathbb{R}_+, [0,1]) \,, \quad y'(t) = c\big(\varphi_\xi(y(t)) - y(t)\big) \,, \text{ et } y(0) = 1 \; . \tag{II.82}$$

   On a vu qu'il n'y a pas unicité de la solution car la fonction constante égale à 1, noté $y_\infty$ et la fonction $y_0(t) = v(1-, t)$, $t \in \mathbb{R}_+$ sont solutions de (II.82). Pour tout $t_0 \in ]0, \infty[$, on définit $y_{t_0} : \mathbb{R}_+ \to [0,1]$ en posant

$$y_{t_0}(t) = 1 \;\text{ si } t \in [0, t_0] \text{ et } \; y_{t_0}(t) = v(1-, t - t_0) \;\text{ si } t \in [t_0, \infty[ \, .$$

   Montrer que $y_{t_0}$ est solution (II.82).
3. Montrer que les seules solutions de (II.82) sont les fonctions $y_{t_0}$, $t_0 \in [0, \infty]$, définies précédemment. □

**Exercice II.20** On considère un processus de branchement de loi de reproduction $\xi$ et de paramètre des temps de vie $c$, noté

$$\big(\,\Omega\,;\, \mathscr{F}\,;\, (\mathscr{F}_t)_{t\in\mathbb{R}_+}\,;\, \mathbf{Z} = (Z_t)_{t\in\mathbb{R}_+}\,;\, G = (q_{i,j})_{i,j\in\mathbb{N}}\,;\, \mathbf{P}_\mu \,,\, \mu \in \mathscr{M}_1(\mathbb{N})\,\big) \; .$$

On utilise les notations introduites précédemment et on pose $v(r,t) = \mathbf{E}_1[r^{Z_t}]$, $r \in ]0,1[$ avec la convention $r^\infty = 0$ et $r^0 = 1$.

1. On pose $Z^*_t = Z_{t/c}$, $t \in \mathbb{R}_+$. Montrer que cela définit un processus de branchement de loi de reproduction $\xi$ et de paramètre de temps de vie 1.
2. On pose $T_{\rm ex} = \inf\{t \in \mathbb{R}_+ \;:\; Z_t = 0\}$, avec la convention que $\inf \emptyset = \infty$. C'est un $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$-temps d'arrêt qui est appelé le temps d'extinction de $\mathbf{Z}$. Montrer que

$$\mathbf{P}_1(T_{\rm ex} \leq t) = \lim_{r\downarrow 0} \downarrow v(r,t) := v(0+, t) \,, \quad t \in \mathbb{R}_+ \; .$$

3. On suppose que $0 < \xi(0) < 1$. Montrer que

$$\int_0^{v(0+,t)} \frac{du}{\varphi_\xi(u) - u} = ct \; .$$

   En déduire que sous $\mathbf{P}_1$ la loi de $T_{\rm ex}$ est donnée par

$$f(t)\,\ell(dt) + (1-q_\xi)\delta_\infty(dt) \,,$$

   où $\ell$ est la mesure de Lebesgue et $f : \mathbb{R}_+ \to \mathbb{R}_+$ est une fonction $C^\infty$ qui satisfait d'intégrale $\int_{\mathbb{R}_+} f(t)\ell(dt) = q_\xi$ qui est égale à $v'(0+,t)$ et qui satisfait l'équation

$$f(t) = c\big(\varphi_\xi(v(0+,t)) - v(0+,t)\big) \,, t \in \mathbb{R}_+ \; .$$

4. Montrer que pour toute loi d'entrée $\mu \in \mathscr{M}_1(\mathbb{N})$, on a

$$\mathbf{P}_\mu(T_{\rm ex} \leq t) = \sum_{i\in\mathbb{N}} \mu(i) v(0+,t)^i \,, \quad t \in \mathbb{R}_+ \; .$$

5. On choisit $\xi(0) = \xi(2) = 1/2$ et $c = 1$. Calculer $v(r,t)$ et calculer la densité de $T_{\rm ex}$. □

**Exercice II.21** On considère un processus de branchement de loi de reproduction $\xi$ et de paramètre des temps de vie $c$, noté

$$\big(\,\Omega\,;\, \mathscr{F}\,;\, (\mathscr{F}_t)_{t\in\mathbb{R}_+}\,;\, \mathbf{Z} = (Z_t)_{t\in\mathbb{R}_+}\,;\, G = (q_{i,j})_{i,j\in\mathbb{N}}\,;\, \mathbf{P}_\mu \,,\, \mu \in \mathscr{M}_1(\mathbb{N})\,\big) \; .$$

On suppose que

$$m_\xi = \sum_{i\in\mathbb{N}} i\xi(i) \;\in ]0,\infty[ \; .$$

Pour tout $t \in \mathbb{R}_+$, on pose $m(t) = \mathbf{E}_1[Z_t]$.

1. Montrer que $t \mapsto m(t)$ est une fonction $C^\infty$ satisfaisant une équation différentielle que l'on précisera et que l'on résoudra.
2. On pose $M_t = Z_t/m(t)$, $t \in \mathbb{R}_+$. Montrer que pour tous $t, s \in \mathbb{R}_+$, et pour toute loi initiale $\mu \in \mathscr{M}_1(\mathbb{N})$, on a

$$\mathbf{P}_\mu\text{-p.s.} \quad \mathbf{E}_\mu\big[M_{t+s} \,|\mathscr{F}_t\big] = M_t \; .$$

□

**Exercice II.22** (*Processus de Yule*) On appelle processus de Yule un processus de branchement de loi de reproduction déterministe binaire $\xi(2) = 1$ et de paramètre des temps de vie $c$.

1. Montrer qu'un processus de Yule est aussi un processus de vie et de mort.
2. On note $\mathbf{Z}$ un processus de Yule et on note $J_n$ son $n$-ème temps saut. On suppose que $\mathbf{P}(Z_0 = 1) = 1$. Calculer $\mathbf{E}[\exp(-\lambda J_n)]$ et donner un équivalent de $\mathbf{E}[J_n]$ lorsque $n$ tend vers l'infini.
3. On pose $U_t^n = Z_{(J_n - t)^+/n}$, $t \in \mathbb{N}$. Montrer que $(U^n)_{t \in \mathbb{R}_+}$ est un processus de Markov à valeur dans $\{1, \dots, n\}$ dont on calculera le générateur. □

**Exercice II.23** On fixe $a \in ]0,1[$. Pour tout $k \in \mathbb{N}$, on pose

$$\xi(0) = \xi(1) = 0 \; , \quad \xi(2) = a \quad \text{et} \quad \xi(k) = a \cdot \frac{(2-a)(3-a)\, \dots (k-1-a)}{k!} \; , \quad k \geq 3 \; .$$

1. On remarque que

$$\log \xi(k) = \sum_{\ell=1}^{k} \log\left(1 - \frac{a+1}{\ell}\right) \; .$$

En déduire qu'il existe deux constantes $C_1, C_2 > 0$ telles que

$$\forall k \geq 3 \; , \quad C_1 k^{-(a+1)} \leq \xi(k) \leq C_2 k^{-(a+1)} \; .$$

2. Montrer que la série entière de terme général $\xi(k)$ est de rayon de convergence 1. Montrer que pour tout $r \in [0,1]$

$$\varphi(r) = \sum_{k \geq 0} r^k \xi(k) = \frac{1 - ar - (1-r)^a}{1-a} \; .$$

En déduire que $\xi$ est une probabilité sur $\mathbb{N}$.

3. Montrer que le processus de branchement $X$ de loi de reproduction $\xi$ et de paramètre des temps de vie $c > 0$ explose avec probabilité 1, c'est-à-dire que pour toute mesure initiale

$$\mathbf{P}_\mu(\zeta < \infty) = 1$$

4. Montrer que

$$\mathbf{P}_1(\zeta > t) \sim_{t \to 0} (ct)^{-\frac{1}{1-a}} \; .$$

5. On prend $a = 1/2$ et on pose

$$\widetilde{\xi}(0) = 1/2 \; , \quad \widetilde{\xi}(1) = 0 \quad \text{et} \quad \widetilde{\xi}(k) = \frac{1}{2}\xi(k) \; , \quad k \geq 2 \; .$$

On considère le processus de branchement $\widetilde{\mathbf{Z}}$ de loi de reproduction $\widetilde{\xi}$ et de paramètre des temps de vie $c$. Calculer $\mathbf{P}_1(\zeta(\widetilde{\mathbf{Z}}) < \infty)$ et calculer $\mathbf{P}_1(\exists t \in \mathbb{R}_+ : \widetilde{Z}_t = 0)$. □

**Exercice II.24** On fixe un paramètre réel $c > 0$. On se donne un processus à valeurs dans $\mathbb{N}$, noté $X = (X_t, t \in \mathbb{R}_+)$ défini sur un espace de probabilité $(\Omega; \mathscr{F}; \mathbf{P})$. On fixe ensuite une loi de probabilité $\xi$ sur $\mathbb{N}$ telle que $\xi(0) < 1$ et $\xi(1) = 0$. On suppose que

$$m = \sum_{k \in \geq 0} k\xi(k) < \infty \; ,$$

et on note $\varphi$ la fonction génératrice de $\xi$. On suppose que tous $s, t \geq 0$ et tout $r \in [0,1]$ on a

$$\mathbf{P}_i - \text{p.s.} \quad \mathbf{E}_i\left[ r^{X_{t+s}} \,\middle|\, \mathscr{F}_s^X \right] = v(r,t)^{X_s} \; , \tag{II.83}$$

où pour tout $r \in [0,1]$, $t \mapsto v(r,t)$ est l'unique solution positive de l'équation différentielle

$$\frac{\partial}{\partial t} v(r,t) = c\,(\,\varphi(v(r,t)) - v(r,t)\,) \quad \text{et} \quad v(r,0) = r \; .$$

1. Montrer que $X$ sous $\mathbf{P}$ a même loi est un processus de branchement de loi de reproduction $\xi$, de paramètre des temps de vie $c$ et de loi initiale la loi de $X_0$ sous $\mathbf{P}$.
2. On suppose que $m > 1$ et on note $q$ la plus petite racine de $\varphi$. On pose

$$\mathrm{Extinction} = \{\exists t \in \mathbb{R}_+ : X_t = 0\} .$$

   Montrer que $X$ sous $\mathbf{P}(\,\cdot\,|\mathrm{Extinction})$ a même loi qu'un processus de branchement de paramètre des temps de vie $c$, de loi initiale la loi de $X_0$ sous $\mathbf{P}$ et de loi de reproduction $\widetilde{\xi}$ donnée par

$$\forall k \in \mathbb{N} , \quad \widetilde{\xi}(k) = q^{k-1}\xi(k) .$$

$\square$

# Chapitre III

# Nuages de Poisson.

Le but de ce chapitre est généraliser la notion de processus de Poisson, de construire des outils de calcul. Ces résultats sont ensuite utilisés pour étudier les processus de Lévy et développer la théorie des excursions des processus Markoviens.

## III.1 Processus de Poisson linéaire homogène.

Dans ce chapitre, sauf mention du contraire, toutes les variables aléatoires sont définies sur un même espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$ qui est supposé suffisamment riche pour que l'on puisse y définir autant de suites de variables aléatoires indépendantes que nécessaire. Il est simple de constater que la phrase "on choisit un point uniformément au hasard sur $\mathbb{R}_+$" n'a pas de sens précis. En effet, raisonnons par l'absurde et supposons que cette phrase ait un sens précis : on note $X$ ce point aléatoire uniforme. C'est donc une variable aléatoire positive définie sur $(\Omega, \mathscr{F}, \mathbf{P})$ et on doit avoir

$$1 = \mathbf{P}(\Omega) = \sum_{n\in\mathbb{N}} \mathbf{P}(X \in [n, n+1[\,) \ .$$

Comme $X$ est de loi "uniforme", on doit avoir $\mathbf{P}(X \in [n, n+1[\,) = \mathbf{P}(X \in [0,1[\,)$, pour tout $n \in \mathbb{N}$, ce qui contredit l'égalité precédente.

Si la phrase précédente n'a pas de sens, il est en revanche possible de définir un ensemble *infini dénombrable* de points aléatoires sur $\mathbb{R}_+$ *uniformément répartis et totalement indépendants*. La définition d'un tel ensemble aléatoire est l'objet de cette section. Nous rappelons d'abord quelques résultats élémentaires sur les lois usuelles (Poisson, exponentielles, Erlang, statistiques d'ordre). Puis nous dégageons heuristiquement une définition d'un tel ensemble aléatoire de point avant d'en donner une construction.

### III.1.a Événements rares et loi de Poisson, lois exponentielles, statistiques d'ordre.

**Définition III.1.1** Une v.a. $X:\Omega\to\mathbb{N}$ est (de loi) de Poisson de paramètre $\theta\in\,]0,\infty[$ si $\mathbf{P}(X=k)=e^{-\theta}\theta^k/k!$, pour tout $k\in\mathbb{N}$. □

Si $X$ est de Poisson de paramètre $\theta$, alors on vérifie facilement que

$$\mathbf{E}\big[r^X\big] = \exp\big(-\theta(1-r)\big), \quad r \in [0,1] \ .$$

En dérivant une puis deux fois cettefonction génératrice en 1,

$$\mathbf{E}[X] = \theta, \quad \mathbf{E}[X^2] = \theta + \theta^2 \quad \text{et donc} \quad \mathbf{var}(X) = \theta \ .$$

Cela justifie la définition suivante des lois de Poisson dégénérées.

**Définition III.1.2** Une v.a. $X$ suit une loi de Poisson de paramètre infini (resp. nul) si $\mathbf{P}(X = \infty) = 1$ (resp. si $\mathbf{P}(X = 0) = 1$). □

**Lemme III.1.3** *Soient $X_1, \ldots, X_n$, des v.a. de Poisson de paramètres respectifs $\theta_1, \ldots, \theta_n$. On les suppose indépendantes. Alors $X_1 + \ldots + X_n$ est une v.a. de Poisson de paramètre $\theta_1 + \ldots + \theta_n$.*

*Plus généralement, soit $(X_n)_{n\in\mathbb{N}}$ des v.a. de Poisson indépendantes. On note $\theta_n$ le paramètre de $X_n$. Alors $X := \sum_{n\in\mathbb{N}} X_n$ est une v.a. de Poisson (éventuellement dégénérée) de paramètre $\theta := \sum_{n\in\mathbb{N}} \theta_n$.*

**Preuve :** on montre d'abord le premier point. Pour cela, on observe que

$$\mathbf{E}\big[r^{X_1+\ldots+X_n}\big] = \mathbf{E}\big[r^{X_1}\big] \ldots \mathbf{E}\Big[r^{X_n}\Big] = \exp\big(-(\theta_1 + \ldots + \theta_n)(1-r)\big) ,$$

ce qui permet de conclure puisque les fonctions génératrices caractérisent les lois sur $\mathbb{N}$.

Montrons le second point du lemme. Pour tout $n\in\mathbb{N}$, on pose $S_n = X_0 + \ldots + X_n$ et $t_n = \theta_0 + \ldots + \theta_n$. On suppose d'abord que $\theta = \sum_{n\in\mathbb{N}} \theta_n < \infty$. On remarque que $S_n \leq S_{n+1}$ et que $\lim_n S_n = X$. Par convergence dominée et d'après le premier point du lemme, on a donc pour tout $r \in [0,1]$,

$$\mathbf{E}\big[r^X\big] = \lim_{n\to\infty} \big[r^{S_n}\big] = \lim_{n\to\infty} \exp\big(-t_n(1-r)\big) = \exp\big(-\theta(1-r)\big) ,$$

ce qui implique que $X$ suit une loi de Poisson de paramètre $\theta$.

On suppose ensuite que $\theta = \infty$. S'il existe $n\in\mathbb{N}$ tel que $\theta_n = \infty$, alors $X_n = \infty$ p.s. et donc $X = \infty$, ce qui montre le résultat voulu. Il reste à examiner le cas où pour tout $n\in\mathbb{N}$, $\theta_n < \infty$. Pour tout $x\in\mathbb{R}_+$ et tout $n\in\mathbb{N}$, une inégalité de Markov implique que

$$\mathbf{P}(S_n \leq x) \leq e^x \mathbf{E}\big[e^{-S_n}\big] = e^{x - t_n(1-e^{-1})} .$$

Donc $\lim_n \mathbf{P}(S_n \leq x) = 0$. Or $\{S < x\} \subset \bigcap_{n\in\mathbb{N}} \{S_n \leq x\}$, donc $\mathbf{P}(X < x) = 0$, pour tout $x \in \mathbb{R}_+$, ce qui implique que p.s. $X = \infty$, ce qui permet de conclure. ■

**Approximation des lois de Poisson par des loi binomiales.** Les lois de Poisson apparaissent comme des loi binomiales dégénérées : plus précisément, on fixe $\theta > 0$ et pour tout $n \in \mathbb{N}^*$ tel que $n > \theta$, on considère $Y_{n,1}, \ldots, Y_{n,n}$, des variables de Bernoulli indépendantes de même paramètre $p_n = \theta/n$. On pose

$$X_n = \sum_{1\leq k\leq n} Y_{n,k} .$$

La probabilité que $Y_{n,k}$ vale 1 est donc $p_n$ : elle devient très petite lorsque $n \to \infty$. En revanche on voit que $\mathbf{E}[X_n] = np_n = \theta$, qui ne varie pas avec $n$. La loi $X_n$ est une binomiale de paramètres $(n, p_n)$. Pour tout $k\in\mathbb{N}$, et tout $n \geq k$, on a

$$\mathbf{P}(X_n = k) = \binom{n}{k} p_n^k (1-p_n)^{n-k} = \frac{\theta^k}{k!}\big(1 - \tfrac{1}{n}\theta\big)^{n-k} \prod_{1\leq j\leq k-1} \big(1 - \tfrac{j}{n}\big)$$

Cela montre que

$$\forall k\in\mathbb{N}, \quad \lim_{n\to\infty} \mathbf{P}(X_n = k) = e^{-\theta}\theta^k/k! .$$

En un sens que nous expliquons précisément plus loin $X_n$ est très proche d'une loi de Poisson. Le théorème suivant donne un résultat plus précis.

**Théorème III.1.4** **(Approximation Binomiale-Poisson)** *Pour tout $n \geq 1$, on se donne une suite $(Y_{n,k})_{1 \leq k \leq n}$ de variables indépendantes de Bernoulli de paramètres respectifs $p_{n,k}$, c'est-à-dire $\mathbf{P}(Y_{n,k} = 1) = p_{n,k}$. On pose*

$$X_n = \sum_{1 \leq k \leq n} Y_{n,k} \quad \text{et} \quad \theta_n = \sum_{1 \leq k \leq n} p_{n,k} \ .$$

*Alors pour tout $\theta > 0$,*

$$\sum_{k \in \mathbb{N}} \left| \mathbf{P}(X_n = k) - e^{-\theta} \frac{\theta^k}{k!} \right| \leq 2|\theta - \theta_n| + 2 \sum_{1 \leq k \leq n} p_{n,k}^2 \ .$$

*Si $\lim_n \theta_n = \theta$ et si $\lim_n \max_{1 \leq k \leq n} p_{n,k} = 0$, alors $\lim_n \sum_{1 \leq k \leq n} p_{n,k}^2 = 0$, et*

$$\lim_{n \to \infty} \sum_{k \in \mathbb{N}} \left| \mathbf{P}(X_n = k) - e^{-\theta} \frac{\theta^k}{k!} \right| = 0 \ .$$

**Exemple d'application.** Avant de prouver ce théorème, voyons un exemple d'application de cette approximation : les statistiques nous disent que 3,2 piétons sont écrasés à Paris chaque mois (c'est un chiffre fantaisiste). Quelle est la probabilité que, le mois prochain, exactement 5 piétons soient écrasés à Paris ? Pour répondre rapidement à cette question on peut établir le modèle suivant : il y a (environ) $n \simeq 2.10^6$ Parisiens qui ont une "chance" très faible de se faire écraser dans le mois suivant, et ce, indépendamment les uns des autres (nous négligeons les parents portant un nourrisson). On numérote les Parisiens de 1 à $n$ et on note $Y_k$ la variable aléatoire qui vaut 1 si le Parisien numero $k$ se fait écraser le mois prochain et qui vaut 0 sinon. D'après nos hypothèses, on peut considérer $(Y_k)_{1 \leq k \leq n}$ comme une suite i.i.d. de variables de Bernoulli de paramètre $p$. Par conséquent $X = Y_1 + \ldots + Y_n$ suit une loi Binomiale de paramètres $(n, p)$ . Sa moyenne vaut $np = 3{,}2$. En appliquant la proposition précédente avec $np = \theta = 3{,}2$, on obtient

$$\sum_{k \in \mathbb{N}} \left| \mathbf{P}(X = k) - e^{-3{,}2} \frac{(3{,}2)^k}{k!} \right| \leq (3{,}2)^2/n.$$

Or $(3{,}2)^2/n$ est de l'ordre de $10^{-5}$. On peut donc assimiler la probabilité que le mois prochain exactement 5 piétons soient écrasés à Paris à la quantité $e^{-3{,}2}(3{,}2)^5/5!$.

**Preuve du théorème d'approximation binomiale-Poisson.** Pour démontrer le théorème III.1.4, on cherche à comparer les mesures de probabilités sur $\mathbb{Z}$. Une mesure de probabilité $\mu$ sur $\mathbb{Z}$ est simplement donnée par $\mu = (\mu(k))_{k \in \mathbb{Z}}$. On note $\mathscr{M}_1(\mathbb{Z})$ l'ensemble des mesures de probabilités sur $\mathbb{Z}$. Pour toutes mesures $\mu, \nu \in \mathscr{M}_1(\mathbb{Z})$, on pose

$$d_{\text{var}}(\mu, \nu) = \sum_{k \in \mathbb{Z}} |\mu(k) - \nu(k)| \ .$$

Cette quantité est appelée *distance en variation de $\mu$ et $\nu$*. On peut voir $\mathscr{M}_1(\mathbb{Z})$ comme un sous-ensemble de l'espace des suite $\ell^1(\mathbb{Z})$ et $d_{\text{var}}$ n'est autre que la norme $\mathrm{L}^1$. Il est facile de vérifier $\mathscr{M}_1(\mathbb{Z})$ est un fermé de $\ell^1(\mathbb{Z})$ ce qui implique que $(\mathscr{M}_1(\mathbb{Z}), d_{\text{var}})$ est un espace métrique séparable complet. En fait $d_{\text{var}}$ métrise la convergence en loi sur $\mathbb{Z}$.

Soient $\mu, \nu \in \mathscr{M}_1(\mathbb{Z})$, on rappelle que $\mu * \nu$ est la convolée de $\mu$ avec $\nu$. Il est clair que $\mu * \nu \in \mathscr{M}_1(\mathbb{Z})$ : précisément, on a $\mu * \nu\,(k) = \sum_{j \in \mathbb{N}} \mu(k - j)\nu(j)$ pour tout $k \in \mathbb{Z}$. On rappelle que si $X$, de loi $\mu$ et $Y$, de loi $\nu$, sont deux v.a. à valeurs dans $\mathbb{Z}$, $\mathscr{F}$-mesurables et indépendantes, alors la loi de $X + Y$ est $\mu * \nu$.

On montre tout d'abord l'inégalité suivante.

$$\forall \mu_1, \mu_2, \nu_1, \nu_2 \in \mathscr{M}_1(\mathbb{Z}), \quad d_{\mathrm{var}}(\mu_1 * \mu_2 \, , \, \nu_1 * \nu_2) \le d_{\mathrm{var}}(\mu_1, \nu_1) + d_{\mathrm{var}}(\mu_2, \nu_2). \tag{III.1}$$

*Preuve :* on a

$$\begin{aligned} d_{\mathrm{var}}(\mu_1 * \mu_2 \, , \, \nu_1 * \nu_2) &\le \sum_{j,k \in \mathbb{Z}} |\mu_2(k-j)\mu_1(j) - \nu_2(k-j)\nu_1(j)| = \sum_{j,k \in \mathbb{Z}} |\mu_2(k)\mu_1(j) - \nu_2(k)\nu_1(j)| \\ &\le \sum_{j,k \in \mathbb{Z}} |\mu_2(k)\mu_1(j) - \mu_2(k)\nu_1(j)| + \sum_{j,k \in \mathbb{Z}} |\mu_2(k)\nu_1(j) - \nu_2(k)\nu_1(j)| \\ &= \sum_{k \in \mathbb{Z}} \mu_2(k) d_{\mathrm{var}}(\mu_1, \nu_1) + \sum_{j \in \mathbb{Z}} \nu_1(j) d_{\mathrm{var}}(\mu_2, \nu_2) = d_{\mathrm{var}}(\mu_1, \nu_1) + d_{\mathrm{var}}(\mu_2, \nu_2) \, , \end{aligned}$$

ce qui termine la preuve de (III.1). □

On introduit la notation suivante

$$\forall p \in [0,1], \quad \mu_p = p\delta_1 + (1-p)\delta_0 \quad \text{et} \quad \forall \theta \in ]0, \infty[ \quad \nu_\theta = \sum_{k \in \mathbb{N}} \tfrac{1}{k!} e^{-\theta} \theta^k \delta_k \, . \tag{III.2}$$

Autrement dit $\mu_p$ est la loi de Bernoulli de paramètre $p$ et $\nu_\theta$ la loi de Poisson de paramètre $\theta$. On montre alors :

$$\forall p \in ]0,1], \quad d_{\mathrm{var}}(\mu_p, \nu_p) \le 2p^2 \quad \text{et} \quad \forall \theta, \theta' \in ]0, \infty[ \, , \quad d_{\mathrm{var}}(\nu_\theta, \nu_{\theta'}) \le 2|\theta - \theta'|. \tag{III.3}$$

*Preuve :* par convexité on rappelle que $0 \le 1 - e^{-p} \le p$. Or

$$\begin{aligned} d_{\mathrm{var}}(\mu_p, \nu_p) &= (e^{-p} - 1 + p) + (p - pe^{-p}) + \sum_{k \ge 2} \frac{p^k}{k!} e^{-p} \\ &= (e^{-p} - 1 + p) + (p - pe^{-p}) + (1 - e^{-p} - pe^{-p}) = 2p(1 - e^{-p}) \le 2p^2, \end{aligned}$$

ce qui montre la première inégalité. Montrons la seconde : sans perte de généralité on suppose que $\theta < \theta'$. On pose $\delta = \theta' - \theta$. Le lemme III.1.3 implique $\nu_{\theta'} = \nu_\theta * \nu_\delta$ et (III.1) implique que

$$d_{\mathrm{var}}(\nu_\theta, \nu_{\theta'}) = d_{\mathrm{var}}(\nu_\theta * \delta_0 \, , \, \nu_\theta * \nu_\delta) \le d_{\mathrm{var}}(\nu_\theta, \nu_\theta) + d_{\mathrm{var}}(\delta_0, \nu_\delta) = d_{\mathrm{var}}(\delta_0, \nu_\delta) \, .$$

Or $d_{\mathrm{var}}(\delta_0, \nu_\delta) = 1 - e^{-\delta} + \sum_{k \ge 1} \nu_\delta(k) = 2(1 - e^{-\delta}) \le 2\delta$, ce qui permet de conclure. □

*Fin de la preuve du théorème III.1.4 :* on fixe $n$. On note $\mu_{X_n}$ la loi de $X_n$. Avec les notations de (III.2), $\mu_{X_n} = \mu_{p_{n,1}} * \ldots * \mu_{p_{n,n}}$. Le lemme III.1.3 implique également que $\nu_{\theta_n} = \nu_{p_{n,1}} * \ldots * \nu_{p_{n,n}}$. Par le lemme précédents et (III.1), on obtient les inégalité suivantes :

$$\begin{aligned} d_{\mathrm{var}}(\nu_\theta, \mu_{X_n}) &\le d_{\mathrm{var}}(\nu_\theta, \nu_{\theta_n}) + d_{\mathrm{var}}(\nu_{\theta_n}, \mu_{X_n}) \le 2|\theta - \theta_n| + d_{\mathrm{var}}\big(\mu_{p_{n,1}} * \ldots * \mu_{p_{n,n}} \, , \, \nu_{p_{n,1}} * \ldots * \nu_{p_{n,n}}\big) \\ &= 2|\theta - \theta_n| + \sum_{1 \le k \le n} 2p_{n,k}^2 \, , \end{aligned}$$

ce qui permet de conclure. ■

**Lois exponentielles.**

**Définition III.1.5** Une v.a. $X : \Omega \to \mathbb{R}_+$ est (de loi) exponentielle de paramètre $c \in ]0, \infty[$ si sa loi admet la densité $t \in \mathbb{R}_+ \longmapsto ce^{-ct}$ par rapport à la mesure de Lebesgue. □

Si $X$ est une exponentielle de paramètre $c$, alors pour toute fonction mesurable $f : \mathbb{R}_+ \to \mathbb{R}_+$,

$$\mathbf{E}[f(X)] = \int_0^\infty f(t)ce^{-ct}\, dt \ .$$

Il est facile de vérifier pour tout $\lambda \in \mathbb{R}_+$ que

$$\mathbf{E}\big[e^{-\lambda X}\big] = \frac{c}{c+\lambda} \ . \tag{III.4}$$

En dérivant une puis deux fois cette transformée de Laplace en 0, on obtient

$$\mathbf{E}[X] = \frac{1}{c}, \quad \mathbf{E}[X^2] = \frac{2}{c^2} \quad \text{et donc} \quad \mathbf{var}(X) = \frac{1}{c^2} \ .$$

Les variables exponentielles sont caractérisée par la propriété d'oubli qui s'énonce comme suit.

**Proposition III.1.6** *Soit $X : \Omega \to \mathbb{R}_+$, une v.a. $\mathscr{F}$-mesurable telle que $\mathbf{P}(X > t) > 0$, pour tout $t \in \mathbb{R}_+$. On suppose qu'elle possède la propriété d'oubli c'est-à-dire que*

$$\forall s, t \in \mathbb{R}_+, \quad \mathbf{P}(X - t > s \,|\, X > t) = \mathbf{P}(X > s) \ .$$

*Alors $X$ est une variable exponentielle pour un certain paramètre $c \in ]0, \infty[$.*

**Preuve :** on pose $f(t) = -\log \mathbf{P}(X > t)$. La fonction $f : \mathbb{R}_+ \to \mathbb{R}_+$ est bien définie. Par ailleurs $f$ est continue à droite car $F_X$ et continue à droite et $f(t) = -\log(1 - F_X(t))$. La propriété d'oubli implique que $f(t) + f(s) = f(t+s)$, pour tous $s, t \in \mathbb{R}_+$. On pose $c = f(1)$. Pour tout entier $n \in \mathbb{N}^*$, $f(n) = cn$ et $nf(1/n) = c$. Donc pour tout $t \in \mathbb{Q}_+$, $f(t) = ct$. La continuité à droite implique que $f(t) = ct$ pour tout $t \in \mathbb{R}_+$. On a donc $\mathbf{P}(X > t) = e^{-ct}$, pour tout $t \in \mathbb{R}_+$ et $c$ ne peut être nul, ce qui termine la preuve de la proposition. ■

**Proposition III.1.7** *Soient $\mathscr{E}_1, \ldots, \mathscr{E}_n$, des exponentielles indépendantes de paramètre $c \in \mathbb{R}_+^*$. Alors $S = \mathscr{E}_1 + \ldots + \mathscr{E}_n$ suit une loi d'Erlang de paramètres $(n, c)$, c'est-à-dire que $S$ a pour densité $f_S(t) = \frac{c^n t^{n-1}}{(n-1)!} e^{-ct}$, $t \in \mathbb{R}_+$.*

**Preuve :** soit $f : \mathbb{R}_+ \to \mathbb{R}_+$ une fonction mesurable bornée. On a

$$\mathbf{E}\left[f(S)\right] = \int_{\mathbb{R}_+^n} c^n e^{-c(x_1 + \ldots + x_n)} f(x_1 + \ldots + x_n)\, dx_1 \ldots dx_n \ .$$

On effectue le changement de variable linéaire :

$$t = x_1 + \ldots + x_n \ , \ t_{n-1} = x_1 + \ldots + x_{n-1} \ , \ \ldots \ , \ t_2 = x_1 + x_2 \ , \ t_1 = x_1$$

La matrice de transition est triangulaire avec des 1 sur la diagonale. Par ailleurs $\mathbb{R}_+^n$ est envoyé sur le domaine $\{0 \le t_1 \le \ldots \le t_{n-1} \le t\}$. On a donc

$$\mathbf{E}\left[f(S)\right] = \int_{\{0 \le t_1 \le \ldots \le t_{n-1} \le t\}} c^n e^{-ct} f(t)\, dt\, dt_{n-1} \ldots dt_1 \ .$$

Une simple récurrence montre que $\int_{\{0 \le t_1 \le \ldots \le t_{n-1} \le t\}} dt_{n-1} \ldots dt_1 = t^{n-1}/(n-1)!$, ce qui implique le résultat par Fubini. ■

**Statistique d'ordre de $n$ v.a. réelles indépendantes de lois diffuses.** Soit $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$. On note $\Lambda_n(\mathbf{x}) = (x_{(k)})_{1 \leq k \leq n}$ le réarrangement croissant des réels $(x_1, \ldots x_n)$, c'est-à-dire que

$$x_{(1)} \leq \ldots \leq x_{(k)} \leq x_{(k+1)} \leq \ldots \leq x_{(n)} \quad \text{et} \quad \{x_k \, ; \, 1 \leq k \leq n\} = \{x_{(k)} \, ; \, 1 \leq k \leq n\} \; .$$

L'application $\Lambda$ est continue donc mesurable.

**Définition III.1.8** Soient $X_1, \ldots, X_n$, des v.a. réelles $\mathscr{F}$ mesurables. On note $(X_{(k)})_{1 \leq k \leq n}$ le réarrangement croissant de ces variables :

$$X_{(1)} \leq \ldots \leq X_{(k)} \leq X_{(k+1)} \leq \ldots \leq X_{(n)} \quad \text{et} \quad \{X_k \, ; \, 1 \leq k \leq n\} = \{X_{(k)} \, ; \, 1 \leq k \leq n\} \; .$$

La v.a. $(X_{(k)})_{1 \leq k \leq n} : \Omega \to \mathbb{R}^n$ est $\mathscr{F}$-mesurable : c'est *statistique d'ordre de la suite* $(X_k)_{1 \leq k \leq n}$. □

La proposition suivante décrit la statistique d'ordre de variables indépendantes de loi diffuse.

**Proposition III.1.9** *Soit $X_1, \ldots, X_n$, des variables réelles indépendantes de même loi $\mu$ qui est supposée diffuse. Alors les assertions suivantes sont vérifiées.*

$(i)$ *On note $\mathbf{S}_n$ le groupe des permutation de $\{1, \ldots, n\}$. Il existe $\boldsymbol{\sigma} : \Omega \to \mathbf{S}_n$, une permutation aléatoire $\mathscr{F}$-mesurable dont la loi est uniforme, qui est indépendante de $(X_{(k)})_{1 \leq k \leq n}$, et qui satisfait*

$$\mathbf{P}\text{-}p.s. \quad X_{(k)} = X_{\boldsymbol{\sigma}(k)}, \quad 1 \leq k \leq n \; . \tag{III.5}$$

$(ii)$ *Pour toute fonction mesurable $f : \mathbb{R}^n \to \mathbb{R}_+$, on a*

$$\mathbf{E}\big[f(X_{(1)}, \ldots, X_{(n)})\big] = n! \int_{\{x_1 < \ldots < x_n\}} f(x_1, \ldots, x_n) \; \mu(dx_1) \ldots \mu(dx_n) \; .$$

$(iii)$ *On note $F(x) = \mu\big(\,]-\infty, x]\big)$, $x \in \mathbb{R}$, la fonction de répartition de $\mu$. Alors, pour tout $1 \leq k \leq n$ et pour toute fonction mesurable $g : \mathbb{R} \to \mathbb{R}_+$,*

$$\mathbf{E}\big[g(X_{(k)})\big] = \int_{\mathbb{R}} g(x) \, k \binom{n}{k} F(x)^{k-1} \big(1 - F(x)\big)^{n-k} \mu(dx) \; .$$

**Preuve :** on note $\Delta := \{(x, x) \, ; \, x \in \mathbb{R}\}$ la diagonale de $\mathbb{R}^2$ qui est un fermé de $\mathbb{R}^2$. Par Fubini et le théorème de transfert, si $i \neq j$,

$$\mathbf{P}(X_i = X_j) = \mu \otimes \mu(\Delta) = \int_{\mathbb{R}} \mu(dx) \mu(\{x\}) = 0$$

car la mesure $\mu$ est diffuse. On pose $A = \{X_j \neq X_k \, , \, 1 \leq j < k \leq n\}$, qui est l'événement que les variables $X_1, \ldots, X_n$ soient toutes distinctes. On a

$$\mathbf{P}(\Omega \backslash A) \leq \sum_{1 \leq i < j \leq n} \mathbf{P}(X_i = X_j) = 0 \; .$$

Donc $\mathbf{P}(A) = 1$. Pour toute permutation $\gamma \in \mathbf{S}_n$, on pose

$$A_\gamma = \big\{ X_{\gamma(1)} < X_{\gamma(2)} < \ldots < X_{\gamma(n)} \big\} \; .$$

Il est clair que les événements $A_\gamma$, $\gamma \in \mathbf{S}_n$ sont deux-à-deux disjoints et que $A = \bigcup_{\gamma\in\mathbf{S}_n} A_\gamma$. On définit alors la permutation aléatoire $\boldsymbol{\sigma}$ en posant $\boldsymbol{\sigma} = \gamma$ sur $A_\gamma$ et $\boldsymbol{\sigma} = \mathrm{Id}$ sur $\Omega\backslash A$. Comme $\mathbf{P}(A) = 1$, $\boldsymbol{\sigma}$ satisfait bien (III.5).

On observe ensuite que pour toute permutation $\gamma \in \mathbf{S}_n$, les v.a. $(X_{\gamma(k)})_{1\leq k\leq n}$ sont indépendantes et de même loi $\mu$. Donc

$$\big(X_{\gamma(1)}, X_{\gamma(2)}, \dots, X_{\gamma(n)}\big) \overset{\text{loi}}{=} (X_1, X_2, \dots, X_n) \ .$$

On en déduit alors que pour toute fonction mesurable $f : \mathbb{R}^n \to \mathbb{R}_+$ et toute permutation $\gamma\in\mathbf{S}_n$,

$$\begin{aligned}
\mathbf{E}\big[f(X_{(1)},\dots,X_{(n)})\mathbf{1}_{\{\boldsymbol{\sigma}=\gamma\}}\big] &= \mathbf{E}\big[f(X_{\gamma(1)},\dots,X_{\gamma(n)})\mathbf{1}_{A_\gamma}\big] \\
&= \mathbf{E}\big[f(X_1,\dots,X_n)\mathbf{1}_{\{X_1<X_2<\ldots<X_n\}}\big] \\
&= \int_{\{x_1<\ldots<x_n\}} f(x_1,\dots,x_n)\ \mu(dx_1)\dots\mu(dx_n) \ . 
\end{aligned} \tag{III.6}$$

En sommant sur $\gamma \in \mathbf{S}_n$, cela montre $(ii)$. On en déduit que $\int_{\{x_1<\ldots<x_n\}} \mu(dx_1)\dots\mu(dx_n) = 1/n!$ et (III.6) implique d'une part que $\boldsymbol{\sigma}$ est de loi uniforme sur $\mathbf{S}_n$ et d'autre part que

$$\mathbf{E}\big[f(X_{(1)},\dots,X_{(n)})\mathbf{1}_{\{\boldsymbol{\sigma}=\gamma\}}\big] = \mathbf{E}\big[f(X_{(1)},\dots,X_{(n)})\big]\mathbf{P}(\boldsymbol{\sigma}=\gamma) \ ,$$

ce qui implique que $\boldsymbol{\sigma}$ est indépendante de $(X_{(k)})_{1\leq k\leq n}$, ce qui termine la preuve de $(i)$.

Montrons $(iii)$ : en raisonnant comme précédemment on montre que pour tout $1\leq k\leq n$ et tout $x\in\mathbb{R}$,

$$\int_{\{x_1<\ldots x_{k-1}<x\}} \mu(dx_1)\dots\mu(dx_{k-1}) = \frac{1}{(k-1)!}\int_{]-\infty,x[^{k-1}} \mu(dx_1)\dots\mu(dx_{k-1}) = \frac{F(x)^{k-1}}{(k-1)!}.$$

De même, on a

$$\int_{\{x<x_{k+1}<\ldots x_n\}} \mu(dx_{k+1})\dots\mu(dx_n) = \frac{1}{(n-k)!}\int_{]x,\infty[^{n-k}} \mu(dx_{k+1})\dots\mu(dx_n) = \frac{(1-F(x))^{n-k}}{(n-k)!}.$$

Par $(ii)$ et Fubini, on a donc

$$\begin{aligned}
\mathbf{E}\big[g(X_{(k)})\big] &= n!\int_{\mathbb{R}} \mu(dx)\, g(x)\Big(\int_{\{x_1<\ldots x_{k-1}<x\}} \mu(dx_1)\dots\mu(dx_{k-1})\Big)\Big(\int_{\{x<x_{k+1}<\ldots x_n\}} \mu(dx_{k+1})\dots\mu(dx_n)\Big) \\
&= \frac{n!}{(k-1)!(n-k)!}\int_{\mathbb{R}} \mu(dx)\, g(x)F(x)^{k-1}\big(1-F(x)\big)^{n-k} \ ,
\end{aligned}$$

ce qui entraîne $(iii)$. ■

En particulier, lorsque les variables $X_1,\dots,X_n$ sont indépendantes et uniformes dans l'intervalle $[0,t]$, pour toute fonction $f : [0,t]^n \to \mathbb{R}$ mesurable bornée, on a

$$\mathbf{E}\left[f\left(U_1^{(n)},\dots,U_n^{(n)}\right)\right] = \frac{n!}{t^n}\int_{\{0\leq x_1<\ldots<x_n\leq t\}} f(x_1,\dots,x_n)\ dx_1\dots dx_n \ . \tag{III.7}$$

### III.1.b Une heuristique.

Pour résoudre le problème consistant à trouver une famille infinie dénombrable de points de $\mathbb{R}_+$ qui soit *totalement aléatoire* et *uniformément répartie*, on peut commencer par élaborer un modéle discret de la manière suivante : on fixe un pas $h > 0$ (qui tendra vers 0) et on discrétise $\mathbb{R}_+$ en la collection d'intervalles $I(h,n) = [nh,(n+1)h[$, $n \in \mathbb{N}$. On note $\Pi_h$ le nuage dénombrable de point "uniformément réparti" en supposant que chaque intervalle $I(n,h)$ contient *au plus un point de* $\Pi_h$. Comme on veut une répartition totalement aléatoire et uniforme, il est naturel de considérer que $\xi_{h,n} = \#(\Pi_h \cap [nh,(n+1)h[\,)$ sont des variables de Bernoulli indépendantes dont paramètre ne dépend que de $h$ : on le note $p(h)$ et on a $0 < p(h) < 1$. On a donc les propriétés suivantes.

(a) Pour tous entiers positifs $n_1 < n_2$, la variable entière $\#(\Pi_h \cap [n_1h, n_2h[\,)$ suit la loi binomiale de paramètres $n_2 - n_1$ et $p(h)$.

(b) Pour tous $n_1 < \ldots < n_p$, les variables $\#(\Pi_h \cap [n_1h, n_2h[\,), \ldots, \#(\Pi_h \cap [n_{p-1}h, n_ph[\,)$ sont indépendantes.

Par ailleurs on veut que $\Pi_h$, soit une version discrétisée d'un ensemble aléatoire dénombrable $\Pi$ sur $\mathbb{R}_+$ ; c'est-à-dire que l'on suppose que lorsque $h$ tend vers 0, $\Pi_h$ "tend" vers $\Pi$. Cela impose que $\Pi_h$ ne dépende que "peu" de $h$ lorsque $h$ est petit. En particulier, si on se fixe un intervalle $[a,b] \subset [0,\infty[$, on veut que la quantité

$$m_h(a,b) = \mathbf{E}\Big[\,\#\big(\,\Pi_h \cap [\,\lfloor a/h \rfloor h\,,\,\lfloor b/h \rfloor h\,[\,\big)\Big] = (\lfloor b/h \rfloor - \lfloor a/h \rfloor)p(h)\ ,$$

ne dépende pas trop de $h$, lorsque $h$ est petit. Cela conduit donc a imposer la chose suivante.

(c) $\quad p(h) = \theta h + o(h)$, où $\theta$ est un réel strictement positif fixé.

L'approximation binomiale-Poisson permet alors de dire que lorsque $h$ est petit, la variable $\#(\Pi_h \cap [a,b])$ est très proche d'une loi de Poisson de paramètre $\theta(a-b)$ et $m_h(a,b)$ est très proche de $\theta(a-b)$.

Intéressons-nous au numéro de l'intervalle où tombe le premier point de $\Pi_h$ : on pose

$$T_1(h) = \inf\{n \geq 0\ :\ \xi_{h,n} = 1\ \}\ .$$

Il est facile de voir que $T_1(h)$ suit une loi géométrique : $\mathbf{P}(T_1(h) = n) = p(h)(1-p(h))^n$, $n \in \mathbb{N}$. On introduit aussi les numéros d'intervalle des points successifs. On les note $(T_k(h))_{k\geq 1}$ et on les définit récursivement par

$$T_{k+1}(h) = \inf\{n > T_k(h)\ :\ \xi_{h,n} = 1\ \}\ .$$

On a clairement les propriétés suivantes.

(d) $\{\lfloor x/h \rfloor h; x \in \Pi_h\}$ est exactement l'ensemble $\{hT_k(h); k \geq 1\}$.

(e) Les variables $\mathscr{E}_1(h) = T_1(h)$, ... , $\mathscr{E}_k(h) = T_k(h) - T_{k-1}(h) - 1$, ... , sont i.i.d.

Il est ensuite facile de constater que $\lim_{h\to 0} \mathbf{P}(h\mathscr{E}_k(h) > t) = \exp(-\theta t)$, $t \in \mathbb{R}_+$. Autrement dit lorsque $h$ est proche de 0, les variables indépendantes $(h\mathscr{E}_k(h))_{k\geq 1}$ sont proches d'une suite i.i.d. de variables exponentielles de paramètre $\theta$.

Résumons les résultats apportés par cette brève heuristique : on cherche à définir un ensemble aléatoire de points $\mathbf{\Pi}$ qui soit infini dénombrable, sans point d'accumulation, "totalement aléatoire" et "uniformément réparti" sur $\mathbb{R}_+$. Si un tel objet existe, alors il est raisonnable de penser qu'il satisfait *nécessairement* les propriétés suivantes.

(i) Il exite $\theta \in \mathbb{R}_+^*$, tel que pour tous réels positifs $a < b$, la variable $\#(\mathbf{\Pi} \cap [a, b])$ comptant le nombre de points de $\mathbf{\Pi}$ tombant dans $[a, b]$, suit une loi de Poisson de paramètre $(b - a)\theta$.

(ii) Pour tous réels positifs $t_1 < \ldots < t_p$, les variables

$$\#(\mathbf{\Pi} \cap [0, t_1[) \ , \ \#(\mathbf{\Pi} \cap [t_1, t_2[) \ , \ \ldots \ , \ \#(\mathbf{\Pi} \cap [t_{p-1}, t_p[)$$

sont indépendantes.

(iii) $\mathbf{\Pi}$ peut être indexé en croissant : $\mathbf{\Pi} = \{T_1 < T_2 < \ldots < T_n < \ldots\}$. De plus les variables $T_1, T_2 - T_1$, ..., $T_n - T_{n-1}$, ... sont des variables exponentielles indépendantes de paramètre $\theta$.

Comme nous le montrons plus loin, ces propriétés sont redondantes : (i) et (ii) sont des conséquences de (iii). La propriété (iii) permet de donner une définition simple de $\mathbf{\Pi}$, qui est le point de départ de la section suivante.

### III.1.c Construction des processus Poisson homogènes sur $\mathbb{R}_+$.

**Définition III.1.10** Soit $\theta \in \mathbb{R}_+^*$ et $(\mathscr{E}_n)_{n\geq 1}$, une suite i.i.d. de variables exponentielles de paramètre $\theta$. On pose

$$\forall n \geq 1, \quad T_n = \mathscr{E}_1 + \ldots + \mathscr{E}_n, \quad \mathbf{\Pi} = \big\{T_n \, ; \, n \in \mathbb{N}^*\big\} \quad \text{et} \quad N_t = \#\big(\mathbf{\Pi} \cap [0, t]\big), \ t \in \mathbb{R}_+.$$

On remarque que $N_t = \sum_{n\geq 1} \mathbf{1}_{[0,t]}(T_n)$, est une variable mesurable. Le processus $(N_t)_{t\geq 0}$ est appelé *processus de Poisson homogène d'intensité $\theta$*. L'ensemble aléatoire discret $\mathbf{\Pi}$ est appelé *nuage Poissonnien homogène d'intensité $\theta$ sur $\mathbb{R}_+$*. □

Par commodité, on pose $\mathbb{N}_\infty = \mathbb{N} \cup \{\infty\}$ et pour tout ensemble Borélien $B \in \mathscr{B}(\mathbb{R}_+)$, on pose

$$N_B(\mathbf{\Pi}) = \# \, \mathbf{\Pi} \cap B = \sum_{n\geq 1} \mathbf{1}_B(T_n) \ \in \ \mathbb{N}_\infty \ .$$

On remarque que $N_B(\mathbf{\Pi})$ est une variable $\mathscr{F}$-mesurable. C'est la *fonction de comptage de $\mathbf{\Pi}$ en $B$*. Par la loi des grands nombre, $\mathbf{P}(\lim_{n\to\infty} T_n = \infty) = 1$. Par conséquent si $B$ est bornée, on a $\mathbf{P}(N_B(\mathbf{\Pi}) < \infty) = 1$.

On constate facilement que pour tout $\omega \in \Omega$, $t \mapsto N_t(\omega)$ est croissant et *continu à droite et admet une limite à gauche en tout point*, c'est-à-dire que $(N_t)_{t\geq 0}$ est càdlàg. On pose $\Delta N(t) = N_t - N_{t-}$, qui représente le saut éventuel au temps $t$ du processus. On a alors, $\mathbf{\Pi} = \{t \in \mathbb{R}_+ : \Delta N(t) \neq 0\}$ et $N_{[0,t]}(\mathbf{\Pi}) = N_t$, $t \in \mathbb{R}_+$. Par conséquent, se donner $(N_t)_{t\geq 0}$ équivaut à se donner $\mathbf{\Pi}$.

Le théorème suivant montre que $\mathbf{\Pi}$ peut se voir comme un ensemble infini dénombrable, complètement aléatoire et uniformément réparti sur $\mathbb{R}_+$.

**Théorème III.1.11** *On note $\ell$ la mesure de Lebesgue sur $\mathbb{R}_+$. Soit $\mathbf{\Pi}$ un nuage Poissonnien homogène d'intensité $\theta$ et soit $(N_t)_{t\geq 0}$. Les propriétés suivantes sont vérifiées.*

*(i) Pour tout $s < t$, $N_t - N_s = N_{]s,t]}(\mathbf{\Pi})$ suit une loi de Poisson de paramètre $\theta(t - s)$.*

*(ii) Pour tous $t_1 < \ldots < t_p$, les variables $N_{t_1}, N_{t_2} - N_{t_1}, \ldots, N_{t_p} - N_{t_{p-1}}$ sont indépendantes.*

*(iii) Pour tout $t \in \mathbb{R}_+^*$ et tout $n \in \mathbb{N}^*$, l'ensemble $\mathbf{\Pi} \cap [0, t]$ sous $\mathbf{P}(\ \cdot \mid N_t = n)$ a même loi que $\{U_1, \ldots, U_n\}$, où $U_1$, ... , $U_n$ sont i.i.d. uniformes sur $[0, t]$.*

*(iv) Soient $A_1, \ldots, A_p \in \mathscr{B}(\mathbb{R}_+)$, disjoints deux-à-deux. Alors, $N_{A_1}(\mathbf{\Pi}), \ldots, N_{A_p}(\mathbf{\Pi})$ sont des variables de Poisson indépendantes de paramètres respectifs $\theta\ell(A_1), \ldots, \theta\ell(A_p)$.*

**Remarque III.1.12** On voit que $(iv)$ implique $(i)$ et $(ii)$, qui sont plus élémentaires et qui ne concernent que $(N_t)_{t\geq 0}$. Puisque dans de nombreuses applications, seul $(N_t)_{t\geq 0}$ intervient, nous avons pris la peine d'expliciter $(i)$ et $(ii)$. □

**Preuve du théorème :** montrons d'abord que pour tout $t \in \mathbb{R}_+$, $N_t$ suit une loi de Poisson de paramètre $\theta t$. Soit $n \in \mathbb{N}^*$. On observe que $T_n$ est indépendante de $\mathscr{E}_{n+1}$ et que $T_n$ suit une loi d'Erlang de paramètres $(n, \theta)$. Par conséquent,

$$\begin{aligned} \mathbf{P}(N_t = n) &= \mathbf{P}(T_n \leq t < T_n + \mathscr{E}_{n+1}) = \int_0^t du\, \theta^n \frac{u^{n-1}}{(n-1)!} e^{-\theta u} \left( \int_{t-u}^{\infty} dv\, \theta e^{-\theta v} \right) \\ &= \frac{\theta^n}{(n-1)!} e^{-\theta t} \int_0^t du\, u^{n-1} = \frac{(t\theta)^n}{n!} e^{-\theta t}, \end{aligned}$$

ce qui montre le résultat désiré.

Nous allons montrer $(iii)$ et ensuite prouver que $(iii)$ implique $(iv)$, ce qui est suffisant d'après la remarque précédente. Soit $n \in \mathbb{N}^*$. Soit $F : \mathbb{R}^n \to \mathbb{R}$, une fonction mesurable bornée. On a les égalités suivantes

$$\begin{aligned} \mathbf{E}\left[F(T_1, \ldots, T_n); N_t = n\right] &= \mathbf{E}\left[F(T_1, \ldots, T_n)\,;\, T_n \leq t < T_n + \mathscr{E}_{n+1}\right] \\ = \int_{\{s_1+\ldots+s_n \leq t\}} ds_1 \ldots ds_n\, \theta^n e^{-\theta(s_1+\ldots+s_n)}\, & F(s_1, s_1+s_2, \ldots, s_1+\ldots+s_n) \int_{t-s_1+\ldots+s_n}^{\infty} ds_{n+1}\, \theta e^{-\theta s_{n+1}} \\ &= \theta^n e^{-\theta t} \int_{\{s_1+\ldots+s_n \leq t\}} ds_1 \ldots ds_n F(s_1, s_1+s_2, \ldots, s_1+\ldots+s_n) \\ &= \theta^n e^{-\theta t} \int_{\{0 \leq x_1 < \ldots < x_n \leq t\}} dx_1 dx_2 \ldots dx_n F(x_1, x_2, \ldots, x_n)\ . \end{aligned}$$

On en déduit que

$$\mathbf{E}\left[F(T_1, \ldots, T_n) \mid N_t = n\right] = \frac{n!}{t^n} \int_{\{0 \leq x_1 < \ldots < x_n \leq t\}} dx_1 dx_2 \ldots dx_n F(x_1, x_2, \ldots, x_n)\ , \tag{III.8}$$

ce qui implique $(iii)$ par (III.7).

Montrons $(iv)$ : soit $t \in \mathbb{R}_+^*$. On suppose tout d'abord que $A_1, \ldots, A_p$ sont des Boréliens disjoints *inclus* dans $[0, t]$. On fixe $n \geq 1$. Le point $(iii)$ implique que

$$\big( N_{A_1}(\mathbf{\Pi}), \ldots, N_{A_p}(\mathbf{\Pi}) \big) \text{ sous } \mathbf{P}(\,\cdot\,|N_t = n) \overset{(loi)}{=} \big( \#(\Pi_n \cap A_1), \ldots, \#(\Pi_n \cap A_p) \big) \text{ sous } \mathbf{P}\ ,$$

où $\Pi_n = \{U_1, \ldots, U_n\}$, avec $(U_k, 1 \leq k \leq n)$ i.i.d. de loi uniforme sur $[0, t]$. Par commodité on pose $A_{p+1} = [0, t] \backslash \bigcup_{j=1}^p A_j$. Soient $u_1, \ldots, u_p \in \mathbb{R}$. Pour tout $1 \leq k \leq n$, on pose :

$$X_k = e^{iu_1} \mathbf{1}_{A_1}(U_k) + \ldots + e^{iu_p} \mathbf{1}_{A_p}(U_k) + \mathbf{1}_{A_{p+1}}(U_k)\ .$$

Les variables $(X_1, \ldots, X_n)$ sont indépendantes et de même loi. De plus on a

$$X_1 \ldots X_n = \prod_{k=1}^{p} \exp\big(iu_k \#(\Pi_n \cap A_k)\big)\ .$$

Par conséquent,

$$\mathbf{E}\Big[ \prod_{1 \leq k \leq p} \exp\left(iu_k N_{A_k}(\mathbf{\Pi})\right) \Big| N_t = n \Big] = \mathbf{E}[X_1 \ldots X_n] = \mathbf{E}[X_1]^n$$

$$= \left(e^{iu_1}\frac{\ell(A_1)}{t} + \ldots + e^{iu_p}\frac{\ell(A_p)}{t} + 1 - \frac{\ell(A_1)}{t} - \ldots \frac{\ell(A_p)}{t}\right)^n$$
$$= t^{-n}\left(t - (1-e^{iu_1})\ell(A_1) - \ldots - (1-e^{iu_p})\ell(A_p)\right)^n .$$

Donc

$$\mathbf{E}\Big[\prod_{1\le k\le p}\exp\left(iu_k N_{A_k}(\mathbf{\Pi})\right)\Big] = \sum_{n\ge 0}\mathbf{P}(N_t=n)t^{-n}\big(t-(1-e^{iu_1})\ell(A_1)-\ldots-(1-e^{iu_p})\ell(A_p)\big)^n$$
$$= e^{-\theta t}\sum_{n\ge 0}\frac{1}{n!}\left(\theta t - (1-e^{iu_1})\theta\ell(A_1) - \ldots - (1-e^{iu_p})\theta\ell(A_p)\right)^n$$
$$= \prod_{1\le k\le p}\exp\left(-\theta\ell(A_k)\left(1-e^{iu_k}\right)\right),$$

ce qui entraîne bien le point $(iv)$, dans le cas où $A_1,\ldots,A_p \subset [0,t]$. Le cas général découle facilement du fait que $\mathbf{P}$-p.s. pour tout $1\le k\le p$, $\lim_{t\to\infty}\uparrow N_{A_k\cap[0,t]}(\mathbf{\Pi}) = N_{A_k}(\mathbf{\Pi})$. ■

## EXERCICES.

**Exercice III.1** *Paradoxe de l'inspection* Soit $(\mathscr{E}_n)_{n\ge1}$, une suite i.i.d. de variables de loi exponentielle de paramètre $\theta\in\mathbb{R}_+^*$. On pose

$$\forall n\in\mathbb{N}^*, \quad T_n = \mathscr{E}_1+\ldots+\mathscr{E}_n \quad \text{et} \quad \mathbf{\Pi} = \left\{T_n\,;\, n\in\mathbb{N}^*\right\},$$

qui est donc un nuage Poissonnien homogène sur $\mathbb{R}_+$ d'intensité $\theta$. Pour tout $t\in\mathbb{R}_+$, on note $N_t = \sum_{n\ge1}\mathbf{1}_{[0,t]}(T_n)$. On s'intéresse aux variables $R_t$ et $D_t$ définies par

$$R_t = t - T_{N_t} \quad \text{et} \quad D_t = T_{N_t+1} - t ,$$

qui sont telles que $R_t + D_t = \mathscr{E}_{N_t+1}$.

1. Montrer que pour tout $t\in\mathbb{R}_+$, $D_t$ suit une loi exponentielle de paramètre $\theta$, c'est-à-dire que $D_t$ a même loi que $\mathscr{E}_1$.
2. Montrer que $R_t$ a même loi que le maximum de $t$ et d'une exponentielle de paramètre $\theta$, c'est-à-dire que $D_t$ a même loi que $t\vee\mathscr{E}_1$.
3. Calculer en fonction de $t$ et de $\theta$, l'espérance $\mathbf{E}[\mathscr{E}_{N_t+1}]$ et remarquer que

$$\mathbf{E}[\mathscr{E}_{N_t+1}] > \frac{1}{\theta} = \mathbf{E}[\mathscr{E}_1] ,$$

   ce qui peut sembler "paradoxal" à première vue.
4. Montrer que $D_t$ et $R_t$ sont indépendantes et prouver que lorsque $t$ tend vers l'infini, $\mathscr{E}_{N_t+1}$ converge vers une loi de densité $s\mapsto\theta^2 s e^{-\theta s}$. □

**Exercice III.2** (*Statistiques d'ordre et processus de Poisson homogènes*) Soit $(U_n)_{n\ge1}$, une suite de variables i.i.d. de loi uniforme sur $[0,1]$. Pour tout $n\ge2$, on rappelle la notation $(U_1^{(n)},\ldots,U_n^{(n)})$ pour la statistique d'ordre associée à $(U_1,\ldots,U_n)$. Pour des raisons pratiques, on pose $U_0^{(n)}=0$ et $U_{n+1}^{(n)}=1$. On se donne également $(\mathscr{E}_n)_{n\ge1}$, une suite i.i.d. de variables de loi exponentielle de paramètre 1, et on pose

$$\forall n\in\mathbb{N}^*, \quad T_n = \mathscr{E}_1+\ldots+\mathscr{E}_n \quad \text{et} \quad \mathbf{\Pi} = \left\{T_n\,;\, n\in\mathbb{N}^*\right\},$$

qui est donc un nuage Poissonnien homogène sur $\mathbb{R}_+$ d'intensité 1.

1. Montrer que pour tout $n\ge2$, on a l'identité en loi suivante

$$(T_1/T_{n+1},\ldots,T_n/T_{n+1}) \overset{(loi)}{=} \left(U_1^{(n)},\ldots,U_n^{(n)}\right) .$$

2. Montrer que pour tout $k\in\mathbb{N}^*$, on a

$$n\,U_k^{(n)} \xrightarrow[n\to\infty]{\text{(loi)}} T_k .$$

3. Montrer que pour tout $x \in \mathbb{R}_+^*$
$$\frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{\{n(U_k^{(n)}-U_{k-1}^{(n)})>x\}} \quad \overset{\text{(proba)}}{\underset{n\to\infty}{\longrightarrow}} \quad e^{-x} \; .$$
4. Montrer que
$$\frac{n}{\log n} \max_{1\leq k\leq n+1} (U_k^{(n)} - U_{k-1}^{(n)}) \quad \overset{\text{(proba)}}{\underset{n\to\infty}{\longrightarrow}} \quad 1 \; .$$
5. Montrer que
$$n^2 \min_{1\leq k\leq n} (U_k^{(n)} - U_{k-1}^{(n)}) \quad \overset{\text{(loi)}}{\underset{n\to\infty}{\longrightarrow}} \quad \text{exponentielle}(1) \; .$$

□

## III.2 Généralités sur les nuages de points.

On détaille dans cette section quelques résultats concernant les nuages de points aléatoires (mesurabilité, loi et indépendence).

### III.2.a Premières définitions.

**Définition III.2.1** (*Nuages, fonctions de comptage*) Soit $(E, \mathscr{E})$, un espace mesurable. On utilise la notation $\mathbb{N}_\infty = \mathbb{N} \cup \{\infty\}$.

(a) $\mathbf{S}_E := \{\pi \subset E : \pi \text{ est dénombrable}\}$ est *l'ensemble des nuages de points sur* $E$. On rappelle que *dénombrable* signifie *fini* (éventuellement vide) ou *en bijection avec* $\mathbb{N}$.

(b) Pour tout $B \in \mathscr{E}$, et tout $\pi \in \mathbf{S}_E$, on pose $N_B(\pi) = \#(\pi \cap B)$, qui est le nombre (éventuellement infini) de points de $\pi$ qui sont dans $B$. La fonction $N_B : \mathbf{S}_E \to \mathbb{N}_\infty$ est appelée la *fonction de comptage en* $B$.

(c) On note $\mathscr{S}_E$ la tribu engendrée sur $\mathbf{S}_E$ par les fonctions de comptage $N_B$, $B \in \mathscr{E}$.

(d) Soit $(\Omega, \mathscr{F})$, un espace mesurable et $\mathbf{\Pi} : \Omega \to \mathbf{S}_E$, une fonction. C'est un *nuage aléatoire* si cette fonction est $(\mathscr{F}, \mathscr{S}_E)$-mesurable, ce qui est équivalement à ce que pour tout $B \in \mathscr{E}$, $N_B(\mathbf{\Pi}) : \Omega \to \mathbb{N}_\infty$ soit $\mathscr{F}$-mesurable. □

**Exemple III.2.2** Soit $(\Omega, \mathscr{F})$, un espace mesurable. Soit $X_n : \Omega \to E$, $n \in \mathbb{N}^*$, une suite de variables $(\mathscr{F}, \mathscr{E})$-mesurables telle que
$$\forall \omega \in \Omega \, , \; \forall n > m \geq 1 \, , \quad X_n(\omega) \neq X_m(\omega) \; . \tag{III.9}$$
Soit $M : \Omega \to \mathbb{N}_\infty$, une variables $\mathscr{F}$-mesurable. On pose :
$$\mathbf{\Pi} = \begin{cases} \emptyset & \text{si } M = 0, \\ \{X_1, \ldots, X_M\} & \text{si } 0 < M < \infty, \\ \{X_n ; n \in \mathbb{N}^*\} & \text{si } M = \infty. \end{cases}$$
Pour tout $B \in \mathscr{E}$, (III.9) implique que
$$N_B(\mathbf{\Pi}) = \sum_{n\geq 1} \mathbf{1}_B(X_n) \mathbf{1}_{\{n \leq M\}}$$
qui est clairement $\mathscr{F}$-mesurable. Cela implique donc que $N_B(\mathbf{\Pi})$ est $\mathscr{F}$-mesurable et donc que $\mathbf{\Pi}$ est un nuage aléatoire. □

Dans le lemme suivant, on définit une caractéristique importante des nuages aléatoires, à savoir leur *intensité.*

**Lemme III.2.3** *Soit* $(\Omega, \mathscr{F}, \mathbf{P})$*, un espace de probabilité et* $\mathbf{\Pi}\!:\!\Omega\!\to\!\mathbf{S}_E$*, un nuage aléatoire. On pose*

$$\forall B \in \mathscr{E} , \quad \mu(B) = \mathbf{E}\big[N_B(\mathbf{\Pi})\big] \; .$$

*Alors,* $\mu\!:\!\mathscr{E}\!\to\![0,\infty]$ *est une mesure positive appelée* **intensité** *de* $\mathbf{\Pi}$.

**Preuve :** on observe que $N_\emptyset(\mathbf{\Pi})=0$, sur $\Omega$, donc $\mu(\emptyset)=0$. Soient $B_n\in\mathscr{E}$, $n\in\mathbb{N}$, deux-à-deux disjoints. On pose $B\!:=\!\bigcup_{n\in\mathbb{N}} B_n$. Par interversion série/intégrale pour la mesure de comptage $\#$ et pour $\mathbf{P}$, on a bien

$$\mu(B) = \mathbf{E}\big[N_B(\mathbf{\Pi})\big] = \mathbf{E}\Big[\sum_{n\in\mathbb{N}} N_{B_n}(\mathbf{\Pi})\Big] = \sum_{n\in\mathbb{N}} \mathbf{E}\big[N_{B_n}(\mathbf{\Pi})\big] = \sum_{n\in\mathbb{N}} \mu(B_n)$$

ce qui implique le résultat désiré. ■

**Définition III.2.4** *(Fonctions de comptage généralisées)* Soit $(E,\mathscr{E})$, un espace mesurable et $f:E\to[0,\infty]$, une fonction $\mathscr{E}$-mesurable. On pose

$$\forall \pi\in\mathbf{S}_E, \quad N_f(\pi) = \sum_{x\in\pi} f(x) \; .$$

On appelle $N_f\!:\!\mathbf{S}_E\!\to\![0,\infty]$ la *fonction de $f$-comptage*. On remarque que $N_B\!=\!N_{\mathbf{1}_B}$, pour tout $B\!\in\!\mathscr{E}$. □

Pour tous $k,n\!\in\!\mathbb{N}$, on pose $B_n\!:=\!\{f\geq n\}$ et $A_{k,n}\!:=\{k2^{-n}\!\leq\! f\!<\!(k+1)2^{-n}[\}$. Il est clair que pour tout $n\!\in\!\mathbb{N}$ fixé, les $A_{k,n}\in\mathscr{E}$, $k\in\mathbb{N}$, sont deux-à-deux disjoints. On pose $f_n\!:=\!n\mathbf{1}_{B_n}+\sum_{0\leq k<n2^n} k2^{-n}\mathbf{1}_{A_k,n}$. On rappelle que $0\!\leq\! f_n\!\leq\! f_{n+1}$ et $\sup_n f_n\!=\!f$. On remarque aussi que

$$N_{f_n} = nN_{B_n} + \sum_{0\leq k\leq n2^n} k2^{-n}N_{A_k,n} \; . \tag{III.10}$$

Cela montre que $N_{f_n}$ est $\mathscr{S}_E$-mesurable. Soit $\pi\!\in\!\mathbf{S}_E$ ; on note $\nu\!=\!\sum_{x\in\pi}\delta_x$ sa mesure empirique. On remarque que $N_f(\pi)\!=\!\int_E f\,d\nu$. Par convergence monotone relativement à $\nu$, on

$$\lim_{n\to\infty}\uparrow N_{f_n}(\pi) = N_f(\pi) \; . \tag{III.11}$$

On en déduit le premier point du lemme suivant.

**Lemme III.2.5** *Soit* $(E,\mathscr{E})$*, un espace mesurable. Soit* $f:E\to[0,\infty]$*, une fonction* $\mathscr{E}$*-mesurable. Alors, les assertions suivantes sont vérifiées.*

*(i) La fonction* $N_f\!:\!\mathbf{S}_E\!\to\![0,\infty]$ *est* $\mathscr{S}_E$*-mesurable.*

*(ii) Soit* $(\Omega,\mathscr{F},\mathbf{P})$*, un espace de probabilité et* $\mathbf{\Pi}:\Omega\to\mathbf{S}_E$*, un nuage aléatoire d'intensité* $\mu$*. Alors* $N_f(\mathbf{\Pi})$ *est une variable* $\mathscr{F}$*-mesurable et* $\mathbf{E}\big[N_f(\mathbf{\Pi})\big]=\int_E f(x)\,\mu(dx)$.

**Preuve :** par (III.10), $\mathbf{E}\big[N_{f_n}(\mathbf{\Pi})\big]=\sum_{0\leq k\leq n2^n} k2^{-n}\mu(A_{k,n})=\int_E f_n(x)\,\mu(dx)$, et on conclut par convergence monotone sous $\mathbf{P}$ et sous $\mu$. ■

Il arrive fréquement de *restreindre* un nuage de point à un certain sous ensemble $B$ et de prendre *l'image* d'un nuage de points par une fonction $\phi$ : modulo un hypothèse, ces deux opérations sont mesurables, comme le montre le lemme suivant.

**Lemme III.2.6** *Soit $(E,\mathscr{E})$, un espace mesurable. Les propriétés suivantes sont vérifiées.*

*(i) Soit $B \in \mathscr{E}$. Alors $R_B : \pi \in \mathbf{S}_E \mapsto \pi \cap B \in \mathbf{S}_E$ est $(\mathscr{S}_E, \mathscr{S}_E)$-mesurable.*

*(ii) Soit $(E',\mathscr{E}')$, un espace mesurable et soit $\phi : E \to E'$, une fonction $(\mathscr{E},\mathscr{E}')$-mesurable supposée injective. Alors $\Phi : \pi \in \mathbf{S}_E \mapsto \phi(\pi) \in \mathbf{S}_{E'}$ est $(\mathscr{S}_E, \mathscr{S}_{E'})$-mesurable.*

**Preuve :** on remarque que pour tout $C \in \mathscr{E}$, $N_C \circ R_B = N_{B\cap C}$ qui est donc $\mathscr{S}_E$-mesurable, ce qui implique $(i)$. Montrons $(ii)$ : puisque $\phi$ est injective, on remarque que pour tout $B'$ et tout nuage $\pi \in \mathbf{S}_E$, on a $\#(\{\phi(x); x \in \pi\} \cap B') = \#\{x \in \pi : \phi(x) \in B'\}$ et donc que $N_{B'}(\phi(\pi)) = N_{\phi^{-1}(B')}(\pi)$. On a $N_{B'} \circ \Phi = N_{\phi^{-1}(B')}$ qui est $\mathscr{S}_E$-mesurable, ce qui implique le résultat voulu. ■

Ces résultats se traduisent immédiatement pour les nuages aléatoires.

**Lemme III.2.7** *Soit $(\Omega,\mathscr{F},\mathbf{P})$, un espace de probabilité. Soit $(E,\mathscr{E})$, un espace mesurable. Soit $\mathbf{\Pi} : \Omega \to \mathbf{S}_E$, un nuage aléatoire. Les assertions suivantes sont vérifiées.*

*(i) Soit $B \in \mathscr{E}$. Alors $\mathbf{\Pi} \cap B$ est un nuage aléatoire. On munit $B$ de la tribu $\mathscr{E}(B) = \{A \cap B; A \in \mathscr{E}\}$. Alors $\mathbf{\Pi} \cap B$ est un nuage aléatoire sur $(B, \mathscr{E}(B))$ au sens de la définition III.2.1 (d).*

*(ii) Soit $(E',\mathscr{E}')$, un espace mesurable et soit $\phi : E \to E'$, une fonction $(\mathscr{E},\mathscr{E}')$-mesurable supposée injective. Alors $\phi(\mathbf{\Pi})$ est un nuage aléatoire sur $E'$ d'intensité $\mu' = \mu \circ \phi^{-1}$ qui est la mesure image de $\mu$ par $\phi$.*

### III.2.b Loi et indépendance de nuages aléatoires.

On voit un nuage aléatoire $\mathbf{\Pi}$ comme le processus de ses fonctions de comptage $(N_B(\mathbf{\Pi}))_{B\in\mathscr{E}}$, processus qui est indexés par $\mathscr{E}$.

**Définition III.2.8** On note $\mathscr{C}_E \subset \mathscr{S}_E$ la classe contenant tous les sous-ensembles $C$ de la forme

$$C = \big\{\pi \in \mathbf{S}_E :\ N_{B_1}(\pi) \geq k_1;\ \dots\ ;\ N_{B_q}(\pi) \geq k_q\ \big\}\ . \tag{III.12}$$

avec $q \in \mathbb{N}^*$, $B_1, \dots, B_q \in \mathscr{E}$, et $k_1, \dots, k_q \in \mathbb{N}$. □

**Lemme III.2.9** *La classe $\mathscr{C}_E$ est un pi-système engendrant $\mathscr{S}_E$ : $\sigma(\mathscr{C}_E) = \mathscr{S}_E$.*

**Preuve :** clairement $\mathscr{C}_E$ est stable par intersection et contient $\mathbf{S}_E = \{\pi \in \mathbf{S}_E : N_E(\pi) \geq 0\}$. De plus, la forme (III.12) des ensembles de $\mathscr{C}_E$ implique que $\mathscr{C}_E \subset \mathscr{S}_E$ et donc que $\sigma(\mathscr{C}_E) \subset \mathscr{S}_E$. On remarque ensuite que pour tout $B \in \mathscr{E}$, et pour tout $k \in \mathbb{N}$, on a $\{\pi \in \mathbf{S}_E : N_B(\pi) \geq k\} \in \mathscr{C}_E$. Par conséquent $N_B : \mathbf{S}_E \to \mathbb{N}_\infty$ est $\sigma(\mathscr{C}_E)$-mesurable. Or $\mathscr{S}_E$ est, par définition, la plus petite tribu sur $\mathbf{S}_E$ rendant mesurable les fonctions de comptage. Par conséquent $\mathscr{S}_E \subset \sigma(\mathscr{C}_E)$, ce qui termine la preuve du lemme. ■

Puisqu'on l'utilise plusieurs fois, on établit sous forme de lemme le résultat élémentaire suivant.

**Lemme III.2.10** *Soit $(E,\mathscr{E})$, un espace mesurable. Soient $B_1, \dots, B_q \in \mathscr{E}$. Alors il existe un entier $p \leq 2^q$ et $A_1, \dots, A_p \in \mathscr{E}$, disjoints deux-à-deux tels que tout $B_k$ est réunion de certains $A_\ell$, c'est-à-dire*

$$\forall k \in \{1, \dots, q\}\ ,\ \exists I_k \subset \{1, \dots, p\}\ :\quad B_k = \bigcup_{\ell \in I_k} A_\ell \quad \text{et donc} \quad N_{B_k} = \sum_{\ell \in I_k} N_{A_\ell}\ . \tag{III.13}$$

**Preuve** : on note $\mathscr{B}$ la classe des ensembles de la forme $\bigcap_{1\leq k\leq q} C_k$, où $C_k$ est soit $B_k$, soit son complémentaire $E\backslash B_k$ ; $\mathscr{B}$ compte au plus $2^q$ ensembles distincts, les ensembles de $\mathscr{B}$ sont dans $\mathscr{E}$ et ils forment une partition de $E$. On se donne explicitement $\mathscr{B}=\{A_1,\ldots,A_p\}$, avec $p=\#\mathscr{B}$. Pour tout $k\in\{1,\ldots,q\}$, on pose $I_k = \{\ell\in\{1,\ldots,p\} : B_k\cap A_\ell\neq\emptyset\}$ et on vérifie (III.13) facilement. ■

La proposition suivante caractérise la loi d'un nuage de points.

**Proposition III.2.11** *Soit $(E,\mathscr{E})$ un espace mesurable. Soient $(\Omega,\mathscr{F},\mathbf{P})$ et $(\Omega,\mathscr{F}',\mathbf{P}')$, des espaces de probabilité. Soient $\mathbf{\Pi}:\Omega\to\mathbf{S}_E$ et $\mathbf{\Pi}':\Omega'\to\mathbf{S}_E$, deux nuages aléatoires. Ils ont la même loi ssi pour tout $p\in\mathbb{N}^*$ et tous $A_1,\ldots,A_p\in\mathscr{E}$ disjoints deux-à-deux , on a*

$$(N_{A_1}(\mathbf{\Pi}),\ldots,N_{A_p}(\mathbf{\Pi})) \text{ sous } \mathbf{P} \overset{(loi)}{=} (N_{A_1}(\mathbf{\Pi}'),\ldots,N_{A_p}(\mathbf{\Pi}')) \text{ sous } \mathbf{P}' . \tag{III.14}$$

**Preuve :** il est clair que $\Psi:\pi\in\mathbf{S}_E\longmapsto(N_{A_1}(\pi),\ldots,N_{A_p}(\pi))\in\mathbb{N}_\infty^p$ est $\mathscr{S}_E$-mesurable car chaque composante de cette application l'est. Si $\mathbf{\Pi}$ et $\mathbf{\Pi}'$ ont même loi, alors il en est de même pour $\Psi(\mathbf{\Pi})$ et $\Psi(\mathbf{\Pi}')$, ce qui est exactement (III.14). Montrons la réciproque : par le lemme III.2.10, l'identité (III.14) implique pour tous $B_1,\ldots,B_q\in\mathscr{E}$, on a

$$(N_{B_1}(\mathbf{\Pi}),\ldots,N_{B_q}(\mathbf{\Pi})) \text{ sous } \mathbf{P} \overset{(loi)}{=} (N_{B_1}(\mathbf{\Pi}'),\ldots,N_{B_p}(\mathbf{\Pi}')) \text{ sous } \mathbf{P}' . \tag{III.15}$$

Cela implique que $\mathbf{P}(\mathbf{\Pi}\in C)=\mathbf{P}'(\mathbf{\Pi}'\in C)$, pour tout $C$ de la forme (III.12). Si on note $\nu$ et $\nu'$ les lois respectives de $\mathbf{\Pi}$ et de $\mathbf{\Pi}'$, cela montre que $\nu(C)=\nu'(C)$, pour tout $C\in\mathscr{C}_E$. Comme $\mathscr{C}_E$ est un pi-système engendrant $\mathscr{S}_E$ (lemme III.2.9), le théorème d'unicité du prolongement des mesures de probabilité implique que $\nu=\nu'$. ■

**Brefs rappels sur la notion d'indépendance.** Rappelons quelques définitions générales sur l'indépendance. (voir la section B.2, page 306, du chapitre B en appendice).

**1–** Soit $(\Omega,\mathscr{F},\mathbf{P})$, un espace de probabilité. Soient $\mathscr{R}_i\subset\mathscr{F}$, $i\in I$, une famille de classes de sous-ensembles mesurables. *Les classes $\mathscr{R}_i$, $i\in I$, sont (mutuellement) indépendantes (sous $\mathbf{P}$)* si pour tout sous-ensemble d'indices $J\subset I$, non-vide et fini, et pour tous $A_j\in\mathscr{R}_j$, $j\in J$, les événements $A_j$, $j\in J$ sont mutuellement indépendants. Autrement dit :

$$\forall J\subset I \ J \text{ fini}, \forall A_j\subset\mathscr{R}_j,\ j\in J, \qquad \mathbf{P}\Big(\bigcap_{j\in J} A_j\Big)=\prod_{j\in J}\mathbf{P}(A_j) .$$

On rappelle le résultat standard suivant.

**Proposition III.2.12** *On conserve les notations précédentes. Soit $\mathscr{P}_i\subset\mathscr{F}$, $i\in I$, une famille de pi-systèmes qui sont mutuellement indépendants. Alors, les tribus engendrées $\sigma(\mathscr{P}_i)$, $i\in I$, sont également mutuellement indépendantes.*

**Preuve :** voir la proposition B.2.5, page B.2.5 en appendice. ■

**2–** Soit $(G,\mathscr{G})$, un espace mesurable. Soit $Y:\Omega\to G$, une variable $(\mathscr{F},\mathscr{G})$-mesurable. On rappelle que la tribu engendrée par $Y$ est simplement donnée par $\sigma(Y)=\{\{Y\in A\};A\in\mathscr{G}\}$. Soit $\mathscr{C}$, un pi-système sur $G$ engendrant $\mathscr{G}$. On pose $\mathscr{P}_Y:=\{\{Y\in C\};C\in\mathscr{C}\}$. On voit immédiatement que $\mathscr{P}_Y$ est un pi-système sur $\Omega$ générant $\sigma(Y)$.

Soit $X_i:\Omega\to G$, $i\in I$, une famille de variables $(\mathscr{F},\mathscr{G})$-mesurables. *Les variables $X_i$, $i\in I$, sont indépendantes* si les tribus $\sigma(X_i)$, $i\in I$ sont indépendantes.

**Lemme III.2.13** *On reprend les notation précédentes. Il y a équivalence entre les deux assertions suivantes.*

*(i) Les variables $X_i$, $i \in I$, sont indépendantes.*

*(ii) $\mathbf{P}\big(X_{i_1} \in C_1; \ldots ; X_{i_n} \in C_n\big) = \mathbf{P}\big(X_{i_1} \in C_1\big) \ldots \mathbf{P}\big(X_{i_n} \in C_n\big)$, pour tous $i_1, \ldots, i_n \in I$, et tous $C_1, \ldots, C_n \in \mathscr{C}$.*

**Preuve :** il est clair que $(i)$ implique $(ii)$. Montrons la réciproque : on remarque que $(ii)$ implique que les pi-systèmes $\mathscr{P}_{X_i}$, $i \in I$ sont mutuellement indépendants. La proposition III.2.12 implique alors que les tribus $\sigma(\mathscr{P}_{X_i}) = \sigma(X_i)$, $i \in I$ sont mutuellement indépendantes, c'est-à-dire $(i)$. ∎

Après ces quelques rappels sur la notion d'indépendance, on peut énoncer le critère d'indépendance suivant pour les nuages aléatoires.

**Proposition III.2.14** *Soit $(E, \mathscr{E})$, un espace mesurable. Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité. Soit $\mathbf{\Pi}_i : \Omega \to \mathbf{S}_E$, $i \in I$, une famille de nuages aléatoires. Il y a équivalence entre les deux assertions suivantes.*

*(i) Les nuages $\mathbf{\Pi}_i$, $i \in I$, sont mutuellement indépendants.*

*(ii) Pour tous $A_1, \ldots A_p \in \mathscr{E}$, deux-à-deux disjoints les vecteurs aléatoires de $\mathbb{N}_\infty^p$, $(N_{A_k}(\mathbf{\Pi}_i))_{1 \leq k \leq p}$, $i \in I$, sont mutuellement indépendants.*

**Preuve :** il est clair que $\Psi : \pi \in \mathbf{S}_E \longmapsto (N_{A_1}(\pi), \ldots, N_{A_p}(\pi)) \in \mathbb{N}_\infty^p$ est $\mathscr{S}_E$-mesurable car chaque composante de cette application l'est. Si les nuages $\mathbf{\Pi}_i$, $i \in I$, sont mutuellement indépendants, alors il en est de même pour la famille de vecteurs aléatoires $\Psi(\mathbf{\Pi}_i)$, $i \in I$. Cela montre que $(i)$ implique $(ii)$.

Réciproquement, par le lemme III.2.10, pour tous $B_1, \ldots, B_q \in \mathscr{E}$, les vecteurs aléatoires $(N_{B_q}(\mathbf{\Pi}_i))_{1 \leq \ell \leq q}$, $i \in I$, sont indépendants. Soient $i_1, \ldots, i_n \in I$, et $C_1, \ldots, C_n \in \mathscr{C}_E$ qui par définition peuvent s'écrire

$$C_m = \big\{ \pi \in \mathbf{S}_E : N_{B_1}(\pi) \geq k_1(m); \ \ldots \ ; \ N_{B_q}(\pi) \geq k_q(m) \big\} ,$$

où $B_1, \ldots B_q \in \mathscr{E}$, et où pour tout $1 \leq m \leq n$, $k_1(m), \ldots, k_q(m) \in \mathbb{N}$. On en déduit donc que $\mathbf{P}\big(\mathbf{\Pi}_{i_1} \in C_1; \ldots ; \mathbf{\Pi}_{i_n} \in C_n\big) = \mathbf{P}\big(\mathbf{\Pi}_{i_1} \in C_1\big) \ldots \mathbf{P}\big(\mathbf{\Pi}_{i_n} \in C_n\big)$, et le lemme III.2.13 permet de conclure. ∎

## III.2.c Mesurabilité de certaines opérations sur les nuages de points.

Le but de cette section technique est de montrer la mesurabilité du *produit*, de l'*intersection* et de l'*union* de nuages aléatoires, ainsi que d'autres fonctions apparaissant naturellement. Or *en général, il n'est du tout clair que ces opérations soient mesurables* et les obstructions sont doubles : les espaces $(E, \mathscr{E})$ et $(E', \mathscr{E}')$ doivent satisfaites des conditions de régularité ; il est aussi nécessaire de se restreindre à des nuages finis sur chaque ensemble d'une partition fixée. Ces deux types de restrictions techniques ne sont pas très contraignantes dans la pratique.

**Hypothèses sur les espace mesurables considérés.** Précisons nos hypothèses sur les espaces sur lequels sont définis les nuages de points. Nous commençons par des hypothèses de mesurabilité simples.

**Définition III.2.15** (*Isomorphisme d'espaces mesurables*) Soient $(E, \mathscr{E})$ et $(E', \mathscr{E}')$, deux espaces mesurables. Ils sont dits *isomorphes* s'il existe une *bijection* $\phi : E \to E'$ qui est est $(\mathscr{E}, \mathscr{E}')$-mesurable et telle que sa réciproque $\phi^{-1} : E' \to E$ soit $(\mathscr{E}', \mathscr{E})$-mesurable. □

Une telle fonction $\phi$ est un *isomorphisme d'espaces mesurables* et l'application $B \in \mathscr{E} \mapsto \phi(B) \in \mathscr{E}'$ est une bijection de $\mathscr{E}$ sur $\mathscr{E}'$.

**Définition III.2.16** Soit $(E, \mathscr{E})$, un espace mesurable. On introduit les notions suivantes.

(a) $(E, \mathscr{E})$ contient les singletons si $\{x\} \in \mathscr{E}$ pour tout $x \in \mathscr{E}$.

(b) $(E, \mathscr{E})$ est dit *diagonal* si $\Delta \in \mathscr{E}^{\otimes 2}$, où $\Delta = \{(x,x); x \in E\}$ est la diagonale de $E^2$.

(c) $(E, \mathscr{E})$ est dit *mesurable séparable* s'il existe $A_n \in \mathscr{E}$, $n \in \mathbb{N}$, engendrant $\mathscr{E}$ : $\mathscr{E} = \sigma\big(\{A_n; n \in \mathbb{N}\}\big)$.

(d) $(E, \mathscr{E})$ est dit *mesurable séparé* si pour tous $x, y \in E$ distincts, il existe $A \in \mathscr{E}$ tel que $x \in A$ et $y \in E \backslash A$. □

Le théorème A.5.3, page 288, montre d'une part que

$$(E, \mathscr{E}) \textit{ séparable et séparé} \Longrightarrow (E, \mathscr{E}) \textit{ mesurable diagonal} \Longrightarrow (E, \mathscr{E}) \textit{ contient les singletons.}$$

Le théorème A.5.3, page 288, montre également le fait suivant :

> *Soit $(E, \mathscr{E})$, espace mesurable. Il est séparable et séparé si et seulement s'il existe $C \subset \mathbb{R}$ (pas nécessairement Borélien), tel que $(E, \mathscr{E})$ est isomorphe à $(C, \mathscr{B}(C))$, où $\mathscr{B}(C)$ est la tribu trace sur $C$ des Boréliens de $\mathbb{R}$ (voir la définition A.1.7 et le lemme A.1.8, page 270, pour cette notion).*

On suppose dans la suite que l'espace mesurable $(E, \mathscr{E})$ sur lequel sont définis les nuages de points sont *séparables et séparés* (définition III.2.16, page 239). Le type de restriction sur les nuages de points est donné par le lemme suivant.

**Définition III.2.17** Soit $(E, \mathscr{E})$, un ensemble mesurable. Soit $B_p \in \mathscr{E}$, $p \in \mathbb{N}$, une *partition de $E$*, que l'on note $\mathbf{B} := (B_p)_{p \in \mathbb{N}}$. On pose

$$\mathbf{S}_E^{\mathbf{B}} = \big\{\pi \in \mathbf{S}_E : \forall p \in \mathbb{N},\ N_{B_p}(\pi) < \infty\big\} \quad \text{et} \quad \mathscr{S}_E^{\mathbf{B}} = \big\{D \cap \mathbf{S}_E^{\mathbf{B}}\,;\ D \in \mathscr{S}_E\big\}.$$

$\mathbf{S}_E^{\mathbf{B}}$ est l'ensemble des nuages finis sur chaque ensemble de la partition $\mathbf{B}$. Clairement que $\mathbf{S}_E^{\mathbf{B}} \in \mathscr{S}_E$. La tribu $\mathscr{S}_E^{\mathbf{B}}$ est la tribu trace de $\mathscr{S}_E$ sur $\mathbf{S}_E^{\mathbf{B}}$. Comme $\mathbf{S}_E^{\mathbf{B}} \in \mathscr{S}_E$, on a bien $\mathscr{S}_E^{\mathbf{B}} \subset \mathscr{S}_E$. □

**Résultats de mesurabilité.** Commençons par résoudre le problème consistant à localiser de manière mesurable les points d'un nuage.

**Lemme III.2.18** ***(Localisation des points)*** *Soit $(E, \mathscr{E})$, un espace mesurable séparable et séparé. Soit $\mathbf{B} = (B_p)_{p \geq 1}$, une partition de $E$ en ensembles de $\mathscr{E}$. Il existe une famille de fonctions $Y_k^{(p)} : \mathbf{S}_E^{\mathbf{B}} \to E$, $p, k \geq 1$, qui sont $(\mathscr{S}_E, \mathscr{E})$-mesurables telles que*

$$\forall p \in \mathbb{N},\ \forall \pi \in \mathbf{S}_E^{\mathbf{B}}, \quad \pi \cap B_p = \big\{Y(p)_1(\pi), Y_2^{(p)}(\pi), \ldots, Y_{N_{B_p}(\pi)}^{(p)}(\pi)\big\} \quad \textit{dès que } N_{B_p}(\pi) \geq 1. \tag{III.16}$$

**Preuve :** comme $(E, \mathscr{E})$ est séparable et séparé, il existe $C \subset \mathbb{R}$ et $\phi : E \to C$ isomorphisme entre les espaces mesurables $(E, \mathscr{E})$ et $(C, \mathscr{B}(C))$. On fixe $p \in \mathbb{N}^*$ et $x^* \in B_p$. Soit $\pi \in \mathbf{S}_E^{\mathbf{B}}$ ; on pose $\ell = N_{B_p}(\pi)$ et on se donne $\pi$ comme l'ensemble $\{x_1, \ldots, x_\ell\}$ où l'indexation est telle que $\phi(x_1) < \ldots < \phi(x_\ell)$. On pose alors pour tout $k \in \mathbb{N}^*$, $Y_k^{(p)}(\pi) := x_k$ si $k \leq \ell$ et $Y_k^{(p)}(\pi) := x^*$ si $k > \ell$. On a clairement (III.16). Puisque $\phi$ est un isomorphisme, $Y_k^{(p)}$ est $(\mathscr{S}_E, \mathscr{E})$-mesurable ssi $\phi \circ Y_k^{(p)}$ est $(\mathscr{S}_E^{\mathbf{B}}, \mathscr{B}(C))$-mesurable. Pour montrer cela, on fixe $y \in \mathbb{R}$ et on pose $D := \phi^{-1}\big(\,]-\infty, y] \cap \phi(B_p)\big)$, qui est dans $\mathscr{E}$ ; on remarque que

$$\{\pi \in \mathbf{S}_E^{\mathbf{B}} : N_{B_p}(\pi) \geq k\} \cap \{\pi \in \mathbf{S}_E^{\mathbf{B}} : \phi \circ Y_k^{(p)} \leq y\} = \{\pi \in \mathbf{S}_E^{\mathbf{B}} : N_D(\pi) \geq k\} \in \mathscr{S}_E^{\mathbf{B}}.$$

D'autre part

$$\{\pi\in\mathbf{S}_E^{\mathbf{B}} : N_{B_p}(\pi)<k\}\cap\{\pi\in\mathbf{S}_E^{\mathbf{B}} : \phi\circ Y_k^{(p)}\leq y\} = \emptyset \text{ ou bien } \{\pi\in\mathbf{S}_E^{\mathbf{B}} : N_{B_p}(\pi)<k\}$$

selon que $\phi(x^*)>y$ ou que $\phi(x^*)\leq y$. Dans les deux cas, les ensembles sont dans $\mathscr{S}_E^{\mathbf{B}}$, ce qui entraîne que $\{\pi\in\mathbf{S}_E^{\mathbf{B}} : \phi\circ Y_k^{(p)}\leq y\}\in\mathscr{S}_E^{\mathbf{B}}$ pour tout $y\in\mathbb{R}$. Donc $\phi\circ Y_k^{(p)}$ est $(\mathscr{S}_E^{\mathbf{B}},\mathscr{B}(C))$-mesurable, ce qui permet de conclure. ■

**Lemme III.2.19** ***(produit, intersection, union de nuages)*** *Soient $(E,\mathscr{E})$ et $(E',\mathscr{E}')$, des espaces mesurables séparables et séparés. On munit $E\times E'$ de la tribu produit $\mathscr{E}\otimes\mathscr{E}'$, qui est également séparable et séparée. Soit $\mathbf{B}=(B_p)_{p\in\mathbb{N}}$, une partition de $E$ en ensembles de $\mathscr{E}$. Soit $\mathbf{B}'=(B'_q)_{q\in\mathbb{N}}$, une partition de $E'$ en ensembles de $\mathscr{E}'$. Alors, les assertions suivantes sont vérifiées.*

*$(i)$ L'application $(\pi,\pi')\in\mathbf{S}_E^{\mathbf{B}}\times\mathbf{S}_{E'}^{\mathbf{B}'}\longmapsto\pi\times\pi'\in\mathbf{S}_{E\times E'}$ est $\big(\mathscr{S}_E^{\mathbf{B}}\otimes\mathscr{S}_{E'}^{\mathbf{B}'},\mathscr{S}_{E\times E'}\big)$-mesurable.*

*On suppose ensuite que $(E,\mathscr{E})=(E',\mathscr{E}')$.*

*$(ii)$ L'application $(\pi,\pi')\in\mathbf{S}_E^{\mathbf{B}}\times\mathbf{S}_E^{\mathbf{B}'}\longmapsto\pi\cap\pi'\in\mathbf{S}_E$ est $\big(\mathscr{S}_E^{\mathbf{B}}\otimes\mathscr{S}_E^{\mathbf{B}'},\mathscr{S}_E\big)$-mesurable.*

*$(iii)$ L'application $(\pi,\pi')\in\mathbf{S}_E^{\mathbf{B}}\times\mathbf{S}_E^{\mathbf{B}'}\longmapsto\pi\cup\pi'\in\mathbf{S}_E$ est $\big(\mathscr{S}_E^{\mathbf{B}}\otimes\mathscr{S}_E^{\mathbf{B}'},\mathscr{S}_E\big)$-mesurable.*

**Preuve :** pour tous $p,k\in\mathbb{N}^*$, soient $Y_k^{(p)}:\mathbf{S}_E^{\mathbf{B}}\to E$ qui est $(\mathscr{S}_E,\mathscr{E})$-mesurable et soit $Z_k^{(p)}:\mathbf{S}_{E'}^{\mathbf{B}'}\to E'$, qui est $(\mathscr{S}_{E'},\mathscr{E}')$-mesurable telles que $\pi\cap B_p=\{Y_k^{(p)}(\pi);1\leq k\leq N_{B_p}(\pi)\}$ et $\pi'\cap B'_p=\{Z_k^{(p)}(\pi');1\leq k\leq N_{B'_p}(\pi')\}$. Il est clair que $(\pi,\pi')\in\mathbf{S}_E^{\mathbf{B}}\times\mathbf{S}_{E'}^{\mathbf{B}'}\mapsto(Y_k^{(p)}(\pi),Z_{k'}^{(p')}(\pi'))\in E\times E'$ est $(\mathscr{S}_E^{\mathbf{B}}\otimes\mathscr{S}_{E'}^{\mathbf{B}'},\mathscr{E}\otimes\mathscr{E}')$-mesurable. Donc pour tout $A\in\mathscr{E}\otimes\mathscr{E}'$

$$(\pi,\pi')\in\mathbf{S}_E^{\mathbf{B}}\times\mathbf{S}_{E'}^{\mathbf{B}'}\longmapsto N_A(\pi\times\pi')=\sum_{p,p',k,k'\geq1}\mathbf{1}_{\{k\leq N_{B_p}(\pi);k'\leq N_{B'_p}(\pi')\}}\mathbf{1}_A\big(Y_k^{(p)}(\pi),Z_{k'}^{(p')}(\pi')\big)$$

est $\mathscr{S}_E^{\mathbf{B}}\otimes\mathscr{S}_{E'}^{\mathbf{B}'}$-mesurable, ce qui entraîne $(i)$.

On suppose ensuite $(E,\mathscr{E})=(E',\mathscr{E}')$ mais on garde les notations précédentes pour la localisation des points. On rappelle que $\Delta:=\{(x,x);\,x\in E\}\in\mathscr{E}^{\otimes2}$ car $(E,\mathscr{E})$ est séparable et séparé. Soit $B\in\mathscr{E}$, on remarque alors que

$$(\pi,\pi')\in\mathbf{S}_E^{\mathbf{B}}\times\mathbf{S}_E^{\mathbf{B}'}\longmapsto N_B(\pi\cap\pi')=\sum_{p,p',k,k'\geq1}\mathbf{1}_{\{k\leq N_{B_p}(\pi);k'\leq N_{B'_p}(\pi')\}}\mathbf{1}_B(Y_k^{(p)}(\pi))\mathbf{1}_\Delta\big(Y_k^{(p)}(\pi),Z_{k'}^{(p')}(\pi')\big)$$

est $\mathscr{S}_E^{\mathbf{B}}\otimes\mathscr{S}_{E'}^{\mathbf{B}'}$-mesurable, ce qui entraîne $(ii)$.

Le point $(iii)$ se déduit de $(ii)$ comme suit : pour tout $C\in\mathscr{E}$, tout $p\in\mathbb{N}$, et tous $(\pi,\pi')\in\mathbf{S}_E^{\mathbf{B}}\times\mathbf{S}_E^{\mathbf{B}'}$, on peut écrire $N_{C\cap B_p}(\pi\cup\pi')=N_{C\cap B_p}(\pi)+N_{C\cap B_p}(\pi')-N_{C\cap B_p}(\pi\cap\pi')$ car $N_{C\cap B_p}(\pi\cap\pi')\leq N_{B_p}(\pi)<\infty$, et $(ii)$ implique que $(\pi,\pi')\in\mathbf{S}_E^{\mathbf{B}}\times\mathbf{S}_E^{\mathbf{B}}\mapsto N_{C\cap B_p}(\pi\cup\pi')$ est $\mathscr{S}_E^{\mathbf{B}}\otimes\mathscr{S}_E^{\mathbf{B}'}$-mesurable. Or $N_C(\pi\cup\pi')=\sum_{p\in\mathbb{N}}N_{C\cap B_p}(\pi\cup\pi')$. Donc $(\pi,\pi')\mapsto N_C(\pi\cup\pi')$ est $\mathscr{S}_E^{\mathbf{B}}\otimes\mathscr{S}_E^{\mathbf{B}'}$-mesurable. ■

**Lemme III.2.20** *Soit $(E,\mathscr{E})$, un espace mesurable séparable et séparé. Soit $\mathbf{B}=(B_p)_{p\in\mathbb{N}}$, une partition $\mathscr{E}$-mesurable de $E$. Les assertions suivantes sont vérifiées.*

*$(i)$ La fonction $x\in E\mapsto\{x\}\in\mathbf{S}_E^{\mathbf{B}}$ est $(\mathscr{E},\mathscr{S}_E^{\mathbf{B}})$-mesurable.*

*$(ii)$ La fonction $(x,\pi)\in E\times\mathbf{S}_E^{\mathbf{B}}\mapsto\pi\cup\{x\}\in\mathbf{S}_E$ est $(\mathscr{E}\otimes\mathscr{S}_E^{\mathbf{B}},\mathscr{S}_E)$-mesurable.*

$(iii)$ *La fonction* $(x_1,\ldots,x_n,\pi)\in E^n\times \mathbf{S}_E^{\mathbf{B}} \mapsto \pi\cup\{x_1,\ldots,x_n\}\in \mathbf{S}_E$ *est* $(\mathscr{E}^{\otimes n}\otimes \mathscr{S}_E^{\mathbf{B}},\mathscr{S}_E)$*-mesurable.*

$(iv)$ ***(Mesurabilité des fonctionnelles de type Palm)*** *Soit* $(E_0,\mathscr{E}_0)$*, un espace mesurable. Soit* $n\in\mathbb{N}^*$ *et* $F\colon E_0\times E^n\times \mathbf{S}_E\to[0,\infty]$*, une fonction* $\mathscr{E}_0\otimes\mathscr{E}^{\otimes n}\otimes\mathscr{S}_E$*-mesurable. Pour tous* $(z,\pi)\in E_0\times \mathbf{S}_E$*, on pose*

$$\Phi_F^{(n)}(z,\pi):=\sum_{\substack{x_1,\ldots,x_n\in\pi\\ \text{distincts}}}F(z;\,x_1,\ldots,x_n,\,\pi\backslash\{x_1,\ldots,x_n\})\quad \textit{and}\quad \Lambda_F^{(n)}(z,\pi)=\sum_{\substack{x_1,\ldots,x_n\in\pi\\ \text{distincts}}}F(z;\,x_1,\ldots,x_n,\,\pi)$$

*avec* $\Phi_F^{(n)}(z,\pi)=\Lambda_F^{(n)}(z,\pi)=0$*, pour tout* $z\in E_0$ *si* $\#\pi<n$*. Alors, les restrictions de* $\Phi_F^{(n)}$ *et de* $\Lambda_F^{(n)}$ *à* $E_0\times\mathbf{S}_E^{\mathbf{B}}$ *sont* $(\mathscr{E}_0\otimes\mathscr{S}_E^{\mathbf{B}})$*-mesurables.*

**Preuve :** on note $\psi$ la fonction du $(i)$. Pour tout $B\in\mathscr{E}$, $(N_B\circ\psi)(x)=\mathbf{1}_B(x)$. Cela montre que $N_B\circ\psi$ est $\mathscr{E}$-mesurable pour tout $B\in\mathscr{E}$, ce qui prouve $(i)$. Montrons $(ii)$ : par $(i)$, $(x,\pi)\in E\times\mathbf{S}_E^{\mathbf{B}}\mapsto(\{x\},\pi)\in\mathbf{S}_E^{\mathbf{B}}\times\mathbf{S}_E^{\mathbf{B}}$ est mesurable. Par le lemme III.2.19 $(iii)$, $(\{x\},\pi)\in\mathbf{S}_E^{\mathbf{B}}\times\mathbf{S}_E^{\mathbf{B}}\mapsto\pi\cup\{x\}$ est également mesurable, ce qui prouve $(ii)$. Par $(i)$ et $(ii)$, combinés à une simple récurrence, $(x_1,\ldots,x_n)\mapsto\{x_1,\ldots,x_n\}$ est $(\mathscr{E}^{\otimes n},\mathscr{S}_E^{\mathbf{B}})$-mesurable et le lemme III.2.19 $(iii)$ termine la preuve du point $(iii)$.

Montrons $(iv)$ pour cela on se donne pour tous $p,k\in\mathbb{N}^*$, soient $Y_k^{(p)}:\mathbf{S}_E^{\mathbf{B}}\to E$ comme dans le lemme III.2.18 de localisation des points. On pose $B_p^*=E\backslash B_p$ : le lemme III.2.6 $(i)$ de restriction implique que $\pi\mapsto\pi\cap B_p^*$ est mesurable. Soient $1\leq i_1<\ldots<i_n\leq r$. On pose $S=\{1,\ldots,r\}\backslash\{i_1,\ldots,i_n\}$ et

$$f^p_{i_1,\ldots,i_n;r}(z,\pi)=F\big(z\,;\,Y_{i_1}^{(p)}(\pi),Y_{i_1}^{(p)}(\pi),Y_{i_n}^{(p)}(\pi)\,,\,(\pi\cap B_p^*)\cup\{Y_j^{(p)};j\in S\}\big)$$

Le point $(iii)$ et ce qui précède implique que $f^p_{i_1,\ldots,i_n;r}:E_0\times\mathbf{S}_E^{\mathbf{B}}\to[0,\infty]$ est $(\mathscr{E}_0\otimes\mathscr{S}_E^{\mathbf{B}})$-mesurable. On remarque ensuite que si $N_{B_p}(\pi)=r$, on a

$$f^p_{i_1,\ldots,i_n;r}(z,\pi)=F\big(z\,;\,Y_{i_1}^{(p)}(\pi),\ldots,Y_{i_n}^{(p)}(\pi)\,,\,\pi\backslash\{Y_{i_1}^{(p)}(\pi),\ldots,Y_{i_n}^{(p)}(\pi)\}\big)$$

et donc $\Phi_F^{(n)}(z,\pi)=\sum_{p\geq1}\sum_{r\geq i_n>\ldots>i_1\geq1}\mathbf{1}_{\{N_{B_p}(\pi)=r\}}f^p_{i_1,\ldots,i_n;r}(z,\pi)$, ce qui entraîne la mesurabilité de $\Phi_F^{(n)}$. La mesurabilité de $\Lambda_F^{(n)}$ en découle par le point $(iii)$. ■

### III.2.d Intersections et unions de nuages aléatoires.

Sauf mention explicite du contraire, on fait les hypothèses suivantes :

$(\Omega,\mathscr{F},\mathbf{P})$ supposé *complet* et $(E,\mathscr{E})$ est supposé séparable et séparé.

La complétude de $(\Omega,\mathscr{F},\mathbf{P})$ simplifie les énoncés où le produit, l'union ou l'intersection de nuages aléatoires interviennent. En effet, considérons $\mathbf{\Pi}:\Omega\to\mathbf{S}_E$, un nuage aléatoire d'intensité $\mu$ supposée *sigma-finie* : il existe une partition $\mathscr{E}$-mesurable $\mathbf{B}=(B_p)_{p\geq1}$ de $E$ telle que $\mathbf{E}[N_{B_p}(\mathbf{\Pi})]=\mu(B_p)<\infty$ pour tout $p\geq1$. Par conséquent, p.s. $\mathbf{\Pi}\in\mathbf{S}_E^{\mathbf{B}}$. On peut appliquer le lemme III.2.18, page 239, de localisation des points des nuages à $\mathbf{\Pi}$ : pour tout $k,p\geq1$, il existe des fonctions $Y_k^{(p)}:\mathbf{S}_E\to E$, qui sont $(\mathscr{S}_E,\mathscr{E})$-mesurables et telles que

$$\mathbf{P}\text{-p.s.}\quad\forall p\geq1,\quad \mathbf{\Pi}\cap B_p=\big\{Y_1^{(p)}(\mathbf{\Pi}),Y_2^{(p)}(\mathbf{\Pi}),\ldots,Y^{(p)}_{N_{B_p}(\mathbf{\Pi})}(\mathbf{\Pi})\big\}\ .$$

**Proposition III.2.21** *Soient* $(E,\mathscr{E})$ *et* $(E',\,\mathscr{E}')$*, des espaces mesurables séparables et séparés. On munit* $E\times E'$ *de la tribu produit* $\mathscr{E}\otimes\mathscr{E}'$*. Soient* $\mathbf{\Pi}:\Omega\to\mathbf{S}_E$ *et* $\mathbf{\Pi}':\Omega\to\mathbf{S}_{E'}$*, des nuages aléatoires d'intensités respectives* $\mu$ *et* $\mu'$ *supposées, sigma-finies. Alors, les assertions suivantes sont vérifiées.*

*(i)* $\mathbf{\Pi}\times\mathbf{\Pi}' : \Omega \to \mathbf{S}_{E\times E'}$ *est un nuage aléatoire.*

*(ii) Si* $\mathbf{\Pi}$ *et* $\mathbf{\Pi}'$ *sont indépendants, alors l'intensité de* $\mathbf{\Pi}\times\mathbf{\Pi}'$ *est* $\mu\otimes\mu'$.

**Preuve :** d'après le fait détaillé ci-dessus, il existe $\mathbf{B} = (B_p)_{p\in\mathbb{N}}$ et $\mathbf{B}' = (B'_q)_{q\in\mathbb{N}}$ des partitions mesurables de resp. $E$ et $E'$ telles que p.s. $\mathbf{\Pi} \in \mathbf{S}_E^{\mathbf{B}}$ et $\mathbf{\Pi}' \in \mathbf{S}_{E'}^{\mathbf{B}'}$. Le lemme III.2.19 $(i)$ implique alors que $\mathbf{\Pi}\times\mathbf{\Pi}'$ est p.s. égal à une variable $(\mathscr{F}, \mathscr{S}_{E\times E'})$-mesurable. Comme $(\Omega, \mathscr{F}, \mathbf{P})$ est supposé complet, $\mathbf{\Pi}\times\mathbf{\Pi}'$ est donc $(\mathscr{F}, \mathscr{S}_{E\times E'})$-mesurable : c'est un nuage aléatoire.

Montrons $(ii)$ : on note $\nu : \mathscr{E}\otimes\mathscr{E}' \to [0,\infty]$ l'intensité de $\mathbf{\Pi}\times\mathbf{\Pi}'$. Comme $\mathbf{\Pi}$ et $\mathbf{\Pi}'$ sont indépendants, pour tous $B\in\mathscr{E}$ et $B'\in\mathscr{E}'$

$$\nu(B\times B') = \mathbf{E}\big[N_{B\times B'}(\mathbf{\Pi}\times\mathbf{\Pi}')\big] = \mathbf{E}\big[N_B(\mathbf{\Pi})N_{B'}(\mathbf{\Pi}')\big] = \mathbf{E}\big[N_B(\mathbf{\Pi})\big]\mathbf{E}\big[N_{B'}(\mathbf{\Pi}')\big] = \mu(B)\mu'(B'). \quad \text{(III.17)}$$

Comme $\mu$ et $\mu'$ sont sigma-finies, $\mu\otimes\mu'$ est l'unique mesure $\nu$ satisfaisant (III.17). ■

**Proposition III.2.22** *Soient* $(E, \mathscr{E})$*, un espace mesurable séparable et séparé. Soient* $\mathbf{\Pi} : \Omega \to \mathbf{S}_E$ *et* $\mathbf{\Pi}' : \Omega \to \mathbf{S}_E$*, deux nuages aléatoires d'intensités respectives* $\mu$ *et* $\mu'$*, supposées sigma-finies. Alors, les assertions suivantes sont vérifiées.*

*(i)* $\mathbf{\Pi}\cap\mathbf{\Pi}' : \Omega \to \mathbf{S}_E$ *est un nuage aléatoire.*

*(ii)* **(Principe de disjonction)** *Si* $\mathbf{\Pi}$ *et* $\mathbf{\Pi}'$ *sont indépendants et si* $\mu$ *(ou* $\mu'$*) est diffuse, alors p.s.* $\mathbf{\Pi}\cap\mathbf{\Pi}' = \emptyset$.

**Preuve :** on raisonne comme dans la preuve du point $(i)$ de la proposition précédente III.2.21 en utilisant le lemme III.2.19 $(ii)$ pour prouver $(i)$. Montrons $(ii)$ : la proposition III.2.21 implique que l'intensité $\mathbf{\Pi}\times\mathbf{\Pi}'$ est $\mu\otimes\mu'$. Donc,

$$\mathbf{E}\big[N_\Delta(\mathbf{\Pi}\times\mathbf{\Pi}')\big] = (\mu\otimes\mu')(\Delta) = \int_E \mu(\{x\})\mu'(dx) = 0 \ ,$$

si $\mu$ est diffuse. Donc p.s. $(\mathbf{\Pi}\times\mathbf{\Pi}')\cap\Delta = \emptyset$, c'est-à-dire $\mathbf{\Pi}\cap\mathbf{\Pi}' = \emptyset$. ■

**Proposition III.2.23** *Soient* $(E, \mathscr{E})$*, un espace mesurable séparable et séparé. Soit* $\mathbf{\Pi}_n : \Omega \to \mathbf{S}_E$*,* $n\in\mathbb{N}$*, une suite de nuages aléatoires d'intensités respectives* $\mu_n$*,* $n\in\mathbb{N}$*, toutes supposées sigma-finies. Alors, les assertions suivantes sont vérifiées.*

*(i)* $\mathbf{\Pi} := \bigcup_{n\in\mathbb{N}} \mathbf{\Pi}_n$ *est un nuage aléatoire.*

*(ii) Si les* $(\mathbf{\Pi}_n)_{n\in\mathbb{N}}$ *sont indépendants et si les* $(\mu_n)_{n\in\mathbb{N}}$ *sont diffuses, alors*

$$\mathbf{P}\text{-}p.s. \quad \forall B\in\mathscr{E}\,, \quad N_B(\mathbf{\Pi}) = \sum_{n\in\mathbb{N}} N_B(\mathbf{\Pi}_n) \quad \text{(III.18)}$$

*et l'intensité de* $\mathbf{\Pi}$ *est* $\mu = \sum_{n\in\mathbb{N}} \mu_n$.

**Preuve :** en utilisant le fait que les intensités sont sigma-finies et en appliquant de manière répétée le lemme III.2.19 $(iii)$, page 240, on montre que $\mathbf{\Pi}_0\cup\ldots\cup\mathbf{\Pi}_n$ est $(\mathscr{F}, \mathscr{S}_E)$-mesurable. Par convergence monotone pour la mesure de comptage, partout sur $\Omega$, et pour tout $B\in\mathscr{E}$, $N_B(\mathbf{\Pi}) = \lim_{n\to\infty} N_B(\mathbf{\Pi}_0\cup\ldots\cup\mathbf{\Pi}_n)$. Donc, $N_B(\mathbf{\Pi})$ est $\mathscr{F}$-mesurable pour tout $B\in\mathscr{E}$, ce qui prouve $(i)$.

Montrons $(ii)$ : la proposition III.2.22 $(ii)$ implique que p.s. pour tout $m<n$, $\mathbf{\Pi}_m\cap\mathbf{\Pi}_n = \emptyset$, ce qui implique (III.18). Par interversion positive série/espérance, on a donc $\mu(B) = \mathbf{E}[N_B(\mathbf{\Pi})] = \sum_{n\in\mathbb{N}} \mathbf{E}[N_B(\mathbf{\Pi}_n)] = \sum_{n\in\mathbb{N}} \mu_n(B)$, pour tout $B\in\mathscr{E}$, ce qui termine la preuve. ■

## III.3 Nuages Poissonniens.

### III.3.a Définition, premières propriétés.

Sauf mention explicite du contraire, on fait les hypothèses suivantes :

$(\Omega, \mathscr{F}, \mathbf{P})$ supposé *complet* et $(E, \mathscr{E})$ est supposé *séparable et séparé.*

Sauf mention explicite du contraire, on fait les hypothèses suivantes : On rappelle qu'une variable à valeurs dans $\mathbb{N}_\infty$ est une variable de Poisson *étendue* si son paramètre appartient à $[0, \infty]$ : notamment son paramètre est nul (resp. infini) ssi la variable est p.s. nulle (resp. p.s. infinie).

**Définition. Caractérisations diverses.** En s'inspirant du cas des nuages homogènes sur $\mathbb{R}_+$, on introduit la définition suivante.

**Définition III.3.1** (*Nuages Poissonniens*) Soit $(E, \mathscr{E})$, un espace mesurable séparable et séparé. Soit $\mathbf{\Pi} : \Omega \to \mathbf{S}_E$, un nuage aléatoire. C'est un *nuage Poissonnien* s'il satisfait les deux hypothèses suivantes.

(**Chaos**) Pour tous $A_1, \dots, A_p \in \mathscr{E}$ deux-à-deux disjoints, les variables $N_{A_1}(\mathbf{\Pi}), \dots, N_{A_p}(\mathbf{\Pi})$ sont indépendantes.

(**Poisson**) Pour tout $A \in \mathscr{E}$, $N_A(\mathbf{\Pi})$ suit une loi de Poisson étendue. □

**Théorème III.3.2** *Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité complet. Soit $(E, \mathscr{E})$, un espace mesurable séparable et séparé. Soit $\mathbf{\Pi} : \Omega \to \mathbf{S}_E$, un nuage Poissonnien d'intensité $\mu$. Alors, $\mu$ est diffuse et la loi de $\mathbf{\Pi}$ est entièrement déterminée par $\mu$.*

**Preuve :** pour tout $x \in E$, $N_{\{x\}}(\mathbf{\Pi}) = 0$ ou 1. Or, $N_{\{x\}}(\mathbf{\Pi})$ suit une loi de Poisson de paramètre $\mu(\{x\})$. On a donc $\mu(\{x\}) = 0$. Montrons le second point de la proposition : supposons que $\mathbf{\Pi}'$ soit un nuage Poissonnien d'intensité $\mu$ défini sur un espace de probabilité $(\Omega', \mathscr{F}', \mathbf{P}')$. Par définition, $\mathbf{\Pi}$ et $\mathbf{\Pi}'$ satisfont l'identité en loi (III.14) de la proposition III.2.11. Cette même proposition implique que $\mathbf{\Pi}$ et $\mathbf{\Pi}'$ ont même loi. ■

Le théorème suivant montre que l'hypothèse de chaos pour un nuage aléatoire implique en réalité son caractère Poissonnien. Sa preuve est placée en fin de section.

**Théorème III.3.3** *Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité complet. Soit $(E, \mathscr{E})$, un espace séparable et séparé. Soit $\mathbf{\Pi} : \Omega \to \mathbf{S}_E$, un nuage de points $(\mathscr{F}, \mathscr{S}_E)$-mesurable d'intensité $\mu$, supposée* sigma-finie et diffuse. *On suppose que $\mathbf{\Pi}$ satisfait également* l'hypothèse de chaos *:*

(**Chaos**) *Pour tous $A_1, \dots, A_p \in \mathscr{E}$ deux-à-deux disjoints, les v.a. $N_{A_1}(\mathbf{\Pi}), \dots, N_{A_p}(\mathbf{\Pi})$ sont indépendantes.*

*Alors, $\mathbf{\Pi}$ est un nuage Poissonnien.*

**Preuve :** voir la fin de la section, page 246. ■

En application de ce résultat on peut montrer le théorème suivant qui simplifie la propriété de chaos et qui caractérise les nuages de Poisson par indépendance des événements d'évitement.

**Théorème III.3.4** **(Caractérisation par évitement indépendants)** *Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité complet. Soit $(E, \mathscr{E})$, un espace séparable et séparé. Soit $\mathbf{\Pi} : \Omega \to \mathbf{S}_E$, un nuage de points $(\mathscr{F}, \mathscr{S}_E)$-mesurable d'intensité $\mu$, supposée* sigma-finie et diffuse. *On suppose que $\mathbf{\Pi}$ satisfait la propriété suivante.*

(**Évitements indépendants**) *Pour tous $A, B \in \mathscr{E}$ disjoints, les événements $\{\mathbf{\Pi} \cap A = \emptyset\}$ et $\{\mathbf{\Pi} \cap B = \emptyset\}$ sont indépendants.*

*Alors $\mathbf{\Pi}$ est un nuage de Poisson.*

**Preuve :** voir l'exercice III.12, page 266. ■

En application du résultat précédent, on peut ensuite montrer le théorème de Rényi qui caractérise les nuages de Poisson par leurs probabilités d'évitement.

**Théorème III.3.5** **(Théorème de Rényi)** *Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité complet. Soit $(E, \mathscr{E})$, un espace séparable et séparé. Soit $\mathbf{\Pi} : \Omega \to \mathbf{S}_E$, un nuage de points $(\mathscr{F}, \mathscr{S}_E)$-mesurable. On suppose que*

$$\forall A \in \mathscr{E}, \quad \mathbf{P}(\mathbf{\Pi} \cap A = \emptyset) = e^{-\mu(A)} ,$$

*où $\mu : \mathscr{E} \to [0, \infty]$ est une mesure* sigma finie et diffuse *(avec la convention que $e^{-\infty} = 0$). Alors $\mathbf{\Pi}$ est alors un nuage de Poisson.*

**Preuve :** voir l'exercice III.13, page 267. ■

**Stabilité des nuages Poisson par restriction, image et superposition.**

**Théorème III.3.6** **(Restriction)** *Soit $(E, \mathscr{E})$, un espace mesurable séparable et séparé. Soit $\mathbf{\Pi}$, un nuage Poissonnien sur $E$ d'intensité $\mu$. Soient $B_1, B_2 \in \mathscr{E}$, disjoints. Alors $\mathbf{\Pi} \cap B_1$ et $\mathbf{\Pi} \cap B_2$ sont deux nuages Poissonniens indépendants d'intensités respectives $\mu(\,\cdot\, \cap B_1)$ et $\mu(\,\cdot\, \cap B_2)$.*

*Plus généralement, soit $B_i \in \mathscr{E}$, $i \in I$, une famille d'ensembles mesurables deux-à-deux disjoints. Alors, $\mathbf{\Pi} \cap B_i$, $i \in I$, est une famille de nuages Poissonniens indépendants d'intensités respectives $\mu(\,\cdot\, \cap B_i)$, $i \in I$.*

**Preuve :** la définition même des nuages Poissonniens implique que pour tout $A \in \mathscr{E}$, $\mathbf{\Pi} \cap A$ est un nuage Poissonnien d'intensité $\mu(\,\cdot\, \cap A)$. Pour tout $i \in I$, on pose $\mathbf{\Pi}_i := \mathbf{\Pi} \cap B_i$ et on a pour tout $A \in \mathscr{E}$, $N_A(\mathbf{\Pi}_i) = N_{A \cap B_i}(\mathbf{\Pi})$. Comme les $B_i \in \mathscr{E}$, $i \in I$, deux-à-deux disjoints, pour tous $A_1, \ldots, A_p \in \mathscr{E}$ deux-à-deux disjoints, les vecteurs aléatoires $(N_{A_k}(\mathbf{\Pi}_i))_{1 \le k \le p}$, $i \in I$, sont indépendants (on utilise la propriété de Chaos des nuages Poissonniens car les ensembles $A_k \cap B_i$, $1 \le k \le p$, $i \in I$, sont deux-à-deux disjoints). On conclut grâce à la proposition III.2.14 $(ii)$. ■

**Théorème III.3.7** **(Superposition croissante ou indépendante)** *Soit $(E, \mathscr{E})$, un espace mesurable séparable et séparé. Soit $(\mathbf{\Pi}_n)_{n \in \mathbb{N}}$, une suite de nuages Poissonniens sur $E$, d'intensités respectives $(\mu_n)_{n \in \mathbb{N}}$ toutes supposées sigma-finies. On pose $\mathbf{\Pi} := \bigcup_{n \in \mathbb{N}} \mathbf{\Pi}_n$. Alors, c'est un nuage aléatoire dont l'intensité est notée $\mu$ et les assertions suivantes sont vérifiées.*

*(i) Si on suppose que $\mathbf{P}$-p.s. $\mathbf{\Pi}_n \subset \mathbf{\Pi}_{n+1}$, pour tout $n \in \mathbb{N}$, alors $\mathbf{\Pi}$ est un nuage de Poisson d'intensité $\mu$ et pour tout $A \in \mathscr{E}$, on a $\mu(A) = \lim_n \uparrow \mu_n(A)$.*

*(ii) Si on suppose que les $(\mathbf{\Pi}_n)_{n \in \mathbb{N}}$ sont indépendants, alors $\mathbf{\Pi}$ est un nuage de Poisson d'intensité $\mu$ et on a $\mu = \sum_{n \in \mathbb{N}} \mu_n$.*

**Preuve :** par la proposition III.2.23 $(i)$, $\mathbf{\Pi}$ est $(\mathscr{F}, \mathscr{S}_E)$-mesurable. Montrons $(i)$. Soit $A \in \mathscr{E}$ ; le théorème de convergence monotone pour la mesure de comptage implique que $\mathbf{P}$-p.s. $N_A(\mathbf{\Pi}) = \lim_n N_A(\mathbf{\Pi}_n)$. En prenant l'espérance, le théorème de convergence monotone sous $\mathbf{P}$ implique alors que $\mu(A) = \lim_n \mu_n(A)$. Alors, clairement, pour tout $r \in [0,1]$, par convergence dominée

$$\mathbf{E}\big[r^{N_A(\mathbf{\Pi})}\big] = \lim_{n\to\infty} \mathbf{E}\big[r^{N_A(\mathbf{\Pi}_n)}\big] = \lim_{n\to\infty} e^{-\mu_n(A)(1-r)} = e^{-\mu(A)(1-r)},$$

ce qui implique que $N_A(\mathbf{\Pi})$ suit une loi de Poisson (étendue) de paramètre $\mu(A)$. Soient $A_1, \ldots, A_p \in \mathscr{E}$, deux-à-deux disjoints. Pour tout $n \in \mathbb{N}$ fixé les v.a. $N_{A_1}(\mathbf{\Pi}_n), \ldots, N_{A_p}(\mathbf{\Pi}_n)$ sont indépendantes et puisque

$$\mathbf{P}\text{-p.s.} \quad \lim_{n\to\infty} \big(N_{A_1}(\mathbf{\Pi}_n), \ldots, N_{A_p}(\mathbf{\Pi}_n)\big) = \big(N_{A_1}(\mathbf{\Pi}), \ldots, N_{A_p}(\mathbf{\Pi})\big),$$

les v.a. limites $N_{A_1}(\mathbf{\Pi}), \ldots, N_{A_p}(\mathbf{\Pi})$ sont nécessairement indépendantes. Cela montre donc que $\mathbf{\Pi}$ est un nuage de Poisson.

Montrons $(ii)$ : par la proposition III.2.23, page 242, $\mathbf{\Pi}$ est un nuage aléatoire $(\mathscr{F}, \mathscr{S}_E)$-mesurable d'intensité $\mu = \sum_{n\in\mathbb{N}} \mu_n$. Soit $B \in \mathscr{E}$. On remarque que les v.a. $(N_B(\mathbf{\Pi}_n))_{n\in\mathbb{N}}$ sont des v.a. de Poisson indépendante. Par conséquent (III.18) et le lemme III.1.3 page 224 impliquent que $N_B(\mathbf{\Pi})$ est de Poisson.

Soient $A_1, \ldots, A_p \in \mathscr{E}$, deux-à-deux disjoints. On vérifie facilement que les v.a. $(N_{A_i}(\mathbf{\Pi}_n))_{1\leq i\leq p; n\in\mathbb{N}}$ sont mutuellement indépendantes. Notamment les suites $(N_{A_1}(\mathbf{\Pi}_n))_{n\in\mathbb{N}}, \ldots, (N_{A_p}(\mathbf{\Pi}_n))_{n\in\mathbb{N}}$ sont indépendantes. Or par (III.18), $N_{A_i}(\mathbf{\Pi}) = \sum_{n\in\mathbb{N}} N_{A_i}(\mathbf{\Pi}_n)$, pour tout $1 \leq i \leq p$. Cela montre que les v.a. $N_{A_1}(\mathbf{\Pi}), \ldots, N_{A_p}(\mathbf{\Pi})$ sont indépendantes. Cela termine la preuve de $(ii)$. ■

**Proposition III.3.8** *Soient $(E, \mathscr{E})$ et $(E', \mathscr{E}')$, des espaces mesurables séparés et séparables. Soit $\mathbf{\Pi}$, un nuage Poissonnien sur $E$ d'intensité $\mu$. Soit $\phi : E \to E'$, supposée $(\mathscr{E}, \mathscr{E}')$-mesurable. On note $\mu'$ la mesure image de $\mu$ par $\phi$ et on pose $\mathbf{\Pi}' := \phi(\mathbf{\Pi})$. Si $\phi$ est injective, alors $\mathbf{\Pi}'$ est un nuage Poissonnien sur $E'$ d'intensité $\mu'$.*

**Preuve :** si $\phi$ est injective, alors

$$\forall \omega \in \Omega\,,\ \forall B' \in \mathscr{E}'\,, \quad N_{B'}(\mathbf{\Pi}'(\omega)) = N_{\phi^{-1}(B')}(\mathbf{\Pi}(\omega))\ . \tag{III.19}$$

Cela montre que $\mathbf{\Pi}'$ est un nuage aléatoire. Par ailleurs $N_{B'}(\mathbf{\Pi}')$ est une variable de Poisson d'intensité $\mu(\phi^{-1}(B')) = \mu'(B')$. Enfin, pour tous $A'_1, \ldots, A'_p \in \mathscr{E}'$, deux-à-deux disjoints, $\phi^{-1}(A'_1), \ldots, \phi^{-1}(A'_p) \in \mathscr{E}$ sont également deux-à-deux disjoints et (III.19) implique que les variables $N_{A'_1}(\mathbf{\Pi}'), \ldots, N_{A'_p}(\mathbf{\Pi}')$ sont indépendantes. ■

Le théorème suivant est une version plus sophistiquée de la proposition précédente.

**Théorème III.3.9** **(Image)** *Soient $(E, \mathscr{E})$ et $(E', \mathscr{E}')$, deux espaces mesurables séparables et séparés. Soit $\mathbf{\Pi}$, un nuage Poissonnien sur $E$ d'intensité $\mu$, supposée sigma-finie. Soit $\phi : E \to E'$, une fonction $(\mathscr{E}, \mathscr{E}')$-mesurable. On note $\mu'$ la mesure image de $\mu$ par $\phi$ et on pose $\mathbf{\Pi}' := \phi(\mathbf{\Pi})$. On suppose $\mu'$ diffuse.*

*Alors, $\mathbf{\Pi}'$ est un nuage Poissonnien d'intensité $\mu'$ et p.s. $\phi$ est injective sur $\mathbf{\Pi}$, c'est-à-dire*

$$\mathbf{P}\text{-}p.s. \quad \forall X, Y \in \mathbf{\Pi}\,, \quad \Big(X \neq Y\Big) \implies \Big(\phi(X) \neq \phi(Y)\Big)\,, \tag{III.20}$$

*ce qui entraîne que*

$$\mathbf{P}\text{-}p.s. \quad \forall x' \in E'\,, \quad N_{\phi^{-1}(\{x'\})}(\mathbf{\Pi}) \in \{0, 1\} \tag{III.21}$$

**Preuve :** on suppose d'abord que l'on a montré (III.20). Cela implique immédiatement (III.21). Il existe alors $\Omega^* \in \mathscr{F}$ tel que $\mathbf{P}(\Omega^*) = 1$ et

$$\forall \omega \in \Omega^*\,,\ \forall B' \in \mathscr{E}'\,, \quad N_{B'}(\mathbf{\Pi}'(\omega)) = N_{\phi^{-1}(B')}(\mathbf{\Pi}(\omega))\ , \tag{III.22}$$

On raisonne ensuite comme à la proposition III.3.8 pour montrer que $\mathbf{\Pi}'$ est Poissonnien d'intensité $\mu'$.

Il reste donc à prouver (III.20). Pour cela on considère $\mathbf{\Pi}^* = \mathbf{\Pi}\times\mathbf{\Pi}$, qui d'après la proposition III.2.21, est un nuage aléatoire sur l'espace $E\times E$ muni de la tribu produit $\mathscr{E}\otimes\mathscr{E}$. On note $\rho$ son intensité, que l'on calcule comme suit : on fixe $A, B\in\mathscr{E}$ et on pose $A_0 := A\backslash(A\cap B)$ et $B_0 := B\backslash(A\cap B)$, si bien que $A_0, A\cap B, B_0$ sont disjoints deux-à-deux. On a donc

$$\begin{aligned}\rho(A\times B) &= \mathbf{E}\big[N_{A_0}(\mathbf{\Pi})N_{B_0}(\mathbf{\Pi})\big] + \mathbf{E}\big[N_{A_0}(\mathbf{\Pi})N_{A\cap B}(\mathbf{\Pi})\big] \\ &\qquad + \mathbf{E}\big[N_{A\cap B}(\mathbf{\Pi})N_{B_0}(\mathbf{\Pi})\big] + \mathbf{E}\big[(N_{A\cap B}(\mathbf{\Pi})\,)^2\big] \\ &= \mu(A_0)\mu(B_0) + \mu(A_0)\mu(A\cap B) + \mu(A\cap B)\mu(B_0) + \mathbf{E}\big[(N_{A\cap B}(\mathbf{\Pi})\,)^2\big].\end{aligned}$$

Comme une variable de Poisson de paramètre $\theta$ a un moment d'ordre 2 est égal à $\theta+\theta^2$, on a

$$\begin{aligned}\rho(A\times B) &= \mu(A_0)\mu(B_0) + \mu(A_0)\mu(A\cap B) + \mu(A\cap B)\mu(B_0) + \mu(A\cap B)^2 + \mu(A\cap B) \\ &= \mu(A)\mu(B) + \mu(A\cap B) = (\mu\otimes\mu)(A\times B) + \mu(A\cap B)\;.\end{aligned}$$

On introduit ensuite la fonction $\jmath : x\in E\mapsto (x,x)\in E\times E$. On rappelle que $\Delta\in\mathscr{E}\otimes\mathscr{E}$. Donc $\jmath$ est $(\mathscr{E},\mathscr{E}\otimes\mathscr{E})$-mesurable : c'est même une bijection $E$ sur $\Delta$. On note $\mu_\Delta$ la mesure induite sur $\Delta$ par $\mu$ via $\jmath$. Comme $\jmath^{-1}((A\times B)\cap\Delta) = A\cap B$, on a donc montré que pour tous $A, B\in\mathscr{E}$, $\rho(A\times B) = \mu_\Delta(A\times B) + (\mu\otimes\mu)(A\times B)$. Comme $\rho$ et $\mu_\Delta + \mu\otimes\mu$ coïncident sur le pi-système $\mathscr{P} := \{A\times B\,;\, A, B\in\mathscr{E}\}$ engendrant $\mathscr{E}\otimes\mathscr{E}$, le théorème d'unicité du prolongement des mesures sigma-finies s'applique facilement pour montrer que $\rho = \mu_\Delta + \mu\otimes\mu$.

On pose $\Phi : (x,y)\in E\times E\mapsto (\phi(x),\phi(y))\in E'\times E'$. Il est facile de montrer que $\Phi$ est $(\mathscr{E}\otimes\mathscr{E}, \mathscr{E}'\otimes\mathscr{E}')$-mesurable. On note $\Delta'$ la diagonale de $E'\times E'$, diagonale qui appartient à $\mathscr{E}'\otimes\mathscr{E}'$. On s'intéresse à l'ensemble $U := \Phi^{-1}(\Delta')\backslash\Delta = \big\{(x,y)\in E\times E : x\neq y \text{ et } \phi(x)=\phi(y)\big\}$. Comme $\mu_\Delta$ est concentrée sur $\Delta$ (par définition), on a $\mu_\Delta(U) = 0$. Par Fubini, on a également,

$$\begin{aligned}(\mu\otimes\mu)(U) &= \int_E \mu(dx)\int_E \mu(dy)\mathbf{1}_{\{y\in E: y\neq x \text{ et } \phi(y)=\phi(x)\}} \\ &\leq \int_E \mu(dx)\mu\big(\phi^{-1}(\{\phi(x)\})\big) = \int_E \mu(dx)\mu'(\{\phi(x)\})\;.\end{aligned}$$

Comme $\mu'$ est diffuse, on a $\mu'(\{\phi(x)\}) = 0$, pour tout $x\in E$ et donc $(\mu\otimes\mu)(U) = 0$. On en déduit donc que $\rho(U) = \mathbf{E}[N_U(\mathbf{\Pi}^*)] = 0$ et donc que $\mathbf{P}$-p.s. $\mathbf{\Pi}^*\cap U = \emptyset$, c'est-à-dire (III.20), ce qui termine la preuve. ■

**Preuve du théorème III.3.3.** On observe tout d'abord que les nuages aléatoires satisfaisant la propriété de chaos sont stables par restriction et que leurs restrictions à des sous-ensembles deux-à-deux disjoints sont mutuellement indépendantes. Cette observation, combinée au principe de superposition indépendante (théorème III.3.7 $(ii)$), montre que l'on peut se ramener aux cas de nuages aléatoires d'intensité finie.

On considère donc un $\mathbf{\Pi} : \Omega\to\mathbf{S}_E$ un nuage aléatoire $(\mathscr{F},\mathscr{S}_E)$-mesurable, satisfaisant la propriété de chaos et d'intensité $\mu$ supposée finie : $\mu(E) < \infty$. Comme $(E,\mathscr{E})$ est séparable et séparé, il existe $C\subset[0,1]$ et $\phi : E\to C$ bijective $(\mathscr{E},\mathscr{B}(C))$-mesurable telle que la réciproque $\phi^{-1} : C\to E$ soit $(\mathscr{B}(C),\mathscr{E})$-mesurable (on rappelle que $\mathscr{B}(C)$ est la tribu trace des Boréliens de $\mathbb{R}$ sur $C$). Le lemme III.2.7, page 236, montre que $\mathbf{\Pi}' := \phi(\mathbf{\Pi})$ est un nuage aléatoire d'intensité $\mu' := \mu\circ\phi^{-1}$. Comme $\phi$ est bijective, il est clair que $\mu'$ est diffuse et que $\mathbf{\Pi}'$ satisfait la propriété de chaos. Si on montre que $\mathbf{\Pi}'$ est Poissonnien, alors la proposition III.3.8, page 245, implique que $\mathbf{\Pi} = \phi^{-1}(\mathbf{\Pi}')$ est Poissonnien également.

Il suffit donc de montrer que $\mathbf{\Pi}'$ est Poissonnien : comme $\mathbf{\Pi}'$ satisfait la propriété de chaos, il suffit de montrer que pour tout $A\in\mathscr{B}(C)$, $N_A(\mathbf{\Pi}')$ suit une loi de Poisson. Soient $n\in\mathbb{N}$ et $0\leq k < 2^n$ ; on pose

$A_{n,k} = A \cap [k2^n, (k+1)2^n[$ et

$$\xi_{n,k} = \mathbf{1}_{\{N_{A_{n,k}}(\mathbf{\Pi}') \geq 1\}}, \quad p_{n,k} = \mathbf{P}(N_{A_{n,k}}(\mathbf{\Pi}') \geq 1) \quad \text{et} \quad X_n = \sum_{0 \leq k < 2^n} \xi_{n,k}.$$

L'hypothèse de chaos implique que pour tout $n \in \mathbb{N}$, $(\xi_{n,k})_{0 \leq k < 2^n}$ sont des v.a. de Bernoulli indépendantes de paramètre respectifs $(p_{n,k})_{0 \leq k < 2^n}$. Il est ensuite facile de voir que p.s. pour tout $n \in \mathbb{N}$

$$X_n \leq X_{n+1} \quad \text{et} \quad \lim_{n \to \infty} X_n = \sup_{n \geq 0} X_n = N_A(\mathbf{\Pi}') < \infty. \tag{III.23}$$

Par convergence monotone, cela implique que

$$\lim_{n \to \infty} \sum_{0 \leq k < 2^n} p_{n,k} = \lim_{n \to \infty} \mathbf{E}[X_n] = \mathbf{E}[N_A(\mathbf{\Pi}')] = \mu'(A) \ . \tag{III.24}$$

On remarque ensuite que $p_{n,k} \leq \mathbf{E}[N_{A_{n,k}}(\mathbf{\Pi}')] = \mu'(A_{n,k})$ et donc que

$$\sum_{0 \leq k < 2^n} p_{n,k}^2 \leq \sum_{0 \leq k < 2^n} \mu'(A_{n,k})^2 \leq \mu'(A) \max_{0 \leq k < 2^n} \mu'(A_{n,k}) \ . \tag{III.25}$$

Raisonnons par l'absurde en supposant que $\max_{0 \leq k < 2^n} \mu'(A_{n,k})$ ne tend pas vers 0 lorsque $n$ tend vers l'infini : il existe donc $\varepsilon > 0$ et une suite strictement croissante d'entiers naturels $(n_\ell)_{\ell \geq 0}$ telle que

$$\max_{0 \leq k < 2^{n_\ell}} \mu'(A_{n_\ell,k}) \geq \varepsilon \ .$$

On introduit les ensembles *compacts* suivants :

$$I_{n,k} = [k2^n, (k+1)2^n], \quad 0 \leq k < 2^n, \qquad \text{et} \qquad J_\ell = \bigcup_{\substack{1 \leq k \leq 2^{n_\ell}: \\ \mu'(A_{n_\ell,k}) \geq \varepsilon}} I_{n_\ell,k} \ .$$

Clairement, $J_{\ell+1} \subset J_\ell$. On pose $K = \bigcap_{\ell \geq 0} J_\ell$, qui est un compact non-vide, par ce qui précède et puisque les $J_\ell$ sont également compacts non-vides. Comme $\mu'(A \cap J_\ell) \geq \varepsilon$, on a $\mu'(K \cap A) \geq \varepsilon$. Soient $x_1, \ldots, x_q \in K \cap A$ distincts ; alors pour tout $\ell$ assez grand, il existe $k_1, \ldots, k_q \in \{0, \ldots, 2^{n_\ell} - 1\}$ *distincts* tels que pour tout $p \in \{1, \ldots, q\}$, on ait $x_p \in A_{n_\ell,k_p}$ et $\mu'(A_{n_\ell,k_p}) \geq \varepsilon$. Donc les intervalles $I_{n_\ell,k_p}$, $1 \leq p \leq q$ sont disjoints. Cela implique que $q\varepsilon \leq \sum_{1 \leq p \leq q} \mu'(A_{n_\ell,k_p}) \leq \mu'(A)$. On a montré que $K \cap A$ contient au plus $\mu'(A)/\varepsilon$ points. Or on a également prouvé que $\mu'(K \cap A) \geq \varepsilon$, ce qui contredit le fait que $\mu'$ soit diffuse.

On a donc montré par l'absurde que $\lim_{n \to \infty} \max_{0 \leq k < 2^n} \mu'(A_{n,k}) = 0$. Par (III.25) et (III.24), l'approximation Binomiale-Poisson généralisée établie au théorème III.1.4 s'applique (voir page 225, en appendice) : les variables $X_n$ tendent en loi vers une loi de Poisson de paramètre $\mu'(A)$ et (III.23) implique que $N_A(\mathbf{\Pi}')$ suit une loi de Poisson, ce qui permet de conclure. ■

## III.3.b Construction.

**Les cas des nuages Poissonniens d'intensité finie.** On se donne les objets suivants.

(a) $(E, \mathscr{E})$, un espace séparable et séparé ; $\mu : \mathscr{E} \to \mathbb{R}_+$, une mesure finie, non-nulle et diffuse.

(b) $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité complet ; $X_n : \Omega \to E$, $n \in \mathbb{N}^*$, des v.a. $(\mathscr{F}, \mathscr{E})$-mesurables, indépendantes de même loi $\mu(\,\cdot\,)/\mu(E)$.

(c) $M:\Omega\to\mathbb{N}$, une v.a. $\mathscr{F}$-mesurable, indépendante de $(X_n)_{n\geq 1}$, loi de Poisson de paramètre $\mu(E)$.

**Théorème III.3.10** *Sous les hypothèses et notations précédentes, on définit* $\mathbf{\Pi}:\Omega\to\mathbf{S}_E$ *par*

$$\mathbf{\Pi}=\{X_1,\ldots,X_M\}\quad \text{si}\quad M\geq 1\quad \text{et}\quad \mathbf{\Pi}=\emptyset\quad \text{si}\quad M=0\;. \tag{III.26}$$

*Alors,* $\mathbf{\Pi}$ *est un nuage Poissonnien d'intensité* $\mu$.

**Preuve :** l'exemple III.2.2 permet d'affirmer que $\mathbf{\Pi}$ est un nuage aléatoire. Soient $p\in\mathbb{N}^*$ et $A_1,\ldots,A_p\in\mathscr{E}$, deux-à-deux disjoints. On pose $A_{p+1}:=E\backslash(A_1\cup\ldots\cup A_p)$, si bien que $A_1,\ldots,A_{p+1}$ est une partition de $E$. On fixe $u_1,\ldots,u_p\in\mathbb{R}$ et pour tout $n\in\mathbb{N}^*$, on pose $Z_n:=e^{iu_1}\mathbf{1}_{A_1}(X_n)+\ldots+e^{iu_p}\mathbf{1}_{A_p}(X_n)+\mathbf{1}_{A_{p+1}}(X_n)$. Les variables $(Z_n)_{n\geq 1}$ sont i.i.d. et on vérifie facilement que

$$\mu(E)\big(1-\mathbf{E}[Z_n]\big)=\mu(E)\big(1-\mathbf{E}[Z_1]\big)=\mu(A_1)(1-e^{iu_1})+\ldots+\mu(A_p)(1-e^{iu_p})\;. \tag{III.27}$$

Comme $\mu$ est sans atome, $\mathbf{P}$-p.s. pour tous $m<n$, $X_n\neq X_m$, ce qui entraîne

$$\mathbf{P}\text{-p.s.}\quad \prod_{1\leq j\leq p}\exp\big(iu_jN_{A_j}(\Pi)\big)=\prod_{1\leq j\leq M}Z_j\;,$$

avec la convention qu'un produit sur un ensemble d'indices vide est égal à 1. On en déduit les égalités suivantes :

$$\begin{aligned}\mathbf{E}\Big[\prod_{1\leq j\leq p}\exp\big(iu_jN_{A_j}(\mathbf{\Pi})\big)\Big] &= \mathbf{E}\Big[\prod_{1\leq j\leq M}Z_j\Big]=1+\sum_{n\geq 1}\mathbf{E}\Big[\mathbf{1}_{\{M=n\}}\prod_{1\leq j\leq n}Z_j\Big]\\ &= 1+\sum_{n\geq 1}\mathbf{P}(M=n)\mathbf{E}\Big[\prod_{1\leq j\leq n}Z_j\Big]=\sum_{n\geq 0}e^{-\mu(E)}\frac{\mu(E)^n}{n!}\mathbf{E}[Z_1]^n\\ &= \exp\left(-\mu(E)(1-\mathbf{E}[Z_1])\right)=\prod_{1\leq j\leq p}\exp\big(-\mu(A_j)(1-e^{iu_j})\big)\;.\end{aligned}$$

L'injectivité de la fonction caractéristique montre que le vecteurs $(N_{A_1}(\mathbf{\Pi}),\ldots,N_{A_p}(\mathbf{\Pi}))$ a des composantes indépendantes suivant des lois de Poisson, ce qui implique le résultat désiré. ■

**Construction pour les intensité sigma-finies.** On procède de la même façon en utilisant le principe de superposition. Plus précisément, on se donne les objets suivants.

(a) $(E,\mathscr{E})$, un espace mesurable séparable et séparé. $\mu:\mathscr{E}\to[0,\infty]$, une mesure sigma-finie, non-nulle et diffuse. Il existe donc $B_p\in\mathscr{E}$, $p\geq 1$, une partition de $E$ telle que pour tout $p\geq 1$, on ait $0<\mu(B_p)<\infty$.

(b) $(\Omega,\mathscr{F},\mathbf{P})$, un espace de probabilité complet ; $X_n^{(p)}:\Omega\to E$, $p,n\geq 1$, des v.a. $\mathscr{F}$-mesurables indépendantes telles que pour tout $p\geq 1$, la suite $(X_n^{(p)})_{n\geq 1}$ est i.i.d. de loi $\mu(\,\cdot\,\cap B_p)/\mu(B_p)$. On pose $T:=(X_m^{(p)})_{p,n\geq 1}$.

(c) $M_p:\Omega\to\mathbb{N}$, $p\geq 1$, des v.a. $\mathscr{F}$-mesurables indépendantes ; $M_p$ suit une loi de Poisson de paramètre $\mu(B_p)$. On suppose que la suite $(M_p)_{p\geq 1}$ est indépendante du tableau de variables $T$.

**Théorème III.3.11** *Sous les hypothèses et notations précédentes, pour tout* $p\geq 1$, *on pose*

$$\mathbf{\Pi}_p=\big\{X_1^{(p)},\ldots,X_{M_p}^{(p)}\big\}\ \text{si}\ M_p\geq 1\quad \text{et}\quad \mathbf{\Pi}_p=\emptyset\ \text{si}\ M_p=0\;.$$

*On pose également* $\mathbf{\Pi}=\bigcup_{p\geq 1}\mathbf{\Pi}_p$. *Alors,* $\mathbf{\Pi}$ *est un nuage Poissonnien sur* $E$ *d'intensité* $\mu$.

**Preuve :** par le théorème III.3.10, $\mathbf{\Pi}_p$ est un nuage Poissonnien d'intensité $\mu(\,\cdot\,\cap B_p)$. La construction garantit que les $\mathbf{\Pi}_p$, $p\geq 1$, sont indépendants. Par la propriété de superposition (proposition III.3.7, page 244), $\mathbf{\Pi}$ est un nuage Poissonnien d'intensité $\sum_{p\geq 1}\mu(\,\cdot\,\cap B_p)=\mu$. ■

**Représentation des nuages Poissonniens.** On montre, à l'aide du lemme de localisation III.2.18 (page 239) que tout nuage Poissonnien, sur un espace séparable et séparé et d'intensité sigma-finie, peut se représenter de manière mesurable comme dans la construction générale qui précède. Pour simplifier quelques arguments on utilise le lemme suivant dont la preuve se trouve en fin de section.

**Lemme III.3.12** *Soit $(E, \mathscr{E})$, un espace séparable et séparé. Soit $\mu : \mathscr{E} \to [0, \infty]$, une mesure sigma finie et diffuse. Soit un entier $n \geq 2$. On note $O$ l'ensembles des $(x_1, \ldots, x_n) \in E$ tels que les $x_i$ soient distincts. Soit $\nu : \mathscr{E}^{\otimes n} \to [0, \infty]$, une mesure telle que pour tous $A_1, \ldots, A_n \in \mathscr{E}$, disjoints deux-à-deux, on ait*

$$\nu\big(A_1 \times \ldots \times A_n\big) = \mu(A_1) \ldots \mu(A_n) .$$

*Alors, $O \in \mathscr{E}^{\otimes n}$ et $\nu(\cdot \cap O) = \mu^{\otimes n}$.*

**Preuve.** Voir la fin de la section. ■

**Remarque III.3.13** On observe que le lemme précédent est faux si $\mu$ n'est pas diffuse. □

On commence par considérer les nuages d'intensité finie. Plus précisément, on se donne les objets suivants.

($\alpha$) $(E, \mathscr{E})$, un espace séparable et séparé.

($\beta$) $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité complet sur lequel est défini $\mathbf{\Pi} : \Omega \to \mathbf{S}_E$, un nuage Poissonnien d'intensité finie : $\mu(E) < \infty$.

($\gamma$) $U : \Omega \to [0, 1]$, une v.a. $\mathscr{F}$-mesurable indépendante de $\mathbf{\Pi}$ et de loi uniforme.

À l'aide de $U$, on construit :
– une suite $(X'_n)_{n\geq 1}$ de v.a. i.i.d. de loi $\mu(\,\cdot\,)/\mu(E)$ ;
– une suite de permutations indépendantes $(\Sigma_n)_{n\geq 1}$ telle que d'une part, $\Sigma_n : \{1, \ldots, n\} \to \{1, \ldots, n\}$ suit la loi uniforme et telle que d'autre part, la suite $(\Sigma_n)_{n\geq 1}$ est indépendante de $(X'_n)_{n\geq 1}$.
Comme $(E, \mathscr{E})$ est séparable et séparé, il existe $C \subset \mathbb{R}$ et une bijection $\phi : E \to C$ qui est $(\mathscr{E}, \mathscr{B}(C))$-mesurable et telle que la réciproque $\phi^{-1}$ soit $(\mathscr{B}(C), \mathscr{E})$-mesurable également. Par lemme III.2.18 (page 239), il existe une suite $Y_k : \mathbf{S}_E \to E$, $k \geq 1$, de fonctions $(\mathscr{S}_E, \mathscr{E})$-mesurables telles que pour tout nuage $\pi \in \mathbf{S}_E$ comptant $n$ points,

$$\pi = \big\{Y_1(\pi), \ldots, Y_n(\pi)\big\} \quad \text{et} \quad \phi(Y_1(\pi)) < \ldots < \phi(Y_n(\pi)) .$$

Pour tout $n \geq 1$, sur $\{\#\mathbf{\Pi} = n\}$, on pose :

$$X_k = Y_{\Sigma_n(k)}(\mathbf{\Pi}) \text{ si } 1 \leq k \leq n \quad \text{et} \quad X_k = X'_{k-n} \text{ si } k > n.$$

Sur $\{\#\mathbf{\Pi} = 0\} \cup \{\#\mathbf{\Pi} = \infty\}$, pour tout $k \geq 1$, on pose $X_k = X'_k$. On montre tout d'abord que pour tout $n \geq 1$,

$$\text{sous } \mathbf{P}(\,\cdot\,|\,\#\mathbf{\Pi} = n), \text{ les v.a. } X_1, \ldots, X_n \text{ sont i.i.d. de loi } \mu(\,\cdot\,)/\mu(E). \tag{III.28}$$

**Preuve de (III.28).** Soient $A_1, \ldots, A_n \in \mathscr{E}$, deux-à-deux disjoints. Observons que

$$\{N_{A_1}(\mathbf{\Pi}) = 1; \ldots; N_{A_n}(\mathbf{\Pi}) = 1\} \cap \{\#\mathbf{\Pi} = n\} = \bigcup_{\sigma \in \mathbf{S}_n} \{X_{\sigma(1)} \in A_1; \ldots; X_{\sigma(n)} \in A_n\} \cap \{\#\mathbf{\Pi} = n\},$$

ces événements étant deux-à-deux disjoints. Ici, $\mathbf{S}_n$ désigne l'ensemble des permutations de $\{1, \ldots, n\}$. On remarque ensuite que pour tout $\sigma \in \mathbf{S}_n$, $\Sigma_n \circ \sigma$ est une permutation de loi uniforme. Cela implique que sous $\mathbf{P}(\,\cdot\,|\,\#\mathbf{\Pi} = n)$, $(X_1, \ldots, X_n)$ à même loi que $(X_{\sigma(1)}, \ldots, X_{\sigma(n)})$ et donc

$$\mathbf{P}\big(N_{A_1}(\mathbf{\Pi}) = 1; \ldots; N_{A_n}(\mathbf{\Pi}) = 1 \,\big|\, \#\mathbf{\Pi} = n\big) \;=\; \sum_{\sigma \in \mathbf{S}_n} \mathbf{P}\big(X_{\sigma(1)} \in A_1; \ldots; X_{\sigma(n)} \in A_n \,\big|\, \#\mathbf{\Pi} = n\big)$$

$$= \quad n!\,\mathbf{P}\big(X_1 \in A_1; \ldots ; X_n \in A_n \,\big|\, \#\mathbf{\Pi} = n\big) \ . \qquad \text{(III.29)}$$

On pose ensuite $A_{n+1} := E \backslash (A_1 \cup \ldots \cup A_n)$ et on remarque que

$$\begin{aligned} \mathbf{P}\big(N_{A_1}(\mathbf{\Pi}) = 1; \ldots ; N_{A_n}(\mathbf{\Pi}) = 1; \#\mathbf{\Pi} = n\big) &= \mathbf{P}\big(N_{A_1}(\mathbf{\Pi}) = 1; \ldots ; N_{A_n}(\mathbf{\Pi}) = 1; N_{A_{n+1}}(\mathbf{\Pi}) = 0\big) \\ &= \mu(A_1) \ldots \mu(A_n) e^{-\mu(E)}. \end{aligned}$$

Comme $\#\mathbf{\Pi}$ est une v.a. de Poisson de paramètre $\mu(E)$, on en déduit que

$$\mathbf{P}\big(N_{A_1}(\mathbf{\Pi}) = 1; \ldots ; N_{A_n}(\mathbf{\Pi}) = 1 \,\big|\, \#\mathbf{\Pi} = n\big) = n!\, \frac{\mu(A_1)}{\mu(E)} \cdots \frac{\mu(A_1)}{\mu(E)}.$$

Si on note $\overline{\mu}(\cdot) = \mu(\cdot)/\mu(E)$ et $\nu$ la loi de $(X_1, \ldots, X_n)$ sous $\mathbf{P}(\,\cdot \,|\, \#\mathbf{\Pi} = n)$. On veut appliquer le lemme III.3.12 : on rappelle la définition de l'ensemble $O$ de l'énoncé ; par construction, $\nu(E^n \backslash O) = 0$ et donc $\nu = \nu(\cdot \cap O)$. Alors ce qui précède, (III.29) et le lemme III.3.12, impliquent que $\nu = \overline{\mu}^{\otimes n}$, c'est-à-dire (III.28). ■

On déduit facilement de (III.28) la proposition suivante.

**Proposition III.3.14** *On suppose $(\alpha)$, $(\beta)$ et $(\gamma)$ ci-dessus. On note $\mathscr{G}$ la tribu engendrée par $U$ et $\mathbf{\Pi}$. On pose $M := \#\mathbf{\Pi}$. Alors, il existe une famille de variables $(X_n)_{n\geq 1}$ satisfaisant les conditions suivantes.*

*(i) Pour tout $n \geq 1$, $X_n : \Omega \to E$ est $(\mathscr{G}, \mathscr{E})$-mesurable.*

*(ii) Les variables $(X_n)_{n\geq 1}$ sont i.i.d. de loi $\mu(\,\cdot\,)/\mu(E)$.*

*(iii) La suite $(X_n)_{n\geq 1}$ est indépendante de $M$.*

*(iv) $\mathbf{P}$-p.s. $\mathbf{\Pi} = \big\{X_1, \ldots, X_M\big\}$ dès que $M \geq 1$.*

Cette proposition se généralise immédiatement au cas sigma-fini.

**Proposition III.3.15** *On suppose $(\alpha)$, $(\beta)$ et $(\gamma)$ ci-dessus, excepté que $\mu$ est seulement supposée sigma-finie ; il existe donc une partition $\mathscr{E}$-mesurable de $E$, notée $\mathbf{B} = (B_p)_{p\geq 1}$ telle que pour tout $p \geq 1$, on ait $0 < \mu(B_p) < \infty$. Pour tout $p \geq 1$, on note $\mathscr{G}_p$ la tribu engendrée par $U$ et $\mathbf{\Pi} \cap B_p$ et on note $M_p = N_{B_p}(\mathbf{\Pi})$. Alors, il existe un tableau de variables $(X_n^{(p)})_{p,n\geq 1}$ qui satisfont les conditions suivantes.*

*(i) Pour tous $n, p \geq 1$, $X_n^{(p)} : \Omega \to B_p$ est $(\mathscr{G}_p, \mathscr{E})$-mesurable.*

*(ii) Les variables $(X_n^{(p)})_{p,n\geq 1}$ sont indépendantes et pour tous $p, n \geq 1$, $X_n^{(p)}$ a pour loi $\mu(\,\cdot \cap B_p)/\mu(B_p)$.*

*(iii) Le tableau de variables $(X_n^{(p)})_{p,n\geq 1}$ est indépendant de la suite $(M_p)_{p\geq 1}$.*

*(iv) $\mathbf{P}$-p.s. pour tout $p \geq 1$, $\mathbf{\Pi} \cap B_p = \big\{X_1^{(p)}, \ldots, X_{M_p}^{(p)}\big\}$ dès que $M_p \geq 1$.*

**Preuve du lemme III.3.12.** Comme $(E, \mathscr{E})$ est séparable et séparé, il existe $C \subset \mathbb{R}$ tel que $(E, \mathscr{E})$ et $(C, \mathscr{B}(C))$ soient isomorphes en tant qu'espaces mesurables et il suffit de montrer le lemme dans le cas où $(E, \mathscr{E}) = (C, \mathscr{B}(C))$. On rappelle que $\mathscr{B}(C)$ est la tribu trace de $\mathscr{B}(\mathbb{R})$ sur $C$ mais aussi la tribu borélienne de la topologie induite sur $C$ par la topologie usuelle sur $\mathbb{R}$.

On suppose tout d'abord que $(E, \mathscr{E}) = (\mathbb{R}, \mathscr{B}(\mathbb{R}))$. On note $O_0$ l'ensemble des vecteurs de $\mathbb{R}^n$ dont les coordonnées sont distinctes. Il est clair que $O_0$ est un ouvert donc $O_0 \in \mathscr{B}(\mathbb{R}^n) = \mathscr{B}(\mathbb{R})^{\otimes n}$. Pour tout $k \in \mathbb{Z}$ et tout $p \in \mathbb{N}$, on pose $I_{p,k} = [k2{-}p, (k{+}1)2^{-p}[$ et pour tout $\mathbf{k} = (k_1, \ldots, k_n) \in \mathbb{Z}^n$, on note $C_{\mathbf{k},p} = I_{k_1,p} \times \ldots \times I_{k_n,p}$, le cube dyadique de « coin » $2^{-p}\mathbf{k}$. Si les $k_n$ sont distincts, alors $\nu(C_{\mathbf{k},p}) = \mu^{\otimes n}(C_{\mathbf{k},p})$. Soit un ouvert $U \subset \mathbb{R}^n$ ;

$U \cap O_0$ est également un ouvert et on voit qu'il existe une suite $\mathbf{k}_m \in \mathbb{Z}^n$, $p_m \in \mathbb{N}$, $m \in \mathbb{N}$, telle que les cubes $(C_{\mathbf{k}_m,p_m})_{m\in\mathbb{N}}$ forment une partition de $U \cap O_0$ ; comme $C_{\mathbf{k}_m,p_m} \subset O_0$, les coordonnées de $\mathbf{k}_m$ sont distinctes, ce qui implique que $\nu(U \cap O_0) = \sum_{m\in\mathbb{N}} \nu(C_{\mathbf{k}_m,p_m}) = \sum_{m\in\mathbb{N}} \mu^{\otimes n}(C_{\mathbf{k}_m,p_m}) = \mu^{\otimes n}(U \cap O_0)$. On observe que $\mu^{\otimes n}(U \cap O_0) = \mu^{\otimes n}(U)$ car $\mu$ est diffuse. Alors, la restriction à $O_0$ de $\nu$ coïncide avec $\mu^{\otimes n}$ sur le pi-système des ouverts de $\mathbb{R}^n$ ; par unicité du prolongement des mesures, ces mesures sont égales.

On fixe ensuite $C \subset \mathbb{R}$ et on considère le cas où $(E, \mathscr{E}) = (C, \mathscr{B}(C))$, $\mathscr{B}(C)$. Pour tout $B' \in \mathscr{B}(\mathbb{R})^{\otimes n}$, on pose $\nu'(B') = \nu(B' \cap C^n)$ et pour tout $A' \in \mathscr{B}(\mathbb{R})$, on pose $\mu'(A') = \mu(A' \cap C)$. On vérifie que $\nu'$ et $\mu'$ sont des mesures sigma finies, que $\mu'$ est diffuse et que pour tous $A'_1, \ldots, A'_n \in \mathscr{B}(\mathbb{R})$, deux à deux-disjoints, on a $\nu'(A'_1 \times \ldots \times A'_n) = \mu'(A'_1) \ldots \mu'(A'_n)$. Par ce qui précède, cela implique que $\nu'(\cdot \cap O_0) = \mu'^{\otimes n}$. Soit $B \in \mathscr{B}(C)^{\otimes n}$. Comme $\mathscr{B}(C)^{\otimes n}$ est aussi la tribu borélienne de la topologie induite sur $C^n$ par la topologie de la norme sur $\mathbb{R}^n$, il existe $B' \in \mathscr{B}(\mathbb{R})^{\otimes n}$ tel que $B = B' \cap C^n$. Puisque $O = O_0 \cap C^n$, on a donc $\nu(B \cap O) = \nu'(B' \cap O_0) = \mu'^{\otimes n}(B')$. On vérifie ensuite que $\mu'^{\otimes n}(B') = \mu'^{\otimes n}(B)$, ce qui permet de conclure. ■

# III.4 Outils de calculs.

## III.4.a Formules de Palm.

Toutes les variables sont définies sur un même espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$ supposé complet. Nous renvoyons au lemme III.2.20 pour les questions de mesurabilité soulevées implicitement par le théorème suivant.

**Théorème III.4.1** **(Formule de Palm)** *Soit $(E, \mathscr{E})$, un espace séparable et séparé. Soit $\mathbf{\Pi} : \Omega \to \mathbf{S}_E$, un nuage Poissonnien d'intensité $\mu$, supposée sigma-finie et non-nulle. Soit $(E_0, \mathscr{E}_0)$, un espace mesurable et soit $Z : \Omega \to E_0$, une variable $(\mathscr{F}, \mathscr{E}_0)$-mesurable supposée indépendante de $\mathbf{\Pi}$. Soit $F : E_0 \times E \times \mathbf{S}_E \to [0, \infty]$, une fonction $\mathscr{E}_0 \otimes \mathscr{E} \otimes \mathscr{S}_E$-mesurable. Alors,*

$$\mathbf{E}\left[\sum_{X\in\mathbf{\Pi}} F(Z, X, \mathbf{\Pi}\backslash\{X\})\right] = \int_E \mu(dx)\, \mathbf{E}[F(Z, x, \mathbf{\Pi})] \,, \tag{III.30}$$

*avec la convention qu'une somme sur un ensemble d'indices vide est nulle.*

**Preuve :** on remarque que (III.30) ne dépend que de la loi jointe de $(Z, \mathbf{\Pi})$. On peut donc choisir $\mathbf{\Pi}$ comme dans le théorème de construction III.3.11 et $Z$ indépendante des variables $X_n^{(p)}$ et $M_p$ (ou bien on applique la proposition de représentation III.3.15). On pose alors

$$a = \mathbf{E}\left[\sum_{X\in\mathbf{\Pi}} F(Z, X, \mathbf{\Pi}\backslash\{X\})\right] \quad \text{et} \quad \mathbf{\Pi}_p^* = \bigcup_{p'\in\mathbb{N}\backslash\{p\}} \mathbf{\Pi}_{p'} \,.$$

Par interversion positive série/intégrale,

$$a = \sum_{p\geq 1}\sum_{n\geq 1} \mathbf{P}(M_p = n) \sum_{1\leq k\leq n} \mathbf{E}\big[F\big(Z, X_k^{(p)}, \mathbf{\Pi}_p^* \cup \{X_1^{(p)}, \ldots, X_n^{(p)}\}\backslash\{X_k^{(p)}\}\big)\big] \,.$$

On observe que pour toute permutation $\gamma$ de $\{1, \ldots, n\}$, on a

$$\big(X_1^{(p)}, \ldots, X_n^{(p)}\big) \overset{\text{(loi)}}{=} \big(X_{\gamma(1)}^{(p)}, \ldots, X_{\gamma(n)}^{(p)}\big) \,. \tag{III.31}$$

On pose $b_{p,n,k} = \mathbf{E}\big[F\big(Z, X_k^{(p)}, \mathbf{\Pi}_p^* \cup \{X_1^{(p)}, \ldots, X_n^{(p)}\}\backslash\{X_k^{(p)}\}\big)\big]$ et on déduit de (III.31) que

$$b_{p,n,k} \;=\; \mathbf{E}\big[F\big(Z, X_n^{(p)}, \mathbf{\Pi}_p^* \cup \{X_1^{(p)}, \ldots, X_n^{(p)}\}\backslash\{X_n^{(p)}\}\big)\big]$$

$$
\begin{aligned}
&= \mathbf{E}\big[F\big(Z, X_n^{(p)}, \mathbf{\Pi}_p^* \cup \{X_1^{(p)}, \ldots, X_{n-1}^{(p)}\}\big)\big] \\
&= \int_{B_p} \frac{\mu(dx)}{\mu(B_p)} \mathbf{E}\big[F\big(Z, x, \mathbf{\Pi}_p^* \cup \{X_1^{(p)}, \ldots, X_{n-1}^{(p)}\}\big)\big] \ ,
\end{aligned}
$$

avec la convention que $\{X_1^{(p)}, \ldots, X_{n-1}^{(p)}\} = \emptyset$ si $n = 1$. Par conséquent

$$
\begin{aligned}
a &= \sum_{p\geq 1}\sum_{n\geq 1} \mathbf{P}(M_p = n)\, n \int_{B_p} \frac{\mu(dx)}{\mu(B_p)} \mathbf{E}\big[F\big(Z, x, \mathbf{\Pi}_p^* \cup \{X_1^{(p)}, \ldots, X_{n-1}^{(p)}\}\big)\big] \\
&= \sum_{p\geq 1}\sum_{n\geq 1} e^{-\mu(B_p)} \frac{\mu(B_p)^n}{n!}\, n \int_{B_p} \frac{\mu(dx)}{\mu(B_p)} \mathbf{E}\big[F\big(Z, x, \mathbf{\Pi}_p^* \cup \{X_1^{(p)}, \ldots, X_{n-1}^{(p)}\}\big)\big] \\
&= \sum_{p\geq 1}\sum_{n\geq 1} e^{-\mu(B_p)} \frac{\mu(B_p)^{n-1}}{(n-1)!} \int_{B_p} \mu(dx) \mathbf{E}\big[F\big(Z, x, \mathbf{\Pi}_p^* \cup \{X_1^{(p)}, \ldots, X_{n-1}^{(p)}\}\big)\big] \\
&= \sum_{p\geq 1}\sum_{n'\geq 0} \mathbf{P}(M_p = n') \int_{B_p} \mu(dx) \mathbf{E}\big[F\big(Z, x, \mathbf{\Pi}_p^* \cup \{X_1^{(p)}, \ldots, X_{n'}^{(p)}\}\big)\big] \\
&= \sum_{p\geq 1} \int_{B_p} \mu(dx) \sum_{n'\geq 0} \mathbf{P}(M_p = n') \mathbf{E}\big[F\big(Z, x, \mathbf{\Pi}_p^* \cup \{X_1^{(p)}, \ldots, X_{n'}^{(p)}\}\big)\big] \\
&= \sum_{p\geq 1} \int_{B_p} \mu(dx) \mathbf{E}\big[F\big(Z, x, \mathbf{\Pi}_p^* \cup \mathbf{\Pi}_p\}\big)\big] \\
&= \sum_{p\geq 1} \int_{B_p} \mu(dx) \mathbf{E}[F(Z, x, \mathbf{\Pi})] = \int_E \mu(dx)\, \mathbf{E}[F(Z, x, \mathbf{\Pi})],
\end{aligned}
$$

qui est bien le résultat voulu. ■

**Remarque III.4.2** On se donne $G : \mathbf{S}_E \to \mathbb{R}_+$ et $g : E \to \mathbb{R}_+$ qui sont respectivement $\mathscr{S}_E$-mesurable et $\mathscr{E}$-mesurable. La formule de Palm appliquée à $F(x, \pi) = g(x)G(\pi)$, donne

$$
\mathbf{E}\left[\sum_{X\in\mathbf{\Pi}} g(X)G(\mathbf{\Pi}\backslash\{X\})\right] = \mathbf{E}[G(\mathbf{\Pi})] \int_E g(x)\mu(dx) \ .
$$

Cette formule peut s'interpréter de la manière suivante : si on choisit un point $X$ de $\mathbf{\Pi}$ "uniformément au hasard" (c'est-à-dire selon la mesure de comptage) alors ce point a pour "loi" $\mu$, il est indépendant du nuage $\mathbf{\Pi}\backslash\{X\}$ qui est Poissonnien d'intensité $\mu$. Autrement dit, si on observe $\mathbf{\Pi}$ depuis un point typique de $\mathbf{\Pi}$ (qui est enlèvé de $\mathbf{\Pi}$, on observe toujours un nuage Poissonnien. De plus, le nouveau nuage observé ne donne pas d'information sur la position d'observation. □

**Remarque III.4.3** On conserve les mêmes hypothèses que le théorème précédent en ce qui concerne $\mathbf{\Pi}$ et $Z$ mais on considère une application $F : E_0 \times E \times \mathbf{S}_E \to \mathbb{R}$ ou $\mathbb{C}$ qui est $\mathscr{E}_0 \otimes \mathscr{E} \otimes \mathscr{S}_E$-mesurable et telle que

$$
\int_E \mu(dx)\mathbf{E}[\,|F(Z, x, \mathbf{\Pi})|\,] < \infty \ . \tag{III.32}
$$

Alors on déduit facilement du théorème III.4.1, qui traite le cas des fonctions positives, que

$$
\mathbf{E}\left[\sum_{X\in\mathbf{\Pi}} F(Z, X, \mathbf{\Pi}\backslash\{X\})\right] = \int_E \mu(dx)\mathbf{E}\left[F(Z, x, \mathbf{\Pi})\right] \ .
$$

En effet, si $F$ est réelle, il suffit de séparer $F$ en sa partie positive et sa partie négative, de montrer que toutes les expressions ont un sens grâce à l'hypothèse (III.32) et au théorème III.4.1 et de déduire le résultat. Dans le cas où $F$ est à valeurs complexes, il faut raisonner sur la partie réelle et la partie imaginaire. Nous laissons les détails au lecteur. $\square$

En itérant la formule de Palm, on prouve le résultat suivant.

**Théorème III.4.4** **(Formule de Palm à $n$ points)** *Soit $(E, \mathscr{E})$, un espace séparable et séparé. Soit $\mathbf{\Pi}: \Omega \to \mathbf{S}_E$, un nuage Poissonnien d'intensité $\mu$, supposée sigma-finie et non-nulle. Soit $(E_0, \mathscr{E}_0)$, un espace mesurable et soit $Z: \Omega \to E_0$, une variable $(\mathscr{F}, \mathscr{E}_0)$-mesurable supposée indépendante de $\mathbf{\Pi}$. Soit $F: E_0 \times E^n \times \mathbf{S}_E \to [0, \infty]$, une fonction $\mathscr{E}_0 \otimes \mathscr{E}^{\otimes n} \otimes \mathscr{S}_E$-mesurable. On rappelle la notation*

$$\Phi_F^{(n)}(Z, \mathbf{\Pi}) = \sum_{\substack{X_1, \ldots, X_n \in \mathbf{\Pi} \\ \text{distincts}}} F\big(Z, X_1, \ldots, X_n,\, \mathbf{\Pi} \backslash \{X_1, \ldots, X_n\}\big) .$$

*avec la convention qu'une somme sur un ensemble d'indices vide est nulle. Alors,*

$$\mathbf{E}\Big[\Phi_F^{(n)}(Z, \mathbf{\Pi})\Big] = \int_{E^n} \mu(dx_1) \ldots \mu(dx_n)\, \mathbf{E}[F(Z, x_1, \ldots, x_n,\, \mathbf{\Pi})] . \tag{III.33}$$

*De même, on rappelle que $\Lambda_F^{(n)}(Z, \mathbf{\Pi}) = \sum_{X_1, \ldots, X_n \in \mathbf{\Pi}, \, \text{distincts}} F\big(Z, X_1, \ldots, X_n,\, \mathbf{\Pi}\big)$. Alors*

$$\mathbf{E}\Big[\Lambda_F^{(n)}(Z, \mathbf{\Pi})\Big] = \int_{E^n} \mu(dx_1) \ldots \mu(dx_n)\, \mathbf{E}[F(Z, x_1, \ldots, x_n,\, \mathbf{\Pi} \cup \{x_1, \ldots, x_n\})] . \tag{III.34}$$

**Preuve :** on fait une récurrence. Le théorème III.4.1 montre (III.33) pour $n = 1$. On suppose que (III.33) est vérifiée au rang $n$. Soit $F : E_0 \times E^{n+1} \times \mathbf{S}_E \to [0, \infty]$, une fonction $\mathscr{E}_0 \otimes \mathscr{E}^{\otimes n+1} \otimes \mathscr{S}_E$-mesurable. On pose $E_1 = E_0 \times E$ muni de la tribu $\mathscr{E}_1 = \mathscr{E}_0 \otimes \mathscr{E}$ et on définit la fonction $G : E_1 \times E^n \times \mathbf{S}_E \to [0, \infty]$, par $G\big((z, x), x_2, \ldots, x_{n+1}, \pi\big) = F(z, x, x_2, \ldots, x_{n+1}, \pi)$. C'est une fonction $\mathscr{E}_1 \otimes \mathscr{E}^{\otimes n} \otimes \mathscr{S}_E$-mesurable. On pose $F_0 = \Phi_G^{(n)}$. La formule de Palm au rang $n = 1$ (théorème III.4.1) implique que

$$\begin{aligned} \mathbf{E}\Big[\sum_{X_1 \in \mathbf{\Pi}} F_0(Z, X_1, \mathbf{\Pi} \backslash \{X_1\})\Big] &= \int_E \mu(dx_1)\, \mathbf{E}[F_0(Z, x_1, \mathbf{\Pi})] \\ &= \int_E \mu(dx_1) \mathbf{E}\big[\Phi_G^{(n)}(Z, x_1, \mathbf{\Pi})\big]. \end{aligned} \tag{III.35}$$

On fixe $x_1 \in E$. L'hypothèse de récurrence implique au rang $n$ que

$$\begin{aligned} \mathbf{E}\big[\Phi_G^{(n)}(Z, x_1, \mathbf{\Pi})\big] &= \mathbf{E}\Big[\sum_{\substack{X_2, \ldots, X_{n+1} \in \mathbf{\Pi} \\ \text{distincts}}} F\big(Z, x_1, X_2 \ldots, X_{n+1},\, \mathbf{\Pi} \backslash \{X_2, \ldots, X_{n+1}\}\big)\Big] \\ &= \int_{E^n} \mu(dx_2) \ldots \mu(dx_n)\, \mathbf{E}[F(Z, x_1, x_2, \ldots, x_{n+1},\, \mathbf{\Pi})] . \end{aligned} \tag{III.36}$$

On remarque enfin que

$$\sum_{X_1 \in \mathbf{\Pi}} F_0(Z, X_1, \mathbf{\Pi} \backslash \{X_1\}) = \sum_{X_1 \in \mathbf{\Pi}} \sum_{\substack{X_2, \ldots, X_{n+1} \in \mathbf{\Pi} \backslash \{X_1\} \\ \text{distincts}}} F\big(Z, X_1,\, X_2, \ldots, X_{n+1}, \mathbf{\Pi} \backslash \{X_1, X_2, \ldots, X_{n+1}\}\big)$$

$$= \Phi_F^{(n+1)}(Z, \mathbf{\Pi}). \tag{III.37}$$

Les égalités (III.35), (III.36) et (III.37) permettent alors de montrer (III.33) au rang $n+1$, ce qui conclut la preuve de (III.33). Pour démontrer la formule (III.34), il suffir d'appliquer la formule (III.33) à la fonction $G(z, x_1, \ldots, x_n, \pi) = F(z, x_1, \ldots, x_n, \pi \cup \{x_1, \ldots, x_n\})$. ■

En application nous donnons une formule de moment d'ordre 2 pour $\Phi_F(\mathbf{\Pi})$.

**Proposition III.4.5** *Soit $(E, \mathscr{E})$, un espace séparable et séparé. Soit $\mathbf{\Pi} : \Omega \to \mathbf{S}_E$, un nuage Poissonnien d'intensité $\mu$, supposée sigma-finie et non-nulle. Soit $(E_0, \mathscr{E}_0)$, un espace mesurable et soit $Z : \Omega \to E_0$, une variable $(\mathscr{F}, \mathscr{E}_0)$-mesurable supposée indépendante de $\mathbf{\Pi}$. Soit $F : E_0 \times E \times \mathbf{S}_E \to [0, \infty]$, une fonction $\mathscr{E}_0 \otimes \mathscr{E} \otimes \mathscr{S}_E$-mesurable. On rappelle que $\Phi_F(\mathbf{\Pi}) = \sum_{X \in \mathbf{\Pi}} F(Z, X, \mathbf{\Pi} \backslash \{X\})$. Alors, on a*

$$\mathbf{E}\Big[\Phi_F(\mathbf{\Pi})^2\Big] = \int_E \mu(dx)\, \mathbf{E}\big[F(Z, x, \mathbf{\Pi})^2\big] + \int_{E^2} \mu(dx)\mu(dy)\, \mathbf{E}\big[F(Z, x, \mathbf{\Pi} \cup \{y\}) F(Z, y, \mathbf{\Pi} \cup \{x\})\big]. \tag{III.38}$$

**Preuve :** pour tout sous-ensemble fini $J \subset \mathbf{\Pi}$, on a bien

$$\Big(\sum_{X \in J} F(Z, X, \mathbf{\Pi} \backslash \{X\})\Big)^2 = \sum_{X \in J} F(Z, X, \mathbf{\Pi} \backslash \{X\})^2 + \sum_{\substack{X_1, X_2 \in J \\ \text{distincts}}} G(Z, X_1, X_1, \mathbf{\Pi} \backslash \{X_1, X_2\}) .$$

où $G(z, x, y, \pi) = F(z, x, \pi \cup \{y\}) F(z, y, \pi \cup \{x\})$. En effet, il s'agit juste du developpement du carré d'une somme finie. En prenant le supremum sur tous les sous-ensembles finis $J \subset \mathbf{\Pi}$, on peut remplacer $J$ par $\mathbf{\Pi}$ dans l'expression précédente : on a $\Phi_F(\mathbf{\Pi})^2 = \Phi_{F^2}(\mathbf{\Pi}) + \Phi_G^{(2)}(\mathbf{\Pi})$ et on conclut par le théorème III.4.4. ■

### III.4.b Formules exponentielles.

Toutes les variables sont définies sur un même espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$ supposé complet. Soit $(E, \mathscr{E})$, un espace mesurable et $f : E \to [0, \infty]$, une fonction $\mathscr{E}$-mesurable. On rappelle la définition III.2.4 (page 235) de la *$f$-fonction de comptage* $N_f(\pi) = \sum_{x \in \pi} f(x)$, pour tout $\pi \in \mathbf{S}_E$. On rappelle également l'approximation $f_n$ de $f$, (III.10) et (III.11) page 235.

**Théorème III.4.6** **(Formule exponentielle positive)** *Soit $(E, \mathscr{E})$, un espace mesurable séparable et séparé. Soit $\mathbf{\Pi} : \Omega \to \mathbf{S}_E$, un nuage Poissonnien d'intensité $\mu$. Soit $f : E \to [0, \infty]$ une application $\mathscr{E}$-mesurable. Alors*

$$\mathbf{E}\big[\exp(-N_f(\mathbf{\Pi}))\big] = \exp\Big(-\int_E \mu(dx)\big(1 - e^{-f(x)}\big)\Big) , \tag{III.39}$$

*avec la convention* $\exp(-\infty) = 0$. *De plus, on a*

$$\mathbf{E}\big[N_f(\mathbf{\Pi})\big] = \int_E \mu(dx) f(x) \quad \text{et} \quad \mathbf{E}\big[N_f(\mathbf{\Pi})^2\big] = \int_E \mu(dx)\, f(x)^2 + \Big(\int_E \mu(dx) f(x)\Big)^2 ,$$

**Preuve :** par (III.10), on a

$$\mathbf{E}\big[\exp(-N_{f_n}(\mathbf{\Pi}))\big] = \prod_{0 \leq k < n2^n} \exp\big(-\mu(A_{k,n})\big(1 - e^{-k2^{-n}}\big)\big) = \exp\Big(-\int_E \mu(dx)\big(1 - e^{-f_n(x)}\big)\Big) ,$$

ce qui entraîne (III.39) par (III.11), par convergence dominée et par convergence monotone. Le premier moment de $N_f(\mathbf{\Pi})$ est donné par le lemme III.2.5 $(ii)$. La formule donnant le moment d'ordre 2 découle directement de la proposition III.4.5 avec $F(z, x, \pi) = f(x)$. ■

**Lemme III.4.7** **(Caractérisation)** *Soit $(E, \mathscr{E})$, un espace mesurable séparable et séparé. Soit $\mathbf{\Pi} : \Omega \to \mathbf{S}_E$, un nuage aléatoire d'intensité $\mu$. C'est un nuage Poissonnien ssi pour toute fonction $f$ positive $\mathscr{E}$-mesurable bornée (III.39) est vérifiée.*

**Preuve :** si $\mathbf{\Pi}$ est Poissonnien, alors le théorème III.4.6 montre que (III.39) est vérifié pour toute fonction $f$ positive, $\mathscr{E}$-mesurable bornée. Réciproquement, soient $A_1, \ldots, A_p \in \mathscr{E}$, disjoints deux-à-deux et soient $\lambda_1, \ldots, \lambda_p \in \mathbb{R}_+$. En choisissant $f = \lambda_1 \mathbf{1}_{A_1} + \ldots + \lambda_p \mathbf{1}_{A_p}$ dans (III.39), on obtient

$$\begin{aligned}\mathbf{E}\big[e^{-\lambda_1 N_{A_1}(\mathbf{\Pi}) - \ldots - \lambda_1 N_{A_1}(\mathbf{\Pi})}\big] &= \mathbf{E}\big[e^{-N_f(\mathbf{\Pi})}\big] = \exp\Big( - \int_E \mu(dx)\big(1 - e^{-f(x)}\big)\Big) \\ &= \prod_{1 \le k \le p} \exp(-\mu(A_k)(1 - e^{-\lambda_k}))\,.\end{aligned}$$

Par injectivité de la transformée de Laplace, cela entraîne que $N_{A_1}(\mathbf{\Pi}), \ldots, N_{A_p}(\mathbf{\Pi})$ sont des variables de Poisson indépendantes si $\mu(A_1), \ldots, \mu(A_p)$ sont des quantités finies. Le cas général est facile à déduire car si $\mu(A) = \infty$, $N_A(\Pi)$ est déterministe et vaut l'infini. ■

**Remarque III.4.8** On observe également que si $Y : \Omega \to \mathbb{N}$ suit une loi de Poisson de paramètre $\theta \in ]0, \infty[$, alors $Y$ admet des moment exponentiels de tous ordres :

$$\forall \lambda \in \mathbb{R}_+, \quad \mathbf{E}\big[e^{\lambda Y}\big] = \sum_{n \in \mathbb{N}} \frac{e^{-\theta}\theta^n}{n!} e^{-\theta} e^{\lambda n} = \exp\big(\theta(e^\lambda - 1)\big)\,.$$

En reprenant l'approximation de la preuve du théorème III.4.6, on montre sous les mêmes hypothèses que ce théorème que

$$\mathbf{E}\big[\exp\big(N_f(\mathbf{\Pi})\big)\big] = \exp\Big(\int_E \mu(dx)\,\big(e^{f(x)} - 1\big)\Big)\,, \tag{III.40}$$

ces quantités pouvant être infinies, avec la convention $\exp(\infty) := \infty$. □

**Proposition III.4.9** *Soit $(E, \mathscr{E})$, un espace mesurable séparable et séparé. Soit $\mathbf{\Pi} : \Omega \to \mathbf{S}_E$, un nuage Poissonnien d'intensité $\mu$. Soit $f : E \to \mathbb{R}_+$ une application $\mathscr{E}$-mesurable ($f$ ne prend pas la valeur $\infty$). Alors, on a l'alternative suivante.*

**(a)** *Si $\int_E (1 \wedge f(x))\, \mu(dx) < \infty$, alors p.s. $N_f(\mathbf{\Pi}) < \infty$.*

**(b)** *Si $\int_E (1 \wedge f(x))\, \mu(dx) = \infty$, alors p.s. $N_f(\mathbf{\Pi}) = \infty$.*

**Preuve :** on pose $c = 1 - e^{-1}$. On a $c(1 \wedge y) \le 1 - e^{-y} \le 1 \wedge y$, pour tout $y \in \mathbb{R}_+$, car la fonction $y \mapsto 1 - e^{-y}$ est concave. Donc pour tout $\lambda \in \mathbb{R}_+$,

$$c \int_E \mu(dx)\,\big(1 \wedge (\lambda f(x))\big) \le \int_E \mu(dx)\,\big(1 - e^{-\lambda f(x)}\big) \le \int_E \mu(dx)\,\big(1 \wedge (\lambda f(x))\big)\,. \tag{III.41}$$

Supposons que $\mathbf{P}\big(N_f(\mathbf{\Pi}) < \infty\big) > 0$, alors $\mathbf{E}[\exp(-N_f(\mathbf{\Pi}))] > 0$, et la formule exponentielle implique que $\int_E \mu(dx)(1 - e^{-f(x)}) < \infty$. L'inégalité (III.41) avec $\lambda = 1$ implique donc que $\int_E \mu(dx)\,(1 \wedge f(x)) < \infty$.

Réciproquement, on suppose que $\int_E \mu(dx)\,(1 \wedge f(x)) < \infty$. On observe que pour tout $\lambda \in [0, 1]$, $1 - e^{-\lambda f(x)} \le 1 \wedge f(x)$. Comme $f$ ne peut pas prendre la valeur $\infty$, on a $\lim_{\lambda \to 0} 1 - e^{-\lambda f(x)} = 0$. Par convergence dominée, on obtient $\lim_{\lambda\to 0}\int_E \mu(dx)\left(1-e^{-\lambda f(x)}\right)=0$ et un argument élémentaire combiné à la formule exponentielle montre que

$$\mathbf{P}\big(N_f(\mathbf{\Pi})<\infty\big)=\lim_{\lambda\to 0}\mathbf{E}[\exp(-\lambda N_f(\mathbf{\Pi}))]=\lim_{\lambda\to 0}\exp\Big(\int_E \mu(dx)\left(1-e^{-\lambda f(x)}\right)\Big)=1\ ,$$

ce qui montre **(a)**. Supposons que $\int_E \mu(dx)(1\wedge f(x))=\infty$. Alors, par (III.41) $\int_E \mu(dx)(1-e^{-f(x)})=\infty$, et la formule exponentielle entraîne que $\mathbf{E}[\exp(-N_f(\mathbf{\Pi}))]=0$, ce qui est équivalent à dire que $\mathbf{P}(N_f(\mathbf{\Pi})=\infty)=1$. Cela montre une implication de **(b)**. L'implication réciproque suit immédiatement de **(a)**. ■

On se place sous les hypothèses du théorème III.4.6. Soit $f:E\to\mathbb{R}$, une fonction $\mathscr{E}$-mesurable telle que

$$\int_E \mu(dx)\ (1\wedge|f(x)|\,)\ <\ \infty\ . \tag{III.42}$$

Cela implique que $\int_E \mu(dx)\ (1\wedge f^+(x)\,)$ et $\int_E \mu(dx)\ (1\wedge f^-(x)\,)$ sont des quantités finies. Ici $f^+$ et $f^-$ sont les parties positives et négatives de $f$. La proposition (III.4.9) entraîne alors que $N_{f^+}(\mathbf{\Pi})$ et $N_{f^-}(\mathbf{\Pi})$ sont des variables finies presque sûrement et on a

$$N_{|f|}(\mathbf{\Pi})=\sum_{X\in\mathbf{\Pi}}|f(X)|=N_{f^+}(\mathbf{\Pi})+N_{f^-}(\mathbf{\Pi})\ <\infty\ .$$

On pose alors

$$N_f(\mathbf{\Pi})=N_{f^+}(\mathbf{\Pi})-N_{f^-}(\mathbf{\Pi}):=\sum_{X\in\mathbf{\Pi}}f(X)\ . \tag{III.43}$$

**Théorème III.4.10** **(Formule exp. réelle)** *Soit $(E,\mathscr{E})$, un espace mesurable séparable et séparé. Soit $\mathbf{\Pi}:\Omega\to\mathbf{S}_E$, un nuage Poissonnien d'intensité $\mu$. Soit $f:E\to\mathbb{R}$ une application $\mathscr{E}$-mesurable satisfaisant (III.42), ce qui permet de définir $N_f(\mathbf{\Pi})$ par (III.43). Alors,*

$$\mathbf{E}\big[\exp(iN_f(\mathbf{\Pi}))\big]=\exp\Big(-\int_E \mu(dx)\big(1-e^{i\,f(x)}\big)\Big)\ . \tag{III.44}$$

*Si $\int_E \mu(dx)|f(x)|<\infty$, alors $N_f(\mathbf{\Pi})$ est une variable intégrable et on a*

$$\mathbf{E}\Big[\sum_{X\in\mathbf{\Pi}}f(X)\Big]=\int_E \mu(dx)f(x)\ . \tag{III.45}$$

*Si $\int_E \mu(dx)|f(x)|<\infty$ et $\int_E \mu(dx)f(x)^2<\infty$, alors $N_f(\mathbf{\Pi})$ a un moment d'ordre deux et*

$$\mathbf{var}(N_f(\mathbf{\Pi}))=\int_E \mu(dx)f(x)^2\ .$$

**Preuve :** soit $g:E\to\mathbb{R}_+$, une fonction $\mathscr{E}$-mesurable telle que $\int_E \mu(dx)(1\wedge g(x))<\infty$. Par la proposition III.4.9, on a $\mathbf{P}(N_g(\mathbf{\Pi})<\infty)=1$. Comme dans la preuve de la formule exponentielle positive (théorème III.4.6) on démontre facilement que

$$\forall u\in\mathbb{R}\,,\quad \mathbf{E}[\exp(iu\,N_g(\mathbf{\Pi}))\,]=\exp\big(-\int_E \mu(dx)\big(1-e^{iu\,g(x)}\big)\big)\ . \tag{III.46}$$

On pose $E_+:=\{f\geq 0\}$ et $E_-:=\{f<0\}$, qui sont deux Boréliens disjoints. On a donc $f=f^+$ sur $E_+$ et $-f=f^-$ sur $E_-$ . On pose également $\mathbf{\Pi}_+:=\mathbf{\Pi}\cap E_+$ et $\mathbf{\Pi}_-:=\mathbf{\Pi}\cap E_-$. Le principe de restriction implique

que $\mathbf{\Pi}_+$ et $\mathbf{\Pi}_-$ sont deux nuages Poissonniens indépendants d'intensités respectives $\mu(\cdot \cap E_+)$ et $\mu(\cdot \cap E_-)$. En utilisant (III.46) avec $\mathbf{\Pi}_{+/-}$, $u = +1$ ou $-1$ et $g = |f|$, on obtient

$$\begin{aligned}
\mathbf{E}[\exp(iN_f(\mathbf{\Pi}))] &= \mathbf{E}[\exp(iN_f(\mathbf{\Pi}_+))]\ \mathbf{E}[\exp(-iN_{-f}(\mathbf{\Pi}_-))] \\
&= \exp\Big(-\int_{E_+}\mu(dx)\big(1-e^{i\,f(x)}\big)\Big)\exp\Big(-\int_{E_-}\mu(dx)\big(1-e^{i\,f(x)}\big)\Big) \\
&= \exp\Big(-\int_{E}\mu(dx)\big(1-e^{i\,f(x)}\big)\Big)\,,
\end{aligned}$$

ce qui montre le point (III.44). On a également, $\mathbf{E}[N_{f^{+/-}}(\mathbf{\Pi})] = \int_E \mu(dx) f^{+/-}(x)$ et on en déduit (III.45). On rappelle que le théorème III.4.6 implique

$$\mathbf{E}[N_{|f|}(\mathbf{\Pi}_{+/-})^2] = \int_{E_{+/-}} \mu(dx)\,|f(x)|^2 + \Big(\int_{E_{+/-}} \mu(dx)\,|f(x)|\Big)^2 .$$

Donc si $\int_E \mu(dx)|f(x)| < \infty$ et $\int_E \mu(dx) f(x)^2 < \infty$, $N_f(\mathbf{\Pi}) = N_{|f|}(\mathbf{\Pi}_+) - N_{|f|}(\mathbf{\Pi}_-)$ admet un moment d'ordre deux et par indépendance de $\mathbf{\Pi}_+$ et de $\mathbf{\Pi}_-$, on a

$$\mathbf{var}(N_f(\mathbf{\Pi})) = \mathbf{var}(N_f(\mathbf{\Pi}_+)) + \mathbf{var}(N_f(\mathbf{\Pi}_-)) = \int_E \mu(dx)(f^+(x))^2 + \int_E \mu(dx)(f^-(x))^2\,,$$

ce qui entraîne le résultat voulu. ■

## III.5 Nuages Poissonniens sur $\mathbb{R}_+ \times E$.

### III.5.a Premières propriétés.

Toutes les variables sont définies sur un même espace de probabilité noté $(\Omega, \mathscr{F}, \mathbf{P})$ supposé complet. Soit $(E, \mathscr{E})$, un espace mesurable séparable et séparé. On munit $\mathbb{R}_+ \times E$ de la tribu $\mathscr{B}(\mathbb{R}_+) \otimes \mathscr{E}$. Il est facile de vérifier que $(\mathbb{R}_+ \times E, \mathscr{B}(\mathbb{R}_+) \otimes \mathscr{E})$ est séparable et séparé. On rappelle ensuite que $\ell$ désigne la mesure de Lebesgue sur $\mathbb{R}_+$. Soit une mesure sigma-finie $\mu$ sur $(E, \mathscr{E})$ : la mesure produit $\ell \otimes \mu$ est également sigma-finie et elle est diffuse, même lorsque $\mu$ n'est pas diffuse.

**Proposition III.5.1** *Soit $\theta \in ]0, \infty[$. Soit $\nu : \mathscr{E} \to [0, \infty]$, une mesure de probabilité. Soit $(\mathscr{E}_n, X_n)$, $n \geq 1$, une suite de variables telles que $(\mathscr{E}_n)_{n\geq 1}$ soient i.i.d. de loi exponentielle de paramètre $\theta$, telles que $(X_n)_{n\geq 1}$ soient i.i.d. à valeurs dans $E$ de loi $\nu$, et telles que $(\mathscr{E}_n)_{n\geq 1}$ et $(X_n)_{n\geq 1}$ soient indépendantes. On pose*

$$\forall n \in \mathbb{N}^*,\ T_n := \mathscr{E}_1 + \ldots + \mathscr{E}_n \quad \text{et} \quad \mathbf{\Pi} := \big\{(T_n, X_n)\,;\, n \in \mathbb{N}^*\big\}.$$

*Alors, $\mathbf{\Pi}$ est un nuage Poissonnien sur $\mathbb{R}_+ \times E$ d'intensité $\theta\, \ell \otimes \nu$.*

**Preuve :** par le lemme III.4.7 page 255, il suffit de montrer que pour toute fonction mesurable $f : \mathbb{R}_+ \times E \to \mathbb{R}_+$,

$$\mathbf{E}\big[\exp\big(-N_f(\mathbf{\Pi})\big)\big] = \exp\Big(-\theta \int_{\mathbb{R}_+\times E} \ell(ds)\,\nu(dx)\big(1 - e^{-f(s,x)}\big)\Big)\,. \qquad \text{(III.47)}$$

On rappelle que $\mathbf{\Pi}_0 = \{T_n; n \in \mathbb{N}^*\}$ est un nuage Poissonnien d'intensité $\theta\ell$ (voir le théorème III.1.11, page 231 en appendice). Il est clairement indépendant de $(X_n)_{n\geq 1}$. On fixe $t > 0$ et on définit $f_t$ en posant $f_t(s, x) := \mathbf{1}_{[0,t]}(s) f(s, x)$. Par le théorème III.1.11 $(iii)$, on montre facilement que pour tout $n \in \mathbb{N}$,

$$\mathbf{E}\big[\exp\big(-N_{f_t}(\mathbf{\Pi})\big)\,\big|\, N_{[0,t]}(\mathbf{\Pi}_0) = n\big] = \Big(\frac{1}{t}\int_{[0,t]\times E} \ell(ds)\nu(dx) e^{-f(s,x)}\Big)^n .$$

On obtient alors

$$\begin{aligned}\mathbf{E}\big[\exp\big(-N_{f_t}(\mathbf{\Pi})\big)\big] &= \sum_{n\in\mathbb{N}} \mathbf{P}\big(N_{[0,t]}(\mathbf{\Pi}_0)=n\big)\Big(\tfrac{1}{t}\int_{[0,t]\times E}\ell(ds)\nu(dx)e^{-f(s,x)}\Big)^n \\ &= e^{-\theta t}\sum_{n\in\mathbb{N}}\frac{1}{n!}\Big(\theta\int_{[0,t]\times E}\ell(ds)\nu(dx)e^{-f(s,x)}\Big)^n \\ &= \exp\Big(-\theta\int_{[0,t]\times E}\ell(ds)\nu(dx)\big(1-e^{-f(s,x)}\big)\Big),\end{aligned}$$

Par convergence monotone, on a $\lim_{t\to\infty} N_{f_t}(\mathbf{\Pi}) = N_f(\mathbf{\Pi})$ et

$$\lim_{t\to\infty}\int_{[0,t]\times E}\ell(ds)\nu(dx)\big(1-e^{-f(s,x)}\big) = \int_{\mathbb{R}_+\times E}\ell(ds)\nu(dx)\big(1-e^{-f(s,x)}\big).$$

Par convergence dominée, on a $\lim_{t\to\infty}\mathbf{E}\big[\exp\big(-N_{f_t}(\mathbf{\Pi})\big)\big] = \mathbf{E}\big[\exp\big(-N_f(\mathbf{\Pi})\big)\big]$, ce qui montre (III.47) et donc la proposition. ■

On veut montrer que tous les nuages d'intensité $\ell\otimes\mu$ avec $\mu$ finie se représentent comme dans la proposition III.5.1. Pour cela, on introduit la notation suivante

> *$R$ est l'ensemble des nuages $\pi$ sur $\mathbb{R}_+\times E$ de la forme $\{(t_n,x_n); n\in\mathbb{N}^*\}$ où la suite $(t_n)_{n\geq 1}$ croît strictement vers $\infty$.*

Le résultat suivant est un résultat technique de mesurabilité dont la preuve peut être passée à première lecture.

**Lemme III.5.2** *$R$ est un ensemble de $\mathscr{S}_{\mathbb{R}_+\times E}$. De plus, pour tout $n\in\mathbb{N}^*$, il existe des fonctions $\mathscr{E}_n:\mathbf{S}_{\mathbb{R}_+\times E}\to\, ]0,\infty[$ qui sont $\mathscr{S}_{\mathbb{R}_+\times E}$-mesurables et des fonctions $X_n:\mathbf{S}_{\mathbb{R}_+\times E}\to E$ qui sont $(\mathscr{S}_{\mathbb{R}_+\times E},\mathscr{E})$-mesurables, et qui satisfont les propriétés suivantes.*

- *Si $T_n := \mathscr{E}_1+\ldots+\mathscr{E}_n$, $n\in\mathbb{N}^*$, alors pour tout $\pi\in\mathbf{S}_{\mathbb{R}_+\times E}$, $\lim_{n\to\infty}T_n(\pi)=\infty$.*
- *Pour tout $\pi\in R$, on a $\pi=\big\{(T_n(\pi),X_n(\pi))\,;\,n\in\mathbb{N}^*\big\}$.*

**Preuve :** on pose $R_0 := N^{-1}_{\mathbb{R}_+\times E}(\{\infty\})\cap\bigcap_{p\geq 1}N^{-1}_{[0,p]\times E}(\mathbb{N})$, qui est dans la tribu $\mathscr{S}_{\mathbb{R}_+\times E}$. Soit $f:E\to\mathbb{R}_+$, une fonction $\mathscr{E}$-mesurable bornée. On pose

$$\forall (t,\pi)\in\mathbb{R}_+\times\mathbf{S}_{\mathbb{R}_+\times E}\,,\quad S_f(t,\pi)=\sum_{(s,x)\in\pi}(1+f(x))\mathbf{1}_{[0,t]}(s)\ \in[0,\infty]\,.$$

Pour tout $t\in\mathbb{R}_+$, on pose $f_t(s,x)=(1+f(x))\mathbf{1}_{[0,t]}(s)$. Clairement, $f_t:\mathbb{R}_+\times E\to\mathbb{R}_+$ est $\mathscr{B}(\mathbb{R}_+)\otimes\mathscr{E}$-mesurable et on voit que $S_f(t,\pi)=N_{f_t}(\pi)$. Cela montre que pour $t$ fixé, $\pi\mapsto S_f(t,\pi)$ est mesurable. Par ailleurs, on remarque que à $\pi$ fixé, $t\mapsto S_f(t,\pi)$ est croissante. Si $S_f(t_0,\pi)<\infty$, alors $t\mapsto S_f(t,\pi)$ est càd sur $[0,t_0[$. Comme $f$ est bornée, pour tout nuage $\pi\in R_0$, $t\mapsto S_f(t,\pi)$ est donc croissante càd.

On pose $\mathscr{S}_{R_0}=\{C\cap R_0; C\in\mathscr{S}_{\mathbb{R}_+\times E}\}$, la tribu trace de $\mathscr{S}_{\mathbb{R}_+\times E}$ sur $R_0$. Comme $R_0$ est dans la tribu $\mathscr{S}_{\mathbb{R}_+\times E}$, $\mathscr{S}_{R_0}$ est une sous-tribu de $\mathscr{S}_{\mathbb{R}_+\times E}$. Montrons que la fonction

$$(t,\pi)\in\mathbb{R}_+\times R_0\longmapsto S_f(t,\pi)\in\mathbb{R}_+ \text{ est } \mathscr{B}(\mathbb{R}_+)\otimes\mathscr{S}_{R_0}\text{-mesurable.} \tag{III.48}$$

*En effet*, pour tout $t\in\mathbb{R}_+$ et tout $n\in\mathbb{N}$, on pose $(t)_n := 2^{-n}\lceil 2^n t\rceil$, qui tend en décroissant vers $t$ lorsque $n$ tend vers l'infini. On remarque que $S_f((t)_n,\pi)=\sum_{k\geq 1}\mathbf{1}_{[(k-1)2^{-n},k2^{-n}[}(t)S_f(k2^{-n},\pi)$. Ce qui précède implique que $(t,\pi) \mapsto \mathbf{1}_{[(k-1)2^{-n},k2^{-n}[}(t) S_f(k2^{-n},\pi)$ est $\mathscr{B}(\mathbb{R}_+)\otimes\mathscr{S}_{R_0}$-mesurable. Il en est donc de même pour $(t,\pi)\mapsto S_f((t)_n,\pi)$. Or $\lim_n S_f((t)_n,\pi)=S_f(t,\pi)$, sur $\mathbb{R}_+\times R_0$, ce qui permet de conclure.

Si $f$ est nulle, alors $S_0(t,\pi) = N_{[0,t]\times E}(\pi)$. La fonction $(t,\pi) \in \mathbb{R}_+ \times R_0 \to N_{[0,t]\times E}(\pi) \in \mathbb{N}$ est donc $\mathscr{B}(\mathbb{R}_+)\otimes\mathscr{S}_{R_0}$-mesurable. On remarque qu'à $\pi$ fixé, les sauts de $t \mapsto N_{[0,t]\times E}(\pi)$ peuvent être strictement supérieurs à 1, si par exemple $(t,x)$ et $(t,y)$ appartiennent à $\pi$ avec $x$ et $y$ distincts. On fixe $\pi\in R_0$ et pour tout $n\in\mathbb{N}$, on pose $T_0(\pi):=0$ et $T_{n+1}(\pi):=\inf\{t\in\mathbb{R}_+ : S_0(t,\pi)>S_0(T_n(\pi),\pi)\}$. Par définition de $R_0$, la suite $T_n(\pi)$ croît strictement vers $\infty$ et on a $S_0(T_n(\pi),\pi)\geq n$. Supposons que $T_n : R_0\to\mathbb{R}_+$ soit $\mathscr{S}_{R_0}$-mesurable ; alors il en est de même pour $\pi \in R_0 \mapsto S_0(T_n(\pi),\pi)$, par le résultat de mesurabilité conjointe (III.48). On constate ensuite que pour tout $t\in\mathbb{R}_+$, $T_{n+1}^{-1}(\,]t,\infty[\,) = \big\{\pi \in R_0 : S_0(T_n(\pi),\pi) = S_0(t,\pi)\big\}$, qui est bien un ensemble de $\mathscr{S}_{R_0}$. Cela montre que $T_{n+1}$ est $\mathscr{S}_{R_0}$-mesurable.

Par récurrence, on a donc montré que les $T_n$, $n \in \mathbb{N}^*$, sont des fonctions $\mathscr{S}_{R_0}$-mesurables. Il en est donc de même pour les fonctions $\pi\mapsto S_0(T_n(\pi),\pi)$. On remarque alors que $R$ est le sous-ensemble des $\pi\in R_0$ où $t\mapsto N_{[0,t]\times E}(\pi)$ ne progresse que de 1 en 1, c'est-à-dire $R=\bigcap_{n\in\mathbb{N}}\{\pi\in R_0 : S_0(T_{n+1}(\pi),\pi)=1+S_0(T_n(\pi),\pi)\}$, qui est donc un ensemble dans la tribu $\mathscr{S}_{R_0}$, donc dans $\mathscr{S}_{\mathbb{R}_+\times E}$.

On note $\mathscr{S}_R$ la tribu trace de $\mathscr{S}_{\mathbb{R}\times E}$ sur $R$. Si $\pi\in R$, alors il existe une suite $X_n(\pi)\in E$, $n\in\mathbb{N}^*$, telle que $\pi = \{(T_n(\pi),X_n(\pi)); n\in\mathbb{N}^*\}$. On remarque ensuite que $f(X_n(\pi)) = S_f(T_n(\pi),\pi) - S_f(T_{n-1}(\pi),\pi) - 1$ et les résultats de mesurabilités conjointes qui précèdent impliquent que pour tout $n\in\mathbb{N}^*$, la fonction $\pi\in R\mapsto f(X_n(\pi)) \in \mathbb{R}_+$ est $\mathscr{S}_R$-mesurable, pour toute fonction Borélienne $f$. Par conséquent, $\pi \in R \mapsto X_n(\pi)$ est $(\mathscr{S}_R,\mathscr{E})$-mesurable.

On fixe $x_0 \in E$, qui ne joue aucun rôle spécifique. On étend $X_n$ de façon mesurable à $\mathbf{S}_{\mathbb{R}_+\times E}$ en posant $X_n(\pi) := x_0$ si $\pi \in \mathbf{S}_{\mathbb{R}_+\times E}\backslash R$ : la fonction obtenue $X_n : \mathbf{S}_{\mathbb{R}_+\times E} \to E$ est donc $(\mathscr{S}_{\mathbb{R}_+\times E},\mathscr{E})$-mesurable. On pose ensuite $\mathscr{E}_n(\pi) := 1$ si $\pi \in \mathbf{S}_{\mathbb{R}_+\times E}\backslash R$ et $\mathscr{E}_n(\pi) := T_n(\pi) - T_{n-1}(\pi)$ si $\pi\in R$. On vérifie immédiatement que $\mathscr{E}_n : \mathbf{S}_{\mathbb{R}_+\times E} \to ]0,\infty[$ est $\mathscr{S}_{\mathbb{R}_+\times E}$-mesurable et que les suites de fonctions $\mathscr{E}_n$, $X_n$, $n\in\mathbb{N}$, satisfont bien les propriétés désirées. ∎

**Proposition III.5.3** *Soit $(E,\mathscr{E})$, un espace séparable et séparé. Soit $\mathbf{\Pi} : \Omega \to \mathbf{S}_{\mathbb{R}_+\times E}$ un nuage Poissonnien d'intensité $\ell\otimes\mu$. La mesure $\mu$ est supposée de masse finie. On rappelle de la proposition III.5.2, les définitions de $X_n$ et $\mathscr{E}_n$. Alors, on a les propriétés suivantes.*

*(i)* $\mathbf{P}(\mathbf{\Pi} \in R)=1$.

*(ii) Les deux suites $(\mathscr{E}_n(\mathbf{\Pi}))_{n\geq1}$ et $(X_n(\mathbf{\Pi}))_{n\geq1}$ sont indépendantes.*

*(iii) Les variables $(\mathscr{E}_n(\mathbf{\Pi}))_{n\geq1}$ sont i.i.d. de loi exponentielle de paramètre $\mu(E)$.*

*(iv) Les variables $(X_n(\mathbf{\Pi}))_{n\geq1}$ sont i.i.d. de loi $\mu(\,\cdot\,)/\mu(E)$.*

*(v) Si on pose $T_n(\mathbf{\Pi}) = \mathscr{E}_1(\mathbf{\Pi}) + \ldots + \mathscr{E}_n(\mathbf{\Pi})$, $n \in \mathbb{N}^*$, on a donc*

$$\mathbf{P}\text{-}p.s. \quad \mathbf{\Pi} = \big\{ \big(T_n(\mathbf{\Pi}), X_n(\mathbf{\Pi})\big) \,;\, n \in \mathbb{N}^* \big\} \,.$$

**Preuve :** soit $(\Omega',\mathscr{F}',\mathbf{P}')$, un espace de probabilité sur lequel sont définies deux suites indépendantes de variables, $(\mathscr{E}'_n)_{n\geq1}$ et $(X'_n)_{n\geq1}$, où les variables $(\mathscr{E}'_n)_{n\geq1}$ sont i.i.d. de loi exponentielle de paramètre $\mu(E)$ et les variables $(X'_n)_{n\geq1}$ sont i.i.d. de loi $\mu(\,\cdot\,)/\mu(E)$. On pose $T'_n = \mathscr{E}'_n + \ldots + \mathscr{E}'_n$ et $\mathbf{\Pi}' = \{(T'_n,X'_n); n\in\mathbb{N}^*\}$. La proposition III.5.1 montre que $\mathbf{\Pi}'$ est un nuage Poissonnien sur $\mathbb{R}_+\times E$ d'intensité $\ell\otimes\mu$. Donc $\mathbf{\Pi}'$ a même loi que $\mathbf{\Pi}$. Comme il est clair que $\mathbf{P}'(\mathbf{\Pi}' \in R) = 1$, on a $\mathbf{P}(\mathbf{\Pi} \in R) = 1$. Le lemme III.5.2 implique alors que la suite $(\mathscr{E}_n(\mathbf{\Pi}'),X_n(\mathbf{\Pi}'))_{n\geq1}$ a même loi que la suite $(\mathscr{E}_n(\mathbf{\Pi}),X_n(\mathbf{\Pi}))_{n\geq1}$. Par ailleurs, il est clair que $\mathbf{P}'$-p.s. pour tout $n\in\mathbb{N}^*$, on a $\mathscr{E}'_n = \mathscr{E}_n(\mathbf{\Pi}')$ et $X'_n = X_n(\mathbf{\Pi}')$, ce qui permet de conclure. ∎

On déduit de ce qui précède et du théorème de restriction le résultat d'extraction suivant qui permet de calculer la loi d'un nuage jusqu'au premier temps d'atteint d'un ensemble.

**Proposition III.5.4** **(Extraction)** *Soit $(E, \mathscr{E})$, un espace séparable et séparé. Soit $\mathbf{\Pi} : \Omega \to \mathbf{S}_{\mathbb{R}_+ \times E}$ un nuage Poissonnien d'intensité $\ell \otimes \mu$. La mesure $\mu$ est supposée sigma-finie. Soit $A \in \mathscr{E}$ tel que $0 < \mu(A) < \infty$. Alors, les assertions suivantes sont vérifiées.*

$(i)$ *il existe des suites de v.a. $(T_n, X_n)_{n\geq 1}$ qui satisfont les propriétés suivantes :*

$$\mathbf{P}\text{-}p.s. \quad \mathbf{\Pi} \cap (\mathbb{R}_+ \times A) = \big\{ (T, X) \in \mathbf{\Pi} : X \in A \big\} = \big\{ (T_n, X_n); n \geq 1 \big\},$$

*$(T_n)_{n\geq 1}$ et $(X_n)_{n\geq 1}$ sont indépendantes; les v.a. $(X_n)_{n\geq 1}$ sont i.i.d. de loi $\mu(\cdot \cap A)/\mu(A)$; on pose $T_0 = 0$; alors les v.a. $(T_n - T_{n-1}))_{n\geq 1}$ sont i.i.d. exponentielle de paramètre $\mu(A)$.*

$(ii)$ *On pose $\mathbf{\Pi}' = \mathbf{\Pi} \cap (\mathbb{R}_+ \times (E \backslash A))$. Alors $T_1$, $X_1$ et $\mathbf{\Pi}'$ sont indépendants et*

$$\mathbf{\Pi} \cap ([0, T_1] \times E) = \{(T_1, X_1)\} \cup \big( \mathbf{\Pi}' \cap [0, T_1] \times E \big),$$

*ce qui implique que pour toute fonction $F : \mathbf{S}_{\mathbb{R}_+ \times E} \to \mathbb{R}$ qui est $\mathscr{S}_{\mathbb{R}_+ \times E}$-mesurable bornée on a*

$$\mathbf{E}\big[ F\big( \mathbf{\Pi} \cap ([0, T_1] \times E) \big) \big] = \int_0^\infty dt\, e^{-\mu(A)t} \int_A \mu(dx)\, \mathbf{E}\big[ F\big( \{t, x\} \cup (\, \mathbf{\Pi} \cap ([0, t] \times (E \backslash A)\,) \big) \big].$$

**Preuve.** Par le principe de restriction (théorème III.3.6, page 244), $\mathbf{\Pi}_A := \mathbf{\Pi} \cap (\mathbb{R}_+ \times A)$ est un nuage Poissonnien d'intensité $\ell \otimes \mu(\cdot \cap A)$. Comme $\mu(\cdot \cap A)$ est de masse $\mu(A)$ finie, la proposition III.5.3 s'applique et on a $T_n = \mathscr{E}_1(\mathbf{\Pi}_A) + \ldots + \mathscr{E}_n(\mathbf{\Pi}_A)$ et $X_n = X_n(\mathbf{\Pi}_A)$, $n \geq 1$, ce qui prouve $(i)$. Le point $(ii)$ est une conséquence du $(i)$ et du même principe de restriction (théorème III.3.6, page 244) qui montre que $\mathbf{\Pi}'$ et $\mathbf{\Pi} \cap (\mathbb{R}_+ \times A)$ sont des nuages Poissonniens indépendants. ■

En appliquant le principe de restriction et les résultats précédents, on a immédiatement le résultat suivant.

**Théorème III.5.5** *Soit $(E, \mathscr{E})$, un espace séparable et séparé. Soit $\mathbf{\Pi} : \Omega \to \mathbf{S}_{\mathbb{R}_+ \times E}$ un nuage Poissonnien d'intensité $\ell \otimes \mu$. La mesure $\mu$ est supposée sigma-finie. Soit $B_p \in \mathscr{E}$, $p \in \mathbb{N}^*$, une partition de $E$ telle que $0 < \mu(B_p) < \infty$. On rappelle de la proposition III.5.2, les définitions de $X_n$ et $\mathscr{E}_n$ et on pose pour tous $n, p \geq 1$*

$$\mathscr{E}_n^{B_p} = \mathscr{E}_n(\mathbf{\Pi} \cap B_p) \quad et \quad X_n^{B_p} = X_n(\mathbf{\Pi} \cap B_p)\ .$$

*Alors, les assertions suivantes sont vraies*

(i) *Les variables $\mathscr{E}_n^{B_p}$, $X_n^{B_p}$, $n, p \geq 1$, sont indépendantes.*

(ii) *Pour tout $p \geq 1$, les variables $(\mathscr{E}_n^{B_p})_{n\geq 1}$ suivent une loi exponentielle de paramètre $\mu(B_p)$.*

(iii) *Pour tout $p \geq 1$, les variables $(X_n^{B_p})_{n\geq 1}$ ont pour loi $\mu(\,\cdot\, \cap B_p)/\mu(B_p)$.*

(iv) *Pour tous $n, p \geq 1$, on pose $T_n^{B_p} := \mathscr{E}_1^{B_p} + \ldots + \mathscr{E}_n^{B_p}$. Alors $\mathbf{P}$-p.s. pour tout $p \geq 1$, $\mathbf{\Pi} \cap B_p = \big\{ \big( T_n^{B_p}, X_n^{B_p} \big)\, ;\, n \in \mathbb{N}^* \big\}$, et donc*

$$\mathbf{P}\text{-}p.s. \qquad \mathbf{\Pi} = \big\{ \big( T_n^{B_p}, X_n^{B_p} \big);\ n, p \geq 1 \big\}. \tag{III.49}$$

Ce théorème donne une construction, et une méthode de simulation, des nuages Poissonniens d'intensité $\ell \otimes \mu$ lorsque $\mu$ est sigma-finie. On remarque que (III.49) implique que sous les hypothèses du théorème III.5.5 qui précède,

$$\mathbf{P}\text{-p.s.} \quad \forall t \in \mathbb{R}_+ , \quad N_{\{t\}\times E}(\mathbf{\Pi}) \in \{0,1\} . \tag{III.50}$$

Cela implique le résultat suivant.

**Lemme III.5.6** *Soit* $(E, \mathscr{E})$*, un espace séparé et séparable. Soit* $\mathbf{\Pi} : \Omega \to \mathbf{S}_{\mathbb{R}_+\times E}$*, un nuage Poissonnien d'intensité* $\ell\otimes\mu$*, où* $\mu$ *est sigma finie et de masse infinie. On pose*

$$J(\mathbf{\Pi}) := \big\{ T \in \mathbb{R}_+ : \exists X \in E \text{ tel que } (T,X) \in \mathbf{\Pi} \big\}.$$

*Alors, p.s.* $J(\mathbf{\Pi})$ *est dense dans* $\mathbb{R}_+$.

**Preuve :** pour tous réels positifs $a<b$, on a $(\ell \otimes \mu)([a,b] \times E) = \infty$. Cela montre que p.s. $N_{[a,b]\times E}(\mathbf{\Pi}) = \infty$. Or (III.50) implique que $\mathbf{P}$-p.s. $\mathbf{P}$-p.s. pour tous réels positifs $a < b$, $\#\big(J(\mathbf{\Pi}) \cap [a,b]\big) = N_{[a,b]\times E}(\mathbf{\Pi})$. Donc, p.s. pour tous rationnels positifs $a < b$, $\#\big(J(\mathbf{\Pi}) \cap [a,b]\big) = \infty$, ce qui implique le résultat voulu. ■

Il sera utile d'établir la loi des grands nombres suivante.

**Proposition III.5.7** *Soit* $(E, \mathscr{E})$*, un espace séparé et séparable. Soit* $\mathbf{\Pi} : \Omega \to \mathbf{S}_{\mathbb{R}_+\times E}$*, un nuage Poissonnien d'intensité* $\ell\otimes\mu$*, où* $\mu$ *est sigma finie. Soient* $E_k \in \mathscr{E}_k$*,* $k \in \mathbb{N}$*, tels que*

$$E_k \subset E_{k+1} , \quad E = \bigcup_{k\in\mathbb{N}} E_k , \quad \mu(E_k) < \infty \quad \textit{et} \quad \lim_{k\to\infty} \mu(E_k) = \infty.$$

*Alors,*

$$\mathbf{P}\text{-}p.s. \;\; \forall t \in \mathbb{R}_+, \quad \sup_{s\in[0,t]} \Big| \tfrac{1}{\mu(E_k)} N_{[0,s]\times E_k}(\mathbf{\Pi}) - s \Big| \xrightarrow[k\to\infty]{} 0 . \tag{III.51}$$

**Preuve.** On fixe un réel $t>0$. Soit $(N'_s)_{s\in\mathbb{R}_+}$ un processus de Poisson linéaire homogène d'intensité $t$, que l'on suppose défini sur $(\Omega, \mathscr{F}, \mathbf{P})$ pour simplifier. Les v.a. $(N'_{n+1} - N'_n)_{n\in\mathbb{N}}$ sont donc i.i.d. de loi de Poisson de paramètre $t$. Comme $N'_n = \sum_{0\leq k<n} N'_{k+1} - N'_k$, la loi des grands nombres implique p.s. que $\lim_{n\to\infty} N'_n/n = t$.Comme pour tout réel $s \geq 1$, on a $(s/\lfloor s\rfloor)\lfloor s\rfloor^{-1} N'_{\lfloor s\rfloor} \leq N'_s/s \leq (s/\lceil s\rceil)\lceil s\rceil^{-1} N'_{\lceil s\rceil}$ et donc $\lim_{s\to\infty} N'_s/s = t$.

On observe ensuite que la suite $(N_{[0,t]\times E_k}(\mathbf{\Pi}))_{k\in\mathbb{N}}$ a la même loi que $(N'_{\mu(E_k)})_{k\in\mathbb{N}}$ et on en déduit donc que p.s. $\lim_{k\to\infty} N_{[0,t]\times E_k}(\mathbf{\Pi})/\mu(E_k) = t$. On pose ensuite

$$\Omega_0 = \bigcap_{q\in\mathbb{Q}\cap\mathbb{R}_+} \Big\{ \lim_{k\to\infty} N_{[0,q]\times E_k}(\mathbf{\Pi})/\mu(E_k) = q \Big\} .$$

Clairement $\Omega_0 \in \mathscr{F}$ et on vient de montrer que $\mathbf{P}(\Omega_0) = 1$. On utilise ensuite le lemme déterministe suivant dont la preuve est donnée en fin de section.

**Lemme III.5.8** **(Second lemme de Dini)** *Pour tout* $k \in \mathbb{N}$*, soit* $f_k : \mathbb{R}_+ \to \mathbb{R}_+$*, une fonction croissante (pas nécessairement continue). Soit* $D \subset \mathbb{R}_+$*, un ensemble infini dénombrable. On fait les hypothèses suivantes.*

- *(a) Pour tout* $t \in D$*,* $f(t) := \lim_{k\to\infty} f_k(t)$ *existe dans* $\mathbb{R}_+$.
- *(b) L'ensemble* $\{f(t)\,;\, t \in D\}$ *est dense dans* $\mathbb{R}_+$.

*Alors, $f$ se prolonge de manière unique à $\mathbb{R}_+$ et ce prolongement noté $f : \mathbb{R}_+ \to \mathbb{R}_+$ est continu, croissant, surjectif et tel que*

$$\forall t \in \mathbb{R}_+, \quad \sup_{s\in[0,t]} |f_k(s) - f(s)| \xrightarrow[\ell\to\infty]{} 0 \; .$$

En effet, on conclut en appliquant sur $\Omega_0$ ce lemme à $D = \mathbb{Q} \cap \mathbb{R}_+$, et les fonctions $f_k(t) = N_{[0,t]\times E_k}(\mathbf{\Pi})$ et $f(t) = t$, $t \in \mathbb{R}_+$. ∎

**Preuve du lemme III.5.8.** pour tout $t \in D$, on pose $g(t) := \lim_{k\to\infty} f_k(t)$ et $S := \{g(t) \, ; \, t \in D\}$, supposé dense dans $\mathbb{R}_+$. La fonction $g : D \to S$ est donc surjective et croissante. Pour tout $s \in \mathbb{R}_+$, on pose ensuite $f(s) := \inf\{g(t) \, ; \, t \in D \cap [s, \infty[\,\}$. La fonction $f : \mathbb{R}_+ \to \mathbb{R}_+$ est croissante, elle étend $g$ et donc $S \subset f(\mathbb{R}_+)$, ce qui implique que $f(\mathbb{R}_+)$ soit dense dans $\mathbb{R}_+$. Elle est donc continue. *En effet,* s'il existe $t_0$ tel que $f(t_0) < f(t_0+)$ (resp. $f(t_0-) < f(t_0)$) alors $]f(t_0), f(t_0+)[ \subset \mathbb{R}_+ \backslash f(\mathbb{R}_+)$ (resp. $]f(t_0-), f(t_0)[ \subset \mathbb{R}_+ \backslash f(\mathbb{R}_+)$) puisque $f$ est croissante, ce qui contredit la densité de $f(\mathbb{R}_+)$. Cela montre donc que $f$ est continue croissante surjective de $\mathbb{R}_+$ sur $\mathbb{R}_+$.

Il est clair que $f(0) = 0$. Comme pour tout $t \in \mathbb{R}_+$, on a $0 \leq f_k(0) \leq f_k(t) \to f(t)$, pour tout $t \in D$, on a $\limsup_k f_k(0) \leq f(t)$ On a donc $\limsup_k f_k(0) \leq \inf f(D) = 0$, ce qui implique que $\lim_k f_k(0) = f(0) = 0$. Dans perte de généralité on peut donc supposer que $0 \in D$.

On fixe ensuite un réel $t > 0$. Soit $\varepsilon > 0$. Comme l'adhérence de $\{f(s) \, ; \, s \in D \cap [0,t]\}$ est $f([0,t]) = [0, f(t)]$, il existe une subdivision de $[0,t]$, $s_0 = 0 < s_1 < \ldots < s_{n+1} = t$ telle que pour tout $i \in \{0, \ldots, n\}$, on ait $s_i \in D$ et $0 \leq f(s_i) - f(s_{i-1}) \leq \varepsilon$. Il existe ensuite $k_\varepsilon \geq 1$, tel que pour tout $k \geq k_\varepsilon$, et pour tout $i \in \{0, \ldots, n\}$, on ait $|f_k(s_i) - f(s_i)| \leq \varepsilon$. Soit $s \in [0,t]$. Il existe $i \in \{1, \ldots, n\}$ tel que $s \in [s_{i-1}, s_i]$ et pour tout $k \geq k_\varepsilon$, on a

$$\begin{aligned}
|f_k(s) - f(s)| &\leq |f_k(s) - f_k(s_i)| + |f_k(s_i) - f(s_i)| + |f(s_i) - f(s)| \\
&\leq f_k(s_i) - f_k(s) + \varepsilon + f(s_i) - f(s) \\
&\leq f_k(s_i) - f_k(s_{i-1}) + \varepsilon + f(s_i) - f(s_{i-1}) \\
&\leq f_k(s_i) - f(s_i) + f(s_{i-1}) - f_k(s_{i-1}) + f(s_i) - f(s_{i-1}) + \varepsilon + f(s_i) - f(s_{i-1}) \\
&\leq |f_k(s_i) - f(s_i)| + |f(s_{i-1}) - f_k(s_{i-1})| + 3\varepsilon \;\leq\; 5\varepsilon \; .
\end{aligned}$$

On a donc montré que pour tout $t > 0$ tel que $t \in D$, et pour tout $\varepsilon > 0$, il existe $k_\varepsilon \geq 1$ tel que pour tout $k \geq k_\varepsilon$, on a $\sup_{s\in[0,t]} |f_k(s) - f(s)| \leq 5\varepsilon$. Cela prouve que pour tout $t \in D$, $\lim_k \sup_{s\in[0,t]} |f_k(s) - f(s)| = 0$, ce qui implique le résultat voulu car $D$ est non-borné. ∎

## EXERCICES.

**Exercice III.3** On munit $\mathbb{R}^d$ de sa base canonique et de la norme Euclidienne associée, notée $\|\cdot\|$. On note $\ell_d$ la mesure de Lebesgue sur $\mathbb{R}^d$. On note $v_d$ le volume de la boule unité. Si $B(x,r)$ dénote la boule ouverte de centre $x$ et de rayon $r$, on a donc $\ell_d(B(x,r)) = v_d r^d$. On rappelle la formule de changement de variable radial : si $f : \mathbb{R}_+ \to \mathbb{R}_+$, est mesurable alors

$$\int_{\mathbb{R}^d} f(\|x\|) \, \ell_d(dx) = d v_d \int_{\mathbb{R}_+} f(r) r^{d-1} \, \ell_1(dr) \; .$$

Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité sur lequel est défini un nuage Poissonnien $\mathbf{\Pi} : \Omega \to \mathbf{S}_{\mathbb{R}^d}$ d'intensité $\ell_d$. On pose

$$\mathbf{\Pi}_0 = \big\{ \|X\| \, ; \, X \in \mathbf{\Pi} \big\} \; .$$

1. Montrer que $\mathbf{\Pi}_0$ est presque sûrement égal à un nuage Poissonnien sur $\mathbb{R}_+$ dont on précisera l'intensité.

2. On note $\phi(r) = br^\beta$. Pour quelles valeurs de $b$ et de $\beta$, $\phi(\mathbf{\Pi}_0)$ est un nuage Poissonnien sur $\mathbb{R}_+$ d'intensité $\ell_1$ ?
3. Montrer qu'il existe une suite $(X_n)_{n\geq 1}$ de variables $\mathscr{F}$-mesurables telles que $0 < \|X_n\| < \|X_{n+1}\|$, pour tout $n \in \mathbb{N}^*$ et telles que $\mathbf{\Pi} = \{X_n; n \in \mathbb{N}^*\}$ presque sûrement. Trouver la fonction de répartition de $\|X_n\|$.
4. Montrer qu'il existe des constantes strictement positives $c_d$ et $C_d$ (que l'on calculera) telles que
$$\mathbf{P}\text{-p.s.} \quad \lim_{n\to\infty} n^{-c_d}\|X_n\| = C_d \; .$$
5. Trouver la densité de $\|X_n\|$.
6. On note $\mu$ la loi uniforme sur la sphère unité $S(0,1) = \{x \in \mathbb{R}^d : \|x\| = 1\}$. On rappelle que si $Z : \Omega \to B(0,1)$ a pour loi $(v_d)^{-1}\ell_d(\,\cdot\, \cap B(0,1))$, alors $(\|Z\|)^{-1}Z$ a pour loi $\mu$. Montrer qu'il existe une suite $(\mathscr{E}_n)_{n\geq 1}$, de variable i.i.d. de loi exponentielles de paramètre 1, et une suite $(Z_n)_{n\geq 1}$, de variable i.i.d. de loi $\mu$, telles que $(\mathscr{E}_n)_{n\geq 1}$ et $(Z_n)_{n\geq 1}$ sont indépendantes et
$$\mathbf{P}\text{-p.s.} \quad \mathbf{\Pi} = \left\{a_d(\mathscr{E}_1 + \ldots + \mathscr{E}_n)^{\gamma_d} \cdot Z_n \; ; \; n \in \mathbb{N}^*\right\} ,$$
où on précisera les constantes $a_d$ et $\gamma_d$. □

**Exercice III.4** On note $\mathbb{R}^* = \mathbb{R}\backslash\{0\}$, $\mathbb{R}_+$ l'ensemble des réels positifs, $\mathbb{R}_+^* = \mathbb{R}_+ \cap \mathbb{R}^*$, l'ensemble des réels strictement positifs. Pour tout $\beta \in \mathbb{R}_+^*$, on définit la loi $\mu_\beta$ sur $\mathbb{R}_+^*$ muni des Boréliens, par $\mu_\beta(dt) = \mathbf{1}_{\mathbb{R}_+^*}(t)t^{-\beta}\ell(dt)$, où $\ell$ désigne la mesure de Lebesgue sur $\mathbb{R}$. On rappelle la définition de la fonction $\Gamma$ sur $\mathbb{R}_+^*$ par l'intégrale
$$\Gamma(u) = \int_0^\infty t^{u-1}e^{-t}\ell(dt) \; .$$
Pour tout nuage déterministe $\pi \in \mathbf{S}_{\mathbb{R}_+^*}$, tout $c \in \mathbb{R}_+^*$ et tout $a \in \mathbb{R}^*$, on pose
$$c \cdot \pi = \{cx; x \in \pi\} \quad \text{et} \quad (\pi)^a = \{x^a; x \in \pi\} \; .$$
Toutes les variables sont définies sur un même espace de probabilité $(\Omega, \mathscr{F}, \mathbf{P})$. Pour tout $\beta \in \mathbb{R}_+^*$, on se donne $\mathbf{\Pi}_\beta : \Omega \to \mathbf{S}_{\mathbb{R}_+^*}$, un nuage Poissonnien d'intensité $\mu_\beta$. On pose
$$S_\beta = \sum_{X \in \mathbf{\Pi}_\beta} X \; ,$$
qui est une variable $\mathscr{F}$-mesurable à valeurs dans $[0, \infty]$.

1. Soient $c \in \mathbb{R}_+^*$ et tout $a \in \mathbb{R}^*$. Montrer que $\pi \in \mathbf{S}_{\mathbb{R}_+^*} \mapsto c \cdot \pi \in \mathbf{S}_{\mathbb{R}_+^*}$ et $\pi \in \mathbf{S}_{\mathbb{R}_+^*} \mapsto (\pi)^a \in \mathbf{S}_{\mathbb{R}_+^*}$ sont $(\mathscr{S}_{\mathbb{R}_+^*}, \mathscr{S}_{\mathbb{R}_+^*})$-mesurables.
2. Soient $\beta_1, \beta_2 \in \mathbb{R}_+^*$. Montrer qu'il existe $c \in \mathbb{R}_+^*$ et $a \in \mathbb{R}^*$, qui dépendent de $\beta_1$ et $\beta_2$ tels que
$$\mathbf{\Pi}_{\beta_1} \overset{\text{(loi)}}{=} c \cdot \left(\mathbf{\Pi}_{\beta_2}\right)^a \; .$$
3. Pour quels $\beta \in \mathbb{R}^*$ a-t-on $\mathbf{P}(S_\beta < \infty) = 1$ ? Que vaut $\mathbf{P}(S_\beta < \infty)$ si $\mathbf{P}(S_\beta < \infty) < 1$ ?
4. On suppose que $\mathbf{P}(S_\beta < \infty) = 1$. Pour tout $\lambda \in \mathbb{R}_+$, calculer $\mathbf{E}[\exp(-\lambda S_\beta)]$ en fonction de $\lambda$, $\beta$ et $\Gamma(2-\beta)$.
5. On fixe $\beta \in \mathbb{R}_+^*$ et pour tout $a \in \mathbb{R}^*$, on pose $S_{a,\beta} = \sum_{X \in \mathbf{\Pi}_\beta} X^a$. Pour quels $a$ a-t-on $\mathbf{P}(S_{a,\beta} < \infty) = 1$ ?
6. On fixe $\beta \in \mathbb{R}_+^*$. Trouver $(c, a) \in \mathbb{R}_*^+ \times \mathbb{R}^*$ tel que $c \cdot (\mathbf{\Pi}_\beta)^a$ soit un nuage Poissonnien d'intensité $\ell$.
7. On fixe $\beta \in \mathbb{R}_+^*$. Montrer qu'il existe une suite de variables $(X_n)_{n\geq 1}$ telles que $\mathbf{P}$-presque sûrement,
$$\forall n \in \mathbb{N}^* , \quad X_n > X_{n+1} , \quad \lim_{n\to\infty} X_n = 0 \quad \text{et} \quad \mathbf{\Pi}_\beta = \left\{X_n \; ; \; n \in \mathbb{N}^*\right\}.$$
8. On fixe $\beta \in \mathbb{R}_+^*$ et on reprend les notations de la question précédente. Trouver la fonction de répartition de $X_n$, sa densité et montrer qu'il existe deux constantes strictement positives $c$ et $C$, que l'on précisera, telles que
$$\mathbf{P}\text{-p.s.} \quad \lim_{n\to\infty} n^c X_n = C \; .$$

**Exercice III.5** Soient $\theta_1, \theta_2 \in \mathbb{R}_+^*$. Soit $\mathbf{\Pi}$ un nuage Poissonnien sur $\mathbb{R}_+ \times [0, \theta_1 + \theta_2]$ d'intensité $\mathbf{1}_{\mathbb{R}_+}(dt)\ell(dt) \otimes \mathbf{1}_{[0,\theta_1+\theta_2]}(x)\ell(dx)$.

1. En utilisant le cours, montrer qu'il existe deux suites indépendantes de variable i.i.d. $(\mathscr{E}_n)_{n\geq 1}$ et $(X_n)_{n\geq 1}$ telles que
   — les $\mathscr{E}_n$ suivent une loi exponentielle de paramètre $\theta_1 + \theta_2$,
   — les $X_n$ sont uniformes dans $[0, \theta_1 + \theta_2]$,

— si on pose $T_n := \mathscr{E}_1 + \ldots + \mathscr{E}_n$, alors $\mathbf{\Pi} = \{(T_n, X_n); n \geq 1\}$ presque sûrement.

2. On pose $\mathbf{\Pi}_1 = \{(T_n, X_n) : n \in \mathbb{N}^* \text{ tels que } X_n \in [0, \theta_1]\}$ et $\mathbf{\Pi}_2 = \{(T_n, X_n) : n \in \mathbb{N}^* \text{ tels que } X_n \in ]\theta_1, \theta_1 + \theta_2]\}$. Montrer que ce sont deux nuages de Poisson indépendants dont on précisera les intensités.
3. On pose $S = \inf\{n \geq 1 : X_n \in ]\theta_1, \theta_1 + \theta_2]\}$. Quelle est la loi de $S$ ? Quelle est la loi de $T_S$ ? Conditionnellement à $T_S$, quelle est la loi de $\#\mathbf{\Pi} \cap [0, T_S[$ ? (*Indication : aucun calcul n'est nécessaire pour résoudre cette question.*)
4. Soit $G : \Omega \to \mathbb{N}^*$, une variable géométrique de paramètre $p \in ]0, 1[$, c'est-à-dire $\mathbf{P}(G = n) = p(1-p)^n$, $n \in \mathbb{N}^*$. Soit $(\mathscr{E}')_{n \geq 1}$ une suite i.i.d. de variables indépendante de $G$ et de loi exponentielles de paramètre $\theta$. On pose
$$\mathscr{E} = \sum_{1 \leq n \leq G} \mathscr{E}'_n \ .$$
Quelle est la loi de $\mathscr{E}$ ? Conditionnellement à $\mathscr{E}$, quelle est la loi de $G - 1$ ? (*Indication : aucun calcul n'est nécessaire pour résoudre cette question.*)

**Exercice III.6** Soit $\mathbf{\Pi}$ un nuage Poissonnien sur $\mathbb{R}$ d'intensité la mesure de Lebesgue. On imagine un glouton partant de l'origine 0 qui mange les points de $\mathbf{\Pi}$ de proche en proche : il mange d'abord le point de $\mathbf{\Pi}$ le plus proche de 0, puis le plus proche parmi les points qui restent ... et ainsi de suite. On se demande s'il mange ainsi tous les points de $\mathbf{\Pi}$.

Formellement, on note $Y_n$, $n \geq 0$ la suite des points mangés par le glouton : on a $Y_0 = 0$, $Y_1$ est le point de $\mathbf{\Pi}$ le plus proche de 0 et si $n \geq 1$, $Y_{n+1}$ est le point de $\mathbf{\Pi} \backslash \{Y_1, \ldots, Y_n\}$ le plus proche de $Y_n$.

Au temps $n$, le glouton se trouve au bord du "trou" qu'il a déjà exploré : c'est un intervalle $I_n$ contenant 0 et dont les extrémités sont deux points de $\mathbf{\Pi}$. Si le glouton est dans $\mathbb{R}_+$, alors le trou est sur sa gauche et à sa droite, il reste une partie inexplorée de $\mathbf{\Pi}$. De même si le glouton est dans $\mathbb{R}_-$, il est à gauche de 0 et à sa gauche, il reste une partie inexplorée de $\mathbf{\Pi}$. On dit qu'il effectue une traversée, si à l'instant $n$ il est à gauche du trou (resp. à droite) et qu'à l'instant $n+1$ il se retrouve à droite (resp. à gauche) du trou.

1. Montrer qu'il existe une suite d'exponentielles $(\mathscr{E}_n)_{n \geq 0}$ indépendantes de paramètre 1 telles que $\ell(I_n) = \mathscr{E}_0 + \mathscr{E}_1 + \ldots + \mathscr{E}_n$.
2. Montrer que le glouton effectue une traversée entre $n$ et $n+1$ ssi $\mathscr{E}_{n+1} > \ell(I_n)$.
3. Montrer que $\mathbf{P}$-p.s. le glouton ne visite pas tous les points de $\mathbf{\Pi}$ et qu'il en laisse même une infinité de côté.

**Exercice III.7** Soit $(\Omega, \mathscr{F}, \mathbf{P})$ un espace de probabilité complet.

1. Soit $p$ un entier plus grand que 2. Soient $X_1, \ldots, X_p : \Omega \to \mathbb{N}$ des v.a. *indépendantes* suivant des lois de Poisson non-dégénérées. On note $\mathbf{E}[X_k] = \theta_k$, $1 \leq k \leq p$. Soit $n \in \mathbb{N}$. En utilisant le lemma III.1.3, calculer de deux façons $\mathbf{P}(X_1 + \ldots + X_p = n)$ et montrer la *formule du multinône* :
$$(\theta_1 + \ldots + \theta_p)^n = \sum_{\substack{k_1, \ldots, k_p \in \mathbb{N} \\ k_1 + \ldots + k_p = n}} \frac{n!}{k_1! \ldots k_p!} \theta_1^{k_1} \ldots \theta_p^{k_p} \ .$$
Soient $q_1, \ldots, q_p \in [0, 1]$ tels que $q_1 + \ldots + q_p = 1$. On dit qu'un vecteur aléatoire $\mathscr{F}$-mesurable $\mathbf{N} = (N_1, \ldots, N_p) : \Omega \to \mathbb{N}^p$ suit une *loi multinômiale* de paramètres $n$ et $\mathbf{q} = (q_1, \ldots, q_p)$ si pour tous $k_1, \ldots, k_p \in \mathbb{N}$,
$$\mathbf{P}\big(\mathbf{N} = (k_1, \ldots, k_p)\big) = \begin{cases} \dfrac{n!}{k_1! \ldots k_p!} q_1^{k_1} \ldots q_p^{k_p} & \text{si } k_1 + \ldots + k_p = n, \\ 0 & \text{sinon.} \end{cases}$$
Vérifier que c'est bien une loi de probabilité.
2. Soient $Z_n : \Omega \to \{1, \ldots, p\}$, $n \in \mathbb{N}^*$, une suite de variables $\mathscr{F}$-mesurables i.i.d. Pour tout $k \in \{1, \ldots, p\}$, on pose $N_k = \#\{j \in \{1, \ldots, n\} : Z_j = k\}$. Montrer que $\mathbf{N} = (N_1, \ldots, N_p)$ est un vecteur aléatoire $\mathscr{F}$-mesurable qui suit une loi multinômiale dont on précisera les paramètres.
3. Soit $(E, \mathscr{E}, \mu)$ un espace mesuré. On suppose $(E, \mathscr{E})$ séparable et séparé. Soit $B \in \mathscr{E}$ tel que $0 < \mu(B) < \infty$. Soient $X_n : \Omega \to E$, $n \in \mathbb{N}^*$, une suite de variables aléatoires i.i.d. de loi $\mu(\cdot \cap B)/\mu(B)$. Pour tout $n \in \mathbb{N}^*$, on pose $\mathbf{\Pi}_n = \{X_1, X_2, \ldots, X_n\}$. On suppose $\mu$ diffuse.
   (a) Soient $A_1, \ldots, A_p \in \mathscr{E}$ deux-à-deux disjoints. Quelle est la loi de $\mathbf{N} = (N_{A_1}(\mathbf{\Pi}_n), \ldots, N_{A_p}(\mathbf{\Pi}_n))$ ? Est-ce que ce résultat est toujours vrai si $\mu$ n'est pas diffuse ?
   (b) Soit $\mathbf{\Pi} : \Omega \to \mathbf{S}_E$, un nuage de Poisson $(\mathscr{F}, \mathscr{S}_E)$-mesurable d'intensité $\mu$. Soit $n \in \mathbb{N}^*$. Montrer que $\mathbf{\Pi} \cap B$ sous $\mathbf{P}(\,\cdot \mid N_B(\mathbf{\Pi}) = n)$ a même loi que $\mathbf{\Pi}_n$.

**Exercice III.8** Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité complet. Soit $(E, \mathscr{E}, \mu)$ un espace mesuré. On suppose $(E, \mathscr{E})$ séparable et séparé et on suppose qu'il existe une partition de $E$ en ensembles mesurables $B_p \in \mathscr{E}$, $p \in \mathbb{N}$ tels que $\mu(B_p) < \infty$, $p \in \mathbb{N}$. $\mathbf{\Pi} : \Omega \to \mathbf{S}_E$, un nuage de Poisson $(\mathscr{F}, \mathscr{S}_E)$-mesurable d'intensité $\mu$.

On cherche à donner un sens à *une indexation mesurable de* $\mathbf{\Pi}$, c'est-à-dire à l'existence d'une suite de v.a. $X_n : \Omega \to E$, $n \in \mathbb{N}^*$, $\mathscr{F}$-mesurables telles que $\mathbf{\Pi} = \{X_n \, ; \; 1 \leq n \leq N_E(\mathbf{\Pi})\}$ (avec la convention que $\mathbf{\Pi} = \{X_n \, ; \; n \in \mathbb{N}^*\}$ si $N_E(\mathbf{\Pi}) = \infty$). On rappelle que pour tous $(n,p) \in \mathbb{N}^* \times \mathbb{N}$, il existe des fonctions $Y_n^{(p)} : \mathbf{S}_E \to E$ qui sont $(\mathscr{S}_E, \mathscr{E})$-mesurables et telles que

$$\mathbf{P}\text{-p.s.} \quad \forall p \in \mathbb{N}, \quad \mathbf{\Pi} \cap B_p = \left\{ Y_n^{(p)}(\mathbf{\Pi}); 1 \leq n \leq N_{B_p}(\mathbf{\Pi}) \right\} .$$

Soit $x^* \in E$. On pose $S_0 = 0$ et $S_{p+1} = N_{B_0}(\mathbf{\Pi}) + \ldots + N_{B_p}(\mathbf{\Pi})$, $p \in \mathbb{N}$. Pour tout $n \in \,]S_p, S_{p+1}] \cap \mathbb{N}^*$, on pose $X_n = Y^{(p)}_{n-S_p}(\mathbf{\Pi})$ et si $N_E(\mathbf{\Pi}) < \infty$, on pose $X_n = x^*$ pour tout $n > N_E(\mathbf{\Pi})$.

Montrer que pour tout $n \in \mathbb{N}^*$, $X_n : \Omega \to E$ est $(\mathscr{F}, \mathscr{E})$-mesurable. Montrer que $\mathbf{P}$-p.s. $\mathbf{\Pi} = \{X_n \, ; \; 1 \leq n \leq N_E(\mathbf{\Pi})\}$. Les $X_n$ sont-elles nécessairement indépendantes ? Ont-elles la même loi ?

**Exercice III.9** Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité complet. Soit $(E, \mathscr{E}, \mu)$ un espace mesuré. On suppose $(E, \mathscr{E})$ séparable et séparé et $\mu$ sigma-finie et diffuse. Soit $\mathbf{\Pi} : \Omega \to \mathbf{S}_E$, un nuage de Poisson $(\mathscr{F}, \mathscr{S}_E)$-mesurable d'intensité $\mu$. On se donne une indexation mesurable de $\mathbf{\Pi}$, c'est-à-dire que $\mathbf{\Pi} = \{X_n \, ; \; 1 \leq n \leq N_E(\mathbf{\Pi})\}$ où $X_n : \Omega \to E$, $n \in \mathbb{N}^*$, est une suite de v.a. $(\mathscr{F}, \mathscr{E})$-mesurables (avec les conventions évidentes si $N_E(\mathbf{\Pi}) = \infty$ ou 0).

1. Soit $(E', \mathscr{E}', \nu)$ un espace de de probabilité. On suppose $(E', \mathscr{E}')$ séparable et séparé. Soient $Y_n : \Omega \to E'$, $n \in \mathbb{N}$ une suite i.i.d. de v.a. $(\mathscr{F}, \mathscr{E}')$-mesurables de loi $\nu$. On pose

$$\mathbf{\Pi}' = \left\{ (X_n, Y_n) \, ; \; 1 \leq n \leq N_E(\mathbf{\Pi}) \right\}$$

(avec les conventions évidentes si $N_E(\mathbf{\Pi}) = \infty$ ou 0). On munit $E \times E'$ de la tribu produit. Montrer que $\mathbf{\Pi}' : \Omega \to \mathbf{S}_{E \times E'}$ est un nuage Poissonnien $(\mathscr{F}, \mathscr{S}_{E \times E'})$-mesurable dont on précisera l'intensité.

2. (*Un théorème de Dobrushin*). On suppose ici que $(E, \mathscr{E}) = (E', \mathscr{E}') = (\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ en conservant les mêmes hypothèses que précédemment. On pose

$$\mathbf{\Pi}_* = \left\{ X_n + Y_n \, ; \; 1 \leq n \leq N_E(\mathbf{\Pi}) \right\}$$

(avec les conventions évidentes si $N_E(\mathbf{\Pi}) = \infty$ ou 0). Montrer que $\mathbf{\Pi}_* : \Omega \to \mathbf{S}_{\mathbb{R}^d}$ est un nuage Poissonnien $(\mathscr{F}, \mathscr{S}_{\mathbb{R}^d})$-mesurable dont on précisera l'intensité en fonction de $\mu$ et $\nu$.

3. On suppose ici que $(E, \mathscr{E}) = (\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ en conservant les mêmes hypothèses que précédemment. On suppose que $Y : \Omega \to \mathbb{R}^d$ est un vecteur aléatoire $\mathscr{F}$-mesurable indépendant de $\mathbf{\Pi}$ et on pose

$$\mathbf{\Pi}_0 = \left\{ X_n + Y \, ; \; 1 \leq n \leq N_E(\mathbf{\Pi}) \right\}$$

(avec les conventions évidentes si $N_E(\mathbf{\Pi}) = \infty$ ou 0). Montrer que $\mathbf{\Pi}_* : \Omega \to \mathbf{S}_{\mathbb{R}^d}$ est un nuage $(\mathscr{F}, \mathscr{S}_{\mathbb{R}^d})$-mesurable. Est-il de Poisson ? (*La réponse dépend de* $\mu$.)

4. On suppose que $(E', \mathscr{E}')$ est régulier selon la définition **??** (voir page **??**). Soit $\mathrm{m} : \mathscr{E} \otimes \mathscr{E}' \to \mathbb{R}_+$, une mesure *finie* dont la première marginale est $\mu$, c'est-à-dire $\mu(A) = \mathrm{m}(A \times E')$ pour tout $A \in \mathscr{E}$. On rappelle le théorème **??** de *représentation des couplages* (voir page **??**) qui affirme qu'il existe $\Psi : E \times \,]0,1[ \to E'$, qui est $(\mathscr{E} \otimes \mathscr{B}(\,]0,1[\,), \mathscr{E}')$-mesurable et telle que

$$\int_{E \times E'} g \, d\mathrm{m} = \int_{E \times \,]0,1[} g\big(x, \Psi(x,z)\big) \, \mu(dx) \ell(dz)$$

où $\ell$ désigne la mesure de Lebesgue sur la droite réelle. On se donne une suite $U_n : \Omega \to ]0,1[$, $n \in \mathbb{N}^*$, de v.a. indépendantes uniformes sur $]0,1[$. On pose

$$\mathbf{\Pi}_\bullet = \left\{ \big(X_n, \Psi(X_n, U_n)\big) \, ; \; 1 \leq n \leq N_E(\mathbf{\Pi}) \right\}$$

(avec les conventions évidentes si $N_E(\mathbf{\Pi}) = \infty$ ou 0). Montrer que $\mathbf{\Pi}_\bullet : \Omega \to \mathbf{S}_{E \times E'}$ est un nuage Poissonnien $(\mathscr{F}, \mathscr{S}_{E \times E'})$-mesurable dont on précisera l'intensité.

**Exercice III.10** Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité complet. Soit $(E, \mathscr{E}, \mu)$ un espace mesuré avec $\mu$ diffuse. Soit $g : E \to \mathbb{R}_+$, une fonction $\mathscr{E}$-mesurable telle que $\int_E \min(1, g(x)) \, \mu(dx) < \infty$. Soit $\mathbf{\Pi} : \Omega \to \mathbf{S}_E$, un nuage de Poisson $(\mathscr{F}, \mathscr{S}_E)$-mesurable d'intensité $\mu$.

1. Montrer que $\int_E (1 - e^{-g(x)}) \mu(dx) < \infty$. Si on suppose que $g$ est de plus à valeurs dans $]0, \infty[$, montrer que cela implique que $\mu$ est sigma finie.

2. Pour tout $\omega \in \Omega$, on pose

$$Z_g(\omega) = \begin{cases} e^{-N_g(\mathbf{\Pi}(\omega))} e^{\int_E (1-\exp(-g(x)))\mu(dx)} & \text{si } N_g(\mathbf{\Pi}(\omega)) < \infty, \\ 1 & \text{si } N_g(\mathbf{\Pi}(\omega)) = \infty. \end{cases}$$

Montrer que $Z_g : \Omega \to ]0, \infty[$ est $\mathscr{F}$-mesurable. On note $\mathbf{P}_g : \mathscr{F} \to [0, \infty]$, la mesure ayant $Z_g$ pour densité par rapport à $\mathbf{P}$, c'est-à-dire que

$$\forall A \in \mathscr{F}, \quad \mathbf{P}_g(A) = \mathbf{E}\big[Z_g \mathbf{1}_A\big] .$$

Montrer que $(\Omega, \mathscr{F}, \mathbf{P}_g)$ est un espace de probabilité complet. On notera $\mathbf{E}_g$ l'espérance par rapport à $\mathbf{P}_g$ et on rappelle que $\mathbf{E}_g[X] = \mathbf{E}[Z_g X]$ pour toute variable $\mathscr{F}$-mesurable $X : \Omega \to \mathbb{R}$ bornée.

3. Montrer que $\mathbf{\Pi}$ sous $\mathbf{P}_g$ est un nuage de Poisson dont on précisera l'intensité en fonction de $\mu$ et $g$.

4. On suppose que $g : E \to \mathbb{R}$ est $\mathscr{F}$-mesurable. On note $g_+ = \max(g, 0)$ la partie positive de $g$ et $g_- = \max(-g, 0)$ la partie négative de $g$. On suppose que $\int_E \min(1, g_-(x))\, \mu(dx) < \infty$ et $\int_E (e^{g_+(x)} - 1)\, \mu(dx) < \infty$.

   (a) Montrer que l'on peut définir une probabilité $\mathbf{P}_g : \mathscr{F} \to [0, 1]$ telle que

   $$\forall A \in \mathscr{F}, \quad \mathbf{P}_g(A) = \mathbf{E}\big[e^{N_g(\mathbf{\Pi})} \mathbf{1}_A\big] e^{-\int_E \left(e^{g(x)} - 1\right) \mu(dx)} .$$

   (b) Montrer que sous $\mathbf{P}_g$, $\mathbf{\Pi}$ est un nuage de Poisson dont on préciséra l'intensité.

**Exercice III.11** (*Caractérisation des nuages Poissonniens par la formule de Palm (ou de Mecke)*) Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité complet. Soit $(E, \mathscr{E}, \mu)$ un espace mesuré avec $\mu$ diffuse et sigma-finie. On suppose $(E, \mathscr{E})$ séparable et séparé. Soit $\mathbf{\Pi} : \Omega \to \mathbf{S}_E$, un nuage $(\mathscr{F}, \mathscr{S}_E)$-mesurable d'intensité $\mu$. Soit $F : E \times \mathbf{S}_E \to [0, \infty]$, une fonction $\mathscr{E} \otimes \mathscr{S}_E$-mesurable. On pose $\Phi_F(\mathbf{\Pi}) = \sum_{X \in \mathbf{\Pi}} F(X, \mathbf{\Pi} \backslash \{X\})$.

1. Montrer à l'aide du lemme III.2.20 $(iv)$ (page 240) que $\Phi_F : \Omega \to [0, \infty]$ est $\mathscr{F}$-mesurable.

2. On suppose que $\mathbf{\Pi}$ satisfait la formule de Palm (ou de Mecke), c'est-à-dire que pour toute fonction $F : E \times \mathbf{S}_E \to [0, \infty]$ qui est $\mathscr{E} \otimes \mathscr{S}_E$-mesurable, on a

$$\mathbf{E}\big[\Phi_F(\mathbf{\Pi})\big] = \int_E \mu(dx)\, \mathbf{E}\big[F(x, \mathbf{\Pi})\big] . \tag{III.52}$$

   (a) Soit $f : E \to \mathbb{R}_+$, une fonction $\mathscr{E}$-mesurable telle que telle que $\int_E f(x)\, \mu(dx) < \infty$. Montrer que $\mathbf{P}$-p.s. $N_f(\mathbf{\Pi}) < \infty$ et montrer que pour tout $\lambda \in \mathbb{R}_+$, on a $L_f(\lambda) = \mathbf{E}[\exp(-\lambda N_f(\mathbf{\Pi}))] > 0$. Montrer que $L_f$ est $C^1$ sur $\mathbb{R}_+$ et que $L'_f(\lambda) = -\mathbf{E}[N_f(\mathbf{\Pi}) \exp(-\lambda N_f(\mathbf{\Pi}))]$.

   (b) À l'aide de la formule de Palm (ou de Mecke) (III.52), montrer que

   $$\forall \lambda \in \mathbb{R}_+, \quad L'_f(\lambda) = -L_f(\lambda) \int_E f(x) e^{-\lambda f(x)} \mu(dx) .$$

   (c) Montrer que $\mathbf{\Pi}$ est un nuage Poissonnien (*on résoudra l'équation différentielle précédente puis on étendra le résultat obtenu à toutes les fonctions $f : E \to [0, \infty]$ $\mathscr{E}$-mesurables*).

**Exercice III.12** (*Caractérisation des nuages Poissonniens par évitement indépendants.*) Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité complet. Soit $(E, \mathscr{E})$ un espace mesurable supposé séparable et séparé. Soit $\mathbf{\Pi} : \Omega \to \mathbf{S}_E$, un nuage $(\mathscr{F}, \mathscr{S}_E)$-mesurable d'intensité $\mu$. On suppose que $\mathbf{\Pi}$ satisfait la propriété suivante.

$(*)$ Pour tous $A, B \in \mathscr{E}$ disjoints, les événements $\{\mathbf{\Pi} \cap A = \emptyset\}$ et $\{\mathbf{\Pi} \cap B = \emptyset\}$ sont indépendants.

Le but de cet exercice est de montrer que si $\mathbf{\Pi}$ satisfait $(*)$ et si $\mu$ est sigma finie et diffuse, alors $\mathbf{\Pi}$ est un nuage de Poisson.

1. On suppose que $\mathbf{\Pi}$ satisfait $(*)$ (mais pas nécessairement que $\mu$ soit sigma finie et diffuse). Soient $B_n \in \mathscr{E}$, $n \in \mathbb{N}$, des ensembles deux-à-deux disjoints. Montrer que les événements $\{\mathbf{\Pi} \cap B_n = \emptyset\}$, $n \in \mathbb{N}$, sont indépendants (mutuellement et pas seulement deux-à-deux).

2. On suppose que $\mathbf{\Pi}$ satisfait $(*)$ (mais pas nécessairement que $\mu$ soit sigma finie et diffuse). Comme $(E, \mathscr{E})$ est séparable et séparé, il existe un sous-ensemble $C \subset [0, 1[$ et une bijection $\phi : E \to [0, 1[$ qui est $(\mathscr{E}, \mathscr{B}([0, 1[)$-mesurable et telle que la réciproque $\phi^{-1} : [0, 1[\to E$ est également $(\mathscr{B}([0, 1[), \mathscr{E})$-mesurable. Pour tout $n \in \mathbb{N}$ et tout $0 \leq k < 2^n$, on pose

$$J_{n,k} = \phi^{-1}([k2^{-n}, (k+1)2^{-n}[) .$$

On voit que pour tout $n \in \mathbb{N}$, les $(J_{n,k})_{0 \leq k < 2^n}$ forment une partition $\mathscr{E}$-mesurable de $E$. Soit $A \in \mathscr{E}$ tel que $\mu(A) < \infty$. On pose

$$\forall n \in \mathbb{N}, \quad Z_n = \sum_{0 \leq k < 2^n} \mathbf{1}_{\{\mathbf{\Pi} \cap (A \cap J_{n,k}) \neq \emptyset\}} .$$

Montrer que

$$\mathbf{P}\text{-p.s. pour tout } n \in \mathbb{N}, \quad Z_n \leq Z_{n+1} \quad \text{et} \quad N_A(\mathbf{\Pi}) = \lim_{n \to \infty} Z_n .$$

(*Indication : on considérera les deux cas :* $N_A(\mathbf{\Pi}) < \infty$ *et* $N_A(\mathbf{\Pi}) = \infty$).

3. Soient $X_q^{(n)}$ et $X_q : \Omega \to [0, \infty]$, $q, n \in \mathbb{N}$, des v.a. $\mathscr{F}$-mesurables. On suppose que
   - pour tout $q \in \mathbb{N}$, $\mathbf{P}$-p.s. $\lim_{n \to \infty} X_q^{(n)} = X_q$,
   - pour tout $n \in \mathbb{N}$ fixé, sous $\mathbf{P}$, les v.a. $(X_q^{(n)})_{q \in \mathbb{N}}$ sont mutuellement indépendantes sous $\mathbf{P}$.

   Montrer que les v.a. $(X_q)_{q \in \mathbb{N}}$ sont mutuellement indépendantes sous $\mathbf{P}$.
4. On suppose que $\mathbf{\Pi}$ satisfait $(*)$ (mais pas nécessairement que $\mu$ soit sigma finie et diffuse). Soient $A_1, \ldots, A_p \in \mathscr{E}$ des ensembles deux-à-deux disjoints. Montrer que pour $n \in \mathbb{N}$ fixé les variables aléatoires $\mathbf{1}_{\{\mathbf{\Pi} \cap (A_q \cap J_{n,k}) \neq \emptyset\}}$, $1 \leq q \leq p$, $0 \leq k < 2^n$, sont mutuellement indépendantes. En déduire que les v.a. $N_{A_1}(\mathbf{\Pi}), \ldots, N_{A_p}(\mathbf{\Pi})$ sont indépendantes.
5. On suppose que $\mathbf{\Pi}$ satisfait $(*)$ et que $\mu$ est sigma finie et diffuse. Conclure en utilisant le théorème III.3.3 (page 243).
6. On choisit ici $(E, \mathscr{E}) = (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ et on définit $\mathbf{\Pi} : \Omega \to \mathbf{S}_\mathbb{R}$ en supposant que $\{\mathbf{\Pi} = \emptyset\} \cup \{\mathbf{\Pi} = \{12\}\} = \Omega$, $\{\mathbf{\Pi} = \emptyset\} \in \mathscr{F}$ et $\mathbf{P}(\mathbf{\Pi} = \emptyset) = 2/3$. Montrer que $\mathbf{\Pi}$ est un nuage de points $(\mathscr{F}, \mathscr{S}_\mathbb{R})$-mesurable dont l'intensité (que l'oncalculera explicitement) est sigma finie et qui satisfait $(*)$. Montrer en revanche que $\mathbf{\Pi}$ n'est pas un nuage de Poisson.

**Exercice III.13** (*Théorème de Rényi*) Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité complet. Soit $(E, \mathscr{E})$, un espace séparable et séparé. Soit $\mathbf{\Pi} : \Omega \to \mathbf{S}_E$, un nuage de points $(\mathscr{F}, \mathscr{S}_E)$-mesurable. On suppose que

$$\forall A \in \mathscr{E}, \quad \mathbf{P}(\mathbf{\Pi} \cap A = \emptyset) = e^{-\mu(A)} ,$$

où $\mu : \mathscr{E} \to [0, \infty]$ est une mesure sigma finie et diffuse (avec la convention que $e^{-\infty} = 0$). Le but de cet exercice est de montrer que $\mathbf{\Pi}$ est alors un nuage de Poisson (c'est le *théorème de Rényi*). Pour cela on s'appuie sur l'exercice III.12 qui précède.

1. Montrer que $\mathbf{\Pi}$ satisfait la propriété $(*)$ de l'exercice III.12 qui précède. On note alors $\nu$ l'intensité de $\mathbf{\Pi}$. Par l'exercice III.12 qui précède, il suffit donc de montrer que $\nu$ est diffuse et sigma finie.
2. Soit $A \in \mathscr{E}$. On reprend les notations de la question 2 de l'exercice III.12. On pose $Z_n = \sum_{0 \leq k < 2^n} \mathbf{1}_{\{\mathbf{\Pi} \cap (A \cap J_{n,k}) \neq \emptyset\}}$. Montrer que $\mathbf{P}$-p.s. $Z_n \leq Z_{n+1}$ pour tout $n \in \mathbb{N}$ et que $N_A(\mathbf{\Pi}) = \lim_{n \to \infty} Z_n$.
3. On suppose que $\mu(A) < \infty$. On pose $\delta_n = \max_{0 \leq k < 2^n} \mu(A \cap J_{n,k})$. En s'inspirant de la preuve du théorème III.3.3, montrer que $\mu$ diffuse implique $\lim_{n \to \infty} \delta_n = 0$.
4. On suppose que $\mu(A) < \infty$. Montrer que $0 \leq x - 1 + e^{-x} \leq \frac{1}{2} x^2$ pour tout $x \in \mathbb{R}_+$. Montrer que $0 \leq \mu(A) - \mathbf{E}[Z_n] \leq \frac{1}{2} \mu(A) \delta_n$.
5. Montrer que si $\mu(A) < \infty$, alors $\nu(A) = \mu(A)$.
6. Conclure.

# Annexe A

# Théorie de la mesure et de l'intégration.

## A.1 Rappels sur la mesurabilité.

**Définition A.1.1** Soit $E$ un ensemble non-vide. On note la classe de tous les sous-ensembles de $E$ par $\mathscr{P}(E) := \{A\,;\, A \subset E\}$. On définit les classes d'ensembles suivantes.

(a) $\mathscr{E} \subset \mathscr{P}(E)$ est une *tribu* si elle vérifie les conditions suivantes.

- $E \in \mathscr{E}$ ;
- Si $A \in \mathscr{E}$, alors $E \backslash A \in \mathscr{E}$ ;
- Si $A_n \in \mathscr{E}$, $n \in \mathbb{N}$, alors $\bigcup_{n\in\mathbb{N}} A_n \in \mathscr{E}$.

(b) $\mathscr{P} \subset \mathscr{P}(E)$ est un *pi-système* s'il vérifie les conditions suivantes.

- $E \in \mathscr{P}$ ;
- Si $A, B \in \mathscr{P}$, alors $A \cap B \in \mathscr{P}$.

(c) $\mathcal{L} \subset \mathscr{P}(E)$ est une *classe monotone* si elle vérifie les conditions suivantes.

- $E \in \mathcal{L}$ ;
- Si $A, B \in \mathcal{L}$ sont tels que $A \subset B$, alors $B \backslash A \in \mathcal{L}$ ;
- Si $A_n \in \mathcal{L}$, $n \in \mathbb{N}$, sont tels que $A_n \subset A_{n+1}$, alors $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{L}$.

(d) $\mathscr{T} \subset \mathscr{P}(E)$ est une *topologie* si elle vérifie les conditions suivantes.

- $E \in \mathscr{T}$ et $\emptyset \in \mathscr{T}$.
- Si $U, V \in \mathscr{T}$, alors $U \cap V \in \mathscr{T}$.
- Soit $U_i \in \mathscr{T}$, $i \in I$, une famille quelconque d'ensembles de $\mathscr{T}$. Alors, $\bigcup_{i\in I} U_i \in \mathscr{T}$.

Les ensembles constituant une topologie sont appelés les *ouverts* de cette topologie et un ensemble dont le complémentaire est un ouvert est appelé *fermé*. □

**Exemple A.1.2** La topologie usuelle de $\mathbb{R}$ est décrite comme suit : un ouvert de $\mathbb{R}$ est une réunion dénombrable d'intervalles ouverts deux-à-deux disjoints. □

**Lemme A.1.3** *Soit $(\mathscr{E}_i, i \in I)$, une famille de tribus de $E$ (resp. de classes monotone, de topologies). Alors,*

$$\bigcap_{i\in I} \mathscr{E}_i = \{A \subset E \;:\; \forall i \in I\,,\; A \in \mathscr{E}_i\}$$

*est également une tribu (resp. classe monotone, tribu) sur $E$.*

**Preuve :** exercice élémentaire. ■

**Définition A.1.4** Soit $\mathscr{R} \subset \mathscr{P}(E)$, une classe de sous-ensembles de $E$. On pose

$$\sigma(\mathscr{R}) = \bigcap_{\substack{\mathscr{E}, \text{ tribu} \\ \mathscr{R}\subset\mathscr{E}}} \mathscr{E}\ , \qquad \lambda(\mathscr{R}) = \bigcap_{\substack{\mathcal{L}, \text{ classe monot.} \\ \mathscr{R}\subset\mathcal{L}}} \mathcal{L} \qquad \text{et} \qquad \tau(\mathscr{R}) = \bigcap_{\substack{\mathscr{T}, \text{ topologie} \\ \mathscr{R}\subset\mathscr{T}}} \mathscr{T}$$

Alors, $\sigma(\mathscr{R})$ (resp. $\lambda(\mathscr{R})$, $\tau(\mathscr{R})$) est une tribu (resp. classe monotone, topologie) : c'est la tribu (resp. classe monotone, topologie) engendrée par $\mathscr{R}$. C'est la plus petite tribu (resp. classe monotone, topologie) contenant $\mathscr{R}$. □

**Définition A.1.5** Soit $E$ muni d'une topologie $\mathscr{T}$. La classe des *boréliens* associée à la topologie la tribu $\sigma(\mathscr{T})$. Lorsqu'il n'y a pas ambiguïté sur la topologie considérée, on note souvent cette tribu $\mathscr{B}(E)$. □

**Remarque A.1.6** Si $E$ est un ensemble non-vide, un théorème bien connu de Cantor montre que $\mathrm{Card}\,(E) < \mathrm{Card}\,(\mathscr{P}(E))$, c'est-à-dire qu'il n'existe pas de surjection de $E$ sur $\mathscr{P}(E)$. On pourrait s'attendre à ce que $\mathscr{B}(\mathbb{R})$ et $\mathscr{P}(\mathbb{R})$ soient de même cardinal or ce n'est pas le cas : on peut en effet montrer, par des arguments de récurrence transfinie, que $\mathrm{Card}\,(\mathbb{R}) = \mathrm{Card}\,(\mathscr{B}(\mathbb{R}))$, c'est-à-dire qu'il existe une bijection de $\mathbb{R}$ sur $\mathscr{B}(\mathbb{R})$. D'un point de vue ensembliste, cela signifie qu'il y a peu de Boréliens. Cela montre aussi qu'il existe beaucoup de sous-ensembles de $\mathbb{R}$ qui ne sont pas des Boréliens. Il n'est cependant pas évident d'en exhiber par des méthodes naïves. □

**Définition A.1.7** (*Topologie relative, tribu trace*) Soit $E$ un ensemble non-vide muni d'une topologie $\mathscr{T}$ et d'une tribu $\mathscr{E}$. Soit $A \subset E$ un ensemble non-vide ($A$ n'est ni nécessairement dans $\mathscr{T}$ ni dans $\mathscr{E}$).

(a) On pose $\mathscr{T}_A = \{O \cap A\ ;\ O \in \mathscr{T}\}$. On remarque facilement que $\mathscr{T}_A$ est une topologie sur $A$ : on l'appelle la *topologie relative de $\mathscr{T}$ sur $A$.*

(b) On pose $\mathscr{E}_A := \big\{B \cap A\,;\ B \in \mathscr{E}\big\}$. On remarque facilement que $\mathscr{E}_A$ est une tribu sur $A$ : on l'appelle la *tribu trace de $\mathscr{E}$ sur $A$.*

Dans le cas des tribus Boréliennes, ces notions sont cohérentes, comme le démontre le lemme suivant.

**Lemme A.1.8** *Soit $(E, \mathscr{T})$ un espace topologique. Soit $A \subset E$. Alors, la tribu des Boréliens sur $A$ associée à la topologie relative de $\mathscr{T}$ sur $A$ est la tribu trace sur $A$ des Boréliens associée à $\mathscr{T}$, c'est-à-dire*

$$\sigma(\mathscr{T}_A) = \big\{B \cap A\,;\ B \in \sigma(\mathscr{T})\ \big\}\ .$$

**Preuve :** on pose $\mathscr{G} = \{A \cap B; B \in \sigma(\mathscr{T})\}$ qui est la tribu trace sur $A$ des Boréliens associés à $\mathscr{T}$. Elle contient $\mathscr{T}_A$, on a donc $\sigma(\mathscr{T}_A) \subset \mathscr{G}$.

On pose ensuite $\mathscr{H} = \{B \in \sigma(\mathscr{T}) : A \cap B \in \sigma(\mathscr{T}_A)\}$. Il est facile de montrer que $\mathscr{H}$ est une tribu sur $E$. Par ailleurs elle contient $\mathscr{T}$, ce qui entraîne que $\sigma(\mathscr{T}) \subset \mathscr{H}$. Or par définition $\mathscr{H} \subset \sigma(\mathscr{T})$, donc $\mathscr{H} = \sigma(\mathscr{T})$. Par conséquent $\mathscr{G} \subset \sigma(\mathscr{T}_A)$, ce qui termine la preuve. ■

Le lemme précédent permet de parler sans ambiguïté des Boréliens de $[0, 1[$ ou des Boréliens d'un ouvert de $\mathbb{R}^n$.

## Thèorème de la classe monotone.

**Lemme A.1.9** *Soit $E$ un ensemble non-vide et $\mathcal{L}$ une classe monotone sur $E$. Alors, $\mathcal{L}$ est une tribu ssi $\mathcal{L}$ est stable par intersection finie.*

**Preuve :** supposons que $\mathcal{L}$ soit une classe monotone stable par intersection finie. Par définition, $E \in \mathcal{L}$. Puisque $\mathcal{L}$ est stable par différence propre, $\mathcal{L}$ est stable par passage au complémentaire. Soient $A, B \in \mathcal{L}$. Puisque $A \cup B = E\backslash\left((E\backslash A) \cap (E\backslash B)\right)$ et puisque $\mathcal{L}$ est stable par intersection finie, on a $A \cup B \in \mathcal{L}$. Par conséquent, $\mathcal{L}$ est stable par union finie.

Soient $B_n \in \mathcal{L}$, $n \in \mathbb{N}$, une suite quelconque d'éléments de $\mathcal{L}$. On pose $A_n = \bigcup_{0\leq p\leq n} B_p$ : on a donc $A_n \in \mathcal{L}$, d'après le raisonnement précédent. On remarque que $A_n \subset A_{n+1}$. comme $\mathcal{L}$ est stable par union dénombrable croissante, $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{L}$. Or $\bigcup_{n\in\mathbb{N}} B_n = \bigcup_{n\in\mathbb{N}} A_n$. Cela montre que $\mathcal{L}$ est stable par union dénombrable : c'est donc une tribu. La réciproque est triviale. ■

**Théorème A.1.10 (Classe monotone)** *Soit $E$, un ensemble non-vide. Soit $\mathcal{L}$, une classe monotone sur $E$. Soit $\mathscr{P}$, un pi-système sur $E$. On suppose que $\mathscr{P} \subset \mathcal{L}$. Alors $\sigma(\mathscr{P}) \subset \mathcal{L}$.*

**Preuve :** pour démontrer le théorème, il suffit de démontrer que $\lambda(\mathscr{P})$ est une tribu : en effet, par définition de la classe monotone engendrée, on a $\mathscr{P} \subset \lambda(\mathscr{P}) \subset \mathcal{L}$ et si $\lambda(\mathscr{P})$ est une tribu, on obtient $\sigma(\mathscr{P}) \subset \lambda(\mathscr{P})$. D'après le lemme précédent, il suffit donc de montrer que $\lambda(\mathscr{P})$ est stable par intersection finie.

La preuve de cette dernière assertion se fait en plusieurs étapes : pour tout $A \subset E$, on pose tout d'abord $\mathcal{L}_A = \{B \subset E : \; A \cap B \in \lambda(\mathscr{P})\}$. On vérifie facilement l'implication suivante :

$$(A \in \lambda(\mathscr{P})) \quad \Longrightarrow \quad (\mathcal{L}_A \text{ est une classe monotone}) \; . \tag{A.1}$$

Comme $\mathscr{P}$ est un pi-système et comme $\mathscr{P} \subset \lambda(\mathscr{P})$ (par définition), pour tout $A \in \mathscr{P}$, on a $\mathscr{P} \subset \mathcal{L}_A$ ; on déduit de (A.1) que $\lambda(\mathscr{P}) \subset \mathcal{L}_A$ car $\lambda(\mathscr{P})$ est la plus petite classe monotone contenant $\mathscr{P}$. On a donc montré :

$$\forall A \in \mathscr{P} \;, \forall B \in \lambda(\mathscr{P}) \;, \quad A \cap B \in \lambda(\mathscr{P}) \; .$$

Par conséquent $\mathscr{P} \subset \mathcal{L}_B$ pour tout $B$ dans $\lambda(\mathscr{P})$ et donc $\lambda(\mathscr{P}) \subset \mathcal{L}_B$, pour tout $B \in \lambda(\mathscr{P})$. Autrement dit, pour tous $B, C \in \lambda(\mathscr{P})$, $B \cap C \in \lambda(\mathscr{P})$, ce qui montre que $\lambda(\mathscr{P})$ est stable par intersection finie et ce qui termine la preuve du théorème. ■

Nous supposons le lecteur familier avec la notion de mesure positive et nous rappelons le théorème suivant, conséquence du théorème de classe monotone.

**Théorème A.1.11 (Unicité de prolongement des mesures)** *Soit $(E, \mathscr{E})$, un espace mesurable. Soit $\mathscr{P}$, un pi-système engendrant $\mathscr{E}$. Soient $\mu_1$ et $\mu_2$, deux mesures sur $(E, \mathscr{E})$ qui coincident sur $\mathscr{P}$ :*

$$\forall A \in \mathscr{P}, \quad \mu_1(A) = \mu_2(A) \; .$$

*(i) Si $\mu_1(E) = \mu_2(E) < \infty$, alors $\mu_1 = \mu_2$.*

*(ii) S'il existe $E_n \in \mathscr{P}$, $n \in \mathbb{N}$, tels que $\mu_1(E_n) = \mu_2(E_n) < \infty$, $E_n \subset E_{n+1}$ et $E = \bigcup_{n \in \mathbb{N}} E_n$, alors $\mu_1 = \mu_2$.*

**Preuve :** on prouve d'abord $(i)$. On pose $\mathcal{L} = \{B \in \mathscr{E} : \mu_1(B) = \mu_2(B)\}$. Par hypothèse, on a $\mathscr{P} \subset \mathcal{L}$. Montrons ensuite que $\mathcal{L}$ est une classe monotone : on voit tout d'abord que $E \in \mathcal{L}$. Soient $B, C \in \mathcal{L}$ tels que $B \subset C$. Comme $\mu_1$ et $\mu_2$ sont supposées finies, on a

$$\mu_1(C \backslash B) = \mu_1(C) - \mu_1(B) = \mu_2(C) - \mu_2(B) = \mu_2(C \backslash B),$$

qui entraîne bien que $C \backslash B \in \mathcal{L}$. Soit $B_n \in \mathcal{L}$, $n \in \mathbb{N}$, une suite d'ensembles tels que $B_n \subset B_{n+1}$. Par les propriétés élémentaires de mesures positives,

$$\mu_1\Big(\bigcup_{n \in \mathbb{N}} B_n\Big) = \lim_{n \to \infty} \mu_1(B_n) = \lim_{n \to \infty} \mu_2(B_n) = \mu_2\Big(\bigcup_{n \in \mathbb{N}} B_n\Big),$$

ce qui montre que $\bigcup B_n \in \mathcal{L}$. La classe $\mathcal{L}$ est donc une classe monotone et le théorème A.1.10 de la classe monotone entraîne que $\mathscr{E} = \sigma(\mathscr{P}) \subset \mathcal{L} \subset \mathscr{E}$. Donc $\mathcal{L} = \mathscr{E}$, qui prouve $(i)$.

Montrons $(ii)$ : pour tout $i \in \{1, 2\}$, on pose $\mu_{i,n}(B) = \mu_i(B \cap E_n)$, $B \in \mathscr{E}$. Clairement, $\mu_{i,n}$ est une mesure. Comme $E_n \in \mathscr{P}$, $\mu_{1,n}$ et $\mu_{2,n}$ coincident sur $\mathscr{P}$ et par le point $(i)$ elles sont égales. Soit $B \in \mathscr{E}$. On a $B \cap E_n \subset B \cap E_{n+1}$, et que $\bigcup B \cap E_n = B$. Clairement on a

$$\mu_1(B) = \lim_{n \to \infty} \mu_1(B \cap E_n) = \lim_{n \to \infty} \mu_{1,n}(B) = \lim_{n \to \infty} \mu_{2,n}(B) = \lim_{n \to \infty} \mu_2(B \cap E_n) = \mu_2(B) \; ,$$

ce qui termine la preuve. ■

## Produit de tribu Boréliennes.

On rappelle la notion de tribu produit en général.

**Définition A.1.12** Soient $(E_1, \mathscr{E}_1)$ et $(E_2, \mathscr{E}_2)$ deux espaces mesurables.

(a) On pose $\mathscr{P} = \{A \times B \; ; \; A \in \mathscr{E}_1 \; , \; B \in \mathscr{E}_2 \; \}$, appelée classe des *rectangles* mesurables de $E_1 \times E_2$. C'est un pi-système car $\big(A \times B\big) \cap \big(A' \times B'\big) = \big(A \cap A'\big) \times \big(B \cap B'\big)$.

(b) On note ensuite $\mathscr{E}_1 \otimes \mathscr{E}_2$ la tribu engendrée par $\mathscr{P}$ sur $E_1 \times E_2$ : $\mathscr{E}_1 \otimes \mathscr{E}_2 := \sigma(\mathscr{P})$. La tribu $\mathscr{E}_1 \otimes \mathscr{E}_2$ est appelée *tribu produit* de $\mathscr{E}_1$ et $\mathscr{E}_2$. □

On note $\pi_1$ et $\pi_2$ les *projections canoniques* de $E_1 \times E_2$ sur respectivement $E_1$ et $E_2$ :

$$\forall (x,y) \in E_1 \times E_2 \ : \quad \pi_1(x,y) = x \quad \text{et} \quad \pi_2(x,y) = y \ .$$

Comme $\pi_1^{-1}(A) = A \times E_2 \in \mathscr{E}_1 \otimes \mathscr{E}_2$, la première projection canonique $\pi_1$ est $(\mathscr{E}_1 \otimes \mathscr{E}_2, \mathscr{E}_1)$-mesurable. De même, $\pi_2$ est $(\mathscr{E}_1 \otimes \mathscr{E}_2, \mathscr{E}_2)$-mesurable. Il est facile de montrer également que $\mathscr{E}_1 \otimes \mathscr{E}_2$ est la plus petite tribu sur $E_1 \times E_2$ rendant mesurables $\pi_1$ et $\pi_2$, c'est-à-dire, selon la définition A.1.21 page 274,

$$\mathscr{E}_1 \otimes \mathscr{E}_2 = \sigma(\pi_1, \pi_2)$$

Soit $(E,d)$, un espace métrique. On note $B_d(x,r) = \{ y \in E : d(x,y) < r \}$ la boule ouverte de centre $x \in E$ et de rayon $r \in \mathbb{R}_+$. On vérifie que la topologie $\mathscr{T}_d$ sur $E$ correspondant à la convergence des suites et à la continuité des fonctions pour la distance $d$, c'est-à-dire la topologie associée à la distance $d$, est la topologie engendrée par la classe des boules ouvertes : $\mathscr{T}_d = \tau\big(\{ B_d(x,r); x \in E, r \in \mathbb{R}_+\}\big)$. Il est facile de voir que

$$\mathscr{T}_d = \big\{ U \subset E : \forall x \in U \, , \exists r > 0 : B_d(x,r) \subset U \big\}. \tag{A.2}$$

Attention, en général, la tribu des Boréliens associée à $\mathscr{T}_d$ n'est pas nécessairement la tribu engendrée par les boules ouvertes de $(E,d)$ : cela vient du fait qu'une topologie est stable par union quelconque, alors qu'une tribu n'est stable que par union dénombrable. Cependant c'est le cas lorsque l'espace est séparable comme le montre le lemme suivant.

**Lemme A.1.13** *Soit $(E,d)$, un espace métrique séparable, c'est-à-dire admettant une suite $x_n \in E$, $n \in \mathbb{N}$, qui est dense. On note $\mathscr{T}_d$, la topologie métrique de $(E,d)$ et $\mathscr{B}(E) = \sigma(\mathscr{T}_d)$, la tribu Borélienne correspondante. On pose*

$$\mathscr{R} = \big\{ B_d(x_n, q) \, ; \, n \in \mathbb{N}, \, q \in \mathbb{Q}_+ \big\} \ .$$

*Alors, pour tout ouvert $U \in \mathscr{T}_d$, il existe une suite de boules $B_n \in \mathscr{R}$, $n \in \mathbb{N}$, telle que $U = \bigcup_{n \in \mathbb{N}} B_n$. Cela implique $\mathscr{T}_d = \tau(\mathscr{R})$ et $\mathscr{B}(E) = \sigma(\mathscr{R})$.*

**Preuve :** comme les boules ouvertes sont des ouverts, on a $\tau(\mathscr{R}) \subset \mathscr{T}_d$ et $\sigma(\mathscr{R}) \subset \mathscr{B}(E)$. Soit $U$ un ouvert. Pour tout $x \in U$, il existe $r_x > 0$ tel que $B_d(x, r_x) \subset U$. Comme la suite $(x_n)_{n \in \mathbb{N}}$ est dense, il existe $n(x) \in \mathbb{N}$ et $q_x \in \mathbb{Q}_+^*$ tel que $d(x, x_{n(x)}) < q_x \leq r_x/2$, ce qui implique que $x \in B_d(x_{n(x)}, q_x) \subset B_d(x, r_x) \subset U$. On en déduit que $U = \bigcup_{x \in U} B_d(x_{n(x)}, q_x)$. Comme $\{x_n \, ; \, n \in \mathbb{N}\}$ et $\mathbb{Q}$ sont dénombrables, cela montre que $U$ est une union dénombrable d'éléments de $\mathscr{R}$. Cela implique que $\mathscr{T}_d \subset \tau(\mathscr{R})$ et $\mathscr{B}(E) \subset \sigma(\mathscr{R})$, ce qui permet de conclure. ■

**Définition A.1.14** Soient $(E, \mathscr{T})$ et $(E', \mathscr{T}')$, deux espaces topologiques : la *topologie produit* sur $E \times E'$ est définie par $\mathscr{T}_{E \times E'} := \tau\big(\{ U \times U' \, ; \, U \in \mathscr{T}, U' \in \mathscr{T}' \}\big)$. □

**Proposition A.1.15** *Soient $(E,d)$ et $(E',d')$, deux espaces métriques dont les tribusBoréliennes sont notées resp. $\mathscr{B}(E)$ et $\mathscr{B}(E')$. On munit $E \times E'$ de la topologie produit $\mathscr{T}_{E \times E'}$ et on note $\mathscr{B}(E \times E')$ la tribu Borélienne correspondante. Pour tous $(x_1,y_1), (x_2,y_2) \in E \times E'$, on pose*

$$D((x_1,y_1),(x_2,y_2)) = \max(d(x_1,x_2), d'(y_1,y_2)) \ .$$

*Les assertions suivantes sont vérifiées.*

*(i) $D$ est une distance sur $E \times E'$, telle que $\mathscr{T}_{E \times E'} = \mathscr{T}_D$. Si de plus, $E$ et $E'$ sont séparables, alors $(E \times E', D)$ est séparable.*

*(ii) Si $E$ et $E'$ sont séparables, alors*

$$\mathscr{B}(E \times E') = \mathscr{B}(E) \otimes \mathscr{B}(E') \ .$$

**Preuve :** il est facile de vérifier que $D$ est une distance sur $E \times E'$ et que $B_D((x,y),r) = B_d(x,r) \times B_{d'}(y,r)$, qui est un produit d'ouverts. Cela implique que $\mathscr{T}_D \subset \mathscr{T}_{E \times E'}$. Soient $U \in \mathscr{T}_d$ et $U' \in \mathscr{T}_{d'}$. Pour tout $(x,y) \in U \times U'$, il existe $r, r' > 0$ tels que $B_d(x,r) \subset U$ et $B_{d'}(y,r') \subset U'$, donc

$$B_D((x,y), r \wedge r')) = B_d(x, r \wedge r') \times B_{d'}(y, r \wedge r') \subset U \times U' \ .$$

Par (A.2) appliqué à $D$, cela montre que $U \times U' \in \mathscr{T}_D$ et donc que $\mathscr{T}_{E \times E'} \subset \mathscr{T}_D$. On a donc montré que $\mathscr{T}_D = \mathscr{T}_{E \times E'}$.

Supposons $E$ et $E'$ séparables : il existe $x_n \in E$, $y_n \in E'$, $n \in \mathbb{N}$, des suites denses. Alors, on vérifie facilement que $(x_p, y_q) \in E \times E'$, $p, q \in \mathbb{N}$, est une famille dénombrable de points qui est dense pour $D$ : donc $(E \times E', D)$ est séparable. Cela montre $(i)$.

Montrons $(ii)$ : on pose $\mathscr{R} := \big\{ B_D((x_m, y_n), q) \, ; \, m, n \in \mathbb{N}, \, q \in \mathbb{Q}_+^* \big\}$. La forme des boules pour $D$ implique que $\mathscr{R} \subset \mathscr{B}(E) \otimes \mathscr{B}(E')$. Soit $V$, un ouvert pour la topologie produit sur $E \times E'$. Par $(i)$, $V \in \mathscr{T}_D$ et le lemme A.1.13 appliqué à $D$ implique qu'il existe une suite $B_n \in \mathscr{R}$, $n \in \mathbb{N}$, telle que $V = \bigcup_{n\in\mathbb{N}} B_n$. Cela entraîne que la topologie produit est incluse dans $\mathscr{B}(E)\otimes\mathscr{B}(E')$ et donc que $\mathscr{B}(E \times E') \subset \mathscr{B}(E) \otimes \mathscr{B}(E')$.

Montrons l'inclusion contraire. Pour cela posons, $\mathscr{E} = \{A \times E' \, ; \, A \in \mathscr{B}(E)\}$. Il est facile de vérifier que c'est une tribu sur $E\times E'$ et qu'elle est engendrée par $\{U\times E' \, ; \, U \in \mathscr{T}_d\}$, qui est un ensemble d'ouverts produits. Cela implique que $\mathscr{E} \subset \mathscr{B}(E\times E')$. En raisonnant de même en l'autre coordonnée, on montre que pour tout $A \in \mathscr{B}(E)$ et tout $B \in \mathscr{B}(E')$, on a $A\times E'$ et $E\times B \in \mathscr{B}(E\times E')$. Par conséquent, $(A \times E') \cap (E \times B) = A \times B \in \mathscr{B}(E \times E')$. Cela entraîne donc que $\mathscr{B}(E) \otimes \mathscr{B}(E') \subset \mathscr{B}(E \times E')$, ce qui permet de conclure. ■

Si $E$ et $E'$ sont deux $\mathbb{R}$-espaces vectoriels normés, de normes respectives $\|\cdot\|$ et $\|\cdot\|'$, pour tout $(x, y) \in E \times E'$ on pose $\|(x, y)\|'' = \max(\|x\|, \|y\|')$. On vérifie que $\|\cdot\|''$ est une norme sur le $\mathbb{R}$-espace vectoriel $E \times E'$ et que le produit de la $\|\cdot\|$-topologie avec la $\|\cdot\|'$-topologie est la $\|\cdot\|''$-topologie sur $E \times E'$.

Comme il n'existe qu'une seule topologie de norme sur les espaces de dimension finie, si $E$ et $E'$ sont des $\mathbb{R}$-espaces vectoriels de dimension finie, alors la topologie de norme sur $E \times E'$ est le produit des topologies de norme sur $E$ et $E'$. Ces faits combinés avec la proposition précédente entraînent le corollaire suivant.

**Corollaire A.1.16** *Pour tous $m, n \in \mathbb{N}$, $\mathscr{B}(\mathbb{R}^{m+n}) = \mathscr{B}(\mathbb{R}^n) \otimes \mathscr{B}(\mathbb{R}^n)$.*

## Approximation des fonctions mesurables par les fonctions étagées.

**Définition A.1.17** (*Fonctions étagées*) Soit $(E, \mathscr{E})$, un espace mesurable. Une fonction étagée est une combinaison linéaire de fonctions indicatrices d'ensembles dans $\mathscr{E}$. Autrement dit $s : E \to \mathbb{R}$ est étagée si $s = \sum_{1\leq k\leq n} c_k \mathbf{1}_{A_k}$, où $c_1, \ldots, c_n \in \mathbb{R}$ et $A_1, \ldots, A_n \in \mathscr{E}$. □

**Proposition A.1.18** *Soit $f : E \to [0, \infty]$, une fonction $\mathscr{E}$-mesurable. Il existe $s_n : E \to \mathbb{R}_+$, $n \in \mathbb{N}$, une suite de fonctions étagées $\mathscr{E}$-mesurables positives telles que*

$$\lim_{n\to\infty} s_n = f \text{ ponctuellement} \quad \text{et} \quad 0 \leq s_n \leq s_{n+1} \leq f \ , \ n \in \mathbb{N} \ . \tag{A.3}$$

**Preuve :** pour tout $n \in \mathbb{N}$, on pose $s_n = n\mathbf{1}_{\{f\geq n\}} + \sum_{0\leq k<n2^n} \mathbf{1}_{\{k2^{-n}\leq f<(k+1)2^{-n}\}}$. On vérifie que les $s_n$ satisfont bien les propriétés désirées. ■

Comme conséquence de la proposition précédente, nous démontrons le lemme suivant qui s'avère être d'un usage très pratique.

**Lemme A.1.19** *Soit $(E, \mathscr{E})$, un espace mesurable. Soit $f : E \to [0, \infty]$, une fonction $\mathscr{E}$-mesurable. Alors, il existe $B_n \in \mathscr{E}$ et $c_n \in \mathbb{R}_+$, $n \in \mathbb{N}$, tels que $f = \sum_{n\in\mathbb{N}} c_n \mathbf{1}_{B_n}$.*

**Preuve :** on considère la suite de fonctions étagées $s_n$, $n \in \mathbb{N}$, donnée comme précédemment. On pose $t_0 = s_0$ et $t_n = s_n - s_{n-1}$, pour tout $n \in \mathbb{N}^*$. Les fonctions $t_n$ sont étagées positives : elles sont donc combinaison linéaire à coefficients positifs de fonctions indicatrices d'ensembles mesurables. De plus on a $f = \sum_{n\in\mathbb{N}} t_n$, ce qui permet de conclure. ■

**Version fonctionnelle de la classe monotone.** Une première application de cette approximation est donnée par la version fonctionnelle du théorème de la classe monotone.

**Théorème A.1.20 (Classe monotone, version fonctionnelle)** *Soit $(E, \mathscr{E})$, un espace mesurable. Soit $\mathscr{P}$, un pi-système sur $E$ tel que $\sigma(\mathscr{P}) = \mathscr{E}$. Soit $\mathscr{H}$, un ensemble d'applications de $E$ dans $\mathbb{R}$. On suppose que l'ensemble de fonctions $\mathscr{H}$ satisfait les propriétés suivantes.*

*(a) $\mathscr{H}$ est un $\mathbb{R}$-espace vectoriel (il contient donc la fonction nulle).*

*(b) Pour tout $C \in \mathscr{P}$, on a $\mathbf{1}_C \in \mathscr{H}$.*

*(c) Si $f_n \in \mathscr{H}$, $n \in \mathbb{N}$, et $M \in \mathbb{R}_+$ sont tels que $0 \leq f_n \leq f_{n+1} \leq M$, alors, $\sup_{n\in\mathbb{N}} f_n \in \mathscr{H}$.*

*Alors, $\mathscr{H}$ contient toutes les fonctions réelles $\mathscr{E}$-mesurables bornées.*

**Preuve :** on pose $\mathcal{L} = \{B \subset E : \mathbf{1}_B \in \mathscr{H}\}$. Les propriétés $(a)$, $(b)$ et $(c)$ permettent facilement de vérifier que $\mathcal{L}$ est une classe monotone contenant $\mathscr{P}$. Le théorème de la classe monotone A.1.10 (page 270) implique que $\mathscr{E} = \sigma(\mathscr{P}) \subset \mathcal{L}$. Donc pour tout $B \in \mathscr{E}$, $\mathbf{1}_B \in \mathscr{H}$.

Soit $f : E \to [0, M]$, une fonction $\mathscr{E}$-mesurable. Le lemme A.1.19 implique qu'il existe $c_n \in \mathbb{R}_+$ et $B_n \in \mathscr{E}$, $n \in \mathbb{N}$, tels que $f = \sum_{n\in\mathbb{N}} c_n \mathbf{1}_{B_n}$. Pour tout $n \in \mathbb{N}$, on pose $f_n = \sum_{0\leq k\leq n} c_k \mathbf{1}_{B_k}$. Comme $\mathscr{H}$ est un espace vectoriel, $f_n \in \mathscr{H}$, pour tout $n \in \mathbb{N}$. De plus on a $0 \leq f_n \leq f_{n+1} \leq M$. Donc $f = \sup_{n\in\mathbb{N}} f_n \in \mathscr{H}$, par $(c)$. L'espace vectoriel $\mathscr{H}$ contient donc toutes les fonctions $\mathscr{E}$-mesurables positives bornées. Si $f : E \to \mathbb{R}$ est $\mathscr{E}$-mesurable bornée, alors que qui précède montre que $f_+, f_- \in \mathscr{H}$ et puisque $\mathscr{H}$ est un espace vectoriel $f = f_+ - f_-$ est dans $\mathscr{H}$. ■

### Tribu engendrée par des fonctions et théorème de représentation.

**Définition A.1.21** Soit $E$, un ensemble ; soit $(E_i, \mathscr{E}_i)$, $i \in I$, une famille d'espaces mesurables ; soient $f_i : E \to E_i$, $i \in I$, des fonctions. On pose

$$\sigma(f_i, i \in I) = \bigcap \big\{ \mathscr{E} \subset \mathscr{P}(E) : \mathscr{E} \text{ tribu} : \forall i \in I,\ f_i \text{ est } (\mathscr{E}, \mathscr{E}_i)\text{-mesurable} \big\}.$$

La tribu $\sigma(f_i, i \in I)$ sur $E$ est appelée la *tribu engendrée par les fonctions* $(f_i)_{i\in I}$. C'est la plus petite tribu sur $E$ rendant mesurables toutes les fonctions $f_i$. La tribu $\sigma(f_i, i \in I)$ dépend des tribus $\mathscr{E}_i$, $i \in I$, bien que cela n'apparaisse pas dans la notation. □

Il est facile de vérifier que

$$\sigma(f_i, i \in I) = \sigma\big(\big\{ f_i^{-1}(B) \, ; \, i \in I,\ B \in \mathscr{E}_i \big\}\big) .$$

Lorsqu'on ne considère qu'une seule fonction, la tribu qu'elle engendre est particulièrement simple.

**Lemme A.1.22** *Soit $E$, un ensemble non-vide. Soit $(E', \mathscr{E}')$, un espace mesurable. Soit une fonction $f : E \to E'$. Alors,*

$$\sigma(f) = \big\{ f^{-1}(B) \; ; \; B \in \mathscr{E}' \big\} . \tag{A.4}$$

**Preuve :** exercice. ■

Montrons le théorème suivant qui trouve quelques applications en probabilité.

**Théorème A.1.23 (Représentation)** *Soit $E$, un ensemble. Soit $(E', \mathscr{E}')$, un espace mesurable. Soit $f : E \to E'$. On considère $h : E \to \mathbb{R}$. Alors, $h$ est $\sigma(f)$-mesurable ssi il existe une fonction $\varphi : E' \to \mathbb{R}$ qui soit $\mathscr{E}'$-mesurable et telle que $h = \varphi \circ f$.*

**Preuve :** on suppose d'abord que $h$ est $\sigma(f)$-mesurable à valeurs dans $\mathbb{R}_+$. Par le lemme A.1.19, il existe $c_n \in \mathbb{R}_+$ et $B_n \in \sigma(f)$, $n \in \mathbb{N}$, tels que $h = \sum_{n\in\mathbb{N}} c_n \mathbf{1}_{B_n}$. Par le lemme A.1.22, il existe $A_n \in \mathscr{E}'$ tel que $B_n = f^{-1}(A_n)$. On pose alors $\varphi_* = \sum_{n\in\mathbb{N}} c_n \mathbf{1}_{A_n}$. Il est clair que $\varphi_* : E' \to [0, \infty]$ est $\mathscr{E}'$-mesurable et on vérifie que $h = \varphi_* \circ f$. Comme $h(x) < \infty$, pour tout $x \in E$, on peut poser $\varphi = \varphi_* \mathbf{1}_{\varphi_*^{-1}([0,\infty[)}$, avec la convention $0 \times \infty = 0$. On voit alors que $\varphi : E' \to \mathbb{R}_+$ est $\mathscr{E}'$-mesurable et telle que $h = \varphi \circ f$, ce qui prouve l'implication directe du théorème dans le cas des fonctions positives. Le cas d'une fonction $h$ réelle s'obtient simplement en considérant les parties positives et négatives de $h$. La réciproque du théorème est élémentaire : nous laissons sa preuve au lecteur. ■

## A.2 Ensemble négligeables et complétion d'un espace mesuré.

**Définition A.2.1** Soit $(E, \mathscr{E}, \mu)$, un espace mesuré.

(a) On note

$$\mathscr{N}_\mu = \big\{ A \subset E : \exists B \in \mathscr{E} \text{ tel que } A \subset B \text{ et } \mu(B) = 0 \big\} .$$

L'ensemble $\mathscr{N}_\mu$ est appelé *classe des ensembles $\mu$-négligeables*. On remarque qu'un ensemble $\mu$-négligeable n'est pas nécessairement dans la tribu $\mathscr{E}$.

(b) On note $\mathscr{E}^\mu$ la tribu engendrée par $\mathscr{E}$ et $\mathscr{N}_\mu$ : $\mathscr{E}^\mu = \sigma(\mathscr{E}, \mathscr{N}_\mu)$. On appelle $\mathscr{E}^\mu$ la *tribu $\mu$-complétée de $\mathscr{E}$*.

(c) L'espace mesuré $(E, \mathscr{E}, \mu)$ est dit *complet* si $\mathscr{E}^\mu = \mathscr{E}$, autrment dit si $\mathscr{N}_\mu \subset \mathscr{E}$. □

**Proposition A.2.2** *Soit $(E, \mathscr{E}, \mu)$, un espace mesuré. Alors, $\mathscr{N}_\mu$ est stable par union dénombrable.*

**Preuve :** soient $N_n \in \mathscr{N}_\mu$. Il existe $B_n \in \mathscr{E}$ tel que $N_n \subset B_n$ et $\mu(B_n) = 0$. Donc $\bigcup_{n\in\mathbb{N}} N_n \subset \bigcup_{n\in\mathbb{N}} B_n \in \mathscr{E}$, et par sigma sous-additivité, $\mu(\bigcup_{n\in\mathbb{N}} B_n) \leq \sum_{n\in\mathbb{N}} \mu(B_n) = 0$. ■

**Théorème A.2.3 (Complétion d'un espace mesuré)** *Soit $(E, \mathscr{E}, \mu)$, un espace mesuré. On note $\mathscr{N}_\mu$ l'ensemble des $\mu$-négligeables de $E$ et on pose*

$$\mathscr{E}^\mu = \sigma(\mathscr{E}, \mathscr{N}_\mu) .$$

*Les assertions suivantes sont vraies.*

*(i) Si $B \in \mathscr{E}^\mu$, alors il existe $A \in \mathscr{E}$ tel que $A \subset B$ et $B \backslash A \in \mathscr{N}_\mu$.*

*(ii) Il existe une unique mesure positive $\overline{\mu} : \mathscr{E}^\mu \to [0, \infty]$ prolongeant $\mu$.*

*(iii) Les ensembles $\overline{\mu}$-négligeables sont exactement les $\mu$-négligeables. Par conséquent, l'espace mesuré $(E, \mathscr{E}^\mu, \overline{\mu})$ est complet : on l'appelle le* **complété** *de $(E, \mathscr{E}, \mu)$.*

**Preuve :** on pose $\mathscr{F} = \{B \subset E : \exists A \in \mathscr{E}$ tel que $A \subset B$ et $B \backslash A \in \mathscr{N}_\mu\}$. On a clairement $\mathscr{E} \subset \mathscr{F}$ et $\mathscr{N}_\mu \subset \mathscr{F}$. De plus, la définition de $\mathscr{F}$ montre que $\mathscr{F} \subset \mathscr{E}^\mu$. Pour montrer le point $(i)$, il suffit donc de montrer que $\mathscr{F}$ est une tribu. Clairement $E \in \mathscr{F}$. Vérifions ensuite la stabilité par union dénombrable : soient $B_n \in \mathscr{F}$, $n \in \mathbb{N}$. Il existe $A_n \in \mathscr{E}$, $n \in \mathbb{N}$, tels que $A_n \subset B_n$ et $N_n := B_n \backslash A_n \in \mathscr{N}_\mu$. Pour simplifier les notations on pose $A = \bigcup A_n$ et $B = \bigcup B_n$. On a donc $A \in \mathscr{E}$, $A \subset B$, et on remarque que $B \backslash A \subset \bigcup_{n \in \mathbb{N}} N_n \in \mathscr{N}_\mu$, par la proposition A.2.2. Donc $B \backslash A \in \mathscr{N}_\mu$, ce qui montre que $B \in \mathscr{F}$. Montrons ensuite la stabilité par passage au complémentaire : on se donne $B \in \mathscr{F}$. Il existe donc $A \in \mathscr{E}$ tel que $A \subset B$ et $N := B \backslash A \in \mathscr{N}_\mu$. On a donc $B = A \cup N$. Par définition des $\mu$-négligeables, il existe $N_0 \in \mathscr{E}$ tel que $N \subset N_0$ et $\mu(N_0) = 0$. On remarque alors que :

$$E \backslash B = (E \backslash A) \cap (E \backslash N) = \big( (E \backslash N_0) \cap (E \backslash A) \big) \cup \big( N_0 \cap (E \backslash A) \cap (E \backslash N) \big) ,$$

Comme $(E \backslash N_0) \cap (E \backslash A) \in \mathscr{E}$ et $N_0 \cap (E \backslash A) \cap (E \backslash N) \in \mathscr{N}_\mu$, on a $E \backslash B \in \mathscr{F}$. La classe d'ensembles $\mathscr{F}$ est donc bien une tribu, ce qui montre le point $(i)$ du théorème.

Montrons le point $(ii)$ : pour tout $B$ de la forme $A \cup N$, où $A \in \mathscr{E}$ et $N \in \mathscr{N}_\mu$, il est naturel de poser $\overline{\mu}(B) = \mu(A)$. Il faut tout d'abord vérifier que cela est cohérent, c'est-à-dire que cela ne dépend pas du choix de $A$ et de $N$ : supposons que $B = A' \cup N'$, avec $A' \in \mathscr{E}$ et $N' \in \mathscr{N}_\mu$. On pose $A'' = A \cap A'$. Alors

$$A \backslash A'' \subset N \cup N' \quad \text{et} \quad A' \backslash A'' \subset N \cup N' .$$

Or $N \cup N'$ est un $\mu$-négligeable. Il existe donc $N_0 \in \mathscr{E}$ tel que $N \cup N' \subset N_0$ et $\mu(N_0) = 0$ et comme $A \backslash A''$ et $A' \backslash A''$ sont contenus dans $N_0$ (et appartiennent à $\mathscr{E}$), on a $\mu(A \backslash A'') = \mu(A' \backslash A'') = 0$. Or

$$\mu(A) = \mu(A'') + \mu(A \backslash A'') = \mu(A'') \quad \text{et} \quad \mu(A') = \mu(A'') + \mu(A' \backslash A'') = \mu(A'') .$$

Donc $\mu(A) = \mu(A')$ et la définition de $\overline{\mu}(B)$ ne dépend pas du choix de $A$ et $N$.

Il est clair que $\overline{\mu}(\emptyset) = 0$. Il faut montrer ensuite que $\overline{\mu}$ est sigma-additive : soient $B_n = A_n \cup N_n$, $n \in \mathbb{N}$, une suite d'éléments deux-à-deux disjoints de $\mathscr{E}^\mu$, où $A_n \in \mathscr{E}$ et $N_n \in \mathscr{N}_\mu$, pour tout $n \in \mathbb{N}$. Pour simplifier les notations, on pose $B = \bigcup B_n$, $A = \bigcup A_n$ et $N = \bigcup N_n$. On a donc $B = A \cup N$. Il est clair que $A \in \mathscr{E}$ et la proposition A.2.2 implique que $N \in \mathscr{N}_\mu$. Par définition, on a $\overline{\mu}(B) = \mu(A)$. Par ailleurs, on observe que les $A_n$ sont disjoints deux-à-deux. Cela, combiné avec la sigma-additivité de $\mu$, implique que $\sum_{n \in \mathbb{N}} \overline{\mu}(B_n) = \sum_{n \in \mathbb{N}} \mu(A_n) = \mu(A) = \overline{\mu}(B)$, ce qui montre bien la sigma-additivité de $\overline{\mu}$.

Il reste ensuite à démontrer l'unicité : supposons que $\nu$ soit une mesure positive sur l'espace mesurable $(E, \mathscr{E}^\mu)$ telle que $\mu(A) = \nu(A)$, pour tout $A \in \mathscr{E}$. Cela implique que $\mathscr{N}_\mu \subset \mathscr{N}_\nu$. Soit $B \in \mathscr{E}^\mu$. Le point $(i)$ implique l'existence de $A \in \mathscr{E}$ et de $N \in \mathscr{N}_\mu$ tels que $B = A \cup N$. Comme $N \in \mathscr{N}_\nu$, $\nu(B) = \nu(A)$ et donc $\nu(B) = \nu(A) = \mu(A) = \overline{\mu}(B)$, ce qui montre que $\nu = \overline{\mu}$.

Montrons enfin la complétude de l'espace mesurable $(E, \mathscr{E}^\mu, \overline{\mu})$ : tout d'abord on a clairement $\mathscr{N}_\mu \subset \mathscr{N}_{\overline{\mu}}$. De plus, si $N \in \mathscr{N}_{\overline{\mu}}$, il existe $N_1 \in \mathscr{E}^\mu$ tel que $N \subset N_1$. Cela implique que $N \in \mathscr{N}_\mu$. Donc $\mathscr{N}_\mu = \mathscr{N}_{\overline{\mu}}$, ce qui montre le résultat voulu. ■

La définition de l'intégrale contre une mesure positive implique immédiatement le résultat suivant : si $g : E \to \overline{\mathbb{R}}$ est une application $\mathscr{E}$-mesurable qui est soit $\mu$-intégrable soit à valeurs dans $[0, \infty]$, alors :

$$\int_E g \, d\mu = \int_E g \, d\overline{\mu} . \tag{A.5}$$

Le théorème suivant discute des liens entre les fonctions $\mathscr{E}$-mesurables et les fonctions $\mathscr{E}^\mu$-mesurables ainsi que de leurs intégrales éventuelles contre respectivement $\mu$ et $\overline{\mu}$.

**Théorème A.2.4** *Soit $(E, \mathscr{E}, \mu)$, un espace mesuré. Soit $(E, \mathscr{E}^\mu, \overline{\mu})$, son complété. Alors les assertions suivantes sont vraies.*

*(i) Pour tout $B \in \mathscr{E}^\mu$, il exite $A_1, A_2 \in \mathscr{E}$ tels que*

$$A_1 \subset B \subset A_2 \ , \ \mu(A_1) = \overline{\mu}(B) = \mu(A_2) \quad \text{et} \quad \mu(A_2 \backslash A_1) = 0 .$$

*(ii) Soit $h : E \to [0, \infty]$, une application $\mathscr{E}^\mu$-mesurable. Il existe $h_1, h_2 : E \to [0, \infty]$, deux applications $\mathscr{E}$-mesurables telles que $h_1 \leq h \leq h_2$ et telles que*

$$\int_E h_1 \, d\mu = \int_E h \, d\overline{\mu} = \int_E h_2 \, d\mu \quad \text{et} \quad \mu(\{h_1 < h_2\}) = 0 .$$

*(iii) Soit $h : E \to \overline{\mathbb{R}}$, une application $\mathscr{E}^\mu$-mesurable. Il existe $h_1, h_2 : E \to \overline{\mathbb{R}}$, $\mathscr{E}$-mesurables telles que $h_1 \leq h \leq h_2$ et $\mu(\{h_1 < h_2\}) = 0$. De plus, $h$ est $\overline{\mu}$-intégrable ssi $h_1$ et (ou) $h_2$ sont $\mu$-intégrables, et dans ce cas on a*

$$\int_E h_1 \, d\mu = \int_E h \, d\overline{\mu} = \int_E h_2 \, d\mu . \tag{A.6}$$

**Preuve :** montrons $(i)$. Pour cela on se donne $B \in \mathscr{E}^\mu$. On sait qu'il existe $A \in \mathscr{E}$ et $N \in \mathscr{N}_\mu$ tels que $A \cup N = B$. Comme $N$ est $\mu$-négligeable, il existe $N_0 \in \mathscr{E}$ tel que $N \subset N_0$ et $\mu(N_0) = 0$. On pose alors $A_1 = A$ et $A_2 = A \cup N_0$ qui satisfont bien les propriétés désirées.

Montrons $(ii)$ : le lemme A.1.19 page 273, implique qu'il existe $c_n \in \mathbb{R}_+$, $B_n \in \mathscr{E}^\mu$, $n \in \mathbb{N}$ tels que $h = \sum_{n\in\mathbb{N}} c_n \mathbf{1}_{B_n}$. Le point $(i)$ implique qu'il existe $A_n^1, A_n^2 \in \mathscr{E}$ tels que $A_n^1 \subset B \subset A_n^2$, $\mu(A_n^1) = \overline{\mu}(B_n) = \mu(A_n^2)$ et $\mu(A_n^2 \backslash A_2^1) = 0$. Pour $i = 1$ et 2, on pose $h_i = \sum_{n\in\mathbb{N}} c_n \mathbf{1}_{A_n^i}$, qui est $\mathscr{E}$-mesurable. On a bien $h_1 \leq h \leq h_2$. De plus $\{h_1 < h_2\} \subset \bigcup_{n\in\mathbb{N}} A_n^2 \backslash A_n^1$, donc $\mu(\{h_1 < h_2\}) = 0$. Pour $i = 1$ et 2, la proposition **??** implique $\int_E h_i d\mu = \sum_{n\in\mathbb{N}} c_n \mu(A_n^i) = \sum_{n\in\mathbb{N}} c_n \overline{\mu}(B_n) = \int_E h d\overline{\mu}$, ce qui montre le point $(ii)$. Le point $(iii)$ se déduit facilement en appliquant $(ii)$ à la partie positive et négative de la fonction $h : E \to \overline{\mathbb{R}}$. ■

**Corollaire A.2.5** *Soit $(E, \mathscr{E}, \mu)$, une espace mesuré. Soit $(E, \mathscr{E}^\mu, \overline{\mu})$, son complété. Soit $h : E \to I$, une application $\mathscr{E}^\mu$-mesurable, où $I$ désigne $\overline{\mathbb{R}}$ ou $\mathbb{C}$. Alors, il existe deux fonctions $h^*, g : E \to I$ telles que $h = h^* + g$, avec $h^*$ $\mathscr{E}$-mesurable et $g$ nulle $\mu$-p.p. et donc $\mathscr{E}^\mu$-mesurable.*

*De plus, si l'intégrale $\int_E h\, d\overline{\mu}$ est bien définie (soit que $h$ soit positive, soit que $h$ soit $\overline{\mu}$-intégrable), alors il en est de même pour l'intégrale $\int_E h^*\, d\mu$ et on a bien $\int_E h\, d\overline{\mu} = \int_E h^*\, d\mu$.*

**Preuve :** lorsque $I = \overline{\mathbb{R}}$, il suffit de prendre $h^* = h_1$ et $g = h - h_1$, où $h_1$ est comme dans le théorème A.2.4 $(iii)$. Lorsque $I = \mathbb{C}$, on se ramène au cas réel en considérant la partie réelle de $f$ et sa partie imaginaire. ■

**Tribu universellement complétées** On rappelle la notion de tribu des universellement mesurable.

**Définition A.2.6** Soit $(E, \mathscr{E})$, un espace mesurable. On note $\mathscr{M}_1(E)$ l'ensemble des mesures de probabilités $\mu : \mathscr{E} \to [0,1]$. La tribu des universellement mesurable est définie par

$$\mathscr{E}^{\mathrm{uni}} = \bigcap_{\mu\in\mathscr{M}_1(E)} \mathscr{E}^\mu .$$

On a clairement $\mathscr{E} \subset \mathscr{E}^{\mathrm{uni}} \subset \mathscr{E}^\mu$, pour toute mesure $\mu \in \mathscr{M}_1(E)$. □

Il est clair que si $f : E \to \mathbb{R}$ est une fonction $\mathscr{E}^{\mathrm{uni}}$-mesurable, alors pour toute mesure $\mu \in \mathscr{M}_1(E)$, l'intégrale $\int_E f\, d\mu$ est bien définie.

## A.3 Régularité des mesures.

On considère un espace métrique $(E, d)$, muni de sa tribu des Boréliens notée $\mathscr{B}(E)$. Pour tout $A \subset E$, on note $\mathring{A}$ *l'intérieur de* $A$ qui est le plus grand ouvert contenu dans $A$ et $\overline{A}$, *l'adhérence de* $A$, qui est le plus petit fermé contenant $A$. Pour tout $x \in E$ et tout $r \in \mathbb{R}_+$, on note $B(x, r)$ (resp. $\overline{B}(x, r)$) la boule ouverte (resp. fermée) de centre $x$ et de rayon $r$. Pour tout $A \subset E$, non-vide, on définit

$$\forall x \in E, \quad d(x, A) = \inf_{y\in A} d(x, y) ,$$

qui est fonction *distance à* $A$. On voit facilement que $d(\cdot, A)$ est 1-Lipschitzienne, c'est-à-dire que

$$\forall x, x' \in E, \quad |d(x, A) - d(x', A)| \leq d(x, x') .$$

Cela implique en particulier que $d(\cdot, A)$ est continue. Il est facile de voir que l'ensemble de ses zéros est l'adhérence de $A$ :

$$\{x \in E : d(x, A) = 0\} = \overline{A} .$$

Donc, si $F$ est un fermé $d(x, F) = 0$ *ssi* $x \in F$.

**Définition A.3.1** Soit $(E, \mathscr{T})$, un espace topologique, dont la tribu Borélienne est notée $\mathscr{B}(E)$. Soit $\mu : \mathscr{B}(E) \to [0, \infty]$, une mesure positive.

(a) La mesure $\mu$ est dite *régulière extérieurement pour les ouverts* si pour tout $A \in \mathscr{B}(E)$

$$\mu(A) = \inf\{\mu(U) ; A \subset U \text{ et } U \text{ ouvert}\} . \tag{A.7}$$

(b) La mesure $\mu$ est dite *régulière intérieurement pour les fermés* si pour tout $A \in \mathscr{B}(E)$

$$\mu(A) = \sup\{\mu(F) ; F \subset A \text{ et } F \text{ fermé}\} . \tag{A.8}$$

(c) La mesure $\mu$ est dite *régulière intérieurement pour les compacts* si pour tout $A \in \mathscr{B}(E)$

$$\mu(A) = \sup\big\{\mu(K)\,;\ K \subset A \text{ et } K \text{ compact}\big\}\ . \tag{A.9}$$

(d) Une mesure $\mu : \mathscr{B}(E) \to [0, \infty]$ est dite *tendue*, si pour tout $\varepsilon > 0$, il existe un compact $K_\varepsilon$ tel que $\mu(E\backslash K_\varepsilon) < \varepsilon$. □

**Théorème A.3.2** *Soit $(E, d)$, un espace métrique. Soit $\mu : \mathscr{B}(E) \to \mathbb{R}_+$, une mesure finie. Alors, $\mu$ est régulière extérieurement pour les ouverts et intérieurement pour les fermés.*

**Preuve :** on note $\mathscr{E}$ la classe des ensembles $A \in \mathscr{B}(E)$ satisfaisant (A.7) et (A.8). Soit $U$, un ouvert de $E$ : (A.7) est évidente pour $A = U$. Pour tout $n \in \mathbb{N}$, on pose $F_n = \{x \in E : d(x, E\backslash U) \geq 2^{-n}\}$. Alors, $F_n$ est un fermé ; on a $F_n \subset F_{n+1} \subset U$ et $U = \bigcup_{n\in\mathbb{N}} F_n$. On a donc $\lim_n \mu(F_n) = \mu(U)$, ce qui implique que $U \in \mathscr{E}$. Donc $\mathscr{E}$ contient les ouverts. Par conséquent, pour prouver le théorème, il suffit de montrer que $\mathscr{E}$ est une tribu.

Il est clair que $E \in \mathscr{E}$ et, comme $\mu$ est finie, $\mathscr{E}$ est stable par passage au complémentaire. Soit $A_n \in \mathscr{E}$, $n \in \mathbb{N}$. Soit $\varepsilon > 0$. Il existe des ouverts $U_n$ et des fermés $F_n$ tels que $F_n \subset A_n \subset U_n$, $\mu(A_n\backslash F_n) < 2^{-n-1}\varepsilon$ et $\mu(U_n\backslash A_n) < 2^{-n-1}\varepsilon$. On pose $U = \bigcup_{n\in\mathbb{N}} U_n$, qui est un ouvert et $A = \bigcup_{n\in\mathbb{N}} A_n$. On a $A \subset U$ et $U\backslash A \subset \bigcup_{n\in\mathbb{N}} U_n\backslash A_n$. Comme $\mu$ est sigma sous-additive,

$$\mu(U) = \mu(A) + \mu(U\backslash A) \leq \mu(A) + \sum_{n\in\mathbb{N}} \mu(U_n\backslash A_n) \leq \mu(A) + \varepsilon\ .$$

Comme $\lim_N \mu(\bigcup_{0\leq n\leq N} A_n) = \mu(A)$ et comme $\mu$ est finie, il existe $N$ tel que $\mu(\bigcup_{0\leq n\leq N} A_n) \geq \mu(A) - \varepsilon$. On pose $B = \bigcup_{0\leq n\leq N} A_n$ et $F = \bigcup_{0\leq n\leq N} F_n$, qui est un fermé tel que $F \subset B \subset A$. Comme $B\backslash F \subset \bigcup_{0\leq n\leq N} A_n\backslash F_n$, on a

$$\mu(A) - \varepsilon \leq \mu(B) = \mu(F) + \mu(B\backslash F) \leq \mu(F) + \sum_{0\leq n\leq N} \mu(A_n\backslash F_n) \leq \mu(F) + \varepsilon\ .$$

On a donc montré que pour tout $\varepsilon > 0$, il existe un ouvert $U$ et un fermé $F$ tels que $F \subset A \subset U$ et $\mu(U) - \varepsilon \leq \mu(A) \leq \mu(F) + 2\varepsilon$, ce qui permet de montrer que $A \in \mathscr{E}$. La classe $\mathscr{E}$ est donc une tribu contenant les ouverts, ce qui termine la preuve du théorème. ■

Lorsque l'espace est séparable complet, il est possible d'être plus précis. On montre tout d'abord le résultat suivant, dû à Ulam.

**Théorème A.3.3** **(Théorème d'Ulam)** *Soit $(E, d)$, un espace métrique séparable et complet. Soit $\mu : \mathscr{B}(E) \to \mathbb{R}_+$, une mesure finie. Alors $\mu$ est tendue.*

**Preuve :** on fixe $\varepsilon \in ]0, 1[$ et on se donne une suite $x_n \in E$, $n \in \mathbb{N}$ qui est dense dans $E$, ce qui implique pour tout $p \in \mathbb{N}$ que $\bigcup_{n\in\mathbb{N}} \overline{B}(x_n, 2^{-p}) = E$. Puisque $\lim_n \mu(\bigcup_{0\leq k\leq n} \overline{B}(x_k, 2^{-p})) = \mu(E) < \infty$ il existe $N_p \in \mathbb{N}$ tel que

$$\mu\Big(E \setminus \bigcup_{0\leq n\leq N_p} \overline{B}(x_n, 2^{-p})\Big) < \varepsilon 2^{-p-1}\ .$$

On pose alors $K_\varepsilon = \bigcap_{p\in\mathbb{N}} \bigcup_{0\leq n\leq N_p} \overline{B}(x_n, 2^{-p})$. Il est clair que c'est un fermé et qu'il est pré-compact : c'est donc un compact. Par ailleurs, $E\backslash K_\varepsilon = \bigcup_{p\in\mathbb{N}} \big(E \setminus \bigcup_{0\leq n\leq N_p} \overline{B}(x_n, 2^{-p})\big)$ et par sigma sous-additivité de $\mu$, on a $\mu(E\backslash K_\varepsilon) < \sum_{p\in\mathbb{N}} \varepsilon 2^{-p-1} = \varepsilon$. ■

**Théorème A.3.4** *Soit $(E, d)$, un espace métrique séparable et complet. Soit $\mu : \mathscr{B}(E) \to \mathbb{R}_+$, une mesure finie. Alors, $\mu$ est régulière intérieurement pour les compacts.*

**Preuve :** soit $A \in \mathcal{B}(E)$. Soit $\varepsilon > 0$. Par le théorème A.3.2, il existe un fermé $F \subset A$ tel que $\mu(A\backslash F) < \varepsilon$. Par le théorème A.3.3 d'Ulam, il existe $K_\varepsilon \subset E$, compact tel que $\mu(E\backslash K_\varepsilon) < \varepsilon$. On pose $K = K_\varepsilon \cap F$, qui est compact tel que $K \subset F \subset A$. On remarque que $F\backslash K \subset E\backslash K_\varepsilon$ et on a

$$\mu(A) = \mu(A\backslash F) + \mu(F\backslash K) + \mu(K) < \varepsilon + \mu(E\backslash K_\varepsilon) + \mu(K) < 2\varepsilon + \mu(K)\ ,$$

ce qui implique le résultat voulu. ■

On étend ces résultats aux mesures qui ne sont pas finies, avec des restrictions sur l'espace métrique. Tout d'abord deux définitions.

**Définition A.3.5** Soit $(E, d)$, un espace métrique et $\mu : \mathscr{B}(E) \to [0, \infty]$, une mesure positive.

(a) La mesure $\mu$ est dite de *Radon* si pour tout compact $K \subset E$, on a $\mu(K) < \infty$.

(b) L'espace $(E, d)$ est dit *localement compact* si pour tout $x \in E$, il existe $r>0$ tel que la boule fermée de centre $x$ et de rayon $r$ soit compacte. □

**Lemme A.3.6** *Soit $(E, d)$, un espace métrique localement compact et séparable. Alors, il existe une suite de compacts $K_n \subset E$, $n \in \mathbb{N}$, tels que $K_n \subset \mathring{K}_{n+1}$ et $E = \bigcup_{n\in\mathbb{N}} K_n = \bigcup_{n\in\mathbb{N}} \mathring{K}_n$.*

**Preuve :** soit $x_q \in E$, $q \in \mathbb{N}$, une suite dense. On pose $S = \{(q,p) \in \mathbb{N}^2 : \overline{B}(x_q, 2^{-p}) \text{ est compacte }\}$. Comme $E$ est localement compact, pour tout $x \in E$, il existe $r_x > 0$ tel que $\overline{B}(x, r_x)$ soit compacte. Comme $(x_q)_{n\in\mathbb{N}}$ est dense, il existe $q_x, p_x \in \mathbb{N}$ tel que $d(x_{q_x}, x) < 2^{-p_x} < r_x/2$. On a donc $x \in \overline{B}(x_{q_x}, 2^{-p_x}) \subset \overline{B}(x, r)$ . Cela montre d'une part que $(q_x, p_x) \in S$, et comme $E = \bigcup_{x\in E} \overline{B}(x_{q_x}, 2^{-p_x})$, on a

$$E = \bigcup_{(q,p)\in S} \overline{B}(x_q, 2^{-p}) . \tag{A.10}$$

Pour tout $n \in \mathbb{N}$, on pose alors $C_n = \bigcup_{(q,p)\in S:p,q\leq n} \overline{B}(x_q, 2^{-p})$. C'est un compact, car union finie de compacts. On a clairement $C_n \subset C_{n+1}$ et (A.10) implique que $\bigcup_{n\in\mathbb{N}} C_n = E$.

On définit les $K_n$ à partir des $C_n$ par récurrence : on pose $K_0 = C_0$. Supposons que $K_n$ soit construit comme dans l'énoncé. On effectue un recouvrement de $K_n \cup C_{n+1}$ par des boules ouvertes d'adhérence compacte, ce qui est toujours possible car $(E, d)$ est supposé localement compact. On extrait de ce recouvrement un recouvrement fini $B_1 \cup \ldots \cup B_p$, ce qui est possible car $K_n \cup C_{n+1}$ est compact ; on pose alors $K_{n+1} = \overline{B_1} \cup \ldots \cup \overline{B_p}$, qui vérifie bien $K_n \subset B_1 \cup \ldots \cup B_p \subset \mathring{K}_{n+1}$. ■

On en déduit le théorème suivant.

**Théorème A.3.7** *Soit $(E, d)$, un espace métrique localement compact séparable. Soit $\mu : \mathscr{B}(E) \to [0, \infty]$, une mesure de Radon. Alors $\mu$ est régulière extérieurement pour les ouverts et intérieurement pour les compacts.*

**Preuve :** on se donne $K_n$ et $\mathring{K}_n$ comme dans le lemme A.3.6. Soit $A \in \mathscr{B}(E)$. Montrons la régularité extérieure. Si $\mu(A) = \infty$, alors $\mu(A) = \mu(E)$, ce qui montre la régularité extérieure dans ce cas car $E$ est ouvert.

On suppose que $\mu(A) < \infty$. Soit $\varepsilon > 0$. Par le théorème A.3.2 appliqué à $\mu(\cdot \cap \mathring{K}_n)$, il existe un ouvert $U_n$ tel que $A \cap \mathring{K}_n \subset U_n$ et tel que $\mu(U_n \cap \mathring{K}_n) - 2^{-n-1}\varepsilon \leq \mu(A \cap \mathring{K}_n)$. On pose $V_n = U_n \cap \mathring{K}_n$ et $V = \bigcup_{n\in\mathbb{N}} V_n$. Alors $V$ est un ouvert contenant $A$ et on a $V\backslash A \subset \bigcup_{n\in\mathbb{N}} V_n \backslash (A \cap \mathring{K}_n)$. Par sigma sous-additivité $\mu(V\backslash A) \leq \sum_{n\in\mathbb{N}} \mu(V_n \backslash (A \cap \mathring{K}_n)) < \varepsilon$. Donc $\mu(V) - \varepsilon \leq \mu(A)$. Comme $\varepsilon$ est arbitrairement petit, on obtient la régularité extérieure.

Montrons la régularité intérieure. On remarque que $(K_n, d)$ est un espace métrique séparable complet et on note $\mu_n = \mu(\cdot \cap K_n)$, la mesure trace de $\mu$ sur $K_n$. Soit $A \in \mathscr{B}(E)$. Le théorème A.3.4 appliqué à $\mu_n$, montre qu'il existe un compact $C_n \subset A \cap K_n$ tel que $\mu(A \cap K_n) \leq 2^{-n} + \mu(C_n)$. Donc

$$\mu(A \cap K_n) - 2^{-n} \leq \mu(C_n) \leq \mu(A) .$$

Or $\lim_n \uparrow \mu(A \cap K_n) = \mu(A)$, donc $\mu(A) = \lim_n \mu(C_n)$, et on conclut. ■

## A.4 Constructions de mesures.

### Mesures extérieures

**Définition A.4.1** Soit $E$ un ensemble non-vide. Une fonction $\mu : \mathscr{P}(E) \to [0, \infty]$ est appelée *mesure extérieure* si $\mu(\emptyset) = 0$ et si

$$\mu(A) \leq \sum_{n\in\mathbb{N}} \mu(B_n) ,$$

pour tout $A \subset E$ et pour tous $B_n \subset E$, $n \in \mathbb{N}$, tels que $A \subset \bigcup_{n\in\mathbb{N}} B_n$. □

Il est clair que si $A \subset B \subset E$, alors $\mu(A) \leq \mu(B)$. Il est également clair qu'une mesure extérieure est sigma sous-additive. Une mesure extérieure est définie pour *tous* les sous-ensembles de $E$. Contrairement à ce que son nom peut suggérer, *une mesure extérieure n'est pas, en général, une mesure positive*. Il s'avère en revanche que sa restriction à une certaine classe d'ensembles pour laquelle elle est additive, est une mesure positive. On rappelle que $B\backslash A = B \cap (E\backslash A) = \{x \in B : x \notin A\}$ et qu'il n'est pas nécessaire de supposer que $B \subset A$ pour utiliser cette notation.

**Définition A.4.2** (*Ensembles mesurables*) Soit $E$ un ensemble non-vide et $\mu$, une mesure extérieure sur $E$. Un sous-ensemble $B \subset E$ est dit *$\mu$-mesurable* si

$$\forall X \subset E , \quad \mu(X) = \mu(X \cap B) + \mu(X\backslash B) .$$

La classe des ensembles $\mu$-mesurables est notée $\mathscr{E}(\mu)$. □

Il est clair que $\mathscr{E}(\mu)$ est stable par passage au complémentaire, qu'elle contient $E$ et $\emptyset$. Par la propriété de sous-additivité de $\mu$, on voit que

$$B \in \mathscr{E}(\mu) \quad \Longleftrightarrow \quad \forall X \subset E, \quad \mu(X) \geq \mu(X \cap B) + \mu(X \backslash B) \ . \tag{A.11}$$

**Théorème A.4.3** *Soit $E$, un ensemble non-vide et $\mu : \mathscr{P}(E) \to [0, \infty]$, une mesure extérieure.*

*(i) $\mathscr{E}(\mu)$ est une tribu contenant les ensembles $N \subset E$ tels que $\mu(N) = 0$.*

*(ii) La restriction de $\mu$ à $\mathscr{E}(\mu)$ est une mesure positive.*

**Preuve :** soit $N \subset E$ tel que $\mu(N) = 0$. Alors pour tout $X \subset E$, on a $\mu(X \cap N) \leq \mu(N) = 0$. Donc $\mu(X \cap N) + \mu(X \backslash N) = \mu(X \backslash N) \leq \mu(X)$ et (A.11) implique que $N \in \mathscr{E}(\mu)$.

Montrons ensuite que $\mathscr{E}(\mu)$ est stable par union finie : soient $B, C \in \mathscr{E}(\mu)$ et $X \subset E$. On a donc

$$\begin{aligned} \mu(X) &= \mu(X \cap B) + \mu(X \backslash B) = \mu(X \cap B) + \mu\big((X \backslash B) \cap C\big) + \mu\big(X \backslash (B \cup C)\big) \\ &\geq \mu\big((X \cap B) \cup ((X \backslash B) \cap C)\big) + \mu\big(X \backslash (B \cup C)\big) \\ &\geq \mu\big(X \cap (B \cup C)\big) + \mu\big(X \backslash (B \cup C)\big) \end{aligned}$$

et (A.11) implique que $B \cup C \in \mathscr{E}(\mu)$. On remarque ensuite la propriété suivante découlant directement de la définition de $\mathscr{E}(\mu)$ :

$$\forall X \subset E, \ \ \forall B, C \in \mathscr{E}(\mu) \ \text{ tels que } B \cap C = \emptyset, \quad \mu\big(X \cap (B \cup C)\big) = \mu(X \cap B) + \mu(X \cap C) \ . \tag{A.12}$$

Montrons ensuite que $\mathscr{E}(\mu)$ est stable par union dénombrable : soient $B_n \in \mathscr{E}(\mu)$, $n \in \mathbb{N}$. On pose $C_0 = B_0$ et $C_n = \big(\bigcup_{0 \leq k \leq n} B_k\big) \backslash \big(\bigcup_{0 \leq k \leq n-1} B_k\big)$, pour tout $n \geq 1$. Comme $\mathscr{E}(\mu)$ est stable par union finie et par passage au complémentaire, elle est stable par intersection finie et donc $C_n \in \mathscr{E}(\mu)$. On observe aussi que les $C_n$ sont disjoints deux-à-deux et que $B := \bigcup_{n \in \mathbb{N}} B_n = \bigcup_{n \in \mathbb{N}} C_n$. Par une application répétée de (A.12), on a

$$\begin{aligned} \mu(X) &= \mu\Big(X \backslash \bigcup_{0 \leq k \leq n} C_k\Big) + \mu\Big(X \cap \bigcup_{0 \leq k \leq n} C_k\Big) = \mu\Big(X \backslash \bigcup_{0 \leq k \leq n} C_k\Big) + \sum_{0 \leq k \leq n} \mu(X \cap C_k) \\ &\geq \mu(X \backslash B) + \sum_{0 \leq k \leq n} \mu(X \cap C_k) \ . \end{aligned}$$

Lorsque $n \to \infty$, on a $\mu(X) \geq \mu(X \backslash B) + \sum_{n \in \mathbb{N}} \mu(X \cap C_n)$. Or $X \cap B = \bigcup_{n \in \mathbb{N}} X \cap C_n$. Par sigma sous-additivité de $\mu$, on en déduit que

$$\mu(X) \geq \mu(X \backslash B) + \sum_{n \in \mathbb{N}} \mu(X \cap C_n) \geq \mu(X \backslash B) + \mu(X \cap B) \tag{A.13}$$

et donc $B \in \mathscr{E}(\mu)$ par (A.11). Cela montre que $\mathscr{E}(\mu)$ est une tribu et cela termine la preuve de $(i)$.

Il reste à prouver que $\mu$ est sigma-additive sur $\mathscr{E}(\mu)$ : soient $C_n \in \mathscr{E}(\mu)$, $n \in \mathbb{N}$, deux-à-deux disjoints. Il suffit alors d'appliquer (A.13) à $X = B = \bigcup_{n \in \mathbb{N}} C_n$. ■

On prouve ici un lemme technique utilisé plus loin.

**Lemme A.4.4** *Soit $E$, un ensemble non-vide. Soit $\mu : \mathscr{P}(E) \to [0, \infty]$, une mesure extérieure. On note $\mathscr{E}(\mu)$ la classe des ensembles $\mu$-mesurables. Soit $f : E \to [-\infty, \infty]$. Alors, $f$ est $\mathscr{E}(\mu)$-mesurable si elle satisfait*

$$\forall\, X \subset E, \ \forall\, a, b \in \mathbb{R} \textit{ tels que } a < b, \qquad \mu(X) \geq \mu\big(X \cap \{f \leq a\}\big) + \mu\big(X \cap \{f \geq b\}\big) \ . \tag{A.14}$$

**Preuve :** on suppose (A.14). Il faut montrer que $\{f \leq x\} \in \mathscr{E}(\mu)$ pour tout $x \in \mathbb{R}$. On fixe $X \subset E$ tel que $\mu(X) < \infty$ et $x \in \mathbb{R}$. Soient $a_1 \leq b_1 < a_2 \leq b_2 \ldots < a_n \leq b_n$. On applique (A.14) à $a = b_{n-1}$, $b = a_n$ et $X^* = X \cap \bigcup_{1 \leq k \leq n} \{f \in [a_k, b_k]\}$, pour obtenir

$$\begin{aligned} \mu(X) \geq \mu\Big(X \cap \bigcup_{1 \leq k \leq n} \{f \in [a_k, b_k]\}\Big) &= \mu(X^*) \geq \mu\big(X^* \cap \{f \leq b_{n-1}\}\big) + \mu\big(X^* \cap \{f \geq a_n\}\big) \\ &= \mu\Big(X \cap \bigcup_{1 \leq k \leq n-1} \{f \in [a_k, b_k]\}\Big) + \mu\big(X \cap \{f \in [a_n, b_n]\}\big) \\ &\geq \sum_{1 \leq k \leq n} \mu\big(X \cap \{f \in [a_k, b_k]\}\big) \ , \end{aligned}$$

par une récurrence immédiate. On en déduit que $\mu(X) \geq \sum_{n\in\mathbb{N}} \mu\big(X \cap \{f \in [x+2^{-2n-1}, x+2^{-2n}]\}\big)$ et que $\mu(X) \geq \sum_{n\in\mathbb{N}} \mu\big(X \cap \{f \in [x+2^{-2n-2}, x+2^{-2n-1}]\}\big)$. En sommant ces deux inégalités, on obtient que $\sum_{n\in\mathbb{N}} \mu\big(X \cap \{f \in [x+2^{-n-1}, x+2^{-n}]\}\big) \leq 2\mu(X) < \infty$. Pour tout $\varepsilon > 0$, il existe $n_\varepsilon$ tel que

$$\varepsilon > \sum_{n \geq n_\varepsilon} \mu\big(X \cap \{f \in [x+2^{-n-1}, x+2^{-n}]\}\big) \geq \mu\big(X \cap \{f \in ]x, x+2^{-n_\varepsilon}]\}\big) , \tag{A.15}$$

la dernière inégalité étant une conséquence de la sigma sous-additivité de $\mu$. On a donc

$$\begin{aligned} \mu\big(X \cap \{f \leq x\}\big) + \mu\big(X \cap \{f > x\}\big) &\leq \mu\big(X \cap \{f \leq x\}\big) + \mu\big(X \cap \{f \in ]x, x+2^{-n_\varepsilon}]\}\big) + \mu\big(X \cap \{f \geq x+2^{-n_\varepsilon}\}\big) \\ &\leq \mu\big(X \cap \{f \leq x\}\big) + \mu\big(X \cap \{f \geq x+2^{-n_\varepsilon}\}\big) + \varepsilon \\ &\leq \mu(X) + \varepsilon , \end{aligned}$$

par sous-additivité de $\mu$ pour la première inégalité, par (A.15) pour la deuxième et par (A.14) pour la troisième. Comme $\varepsilon$ peut être arbitrairement petit, on a donc $\mu\big(X \cap \{f \leq x\}\big) + \mu\big(X \cap \{f > x\}\big) \leq \mu(X)$, et cela montre que $\{f \leq x\} \in \mathscr{E}(\mu)$ par (A.11). Comme $x$ est arbitraire dans $\mathbb{R}$, cela montre que $f$ est $\mathscr{E}(\mu)$-mesurable. ■

**Extension de Carathéodory. Construction de la mesure de Lebesgue sur $\mathbb{R}$.** Cette section se propose, à partir d'une fonction additive d'ensembles, de construire un prolongement qui est une mesure. Plus précisément, on introduit les définitions suivantes.

**Définition A.4.5** Soit $E$, un ensemble non-vide.

(a) $\mathscr{A} \subset \mathscr{P}(E)$ est une *algèbre* si $E \in \mathscr{A}$ et si $\mathscr{A}$ est stable par passage au complémentaire et par union finie.

(b) Soit $\mathscr{A}$, une algèbre sur $E$. Une fonction d'ensemble $\mu : \mathscr{A} \to [0, \infty]$ est dite *additive* si

$$\forall A, B \in \mathscr{A} \text{ tels que } A \cap B = \emptyset, \quad \mu(A \cup B) = \mu(A) + \mu(B) .$$

(c) Soit $\mathscr{A}$, une algèbre sur $E$. Une fonction d'ensemble $\mu : \mathscr{A} \to [0, \infty]$ est dite *sigma additive* si

$$\forall A_n \in \mathscr{A},\ n \in \mathbb{N}, \text{ deux-à-deux disjoints et tels que } \bigcup_{n\in\mathbb{N}} A_n \in \mathscr{A} , \quad \mu\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) = \sum_{n\in\mathbb{N}} \mu(A_n) .$$

On remarque qu'il est nécessaire d'imposer $\bigcup_{n\in\mathbb{N}} A_n \in \mathscr{A}$ dans cette définition car a priori $\mathscr{A}$ n'est pas stable par union dénombrable.

(d) Soit $\mathscr{A}$, une algèbre sur $E$. Une fonction d'ensemble $\mu : \mathscr{A} \to [0, \infty]$ est dite *sigma finie* s'il existe $E_n \in \mathscr{A}$, $n \in \mathbb{N}$, tels que $E = \bigcup_{n\in\mathbb{N}} E_n$ et $\mu(E_n) < \infty$, pour tout $n \in \mathbb{N}$. □

Il est clair que si $\mu$ est additive sur une algèbre $\mathscr{A}$, alors $\mu(\emptyset) = 0$ et pour tous $A, B \in \mathscr{A}$ tels que $A \subset B$, on a $\mu(A) \leq \mu(B)$. En général, une fonction additive d'ensembles n'est pas sigma additive. La proposition suivante donne un critère utile de sigma additivité.

**Proposition A.4.6** *Soit $E$, un ensemble non-vide. Soit $\mathscr{A}$, une algèbre sur $E$. Soit $\mu : \mathscr{A} \to \mathbb{R}_+$, une fonction additive. On suppose que $\mu(E) < \infty$. Alors, il y a équivalence entre les deux assertions suivantes.*

*(i) $\mu$ est sigma additive sur $\mathscr{A}$.*

*(ii) Pour tous $A_n \in \mathscr{A}$, $n \in \mathbb{N}$, tels que $A_{n+1} \subset A_n$ et $\bigcap_{n\in\mathbb{N}} A_n = \emptyset$, on a $\lim_n \mu(A_n) = 0$.*

**Preuve :** montrons d'abord $(i) \Longrightarrow (ii)$. Soit $(A_n)_{n\in\mathbb{N}}$ comme dans $(ii)$. On pose $B_n = A_n \backslash A_{n+1}$, $n \in \mathbb{N}$. Les $B_n$ sont des ensembles de $\mathscr{A}$ deux-à-deux disjoints tel que $\bigcup_{p\geq n} B_p = A_n \in \mathscr{A}$. Donc $\mu(A_n) = \sum_{p\geq n} \mu(B_p)$, qui est le reste d'une série convergente car $\mu(E) < \infty$. Donc $\lim_n \mu(A_n) = 0$.

Montrons $(ii) \Longrightarrow (i)$ : soient $B_n \in \mathscr{A}$, $n \in \mathbb{N}$, deux-à-deux disjoints tels que $B := \bigcup_{n\in\mathbb{N}} B_n \in \mathscr{A}$. On pose $A_n = B \backslash (B_0 \cup \ldots \cup B_n)$. On voit que les $A_n$ satisfont les conditions de $(ii)$. Comme $\mu$ est finie

$$\mu(A_n) = \mu(B) - \mu(B_0 \cup \ldots \cup B_n) = \mu(B) - \sum_{0\leq k\leq n} \mu(B_k) ,$$

et $(ii)$ implique que $\mu(B) = \sum_{n\in\mathbb{N}} \mu(B_n)$, ce qui prouve que $\mu$ est sigma additive. ■

Tout comme les mesures positives, les fonctions sigma additives sont stables par restriction et sommation, comme le montre la proposition suivante dont la preuve, laissée au lecteur.

**Proposition A.4.7** *Soit $E$, un ensemble non-vide. Soit $\mathscr{A}$ une algèbre sur $E$.*

*(i) Soit $\mu : \mathscr{A} \to [0,\infty]$, une fonction sigma additive. Soit $A \in \mathscr{A}$. Alors $\mu(\cdot \cap A) : \mathscr{A} \to [0,\infty]$ est sigma additive.*

*(ii) Soient $\mu_n : \mathscr{A} \to [0,\infty]$, $n \in \mathbb{N}$, des fonctions sigma additives. Alors $\sum_{n\in\mathbb{N}} \mu_n : \mathscr{A} \to [0,\infty]$ est sigma additive.*

La proposition suivante montre l'existence d'une fonction sigma-additive sur les unions finies d'intervalles qui correspond à la notion intuitive de longueur.

**Proposition A.4.8** *On note $\mathscr{P}$ la classe des intervalles de $\mathbb{R}$ et $\mathscr{A}$ la classe des unions finies d'intervalles deux-à-deux disjoints. Alors, les assertions suivantes sont vraies.*

*(i) $\mathscr{P}$ est un pi-système engendrant les Boréliens de $\mathbb{R}$.*

*(ii) $\mathscr{A}$ est une algèbre contenant $\mathscr{P}$, donc $\sigma(\mathscr{A}) = \mathscr{B}(\mathbb{R})$.*

*(iii) Il existe une fonction $\mathrm{long} : \mathscr{A} \to [0,\infty]$ qui est sigma additive, sigma finie et telle que*

$$\forall a, b \in \mathbb{R} \textit{ tels que } a \le b, \qquad b - a = \mathrm{long}\big([a,b]\big) = \mathrm{long}\big(\,]a,b]\big) = \mathrm{long}\big([a,b[\,\big) = \mathrm{long}\big(\,]a,b[\,\big) \; .$$

**Preuve :** $(i)$ est un fait élémentaire. Il est clair que $\mathscr{P} \subset \mathscr{A}$, donc $\mathbb{R} \in \mathscr{A}$. Il est facile de vérifier que $\mathscr{A}$ est stable par union finie. Comme le complémentaire d'un intervalle est la réunion d'au plus deux intervalles disjoints, $\mathscr{A}$ est stable par passage au complémentaire, ce qui prouve $(ii)$.

Soit $I$ un intervalle d'extrémités $a$ et $b$, avec $a \le b$. On pose $\mathrm{long}(I) = b - a$. Si $I$ est un intervalle non-borné, on pose $\mathrm{long}(I) = \infty$.

Supposons que l'on ait montré que pour tout $q \in \mathbb{Z}$, il existe $\mathrm{long}_q : \mathscr{A} \to [0,\infty]$ qui est sigma-additive telle que pour tout $I \in \mathscr{P}$, $\mathrm{long}_q(I) = \mathrm{long}(I \cap [q, q+1[)$. On pose alors $\lambda = \sum_{q\in\mathbb{Z}} \mathrm{long}_q$. Par la proposition A.4.7, $\lambda : \mathscr{A} \to [0,\infty]$ est sigma additive. Soit $I$ un intervalle borné d'extrémités $a, b \in \mathbb{R}$ telles que $a \le b$. Alors, $I \cap [q, q+1[$ est vide ou réduit à un point si $b \le q$ ou $q+1 \le a$ ; sinon, $I \cap [q, q+1[$ est un intervalle d'extrémité gauche $q \vee a$ et d'extrémité droite $(q+1) \wedge b$. Donc $\mathrm{long}(I \cap [q, q+1[) = \big((q+1) \wedge b - q \vee a\big)_+$. On voit alors que

$$\lambda(I) = \sum_{q\in\mathbb{Z}} \mathrm{long}_q(I) = \sum_{q\in\mathbb{Z}} \big((q+1) \wedge b - q \vee a\big)_+ = b - a$$

et $\lambda$ satisfait bien les conditions de $(iii)$. Cela montre qu'il suffit de prouver le résultat suivant.

$(*)$ Soit $q \in \mathbb{Z}$ ; on note $\mathscr{P}_q$ les intervalles de $[q, q+1[$ et $\mathscr{A}_q$ les unions finies d'intervalles de $\mathscr{P}_q$ deux-à-deux disjoints. On a clairement $\mathscr{A}_q = \{A \cap [q, q+1[\, ; A \in \mathscr{A}\}$ qui est une algèbre sur $[q, q+1[$. Il existe $\mathrm{long}_q : \mathscr{A}_q \to \mathbb{R}_+$, une fonction sigma-additive telle que $\mathrm{long}_q(I) = \mathrm{long}(I)$ pour tout $I \in \mathscr{P}_q$.

Prouvons ce résultat : soient $I_1, \ldots, I_n \in \mathscr{P}_q$, disjoints deux-à-deux. On pose $A = \bigcup_{1\le k\le n} I_k$ et on définit

$$\mathrm{long}_q(A) := \sum_{1\le k\le n} \mathrm{long}_q(I_k) \; .$$

Il est facile de vérifier que cette définition est cohérente, c'est-à-dire qu'elle ne dépend pas de la façons dont $A$ s'écrit comme union finie d'intervalles de $\mathscr{P}_q$ disjoints deux -à-deux. Par ailleurs, cette définition implique immédiatement que $\mathrm{long}_q$ *est additive*.

Montrons la sigma-additivité de $\mathrm{long}_q$. Soient $I_1, \ldots, I_n \in \mathscr{P}_q$, disjoints deux-à-deux et non-vides. On note $a_k$ l'extrémité gauche de $I_k$ et $b_k$ son extrémité droite. Soit $\varepsilon > 0$. Si $\varepsilon \le 2^{k+1}(b_k - a_k)$, on pose $J_k = ]a_k + 2^{-k-2}\varepsilon, b_k - 2^{-k-2}\varepsilon[$ ; sinon on pose $J_k = \emptyset$. Alors, les $J_k \in \mathscr{P}_q$ sont disjoints deux-à-deux, $\bigcup \overline{J_k} \subset \bigcup I_k$ et $\mathrm{long}_q(\bigcup I_k) = \mathrm{long}_q(\bigcup J_k) + \varepsilon(2^{-1} + \ldots + 2^{-n-1})$. Cela montre que

$$\forall A \in \mathscr{A}_q,\ \forall \varepsilon > 0, \quad \exists B_\varepsilon \in \mathscr{A}_q \ : \ \overline{B_\varepsilon} \subset A \quad \text{et} \quad \mathrm{long}_q(A) \le \mathrm{long}_q(B_\varepsilon) + \varepsilon \; . \tag{A.16}$$

On utilise le critère de la proposition A.4.6. Soient $A_n \in \mathscr{A}_q$ tels que $A_{n+1} \subset A_n$ et $\bigcap_{n\in\mathbb{N}} A_n = \emptyset$. Soit $\varepsilon > 0$. Par (A.16), il existe $B_n \in \mathscr{A}_q$ tels que $\overline{B_n} \subset A_n$ et $\mathrm{long}_q(A_n) \le 2^{-n-1}\varepsilon + \mathrm{long}_q(B_n)$. On pose $C_n = B_0 \cap \ldots \cap B_n$. On a clairement $C_n \in \mathscr{A}_q$, $C_n \subset C_{n+1}$, $\overline{C_n} \subset A_n$ et $A_n \backslash C_n \subset \bigcup_{0\le k\le n} A_k \backslash B_k$, donc

$$\mathrm{long}_q(A_n \backslash C_n) \le \sum_{0\le k\le n} \mathrm{long}_q(A_k \backslash B_k) \le \varepsilon(2^{-1} + \ldots 2^{-n-1}) < \varepsilon \; .$$

Comme $\bigcap_{n\in\mathbb{N}} A_n = \emptyset$, on également $\bigcap_{n\in\mathbb{N}} \overline{C_n} = \emptyset$. Par compacité, il existe $n_\varepsilon \in \mathbb{N}$ tel que $C_n = \emptyset$, pour tout $n \ge n_\varepsilon$. Cela implique que $\mathrm{long}_q(A_n) \le \varepsilon$ pour tout $n \ge n_\varepsilon$. Cela prouve que $\lim_n \mathrm{long}_q(A_n) = 0$. La proposition A.4.6 implique que $\mathrm{long}_q : \mathscr{A} \to \mathbb{R}_+$ est sigma additive, ce qui prouve $(*)$ et donc la proposition. ■

Le théorème suivant, dû à Carathéodory, associe à une fonction sigma additive sur une algèbre, une mesure extérieure qui la prolonge.

**Théorème A.4.9 (Théorème d'extension de Caratheodory)** *Soit $E$, un ensemble non-vide. Soit $\mathscr{A}$, une algèbre sur $E$. Soit $\mu_0 : \mathscr{A} \to [0, \infty]$, supposée sigma additive. Pour tout $X \subset E$, on pose*

$$\mu(X) = \inf \Big\{ \sum_{n \in \mathbb{N}} \mu_0(A_n) \; ; \; A_n \in \mathscr{A}, \, n \in \mathbb{N} \, : \, X \subset \bigcup_{n \in \mathbb{N}} A_n \Big\} . \tag{A.17}$$

*Alors, les assertions suivantes sont vérifiées.*

*(i)* $\mu : \mathscr{P}(E) \to [0, \infty]$ *est une mesure extérieure.*

*(ii) Pour tout* $A \in \mathscr{A}$, $\mu(A) = \mu_0(A)$.

*(iii)* $\sigma(\mathscr{A}) \subset \mathscr{E}(\mu)$, *où* $\mathscr{E}(\mu)$ *désigne la classe des ensembles* $\mu$*-mesurables.*

*(iv) Pour tout* $X \subset E$, *il existe* $B \in \sigma(\mathscr{A})$ *tel que* $X \subset B$ *et* $\mu(X) = \mu(B)$.

**Preuve :** on vérifie immédiatement que $\mu(\emptyset) = 0$ et que pour tous $B \subset C \subset E$, on a bien $\mu(B) \leq \mu(C)$. On fixe ensuite $B_n \subset E$, $n \in \mathbb{N}$. Par définition de $\mu$, pour tout $\varepsilon > 0$, il existe des ensembles $A_{n,p} \in \mathscr{A}$, $n, p \in \mathbb{N}$, tels que $B_n \subset \bigcup_{p \in \mathbb{N}} A_{n,p}$ et $\sum_{p \in \mathbb{N}} \mu_0(A_{p,n}) \leq \mu(B_n) + 2^{-n-1}\varepsilon$, pour tout $n \in \mathbb{N}$. On constate donc que $\bigcup_{n \in \mathbb{N}} B_n \subset \bigcup_{p,n \in \mathbb{N}} A_{p,n}$ et donc que

$$\mu\big( \bigcup_{n \in \mathbb{N}} B_n \big) \leq \sum_{p,n \in \mathbb{N}} \mu_0(A_{p,n}) \leq \sum_{n \in \mathbb{N}} \mu(B_n) \; + \; \varepsilon$$

Comme $\varepsilon$ peut être arbitrairement petit, on a $\mu(\bigcup_{n \in \mathbb{N}} B_n) \leq \sum_{n \in \mathbb{N}} \mu(B_n)$, ce qui montre que $\mu$ une mesure extérieure.

Montrons $(ii)$. On fixe $A \in \mathscr{A}$. La définition de $\mu$ implique que $\mu(A) \leq \mu_0(A)$. Montrons l'inégalité contraire : on se donne $A_n \in \mathscr{A}$, $n \in \mathbb{N}$, tels que $A \subset \bigcup_{n \in \mathbb{N}} A_n$. Pour tout $n \in \mathbb{N}$, on pose

$$B_n = \bigcup_{k=0}^{n} A \cap A_k \; , \quad C_0 = B_0 \quad \text{et} \quad C_{n+1} = B_{n+1} \backslash B_n \quad n \in \mathbb{N} \; .$$

Il est clair que $B_n, C_n \in \mathscr{A}$, que les $C_n$ sont deux-à-deux disjoints, que $C_n \subset A \cap A_n \subset A_n$, et que $\bigcup_{n \in \mathbb{N}} C_n = A \in \mathscr{A}$. Comme $\mu_0$ est une pré-mesure, on a donc

$$\mu_0(A) = \sum_{n \in \mathbb{N}} \mu_0(C_n) \leq \sum_{n \in \mathbb{N}} \mu_0(A \cap A_n) \leq \sum_{n \in \mathbb{N}} \mu_0(A_n) \; .$$

Cette inégalité étant vraie pour toute suite $A_n \in \mathscr{A}$, $n \in \mathbb{N}$, telle que $A \subset \bigcup_{n \in \mathbb{N}} A_n$, on en déduit que $\mu_0(A) \leq \mu(A)$, ce qui achève la preuve de $(ii)$.

Montrons $(iii)$. Comme $\mathscr{E}(\mu)$ est une tribu par le théorème A.4.3 $(i)$ page 279, il suffit de montrer que $\mathscr{A} \subset \mathscr{E}(\mu)$, c'est-à-dire qu'il suffit, par (A.11), de montrer que pour tout $A \in \mathscr{A}$ et tout $X \subset E$, on a $\mu(X) \geq \mu(X \cap A) + \mu(X \backslash A)$ : si $\mu(X) = \infty$, c'est trivialement vrai. Supposons $\mu(X) < \infty$. Pour tout $\varepsilon > 0$, il existe donc $B_n \in \mathscr{A}$, $n \in \mathbb{N}$, tels que $X \subset \bigcup_{n \in \mathbb{N}} B_n$ et $\varepsilon + \mu(X) \geq \sum_{n \in \mathbb{N}} \mu_0(B_n)$. On a donc

$$\varepsilon + \mu(X) \geq \sum_{n \in \mathbb{N}} \mu_0(B_n \cap A) + \sum_{n \in \mathbb{N}} \mu_0(B_n \backslash A) \geq \mu(X \cap A) + \mu(X \backslash A) \; ,$$

ce qui entraîne bien le résultat voulu car $\varepsilon > 0$ peut être arbitrairement petit.

Montrons $(iv)$. On fixe $X \subset E$. Si $\mu(X) = \infty$, alors $\mu(X) = \mu(E)$ et $(iv)$ est trivialement vérifiée. Supposons que $\mu(X) < \infty$. Il existe donc $A_{p,n} \in \mathscr{A}$, $p, n \in \mathbb{N}$, tels que $X \subset \bigcup_{n \in \mathbb{N}} A_{p,n} =: B_p$ et $2^{-p} + \mu(X) \geq \sum_{n \in \mathbb{N}} \mu_0(A_{p,n}) \geq \mu(B_p)$. Il est clair que $B_p \in \sigma(\mathscr{A})$ et donc $B := \bigcap_{p \in \mathbb{N}} B_p \in \sigma(\mathscr{A})$. De plus, $X \subset B$. Enfin, pour tout $p \in \mathbb{N}$, on a $B \subset B_p$ et donc $\mu(B) \leq \mu(B_p) \leq \mu(X) + 2^{-p} \leq \mu(B) + 2^{-p}$, ce qui implique que $\mu(B) = \mu(X)$. ■

Le théorème d'extension de Carathéodory et le théorème d'unicité du prolongement des mesures impliquent immédiatement le corollaire suivant.

**Théorème A.4.10 (Théorème d'extension de Carathéodory)** *Soit $E$, un ensemble non-vide. Soit $\mathscr{A}$, une algèbre sur $E$. Soit $\mu_0 : \mathscr{A} \to [0, \infty]$, supposée sigma additive. Alors, il existe une mesure positive $\mu : \sigma(\mathscr{A}) \to [0, \infty]$ prolongeant $\mu_0$. De plus, ce prolongement est unique si $\mu_0$ est sigma finie.*

On déduit également des résultats précédents l'existence de la mesure de Lebesgue sur $\mathbb{R}$.

**Théorème A.4.11** **(Existence de la mesure de Lebesgue)** *Soit $I \subset \mathbb{R}$, un intervalle. Si $I$ est borné, on pose $\mathrm{long}(I) = b - a$ où $a$ est l'extrémité gauche de $I$ et $b$ son extrémité droite, et si $I$ est non-borné, on pose $\mathrm{long}(I) = \infty$. On on définit alors $\ell : \mathscr{P}(\mathbb{R}) \to [0, \infty]$ par*

$$\forall X \subset \mathbb{R}, \quad \ell(X) = \inf \Big\{ \sum_{n \in \mathbb{N}} \mathrm{long}\,(I_n)\,; \quad (I_n)_{n \in \mathbb{N}} \textit{ intervalles tels que } X \subset \bigcup_{n \in \mathbb{N}} I_n \Big\} .$$

*C'est la* **mesure (extérieure) de Lebesgue sur** $\mathbb{R}$. *La tribu de ses ensembles mesurables, appelée* **tribu Lebesguienne**, *est notée $\mathcal{L}(\mathbb{R})$. Elle contient $\mathscr{B}(\mathbb{R})$. La restriction de $\ell$ à $\mathscr{B}(\mathbb{R})$ est l'unique mesure positive telle que*

$$\forall a, b \in \mathbb{R} \textit{ tels que } a \leq b, \qquad b - a = \ell\big([a, b]\big) = \ell\big(\,]a, b]\big) = \ell\big([a, b[\,\big) = \ell\big(\,]a, b[\,\big) .$$

**Extension, complétion, approximation.** Soit $(E, \mathscr{E}, \mu)$, un espace mesuré. On rappelle que $\mathscr{N}_\mu$ désigne l'ensemble des $\mu$-négligeable, que $\mathscr{E}^\mu = \sigma(\mathscr{E}, \mathscr{N}_\mu)$ et que $(E, \mathscr{E}^\mu, \overline{\mu})$ est le complété de $(E, \mathscr{E}, \mu)$. Soit $\mathscr{A}$, un algèbre sur $E$. On suppose que

$$\sigma(\mathscr{A}) = \mathscr{E}\,, \quad \exists E_n \in \mathscr{A}, n \in \mathbb{N}, \text{ satisfaisant : } \quad \mu(E_n) < \infty \quad \text{et} \quad \bigcup_{n \in \mathbb{N}} E_n = E\,. \tag{A.18}$$

Cela implique notamment que $\mu$ *est sigma finie*. Pour tout $X \subset E$, on pose

$$\mu^*(X) = \inf \Big\{ \sum_{n \in \mathbb{N}} \mu(A_n)\,; \quad A_n \in \mathscr{A}, n \in \mathbb{N} \;:\; X \subset \bigcup_{n \in \mathbb{N}} A_n \Big\} . \tag{A.19}$$

On rappelle que $\mathscr{E}(\mu^*)$ désigne la tribu des ensembles $\mu^*$-mesurables. Le théorème A.4.3 d'extension de Carathéodory montre que

$$\mathscr{E} = \sigma(\mathscr{A}) \subset \mathscr{E}(\mu^*)\,, \quad \forall A \in \mathscr{E}\,, \quad \mu(A) = \mu^*(A) \quad \text{et} \quad \big\{N \subset E : \mu^*(N) = 0\big\} \subset \mathscr{E}(\mu^*). \tag{A.20}$$

En effet, le théorème A.4.9 d'extension de Carathéodory (page 282) implique seulement que pour tout $A \in \mathscr{A}$, on a $\mu(A) = \mu^*(A)$, mais l'hypothèse (A.18) permet d'appliquer l'unicité du prolongement des mesures.

On rappelle ensuite que la *différence symétrique de deux ensembles $A, B \subset E$* est définie par

$$A \Delta B = \big(A \backslash B\big) \cup \big(B \backslash A\big) .$$

C'est-à-dire que $(A \cup B) \backslash (A \Delta B) = A \cap B$, et donc, $\mathbf{1}_{A\Delta B} = |\mathbf{1}_A - \mathbf{1}_B|$.

**Théorème A.4.12** *Soit $(E, \mathscr{E}, \mu)$, un espace mesuré. On suppose $\mu$ sigma finie. Soit $\mathscr{A}$, une algèbre sur $E$ satisfaisant (A.18). Soit $\mu^*$, la mesure extérieure définie par (A.19). Soit $\mathscr{E}(\mu^*)$, la classe des ensembles $\mu^*$-mesurables associés. Alors, les assertions suivantes sont vraies.*

*(i) On a $\mathscr{N}_\mu = \{N \subset E : \mu^*(N) = 0\}$ et donc $(E, \mathscr{E}^\mu, \overline{\mu}) = (E, \mathscr{E}(\mu^*), \mu^*)$.*

*(ii) $N \in \mathscr{N}_\mu$ ssi pour tout $\varepsilon > 0$, il existe $A_n \in \mathscr{A}$, $n \in \mathbb{N}$, tels que*

$$\textit{les } A_n \textit{ sont disjoints deux-à-deux,} \quad N \subset \bigcup_{n \in \mathbb{N}} A_n \quad \text{et} \quad \sum_{n \in \mathbb{N}} \mu(A_n) \; < \varepsilon\, .$$

*(iii) Soit $B \in \mathscr{E}^\mu$. Pour tout $\varepsilon > 0$, il existe $A_n \in \mathscr{A}$, $n \in \mathbb{N}$, tels que*

$$\textit{les } A_n \textit{ sont disjoints deux-à-deux,} \quad B \subset \bigcup_{n \in \mathbb{N}} A_n \quad \text{et} \quad \overline{\mu}\Big( \bigcup_{n \in \mathbb{N}} A_n \backslash B \Big) < \varepsilon\, .$$

*(iv) Soit $B \in \mathscr{E}^\mu$ tel que $\mu(B) < \infty$. Alors, pour tout $\varepsilon > 0$, il existe $A \in \mathscr{A}$, tel que $\mu(A \Delta B) < \varepsilon$.*

**Preuve :** (A.20) implique que $\mathscr{N}_\mu \subset \{N \subset E : \mu^*(N) = 0\}$, et le théorème A.4.9 $(iv)$ implique $\{N \subset E : \mu^*(N) = 0\} \subset \mathscr{N}_\mu$. Donc $\mathscr{N}_\mu = \{N \subset E : \mu^*(N) = 0\}$. Donc $\mathscr{E}^\mu \subset \mathscr{E}(\mu^*)$. Par ailleurs, le théorème A.4.9 $(iv)$ et ce qui précède entraînent également que $\mathscr{E}(\mu^*) \subset \mathscr{E}^\mu$, et donc $\mathscr{E}(\mu^*) = \mathscr{E}^\mu$. Comme $\overline{\mu}$ et $\mu^*$ sont deux prolongements de $\mu$ sur $\mathscr{E}^\mu$ et que par le théorème A.2.3 de complétion, il n'en existe qu'un seul, on a $\overline{\mu} = \mu^*$ sur $\mathscr{E}^\mu$, ce qui montre $(i)$.

Montrons $(ii)$ : le point $(i)$ implique que $N$ est un $\mu$-négligeable ssi $\mu^*(N) = 0$. D'après la définition (A.19), pour tout $\varepsilon > 0$, il existe $A'_n \in \mathscr{A}$, $n \in \mathbb{N}$, tels que $N \subset \bigcup_{n \in \mathbb{N}} A'_n$ et $\sum_{n \in \mathbb{N}} \mu(A'_n) < \varepsilon$. On pose ensuite $A_0 = A'_0$ et $A_{n+1} = (\bigcup_{0 \leq k \leq n+1} A'_k) \backslash (\bigcup_{0 \leq k \leq n} A'_k)$. On voit que les $A_n$, $n \in \mathbb{N}$, sont dans $\mathscr{A}$, disjoints deux-à-deux et que $\bigcup_{n \in \mathbb{N}} A_n = \bigcup_{n \in \mathbb{N}} A'_n$. Donc

$$\sum_{n \in \mathbb{N}} \mu(A_n) = \mu\big( \bigcup_{n \in \mathbb{N}} A_n \big) = \mu\big( \bigcup_{n \in \mathbb{N}} A'_n \big) \leq \sum_{n \in \mathbb{N}} \mu(A'_n) < \varepsilon\, .$$

La réciproque est triviale. Cela montre le point $(ii)$.

Montrons le point $(iii)$ : soit $B \in \mathscr{E}^\mu$ et $\varepsilon > 0$. On suppose d'abord que $\overline{\mu}(B) < \infty$. Comme on a montré que $\overline{\mu}(B) = \mu^*(B)$, la définition (A.19) implique qu'il existe $A'_n \in \mathscr{A}$, $n \in \mathbb{N}$ tels que $B \subset \bigcup_{n\in\mathbb{N}} A'_n$ et $\sum \mu(A'_n) \leq \overline{\mu}(B) + \varepsilon$. On raisonne comme précédemment en posant $A_0 = A'_0$ et $A_{n+1} = (\bigcup_{0\leq k\leq n+1} A'_k) \backslash (\bigcup_{0\leq k\leq n} A'_k)$ : on voit que les $A_n$, $n \in \mathbb{N}$, sont dans $\mathscr{A}$, disjoints deux-à-deux, et que $\bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n\in\mathbb{N}} A'_n$. Donc

$$\overline{\mu}(B) \leq \mu\big(\bigcup_{n\in\mathbb{N}} A_n\big) = \mu\big(\bigcup_{n\in\mathbb{N}} A'_n\big) \leq \sum_{n\in\mathbb{N}} \mu(A'_n) < \overline{\mu}(B) + \varepsilon \ .$$

Comme $\overline{\mu}(B) < \infty$, cela implique le point $(iii)$ dans ce cas.

On suppose ensuite que $\overline{\mu}(B) = \infty$. On rappelle que les $E_n \in \mathscr{A}$, $n \in \mathbb{N}$, satisfont (A.18). D'après le cas fini, pour tout $n \in \mathbb{N}$, il existe $A_{n,p} \in \mathscr{A}$, deux-à-deux disjoints tels que

$$\forall n \in \mathbb{N} \ , \quad B \cap E_n \subset \bigcup_{p\in\mathbb{N}} A_{n,p} \quad \text{et} \quad \overline{\mu}\big(\bigcup_{p\in\mathbb{N}} A_{n,p} \backslash (B \cap E_n)\big) < 2^{-n-1}\varepsilon \ .$$

Quitte à remplacer $A_{n,p}$ par $A_{n,p} \cap E_n$, qui est dans $\mathscr{A}$, on peut supposer que $A_{n,p} \subset E_n$ et la famille $(A_{n,p})_{n,p\in\mathbb{N}}$ est alors formée d'ensembles deux-à-deux disjoints ; de plus, $B \subset \bigcup_{n,p\in\mathbb{N}} A_{n,p}$ et

$$\bigcup_{n,p\in\mathbb{N}} A_{n,p} \backslash B \subset \bigcup_{n\in\mathbb{N}} \big(\bigcup_{p\in\mathbb{N}} A_{n,p} \backslash (B \cap E_n)\big) \ .$$

Donc $\overline{\mu}\big(\bigcup_{n,p\in\mathbb{N}} A_{n,p} \backslash B\big) < \sum_{n\in\mathbb{N}} 2^{-n-1}\varepsilon = \varepsilon$, ce qui entraîne le point $(iii)$ lorsque $\overline{\mu}(B) = \infty$.

Montrons le point $(iv)$. On fixe $B \in \mathscr{E}^\mu$ tel que $\overline{\mu}(B) < \infty$. Soit $\varepsilon > 0$ et soient des ensembles $A_n \in \mathscr{A}$, $n \in \mathbb{N}$ qui satisfont les conditions du point $(iii)$ :

$$\text{les } A_n \text{ sont disjoints deux-à-deux,} \quad B \subset \bigcup_{n\in\mathbb{N}} A_n \quad \text{et} \quad \overline{\mu}\Big(\bigcup_{n\in\mathbb{N}} A_n \backslash B\Big) < \varepsilon/2 \ .$$

On pose $B_n = B \cap (E \backslash \bigcup_{0\leq k\leq n} A_k)$. Il est clair que $B_{n+1} \subset B_n \subset B$ et $\bigcap_{n\in\mathbb{N}} B_n = \emptyset$. Comme $\overline{\mu}(B) < \infty$, on a $\lim_n \downarrow \overline{\mu}(B_n) = 0$. Il existe donc $n_0 \in \mathbb{N}$ tel que $\overline{\mu}(B_{n_0}) < \varepsilon/2$. On pose alors $A = \bigcup_{0\leq k\leq n_0} A_k$, On a clairement $A \in \mathscr{A}$, $B_{n_0} = B \cap (E \backslash A)$, donc $\overline{\mu}(B \cap (E\backslash A)) < \varepsilon/2$. De plus $A \cap (E\backslash B) \subset \bigcup_{n\in\mathbb{N}} A_n \backslash B$ et donc $\overline{\mu}(A \cap (E\backslash B) < \varepsilon/2$, on a donc $\overline{\mu}(A \Delta B) = \overline{\mu}(B \cap (E\backslash A)) + \overline{\mu}(A \cap (E\backslash B) < \varepsilon$, ce qui termine la preuve de $(iv)$. ■

**Le théorème de Daniell-Stone.** Jusqu'ici, nous avons d'abord construit des mesures, puis les intégrales de certaines fonctions contre ces mesures, comme expliqué dans le premier chapitre. Cette section expose la méthode contraire qui consiste à construire les mesures à partir des intégrales, vues comme des *fonctionnelles additives* sur certaines classes de fonctions. Ces classes de fonctions doivent, bien entendu, satisfaire quelques hypothèses minimales, comme cela est détaillé dans la définition ci-dessous.

**Définition A.4.13** Soit $E$, un ensemble non-vide. Soit $\mathcal{L}$, un ensemble de fonctions de $E$ dans $\mathbb{R}$. C'est une *classe de Stone* si les propriétés suivantes sont vérifiées.

(a) Si $f, g \in \mathcal{L}$, alors $f + g \in \mathcal{L}$ et $f \wedge g \in \mathcal{L}$.

(b) Si $f \in \mathcal{L}$ et si $c \in \mathbb{R}_+$, alors $cf \in \mathcal{L}$ et $c \wedge f \in \mathcal{L}$

(c) Si $f, g \in \mathcal{L}$ sont telles que $f \leq g$, alors $g - f \in \mathcal{L}$. □

Une "intégrale" se définit alors comme une application linéaire sur la classe de fonction $\mathcal{L}$. Plus précisément, on pose la définition suivante.

**Définition A.4.14** Soit $E$, un ensemble non-vide. Soit $\mathcal{L}$, un ensemble de fonctions de $E$ dans $\mathbb{R}$, qui est classe de Stone. Soit $I : \mathcal{L} \to \mathbb{R}$. C'est une *integrale de Daniell* si les propriétés suivantes sont vérifiées.

(a) Si $f, g \in \mathcal{L}$, alors $I(f + g) = I(f) + I(g)$.

(b) Si $f \in \mathcal{L}$ et si $c \in \mathbb{R}_+$, alors $I(cf) = cI(f)$.

(c) Si $f, g \in \mathcal{L}$ sont telles que $f \leq g$, alors $I(f) \leq I(g)$.

(d) Si $h, h_n \in \mathcal{L}$, $n \in \mathbb{N}$, sont telles que $h_n \leq h_{n+1}$ et $\lim_n h_n = h$ ponctuellement, alors, $\lim_n I(h_n) = I(h)$. □

Soit $(E, \mathscr{E}, \mu)$, un espace mesuré. On voit que $\mathcal{L}^1(E, \mathscr{E}, \mu)$ est une classe de Stone et que $I(f) := \int_E f\, d\mu$ est une intégrale de Daniell, le point (d) étant une conséquence du théorème de converge dominée. Le théorème suivant montre qu'il est toujours possible de représenter ainsi une intégrale de Daniell. On note que dans cet énoncé, l'unicité de la mesure n'est pas garantie en général : cela dépend de la richesse de la classe de Stone sur laquelle elle est définie.

**Théorème A.4.15 (Théorème de Daniell-Stone)** *Soit $E$, un ensemble non-vide. Soit $\mathcal{L}$, un ensemble de fonctions de $E$ dans $\mathbb{R}$, qui est classe de Stone. Soit $I : \mathcal{L} \to \mathbb{R}$, une integrale de Daniell. On note $\mathscr{C}$ la classe de sous-ensembles de $E$ de la forme $\{f > x\}$, avec $f \in \mathcal{L}$ et $x \in \mathbb{R}_+$ et on pose $\mathscr{E} = \sigma(\mathscr{C})$. Alors les assertions suivantes sont vérifiées.*

*(i) Il existe une mesure positive $\mu : \mathscr{E} \to [0, \infty]$ telle que toute fonction $f \in \mathcal{L}$ est $\mu$-intégrable et*

$$\forall f \in \mathcal{L}\, , \quad I(f) = \int_E f\, d\mu\ . \tag{A.21}$$

*(ii) Si $\nu$ est une mesure positive satisfaisant aussi (A.21), alors $\nu(C) = \mu(C)$, pour tout $C \in \mathscr{C}$.*

**Preuve :** on note $\mathcal{L}^+ = \big\{ f \in \mathcal{L} : f \geq 0 \big\}$. Soit $X \subset E$ et soit $f_n \in \mathcal{L}^+$, $n \in \mathbb{N}$. Le symbole

$$(f_n)_{n\in\mathbb{N}} \succeq X$$

signifie que $f_n \leq f_{n+1}$ et $\sup_n f_n \geq \mathbf{1}_X$. On remarque que $\sup_n f_n = \lim_n f_n$ et que cette fonction n'est pas nécessairement dans $\mathcal{L}$, ni même finie. On constate aussi que $(I(f_n))_{n\in\mathbb{N}}$ est une suite croissante, donc $\lim_n I(f_n) = \sup_n I(f_n)$ qui existe dans $[0, \infty]$. On pose alors

$$\mu(X) = \inf \big\{ \lim_{n\to\mathbb{N}} I(f_n)\, ;\ (f_n)_{n\in\mathbb{N}} \succeq X \big\}\ .$$

Montrons d'abord que $\mu : \mathscr{P}(E) \to [0, \infty]$ est une mesure extérieure. On constate tout d'abord que la fonction nulle, notée 0 est dans $\mathcal{L}$ et que $I(0) = 0$. Si on prend $f_n \equiv 0$, on a alors $(f_n)_{n\in\mathbb{N}} \succeq \emptyset$ et donc $\mu(\emptyset) = 0$.

Montrons ensuite la sigma sous-additivité de $\mu$ : soit $X, A_p \subset E$, $p \in \mathbb{N}$, tels que $X \subset \bigcup_{p\in\mathbb{N}} A_p$. S'il existe $p \in \mathbb{N}$, tel que $\mu(A_p) = \infty$, alors trivialement $\mu(X) \leq \sum_{p\in\mathbb{N}} \mu(A_p)$. Supposons que $\mu(A_p) < \infty$, pour tout $p \in \mathbb{N}$. Soit $\varepsilon > 0$. Il existe $(f_n^p)_{n\in\mathbb{N}} \succeq A_p$ tels que $\lim_n I(f_n^p) \leq \mu(A_p) + 2^{-p-1}\varepsilon$. On pose alors $g_n = \sum_{0\leq p\leq n} f_n^p$. Clairement $g_n \in \mathcal{L}^+$ et on vérifie facilement que $(g_n)_{n\in\mathbb{N}} \succeq X$. D'autre part, on a

$$I(g_n) = \sum_{0\leq p\leq n} I(f_n^p) \leq \sum_{p\in\mathbb{N}} \lim_{n\to\infty} I(f_n^p) \leq \varepsilon + \sum_{p\in\mathbb{N}} \mu(A_p).$$

Donc $\mu(X) \leq \lim_n I(g_n) \leq \varepsilon + \sum_{p\in\mathbb{N}} \mu(A_p)$. Comme $\varepsilon$ peut être arbitrairement petit, cela montre que $\mu$ est bien une mesure extérieure.

On note $\mathscr{E}(\mu)$ la classe des ensemble $\mu$-négligeables. On montre ensuite que toute fonction $f \in \mathcal{L}^+$ est $\mathscr{E}(\mu)$-mesurable. D'après le lemme A.4.4, page 279, il suffit de vérifier que pour tout $X \subset E$, et tous $a, b \in \mathbb{R}$ tels que $a < b$, on a

$$\mu(X) \geq \mu\big(X \cap \{f \leq a\}\big) + \mu\big(X \cap \{f \geq b\}\big)\ . \tag{A.22}$$

Le cas $a \leq 0$ est immédiat. On suppose donc que $a > 0$. On pose d'abord

$$h = \tfrac{1}{b-a}\big((b-a) \wedge (f-a)_+\big) = \tfrac{1}{b-a}(b \wedge f - a \wedge f)\ .$$

Le dernier membre implique que $h \in \mathcal{L}^+$ et que $\mathbf{1}_{\{f\geq b\}} \leq h \leq \mathbf{1}_{\{f>a\}}$.

Soit $(g_n)_{n\in\mathbb{N}} \succeq X$. On pose ensuite $f_n = g_n \wedge h$. On a donc $f_n \in \mathcal{L}^+$. On vérifie facilement que $f_n \leq f_{n+1}$ et que $(f_n)_{n\in\mathbb{N}} \succeq X \cap \{f \geq b\}$. Comme $g_n \geq f_n$, on a $g_n - f_n \in \mathcal{L}^+$. On observe ensuite que $g_n - f_n = (g_n - h)\mathbf{1}_{\{g_n \geq h\}}$, ce qui implique que $g_n - f_n \leq g_{n+1} - f_{n+1}$. Si $x \in X \cap \{f \leq a\}$, on a $h(x) = 0$ et $\lim_n (g_n(x) - h(x))\mathbf{1}_{\{g_n\geq h\}}(x) = \lim_n g_n(x) \geq 1$. Cela montre donc que $(g_n - f_n)_{n\in\mathbb{N}} \succeq X \cap \{f \leq a\}$. On a donc

$$\lim_{n\to\infty} I(g_n) \geq \lim_{n\to\infty} I(g_n - f_n) + \lim_{n\to\infty} I(f_n) \geq \mu\big(X \cap \{f \leq a\}\big) + \mu\big(X \cap \{f \geq b\}\big)\ ,$$

ce qui implique (A.22) en prenant l'infimum sur les $(g_n)_{n\in\mathbb{N}} \succeq X$. On a donc montré que toute fonction $f \in \mathcal{L}^+$ est $\mathscr{E}(\mu)$-mesurable.

Par définition de $\mu$, on a ensuite

$$\forall X \subset E,\ \forall f \in \mathcal{L}^+ \text{ tel que } f \geq \mathbf{1}_X, \qquad \mu(X) \leq I(f)\ . \tag{A.23}$$

On montre aussi le résultat suivant.

$$\forall X \subset E,\ \forall f \in \mathcal{L}^+ \text{ tel que } f \leq \mathbf{1}_X, \qquad I(f) \leq \mu(X)\ . \tag{A.24}$$

En effet, soit $(g_n)_{n\in\mathbb{N}} \succeq X$. On pose $f_n = f \wedge g_n$. On a donc $f_n \leq f_{n+1}$ et puisque $f \leq \mathbf{1}_X$, on a $\lim_n f_n = f$. Donc $\lim_n I(f_n) = I(f)$, car $I$ est une intégrale de Daniell. Or $g_n \geq f_n$, donc $I(g_n) \geq I(f_n)$. Par conséquent, $I(f) \leq \lim_n I(g_n)$, ce qui entraîne (A.24) en prenant l'infimum sur $(g_n)_{n\in\mathbb{N}} \succeq X$.

Montrons ensuite que pour toute $f \in \mathcal{L}^+$, on a $I(f) = \int_E f\, d\mu$, l'intégrale ayant bien un sens car on a montré que $f$ est $\mathscr{E}(\mu)$-mesurable et que la restriction de $\mu$ à $\mathscr{E}(\mu)$ est une mesure positive par le théorème A.4.3, page 279.

Pour cela, on fixe$f \in \mathcal{L}^+$ et pour tout $r \in \mathbb{R}_+$, on pose $f_r = r \wedge f \in \mathcal{L}^+$. On fixe $\varepsilon > 0$ et pour tout $n \in \mathbb{N}$, on remarque que $\varepsilon \mathbf{1}_{\{f \geq n\varepsilon\}} \geq f_{(n+1)\varepsilon} - f_{n\varepsilon} = \varepsilon \wedge (f - n\varepsilon)_+ \geq \varepsilon \mathbf{1}_{\{f \geq (n+1)\varepsilon\}}$. Par le point (c) de la définition d'une classe de Stone, $f_{(n+1)\varepsilon} - f_{n\varepsilon} \in \mathcal{L}^+$ et donc, par (A.23) et (A.24)

$$\varepsilon\mu\big(\{f \geq n\varepsilon\}\big) \geq I\big(f_{(n+1)\varepsilon} - f_{n\varepsilon}\big) \geq \varepsilon\mu\big(\{f \geq (n+1)\varepsilon\}\big) .$$

Par ailleurs, on a également

$$\varepsilon\mu\big(\{f \geq n\varepsilon\}\big) \geq \int_E \big(f_{(n+1)\varepsilon} - f_{(n+1)\varepsilon}\big)\, d\mu \geq \varepsilon\mu\big(\{f \geq (n+1)\varepsilon\}\big) .$$

On en déduit donc que

$$I\big(f_{(n+1)\varepsilon} - f_{n\varepsilon}\big) \geq \varepsilon\mu\big(\{f \geq (n+1)\varepsilon\}\big) \geq \int_E \big(f_{(n+2)\varepsilon} - f_{(n+1)\varepsilon}\big)\, d\mu \geq \varepsilon\mu\big(\{f \geq (n+2)\varepsilon\}\big) \geq I\big(f_{(n+3)\varepsilon} - f_{(n+2)\varepsilon}\big).$$

En sommant, il vient $I(f_{n\varepsilon}) \geq \int_E (f_{(n+1)\varepsilon} - f_\varepsilon)\, d\mu \geq I(f_{(n+2)\varepsilon} - f_{2\varepsilon})$. Comme $\lim_n \uparrow f_{n\varepsilon} = f$, le point (d) de la définition l'intégrale de Daniell, et le théorème de convergence monotone pour $\mu$, impliquent $I(f) \geq \int_E (f - f_\varepsilon)\, d\mu \geq I(f - f_{2\varepsilon})$. On remarque ensuite que $\lim \uparrow f - f_\varepsilon = lim \uparrow f - f_{2\varepsilon} = f$ lorsque $\varepsilon \downarrow 0$, et par le point (d) de la définition de l'intégrale de Daniell, on obtient finalement que $I(f) = \int_E f\, d\mu$, pour toute $f \in \mathcal{L}^+$.

Il reste à considérer des $f \in \mathcal{L}$ de signe quelconque : clairement $f_+ = 0 \wedge f$ est dans $\mathcal{L}^+$. Comme $f \leq f_+$, la propriété (c) des classes de Stone implique que $f_+ - f = f_- \in \mathcal{L}^+$. Donc

$$\int_E f_+\, d\mu = I(f_+) = I(f + f_-) = I(f) + I(f_-) = I(f) + \int_E f_-\, d\mu .$$

Cela implique donc que $I(f) = \int_E f\, d\mu$, ce qui termine la preuve de $(i)$.

Pour montrer $(ii)$, on procède comme suit : on fixe $f \in \mathcal{L}$ et $x \in \mathbb{R}_+$ et pour tout $n \in \mathbb{N}$, on pose

$$h_n = 2^n\big((x + 2^{-n}) \wedge f - x \wedge f\big) = 2^n\big(2^{-n} \wedge (f - x)_+\big)$$

La première égalité montre que $h_n \in \mathcal{L}^+$. De plus $\lim_n h_n = \mathbf{1}_{\{f > x\}}$. Enfin, $\mathbf{1}_{\{f > x + 2^{-n}\}} \leq h_n \leq \mathbf{1}_{\{f > x\}}$. Par convergence monotone $\lim_n \mu(\{f > x + 2^{-n}\}) = \mu(\{f > x\})$. Ceci combiné avec l'inégalité précédente et le fait que $I(h_n) = \int_E h_n\, d\mu$, impliquent que $\lim_n I(h_n) = \mu\big(\{f > x\}\big)$, ce qui entraîne $(ii)$. ■

Comme il a été mentionné avant l'énoncé du théorème, l'unicité de la mesure représentant une intégrale de Daniell n'est pas garantie si la classe de Stone est trop pauvre : considérons par exemple un ensemble ayant au moins deux points distincts $x_0$ et $x_1$. On prend $\mathcal{L}$ comme l'ensemble des fonctions $f : E \to \mathbb{R}$ qui sont nulles partout, sauf éventuellement en $x_0$ : c'est clairement une classe de Stone. Pour tout $f \in \mathcal{L}$, on pose $I(f) = f(x_0)$. On voit facilement que $I$ est une intégrale de Daniell et il est clair que les mesures $\delta_{x_0}$ et $\delta_{x_0} + \delta_{x_1}$ représentent toutes les deux $I$.

**Le théorème de Riesz pour les formes linéaires positives.** Rappelons tout d'abord quelques résultats élémentaires sur les espaces métriques séparables et localement compacts. Soit $(E, d)$, un tel espace. On note $C_c(E)$, le $\mathbb{R}$-espace vectoriel des fonctions *continues à support compact* que l'on munit de la norme du supremum notée $\|\cdot\|_\infty$.

**Lemme A.4.16** *Soit $(E, d)$, un espace métrique séparable localement compact. Soit $K \subset E$, un compact. Il existe $g_K \in C_c(E)$ telle que $\mathbf{1}_K \leq g_K \leq 1$.*

**Preuve :** on rappelle le lemme A.3.6, page 278, qui affirme l'existence d'une suite de compacts $(K_n)_{n\in\mathbb{N}}$, tels que $K_n \subset \mathring{K}_{n+1}$ et $E = \bigcup_{n\in\mathbb{N}} \mathring{K}_n = \bigcup_{n\in\mathbb{N}} K_n$. Cela montre en particulier que $K$ est recouvert par les ouverts $\mathring{K}_n$ ; comme $K$ est compact on en extrait un recouvrement fini, c'est-à-dire qu'il existe $n_0$ tel que $K \subset \mathring{K}_{n_0}$. On pose $F = E \backslash \mathring{K}_{n_0}$, qui est fermé. La fonction continue $d(x, F)$ atteint son minimum $r$ sur le compact $K$ en $x_0 \in K$ et $r = d(x_0, K) = \mathrm{dist}(K, F)$. Or si $r = 0$, alors $x_0 \in F$, ce qui est impossible car $K \cap F = \emptyset$. Donc $r = \mathrm{dist}(K, F) > 0$. On pose alors $g_K(x) = 1 \wedge (d(x, F)/r)$, qui est continue, telle que $g_K \geq \mathbf{1}_K$. Par ailleurs le support de $g_K$ est inclus dans $K_{n_0}$. ■

Nous rappelons ici le lemme de Dini.

**Lemme A.4.17 (Lemme de Dini)** *Soit $(E,d)$, un espace métrique localement compact. Soient $f_n \in C_c(E)$, $n \in \mathbb{N}$, telles que pour tout $x \in E$, $0 \leq f_{n+1}(x) \leq f_n(x)$ et $\lim_{n\in\mathbb{N}} f_n(x) = 0$. Alors, $\lim_{n\to\infty} \|f_n\|_\infty = 0$.*

**Preuve :** comme les $f_n$ décroissent, il est clair d'une part qu'elles sont toutes à support dans le même compact $K$ qui est le support de $f_0$, d'autre part, il est clair que les normes $\|f_n\|_\infty$ décroissent. Donc $c := \lim_{n\to\infty}\|f_n\|_\infty$ existe. Raisonnons par l'absurde en supposant que $c>0$. Il existe donc une suite $x_n \in K$, $n\in\mathbb{N}$, telle que $f_n(x_n) \geq c$, pour tout $n\in\mathbb{N}$. Comme $K$ est compact, on extrait de $(x_n)_{n\in\mathbb{N}}$, une suite convergeant vers $x\in K$. Il existe ensuite $n_0\in\mathbb{N}$ tel que $f_{n_0}(x) \leq c/3$ et comme $f_{n_0}$ est continue, il existe $\delta>0$ tel que pour tout $y \in B(x,\delta)$, $f_{n_0}(y) \leq 2c/3$. Comme les $f_n$ décroissent, on a montré que

$$\forall n \geq n_0,\ \forall y \in B(x,\delta)\,, \quad f_n(y) \leq 2c/3\ .$$

Or il y a une infinité d'indices $n \geq n_0$ tels que $x_n \in B(x,\delta)$, et donc tels que $f_n(x_n) \leq 2c/3$, ce qui contredit $f(x_n) \geq c$. On a donc $c = 0$. ■

Le théorème suivant admet des généralisations pour des espaces localement compacts non-séparables.

**Théorème A.4.18 (Théorème de Riesz pour les formes positives)** *Soit $(E,d)$, un espace métrique localement compact et séparable. Soit $C_c(E)$, le $\mathbb{R}$-espace vectoriel des fonctions continues à support compact. Soit $\phi : C_c(E) \to \mathbb{R}$ qui satisfait les conditions suivantes.*

$(a)$ *$\phi$ est linéaire : pour toutes $f,g \in C_c(E)$ et tout $c \in \mathbb{R}$, $\phi(f+cg) = \phi(f) + c\phi(g)$.*

$(b)$ *$\phi$ est positive, c'est-à-dire que pour toute $f \in C_c(E)$ telles que $f \geq 0$, on a $\phi(f) \geq 0$.*

*Alors, il existe une unique mesure positive $\mu : \mathscr{B}(E) \to [0,\infty]$ telle que*

$$\forall f \in C_c(E)\,, \quad \phi(f) = \int_E f\, d\mu\ . \tag{A.25}$$

*De plus, $\mu$ est une mesure de Radon et elle est régulière extérieurement pour les ouverts et intérieurement pour les compacts.*

**Preuve :** $C_c(E)$ est clairement une classe de Stone et $\phi$ satisfait les conditions (a), (b) et (c) de la définition d'une intégrale de Daniell. Montrons qu'elle satisfait également (d) : soient $h, h_n \in C_c(E)$, $n \in \mathbb{N}$ telles que $h_n \leq h_{n+1}$ et $\lim_n h_n = h$ ponctuellement. On voit que les $h_n$ et $h$ sont à support dans le même compact $K$ qui est le support de $h$, par exemple. Il en est de même pour les fonctions $f_n := h - h_n$. Les $f_n$ satisfont les hypothèses du lemme A.4.17 de Dini : on a donc $\lim_{n\to\infty}\|f_n\|_\infty = 0$. Par ailleurs, le lemme A.4.16 montre l'existence de $g_K \in C_c(E)$ telle que $g_K \geq \mathbf{1}_K$. On a donc $0 \leq f_n \leq \|f_n\|_\infty \mathbf{1}_K \leq \|f_n\|_\infty g_K$ et donc $0 \leq \phi(f_n) \leq \|f_n\|_\infty \phi(g_K)$, ce qui implique que $\lim_n \phi(f_n) = 0$, c'est-à-dire $\lim_n \phi(h_n) = \phi(h)$. Cela montre que $\phi$ est une intégrale de Daniell.

On note $\mathscr{C}$ la classe des sous-ensembles de la forme $\{f > x\}$, avec $f \in C_c(E)$ et $x \in \mathbb{R}_+$. Pour tout compact $K$ de $E$, on pose $f = d(\cdot, E\backslash K)$ et on a $K = \{f > 0\}$. Donc, $\mathscr{C}$ contient tous les compacts et $\sigma(\mathscr{C}) = \mathscr{B}(E)$. Le théorème A.4.15 de Daniell-Stone implique alors l'existence d'une mesure positive $\mu : \mathscr{B}(E) \to [0,\infty]$ satisfaisant (A.25). Pour tout compact $K$, le lemme A.4.16 implique l'existence de $g_K \in C_c(E)$ tel que $g_K \geq \mathbf{1}_K$, ce qui implique que $\mu(K) \leq \phi(g_K) < \infty$. Cela montre que $\mu$ est une mesure de Radon.

Si $\nu$ est une autre mesure positive sur $\mathscr{B}(E)$ représentant $\phi$, le point $(ii)$ du théorème A.4.15 de Daniell-Stone implique qu'elle coincide avec $\mu$ sur $\mathscr{C}$ et donc que $\nu(K) = \mu(K)$, pour tout $K$ compact. Or $E$ peut être recouvert par une suite de compacts : le théorème A.1.11 d'unicité du prolongement des mesures s'applique et montre que $\mu = \nu$.

Enfin, le théorème A.3.7, page 278, montre que toute mesure de Radon sur un espace métrique localement compact et séparable, est régulière extérieurement pour les ouverts et intérieurement pour les compacts. ■

**Remarque A.4.19** Le théorème de Riesz implique l'existence de la mesure de Lebesgue sur $\mathbb{R}$, comme prolongement de l'intégrale usuelle des fonctions continues à support compact de $\mathbb{R}$. C'est une autre construction de la mesure de Lebesgue sur $\mathbb{R}$. □

# A.5 Compléments sur la mesurabilité.

### Les espaces mesurables séparables et séparés.

**Définition A.5.1** Soient $(E,\mathscr{E})$ et $(E',\mathscr{E}')$, deux espace mesurables. Ils sont dits *isomorphes* s'il existe une fonction $\phi : E \to E'$ satisfaisant les propriétés suivantes.

(a) $\phi$ est une bijection.

(b) $\phi$ est $(\mathscr{E}, \mathscr{E}')$-mesurable.

(c) La fonction réciproque $\phi^{-1} : E' \to E$ est $(\mathscr{E}', \mathscr{E})$-mesurable.

Une telle fonction $\phi$ est un *isomorphisme d'espaces mesurables* et l'application $B \in \mathscr{E} \mapsto \phi(B) \in \mathscr{E}'$ est une bijection de $\mathscr{E}$ sur $\mathscr{E}'$. □

**Définition A.5.2** Soit $(E, \mathscr{E})$, un espace mesurable. On introduit les notions suivantes.

(a) $(E, \mathscr{E})$ est un *espace mesurable séparable* s'il existe une suite d'ensembles $A_n \in \mathscr{E}, n \in \mathbb{N}$ engendrant $\mathscr{E}$ : $\mathscr{E} = \sigma\big(\{A_n ; n \in \mathbb{N}\}\big)$.

(b) $(E, \mathscr{E})$ est un *espace mesurable séparé* si pour tous $x, y \in E$ distincts, il existe $A \in \mathscr{E}$ tel que $x \in A$ et $y \in E \backslash A$. □

On adopte les conventions suivantes : si $E$ est muni d'une topologie $\mathscr{T}$, on note $\mathscr{B}(E)$ la tribu Borélienne associée. Si $X \subset E$ est un sous-ensemble de $E$, pas nécessairement Borélien, on note $\mathscr{T}_X = \{X \cap U ; U \in \mathscr{T}\}$ la topologie relative sur $X$ et on note $\mathscr{B}(X) = \sigma(\mathscr{T}_X)$ la tribu Borélienne associée. On rappelle le lemme A.1.8 page 270 qui affirme que $\mathscr{B}(X)$ est la tribu Borélienne trace sur $X$, c'est-à-dire

$$\mathscr{B}(X) = \big\{X \cap B \, ; \, B \in \mathscr{B}(E)\big\} .$$

Le théorème suivant donne une description assez générale des espace mesurables séparables et séparés, ainsi que quelques propriétés utiles.

**Théorème A.5.3** *Soit $(E, \mathscr{E})$, un espace séparable et séparé. Alors, les assertions suivantes sont vérifiées.*

$(i)$ *Il existe un pi-système $\mathscr{P} = \{A_n ; n \in \mathbb{N}\}$ dénombrable tel que $\sigma(\mathscr{P}) = \mathscr{E}$.*

$(ii)$ *Avec les notations du $(i)$, pour tout $x \in E$ et tout $n \in \mathbb{N}$, on pose*

$$B_n^x = A_n, \text{ si } x \in A_n \quad \text{et} \quad B_n^x = E \backslash A_n, \text{ si } x \notin A_n.$$

*Alors, $\{x\} = \bigcap_{n \in \mathbb{N}} B_n^x$. Cela implique que $\mathscr{E}$ contienne tous les singletons.*

$(iii)$ *On pose $\Delta = \{(x,x) \, ; \, x \in E\}$, la diagonale de $E \times E$. Alors $\Delta \in \mathscr{E} \otimes \mathscr{E}$.*

$(iv)$ *Il existe $X \subset \mathbb{R}$ (pas nécessairement Borélien) tel que $(E, \mathscr{E})$ soit isomorphe à $(X, \mathscr{B}(X))$.*

$(v)$ *La tribu $\mathscr{E}$ est soit finie, soit en bijection avec $\mathbb{R}$.*

$(vi)$ *$\mathscr{E}$ est la tribu Borélienne d'une topologie séparable et séparée sur $E$.*

**Preuve :** on montre d'abord $(i)$. Comme $\mathscr{E}$ est séparable, il existe une suite d'ensembles $C_n \in \mathscr{E}$, $n \in \mathbb{N}$ engendrant $\mathscr{E}$. On pose $\mathscr{P} = \{E\} \cup \big\{ \bigcap_{n \in S} C_n \, ; \, S \subset \mathbb{N} \text{ et } S \text{ fini}\big\}$. On vérifie facilement que $\mathscr{P}$ est un pi système dénombrable. On a clairement $\sigma(\mathscr{P}) = \mathscr{E}$. On se donne une énumération de $\mathscr{P}$ sous la forme $\mathscr{P} = \{A_n \, ; \, n \in \mathbb{N}\}$, quitte à prendre des $A_n$ vides si $\mathscr{P}$ est fini. Cela prouve $(i)$.

Montrons $(ii)$. Pour tout $x \in \mathscr{E}$, on pose $A_x = \bigcap_{n \in \mathbb{N}} B_n^x$. Il est clair que $x \in A_x$ et que $A_x \in \mathscr{E}$. Supposons que $y \in A_x$. Si $x \in A_n$, alors on a $B_n^x = A_n$, donc $y \in A_n$ et $B_n^y = A_n$. De même si $x \notin A_n$, alors $B_n^x = E \backslash A_n$, donc $y \notin A_n$ et $B_n^y = E \backslash A_n$. Dans les deux cas, si $y \in A_x$, pour tout $n \in \mathbb{N}$, $B_n^x = B_n^y$ donc $A_x = A_y$ et aussi $x \in A_y$.

Soient $x, y \in E$, si $A_x \cap A_y \neq \emptyset$, il existe $z \in A_x \cap A_y$ et ce qui précède montre que $A_z = A_x = A_y$. On a donc montré que pour tous $x, y \in E$, ou bien $A_x \cap A_y = \emptyset$, ou bien $A_x = A_y$.

On fixe ensuite $x \in E$ et $y \in A_x$. Ce qui précède implique que $A_x = A_y$. On pose

$$\mathcal{L} = \big\{B \in \mathscr{E} : \mathbf{1}_B(x) = \mathbf{1}_B(y)\big\} .$$

On vérifie facilement que $\mathcal{L}$ est une classe monotone qui contient $\mathscr{P}$. Par le théorème A.1.10 page 270, de la classe monotone, $\mathcal{L} = \mathscr{E}$. Donc, pour tout $B \in \mathscr{E}$, $\mathbf{1}_B(x) = \mathbf{1}_B(y)$. Si $x \neq y$, comme $E$ est séparé, il existe $A \in \mathscr{E}$, tel que $\mathbf{1}_A(x) = 1$ et $\mathbf{1}_A(y) = 0$, ce qui contradictoire. Donc $x = y$, nécessairement. On a donc montré que $A_x = \{x\}$, ce qui prouve $(ii)$.

Montrons $(iii)$ : soient $x, y \in E$, distincts. On vérifie facilement qu'il existe $n \in \mathbb{N}$ tel que $\mathbf{1}_{A_n}(x) \neq \mathbf{1}_{A_n}(y)$, sinon le $(ii)$ entraîne que $x = y$, ce qui n'est pas le cas. est un ensemble infini d'entiers. On vérifie alors que

$$E^2 \backslash \Delta = \bigcup_{n \in \mathbb{N}} \big(A_n \times (E \backslash A_n)\big) \cup \big((E \backslash A_n) \times A_n\big) ,$$

ce qui montre que $\Delta \in \mathscr{E}^{\otimes 2}$.

Montrons $(iv)$. Pour tout $x \in E$, on pose

$$\phi(x) = \sum_{n \in \mathbb{N}} 3^{-n-1} \mathbf{1}_{A_n}(x) .$$

Il est clair que $\phi : E \to [0, \frac{1}{2}]$ est $\mathscr{E}$-mesurable. On pose $X = \phi(E)$. On voit que $(\mathbf{1}_{A_n}(x))_{n\in\mathbb{N}}$ est le développement triadique de $\phi(x)$. Si $\phi(x) = \phi(y)$, par unicité du développement triadique, $\mathbf{1}_{A_n}(x) = \mathbf{1}_{A_n}(y)$, pour tout $n \in \mathbb{N}$. Cela implique que $A_x = A_y$ et donc $x = y$ par $(ii)$. Cela montre que $\phi : E \to X$ est bijective et $(\mathscr{E}, \mathscr{B}(X))$-mesurable.

Il reste à montrer que la fonction réciproque $\phi^{-1} : X \to E$ est $(\mathscr{B}(X), \mathscr{E})$-mesurable. Pour tout $n \in \mathbb{N}$, on pose $C_n^0 = E \backslash A_n$ et $C_n^1 = A_n$. On fixe $\varepsilon_0, \dots, \varepsilon_n \in \{0, 1\}$ et on pose $C = \bigcap_{0\le k\le n} C_k^{\varepsilon_k}$. Comme $\phi$ est bijective, l'image réciproque par $\phi^{-1}$ de $C$ est $\phi(C)$. On veut montrer que $\phi(C) \in \mathscr{B}(X)$. Si $C$ est vide, c'est trivialement vrai. Supposons $C$ non-vide. Alors, pour tout $x \in C$, on a $\mathbf{1}_{A_k}(x) = \varepsilon_k$, pour tout $0 \le k \le n$. On pose alors $a = \sum_{0\le k\le n} 3^{-k-1}\varepsilon_k$. Ce qui précède montre que

$$\phi(C) \subset [a, a + \tfrac{1}{2}3^{-n-1}] =: I\ . \tag{A.26}$$

Supposons que $b \in I \cap X$. Alors, il existe $x \in E$ tel que $\phi(x) = b$. Par unicité du développement triadique, on a $\mathbf{1}_{A_k}(x) = \varepsilon_k$, pour tout $0 \le k \le n$ et donc $x \in C$ par conséquent $b \in \phi(C)$. Autrement dit $\phi(C) = X \cap I$ et donc $\phi(C) \in \mathscr{B}(X)$.

On définit on pose ensuite $\mathscr{A} = \bigcup_{n\in\mathbb{N}} \big\{ \bigcap_{0\le k\le n} C_k^{\varepsilon_k}\ ;\ \varepsilon_0, \dots, \varepsilon_n \in \{0, 1\}\big\}$. On a montré que pour tout $C \in \mathscr{A}$, $(\phi^{-1})^{-1}(C) \in \mathscr{B}(X)$. Or il est facile de vérifier que $\sigma(\mathscr{A}) = \mathscr{E}$. Cela montre donc que $\phi^{-1}$ est $(\mathscr{B}(X), \mathscr{E})$-mesurable, ce qui prouve $(iii)$.

Montrons $(iv)$ : si $E$ est fini et possède $p$ éléments distincts, alors $\mathscr{E}$ est finie car $\mathscr{E} \subset \mathscr{P}(E)$ et $\mathscr{P}(E)$ compte $2^p$ éléments. Supposons que $E$ soit infini. Il existe une suite $x_n \in E$, $n \in \mathbb{N}$, d'éléments distincts. Puisque $\mathscr{E}$ contient les singletons, pour tout sous-ensemble $S \subset \mathbb{N}$, on a $B_S = \{x_n\ ;\ n \in S\} \in \mathscr{E}$. Donc $\mathrm{Card}\,(\mathscr{E}) \ge \mathrm{Card}\,(\mathscr{P}(\mathbb{N})) = \mathrm{Card}\,(\mathbb{R})$. Mais le point $(iii)$ implique que $\mathrm{Card}\,(\mathscr{E}) \le \mathrm{Card}\,(\mathscr{B}(\mathbb{R})) = \mathrm{Card}\,(\mathbb{R})$ donc $\mathrm{Card}\,(\mathscr{E}) = \mathrm{Card}\,(\mathbb{R})$, ce qui prouve $(iv)$.

Montrons $(v)$. La topologie relative $\mathscr{T}_X = \{X \cap U\ ;\ U \text{ ouvert de } \mathbb{R}\}$ est séparable et séparée. On pose $\mathscr{T} = \{\phi^{-1}(V)\ ;\ V \in \mathscr{T}_X\}$. C'est une topologie sur $E$ : c'est la topologie induite par $\phi$. Il est clair qu'elle est séparable et séparée tout comme $\mathscr{T}_X$. Il est aussi clair que $\sigma(\mathscr{T}) \subset \mathscr{E}$. On a vu précédemment que pour tout $C \in \mathscr{A}$, il existe un intervalle fermé $I$ tel que $\phi^{-1}(X \cap I) = C$. Cela montre que $\mathscr{A} \subset \sigma(\mathscr{T})$, ce qui implique que $\mathscr{E} = \sigma(\mathscr{T})$, ce qui termine la preuve du $(v)$. ■

**Espace mesurables réguliers. Isomorphisme de Borel (énoncé et preuve partielle).** Soit $(E, d)$, un espace métrique séparable et soit $A \in \mathscr{B}(E)$. Alors $(A, d)$ est aussi un espace métrique séparable. On en déduit que $\mathscr{B}(A)$ est séparable et séparé : $(A, \mathscr{B}(A))$ est donc isomorphe à un espace $(X, \mathscr{B}(X))$, avec $X \subset \mathbb{R}$. En général, on ne peut pas en dire beaucoup plus sur $X$, notamment on ne sait pas s'il est possible de choisir $X$ Borélien de $\mathbb{R}$. Lorsque l'espace $(E, d)$ est séparable complet, il est possible de montrer le résultat suivant, beaucoup plus précis, qui est connu sous le nom de théorème d'isomorphisme de Borel.

**Théorème A.5.4 (Théorème d'isomorphisme de Borel)** *Soit $(E, d)$, un espace métrique séparable et complet. Soit un Borélien $B \in \mathscr{B}(E)$. Alors,*

— *ou bien $B$ est dénombrable,*

— *ou bien l'espace mesurable $(B, \mathscr{B}(B))$ est isomorphe à $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$.*

Nous ne prouvons pas ce théorème mais un résultat plus faible, qui est pourtant suffisant pour la plupart des applications. Avant cela, quelques commentaires.

**1)** Les Boréliens des espaces métriques séparables complets satisfont l'hypothèse du continu : ou bien ils sont dénombrables, ou bien ils ont la puissance du continu, ce qui est remarquable car la preuve du théorème d'isomorphisme de Borel ne nécessite pas de supposer l'hypothèse du continu.

**2)** Du point de vue de la théorie de la mesure, un espace $(E, d)$ métrique séparable complet non-dénombrable est le même objet que $\mathbb{R}$. Cela permet de montrer que sur un tel espace il existe une mesure de probabilité $\mu : \mathscr{B}(E) \to [0, 1]$ qui est diffuse. En effet, si $\phi : E \to \mathbb{R}$ est un isomorphisme d'espaces mesurables, on note prend $\mu$ la mesure image par $\phi$ de la mesure $\nu = f\ell$, où $f : \mathbb{R} \to \mathbb{R}_+$ est n'importe quelle fonction Borel-mesurable $\ell$-intégrable telle que $\int_{\mathbb{R}} f\, d\ell = 1$.

**3)** Le théorème d'isomorphisme de Borel implique immédiatement le (et est équivalent au) corollaire suivant.

**Corollaire A.5.5** *Soient $(E, d)$ et $(E', d')$, deux espaces métriques séparables qui sont des Boréliens de leurs complétés. Alors ces espaces sont soit dénombrables, soit en bijection avec $\mathbb{R}$. De plus, $(E, \mathscr{B}(E))$ et $(E, \mathscr{B}(E'))$ sont isomorphes ssi ils sont en bijection.*

On introduit la notion d'espace mesurable régulier.

**Définition A.5.6** Soit $(E, \mathscr{E})$, un espace mesurable. Il est dit *régulier* s'il existe $C \in \mathscr{B}(\mathbb{R})$ tel que $(E, \mathscr{E})$ et $(C, \mathscr{B}(C))$ sont isomorphes. □

Au lieu de prouver le théorème d'isomorphisme, on montre le théorème suivant qui est en général suffisant pour les applications.

**Théorème A.5.7** *Soit $(E, d)$, un espace métrique séparable complet. Soit $A \in \mathscr{B}(E)$. Alors il existe $C \in \mathscr{B}(\mathbb{R})$ tel que $(A, \mathscr{B}(A))$ et $(C, \mathscr{B}(C))$ soient isomorphes. Autrement dit, tout Borélien d'un Polonais muni la tribu Borélienne trace est un espace mesurable régulier.*

**Preuve du théorème A.5.7.** On mentionne d'abord le résultat suivant qui est relativement élémentaire.

**Proposition A.5.8** *Soient $(E_n, d_n)$, $n \in \mathbb{N}$, une suite d'espaces métriques. On pose $E = \prod_{n\in\mathbb{N}} E_n$ et pour tout $x = (x_n)_{n\in\mathbb{N}} \in E$ et tout $y = (y_n)_{n\in\mathbb{N}} \in E$, on définit*

$$d(x, y) = \sum_{n\in\mathbb{N}} 2^{-n-1}\big(1 \wedge d_n(x_n, y_n)\big) \ \in [0, 1].$$

*Alors les assertions suivantes sont vérifiées.*

$(i)$ *$d$ est une distance et la topologie-produit sur $E$ est la topologie de la métrique $d$.*

$(ii)$ *Si les espaces $(E_n, d_n)$, $n \in \mathbb{N}$, sont séparables complets, il en est de même pour $(E, d)$.*

$(iii)$ *Si les espaces $(E_n, d_n)$, $n \in \mathbb{N}$, sont compacts, il en est de même pour $(E, d)$.*

**Preuve :** élémentaire (exercice ou voir un cours de topologie). ■

On introduit les espaces métriques suivants.

• $\mathbf{T} = \{0, 1\}^{\mathbb{N}}$, muni de la topologie-produit. C'est un espace compact métrisable.

• $\mathbf{U} = \mathbb{N}^{\mathbb{N}}$, muni de la topologie-produit. C'est un espace métrisable séparable et complet.

• $\mathbf{J} = [0, 1]^{\mathbb{N}}$, muni de la topologie-produit. C'est un espace compact métrisable.

Pour être explicite, ces notations n'apparaissent que dans les preuves et pas dans les énoncés.

**Proposition A.5.9** *Les assertions suivantes sont vérifiées.*

$(i)$ *Il existe $A \in \mathscr{B}\big(\{0, 1\}^{\mathbb{N}}\big)$ tel que $(A, \mathscr{B}(A))$ soit isomorphe à $([0, 1], \mathscr{B}([0, 1]))$.*

$(ii)$ *Il existe $B \in \mathscr{B}\big(\{0, 1\}^{\mathbb{N}}\big)$ tel que $B \subset A$ et $(B, \mathscr{B}(B))$ soit isomorphe à $\big([0, 1]^{\mathbb{N}}, \mathscr{B}([0, 1]^{\mathbb{N}})\big)$.*

$(iii)$ *Soit $(E, d)$, un espace métrique séparable complet. Il existe $C \in \mathscr{B}\big([0, 1]^{\mathbb{N}}\big)$ tel que $(C, \mathscr{B}(C))$ soit isomorphe à $(E, \mathscr{B}(E))$.*

**Preuve :** on rappelle les notations $\mathbf{T} = \{0, 1\}^{\mathbb{N}}$ et $\mathbf{J} = [0, 1]^{\mathbb{N}}$. On prouve d'abord $(i)$. Soit $A$, l'ensemble des $u = (u_n)_{n\in\mathbb{N}} \in \mathbf{T}$ tels que ou bien $u_n = 0$ pour une infinité de $n$, ou bien $u_n = 1$ pour tout $n$. Il est clair que $\mathbf{T}\backslash A$ est dénombrable, donc $A \in \mathscr{B}(\mathbf{T})$. On pose $\phi(u) = \sum_{n\in\mathbb{N}} 2^{-n-1} u_n$. On vérifie facilement que $\phi : A \to [0, 1]$ est un isomorphisme d'espaces mesurables.

Montrons $(ii)$. On munit $\mathbf{T}^{\mathbb{N}}$ de la topologie-produit, métrisable compacte. Soit $v = ((v_{(p,n)})_{n\in\mathbb{N}})_{p\in\mathbb{N}} \in \mathbf{T}^{\mathbb{N}}$, c'est-à-dire que $v_{(p,n)} \in \{0, 1\}$, pour tous $p, n \in \mathbb{N}$. On se donne une bijection $g : \mathbb{N} \to \mathbb{N}\times\mathbb{N}$ et on définit une suite $\varphi(v) = u = (u_k)_{k\in\mathbb{N}}$ en posant $u_k = v_{g(k)}$. Il est facile de démontrer que $\varphi : \mathbf{T}^{\mathbb{N}} \to \mathbf{T}$ est bijective continue et que $\varphi^{-1}$ est également continue : il suffit de vérifier la continuité coordonnée par coordonnée car les espaces sont munis de la topologie produit. On voit donc que $\varphi$ est un isomorphisme d'espaces mesurables. Il est clair que $A^{\mathbb{N}} \in \mathscr{B}(\mathbf{T}^{\mathbb{N}})$ et par $(i)$, $(\mathbf{J}, \mathscr{B}(\mathbf{J}))$ est isomorphe à $(A^{\mathbb{N}}, \mathscr{B}(A^{\mathbb{N}}))$. On pose $B = \varphi(A^{\mathbb{N}}) \in \mathscr{B}(\mathbf{T})$. On vérifie que $B \subset A$ et ce qui précède montre $(ii)$.

Montrons $(iii)$. On munit $\mathbf{J}$ de la métrique $\delta(u, v) = \sum_{n\in\mathbb{N}} 2^{-n-1}|u_n - v_n|$ si bien que $(\mathbf{J}, \delta)$ est métrique compact. Soit $(E, d)$, métrique séparable complet. Soit $q_n \in E$, $n \in E$, une suite dense dans $E$. On définit $\psi : E \to \mathbf{J}$ par

$$\forall x \in E\,, \quad \psi(x) = \big(1 \wedge d(x, q_n)\big)_{n\in\mathbb{N}}\ .$$

Comme pour tout $n \in E$, la fonction $x \in E \mapsto 1 \wedge d(x, q_n)$ est continue et puisque $\mathbf{J}$ est muni de la topologie-produit, $\psi$ est continue. Supposons que $\psi(x) = \psi(y)$, c'est-à-dire que $1 \wedge d(x, q_n) = 1 \wedge d(y, q_n)$, pour tout $n \in \mathbb{N}$. Comme la suite $(q_n)_{n\in\mathbb{N}}$ est dense dans $E$, il existe une suite d'entiers positifs $(n_k)_{k\in\mathbb{N}}$ croissant vers l'infini telle que $\lim_k d(x, q_{n_k}) = 0$. Donc, $\lim_k d(y, q_{n_k}) = 0$, et on a $x = y$. Cela montre que $\psi$ est injective.

On pose $C = \psi(E)$ et on note $\phi = \psi^{-1} : C \to E$, la réciproque de $\psi$. On montre que $\phi$ est continue comme suit. Soient $x_p \in E$, $p \in \mathbb{N} \cup \{\infty\}$ tels que $\lim_p \delta(\psi(x_p), \psi(x_\infty)) = 0$. L'inégalité $1 \wedge (a + b) \leq 1 \wedge a + 1 \wedge b$, et l'inégalité triangulaire pour $d$, impliquent pour tout $n \in \mathbb{N}$,

$$\begin{aligned} 1 \wedge d(x_p, x_\infty) &\leq 1 \wedge d(x_p, q_n) + 1 \wedge d(x_\infty, q_n) \\ &\leq 2\big(1 \wedge d(x_\infty, q_n)\big) + \big|1 \wedge d(x_p, q_n) - 1 \wedge d(x_\infty, q_n)\big| \end{aligned}$$

$$\leq \quad 2\big(1 \wedge d(x_\infty, q_n)\big) + 2^{n+1}\delta(\psi(x_p), \psi(x_\infty)).$$

Pour tout $n\in\mathbb{N}$, on a donc $\limsup_p 1 \wedge d(x_p, x_\infty) \leq 2\big(1 \wedge d(x_\infty, q_n)\big)$, Comme la suite $(q_n)_{n\in\mathbb{N}}$ est dense dans $E$, il est possible de trouver des $q_n$ arbitrairement proches de $x_\infty$, et la limite supérieure précédente, qui ne dépend pas de $n$, est nulle, ce qui implique que $\lim_p d(x_p, x_\infty) = 0$. Cela montre bien que $\phi : C \to E$ est continue.

Cela montre que $\psi$ est un isomorphisme entre les espaces mesurables $(E, \mathscr{B}(E))$ et $(C, \mathscr{B}(C))$ et que $\psi$ et $\psi^{-1}$ sont continues. Montrons que $C \in \mathscr{B}(\mathbf{J})$. Le fait que $\phi = \psi^{-1}$ soit continue implique que

$$\forall x \in E\,,\ \forall \varepsilon > 0\,,\ \exists \eta > 0\ :\ \Big(u \in C \cap B_\delta\big(\psi(x), \eta\big)\Big) \Longrightarrow \Big(d(\phi(u), x) < \varepsilon\Big)\,, \tag{A.27}$$

Pour tout $p\in\mathbb{N}$, on pose alors

$$U_p = \big\{v \in J : \exists \eta_v \in ]0, 3^{-p}[\ \text{ tel que }\ \forall u, u' \in C \cap B_\delta(v, \eta_v)\,,\ \ d(\phi(u), \phi(u')) < 3^{-p}\big\}\,.$$

Soit $v \in U_p$ et soit $\eta_v$ comme ci-dessus. Soit $v' \in B_\delta(v, \eta_v)$. On pose $\eta_{v'} = \eta_v - \delta(v, v') < 3^{-p}$. Alors, $B_\delta(v', \eta_{v'}) \subset B_\delta(v, \eta_v)$ et donc $v' \in U_p$. Cela montre que $U_p$ est un ouvert de $\mathbf{J}$.

On note $\overline{C}$ l'adhérence de $C$ dans $\mathbf{J}$ et on pose

$$D = \overline{C} \cap \bigcap_{p\in\mathbb{N}} U_p\,.$$

On observe que $D \in \mathscr{B}(\mathbf{J})$, et (A.27) implique que $C \subset D$. Pour terminer la preuve de $(iii)$, il suffit donc de montrer que $D \subset C$.

Soit $v \in D$. Pour tout $p \in \mathbb{N}$, il existe donc $0 < \eta_v(p) < 3^{-p}$ tel que pour tous $u, u' \in C \cap B(v, \eta_v(p))$, on ait $d(\phi(u), \phi(u')) < 3^{-p}$. On se donne une suite décroissante $(r_p)_{p\in\mathbb{N}}$ telle que $0 < r_p < \frac{1}{2}\eta_v(p)$. Comme $v$ est dans l'adhérence de $C$, on construit par récurrence une suite $x_p \in E$, $p \in \mathbb{N}$, telle que $\delta(v, \psi(x_p)) < r_p$. On observe que $\delta(\psi(x_{p+1}), \psi(x_p)) < r_p + r_{p+1} < 2r_p < \eta_v(p)$, et par définition de $U_p$, $d(x_{p+1}, x_p) < 3^{-p}$. Cela implique que $(x_p)_{p\in\mathbb{N}}$ est de Cauchy dans $E$, qui est complet. Il existe donc $x \in E$ tel que $\lim_p d(x, x_p) = 0$. Par continuité de $\psi$, $v = \psi(x)$ et on a bien $v \in C$, ce qui termine la preuve de $(iii)$. ■

**Fin de la preuve du théorème A.5.7.** Soit $(E, d)$, un espace métrique séparable complet. Soit $A \in \mathscr{B}(E)$. Par la proposition A.5.9 $(iii)$, il existe $B \in \mathscr{B}(\mathbb{R})$ et $\phi : E \to B$ qui est un isomorphisme entre les espaces mesurables $(E, \mathscr{B}(E))$ et $(B, \mathscr{B}(B))$. On note $\varphi$ la réciproque de $\phi$ et on pose $C = \phi(A)$. Comme $\varphi$ est $(\mathscr{B}(B), \mathscr{B}(E))$-mesurable et comme $\varphi^{-1}(A) = C$, on en déduit que $C \in \mathscr{B}(B)$, c'est-à-dire qu'il existe un Borélien $D \in \mathscr{B}(\mathbb{R})$ tel que $D \cap B = C$. Comme $B \in \mathscr{B}(\mathbb{R})$, cela entraîne que $C \in \mathscr{B}(\mathbb{R})$. On vérifie ensuite immédiatement que la restriction $\phi_{|A} : A \to C$ est un isomorphisme entre les espaces mesurables $(A, \mathscr{B}(A))$ et $(C, \mathscr{B}(C))$. ■

## A.6 Métrisabilité des espaces localement compacts à base dénombrable.

**Définition A.6.1** (*Espaces normaux*) Soit $E$, un espace muni d'une topologie $\mathscr{T}$.

$(a)$ $E$ est dit *séparé* si pour tous $x, y \in E$ distincts il existe $U \in \mathscr{T}$ tel que $x \in U$ et $y \notin U$.

$(b)$ $E$ satisfait $T_4$ si pour tous fermés disjoints $F$ et $F'$ de $E$, il existe $U, U' \in \mathscr{T}$ tels que $U \cap U' = \emptyset$ et $F \subset U$ et $F' \subset U'$.

$(c)$ $E$ est dit *normal* s'il est séparé et satisfait $T_4$.

$(d)$ $E$ est dit *compact* s'il est séparé et si de tout recouvrement de $E$ par des ouverts, on extrait un recouvrement fini. □

**Lemme A.6.2** *Tout espace compact est normal.*

**Preuve :** soit $(E, \mathscr{T})$ espace topologique ; soient $F, F' \subset E$, deux fermés disjoints. On voit immédiatement que ce sont des compacts (car un recouvrement de $F$ (resp. de $F'$) par des ouverts induit en y ajoutant l'ouvert $E \backslash F$ un recouvrement de $E$, dont on extrait un recouvrement fini). On fixe $x \in F$. Pour tout $y \in F'$, il existe $U_{x,y}, U'_{x,y} \in \mathscr{T}$ disjoints tels que $y \in U'_{x,y}$ et $x \in U_{x,y}$ (puisque $E$ est séparé et que $x$ et $y$ sont nécessairement distincts, $F$ et $F'$ étant disjoints). comme $F' \subset \bigcup_{y\in F'} U'_{x,y}$, il existe $y_1, \ldots, y_n \in F'$ tels que $F' \subset \bigcup_{1\leq k\leq n} U'_{x,y_n}$. On pose $V'_x = \bigcup_{1\leq k\leq n} U'_{x,y_n}$ et $V_x = \bigcap_{1\leq k\leq n} U_{x,y_n}$. On voit que $V_x$ est un ouvert contenant $x$, que $V'_x$ est un ouvert contenant $F'$ et que $V_x \cap V'_x = \emptyset$. Comme $F \subset \bigcup_{x\in F} V_x$, il existe $x_1, \ldots, x_m \in F$ tels que $F \subset \bigcup_{1\leq l\leq m} V_{x_m}$. On pose alors $V = \bigcup_{1\leq l\leq m} V_{x_m}$ et $V' = \bigcap_{1\leq l\leq m} V'_{x_m}$. On constate que $V$ et $V'$ sont des ouverts disjoints tels que $F \subset V$ et $F' \subset V'$. Cela montre que $(E, \mathscr{T})$ satisfait $T_4$ : il est donc normal. ■

**Lemme A.6.3** ***(Urysohn)*** *Soit $(E, \mathscr{T})$ un espace topologique normal. Soient $F, F'$ deux fermés disjoints de $E$. Alors, il existe $f : E \to [0, 1]$ qui est continue et telle que $f(x) = 1$ pour tout $x \in F$ et $f(x) = 0$ pour tout $x \in F'$.*

**Preuve :** soit $(r_n)_{n\in\mathbb{N}}$, une énumération des rationnels de $]0,1[$. On définit récursivement une suite d'ouverts $U_n$ tels que (1) : pour tous $F \subset U_n$ et $\overline{U}_n \subset E\backslash F'$ (où $\overline{A}$ est l'adhérence de $A$, c'est-à-dire l'intersection de tous les fermés contenant $A$) et tels que (2) : pour tous $n,m \in \mathbb{N}$ distinct tels que $r_m < r_n$, on ait $\overline{U}_m \subset U_n$. *En effet*, supposons $U_0,\dots,U_n$ construits : on pose $F_n = F\cup\bigcup_{1\leq k\leq n: r_k<r_{n+1}} \overline{U}_k$ et $F'_n = F'\cup\bigcup_{1\leq k\leq n: r_k>r_{n+1}}(E\backslash U_k)$, qui sont deux fermés disjoints ; comme $E$ est normal, il existe deux ouverts $U_{n+1}$ et $U'$ tels que $F_n \subset U_{n+1}$ et $F'_n \subset U'$ ; on en déduit que $\overline{U}_{n+1} \subset E\backslash U' \subset E\backslash F'_n$. On vérifie que $U_{n+1}$ satisfait bien les conditions voulues.

Pour tout $x\in E$, on pose alors $f(x) = \inf(\{1\} \cup\{r_n; n\in\mathbb{N} : x \in U_n\})$. On remarque aussi que $f(x) = \sup(\{0\} \cup\{r_n; n \in \mathbb{N} : x \notin \overline{U}_n\})$. On voit ensuite que $\{y \in E : f(x) < r\} = \bigcup_{n:r_n<r} U_n$ est ouvert pour tout $r\in[0,1]$. Donc $f$ est semi-continue supérieurement. On montre de même que $\{y\in E : f(x)>r\} = \bigcup_{n:r_n>r} E\backslash\overline{U}_n$ est ouvert ; donc $f$ est semi-continue inférieurement : la fonction $f$ est donc continue. On voit également que $f$ vaut 1 sur $F$ et 0 sur $F'$. ■

On s'intéresse à la *métrisabilité* dont on rappelle la définition. On rappelle que si $(E,d)$ est un espace métrique, la topologie métrique est la collection des unions (quelconques) de boules ouvertes.

**Définition A.6.4** Soit $E$, un espace muni d'une topologie. On dit que la topologie est *métrisable* s'il existe une distance $d$ sur $E$ dont la topologie est $\mathscr{T}$. □

On rappelle que des distances distinctes peuvent générer la même topologie. Certains concepts métriques, comme la complétude, ne sont pas topologiques, c'est-à-dire qu'il ne se sont pas invariants par homéomorphisme. En revanche la métrisabilité l'est, comme l'énonce le lemme suivant qui est élémentaire.

**Lemme A.6.5** *Soient $(E,\mathscr{T})$ et $(E',\mathscr{T}')$, des espaces topologiques. On suppose $(E',\mathscr{T}')$ métrisable.*

- $(i)$ *On suppose que $E \subset E'$ et que $\mathscr{T}$ est la topologie induite sur $E$ par $\mathscr{T}'$ : $\mathscr{T} = \{E\cap U'; U'\in\mathscr{T}'\}$. Alors $(E,\mathscr{T})$ est métrisable.*
- $(ii)$ *On suppose que $E$ et $E'$ sont homéomorphes, c'est-à-dire qu'il existe une bijection $\Phi : E\to E'$ continue dont la réciproque est aussi continue. Alors $(E,\mathscr{T})$ est aussi métrisable.*

**Preuve :** soit $d'$ la distance sur $E'$ générant $\mathscr{T}'$ ; On note $d$ la restriction de $d'$ à $E\times E$. Soit $r\in]0,\infty[$ ; pour tout $x\in E$, on note $B_d(x,r)$ la $d$-boule ouverte de $E$ de centre $x$ et de rayon $r$ et on note $B'_{d'}(x',r)$ la $d'$-boule ouverte de $E'$ de centre $x'$ et de rayon $r$. Soit $x'\in E'$ ; pour tout point $x\in E\cap B'(x',r)$ (s'il en existe) on note $r_x = \frac{1}{2}(r - d'(x,x')) > 0$ et on observe que

$$B_d(x,r_x) = B'_{d'}(x,r_x)\cap E \subset B'_{d'}(x',r)\cap E \; .$$

Donc $E\cap B'_{d'}(x',r) = \bigcup_{x\in E\cap B'd'(x',r)} B_d(x,r_x)$, ce qui entraîne que $E\cap B'_d(x',r)\in\mathscr{T}_d$, où $\mathscr{T}_d$ est la topologie induite par $d$ sur $E$. Comme tout ouvert de $\mathscr{T}$ est la réunion de boules du type $E\cap B'_d(x',r)$, cela montre que $\mathscr{T}\subset\mathscr{T}_d$. Ensuite, il est clair que $B_d(x,r) = E\cap B'_{d'}(x,r)$ et donc toute $d$-boule ouverte appartient à $\mathscr{T}$, ce qui permet de conclure que $\mathscr{T} = \mathscr{T}_d$.

Montrons $(ii)$ : on note $d'$ une distance générant $\mathscr{T}'$. Pour tous $x,y\in E$, on pose $d(x,y) = d'(\Phi(x),\Phi(y))$ ; on vérifie immédiatement qu'il s'agit d'une distance sur $E$ et que $B_d(x,r) = \Phi^{-1}(B'_{d'}(\Phi(x),r))$, ce qui montre que $B_d(x,r)\in\mathscr{T}$, car c'est l'image réciproque d'un ouvert par une application $(\mathscr{T},\mathscr{T}')$-continue. Les $d$-boules ouvertes sont donc des ouverts de $\mathscr{T}$ ce qui implique que $\mathscr{T}_d\subset\mathscr{T}$. Soit $U$ un ouvert de $\mathscr{T}$ ; on note $\Psi$ la réciproque de $\Phi$ et on voit que $U' = \Phi(U) = \Psi^{-1}(U)$ est un ouvert de $\mathscr{T}'$ car il est l'image réciproque par la fonction $(\mathscr{T}',\mathscr{T})$-continue $\Psi$ d'un ouvert de $\mathscr{T}$. Comme $\mathscr{T}'$ est métrisé par $d$, $U' = \bigcup_{i\in I} B'_{d'}(x'_i,r_i)$ et donc $U = \Phi^{-1}(U') = \bigcup_{i\in I}\Phi^{-1}\big(B'_{d'}(x'_i,r_i)\big) = \bigcup_{i\in I} B_d(\Psi(x'_i),r_i)$ et $U\in\mathscr{T}_d$. Cela montre donc que $\mathscr{T} = \mathscr{T}_d$ et donc que $(E,\mathscr{T})$ est métrisable. ■

On s'intéresse plus particulièrement à la métrisabilité des espaces compacts. Le lemme suivant montre que les espaces compacts de cardinal trop grand ne sont pas métrisables.

**Lemme A.6.6** *Soit $(E,\mathscr{T})$, un espace topologique compact. On le suppose métrisable. Alors il admet une suite dense, c'est-à-dire un ensemble $D$ dénombrable tel que $\overline{D} = E$ et par conséquent* $\mathrm{Card}(E)\leq\mathrm{Card}(\mathbb{R})$.

**Preuve :** soit $d$ une distance générant la topologie $\mathscr{T}$. Pour tout $n\in\mathbb{N}$, on extrait du recouvrement d'ouvert $E = \bigcup_{x\in E} B_d(x,2^{-n})$, un recouvrement fini : $E = \bigcup_{1\leq k\leq p_n} B_d(x_k^{(n)},2^{-n})$. Il est facile de vérifier que tout point de $E$ est limite d'une suite à valeurs dans l'ensemble dénombrable $D = \{x_k^{(n)}; 1\leq k\leq p_n, n\in\mathbb{N}\}$ et donc $\overline{D} = E$.

Soit $y_n\in D$, $n\in\mathbb{N}$ telle que $\{y_n; n\in\mathbb{N}\} = D$. Pour tout $x\in E$, il existe un ensemble infini $I_x\subset\mathbb{N}$ tel que si $(n_k)_{k\in\mathbb{N}}$ est une indexation strictement croissante des indices $I_x$, $\lim_{k\to\infty} d(y_{n_k},x) = 0$ ; par conséquent, $x\mapsto I_x$ est une injection de $E$ dans $\mathscr{P}(\mathbb{N})$, la classe des sous-ensembles de $\mathbb{N}$, qui a la puissance du continu. ■

L'espace $[0,1]^{\mathbb{R}_+}$ des fonctions de $\mathbb{R}_+$ dans $[0,1]$ muni de la convergence ponctuelle (c'est-à-dire la topologie produit) est un espace compact de cardinal strictement plus grand que $\mathbb{R}$ : il ne peut pas être métrisé.

**Définition A.6.7** Soit $E$ muni de la topologie $\mathscr{T}$. Une *base* de cette topologie est un sous-ensemble $\mathscr{B} \subset \mathscr{T}$ de cette topologie telle que tout ouvert de $\mathscr{T}$ est union d'ouverts de $\mathscr{B}$. On dit qu'un espace topologique est *complétement séparable* si sa topologie admet une base dénombrable. On dit aussi, plus simplement, que l'espace topologique est à base (topologique) dénombrable. Cela correspond au terme anglais "*second countable*". □

Le lemme suivant donne un exemple d'espace completement séparable (ou à base dénombrable).

**Lemme A.6.8** *Soit $(E, d)$ un espace métrique muni de la topologie métrique $\mathscr{T}_d$. On suppose que $(E, d)$ est séparable en tant qu'espace métrique c'est-à-dire qu'il existe une suite dense. Alors, $(E, \mathscr{T}_d)$ est compétement séparable (à base dénombrable).*

**Preuve :** $\mathscr{T}_d$ est contitué des ensembles qui sont des unions quelconques de boules ouvertes. Soit $x_n \in E$, $n \in \mathbb{N}$, une suite dense dans $E$, c'est-à-dire que tout point de $E$ est limite d'une suite à valeurs dans $\{x_n; n \in \mathbb{N}\}$. Clairement $\mathscr{B} = \{B_d(x_n, q); q \in \mathbb{Q}_+, n \in \mathbb{N}\}$ est une base dénombrable de $\mathscr{T}_d$ (ici $B_d(x, r)$ désigne la boule ouverte de centre $x$ et de rayon $r$). ■

**Théorème A.6.9** *Un espace compact est métrisable si et seulement si sa topologie est à base dénombrable.*

**Preuve :** si $(E, \mathscr{T})$ est un espace compact métrisable, le lemme A.6.6 montre qu'il admet une suite dense et le lemme A.6.8 montre qu'il est complément séparable (qu'il admet une base topologique dénombrable).

Montrons la réciproque : on suppose que $(E, \mathscr{T})$ est compact à base dénombrable ; soient $U_n \in \mathscr{T}$, $n \in \mathbb{N}$, une base de $\mathscr{T}$. On montre d'abord le fait suivant :

$$\forall x \in E,\ \forall U \in \mathscr{T} \text{ tel que } x \in U, \quad \exists n \in \mathbb{N} : x \in U_n \subset \overline{U}_n \subset U \ . \tag{A.28}$$

*En effet,* soit $(n_k)$ la suite strictement croissante d'entiers définie par $\{n_k; k \in \mathbb{N}\} = \{n \in \mathbb{N} : x \in U_n\}$ ; soit $y$ distinct de $x$ ; comme $E$ est séparé (c'est un espace compact), il existe $k$ tel que $x \in U_{n_k}$ et un ouvert $V$ contenant $y$ tels que $U_{n_k} \cap V = \emptyset$, c'est-à-dire que $U_{n_k}$ est inclu dans le fermé $E \backslash V$, ce qui implique que $\overline{U}_{n_k} \subset E \backslash V$ et donc $\overline{U}_{n_k} \cap V = \emptyset$. En particulier, il existe $k$ tel que $y \notin \overline{U}_{n_k}$. Cela montre que $\{x\} = \bigcap_{k \in \mathbb{N}} \overline{U}_{n_k}$. Donc l'intersection de compacts $\bigcap_{k \in \mathbb{N}} (E \backslash U) \cap \overline{U}_{n_k}$ est vide ; en passant au complémentaire et en extrayant un recouvrement fini, cela montre qu'il existe $k_0$ tel que $\bigcap_{0 \leq k \leq k_0} (E \backslash U) \cap \overline{U}_{n_k} = 0$ ; on pose $V = \bigcap_{0 \leq k \leq k_0} U_{n_k}$, qui est un ouvert contenant $x$ tel que $\overline{V} \subset \bigcap_{0 \leq k \leq k_0} \overline{U}_{n_k}$ et donc tel que $(E \backslash U) \cap \overline{V} = \emptyset$, c'est-à-dire $\overline{V} \subset U$. Soit $p$ tel que $U_{n_p} \subset V$ : alors on a bien $x \in U_{n_p} \subset \overline{U}_{n_p} \subset U$, ce qui montre (A.28).

On introduit ensuite $I = \{(m, n) \in \mathbb{N}^2 : \overline{U}_m \cap \overline{U}_n = \emptyset\}$. On voit que

$$\forall x, y \in E \text{ distincts}, \quad \exists (m, n) \in I : \ x \in \overline{U}_m \quad \text{et} \quad y \in \overline{U}_n \tag{A.29}$$

*En effet,* comme $E$ est normal (Lemma A.6.2), il existe $U, U'$ deux ouverts disjoints tels que $x \in U$ et $y \in U'$ ; par (A.28) il existe $m, n \in \mathbb{N}$ tel que $x \in U_m \subset \overline{U}_m \subset U$ et $y \in U_n \subset \overline{U}_n \subset U'$ et on bserve que $\overline{U}_m$ et $\overline{U}_n$ sont nécessairement disjoints, c'est-à-dire que $(m, n) \in I$.

On note $(m_p, n_p)$, $p \in \mathbb{N}$, une énumération de $I$ et pour tout $p \in \mathbb{N}$ on note $f_p$ une fonction continue de $E$ dans $[0, 1]$ valant 1 sur $\overline{U}_{m_p}$ et 0 sur $\overline{U}_{n_p}$, dont l'existence est garantie par le lemme d'Urysohn A.6.3. On voit immédiatement que (A.29) implique que la suite de fonctions $(f_p)_{p \in \mathbb{N}}$ sépare les points de $E$, c'est-à-dire

$$\forall x, y \in E \text{ distincts}, \quad \exists p \in \mathbb{N} : \quad f_p(x) \neq f_p(y) \ . \tag{A.30}$$

On rappelle la proposition A.5.8 (élémentaire) qui affirme que l'espace $\mathbf{J} = [0, 1]^{\mathbb{N}}$ des suites à valeurs dans $[0, 1]$ muni de la distance $\delta(\mathbf{x}, \mathbf{y}) = \sum_{n \in \mathbb{N}} 2^{-n} |x_n - y_n|$ où $\mathbf{x} = (x_n)_{n \in \mathbb{N}}$ et $\mathbf{y} = (y_n)_{n \in \mathbb{N}}$, est un espace métrique compact et que la topologie induite par $\delta$ est la topologie produit, c'est-à-dire que $\delta$-$\lim_{p \to \infty} (x_n^{(p)})_{n \in \mathbb{N}} = (x_n)_{n \in \mathbb{N}}$ ssi pour tout $n \in \mathbb{N}$, $\lim_{p \to \infty} x_n^{(p)} = x_n$ dans $[0, 1]$. On définit alors $\Phi : E \to \mathbf{J}$ par $\Phi(x) = (f_n(x))_{n \in \mathbb{N}}$. Comme chaque $f_n$ est continue, $\Phi$ est continue. Comme les $(f_n)_{n \in \mathbb{N}}$ séparent les points (c'est-à-dire (A.30), $\Phi$ est injective : donc $\Phi$ est un homéomorphisme de $E$ sur $K := \Phi(E)$ (voir le lemme **??**). Comme $\mathbf{J}$ est compact métrisable et que la toplogie de $K$ est celle induite par celle de $\mathbf{J}$, le lemme A.6.5 $(i)$ implique que $K$ est métrisable. Le point $(ii)$ du même lemme A.6.5 implique alors que $E$ est métrisable, ce qui termine la preuve. ■

On s'intéresse à la compactification des espaces localement compacts.

**Définition A.6.10** Soit $E$, un espace muni d'une topologie $\mathscr{T}$. Cette topologie est est dite localement compacte si pour tout $x \in E$ et tout ouvert $U \in \mathscr{T}$ contenant $x$, il existe un ouvert $V$ tel que $x \in V \subset U$ et tel que $\overline{V}$ soit compact. □

**Lemme A.6.11** . *Soit $(E, \mathscr{T})$ un espace topologique localement compact. On le suppose à base dénombrable. Alors il existe une base d'ouverts $(V_n)_{n \in \mathbb{N}}$, tels que $\overline{V}_n$ est compact pour tout $n \in \mathbb{N}$.*

**Preuve :** soit $(V_n)_{n\in\mathbb{N}}$ une base de $\mathscr{T}$. On pose $I=\{n\in\mathbb{N} : \overline{U}_n$ est compact$\}$ et on montre que $(U_n)_{n\in I}$ est une base de $\mathscr{T}$. *En effet,* soit $V$ un ouvert non-vide de $\mathscr{T}$. Soit $x\in V$ ; il existe $U$, un ouvert contenant $x$ tel que $\overline{U}$ compact ; par ailleurs comme $(U_n)_{n\in\mathbb{N}}$ est une base de $\mathscr{T}$, il existe $n_x$ tel que $x\in U_{n_x}\subset U$. On en déduit que l'adhérence de $U_{n_x}$ est compacte, c'est-à-dire que $n_x\in I$. On a donc $V=\bigcup_{x\in V} U_{n_x}$. Cela montre que tout ouvert est réunion d'ouverts $(U_n)_{n\in I}$, qui est donc est une base de $\mathscr{T}$. ■

**Proposition A.6.12** *Soit $(E,\mathscr{T})$ un espace topologique localement séparable à base dénombrable (un espace LCBD). Soit $\partial\notin E$. Il existe une topologie compacte à base dénombrable $\mathscr{T}'$ sur $E_\partial:=E\cup\{\partial\}$ induisant $\mathscr{T}$ sur $E$.*

**Preuve :** si $E$ est déjà compact, la proposition est élémentaire : on ajoute un point isolé. On suppose dans la suite que $E$ n'est pas compact.

On définit $\mathscr{T}'$ comme l'ensemble des unions quelconques d'ensembles de $\mathscr{T}$ et d'ensembles de la forme $\{\partial\}\cup(E\backslash K)$ où $K$ est un compact de $E$. Il est clair que $\emptyset$ et $E_\partial$ appartiennent à $\mathscr{T}'$ et que $\mathscr{T}'$ est stable par union quelconque ; on remarque ensuite que $(E\backslash K_1)\cap(E\backslash K_2)=E\backslash(K_1\cup K_2)$, et si $K_1$ et $K_2$ sont compact, il en est de même de $K_1\cup K_2$ ; combiné à un argument simple, cela implique que $\mathscr{T}'$ est stable par intersection simple. Cela montre donc que $\mathscr{T}'$ est une topologie. Elle est séparée. Il est aussi facile de voir que $\mathscr{T}'$ induit $\mathscr{T}$ sur $E$.

Montrons que $\mathscr{T}'$ est compacte : supposons que $E_\partial=(\bigcup_{i\in I}V_i)\cup\bigcup_{j\in J}(\{\partial\}\cup(E\backslash K_j))$, avec les $K_j$ compacts de $E$ et les $V_i$ ouverts de $E$. L'ensemble d'indices $J$ ne peut être vide ; soit $j_0\in J$ ; on pose $O=\{\partial\}\cup(E\backslash K_{j_0})$ ; on vérifie que $K_{j_0}=E_\partial\backslash O\subset(\bigcup_{i\in I}V_i)\cup\bigcup_{j\in J\backslash\{j_0\}}(E\backslash K_j)$. Comme $K_{j_0}$ est compact et que les ensemble $U_i$ et $E\backslash K_j$ sont des ouverts de $E$, il existe $I'\subset I$ et $J'\subset J\backslash\{j_0\}$, deux sous-ensembles finis d'indices tels que $K_{j_0}\subset(\bigcup_{i\in I'}V_i)\cup\bigcup_{j\in J'}(E\backslash K_j)$ et donc $E_\partial=(\bigcup_{i\in I'}V_i)\cup\bigcup_{j\in J'\cup\{j_0\}}(\{\partial\}\cup(E\backslash K_j))$, ce qui montre bien que $\mathscr{T}'$ est une topologie compacte.

On constate ensuite que tout ouvert $V\in\mathscr{T}'$ est soit un ouvert de $\mathscr{T}$ soit de la forme $U\cup\{\partial\}\cup(E\backslash K)$, avec $K$ compact de $E$ ou bien $K$ vide (on rappelle que les compacts sont, par définition, non-vides). Cela provient immédiatement du fait que $\bigcup_{j\in J}(\{\partial\}\cup(E\backslash K_j))=\{\partial\}\cup(E\backslash K)$ où $K=\bigcap_{j\in J}K_j$ qui est compact s'il n'est pas vide.

Soit $(U_n)_{n\in\mathbb{N}}$ une base dénombrable de $\mathscr{T}$ ; grâce au lemme A.6.11, on peut supposer sans perte de généralité que les $\overline{U}_n$ sont compacts. On montre que $\mathscr{B}=\{U_n,\{\partial\}\cup(E\backslash\overline{U}_n); n\in\mathbb{N}\}$ est une base de $\mathscr{T}'$. D'après ce qui précéde, il suffit de montrer que pour tout compact $K$ de $E$, $\{\partial\}\cup(E\backslash K)$ est une union d'ouverts de $\mathscr{B}$ : comme $K$ est compact dans $E$, il existe $m$ tel que $K\subset U_1\cup\ldots\cup U_m$ ; comme $E$ n'est pas compact, l'ouvert $E\backslash(\overline{U}_1\cup\cup\ldots\cup\overline{U}_m)$ n'est pas vide et il existe $n$ tel que $U_n\subset E\backslash(\overline{U}_1\cup\ldots\cup\overline{U}_m)$ ; comme $U_n$ est inclus dans le fermé $E\backslash(U_1\cup\ldots\cup U_m)$, il en est de même pour $\overline{U}_n$ ; on a donc montré l'existence de $n$ tel que les deux compacts $K$ et $\overline{U}_n$ soient disjoints. Comme $E\backslash(K\cup\overline{U}_n)$ est un ouvert non-vide de $E$, et que les $U_k$ forment une base de $\mathscr{T}$, il existe $I\subset\mathbb{N}$ tel que $E\backslash(K\cup\overline{U}_n)=\bigcup_{k\in I}U_k$. Comme $K$ et $\overline{U}_n$ soient disjoints on en déduit que $\{\partial\}\cup(E\backslash K)=(\{\partial\}\cup(E\backslash\overline{U}_n)\cup\bigcup_{k\in I}U_k$. Cela montre bien que $\mathscr{B}$ est une base de $\mathscr{T}'$. Clairement $\mathscr{B}$ est dénombrable. ■

**Théorème A.6.13** *Soit $(E,\mathscr{T})$, un espace topologique localement compact à base dénombrable (LCBD). Soit $\partial\notin E$. On pose $E_\partial=E\cup\{\partial\}$. Alors il existe une distance $d$ sur $E_\partial$, telle que*

$(a)$ $(E_\partial,d)$ *est un espace métrique compact (donc séparable complet),*

$(b)$ *la topologie induite par $d$ sur $E$ est $\mathscr{T}$.*

**Preuve :** cela suit immédiatement de la proposition A.6.12 et du théorème qui précède et du théoreme A.6.9. ■

## A.7 Ensembles analytiques, preuve du théorème du "Début".

La preuve exposée ici suit en partie celle du livre *Probabilités et Potentiel* de C. Dellacherie et P-A. Meyer et elle requiert l'introduction des ensembles analytiques dans leur généralité.

### Préliminaires.

Commençons par quelques notations standard sur les ensembles.

**Définition A.7.1** Soit $E$, un ensemble non-vide et $\mathscr{P}(E)$, la classe de tous ses sous-ensembles. Soit $\mathscr{R}$, une classe de sous-ensembles de $E$ : $\mathscr{R}\subset\mathscr{P}(E)$. On utilise les notations suivantes.

(a) On note $\mathscr{R}_\sigma$ la classe des unions dénombrables d'ensembles de $\mathscr{R}$ :

$$\mathscr{R}_\sigma=\Big\{\bigcup_{n\in\mathbb{N}}A_n\,;\,A_n\in\mathscr{R},\,n\in\mathbb{N}\Big\}\,.$$

(b) On note $\mathscr{R}_\delta$ la classe des intersections dénombrables d'ensembles de $\mathscr{R}$ :

$$\mathscr{R}_\delta = \Big\{ \bigcap_{n\in\mathbb{N}} A_n \, ; \, A_n \in \mathscr{R}, \, n\in\mathbb{N} \Big\} .$$

(c) On note simplement $\mathscr{R}_{\sigma\delta}$ la classe d'ensemble $(\mathscr{R}_\sigma)_\delta$. De même, on note simplement $\mathscr{R}_{\delta\sigma}$ la classe d'ensemble $(\mathscr{R}_\delta)_\sigma$.

(d) Soit $E'$, un ensemble non-vide et soit $\mathscr{R}' \subset \mathscr{P}(E')$. On note

$$\mathscr{R} \times \mathscr{R}' = \big\{ A\times B \, ; \, A\in\mathscr{R}, \, B\in\mathscr{R}' \big\}$$

qui est la classe produit (à ne pas confondre, lorsque $\mathscr{R}$ et $\mathscr{R}'$ sont des tribus avec la tribu produit).

(e) Soit $S$ un espace topologique. On note $\mathscr{K}(S)$ la classe de ses sous-ensembles compacts. □

Soient $E$ et $E'$, deux ensembles non-vide. Soit $f : E \to E'$, une fonction. Soit $A \subset E$ et soit $B \subset E'$. On rappelle les notations suivantes :

$$f(A) = \{f(x) \, ; \, x\in A\} \quad \text{et} \quad f^{-1}(B) = \{x\in E : f(x)\in B\} .$$

On rappelle également les faits suivants : soient $B_i \subset E'$, $i\in I$, une famille de sous-ensembles de $E'$. Alors

$$f^{-1}\Big( \bigcup_{i\in I} B_i \Big) = \bigcup_{i\in I} f^{-1}(B_i) \quad \text{et} \quad f^{-1}\Big( \bigcap_{i\in I} B_i \Big) = \bigcap_{i\in I} f^{-1}(B_i) . \tag{A.31}$$

On rappelle également que pour toute famille $A_j \subset E$, $j\in J$, de sous-ensembles de $E$, on a

$$f\Big( \bigcup_{j\in j} A_j \Big) = \bigcup_{j\in J} f(A_j) \quad \text{et} \quad f\Big( \bigcap_{j\in J} A_j \Big) \subset \bigcap_{j\in J} f(A_j) , \tag{A.32}$$

*Attention :* l'inclusion dans (A.32) est stricte en général (il suffit de prendre $f$ constante à la valeur $y$, $A_1$ et $A_2$ non-vides et disjoints et on a alors $f(A_1\cap A_2)=\emptyset$, alors que $f(A_1)=f(A_2)=\{y\}$). Les relations (A.31) montrent que les opérations ensemblistes élémentaires commutent avec la pré-image et lorsqu'on veut transporter une structure (tribu, mesure ou topologie par exemple), on observe que c'est la notion de pré-image qui est utilisée : voir la définition d'une fonction mesurable, d'une mesure image, d'une fonction continue, etc. En revanche, le fait que l'image d'une intersection d'ensembles ne coïncide pas avec l'intersection des images a de nombreuses conséquences : l'image d'un ouvert par une fonction continue n'est pas un ouvert, l'image d'un borélien par une fonction mesurable n'est pas nécessairement mesurable, etc.

Les *ensembles analytiques*, introduits plus bas, permettent de comprendre les propriétés de l'image directe d'ensembles mesurables. Une idée importante qui sous-tend les définitions liées aux ensembles analytiques est d'utiliser les ensembles compacts qui se comportent relativement bien par image directe : rappelons par exemple que l'image continue d'un compact est un compact. Aussi avant d'introduire les ensembles analytiques, nous rappelons brièvement quelques propriétés des compacts que nous utilisons dans la suite de cette section.

La propriété essentielle qui est utilisée est celle des "compacts emboîtés" que nous rappelons comme suit : soit $E$, un espace topologique métrisable ; soient $K_n \subset E$, $n\in\mathbb{N}$, des ensembles compacts ; alors

$$\Big( \forall n\in\mathbb{N}, \quad K_0 \cap \ldots \cap K_n \neq \emptyset \Big) \Longrightarrow \Big( \bigcap_{n\in\mathbb{N}} K_n \neq \emptyset \Big) \tag{A.33}$$

Pour exploiter cette propriété pour des ensembles qui ne sont pas nécessairement compacts, on considére des ensembles produits avec un espace compact : le lemme suivant, qui est utilisé dans la suite, illustre cette idée, à la base de la définition des ensembles analytiques.

**Lemme A.7.2** *Soit $E$, un ensemble non-vide. Soit $\mathscr{R} \subset \mathscr{P}(E)$, une classe de sous-ensembles de $E$. On suppose que $\emptyset \in \mathscr{R}$ et que $\mathscr{R}$ est stable par unions finies. Soit $S$, un espace topologique compact métrisable et on rappelle que $\mathscr{K}(S)$ désigne la classe de ses sous-ensembles compacts. On pose*

$$\mathscr{H} = \Big\{ \bigcup_{0\leq p\leq N} K_p \times A_p \; ; \; N\in\mathbb{N} \, , \, \forall p\in\{0,\ldots,N\} \, , \; K_p\in\mathscr{K}(S) \, , \; A_p\in\mathscr{R} \Big\} \tag{A.34}$$

*qui est la classe des unions finies d'ensembles de $\mathscr{K}(S)\times\mathscr{R}$. On note $\pi : (x,y)\in S\times E \mapsto y\in E$, la projection canonique sur $E$. Alors,*

$$\forall B_n \in \mathscr{H} \, , \; n\in\mathbb{N} \; \textit{tels que } B_{n+1} \subset B_n, \qquad \bigcap_{n\in\mathbb{N}} \pi(B_n) = \pi\Big( \bigcap_{n\in\mathbb{N}} B_n \Big) . \tag{A.35}$$

**Preuve :** par (A.32), on a $\pi(\bigcap_{n\in\mathbb{N}} B_n)\subset\bigcap_{n\in\mathbb{N}}\pi(B_n)$. Si $\bigcap_{n\in\mathbb{N}}\pi(B_n)=\emptyset$, alors (A.35) est vérifiée trivialement. On suppose donc $\bigcap_{n\in\mathbb{N}}\pi(B_n)$ non-vide et on fixe $y\in\bigcap_{n\in\mathbb{N}}\pi(B_n)$. Comme $B_n\in\mathscr{H}$, il existe $N_n\in\mathbb{N}$, des compacts $K_{n,p}\in\mathscr{K}(S)$ et des ensembles $A_{n,p}\in\mathscr{R}$, $0\le p\le N_n$ tels que $B_n=\bigcup_{0\le p\le N_n} K_{n,p}\times A_{n,p}$. On fixe ensuite $n\in\mathbb{N}$ et on pose

$$U_n=\Big\{u=(u_0,\dots,u_n)\in\mathbb{N}^{n+1}:\forall j\in\{0,\dots,n\},\ u_j\le N_j\ ,\ y\in\bigcap_{0\le j\le n}A_{j,u_j}\ \text{et}\ \bigcap_{0\le j\le n}K_{j,u_j}\neq\emptyset\Big\}\ .$$

Comme $y\in\pi(B_n)$, il existe $x_n\in S$ tel que $(x_n,y)\in B_n$ si bien qu'il existe $u_n\in\{0,\dots,N_n\}$ tel que $y\in A_{n,u_n}$ et $x_n\in K_{n,u_n}$. Soit $j\in\{0,\dots,n\}$, comme $B_n\subset B_j$, $(x_n,y)\in B_j$ et il existe donc $u_j\in\{0,\dots,N_j\}$ tel que $y\in A_{j,u_j}$ et $x_n\in K_{j,u_j}$. Donc $x_n\in\bigcap_{0\le j\le n}K_{j,u_j}$, qui n'est pas vide et $y\in\bigcap_{0\le j\le n}A_{j,u_j}$. Autrement dit $(u_0,\dots,u_n)\in U_n$.

On a donc montré que pour tout $n\in\mathbb{N}$, $U_n$ n'est pas vide. On pose $U:=\bigcup_{n\in\mathbb{N}}U_n$, qui est donc un ensemble infini. On construit une suite d'entiers $(v_n)_{n\in\mathbb{N}}$ récursivement de la manière suivante : comme $U$ est infini, il existe nécessairement un entier $v_0\in\{0,\dots,N_0\}$ tel qu'une infinité de suites de $U$ commencent par le terme $v_0$ ; supposons construits $v_0,\dots,v_n$ : il existe nécessairement un entier $v_{n+1}\in\{0,\dots,N_{n+1}\}$ tel qu'une infinité de suites de $U$ aient pour premiers termes $(v_0,\dots,v_{n+1})$. La suite $(v_n)_{n\in\mathbb{N}}$ ainsi construite est telle que pour tout $n\in\mathbb{N}$, $(v_0,\dots,v_n)\in U_n$, c'est-à-dire

$$\forall n\in\mathbb{N},\qquad \bigcap_{0\le j\le n}K_{j,u_j}\neq\emptyset\quad\text{et}\quad y\in\bigcap_{0\le j\le n}A_{j,u_j}\ .$$

On a donc $y\in\bigcap_{n\in\mathbb{N}}A_{n,v_n}$ et par le principe (A.33) des compacts emboîtés, $\bigcap_{n\in\mathbb{N}}K_{n,v_n}\neq\emptyset$. On peut donc choisir $x\in\bigcap_{n\in\mathbb{N}}K_{n,v_n}$ et on voit que $(x,y)\in\bigcap_{n\in\mathbb{N}}B_n$ et on a bien $\pi(x,y)=y$. Donc $y\in\pi(\bigcap_{n\in\mathbb{N}}B_n)$. Cela montre que $\bigcap_{n\in\mathbb{N}}\pi(B_n)\subset\pi(\bigcap_{n\in\mathbb{N}}B_n)$, ce qui permet de conclure. ■

On utilise le résultat suivant sur la topologie compacte métrisable.

**Lemme A.7.3** *Soient $(S_n,d_n)$, $n\in\mathbb{N}$, des espaces métriques compacts. On pose $S=\prod_{n\in\mathbb{N}}S_n$ et pour toutes suites $s=(s_n)_{n\in\mathbb{N}}, t=(t_n)_{n\in\mathbb{N}}$ dans $S$, on pose*

$$d(s,t)=\sum_{n\in\mathbb{N}}2^{-n}\big(1\wedge d_n(s_n,t_n)\big)\ .$$

*Alors $(S,d)$ est un espace métrique compact et la topologie engendrée par $d$ est la topologie produit sur $S$. De plus, si pour tout $n\in\mathbb{N}$, $K_n$ est un compact de $S_n$, alors $\prod_{n\in\mathbb{N}}K_n$ est un compact de $S$.*

**Preuve :** il s'agit d'un résultat élémentaire de topologie métrique dont la preuve est laissée en exercice. □

**Ensembles analytiques.** Les ensembles analytiques en général sont introduits de la manière suivante.

**Définition A.7.4** (*Ensembles analytiques*) Soit $E$, un ensemble non-vide. Soit $\mathscr{R}\subset\mathscr{P}(E)$, une classe de sous-ensembles de $E$. On suppose que $\emptyset\in\mathscr{R}$. On rappelle les notation de la définition A.7.1.

• Un sous-ensemble $A\subset E$ est dit $\mathscr{R}$-analytique s'il existe un espace topologique $S$ compact métrisable et un ensemble $B\in(\mathscr{K}(S)\times\mathscr{R})_{\sigma\delta}$ tel que $\pi(B)=A$, où $\pi:(s,x)\in S\times E\mapsto x\in E$ désigne la projection canonique de $S\times E$ sur $E$.

• On note $\Sigma(\mathscr{R})$ la classe des ensembles $\mathscr{R}$-analytiques. □

On observe que dans cette définition, l'espace $S$ dépend de $A$. Comme déjà mentionné, la définition des ensembles analytiques permet en quelque sorte d'importer sur des espaces dépourvus de topologie certaines propriétés compacts.

**Remarque A.7.5** Soit $A\in\Sigma(\mathscr{R})$. Il existe $D\in\mathscr{R}_\sigma$ tel que $A\subset D$. *En effet,* il existe un espace topologique $S$ compact métrisable et il existe $C_{n,m}\times A_{n,m}$, $m,n\in\mathbb{N}$, où $C_{n,m}$ est un compact de $S$ et $A_{n,m}\in\mathscr{R}$ tel que si on pose $B=\bigcap_{n\in\mathbb{N}}\bigcup_{m\in\mathbb{N}}C_{n,m}\times A_{n,m}$, alors $A=\pi(B)$ où $\pi$ désigne la projection canonique de $S\times E$ sur $E$. On a donc pour tout $n_0\in\mathbb{N}$, $B\subset\bigcup_{m\in\mathbb{N}}C_{n_0,m}\times A_{n_0,m}$ et donc

$$A=\pi(B)\subset\ \pi\Big(\bigcup_{m\in\mathbb{N}}C_{n_0,m}\times A_{n_0,m}\Big)=\bigcup_{m\in\mathbb{N}}\pi(C_{n_0,m}\times A_{n_0,m})=\bigcup_{m\in\mathbb{N}}A_{n_0,m}\ \in\mathscr{R}_\sigma\ ,$$

ce qui montre le résultat désiré. □

Dans la suite du paragraphe, nous discutons de quelques propriétés des ensembles analytiques.

**Proposition A.7.6** *Soit $E$, un ensemble non-vide. Soit $\mathscr{R}\subset\mathscr{P}(E)$, une classe de sous-ensembles de $E$. On suppose que $\emptyset\in\mathscr{R}$. Alors, d'une part $\mathscr{R}\subset\Sigma(\mathscr{R})$ et $\Sigma(\mathscr{R})$ est stable par union dénombrable et par intersection dénombrable.*

**Preuve :** soit $A \in \mathscr{R}$. Soit $S$ n'importe quel espace topologique compact métrisable. Il est clair que $S \times A \in (\mathscr{K}(S) \times \mathscr{R})_{\sigma\delta}$. Or la projection sur $E$ de $S \times A$ est $A$. cela montre donc que $A \in \Sigma(\mathscr{R})$ et donc que $\mathscr{R} \subset \Sigma(\mathscr{R})$.

Montrons ensuite que $\Sigma(\mathscr{R})$ est stable par intersection dénombrable. Soient $A_n \in \Sigma(\mathscr{R})$, $n \in \mathbb{N}$. Pour tout $n \in \mathbb{N}$, il existe donc un espace topologique $S_n$, compact métrisable, et $B_n \in (\mathscr{K}(S_n) \times \mathscr{R})_{\sigma\delta}$ tel que $A_n$ est la projection de $B_n$ sur $E$. On écrit $B_n = \bigcap_{k\in\mathbb{N}} \bigcup_{\ell\in\mathbb{N}} K^n_{k,\ell} \times A^n_{k,\ell}$, avec $A^n_{k,\ell} \in \mathscr{R}$ et $K^n_{k,\ell} \in \mathscr{K}(S_n)$. On pose $S := \prod_{n\in\mathbb{N}} S_n$, muni de la topologie produit. Le lemme A.7.3 implique que $S$ est un espace topologique compact métrisable. On pose $C_n = \bigcap_{k\in\mathbb{N}} \bigcup_{\ell\in\mathbb{N}} (K^n_{k,\ell} \times \prod_{q\in\mathbb{N}\setminus\{n\}} S_q) \times A^n_{k,\ell}$ qui appartient clairement à $(\mathscr{K}(S) \times \mathscr{R})_{\sigma\delta}$. Par conséquent, $\bigcap_{n\in\mathbb{N}} C_n \in (\mathscr{K}(S) \times \mathscr{R})_{\sigma\delta}$. On note $\pi$ la projection canonique de $S \times E$ sur $E$ et on vérifie ensuite l'égalité suivante qui, combinée avec l'argument précédent, entraîne immédiatement que $\bigcap_{n\in\mathbb{N}} A_n \in \Sigma(\mathscr{R})$ :

$$\pi\Big(\bigcap_{n\in\mathbb{N}} C_n\Big) = \bigcap_{n\in\mathbb{N}} A_n \tag{A.36}$$

*Preuve :* soit $y \in \bigcap_{n\in\mathbb{N}} A_n$ ; pour tout $n \in \mathbb{N}$, $y \in A_n$ et il existe $x_n \in S_n$ tel que pour tout $k \in \mathbb{N}$,on trouve $\ell_{k,n}$ tel que $x_n \in K^n_{k,\ell_{k,n}}$ et $y \in A^n_{k,\ell_{k,n}}$. Pour tout $n \in \mathbb{N}$ et tout $k \in \mathbb{N}$, on a donc $((x_q)_{q\in\mathbb{N}}, y) \in K^n_{k,\ell_{k,n}} \times A^n_{k,\ell_{k,n}}$ et donc pour tout $n \in \mathbb{N}$, $((x_q)_{q\in\mathbb{N}}, y) \in C_n$, c'est-à-dire $((x_q)_{q\in\mathbb{N}}, y) \in \bigcap_{n\in\mathbb{N}} C_n$. Donc $y \in \pi\big(\bigcap_{n\in\mathbb{N}} B_n\big)$. Cela montre que $\bigcap_{n\in\mathbb{N}} \pi(B_n) = \bigcap_{n\in\mathbb{N}} A_n \subset \pi(\bigcap_{n\in\mathbb{N}} B_n)$. Par (A.32), on en déduit (A.36). □

Montrons la stabilité par union dénombrable : on conserve les mêmes notations. On note $S^o = \coprod_{n\in\mathbb{N}} S_n$, l'union disjointe des ensembles $S_n$, en effectuant les identifications ensemblistes nécessaires pour que $S_n$ soit vu comme un sous-ensemble de $S^o$. On munit $S^o$ de la topologie engendrée par les topologies des $S_n$ : c'est la plus petite topologie contenant les ouverts des $S_n$. On vérifie que $S^o$ est un espace topologie localement compact à base dénombrable. Soit $\partial \notin S^o$ : il existe une distance $d$ sur $S := \{\partial\} \cup S^o$ telle que $(S, d)$ soit un espace métrique compact et telle que la topologie induite sur $S^o$ soit celle d'origine ; cela implique notamment que la topologie d'origine de chaque $S_n$ soit la topologie induite par $d$ sur $S_n$. Un compact de $S_n$ est donc un compact de $S$. On rappelle que $B_n$ s'exprime par comme $B_n = \bigcap_{m\in\mathbb{N}} B_{n,m}$ avec $B_{n,m} \in (\mathscr{K}(S_n) \times \mathscr{R})_\sigma \subset (\mathscr{K}(S) \times \mathscr{R})_\sigma$. On peut voir $B_n$ comme un sous-ensemble de $S \times E$. Comme les $S_n$ sont deux-à-deux disjoints dans $S$, les $B_n$ sont deux-à-deux disjoints dans $S \times E$, ce qui implique que

$$\bigcup_{n\in\mathbb{N}} B_n = \bigcup_{n\in\mathbb{N}} \bigcap_{m\in\mathbb{N}} B_{n,m} = \bigcap_{m\in\mathbb{N}} \bigcup_{n\in\mathbb{N}} B_{n,m} \ .$$

Or $\bigcup_{n\in\mathbb{N}} B_{n,m} \in (\mathscr{K}(S) \times \mathscr{R})_\sigma$. Donc $\bigcup_{n\in\mathbb{N}} B_n \in (\mathscr{K}(S) \times \mathscr{R})_{\sigma\delta}$. On note $\pi$ la projection canonique de $S \times E$ sur $E$ et par (A.32), $\pi(\bigcup_{n\in\mathbb{N}} B_n) = \bigcup_{n\in\mathbb{N}} \pi(B_n) = \bigcup_{n\in\mathbb{N}} A_n$, ce qui montre que $\bigcup_{n\in\mathbb{N}} A_n$ est analytique. ■

**Lemme A.7.7** *Soient $E, E'$ deux ensembles non-vides. Soit $\mathscr{R} \subset \mathscr{P}(E)$, tel que $\emptyset \in \mathscr{R}$. Soit $\mathscr{R}' \subset \mathscr{P}(E')$, tel que $\emptyset \in \mathscr{R}'$. Alors on a*

$$\Sigma(\mathscr{R}) \times \Sigma(\mathscr{R}') \subset \Sigma(\mathscr{R} \times \mathscr{R}') \ .$$

**Preuve :** on vérifie immédiatement que $\Sigma(\mathscr{R}) \times \mathscr{R}' \subset \Sigma(\mathscr{R} \times \mathscr{R}')$. De même, $\mathscr{R} \times \Sigma(\mathscr{R}') \subset \Sigma(\mathscr{R} \times \mathscr{R}')$. La stabilité des ensembles analytiques par union dénombrable (proposition A.7.6) implique que

$$\Sigma(\mathscr{R}) \times \mathscr{R}'_\sigma \subset \Sigma(\mathscr{R} \times \mathscr{R}') \quad \text{et} \quad \mathscr{R}_\sigma \times \Sigma(\mathscr{R}') \subset \Sigma(\mathscr{R} \times \mathscr{R}'). \tag{A.37}$$

Soient $A \in \Sigma(\mathscr{R})$ et $A' \in \Sigma(\mathscr{R}')$. La remarque A.7.5 implique l'existence de $B \in \mathscr{R}_\sigma$ et de $B' \in \mathscr{R}'_\sigma$ tels que $A \subset B$ et $A' \subset B'$. On a donc $A \times B' \in \Sigma(\mathscr{R}) \times \mathscr{R}'_\sigma$ et $B \times A' \in \mathscr{R}_\sigma \times \Sigma(\mathscr{R}')$. Donc par (A.37), on a $A \times B' \in \Sigma(\mathscr{R} \times \mathscr{R}')$ et $B \times A' \in \Sigma(\mathscr{R} \times \mathscr{R}')$. Comme les ensembles analytiques sont stable par intersections dénombrables (proposition A.7.6), on a

$$A \times A' = (A \times B') \cap (B \times A') \in \Sigma(\mathscr{R} \times \mathscr{R}') \ ,$$

ce qui permet de conclure. ■

**Lemme A.7.8** *Soit $(S_0, d_0)$, un espace métrique compact. Soit $E$, un ensemble non-vide et soit $\mathscr{R} \subset \mathscr{P}(E)$, tel que $\emptyset \in \mathscr{R}$. On note $\pi$ la projection canonique de $S_0 \times E$ sur $E$. Alors, pour tout $A \in \Sigma(\mathscr{K}(S_0) \times \mathscr{R})$, $\pi(A) \in \Sigma(\mathscr{R})$.*

**Preuve :** soit $A \in \Sigma(\mathscr{K}(S_0) \times \mathscr{R})$. Il existe $(S, d)$, un espace métrique compact et il existe $B \in (\mathscr{K}(S) \times (\mathscr{K}(S_0) \times \mathscr{R}))_{\sigma\delta}$ tels que $A = \pi'(B)$, où $\pi'$ est la projection canonique de $S \times S_0 \times E$ sur $S_0 \times E$. On note $\pi''$ la projection canonique $S \times S_0 \times E$ sur $E$ : on a $\pi''(B) = \pi(A)$. On observe ensuite que $\mathscr{K}(S) \times \mathscr{K}(S_0) \subset \mathscr{K}(S \times S_0)$, donc $B \in (\mathscr{K}(S \times S_0) \times \mathscr{R})_{\sigma\delta}$, ce qui implique donc que $\pi(A) \in \Sigma(\mathscr{R})$ car $S \times S_0$ est métrisable compact. ■

**Lemme A.7.9** *Soit $E$, un ensemble non-vide et soit $\mathscr{R} \subset \mathscr{P}(E)$, tel que $\emptyset \in \mathscr{R}$. Soit $\mathscr{R} \subset \mathscr{S} \subset \Sigma(\mathscr{R})$. Alors $\Sigma(\mathscr{S}) = \Sigma(\mathscr{R})$.*

**Preuve :** clairement $\Sigma(\mathscr{R}) \subset \Sigma(\mathscr{S}) \subset \Sigma(\Sigma(\mathscr{R}))$. Soit $A \in \Sigma(\Sigma(\mathscr{R}))$. Il existe $S$, un espace compact métrisable et $B \in (\mathscr{K}(S) \times \Sigma(\mathscr{R}))_{\sigma\delta}$ tel que $A$ est la projection de $B$ sur $E$. Or la proposition A.7.6 montre que $\mathscr{K}(S) \subset \Sigma(\mathscr{K}(S))$, le lemme A.7.7 implique que $\Sigma(\mathscr{K}(S)) \times \Sigma(\mathscr{R}) \subset \Sigma(\mathscr{K}(S) \times \mathscr{R})$. Donc $\mathscr{K}(S) \times \Sigma(\mathscr{R}) \subset \Sigma(\mathscr{K}(S) \times \mathscr{R})$. La proposition A.7.6 montre que les ensembles analytiques sont stables par unions et intersections dénombrables. Donc $(\mathscr{K}(S) \times \Sigma(\mathscr{R}))_{\sigma\delta} \subset \Sigma(\mathscr{K}(S) \times \mathscr{R})$. Donc $B \in \Sigma(\mathscr{K}(S) \times \mathscr{R})$. Le lemme A.7.8 implique alors que la projection de $B$ sur $E$, qui est $A$, est un ensemble $\mathscr{R}$-analytique. Donc $\Sigma(\Sigma(\mathscr{R})) = \Sigma(\mathscr{R})$, ce qui implique le résultat voulu. ∎

**Proposition A.7.10** *Soit $E$, un ensemble non-vide et soit $\mathscr{R} \subset \mathscr{P}(E)$, tel que $\emptyset \in \mathscr{R}$. On suppose que*

$$\forall A \in \mathscr{R}, \quad E \backslash A \in \Sigma(\mathscr{R}) .$$

*Alors $\sigma(\mathscr{R}) \subset \Sigma(\mathscr{R})$.*

**Preuve :** on pose $\mathscr{E} = \{A \in \Sigma(\mathscr{R}) : E\backslash A \in \Sigma(\mathscr{R})\}$. Par hypothèse, $\mathscr{R} \subset \mathscr{E}$. Grâce à la proposition A.7.6, on vérifie que $\mathscr{E}$ est une tribu, ce qui permet de conclure. ∎

**Théorème A.7.11** *On note $\mathscr{B}([0,\infty])$ les boréliens de $[0,\infty]$. Soit $(\Omega, \mathscr{F})$, un espace mesurable. Alors*

$$\mathscr{B}([0,\infty]) \otimes \mathscr{F} \subset \Sigma(\mathscr{K}([0,\infty]) \times \mathscr{F}) \tag{A.38}$$

*Soit $\pi : [0,\infty] \times \Omega \to \Omega$ la projection canonique. Alors,*

$$\forall B \in \mathscr{B}([0,\infty]) \otimes \mathscr{F}, \quad \pi(B) \in \Sigma(\mathscr{F}) . \tag{A.39}$$

**Preuve :** on a $\sigma(\mathscr{K}([0,\infty]) \times \mathscr{F}) = \mathscr{B}([0,\infty]) \otimes \mathscr{F}$. Le complémentaire de tout compact de $[0,\infty]$ est une union dénombrable de compacts. Soit $K \times B \in \mathscr{K}([0,\infty]) \times \mathscr{F}$. Donc

$$([0,\infty] \times \Omega) \backslash (K \times B) = (([0,\infty] \backslash K) \times \Omega) \cup ([0,\infty] \times (\Omega \backslash B)) .$$

On en déduit que $([0,\infty] \times \Omega) \backslash (K \times B) \in (\mathscr{K}([0,\infty]) \times \mathscr{F})_\sigma$. Or par la proposition A.7.6 et le lemme A.7.7 on a

$$\mathscr{K}([0,\infty]) \times \Sigma(\mathscr{K}([0,\infty])) \times \Sigma(\mathscr{F}) \subset \Sigma(\mathscr{K}([0,\infty]) \times \mathscr{F})$$

et comme $\Sigma(\mathscr{K}([0,\infty]) \times \mathscr{F})$ est stable par union dénombrable (proposition A.7.6), on en déduit que $(\mathscr{K}([0,\infty]) \times \mathscr{F})_\sigma \subset \Sigma(\mathscr{K}([0,\infty]) \times \mathscr{F})$. Par conséquent les ensembles de $\mathscr{K}([0,\infty]) \times \mathscr{F}$ et leurs complémentaires sont dans $\Sigma(\mathscr{K}([0,\infty]) \times \mathscr{F})$ : la proposition A.7.10 implique que $\sigma(\mathscr{K}([0,\infty]) \times \mathscr{F}) \subset \Sigma(\mathscr{K}([0,\infty]) \times \mathscr{F})$. Or $\sigma(\mathscr{K}([0,\infty]) \times \mathscr{F}) = \mathscr{B}([0,\infty]) \otimes \mathscr{F}$, ce qui implique (A.38).Enfin, (A.39)est une conséquence immédiate du lemme A.7.8 avec $S_0 = [0,\infty]$. ∎

**Capacités de Choquet.** Nous donnons ici la définition d'une capacité de Choquet, qui étend dans un cadre général la notion de capacité newtonienne en théorie du potentiel classique, issue de l'électrostatique.

**Définition A.7.12** (*Capacité de Choquet.*) Soit $E$, un ensemble non-vide et soit $\mathscr{R} \subset \mathscr{P}(E)$, tel que $\emptyset \in \mathscr{R}$ et qui est stable par intersections et unions finies. Soit $I : \mathscr{P}(E) \to [-\infty, \infty]$, une fonction d'ensembles. On dit que c'est une *$\mathscr{R}$-capacité de Choquet* si elle satisfait les conditions suivantes.

(a) $I$ est croissante pour l'inclusion : pour tous $A \subset B \subset E$, on a $I(A) \leq I(B)$.

(b) Pour tous $A_n \subset E$, tels que $A_n \subset A_{n+1}$, $n \in \mathbb{N}$, on a $I\big(\bigcup_{n\in\mathbb{N}} A_n\big) = \sup_{n\in\mathbb{N}} I(A_n)$.

(c) Pour tous $B_n \in \mathscr{R}$ tels que $B_{n+1} \subset B_n$, $n \in \mathbb{N}$, on a $I\big(\bigcap_{n\in\mathbb{N}} B_n\big) = \inf_{n\in\mathbb{N}} I(B_n)$.

Un sous-ensemble $A$ de $E$ tel que

$$I(A) = \sup_{\substack{B \in \mathscr{R}_\delta \\ B \subset A}} I(B) , \tag{A.40}$$

est dit *$I$-capacitable*. □

L'exemple suivant est celui qui est utilisé par la suite.

**Exemple A.7.13** Le lecteur est invité à lire les sous-sections A.4 (page 278) et A.4 (page 283) sur les mesures extérieures associées à des mesures positives et la section A.2, en appendice page 274 sur la complétion d'espaces mesurés. Résumons le contenu de ces sections. Soit $(E, \mathscr{E}, \mu)$, un espace mesuré.

• *Complétion.* L'ensemble des $\mu$-négligeables est défini par $\mathscr{N}_\mu := \big\{ A \subset E : \exists B \in \mathscr{E},\ A \subset B,\ \mu(B) = 0 \big\}$. La $\mu$-complétion de la tribu $\mathscr{E}$ est la tribu $\mathscr{E}^\mu := \sigma\big(\mathscr{E}, \mathscr{N}_\mu\big)$. Le théorème A.2.3, page 274, montre qu'il existe une unique mesure positive $\overline{\mu} : \mathscr{E}^\mu \to [0, \infty]$ telle que pour tout $A \in \mathscr{E}$, $\overline{\mu}(A) = \mu(A)$. Par la suite, on note $\mu$ l'unique extension de $\mu$ à $\mathscr{E}^\mu$, mais pour cette section, nous gardons la notation $\overline{\mu}$ pour cette extension, par souci de clarté.

• *Mesure extérieure associée à $\mu$.* Pour tout $X \subset E$, on pose ensuite

$$\mu^*(X) = \inf \Big\{ \sum_{n \in \mathbb{N}} \mu(A_n)\ ; \quad A_n \in \mathscr{E},\ n \in \mathbb{N}\ :\ X \subset \bigcup_{n \in \mathbb{N}} A_n \Big\} . \tag{A.41}$$

Le théorème A.4.9 de Carathéodory (où l'algèbre $\mathscr{A}$ est la sigma-algèbre $\mathscr{E}$) page 282, implique que $\mu^*$ est une mesure extérieure : c'est-à-dire, selon la définition A.4.1 des mesures extérieures, page 278, $\mu^*(\emptyset) = 0$ et

$$\mu^*(A) \leq \sum_{n \in \mathbb{N}} \mu^*(B_n), \quad A \subset E,\ B_n \subset E\ ,\ n \in \mathbb{N}\ :\ A \subset \bigcup_{n \in \mathbb{N}} B_n\ .$$

On rappelle la définition A.4.2, page 278, des ensembles mesurable $\mu^*$-mesurables : $B \subset E$ est dit $\mu^*$-mesurable si

$$\forall X \subset E, \quad \mu^*(X) = \mu^*(X \cap B) + \mu^*(X \cap (E \backslash B))\ .$$

On note $\mathscr{E}(\mu^*)$ la classe des ensembles mesurable. Le théorème A.4.3, page 279, implique que $\mathscr{E}(\mu^*)$ est une tribu et que la restriction de $\mu^*$ à $\mathscr{E}(\mu^*)$ est une mesure positive. Le théorème A.4.12, page 283, implique que

$$(E, \mathscr{E}^\mu, \overline{\mu}) = (E, \mathscr{E}(\mu^*), \mu^*)\ . \tag{A.42}$$

Par ailleurs, $\mu^*$ est une mesure extérieure régulière, c'est-à-dire que

$$\forall X \subset E,\ \exists B \in \mathscr{E}\ : \quad X \subset B \quad \text{et} \quad \mu^*(X) = \mu^*(B) = \mu(B)\ . \tag{A.43}$$

*En effet,* la définition (A.41) implique l'existence $A_{p,n} \in \mathscr{E}$, $p, n \in \mathbb{N}$ tels que $X \subset \bigcup_{n \in \mathbb{N}} A_{p,n}$ pour tout $p \in \mathbb{N}$ et $\mu^*(X) = \lim_{p \to \infty} \sum_{n \in \mathbb{N}} \mu(A_{p,n})$. Pour tout $p \in \mathbb{N}$, on pose $C_p := \bigcup_{n \in \mathbb{N}} A_{p,n}$, $B_p = C_0 \cap \ldots \cap C_p$ et $B = \bigcap_{p \in \mathbb{N}} C_p$. On a donc $X \subset B \subset B_p \subset C_p$ et donc

$$\mu^*(X) \leq \mu^*(B) = \mu(B) \leq \mu(C_p) \leq \sum_{n \in \mathbb{N}} \mu(A_{p,n}) \ \xrightarrow[p \to \infty]{}\ \mu^*(X)$$

et donc $\mu^*(X) = \mu^*(B) = \mu(B)$, ce qui montre (A.43).

On en déduit le lemme suivant.

**Lemme A.7.14** *Soit $(E, \mathscr{E}, \mu)$, un espace mesuré. Soit $\mu^* : \mathscr{P}(E) \to [0, \infty]$ la mesure extérieure définie par (A.41). Alors $\mu^*$ est une $\mathscr{E}$-capacité de Choquet.*

**Preuve :** $\mu^*$ est croissante pour l'inclusion. Soient $B_n \in \mathscr{E}$ tels que $B_{n+1} \subset B_n$, $n \in \mathbb{N}$. Alors $B := \bigcap_{n \in \mathbb{N}} B_n \in \mathscr{E}$. Comme $\mu^*$ restreinte à $\mathscr{E}$ est la mesure $\mu$, on a bien $\mu^*(B) = \mu(B) = \inf_{n \in \mathbb{N}} \mu(B_n) = \inf_{n \in \mathbb{N}} \mu^*(B_n)$, ce qui est propriété (c) des capacité de Choquet.

Montrons la propriété (b) : soient $X_n \subset E$, tels que $X_n \subset X_{n+1}$, $n \in \mathbb{N}$. On pose $X := \bigcup_{n \in \mathbb{N}} X_n$. Comme $\mu^*$ est croissante pour l'inclusion, $\mu^*(X_n) \leq \mu^*(X_{n+1}) \leq \mu^*(X)$ et on a donc $\sup_{n \in \mathbb{N}} \mu^*(X_n) \leq \mu^*(X)$. Comme $\mu^*$ est régulière (c'est-à-dire que $\mu^*$ satisfait (A.43) il existe $B_n \in \mathscr{E}$, $n \in \mathbb{N}$ tels que $\mu^*(X_n) = \mu(B_n)$ et $X_n \subset B_n$. On pose $A_n = \bigcap_{p \geq n} B_p$. Comme $\mathscr{E}$ est une tribu $A_n \in \mathscr{E}$. Par ailleurs, on a $X_n \subset A_n \subset B_n$. On pose $A := \bigcup_{n \in \mathbb{N}} A_n$. On a clairement $X \subset A$. Comme $\mu^*$ est croissante pour l'inclusion, on en déduit que $\mu^*(X_n) = \mu^*(A_n) = \mu(A_n)$. Les propriété s élémentaires des mesures positives impliques que $\sup_{n \in \mathbb{N}} \mu(A_n) = \mu(A) \geq \mu^*(X)$, et donc $\sup_{n \in \mathbb{N}} \mu^*(X_n) \geq \mu^*(X)$. On a donc donc $\sup_{n \in \mathbb{N}} \mu^*(X_n) = \mu^*(X)$. Cela termine la preuve du fait que $\mu^*$ est une capacité de Choquet. ■

Ce lemme montre qu'une capacité de Choquet généralise en quelques sorte la notion de mesure extérieure. □

Nous montrons ensuite quelques propriétés générales des capacités de Choquet et des ensembles capacitables.

**Lemme A.7.15** *Soit $E$, un ensemble non-vide et soit $\mathscr{R} \subset \mathscr{P}(E)$, tel que $\emptyset \in \mathscr{R}$ et qui est stable par intersections et unions finies. Soit $I : \mathscr{P}(E) \to [-\infty, \infty]$, une $\mathscr{R}$-capacité de Choquet. Alors les ensembles de $\mathscr{R}_{\sigma\delta}$ sont $I$-capacitables.*

**Preuve :** soit $A \in \mathscr{R}_{\sigma\delta}$. Si $I(A) = -\infty$, alors il est clairement $I$-capacitable. Supposons que $I(A) > -\infty$. Alors il existe $A_n \in \mathscr{R}_\sigma$, $n \in \mathbb{N}$ tels que $A = \bigcap_{n\in\mathbb{N}} A_n$. Par ailleurs, pour tout $n \in \mathbb{N}$, il existe $A'_{n,p}$, $p \in \mathbb{N}$ tels que $A_n = \bigcup_{p\in\mathbb{N}} A'_{n,p}$. Si on pose $A_{n,p} := A'_{n,0} \cup \ldots \cup A'_{n,p}$, on a d'une part $A_{n,p} \subset A_{n,p+1}$, $A_n = \bigcup_{p\in\mathbb{N}} A_{n,p}$ et puisque $\mathscr{R}$ est stable par unions finies, on a $A_{n,p} \in \mathscr{R}$, pour tout $n, p \in \mathbb{N}$. on a donc

$$A_{n,p} \in \mathscr{R},\ A_{n,p} \subset A_{n,p+1},\ A_n = \bigcup_{p\in\mathbb{N}} A_{n,p} \quad \text{et} \quad A = \bigcap_{n\in\mathbb{N}} A_n \ .$$

On fixe $a \in ]-\infty, I(A)[$. On montre par récurrence le résultat suivant :

$$\exists B_n \in \mathscr{R},\ n\in\mathbb{N} \ : \quad B_n \subset A_n \quad \text{et} \quad I(C_n) > a \quad \text{où} \quad C_n := A \cap B_0 \cap \ldots \cap B_n \ . \tag{A.44}$$

Construisons d'abord $B_0$. Comme $A \subset A_0$, on a $A \cap A_0 = A$ et donc $I(A) = I(A \cap A_0) = \sup_{p\in\mathbb{N}} I(A \cap A_{0,p})$, par la propriété (b) des capacités de Choquet. Il existe donc $p_0 \in \mathbb{N}$ tel que $I(A \cap A_{0,p_0}) > a$. On peut donc poser $B_0 := A_{0,p_0}$. Supposons construits $B_0, \ldots, B_n$ satisfaisant (A.44). On a donc $C_n \cap A_{n+1} = C_n$ et $I(C_n) = I(C_n \cap A_{n+1}) = \sup_{p\in\mathbb{N}} I(C_n \cap A_{n+1,p})$, par la propriété (b) des capacités de Choquet. Il existe donc $p_{n+1} \in \mathbb{N}$ tel que $I(C_n \cap A_{n+1,p_n}) > a$ et pn peut poser $B_{n+1} := A_{n+1,p_n}$. Cela montre (A.44) par récurrence.

On pose ensuite $B'_n := B_0 \cap \ldots \cap B_n$ et $B := \bigcap_{n\in\mathbb{N}} B_n$. On a bien $B \in \mathscr{R}_\delta$. Comme $B_n \subset A_n$, on a $B = \bigcap_{n\in\mathbb{N}} B_n \subset \bigcap_{n\in\mathbb{N}} A_n = A$. Comme $C_n \subset B'_n$, on a $I(B'_n) > a$ ; comme $\mathscr{R}$ est stable par intersections finies on a également $B'_n \in \mathscr{R}$, $n \in \mathbb{N}$ et la propriété (c) des capacités de Choquet implique que $I(B) = \inf_{n\in\mathbb{N}} I(B'_n) \geq a$. On a montré que pour tout $A \in \mathscr{R}_{\sigma\delta}$, et tout $a \in ]-\infty, I(A)[$ il existe $B \in \mathscr{R}_\sigma$ tel que $B \subset A$ et $I(B) \geq a$. Cela entraîne bien (A.40), c'est-à-dire cela montre que $A$ est $I$-capacitable. ∎

On rappelle ensuite le lemme A.7.2 qui est utilisé dans la preuve du lemme suivant.

**Lemme A.7.16** *Soit $E$, un ensemble non-vide. Soit $\mathscr{R} \subset \mathscr{P}(E)$, une classe de sous-ensembles de $E$. On suppose que $\emptyset \in \mathscr{R}$ et que $\mathscr{R}$ est stable par unions et intersection finies. Soit $I : \mathscr{P}(E) \to [-\infty, \infty]$, une $\mathscr{R}$-capacité de Choquet.*

*Soit $S$, un espace topologique compact métrisable et on rappelle que $\mathscr{K}(S)$ désigne la classe des sous-ensembles compacts. On note $\mathscr{H}$ la classe des unions finies d'ensembles de $\mathscr{K}(S) \times \mathscr{R}$ comme définie en (A.34). On vérifie que $\mathscr{H}$ est stable par unions et intersections finies. On note $\pi : (x, y) \in S \times E \mapsto y \in E$, la projection canonique sur $E$.*

*On définit $J : \mathscr{P}(S \times E) \to [-\infty, \infty]$, une fonction d'ensembles sur $S \times E$ en posant*

$$\forall B \subset S \times E, \quad J(B) = I(\pi(B)) \ . \tag{A.45}$$

*Alors $J$ est une $\mathscr{H}$-capacité sur $S \times E$.*

**Preuve :** il est clair que $J$ est croissante. Soient $A_n \subset A_{n+1} \subset S \times E$, $n \in \mathbb{N}$. Par la première relation de (A.32), on a $\pi(\bigcup_{n\in\mathbb{N}} A_n) = \bigcup_{n\in\mathbb{N}} \pi(A_n)$ et la propriété (b) des capacités appliquées à $I$ entraîne que $J(\bigcup_{n\in\mathbb{N}} A_n) = \sup_{n\in\mathbb{N}} J(A_n)$, c'est-à-dire que $J$ satisfait également la propriété (b) des capacités. Soient $B_n \in \mathscr{H}$, $n \in \mathbb{N}$ tels que $B_{n+1} \subset B_n$. Comme $B_n$ est une union finie d'ensembles dans $\mathscr{K}(S) \times \mathscr{R}$ et puisque $\mathscr{R}$ est stable par unions finies, il est clair, par la première relation de (A.32), que $\pi(B_n) \in \mathscr{R}$. La propriété (c) des capacités appliquée à $I$ implique alors que $\inf_{n\in\mathbb{N}} J(B_n) = \inf_{n\in\mathbb{N}} I(\pi(B_n)) = I(\bigcap_{n\in\mathbb{N}} \pi(B_n))$. Or le lemme A.7.2 implique par ailleurs que $\bigcap_{n\in\mathbb{N}} \pi(B_n) = \pi(\bigcap_{n\in\mathbb{N}} B_n)$ et on a donc $\inf_{n\in\mathbb{N}} J(B_n) = I(\pi(\bigcap_{n\in\mathbb{N}} B_n)) = J(\bigcap_{n\in\mathbb{N}} B_n)$. Cela montre que $J$ satisfait la propriété (c) des capacités et cela permet de conclure. ∎

On peut maintenant montrer le principal théorème sur les capacités de Choquet, à savoir le théorème de Choquet.

**Théorème A.7.17** *Soit $E$, un ensemble non-vide. Soit $\mathscr{R} \subset \mathscr{P}(E)$, une classe de sous-ensembles de $E$. On suppose que $\emptyset \in \mathscr{R}$ et que $\mathscr{R}$ est stable par unions et intersection finies. Soit $I : \mathscr{P}(E) \to [-\infty, \infty]$, une $\mathscr{R}$-capacité de Choquet. Alors les ensembles $\mathscr{R}$-analytiques sont $I$-capacitables :*

$$\forall A \in \Sigma(\mathscr{R}), \quad I(A) = \sup_{\substack{B \in \mathscr{R}_\delta \\ B \subset A}} I(B) \ .$$

**Preuve :** soit $A \in \Sigma(\mathscr{R})$. Par définition, il existe un espace topologique $S$, compact métrisable, et il existe $B \in (\mathscr{K}(S) \times \mathscr{R})_{\sigma\delta}$ tel que $\pi(B) = A$, où $\pi$ désigne la projection canonique de $S \times E$ sur $E$. On note $\mathscr{H}$ la classe des unions dénombrables d'ensembles de $\mathscr{K}(S) \times E$. Il est clair que $\mathscr{H}_{\sigma\delta} = (\mathscr{K}(S) \times \mathscr{R})_{\sigma\delta}$ et donc $B \in \mathscr{H}_{\sigma\delta}$. On définit la fonction d'ensembles $J$ sur $S \times E$ par (A.45). Le lemme A.7.16 imlique que $J$ est une $\mathscr{H}$-capacité. Le lemme A.7.15 appliqué à $J$ et $\mathscr{H}$, implique que tout ensemble de $\mathscr{H}_{\sigma\delta}$ est $J$-capacitable : en particulier $B$ est $J$-capacitable. Donc pour tout $\varepsilon > 0$, il existe $D \in \mathscr{H}_\delta$ tel que $D \subset B$ et $J(D) \geq J(B) - \varepsilon$, c'est-à-dire $I(\pi(D)) \geq I(A) - \varepsilon$ car par définition $J(B) = I(\pi(D))$ et $J(B) = I(\pi(B)) = I(A)$.

Puisque $D \in \mathscr{H}_\delta$, il existe $D^*_n \in \mathscr{H}$, $n \in \mathbb{N}$, tels que $D = \bigcap_{n\in\mathbb{N}} D^*_n$. On pose alors $D_n := D^*_0 \cap \ldots \cap D^*_n$. Puisque $\mathscr{H}$ est stable par intersection finies (car $\mathscr{R}$ l'est, une intersection de compacts est compacte et une intersection d'ensembles produits est un ensemble produit), on a $D_n \in \mathscr{H}$, $D_{n+1} \subset D_n$ et $\bigcap_{n\in\mathbb{N}} D_n = D$. D'une part, le lemme A.7.2 appliqué aux $D_n$ implique donc que

$\pi(D)=\pi(\bigcap_{n\in\mathbb{N}} D_n)=\bigcap_{n\in\mathbb{N}}\pi(D_n)$. D'autre par comme $D_n \in \mathscr{H}$, il s'écrit $D_n=\bigcup_{1\leq p\leq N} K_p\times A_p$, avec $K_p$ compact de $S$ et $A_p\in\mathscr{R}$. Donc $\pi(D_n)=\bigcup_{0\leq n\leq N} A_p$, ce qui montre que $\pi(D_n)\in\mathscr{R}$ car $\mathscr{R}$ est stable par unions finies. Comme $\pi(D)=\bigcap_{n\in\mathbb{N}}\pi(D_n)$, et $\pi(D_n)\in\mathscr{R}$, on en déduit que $\pi(D)\in\mathscr{R}_\delta$. Par ailleurs comme $D\subset B$, on a $\pi(D)\subset\pi(B)=A$.

En résumé, pour tout $\varepsilon>0$, on a trouvé $\pi(D)\in\mathscr{R}_\delta$ tel que $\pi(D)\subset A$ tel que $I(\pi(D))\geq I(A)-\varepsilon$, ce qui implique que $A$ est $I$-capacitable. ■

On déduit du théorème de Choquet le théorème suivant montrant que les ensembles analytiques sont dans la tribu complétée par n'importe quelle mesure sigma-finie.

**Théorème A.7.18 (Mesurabilité universelle des ensembles analytiques)** *Soit $(E,\mathscr{E},\mu)$, un espace mesuré. On suppose $\mu$ finie. On note $\mathscr{E}^\mu$, la tribu $\mathscr{E}$ complétée par les $\mu$-négligeables. Alors, les ensembles $\mathscr{E}$-analytiques sont dans $\mathscr{E}^\mu$ :*

$$\Sigma(\mathscr{E})\subset\mathscr{E}^\mu\ .$$

**Preuve :** soit $A\in\Sigma(\mathscr{E})$. On note $\mu^*$ la mesure extérieure associée à $\mu$ comme définie par (A.41). C'est une mesure extérieure régulière, comme montré au (A.43) : il existe donc $B\in\mathscr{E}$ tel que $A\subset B$ et $\mu^*(A)=\mu(B)$. Le lemme A.7.14 montre que $\mu^*$ est une $\mathscr{E}$-capacité de Choquet. Le théorème A.7.17 de Choquet implique alors que $A$ est $\mu^*$-capacitable : il existe donc $C_n\in\mathscr{E}_\delta$, $n\in\mathbb{N}$ tels que $C_n\subset A$ et $\mu^*(C_n)+2^{-n}\geq\mu^*(A)$. Or d'une part, comme $\mathscr{E}$ est une tribu, elle est stable par intersections dénombrables, donc $\mathscr{E}_\delta=\mathscr{E}$, et donc $C_n\in\mathscr{E}$. D'autre part comme $\mu^*$ coïncide avec $\mu$ sur $\mathscr{E}$, on a $\mu^*(C_n)=\mu(C_n)$, et donc $\mu(C_n)+2^{-n}\geq\mu^*(A)$ pour tout $n\in\mathbb{N}$. On pose alors $C:=\bigcup_{n\in\mathbb{N}} C_n$. On a donc $C\in\mathscr{E}$ et $C\subset A$. Comme $C_n\subset C$, on a $\mu(C)+2^{-n}\geq\mu(C_n)+2^{-n}\geq\mu^*(A)$, pour tout $n\in\mathbb{N}$. Donc $\mu(C)\geq\mu^*(A)$. Or comme $\mu^*$ est croissante pour l'inclusion, comme $\mu^*$ coïncide avec $\mu$ sur $\mathscr{E}$ et comme $C\subset A$, on a $\mu(C)=\mu^*(C)\leq\mu^*(A)$. Cela montre finalement que $\mu(C)=\mu^*(A)$.

Pour résumer, on a montré qu'il existe $C,B\in\mathscr{E}$ tels que $C\subset A\subset B$ tels que $\mu(C)=\mu^*(A)=\mu(B)$. On pose $N=A\backslash C$. On a clairement $N\subset B\backslash C$. Or $B\backslash C\in\mathscr{E}$ et comme $\mu$ est finie, $\mu(B\backslash C)=\mu(B)-\mu(C)=0$. Donc $N$ est un $\mu$-négligeable. Or $A=C\cup N$, donc $A\in\mathscr{E}^\mu$, ce qui termine la preuve. ■

On utilise le théorème précédent de mesurabilité des ensembles analytiques sous la forme suivante.

**Théorème A.7.19 (Mesurabilité universelle des projections)** *On note $\mathscr{B}([0,\infty])$ la tribu des boréliens de $[0,\infty]$. Soit $(\Omega,\mathscr{F},\mathbf{P})$, un espace mesurable. Soit $\pi$ la projection canonique de $[0,\infty]\times\Omega$ sur $\Omega$. On note $\mathscr{F}^{\mathbf{P}}$, la tribu $\mathscr{F}$ complétée par les $\mathbf{P}$-négligeables. Alors*

$$\forall B\in\mathscr{B}([0,\infty])\otimes\mathscr{F},\quad \pi(B)\in\mathscr{F}^{\mathbf{P}}\ . \tag{A.46}$$

**Preuve :** soit $B\in\mathscr{B}([0,\infty])\otimes\mathscr{F}$. Le théorème A.7.11 affirme que $\pi(B)\in\Sigma(\mathscr{F})$ et le théorème A.7.18 qui précède montre que $\Sigma(\mathscr{F})\subset\mathscr{F}^{\mathbf{P}}$. ■

**Preuve du théorème ?? du "Début".** Nous sommes maintenant à même de prouver les théorème du "Début". Soit $(\Omega,\mathscr{F},\mathbf{P})$, un espace de probabilité. Soit $(\mathscr{G}_t)_{t\in\mathbb{R}_+}$, une filtration sur $(\Omega,\mathscr{F})$. Soit $A\subset\mathbb{R}_+\times\Omega$, que l'on suppose progressivement mesurable par rapport à $(\mathscr{G}_t)_{t\in\mathbb{R}_+}$. On rappelle la définition du début de $A$ : On pose

$$\forall\,\omega\in\Omega,\quad D_A(\omega)=\inf\big\{t\in\mathbb{R}_+ : (t,\omega)\in A\big\}$$

avec la convention que $\inf\emptyset=\infty$. On rappelle que $(\mathscr{G}^{\mathbf{P}}_{t+})_{t\in\mathbb{R}_+}$ est la filtration $(\mathscr{G}_{t+})_{t\in\mathbb{R}_+}$ qui est complétée avec les $\mathbf{P}$-négligeables. On note $\pi$ la projection de $[0,\infty]\times\Omega$ sur $\Omega$.

On fixe $t\in]0,\infty[$. Comme $A$ est progressivement mesurable par rapport à $(\mathscr{G}_t)_{t\in\mathbb{R}_+}$, l'ensemble $B'=\{(s,\omega)\in[0,t]\times\Omega :(s,\omega)\in A\}$ est un ensemble $\mathscr{B}([0,t])\otimes\mathscr{G}_t$-mesurable. Cela implique immédiatement que

$$B:=\big\{(s,\omega)\in[0,\infty]\times\Omega\ :\ s<t \text{ et } (s,\omega)\in A\big\}=\big([0,t[\times\Omega\big)\cap B'\ \in\ \mathscr{B}([0,\infty])\otimes\mathscr{G}_t\ .$$

Or on remarque que

$$\{\omega\in\Omega\ :\ D_A(\omega)<t\}=\pi(B)\ .$$

Le théorème A.7.19 implique donc que $\{D_A<t\}\in\mathscr{G}^{\mathbf{P}}_t$, pour tout $t\in]0,\infty[$. Le lemme **??** $(i)$ (page **??**) implique que $D_A$ est un temps d'arrêt par rapport à la filtration continue à droite associée à $(\mathscr{G}^{\mathbf{P}}_{t+})_{t\in\mathbb{R}_+}$. Cela termine la preuve du théorème du "Début". ■

# Annexe B

# Bases de la théorie des probabilités.

## B.1 Brefs rappels.

On suppose le lecteur familier avec le formalisme des probabilités qui s'appuie sur la théorie de la mesure exposée au chapitre précédent. Dans la suite on note $(\Omega, \mathscr{F}, \mathbf{P})$ l'espace de probabilité sur lequel on travaille.

Les deux définitions suivantes n'en sont pas vraiement une : il s'agit plutôt d'introduire le vocabulaire probabiliste.

**Définition B.1.1** Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité. Soit $(E, \mathscr{E})$, un espace mesurable. Une fonction $X : \Omega \to E$ qui est $(\mathscr{F}, \mathscr{E})$-mesurable est plutôt appelée une *variable aléatoire* $(\mathscr{F}, \mathscr{E})$-mesurable. On utilise l'abréviation v.a. pour variable aléatoire. $\square$

**Notations.** Les variables aléatoires sont souvent notées par des lettres capitales scriptes de la fin de l'alphabet $X$, $Y$, $Z$. Soit $X : \Omega \to E$, unev.a. $(\mathscr{F}, \mathscr{E})$-mesurable. On utilise la notation suivante

$$\{X \in C\} := X^{-1}(C) = \{\omega \in \Omega : X(\omega) \in C \ , \} \ .$$

On omet les accolades { } le plus souvent possible. Par exemple on écrit $\mathbf{P}(X \in C)$ plutôt que $\mathbf{P}(\{X \in C\})$ et on écrit rarement $\mathbf{P}(X^{-1}(C))$. Si $Y$ est une autre variable aléatoire $(\mathscr{F}, \mathscr{E})$-mesurable, pour tout $D \in \mathscr{E}$, on écrit

$$\mathbf{P}(X \in C \, ; \, Y \in D) \text{ au lieu de } \mathbf{P}(X^{-1}(C) \cap Y^{-1}(D)).$$

Lorsque $E = \mathbb{R}$, $\mathbb{C}$, $[0, \infty]$, $[-\infty, \infty]$, $\mathbb{R}^d$, et que $\mathscr{E}$ est la tribu des Boréliens correspondante, on omet de préciser la tribu de l'espace d'arrivée : on parle simplement d'une variable aléatoire $\mathscr{F}$-mesurable (réelle, complexe, positive ...).

**Définition B.1.2** Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité.

(a) Soit $X : \Omega \to [0, \infty]$, une v.a. $\mathscr{F}$-mesurable. Son *espérance* est son intégrale par rapport à $\mathbf{P}$, notée $\mathbf{E}[X]$ :

$$\mathbf{E}[X] = \int_{\Omega} X(\omega) \, d\mathbf{P}(\omega) \ . \tag{B.1}$$

(b) Soit $X : \Omega \to \mathbb{R}$ ou $\mathbb{C}$, v.a. $\mathscr{F}$-mesurable est dite intégrable si elle est $\mathbf{P}$-intégrable, c'est-à-dire que $\mathbf{E}\big[|X|\big] < \infty$. Son *espérance* est son intégrale par rapport à $\mathbf{P}$. On note cette intégrale $\mathbf{E}[X]$, c'est-à-dire que l'on a (B.1). $\square$

L'espérance est linéaire positive, on a $\big|\mathbf{E}[X]\big| \leq \mathbf{E}\big[|X|\big]$. On rappelle, avec les notations probabilistes, le fait élémentaire suivant

$$\forall A \in \mathscr{F}, \quad \mathbf{E}[\mathbf{1}_A] = \mathbf{P}(A) \ .$$

Nous ne répétons pas intégralement les propriétés de l'espèrance puisque ce sont celle de toute intégrale contre une mesure positive. Nous rappelons qu *presque sûrement* signifie presque partout. Nous énonçons les principaux théorèmes de passage à la limite sous l'intégrale avec les notations probabilistes. Ici, $(X_n)_{n \geq 0}$ est une suite de v.a. $\mathscr{F}$-mesurables, à valeurs dans $[0, \infty]$, $\mathbb{R}$ ou $\mathbb{C}$.

**Convergence monotone.** On suppose les v.a. $X_n$ à valeurs dans $[0, \infty]$ et telles que que $X_n \leq X_{n+1}$, pour tout $n \in \mathbb{N}$. On pose $X_\infty = \sup_n X_n$. Alors, $\lim_n \mathbf{E}[X_n] = \mathbf{E}[X_\infty]$.

**Interversion positive.** On suppose les v.a. $X_n$ à valeurs dans $[0,\infty]$. Alors,

$$\mathbf{E}\Big[\sum_{n\geq 0} X_n\Big] = \sum_{n\geq 0} \mathbf{E}[X_n] \; .$$

**Fatou.** On suppose les v.a. $X_n$ à valeurs dans $[0,\infty]$. Alors,

$$\mathbf{E}[\liminf_n X_n] \leq \liminf_n \mathbf{E}[X_n] \; .$$

**Convergence dominée.** On suppose les v.a. $X_n$ à valeurs dans $\mathbb{C}$. On fait les hypothèses suivantes.

(a) Il existe une v.a. réelle $X_\infty$ telle que $\lim_n X_n = X_\infty$ p.s.

(b) Il existe une v.a. intégrable $Z$ telle que p.s. $|X_n| \leq Z$, pour tout $n \geq 0$.

Alors, $X_\infty$ est intégrable, $\lim_n \mathbf{E}\big[\,|X_n - X_\infty|\,\big] = 0$, et $\lim_n \mathbf{E}[X_n] = \mathbf{E}[X_\infty]$.

**Interversion $\mathrm{L}^1$.** On suppose les v.a. $X_n$ à valeurs dans $\mathbb{C}$. On suppose que $\sum_{n\geq 0} \mathbf{E}\big[|X_n|\big] < \infty$ Alors, $\sum_{n\geq 0} X_n$ est une v.a. bie définie p.s. et intégrable. On a,

$$\mathbf{E}\Big[\sum_{n\geq 0} X_n\Big] = \sum_{n\geq 0} \mathbf{E}[X_n] \; .$$

On utilisera souvent le lemme technique suivant qui est une reformulation du lemme A.1.19, page 273.

**Lemme B.1.3** *Soit $(\Omega, \mathscr{F})$, un espace mesurable et $\mathscr{G}$ une sous-tribu de $\mathscr{F}$. Soit $Z$, une v.a. positive $\mathscr{G}$-mesurable. Alors, Il existe $B_n \in \mathscr{G}$, $c_n \in \mathbb{R}_+$, $n \in \mathbb{N}$, tels que $Z = \sum_{n\geq 0} c_n \mathbf{1}_{B_n}$.*

**Définition B.1.4** (*Loi d'une variable aléatoire*) Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité. Soit $(E, \mathscr{E})$, un espace mesurable. Soit $X : \Omega \to E$, une v.a. $(\mathscr{F}, \mathscr{E})$-mesurable. *La loi de $X$* est la mesure image de $\mathbf{P}$ par $X$. Dans ce cours, on la note $\mu_X$ et on a donc

$$\forall C \in \mathscr{E}\,, \quad \mu_X(C) = \mathbf{P}(X \in C) \; .$$

Le théorème de *transfert* implique que pour toute fonction $f : E \to [0,\infty]$ qui est $\mathscr{E}$-mesurable, on a

$$\mathbf{E}\big[f(X)\big] = \int_E f \, d\mu_X \; .$$

Cette égalité reste vraie lorsque $f$ est à valeurs complexes si $f(X)$ est $\mathscr{F}$-mesurable intégrable, ce qui est équivalent à ce que $f$ soit $\mu_X$-intégrable. □

Par exemple si $X$ est une v.a. complexe $\mathscr{F}$-mesurable et intégrable, on a $\mathbf{E}[X] = \int_{\mathbb{C}} y \mu_X(dy)$.

On utilisera souvent le théorème suivant qui n'est qu'une réécriture avec des notations probabilistes du théorème A.1.23, page 274.

**Théorème B.1.5** *Soit $(\Omega, \mathscr{F})$, et $(E, \mathscr{E})$ de espaces mesurable. Soit $Y : \Omega \to E$, une v.a. $(\mathscr{F}, \mathscr{E})$-mesurable. Soit $X : \Omega \to \mathbb{R}$, une v.a. qui est bornée et mesurable par rapport à $\sigma(Y)$, la tribu engendrée par $Y$. Alors il existe une fonction $F : E \to \mathbb{R}$ qui est $\mathscr{E}$-mesurable bornée telle que $X = F(Y)$.*

**Preuve :** il s'agit du théorème A.1.23, page 274. ■

## Fonction caractéristique (transformée de Fourier des lois).

**Définition B.1.6** (*Fonction caractéristique d'une v.a.*) oit $\mathbf{X} : \Omega \to \mathbb{R}^n$, un vecteur aléatoire $\mathscr{F}$-mesurable. Sa *fonction caractéristique* est la transformée de Fourier de sa loi, c'est-à-dire

$$\forall \mathbf{u} \in \mathbb{R}^n, \quad \widehat{\mu}_{\mathbf{X}}(\mathbf{u}) = \mathbf{E}\big[e^{i\langle \mathbf{u}, \mathbf{X}\rangle}\big] \; .$$

On rappelle que $\widehat{\mu}_{\mathbf{X}} : \mathbb{R}^n \to \mathbb{C}$ est continue et que $|\widehat{\mu}_{\mathbf{X}}(\mathbf{u})| \leq \widehat{\mu}_{\mathbf{X}}(0) = 1$. □

On rappelle les propriétés de la transformée de Fourier, et donc des fonctions caractéristiques, en termes probabilistes.

**Théorème B.1.7** ***(Injectivité)*** *Soit $(\Omega', \mathscr{F}', \mathbf{P}')$, un espace de probabilité et soit $\mathbf{Y} : \Omega' \to \mathbb{R}^n$, un vecteur aléatoire $\mathscr{F}'$-mesurable. Alors $\mathbf{X}$ et $\mathbf{Y}$ ont même loi ssi leurs fonctions caractéristiques sont égales, c'est-à-dire*

$$\widehat{\mu}_{\mathbf{X}} = \widehat{\mu}_{\mathbf{Y}} \Longleftrightarrow \mu_{\mathbf{X}} = \mu_{\mathbf{Y}} \; .$$

**Théorème B.1.8** ***(Inversion)*** *Soit* $\mathbf{X}:\Omega\to\mathbb{R}^n$*, un vecteur aléatoire* $\mathscr{F}$*-mesurable. On suppose que* $\int_{\mathbb{R}^n}|\widehat{\mu}_{\mathbf{X}}|\,d\ell_n<\infty$*. Alors la loi de* $\mathbf{X}$ *admet la densité suivante*

$$\forall \mathbf{x}\in\mathbb{R}^n,\quad p(\mathbf{x})=(2\pi)^{-n}\int_{\mathbb{R}^n} e^{-i\langle \mathbf{u},\mathbf{x}\rangle}\widehat{\mu}_{\mathbf{X}}(\mathbf{u})\,d\ell_n(\mathbf{u})\;,$$

*qui est continue et tend vers* $0$ *en l'infini.*

## Variance et covariance.

**Définition B.1.9** (*Covariance de deux v.a. réelles*) Soient $X, Y$, deux v.a. réelles $\mathscr{F}$-mesurables. On suppose qu'elles admettent un moment d'ordre 2. L'inégalité de Hölder implique que $XY$ est intégrable et on pose

$$\mathbf{cov}(X,Y)=\mathbf{E}[XY]-\mathbf{E}[X]\mathbf{E}[Y]\;.$$

• Cette quantité est appelée la *covariance de* $X$ *et* $Y$. On vérifie facilement que

$$\mathbf{cov}(X,Y)=\mathbf{E}\big[(X-\mathbf{E}[X])(Y-\mathbf{E}[Y])\big]\;.$$

• La *variance de* $X$ est $\mathbf{var}(X):=\mathbf{cov}(X,X)=\mathbf{E}[(X-\mathbf{E}[X])^2]=\mathbf{E}[X^2]-(\mathbf{E}[X])^2$.

• (*Matrice de covariance d'un vecteur aléatoire*) Plus généralement, soit $\mathbf{X}=(X_1,\ldots,X_n):\Omega\to\mathbb{R}^n$, un vecteur aléatoire $\mathscr{F}$-mesurable. On dit qu'il admet un moment d'ordre 2 si $\mathbf{E}\big[\|\mathbf{X}\|^2\big]<\infty$. Cela implique que les $X_k$ admettent des moments d'ordre 2. Cela permet de poser

$$\Gamma_{\mathbf{X}}=\big(\mathrm{cov}(X_i,X_j)\big)_{1\le i,j\le n}\;.$$

Cette matrice carrée $n\times n$ est appelée la *matrice de covariance de* $\mathbf{X}$. □

**Proposition B.1.10** $\mathbf{X}=(X_1,\ldots,X_n):\Omega\to\mathbb{R}^n$*, un vecteur aléatoire* $\mathscr{F}$*-mesurable admettant un moment d'ordre* $2$*. On note* $\Gamma$ *sa matrice de covariance et on note*

$$\mathbf{v}=\mathbf{E}[\mathbf{X}]=\big(\mathbf{E}[X_1],\ldots,\mathbf{E}[X_n]\big)\in\mathbb{R}^n\;,$$

*son vecteur moyen. Alors les propriétés suivantes sont vérifiées.*

$(i)$ *Pour tout* $\mathbf{u}\in\mathbb{R}^n$*,*

$$\mathbf{E}\big[\langle\mathbf{u},\mathbf{X}\rangle\big]=\langle\mathbf{u},\mathbf{v}\rangle\quad \textit{et}\quad \mathbf{var}\big(\langle\mathbf{u},\mathbf{X}\rangle\big)=\langle\mathbf{u},\Gamma\mathbf{u}\rangle\;.$$

*Cela implique que* $\langle\mathbf{u},\Gamma\mathbf{u}\rangle\ge 0$*, pour tout* $\mathbf{u}\in\mathbb{R}^n$ *et donc que* $\Gamma$ *est une matrice symétrique positive.*

$(ii)$ *Soit* $M$*, une matrice réelle* $n\times n$*. On note* $M^*$ *sa transposée. Alors*

$$\Gamma_{M\mathbf{X}}=M\Gamma M^*\;.$$

**Preuve :** il est facile de vérifier que $\mathbf{E}\big[\langle\mathbf{u},\mathbf{X}\rangle\big]=\langle\mathbf{u},\mathbf{v}\rangle$ et que $\mathbf{var}\big(\langle\mathbf{u},\mathbf{X}\rangle\big)=\mathbf{var}\big(\langle\mathbf{u},\mathbf{X}-\mathbf{v}\rangle\big)$. On peut donc se ramener au cas où $\mathbf{E}[\mathbf{X}]=0$. Dans ce cas $\mathbf{cov}(X_i,X_j)=\mathbf{E}[X_iX_j]$. Pour tout $\mathbf{u}=(u_1,\ldots,u_n)\in\mathbb{R}^n$, on a donc

$$\mathbf{var}\big(\langle\mathbf{u},\mathbf{X}\rangle\big)=\mathbf{E}\big[\langle\mathbf{u},\mathbf{X}\rangle^2\big]=\sum_{1\le j,k\le n}u_ju_k\mathbf{E}[X_jX_k]=\langle\mathbf{u},\Gamma\mathbf{u}\rangle\;,$$

en développant le carré. Cela montre $(i)$. Montrons $(ii)$ : par définition de la transposée, $\langle\mathbf{x},M\mathbf{y}\rangle=\langle M^*\mathbf{x},\mathbf{y}\rangle$, pour tous $\mathbf{x},\mathbf{y}\in\mathbb{R}^n$. Le $(i)$ implique donc pour tout $\mathbf{u}\in\mathbb{R}^n$,

$$\langle\mathbf{u},\Gamma_{M\mathbf{X}}.\mathbf{u}\rangle=\mathbf{var}\big(\langle\mathbf{u},M\mathbf{X}\rangle\big)=\mathbf{var}\big(\langle M^*\mathbf{u},\mathbf{X}\rangle\big)=\langle M^*\mathbf{u},\Gamma M^*\mathbf{u}\rangle=\langle\mathbf{u},M\Gamma M^*\mathbf{u}\rangle$$

et il est facile de voir que cela implique $(ii)$. ■

**Proposition B.1.11** *Soit* $\mathbf{X}=(X_1,\ldots,X_n):\Omega\to\mathbb{R}^n$*, un vecteur aléatoire* $\mathscr{F}$*-mesurable admettant un moment d'ordre* $2$*. Alors, sa fonction caractéristique est* $C^2$ *est on a le développement de Taylor suivant au voisinage de* $0$ *:*

$$\forall\mathbf{u}=(u_1,\ldots,u_n)\in\mathbb{R}^n,\quad \widehat{\mu}_{\mathbf{X}}(\mathbf{u})=1+i\sum_{1\le k\le n}\mathbf{E}[X_k]\,u_k\;-\tfrac{1}{2}\sum_{1\le j,k\le n}\mathbf{E}[X_jX_k]\,u_ju_k\;+\mathrm{o}\big(\|\mathbf{u}\|^2\big)\;.$$

**Preuve :** les $X_k$ on des moments d'ordre 2. Ce sont des v.a. intégrables. Comme déjà mentionné, $X_j X_k$ sont aussi intégrables. Or on observe que

$$\frac{\partial}{\partial u_k} e^{i\langle \mathbf{u},\mathbf{X}\rangle} = i u_k X_k e^{i\langle \mathbf{u},\mathbf{X}\rangle} \quad \text{et} \quad \frac{\partial^2}{\partial u_j \partial u_k} e^{i\langle \mathbf{u},\mathbf{X}\rangle} = -u_j u_k X_j X_k e^{i\langle \mathbf{u},\mathbf{X}\rangle} .$$

Le théorème de dérivation sous l'intégrale montre alors que

$$\frac{\partial \widehat{\mu}_{\mathbf{X}}}{\partial u_k}(\mathbf{u}) = i u_k \mathbf{E}\big[X_k e^{i\langle \mathbf{u},\mathbf{X}\rangle}\big] \quad \text{et} \quad \frac{\partial^2 \widehat{\mu}_{\mathbf{X}}}{\partial u_j \partial u_k}(\mathbf{u}) = -u_j u_k \mathbf{E}\big[X_j X_k e^{i\langle \mathbf{u},\mathbf{X}\rangle}\big] ,$$

ce qui implique le résultat voulu. ■

## B.2 Indépendance.

**Définition B.2.1** (*Indépendance d'événements*) Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité. On introduit les notions suivantes.

(a) Soient $A, B \in \mathscr{F}$. Les événements $A$ et $B$ sont dits *indépendants sous* $\mathbf{P}$ si $\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$.

(b) Soient $A_1, \ldots, A_n \in \mathscr{F}$. Les événements $A_1, \ldots, A_n$ sont dits *mutuellement indépendants sous* $\mathbf{P}$ si pour tout $2 \leq k \leq n$ et tous $1 \leq j_1 < \ldots < j_k \leq n$, on

$$\mathbf{P}(A_{j_1} \cap A_{j_2} \cap \ldots \cap A_{j_k}) = \mathbf{P}(A_{j_1})\mathbf{P}(A_{j_2}) \ldots \mathbf{P}(A_{j_k}) .$$

Il y a donc $2^n - 1 - n$ égalités à vérifier.

(c) Soit $A_i \in \mathscr{F}$, $i \in I$, une famille d'événements. On dit qu'ils sont *mutuellement indépendants sous* $\mathbf{P}$ si pour tout sous-ensemble d'indices $J \subset I$ qui est *fini et non-vide*, on a

$$\mathbf{P}\Big(\bigcap_{j \in J} A_j\Big) = \prod_{j \in J} \mathbf{P}(A_j) .$$

On remarque que les $(A_i)_{i \in I}$ sont mutuellement indépendants sous $\mathbf{P}$ ssi si pour tout sous-ensemble d'indices $J \subset I$ qui est fini et qui compte au mois deux éléments, les événements $(A_j)_{j \in J}$ sont mutuellement indépendants sous $\mathbf{P}$ au sens du (b). □

**Définition B.2.2** (*Indépendance de classe d'ensembles*) Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité. On introduit les notions suivantes.

(a) Soient $\mathscr{R}_1, \ldots, \mathscr{R}_n \subset \mathscr{F}$, des classes non-vides d'événements. *Les classes d'événements* $\mathscr{R}_1, \ldots, \mathscr{R}_n$ *sont mutuellement indépendantes sous* $\mathbf{P}$ si

$$\forall A_1 \in \mathscr{R}_1, \ \ldots \ \forall A_n \in \mathscr{R}_n, \quad \text{les événements } A_1, \ldots, A_n \text{ sont indépendants.}$$

(b) Soit $\mathscr{R}_i \subset \mathscr{F}$, $i \in I$, une famille de classes non-vides d'événements. *Les classes* $\mathscr{R}_i$, $i \in I$, *sont mutuellement indépendantes sous* $\mathbf{P}$ si pour tout sous-ensemble d'indices $J \subset I$, qui est *fini* et qui contient au moins deux éléments, les classes $(\mathscr{R}_j)_{j \in J}$ sont mutuellement indépendantes au sens du (b). □

Au vu du point (c) de la définition B.2.1, les classes d'événéments $(\mathscr{R}_i)_{i \in I}$ sont mutuellement indépendantes sous $\mathbf{P}$ ssi

$$\forall J \subset I, \ J \text{ fini et non-vide}, \forall A_j \in \mathscr{R}_j \, , j \in J \, , \quad \mathbf{P}\Big(\bigcap_{j \in J} A_j\Big) = \prod_{j \in J} \mathbf{P}(A_j) . \tag{B.2}$$

La proposition suivante est une conséquence immédiate des définitions précédentes.

**Proposition B.2.3** *Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité. Soit $\mathscr{R}_i \subset \mathscr{F}$, $i \in I$, une famille de classes non-vides d'événements. Soit $I' \subset I$ et soient $\mathscr{R}'_i \subset \mathscr{R}_i$, $i \in I'$, des classes non-vides d'événements. On suppose que les classes d'événements $\mathscr{R}_i$, $i \in I$ sont mutuellement indépendantes sous $\mathbf{P}$. Alors il en est de même pour les classes d'événements $\mathscr{R}'_i$, $i \in I'$.*

**Lemme B.2.4** *Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité. Soient $\mathscr{R}_1, \ldots, \mathscr{R}_n \subset \mathscr{F}$, des classes non-vides d'événements. On suppose que $\Omega \in \mathscr{R}_k$ pour tout $1 \leq k \leq n$. Alors,*

$$\mathscr{R}_1, \ldots, \mathscr{R}_n \text{ indépendants} \iff \forall A_1 \in \mathscr{R}_1, \ldots, \forall A_n \in \mathscr{R}_n, \ \mathbf{P}(A_1 \cap \ldots \cap A_n) = \mathbf{P}(A_n) \ldots \mathbf{P}(A_n).$$

**Preuve :** l'implication $\Longrightarrow$ est claire. Montrons l'implication contraire. On suppose donc que

$$\forall A_1 \in \mathscr{R}_1, \ldots, \forall A_n \in \mathscr{R}_n,\ \mathbf{P}(A_1 \cap \ldots \cap A_n) = \mathbf{P}(A_n) \ldots \mathbf{P}(A_n). \tag{B.3}$$

Soient $B_1 \in \mathscr{R}_1, \ldots, B_n \in \mathscr{R}_n$. Il faut montrer que les événements $B_1, \ldots, B_n$ sont mutuellement indépendants sous $\mathbf{P}$ : soient $2 \leq k \leq n$ et $1 \leq j_1 < \ldots < j_k \leq n$. Pour tout $1 \leq \ell \leq k$, on pose $A_{j_\ell} := B_{j_\ell}$ et si $j \notin \{j_1, \ldots, j_k\}$, on pose $A_j := \Omega$, qui appartient à $\mathscr{R}_j$ par hypothèse. Par (B.3), on a donc

$$\begin{aligned} \mathbf{P}(B_{j_1} \cap \ldots \cap B_{j_k}) &= \mathbf{P}(A_1 \cap A_2 \cap \ldots \cap A_n) \\ &= \mathbf{P}(A_1)\mathbf{P}(A_2) \ldots \mathbf{P}(A_n) = \mathbf{P}(B_{j_1}) \ldots \mathbf{P}(B_{j_k}). \end{aligned}$$

Cela implique que pour tous $B_1 \in \mathscr{R}_1, \ldots, B_n \in \mathscr{R}_n$, les événements $B_1, \ldots, B_n$ sont mutuellement indépendants sous $\mathbf{P}$. Les classes d'événements $\mathscr{R}_1, \ldots, \mathscr{R}_n$ sont donc mutuellement indépendantes sous $\mathbf{P}$. ■

**Théorème B.2.5** *Soit $\mathscr{P}_i \subset \mathscr{F}$, $i \in I$, une famille de pi-systèmes. On suppose que les pi-systèmes $\mathscr{P}_i$, $i \in I$, sont mutuellement indépendants sous $\mathbf{P}$. Alors, les tribus engendrées $\sigma(\mathscr{P}_i)$, $i \in I$, sont également mutuellement indépendantes sous $\mathbf{P}$.*

**Preuve :** au vu de la définition B.2.2 (b), il suffit de prouver le théorème dans le cas où $I$ est fini. On choisit donc $I = \{1, \ldots, n\}$. On montre d'abord l'assertion suivante.

- Pour *tous* pi-systèmes $\mathscr{P}_i^* \subset \mathscr{F}$, $1 \leq i \leq n$, on a l'implication suivante :

$$\big(\mathscr{P}_1^*, \ldots, \mathscr{P}_n^* \text{ mut. ind. sous } \mathbf{P}\big) \implies \big(\sigma(\mathscr{P}_1^*), \mathscr{P}_2^*, \ldots, \mathscr{P}_n^* \text{ mut. ind. sous } \mathbf{P}\big) . \tag{B.4}$$

En effet, on suppose $\mathscr{P}_1^*, \ldots, \mathscr{P}_n^*$ indépendants, on fixe $A_2 \in \mathscr{P}_2^*, \ \ldots, \ A_n \in \mathscr{P}_n^*$ et on pose

$$\mathcal{L} = \big\{ B \in \sigma(\mathscr{P}_1^*) : \mathbf{P}(B \cap A_2 \cap \ldots \cap A_n) = \mathbf{P}(B)\mathbf{P}(A_2 \cap \ldots \cap A_n) \big\} .$$

On a donc $\mathscr{P}_1^* \subset \mathcal{L}$. On montre facilement que $\mathcal{L}$ est une classe monotone (exercice). Le théorème de la classe monotone implique alors que $\sigma(\mathscr{P}_1^*) \subset \mathcal{L}$ et donc que $\sigma(\mathscr{P}_1^*) = \mathcal{L}$. Donc, pour tout $B \in \sigma(\mathscr{P}_1^*)$, $\mathbf{P}(B \cap A_2 \cap \ldots \cap A_n) = \mathbf{P}(B)\mathbf{P}(A_2 \cap \ldots \cap A_n)$. Mais comme $\mathscr{P}_2^*, \ldots, \mathscr{P}_n^*$ sont indépendants, on a $\mathbf{P}(A_2 \cap \ldots \cap A_n) = \mathbf{P}(A_2) \ldots \mathbf{P}(A_n)$. On a a donc montré que

$$\forall B \in \sigma(\mathscr{P}_1^*),\ \forall A_2 \in \mathscr{P}_2^*, \ldots ,\ \forall A_n \in \mathscr{P}_n^*, \quad \mathbf{P}(B \cap A_2 \cap \ldots \cap A_n) = \mathbf{P}(B)\mathbf{P}(A_2) \ldots \mathbf{P}(A_n) .$$

Le lemme B.3 implique ensuite que les classes d'événements $\sigma(\mathscr{P}_1^*), \mathscr{P}_2^* \ldots, \mathscr{P}_n^*$ sont mutuellement indépendantes sous $\mathbf{P}$, ce qui termine la preuve de (B.4).

En appliquant (B.4) à $(\mathscr{P}_1^*, \ldots, \mathscr{P}_n^*) = (\mathscr{P}_1, \ldots, \mathscr{P}_n)$, on montre que $\sigma(\mathscr{P}_1), \mathscr{P}_2, \ldots, \mathscr{P}_n$ sont mutuellement indépendants sous $\mathbf{P}$. Or $\sigma(\mathscr{P}_1)$ est une tribu donc un pi-système. On applique alors (B.4) à la suite de pi-systèmes

$$(\mathscr{P}_1^*, \ldots, \mathscr{P}_n^*) = (\mathscr{P}_2, \ldots, \mathscr{P}_n, \sigma(\mathscr{P}_1)) ,$$

ce qui montre la mutuelle indépendance des $\sigma(\mathscr{P}_2), \mathscr{P}_3, \ldots, \mathscr{P}_n, \sigma(\mathscr{P}_1)$. Ainsi de suite jusqu'à montrer que $\sigma(\mathscr{P}_1), \ldots, \sigma(\mathscr{P}_n)$ sont mutuellement indépendantes sous $\mathbf{P}$, ce qui termine la preuve. ■

**Proposition B.2.6** **(Indépendance par paquets)** *Soit $I$ un ensemble d'indices. Soit $I_k \subset I$, $k \in K$, une partition de $I$. Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité. Soit $\mathscr{F}_i \subset \mathscr{F}$, $i \in I$, une famille de tribus. Pour tout $k \in K$, on pose $\mathscr{G}_k = \sigma(\mathscr{F}_i; i \in I_k)$. Alors,*

$$\mathscr{F}_i,\ i \in I \text{ mut. ind. sous } \mathbf{P} \Longrightarrow \mathscr{G}_k,\ k \in K \text{ mut. ind. sous } \mathbf{P} .$$

**Preuve :** pour tout $k \in K$, on note $\mathscr{P}_k$ la classe des intersections finies d'événements dans $\bigcup_{i \in I_k} \mathscr{F}_i$ :

$$\mathscr{P}_k = \big\{ A_1 \cap \ldots \cap A_n \,;\, n \in \mathbb{N},\ A_1, \ldots, A_n \in \bigcup_{i \in I_k} \mathscr{F}_i \,\big\}.$$

Il est facile de voir que $\mathscr{P}_k$ est un pi-système tel que $\sigma(\mathscr{P}_k) = \mathscr{G}_k$. L'indépendance des tribus $\mathscr{F}_i$, $i \in I$, entraîne facilement celle des $\mathscr{P}_k$, $k \in K$. Le théorème B.2.5 permet de conclure. ■

Soit $(E, \mathscr{E})$, un espace mesurable et $X : \Omega \to E$, une v.a. $(\mathscr{F}, \mathscr{E})$-mesurable. On rappelle que la tribu engendrée par $X$ est donnée par la tribu pré-image de $\mathscr{E}$ par $X$, c'est-à-dire

$$\sigma(X) = \big\{ \{X \in B\} \,;\, B \in \mathscr{E} \big\} \tag{B.5}$$

qui est la plus petite tribu sur $\Omega$ rendant $X$ mesurable. On introduit la définition suivante.

**Définition B.2.7** (*Indépendance de v.a.*) Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité. Soit $I$, un ensemble d'indices. Pour tout $i \in I$, soit $(E_i, \mathscr{E}_i)$, un espace mesurable et $X_i : \Omega \to E_i$, une v.a. $(\mathscr{F}, \mathscr{E}_i)$-mesurable.

On dit que *les variables $X_i$, $i \in I$, sont mutuellement indépendantes sous* $\mathbf{P}$ si les tribus $\sigma(X_i) \subset \mathscr{F}$, $i \in I$, sont mutuellement indépendantes sous $\mathbf{P}$. □

La proposition suivante donne plusieurs reformulations de la définition de l'indépendance des variables aléatoires.

**Proposition B.2.8** *Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité. Soit $I$, un ensemble d'indices. Pour tout $i \in I$, on fixe les objets suivants.*

• *Un espace mesurable $(E_i, \mathscr{E}_i)$ et un pi-système $\mathscr{C}_i \subset \mathscr{E}_i$ tel que $\sigma(\mathscr{C}_i) = \mathscr{E}_i$.*

• *Une v.a. $X_i : \Omega \to E_i$ qui est $(\mathscr{F}, \mathscr{E}_i)$-mesurable et dont la loi sous $\mathbf{P}$ est notée $\mu_i$ : $\mu_i(B) = \mathbf{P}(X_i \in B)$, $B \in \mathscr{E}_i$.*

*Alors, les assertions suivantes sont équivalentes.*

$(i)$ *Les variables $X_i$, $i \in I$, sont mutuellement indépendantes sous* $\mathbf{P}$.

$(ii)$ *Pour tout $J \subset I$, fini non-vide et pour tous $B_j \in \mathscr{E}_j$, $j \in J$,*

$$\mathbf{P}\Big( \bigcap_{j \in J} \{X_j \in B_j\} \Big) = \prod_{j \in J} \mathbf{P}(X_j \in B_j) \; .$$

$(iii)$ *Pour tout $J \subset I$, fini non-vide et pour tous $C_j \in \mathscr{C}_j$, $j \in J$,*

$$\mathbf{P}\Big( \bigcap_{j \in J} \{X_j \in C_j\} \Big) = \prod_{j \in J} \mathbf{P}(X_j \in C_j) \; .$$

$(iv)$ *Pour tout $J = \{j_1, \ldots, j_n\} \subset I$, où les $j_k$ sont distincts,*

$$\textit{la loi de } (X_{j_1}, \ldots, X_{j_n}) \textit{ sous } \mathbf{P} \textit{ est } \mu_{j_1} \otimes \ldots \otimes \mu_{j_n}.$$

$(v)$ *Pour tout $J \subset I$, fini non-vide et pour toutes fonctions $\mathscr{E}_j$-mesurables bornées $h_j : E_j \to \mathbb{R}$, $j \in J$,*

$$\mathbf{E}\Big[ \prod_{j \in J} h_j(X_j) \Big] = \prod_{j \in J} \mathbf{E}\big[ h_j(X_j) \big] \; .$$

**Preuve :** au vu des définitions B.2.7 et B.2.2, il suffit de traiter le cas où $I$ est fini. On prend donc $I = \{1, \ldots, n\}$. Par définition, $(i)$ est équivalent à l'indépendance des tribus $\sigma(X_i)$, $i \in I$. Or (B.5) et (B.2) montrent que $(i) \Leftrightarrow (ii)$. L'implication $(ii) \Rightarrow (iii)$ est triviale. Supposons $(iii)$ : on pose

$$\mathscr{P}_i = \big\{ \{X_i \in C\} \, ; \, C \in \mathscr{C}_i \big\} \; .$$

C'est clairement un pi-système tel que $\sigma(X_i) = \sigma(\mathscr{P}_i)$. Or $(iii)$ signifie que les pi-systèmes $\mathscr{P}_i$, $i \in I$ sont indépendants : le théorème B.2.5 implique alors que les tribus $\sigma(\mathscr{P}_i)$, $i \in I$, sont indépendantes, ce qui implique $(i)$. On a donc montré que $(iii) \Rightarrow (i)$ et donc que $(i)$, $(ii)$ et $(iii)$ sont équivalents.

On suppose $(ii)$. On rappelle ensuite que $\mu_1 \otimes \ldots \otimes \mu_n$ est l'unique mesure de probabilité sur $\mathscr{E}_1 \otimes \ldots \otimes \mathscr{E}_n$ telle que $(\mu_1 \otimes \ldots \otimes \mu_n)(B_1 \times \ldots \times B_n) = \mu_1(B_1) \ldots \mu_n(B_n)$, pour tous $B_1 \in \mathscr{E}_1$, ..., $B_n \in \mathscr{E}_n$. On note $\mu$ la loi de $X = (X_1, \ldots, X_n)$. Il est clair que pour tous $B_1 \in \mathscr{E}_1, \ldots, B_n \in \mathscr{E}_n$

$$\mu(B_1 \times \ldots \times B_n) = \mathbf{P}\big(X \in B_1 \times \ldots \times B_n\big) = \mathbf{P}\Big( \bigcap_{1 \leq j \leq n} \{X_j \in B_j\} \Big).$$

Or, par $(ii)$, $\mathbf{P}(\bigcap_{1 \leq j \leq n} \{X_j \in B_j\}) = \prod_{1 \leq j \leq n} \mathbf{P}(X_j \in B_j) = \mu_1(B_1) \ldots \mu_n(B_n)$. Cela entraîne donc $\mu = \mu_1 \otimes \ldots \otimes \mu_n$, et donc $(iv)$. On a montré que $(ii) \Rightarrow (iv)$. L'implication $(iv) \Rightarrow (v)$ est une conséquence directe du théorème de Fubini. L'implication $(v) \Rightarrow (ii)$ est triviale. ■

Pour des variables aléatoires réelles, on obtient le corollaire suivant

**Corollaire B.2.9** *Soit $(X_i)_{i \in I}$, une famille de v.a. réelles $\mathscr{F}$-mesurables. Elles sont indépendantes ssi pour tout $\{j_1, \ldots, j_n\} \subset I$ où les $j_k$ sont distincts, et pour tous $x_1, \ldots, x_n \in \mathbb{R}$, on ait*

$$\mathbf{P}\big(X_{j_1} \leq x_1; \ldots; X_{j_n} \leq x_n\big) = \mathbf{P}(X_{j_1} \leq x_1) \ldots \mathbf{P}(X_{j_n} \leq x_n) \; .$$

**Preuve :** on applique l'équivalence de $(i)$ et de $(iii)$ de la proposition B.2.8 à $(E_i, \mathscr{E}_i) = (\mathbb{R}, \mathscr{B}(\mathbb{R})$ et $\mathscr{C}_i = \{\mathbb{R}\} \cup \{\, ]-\infty, a] \, ; \, a \in \mathbb{R}\}$, qui est un pi-système générant $\mathscr{B}(\mathbb{R})$. ■

Les fonctions caractéristiques sont plus adaptées aux sommes de variables indépendantes.

**Proposition B.2.10** *Soient $\mathbf{X}_k : \Omega \to \mathbb{R}^n$, $1 \leq k \leq p$, des vecteurs aléatoires $\mathscr{F}$-mesurables que l'on suppose indépendants. On note $\mu_1, \ldots \mu_p$ leurs lois respectives. Alors, la loi $\mu$ de $\mathbf{X} := \mathbf{X}_1 + \ldots + \mathbf{X}_p$ est $\mu_1 * \ldots * \mu_p$ et on a*

$$\forall \mathbf{u} \in \mathbb{R}^n, \quad \widehat{\mu}(\mathbf{u}) = \mathbf{E}\big[e^{i\langle \mathbf{u}, \mathbf{X}\rangle}\big] = \mathbf{E}\big[e^{i\langle \mathbf{u}, \mathbf{X}_1\rangle}\big] \ldots \mathbf{E}\big[e^{i\langle \mathbf{u}, \mathbf{X}_p\rangle}\big] = \widehat{\mu}_1(\mathbf{u}) \ldots \widehat{\mu}_p(\mathbf{u}) \; .$$

**Preuve :** on note $\phi : (\mathbf{x}_1, \ldots, \mathbf{x}_p) \in (\mathbb{R}^n)^p \longmapsto \mathbf{x}_1 + \ldots + \mathbf{x}_p$, qui est mesurable car continue. Par définition du produit de convolution $\mu_1 * \ldots * \mu_p$ est la mesure image de $\mu_1 \otimes \ldots \otimes \mu_p$ par $\phi$. Comme les v.a. $\mathbf{X}_k$ sont indépendantes, la proposition B.2.8 $(iv)$ implique que la loi de $(\mathbf{X}_1, \ldots, \mathbf{X}_p)$ est $\mu_1 \otimes \ldots \otimes \mu_p$. Par conséquent la loi de $\mathbf{X}$, notée $\mu$ est bien $\mu_1 * \ldots * \mu_p$. Le reste du lemme est une conséquence de propriétés de la transformée de Fourier des mesures et de la convolution. ■

On prouve ensuite le critère d'indépendance suivant.

**Proposition B.2.11** *Soient $X_1, \ldots, X_n$, des v.a. réelles $\mathscr{F}$-mesurables. Elles sont indépendantes ssi on*

$$\forall u_1, \ldots, u_n \in \mathbb{R}, \quad \mathbf{E}\big[e^{iu_1 X_1 + \ldots + iu_n X_n}\big] = \mathbf{E}\big[e^{iu_1 X_1}\big] \ldots \mathbf{E}\big[e^{iu_n X_n}\big] \; . \tag{B.6}$$

**Preuve :** si les variables sont indépendantes, alors la proposition B.2.10 implique (B.6). Supposons (B.6). On note $\mu_k$ la loi de $X_k$ et on note $\mu$ la loi de $\mathbf{X} = (X_1, \ldots, X_n)$, qui est vecteur aléatoire $\mathscr{F}$-mesurable de $\mathbb{R}^n$. Par injectivité de la transformée de Fourier, (B.6) implique que $\mu = \mu_1 \otimes \ldots \otimes \mu_n$, ce qui entraîne l'indépendances des $X_k$ par la proposition B.2.8. ■

Donnons une conséquence très utile de la proposition B.2.8.

**Proposition B.2.12** *Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité. Soient $X_1, \ldots, X_n$, des v.a. complexes $\mathscr{F}$-mesurables intégrables indépendantes. Alors le produit $X_1 \ldots X_n$ est également intégrable et*

$$\mathbf{E}[X_1 \ldots X_n] = \mathbf{E}[X_1] \ldots \mathbf{E}[X_n] \; . \tag{B.7}$$

*De même si $X_1, \ldots, X_n : \Omega \to [0, \infty]$ sont $\mathscr{F}$-mesurables indépendantes, on a (B.7).*

**Preuve :** par la proposition B.2.8 la loi de $X = (X_1, \ldots, X_n)$ est le produit des lois des $X_i$ : $\mu_X = \mu_{X_1} \otimes \ldots \otimes \mu_{X_n}$. Par le théorème de transfert et Fubini positif

$$\begin{aligned} \mathbf{E}\big[\, |X_1| \ldots |X_n| \,\big] &= \int_{\mathbb{R}^n} |x_1| \ldots |x_n| \, \mu_{X_1}(dx_1) \ldots \mu_{X_n}(dx_n) \\ &= \Big( \int_{\mathbb{R}} |x_1| \, \mu_{X_1}(dx_1) \Big) \ldots \Big( \int_{\mathbb{R}} |x_n| \, \mu_{X_n}(dx_n) \Big) \\ &= \mathbf{E}\big[\, |X_1| \,\big] \ldots \mathbf{E}\big[\, |X_n| \,\big] \; . \end{aligned}$$

Cela démontre le résultat voulu lorsque les variables sont à valeurs dans $[0, \infty]$. Si les variables $X_1, \ldots X_n$ sont intégrables, il en est de même pour $X_1 \ldots X_n$ et par transfert et Fubini, on a

$$\begin{aligned} \mathbf{E}\big[X_1 \ldots X_n\,\big] &= \int_{\mathbb{R}^n} x_1 \ldots x_n \, \mu_{X_1}(dx_1) \ldots \mu_{X_n}(dx_n)(dx_n) \\ &= \Big( \int_{\mathbb{R}} x_1 \, \mu_{X_1}(dx_1) \Big) \ldots \Big( \int_{\mathbb{R}} x_n \, \mu_{X_n}(dx_n) \Big) \\ &= \mathbf{E}[X_1] \ldots \mathbf{E}[X_n] \; , \end{aligned}$$

qui est le résultat voulu. ■

# B.3 Variables et vecteurs Gaussiens.

**Définition B.3.1** (*Lois Gaussiennes (normales) en dimension* 1) Soient $m \in \mathbb{R}$ et $\sigma \in ]0, \infty[$ Soit $X : \Omega \to \mathbb{R}$, une variable $\mathscr{F}$-mesurable. Cette variable est dite (de loi) Gaussienne, ou encore normale, de moyenne $m$ et de varaince $\sigma^2$ si sa loi admet la densité $g_{m,\sigma}$ donnée par

$$g_{m,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} \; , \quad x \in \mathbb{R} \; . \tag{B.8}$$

On note cela $X \sim \mathscr{N}(m, \sigma^2)$. Si $m = 0$, $X$ est dite centrée et une variable Gaussienne $\mathscr{N}(0,1)$ est appelée Gaussienne standard.

Par commodité, on dit également que $X$ est une loi Gaussienne de moyenne $m$ et de variance nulle si $X$ est presque sûrement égal à $m$. On continue de noter cela $X \sim \mathscr{N}(m, 0)$. On parle alors de variable *Gaussienne dégénérée*. □

On vérifie facilement que si $X \sim \mathscr{N}(m, \sigma^2)$,

$$\mathbf{E}[X] = \int_{\mathbb{R}} x g_{m,\sigma}(x) dx = m \quad \text{et} \quad \mathbf{var}(X) = \int_{\mathbb{R}} (x-m)^2 g_{m,\sigma}(x) dx = \sigma^2 .$$

Autrement dit la loi de $X$ est entièrement déterminée par sa moyenne et sa variance. On rappelle que $\widehat{g_{0,\sigma}}(u) = \exp(-\frac{1}{2}\sigma^2 u^2)$. Cela entraîne facilement que si $X \sim \mathscr{N}(m, \sigma)$,

$$\forall u \in \mathbb{R}, \quad \mathbf{E}\big[e^{iuX}\big] = \widehat{\mu}_X(u) = \widehat{g}_{m,\sigma}(u) = e^{imu - \frac{1}{2}\sigma^2 u^2} .$$

Il est facile de montrer que la fonction caractéristique de $X$ s'étend à $\mathbb{C}$, c'est-à-dire que

$$\forall z \in \mathbb{C}, \quad \mathbf{E}\big[e^{zX}\big] = e^{mz + \frac{1}{2}\sigma^2 z^2} .$$

Le membre de droite est une fonction analytique sur $\mathbb{C}$ dont le rayon de convergence est infini (c'est une fonction entière). Si on suppose pour simplifier que $X \sim \mathscr{N}(0, \sigma)$, on vérifie alors

$$\mathbf{E}\big[e^{zX}\big] = \sum_{n \geq 0} \frac{\mathbf{E}[X^n]}{n!} z^n = \sum_{n \geq 0} \frac{\sigma^{2n}}{2^n n!} z^{2n} .$$

Cela implique que si $X \sim \mathscr{N}(0, \sigma)$, pour tout $n \in \mathbb{N}$,

$$\mathbf{E}[X^{2n+1}] = 0 \quad \text{et} \quad \mathbf{E}[X^{2n}] = \frac{(2n)!}{n!\, 2^n} \sigma^{2n} = \sigma^{2n} (2n-1) \times (2n-3) \times \ldots \times 5 \times 3 \times 1 .$$

**Lemme B.3.2** *Soient $X_1 \sim \mathscr{N}(m_1, \sigma_1^2), \ldots, X_n \sim \mathscr{N}(m_n, \sigma_n^2)$ des variables Gaussiennes indépendantes. Soient $a_1, \ldots, a_n \in \mathbb{R}$. Alors $a_1 X_1 + \ldots + a_n X_n \sim \mathscr{N}(m, \sigma^2)$ où*

$$m = a_1 m_1 + \ldots + a_n m_n \quad \text{et} \quad \sigma^2 = (a_1 \sigma_1)^2 + \ldots + (a_n \sigma_n)^2 .$$

**Preuve :** on pose $X = a_1 X_1 + \ldots + a_n X_n$ et on a pour tout $u \in \mathbb{R}$,

$$\mathbf{E}\big[e^{iuX}\big] = \mathbf{E}\big[e^{i a_1 u X_1}\big] \ldots \mathbf{E}\big[e^{i a_n u X_n}\big] = e^{imu - \frac{1}{2}\sigma^2 u^2} ,$$

et l'injectivité de la transformée de Fourier permet de conclure. ■

**Définition B.3.3** (*Vecteurs Gaussiens*) Soit $\mathbf{X} = (X_1, \ldots, X_d) : \Omega \to \mathbb{R}^d$, un vecteur $\mathscr{F}$-mesurable. C'est un *vecteur Gaussien* si pour tout $\mathbf{u} = (u_1, \ldots, u_d) \in \mathbb{R}^d$, $\langle u, \mathbf{X} \rangle = u_1 X_1 + \ldots + u_d X_d$ est une variable Gaussienne réelle (possiblement dégénérée). □

Le lemme suivant découle immédiatement de la définition.

**Lemme B.3.4** *Soit $\mathbf{X} = (X_1, \ldots, X_d) : \Omega \to \mathbb{R}^d$, un vecteur Gaussien. Soit $M$, une matrice réelle carrée de taille $d \times d$. Alors, le vecteur aléatoire $M\mathbf{X}$ est Gaussien.*

**Exemple B.3.5** Soient $X_1, \ldots, X_d$, des v.a. réelles indépendantes de loi $\mathscr{N}(0,1)$. Alors, $\mathbf{X} := (X_1, \ldots, X_d)$ est un vecteur Gaussien. En effet, On fixe $\mathbf{u} = (u_1, \ldots, u_d) \in \mathbb{R}^d$ et $v \in \mathbb{R}$. L'indépendance implique que

$$\widehat{\mu}_{\langle \mathbf{u}, \mathbf{X} \rangle}(v) = \prod_{1 \leq k \leq d} \mathbf{E}[e^{ivu_k X_k}] = \exp\big(-\tfrac{1}{2} v^2 (u_1^2 + \ldots + u_d^2)\big) = \exp(-\tfrac{1}{2} v^2 \|\mathbf{u}\|^2) .$$

Cela montre que pour tout $\mathbf{u} \in \mathbb{R}^d$, $\langle \mathbf{u}, \mathbf{X} \rangle$ est une variable Gaussienne $\mathscr{N}(0, \|\mathbf{u}\|^2)$ par injectivité de la fonction caractéristique. □

On rappelle les définitions du vecteur moyenne et de la matrice de covariance, ainsi que des formules élémentaires de la proposition B.1.10 page 305 qui entraînent immédiatement le lemme suivant.

**Lemme B.3.6** *Soit $\mathbf{X} : \Omega \to \mathbb{R}^d$, un vecteur Gaussien $\mathscr{F}$-mesurable. On note $\mathbf{v}$, le vecteur moyenne de $\mathbf{X}$ et $\Gamma$, sa matrice de covariance. Alors pour tout $\mathbf{u} \in \mathbb{R}^d$,*

$$\langle \mathbf{u}, \mathbf{X}\rangle \textit{ est une v.a. réelle Gaussienne de moyenne } \langle \mathbf{u}, \mathbf{v}\rangle \textit{ et de variance } \langle \mathbf{u}, \Gamma\mathbf{u}\rangle.$$

*et donc*

$$\widehat{\mu}_{\mathbf{X}}(\mathbf{u}) = \mathbf{E}\big[e^{i\langle \mathbf{u},\mathbf{X}\rangle}\big] = \widehat{\mu}_{\langle \mathbf{u},\mathbf{X}\rangle}(1) = \exp\big(i\langle \mathbf{u}, \mathbf{v}\rangle - \tfrac{1}{2}\langle \mathbf{u}, \Gamma\mathbf{u}\rangle\big)\ . \tag{B.9}$$

**Définition B.3.7** Le lemme précédent montre que la loi d'un vecteur Gaussien $\mathbf{X}$ est entièrement caractérisée par son vecteur moyenne $\mathbf{v}$ et sa matrice de covariance $\Gamma$. On note le fait que $\mathbf{X}$ est un vecteur Gaussien de vecteur moyen $\mathbf{v}$ et de matrice de covariance $\Gamma$ par $\mathbf{X} \sim \mathscr{N}\big(\mathbf{v}, \Gamma\big)$. □

Soit $\mathbf{X} : \Omega \to \mathbb{R}^d$, un vecteur Gaussien $\mathscr{F}$-mesurable de vecteur moyen $\mathbf{v}$ et de matrice de covariance $\Gamma$. On note $r$ le rang de $\Gamma$ et $\lambda_1, \dots, \lambda_r$ ses valeurs propres non-nulles. Comme $\Gamma$ est symétrique positive, $\lambda_1, \dots, \lambda_r > 0$. On note $\Delta$ et $\widetilde{\Delta}$ les matrices $d \times d$, diagonales données par

$$\Delta = \begin{pmatrix} \sqrt{\lambda_1} & 0 & \dots & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & \sqrt{\lambda_r} & \ddots & & \vdots \\ \vdots & & \ddots & 0 & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \dots & & \dots & 0 & 0 \end{pmatrix} \quad \text{et} \quad \widetilde{\Delta} = \begin{pmatrix} 1/\sqrt{\lambda_1} & 0 & \dots & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & 1/\sqrt{\lambda_r} & \ddots & & \vdots \\ \vdots & & \ddots & 0 & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \dots & & \dots & 0 & 0 \end{pmatrix}$$

$\Delta\widetilde{\Delta}$ est donc la matrice $\mathbf{I}_r$ qui est diagonale et dont les $r$ premiers coefficients diagonaux sont égaux à 1 et les $d-r$ suivants sont nuls. Comme $\Gamma$ est symétrique, il existe une matrice orthogonale $O$ telle que $O^*\Gamma O = \Delta^2$, où $O^*$ désigne la transposée de $O$.

**Proposition B.3.8** *On reprend les notations ci-dessus. On pose $\mathbf{Y} := \widetilde{\Delta}O^*(\mathbf{X} - \mathbf{v})$. Alors*

$$\mathbf{Y} = (Y_1, \dots, Y_r, 0, \dots, 0)\ ,$$

*où $Y_1, \dots, Y_r$ sont des v.a. réelles indépendantes de loi $\mathscr{N}(0,1)$, qui sont appelées les composantes indépendantes de $\mathbf{X}$. De plus $\mathbf{X} = \mathbf{v} + O\Delta\mathbf{Y}$.*

**Preuve :** $\mathbf{Y}$ est un vecteur Gaussien de vecteur moyen nul. Par la proposition B.1.10 $(ii)$, page 305, sa matrice de covariance $\Gamma'$ est donnée par

$$\Gamma' = \widetilde{\Delta}O^*\Gamma(\widetilde{\Delta}O^*)^* = \widetilde{\Delta}O^*\Gamma O\widetilde{\Delta} = \widetilde{\Delta}\Delta^2\widetilde{\Delta} = \mathbf{I}_r\ .$$

On a donc $\mathbf{Y} \sim \mathscr{N}(0, \mathbf{I}_r)$, ce qui montre la proposition. ■

Cette proposition implique que toute matrice symétrique positive peut être la matrice de covariance d'un vecteur Gaussien. Dans le cas où cette matrice de covariance est inversible, la loi du vecteur Gaussien correspondant admet une densité par rapport à la mesure de Lebesgue, donnée dans la proposition suivante.

**Proposition B.3.9 (Densité des vecteurs Gaussiens)** *Soit $\mathbf{X} : \Omega \to \mathbb{R}^d$, un vecteur Gaussien $\mathscr{F}$-mesurable, de vecteur moyenne nul et de matrice de covariance $\Gamma$ supposée inversible. La loi de $\mathbf{X}$ admet la densité suivante par rapport à la mesure de Lebesgue sur $\mathbb{R}^d$ :*

$$\forall \mathbf{x} \in \mathbb{R}^d, \quad g_\Gamma(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}\sqrt{\det(\Gamma)}} \exp\big(-\tfrac{1}{2}\langle \mathbf{x}, \Gamma^{-1}\mathbf{x}\rangle\big)\ .$$

**Preuve :** on reprend les notation précédentes : on pose $M = O\Delta$, si bien que $\Gamma = MM^*$ et la proposition précédente implique que $\mathbf{X} = M\mathbf{Y}$, où $\mathbf{Y} \sim \mathscr{N}(0, \mathrm{Id})$. Soit $f : \mathbb{R}^d \to \mathbb{R}_+$, une fonction mesurable. On a donc

$$\mathbf{E}\big[f(M.\mathbf{Y})\big] = (2\pi)^{-d/2}\int_{\mathbb{R}^d} e^{-\frac{1}{2}\|\mathbf{y}\|^2} f(M\mathbf{y})\,\ell_d(d\mathbf{y})\ .$$

Comme $\Gamma$ est supposée inversible il en est de même pour $M$ et on a $\det(M) = \sqrt{\det(\Gamma)}$ et $\Gamma^{-1} = (M^{-1})^*M^{-1}$. Donc $\|M^{-1}\mathbf{x}\|^2 = \langle \mathbf{x}, \Gamma^{-1}\mathbf{x}\rangle$. Le théorème de changement de variable linéaire implique alors que

$$\mathbf{E}\big[f(M.\mathbf{Y})\big] = \frac{1}{(2\pi)^{d/2}\det(M)}\int_{\mathbb{R}^d} e^{-\frac{1}{2}\|M^{-1}\mathbf{x}\|^2} f(\mathbf{x})\,\ell_d(d\mathbf{x}) = \frac{1}{(2\pi)^{d/2}\sqrt{\det(\Gamma)}}\int_{\mathbb{R}^d} e^{-\frac{1}{2}\langle \mathbf{x}, \Gamma^{-1}\mathbf{x}\rangle} f(\mathbf{x})\,\ell_d(\mathbf{x})\ ,$$

ce qui implique le résultat voulu. ■

## B.4 Convergence de variables aléatoires.

**Convergence en probabilité.** On note $(\Omega, \mathscr{F}, \mathbf{P})$, l'espace de probabilité de référence pour cette section, espace sur lequel toutes les v.a. mentionnées sont a priori définies, sauf mention explicite du contraire. Soit $(E, d)$, un espace métrique séparable. On note $\mathscr{B}(E)$ la tribu Borélienne. On pose $\mathcal{L}_E(\Omega, \mathscr{F}) = \big\{X : \Omega \to E,\ (\mathscr{F}, \mathscr{B}(E))\text{-mesurable}\big\}$.

**Définition B.4.1** (*Convergence p.s. et en probabilité*) Soient $X_n \in \mathcal{L}_E(\Omega, \mathscr{F})$, $n \in \mathbb{N}$ et $X \in \mathcal{L}_E(\Omega, \mathscr{F})$.

(a) *Convergence en probabilité.* La suite $(X_n)_{n\geq 0}$ converge en $\mathbf{P}$-probabilité vers $X$ ssi

$$\forall \varepsilon > 0\,, \quad \lim_{n\to\infty} \mathbf{P}\big(d(X_n, X) > \varepsilon\big) = 0\ .$$

(b) *Convergence presque sûre.* La suite $(X_n)_{n\geq 0}$ converge $\mathbf{P}$-presque sûrement vers $X$ ssi il existe $N \in \mathscr{F}$ tel que $\mathbf{P}(N) = 0$ et

$$\forall \omega \in \Omega \backslash N\,, \quad \lim_{n\to\infty} d\big(X_n(\omega), X(\omega)\big) = 0\ .$$

Lorsqu'il n'y a pas ambiguïté, on parle seulement de convergence en probabilité plutôt que de convergence en $\mathbf{P}$-probabilité et de convergence p.s.ãulieu de convergence $\mathbf{P}$-presque sûre. □

Le théorème suivant fait le lien entre la convergence en probabilité et la convergence presque sûre.

**Théorème B.4.2** *Soient $X_n \in \mathcal{L}_E(\Omega, \mathscr{F})$, $n \in \mathbb{N}$ et $X \in \mathcal{L}_E(\Omega, \mathscr{F})$. Les assertions suivantes sont vérifiées*

$(i)$ *Si $\lim_n X_n = X$ p.s., alors $\lim_n X_n = X$ en probabilité.*

$(ii)$ *Si $\lim_n X_n = X$ en probabilité, il existe une suite strictement croissante d'indices $(n_k)_{k\geq 0}$ tels que $\lim_k X_{n_k} = X$ presque sûrement.*

$(iii)$ *On suppose que $\lim_n X_n = X$ en probabilité. Soit $Y \in \mathcal{L}_E(\Omega, \mathscr{F})$ tel que l'on ait aussi $\lim_n X_n = Y$ en probabilité. Alors $X = Y$ presque sûrement.*

**Preuve :** on montre $(i)$. Supposons que $\lim_n X_n = X$ presque sûrement. On pose $Y_n = \mathbf{1}_{\{d(X_n, X) > \varepsilon\}}$. Alors $0 \leq Y_n \leq 1$ et p.s $\lim_n Y_n = 0$. Par convergence dominée, $\lim_n \mathbf{E}[Y_n] = 0$, c'est-à-dire $\lim_n \mathbf{P}(d(X_n, X) > \varepsilon) = 0$. Comme $\varepsilon$ peut être arbitrairement petit, cela montre que $\lim_n X_n = X$ en probabilité.

Montrons $(ii)$. Supposons que $\lim_n X_n = X$ en probabilité. Par définition, pour tout $k \in \mathbb{N}$, il existe $p_k \in \mathbb{N}$, tel que pour tout $n \geq p_k$, on ait $\mathbf{P}(d(X_n, X) > 2^{-k}) \leq 2^{-k}$. On pose $n_k = p_1 + \ldots + p_k$ et on a bien $\mathbf{P}(d(X_{n_k}, X) > 2^{-k}) \leq 2^{-k}$. Donc $\sum_{k\in\mathbb{N}} \mathbf{P}\big(d(X_{n_k}, X) > 2^{-k}\big) < \infty$. Le lemme de Borel-Cantelli implique que presque sûrement pour tout $k$ suffisamment grand, $d(X_{n_k}, X) \leq 2^{-k}$ et donc $\lim_k X_{n_k} = X$ presque sûrement, ce qui prouve $(ii)$. Le point $(iii)$ est une conséquence immédiate du $(ii)$. Une preuve directe sans utiliser $(ii)$ est possible. ■

**Proposition B.4.3** *Soient $(E, d)$ et $(E', d')$, deux espace métriques séparables. On munit $E \times E'$ de la distance $D\big((x, x'), (y, y')\big) := \max\big(d(x, y), d(x', y')\big)$. On rappelle que $(E \times E', D)$ est séparable et que la topologie de la métrique $D$ est la topologie produit sur $E \times E'$.*

*Soient $X_n \in \mathcal{L}_E(\Omega, \mathscr{F})$, $X'_n \in \mathcal{L}_{E'}(\Omega, \mathscr{F})$, $n \in \mathbb{N}$, et soient $X' \in \mathcal{L}_{E'}(\Omega, \mathscr{F})$ et $X' \in \mathcal{L}_{E'}(\Omega, \mathscr{F})$. Alors,il y a équivalence entre les deux assertions suivantes.*

$(a)$ $\lim_n X_n = X$ *et* $\lim_n X'_n = X'$ *en probabilité.*

$(b)$ $\lim_n (X_n, X'_n) = (X, X')$ *en probabilité dans* $(E \times E', D)$.

**Preuve :** pour tout $\varepsilon > 0$, on remarque d'abord que

$$\mathbf{P}\big(D((X_n, X'_n), (X, X')) > \varepsilon\big) \leq \mathbf{P}\big(d(X_n, X) > \varepsilon\big) + \mathbf{P}\big(d'(X'_n, X') > \varepsilon\big)\ ,$$

ce qui montre que $(a) \Longrightarrow (b)$. On vérifie également que

$$\max\big(\mathbf{P}\big(d(X_n, X) > \varepsilon\big), \mathbf{P}\big(d'(X'_n, X') > \varepsilon\big)\big) \leq \mathbf{P}\big(D((X_n, X'_n), (X, X')) > \varepsilon\big)\ ,$$

qui entraîne que $(b) \Longrightarrow (a)$. ■

On s'intéresse ensuite aux propriétés de continuité de la convergence en probabilité. Plus précisément, soient $(E, d)$ et $(E', d')$, deux espaces métriques séparables ; soit $f : E \to E'$, une fonction $(\mathscr{B}(E), \mathscr{B}(E'))$-mesurable. On définit le module de continuité de $f$ en $x$ de la manière suivante : pour tout $x \in E$ et tout réel $\delta > 0$, on pose

$$w_f(x, \delta) := \sup \big\{ d'(f(y), f(z)) \, ; \, y, z \in B(x, \delta) \big\} \ ,$$

où $B(x, \delta) := \{y \in E \, : \, d(x, y) < \delta\}$ est la boule ouverte de centre $x$ et de rayon $\delta$. On rappelle que $f$ est continue en $x$ ssi $\lim_{\delta \to 0} w_f(x, \delta) = 0$. On montre le lemme suivant.

**Lemme B.4.4** *Soient $(E, d)$ et $(E', d')$, deux espaces métriques séparables. Soit $f : E \to E'$, une fonction $(\mathscr{B}(E), \mathscr{B}(E'))$-mesurable. Alors les assertions suivantes sont vérifiées.*

$(i)$ *Pour tout $\delta > 0$, la fonction $x \in E \mapsto w_f(x, \delta) \in [0, \infty]$ est $\mathscr{B}(E)$-mesurable.*

$(ii)$ $C := \{x \in E \, : \, f$ *est continue en* $x\} \in \mathscr{B}(E)$.

**Preuve :** pour tous réels $a, \delta > 0$, on pose

$$U_{a,\delta} = \big\{ x \in E \, : \, \exists y, z \in B(x, \delta) \text{ tels que } d'\big(f(y), f(z)\big) > a \big\}.$$

On voit que $U_{a,\delta} = \{x \in E \, : \, w_f(x, \delta) > a\}$. On vérifie que $U_{a,\delta}$ est un ouvert. En effet, soit $x \in U_{a,\delta}$. Il existe $y, z \in B(x, \delta)$ tels que $d'\big(f(y), f(z)\big) > a$. On pose $\eta = \min(\delta - d(x, y), \delta - d(x, z))$. Donc $\eta > 0$, et pour tout $x_* \in B(x, \eta)$, on a bien $d(x_*, y) \leq d(x_*, x) + d(x, y) < \delta$. De même, $d(x_*, z) < \delta$. Donc, $y, z \in B(x_*, \delta)$ et on a bien $B(x_*, \eta) \subset U_{a,\delta}$. Cela montre $(i)$.

Pour montrer $(ii)$, on pose pour tout $x \in E$, $g(x) = \limsup_p w_f(x, 2^{-p})$. Par $(i)$, c'est une application $\mathscr{B}(E)$-mesurable. On remarque ensuite que pour tout $p \in \mathbb{N}$, et tout $x \in E$, $w_f(x, 2^{-p-1}) \leq w_f(x, 2^{-p})$. Donc

$$C = \{x \in E \, : \, f \text{ est continue en } x\} = \{x \in E : \lim_{\delta \to 0} w_f(x, \delta) = 0\} = \{g = 0\} \ ,$$

ce qui termine la preuve. ■

**Proposition B.4.5** *Soient $(E, d)$ et $(E', d')$, deux espace métriques séparables. Soit $f : E \to E'$, une fonction $(\mathscr{B}(E), \mathscr{B}(E'))$-mesurable. Soient $X_n \in \mathcal{L}_E(\Omega, \mathscr{F})$, $n \in \mathbb{N}$ et $X \in \mathcal{L}_E(\Omega, \mathscr{F})$. On fait les hypothèses suivantes.*

$(a)$ $\lim_n X_n = X$ *en probabilité.*

$(b)$ *Presque sûrement $f$ est continue en $X$.*

*Alors* $\lim_n f(X_n) = f(X_\infty)$ *en probabilité.*

**Preuve :** pour tous réels $\varepsilon, \delta > 0$, on a

$$\begin{aligned} \mathbf{P}\big(d'\big(f(X_n), f(X)\big) > \varepsilon\big) &= \mathbf{P}\big(d'\big(f(X_n), f(X)\big) > \varepsilon ; d(X_n, X) < \delta\big) + \mathbf{P}\big(d'\big(f(X_n), f(X)\big) > \varepsilon ; d(X_n, X) \geq \delta\big) \\ &\leq \ \mathbf{P}\big(w_f(X, \delta) > \varepsilon\big) + \mathbf{P}\big( d(X_n, X) > \delta/2\big). \end{aligned}$$

Donc pour tous $\varepsilon, \delta > 0$, l'hypothèse $(a)$ implique que $\limsup_{n \to \infty} \mathbf{P}\big(d'\big(f(X_n), f(X)\big) > \varepsilon\big) \leq \mathbf{P}\big(w_f(X, \delta) > \varepsilon\big)$. Pour tout $p \in \mathbb{N}$, on pose $A_p = \{w_f(X, 2^{-p}) > \varepsilon\}$. On a donc $A_{p+1} \subset A_p$ et $A := \bigcap_{p \in \mathbb{N}} A_p \subset \{f \text{ n'est pas continue en } X\}$. Donc $\lim_p \mathbf{P}(A_p) = \mathbf{P}(A) = 0$ par l'hypothèse $(b)$, ce qui permet de conclure. ■

On déduit des deux proposition précédentes les propriétés de continuité suivantes de la convergence en probabilité (bien d'autres énoncés du même genre en découlent).

**Proposition B.4.6** *Soit $(E, d)$, un espace métrique séparable. Soient $X_n, Y_n, X, Y \in \mathcal{L}_E(\Omega, \mathscr{F})$. On suppose que* $\lim_{n \to \infty} X_n = X$ *en probabilité et* $\lim_{n \to \infty} Y_n = Y$ *en probabilité. Alors les assertions suivantes sont vérifiées.*

$(i)$ *Si $E = \mathbb{R}^k$, alors* $\lim_n X_n + Y_n = X + Y$ *en probabilité.*

$(ii)$ *Si $E = \mathbb{R}$ ou $\mathbb{C}$, alors* $\lim_n X_n Y_n = XY$ *en probabilité.*

$(iii)$ *Si $E = \mathbb{R}$ ou $\mathbb{C}$ et si p.s. $Y \neq 0$, alors* $\lim_n X_n / Y_n = X/Y$ *en probabilité.*

$(iv)$ *Si $E = [-\infty, \infty]$, alors* $\lim_n X_n \wedge Y_n = X \wedge Y$ *et* $\lim_n X_n \vee Y_n = X \vee Y$ *en probabilité.*

**Preuve :** la proposition B.4.3 implique que $\lim_n (X_n, Y_n) = (X, Y)$ en probabilité dans $E^2$. On applique ensuite la proposition B.4.5 à $E = \mathbb{R}^k$, et $f(x, y) = x + y$ pour obtenir $(i)$, à $E = \mathbb{R}$ ou $\mathbb{C}$ et $f(x, y) = xy$ pour obtenir $(ii)$, à $f(x, y) = x/y$ qui est continue sur $\mathbb{C} \times \mathbb{C}^*$ pour obtenir $(iii)$, et enfin à $E = [-\infty, \infty]$ et $f(x, y) = x \wedge y$ et $f(x, y) = x \vee y$ pour obtenir $(iv)$. ■

**Proposition B.4.7** *Soit $(E, d)$, un espace séparable que l'on suppose complet. Soit $\varepsilon_n \in ]0, \infty[$, $n \in \mathbb{N}$, une suite telle que $\sum_{n \geq 0} \varepsilon_n < \infty$. Soient $X_n \in \mathcal{L}_E(\Omega, \mathscr{F})$, $n \in \mathbb{N}$. On suppose $\sum_{n \geq 0} \mathbf{P}\big( d(X_{n+1}, X_n) > \varepsilon_n\big) < \infty$. Alors, il existe $X \in \mathcal{L}_E(\Omega, \mathscr{F})$ telle que $\lim_n X_n = X$ presque sûrement.*

**Preuve :** le lemme de Borel-Cantelli implique l'existence de $B \in \mathscr{F}$ tel que $\mathbf{P}(B) = 1$ et tel pour tout $\omega \in B$, on ait $d(X_{n+1}(\omega), X_n(\omega)) \leq \varepsilon_n$ pour tout $n \geq n_0(\omega)$. L'inégalité triangulaire implique que pour tous $m, n \geq p \geq n_0(\omega)$, on a $d\big(X_n(\omega), X_m(\omega)\big) \leq \sum_{k \geq p} \varepsilon_k$. Comme $\sum_{n \in \mathbb{N}} \varepsilon_n < \infty$, pour tout $\omega \in B$, la suite $(X_n(\omega))_{n \in \mathbb{N}}$ est de Cauchy. Puisque $E$ est supposé complet, il existe $X(\omega) \in E$ tel que $\lim_n d(X_n(\omega), X(\omega)) = 0$. On fixe ensuite $x_0 \in E$ et pour tout $\omega \in \Omega \backslash B$, on pose $X(\omega) = x_0$. On a donc montré que $\lim_n X_n = X$ presque sûrement car $\mathbf{P}(B) = 1$.

Il reste à montrer que $X$ est $(\mathscr{F}, \mathscr{B}(E))$-mesurable. Pour cela, on rappelle que les fonctions continues bornées sur $E$ engendre la tribu Borélienne $\mathscr{B}(E)$. Cela implique que si pour toute fonction continue bornée $f : E \to \mathbb{R}$, $f(X)$ est $\mathscr{F}$-mesurable alors $X$ est $(\mathscr{F}, \mathscr{B}(E))$-mesurable. Or on a $\mathbf{P}$-p.s. $\lim_n f(X_n) = f(X)$, ce qui implique que $f(X)$ est $\mathscr{F}$-mesurable une limite presque sûr de v.a. réelles $\mathscr{F}$-mesurable (les $f(X_n)$ le sont) est $\mathscr{F}$-mesurable. Cela termine la preuve de la proposition. ■

Deux variables aléatoires $X, Y \in \mathcal{L}_E(\Omega, \mathscr{F})$ sont égales p.s. ssi $\mathbf{P}(d(X, Y) = 0) = 1$. On note cela $X \sim Y$. L'inégalité triangulaire implique que $\sim$ est transitive. Elle est clairement symétrique et réflexive. C'est donc une relation d'équivalence. On considère le quotient

$$\mathrm{L}_E(\Omega, \mathscr{F}, \mathbf{P}) = \mathcal{L}_E(\Omega, \mathscr{F}) / \sim \ .$$

Bien que les éléments de $\mathrm{L}_E(\Omega, \mathscr{F}, \mathbf{P})$ ne soient pas des v.a. mais des classes de v.a., on confond un élément de $\mathrm{L}_E(\Omega, \mathscr{F}, \mathbf{P})$ avec n'importe lequel de ses représentants. Le résultat suivant qui montre que la convergence en probabilité correspond à une distance sur $\mathrm{L}_E(\Omega, \mathscr{F}, \mathbf{P})$.

**Théorème B.4.8** *Soit $(E, d)$, un espace métrique séparable. Pour tous $X, Y \in \mathrm{L}_E(\Omega, \mathscr{F}, \mathbf{P})$, on pose*

$$d_{\mathbf{P}}(X, Y) = \mathbf{E}\big[1 \wedge d(X, Y)\big] \ .$$

*Alors, $d_{\mathbf{P}}$ est une distance sur $\mathrm{L}_E(\Omega, \mathscr{F}, \mathbf{P})$ telle que $\lim_n d_{\mathbf{P}}(X_n X)$ ssi $\lim_n X_n = X$ en probabilité. De plus, l'espace métrique $\big(\mathrm{L}_E(\Omega, \mathscr{F}, \mathbf{P}), d_{\mathbf{P}}\big)$ est complet si $E$ est complet.*

**Preuve :** on vérifie que pour tous $a, b \in \mathbb{R}_+$, on a $1 \wedge (a + b) \leq 1 \wedge a + 1 \wedge b$. On en déduit que pour toutes $X, Y, Z \in \mathrm{L}_E(\Omega, \mathscr{F}, \mathbf{P})$, on a

$$1 \wedge d(X, Z) \leq 1 \wedge \big(d(X, Y) + d(Y, Z)\big) \leq 1 \wedge d(X, Y) + 1 \wedge d(Y, Z) \ .$$

En prenant l'espérance on a $d_{\mathbf{P}}(X, Z) \leq d_{\mathbf{P}}(X, Y) + d_{\mathbf{P}}(Y, Z)$. Par ailleurs, il est clair que $d_{\mathbf{P}}(X, Y) = d_{\mathbf{P}}(Y, X)$. Enfin si $d_{\mathbf{P}}(X, Y) = 0$, alors $X = Y$ p.s. c'est-à-dire que $X = Y$ dans $\mathrm{L}_E(\Omega, \mathscr{F}, \mathbf{P})$. Cela montre que $d_{\mathbf{P}}$ est une distance.

Soient $X_n \in \mathcal{L}_E(\Omega, \mathscr{F})$, $n \in \mathbb{N}$ et soit $X \in \mathcal{L}_E(\Omega, \mathscr{F})$. On remarque d'abord que pour tout $a \in \mathbb{R}_+$, $(a/\varepsilon) \wedge 1 = (a \wedge \varepsilon)/\varepsilon \leq (a \wedge 1)/\varepsilon$. Par conséquent,

$$\mathbf{1}_{\{d(X_n, X) > \varepsilon\}} \leq \big(\varepsilon^{-1} d(X_n, X)\big) \wedge 1 \leq \varepsilon^{-1} \big(d(X_n, X) \wedge 1\big) \ .$$

En prenant l'espérance, on obtient $\mathbf{P}\big(d(X_n, X) > \varepsilon\big) \leq \varepsilon^{-1} d_{\mathbf{P}}(X_n, X)$, pour tout $\varepsilon \in ]0, 1]$. Par conséquent, si $\lim_n d_{\mathbf{P}}(X_n, X) = 0$, alors, $\lim_n X_n = X$ en probabilité.

On observe ensuite que

$$\begin{aligned} d_{\mathbf{P}}(X_n, X) &= \mathbf{E}\big[\mathbf{1}_{\{d(X_n, X) \leq \varepsilon\}} \ 1 \wedge d(X_n, X)\big] + \mathbf{E}\big[\mathbf{1}_{\{d(X_n, X) > \varepsilon\}} \ 1 \wedge d(X_n, X)\big] \\ &\leq \varepsilon + \mathbf{P}\big(d(X_n, X) > \varepsilon\big) \ . \end{aligned}$$

Si $\lim_n X_n = X$ en probabilité, alors $\limsup_n d_{\mathbf{P}}(X_n, X) \leq \varepsilon$, pour tout $\varepsilon > 0$, et donc $\lim_n d_{\mathbf{P}}(X_n, X) = 0$.

On suppose $(E, d)$ complet. Soit $X_n \in \mathrm{L}_E(\Omega, \mathscr{F}, \mathbf{P})$, $n \in \mathbb{N}$ qui est de Cauchy pour $d_{\mathbf{P}}$. Pour montrer qu'elle est convergente, il suffit d'en extraire une suite qui est convergente. Pour cela on rappelle l'inégalité précédente : pour tout $\varepsilon \in ]0, 1]$, pour tous $m, n$, on a

$$\mathbf{P}\big(d(X_n, X_m) > \varepsilon\big) \leq \varepsilon^{-1} d_{\mathbf{P}}(X_n, X_m) \ .$$

Comme $(X_n)_{n \in \mathbb{N}}$, est de Cauchy, il est facile d'en déduire qu'il existe une suite strictement croissante d'entiers $(n_k)_{k \in \mathbb{N}}$ telle que

$$\forall k \in \mathbb{N}, \quad \mathbf{P}\big(d(X_{n_{k+1}}, X_{n_k}) > 2^{-k}\big) \leq 2^{-k} \ .$$

La proposition B.4.7 implique qu'il existe $X \in \mathrm{L}_E(\Omega, \mathscr{F}, \mathbf{P})$ telle que $\lim_k X_{n_k} = X$ p.s. et le théorème B.4.2 $(i)$ implique que $\lim_k X_{n_k} = X$ en probabilité, ce qui implique que $\lim_k d_{\mathbf{P}}(X_{n_k}, X) = 0$, comme déjà démontré plus haut. Cela termine la preuve du théorème. ■

**Convergence dans $\mathrm{L}^p$. Uniforme intégrabilité.** Pour simplifier les notations on pose $\mathcal{L}(\Omega, \mathscr{F}) = \mathcal{L}_{\mathbb{R}}(\Omega, \mathscr{F})$, c'est-à-dire $\mathcal{L}(\Omega, \mathscr{F}) = \{X : \Omega \to \mathbb{R} \ \mathscr{F}\text{-mesurable}\}$. On utilise la notation $X \sim Y$ pour dire que $X = Y$ presque sûrement. On rappelle la notation

$$\mathrm{L}(\Omega, \mathscr{F}, \mathbf{P}) = \mathcal{L}(\Omega, \mathscr{F}) / \sim \ .$$

Bien que les éléments de $\mathrm{L}(\Omega, \mathscr{F}, \mathbf{P})$ ne soient pas des v.a. mais des classes de v.a., on confond un élément de $\mathrm{L}(\Omega, \mathscr{F}, \mathbf{P})$ avec n'importe quel de ses représentants. Pour tout $p \in [1, \infty[$ et toute v.a. $X \in \mathcal{L}(\Omega, \mathscr{F})$, on pose

$$\|X\|_p = \big(\mathbf{E}\big[X^p\big]\big)^{1/p} ,$$

avec la convention que $\infty^{1/p} = \infty$. On pose également

$$\|X\|_\infty = \inf\big\{a \in [0, \infty] : \mathbf{P}(|X| > a) = 0\big\} .$$

Pour tout $p \in [1, \infty]$, on pose alors

$$\mathrm{L}^p(\Omega, \mathscr{F}, \mathbf{P}) = \big\{X \in \mathrm{L}(\Omega, \mathscr{F}, \mathbf{P}) : \|X\|_p < \infty\big\} .$$

On rappelle que $(\mathrm{L}^p(\Omega, \mathscr{F}, \mathbf{P}), \|\cdot\|_p)$ est un $\mathbb{R}$-espace vectoriel normé complet, c'est-à-dire un espace de Banach. On montre tout d'abord la proposition suivante faisant le lien entre convergence $\mathrm{L}^p$ et en probabilité.

**Proposition B.4.9** *Soit $p \in [1, \infty]$. Soient $X_n \in \mathrm{L}^p(\Omega, \mathscr{F}, \mathbf{P})$, $n \in \mathbb{N}$ et $X \in \mathrm{L}^p(\Omega, \mathscr{F}, \mathbf{P})$. Si $\lim_n \|X_n - X\|_p = 0$, alors $\lim_n X_n = X$ en probabilité.*

**Preuve :** on suppose d'abord que $p < \infty$. Une inégalité de type Markov implique que pour tout $\varepsilon > 0$ et tout $n \in \mathbb{N}$,

$$\mathbf{P}\big(|X_n - X| > \varepsilon\big) \leq \varepsilon^{-p} \mathbf{E}\big[|X_n - X|^p\big] = \varepsilon^{-p} \|X_n - X\|_p^p ,$$

ce qui implique le résultat pour $p \in [1, \infty[$. On suppose ensuite que $\lim_n \|X_n - X\|_\infty = 0$. Pour tout $\varepsilon > 0$, il existe donc $n_\varepsilon \in \mathbb{N}$ tel que pour tout $n \geq n_\varepsilon$, on ait $\|X_n - X\|_\infty < \varepsilon$ et donc $\mathbf{P}\big(|X_n - X| > \varepsilon\big) = 0$, ce qui permet de conclure. ■

En combinant le résultat précédent et le théorème B.4.2 $(ii)$, on obtient immédiatement

**Proposition B.4.10** *Soit $p \in [1, \infty]$. Soient $X_n \in \mathrm{L}^p(\Omega, \mathscr{F}, \mathbf{P})$, $n \in \mathbb{N}$ et $X \in \mathrm{L}^p(\Omega, \mathscr{F}, \mathbf{P})$. On suppose que $\lim_n \|X_n - X\|_p = 0$. Alors, il existe une suite strictement croissante d'indices $(n_k)_{k \geq 0}$ tels que $\lim_k X_{n_k} = X$ presque sûrement.*

**Définition B.4.11** (*Famille de v.a. uniformément intégrables*) Soient $X_i \in \mathcal{L}(\Omega, \mathscr{F})$, $i \in I$. La famille de v.a. $(X_i)_{i \in I}$ est *uniformément intégrable* si

$$\lim_{a \to \infty} \sup_{i \in I} \mathbf{E}\big[|X_i| \mathbf{1}_{\{|X_i| \geq a\}}\big] = 0 . \tag{B.10}$$

On observe que cette notion ne concerne pas les variables en elles-même mais leur loi. □

**Lemme B.4.12** *Soient $X_i \in \mathcal{L}(\Omega, \mathscr{F})$, $i \in I$. Alors, $\sup_{i \in I} \|X_i\|_1 < \infty$. En particulier, les $X_i$ sont des v.a. intégrables.*

**Preuve :** (B.10) implique qu'il existe $a > 0$ tel que $C := \sup_{i \in I} \mathbf{E}\big[|X_i| \mathbf{1}_{\{|X_i| \geq a\}}\big] < \infty$. Donc

$$\forall i \in I, \quad \mathbf{E}\big[|X_i|\big] = \mathbf{E}\big[|X_i| \mathbf{1}_{\{|X_i| \leq a\}}\big] + \mathbf{E}\big[|X_i| \mathbf{1}_{\{|X_i| \geq a\}}\big] \leq a + C ,$$

ce qui permet de conclure. ■

**Lemme B.4.13** *Soit $(X_i)_{i \in I}$, une famille de v.a. réelles intégrables.*

*$(i)$ Si $I = J_1 \cup J_2$, si $(X_i)_{i \in J_1}$ est uniformément intégrable et si $(X_i)_{i \in J_2}$ est uniformément intégrable, alors $(X_i)_{i \in I}$ est uniformément intégrable.*

*$(ii)$ Si $I$ est fini, alors, la famille de variables $(X_i)_{i \in I}$ est uniformément intégrable.*

*$(iii)$ Soit $Z$, une v.a. réelle intégrable telle que pour tout $i \in I$, on ait presque sûrement $|X_i| \leq Z$. Alors, la famille de variables $(X_i)_{i \in I}$ est uniformément intégrable.*

*$(iv)$ Soit $g : \mathbb{R}_+ \to \mathbb{R}_+$, une application mesurable telle que $\lim_{x \to \infty} g(x)/x = \infty$. On suppose que $\sup_{i \in I} \mathbf{E}\big[g(|X_i|)\big] < \infty$. Alors, la famille de variables $(X_i)_{i \in I}$ est uniformément intégrable.*

**Preuve :** le premier point découle du fait que pour tout $a\in\mathbb{R}_+$ on ait

$$\sup_{i\in I}\mathbf{E}\big[|X_i|\mathbf{1}_{\{|X_i|\geq a\}}\big] \leq \sup_{i\in J_1}\mathbf{E}\big[|X_i|\mathbf{1}_{\{|X_i|\geq a\}}\big] + \sup_{i\in J_2}\mathbf{E}\big[|X_i|\mathbf{1}_{\{|X_i|\geq a\}}\big] \; .$$

Pour montrer $(ii)$, on remarque d'abord que que si $Z$ est une variable réelle positive, alors p.s. $\lim_{a\to\infty} Z\mathbf{1}_{\{Z>a\}} = 0$. Or $0\leq Z\mathbf{1}_{\{Z>a\}}\leq Z$. Donc, si $Z$ est intégrable, le théorème de convergence dominée entraîne que

$$\lim_{a\to\infty}\mathbf{E}[Z\mathbf{1}_{\{Z>a\}}] = 0 \; .$$

Le point $(ii)$ découle alors de cette constatation et du $(i)$. Le point $(iii)$ est une conséquence de l'inégalité suivante

$$\sup_{i\in I}\mathbf{E}\big[\,|X_i|\mathbf{1}_{\{|X_i|\geq a\}}\big] \leq \mathbf{E}[Z\mathbf{1}_{\{Z>a\}}] \xrightarrow[a\to\infty]{} 0 \, .$$

Montrons $(iv)$ : il existe $a_0>0$, tel que $g(x)>0$, pour tout $x>a_0$. Pour tout $a>a_0$, on peut poser $w(a) = \sup_{x>a} x/g(x)$. La fonction $w$ est donc décroissante et $\lim_{a\to\infty} w(a) = 0$. On pose $c = \sup_{i\in I}\mathbf{E}\big[g(|X_i|)\big]$. Pour tout $a > a_0$ et tout $i\in I$, on a donc

$$\mathbf{E}[|X_i|\mathbf{1}_{\{|X_i|>a\}}] = \mathbf{E}[g(|X_i|)|X_i|g(|X_i|)^{-1}\mathbf{1}_{\{|X_i|>a\}}] \leq w(a)\mathbf{E}\big[g(|X_i|)\big] \leq cw(a) \;\; ,$$

ce qui implique bien que $\lim_{a\to\infty}\sup_{i\in I}\mathbf{E}[|X_i|\mathbf{1}_{\{|X_i|>a\}}] = 0$. ■

**Corollaire B.4.14** *Soit $1<p<\infty$ et soit $(X_n)_{n\geq 0}$, une suite bornée en norme $\mathrm{L}^p$, c'est-à-dire que $\sup_{n\geq 0}\mathbf{E}[|X_n|^p]<\infty$. Alors, la suite $(X_n)_{n\geq 0}$ est uniformément intégrable.*

L'utilité de notion d'uniforme intégralité apparaît dans le théorème suivant qui est une généralisation du théorème de convergence dominée.

**Théorème B.4.15** *Soit $(X_n)_{n\geq 0}$, une suite de v.a. réelles $\mathscr{F}$-mesurables. Soit $X$, une v.a. réelle $\mathscr{F}$-mesurable. Il y a équivalence entre les assertions suivantes.*

$(a)$ *La suite $(X_n)_{n\in\mathbb{N}}$ est uniformément intégrable et $\lim_n X_n = X$ en probabilité.*

$(b)$ $\lim_n\|X_n - X\|_1 = 0$.

**Preuve :** on prouve d'abord $(a) \implies (b)$. Le lemme B.4.12 implique que $\sup_{n\in\mathbb{N}}\|X_n\|_1 < \infty$. Par Fatou, on a donc $\mathbf{E}[|X|] \leq \liminf_n \mathbf{E}[|X_n|]<\infty$. Il est clair qu'en ajoutant une variable intégrable à une famille uniformément intégrable, on obtient une famille uniformément intégrable : donc, si on pose

$$w(a) = \max\big(\mathbf{E}[\,|X|\mathbf{1}_{\{|X|>a\}}]\, , \sup_{n\in\mathbb{N}}\mathbf{E}[\,|X_n|\mathbf{1}_{\{|X_n|>a\}}]\big) \; ,$$

on a $\lim_\infty w(a) = 0$. On vérifie ensuite l'inégalité suivante pour tous $x,y\in\mathbb{R}$ et tout $a\in\mathbb{R}_+$ :

$$\forall x,y\in\mathbb{R},\ \forall a\in\mathbb{R}_+,\quad |x-y| \;\leq\; (2a)\wedge|x-y| \;+\; |x|\mathbf{1}_{\{|x|>a\}} \;+\; |y|\mathbf{1}_{\{|y|>a\}}$$

Par ailleurs, pour tout $\varepsilon\in\,]0,2a]$, on a également

$$(2a)\wedge|x-y| = \big((2a)\wedge|x-y|\big)\mathbf{1}_{\{|x-y|\leq\varepsilon\}} \;+\; \big((2a)\wedge|x-y|\big)\mathbf{1}_{\{|x-y|>\varepsilon\}} \;\leq\; \varepsilon + 2a\mathbf{1}_{\{|x-y|>\varepsilon\}} \; .$$

Donc pour tous $x,y\in\mathbb{R}$ et tous $a\geq\varepsilon>0$, on a

$$|x-y| \;\leq\; \varepsilon + 2a\mathbf{1}_{\{|x-y|>\varepsilon\}} + |x|\mathbf{1}_{\{|x|>a\}} + |y|\mathbf{1}_{\{|y|>a\}}$$

Par conséquent,

$$\mathbf{E}[\,|X_n - X|\,] \;\leq\; \varepsilon + 2a\mathbf{P}\big(|X_n-X|>\varepsilon\big) + 2w(a) \; .$$

et on a $\limsup_n\mathbf{E}[\,|X_n-X|\,]\leq\varepsilon + 2w(a)$, ce qui implique le résultat lorsque $a\to\infty$ et $\varepsilon\to 0$.

Montrons que $(b) \implies (a)$ : on suppose donc que $\lim_n\|X_n - X\|_1 = 0$. Les normes $\mathrm{L}^1$ des $X_n$ convergent et forment une suite bornée : on pose $c = \sup_{n\geq 0}\mathbf{E}\big[|X_n|\big]$. Pour tous réels $a,b>0$, et pour tout $n$, on a les inégalités suivantes

$$\begin{aligned} |X_n|\mathbf{1}_{\{|X_n|>a\}} &\leq |X_n - X|\mathbf{1}_{\{|X_n|>a\}} + |X|\mathbf{1}_{\{|X_n|>a\}} \leq |X_n - X| + |X|\mathbf{1}_{\{|X_n|>a\}} \\ &\leq |X_n - X| + |X|\mathbf{1}_{\{|X|\leq b\,;\,|X_n|>a\}} + |X|\mathbf{1}_{\{|X|>b\,;\,|X_n|>a\}} \\ &\leq |X_n - X| + b\mathbf{1}_{\{|X_n|>a\}} + |X|\mathbf{1}_{\{|X|>b\}} \end{aligned}$$

$$\leq \quad |X_n - X| + \tfrac{b}{a}|X_n| + |X|\mathbf{1}_{\{|X|>b\}}.$$

On prend $b = \sqrt{a}$ et on intègre l'inégalité précédente pour obtenir

$$\mathbf{E}\big[|X_n|\mathbf{1}_{\{|X_n|>a\}}\big] \leq \mathbf{E}\big[|X_n - X|\big] + \tfrac{c}{\sqrt{a}} + \mathbf{E}\big[|X|\mathbf{1}_{\{|X|>\sqrt{a}\}}\big] \; .$$

On pose $\phi(a) = \sup_{n\geq 0} \mathbf{E}\big[|X_n|\mathbf{1}_{\{|X_n|>a\}}\big]$ et on fixe $p \in \mathbb{N}$. On a donc

$$\begin{aligned} \phi(a) &\leq \sum_{0\leq k\leq p} \mathbf{E}\big[|X_k|\mathbf{1}_{\{|X_k|>a\}}\big] + \sup_{n\geq p} \mathbf{E}\big[|X_n|\mathbf{1}_{\{|X_n|>a\}}\big] \\ &\leq \sum_{0\leq k\leq p} \mathbf{E}\big[|X_k|\mathbf{1}_{\{|X_k|>a\}}\big] + \sup_{n\geq p} \mathbf{E}\big[|X_n - X|\big] + \tfrac{c}{\sqrt{a}} + \mathbf{E}\big[|X|\mathbf{1}_{\{|X|>\sqrt{a}\}}\big] \; . \end{aligned}$$

On a donc

$$\limsup_{a\to\infty} \phi(a) \; \leq \; \sup_{n\geq p} \mathbf{E}\big[|X_n - X|\big] \xrightarrow[p\to\infty]{} 0 \; ,$$

ce qui montre que $\lim_{a\to\infty} \phi(a) = 0$, et donc que $(X_n)_{n\geq 0}$ est uniformément intégrable. ■

## B.5 Espérance conditionnelle.

**Définition B.5.1** Soit $(\Omega, \mathscr{F}, \mathbf{P})$ un espace de probabilité et soit $\mathscr{G}$, une sous-tribu de $\mathscr{F}$.

**(Variables positives)** Soit $X$ une v.a. $\mathscr{F}$-mesurable à valeurs dans $[0,\infty]$. La fonction $Y : \Omega \to [0,\infty]$ est *une version de l'espérance conditionnelle de $X$ sachant $\mathscr{G}$* si

- la variable $Y$ est $\mathscr{G}$-mesurable,
- pour tout $B \in \mathscr{G}$, on a $\mathbf{E}[Y\mathbf{1}_B] = \mathbf{E}[X\mathbf{1}_B]$.

**(Variables intégrables)** Soit $X$ une v.a. $\mathscr{F}$-mesurable à valeurs réelles intégrable. La fonction $Y : \Omega \to \mathbb{R}$ est *une version de l'espérance conditionnelle de $X$ sachant $\mathscr{G}$* si

- la variable $Y$ est $\mathscr{G}$-mesurable et intégrable,
- pour tout $B \in \mathscr{G}$, on a $\mathbf{E}[Y\mathbf{1}_B] = \mathbf{E}[X\mathbf{1}_B]$. □

Pour que l'on puisse définir l'espérance conditionnelle d'une variable, il faut que son espérance soit bien définie, c'est-à-dire que la variable soit intégrable ou à valeurs dans $[0,\infty]$. Pour simplifier, on dira qu'une variable (ou une fonction) est positive si elle est à valeurs dans $[0,\infty]$. On montre d'abord l'unicité $\mathbf{P}$-preque sûre des espérances conditionnelles grâce à la proposition suivante.

**Proposition B.5.2** *Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité et soit $\mathscr{G}$, une sous-tribu de $\mathscr{F}$. Soient $Y$ et $Y'$, deux v.a. $\mathscr{G}$-mesurables qui sont soit réelles intégrables, soit positives. On suppose que pour tout $B \in \mathscr{G}$, $\mathbf{E}[Y\mathbf{1}_B] \leq \mathbf{E}[Y'\mathbf{1}_B]$ Alors $Y \leq Y'$ presque sûrement.*

**Preuve :** pour tous réels $a < b$, on pose $C_{a,b} = \{Y' \leq a < b \leq Y\}$. On a bien $C_{a,b} = (Y')^{-1}(]-\infty, a]) \cap Y^{-1}([b,\infty]) \in \mathscr{G}$. On remarque ensuite que $Y'\mathbf{1}_{C_{a,b}} \leq a\mathbf{1}_{C_{a,b}} \leq b\mathbf{1}_{C_{a,b}} \leq Y\mathbf{1}_{C_{a,b}}$. En intégrant cette inégalité on a donc $\mathbf{E}[Y'\mathbf{1}_{C_{a,b}}] \leq a\mathbf{P}(C_{a,b}) \leq b\mathbf{P}(C_{a,b}) \leq \mathbf{E}[Y\mathbf{1}_{C_{a,b}}]$. Mais par hypothèse, $\mathbf{E}[Y\mathbf{1}_{C_{a,b}}] \leq \mathbf{E}[Y'\mathbf{1}_{C_{a,b}}]$. Cela montre que $a\mathbf{P}(C_{a,b}) = b\mathbf{P}(C_{a,b})$, et donc que $\mathbf{P}(C_{a,b}) = 0$. Or $\{Y' < Y\} = \bigcup_{a,b\in\mathbb{Q}, a<b} C_{a,b}$. Comme l'ensemble des couples $(a,b) \in \mathbb{Q}^2$ tel que $a < b$ est dénombrable et comme une union dénombrable de $\mathbf{P}$-négligeables est un ensemble $\mathbf{P}$-négligeable, on a $\mathbf{P}(Y' < Y) = 0$, ce qu'il fallait démontrer. ■

**Corollaire B.5.3** *Si $Y$ et $Y'$ sont deux versions de l'espérance conditionnelle de $X$ sachant $\mathscr{G}$, alors $Y = Y'$ presque sûrement.*

**Théorème B.5.4** **(Existence de l'espérance conditionnelle)** *Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité et soit $\mathscr{G}$, une sous-tribu de $\mathscr{F}$. Soit $X$, une v.a. réelle intégrable ou positive. Elle admet une (version de son) espérance conditionnelle sachant $\mathscr{G}$.*

**Preuve :** on fixe $X : \Omega \to [0,\infty]$, une fonction $\mathscr{F}$-mesurable. Pour tout $B \in \mathscr{G}$, on pose $\nu(B) = \mathbf{E}[X\mathbf{1}_B] = \int_B X\, d\mathbf{P}$. Clairement, $\nu : \mathscr{G} \to [0,\infty]$ est une somme de mesures de masse finie telle que $\nu \ll \mathbf{P}$. Le théorème de Radon-Nikodym s'applique et montre qu'il existe $Y : \Omega \to [0,\infty]$, $\mathscr{G}$-mesurable telle que $\nu(B) = \int_B Y\, d\mathbf{P} = \mathbf{E}[Y\mathbf{1}_B]$, ce qui montre le résultat voulu dans le cas des v.a. positives.

Supposons que $X$ soit une v.a. réelle intégrable : on note $X^+$ et $X^-$ ses parties positives et négatives. On note $Y^+$ et $Y^-$ les v.a. $\mathscr{G}$-mesurables telles que $\mathbf{E}[X^{+/-}\mathbf{1}_B] = \mathbf{E}[Y^{+/-}\mathbf{1}_B]$, $B \in \mathscr{G}$. Par conséquent, $\mathbf{E}[Y^{+/-}] = \mathbf{E}[X^{+/-}] < \infty$ et donc p.s. $Y^{+/-} < \infty$. Sans perte de généralité on peut supposer les variables $Y^+$ et $Y^-$ finies partout et on pose $Y = Y^+ - Y^-$, qui est bien (une version de) l'espérance conditionnelle de $X$ sachant $\mathscr{G}$. ■

**Proposition B.5.5** *Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité et soit $\mathscr{G}$, une sous-tribu de $\mathscr{F}$. Soient $X, X_1, X_2$, des v.a. réelles intégrables (resp. positives). Les assertions suivantes sont vérifiées.*

*(i) Si $\mathscr{G} = \{\emptyset, \Omega\}$ (qui est la tribu grossière), alors $\mathbf{E}[X|\mathscr{G}] = \mathbf{E}[X]$.*

*(ii) Si $X$ est $\mathscr{G}$-mesurable, $\mathbf{E}[X|\mathscr{G}] = X$.*

*(iii) Pour tout $c \in \mathbb{R}$ (resp. tout $c \in \mathbb{R}_+$), $\mathbf{E}[X_1 + cX_2|\mathscr{G}] = \mathbf{E}[X_1|\mathscr{G}] + c\mathbf{E}[X_2|\mathscr{G}]$.*

*(iv) Si $X_1 \leq X_2$ p.s., alors $\mathbf{E}[X_1|\mathscr{G}] \leq \mathbf{E}[X_2|\mathscr{G}]$ presque sûrement. Notamment, si $X$ est positive p.s., $\mathbf{E}[X|\mathscr{G}]$ l'est aussi.*

*(v) (Inégalité triangulaire) $\big|\mathbf{E}[X \mid \mathscr{G}]\big| \leq \mathbf{E}\big[|X| \,\big|\, \mathscr{G}\big]$, p.s.*

*(vi) $\mathbf{E}\big[\mathbf{E}[X|\mathscr{G}]\big] = \mathbf{E}[X]$.*

**Preuve :** on note $c = \mathbf{E}[X]$. La v.a. constante à $c$ est clairement mesurable par rapport à la tribu grossière. On vérifie ensuite que $\mathbf{E}[c\mathbf{1}_\emptyset] = 0 = \mathbf{E}[X\mathbf{1}_\emptyset]$ et que $\mathbf{E}[c\mathbf{1}_\Omega] = \mathbf{E}[X\mathbf{1}_\Omega]$. Donc $c = \mathbf{E}[X|\mathscr{G}]$ si $\mathscr{G}$ est la tribu grossière, ce qui prouve $(i)$.

Montrons $(ii)$ : on suppose que $X$ est $\mathscr{G}$-mesurable ; comme $\mathbf{E}[X\mathbf{1}_B] = \mathbf{E}[X\mathbf{1}_B]$, pour tout $B \in \mathscr{G}$, on a donc donc $\mathbf{E}[X|\mathscr{G}] = X$.

Montrons $(iii)$ : on pose $Y = \mathbf{E}[X_1|\mathscr{G}] + c\mathbf{E}[X_2|\mathscr{G}]$. C'est une v.a. $\mathscr{G}$-mesurable intégrable (resp positive). Par linéarité de l'espérance et par définition de l'espérance conditionnelle, on a pour tout $B \in \mathscr{G}$,

$$\mathbf{E}[Y\mathbf{1}_B] = \mathbf{E}\big[\mathbf{E}[X_1|\mathscr{G}]\mathbf{1}_B\big] + c\mathbf{E}\big[\mathbf{E}[X_2|\mathscr{G}]\mathbf{1}_B\big] = \mathbf{E}[X_1\mathbf{1}_B] + c\mathbf{E}[X_2\mathbf{1}_B] = \mathbf{E}\big[(X_1 + cX_2)\mathbf{1}_B\big],$$

ce qui entraîne le point $(iii)$.

Montrons $(iv)$ : on pose $Y = \mathbf{E}[X_1|\mathscr{G}]$ et $Y' = \mathbf{E}[X_2|\mathscr{G}]$. Par définition de l'espérance conditionnelle, pour tout $B \in \mathscr{G}$, on a $\mathbf{E}[Y\mathbf{1}_B] = \mathbf{E}[X_1\mathbf{1}_B]$ et $\mathbf{E}[Y'\mathbf{1}_B] = \mathbf{E}[X_2\mathbf{1}_B]$. Comme $X_1 \leq X_2$ p.s., ce qui précède montre que $\mathbf{E}[Y\mathbf{1}_B] \leq \mathbf{E}[Y'\mathbf{1}_B]$ et a proposition B.5.2 implique donc que $Y \leq Y'$, ce qui est le point $(iv)$.

On en déduit $(v)$ : on note $X^+ = \max(0, X)$ et $X^- = \max(0, -X)$, les parties positive et négative de $X$. On a donc $X = X^+ - X^-$ et $|X| = X^+ + X^-$. On a de plus $\mathbf{E}[X^{+/-}|\mathscr{G}] \geq 0$, p.s. Par conséquent,

$$\big|\mathbf{E}[X|\mathscr{G}]\big| = \big|\mathbf{E}[X^+|\mathscr{G}] - \mathbf{E}[X^-|\mathscr{G}]\big| \leq \mathbf{E}[X^+|\mathscr{G}] + \mathbf{E}[X^-|\mathscr{G}] = \mathbf{E}\big[|X|\,\big|\mathscr{G}\big] \;.$$

Le point $(vi)$ est une conséquence de la définition B.5.1 de l'espérance avec $B = \Omega$. ■

**Lemme B.5.6** *Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité et soit $\mathscr{G}$, une sous-tribu de $\mathscr{F}$. Soit $X$, une v.a. réelle intégrable (resp. positive). On pose $Y = \mathbf{E}[X|\mathscr{G}]$. Alors, pour toute v.a. $Z$ qui est $\mathscr{G}$-mesurable bornée (resp. positive), on a $\mathbf{E}[ZY] = \mathbf{E}[ZX]$.*

**Preuve :** on suppose que $X$ est positive. Par croissance de l'espérance conditionnelle (proposition B.5.5 $(iv)$), $Y$ est positive. Soit $Z$, une v.a. $\mathscr{G}$-mesurable positive. Par le lemme B.1.3, il existe une suite de constantes positive $(c_n)_{n\geq 0}$ et une suite d'événements $B_n \in \mathscr{G}$, $n \in \mathbb{N}$, telles que $Z = \sum_{n\geq 0} c_n\mathbf{1}_{B_n}$. Par interversion série / espérance positive (appliquée deux fois) et par la définition de l'espérance conditionnelle, on a donc

$$\mathbf{E}[ZY] = \mathbf{E}\Big[\sum_{n\geq 0} c_n\mathbf{1}_{B_n}Y\Big] = \sum_{n\geq 0} c_n\mathbf{E}[\mathbf{1}_{B_n}Y] = \sum_{n\geq 0} c_n\mathbf{E}[\mathbf{1}_{B_n}X] = \mathbf{E}\Big[\sum_{n\geq 0} c_n\mathbf{1}_{B_n}X\Big] = \mathbf{E}[ZX] \;,$$

ce qui montre le résultat voulu pour les v.a. positives. Le cas réel intégrable se traite en considérant la partie positive et la partie négative de la variable et en se ramenant au cas positif (les détails sont laissés au lecteur). ■

La proposition B.5.5 discute des propriétés de linéarité et de croissance de l'espérance conditionnelle. La proposition suivante a trait à l'espérance conditionnelle d'un *produit de variables*.

**Proposition B.5.7** *Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité et soit $\mathscr{G}$, une sous-tribu de $\mathscr{F}$. Soit $X$, une v.a. réelle intégrable (resp. positive). Alors, pour toute v.a. $Z$ qui $\mathscr{G}$-mesurable bornée (resp. positive), on a*

$$\text{p.s.} \quad \mathbf{E}[ZX|\mathscr{G}] = Z\mathbf{E}[X|\mathscr{G}] \;. \tag{B.11}$$

*Cette égalité reste vraie si $Z$ est $\mathscr{G}$-mesurable et si $ZX$ et $X$ sont intégrables.*

**Preuve :** on pose $Y = Z\mathbf{E}[X|\mathscr{G}]$, qui est bien $\mathscr{G}$-mesurable. Soit $B \in \mathscr{G}$. On observe que $\mathbf{1}_B Z$ est intégrable (resp. positive). Le lemme B.5.6 implique donc que $\mathbf{E}[\mathbf{1}_B ZX] = \mathbf{E}\big[\mathbf{1}_B Z\mathbf{E}[X|\mathscr{G}]\big]$. Par définition de $Y$, $\mathbf{E}\big[\mathbf{1}_B Z\mathbf{E}[X|\mathscr{G}]\big] = \mathbf{E}[\mathbf{1}_B Y]$, donc $\mathbf{E}[\mathbf{1}_B ZX] = \mathbf{E}[\mathbf{1}_B Y]$, pour tout $B \in \mathscr{G}$, ce qui implique que $Y = \mathbf{E}[ZX|\mathscr{G}]$, et cela prouve (B.11). On suppose que $ZX$ et $X$ sont intégrables. On montre (B.11) dans ce cas en considérant les parties positives et négatives de $X$ et de $Z$ et en utilisant le cas positif déjà démontré. ■

Dans les propositions précédentes la tribu par rapport à laquelle on conditionne est fixée. Les propositions suivantes discutent de quelques propriétés de l'espérance conditionnelle lorsque la tribu de conditionnement varie, ou lorsqu'elle est donnée sous une certaine forme.

**Proposition B.5.8** *Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité. Soient $\mathscr{G}_1$ et $\mathscr{G}_2$, deux sous-tribus de $\mathscr{F}$. On suppose que $\mathscr{G}_1 \subset \mathscr{G}_2$. Alors pour toute v.a. $X$, réelle intégrable ou positive, on a*

$$\text{p.s.} \quad \mathbf{E}\big[\mathbf{E}[X|\mathscr{G}_2]\,|\mathscr{G}_1\big] = \mathbf{E}[X|\mathscr{G}_1] \;.$$

**Preuve :** on pose $Y = \mathbf{E}\big[\mathbf{E}[X|\mathscr{G}_2]\,|\mathscr{G}_1\big]$. C'est une v.a. $\mathscr{G}_1$-mesurable. Soit $B \in \mathscr{G}_1$. La proposition B.5.7 implique que $\mathbf{1}_B Y = \mathbf{E}\big[\mathbf{1}_B \mathbf{E}[X|\mathscr{G}_2]\,|\mathscr{G}_1\big]$. Comme, $\mathscr{G}_1 \subset \mathscr{G}_2$, on a également, $B \in \mathscr{G}_2$. La proposition B.5.7 implique que$\mathbf{1}_B\mathbf{E}[X|\mathscr{G}_2] = \mathbf{E}[\mathbf{1}_B X|\mathscr{G}_2]$ et ce qui précède montre que $\mathbf{1}_B Y = \mathbf{E}\big[\mathbf{E}[\mathbf{1}_B X|\mathscr{G}_2]\,|\mathscr{G}_1\big]$. En appliquant ensuite la proposition B.5.5 $(v)$ deux fois, on a donc

$$\mathbf{E}[\mathbf{1}_B Y] = \mathbf{E}\big[\mathbf{E}\big[\mathbf{E}[\mathbf{1}_B X|\mathscr{G}_2]\,|\mathscr{G}_1\big]\big] = \mathbf{E}\big[\mathbf{E}[\mathbf{1}_B X|\mathscr{G}_2]\,\big] = \mathbf{E}[\mathbf{1}_B X] \;.$$

Comme cela est vrai pour tout $B \in \mathscr{G}_1$, on a bien $Y = \mathbf{E}[X|\mathscr{G}_1]$, qui est le résultat voulu. ■

**Proposition B.5.9** *Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité. Soit $\mathscr{G}_i$, $i \in I$, une famille de sous-tribus de $\mathscr{F}$. On pose $\mathscr{G} = \sigma(\mathscr{G}_i, i \in I)$, la tribu engendrée par les $\mathscr{G}_i$, $i \in I$. Soit $X$, une v.a. réelle intégrable (resp. positive) Il y a équivalence entre les assertions suivantes.*

*(i)* $Y = \mathbf{E}[X|\mathscr{G}]$.

*(ii)* *$Y$ est $\mathscr{G}$-mesurable intégrable (resp. positive) et pour tout sous-ensemble* **fini** *d'indices $J \subset I$, pour tous $B_j \in \mathscr{G}_j$, $j \in J$, on a*

$$\mathbf{E}\big[X\mathbf{1}_{\bigcap_{j\in J} B_j}\big] = \mathbf{E}\big[Y\mathbf{1}_{\bigcap_{j\in J} B_j}\big] \;.$$

**Preuve :** $(i) \implies (ii)$ est trivial. Montrons $(ii) \implies (i)$ dans le cas où $X$ est réelle intégrable. On pose $\mathscr{P} = \{\bigcap_{j\in J} B_j ; B_j \in \mathscr{G}_j \;,\; J \subset I, \; J \text{ fini}\}$. Il est clair que $\Omega \in \mathscr{P}$ et que $\mathscr{P}$ est stable par intersection finie. C'est donc un pi-système. Par ailleurs, il est clair que $\mathscr{P} \subset \mathscr{G}$ et que $\mathscr{G}_i \subset \mathscr{P}$, pour tout $i \in I$. Donc, $\sigma(\mathscr{P}) = \mathscr{G}$. On pose

$$\mathcal{L} = \big\{B \in \mathscr{G} : \mathbf{E}[X\mathbf{1}_B] = \mathbf{E}[Y\mathbf{1}_B]\big\} \;.$$

Le point $(ii)$ signifie que $\mathscr{P} \subset \mathcal{L}$. Montrons que $\mathcal{L}$ est une classe monotone : il est d'abord clair que $\Omega \in \mathcal{L}$, car $\Omega \in \mathscr{P}$. Soient $A, B \in \mathcal{L}$ tels que $A \subset B$. Comme $X$ et $Y$ sont intégrables, pour tout $C \in \mathscr{F}$, $\mathbf{E}[X\mathbf{1}_C]$ et $\mathbf{E}[Y\mathbf{1}_C]$ sont des réels bien définis et on a bien

$$\mathbf{E}[X\mathbf{1}_{B\backslash A}] = \mathbf{E}[X\mathbf{1}_B] - \mathbf{E}[X\mathbf{1}_A] = \mathbf{E}[Y\mathbf{1}_B] - \mathbf{E}[Y\mathbf{1}_A] = \mathbf{E}[Y\mathbf{1}_{B\backslash A}] \;,$$

ce qui montre que $B\backslash A \in \mathcal{L}$. La classe d'événements $\mathcal{L}$ est donc stable par différence propre. Soient $B_n \in \mathcal{L}$, $n \geq 0$, tels que $B_n \subset B_{n+1}$. On pose $B = \bigcup_{n\geq 0} B_n$. Il est clair que $\mathbf{1}_B = \lim_n \mathbf{1}_{B_n}$, et donc $X\mathbf{1}_B = \lim_n X\mathbf{1}_{B_n}$, $Y\mathbf{1}_B = \lim_n Y\mathbf{1}_{B_n}$. De plus $|X\mathbf{1}_{B_n}| \leq |X|$ et $|Y\mathbf{1}_{B_n}| \leq |Y|$, pour tout $n \geq 0$. Comme $X$ et $Y$ sont intégrable, le théorème de convergence dominée s'applique et on a

$$\mathbf{E}[X\mathbf{1}_B] = \lim_n \mathbf{E}[X\mathbf{1}_{B_n}] = \lim_n \mathbf{E}[Y\mathbf{1}_{B_n}] = \mathbf{E}[Y\mathbf{1}_B] \;,$$

ce qui montre que $B \in \mathcal{L}$. La classe d'événements $\mathcal{L}$ est donc stable par union dénombrable croissante. Cela prouve que $\mathcal{L}$ est une classe monotone.

Le théorème de la classe monotone (voit le théorème A.1.10 en appendice) entraîne que $\sigma(\mathscr{P}) \subset \mathcal{L}$. Mais puisque, par définition, $\mathcal{L} \subset \mathscr{G}$ et puisque $\sigma(\mathscr{P}) = \mathscr{G}$, on a donc $\mathcal{L} = \mathscr{G}$, ce qui signifie que pour tout $B \in \mathscr{G}$, on a $\mathbf{E}[X\mathbf{1}_B] = \mathbf{E}[Y\mathbf{1}_B]$. Comme $Y$ est supposée intégrable et $\mathscr{G}$-mesurable, cela entraîne bien $Y = \mathbf{E}[X|\mathscr{G}]$. ■

## Convergence sous l'espérance conditionnelle.

**Proposition B.5.10** *Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité. Soit $\mathscr{G}$, une sous-tribu de $\mathscr{F}$. Soit $(X_n)_{n\geq 0}$, une suite de v.a. positives. Les assertions suivantes sont vérifiées.*

**(Convergence monotone)** *On suppose que $X_n \leq X_{n+1}$, pour tout $n \geq 0$. On pose $X_\infty = \sup_n X_n$. Alors, p.s. $\lim_n \mathbf{E}[X_n|\mathscr{G}] = \mathbf{E}[X_\infty|\mathscr{G}]$.*

**(Interversion positive)** $\mathbf{E}\Big[\sum_{n\geq 0} X_n \,|\, \mathscr{G}\Big] = \sum_{n\geq 0} \mathbf{E}[X_n|\mathscr{G}]$ *p.s.*

**(Fatou)** $\mathbf{E}[\liminf_n X_n \,|\, \mathscr{G}] \leq \liminf_n \mathbf{E}[X_n|\mathscr{G}]$ *p.s.*

**Preuve :** on voit que $X_n \leq X_{n+1}$ implique que $\mathbf{E}[X_n|\mathscr{G}] \leq \mathbf{E}[X_{n+1}|\mathscr{G}]$ (par la proposition B.5.5 $(iv)$). On pose $Y = \sup_n \mathbf{E}[X_n|\mathscr{G}]$, qui est une v.a. positive $\mathscr{G}$-mesurable. Soit $B \in \mathscr{G}$. Le théorème de convergence monotone non-conditionnel ainsi que la définition de l'espérance conditionnelle implique donc que $\mathbf{E}[Y\mathbf{1}_B] = \lim_n \mathbf{E}\big[\mathbf{E}[X_n|\mathscr{G}]\mathbf{1}_B\big] = \lim_n \mathbf{E}[X_n\mathbf{1}_B] = \mathbf{E}[X_\infty\mathbf{1}_B]$, ce qui implique que $Y = \mathbf{E}[X_\infty|\mathscr{G}]$.

L'interversion et Fatou conditionnels se déduisent de la convergence monotone conditionnelle exactement comme l'interversion et Fatou non-conditionnels se déduisent de la convergence monotone usuelle. ■

**Proposition B.5.11 (Convergence dominée)** *Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité. Soit $\mathscr{G}$, une sous-tribu de $\mathscr{F}$. Soit $(X_n)_{n\geq 0}$, une suite de v.a. réelles. On fait les hypothèses suivantes.*

- *Il existe une v.a. réelle $X_\infty$ telle que $\lim_n X_n = X_\infty$ p.s.*
- *Il existe une v.a. positive $Z$ telle que $\mathbf{E}[Z] < \infty$ et $|X_n| \leq Z$, pour tout $n \geq 0$, p.s.*

*Alors, $X_\infty$ est intégrable, p.s. $\lim_n \mathbf{E}\big[\,|X_n - X_\infty|\,\big|\,\mathscr{G}\big] = 0$ et $\lim_n \mathbf{E}[X_n|\mathscr{G}] = \mathbf{E}[X_\infty|\mathscr{G}]$.*

**Preuve :** il est clair que p.s. $|X_\infty| \leq Z$, donc $\mathbf{E}\big[|X_\infty|\big] < \infty$. On pose $Y_n = 2Z - |X_\infty - X_n|$. On a bien $0 \leq Y_n \leq 2Z$ et $\lim_n Y_n = 2Z$ p.s. Par Fatou conditionnel on a

$$\text{p.s.}\quad \mathbf{E}[2Z\,|\mathscr{G}] = \mathbf{E}\big[\liminf_{n\to\infty} Y_n\,|\mathscr{G}\big] \leq \liminf_{n\to\infty} \mathbf{E}[Y_n|\mathscr{G}] = \mathbf{E}[2Z\,|\mathscr{G}] - \limsup_{n\to\infty} \mathbf{E}\big[|X_\infty - X_n|\,\big|\,\mathscr{G}\big]\,,$$

ce qui implique que presque sûrement $\limsup_n \mathbf{E}\big[\,|X_n - X_\infty|\,\big|\,\mathscr{G}\big] \leq 0$. Or pour tout $n \geq 0$, $\mathbf{E}\big[\,|X_n - X_\infty|\,\big|\,\mathscr{G}\big] \geq 0$, par la proposition B.5.5 $(iii)$. Donc $\limsup_n \mathbf{E}\big[\,|X_n - X_\infty|\,\big|\,\mathscr{G}\big] = \lim_n \mathbf{E}\big[\,|X_n - X_\infty|\,\big|\,\mathscr{G}\big] = 0$. Comme $\big|\mathbf{E}[X_\infty|\mathscr{G}] - \mathbf{E}[X_n|\mathscr{G}]\big| \leq \mathbf{E}\big[\,|X_n - X_\infty|\,\big|\,\mathscr{G}\big]$ (voir proposition B.5.5 $(v)$), on en déduit le dernier point de la proposition. ∎

**Corollaire B.5.12 (Interversion $\mathrm{L}^1$)** *Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité. Soit $\mathscr{G}$, une sous-tribu de $\mathscr{F}$. Soit $(X_n)_{n\geq 0}$, une suite de v.a. réelles telle que $\sum_{n\geq 0} \mathbf{E}\big[|X_n|\big] < \infty$. Alors $\sum_{n\geq 0} X_n$ est une v.a. intégrable et on a l'égalité presque-sûre suivante*

$$\mathbf{E}\Big[\sum_{n\geq 0} X_n\,\Big|\,\mathscr{G}\Big] = \sum_{n\geq 0} \mathbf{E}[X_n|\mathscr{G}]\,.$$

**Preuve :** on pose $K = \sum_{n\geq 0} |X_n|$. On a $\mathbf{E}[K] = \sum_{n\geq 0} \mathbf{E}\big[|X_n|\big] < \infty$, par interversion série / espérance positive. Donc, presque sûrement, la série de terme général $X_n$ est absolument convergente et donc convergente : on note $S_\infty = \sum_{n\geq 0} X_n$ sa somme. On pose ensuite $S_n = \sum_{0\leq k\leq n} X_k$. On a bien $|S_n| \leq K$ et p.s. $\lim_n S_n = S_\infty$. On conclut en appliquant le théorème de convergence dominée conditionnel à la suite $(S_n)_{n\geq 0}$. ∎

## Quelques inégalités.

Nous établissons une version "conditionnelle" des inégalités de Jensen et de Hölder. Comme dans le cas déterministe, ces inégalités sont démontrées à l'aide de fonctions convexes. Rappelons qu'une fonction $\varphi : I \to \mathbb{R}$, définie sur un intervalle $I$ de $\mathbb{R}$ est dite *convexe* ssi

$$\forall\, x, y \in I,\ \forall\, \theta \in [0,1] \quad \varphi\big(\theta x + (1-\theta)y\big) \leq \theta\varphi(x) + (1-\theta)\varphi(y)\,.$$

On rappelle que si $\varphi : I \to \mathbb{R}$ est dérivable sur l'intérieur de $I$ et de dérivée croissante, alors $\varphi$ est convexe. Notamment, $\exp$ et $-\log$ sont des fonctions convexes. On rappelle qu'une fonction convexe est l'enveloppe supérieure des droites passant sous son graphe. Plus précisément, on décrit une droite non-verticale par son équation cartésienne qui est de la forme $y = ax + b$, où $a, b \in \mathbb{R}$. On note $S$ l'ensemble des couples $(a,b)$ tels que $ax + b \leq \varphi(x)$ pour tout $x \in I$. On vérifie facilement que $\varphi(x) = \sup_{(a,b)\in S} ax + b$, $x \in I$. On aura besoin d'un résultat légèrement plus fin mais qui s'en déduit très facilement : on peut trouver un ensemble $S_0 \subset S$, qui est **dénombrable** et tel que

$$\forall x \in I\,,\quad \varphi(x) = \sup_{(a,b)\in S_0} ax + b\,. \tag{B.12}$$

**Lemme B.5.13 (Jensen conditionnelle)** *Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité. Soit $X$, une v.a. réelle intégrable. Soit $\varphi : \mathbb{R} \to \mathbb{R}$, une fonction convexe. On suppose que $X$ et $\varphi(X)$ sont intégrables. Alors pour toute sous-tribu $\mathscr{G}$ de $\mathscr{F}$, on a*

$$\textit{p.s.}\qquad \varphi\big(\mathbf{E}[X|\mathscr{G}]\big) \leq \mathbf{E}[\varphi(X)|\mathscr{G}]\,.$$

**Preuve :** on utilise la notation $S_0$ introduite avant l'énoncé du lemme. Si $(a,b) \in S_0$, alors $aX + b \leq \varphi(X)$. La croissante et la linéarité de l'espérance conditionnelle (voir la proposition B.5.5 $(iii)$ et $(iv)$) impliquent que $a\mathbf{E}[X|\mathscr{G}] + b \leq \mathbf{E}[\varphi(X)|\mathscr{G}]$, presque sûrement. Comme $S_0$ est dénombrable, on en déduit que

$$\mathbf{P}\text{-p.s.}\quad \forall (a,b) \in S_0\,,\ a\mathbf{E}[X|\mathscr{G}] + b \leq \mathbf{E}[\varphi(X)|\mathscr{G}]$$

et par (B.12) on a donc $\mathbf{P}$-p.s. $\varphi\big(\mathbf{E}[X|\mathscr{G}]\big) = \sup_{(a,b)\in S_0}\big(a\mathbf{E}[X|\mathscr{G}] + b\big) \leq \mathbf{E}[\varphi(X)|\mathscr{G}]$, ce qui permet de conclure. ∎

**Corollaire B.5.14** *Soit $(\Omega, \mathscr{F}, \mathbf{P})$, un espace de probabilité et soit $\mathscr{G}$, une sous-tribu. Soit $X$, une v.a. et soit $p \in [1, \infty[$. Si $\mathbf{E}[\,|X|^p\,] < \infty$, alors $\mathbf{E}[|X|] < \infty$. De plus, p.s. on a*

$$|\mathbf{E}[X|\mathscr{G}]\,|^p \leq \mathbf{E}\big[|X|^p\,|\mathscr{G}\big]\,, \tag{B.13}$$

*ce qui implique que $\mathbf{E}[X|\mathscr{G}] \in \mathrm{L}^p$.*

**Preuve :** on remarque que la fonction $x \mapsto |x|^p$ est convexe. ∎

**Espérance conditionnelle et indépendance.**

**Proposition B.5.15** *Soit $\mathscr{G}$, une sous-tribu de $\mathscr{F}$. Soit $X$, v.a. réelle intégrable ou positive indépendante de $\mathscr{G}$. Alors p.s.* $\mathbf{E}[X|\mathscr{G}] = \mathbf{E}[X]$.

**Preuve :** on traite le cas réel intégrable, le cas positif étant similaire. On pose $c = \mathbf{E}[X]$ et on fixe $B \in \mathscr{G}$. Dire que $X$ est indépendante de $\mathscr{G}$ signifie que $\sigma(X)$ est indépendante de $\mathscr{G}$, et donc $X$ et $\mathbf{1}_B$ sont deux v.a. indépendantes. Par le lemme B.2.12, on a $\mathbf{E}[\mathbf{1}_B X] = \mathbf{E}[\mathbf{1}_B]\mathbf{E}[X] = \mathbf{E}[c\mathbf{1}_B]$. Comme la variable constante $c$ est $\mathscr{G}$-mesurable on en déduit le résultat. ■

**Corollaire B.5.16** *Soit $\mathscr{G}$, une sous-tribu de $\mathscr{F}$. Soient $X_1, \ldots, X_n$, des v.a. réelles intégrables. On les suppose mutuellement indépendantes. On suppose également que le vecteur aléatoire $(X_1, \ldots, X_n)$ est indépendant de $\mathscr{G}$. Alors*

$$\textit{p.s.} \quad \mathbf{E}\big[X_1 \ldots X_n|\mathscr{G}\big] = \mathbf{E}[X_1] \ldots \mathbf{E}[X_n] \; .$$

**Preuve :** par définition, dire que le vecteur $(X_1, \ldots, X_n)$ est indépendant de $\mathscr{G}$ signifie que les tribus $\sigma(X_1, \ldots, X_n)$ et $\mathscr{G}$ sont indépendantes. Comme le produit $X := X_1 \ldots X_n$ est une fonction déterministe du vecteur $(X_1, \ldots, X_n)$, il est $\sigma(X_1, \ldots, X_n)$-mesurable. Cela implique que $X$ est indépendante de $\mathscr{G}$. Le lemme B.2.12 montre que $X$ est intégrable telle que $\mathbf{E}[X] = \mathbf{E}[X_1] \ldots \mathbf{E}[X_n]$ et la proposition B.5.15 précédente montre que $\mathbf{E}[X|\mathscr{G}] = \mathbf{E}[X]$, ce qui permet de conclure. ■

On généralise la proposition B.5.15 de la façon suivante.

**Proposition B.5.17** *Soient $\mathscr{G}_1$ et $\mathscr{G}_1$, deux sous-tribus de $\mathscr{F}$. Soit $X$, une v.a. réelle intégrable. On suppose que $\mathscr{G}_2$ est indépendante de $\sigma(\sigma(X), \mathscr{G}_1)$. Alors, presque sûrement*

$$\mathbf{E}[X \mid \sigma(\mathscr{G}_1, \mathscr{G}_2)] = \mathbf{E}[X|\mathscr{G}_1] \; .$$

**Preuve :** soit $B_1 \in \mathscr{G}_1$ et $B_2 \in \mathscr{G}_2$. Comme $\mathscr{G}_2$ est indépendante de $\sigma(\sigma(X), \mathscr{G}_1)$, la variable $\mathbf{1}_{B_2}$ est indépendante de $X\mathbf{1}_{B_1}$ et le lemme B.2.12 implique que $\mathbf{E}[X\mathbf{1}_{B_1}\mathbf{1}_{B_2}] = \mathbf{E}[X\mathbf{1}_{B_1}]\mathbf{E}[\mathbf{1}_{B_2}]$. La proposition B.5.7 entraîne ensuite que $\mathbf{E}[X\mathbf{1}_{B_1}] = \mathbf{E}\big[\mathbf{E}[X|\mathscr{G}_1]\mathbf{1}_{B_1}\big]$. Par ailleurs, $\mathscr{G}_2$ est indépendante de $\mathscr{G}_1$, donc $\mathbf{1}_{B_2}$ est indépendante de la variable $\mathscr{G}_1$-mesurable $\mathbf{1}_{B_1}\mathbf{E}[X|\mathscr{G}_1]$ et le lemme B.2.12 entraîne que $\mathbf{E}\big[\mathbf{1}_{B_1}\mathbf{E}[X|\mathscr{G}_1]\big]\mathbf{E}[\mathbf{1}_{B_2}] = \mathbf{E}\big[\mathbf{1}_{B_1}\mathbf{E}[X|\mathscr{G}_1]\mathbf{1}_{B_2}\big]$. On a donc montré que

$$\forall B_1 \in \mathscr{G}_1 \, , \; \forall B_2 \in \mathscr{G}_2 \, , \quad \mathbf{E}[X\mathbf{1}_{B_1 \cap B_2}] = \mathbf{E}\big[\mathbf{E}[X|\mathscr{G}_1]\mathbf{1}_{B_1 \cap B_2}\big] \; .$$

Comme $\mathbf{E}[X|\mathscr{G}_1]$ est $\sigma(\mathscr{G}_1, \mathscr{G}_2)$-mesurable la proposition B.5.9 permet de conclure. ■

# B.6 Martingales à temps discret.

Toutes les variables mentionnées dans cette section sont définies sur le même espace de probabilité $\big(\Omega, \mathscr{F}, \mathbf{P}\big)$. Pour tous $m, n \in \mathbb{Z}$ tels que $m < n$, on pose $[\![m, n]\!] := [m, n] \cap \mathbb{Z}$, avec une notation évident pour $]\!]m, n]\!]$, $]\!]m, n[\![$ et $[\![m, n[\![$. On rappelle la définition des martingales, sur-martingales et sous-martingales à temps discret.

**Définition B.6.1** (*Martingales à temps discret*) Soit $(\mathscr{F}_n)_{n \in \mathbb{N}}$ une *filtration en temps discret* définie sur $(\Omega, \mathscr{F})$, (c'est-à-dire que pour tout $n \in \mathbb{N}$, $\mathscr{F}_n$ est une sous-tribu de $\mathscr{F}$ et $\mathscr{F}_n \subset \mathscr{F}_{n+1}$). Pour tout $n \in \mathbb{N}$, soit $M_n$, une v.a. réelle. Alors $(M_n)_{n \in \mathbb{N}}$ est une *martingale* (resp. *sur-martingale*, *sous-martingale*) *relativement à* $(\mathscr{F}_n)_{n \in \mathbb{N}}$ si pour tout $n \in \mathbb{N}$, (a) : $M_n$ est $\mathscr{F}_n$-mesurable, (b) : $\mathbf{E}\big[|M_n|\big] < \infty$ et (c) : $\mathbf{P}$-p.s.

$$\mathbf{E}\big[M_{n+1}\big|\mathscr{F}_n\big] = M_n \quad \text{(resp. } \mathbf{E}\big[M_{n+1}\big|\mathscr{F}_n\big] \leq M_n, \; \mathbf{E}\big[M_{n+1}\big|\mathscr{F}_n\big] \geq M_n \text{ ).}$$

Lorsque la suite $(M_n)_{n \in [\![0, n_0]\!]}$ est finie, on parle de martingale (resp. sur-martinagle, sous-martingale) *à horizon fini*, l'horizon (de temps) étant $n_0$. □

Le lemme suivant fait la liste des propriétés élémentaires des martingales à temps discret.

**Lemme B.6.2** *Soit $(\mathscr{F}_n)_{n \in \mathbb{N}}$ une filtration sur $(\Omega, \mathscr{F})$. Soit $M := (M_n)_{n \in \mathbb{N}}$, une suite de v.a. réelles. Les propriétés suivantes sont vérifiées.*

*(i) Si $M$ est une sur-martingale (reps. une sous-martingale) relativement à $(\mathscr{F}_n)_{n \in \mathbb{N}}$, alors $-M$ est une sous-martingale (resp. sur-martingale) relativement à la même filtration.*

(ii) *Soit $\varphi : \mathbb{R} \to \mathbb{R}$, une fonction convexe telle que les v.a. $\varphi(M_n)$ soient intégrables. Si $M$ est une $(\mathscr{F}_n)_{n\in\mathbb{N}}$-martingale, alors $\varphi(M) := (\varphi(M_n))_{n\in\mathbb{N}}$ est une sous-martingale. De plus, si $\varphi$ est convexe croissante et si $M$ est une $(\mathscr{F}_n)_{n\in\mathbb{N}}$-sous-martingale, alors $\varphi(M)$ est également une $(\mathscr{F}_n)_{n\in\mathbb{N}}$-sous-martingale. Des énoncés analogues sont vrais pour les fonctions concaves et les sur-martingales.*

(iii) *On suppose que $M$ est une martingale (resp. sur-martingale, sous-martingale) relativement à $(\mathscr{F}_n)_{n\in\mathbb{N}}$. Alors pour tout $m, n \in \mathbb{N}$, on a*

$$\mathbf{E}\big[M_{n+m}\big|\mathscr{F}_n\big] = M_n \quad (\textit{resp. } \mathbf{E}\big[M_{n+m}\big|\mathscr{F}_n\big] \leq M_n, \ \mathbf{E}\big[M_{n+m}\big|\mathscr{F}_n\big] \geq M_n \text{ ).}$$

*En prenant l'espérance, cela montre que $(\mathbf{E}[M_n])_{n\in\mathbb{N}}$ est une suite réelle constante (resp. décroissante, croissante).*

**Preuve :** le point $(i)$ est une conséquence évident de la linéarité de l'espérance conditionnelle. Le point $(ii)$ est une conséquence immédiate de l'inégalité de Jensen conditionnelle : si $M$ est une $(\mathscr{F}_n)_{n\in\mathbb{N}}$-martingale, alors, pour tout $n \in \mathbb{N}$, $\mathbf{P}$-p.s.

$$\mathbf{E}\big[\varphi(M_{n+1})\big|\mathscr{F}_n\big] \geq \varphi\big(\mathbf{E}[M_{n+1}|\mathscr{F}_n]\big) = \varphi(M_n) \ .$$

Le cas où $\varphi$ est croissante et convexe et où $M$ est une sous-martingale se traite de la même manière. Le point $(iii)$ découle d'une récurrence facile. ■

**Proposition B.6.3 (Décomposition de Doob)** *Soit $(X_n)_{n\in\mathbb{N}}$, une suite de v.a. réelles intégrables. Soit une filtration $(\mathscr{F}_n)_{n\in\mathbb{N}}$ sur $(\Omega, \mathscr{F})$. On suppose que pour tout $n \in \mathbb{N}$, $X_n$ est $\mathscr{F}_n$-mesurable, Il existe une* **unique** *paire de suites de v.a. $(M_n)_{n\in\mathbb{N}}$, $(V_n)_{n\in\mathbb{N}}$ satisfaisant les propriétés suivantes.*

(i) *$X_n = X_0 + M_n + V_n$, $n \in \mathbb{N}$.*

(ii) *$(M_n)_{n\in\mathbb{N}}$ est une $(\mathscr{F}_n)_{n\in\mathbb{N}}$-martingale.*

(iii) *$V_0 = 0$ et pour tout $n \in \mathbb{N}^*$, $V_n$ est $\mathscr{F}_{n-1}$-mesurable et intégrable.*

*De plus, $(X_n)_{n\in\mathbb{N}}$ est une martingale resp. sur, sous-martingale) ssi $(V_n)_{n\in\mathbb{N}}$ est nul (resp. croissant, décroissant).*

**Preuve :** on suppose d'abord qu'il existe $(M_n)_{n\in\mathbb{N}}$, $(V_n)_{n\in\mathbb{N}}$ satisfaisant $(i)$, $(ii)$ et $(iii)$. On remarque que $V_{n+1} - V_n$ est $\mathscr{F}_n$-mesurable. Pour tout $n \in \mathbb{N}$, on a donc

$$\begin{aligned}\mathbf{E}[X_{n+1} - X_n \,|\, \mathscr{F}_n] &= \mathbf{E}[M_{n+1} - M_n \,|\, \mathscr{F}_n] + \mathbf{E}[V_{n+1} - V_n \,|\, \mathscr{F}_n] \\ &= \mathbf{E}[M_{n+1}|\mathscr{F}_n] - M_n + V_{n+1} - V_n = V_{n+1} - V_n.\end{aligned}$$

Par conséquent, on a nécessairement $M_0 = V_0 = 0$ et pour tout $n \in \mathbb{N}^*$,

$$V_n = \sum_{1\leq k\leq n} \mathbf{E}[X_k - X_{k-1} \,|\, \mathscr{F}_{k-1}] \quad \text{et} \quad M_n = X_n - V_n \ . \tag{B.14}$$

Cela montre l'unicité. Par ailleurs si on définit $M$ et $V$ par (B.14), on vérifie facilement $(i)$, $(ii)$ et $(iii)$. Le reste de la proposition est élémentaire. ■

**Inégalités fondamentales sur les martingales.** On rappelle la définition **??**, page **??** du nombre de montées.

**Théorème B.6.4 (Inégalité de Doob)** *Soit $(M_n)_{n\in\mathbb{N}}$ une sur-martingale relativement à la filtration $(\mathscr{F}_n)_{n\in\mathbb{N}}$ Alors, pour tout $a, b \in \mathbb{R}$ tels que $a < b$, on a*

$$(b-a)\mathbf{E}\big[U_{[\![0,n_0]\!]}(M, [a,b])\big] \leq \mathbf{E}\big[(M_{n_0} - a)_-\big] \tag{B.15}$$

*où $U_{[\![0,n_0]\!]}(M, [a,b])$ désigne le nombre de montées de $M : [\![0, n_0]\!] \to \mathbb{R}$ de l'intervalle $[a, b]$ (voir la définition **??**, page **??**) et où $(\cdot)_-$ désigne la fonction partie négative.*

**Preuve :** on pose $\tau_0 = -1$ et pour tout $p \in \mathbb{N}^*$, on définit les variables aléatoires $\sigma_p$ et $\tau_p$ récursivement par

$$\sigma_p = \inf\big\{n \in ]\!]\tau_{p-1}, n_0]\!] : M_n < a\big\} \quad \text{et} \quad \tau_p = \inf\big\{n \in ]\!]\sigma_p, n_0]\!] : M_n > b\big\}$$

avec la convention que $\inf \emptyset = \infty$. On voit que $U_{[\![0,n_0]\!]}(M, [a,b]) = \inf\{p \in \mathbb{N} : \tau_p < \infty\}$. Pour simplifier les notations, on pose $U := U_{[\![0,n_0]\!]}(M, [a,b])$. Pour tout $p \leq U$, entre les instants $\sigma_p$ et $\tau_p$, $M$ effectue une montée de l'intervalle $[a, b]$. Par ailleurs, il se peut que $\sigma_{U+1} \leq n_0$ mais, par définition on a $\tau_{U+1} = \infty$. On pose ensuite

$$\forall n \in [\![1, n_0]\!], \quad C_n = \sum_{p\in\mathbb{N}^*} \mathbf{1}_{\{\sigma_p < n \leq \tau_p\}} \ . \tag{B.16}$$

Cette somme est en réalité finie car $C_n = \sum_{1\leq p\leq U} \mathbf{1}_{\{\sigma_p<n\leq\tau_p\}} \quad + \mathbf{1}_{\{\sigma_{U+1}<n\leq n_0\}}$. On voit que $C_n = 1$ si $n$ est dans intervalle de temps pendant lequel $M$ effectue une montée de $[a, b]$ et $C_n = 0$ sinon. On pose ensuite

$$Z_0 := 0 \quad \text{et} \quad \forall n \in [\![1, n_0]\!], \quad Z_n := \sum_{1\leq k\leq n} C_k\big(M_k - M_{k-1}\big) \; .$$

On observe alors que

$$\begin{aligned} Z_{n_0} &= \sum_{1\leq p\leq U} \sum_{1\leq k\leq n_0} (M_k - M_{k-1})\mathbf{1}_{\{\sigma_p<k\leq\tau_p\}} \quad + \sum_{1\leq k\leq n_0} (M_k - M_{k-1})\mathbf{1}_{\{\sigma_{U+1}<k\leq n_0\}} \\ &= \sum_{1\leq p\leq U} \big(M_{\tau_p} - M_{\sigma_p}\big) + (M_{n_0} - M_{\sigma_{U+1}})\mathbf{1}_{\{\sigma_{U+1}\leq n_0\}}. \end{aligned} \tag{B.17}$$

On remarque ensuite que si $p \in [\![1, U]\!]$, $M_{\tau_p} - M_{\sigma_p} \geq b - a$. De plus si $\sigma_{U+1} \leq n_0$, $M_{\sigma_{U+1}} < a$. On a donc

$$(M_{n_0} - M_{\sigma_{U+1}})\mathbf{1}_{\{\sigma_{U+1}\leq n_0\}} \leq -(M_{n_0} - a)_- \; .$$

On déduit donc de (B.17)

$$Z_{n_0} \geq (b-a)U - (M_{n_0} - a)_- \; . \tag{B.18}$$

On revient à la définition des temps $\sigma_p$ et $\tau_p$ : il est facile de montrer par récurrence sur $p$ que

$$\forall p \in \mathbb{N}^*, \forall n \in [\![0, n_0]\!], \quad \{\sigma_p \leq n\} \in \mathscr{F}_n \quad \text{et} \quad \{\tau_p \leq n\} \in \mathscr{F}_n \; . \tag{B.19}$$

On remarque que $\{\sigma_p < n \leq \tau_p\} = \{\sigma_p \leq n-1\} \cap (\Omega \backslash \{\tau_p \leq n-1\}) \in \mathscr{F}_{n-1}$. Ce fait combiné avec la définition (B.16) implique que $C_n$ est $\mathscr{F}_{n-1}$ mesurable. On en déduit que pour tout $n \in [\![1, n_0]\!]$, $\mathbf{P}$-p.s.

$$\mathbf{E}\big[Z_n\big|\mathscr{F}_{n-1}\big] = Z_{n-1} + \mathbf{E}\big[C_n(M_n - M_{n-1})\big|\mathscr{F}_{n-1}\big] = Z_{n-1} + C_n\big(\mathbf{E}[M_{n-1}|\mathscr{F}_{n-1}] - M_{n-1}\big) \leq Z_{n-1}$$

car $C_n \geq 0$, et $\mathbf{P}$-p.s. $\mathbf{E}[M_{n-1}|\mathscr{F}_{n-1}] \leq M_{n-1}$, puisque $M$ est une sur-martingale. On en déduit donc que $\mathbf{E}[Z_n] \leq \mathbf{E}[Z_{n-1}]$ : la suite $n \mapsto \mathbf{E}[Z_n]$ est décroissante et on a $\mathbf{E}[Z_{n_0}] \leq \mathbf{E}[Z_0] = 0$, ce qui implique le résultat voulu en prenant l'espérance de (B.18). ■

**Théorème B.6.5 (Inégalité maximale des sous-martingales positives)** *Soit $(M_n)_{n\in\mathbb{N}}$ une sous-martingale relativement à la filtration $(\mathscr{F}_n)_{n\in\mathbb{N}}$. On suppose également que pour tout $n \in \mathbb{N}$, $\mathbf{P}$-p.s. $M_n \geq 0$. Alors, pour tout $a \in \mathbb{R}_+$ et tout $n_0 \in \mathbb{N}$,*

$$a\,\mathbf{P}\big(\sup_{0\leq n\leq n_0} M_n \geq a\big) \leq \mathbf{E}\big[M_{n_0}\mathbf{1}_{\{\sup_{0\leq n\leq n_0} M_n\geq a\}}\big] \; .$$

**Preuve :** on définit ensuite les événements suivants

$$\forall n \in [\![0, n_0]\!], \quad B_n = \{M_0 < a\} \cap \{M_1 < a\} \cap \ldots \cap \{M_{n-1} < a\} \cap \{M_n \geq a\}.$$

Il est clair que $B_n \in \mathscr{F}_n$. Le lemme B.6.2 $(iii)$ implique $\mathbf{P}$-p.s. que

$$\mathbf{E}\big[M_{n_0}\mathbf{1}_{B_n}|\mathscr{F}_n\big] = \mathbf{1}_{B_n}\mathbf{E}\big[M_{n_0}\,|\,\mathscr{F}_n\big] \geq \mathbf{1}_{B_n}M_n \geq a\mathbf{1}_{B_n}.$$

En prenant l'espérance de cette inégalité, obtient

$$\forall n \in [\![0, n_0]\!], \quad \mathbf{E}\big[M_{n_0}\mathbf{1}_{B_n}\big] \geq a\mathbf{P}(B_n). \tag{B.20}$$

On pose $B = \bigcup_{0\leq n\leq n_0} B_n$. On remarque que $B = \big\{\sup_{0\leq n\leq n_0} M_n \geq a\big\}$. Comme les $B_n$ sont disjoint-deux-à-deux, on a $\mathbf{1}_B = \sum_{0\leq n\leq n_0} \mathbf{1}_{B_n}$, ce qui, combiné avec (B.20), implique

$$\begin{aligned} \mathbf{E}\big[M_{n_0}\mathbf{1}_{\{\sup_{0\leq n\leq n_0} M_n\geq a\}}\big] &\geq \mathbf{E}\Big[M_{n_0}\sum_{0\leq n\leq n_0}\mathbf{1}_{B_n}\Big] = \sum_{0\leq n\leq n_0} \mathbf{E}\big[M_{n_0}\mathbf{1}_{B_n}|\big] \\ &\geq \sum_{0\leq n\leq n_0} a\mathbf{P}(B_n) = a\mathbf{P}\Big(\bigcup_{0\leq n\leq n_0} B_n\Big) = a\,\mathbf{P}\big(\sup_{0\leq n\leq n_0} M_n \geq a\big) \; , \end{aligned}$$

ce qui permet de conclure. ■

On déduit de l'inégalité maximale pour les sous-martingales positive, les inégalité $\mathrm{L}^p$ de Doob. On note $\|X\|_p := \mathbf{E}\big[|X|^p\big]^{1/p}$, la norme $\mathrm{L}^p$ sur l'espace $(\Omega, \mathscr{F}, \mathbf{P})$.

**Théorème B.6.6 (Inégalité $\mathrm{L}^p$)** *Soit $p \in ]1, \infty[$. Soit $(M_n)_{n\in\mathbb{N}}$, une martingale relativement à la filtration $(\mathscr{F}_n)_{n\geq 0}$. On suppose qu'elle est bornée en norme $\mathrm{L}^p$, c'est-à-dire que $\sup_{n\in\mathbb{N}} \mathbf{E}[\,|M_n|^p] < \infty$. Alors,*

$$\mathbf{E}\big[\sup_{n\in\mathbb{N}} |M_n|^p\big] \leq \big(\tfrac{p}{p-1}\big)^p \sup_{n\in\mathbb{N}} \mathbf{E}\big[\,|M_n|^p\,\big] < \infty\ , \tag{B.21}$$

*c'est-à-dire $\|S\|_p \leq \frac{p}{p-1} \sup_{n\geq 0} \|M_n\|_p$, où on a posé $S = \sup_{n\geq 0} |M_n|$.*

**Preuve :** comme la fonction valeur absolue est convexe, le lemme B.6.2 $(ii)$ implique que $(|M_n|)_{n\geq 0}$ est une sous-martingale positive formée de variables intégrables. On pose $S_n = \sup_{0\leq m\leq n} |M_m|$, et l'inégalité maximale (théorème B.6.5) entraîne que

$$\forall a \in \mathbb{R}_+\,,\ \forall n \in \mathbb{N}\,, \quad a\, \mathbf{E}[\mathbf{1}_{\{S_n > a\}}] \leq \mathbf{E}[\,|M_n|\mathbf{1}_{\{S_n>a\}}]\ . \tag{B.22}$$

Par Fubini-positif, on a

$$\int_0^\infty a\, \mathbf{E}[\mathbf{1}_{\{S_n>a\}}]\, pa^{p-2}\, da = \mathbf{E}\Big[\int_0^{S_n} pa^{p-1}\, da\Big] = \mathbf{E}[S_n^p].$$

De même, on a

$$\int_0^\infty \mathbf{E}[\,|M_n|\mathbf{1}_{\{S_n>a\}}]\, pa^{p-2}\, da = \mathbf{E}\Big[\,|M_n| \int_0^{S_n} pa^{p-2}\, da\Big] = \tfrac{p}{p-1} \cdot \mathbf{E}[\,|M_n| S_n^{p-1}]\ .$$

Les deux égalités précédentes et (B.22) entraînent que

$$\mathbf{E}[S_n^p] \leq \tfrac{p}{p-1} \mathbf{E}\big[|M_n|\, S_n^{p-1}\big]\ .$$

On note $q$ l'exposant conjugué de $p$ : $q = p/(p-1)$. Par Hölder et l'inégalité précédente,

$$\|S_n\|_p^p \leq \tfrac{p}{p-1} \|M_n\|_p \big(\mathbf{E}[\,S_n^{q(p-1)}]\big)^{\frac{1}{q}} = \tfrac{p}{p-1} \|M_n\|_p \big(\mathbf{E}[\,S_n^p]\big)^{1-\frac{1}{p}} = \tfrac{p}{p-1} \|M_n\|_p \|S_n\|_p^{p-1}. \tag{B.23}$$

Comme $S_n \leq |M_0| + \ldots + |M_n|$, on a $\|S_n\|_p \leq \|M_0\|_p + \ldots + \|M_n\|_p < \infty$, et par (B.23), on a $\|S_n\|_p \leq \frac{p}{p-1} \|M_n\|_p$. On conclut par convergence monotone car $\lim_n \uparrow S_n = S$. ■

Nous prouvons ensuite une inégalité similaire pour les sur-martingales.

**Théorème B.6.7 (Inégalité maximale des sur-martingales)** *Soit $(M_n)_{n\in\mathbb{N}}$ une sur-martingale relativement à la filtration $(\mathscr{F}_n)_{n\in\mathbb{N}}$. Alors, pour tout $a \in \mathbb{R}_+$ et tout $n_0 \in \mathbb{N}$,*

$$a\, \mathbf{P}\big(\sup_{0\leq n\leq n_0} |M_n| \geq 3a\big) \leq 4\mathbf{E}\big[|M_0|\big] + 3\mathbf{E}\big[|M_{n_0}|\big]\ .$$

**Preuve :** on pose $V_0 := 0$ et $M_0^* := 0$ et pour tout $n \in \mathbb{N}^*$,

$$V_n := \sum_{1\leq k\leq n} \mathbf{E}\big[M_k - M_{k-1} | \mathscr{F}_{k-1}\big] \quad \text{et} \quad M_n^* := M_n - M_0 - V_n\ .$$

La proposition B.6.3 montre que $(M_n^*)_{n\in\mathbb{N}}$ est une $(\mathscr{F}_n)_{n\in\mathbb{N}}$, que $(V_n)_{n\in\mathbb{N}}$ est décroissante et $V_n$ est $\mathscr{F}_{n-1}$-mesurable intégrable, pour tout $n \in \mathbb{N}^*$. Autrement dit la paire $(M_n^*)_{n\in\mathbb{N}}$, $(V_n)_{n\in\mathbb{N}}$ est la décomposition de Doob de la sur-martingale $(M_n)_{n\in\mathbb{N}}$. Comme $M_n = M_0 + M_n^* + V_n$ et comme $(V_n)_{n\in\mathbb{N}}$ est $\mathbf{P}$-p.s. décroissante

$$\sup_{0\leq n\leq n_0} |M_n| \leq |M_0| + \sup_{0\leq n\leq n_0} |M_n^*| + |V_{n_0}|$$

On a donc

$$\begin{aligned} a\, \mathbf{P}\big(\sup_{0\leq n\leq n_0} |M_n| \geq 3a\big) &\leq a\mathbf{P}(|M_0| \geq a) + a\, \mathbf{P}\big(\sup_{0\leq n\leq n_0} |M_n^*| \geq a\big) + a\mathbf{P}(|V_{n_0}| \geq a) \\ &= \mathbf{E}\big[|M_0|\big] + a\, \mathbf{P}\big(\sup_{0\leq n\leq n_0} |M_n^*| \geq a\big) + \mathbf{E}\big[|V_{n_0}|\big], \end{aligned} \tag{B.24}$$

par l'inégalité de Markov : $\mathbf{P}(|Y| > a) \leq \mathbf{E}\big[|Y|\big]/a$, qui est valable pour tout réel $a$ strictement positif et toute v.a. $Y$ intégrable. Comme $M^*$ est une martingale, le lemme B.6.2 $(iii)$ implique que $\mathbf{P}$-p.s. $\mathbf{E}[M_{n_0}^*] = 0$ car $M_0^* = 0$, par définition. Comme $V$ est $\mathbf{P}$-p.s. décroissant négatif, cela implique :

$$\mathbf{E}\big[|V_{n_0}|\big] = -\mathbf{E}[V_{n_0}] = -\mathbf{E}\big[M_{n_0} - M_{n_0}^* - M_0\big] = -\mathbf{E}[M_{n_0}] + \mathbf{E}[M_0] \leq \mathbf{E}\big[|M_0|\big] + \mathbf{E}\big[|M_{n_0}|\big]. \tag{B.25}$$

La fonction valeur absolue étant convexe et $M^*$ étant une martingale, le lemme B.6.2 $(ii)$ implique que $(|M_n^*|)_{n\in\mathbb{N}}$ est une sous-martinagle positive. L'inégalité maximale des sous-martingales positives (théorème B.6.5) implique que

$$a\,\mathbf{P}\big(\sup_{0\leq n\leq n_0}|M_n^*|\geq a\big)\leq \mathbf{E}\big[|M_{n_0}^*|\big]. \tag{B.26}$$

On remarque ensuite que

$$\begin{aligned}\mathbf{E}\big[|M_{n_0}^*|\big] &= \mathbf{E}\big[|M_{n_0}-M_0-V_{n_0}|\big]\leq \mathbf{E}\big[|M_{n_0}|\big]+\mathbf{E}\big[|M_0|\big]+\mathbf{E}\big[|V_{n_0}|\big]\\ &\leq 2\mathbf{E}\big[|M_0|\big]+2\mathbf{E}\big[|M_{n_0}|\big],\end{aligned} \tag{B.27}$$

par (B.25). En combinant (B.24), (B.25), (B.26) et (B.27), on obtient le résultat voulu. ■

**Convergence des martingales.** Nous déduisons des inégalités précédentes les théorèmes classiques de convergence des martingales : convergence presque-sûre, en norme $\mathrm{L}^1$ et en norme $\mathrm{L}^p$.

**Théorème B.6.8 (Conv. p.s. des sur/sous-martingales).** *Soit $M=(M_n)_{n\geq 0}$ une sous-martingale ou une sur-martingale relativement à $(\mathscr{F}_n)_{n\geq 0}$. On suppose que*

$$c:=\sup_{n\in\mathbb{N}}\mathbf{E}\big[|M_n|\big]<\infty\ . \tag{B.28}$$

*Alors, il existe une v.a. réelle $M_\infty$, qui est $\mathscr{F}_\infty$-mesurable et intégrable telle que $\mathbf{P}$-p.s. $\lim_{n\to\infty}M_n=M_\infty$.*

**Preuve :** quitte à considérer $-M$, on se ramène sans perte de généralité à suppose que $M$ est une sur-martingale. Soient $a,b\in\mathbb{R}$ tels que $a<b$. On rappelle de la définition **??** que $U_{[\![0,n]\!]}(M,[a,b])$ désigne le nombre de montées de $M$ de l'intervalle $[a,b]$ entre les instants 0 et $n$. L'inégalité de Doob (théorème B.6.4) combinée à une inégalité élémentaire et l'hypothèse (B.28), implique implique que

$$\forall n\in\mathbb{N},\quad (b-a)\mathbf{E}\big[U_{[\![0,n]\!]}(M,[a,b])\big]\leq \mathbf{E}\big[(M_n-a)_-\big]\leq |a|+\mathbf{E}\big[|M_n|\big]\leq |a|+c\ .$$

On vérifie que $n\mapsto U_{[\![0,n]\!]}(M,[a,b])$ est croissante et que $\lim_{n\to\infty}U_{[\![0,n]\!]}(M,[a,b])=U_{\mathbb{N}}(M,[a,b])$ qui est le nombre total de montées de $M$ de l'intervalle $[a,b]$. On a donc

$$\forall a,b\in\mathbb{R}\ :\ a<b,\quad \mathbf{E}\big[U_{\mathbb{N}}(M,[a,b])\big]\leq\frac{|a|+c}{b-a}<\infty\ .$$

On pose alors

$$\Omega_0=\bigcap_{\substack{a,b\in\mathbb{Q}\\ a<b}}\big\{U_{\mathbb{N}}(M,[a,b])<\infty\big\}\ . \tag{B.29}$$

Il est clair que $\Omega_0\in\mathscr{F}$ et (B.29) implique que $\mathbf{P}(\Omega_0)=1$. En raisonnant comme dans le théorème **??**, on voit que pour tout $\omega\in\Omega_0$, $\limsup_n M_n(\omega)=\liminf_n M_n(\omega)$. On pose donc $M_\infty^*=\limsup_n M_n$. Alors $M_\infty$ est $\mathscr{F}_\infty$-mesurable et $\mathbf{P}$-p.s. $\lim_n M_n=M_\infty$. A priori $M_\infty$ est à valeurs dans $[-\infty,\infty]$ mais par Fatou on a $\mathbf{E}\big[|M_\infty|\big]=\mathbf{E}\big[\liminf_n|M_n|\big]\leq\liminf_n\mathbf{E}\big[|M_n|\big]\leq c$. Quitte à modifier $M_\infty$ sur un ensemble $\mathbf{P}$-négligeable, on peut supposer cette v.a. à valeurs réelles. ■

On examine ensuite la convergence en norme $\mathrm{L}^1$ des martingales. Pour cela on rappelle la définition B.4.11, page 315, d'une famille de v.a. uniformément intégrables : soit $(X_i)_{i\in I}$, une famille de v.a. réelles intégrables. La famille de variables $(X_i)_{i\in I}$ est *uniformément intégrable* si

$$\lim_{a\to\infty}\sup_{i\in I}\mathbf{E}\big[|X_i|\mathbf{1}_{\{|X_i|\geq a\}}\big]=0\ . \tag{B.30}$$

On rappelle le point $(iv)$ su lemme B.4.13, page 315 : soit $(X_i)_{i\in I}$ une famille de v.a. réelles. Soit $g:\mathbb{R}_+\mathbb{R}_+\to\mathbb{R}_+$ telle que $\lim_{x\to\infty}g(x)/x=\infty$. Alors

$$\sup_{i\in I}\mathbf{E}\big[g(X_i)\big]<\infty\Longrightarrow (X_i)_{i\in I}\ \text{est uniformément intégrable.} \tag{B.31}$$

On rappelle également le théorème B.4.15, page 316, qui montre que l'uniforme intégrabilité combinée à une convergence en probabilité implique est équivalent à une convergence $\mathrm{L}^1$. Plus précisément, soit $(X_n)_{n\in\mathbb{N}}$ une suite de v.a. réelles. Alors,

$$(X_n)_{n\in\mathbb{N}}\ \text{uniformément intégrable et}\ \lim_{n\to\infty}X_n=X_\infty\ \text{en probabilité}\iff\lim_{n\to\infty}\mathbf{E}\big[|X_n-X_\infty|\big]=0 \tag{B.32}$$

**Lemme B.6.9** *Soit $X$, une v.a. réelle intégrable. Soit $\mathscr{G}_i$, $i\in I$, une famille de sous-tribus de $\mathscr{F}$. Pour tout $i\in I$, on pose $X_i=\mathbf{E}[X|\mathscr{G}_i]$. Alors, la famille de variables $(X_i)_{i\in I}$ est uniformément intégrable.*

**Preuve :** on fixe $a > 0$. Par Jensen, $|X_i| = |\mathbf{E}[X|\mathscr{G}_i]| \leq \mathbf{E}[\,|X|\,|\mathscr{G}_i]$ et donc

$$\mathbf{E}\big[\,|X_i|\mathbf{1}_{\{|X_i|\geq a\}}\big] \leq \mathbf{E}\big[\mathbf{E}[\,|X|\,|\mathscr{G}_i].\mathbf{1}_{\{\mathbf{E}[\,|X|\,|\mathscr{G}_i]\geq a\}}\big] = \mathbf{E}\big[\,|X|\mathbf{1}_{\{\mathbf{E}[\,|X|\,|\mathscr{G}_i]\geq a\}}\big], \tag{B.33}$$

car $\{\mathbf{E}[\,|X|\,|\mathscr{G}_i] \geq a\} \in \mathscr{G}_i$. On fixe $b > 0$. On remarque que

$$\begin{aligned} \mathbf{E}\big[\,|X|\mathbf{1}_{\{|X|\leq b\,;\,\mathbf{E}[\,|X|\,|\mathscr{G}_i]\geq a\}}\,\big] &\leq b\mathbf{E}\big[\mathbf{1}_{\{|X|\leq b\,;\,\mathbf{E}[\,|X|\,|\mathscr{G}_i]\geq a\}}\,\big] \leq b\mathbf{E}\big[\mathbf{1}_{\{\mathbf{E}[\,|X|\,|\mathscr{G}_i]\geq a\}}\,\big] \\ &\leq b\mathbf{P}\big(\mathbf{E}[\,|X|\,|\mathscr{G}_i] \geq a\big) \leq \tfrac{b}{a}\mathbf{E}\big[\mathbf{E}[\,|X|\,|\mathscr{G}_i]\big] = \tfrac{b}{a}\mathbf{E}[\,|X|\,]. \end{aligned}$$

On remarque ensuite que $\mathbf{E}\big[\,|X|\mathbf{1}_{\{|X|>b\,;\,\mathbf{E}[\,|X|\,|\mathscr{G}_i]\geq a\}}\,\big] \leq \mathbf{E}\big[\,|X|\mathbf{1}_{\{|X|>b\}}\,\big]$, et on obtient finalement pour tous $a, b > 0$,

$$\begin{aligned} \mathbf{E}\big[\,|X|\mathbf{1}_{\{\mathbf{E}[\,|X|\,|\mathscr{G}_i]\geq a\}}\big] &= \mathbf{E}\big[\,|X|\mathbf{1}_{\{|X|\leq b\,;\,\mathbf{E}[\,|X|\,|\mathscr{G}_i]\geq a\}}\,\big] + \mathbf{E}\big[\,|X|\mathbf{1}_{\{|X|>b\,;\,\mathbf{E}[\,|X|\,|\mathscr{G}_i]\geq a\}}\,\big] \\ &\leq \tfrac{b}{a}\mathbf{E}[\,|X|\,] + \mathbf{E}\big[\,|X|\mathbf{1}_{\{|X|>b\}}\,\big]. \end{aligned}$$

On choisit $b = \sqrt{a}$ dans l'inégalité précédente qui, combinée avec (B.33), donne pour tout $a > 0$,

$$\sup_{i\in I}\mathbf{E}\big[\,|X_i|\mathbf{1}_{\{|X_i|\geq a\}}\big] \leq \tfrac{1}{\sqrt{a}}\mathbf{E}[\,|X|\,] + \mathbf{E}\big[\,|X|\mathbf{1}_{\{|X|>\sqrt{a}\}}\,\big] \xrightarrow[a\to\infty]{} 0,$$

car $\lim_{a\to\infty}\mathbf{E}[\,|X|\mathbf{1}_{\{|X|>\sqrt{a}\}}] = 0$, par convergence dominée. ■

Le théorème de convergence $\mathrm{L}^1$ des martingales s'énonce alors comme suit.

**Théorème B.6.10** ***(Convergence*** $\mathbf{L}^1$***)*** *Soit $(M_n)_{n\in\mathbb{N}}$, une $(\mathscr{F}_n)_{n\in\mathbb{N}}$-martingale. Alors, les conditions suivantes sont équivalentes.*

*(i) La martingale $(M_n)_{n\in\mathbb{N}}$ converge dans $\mathrm{L}^1$.*

*(ii) On a $\sup_{n\in\mathbb{N}}\mathbf{E}[\,|M_n|\,] < \infty$ et*

$$\mathbf{P}\text{-}p.s.\quad \forall n\in\mathbb{N}\,,\quad \mathbf{E}[M_\infty|\mathscr{F}_n]=M_n\ ,$$

*où $M_\infty$ désigne la limite presque sûre de $(M_n)_{n\in\mathbb{N}}$ (dont l'existence est garantie par le théorème B.6.8).*

*(iii) Il existe une v.a. réelle intégrable $X$ telle que*

$$\mathbf{P}\text{-}p.s.\quad \forall n \in \mathbb{N}\,,\quad \mathbf{E}[X|\mathscr{F}_n] = M_n\ .$$

*(iv) La famille de variables $(M_n)_{\in\mathbb{N}}$ est uniformément intégrable.*

**Preuve :** montrons $(i) \Rightarrow (ii)$. On note $M_\infty$ la limite dans $\mathrm{L}^1$ de $(M_n)_{n\in\mathbb{N}}$. Puisqu'il y a convergence dans $\mathrm{L}^1$, il y a convergence des normes $\mathrm{L}^1$ : $\lim_n\mathbf{E}[\,|M_n|\,]=\mathbf{E}[\,|M_\infty|\,]$, ce qui implique que $\sup_{n\in\mathbb{N}}\mathbf{E}[\,|M_n|\,]<\infty$. Par le théorème B.6.8, $(M_n)_{n\in\mathbb{N}}$ converge p.s. et comme de toute suite convergente pour la norme $\mathrm{L}^1$, on extrait une suite convergeant p.s., il est clair que la limite dans $\mathrm{L}^1$ de $(M_n)_{n\in\mathbb{N}}$ est la même que sa limite presque sûre. On a donc $\lim_n M_n = M_\infty$ p.s. et $\lim_n\mathbf{E}[\,|M_n - M_\infty|\,]=0$. Le lemme B.6.2 $(iii)$ implique ensuite que pour tous $m, n\in\mathbb{N}$, on a $\mathbf{E}[M_{n+m}|\mathscr{F}_n] = M_n$. Cela entraîne que

$$\begin{aligned} \mathbf{E}\big[\,|\mathbf{E}[M_\infty|\mathscr{F}_n]-M_n|\,\big] &= \mathbf{E}\big[\,|\mathbf{E}[M_\infty - M_{n+m}|\mathscr{F}_n]\,|\,\big] \\ &\leq \mathbf{E}\big[\,\mathbf{E}[\,|M_\infty-M_{n+m}|\,|\mathscr{F}_n\,]\,\big]=\mathbf{E}[\,|M_\infty-M_{n+m}|\,]. \end{aligned}$$

Lorsque $m$ tend vers l'infini, on a $\mathbf{E}\big[\,|\mathbf{E}[M_\infty|\mathscr{F}_n]-M_n|\,\big] = 0$, ce qui montre $(ii)$.

L'implication $(ii) \Rightarrow (iii)$ est immédiate. L'implication $(iii) \Rightarrow (iv)$ est une conséquence directe du lemme B.6.9. Montrons que $(iv) \Rightarrow (i)$ : la propriété d'uniforme intégrabilité implique que pour un certain $a > 0$, il existe une constante $c$, positive et finie telle que pour tout $n\in\mathbb{N}$, on ait $\mathbf{E}[\,|M_n|\mathbf{1}_{\{|M_n|\geq a\}}\,]\leq c$. Donc pour tout $n\in\mathbb{N}$,

$$\mathbf{E}[\,|M_n|\,] = \mathbf{E}[\,|M_n|\mathbf{1}_{\{|M_n|<a\}}\,] + \mathbf{E}[\,|M_n|\mathbf{1}_{\{|M_n|\geq a\}}\,] \leq a + c.$$

On en déduit donc que $\sup_{n\in\mathbb{N}}\mathbf{E}\big[\,|M_n|\,\big] < \infty$. Par le théorème (B.6.8), la suite $(M_n)_{n\in\mathbb{N}}$ converge presque sûrement, donc en probabilité, mais comme elle est uniformément intégrable, le lemme B.6.9 implique qu'elle converge dans $\mathrm{L}^1$ également, ce qui montre $(i)$. ■

Nous prouvons ensuite une conséquence utile du théorème de convergence $\mathrm{L}^1$ des martingales.

**Théorème B.6.11** *Soit $X$, une v.a. intégrable. Soit $(\mathscr{F}_n)_{n\in\mathbb{N}}$, une filtration sur $(\Omega, \mathscr{F})$. On rappelle que $\mathscr{F}_\infty = \sigma(\mathscr{F}_n; n\in\mathbb{N})$. Alors*

$$\mathbf{E}[X\,|\mathscr{F}_n] \xrightarrow[n\to\infty]{} \mathbf{E}[X\,|\mathscr{F}_\infty] \quad \textit{p.s. et dans } \mathrm{L}^1 .$$

**Preuve :** on pose $M_n = \mathbf{E}[X|\mathscr{F}_n]$, $n\in\mathbb{N}$. C'est une martingale uniformément intégrable : par le théorème B.6.10, elle converge p.s. et dans $\mathrm{L}^1$ vers une variable notée $M_\infty$. Il faut donc montrer que $\mathbf{P}$-p.s. $M_\infty = \mathbf{E}[X|\mathscr{F}_\infty]$. Pour cela, on fixe $p\in\mathbb{N}$ et $A\in\mathscr{F}_p$. Pour tout $n\in\mathbb{N}$, on a d'abord

$$\big|\mathbf{E}[M_n\mathbf{1}_A]-\mathbf{E}[M_\infty\mathbf{1}_A]\big| = \big|\mathbf{E}[(M_n-M_\infty)\mathbf{1}_A]\big| \leq \mathbf{E}\big[|M_n-M_\infty|\mathbf{1}_A\big] \leq \mathbf{E}\big[|M_n-M_\infty|\big].$$

La convergence $\mathrm{L}^1$ des $M_n$ vers $M_\infty$ implique que $\mathbf{E}[M_n\mathbf{1}_A] \to \mathbf{E}[M_\infty\mathbf{1}_A]$. Or pour tout $n\geq p$, comme $\mathscr{F}_p\subset\mathscr{F}_n$, on a $\mathbf{E}[M_n\mathbf{1}_A] = \mathbf{E}[\mathbf{E}[X|\mathscr{F}_n]\mathbf{1}_A] = \mathbf{E}[X\mathbf{1}_A]$. Cela prouve donc que

$$\forall p \in \mathbb{N}\,,\ \forall A \in \mathscr{F}_p\,, \quad \mathbf{E}[X\mathbf{1}_A] = \mathbf{E}[M_\infty\mathbf{1}_A]\ . \tag{B.34}$$

On pose alors

$$\mathcal{L} := \big\{A \in \mathscr{F}_\infty : \mathbf{E}[X\mathbf{1}_A] = \mathbf{E}[M_\infty\mathbf{1}_A]\big\} \quad \text{et} \quad \mathscr{P} := \bigcup_{p\in\mathbb{N}} \mathscr{F}_p\ .$$

Il est facile de voir que $\mathscr{P}$ est un pi-système. On a clairement $\sigma(\mathscr{P}) = \mathscr{F}_\infty$ et par le théorème de convergence dominée on vérifie aisément que $\mathcal{L}$ est une classe monotone. Or (B.34) implique que $cP \subset \mathcal{L}$. Le théorème A.1.10 de la classe monotone, page 270, montre que $\sigma(\mathscr{P}) \subset \mathcal{L}$, ce qui implique que $\mathcal{L} = \mathscr{F}_\infty$. Autrement dit, pour tout $A\in\mathscr{F}_\infty$, $\mathbf{E}[M_\infty\mathbf{1}_A] = \mathbf{E}[X\mathbf{1}_A]$. Comme $M_\infty$ est $\mathscr{F}_\infty$-mesurable, on en déduit que p.s. $X_\infty = \mathbf{E}[X|\mathscr{F}_\infty]$. ■

Supposons que $(M_n)_{n\in\mathbb{N}}$ soit une martingale bornée en norme $\mathrm{L}^p$ : $\sup_{n\in\mathbb{N}} \mathbf{E}\big[|M_n|^p\big] < \infty$. Si $p>1$, alors (B.31) implique que $(M_n)_{n\in\mathbb{N}}$ est uniformément intégrable et celle martingale converge dans $\mathrm{L}^1$ et presque sûrement par le théorème précédent. On a en fait une converence dans $\mathrm{L}^p$, comme le montre le théorème suivant.

**Théorème B.6.12 (Convergence $\mathrm{L}^p$ des martingales)** *Soit $p \in ]1,\infty[$. Soit $(M_n)_{n\geq 0}$, une martingale relativement à la filtration $(\mathscr{F}_n)_{n\geq 0}$. On suppose que*

$$\sup_{n\in\mathbb{N}} \mathbf{E}[\,|M_n|^p] < \infty\ .$$

*Alors, il existe une v.a. $M_\infty$ dans $\mathrm{L}^p$ telle que*

*(i)* $\mathbf{P}$*-p.s.* $\lim_n M_n = M_\infty$ *presque sûrement;*

*(ii)* $(M_n)_{n\geq 0}$ *converge vers $M_\infty$ dans $\mathrm{L}^p$ :* $\lim_n \mathbf{E}[\,|M_\infty - M_n|^p\,] = 0$*;*

*(iii)* $\mathbf{P}$*-p.s.* $\mathbf{E}[M_\infty\,|\,\mathscr{F}_n] = M_n$, $n\in\mathbb{N}$.

**Preuve :** par Jensen, $(\mathbf{E}[\,|M_n|\,])^p \leq \mathbf{E}[\,|M_n|^p\,]$, donc $\sup_{n\geq 0}\mathbf{E}[\,|M_n|\,] < \infty$, et le théorème B.6.8 de convergence p.s. des martingales implique que $(M_n)_{n\geq 0}$ converge p.s. On note $M_\infty$ sa limite p.s. et on pose $S = \sup_{n\geq 0}|M_n|$. On a donc $|M_\infty|\leq S$. Pour tout $n\in\mathbb{N}$, on a donc $|M_\infty - M_n|\leq 2S$, et donc $|M_\infty - M_n|^p \leq 2^pS^p$. L'inégalité $\mathrm{L}^p$ des sous-martingales positives (théorème B.6.6) montre que $\mathbf{E}[S^p] < \infty$, or $\lim_n |M_\infty - M_n|^p = 0$ p.s. Par convergence dominée on a donc $(ii)$. Enfin, par Jensen, on a $(\mathbf{E}[\,|M_\infty - M_n|\,])^p \leq \mathbf{E}[\,|M_\infty - M_n|^p\,]$, donc $(M_n)_{n\in\mathbb{N}}$ converge vers $M_\infty$ dans $\mathrm{L}^1$ et le théorème B.6.10 entraîne $(iii)$. ■

## Martingales rétrogrades.

On introduit la notion de sur-martingales rétrograde (ou inverse) comme suit.

**Définition B.6.13** (*Sur-martingales rétrogrades*) On introduit les notions suivantes.

- Soit $\mathscr{F}_n$, $n \in \mathbb{N}$, une suite de sous-tribus de l'espace mesurable $(\Omega, \mathscr{F})$. On dit que $(\mathscr{F}_n)_{n\geq 0}$ est une *filtration rétrograde* ssi $\mathscr{F}_{n+1}\subset\mathscr{F}_n$, pour tout $n\geq 0$. On définit alors
$$\mathscr{F}_\infty := \bigcap_{n\in\mathbb{N}} \mathscr{F}_n\ .$$
- Soit $(M_n)_{n\in\mathbb{N}}$, une suite de v.a. réelles. C'est une *martingale rétrograde* (resp. *sur-martingale rétrograde*) *relativement à la filtration rétrograde*$(\mathscr{F}_n)_{n\geq 0}$ si pour tout $n\in\mathbb{N}$, $M_n$ est $\mathscr{F}_n$-mesurable, intégrable et $\mathbf{E}[M_n|\mathscr{F}_{n+1}] = M_{n+1}$ (resp. $\mathbf{E}[M_n|\mathscr{F}_{n+1}] \leq M_{n+1}$). □

On notera bien la différence avec la définition de martingale : les tribus $\mathscr{F}_n$ d'une filtration rétrograde décroissent, les variables $M_n$ sont donc *"de moins en moins aléatoires"*. Il est donc normal de s'attendre à ce que les martingales rétrogrades se comportent plus régulièrement. On remarque par exemple que $(M_n)_{n\geq 0}$ est une martingale rétrograde **ssi**

$$\forall n \in \mathbb{N}\,, \quad \mathbf{E}[M_0|\mathscr{F}_n] = M_n \;. \tag{B.35}$$

Le lemme suivant fait le lien entre les notions de martingale et martingale inverse. Sa preuve est immédiate.

**Lemme B.6.14** *Soit $(\mathscr{F}_n)_{n\geq 0}$, une filtration rétrograde et soit $(M_n)_{n\in\mathbb{N}}$, une $(\mathscr{F}_n)_{n\geq 0}$-sur-martingale (resp. martingale) Pour tout $n_0 \in \mathbb{N}$, on pose $M_n^{n_0} = M_{(n_0-n)^+}$ et $\mathscr{F}_n^{n_0} = \mathscr{F}_{(n_0-n)^+}$, c'est-à-dire*

$$(M_n^{n_0})_{n\in\mathbb{N}} = \big(M_{n_0}, M_{n_0-1}, \ldots, M_2, M_1, M_0, M_0, M_0, \ldots\big) \quad \text{et}$$

$$(\mathscr{F}_n^{n_0})_{n\in\mathbb{N}} = \big(\mathscr{F}_{n_0}, \mathscr{F}_{n_0-1}, \ldots, \mathscr{F}_2, \mathscr{F}_1, \mathscr{F}_0, \mathscr{F}_0, \mathscr{F}_0, \ldots\big)$$

*Alors $(\mathscr{F}_n^{n_0})_{n\in\mathbb{N}}$ est une filtration, au sens usuel, et $(M_n^{n_0})_{n\in\mathbb{N}}$ est une sur-martingale (resp. martingale) relativement à $(\mathscr{F}_n^{n_0})_{n\in\mathbb{N}}$, au sens usuel.*

**Théorème B.6.15 (Convergence des sur-martingales rétrogrades)** *Soit $(\mathscr{F}_n)_{n\geq 0}$, une filtration rétrograde de $(\Omega, \mathscr{F})$. On pose $\mathscr{F}_\infty = \bigcap_{n\in\mathbb{N}} \mathscr{F}_n$. Soit $(M_n)_{n\in\mathbb{N}}$ une sur-martingale rétrograde relativement à $(\mathscr{F}_n)_{n\in\mathbb{N}}$. On obseve que cela implique que $n \mapsto \mathbf{E}[M_n]$ croît et on suppose que*

$$\sup_{n\in\mathbb{N}} \mathbf{E}[M_n] < \infty \;. \tag{B.36}$$

*Alors, il existe une v.a. réelle $M_\infty$ intégrable $\mathscr{F}_\infty$-mesurable, intégrable telle que $\lim_n M_n = M_\infty$ p.s. et dans $\mathrm{L}^1$ et telle que*

$$\forall n \in \mathbb{N}, \quad \mathbf{P}\text{-}p.s. \quad \mathbf{E}[M_n|\mathscr{F}_\infty] \leq M_\infty \;.$$

*Lorsque $M$ est une martingale rétrograde, la suite $n \mapsto \mathbf{E}[M_n]$ et (B.36) est toujours vérifiée de plus on a pour tout $n \in \mathbb{N}$, $\mathbf{P}$-p.s. $\mathbf{E}[M_n|\mathscr{F}_\infty] = M_\infty$.*

**Preuve :** on commence par prouver le théorème dans le cas ou $M$ est une martingale rétrograde. Soit $n_0 \in \mathbb{N}$. Soient $a, b \in \mathbb{R}$, tels que $a < b$. On reprend les notations du lemme B.6.14. Un simple dessin montre que $U_{[\![0,n_0]\!]}(-M, [-b,-a]) = U_{[\![0,n_0]\!]}(M^{n_0}, [a,b])$. L'inégalité de Doob (théorème B.6.4) appliqué à la martingale usuelle $M^{n_0}$ (lemme B.6.14) implique que

$$\begin{aligned}(b-a)\mathbf{E}\big[U_{[\![0,n_0]\!]}(-M,[-b,-a])\big] \;&=\; (b-a)\mathbf{E}\big[U_{[\![0,n_0]\!]}(M^{n_0},[a,b])\big] \\ &\leq\; \mathbf{E}\big[(M_{n_0}^{n_0}-a)_-\big] = \mathbf{E}\big[(M_0-a)_-\big] \leq |a| + \mathbf{E}\big[|M_0|\big] \;.\end{aligned}$$

On raisonne ensuite sur $-M$ comme dans la preuve du théorème B.6.8 de convergence presque sûre des martingales pour montrer qu'il existe une v.a. $M_\infty$ telle que $\mathbf{P}$-p.s. $\lim_n M_n = M_\infty$. La remarque (B.35) combinée au lemme B.6.9 implique que $(M_n)_{n\in\mathbb{N}}$ est uniformément intégrable et par (B.32) on a $\lim_n \mathbf{E}\big[|M_\infty - M_n|\big] = 0$. Pour tout $n \in \mathbb{N}$, et tout $p \geq n$, $M_p$ est $\mathscr{F}_n$-mesurable et donc $M_\infty$ est $\mathscr{F}_n$-mesurable. Donc pour tout $a \in \mathbb{R}$, $\{M_\infty \leq a\} \in \mathscr{F}_n$, pour tout $n \in \mathbb{N}$, et donc que $\{M_\infty \leq a\} \in \bigcap_{n\in\mathbb{N}} \mathscr{F}_n = \mathscr{F}_\infty$. Donc $M_\infty$ est $\mathscr{F}_\infty$-mesurable. Soit $A \in \mathscr{F}_\infty$. Pour tous $m, n \in \mathbb{N}$, comme $A \in \mathscr{F}_{n+m}$ comme $\mathbf{E}[M_n|\mathscr{F}_{n+m}] = M_{m+n}$ (conséquence immédiate de (B.35)), on a $\mathbf{E}[M_n \mathbf{1}_A] = \mathbf{E}[M_{m+n}\mathbf{1}_A]$. Comme les $M_n$ convergent dans $\mathrm{L}^1$ vers $M_\infty$, $\lim_{m\to\infty} \mathbf{E}[M_{m+n}\mathbf{1}_A] = \mathbf{E}[M_\infty \mathbf{1}_A]$ et donc $\mathbf{E}[M_n\mathbf{1}_A] = \mathbf{E}[M_\infty\mathbf{1}_A]$. On en déduit donc que $\mathbf{P}$-p.s. $\mathbf{E}[M_n|\mathscr{F}_\infty] = M_\infty$, ce qui termine la preuve du théorème dans le cas des martingales rétrogrades.

Supposons que $M$ soit une sur-martingale rétrograde. Pour tout $n \in \mathbb{N}$, on $\Delta_n := M_{n+1} - \mathbf{E}[M_n|\mathscr{F}_{n+1}]$. On constate que $\Delta_n$ est $\mathscr{F}_{n+1}$ mesurable et positive. Pour tout $n \in \mathbb{N}^*$, on pose $V_n := \sum_{p\geq n} \Delta_p$, qui est a priori à valeurs dans $[0, \infty]$ et qui est $\mathscr{F}_{n+1}$-mesurable et décroissante en $n$. On observe que

$$\mathbf{E}[|V_n|] = \mathbf{E}[V_n] = \sum_{p\geq n} \mathbf{E}[\Delta_p] = \sum_{p\geq n} \mathbf{E}[M_p] - \mathbf{E}[M_{p-1}] = \sup_{p\geq 0} \mathbf{E}[M_p] - \mathbf{E}[M_n] < \infty \;.$$

Cela implique que chaque $V_n$ est une v.a. intégrable et comme la suite $n \mapsto \mathbf{E}[M_n]$ est croissante, on en déduit que $\lim_n \mathbf{E}[V_n] = 0$. Comme les $V_n$ décroissent cela montre que $\lim_n V_n = 0$ p.s. et dans $\mathrm{L}^1$. On vérifie ensuite immediatement que $M_n^* = M_n + V_n$ est une $(\mathscr{F}_n)_{n\in\mathbb{N}}$-martingale rétrograde. En lui appliquant le théorème prouvé dans le cas des martingales, on voit que $\lim_n M_n = \lim_n M_n^* =: M_\infty$ p.s. et dans $\mathrm{L}^1$, qui est une v.a. $\mathscr{F}_\infty$-mesurable intégrable et pour tout $n \in \mathbb{N}$, on a $\mathbf{P}$-p.s.

$$M_\infty = \mathbf{E}[M_n^*|\mathscr{F}_\infty] = \mathbf{E}[M_n|\mathscr{F}_\infty] + \mathbf{E}[V_n|\mathscr{F}_\infty] \geq \mathbf{E}[M_n|\mathscr{F}_\infty] \;,$$

ce qui termine la preuve. ∎