Chapitre 13

Chaînes de Markov

13.1 Définition et premières propriétés

Dans tout ce chapitre, E est un espace fini ou dénombrable, qui est muni comme d'habitude de la tribu $\mathcal{P}(E)$. Une matrice stochastique sur E est une famille $(Q(x,y), x, y \in E)$ de nombres réels satisfaisant les deux conditions :

(i) $0 \leq Q(x,y) \leq 1$ pour tous $x, y \in E$;

(ii) pour tout $x \in E$, $\displaystyle\sum_{y \in E} Q(x,y) = 1$.

Cette notion est équivalente à celle de probabilité de transition de E dans E : si on pose

$$\nu(x, A) = \sum_{y \in A} Q(x,y)\ , \qquad x \in E,\ A \subset E,$$

on voit que ν est une probabilité de transition de E dans E (voir le Chapitre 11), et inversement si on part d'une telle probabilité de transition ν, la formule $Q(x,y) = \nu(x, \{y\})$ définit une matrice stochastique sur E.

Pour tout entier $n \geq 1$, on peut définir $Q_n = (Q)^n$: $Q_1 = Q$, et ensuite par récurrence,

$$Q_{n+1}(x,y) = \sum_{z \in E} Q_n(x,z) Q(z,y).$$

On vérifie que Q_n est encore une matrice stochastique sur E. On pose aussi $Q_0(x,y) = \mathbf{1}_{\{x=y\}}$.

Pour toute fonction $f : E \longrightarrow \mathbb{R}_+$, on notera Qf la fonction définie par

$$Qf(x) = \sum_{y \in E} Q(x,y) f(y).$$

Définition 13.1.1 *Soit Q une matrice stochastique sur E, et soit $(X_n)_{n \in \mathbb{N}}$ un processus aléatoire à valeurs dans E. On dit que $(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov de matrice de transition Q si pour tout entier $n \geq 0$, la loi conditionnelle de X_{n+1} connaissant $(X_0, X_1, \ldots, X_n)$ est $Q(X_n, y)$. De manière équivalente, cela signifie que*

$$P(X_{n+1} = y \mid X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n) = Q(x_n, y),$$

pour tous $x_0, x_1, \ldots, x_n, y \in E$ tels que $P(X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n) > 0$.

Remarques. (i) En général, la loi conditionnelle de X_{n+1} connaissant $X_0, X_1, \ldots, X_n$ dépend de toutes les variables $X_0, X_1, \ldots, X_n$ et pas seulement de la dernière X_n. Le fait qu'ici cette loi conditionnelle ne dépende que de X_n est ce qu'on appelle la **propriété de Markov** : pour prédire le futur (X_{n+1}) la connaissance du passé ($X_0, X_1, \ldots, X_n$) ne donne pas plus d'information que celle du présent (X_n). Nous verrons plus tard d'autres formes plus précises de la propriété de Markov, qui correspondent à la même idée.

(ii) La fonction $Q(x, \cdot)$ donnant la loi conditionnelle de X_{n+1} sachant que $X_n = x$ ne dépend pas de l'entier n : c'est le caractère homogène de la chaîne de Markov. On pourrait aussi considérer des chaînes de Markov inhomogènes, pour lesquelles le mécanisme de transition entre les instants n et $n+1$ dépend de n.

Proposition 13.1.1 *Un processus $(X_n)_{n\in\mathbb{N}}$ à valeurs dans E est une chaîne de Markov de matrice de transition Q ssi, pour tout $n \geq 0$ et pour tous $x_0, x_1, \ldots, x_n \in E$,*

$$P(X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n) = P(X_0 = x_0)Q(x_0, x_1)Q(x_1, x_2)\cdots Q(x_{n-1}, x_n). \quad (13.1)$$

En particulier, on a si $P(X_0 = x_0) > 0$,

$$P(X_n = x_n \mid X_0 = x_0) = Q_n(x_0, x_n).$$

Preuve. Si $(X_n)_{n\in\mathbb{N}}$ est une chaîne de Markov de matrice de transition Q la formule donnée est immédiate par récurrence sur n en écrivant

$$\begin{aligned} &P(X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n, X_{n+1} = x_{n+1}) = \\ &= P(X_0 = x_0, \ldots, X_n = x_n) \times P(X_{n+1} = x_{n+1} \mid X_0 = x_0, \ldots, X_n = x_n). \end{aligned}$$

Inversement, si la formule donnée est vraie, on vérifie immédiatement que

$$\begin{aligned} P(X_{n+1} = y \mid X_0 = x_0, \ldots, X_n = x_n) &= \frac{P(X_0 = x_0)Q(x_0, x_1)\cdots Q(x_{n-1}, x_n)Q(x_n, y)}{P(X_0 = x_0)Q(x_0, x_1)\cdots Q(x_{n-1}, x_n)} \\ &= Q(x_n, y). \end{aligned}$$

La dernière assertion s'obtient en remarquant que

$$Q_n(x_0, x_n) = \sum_{x_1, x_2, \ldots, x_{n-1} \in E} Q(x_0, x_1)Q(x_1, x_2)\cdots Q(x_{n-1}, x_n).$$

Remarque. La formule (13.1) montre que pour une chaîne de Markov $(X_n)_{n\in\mathbb{N}}$, la loi de $(X_0, X_1, \ldots, X_n)$ est complètement déterminée par la connaissance de la loi initiale (la loi de X_0) et de la matrice de transition Q.

La proposition suivante rassemble d'autres propriétés simples des chaînes de Markov. Dans (ii) ci-dessous, on utilise la notation $P(A \mid Z)$ pour désigner l'espérance conditionnelle $E[\mathbf{1}_A \mid Z]$.

Proposition 13.1.2 *Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov de matrice de transition Q.*

(i) *Pour tout entier $n \geq 0$ et toute fonction mesurable $f : E \longrightarrow \mathbb{R}_+$,*

$$E[f(X_{n+1}) \mid X_0, X_1, \ldots, X_n] = E[f(X_{n+1}) \mid X_n] = Qf(X_n).$$

Plus généralement, pour tout sous-ensemble fini $\{i_1, \ldots, i_k\}$ de $\{0, 1, \ldots, n-1\}$, on a

$$E[f(X_{n+1}) \mid X_{i_1}, \ldots, X_{i_k}, X_n] = Qf(X_n).$$

(ii) *Pour tous les entiers $n \geq 0, p \geq 1$ et pour tous $y_1, \ldots, y_p \in E$,*

$$P(X_{n+1} = y_1, \ldots, X_{n+p} = y_p \mid X_0, \ldots, X_n) = Q(X_n, y_1)Q(y_1, y_2) \ldots Q(y_{p-1}, y_p),$$

et donc

$$P(X_{n+p} = y_p \mid X_n) = Q_p(X_n, y_p).$$

Si on pose $Y_p = X_{n+p}$ pour tout $p \in \mathbb{N}$, le processus $(Y_p)_{p \in \mathbb{N}}$ est encore une chaîne de Markov de matrice de transition Q.

Preuve. (i) D'après la définition,

$$E[f(X_{n+1}) \mid X_0, X_1, \ldots, X_n] = \sum_{y \in E} Q(X_n, y) f(y) = Qf(X_n).$$

Ensuite, si $\{i_1, \ldots, i_k\}$ est un sous-ensemble fini de $\{0, 1, \ldots, n-1\}$, on a

$$\begin{aligned} E[f(X_{n+1}) \mid X_{i_1}, \ldots, X_{i_k}, X_n] &= E[E[f(X_{n+1}) \mid X_0, X_1, \ldots, X_n] \mid X_{i_1}, \ldots, X_{i_k}, X_n] \\ &= E[Qf(X_n) \mid X_{i_1}, \ldots, X_{i_k}, X_n] \\ &= Qf(X_n). \end{aligned}$$

(ii) Il découle immédiatement de (13.1) que

$$P(X_{n+1} = y_1, \ldots, X_{n+p} = y_p \mid X_0 = x_0, \ldots, X_n = x_n) = Q(x_n, y_1)Q(y_1, y_2) \cdots Q(y_{p-1}, y_p).$$

La formule pour $P(X_{n+p} = y_p \mid X_n)$ en découle en sommant sur les choix possibles de $y_1, \ldots, y_{p-1}$. Enfin, pour la dernière assertion, on déduit de ce qui précède que

$$P(Y_0 = y_0, Y_1 = y_1, \ldots, Y_p = y_p) = P(X_n = y_0)Q(y_0, y_1)Q(y_1, y_2) \ldots Q(y_{p-1}, y_p),$$

et on utilise la caractérisation donnée dans la proposition 13.1.1. □

13.2 Quelques exemples

13.2.1 Variables aléatoires indépendantes

Si $(X_n)_{n \in \mathbb{N}}$ est une suite de v.a. indépendantes à valeurs dans E, de même loi μ, alors $(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov de matrice de transition

$$Q(x, y) = \mu(y), \quad \forall x, y \in E.$$

La vérification est immédiate. Ce n'est pas l'exemple le plus intéressant de chaîne de Markov !

13.2.2 Marches aléatoires sur $\mathbb{Z}^d$

Soient $\eta, \xi_1, \xi_2, \ldots, \xi_n, \ldots$ des v.a. indépendantes à valeurs dans $\mathbb{Z}^d$. On suppose que $\xi_1, \xi_2, \ldots$ ont même loi μ et on pose pour tout $n \geq 0$,

$$X_n = \eta + \xi_1 + \xi_2 + \cdots + \xi_n.$$

Alors $(X_n)_{n\in\mathbb{N}}$ est une chaîne de Markov de matrice de transition

$$Q(x,y) = \mu(y-x), \quad \forall x, y \in E.$$

En effet, en remarquant que ξ_{n+1} est indépendante de $(X_0, X_1, \ldots, X_n)$, on a

$$\begin{aligned}
&P(X_{n+1} = y \mid X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n) \\
&\quad = P(\xi_{n+1} = y - x_n \mid X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n) \\
&\quad = P(\xi_{n+1} = y - x_n) \\
&\quad = \mu(y - x_n).
\end{aligned}$$

Soit $(e_1, \ldots, e_d)$ la base canonique de $\mathbb{R}^d$. Dans le cas où $\mu(e_i) = \mu(-e_i) = \frac{1}{2d}$ pour tout $i \in \{1, \ldots, d\}$, la chaîne de Markov obtenue est appelée la marche aléatoire simple sur $\mathbb{Z}^d$.

13.2.3 Marche aléatoire simple sur un graphe

Soit $\mathcal{P}_2(E)$ l'ensemble des parties de E à deux éléments, et soit A un sous-ensemble de $\mathcal{P}_2(E)$. Pour tout $x \in E$, on note

$$A_x = \{y \in E : \{x,y\} \in A\}.$$

On suppose que A_x est fini et non vide pour tout $x \in E$. On définit alors une matrice de transition Q sur E en posant pour tous $x, y \in E$,

$$Q(x,y) = \begin{cases} \dfrac{1}{\operatorname{Card} A_x} & \text{si } \{x,y\} \in A \\ 0 & \text{sinon.} \end{cases}$$

Une chaîne de Markov de matrice de transition Q est appelée marche aléatoire simple sur le graphe (E, A).

13.2.4 Processus de branchement

Rappelons la définition de ces processus déjà étudiés dans le chapitre précédent. Si μ est une mesure de probabilité sur $\mathbb{N}$, et $\ell \in \mathbb{N}$, on définit par récurrence une suite (X_n) de v.a. à valeurs dans $\mathbb{N}$ en posant

$$\begin{aligned}
X_0 &= \ell \\
X_{n+1} &= \sum_{j=1}^{X_n} \xi_{n,j}\,, \qquad \forall n \in \mathbb{N}\,,
\end{aligned}$$

où les v.a. $\xi_{n,j},\ n, j \in \mathbb{N}$ sont indépendantes et de loi μ. Alors, $(X_n)_{n\in\mathbb{N}}$ est une chaîne de Markov sur $E = \mathbb{N}$ de matrice de transition

$$Q(x, y) = \mu^{*x}(y), \quad \forall x, y \in \mathbb{N},$$

où μ^{*x} est la convolution de μ x fois avec elle-même, ou de manière équivalente la loi de la somme de x v.a. indépendantes de loi μ (en particulier μ^{*0} est la mesure de Dirac en 0). En effet, en observant que les v.a. $\xi_{n,j},\ j \in \mathbb{N}$ sont indépendantes de $X_0, \ldots, X_n$, on a

$$\begin{aligned}
&P(X_{n+1} = y \mid X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n) \\
&\quad = P(\sum_{j=1}^{x_n} \xi_{n,j} = y \mid X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n) \\
&\quad = P(\sum_{j=1}^{x_n} \xi_{n,j} = y) \\
&\quad = \mu^{*x_n}(y).
\end{aligned}$$

13.3 La chaîne de Markov canonique

Nous commençons par un résultat d'existence de chaîne de Markov associée à une matrice de transition donnée.

Proposition 13.3.1 *Soit Q une matrice stochastique sur E. On peut trouver un espace de probabilité $(\Omega', \mathcal{F}', P')$ sur lequel il existe, pour tout $x \in E$, un processus $(X^x_n)_{n\in\mathbb{N}}$ qui est une chaîne de Markov de matrice de transition Q, issue de $X^x_0 = x$.*

Preuve. On peut prendre $\Omega' = [0, 1[$, muni de la tribu borélienne et de la mesure de Lebesgue. A partir du développement dyadique (propre) d'un réel $\omega \in [0, 1[$,

$$\omega = \sum_{n=0}^{\infty} \varepsilon_n(\omega)\, 2^{-n-1}, \quad \varepsilon_n(\omega) \in \{0, 1\}$$

on construit une suite $(\varepsilon_n)_{n\in\mathbb{N}}$ de v.a. indépendantes de même loi $P(\varepsilon_n = 1) = P(\varepsilon_n = 0) = 1/2$. Si φ est une injection de $\mathbb{N}\times\mathbb{N}$ dans $\mathbb{N}$, les v.a. $\eta_{i,j} = \varepsilon_{\varphi(i,j)}$, $i, j \in \mathbb{N}$ sont (évidemment) encore indépendantes et de même loi. En posant

$$U_i = \sum_{j=0}^{\infty} \eta_{i,j}\, 2^{-j-1}$$

on obtient une suite $U_0, U_1, U_2, \ldots$ de v.a. indépendantes de loi uniforme sur $[0, 1]$ (pour voir que U_i suit la loi uniforme, noter que $\sum_{j=0}^{p} \eta_{i,j}\, 2^{-j-1}$ a même loi que $\sum_{n=0}^{p} \varepsilon_n\, 2^{-n-1}$, pour tout entier p, et faire tendre p vers ∞).

Soit $y_1, y_2, \ldots, y_k, \ldots$ une énumération des éléments de E. Fixons aussi $x \in E$. On pose $X^x_0 = x$ puis

$$X^x_1 = y_k \quad \text{si} \quad \sum_{1\le j<k} Q(x, y_j) < U_1 \le \sum_{1\le j\le k} Q(x, y_j)$$

de sorte qu'il est clair que $P(X_1^x = y) = Q(x,y)$ pour tout $y \in E$. On continue par récurrence en posant

$$X_{n+1}^x = y_k \quad \text{si} \quad \sum_{1 \leq j < k} Q(X_n^x, y_j) < U_{n+1} \leq \sum_{1 \leq j \leq k} Q(X_n^x, y_j).$$

En utilisant l'indépendance des v.a. U_i, on vérifie très facilement que pour tout $k \geq 1$,

$$\begin{aligned}
&P(X_{n+1}^x = y_k \mid X_0^x = x_0, X_1^x = x_1, \dots X_n^x = x_n) \\
&\quad = P(\sum_{1 \leq j < k} Q(x_n, y_j) < U_{n+1} \leq \sum_{1 \leq j \leq k} Q(x_n, y_j) \mid X_0^x = x_0, X_1^x = x_1, \dots X_n^x = x_n) \\
&\quad = P(\sum_{1 \leq j < k} Q(x_n, y_j) < U_{n+1} \leq \sum_{1 \leq j \leq k} Q(x_n, y_j)) \\
&\quad = Q(x_n, y_k),
\end{aligned}$$

de sorte que $(X_n^x)_{n \in \mathbb{N}}$ est une chaîne de Markov de transition Q. □

Dans la suite, il sera utile de faire un choix canonique de l'espace de probabilité sur lequel sera définie la chaîne de Markov étudiée. On prendra

$$\Omega = E^{\mathbb{N}}.$$

Un élément ω de Ω est donc une suite $\omega = (\omega_0, \omega_1, \omega_2, \dots)$ d'élements de E. Les applications coordonnées X_n, $n \in \mathbb{N}$ sont alors définies par

$$X_n(\omega) = \omega_n.$$

On munit Ω de la plus petite tribu, notée $\mathcal{F}$, qui rende mesurables les applications coordonnées. C'est aussi la tribu engendrée par les "cylindres", c'est-à-dire les ensembles C de la forme

$$C = \{\omega \in \Omega : \omega_0 = x_0, \omega_1 = x_1, \dots, \omega_n = x_n\}$$

où $n \in \mathbb{N}$ et $x_0, x_1, \dots x_n \in E$.

Lemme 13.3.2 *Soit $(G, \mathcal{G})$ un espace mesurable, et soit ψ une application de G dans Ω. Alors ψ est mesurable ssi $X_n \circ \psi$ l'est pour tout $n \in \mathbb{N}$.*

Preuve. Il suffit bien sûr de montrer que si $X_n \circ \psi$ est mesurable pour tout n, alors ψ l'est aussi. Or,

$$\{A \in \mathcal{F} : \psi^{-1}(A) \in \mathcal{G}\}$$

est une tribu sur Ω qui par hypothèse contient tous les ensembles de la forme $X_n^{-1}(y)$, $y \in E$, donc rend mesurables toutes les applications coordonnées X_n. Cette tribu est nécessairement $\mathcal{F}$ tout entière. □

Théorème 13.3.3 *Soit Q une matrice stochastique sur E. Pour tout $x \in E$, il existe une unique probabilité, notée $\mathbb{P}_x$, sur $\Omega = E^{\mathbb{N}}$ telle que sous $\mathbb{P}_x$, le processus des coordonnées $(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov de matrice de transition Q, et $\mathbb{P}_x(X_0 = x) = 1$.*

Preuve. Soit $x \in E$. La proposition 13.3.1 permet de construire sur un espace de probabilité $(\Omega', \mathcal{F}', P')$ un processus $(X_n^x)_{n\in\mathbb{N}}$ qui est une chaîne de Markov de transition Q telle que $X_0^x = x$. On définit alors $\mathbb{P}_x$ comme la mesure image de P' par l'application

$$\begin{array}{rcl} \Omega' & \longrightarrow & \Omega \\ \omega' & \longrightarrow & (X_n^x(\omega'))_{n\in\mathbb{N}}. \end{array}$$

Cette application est mesurable grâce au lemme précédent. On a $\mathbb{P}_x(X_0 = x) = P'(X_0^x = x) = 1$ et de plus pour tous $x_0, x_1, \dots, x_n \in E$,

$$\begin{aligned} \mathbb{P}_x(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) &= P'(X_0^x = x_0, X_1^x = x_1, \dots, X_n^x = x_n) \\ &= P'(X_0^x = x_0) Q(x_0, x_1) \dots Q(x_{n-1}, x_n) \\ &= \mathbb{P}_x(X_0 = x_0) Q(x_0, x_1) \dots Q(x_{n-1}, x_n) \end{aligned}$$

ce qui montre que sous $\mathbb{P}_x$ le processus des coordonnées est une chaîne de Markov de transition Q (cf proposition 13.1.1).

Pour l'unicité, on remarque que si $\mathbb{P}'_x$ est une autre mesure de probabilité satisfaisant la propriété du théorème, les mesures $\mathbb{P}_x$ et $\mathbb{P}'_x$ coïncident sur les cylindres. Or les cylindres forment une classe stable par intersection finie et qui engendre la tribu $\mathcal{F}$. Le lemme de classe monotone montre alors que $\mathbb{P}_x = \mathbb{P}'_x$ (cf Corollaire 1.4.2). □

Remarques. (a) De la dernière assertion de la proposition 13.1.1, on déduit que, pour tout $n \geq 0$ et tous $x, y \in E$,

$$\mathbb{P}_x(X_n = y) = Q_n(x, y).$$

(b) Si μ est une mesure de probabilité sur E, on notera

$$\mathbb{P}_\mu = \sum_{x\in E} \mu(x)\, \mathbb{P}_x$$

qui définit une mesure de probabilité sur Ω. En écrivant la formule explicite pour $\mathbb{P}_\mu(X_0 = x_0, \dots, X_n = x_n)$, on vérifie immédiatement que sous $\mathbb{P}_\mu$, $(X_n)_{n\in\mathbb{N}}$ est une chaîne de Markov de transition Q, et X_0 a pour loi μ.

(c) Si $(X'_n)_{n\in\mathbb{N}}$ est une chaîne de Markov de matrice de transition Q et de loi initiale μ, alors pour toute partie mesurable B de $\Omega = E^{\mathbb{N}}$, on a

$$P((X'_n)_{n\in\mathbb{N}} \in B) = \mathbb{P}_\mu(B).$$

En effet cette égalité est vraie lorsque B est un cylindre, et on peut ensuite utiliser le même argument qu'à la fin de la preuve ci-dessus. Cette égalité montre que tous les résultats que nous établirons dans la suite pour la chaîne de Markov canonique (celle fournie par le théorème 13.3.3) se transporteront à une chaîne de Markov quelconque de même matrice de transition.

L'un des avantages importants de la chaîne de Markov canonique est de pouvoir utiliser les *opérateurs de translation.* Pour tout $k \in \mathbb{N}$ on définit l'application $\theta_k : \Omega \longrightarrow \Omega$ en posant

$$\theta_k((\omega_n)_{n\in\mathbb{N}}) = (\omega_{k+n})_{n\in\mathbb{N}}.$$

Le lemme 13.3.2 montre que ces applications sont mesurables.

On note $\mathcal{F}_n = \sigma(X_0, X_1, \dots, X_n)$ la filtration canonique sur Ω. On utilise aussi la notation $\mathbb{E}_x$ pour désigner l'espérance sous la probabilité $\mathbb{P}_x$.

Théorème 13.3.4 (Propriété de Markov simple) *Soient F et G deux fonctions mesurables positives sur Ω et soit $n \geq 0$. Supposons que F est $\mathcal{F}_n$-mesurable. Alors, pour tout $x \in E$,*

$$\mathbb{E}_x[F \cdot G \circ \theta_n] = \mathbb{E}_x[F\, \mathbb{E}_{X_n}[G]].$$

De manière équivalente,

$$\mathbb{E}_x[G \circ \theta_n \mid \mathcal{F}_n] = \mathbb{E}_{X_n}[G],$$

ce qu'on peut traduire en disant que la loi conditionnelle de $\theta_n(\omega)$ connaissant $(X_0, X_1, \dots, X_n)$ est $\mathbb{P}_{X_n}$.

Remarque. Cet énoncé se généralise aussitôt au cas où on remplace $\mathbb{E}_x$ par $\mathbb{E}_\mu$ pour n'importe quelle loi initiale μ. Il en sera de même pour l'énoncé suivant.

Preuve. Il suffit de montrer la première assertion, et pour cela de traiter le cas où

$$F = \mathbf{1}_{\{X_0 = x_0, X_1 = x_1, \dots, X_n = x_n\}}$$

pour $x_0, x_1, \dots, x_n \in E$. Considérons d'abord le cas où G est du même type :

$$G = \mathbf{1}_{\{X_0 = y_0, X_1 = y_1, \dots, X_p = y_p\}}$$

où $p \geq 0$ et $y_0, \dots, y_p \in E$. Dans ce cas, si $y \in E$,

$$\mathbb{E}_y[G] = \mathbf{1}_{\{y_0 = y\}} Q(y_0, y_1) \dots Q(y_{p-1}, y_p)$$

et par ailleurs

$$\begin{aligned} \mathbb{E}_x[F \cdot G \circ \theta_n] &= \mathbb{P}_x(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n, X_n = y_0, X_{n+1} = y_{n+1}, \dots, X_{n+p} = y_p) \\ &= \mathbf{1}_{\{x_0 = x\}} Q(x_0, x_1) \dots Q(x_{n-1}, x_n)\, \mathbf{1}_{\{y_0 = x_n\}} Q(y_0, y_1) \dots Q(y_{p-1}, y_p) \end{aligned}$$

de sorte qu'on obtient facilement le résultat. Un argument de classe monotone montre ensuite que le résultat reste vrai pour toute fonction $G = 1_A$, $A \in \mathcal{F}$, ce qui permet de conclure. $\square$

Le théorème précédent donne une forme générale de la propriété de Markov (simple) : la loi conditionnelle du futur $\theta_n(\omega)$ connaissant le passé $(X_0, X_1, \dots, X_n)$ ne dépend que du présent X_n. Il sera très important de pouvoir étendre cette propriété au cas où n est remplacé par un temps aléatoire T.

Pour illustrer l'intérêt de cette extension, considérons le problème de savoir si partant d'un point x la chaîne y revient infiniment souvent. Autrement dit, en notant

$$N_x = \sum_{n=0}^{\infty} \mathbf{1}_{\{X_n = x\}}$$

a-t-on $\mathbb{P}_x(N_x = \infty) = 1$? Il suffit en fait de vérifier que la chaîne revient au moins une fois en x. Si

$$H_x = \inf\{n \geq 1 : X_n = x\}$$

avec la convention habituelle $\inf \varnothing = +\infty$, on a l'équivalence

$$\mathbb{P}_x(N_x = \infty) = 1 \Leftrightarrow \mathbb{P}_x(H_x < \infty) = 1.$$

L'implication $\Rightarrow$ est triviale. Dans l'autre sens, supposons $\mathbb{P}_x(H_x < \infty) = 1$. Modulo l'extension de la propriété de Markov mentionnée ci-dessus, on sait que $\theta_{H_x}(\omega) = (\omega_{H_x(\omega)+n})_{n\in\mathbb{N}}$ a pour loi $\mathbb{P}_x$. Mais alors, en écrivant

$$N_x(\omega) = 1 + N_x(\theta_{H_x}(\omega))$$

on voit que N_x a même loi que $1 + N_x$ sous $\mathbb{P}_x$, ce qui n'est possible que si $\mathbb{N}_x = \infty$, $\mathbb{P}_x$ p.s.

Le théorème qui suit permet de rendre ce raisonnement rigoureux (le résultat obtenu sera repris et détaillé dans la partie suivante).

Théorème 13.3.5 (Propriété de Markov forte) *Soit T un temps d'arrêt de la filtration $(\mathcal{F}_n)$. Soient F et G deux fonctions mesurables positives sur Ω. Supposons que F est $\mathcal{F}_T$-mesurable. Alors, pour tout $x \in E$,*

$$\mathbb{E}_x[\mathbf{1}_{\{T<\infty\}}\, F \cdot G \circ \theta_T] = \mathbb{E}_x[\mathbf{1}_{\{T<\infty\}}\, F\, \mathbb{E}_{X_T}[G]].$$

De manière équivalente,

$$\mathbb{E}_x[\mathbf{1}_{\{T<\infty\}}\, G \circ \theta_T \mid \mathcal{F}_T] = \mathbf{1}_{\{T<\infty\}}\, \mathbb{E}_{X_T}[G].$$

Remarque. La v.a. X_T, définie sur l'ensemble $\mathcal{F}_T$-mesurable $\{T < \infty\}$, est $\mathcal{F}_T$-mesurable (cf Proposition 12.2.3 - dans le chapitre précédent on considère des processus à valeurs réelles, mais l'argument reste le même). La v.a. $\mathbb{E}_{X_T}[G]$, définie aussi sur l'ensemble $\{T < \infty\}$, est la composée des applications $\omega \longrightarrow X_T(\omega)$ et $x \to \mathbb{E}_x[G]$.

Preuve. Pour tout entier $n \geq 0$,

$$\mathbb{E}_x[\mathbf{1}_{\{T=n\}}\, F \cdot G \circ \theta_T] = \mathbb{E}_x[\mathbf{1}_{\{T=n\}}\, F \cdot G \circ \theta_n] = \mathbb{E}_x[\mathbf{1}_{\{T=n\}}\, F\, \mathbb{E}_{X_n}[G]]$$

d'après la propriété de Markov simple (théorème 13.3.4) appliquée en observant que $\mathbf{1}_{\{T=n\}}F$ est $\mathcal{F}_n$-mesurable parce que F est $\mathcal{F}_T$-mesurable (cf définition de la tribu $\mathcal{F}_T$ dans le chapitre précédent). Il suffit ensuite de sommer l'égalité obtenue sur toutes les valeurs de $n \in \mathbb{N}$. □

Corollaire 13.3.6 *Soit T un temps d'arrêt tel que $\mathbb{P}_x(T < \infty) = 1$. Supposons qu'il existe $y \in E$ tel que $\mathbb{P}_x(X_T = y) = 1$. Alors sous $\mathbb{P}_x$, $\theta_T(\omega)$ est indépendant de $\mathcal{F}_T$ et de loi $\mathbb{P}_y$.*

Preuve. Avec les notations du théorème, on a

$$\mathbb{E}_x[F \cdot G(\theta_T(\omega))] = \mathbb{E}_x[F\, \mathbb{E}_{X_T}[G]] = \mathbb{E}_x[F\, \mathbb{E}_y[G]] = \mathbb{E}_x[F]\, \mathbb{E}_y[G]$$

d'où les assertions de l'énoncé. □

13.4 La classification des états

A partir de maintenant, on utilise uniquement (sauf exception, notamment dans les exemples) la chaîne de Markov canonique construite dans le paragraphe précédent. Rappelons la notation : pour $x \in E$,

$$H_x = \inf\{n \geq 1 : X_n = x\}$$
$$N_x = \sum_{n=0}^{\infty} \mathbf{1}_{\{X_n = x\}}.$$

Proposition 13.4.1 (et définition) *Soit $x \in E$. On a :*

- *ou bien $\mathbb{P}_x(H_x < \infty) = 1$, et alors*

$$N_x = \infty\,, \quad \mathbb{P}_x \ p.s.$$

dans ce cas x est dit récurrent;

- *ou bien $\mathbb{P}_x(H_x < \infty) < 1$, et alors*

$$N_x < \infty\,, \quad \mathbb{P}_x \ p.s.$$

et plus précisément $\mathbb{E}_x[N_x] = 1/\mathbb{P}_x(H_x = \infty) < \infty$; dans ce cas x est dit transitoire.

Preuve. Pour tout entier $k \geq 1$, la propriété de Markov forte montre que

$$\begin{aligned}
\mathbb{P}_x(N_x \geq k+1) &= \mathbb{E}_x[\mathbf{1}_{\{H_x<\infty\}}\,\mathbf{1}_{\{N_x\geq k\}} \circ \theta_{H_x}] \\
&= \mathbb{E}_x[\mathbf{1}_{\{H_x<\infty\}}\,\mathbb{E}_x[\mathbf{1}_{\{N_x\geq k\}}]] \\
&= \mathbb{P}_x(H_x < \infty)\,\mathbb{P}_x(N_x \geq k).
\end{aligned}$$

Puisque $\mathbb{P}_x(N_x \geq 1) = 1$, une récurrence immédiate donne $\mathbb{P}_x(N_x \geq k) = \mathbb{P}_x(H_x < \infty)^{k-1}$. Si $\mathbb{P}_x(H_x < \infty) = 1$ il en découle aussitôt que $\mathbb{P}_x(N_x = \infty) = 1$. Si $\mathbb{P}_x(H_x < \infty) < 1$, on trouve

$$\mathbb{E}_x[N_x] = \sum_{k=1}^{\infty} \mathbb{P}_x(N_x \geq k) = \frac{1}{\mathbb{P}_x(H_x = \infty)} < \infty.$$

Définition 13.4.1 *Le noyau potentiel de la chaîne est la fonction $U : E \times E \longrightarrow [0,\infty]$ définie par*

$$U(x,y) = \mathbb{E}_x[N_y].$$

Proposition 13.4.2 (i) *Pour tous $x, y \in E$,*

$$U(x,y) = \sum_{n=0}^{\infty} Q_n(x,y).$$

(ii) *$U(x,x) = \infty$ si et seulement si x est récurrent.*
(iii) *Pour tous $x, y \in E$, avec $x \neq y$,*

$$U(x,y) = \mathbb{P}_x(H_y < \infty)\,U(y,y).$$

Preuve. La propriété (i) est obtenue en écrivant :

$$U(x,y) = \mathbb{E}_x\Big[\sum_{n=0}^{\infty} \mathbf{1}_{\{X_n=y\}}\Big] = \sum_{n=0}^{\infty} \mathbb{P}_x(X_n = y) = \sum_{n=0}^{\infty} Q_n(x,y).$$

La propriété (ii) est une conséquence immédiate de la proposition 13.4.1 et de la définition de U.

Enfin (iii) découle de la propriété de Markov forte :

$$\mathbb{E}_x[N_y] = \mathbb{E}_x[\mathbf{1}_{\{H_y<\infty\}}\, N_y \circ \theta_{H_y}] = \mathbb{E}_x[\mathbf{1}_{\{H_y<\infty\}}\, \mathbb{E}_y[N_y]] = \mathbb{P}_x(H_y < \infty)\, U(y,y).$$

Exemple. Considérons la chaîne de Markov sur $\mathbb{Z}^d$ de matrice de transition

$$Q((x_1,\ldots,x_d),(y_1,\ldots,y_d)) = \frac{1}{2^d} \prod_{i=1}^{d} \mathbf{1}_{\{|y_i-x_i|=1\}}$$

(c'est un cas particulier de marche aléatoire sur $\mathbb{Z}^d$). Cette chaîne de Markov issue de 0 a même loi que $(Y_n^1,\ldots,Y_n^d)_{n\in\mathbb{N}}$, où les processus $Y^1,\ldots,Y^d$ sont des copies indépendantes de la marche aléatoire simple (pile ou face) sur $\mathbb{Z}$, issue de 0. En conséquence,

$$Q_n(0,0) = P(Y_n^1 = 0,\ldots,Y_n^d = 0) = P(Y_n^1 = 0)^d.$$

Or $P(Y_n^1 = 0) = 0$ si n est impair, et si $n = 2k$ est pair, un argument de dénombrement simple montre que

$$P(Y_{2k}^1 = 0) = 2^{-2k}\, C_{2k}^k.$$

En conséquence,

$$U(0,0) = \sum_{k=0}^{\infty} Q_{2k}(0,0) = \sum_{k=0}^{\infty} (2^{-2k} C_{2k}^k)^d.$$

La formule de Stirling montre que

$$2^{-2k} C_{2k}^k \underset{k\to\infty}{\sim} \frac{(\frac{2k}{e})^{2k}\sqrt{4\pi k}}{2^{2k}((\frac{k}{e})^k\sqrt{2\pi k})^2} \underset{k\to\infty}{\sim} \sqrt{\frac{1}{\pi k}}.$$

Donc 0 est récurrent si $d = 1$ ou 2, et transitoire si $d \geq 3$.

On note R l'ensemble des états (points) récurrents.

Lemme 13.4.3 *Soit $x \in R$ et soit y un autre point de E tel que $U(x,y) > 0$. Alors $y \in R$ et $\mathbb{P}_y(H_x < \infty) = 1$, donc en particulier $U(y,x) > 0$.*

Preuve. Montrons d'abord que $\mathbb{P}_y(H_x < \infty) = 1$. Pour cela on écrit

$$\begin{aligned} 0 = \mathbb{P}_x(N_x < \infty) &\geq \mathbb{P}_x(H_y < \infty, H_x \circ \theta_{H_y} = \infty) \\ &= \mathbb{E}_x[\mathbf{1}_{\{H_y<\infty\}}\, \mathbf{1}_{\{H_x=\infty\}} \circ \theta_{H_y}] \\ &= \mathbb{E}_x[\mathbf{1}_{\{H_y<\infty\}}\, \mathbb{P}_y(H_x = \infty)] \\ &= \mathbb{P}_x(H_y < \infty)\, \mathbb{P}_y(H_x = \infty). \end{aligned}$$

L'hypothèse $U(x,y)>0$ entraîne $\mathbb{P}_x(H_y<\infty)>0$. On conclut que $\mathbb{P}_y(H_x=\infty)=0$.

Ensuite, on peut trouver des entiers $n_1,n_2\geq 1$ tels que $Q_{n_1}(x,y)>0$, et $Q_{n_2}(y,x)>0$. Pour tout entier $p\geq 0$, on a alors

$$Q_{n_2+p+n_1}(y,y)\geq Q_{n_2}(y,x)Q_p(x,x)Q_{n_1}(x,y)$$

et donc

$$U(y,y)\geq \sum_{p=0}^{\infty} Q_{n_2+p+n_1}(y,y)\geq Q_{n_2}(y,x)\Big(\sum_{p=0}^{\infty} Q_p(x,x)\Big)Q_{n_1}(x,y)=\infty$$

puisque $x\in R$ entraîne $\sum_{p=0}^{\infty} Q_p(x,x)=U(x,x)=\infty$. □

En conséquence du lemme, si $x\in R$ et $y\in E\backslash R$ on a $U(x,y)=0$: on ne peut pas passer d'un point récurrent à un point transitoire. Cette propriété joue un rôle important dans le théorème suivant.

Théorème 13.4.4 (Classification des états) *Il existe une partition de R*

$$R=\bigcup_{i\in I} R_i$$

telle qu'on ait les propriétés suivantes :

- *si $x\in R$, et si $i\in I$ est tel que $x\in R_i$, on a $\mathbb{P}_x$ p.s.*
  - — $N_y=+\infty\,,\qquad \forall y\in R_i$;
  - — $N_y=0\,,\qquad \forall y\in E\backslash R_i$;
- *si $x\in E\backslash R$ et $T=\inf\{n\geq 0: X_n\in R\}$, on a $\mathbb{P}_x$ p.s.*
  - *— ou bien $T=\infty$ et $N_y<\infty$, $\forall y\in E$;*
  - *— ou bien $T<\infty$ et il existe un indice (aléatoire) $j\in I$ tel que : $\forall n\geq T,\ X_n\in R_j$.*

Preuve. Pour $x,y\in R$, notons $x\sim y$ si $U(x,y)>0$. Il découle du lemme précédent qu'on ainsi défini une relation d'équivalence sur R (pour la transitivité, on observe que $Q_n(x,y)>0$ et $Q_m(y,z)>0$ entraînent $Q_{n+m}(x,z)>0$. La partition du théorème correspond alors aux classes d'équivalence pour cette relation d'équivalence, qu'on appelle aussi les classes de récurrence de la chaîne de Markov.

Soit $i\in I$ et $x\in R_i$. On a $U(x,y)=0$ pour tout $y\in E\backslash R_i$ (dans le cas $y\in E\backslash R$ on utilise le lemme) et donc $N_y=0$, $\mathbb{P}_x$ p.s. pour tout $y\in E\backslash R_i$. En revanche, si $y\in R_i$, on a $\mathbb{P}_x(H_y<\infty)=1$ d'après le lemme, et la propriété de Markov forte montre que

$$\mathbb{P}_x(N_y=\infty)=\mathbb{E}_x(\mathbf{1}_{\{H_y<\infty\}}\,\mathbf{1}_{\{N_y=\infty\}}\circ\theta_{H_y}]=\mathbb{P}_x(H_y<\infty)\,\mathbb{P}_y(N_y=\infty)=1.$$

Si $x\in E\backslash R$ et $T=\infty$, alors on déduit facilement de la propriété de Markov forte que $N_y<\infty$ pour tout $y\in E\backslash R$. Si $T<\infty$, notons j l'indice (aléatoire) tel que $X_T\in R_j$. En appliquant la propriété de Markov forte en T, et la première partie de l'énoncé, on obtient aisément que $X_n\in R_j$ pour tout $n\geq T$. □

Définition 13.4.2 *La chaîne est dite irréductible si $U(x, y) > 0$ pour tous $x, y \in E$.*

Corollaire 13.4.5 *Si la chaîne est irréductible :*

- *ou bien tous les états sont récurrents, il existe une seule classe de récurrence et on a pour tout $x \in E$,*
$$\mathbb{P}_x(N_y = \infty,\ \forall y \in E) = 1.$$
- *ou bien tous les états sont transitoires et alors, pour tout $x \in E$,*
$$\mathbb{P}_x(N_y < \infty,\ \forall y \in E) = 1.$$

Lorsque E est fini, seul le premier cas peut se produire.

Preuve. S'il existe un état récurrent, le lemme 13.4.3 montre aussitôt que tous les états sont récurrents, et puisque $U(x, y) > 0$ pour tous $x, y \in E$, on voit aussi qu'il y a une seule classe de récurrence. Le reste découle du théorème, à l'exception de la dernière assertion : si E est fini et si on suppose que tous les états sont transitoires, on a

$$\mathbb{P}_x \text{ p.s. },\quad \sum_{y \in E} N_y < \infty$$

ce qui est absurde puisque

$$\sum_{y \in E} N_y = \sum_{y \in E} \sum_{n=0}^{\infty} \mathbf{1}_{\{X_n = y\}} = \sum_{n=0}^{\infty} \sum_{y \in E} \mathbf{1}_{\{X_n = y\}} = \infty.$$

$\square$

Une chaîne de Markov irréductible dont les états sont récurrents sera dite récurrente irréductible.

Exemples. Nous reprenons maintenant les différents exemples introduits ci-dessus pour discuter dans chaque cas la classification des états. Avant cela, insistons sur le fait que les résultats obtenus pour la chaîne de Markov canonique se traduisent immédiatement pour une chaîne de Markov quelconque $(Y_n)_{n \in \mathbb{N}}$ de transition Q (et inversement). Par exemple, si $Y_0 = y$, en notant $N_x^Y = \sum_{n=0}^{\infty} \mathbf{1}_{\{Y_n = x\}}$, on a pour tout $k \in \overline{\mathbb{N}}$,

$$P(N_x^Y = k) = \mathbb{P}_y(N_x = k))$$

puisque le terme de gauche s'écrit aussi bien

$$P((Y_n)_{n \in \mathbb{N}} \in B)$$

avec $B = \{\omega \in E^{\mathbb{N}} : N_x(\omega) = k\}$, et il suffit d'utiliser la remarque (b) suivant le théorème 13.3.3.

(1) **Cas de variables aléatoires indépendantes de loi** μ. Dans ce cas $Q(x, y) = \mu(y)$. On voit facilement que y est récurrent ssi $\mu(y) > 0$, et il y a une seule classe de récurrence. La chaîne est irréductible ssi $\mu(y) > 0$ pour tout $y \in E$.

(2) **Marche aléatoire sur** $\mathbb{Z}$. On a

$$Y_n = Y_0 + \sum_{i=1}^{n} \xi_i$$

où les v.a. ξ_i, à valeurs dans $\mathbb{Z}$, sont indépendantes et de loi μ (et indépendantes de Y_0). Dans ce cas, puisque $Q(x, y) = \mu(y - x)$, on voit aisément que $U(x, y)$ est fonction de $y - x$, et donc tous les états sont du même type, récurrent ou transitoire.

Théorème 13.4.6 *Supposons $E[|\xi_1|] < \infty$ et soit $m = E[\xi_1]$.*

(i) *Si $m \neq 0$, tous les états sont transitoires.*

(ii) *Si $m = 0$, tous les états sont récurrents. De plus, la chaîne est irréductible ssi le sous-groupe engendré par $\{y \in \mathbb{Z} : \mu(y) > 0\}$ est $\mathbb{Z}$ tout entier.*

Preuve. (i) Si $m \neq 0$, la loi forte des grands nombres montre aussitôt que $|Y_n| \longrightarrow \infty$ p.s. et donc tous les états sont transitoires.
(ii) Supposons que $m = 0$ et que 0 est transitoire, donc $U(0, 0) < \infty$. Nous allons voir que ceci conduit à une contradiction. Sans perte de généralité, on suppose dans la suite que $Y_0 = 0$. On observe que, pour tout $x \in \mathbb{Z}$,

$$U(0, x) \leq U(x, x) = U(0, 0)$$

la première inégalité découlant de la proposition 13.4.2(iii). En conséquence, pour tout $n \geq 1$,

$$\sum_{|x| \leq n} U(0, x) \leq (2n + 1)U(0, 0) \leq Cn \tag{13.2}$$

avec $C = 3U(0, 0) < \infty$.

D'autre part, on sait que $n^{-1}Y_n$ converge p.s., donc aussi en probabilité, vers 0. Si on pose $\varepsilon = (4C)^{-1}$, on peut trouver N assez grand pour que, pour tout $n \geq N$,

$$P(|Y_n| \leq \varepsilon n) > \frac{1}{2},$$

ou de manière équivalente,

$$\sum_{|x| \leq \varepsilon n} Q_n(0, x) > \frac{1}{2}.$$

Si $n \geq p \geq N$, on a aussi

$$\sum_{|x| \leq \varepsilon n} Q_p(0, x) \geq \sum_{|x| \leq \varepsilon p} Q_p(0, x) > \frac{1}{2}$$

puis en sommant sur p,

$$\sum_{|x| \leq \varepsilon n} U(0, x) \geq \sum_{p=N}^{n} \sum_{|x| \leq \varepsilon p} Q_p(0, x) > \frac{n - N}{2}.$$

Mais d'autre part, d'après (13.2), si $\varepsilon n \geq 1$,

$$\sum_{|x| \leq \varepsilon n} U(0,x) \leq C\varepsilon n = \frac{n}{4}.$$

On obtient une contradiction dès que n est assez grand.

Il reste à établir la dernière assertion. Notons G le sous-groupe engendré par $\{x \in \mathbb{Z} : \mu(x) > 0\}$. Il est immédiat que

$$P(Y_n \in G, \forall n \in \mathbb{N}) = 1$$

(rappelons que nous avons pris $Y_0 = 0$). Cela montre que si $G \neq \mathbb{Z}$, la chaîne n'est pas irréductible. Inversement, supposons que $G = \mathbb{Z}$. Alors, notons

$$H = \{x \in \mathbb{Z} : U(0,x) > 0\}$$

et observons que H est un sous-groupe de $\mathbb{Z}$:

- si $x, y \in H$, l'inégalité

$$Q_{n+p}(0, x+y) \geq Q_n(0,x)\, Q_p(x, x+y) = Q_n(0,x)\, Q_p(0,y)$$

  montre que $x + y \in H$;

- si $x \in H$, comme 0 est récurrent, la condition $U(0,x) > 0$ entraîne $U(x,0) > 0$ (lemme 13.4.3) et puisque $U(x,0) = U(0,-x)$ on a bien $-x \in H$.

Finalement, puisque H contient $\{x \in \mathbb{Z} : \mu(x) > 0\}$, on a forcément $H = \mathbb{Z}$. □

Par exemple, si $\mu = \frac{1}{2}\delta_{-2} + \frac{1}{2}\delta_2$, tous les états sont récurrents, mais il y a deux classes de récurrence, les entiers pairs et les entiers impairs.

(3) **Marche aléatoire sur un graphe**. On considère ici le cas d'un graphe fini : E est fini et A est un sous-ensemble de $\mathcal{P}_2(E)$ tel que, pour tout $x \in E$, $A_x := \{y \in E : \{x,y\} \in A\}$ est non vide. Le graphe est dit connexe si pour tous $x, y \in E$, on peut trouver un entier $p \geq 0$ et des élements $x_0 = x, x_1, \ldots, x_{p-1}, x_p = y$ de E tels que $\{x_{i-1}, x_i\} \in A$ pour tout $i \in \{1, \ldots, p\}$.

Proposition 13.4.7 *La marche aléatoire simple sur un graphe fini connexe est récurrente irréductible.*

Preuve. Le caractère irréductible de la chaîne découle de la connexité du graphe. Il suffit ensuite d'appliquer le corollaire 13.4.5. □

(4) **Processus de branchement**. Dans ce cas $E = \mathbb{N}$ et $Q(x,y) = \mu^{*x}(y)$. On remarque que l'état 0 est toujours absorbant, au sens où

$$\mathbb{P}_0(\forall n \in \mathbb{N}\,,\, X_n = 0) = 1.$$

En conséquence 0 est aussi récurrent.

Dans la proposition suivante, nous écartons le cas trivial $\mu = \delta_1$, où tous les états sont absorbants.

Proposition 13.4.8 0 *est le seul état récurrent. En conséquence, on a p.s.*

- *ou bien* $\exists N : \forall n \geq N \, , \, X_n = 0$.
- *ou bien* $X_n \longrightarrow +\infty$ *quand* $n \to \infty$.

Remarque. On a vu dans le chapitre précédent que le premier cas se produit avec probabilité 1 si $m = \sum k\mu(k) \leq 1$, et que le second cas se produit avec probabilité strictement positive si $m > 1$ (sous l'hypothèse supplémentaire que μ a un moment d'ordre 2).

Preuve. Supposons d'abord que $\mu(0) > 0$. Si $x \geq 1$, $U(x, 0) \geq \mathbb{P}_x(X_1 = 0) = \mu(0)^x > 0$ alors que $U(0, x) = 0$. Cela n'est possible que si x est transitoire. Traitons ensuite le cas où $\mu(0) = 0$. Comme nous excluons le cas $\mu = \delta_1$, il existe alors $k \geq 2$ tel que $\mu(k) > 0$. Alors, pour tout $x \geq 1$, $\mathbb{P}_x(X_1 > x) > 0$, ce qui entraîne qu'il existe $y > x$ tel que $U(x, y) > 0$. Comme on a clairement $U(y, x) = 0$, on conclut encore que x est transitoire. Les autres assertions découlent maintenant du théorème 13.4.4. $\square$

13.5 Mesures invariantes

Définition 13.5.1 *Soit μ une mesure positive sur E, telle que $\mu(x) < \infty$ pour tout $x \in E$ et μ n'est pas la mesure identiquement nulle. On dit que μ est invariante pour la matrice de transition Q (ou simplement invariante s'il n'y a pas ambiguïté) si*

$$\forall y \in E \, , \quad \mu(y) = \sum_{x \in E} \mu(x) Q(x, y).$$

Sous forme matricielle, la condition d'invariance s'écrit $\mu Q = \mu$. Puisque pour tout n, $Q_n = (Q)^n$, on peut itérer cette relation et obtenir que $\mu Q_n = \mu$ pour tout $n \in \mathbb{N}$.

Interprétation. Supposons de plus que $\mu(E) < \infty$ (ce qui sera toujours le cas si E est fini). Quitte à remplacer μ par $\mu(E)^{-1}\mu$, on peut supposer $\mu(E) = 1$. Alors, pour toute fonction $f : E \longrightarrow \mathbb{R}_+$,

$$\mathbb{E}_\mu[f(X_1)] = \sum_{x \in E} \mu(x) \sum_{y \in E} Q(x, y) f(y) = \sum_{y \in E} f(y) \sum_{x \in E} \mu(x) Q(x, y) = \sum_{y \in E} \mu(y) f(y)$$

ce qui montre que sous $\mathbb{P}_\mu$, X_1 a même loi μ que X_0. En utilisant la relation $\mu Q_n = Q$, on obtient de même que pour tout $n \in \mathbb{N}$ la loi de X_n sous $\mathbb{P}_\mu$ est μ. Plus précisément, pour toute fonction $F : \Omega \longrightarrow \mathbb{R}_+$ mesurable,

$$\mathbb{E}_\mu[F \circ \theta_1] = \mathbb{E}_\mu[\mathbb{E}_{X_1}[F]] = \sum_{x \in E} \mu(x) \, \mathbb{E}_x[F] = \mathbb{E}_\mu[F]$$

ce qui montre que sous $\mathbb{P}_\mu$, $(X_{1+n})_{n \in \mathbb{N}}$ a même loi que $(X_n)_{n \in \mathbb{N}}$ (et de même, pour tout entier $k \geq 0$, $(X_{k+n})_{n \in \mathbb{N}}$ a même loi que $(X_n)_{n \in \mathbb{N}}$).

Exemple. Pour toute marche aléatoire sur $\mathbb{Z}^d$ ($Q(x, y) = \gamma(y - x)$ ne dépend que la différence $y - x$), on vérifie immédiatement que la mesure de comptage sur $\mathbb{Z}^d$ est invariante.

Définition 13.5.2 *Soit μ une mesure positive non triviale sur E, telle que $\mu(x) < \infty$ pour tout $x \in E$. On dit que μ est réversible si*

$$\forall x, y \in E\,, \quad \mu(x)Q(x,y) = \mu(y)Q(y,x).$$

Proposition 13.5.1 *Toute mesure réversible est invariante.*

Preuve. Si μ est réversible,

$$\sum_{x \in E} \mu(x)Q(x,y) = \sum_{x \in E} \mu(y)Q(y,x) = \mu(y).$$

□

En revanche, il existe des mesures invariantes qui ne sont pas réversibles : nous avons vu que la mesure de comptage est invariante pour toute marche aléatoire sur $\mathbb{Z}^d$, cependant elle n'est réversible que si la loi de saut γ est symétrique ($\gamma(x) = \gamma(-x)$).

Exemples. (a) **Pile ou face biaisé**. C'est la marche aléatoire sur $\mathbb{Z}$ de matrice de transition

$$\begin{aligned} Q(i, i+1) &= p \\ Q(i, i-1) &= q = 1 - p \end{aligned}$$

où $p \in]0,1[$. Dans ce cas, on vérifie aisément que la mesure

$$\mu(i) = (\frac{p}{q})^i\,, \qquad i \in \mathbb{Z}$$

est réversible, donc invariante. Remarquons que μ est différente de la mesure de comptage (qui est aussi invariante) sauf dans le cas $p = 1/2$.

(b) **Marche aléatoire sur un graphe**. La mesure

$$\mu(x) = \mathrm{Card}(A_x)$$

est réversible. En effet, si $\{x, y\} \in A$,

$$\mu(x)Q(x,y) = \mathrm{Card}(A_x)\,\frac{1}{\mathrm{Card}(A_x)} = 1 = \mu(y)Q(y,x).$$

(c) **Modèle d'urne d'Ehrenfest**. C'est la chaîne de Markov dans $\{0, 1, \ldots, k\}$ de matrice de transition

$$\begin{aligned} Q(j, j+1) &= \tfrac{k-j}{k} \qquad && \text{si } 0 \leq j \leq k-1 \\ Q(j, j-1) &= \tfrac{j}{k} && \text{si } 1 \leq j \leq k. \end{aligned}$$

Une mesure μ est réversible ssi

$$\mu(j)\,\frac{k-j}{k} = \mu(j+1)\,\frac{j+1}{k}$$

pour tout $0 \leq j \leq k-1$. On trouve aisément que

$$\mu(j) = C_k^j$$

convient.

Théorème 13.5.2 *Soit x un point récurrent. La formule*

$$\mu(y) = \mathbb{E}_x\Big[\sum_{k=0}^{H_x-1} \mathbf{1}_{\{X_k=y\}}\Big]$$

définit une mesure invariante. De plus, $\mu(y) > 0$ ssi y appartient à la classe de récurrence de x.

Preuve. Remarquons d'abord que si y n'est pas dans la classe de récurrence de x on a $\mathbb{E}_x[N_y] = U(x,y) = 0$, et donc a fortiori $\mu(y) = 0$.

Ensuite, on écrit pour tout $y \in E$,

$$\begin{aligned}
\mu(y) &= \mathbb{E}_x\Big[\sum_{k=1}^{H_x} \mathbf{1}_{\{X_k=y\}}\Big] \\
&= \sum_{z\in E} \mathbb{E}_x\Big[\sum_{k=1}^{H_x} \mathbf{1}_{\{X_{k-1}=z,\, X_k=y\}}\Big] \\
&= \sum_{z\in E}\sum_{k=1}^{\infty} \mathbb{E}_x\Big[\mathbf{1}_{\{k\le H_x,\, X_{k-1}=z\}}\, \mathbf{1}_{\{X_k=y\}}\Big] \\
&= \sum_{z\in E}\sum_{k=1}^{\infty} \mathbb{E}_x\Big[\mathbf{1}_{\{k\le H_x,\, X_{k-1}=z\}}\Big]\, Q(z,y) \\
&= \sum_{z\in E} \mathbb{E}_x\Big[\sum_{k=1}^{H_x} \mathbf{1}_{\{X_{k-1}=z\}}\Big]\, Q(z,y) \\
&= \sum_{z\in E} \mu(z)Q(z,y).
\end{aligned}$$

Dans la quatrième égalité, on a utilisé le fait que l'événement $\{k \le H_x,\, X_{k-1} = z\}$ est $\mathcal{F}_{k-1}$-mesurable pour appliquer la propriété de Markov à l'instant $k-1$.

On a obtenu l'identité $\mu Q = \mu$, qu'on peut itérer pour avoir $\mu Q_n = \mu$ pour tout entier $n \ge 0$. En particulier, pour tout entier $n \ge 0$,

$$\mu(x) = 1 = \sum_{z\in E} \mu(z)Q_n(z,x).$$

Soit y un point de la classe de récurrence de x. Alors, il existe $n \ge 0$ tel que $Q_n(y,x) > 0$, et la formule précédente montre que $\mu(y) < \infty$. On peut aussi trouver m tel que $Q_m(x,y) > 0$, et on a

$$\mu(y) = \sum_{z\in E} \mu(z)Q_m(z,y) \ge Q_m(x,y) > 0.$$

Remarque. S'il existe plusieurs classes de récurrence R_i, $i \in I$, alors en choisissant pour chaque $i \in I$ un point $x_i \in R_i$ et en posant

$$\mu_i(y) = \mathbb{E}_{x_i}\Big[\sum_{k=0}^{H_{x_i}-1} \mathbf{1}_{\{X_k=y\}}\Big]$$

on construit des mesures invariantes à supports disjoints.

Théorème 13.5.3 *Supposons la chaîne récurrente irréductible. Alors la mesure invariante est unique à une constante multiplicative près.*

Preuve. Soit μ une mesure invariante. On montre par récurrence que, pour tout entier $p \geq 0$, pour tous $x, y \in E$,

$$\mu(y) \geq \mu(x)\,\mathbb{E}_x\Big[\sum_{k=0}^{p\wedge(H_x-1)} \mathbf{1}_{\{X_k=y\}}\Big]. \tag{13.3}$$

D'abord, si $y = x$, l'inégalité est immédiate (avec même une égalité). On suppose donc $y \neq x$. Si $p = 0$, l'inégalité (13.3) est triviale. On suppose que (13.3) est vraie à l'ordre p. Alors,

$$\begin{aligned}
\mu(y) &= \sum_{z\in E} \mu(z)\,Q(z,y)\\
&\geq \mu(x)\sum_{z\in E}\mathbb{E}_x\Big[\sum_{k=0}^{p\wedge(H_x-1)} \mathbf{1}_{\{X_k=z\}}\Big]Q(z,y)\\
&= \mu(x)\sum_{z\in E}\sum_{k=0}^{p}\mathbb{E}_x\Big[\mathbf{1}_{\{X_k=z,\,k\leq H_x-1\}}\Big]Q(z,y)\\
&= \mu(x)\sum_{z\in E}\sum_{k=0}^{p}\mathbb{E}_x\Big[\mathbf{1}_{\{X_k=z,\,k\leq H_x-1\}}\,\mathbf{1}_{\{X_{k+1}=y\}}\Big]\\
&= \mu(x)\mathbb{E}_x\Big[\sum_{k=0}^{p\wedge(H_x-1)} \mathbf{1}_{\{X_{k+1}=y\}}\Big]\\
&= \mu(x)\mathbb{E}_x\Big[\sum_{k=1}^{(p+1)\wedge H_x} \mathbf{1}_{\{X_k=y\}}\Big],
\end{aligned}$$

ce qui donne le résultat voulu à l'ordre $p+1$. De manière analogue à la preuve du théorème précédent, on a utilisé le fait que l'événement $\{X_k = z,\, k \leq H_x - 1\}$ est $\mathcal{F}_k$-mesurable pour appliquer la propriété de Markov à l'instant k.

En faisant tendre p vers $+\infty$ dans (13.3) on trouve

$$\mu(y) \geq \mu(x)\,\mathbb{E}_x\Big[\sum_{k=0}^{H_x-1} \mathbf{1}_{\{X_k=y\}}\Big].$$

Fixons $x \in E$. La mesure

$$\nu_x(y) = \mathbb{E}_x\Big[\sum_{k=0}^{H_x-1} \mathbf{1}_{\{X_k=y\}}\Big]$$

est invariante (théorème 13.5.2), et on a $\mu(y) \geq \mu(x)\nu_x(y)$ pour tout $y \in E$. Donc, pour tout $n \geq 1$,

$$\mu(x) = \sum_{z \in E} \mu(z) Q_n(z,x) \geq \sum_{z \in E} \mu(x)\nu_x(z) Q_n(z,x) = \mu(x)\nu_x(x) = \mu(x),$$

ce qui montre que l'égalité $\mu(z) = \mu(x)\nu_x(z)$ a lieu pour tout z tel que $Q_n(z,x) > 0$. L'irréductibilité assure que pour tout $z \in E$ on peut trouver un entier n tel que $Q_n(z,x) > 0$, et on conlut donc que $\mu = \mu(x)\nu_x$, ce qui termine la preuve. □

Corollaire 13.5.4 *Supposons la chaîne récurrente irréductible. Alors :*

(i) *Ou bien il existe une mesure de probabilité invariante μ, et on a pour tout $x \in E$,*

$$\mathbb{E}_x[H_x] = \frac{1}{\mu(x)}.$$

(ii) *Ou bien toute mesure invariante a une masse totale infinie, et on a pour tout $x \in E$,*

$$\mathbb{E}_x[H_x] = \infty.$$

La chaîne est dite récurrente positive dans le cas (i) *et récurrente nulle dans le cas* (ii).

Remarque. Si E est fini seul le cas (i) se produit.

Preuve. D'après le théorème 13.5.3, toutes les mesures invariantes sont proportionnelles. Donc ou bien elles sont toutes de masse totale infinie (cas (ii)) ou bien elles sont toutes finies, et on peut normaliser pour en trouver une qui soit une mesure de probabilité (cas (i)). Dans le cas (i), soit μ l'unique mesure de probabilité invariante et soit $x \in E$. Alors, si ν_x désigne la mesure invariante fournie par le théorème 13.5.2,

$$\nu_x(y) = \mathbb{E}_x\Big[\sum_{k=0}^{H_x-1} \mathbf{1}_{\{X_k=y\}}\Big],$$

μ est proportionnelle à ν_x : $\mu = C\nu_x$ avec $C > 0$. En écrivant $1 = \mu(E) = C\,\nu_x(E)$, on trouve $C = (\nu_x(E))^{-1}$, d'où

$$\mu(x) = \frac{\nu_x(x)}{\nu_x(E)} = \frac{1}{\nu_x(E)}.$$

Or

$$\nu_x(E) = \sum_{y \in E} \mathbb{E}_x\Big[\sum_{k=0}^{H_x-1} \mathbf{1}_{\{X_k=y\}}\Big] = \mathbb{E}_x\Big[\sum_{k=0}^{H_x-1}\Big(\sum_{y \in E} \mathbf{1}_{\{X_k=y\}}\Big)\Big] = \mathbb{E}_x[H_x].$$

Dans le cas (ii), ν_x est infinie, et donc, par le même calcul,

$$\mathbb{E}_x[H_x] = \nu_x(E) = \infty.$$

Proposition 13.5.5 *Supposons la chaîne irréductible. S'il existe une mesure invariante finie, la chaîne est récurrente (et donc récurrente positive).*

Preuve. Soit γ une mesure invariante finie, et soit $y \in E$ tel que $\gamma(y) > 0$. Pour tout $x \in E$, la proposition 13.4.2(iii) donne l'inégalité

$$\sum_{n=0}^{\infty} Q_n(x,y) = U(x,y) \leq U(y,y).$$

On multiplie les deux membres de cette inégalité par $\gamma(x)$ et on somme sur toutes les valeurs de $x \in E$. Il vient

$$\sum_{n=0}^{\infty} \gamma Q_n(y) \leq \gamma(E)\, U(y,y).$$

Puisque γ est invariante on a $\gamma Q_n(y) = \gamma(y) > 0$ pour tout $n \geq 0$. On conclut donc que

$$\gamma(E)\, U(y,y) = \infty.$$

Comme $\gamma(E) < \infty$, cela entraîne que $U(y,y) = \infty$. Donc y est récurrent et puisque la chaîne est irréductible elle est récurrente (corollaire 13.4.5). $\square$

Remarque. L'existence d'une mesure invariante infinie ne permet pas de conclure : considérer par exemple le pile ou face biaisé (exemple (1) ci-dessus après la proposition 13.5.1) qui n'est récurrent que si $p = 1/2$.

Exemple. Soit $p \in]0,1[$. Considérons la chaîne de Markov sur $E = \mathbb{N}$ de matrice de transition

$$\begin{aligned} &Q(k,k+1) = p\ ,\ \ Q(k,k-1) = 1-p\ , \qquad \text{si } k \geq 1, \\ &Q(0,1) = 1. \end{aligned}$$

Cette chaîne est irréductible. De plus on vérifie immédiatement que la mesure μ définie par

$$\begin{aligned} &\mu(k) = \Big(\frac{p}{1-p}\Big)^{k-1}, \qquad \text{si } k \geq 1, \\ &\mu(0) = 1-p\ , \end{aligned}$$

est réversible donc invariante.

Si $p < \frac{1}{2}$, la mesure μ est finie, et la proposition 13.5.5 entraîne que la chaîne est récurrente positive. (Exercice : Montrer que la chaîne est récurrente nulle si $p = \frac{1}{2}$, et transitoire si $p > \frac{1}{2}$.)

13.6 Comportement asymptotique

Nous continuons à considérer la chaîne de Markov canonique associée à une matrice de transition Q.

Théorème 13.6.1 *Supposons la chaîne récurrente irréductible, et soit μ une mesure invariante. Soient f et g deux fonctions positives sur E telles que $\int f\, d\mu < \infty$ et $0 < \int g\, d\mu < \infty$. Alors, pour tout $x \in E$ on a $\mathbb{P}_x$ p.s.*

$$\frac{\sum_{k=0}^{n} f(X_k)}{\sum_{k=0}^{n} g(X_k)} \underset{n\to\infty}{\longrightarrow} \frac{\int f\, d\mu}{\int g\, d\mu}.$$

Remarque. Le résultat reste vrai si $\mu(f) = \infty$. Il suffit d'utiliser un argument de comparaison en écrivant $f = \lim \uparrow f_k$, avec des fonctions positives f_k telles que $\int f_k d\mu < \infty$.

Corollaire 13.6.2 *Si la chaîne de Markov est irréductible et récurrente positive, et si μ désigne l'unique probabilité invariante, on a $\mathbb{P}_x$ p.s.*

$$\frac{1}{n}\sum_{k=0}^{n} f(X_k) \underset{n\to\infty}{\longrightarrow} \int f\, d\mu.$$

Le corollaire découle immédiatement du théorème en prenant $g = 1$ dans l'énoncé.

Preuve du théorème 13.6.1. On définit les temps d'arrêt

$$T_0 = 0\ ,\ T_1 = H_x$$

et par récurrence

$$T_{n+1} = \inf\{k > T_n : X_k = x\}.$$

Le temps T_n est l'instant du n-ième retour en x de la chaîne. Puisque l'état x est récurrent, tous ces temps d'arrêt sont finis p.s. On pose aussi pour tout $k \geq 0$,

$$Z_k(f) = \sum_{n=T_k}^{T_{k+1}-1} f(X_n).$$

Lemme 13.6.3 *Les v.a. $Z_k(f)$, $k = 0, 1, 2, \ldots$, sont indépendantes et de même loi.*

Preuve. Soient $g_0, g_1, g_2, \ldots$ des fonctions mesurables bornées sur $\mathbb{R}_+$. Il suffit de montrer que, pour tout entier $k \geq 0$, on a

$$\mathbb{E}_x\Big[\prod_{i=0}^{k} g_i(Z_i(f))\Big] = \prod_{i=0}^{k} \mathbb{E}_x[g_i(Z_0(f))].$$

On démontre cette identité par récurrence sur k. Pour $k = 0$ il n'y a rien à montrer. Pour passer de l'ordre $k-1$ à l'ordre k, on observe que :

- les v.a. $Z_0(f), Z_1(f), \ldots, Z_{k-1}(f)$ sont $\mathcal{F}_{T_k}$-mesurables (exercice !);
- la suite translatée $\theta_{T_k}(\omega)$ est indépendante de $\mathcal{F}_{T_k}$ et de loi $\mathbb{P}_x$, d'après le corollaire 13.3.6;
- on a $Z_k(f) = Z_0(f) \circ \theta_{T_k}$, par construction.

Il découle de tout ceci que

$$\mathbb{E}_x\Big[\prod_{i=0}^{k} g_i(Z_i(f))\Big] = \mathbb{E}_x\Big[\Big(\prod_{i=0}^{k-1} g_i(Z_i(f))\Big) g_k(Z_0(f)\circ\theta_{T_k})\Big] = \mathbb{E}_x\Big[\prod_{i=0}^{k-1} g_i(Z_i(f))\Big]\,\mathbb{E}_x[g_k(Z_0(f))],$$

d'où le résultat voulu à l'ordre k. □

Nous revenons à la preuve du théorème. Si ν_x désigne comme précédemment la mesure invariante construite dans le théorème 13.5.2, on a $\mu = \mu(x)\nu_x$ puisque $\nu_x(x) = 1$ et que toutes les mesures invariantes sont proportionnelles (théorème 13.5.3). On observe alors que

$$\mathbb{E}_x[Z_0(f)] = \mathbb{E}_x\Big[\sum_{k=0}^{H_x-1}\sum_{y\in E} f(y)\,\mathbf{1}_{\{X_k=y\}}\Big] = \sum_{y\in E} f(y)\,\nu_x(y) = \frac{\int f\,d\mu}{\mu(x)}.$$

Le lemme 13.6.3 et la loi forte des grands nombres montrent ensuite que $\mathbb{P}_x$ p.s.

$$\frac{1}{n}\sum_{k=0}^{n-1} Z_k(f) \underset{n\to\infty}{\longrightarrow} \frac{\int f\,d\mu}{\mu(x)}. \tag{13.4}$$

Pour tout entier n, notons $N_x(n)$ le nombre de retours en x effectués par la chaîne avant l'instant n, de sorte que $T_{N_x(n)} \le n < T_{N_x(n)+1}$. En écrivant

$$\frac{\sum_{k=0}^{T_{N_x(n)}-1} f(X_k)}{N_x(n)} \le \frac{\sum_{k=0}^{n} f(X_k)}{N_x(n)} \le \frac{\sum_{k=0}^{T_{N_x(n)+1}-1} f(X_k)}{N_x(n)}$$

ce qui équivaut à

$$\frac{\sum_{j=0}^{N_x(n)-1} Z_j(f)}{N_x(n)} \le \frac{\sum_{k=0}^{n} f(X_k)}{N_x(n)} \le \frac{\sum_{j=0}^{N_x(n)} Z_j(f)}{N_x(n)}$$

on déduit de la convergence (13.4) que $\mathbb{P}_x$ p.s.

$$\frac{1}{N_x(n)}\sum_{k=0}^{n} f(X_k) \underset{n\to\infty}{\longrightarrow} \frac{\int f\,d\mu}{\mu(x)}.$$

Il suffit ensuite d'utiliser le même résultat avec f remplacée par g pour finir la preuve. $\square$

Corollaire 13.6.4 *Supposons la chaîne récurrente irréductible. Alors, pour tout $x \in E$,*

(i) *dans le cas récurrent positif,*

$$\frac{1}{n}\sum_{k=0}^{n-1}\mathbf{1}_{\{X_k=x\}} \xrightarrow[n\to\infty]{\text{p.s.}} \mu(x),$$

où μ est l'unique probabilité invariante;

(ii) *dans le cas récurrent nul,*

$$\frac{1}{n}\sum_{k=0}^{n-1}\mathbf{1}_{\{X_k=x\}} \xrightarrow[n\to\infty]{\text{p.s.}} 0.$$

Dans les deux cas la convergence a lieu pour toute loi initiale de la chaîne.

Définition 13.6.1 *Soit x un point récurrent, et*

$$L_x = \{n \geq 0 : Q_n(x,x) > 0\}.$$

La période de x, notée $d(x)$, est le PGCD de L_x.

Remarque. Puisque L_x est stable par addition ($Q_{n+m}(x,x) \geq Q_n(x,x)Q_m(x,x)$), le sous groupe engendré par L_x est $L_x - L_x = d(x)\mathbb{Z}$.

Proposition 13.6.5 *Supposons la chaîne récurrente irréductible.*

(i) *Tous les points ont la même période, appelée la période de la chaîne et notée d.*

(ii) *Si $d = 1$ (la chaîne est alors dite apériodique), pour tous $x, y \in E$, il existe un entier n_0 tel que $Q_n(x,y) > 0$ pour tout $n \geq n_0$.*

Preuve. (i) Soient $x, y \in E$. Puisque la chaîne est irréductible, il existe deux entiers n_1 et n_2 tels que $Q_{n_1}(x,y) > 0$ et $Q_{n_2}(y,x) > 0$. Mais alors, si $n \in L_x$, on a $n_1 + n + n_2 \in L_y$, ce qui entraîne que $L_x - L_x \subset L_y - L_y$ et donc $d(y)$ divise $d(x)$. Par symétrie on a $d(y) = d(x)$.
(ii) Clairement, il suffit de traiter le cas où $y = x$. Puisque $d(x) = 1$, on peut trouver deux entiers $n_1, m_1 \geq 0$ tels que $1 = n_1 - m_1$ et

$$Q_{n_1}(x,x) > 0,\ Q_{m_1}(x,x) > 0.$$

Si $m_1 = 0$, donc $n_1 = 1$ le résultat est évident avec $n_0 = 0$. Si $m_1 \geq 1$, alors, pour tout $j \in \{0, 1, \ldots, m_1 - 1\}$, on a

$$Q_{m_1^2+j}(x,x) = Q_{jn_1+(m_1-j)m_1}(x,x) > 0.$$

Il en découle que, si $n_0 = m_1^2$ on a pour tout entier $j \geq 0$,

$$Q_{n_0+j}(x,x) > 0.$$

Théorème 13.6.6 *Supposons la chaîne irréductible, récurrente positive et apériodique. Alors, si μ désigne l'unique probabilité invariante, on a pour tout $x \in E$,*

$$\sum_{y \in E} |\mathbb{P}_x(X_n = y) - \mu(y)| \underset{n\to\infty}{\longrightarrow} 0.$$

Preuve. La formule
$$\overline{Q}((x_1,x_2),(y_1,y_2)) = Q(x_1,y_1)Q(x_2,y_2)$$
définit une matrice stochastique sur le $E \times E$. On note $((X_n^1, X_n^2)_{n\in\mathbb{N}}, (\overline{P}_{(x_1,x_2)})_{(x_1,x_2)\in E\times E})$ la chaîne de Markov canonique associée.

Remarquons que Q est irréductible : si $(x_1,x_2),(y_1,y_2) \in E \times E$, la proposition 13.6.5(ii) permet de trouver deux entiers n_1 et n_2 tels que $Q_n(x_1,y_1) > 0$ pour tout $n \geq n_1$, et $Q_n(x_2,y_2) > 0$ pour tout $n \geq n_2$. Si $n \geq n_1 \vee n_2$, on a par définition $\overline{Q}_n((x_1,x_2),(y_1,y_2)) > 0$.

De plus la mesure produit $\mu\otimes\mu$ est invariante pour $\overline{Q}$:

$$\begin{aligned}\sum_{(x_1,x_2)\in E\times E}\mu(x_1)\mu(x_2)Q(x_1,y_1)Q(x_2,y_2) &= \sum_{x_1\in E}\mu(x_1)Q(x_1,y_1)\sum_{x_2\in E}\mu(x_2)Q(x_2,y_2)\\ &= \mu(y_1)\mu(y_2).\end{aligned}$$

La proposition 13.5.5 permet de conclure que la chaîne (X_n^1,X_n^2) est récurrente positive. Observons maintenant que

$$\mathbb{P}_x(X_n=y)-\mu(y)=\overline{\mathbb{P}}_{\mu\otimes\delta_x}(X_n^2=y)-\overline{\mathbb{P}}_{\mu\otimes\delta_x}(X_n^1=y)=\overline{\mathbb{E}}_{\mu\otimes\delta_x}[\mathbf{1}_{\{X_n^2=y\}}-\mathbf{1}_{\{X_n^1=y\}}].$$

Introduisons le temps d'arrêt $T=\inf\{n\geq 0: X_n^1=X_n^2\}$. Alors, l'égalité précédente montre que

$$\begin{aligned}\mathbb{P}_x(X_n=y)-\mu(y) &= \overline{\mathbb{E}}_{\mu\otimes\delta_x}[\mathbf{1}_{\{T>n\}}(\mathbf{1}_{\{X_n^2=y\}}-\mathbf{1}_{\{X_n^1=y\}})]\\ &\quad+\sum_{k=0}^{n}\sum_{z\in E}\overline{\mathbb{E}}_{\mu\otimes\delta_x}[\mathbf{1}_{\{T=k,X_k^1=X_k^2=z\}}(\mathbf{1}_{\{X_n^2=y\}}-\mathbf{1}_{\{X_n^1=y\}})].\end{aligned}\tag{13.5}$$

Mais, pour tout $k\in\{0,1,\ldots,n\}$ et tout $z\in E$, la propriété de Markov entraîne que

$$\begin{aligned}\overline{\mathbb{E}}_{\mu\otimes\delta_x}[\mathbf{1}_{\{T=k,X_k^1=X_k^2=z\}}\mathbf{1}_{\{X_n^2=y\}}] &= \overline{\mathbb{E}}_{\mu\otimes\delta_x}[\mathbf{1}_{\{T=k,X_k^1=X_k^2=z\}}]\,Q_{n-k}(z,y)\\ &= \overline{\mathbb{E}}_{\mu\otimes\delta_x}[\mathbf{1}_{\{T=k,X_k^1=X_k^2=z\}}\mathbf{1}_{\{X_n^1=y\}}],\end{aligned}$$

et donc le deuxième terme de la somme dans (13.5) est nul. On obtient ainsi que

$$\begin{aligned}\sum_{y\in E}|\mathbb{P}_x(X_n=y)-\mu(y)| &= \sum_{y\in E}|\overline{\mathbb{E}}_{\mu\otimes\delta_x}[\mathbf{1}_{\{T>n\}}(\mathbf{1}_{\{X_n^2=y\}}-\mathbf{1}_{\{X_n^1=y\}})]|\\ &\leq \sum_{y\in E}\overline{\mathbb{E}}_{\mu\otimes\delta_x}[\mathbf{1}_{\{T>n\}}(\mathbf{1}_{\{X_n^2=y\}}+\mathbf{1}_{\{X_n^1=y\}})]\\ &= 2\,\overline{\mathbb{P}}_{\mu\otimes\delta_x}(T>n),\end{aligned}$$

qui tend vers 0 quand $n\to\infty$, grâce à la récurrence de la chaîne (X_n^1,X_n^2). □

13.7 Martingales et chaînes de Markov

On considère toujours la chaîne de Markov canonique de matrice de transition Q.

Définition 13.7.1 *Une fonction $f:E\longrightarrow\mathbb{R}_+$ est dite harmonique (resp. surharmonique) si on a pour tout $x\in E$,*

$$f(x)=Qf(x)\qquad(\textit{resp. } f(x)\geq Qf(x)).$$

Plus généralement, si $F\subset E$, on dit que f est harmonique sur F (resp. surharmonique sur F) si la propriété $f(x)=Qf(x)$ (resp. $f(x)\geq Qf(x)$) est vraie pour $x\in F$.

Remarque. On pourrait considérer plus généralement des fonctions harmoniques ou surharmoniques de signe quelconque.

Proposition 13.7.1 (i) *La fonction f est harmonique (resp. surharmonique) ssi, pour tout $x \in E$, le processus $(f(X_n))_{n\in\mathbb{N}}$ est une martingale (resp. une surmartingale) sous $\mathbb{P}_x$, relativement à la filtration $(\mathcal{F}_n)$.*
(ii) *Soit $F \subset E$ et $G = E\backslash F$. On note T_G le temps d'arrêt*

$$T_G = \inf\{n \geq 0 : X_n \in G\}.$$

Alors si f est harmonique (resp. surharmonique) sur F, le processus $(f(X_{n\wedge T_G}))_{n\in\mathbb{N}}$ est une martingale (resp. une surmartingale) sous $\mathbb{P}_x$, pour tout $x \in F$.

Preuve. (i) Supposons d'abord f harmonique. Alors, d'après la proposition 13.1.2(i),

$$\mathbb{E}_x[f(X_{n+1}) \mid \mathcal{F}_n] = Qf(X_n) = f(X_n)$$

et en conséquence $\mathbb{E}_x[f(X_n)] = \mathbb{E}_x[f(X_0)] = f(x)$, donc $f(X_n) \in L^1$.

Inversement, supposons que $f(X_n)$ est une martingale sour $\mathbb{P}_x$. Il vient immédiatement que

$$f(x) = \mathbb{E}_x[f(X_0)] = \mathbb{E}_x[f(X_1)] = Qf(x).$$

Le cas d'une fonction surharmonique est traité de la même façon.

(ii) Traitons le cas d'une fonction harmonique. On écrit pour $x \in F$

$$\begin{aligned}
\mathbb{E}_x[f(X_{(n+1)\wedge T_G}) \mid \mathcal{F}_n] &= \mathbb{E}_x[f(X_{n+1})\,\mathbf{1}_{\{T_G>n\}} \mid \mathcal{F}_n] + \mathbb{E}_x[f(X_{T_G})\,\mathbf{1}_{\{T_G\leq n\}} \mid \mathcal{F}_n] \\
&= \mathbf{1}_{\{T_G>n\}}\,\mathbb{E}_x[f(X_{n+1}) \mid \mathcal{F}_n] + f(X_{T_G})\,\mathbf{1}_{\{T_G\leq n\}} \\
&= \mathbf{1}_{\{T_G>n\}}\,Qf(X_n) + f(X_{T_G})\,\mathbf{1}_{\{T_G\leq n\}} \\
&= \mathbf{1}_{\{T_G>n\}}\,f(X_n) + f(X_{T_G})\,\mathbf{1}_{\{T_G\leq n\}} \\
&= f(X_{n\wedge T_G})
\end{aligned}$$

On a utilisé le fait que $f(X_{T_G})\,\mathbf{1}_{\{T_G\leq n\}} = f(X_{T_G\wedge n})\,\mathbf{1}_{\{T_G\leq n\}}$ est $\mathcal{F}_n$-mesurable. □

Théorème 13.7.2 *Soit F un sous-ensemble non vide de E et $G = E\backslash F$. Soit $g : G \longrightarrow \mathbb{R}_+$ une fonction bornée.*

(i) *La fonction*

$$h(x) = \mathbb{E}_x[g(X_{T_G})\,\mathbf{1}_{\{T_G<\infty\}}], \qquad x \in E$$

est harmonique sur F.

(ii) *Supposons $T_G < \infty$, $\mathbb{P}_x$ p.s. pour tout $x \in F$. Alors la fonction h est l'unique fonction bornée sur E qui*

- *est harmonique sur F,*
- *coïncide avec g sur G.*

Preuve. (i) On remarque que si $x \in F$ on a $\mathbb{P}_x$ p.s.

$$g(X_{T_G})\,\mathbf{1}_{\{T_G<\infty\}} = g(X_{T_G}\circ\theta_1)\,\mathbf{1}_{\{T_G\circ\theta_1<\infty\}}.$$

Autrement dit, si $U(\omega) = g(X_{T_G}(\omega))\,\mathbf{1}_{\{T_G(\omega)<\infty\}}$, on a $U = U\circ\theta_1$, $\mathbb{P}_x$ p.s. Donc, pour $x \in F$, d'après le théorème 13.3.4,

$$h(x) = \mathbb{E}_x[U] = \mathbb{E}_x[U\circ\theta_1] = \mathbb{E}_x[\mathbb{E}_{X_1}[U]] = \mathbb{E}_x[h(X_1)] = Qh(x),$$

ce qui montre que h est harmonique sur F.

(ii) Il est trivial que $h(x) = g(x)$ si $x \in G$. Soit h' une autre fonction harmonique sur F, bornée sur E et coïncidant avec g sur G. Si $x \in F$, d'après la proposition 13.7.1, $Y_n = h'(X_{n\wedge T_G})$ est une martingale sous $\mathbb{P}_x$. Cette martingale est bornée, donc uniformément intégrable, et converge $\mathbb{P}_x$ p.s. vers $h'(X_{T_G}) = g(X_{T_G})$. D'après les résultats du chapitre 12, on a donc

$$h'(x) = \mathbb{E}_x[Y_0] = \mathbb{E}_x[Y_\infty] = \mathbb{E}_x[g(X_{T_G})] = h(x).$$

Exemple. *Problème de Dirichlet discret.* Soit F une partie finie de $\mathbb{Z}^d$. La frontière de F est

$$\partial F = \{y \in \mathbb{Z}^d\backslash F : \exists x \in F,\ |y-x| = 1\}.$$

On note $\overline{F} = F \cup \partial F$.

Une fonction h définie sur $\overline{F}$ est dite harmonique (au sens discret) sur F si pour tout $x \in F$, $h(x)$ est égal à la moyenne des valeurs de h sur les $2d$ plus proches voisins de x. On retrouve la notion précédente en prenant comme chaîne de Markov la marche aléatoire simple sur $\mathbb{Z}^d$: $Q(x, x\pm e_j) = \frac{1}{2d}$ pour $j = 1,\ldots,d$, où $(e_1,\ldots,e_d)$ est la base canonique.

Alors, le théorème précédent conduit au résultat suivant : pour toute fonction (positive) g définie sur ∂F, la seule fonction $h : \overline{F} \longrightarrow \mathbb{R}_+$ telle que :

- h est harmonique sur F,
- $h(y) = g(y)$, $\forall y \in \partial F$,

est donnée par

$$h(x) = \mathbb{E}_x[g(X_{T_{\partial F}})]\,, \qquad x \in F,$$

où

$$T_{\partial F} = \inf\{n \geq 0 : X_n \in \partial F\}.$$

Noter que pour appliquer le théorème 13.7.2, on a a priori besoin de définir g sur $\mathbb{Z}^d\backslash F$ et non pas seulement sur ∂F : cependant le choix des valeurs de g sur $\mathbb{Z}^d\backslash\overline{F}$ n'influe pas sur les valeurs de h sur F.