**Faculté des Sciences et Technologies, Université de Lorraine.**
**Ecole Supérieure des Sciences et Technologies d'Hammam-Sousse.**

**Master 2 mathématiques et applications**
**UE 908 calcul stochastique.**
**Année 2016/2017**

P. Vallois : `pierre.vallois@univ-lorraine.fr`
N. Champagnat : `Nicolas.Champagnat@inria.fr`

# CHAÎNES DE MARKOV

Ce cours est basé sur le livre de P. Vallois, *Modélisations et simultations stochatiques.* Ellipses 2007.

## Table des matières

Les chaînes de Markov jouent un rôle important dans la modélisation de phénomènes d'évolution temporels avec mémoire. Plus précisément une chaîne de Markov est un processus aléatoire qui ne retient du passé que l'instant présent. Il est clair que tous les phénomènes ne rentrent pas dans ce cadre restrictif mais il existe une large classe d'exemples pour lesquels la propriété précédente est réalisée.

# 1 Quelques rappels concernant l'espérance conditionnelle

La notion de chaîne de Markov fait appel de manière essentielle à la notion de loi conditionnelle. Nous rappelons brièvement cette notion dans le cadre des variables aléatoires à valeurs dans un ensemble fini ou dénombrable.
Dans ce qui suit, $U$ désigne une v.a. fixée, à valeurs dans un ensemble $\mathbb{U} = U(\Omega)$ fini ou dénombrable. On s'intéresse aux v.a. $Y$ définies sur le même espace de probabilité que $U$, à valeurs dans un ensemble $\mathbb{Y} = \{y_i ; i \in I\}$ ayant un nombre fini ou dénombrable d'éléments.
**1)** La loi de $Y$ est la probabilité $\mathbb{P}_Y$ sur $\mathbb{Y}$ :

$$\mathbb{P}_Y = \sum_{i\in I} \mathbb{P}(Y = y_i)\delta_{y_i},$$

où $\delta_{y_i}$ dénote la mesure de Dirac au point $y_i$.
Il est commode d'identifier toute probabilité $\mu$ sur $\mathbb{Y}$ à la famille $\big(\mu(y_i)\big)_{i\in I}$ ; en particulier : $\mathbb{P}_Y = \big(\mathbb{P}(Y = y_i)\big)_{i\in I}$.
**2)** La loi de $Y$ sachant $U$ est une famille de probabilités $\Big\{\big(\theta(u, y_i)\big)_{i\in I};\ u \in \mathbb{U}\Big\}$ sur $\mathbb{Y}$ vérifiant :

1. pour tout $u \in \mathbb{U}$, $\big(\theta(u, y_i)\big)_{i\in I}$ est une probabilité sur $\mathbb{Y}$ ;
2. pour tout $u \in \mathbb{U}$ tel que $\mathbb{P}(U = u) > 0$, on a :

$$\theta(u, y_i) = \mathbb{P}(Y = y_i | U = u) = \frac{\mathbb{P}(Y = y_i, U = u)}{\mathbb{P}(U = u)}, \quad \text{pour tout} \ \ y_i \in \mathbb{Y};$$

Notons qu'il n'est rien demandé quant à la valeur de $\big(\theta(u, y_i)\big)_{i\in I}$ lorsque $\mathbb{P}(U = u) = 0$. Il commode de dire que la **loi conditionelle de $Y$ sachant** $U$ est $\theta(U, \cdot)$.
On déduit aisément de la définition la relation clé :

$$\mathbb{P}(Y = y, U = u) = \mathbb{P}(U = u)\theta(u, y), \quad \text{pour tout} \ \ u \in \mathbb{U} \text{ et } \ y \in \mathbb{Y}. \tag{1.1}$$

**3)** Soit $f : \mathbb{Y} \to \mathbb{R}$ positive ou bornée. L'**espérance conditionnelle** de $f(Y)$ sachant $U$ est la v.a. :

$$\mathbb{E}\big[f(Y)|U\big] := \int_{\mathbb{Y}} f(y)\theta(U, dy) = \sum_{i\in I} f(y_i)\theta(U, y_i), \tag{1.2}$$

c'est-à-dire la v.a. (variable aléatoire) valant $\mathbb{E}[f(Y) \mid U = u_i]$ sur l'événement $\{U = u_i\}$, pour tout $i \in I$.

Il est usuel de noter :

$$\mathbb{E}\big[f(Y)|U = u\big] = M_1(f, u), \quad \text{avec} \ \ M_1(f, u) = \sum_{i\in I} f(y_i)\theta(u, y_i), \quad \text{et} \ \ u \in \mathbb{U}.$$

On vérifie facilement :

$$\mathbb{E}\big[h(U)f(Y)|U\big] = h(U)\,\mathbb{E}\big[f(Y)|U\big], \tag{1.3}$$

où $f$ et $h$ sont positives ou bornées. En particulier,

$$\mathbb{E}\big[h(U)|U\big] = h(U). \tag{1.4}$$

**4)** Soient $f : \mathbb{Y} \to \mathbb{R}$, et $g : \mathbb{U} \to \mathbb{R}$ deux fonctions toutes les deux positives ou bornées. Alors :

$$\mathbb{E}\big[f(Y)g(U)\big] = \mathbb{E}\Big[\mathbb{E}\big[f(Y)|U\big]\ g(U)\Big] = \mathbb{E}\Big[M_1(f,U)g(U)\Big]. \tag{1.5}$$

**5)** Énonçons la propriété de **double projection** dans un cadre simplifié mais qui est suffisant pour traiter le cas des chaînes de Markov et également de nombreuses autres situations. Soient $U_1$ et $U_2$ deux v.a. à valeurs dans $\mathbb{U}$ et $f : \mathbb{Y} \to \mathbb{R}$ une application positive ou bornée. Alors :

$$\mathbb{E}\big[f(Y)|U_1\big] = \mathbb{E}\Big\{\mathbb{E}\big[f(Y)|U\big]\ \Big| U_1\Big\}, \quad \text{où} \ \ U = (U_1, U_2). \tag{1.6}$$

**6)** Lorsque $Y$ et $U$ sont deux v.a. indépendantes, on a :

$$\mathbb{E}\big[F(U,Y)|U\big] = F_1(U), \quad \text{avec} \ \ F_1(u) = \mathbb{E}[F(u,Y)] \quad (F \geq 0 \text{ ou bornée}). \tag{1.7}$$

En particulier,

$$\mathbb{E}[F(U)G(Y) \mid U] = F(U)\mathbb{E}[G(Y)]$$

et

$$\mathbb{E}[G(Y) \mid U] = \mathbb{E}[G(Y)].$$

# 2 Définitions, premières propriétés et premiers exemples

## 2.1 Définition d'une chaîne de Markov

Soit $E$ un ensemble fini ou dénombrable. $E$ est appelé l'**espace d'état**. En pratique, lorsque $E$ est dénombrable, $E = \mathbb{N}, \mathbb{Z}, \mathbb{N}^d$, ou $\mathbb{Z}^d$, ou bien un sous-ensemble d'un de ces ensembles. L'ensemble $E$ étant fixé, on identifie toutes les probabilités $\mu = \sum_{x \in E} \mu(x)\delta_x$ à la famille $\big(\mu(x); x \in E\big)$.

**Définition 2.1** 1. *$(\pi(x,y); x, y \in E)$ est une* **probabilité de transition** *si pour tout $x$ de $E$, $\pi(x,\cdot)$ est une probabilité sur $E$ :*

$$\pi(x,y) \geq 0, \text{ pour tout } y \in E \text{ et } \sum_{y \in E} \pi(x,y) = 1.$$

*On note :*

$$\pi(x,f) := \int_E f(y)\pi(x,dy) = \sum_{y \in E} f(y)\pi(x,y), \quad \text{pour tout } x \in E,$$

*ou $f : E \to \mathbb{R}$ désigne une fonction positive ou bornée.*

2. *Soit $\mu$ une probabilité sur $E$. Une suite de v.a. (ou* processus*) $(X_n; n \geq 0)$ définie sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ et à valeurs dans $E$, est une* **chaîne de Markov** *de* **loi initiale** *$\mu$ et de probabilité de transition $\pi$ si :*
   (a) *$\mathbb{P}(X_0 = x) = \mu(x)$, pour tout $x \in E$.*
   (b) *La loi conditionnelle de $X_{n+1}$ sachant $X_0, X_1, \cdots, X_n$ est $\pi(X_n, \cdot)$, pour tout $n \geq 0$.*

   *Lorsque $\mu = \delta_{x_0}$ où $x_0 \in E$, alors $\mathbb{P}(X_0 = x_0) = 1$. On dit alors que la chaîne est* **issue** *de $x_0$.*

Lorsque $E$ est un ensemble fini à $r$ éléments, $\pi$ est aussi appelée **matrice de transition**. Supposons $E = \{1, 2, \cdots, r\}$, alors la famille de probabilités de transition $\big(\pi(i,j); i, j \in E\big)$ s'identifie avec la matrice carré d'ordre $r$ $\big(\pi(i,j)\big)_{1 \leq i,j \leq r}$, dont tous les éléments sont positifs ou nuls, et la somme de tous les éléments d'une ligne vaut 1 pour chaque ligne.

**Remarque 2.2** 1. *La famille de v.a. $(X_n; n \geq 0)$ est indexée par $\mathbb{N}$, $\mathbb{N}$ représente le temps.*

2. *Une chaîne de Markov modélise le mouvement "aléatoire" d'une particule à valeurs dans $E$ suivant les règles suivantes :*
   (a) *à l'instant $0$, la position de la particule est aléatoire avec la loi $\mu$,*
   (b) *Si $X_0 = x_0$, à l'instant $1$, la particule se déplace aléatoirement suivant la probabilité $\pi(x_0, \cdot)$.*
   (c) *Plus généralement, si $X_n$ occupe le site $x_n$ au temps $n$, alors $X_{n+1}$ visite $x_{n+1}$ avec probabilité $\pi(x_n, x_{n+1})$.*

   *La position de la particule à l'instant $n+1$ est aléatoire, mais ne dépend que de sa position à l'instant $n$. En d'autres termes, la particule ne retient du passé que sa position à l'instant précédent.*

3. *La propriété 2. (b) de la définition précédente est équivalente à :*

$$\mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n, \cdots, X_0 = x_0) = \pi(x_n, x_{n+1}), \tag{2.8}$$

*pour tout $n$ et $x_0, x_1, \cdots, x_n, x_{n+1}$ de $E$.*

4. *Il est possible de définir une chaîne de Markov à valeurs dans un ensemble $E$ qui n'est pas dénombrable, par exemple $E = \mathbb{R}$. Nous nous sommes restreint au cas discret car le cadre théorique est facile à développer et tous les exemples que nous considérerons dans la suite restent dans le cadre discret.*

Donnons une version fonctionnelle de (2.8).

**Proposition 2.3** *Soit $(X_n; n \geq 0)$ une suite de v.a. à valeurs dans $E$, $(\pi(x,y); x, y \in E)$ une probabilité de transition, et $\mu$ une probabilité sur $E$. Alors les trois propriétés suivantes sont équivalentes :*

1. *$(X_n; n \geq 0)$ est une chaîne de Markov de loi initiale $\mu$ et de probabilité de transition $\pi$.*
2. *Pour tous $n \geq 1$ et $x_0, x_1, \cdots, x_n$ de $E$, $\mathbb{P}(X_0 = x_0) = \mu(x_0)$ et*

$$\mathbb{P}(X_0 = x_0, \cdots, X_n = x_n) = \mu(x_0)\pi(x_0, x_1) \cdots \pi(x_{n-1}, x_n). \tag{2.9}$$

3. *Pour toutes fonctions $f : E \to \mathbb{R}$ et $g : E^n \to \mathbb{R}$ positives ou bornées, on a :*

$$\mathbb{E}[f(X_0)] = \int_E f(x)\mu(dx) = \sum_{x \in E} f(x)\mu(x) = \mu(f), \tag{2.10}$$

*et*

$$\mathbb{E}\big[f(X_{n+1})g(X_0, \cdots, X_n)\big] = \mathbb{E}\big[\pi(X_n, f)g(X_0, \cdots, X_n)\big]. \tag{2.11}$$

**Preuve.** On montre l'équivalence entre les points 1. et 2. en raisonnant par récurrence sur $n$, et en utilisant l'identité (1.1).
En ce qui concerne 1. $\Leftrightarrow$ 3., on calcule le membre de gauche de (2.11) en utilisant les propiétés (1.2) et (1.5) de l'espérance conditionnelle (avec $Y = X_{n+1}$ et $U = (X_0, \cdots, X_n)$) ; il vient :

$$\begin{aligned}
\mathbb{E}\big[f(X_{n+1})g(X_0, \cdots, X_n)\big] &= \mathbb{E}\Big[\mathbb{E}\big[f(X_{n+1}) | X_0, \cdots, X_n\big]\ g(X_0, \cdots, X_n)\Big] \\
&= \mathbb{E}\big[\pi(X_n, f)g(X_0, \cdots, X_n)\big].
\end{aligned}$$

■

**Remarque 2.4** 1. *La version fonctionnelle de (2.9) est la suivante : pour tout $n \geq 1$,*

$$\begin{aligned}
&\mathbb{E}\big[f_0(X_0) \cdots f_n(X_n)\big] \\
&\qquad = \sum_{x_0, \cdots, x_n} f_0(x_0) \times \cdots \times f_n(x_n)\mu(x_0)\pi(x_0, x_1) \times \cdots \times \pi(x_{n-1}, x_n)
\end{aligned} \tag{2.12}$$

*pour toutes fonctions $f_0, f_1, ..., f_n$ définies sur $E$ et à valeurs dans $\mathbb{R}$, positives ou bornées. On peut aussi écrire*

$$\begin{aligned}
\mathbb{E}\big[f_0(X_0) \cdots f_n(X_n)\big] &= \int_E f_0(x_0)\mu(dx_0) \int_E \pi(x_0, dx_1) f_1(x_1) \\
&\cdots \times \int_E \pi(x_{n-2}, dx_{n-1}) f_{n-1}(x_{n-1}) \int_E \pi(x_{n-1}, dx_n) f_n(x_n).
\end{aligned}$$

*2. Le point 2. de la Proposition 2.3 signifie que la loi de $(X_0, X_1, \cdots, X_n)$ est entièrement déterminée par $\pi$ et $\mu$.*

**Exemple 2.5** *Donnons l'exemple le plus simple, celui où $E = \{1, 2\}$ et*

$$\pi = \begin{pmatrix} \theta_1 & 1-\theta_1 \\ 1-\theta_2 & \theta_2 \end{pmatrix}$$

*avec $0 \leq \theta_1, \theta_2 \leq 1$.*
*Ainsi, lorsque la particule est en 1 (resp. 2), à l'instant suivant,*

1. *elle reste en 1 (resp. 2) avec probabilité $\theta_1$ (resp. $\theta_2$)*
2. *elle va en 2 (resp. 1) avec probabilité $1-\theta_1$ (resp. $1-\theta_2$).*

*Les trois cas "extrêmes" correspondent à $\theta_1, \theta_2 = 0$ ou 1.*

1. $\pi = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$*, alors partant de 1, le processus va ensuite en 2 et y reste. L'état 2 est un état "absorbant". Le cas* $\pi = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$ *est analogue.*
2. $\pi = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$*, alors le comportement de la chaîne est purement déterministe : elle va alternativement de 1 à 2.*
3. $\pi = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$*, dans ce cas la chaîne est constante : $X_n = X_0$.*

**Exercice 2.6** *Soient $(X_n; n \geq 0)$ une suite de v.a. à valeurs dans $E$, indépendantes et de même loi $\nu$.*

1. *Montrer que $(X_n; n \geq 0)$ est une chaîne de Markov de probabilité de transition $\pi(x, y) = \nu(y)$ qui ne dépend pas $x$.*
2. *On pose : $Y_n := (X_n, X_{n+1})$ pour tout $n \geq 0$. Montrer que $(Y_n)_{n\geq 0}$ est une chaine de Markov.*

**Exercice 2.7** *Soit $(\xi_n)_{n\geq 0}$ une suite de variables aléatoires (v.a.), indépendantes à valeurs dans $E := \{1, 2, \cdots, , N\}$ et de loi uniforme sur $E$. On pose : $X_n := \mathrm{Card}\big(\{\xi_1, \cdots, \xi_n\}\big)$. Montrer que $(X_n)_{n\geq 0}$ est une chaine de Markov.*

**Exercice 2.8** *Soit $(X_n, n \geq 0)$ une chaine de Markov à valeurs dans $E$, de loi initiale $\mu$ et de probabilité de transition $\pi$. On considère $f : E \to F$.*

1. *Montrer que $\big((X_n, f(X_n)), n \geq 0\big)$ une chaine de Markov.*
2. *On suppose que $f$ est injective. Montrer que $\big(f(X_n), n \geq 0\big)$ une chaine de Markov.*
3. *On suppose $E := \{-1, 0, 1\}$. Soit $\pi$ la matrice de transition de $(X_n)$. On suppose $\pi(i, j) > 0$ pour tout $i$ et $j$ appartenant à $E$. A quelle condition le processus $(|X_n|, n \geq 0)$ est-il une chaine de Markov ?*

## 2.2 La propriété faible de Markov

Il est commode de considérer qu'une chaîne de Markov est un processus tel que, pour toute date $n \geq 1$, la loi du futur du processus après $n$ sachant le passé avant $n$ ne dépend que de l'état du processus à l'instant présent $n$. Cette propriété d'absence de mémoire, appelée *propriété (faible) de Markov* (on verra la version forte en section 3), est très utile pour développer son intuition sur les chaînes de Markov, mais est délicate à définir et caractériser mathématiquement. L'objectif de cette sous-section est d'en donner plusieurs formes équivalentes.

On considère un processus (ou suite de v.a.) $(X_n; n \geq 0)$ à valeurs dans $E$, une probabilité $\mu$ sur $E$ et une famille de probabilités de transition $\pi = (\pi(x,y), x, y \in E)$ sur $E$. Pour tout $n \geq 0$, on note $\mathcal{F}_n = \sigma(X_0, X_1, \ldots, X_n)$ la tribu engendrée par le passé de la suite avant $n$. On rappelle qu'une v.a. est $\mathcal{F}_n$-mesurable si et seulement si elle s'écrit comme une fonction mesurable du vecteur $(X_0, X_1, \ldots, X_n)$.

**Proposition 2.9** . *Le processus $(X_n; n \geq 0)$ est une chaîne de Markov de loi initiale $\mu$ et de probabilité de transition $\pi$ ssi il satisfait les trois propriétés suivantes*

1. *$X_0$ a pour loi $\mu$ ;*
2. *il satisfait la* ***propriété faible de Markov*** *: pour tout $n \geq 1$ et pour tout $A \subset E$,*

$$\mathbb{P}(X_{n+1} \in A \mid \mathcal{F}_n) = \mathbb{P}(X_{n+1} \in A \mid X_0, \ldots, X_n) = \mathbb{P}(X_{n+1} \in A \mid X_n); \tag{2.13}$$

3. ***(homogénéité)*** *Le membre de droite de* (2.13) *est donné par*

$$\mathbb{P}(X_{n+1} \in A \mid X_n) = \pi(X_n, A), \quad \forall n \geq 1, \ \forall A \subset E \tag{2.14}$$

*(et en particulier il dépend de $n$ seulement à travers $X_n$).*

**Preuve.** Supposons que $(X_n, n \geq 0)$ est une chaîne de Markov. Le point 1. est trivialement vérifié. D'après (2.11),

$$\mathbb{P}(X_{n+1} \in A \mid \mathcal{F}_n) = \pi(X_n, A). \tag{2.15}$$

En utilisant la propriété de double-projection (1.6) et la relation (1.4), on en déduit que

$$\mathbb{P}(X_{n+1} \in A \mid X_n) = \mathbb{E}\left[\mathbb{P}(X_{n+1} \in A \mid X_0, \ldots, X_n) \mid X_n\right] = \mathbb{E}[\pi(X_n, A) \mid X_n] = \pi(X_n, A).$$

Ceci implique le point 2., et le point 3. découle de (2.15).

Réciproquement, la propriété (a) dans la définition d'une chaîne de Markov est exactement le point 1., et la propriété (b) découle directement des points 2. et 3. ■

Il y a beaucoup de formulations équivalentes de la propriété de Markov. L'une ou l'autre peut être utile suivant les applications. Nous en donnons quelques-unes ci-dessous.

**Proposition 2.10** *La propriété faible de Markov : pour tout $n \geq 1$ et pour tout $A \subset E$,*

$$\mathbb{P}(X_{n+1} \in A \mid \mathcal{F}_n) = \mathbb{P}(X_{n+1} \in A \mid X_n) \tag{2.16}$$

*est équivalente à chacun des points ci-dessous :*

1. *Pour tout $n \geq 1$, $x_0, \dots, x_n \in E$ et $A \subset E$,*

$$\mathbb{P}(X_{n+1} \in A \mid X_0 = x_0, \dots, X_n = x_n) = \mathbb{P}(X_{n+1} \in A \mid X_n = x_n).$$

2. *Pour tout $n \geq 1$ et $f : E \to \mathbb{R}$ bornée,*

$$\mathbb{E}(f(X_{n+1}) \mid \mathcal{F}_n) = \mathbb{E}(f(X_{n+1}) \mid X_n).$$

3. *Pour tout $n, k \geq 1$ et $x_1, \dots, x_k \in E$,*

$$\mathbb{P}(X_{n+1} = x_1, \dots, X_{n+k} = x_k \mid \mathcal{F}_n) = \mathbb{P}(X_{n+1} = x_1, \dots, X_{n+k} = x_k \mid X_n).$$

4. *Soit $M$ l'ensemble des fonctions de $\mathbb{N}$ dans $E$. On définit sur $M$ la tribu $\mathcal{M}$ engendrée par tous les événements de la forme $\{f \in M : f(0) = x_0, f(1) = x_1, \dots, f(n) = x_n\}$, pour tout $n \geq 0$ et $x_0, \dots, x_n \in E$. Alors, pour tout $n \geq 1$ et $\Gamma \in \mathcal{M}$,*

$$\mathbb{P}((X_{n+k}, k \geq 0) \in \Gamma \mid \mathcal{F}_n) = \mathbb{P}((X_{n+k}, k \geq 0) \in \Gamma \mid X_n). \tag{2.17}$$

5. *Pour tout $n \geq 1$ et toute fonction $F : M \to \mathbb{R}$ bornée et mesurable par rapport à la tribu $\mathcal{M}$,*

$$\mathbb{E}[F(X_{n+k}, k \geq 0) \mid \mathcal{F}_n] = \mathbb{E}[F(X_{n+k}, k \geq 0) \mid X_n].$$

6. ***(indépendance du futur et du passé conditionnellement au présent)*** *Pour tout $n \geq 1$ et $x_0, \dots, x_{n+1}$,*

$$\begin{aligned}&\mathbb{P}(X_0 = x_0, \dots, X_n = x_n, X_{n+1} = x_{n+1} \mid X_n = x_n)\\ &\quad = \mathbb{P}(X_0 = x_0, \dots, X_{n-1} = x_{n-1} \mid X_n = x_n)\mathbb{P}(X_{n+1} = x_{n+1} \mid X_n = x_n).\end{aligned} \tag{2.18}$$

**Preuve.** L'équivalence de la propriété de Markov avec le point 1. est triviale par la définition des probabilités conditionnelles et avec le point 2. est habituelle en probabilités. De même, les points 4. et 5. sont équivalents. Le fait que la propriété de Markov implique le point 3. découle (par récurrence) de la relation

$$\begin{aligned}\mathbb{P}(X_{n+k} = x_k, &\dots, X_{n+1} = x_1 \mid X_n = y_n, \dots, X_0 = y_0)\\ &= \mathbb{P}(X_{n+k} = x_k \mid X_{n+k-1} = x_{k-1}, \dots, X_{n+1} = x_1, X_n = y_n, \dots, X_0 = y_0)\\ &\quad \times \mathbb{P}(X_{n+k-1} = x_{k-1}, \dots, X_{n+1} = x_1 \mid X_n = y_n, \dots, X_0 = y_0)\\ &= \pi(x_{k-1}, x_k)\mathbb{P}(X_{n+k-1} = x_{k-1}, \dots, X_{n+1} = x_1 \mid X_n = y_n, \dots, X_0 = y_0).\end{aligned}$$

On en déduit que

$$\mathbb{P}(X_{n+k} = x_k, \dots, X_{n+1} = x_1 \mid X_n = y_n, \dots, X_0 = y_0) = \pi(x_{k-1}, x_k) \dots \pi(x_1, x_2)\pi(y_n, x_1).$$

Par définition de l'espérance conditionnelle, ceci signifie que

$$\mathbb{P}(X_{n+k} = x_k, \dots, X_{n+1} = x_1 \mid \mathcal{F}_n) = \pi(X_n, x_1)\pi(x_1, x_2) \dots \pi(x_{k-1}, x_k). \tag{2.19}$$

En prenant l'espérance sachant $X_n$ de chaque côté et en utilisant la propriété de double-projection (1.6), on trouve

$$\mathbb{P}(X_{n+k} = x_k, \dots, X_{n+1} = x_1 \mid X_n) = \pi(X_n, x_1)\pi(x_1, x_2) \dots \pi(x_{k-1}, x_k),$$

ce qui achève la preuve du point 3.

D'autre part, il découle du point 3. que (2.17) est satisfait pour des événements $\Gamma$ de la forme $\{f \in M : f(1) = x_1, f(2) = x_2, \dots, f(k) = x_k\}$. Puisque, par définition, ces événements engendrent $\mathcal{M}$, on en déduit le point 4. par le théorème des classes monotones. Enfin, le point 4. implique évidemment la propriété de Markov (il suffit de prendre $\Gamma = \{f \in M : f(0) \in A\}$).

Il reste à démontrer l'équivalence avec le point 6. Si la propriété de Markov est satisfaite, en utilisant (1.3) et la propriété de Markov,

$$\begin{aligned}\mathbb{P}(X_0 = x_0, \dots, X_{n+1} = x_n \mid \mathcal{F}_n) &= \mathbb{1}_{X_0=x_0} \dots \mathbb{1}_{X_n=x_n} \mathbb{P}(X_{n+1} = x_{n+1} \mid \mathcal{F}_n) \\ &= \mathbb{1}_{X_0=x_0} \dots \mathbb{1}_{X_n=x_n} \mathbb{P}(X_{n+1} = x_{n+1} \mid X_n).\end{aligned}$$

En prenant l'espérance sachant $X_n$ du membre de gauche — en utilisant la propriété de double-projection (1.6) — et du membre de droite — en utilisant de nouveau la relation (1.3) — on obtient

$$\begin{aligned}&\mathbb{P}(X_0 = x_0, \dots, X_{n+1} = x_n \mid X_n) \\ &\qquad = \mathbb{P}(X_0 = x_0, \dots, X_{n-1} = x_{n-1} \mid X_n) \mathbb{1}_{X_n=x_n} \mathbb{P}(X_{n+1} = x_{n+1} \mid X_n).\end{aligned}$$

Cette égalité implique bien (2.18) en se plaçant sur l'événement $\{X_n = x_n\}$.

Réciproquement,

$$\begin{aligned}&\mathbb{P}(X_{n+1} = x_{n+1} \mid X_0 = x_0, \dots, X_n = x_n) \\ &\qquad = \frac{\mathbb{P}(X_0, \dots, X_{n+1} = x_{n+1} \mid X_n = x_n)}{\mathbb{P}(X_0, \dots, X_n = x_n \mid X_n = x_n)} \\ &\qquad = \frac{\mathbb{P}(X_0 = x_0, \dots, X_{n-1} = x_{n-1} \mid X_n = x_n)\mathbb{P}(X_{n+1} = x_{n+1} \mid X_n = x_n)}{\mathbb{P}(X_0, \dots, X_{n-1} = x_{n-1} \mid X_n = x_n)} \\ &\qquad = \mathbb{P}(X_{n+1} = x_{n+1} \mid X_n = x_n),\end{aligned}$$

et on retrouve la formulation équivalente de la propriété de Markov donnée au point 1. ■

Pour finir, nous donnons quelques conséquences utiles de la propriété de Markov faible et de la propriété d'homogénéité (2.14). Pour cela, nous introduisons la notation $\mathbb{P}_x$ pour la loi conditionnelle de la chaîne de Markov $(X_n, n \geq 0)$ sachant $X_0 = x$, et on note $\mathbb{E}_x$ l'espérance associée. Notons que l'on peut définir $\mathbb{P}_x$ même si $\mathbb{P}(X_0 = x) = 0$, puisque la formule (2.9) permet de définir $\mathbb{P}_x$ seulement à partir de la famille de probabilités de transition $\pi$ comme

$$\mathbb{P}_x(X_0 = x_0, \cdots, X_n = x_n) = \mathbb{1}_{x_0=x} \pi(x_0, x_1) \cdots \pi(x_{n-1}, x_n).$$

**Proposition 2.11** *Soit* $(X_n, n \geq 0)$ *une chaîne de Markov de probabilités de transition* $\pi$*. Les propriétés de Markov faible* (2.13) *et d'homogénéité* (2.14) *impliquent les propriétés suivantes :*

1. *Pour tout* $n \geq 1$ *and* $A \subset E$*,*

$$\mathbb{P}(X_{n+1} \in A \mid \mathcal{F}_n) = \mathbb{P}_{X_n}(X_1 \in A).$$

2. *Pour tout* $n \geq 1$ *et* $f : E \to \mathbb{R}$ *bornée,*

$$\mathbb{E}(f(X_{n+1}) \mid \mathcal{F}_n) = \mathbb{E}_{X_n}(f(X_1)).$$

3. *Pour tout* $n, k \geq 1$ *et* $x_1, \ldots, x_k \in E$,
$$\mathbb{P}(X_{n+1} = x_1, \ldots, X_{n+k} = x_k \mid \mathcal{F}_n) = \mathbb{P}_{X_n}(X_1 = x_1, \ldots, X_k = x_k).$$
4. *Pour tout* $n \geq 1$ *et* $\Gamma \in \mathcal{M}$,
$$\mathbb{P}((X_{n+k}, k \geq 0) \in \Gamma \mid \mathcal{F}_n) = \mathbb{P}_{X_n}((X_k, k \geq 0) \in \Gamma).$$
5. *Pour tout* $n \geq 1$ *et toute fonction* $F : M \to \mathbb{R}$ *bornée et mesurable par rapport à la tribu* $\mathcal{M}$,
$$\mathbb{E}[F(X_{n+k}, k \geq 0) \mid \mathcal{F}_n] = \mathbb{E}_{X_n}[F(X_k, k \geq 0)].$$

C'est souvent les propriétés donnée dans cette proposition qui sont les formes les plus utiles de la propriété de Markov. Des exercices sont donnés plus loin.

**Preuve.** Chacune de ces propriétés se déduit de la forme correspondante de la Proposition 2.10 de la même façon. Nous détaillons la preuve du point 3. seulement : d'après (2.19) et (2.9),
$$\begin{aligned}\mathbb{P}(X_{n+k} = x_k, \ldots, X_{n+1} = x_1 \mid \mathcal{F}_n) &= \pi(X_n, x_1)\pi(x_1, x_2) \ldots \pi(x_{k-1}, x_k) \\ &= \mathbb{P}_{X_n}(X_1 = x_1, \ldots, X_k = x_k).\end{aligned}$$
On obtient bien la propriété voulue. ■

## 2.3 Construction d'une chaîne de Markov

La définition d'une chaîne de Markov est une définition *en loi.* Étant donné un processus défini trajectoriellement (c'est-à-dire $\omega$ par $\omega$) sur un certain espace de probabilités, il peut être parfois délicat de vérifier que c'est une chaîne de Markov, c'est-à-dire que sa loi est bien celle d'une chaîne de Markov. Le critère suivant est souvent utile en pratique. Il donne également une façon de simuler une chaîne de Markov sur un ordinateur.

**Proposition 2.12** *Soit* $(\mathbb{V}, \mathcal{V})$ *un espace mesurable quelconque. Soit un processus* $(X_n, n \geq 0)$ *à valeurs dans* $E$ *défini par une v.a.* $X_0 \in E$ *et la relation de récurrence*
$$X_{n+1} = F(X_n, V_n), \tag{2.20}$$
*où* $F : E \times \mathbb{V} \to E$ *est une fonction mesurable et* $(V_n, n \geq 0)$ *est une suite de v.a. i.i.d. à valeurs dans* $\mathbb{V}$, *indépendante de* $X_0$. *Alors* $(X_n, n \geq 0)$ *est une chaîne de Markov.*

*Récirporquement, étant donnés une famille de probabilités de transition* $\pi = (\pi(x, y), x, y \in E)$, *une v.a.* $X_0$ *et une suite* $(V_n, n \geq 0)$ *indépendante de* $X_0$ *et i.i.d. de la loi uniforme sur* $[0, 1]$, *il existe une fonction* $F : E \times [0, 1] \to E$ *mesurable telle que le processus* $(X_n, n \geq 0)$ *défini par* (2.20) *est une chaîne de Markov de probabilité de transition* $\pi$.

**Preuve.** Pour tout $n \geq 1$ et pour tout $x_0, \ldots, x_{n+1} \in E$,
$$\begin{aligned}&\mathbb{P}(X_{n+1} = x_{n+1} \mid X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n) \\ &\qquad = \mathbb{P}(F(X_n, V_n) = x_{n+1} \mid X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n) \\ &\qquad = \mathbb{P}(F(X_n, V_n) = x_{n+1} \mid X_n = x_n) \\ &\qquad = \mathbb{P}(F(x_n, V_n) = x_{n+1}) \\ &\qquad = \mathbb{P}(F(x_n, V_0) = x_{n+1}),\end{aligned}$$

où la seconde égalité découle de (1.7) et de l'indépendance de $X_0, \dots, X_n$ avec $V_n$, évidente puisque $X_0, \dots, X_n$ sont des fonctions de $X_0, V_0, \dots, V_{n-1}$ seulement. Le fait que

$$\mathbb{P}(X_{n+1} = x_{n+1} \mid X_0 = x_0, X_1 = x_1, \dots, X_n = x_n)$$

ne dépende que de $x_n$ et de $x_{n+1}$ et pas de $n$, $x_0, \dots, x_{n-1}$ est exactement la caractérisation d'une chaîne de Markov donnée dans la Proposition 2.9. ■

**Exercice 2.13** *Démontrer la réciproque du lemme précédent en utilisant la propriété bien connue que, pour toute v.a. réelle $X$ de fonction de répartition $F$, $F^{-1}(V)$ a même loi que $X$ si $V$ est une v.a. uniforme sur $[0,1]$, où $F^{-1}$ est la fonction inverse généralisée de $F$, définie par $F^{-1}(y) = \inf\{x \in \mathbb{R} : F(x) \geq y\}$ pour tout $y \in ]0,1[$.*

**Exercice 2.14 Marches** *(ou* **promenades***)* **aléatoires**.
*Soient $(Y_n; n \geq 1)$ une suite de v.a. à valeurs dans $\mathbb{Z}^d$, indépendantes et de même loi $\nu$. On pose :*

$$X_n = X_0 + Y_1 + \cdots + Y_n, \quad n \geq 0,$$

*où $X_0$ est une v.a. à valeurs dans $\mathbb{Z}^d$, de loi $\mu$ et indépendante de la suite $(Y_n; n \geq 1)$.*

1. *De la relation $X_{n+1} = X_n + Y_n$ déduire que $(X_n; n \geq 0)$ est une chaîne de Markov de loi initiale $\mu$ et de probabilité de transition $\pi(x,y) = \nu(y-x)$.*
2. *Montrer :*

$$\int_E f(y)\pi(x,dy) = \mathbb{E}\big[f(x+Y_1)\big],$$

*où $f : E \to \mathbb{R}$ est positive ou bornée.*

En particulier si $d = 1$ et $\mu$ est une probabilité sur $\{-1,0,1\}$, $(X_n; n \geq 0)$ modélise l'évolution de la **fortune d'un joueur**. Un joueur participe à une série de parties indépendantes entre elles. A chaque partie le joueur peut gagner une somme $S$, perdre $S$ ou conserver son gain si la partie est déclarée nulle. Pour simplifier on prend $S = 1$. $X_0$ est la fortune initiale du joueur, et $X_n$ représente la fortune du joueur après $n$ parties.

Si de plus $X_0 = 0$ (i.e. $\mu = \delta_0$) et $\nu = p\delta_1 + (1-p)\delta_{-1}$ avec $0 \leq p \leq 1$ (i.e. $\mathbb{P}(Y_1 = 1) = p$ et $\mathbb{P}(Y_1 = -1) = 1-p$), alors $(X_n; n \geq 0)$ est la **marche aléatoire de Bernoulli de paramètre** $p$ sur $\mathbb{Z}$ issue de 0, et de probabilité de transition $\pi$, avec $\pi(n,n+1) = p$ et $\pi(n,n-1) = 1-p$ pour tout $n \in \mathbb{Z}$. Dans le cas où $p = 1/2$ on parle également de **marche aléatoire simple**. Les exercices suivants portent sur ces processus.

**Exercice 2.15** *On suppose que $E = \mathbb{Z}$. On pose $\mu := \delta_0$ et*

$$\pi(x,x+1) := \frac{1}{2}, \quad \pi(x,x-1) := \frac{1}{2}, \quad \forall x \in \mathbb{Z}.$$

*Soit $(X_n)_{n\geq 0}$ la chaine de Markov de loi initiale $\mu$ et de probabilité de transition $\pi$ et $R_n := \frac{1}{2}X_{2n}$. Montrer que $R_n$ est à valeurs dans $\mathbb{Z}$ et que $(R_n)_{n\geq 0}$ et $(|X_n|)_{n\geq 0}$ sont deux chaines de Markov.*

**Exercice 2.16** *Soit* $(Y_n)_{n\geq 1}$ *une suite de variables aléatoires indépendantes à valeurs dans* $E := \{-1, 1\}$ *et de loi :*

$$\mathbb{P}(Y_n = 1) = p, \quad \mathbb{P}(Y_n = -1) = q = 1 - p, \quad n \geq 1.$$

*Soit* $(X_n)$ *la marche aléatoire associée :*

$$X_n = \sum_{k=1}^{n} Y_k, \quad n \geq 1, \quad X_0 = 0.$$

*Pour tout* $n \geq 0$, *on pose :* $S_n := \max\{X_0, X_1, \cdots, X_n\}$. *Montrer que* $(X_n, S_n)_{n\geq 0}$ *est une chaine de Markov mais que* $(S_n)_{n\geq 0}$ *n'est pas une chaine de Markov.*

**Exercice 2.17** *Soit* $(X_n, n \geq 0)$ *une marche aléatoire de Bernoulli de paramètre* $p$. *On se donne un entier* $N \geq 1$, *et on appelle* $T_0$ *et* $T_N$ *les premiers temps d'atteinte de* $0$ *et de* $N$ *par la marche aléatoire :*

$$T_0 = \inf\{n \geq 0 : X_n = 0\} \quad et \quad T_N = \inf\{n \geq 0 : X_n = N\}.$$

*Pour tout* $k \in \{0, 1, \ldots, N\}$, *on note*

$$p_k = \mathbb{P}_k(T_N < T_0) \quad et \quad e_k = \mathbb{E}_k[T_0 \wedge T_N].$$

1. *En utilisant la propriété de Markov (on pourra par exemple utiliser la Proposition 2.11 4. en choisissant* $n = 1$ *et un événement* $\Gamma$ *convenable), montrer que, pour tout* $1 \leq k \leq N - 1$,
$$p_k = pp_{k+1} + (1-p)p_{k-1} \quad et \quad e_k = 1 + pe_{k+1} + (1-p)e_{k-1}.$$
2. *Donner les valeurs de* $p_0$ *et* $p_N$ *et écrire une relation de récurrence sur les nombres* $q_k = p_{k+1} - p_k$.
3. *En déduire que, si* $p \neq 1/2$, *pour tout* $0 \leq k \leq N$,
$$p_k = \frac{\left(\frac{1-p}{p}\right)^k - 1}{\left(\frac{1-p}{p}\right)^N - 1}.$$
*Calculer* $p_k$ *lorsque* $p = 1/2$.
4. *Déterminer la limite de* $\mathbb{P}_1(T_0 < T_N)$ *lorsque* $N \to +\infty$ *en fonction de* $p$. *Interpréter le résultat.*
5. *Démontrer que, lorsque* $p \neq 1/2$,
$$e_k = C\left[\left(\frac{1-p}{p}\right)^k - \frac{k}{N}\left(\frac{1-p}{p}\right)^N - \frac{N-k}{N}\right]$$
*pour une constante* $C$ *que l'on déterminera. Donner l'expression de* $e_k$ *lorsque* $p = 1/2$.
6. *Déterminer la limite de* $\mathbb{E}_1(T_0 \wedge T_N)$ *lorsque* $N \to +\infty$ *en fonction de* $p$.

**Exercice 2.18** *On considère une double suite $(\xi_{i,n}, i \in \mathbb{N}^*, n \in \mathbb{N})$ de v.a. i.i.d. à valeurs dans $\mathbb{N}$, et une v.a. $Z_0$ indépendante. On définit pour tout $n \geq 0$*

$$Z_{n+1} = \sum_{i=1}^{Z_n} \xi_{i,n}.$$

*Dans ce processus, $Z_n$ représente une taille de population à la génération $n$, et $\xi_{i,n}$ représente le nombre d'enfants du i-ième individu vivant à la date $n$. Chaque individu vivant à chaque génération a donc un nombre i.i.d. d'enfants. Le processus $(Z_n, n \geq 0)$ s'appelle* ***processus de Galton-Watson****.*

1. *Démontrer que $(Z_n, n \geq 0)$ est une chaîne de Markov à valeurs dans $\mathbb{N}$.*
2. *La chaîne de Markov a-t-elle des points absorbants ? Lesquels ?*
3. *On suppose que $\mathbb{P}(\xi = 2) = p = 1 - \mathbb{P}(\xi = 0)$, où $\xi$ est l'un des $\xi_{i,n}$ (ils ont tous même loi). Démontrer que la probabilités de transition $\pi = (\pi(x, y), x, y \in \mathbb{N})$ de cette chaîne de Markov est donnée par*

$$\pi(x,y) = \begin{cases} 0 & \text{si } y \text{ est impair,} \\ \binom{x}{z} p^z (1-p)^{x-z} & \text{si } y = 2z \text{ pour } z \in \mathbb{N}, \end{cases}$$

   *où l'on utilise la convention que le coefficient binomial $\binom{x}{z}$ vaut $0$ si $z > x$.*
4. *Pour tout $s \in [-1, 1]$, démontrer que*

$$\mathbb{E}(s^{Z_{n+1}} \mid Z_n) = (s^2 p + 1 - p)^{Z_n}.$$

   *On définit la fonction $f(s) = s^2 p + 1 - p$ pour tout $s \in [-1, 1]$.*
5. *Pour tout $n \geq 1$ et $s \in [-1, 1]$, montrer que*

$$\mathbb{E}_1(s^{Z_n}) = f^{\circ n}(s),$$

   *où $f^{\circ n} = f \circ \ldots \circ f$ est la composée $n$-ième de $f$.*
6. *Exprimer $\mathbb{P}(Z_n = 0)$ en fonction de $f^{\circ n}$ (on reconnaîtra une suite récurrente). En étudiant la fonction $f$, déterminser la limite de $\mathbb{P}(Z_n = 0)$ lorsque $n \to +\infty$ en fonction de $p$. Justifier le nom de* ***probabilité d'extinction du processus de Galton-Watson*** *pour cette limite.*

## 2.4 Lois marginales d'une chaîne de Markov

**Définition 2.19**

1. *Soient $\pi_1$ et $\pi_2$ deux probabilités de transition sur $E$, $\pi_1\pi_2 = \pi_1 \times \pi_2$ désigne le* **produit** *de $\pi_1$ et $\pi_2$ et est défini de la manière suivante :*

$$\pi_1\pi_2(x,y) = \sum_{z \in E} \pi_1(x,z)\pi_2(z,y), \quad x, y \in E.$$

   *Lorsque $E$ est un ensemble fini, il est clair que $\pi_1\pi_2$ est le produit usuel des deux matrices $\pi_1$ et $\pi_2$.*
2. *En particulier si $\pi$ est une probabilité de transition, et $n$ un entier naturel $n \geq 1$, alors $\pi^n$ est la* **puissance $n$-ième** *de $\pi$ : $\pi^n = \underbrace{\pi \times \cdots \times \pi}_{n \text{ fois}}$.*

3. *Si* $\mu = (\mu(x), x \in E)$ *est une probabilité sur* $E$ *et* $\pi$ *une probabilité de transition également sur* $E$*, soit* $\mu\pi$ *la mesure positive :*

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

*Lorsque* $E$ *est fini, avec nos conventions,* $\mu\pi$ *est le produit de la matrice uniligne* $\mu$ *par la matrice* $\pi$*.*

**Exercice 2.20** *Soient* $\mu$ *une probabilité,* $\pi$*,* $\pi_1$ *et* $\pi_2$ *trois probabilités de transition sur* $E$*. Montrer :*

1. $\pi_1\pi_2$ *est une probabilité de transition sur* $E$ *;*
2. $\mu\pi$ *est une probabilité sur* $E$*.*

Le produit de matrices de transition permet d'exprimer aisément la loi d'une sous-suite finie de $(X_n, n \geq 0)$.

**Proposition 2.21** *Soit* $(X_n)$ *une chaîne de Markov de probabilité de transition* $\pi$*. Alors :*

1. *la loi conditionnelle de* $X_{n+m}$ *sachant* $X_0, \cdots, X_n$ *est* $\pi^m(X_n, \cdot)$ *;*
2. *Si* $X_0$ *a pour loi* $\mu$ *(resp.* $X_0 = x_0$*, avec* $x_0 \in E$*) alors* $X_m$ *a pour loi* $\mu\pi^m$ *(resp.* $\pi^m(x_0, \cdot)$*).*
3. *Si* $X_0$ *a pour loi* $\mu$ *et si, pour tout* $k \geq 1$*,* $0 < m_1 < m_2 < \ldots < m_k \in \mathbb{N}$*,* $x_0, x_1, \ldots, x_k \in E$*, alors*

$$\begin{aligned}\mathbb{P}(X_0 = x_0, X_{m_1} = x_1, X_{m_2} = x_2 \ldots, X_{m_k} = x_k)& \\ = \mu(x_0)\pi^{m_1}(x_0, x_1)\pi^{m_2 - m_1}(x_1, x_2) \ldots \pi^{m_k - m_{k-1}}(x_{k-1}, x_k)&.\end{aligned}$$

Pour la preuve de la Proposition 2.21, voir l'Exercice 2.22.

**Exercice 2.22** *Soit* $(X_n)$ *une chaîne de Markov de probabilité de transition* $\pi$ *et on note* $\mu$ *la loi de* $X_0$*.*

1. *Montrer que* $X_1$ *a pour loi* $\mu\pi$ *et* $X_2$ *a pour loi* $\mu\pi^2$*.*
2. *Montrer par récurrence sur* $m$*, que* $X_m$ *a pour loi* $\mu\pi^m$*.*
3. *On souhaite montrer que la loi conditionnelle de* $X_{n+m}$ *sachant* $X_0, \cdots, X_n$ *est* $\pi^m(X_n, \cdot)$*.*
   (a) *Etudier le cas* $m = 1$*.*
   (b) *On étudie le passage de* $m$ *à* $m + 1$*. Soient :* $f : E^{n+1} \to \mathbb{R}$ *et* $g : E \to \mathbb{R}$ *deux fonctions bornées. Posons :*

$$\Delta := \mathbb{E}\big[g(X_{n+m+1})f(X_0, \cdots, X_n)\big].$$

   - *Montrer :*

$$\Delta = \mathbb{E}\big[\pi(X_{n+m}, g)f(X_0, \cdots, X_n)\big].$$

   *puis*

$$\Delta = \mathbb{E}\big[\pi^m(X_n, g_1)f(X_0, \cdots, X_n)\big],$$

   *avec* $g_1(y) = \pi(y, g)$*,* $y \in E$*.*
   - *Vérifier :*

$$\pi^m(x, g_1) = \sum_{z \in E} g(z)\pi^{m+1}(x, z)$$

   *et ensuite conclure.*

4. *En déduire que*

$$\mathbb{P}(X_0 = x_0, X_{m_1} = x_1, X_{m_2} = x_2 \ldots, X_{m_k} = x_k)$$
$$= \mu(x_0)\pi^{m_1}(x_0, x_1)\pi^{m_2-m_1}(x_1, x_2) \ldots \pi^{m_k-m_{k-1}}(x_{k-1}, x_k).$$

5. *Montrer que le résultat de la question 3 implique celui de la question 2.*

**Exercice 2.23** *Soit $\pi$ la matrice :*

$$\pi := \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 0 & 1/3 & 2/3 \\ 0 & 3/4 & 1/4 \end{pmatrix}$$

1. *Montrer que $\pi$ est une matrice de transition et tracer le graphe associé.*
2. *Soit $(X_n, n \geq 0)$ la chaine de Markov associée, à valeurs dans $\{1, 2, 3\}$. On suppose que*

$$\mathbb{P}(X_0 = 1) = \frac{1}{5}, \quad \mathbb{P}(X_0 = 2) = \frac{3}{5}, \quad \mathbb{P}(X_0 = 3) = \frac{1}{5}.$$

   *Quelle est la loi de $X_2$ ?*

**Exercice 2.24** *Soit $(X_n, n \geq 0)$ une chaine de Markov à valeurs dans $E$, de loi initiale $\mu$ et de probabilité de transition $\pi$. Soient $a$ et $b$ deux entiers supérieurs ou égaux à 1. Montrer que $(X_{n+a}, n \geq 0)$ et $(X_{bn}, n \geq 0)$ sont des chaines de Markov. Pour chacun des 2 processus, préciser la loi initiale et la probabilité de transition.*

**Exercice 2.25** *Soit $(X_n)_{n\geq 0}$ la chaine de Markov à valeurs dans $E = \{1, 2\}$ de loi initiale $(\mu_0, 1 - \mu_0)$ et de probabilité de transition :*

$$\pi := \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$$

*où $0 \leq \alpha, \beta \leq 1$. Montrer :*

$$\mathbb{P}(X_n = 1) = \frac{\beta}{\alpha+\beta} + (1 - \alpha + \beta)^n \Big(\mu_0 - \frac{\beta}{\alpha+\beta}\Big).$$

## 3 Temps d'arrêt et propriété forte de Markov

Le but de cette section est de généraliser la propriété de Markov faible présentée en section 2.2 à des temps aléatoires. La difficulté principale consiste à identifier la bonne classe de temps aléatoire où l'on peut appliquer la propriété de Markov. Si l'on prend l'exemple de la fortune du joueur, l'intuition est que, s'il arrive au casino av ec 100 Euros, après que sa fortune ait atteint 10 Euros pour la première fois, la suite du jeu doit avoir la même loi que si le joueur était arrivé dès la début avec 10 Euros. C'est donc toujours une propriété d'indépendance du futur et du passé conditionnellement au présent, mais à un temps qui est aléatoire (ici, le premier temps d'atteinte de 10 par le processus).

On commence par quelques définitions.

**Définition 3.1**

1. *Une* **filtration** $(\mathcal{F}_n)_{n\in\mathbb{N}}$ *est une suite croissante de sous-tribus de* $\mathcal{F}$ *: pour tout* $n$ *et* $m$ *de* $\mathbb{N}$, $n \leq m \Rightarrow \mathcal{F}_n \subset \mathcal{F}_m \subset \mathcal{F}$. *Soit* $(\xi_n)_{n\in\mathbb{N}}$ *une suite de v. a. à valeurs dans* $E$. *L'exemple classique de filtration associé à cette suite de v.a. est obtenu en prenant* $\mathcal{F}_n = \sigma(\xi_0, \cdots, \xi_n)$ *la tribu engendrée par* $\xi_0, \cdots, \xi_n$. *On rappelle que toute v.a.* $X$, $\mathcal{F}_n$*- mesurable s'écrit comme une fonction mesurable de* $\xi_0, \cdots, \xi_n$ *:* $X = f(\xi_0, \cdots, \xi_n)$.
2. *Un* **processus** $X = (X_k; k \in \mathbb{N})$ *à valeurs dans un espace mesuré* $(U,\mathcal{U})$ *est une collection de fonctions mesurables* $X_k : (\Omega, \mathcal{F}) \to \big(U,\mathcal{U}\big)$. *Dans la suite nous considérerons essentiellement des processus à valeurs dans* $E$.
3. *Un processus* $(X_n : n \in \mathbb{N})$ *est* **adapté** *par rapport à* $(\mathcal{F}_n)_{n\in\mathbb{N}}$ *si pour tout* $n$ *de* $\mathbb{N}$, $X_n$ *est* $\mathcal{F}_n$*-mesurable. Remarquons que si* $\mathcal{F}_n = \sigma(X_0, \cdots, X_n)$ *alors* $(X_n : n \in \mathbb{N})$ *est adapté.*
4. *Une application* $T$ *de* $(\Omega, \mathcal{F})$ *à valeurs dans* $\mathbb{N} \cup \{+\infty\}$ *est un* $(\mathcal{F}_n)_{n\in\mathbb{N}}$ **temps d'arrêt** *si* $\{T \leq n\} \in \mathcal{F}_n \quad \forall n \in \mathbb{N}$.

   *Si* $(\mathcal{F}_n)_{n\in\mathbb{N}}$ *est fixée, on dit simplement que* $T$ *est un temps d'arrêt.*

La notion de temps d'arrêt est fondamentale. Il s'agit d'un temps aléatoire qui n'anticipe pas sur son futur. Autrement dit, pour décider si le temps d'arrêt vaut $n$, on ne se base que sur l'information disponible au temps $n$, c'est-à-dire les valeurs des v.a. $X_0, X_1, \ldots, X_n$. C'est exactement ce que formule le point 3. de la proposition suivante.

**Proposition 3.2**

1. *Les temps constants sont des temps d'arrêt.*
2. *Si* $U_1$ *et* $U_2$ *sont deux temps d'arrêt,* $\sup\{U_1, U_2\}, U_1 + U_2$ *et* $\inf\{U_1, U_2\}$ *sont deux temps d'arrêt.*
3. $U$ *est un temps d'arrêt si et seulement si*
$$\{U = k\} \in \mathcal{F}_k \quad \textit{pour tout} \;\; k \in \mathbb{N}. \tag{3.21}$$
4. *Soient* $(X_k)_{k\in\mathbb{N}}$ *un processus adapté à valeurs dans* $(U,\mathcal{U})$, *et* $F$ *un élément de* $\mathcal{U}$. *On pose :*
$$D_F = \inf\{k \in \mathbb{N}, X_k \in F\}$$
*(avec la convention* $\inf \emptyset = +\infty$*).*

   *Alors* $D_F$ *est un temps d'arrêt.*

**Preuve.** Voir l'Exercice 3.3. ∎

**Exercice 3.3** *1. Montrer que $U$ est un temps d'arrêt si et seulement si*

$$\{U = k\} \in \mathcal{F}_k \quad \textit{pour tout} \;\; k \in \mathbb{N}.$$

2. *Soient $U_1$ et $U_2$ deux temps d'arrêt. Montrer que* $\sup\{U_1, U_2\}, U_1 + U_2$ *et* $\inf\{U_1, U_2\}$ *sont des temps d'arrêt.*
3. *Soient $(X_k)_{k\in\mathbb{N}}$ un processus adapté à valeurs dans $(U,\mathcal{U})$, et $F$ un élément de $\mathcal{U}$. Montrer que*

$$D_F = \inf\{k \in \mathbb{N}, X_k \in F\}$$

*est un temps d'arrêt.*

Pour décider si un événemnt de $\mathcal{F}_n$ est réalisé, il suffit de connaître les valeurs de $X_0, \ldots, X_n$. Autrement dit, la tribu $\mathcal{F}_n$ contient toute l'information disponible au temps $n$. On veut définir une tribu qui contient toute l'information disponible au temps d'arrêt $T$, c'est-à-dire la tribu des événements pour lesquels ils suffit de connaître les valeurs de $X_0, \ldots, X_T$ pour décider si l'événement est réalisé.

**Définition 3.4** *Soit $T$ un $(\mathcal{F}_n)_{n\in\mathbb{N}}$- temps d'arrêt, on pose*

$$\mathcal{F}_T = \{A \in \mathcal{F}; A \cap \{T \le n\} \in \mathcal{F}_n \quad \forall n \in \mathbb{N}\}.$$

**Proposition 3.5** *1. (a) $\mathcal{F}_T$ est une sous-tribu de $\mathcal{F}$, $T$ désignant un temps d'arrêt.*

*(b) Si $(n)$ désigne le temps d'arrêt constant égal à $n$, alors $\mathcal{F}_{(n)} = \mathcal{F}_n$.*

*(c) La tribu $\mathcal{F}_T$ peut être définie d'une manière équivalente : $\mathcal{F}_T = \{A \in \mathcal{F}; A \cap \{T = n\} \in \mathcal{F}_n \quad \forall n \in \mathbb{N}\}$.*

*(d) Si $S$ et $T$ sont deux temps d'arrêt alors,*

$$\mathcal{F}_{S\wedge T} = \mathcal{F}_S \cap \mathcal{F}_T \quad \textit{et}\, \mathcal{F}_S \subset \mathcal{F}_T \textit{ dès que } S \le T.$$

2. *Soient $T$ un temps d'arrêt fini à valeurs dans $\mathbb{N}$ et $(X_n; n \in \mathbb{N})$ un processus adapté. On pose $X_T(\omega) = X_{T(\omega)}(\omega)$. Alors $X_T$ est une v.a. $\mathcal{F}_T$-mesurable. En particulier $T$ est $\mathcal{F}_T$-mesurable.*

**Exercice 3.6** *Soit $T$ un $(\mathcal{F}_n)_{n\in\mathbb{N}}$- temps d'arrêt. Montrer :*

1. *$\mathcal{F}_T$ est une sous-tribu de $\mathcal{F}$.*
2. *Si $(n)$ désigne le temps d'arrêt constant égal à $n$, alors $\mathcal{F}_{(n)} = \mathcal{F}_n$.*
3. *$\mathcal{F}_T = \{A \in \mathcal{F}; A \cap \{T = n\} \in \mathcal{F}_n \quad \forall n \in \mathbb{N}\}$.*
4. *Si $S$ et $T$ sont deux temps d'arrêt alors,*

$$\mathcal{F}_{S\wedge T} = \mathcal{F}_S \cap \mathcal{F}_T \quad \textit{et}\, \mathcal{F}_S \subset \mathcal{F}_T \textit{ dès que } S \le T.$$

5. *Soient $T$ un temps d'arrêt fini à valeurs dans $\mathbb{N}$ et $(X_n; n \in \mathbb{N})$ un processus adapté. On pose $X_T(\omega) = X_{T(\omega)}(\omega)$. Alors :*

*(a) $X_T$ est une v.a. $\mathcal{F}_T$-mesurable. En particulier $T$ est $\mathcal{F}_T$-mesurable.*

*(b)* $(X_{T\wedge t})_{t\in\mathbb{N}}$ *est un processus* $(\mathcal{F}_{T\wedge t})_{t\in\mathbb{N}}$*-adapté (resp.* $(\mathcal{F}_t)_{t\in\mathbb{N}}$*-adapté).*

Nous nous proposons de généraliser la propriété de Markov énoncée dans la Proposition 2.3 en remplaçant le temps constant $n$ par un temps d'arrêt $T$. La filtration utilisée est $\big(\mathcal{F}_n\big)_{n\geq 0}$, avec $\mathcal{F}_n = \sigma(X_0, X_1, \cdots, X_n)$ et $(X_n)$ est une chaine de Markov de probabilité de transition $\pi$.

**Théorème 3.7** *(***Propriété forte de Markov***) Soit $T$ un temps d'arrêt fini. Alors :*

1. *Pour tout entier $n$,*

$$\mathbb{P}(X_{T+1} \in A|\mathcal{F}_T) = \pi(X_T, A), \quad \forall\, A \subset E. \tag{3.22}$$

2. *En particulier, le processus $(X_{T+n}; n \geq 0)$ est une chaîne de Markov de probabilité de transition $\pi$ et de loi initiale $\nu$, la loi de $X_T$.*

**Preuve du Théorème 3.7.**
Montrons (3.22). Cette relation est équivalente à

$$\mathbb{P}\big(\{X_{T+1} \in A\} \cap \Lambda\big) = \mathbb{E}\big(\pi(X_T, A)\mathbb{1}_\Lambda\big) \tag{3.23}$$

pour tout $\Lambda \in \mathcal{F}_T$.
Posons : $\Delta := \mathbb{P}\big(\{X_{T+1} \in A\} \cap \Lambda)$.
Puisque $T$ prend des valeurs entières,

$$\Delta = \sum_{k\geq 0} \Delta_k$$

avec

$$\begin{aligned} \Delta_k &:= \mathbb{P}\big(\{X_{T+1} \in A\} \cap \Lambda \cap \{T = k\}\big) \\ &= \mathbb{P}\big(\{X_{k+1} \in A\} \cap \Lambda \cap \{T = k\}\big) \\ &= \mathbb{P}\big(\{X_{k+1} \in A\} \cap \Lambda_k\big) \end{aligned}$$

avec $\Lambda_k := \Lambda \cap \{T = k\}$.
Mais $\Lambda \in \mathcal{F}_T$, donc $\Lambda_k \in \mathcal{F}_k$. On peut appliquer la propriété de Markov usuelle :

$$\begin{aligned} \Delta_k &= \mathbb{E}\big(\pi(X_k, A)\mathbb{1}_{\Lambda_k}\big) \\ &= \mathbb{E}\big(\pi(X_k, A)\mathbb{1}_{\Lambda\cap\{T=k\}}\big) \\ &= \mathbb{E}\big(\pi(X_k, A)\mathbb{1}_\Lambda \mathbb{1}_{\{T=k\}}\big). \end{aligned}$$

Donc :

$$\begin{aligned} \Delta &= \sum_{k\geq 0} \mathbb{E}\big(\pi(X_k, A)\mathbb{1}_\Lambda \mathbb{1}_{\{T=k\}}\big) \\ &= \sum_{k\geq 0} \mathbb{E}\big(\pi(X_T, A)\mathbb{1}_\Lambda \mathbb{1}_{\{T=k\}}\big) \\ &= \mathbb{E}\big(\pi(X_T, A)\mathbb{1}_\Lambda\big). \end{aligned}$$

■

Il existe beaucoup d'autres formes équivalentes à la propriété de Markov forte, comme dans la section 2.2. Nous en donnons certaines dans les deux propostions ci-dessous. Leurs preuves sont très similaires à celles des propositions 2.10 et 2.11. Nous ne les détaillons pas.

**Proposition 3.8** *Soit $T$ un temps d'arrêt fini. La propriété de Markov forte est équivalente à chacun des points ci-dessous :*

1. *Pour tout $n \geq 1$, $x_0, \ldots, x_n \in E$ et $A \subset E$,*

$$\mathbb{P}(X_{T+1} \in A \mid T = n, X_0 = x_0, \ldots, X_n = x_n) = \mathbb{P}(X_{n+1} \in A \mid X_n = x_n).$$

2. *Pour tout $k \geq 1$ et $x_1, \ldots, x_k \in E$,*

$$\mathbb{P}(X_{T+1} = x_1, \ldots, X_{T+k} = x_k \mid \mathcal{F}_T) = \mathbb{P}(X_{T+1} = x_1, \ldots, X_{T+k} = x_k \mid X_T).$$

3. *En reprenant les notations de la proposition 2.10, pour tout $\Gamma \in \mathcal{M}$,*

$$\mathbb{P}((X_{T+k}, k \geq 0) \in \Gamma \mid \mathcal{F}_T) = \mathbb{P}((X_{T+k}, k \geq 0) \in \Gamma \mid X_T).$$

4. *Pour toute fonction $F : M \to \mathbb{R}$ bornée et mesurable par rapport à la tribu $\mathcal{M}$,*

$$\mathbb{E}[F(X_{T+k}, k \geq 0) \mid \mathcal{F}_T] = \mathbb{E}[F(X_{T+k}, k \geq 0) \mid X_T].$$

5. ***(indépendance du futur et du passé conditionnellement au présent)*** *Pour tout $n \geq 1$ et $x_0, \ldots, x_{n+1}$,*

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

6. *Pour tout $x \in E$, conditionnellement à $\{X_T = x\}$, le processus $(X_{T+k}, k \geq 0)$ est indépendant de la tribu $\mathcal{F}_T$ et a pour loi la loi de la chaîne de Markov de loi initiale $\delta_x$ (la masse de Dirac en $x$) et de probabilité de transition $\pi$.*

Nous utilisons de nouveau la notation $\mathbb{P}_x$ pour la loi de la chaîne de Markov conditionnée à $\{X_0 = x\}$.

**Proposition 3.9** *Soit $T$ un temps d'arrêt fini. La propriété de Markov forte implique :*

1. *Pour tout $A \subset E$, $\mathbb{P}(X_{T+1} \in A \mid \mathcal{F}_T) = \mathbb{P}_{X_T}(X_1 \in A)$.*
2. *Pour tout $k \geq 1$ et $x_1, \ldots, x_k \in E$,*

$$\mathbb{P}(X_{T+1} = x_1, \ldots, X_{T+k} = x_k \mid \mathcal{F}_T) = \mathbb{P}_{X_T}(X_1 = x_1, \ldots, X_k = x_k).$$

3. *Pour tout $n \geq 1$ et $\Gamma \in \mathcal{M}$,*

$$\mathbb{P}((X_{T+k}, k \geq 0) \in \Gamma \mid \mathcal{F}_T) = \mathbb{P}_{X_T}((X_k, k \geq 0) \in \Gamma).$$

4. *Pour tout $n \geq 1$ et toute fonction $F : M \to \mathbb{R}$ bornée et mesurable par rapport à la tribu $\mathcal{M}$,*

$$\mathbb{E}[F(X_{T+k}, k \geq 0) \mid \mathcal{F}_T] = \mathbb{E}_{X_T}[F(X_k, k \geq 0)].$$

**Exercice 3.10** *On considère $(X_n, n \geq 0)$ une marche aléatoire de Bernoulli de paramètre $p$. On suppose que $p < 1/2$. On se propose de calculer $\mathbb{E}_k(T_0)$ avec une méthode différente de celle de l'exercice 2.17.*

1. *Justifier que $\mathbb{E}_1(T_0) = \mathbb{E}_2(T_1)$, où $T_k$ est le premier temps d'atteinte de $k$ par la chaîne de Markov.*
2. *On suppose que $X_0 = 2$. Justifier que $T_1$ est un temps d'arrêt.*
3. *En appliquant la propriété de Markov forte au temps $T_1$, démontrer que $\mathbb{E}_2(T_0) = \mathbb{E}_2(T_1) + \mathbb{E}_1(T_0)$.*
4. *En appliquant la propriété de Markov faible au temps 1, déduire des questions précédentes que*
$$\mathbb{E}_1(T_0) = p(1 + 2\mathbb{E}_1(T_0)) + 1 - p.$$
*En déduire la valeur de $\mathbb{E}_1(T_0)$ (on admettra que ce nombre est fini).*
5. *En utilisant la propriété de Markov forte, démontrer que, pour tout $k \geq 1$,*
$$\mathbb{E}_k(T_0) = \mathbb{E}_k(T_{k-1}) + \mathbb{E}_{k-1}(T_{k-2}) + \ldots + \mathbb{E}_1(T_0).$$
*En déduire la valeur de $\mathbb{E}_k(T_0)$.*

# 4 Potentiel, états récurrents et transients

Dans cette section, $(X_n; n \geq 0)$ désigne une chaîne de Markov, de probabilité de transition $\pi$ et à valeurs dans $E$.

**Définition 4.1** *1. On notera $\mathbb{P}_x$ la probabilité conditionnelle sachant $X_0 = x$, où $x \in E$.*

*2. L'application $U : E \times E \to [0, +\infty]$ :*

$$U = \sum_{n \geq 0} \pi^n = I + \pi + \pi^2 + \ldots + \pi^n + \ldots,$$

*ou*

$$U(x, y) = \sum_{n \geq 0} \pi^n(x, y) = \mathbb{1}_{x=y} + \pi(x, y) + \pi^2(x, y) + \ldots,$$

*s'appelle* **le noyau potentiel** *ou* **potentiel** *de la chaîne, l'application $I$ désignant l'indicatrice de la diagonale de $E \times E$.*

*En utilisant le théorème de Fubini on peut écrire le potentiel sous la forme suivante :*

$$U(x, y) = \sum_{n \geq 0} \pi^n(x, y) = \sum_{n \geq 0} \mathbb{P}_x(X_n = y) = \mathbb{E}_x\Big[\sum_{n \geq 0} \mathbb{1}_{\{X_n = y\}}\Big], \quad x, y \in E. \quad (4.24)$$

*On désigne par $U(x, \cdot)$ la mesure positive associée : $U(x, B) = \sum_{y \in B} U(x, y)$, où $B$ est un sous-ensemble de $E$. Ainsi :*

$$U(x, B) = \sum_{n \geq 0} \pi^n(x, B) = \sum_{n \geq 0} \mathbb{P}_x(X_n \in B) = \mathbb{E}_x\Big[\sum_{n \geq 0} \mathbb{1}_{\{X_n \in B\}}\Big], \quad x \in E.$$

*Remarquons que $U = I + \pi U$. En particulier $I = (I - \pi)U = U(I - \pi)$ ; ce qui signifie que $U$ est "l'inverse" de $I - \pi$.*

*3. Pour tout sous-ensemble $B$ de $E$, on note $N_B$ le nombre de fois où la chaîne $(X_n, n \geq 0)$ visite $B$ :*

$$N_B = \sum_{n \geq 0} \mathbb{1}_{\{X_n \in B\}}.$$

*Lorsque $B = \{x\}$, on note $N_x = N_{\{x\}}$.*

*Remarquons que $U$ admet l'interprétation probabiliste suivante :*

$$U(x, B) = \mathbb{E}_x\big[N_B\big], \quad x \in E, B \subset E. \quad (4.25)$$

*En particulier si $B = \{x\}$, on a :*

$$U(x, x) = \mathbb{E}_x[N_x], \quad x \in E. \quad (4.26)$$

*4. Pour tout $x \in E$, on désigne par $\sigma_x$, le* **premier temps de retour** *en $x$ :*

$$\sigma_x = \inf\{n \geq 1; X_n = x\},$$

*avec la convention* $\inf \emptyset = +\infty$.

*Notons que d'après la Proposition 3.2, $\sigma_x$ est un temps d'arrêt.*

*Il est clair que d'après la définition :*

$$1 \leq \sigma_x \leq \infty, \quad X_{\sigma_x} = x \text{ sur } \{\sigma_x < \infty\}.$$

5. *Un état $x \in E$ est dit* **récurrent** *si* $\mathbb{P}_x(\sigma_x < \infty) = 1$. *Lorsque* $\mathbb{P}_x(\sigma_x < \infty) < 1$, *l'état $x$ est dit* **transient** *ou* **transitoire**.
   *Par conséquent un état est soit transient soit récurrent.*

**Théorème 4.2** *Soit $x \in E$.*

1. *$x$ est récurrent si et seulement si $\mathbb{P}_x(N_x = \infty) = 1$.*
2. *$x$ est transient si et seulement si $\mathbb{P}_x(N_x < \infty) = 1$.*

**Preuve du Théorème 4.2.** **1)** Soit $x \in E$ fixé. On définit par récurrence la suite de temps d'arrêt $(\sigma^{(n)}; n \geq 1)$ :

$$\sigma^{(1)} = \sigma_x, \quad \sigma^{(n+1)} = \inf\{k > \sigma^{(n)}; X_k = x\}. \tag{4.27}$$

Ainsi $\big\{\sigma^{(n)}; n \geq 1\big\}$ est l'ensemble des **temps de passages successifs** en $x$, plus précisément :

$$\big\{n \geq 1; \sigma^{(n)} < \infty\big\} = \big\{n \geq 1; X_n = x\big\}.$$

Remarquons de plus que :

$$\sigma^{(n+1)} = \sigma^{(n)} + \rho^{(n)}, \quad \text{sur} \quad \{\sigma^{(n)} < \infty\} \tag{4.28}$$

où

$$\rho^{(n)} := \inf\{k \geq 1, X_{\sigma^{(n)}+k} = x\}$$

**2)** On pose $a_n = \mathbb{P}_x(\sigma^{(n)} < \infty)$. Montrons :

$$a_n = a_1^n; \quad n \geq 1. \tag{4.29}$$

Soit $k$ un entier et

$$\gamma_k := \mathbb{P}_x\big(\sigma^{(n+1)} \leq k\big) = \mathbb{P}_x\big(\sigma^{(n)} + \rho^{(n)} \leq k, \big).$$

On a :

$$\gamma_k = \sum_{l=1}^{k} \gamma_{k,l}$$

avec

$$\gamma_{k,l} := \mathbb{P}_x\big(\sigma^{(n)} + l \leq k, \rho^{(n)} = l\big).$$

Or :

$$\gamma_{k,l} = \mathbb{P}_x\big(\sigma^{(n)} \leq k - l, Y_1 \neq x, \cdots, Y_{l-1} \neq x, Y_l = x\big)$$

où l'on a posé $Y_m := X_{\sigma^{(n)}+m}$. D'où :

$$\gamma_{k,l} = \sum_{y_1,\cdots,y_{l-1} \in E\setminus\{x\}} \mathbb{P}_x\big(\sigma^{(n)} \leq k - l, Y_1 = y_1, \cdots, Y_{l-1} = y_{l-1}, Y_l = x\big).$$

On applique la propriété de Markov forte (Théorème 3.7) :

$$\gamma_{k,l} = \sum_{y_1,\cdots,y_{l-1} \in E\setminus\{x\}} \mathbb{E}_x\Big(\mathbb{1}_{\{\sigma^{(n)} \leq k-l\}} \pi(X_{\sigma^{(n)}}, y_1)\pi(y_1, y_2) \times \cdots \pi(y_{l-2}, y_{l-1})\pi(y_{l-1}, x)\Big).$$

Mais si $\sigma^{(n)} < \infty$, alors $X_{\sigma^{(n)}} = x$. Ou peut donc simplifier l'expression précédente :

$$\gamma_{k,l} = \left[ \sum_{y_1,\cdots,y_{l-1} \in E \setminus \{x\}} \pi(x,y_1)\pi(y_1,y_2) \times \cdots \pi(y_{l-2},y_{l-1})\pi(y_{l-1},x) \right] \mathbb{P}_x(\sigma^{(n)} \leq k-l)$$
$$= \mathbb{P}_x(\sigma^{(1)} = l)\mathbb{P}_x(\sigma^{(n)} \leq k-l).$$

Par conséquent,

$$\gamma_k := \mathbb{P}\big(\sigma^{(n+1)} \leq k\big) = \sum_{l=1}^{k} \mathbb{P}_x(\sigma^{(1)} = l)\mathbb{P}_x(\sigma^{(n)} \leq k-l)$$
$$= \mathbb{P}_x(\sigma^{(1)} \leq k)\mathbb{P}_x(\sigma^{(n)} \leq k-l)$$

On fait tendre $k$ vers l'infini, on obtient par convergence monotone :

$$a_{n+1} = \mathbb{P}\big(\sigma^{(n+1)} < \infty\big) = \mathbb{P}_x(\sigma^{(1)} < \infty)\mathbb{P}_x(\sigma^{(n)} \leq k-l) = a_1 a_n.$$

**3)** On remarque que $\{\sigma^{(n+1)} < \infty\} \subset \{\sigma^{(n)} < \infty\}$, donc :

$$\mathbb{P}_x\Big(\bigcap_{n\geq 1} \{\sigma^{(n)} < \infty\}\Big) = \lim_{n\to\infty} \mathbb{P}_x(\sigma^{(n)} < \infty) = \lim_{n\to\infty} a_n = \lim_{n\to\infty} a_1^n.$$

Puisque $\bigcap_{n\geq 1} \{\sigma^{(n)} < \infty\} = \{N_x = \infty\}$, on en déduit que $x$ est récurrent si et seulement si $\mathbb{P}_x(N_x = \infty) = 1$.

■

**Corollaire 4.3** *Si $x$ est un état transient, alors, sous $\mathbb{P}_x$, $N_x$ est une v.a. de loi géométrique de paramètre $p = \mathbb{P}_x(\sigma_x = \infty)$ : $\mathbb{P}_x(N_x = k) = p(1-p)^{k-1}$, $k \geq 1$.*

**Preuve.** Soit $k \geq 2$. On introduit, comme dans la preuve du Théorème 4.2, la suite de temps d'arrêt $(\sigma^{(n)}; n \geq 1)$ définie par (4.28). On a : $\{N_x \geq k\} = \{\sigma^{(k-1)} < \infty\}$, $\mathbb{P}_x$ presque sûrement (on rappelle que $X_0 = x$ sous $\mathbb{P}_x$). Donc, d'après (4.29), on a :

$$\mathbb{P}_x(N_x \geq k) = \mathbb{P}_x(\sigma^{(k-1)} < \infty) = a_{k-1} = a_1^{k-1} = (1-p)^{k-1}, \quad k \geq 2.$$

Ce qui implique le résultat, sachant que $\mathbb{P}_x(N_x \geq 1) = 1$. ■

L'objet du théorème qui suit est de donner une caractérisation analytique (i.e. déterministe) des points récurrents, resp. transients.

**Théorème 4.4** *Soit $x$ un point de $E$. Alors :*

1. *$x$ est récurrent si et seulement si $U(x,x) = \infty$ ;*
2. *$x$ est transient si et seulement si $U(x,x) < \infty$ ; dans ce cas : $U(x,x) = \dfrac{1}{\mathbb{P}_x(\sigma_x = \infty)}$.*

**Preuve.** **a)** Si $x$ est récurrent, $\mathbb{P}_x(N_x = \infty) = 1$. Mais $U(x,x) = \mathbb{E}_x[N_x]$. Donc $U(x,x) = \infty$.
**b)** Supposons $x$ transient. Rappelons que si $\xi$ est une v.a. de loi géométrique de paramètre $p$ alors $\mathbb{E}[\xi] = 1/p$. On déduit alors du Corollaire 4.3 : $U(x,x) = \mathbb{E}_x[N_x] = 1/\mathbb{P}_x(\sigma_x = \infty)$. ■

Nous allons à présent étudier quelques exemples.

**Exercice 4.5** *On reprend le cadre de l'exercice 2.5, où* $E = \{1, 2\}$ *et*

$$\pi = \begin{pmatrix} \theta_1 & 1-\theta_1 \\ 1-\theta_2 & \theta_2 \end{pmatrix}$$

*avec* $0 < \theta_1, \theta_2 < 1$.

1. *Montrer que* $\mathbb{P}_1(\sigma_1 \geq k+1) = (1-\theta_1)\theta_2^{k-1}$.
2. *En déduire que l'état 1 est récurrent. Une approche analogue montrerait que 2 est également récurrent.*

Nous étudions à présent les marches aléatoires. Nous conservons les notations introduites dans l'Exemple 2.5.

**Proposition 4.6** *Soient* $(Y_n; n \geq 1)$ *une suite de v.a. à valeurs dans* $\mathbb{Z}^d$, *indépendantes et de même loi* $\nu$. *Soit* $(X_n, n \geq 0)$ *la marche aléatoire associée :* $X_n = X_0 + Y_1 + \cdots + Y_n$, *où* $X_0 \in \mathbb{Z}^d$ *et est indépendante de* $(Y_n; n \geq 1)$.

1. *Pour tout* $x$ *de* $\mathbb{Z}^d$, *on a :* $U(x,x) = U(0,0)$. *En particulier tous les points sont simultanément transients ou récurrents.*
2. *Supposons que* $\nu$ *possède un moment d'ordre 1 fini :* $\displaystyle\sum_{k\in\mathbb{Z}^d} \|k\|\nu(k) < \infty$. *Alors tous les points sont transients si* $m := \displaystyle\sum_{k\in\mathbb{Z}^d} k\nu(k) \neq 0$.

**Preuve.** **a)** Puisque $\pi(x,y) = \nu(y-x)$ il est facile de vérifier que $\pi^n(x,y) = \nu^{\star(n)}(y-x)$, où $\nu^{\star(n)}$ désigne la puissance $n$-ième de convolution. On rappelle que si $\nu_1$ et $\nu_2$, sont deux probabilités sur $\mathbb{Z}^d$, alors $\nu_1 \star \nu_2$ est la probabilité :

$$\nu_1 \star \nu_2(x) = \sum_{y\in\mathbb{Z}^d} \nu_1(y)\nu_2(x-y), \quad \forall x \in \mathbb{Z}^d. \tag{4.30}$$

Par conséquent :

$$U(x,y) = \sum_{n\geq 0} \pi^n(x,y) = \sum_{n\geq 0} \nu^{\star(n)}(y-x) = U(y-x, 0), \quad \forall x, y \in \mathbb{Z}^d.$$

Le point 1. de la Proposition 4.6 est une conséquence directe du Théorème 4.4.
**b)** On suppose les conditions du point 2. vérifiées. D'après ce qui précède, il suffit de montrer que 0 est transient.
Notons $m^{(1)}, \cdots, m^{(d)}$ les coordonnées du vecteur $m$ et $X_n^{(1)}, \cdots, X_n^{(d)}$ celles de $X_n$. Puisque $m$ est non-nul, il existe un indice $i$ tel que $m^{(i)} \neq 0$.
D'après la loi forte des grands nombres :

$$\lim_{n\to\infty} \frac{X_n^{(i)}}{n} = \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} Y_k^{(i)} = m^{(i)}.$$

Ce qui implique $|X_n^{(i)}| \sim nm^{(i)}$, lorsque $n \to \infty$. Il existe un entier (aléatoire) $n_0$, tel que $|X_n^{(i)}| \geq 1$, pour tout $n \geq n_0$. Il est clair que la chaîne ne visite plus 0 après l'instant $n_0$, 0 est bien un état transient. ■

La Proposition 4.6 conduit à étudier le cas des marches à valeurs dans $\mathbb{Z}^d$, à accroissements intégrables et de moyenne nulle : $\mathbb{E}[|Y_1|] < \infty$ et $\mathbb{E}[Y_1] = 0$. Nous allons nous restreindre au cas des marches au plus proche voisin.

**Définition 4.7** 1. *Une marche aléatoire :* $X_n = X_0 + \sum_{k=1}^{n} Y_k$, *à valeurs dans* $\mathbb{Z}$ *est une* **marche de Bernoulli de paramètre** $p$ *si* $\mathbb{P}(Y_n = 1) = p$ *et* $\mathbb{P}(Y_n = -1) = 1 - p$, *pour tout* $n \geq 1$, *où* $p \in [0, 1]$.
*Il est clair que* $(X_n)_{n \geq 0}$ *est une chaîne de Markov de probabilité de transition :*

$$\pi(x, x + 1) = p, \quad \pi(x, x - 1) = 1 - p, \quad \textit{pour tout } \; x \in \mathbb{Z}.$$

2. *Une marche aléatoire :* $X_n = X_0 + \sum_{k=1}^{n} Y_k$, *à valeurs dans* $\mathbb{Z}^d$ *est une* **marche aléatoire simple** *si* $\mathbb{P}(Y_n = \pm e_i) = \dfrac{1}{2d}$ *pour tout* $n \geq 1$ *et* $i \in \{1, 2, \cdots, d\}$, $\{e_1, e_2, \cdots, e_d\}$ *désignant la base canonique de* $\mathbb{Z}^d$.

On commence par la marche de Bernoulli en dimension 1.

**Proposition 4.8** *Soit* $(X_n)_{n \geq 0}$ *une marche aléatoire de Bernoulli sur* $\mathbb{Z}$ *de paramètre* $p$. *Alors :*

1. *tous les états sont récurrents si* $p = 1/2$ *;*
2. *tous les états sont transients si* $p \neq 1/2$.

**Preuve de la Proposition 4.8**. D'après la Proposition 4.6, on peut supposer $X_0 = 0$ et montrer que 0 est un état récurrent (resp. transient) si $p = 1/2$ (resp. $p \neq 1/2$ ).
Observons que $\mathbb{E}_0[Y_1] = p - (1 - p) = 2p - 1$. Ainsi si $p \neq 1/2$, la Proposition 4.6 permet d'affirmer que 0 est transient. Il suffit donc d'étudier le cas $p = 1/2$.
Pour la suite de la preuve, voir l'Exercice 4.9.

■

**Exercice 4.9** *Soit* $(X_n)_{n \geq 0}$ *une marche aléatoire de Bernoulli sur* $\mathbb{Z}$ *de paramètre* $p$. *Le but de l'exercice est de démontrer directement que* 0 *est récurrent si et seulement si* $p = 1/2$. *L'approche consiste à calculer* $\pi^n(0, 0)$, *puis à étudier la finitude de* $U(0, 0)$.

1. *Montrer :* $\pi^n(0, 0) = \mathbb{P}\Big(\sum_{k=1}^{n} Y_k = 0\Big)$.

2. *Posons* $\widetilde{Y}_n = (1 + Y_n)/2$. *En déduire :* $\pi^n(0, 0) = \mathbb{P}\Big(2 \sum_{k=1}^{n} \widetilde{Y}_k = n\Big)$, *puis* $\pi^{2n}(0, 0) = 0$ *et*
$$\pi^{2n}(0, 0) = \frac{(2n)!}{(n!)^2} p^n (1 - p)^n.$$

3. *Rappelons la formule de Stirling*
$$n! \underset{n \to \infty}{\sim} \sqrt{2\pi}\, n^{n+1/2} e^{-n}. \tag{4.31}$$

(a) *Supposons* $p = 1/2$. *Montrer que* $\pi^{2n}(0,0) \underset{n\to\infty}{\sim} \dfrac{1}{\sqrt{\pi n}}$ *puis* $U(0,0) = \infty$.

(b) *On suppose* $p \neq 1/2$. *Vérifier que* $a := 4p(1-p) < 1$. *En déduire* $\pi^{2n}(0,0) \underset{n\to\infty}{\sim} \dfrac{a^n}{\sqrt{\pi n}}$ *et ainsi le point* $0$ *est transient.*

**Exercice 4.10** *On considère la chaine de Markov à valeurs dans* $\mathbb{Z}$ *et de probabilité de transition* $\pi$ *:*

$$\pi(i, i+1) = p, \quad \pi(i, i-1) = q = 1 - p, \quad \forall i \in \mathbb{Z}.$$

1. *Vérifier :*
$$\sum_{n\geq 0} \binom{2n}{n} x^n = (1-4x)^{-1/2}, \quad \forall\, x \in \Big[0, \frac{1}{4}\Big[.$$
2. *Calculer* $U(x,x)$ *pour tout* $x$ *de* $\mathbb{Z}$ *, où* $U$ *désigne le potentiel.*
3. *En déduire que la chaine est récurrente si et seulement si* $p = 1/2$.
4. *Lorsque* $p \neq 1/2$, *calculer la probabilité de non-retour en* $0$, *sachant qu'elle est issue de* $0$.

**Proposition 4.11** *Soit* $X_n = X_0 + \sum_{k=1}^{n} Y_k$ *une marche aléatoire simple et dans* $\mathbb{Z}^d$ *avec* $d \geq 2$. *Alors* $(X_n, n \geq 0)$ *est une marche récurrente (i.e. tous les états sont récurrents) si et seulement si* $d = 2$.

Le lecteur pourra trouver une preuve complète de cette proposition dans [Durrett, 1996] (chapitre 3, section 2) ou [Feller, 1968] (chapitre XIV, section 7).
On revient au cas général de la chaîne $(X_n, n \geq 0)$ à valeurs dans $E$. On termine cette section en montant que le potentiel $U(x,y)$ peut s'exprimer à l'aide de $U(x,x)$.

**Proposition 4.12** *Soient* $x$ *et* $y$ *deux points distincts de* $E$. *Alors :*

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

*avec la convention :* $0 \times (+\infty) = 0$.

**Preuve.** Pour la preuve voir l'exercice 4.13. ■

**Exercice 4.13** *Soit* $x \neq y$.

1. *Montrer :*
$$U(x,y) = \mathbb{E}_x\Big[\Big(\sum_{n\geq\sigma_y} \mathbb{1}_{\{X_n = y\}}\Big)\mathbb{1}_{\{\sigma_y<\infty\}}\Big]$$
*puis :*
$$U(x,y) = \sum_{m\geq 0} \mathbb{P}_x(X_{m+\sigma_y} = y, \sigma_y < \infty).$$
2. *En appliquant la propriété de Markov forte, démontrer :*
$$U(x,y) = \sum_{m\geq 0} \mathbb{E}\Big(\pi^m(y,y)\mathbb{1}_{\{\sigma_y<\infty\}}\Big).$$

*3. Conclure :* $U(x,y) = \mathbb{P}_x(\sigma_y < \infty)U(y,y)$.

**Remarque 4.14** *Soit* $y$ *fixé. Puisque* $\mathbb{P}_x(\sigma_y < \infty) \leq 1$*, la fonction* $x \mapsto U(x,y)$ *atteint son maximum pour* $x = y$.

Donnons une première application de la Proposition 4.12.

**Proposition 4.15** *Si* $E$ *est fini, il existe au moins un état récurrent.*

**Preuve.** En effet
$$\sum_{y \in E} N_y = N_E = \infty.$$
On prend l'espérance de part et d'autre de cette égalité, il vient :
$$\sum_{y \in E} \mathbb{E}_x[N_y] = \sum_{y \in E} U(x,y) = \infty, \quad x \in E.$$
L'ensemble $E$ étant fini, la somme comporte un nombre fini de termes, il existe alors un $y \in E$ tel que $U(x,y) = \infty$.
Si $y = x$, alors $U(x,x) = \infty$, $x$ est alors récurrent.
Étudions à présent le cas où $y \neq x$. On applique la Proposition 4.12 :
$$\infty = U(x,y) = \mathbb{P}_x(\sigma_y < \infty)U(y,y) \leq U(y,y).$$
Par conséquent $U(y,y) = \infty$, ce qui implique que $y$ est récurrent. ∎

## 5 Chaînes de Markov irréductibles

On conserve les notations des sections précédentes, $(X_n; n \geq 0)$ désigne une chaîne de Markov, de probabilité de transition $\pi$ et à valeurs dans $E$. On commence par quelques définitions.

**Définition 5.1** 1. *Un point $x$ de $E$* **conduit à** *$y \in E$, si $x = y$ ou $x \neq y$ et $\mathbb{P}_x(\sigma_y < \infty) > 0$. On note $x \hookrightarrow y$. Ainsi dans le cas où $x \neq y$, $x$ conduit à $y$ si, partant de $x$, la chaîne visite $y$ avec une probabilité strictement positive.*

2. *Deux points $x$ et $y$ de $E$* **communiquent** *si $x \hookrightarrow y$ et $y \hookrightarrow x$. On note $x \leftrightarrows y$.*
3. *Une partie non-vide $A$ de $E$ est dite* **fermée** *ou* **absorbante** *si pour tout $x$ de $A$, la chaîne de Markov reste ensuite dans $A$ : $\mathbb{P}_x\big(X_n \in A, \forall\, n \geq 0\big) = 1$.*
4. *Un état $x$ est dit* **absorbant** *si $A = \{x\}$ est une partie absorbante, i.e. $\mathbb{P}_x\big(X_n = x, \forall\, n \geq 0\big) = 1$.*
5. *La chaîne est dite* **irréductible** *si la seule partie fermée non vide est $E$.*
6. *$\mathcal{R}$ désigne l'ensemble des états récurrents.*

Le lemme qui suit est technique. Il ne possède pas d'intérêt en lui-même. Toutefois il sera utilisé de nombreuses fois dans les preuves des propositions et théorèmes de la suite de ce chapitre.

**Lemme 5.2** *Soient $x$ et $y$ deux éléments distincts de $E$. Les cinq assertions suivantes sont équivalentes :*

1. $x \hookrightarrow y$ ;
2. $\mathbb{P}_x(\sigma_y < \infty) > 0$ ;
3. $\mathbb{E}_x[N_y] = U(x, y) > 0$ ;
4. $\mathbb{P}_x(\exists n \geq 0; X_n = y) > 0$ ;
5. $\exists\, n;\ \pi^n(x, y) = \mathbb{P}_x(X_n = y) > 0$.

*En particulier, $x$ ne conduit pas à $y$ si et seulement si :*

$$\mathbb{P}_x(\sigma_y < \infty) = 0 \Longleftrightarrow U(x, y) = 0 \Longleftrightarrow \mathbb{P}_x\big(X_n \neq y; \forall n \geq 0\big) = 0.$$

**Preuve.** Par définition $x \hookrightarrow y \Longleftrightarrow \mathbb{P}_x(\sigma_y < \infty) > 0$.
La v.a. $N_y$ est positive. Donc :

$$\mathbb{E}_x[N_y] = 0 \quad \Longleftrightarrow \quad N_y = 0 \ \ \mathbb{P}_x \text{ presque sûrement.}$$

Mais $N_y = \sum_{n \geq 0} \mathbb{1}_{\{X_n = y\}}$, d'où :

$$\mathbb{E}_x[N_y] = 0 \iff \mathbb{P}_x \ \text{ p.s. } \ X_n \neq y \text{ pour tout } \ n \geq 0 \quad \Longleftrightarrow \ \mathbb{P}_x\Big(\bigcap_{n \geq 0} A_n\Big) = 1.$$

où l'on a posé $A_n = \{X_n \neq y\}$.
Soit $B_n = A_n^c = \{X_n = y\}$ ; on a :

$$\mathbb{P}_x\Big(\bigcap_{n \geq 0} A_n\Big) = 1 \Longleftrightarrow \mathbb{P}_x\Big(\Big\{\bigcap_{n \geq 0} A_n\Big\}^c\Big) = \mathbb{P}_x\Big(\bigcup_{n \geq 0} B_n\Big) = 0 \Longleftrightarrow \mathbb{P}_x(B_n) = 0, \forall\, n \geq 0.$$

On déduit de l'égalité : $U(x,y) = \mathbb{E}_x[N_y]$ et (4.24) :

$$U(x,y) = 0 \Longleftrightarrow \pi^n(x,y) = \mathbb{P}_x(X_n = y) = \mathbb{P}_x(B_n) = 0, \quad \text{pour tout} \quad n \geq 0.$$

Ce qui prouve la série des cinq équivalences du Lemme 5.2, compte tenu du fait que $\mathbb{P}_x\Big(\bigcap_{n\geq 0} A_n\Big) = \mathbb{P}_x(\sigma_y = \infty)$. ■

**Proposition 5.3** *1. La relation "conduit à" est transitive : si $x \hookrightarrow y$ et $y \hookrightarrow z$ alors $x \hookrightarrow z$.*

*2. La relation "communiquer" est une relation d'équivalence.*

**Preuve.** **1)** On commence par montrer l'item 1.
Il est clair que si $x = y$ ou $y = z$, on a $x \hookrightarrow z$. Nous supposerons donc dans la suite $x \neq y$ et $y \neq z$.
Soit :

$$T := \inf\{n > \sigma_y;\, X_n = z\} = \sigma_y + \rho$$

où $\rho := \inf\{k \geq 1, X_{\sigma_y+k} = z\}$.
Il est clair que $\mathbb{P}_x(T < \infty) = \mathbb{P}_x(\sigma_y < \infty, \rho < \infty)$.
D'après la propriété forte de Markov, sur l'ensemble $\{\sigma_y < \infty\}$, le processus $(X_{\sigma_y+k})_{k\geq 0}$ est une chaine de Markov, issue de $y$ et de probabilité de transition $\pi$. Donc

$$\mathbb{P}_x(\rho < \infty | \sigma_y < \infty) = \mathbb{P}_y(\sigma_z < \infty).$$

On en déduit alors :

$$\begin{aligned} \mathbb{P}_x(T < \infty) &= \mathbb{E}_x\big(\mathbb{1}_{\{\sigma_y<\infty\}}\mathbb{1}_{\{\rho<\infty\}}\big) \\ &= \mathbb{E}_x\Big[\mathbb{1}_{\{\sigma_y<\infty\}}\mathbb{E}_x\big(\mathbb{1}_{\{\rho<\infty\}}\big|\sigma_y < \infty\big)\Big] \\ &= \mathbb{E}_x\big[\mathbb{1}_{\{\sigma_y<\infty\}}\mathbb{P}_y(\sigma_z < \infty)\big] \\ &= \mathbb{P}_x(\sigma_y < \infty)\mathbb{P}_y(\sigma_z < \infty). \end{aligned}$$

Sachant que $x \hookrightarrow y$ et $y \hookrightarrow z$ , alors $\mathbb{P}_x(\sigma_y < \infty) > 0$ et $\mathbb{P}_y(\sigma_z < \infty) > 0$. Par conséquent $\mathbb{P}_x(T < \infty) > 0$.
Puisque $\sigma_z$ est le premier instant plus grand que 1, où la chaîne est en $z$, alors $\mathbb{P}_x$ presque sûrement : $\sigma_z \leq T$. Donc : $0 < \mathbb{P}_x(T < \infty) \leq \mathbb{P}_x(\sigma_z < \infty)$. Ce qui prouve $x \hookrightarrow z$.
**2)** Par construction la relation "communiquer" est réflexive et symétrique. Montrons qu'elle est transitive. Supposons : $x \leftrightarrows y$ et $y \leftrightarrows z$, alors ($x \hookrightarrow y$ et $y \hookrightarrow z$) et ($z \hookrightarrow y$ et $y \hookrightarrow x$). D'après la transitivité de la relation $\hookrightarrow$, on a $x \hookrightarrow z$ et $z \hookrightarrow x$. D'où $x \leftrightarrows z$. ■

La proposition qui suit et en particulier l'assertion 1, joue un rôle important en pratique.

**Proposition 5.4** *Soient $x$ un état récurrent et $y \in E$, $y \neq x$. On suppose que $x$ conduit à $y$.*

*1. Alors $y$ est récurrent et $y \hookrightarrow x$. De plus :*

$$U(x,y) = U(y,x) = \infty \quad \textit{et} \;\; \mathbb{P}_x(\sigma_y < \infty) = \mathbb{P}_y(\sigma_x < \infty) = 1.$$

*2. Soit $E = \bigcup_{i\in I} E_i$, la partition de $E$ en classes d'équivalence pour la relation "communiquer". Alors chaque $E_i$ ne contient que des états récurrents, ou que des états transients.*

**Preuve.** **1)** On commence par le point 1. Soit $x$ un état récurrent conduisant à $y \in E, y \neq x$.
**a)** La première étape de la démonstration consiste à établir que $y$ est récurrent. Voir l'Exercice 5.5
**b)** Montrons que $y$ conduit à $x$, i.e. $\mathbb{P}_y(\sigma_x < \infty) > 0$. On raisonne par l'absurde, supposons : $\mathbb{P}_y(\sigma_x < \infty) = 0$.
D'après la propriété forte de Markov, sur $\{\sigma_y < \infty\}$, $(X_{\sigma_y+k})_{k\geq 0}$ est une chaine de Markov, issue de $y$ et de probabilité de transition $\pi$. Puisque $\mathbb{P}_y(\sigma_x < \infty) = 0$, cette chaine de Markov ne visite pas $x$. Donc, sur $\{\sigma_y < \infty\}$, $(X_k)$ visite un nombre fini de fois $x$. Mais $\mathbb{P}_x(\sigma_y < \infty) > 0$, ce qui génére une contracdiction, cf le Théorème 4.2.
**c)** Résumons les deux étapes précédentes, on a montré que si $x$ récurrent, $x \hookrightarrow y$ alors $y$ récurrent, $y \hookrightarrow x$, $U(x,y) = \infty$ et $\mathbb{P}_y(\sigma_x < \infty) = 1$. On choisit $x' = y$, $y' = x$. D'après ce qui précède, $\mathbb{P}_{y'}(\sigma_{x'} < \infty) = \mathbb{P}_x(\sigma_y < \infty) = 1$ et $U(x', y') = U(y,x) = \infty$. Ce qui achève la preuve du point 1. de la Proposition 5.4.
**2)** Il reste à montrer le point 2. Soit $E_i$ une classe d'équivalence pour la relation "communiquer". Si tous les éléments de $E_i$ sont transients, le résultat est établi. Sinon il existe un état récurrent $x$ appartenant à $E_i$. Soit $y \in E_i$, $y \neq x$, alors $x \hookrightarrow y$. D'après le 1., $y$ est récurrent.
■

**Exercice 5.5** *Soient $x$ un état récurrent et $y \in E$, $y \neq x$. On suppose que $x$ conduit à $y$. On souhaite montrer que $y$ est récurrent.*

1. *Montrer la série d'implication :*

$$\mathbb{P}_x(X_n = y) = 0,\ \forall\, n \quad \Rightarrow \quad \mathbb{P}_x\Big(\bigcup_{n\geq 0}\{X_n = y\}\Big) = 0 \quad \Rightarrow \quad \mathbb{P}_x(\sigma_y < \infty) = 0.$$

   *En déduire qu'il existe $n$ tel que $\pi^n(x,y) > 0$.*

2. *Montrer l'inégalité : $U(x,y) \geq \sum_{k\geq 0} \pi^{k+n}(x,y)$.*

3. *En observant que $\pi^{k+n}(x,y) = \big(\pi^k \times \pi^n\big)(x,y)$ déduire l'inégalité :*

$$\pi^{k+n}(x,y) \geq \pi^k(x,x)\pi^n(x,y).$$

   *puis :*

$$U(x,y) \geq \pi^n(x,y)U(x,x).$$

4. *En déduire : $U(x,y) = \infty$ puis $U(y,y) = \infty$.*

**Théorème 5.6** 1. *La chaîne est irréductible si et seulement si $U(x,y) > 0$, pour tout $x$ et $y$ de $E$.*

2. *Lorsque la chaîne est irréductible, il n'y a que deux cas possibles :*
   (a) *$U(x,y) < \infty$, pour tout $x$ et $y$ de $E$, tous les états sont transients, la chaîne est dite* **transiente**.
   (b) *$U(x,y) = \infty$ pour tout $x$ et $y$ de $E$, tous les états sont récurrents, la chaîne est dite* **récurrente**.

**Preuve.** **1)** Montrons le point 1. Supposons que $U(x,y) > 0$, pour tout $x$ et $y$ de $E$. Soient $A$ une partie fermée et $x \in A$. Il s'agit de montrer que $A = E$. Raisonnons par l'absurde : supposons qu'il existe $y \notin A$.
Remarquons :

$$U(x,y) = \mathbb{E}_x(N_y) > 0 \quad \Rightarrow \quad \mathbb{P}_x(\sigma_y < \infty) > 0 \quad \Leftrightarrow \quad x \hookrightarrow y. \tag{5.32}$$

Mais : $\{\sigma_y < \infty\} \subset \{\exists\, n, X_n \notin A\}$. Ce qui conduit à une contradiction puisque :

$$\mathbb{P}_x\big(\exists\, n, X_n \notin A\big) = 1 - \mathbb{P}_x\big(\forall\, n, X_n \in A\big) = 0 \geq \mathbb{P}_x(\sigma_y < \infty) > 0.$$

Établissons à présent la réciproque : on suppose que $E$ est la seule partie fermée. Soit $x$ un élément fixé de $E$. On note $F = \{y \in E;\, U(x,y) > 0\}$. Puisque $U(x,x) \geq \pi^0(x,x) = 1$, $x$ est élément de $F$. Montrons que $F$ est une partie fermée. On en déduira alors que $E = F = \{y \in E;\, U(x,y) > 0\}$.
Soit $y \in F$. La propriété (5.32) implique que $x \hookrightarrow y$.
Supposons que $\mathbb{P}_y(\forall\, n, X_n \in F) < 1$. Cela signifie que $\mathbb{P}_y(\exists\, n, X_n \notin F) > 0$ et donc qu'il existe $z \notin F$ tel que $\mathbb{P}_y(\sigma_z < \infty) > 0$. En d'autres termes, $y \hookrightarrow z$. On déduit de la Proposition 5.3 que $x \hookrightarrow z$. D'après la définition de $F$ et (5.32), on déduit $z \in F$. Ce qui conduit à une contradiction, donc $\mathbb{P}_y(\forall\, n, X_n \in F) = 1$.
**2)** On suppose que la chaîne est irréductible.
Supposons tous tous les états sont transients. On sait que $U(y,y) < \infty$ pour tout $y \in E$ (cf. le Théorème 4.4). On déduit directement de la Proposition 4.12 la finitude de $U(x,y)$ pour tout $y \neq x$.
On étudie à présent le cas où il existe un état récurrent $x_0$. Soit $F$ la classe d'équivalence, pour la relation communiquer, contenant $x_0$. Soit $x$ un élément de $F$. D'après la la Proposition 5.4 le point $x$ est récurrent.
Soit $z \notin F$. Si $z$ est transient, d'après la Proposition 5.4, $x$ ne conduit pas à $z$, donc $\mathbb{P}_x(\sigma_z < \infty) = 0$. Si $z$ est récurrent, si $x \hookrightarrow z$, la Proposition 5.4 implique que $z \hookrightarrow x$. Donc $z$ et $x$ communiquent, ce qui est exclu puisque $z \notin F$. En conclusion, pour tout $z$ n'appartenant pas à $F$, $\mathbb{P}_x(\sigma_z < \infty) = 0$. Ce qui implique $\mathbb{P}_x\big(\exists\, n, X_n \notin F) = 0$. On a donc montré que $F$ est fermée. Puisque $F$ est non-vide, et la chaine étant irréductible alors $F = E$.
Tous les états sont récurrents, donc $U(x,x) = \infty$. De plus la Proposition 4.12 implique que pour tout $x$ et $y$ de $E$, $U(x,y) = \infty$. ■

**Exercice 5.7** *Déterminer pour chaque chaine de Markov associée aux matrices de transition* $\pi_1, \cdots, \pi_4$ *:*

$$\pi_1 := \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/4 & 1/4 \\ 0 & 1/3 & 2/3 \end{pmatrix} \quad \pi_2 := \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\pi_3 := \begin{pmatrix} 0 & 0 & 1/2 & 1/2 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad \pi_4 := \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 \\ 1/4 & 1/4 & 0 & 0 & 1/2 \end{pmatrix}$$

*les états récurrents, transients et les ensembles fermés.*

# 6 Mesures invariantes et convergence en temps long

## 6.1 Mesures invariantes

**Définition 6.1** *Soit $\nu$ une mesure positive sur $E$, non identiquement nulle. On dit que $\nu$ est une* **mesure stationnaire** *ou* **invariante** *associée à la probabilité de transition $\pi$ si*

$$\nu\pi = \nu. \tag{6.33}$$

*Lorsque $\nu(E) = 1$, on dit que $\nu$ est une* **probabilité stationnaire** *ou* **invariante**.

**Remarque 6.2**

1. *$\nu = \big(\nu(x), x \in E\big)$ est une mesure invariante si $\nu(x)$ est un réel positif ou nul pour tout $x$ de $E$, $\exists\, x_0 \in E$ tel que $\nu(x_0) > 0$ et*

$$\nu(x) = \sum_{y\in E} \nu(y)\pi(y,x), \qquad \text{pour tout } \ x \in E. \tag{6.34}$$

*Lorsque $E$ a $r$ éléments, alors (6.34) est un système de $r$ équations linéaires avec $r$ inconnues.*

2. *Lorsque $E$ est un ensemble fini. Alors la condition (6.33) est équivalente à :$\pi^\star \nu^\star = \nu^\star$. Ce qui signifie que $\nu^\star$ est un vecteur propre de $\pi^\star$ associé à la valeur propre 1 et dont toutes les coordonnées sont positives ou nulles.*
3. *Si $\lambda > 0$ et $\nu$ est une mesure stationnaire alors $\lambda\nu$ est une mesure stationnaire. Si $\nu$ est de masse totale finie, i.e. $\displaystyle\sum_{x\in E} \nu(x) < \infty$, alors $\rho = \displaystyle\sum_{x\in E} \nu(x)$ est l'unique constante strictement positive telle que $\dfrac{1}{\rho}\nu$ est une probabilité invariante.*
4. *Soit $(X_n)$ une chaîne de Markov de probabilité de transition $\pi$ et de probabilité stationnaire $\nu$. Alors :*

$$\nu\pi^n = \nu, \quad \text{pour tout } \ n \geq 1. \tag{6.35}$$

*Cette relation se montre aisément, en raisonnant par récurrence sur $n$.*

*L'interprétation probabiliste de cette identité est la suivante : si $X_0$ a pour loi $\nu$, alors $X_n$ a pour loi $\nu$, pour tout $n \geq 0$. Cette propriété peut se généraliser de la manière suivante : le vecteur aléatoire $X_k, \cdots, X_{k+n}$ a même loi que $X_0, \cdots, X_n$, pour tout $n, k \geq 0$. On dit alors que $(X_n)$ est un* **processus stationnaire** .

*En effet, pour montrer que $(X_n)$ est stationnaire, il suffit d'utiliser la propriété de Markov, et le fait que la v.a. $X_k$ a pour loi $\nu$.*

5. *Soit $(X_n)$ une marche aléatoire à valeurs dans $\mathbb{Z}^d$, alors la mesure uniforme $\lambda$ sur $\mathbb{Z}^d$ (i.e. $\lambda(x) = 1$, pour tout $x \in \mathbb{Z}^d$ ) est une mesure invariante. On conserve les notations de l'Exemple 2.5 : $\pi(x,y) = \nu(y-x)$, $x, y \in \mathbb{Z}^d$. On a :*

$$\lambda\,\pi(x) = \sum_{y\in\mathbb{Z}^d} \lambda(y)\pi(y,x) = \sum_{y\in\mathbb{Z}^d} \nu(y-x) = \sum_{y'\in\mathbb{Z}^d} \nu(y') = 1 = \lambda(x).$$

*Ce qui montre que $\lambda$ est bien une mesure invariante. Remarquons de plus que $\lambda$ n'est pas une mesure finie.*

6. *Il n'y a pas a priori unicité de la mesure invariante.*
*a) Considérons en effet une chaîne de Markov avec deux états absorbants* $a$ *et* $b$ *distincts. Alors* $\delta_a$ *et* $\delta_b$ *sont deux probabilités invariantes.*
*b) Examinons à présent le cas particulier de la marche de Bernoulli sur* $\mathbb{Z}$ *de paramètre* $p$ *(cf. la Définition 4.7). Soit* $\nu_1 = \big(\nu_1(n)\big)_{n\in\mathbb{Z}}$ *la mesure positive :* $\nu_1(n) = \Big(\dfrac{p}{1-p}\Big)^n, n \in \mathbb{Z}$*. Remarquons que* $\nu_1$ *est la mesure uniforme lorsque* $p = 1/2$*. Montrons que* $\nu_1$ *est une mesure stationnaire. En effet :*

$$\begin{aligned}
\nu_1\,\pi(j) &= \sum_{i\in\mathbb{Z}} \nu_1(i)\pi(i,j) = \nu_1(j-1)\pi(j-1,j) + \mu(j+1)\pi(j+1,j)\\
&= \Big(\frac{p}{1-p}\Big)^{j-1} p + \Big(\frac{p}{1-p}\Big)^{j+1}(1-p)\\
&= \Big(\frac{p}{1-p}\Big)^{j}\Big(\frac{1-p}{p}p + \frac{p}{1-p}(1-p)\Big) = \Big(\frac{p}{1-p}\Big)^{j} = \nu_1(j).
\end{aligned}$$

*Soit* $\nu_2 = \big(\nu_2(n)\big)_{n\in\mathbb{Z}}$ *la mesure positive uniforme :* $\nu_2(n) = 1$*, pour* $n$*. On montre facilement que* $\nu_2$ *est stationnaire :*

$$\begin{aligned}
\nu_2\,\pi(j) &= \sum_{i\in\mathbb{Z}} \nu_2(i)\pi(i,j) = \nu_2(j-1)\pi(j-1,j) + \mu(j+1)\pi(j+1,j)\\
&= p + (1-p)\\
&= 1 = \nu_2(j).
\end{aligned}$$

*Ainsi si* $p \neq 1/2$*, la marche aléatoire de Bernoulli sur* $\mathbb{Z}$ *possède deux mesures invariantes* $\nu_1$ *et* $\nu_2$*. De plus ces deux mesures invariantes sont de masse totale infinie. On peut remarquer de plus que la chaîne est transiente.*

7. *Indiquons un procédé pour obtenir une probabilité invariante. Supposons que* $E$ *soit fini et :*

$$\lim_{n\to\infty} \pi^n(i,j) = \nu(j), \quad \textit{pour tout } \; i,j \in E. \tag{6.36}$$

*Il est évident que* $\nu(j) \geq 0$*, pour tout* $j$ *de* $E$*. Puisque* $\pi^n(i,\cdot)$ *est une probabilité :*

$$\sum_{k\in E} \pi^n(i,k) = 1, \quad i \in E.$$

*La somme comportant un nombre fini de termes, en passant à la limite,* $n \to \infty$*, on obtient :* $\displaystyle\sum_{k\in E} \nu(k) = 1$*. Ce qui signifie que* $\nu$ *est une probabilité sur* $E$*. Montrons de plus que* $\nu$ *est invariante. On part de l'égalité :*

$$\pi^{n+1}(i,j) = \big(\pi^n\,\pi\big)(i,j) = \sum_{k\in E} \pi^n(i,k)\pi(k,j),$$

*et on prend la limite, lorsque* $n$ *tend vers l'infini :*

$$\nu(j) = \sum_{k\in E} \nu(k)\pi(k,j) = (\nu\,\pi)(j), \quad \textit{pour tout } \; j \in E.$$

*D'où l'égalité :* $\nu = \nu\,\pi$*.*
*L'interprétation probabiliste de (6.36) est la suivante :* $X_n$ *converge en loi, sous* $\mathbb{P}_i$*, lorsque* $n \to \infty$*, vers une distribution* $\nu$ *indépendante de* $i$*.*

**Exercice 6.3** *Déterminer les mesures invariantes pour les matrices de transition* $\pi_1, \cdots, \pi_4$ *introduites dans l'Exercice* (5.7).

Lorsque la chaîne admet au moins un point récurrent, il est alors possible de construire une mesure invariante comme le montre le résultat suivant.

**Théorème 6.4** *Soient* $x$ *un état récurrent et* $T = \sigma_x = \inf\{n \geq 1; X_n = x\}$. *Alors :*

$$\mu_x(y) = \mathbb{E}_x\Big[\sum_{n=0}^{T-1} \mathbb{1}_{\{X_n=y\}}\Big],$$

*définit une mesure stationnaire.*
*De plus* $\mu_x(x) = 1$.

**Preuve.** Le point $x$ étant fixé, pour alléger les notations nous noterons dans la suite $\mu = \mu_x$. Il est clair que $\mu(y) \geq 0$ pour tout $y$ de $E$.
Puisque $x$ est récurrent, le temps d'arrêt $\sigma_x$ est fini sous $\mathbb{P}_x$ ; donc la somme définissant $\mu(y)$ est finie, mais l'espérance peut être infinie.
**1)** Montrons que $\mu$ vérifie : $(\mu\,\pi)(z) = \mu(z)$, pour tout $z \in E$. D'après la définition de $\mu$, on a :

$$(\mu\,\pi)(z) = \sum_{y\in E} \mu(y)\pi(y,z) = \sum_{y\in E} \mathbb{E}_x\Big[\sum_{n\geq 0} \mathbb{1}_{\{n\leq T-1,\, X_n=y\}}\pi(y,z)\Big].$$

Puisque les quantités qui interviennent sont positives, on peut permuter l'espérance et la somme en $n$ :

$$(\mu\,\pi)(z) = \sum_{y\in E}\Big\{\sum_{n\geq 0}\mathbb{E}_x\Big[\mathbb{1}_{\{n\leq T-1,\, X_n=y\}}\pi(y,z)\Big]\Big\}.$$

Mais $\{n \leq T-1\} = \{n < T\}$ et $\{T \leq n\} = \{n < T\}^c \in \mathcal{F}_n$, donc l'événement $\{n \leq T-1,\, X_n = y\}$ appartient à $\mathcal{F}_n$. De plus sur $\{X_n = y\}$, on a : $\pi(y,z) = \mathbb{P}_y(X_1 = z) = \mathbb{P}_{X_n}(X_1 = z)$. Une application directe de la propriété de Markov conduit à :

$$\mathbb{E}_x\Big[\mathbb{1}_{\{X_{n+1}=z\}}|\mathcal{F}_n\Big] = \mathbb{P}_{X_n}(X_1 = z),$$

D'où :

$$\begin{aligned}
(\mu\,\pi)(z) &= \sum_{y\in E}\Big\{\sum_{n\geq 0}\mathbb{E}_x\Big[\mathbb{1}_{\{n\leq T-1,\, X_n=y,\, X_{n+1}=z\}}\Big]\Big\}\\
&= \sum_{n\geq 0}\mathbb{E}_x\Big[\sum_{y\in E}\mathbb{1}_{\{n\leq T-1,\, X_n=y,\, X_{n+1}=z\}}\Big]\\
&= \mathbb{E}_x\Big[\sum_{n\geq 0}\mathbb{1}_{\{n\leq T-1,\, X_{n+1}=z\}}\Big].
\end{aligned}$$

Puisque $x$ est récurrent, sous $\mathbb{P}_x$, le temps d'arrêt $T$ est fini. On pose $m = n+1$, il vient :

$$(\mu\,\pi)(z) = \mathbb{E}_x\Big[\sum_{m=1}^{T}\mathbb{1}_{\{X_m=z\}}\Big] = \mathbb{E}_x\Big[\sum_{m=0}^{T-1}\mathbb{1}_{\{X_m=z\}} + \xi\Big],$$

avec $\xi = \mathbb{1}_{\{X_T=z\}} - \mathbb{1}_{\{X_0=z\}}$. On utilise à nouveau le fait que $x$ est récurrent, sous $\mathbb{P}_x$ : $\xi = \mathbb{1}_{\{x=z\}} - \mathbb{1}_{\{x=z\}} = 0$. Ce qui prouve que $(\mu\,\pi)(z) = \mu(z)$.

**2)** Par définition de $T$, pour tout $n$ tel que $1 \leq n < T$, on a : $X_n \neq x$. Mais $\mathbb{P}_x(X_0 = x) = 1$, donc $\mu(x) = 1$. $\mu$ est donc non identiquement nulle.
**3)** Montrons que $\mu(y)$ est fini pour tout $y \in E$. Puisque $\mu\,\pi = \mu$, d'après le point 4. de la Remarque 6.2, on a : $\mu\,\pi^n = \mu$. En particulier :

$$1 = \mu(x) = (\mu\,\pi^n)(x) = \sum_{z \in E} \mu(z)\pi^n(z,x) \geq \mu(y)\pi^n(y,x), \quad n \geq 0,$$

car la somme ne comporte que des termes positifs.

1. S'il existe un entier $n \geq 0$, tel que $\pi^n(y,x) > 0$, alors : $\mu(y) \leq \dfrac{1}{\pi^n(y,x)} < \infty$.
2. Sinon, d'après le Lemme 5.2, $y$ ne conduit pas à $x$. Mais $x$ etant récurrent, si $x \hookrightarrow y$, d'après la Proposition 5.4, $y \hookrightarrow x$, ce qui est impossible. Par conséquent $x$ ne conduit pas à $y$ donc $\mathbb{P}_x(\sigma_y < \infty) = 0$. Par conséquent, sous $\mathbb{P}_x$, $\displaystyle\sum_{n=0}^{T-1} \mathbb{1}_{\{X_n = y\}} = 0$. Ce qui implique que $\mu(y) = 0$.

■

Nous énonçons une réciproque partielle du Théorème 6.4.

**Proposition 6.5** *Soit $\mu$ une probabilité invariante. Si $\mu(x) > 0$ alors $x$ est récurrent.*

**Remarque 6.6** *Si $\mu$ est une mesure invariante de masse totale infinie, alors la condition $\mu(x) > 0$ n'implique pas nécessairement que $x$ est récurrent. Il suffit de considérer la marche aléatoire de Bernoulli sur $\mathbb{Z}$ de paramètre $p \neq 1/2$. On a montré (cf. le point 5. de la Remarque 6.2) que la mesure uniforme $\lambda$ sur $\mathbb{Z}$ est une mesure invariante. À l'évidence $\lambda(x) = 1 > 0$, pour tout $x$ de $\mathbb{Z}$, bien que tout état $x$ est transient (cf. la Proposition 4.8).*

**Preuve de la Proposition 6.5.**
Soit $x$ tel que $\mu(x) > 0$. D'après (4.24), on a :

$$\begin{aligned}
\sum_{y \in E} \mu(y) U(y,x) &= \sum_{y \in E} \mu(y) \Big( \sum_{n \geq 0} \pi^n(y,x) \Big) = \sum_{n \geq 0} \Big( \sum_{y \in E} \mu(y)\pi^n(y,x) \Big) \\
&= \sum_{n \geq 0} \big(\mu\,\pi^n\big)(x) = \sum_{n \geq 0} \mu(x) = \infty.
\end{aligned}$$

Rappelons que d'après la Proposition 4.12, $U(y,x) \leq U(x,x)$. Par conséquent :

$$\infty = \sum_{y \in E} \mu(y) U(y,x) \leq \sum_{y \in E} \mu(y) U(x,x) = \Big( \sum_{y \in E} \mu(y) \Big) U(x,x) = U(x,x).$$

Donc $U(x,x) = \infty$, ce qui signifie que $x$ est récurrent. ■

**Théorème 6.7** *Si la chaîne de Markov est récurrente et irréductible alors :*

1. *il existe une mesure invariante $\lambda$ strictement positive ;*
2. *toute mesure positive $\nu$ vérifiant $\nu\pi = \nu$ est proportionelle à $\lambda$. En particulier, si elle n'est pas identiquement nulle, elle est strictement positive.*

**Preuve.** Soit $\lambda$ une mesure positive vérifiant $\lambda\pi = \lambda$. Remarquons que nous ne supposons pas $\lambda$ est invariante pour la chaîne de Markov, i.e. cette mesure peut être identiquement nulle.
**1)** Considérons un élément fixé $a$ de $E$. On pose

$$q_n(y) = \mathbb{P}_a(n < \sigma_a, X_n = y), \quad y \in E,\, n \geq 0.$$

Montrons :

$$q_n\pi \geq q_{n+1}, \quad \text{sur} \;\; E - \{a\}. \tag{6.37}$$

En effet, pour tout $y \neq a$, on a :

$$(q_n\pi)(y) = \sum_{z\in E} q_n(z)\pi(z,y) = \sum_{z\in E} \mathbb{E}_a\big[\mathbb{1}_{\{n<\sigma_a,\, X_n=z\}}\big]\pi(z,y).$$

Puisque $\{n < \sigma_a\} = \{\sigma_a \leq n\}^c \in \mathcal{A}_n$, une application de la propriété de Markov conduit à :

$$\mathbb{E}_a\big[\mathbb{1}_{\{n<\sigma_a, X_n=z\}}\big]\pi(z,y) = \mathbb{E}_a\big[\mathbb{1}_{\{n<\sigma_a,\, X_n=z,\, X_{n+1}=y\}}\big].$$

Par conséquent :

$$(q_n\pi)(y) = \mathbb{P}_a\big(n < \sigma_a,\, X_{n+1} = y\big) \geq \mathbb{P}_a\big(n+1 < \sigma_a,\, X_{n+1} = y\big) = q_{n+1}(y).$$

**2)** Montrons par récurrence sur $m \geq 0$ :

$$\lambda \geq \lambda(a)\Big(\sum_{n=0}^{m} q_n\Big). \tag{6.38}$$

Si $m = 0$, on a $\sum_{n=0}^{m} q_n = q_0$ et

$$q_0(y) = \mathbb{P}_a(X_0 = y) = \delta_a(y), \quad y \in E. \tag{6.39}$$

En particulier, on a : $\lambda(y) \geq \lambda(a)\mathbb{1}_{\{y=a\}}$. D'où (6.38) (avec $n = 0$).
Supposons à présent (6.38), et montrons que cette identité est encore réalisée lorsque $m$ est remplacé par $m+1$. Sachant que $\lambda = \lambda\pi$, on a :

$$\begin{aligned}\lambda(y) &= \sum_{z\in E}\lambda(z)\pi(z,y) \geq \lambda(a)\sum_{z\in E}\Big(\sum_{n=0}^{m} q_n(z)\Big)\pi(z,y)\\ &\geq \lambda(a)\Big(\sum_{n=0}^{m}\Big(\sum_{z\in E} q_n(z)\pi(z,y)\Big) = \lambda(a)\sum_{n=0}^{m}(q_n\pi)(y).\end{aligned}$$

Soit $y \neq a$. On déduit de (6.37) et (6.39) et :

$$\lambda(y) \geq \lambda(a)\sum_{n=0}^{m} q_{n+1}(y) = \lambda(a)\sum_{n=0}^{m+1} q_n(y).$$

D'après la définition de $\sigma_a$, $q_n(a) = 0$ pour tout $n \geq 1$. Par conséquent, l'inégalité précédente est encore réalisée lorsque $y = a$.

**3)** L'état $a$ étant récurrent, la mesure $\mu_a$ définie par le Théorème 6.4 est invariante. Vérifions :

$$\lambda \geq \lambda(a)\mu_a. \tag{6.40}$$

Soit $y \in E$. D'après (6.38) : $\lambda(y) \geq \lambda(a)\Big(\sum_{n=0}^{m} q_n(y)\Big)$. On fait tendre $m \to \infty$, il vient :

$$\lambda(y) \geq \lambda(a)\Big(\sum_{n\geq 0} q_n(y)\Big). \tag{6.41}$$

En utilisant la définition de $\mu_a$ et le théorème de Fubini on a :

$$\begin{aligned}
\mu_a(y) &= \mathbb{E}_a\Big[\sum_{n=0}^{\sigma_a-1} \mathbb{1}_{\{X_n=y\}}\Big] = \mathbb{E}_a\Big[\sum_{n\geq 0} \mathbb{1}_{\{n<\sigma_a,\, X_n=y\}}\Big] \\
&= \sum_{n\geq 0} \mathbb{P}_a(n<\sigma_a, X_n=y) = \sum_{n\geq 0} q_n(y).
\end{aligned}$$

Il est clair que (6.41) implique (6.40).
**4)** La chaîne $(X_n)$ est récurrente et irréductible, donc $\{y \in E;\ a \hookrightarrow y\} = E$. On a montré que $\{y \in E;\ a \hookrightarrow y\} = \{y \in E;\ \mu_a(y) > 0\}$. Donc $\mu_a$ est une mesure invariante qui ne s'annule pas sur $E$. On déduit de (6.40) que s'il existe $a$ tel que $\lambda(a) > 0$, alors $\lambda(y) > 0$ pour tout $y \in E$.
**5)** On suppose la mesure $\lambda$ non-identiquement nulle, $\lambda$ est alors une mesure invariante. On choisit $a$ tel que $\lambda(a) > 0$. Posons $\lambda' := \lambda - \lambda(a)\mu_a$. D'après l'inégalité (6.40), $\lambda'$ est une mesure positive. De plus elle vérifie $\lambda'\pi = \lambda'$, en effet :

$$\lambda'\pi = (\lambda - \lambda(a)\mu_a)\pi = \lambda\pi - \lambda(a)\mu_a\pi = \lambda - \lambda(a)\mu_a = \lambda'.$$

Par ailleurs $\lambda'(a) = \lambda(a) - \lambda(a)\mu_a(a) = \lambda(a) - \lambda(a) = 0$. Appliquons le résultat de l'étape précédente en changeant $\lambda$ en $\lambda'$ : puisque $\lambda'$ s'annule, $\lambda'$ est identiquement nulle. Il est maintenant clair que $\lambda = \lambda(a)\mu_a$. ■

## 6.2 Probabilité invariante et récurrence positive

Rappelons que $x \in E$ est récurrent si $\mathbb{P}_x(\sigma_x < \infty) = 1$ où $\sigma_x$ est le premier temps de retour en $x$ de la chaîne de Markov, et que nous avons démontré dans le Théorème 4.2 que cette propriété est équivalente à $\mathbb{P}_x(X_n = x$ pour un nombre infini de $n \in \mathbb{N}) = 1$.

**Définition 6.8** *Soit $x \in E$ un état récurrent. Si de plus*

$$m(x) = \mathbb{E}_x(\sigma_x) < \infty,$$

*on dit que le point $x$ est **récurrent positif**. Sinon, on dit que $x$ est **récurrent nul**.*

**Proposition 6.9** *Si la chaîne de Markov est récurrente, alors les trois conditions suivantes sont équivalentes :*

1. *Tout état $x \in E$ est récurrent positif.*
2. *Il existe un état $x \in E$ récurrent positif.*

*3. La chaîne de Markov a une probabilité invariante $\lambda$.*

*De plus, lorsque l'une de ces conditions est satisfaite,*

$$m(x) = \frac{1}{\lambda(x)}, \quad \forall x \in E. \tag{6.42}$$

**Preuve.** Le point 1. implique évidemment le point 2.

Supposons que le point 2. est satisfait. Puisque $x$ est récurrent positif, il est récurrent et donc, d'après le théorème 6.4, $\mu_x$ est une mesure invariante. Or, d'après le théorème de Fubini,

$$\sum_{y \in E} \mu_x(y) = \sum_{y \in E} \mathbb{E}_x \left[ \sum_{n=0}^{\sigma_x - 1} \mathbb{1}_{\{X_n = y\}} \right] = \mathbb{E}_x(\sigma_x) = m(x) < \infty.$$

Donc $\lambda(y) = \mu_x(y)/m(x)$ est une probabilité invariante.

Supposons enfin que le point 3. est vérifié. Soit $x \in E$. Puisque $\sum_{y \in E} \lambda(x) = 1$, on a $\lambda(x) = \sum_{y \in E} \lambda(y)\pi^n(y,x)$ pour tout $n \geq 1$, et puisque la chaîne est irréductible, cette quantité est strictement positive pour un certain $n \geq 1$. Donc $\lambda(x) > 0$, et on déduit de la proposition 6.5 que $x$ esr récurrent. Posons $\hat{\mu}_x(y) = \lambda(y)/\lambda(x)$. Puisque $\hat{\mu}_x(x) = 1$ et par unicité à une constante multiplicative près des mesures invariantes dans le cas récurrent et irréductible (théorème 6.7), on en déduit que $\hat{\mu}_x = \mu_x$. Par conséquent,

$$m(x) = \sum_{y \in E} \mu_x(y) = \frac{\sum_{y \in E} \lambda(y)}{\lambda(x)} = \frac{1}{\lambda(x)} < \infty.$$

L'état $x$ est donc récurrent psotif et on a également démontré (6.5). ■

## 6.3 Convergence vers la loi invariante

On vient de voir dans quel cas il existe une probabilité invariante. On va maintenant s'intéresser à la réciproque de la remarque 6.2 7., c'est-à-dire à la convergence en loi de $X_n$ vers cette probabilité invariante. Cette convergence n'a pas toujours lieu, comme le montre l'exemple suivant.

**Exemple 6.10** *On considère la chaîne de Markov sur $E = \{0, 1\}$ de matrice de transition*

$$\pi = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

*On vérifie que $\pi^2 = Id$ (la matrice identité), et donc $\pi^{2m} = Id$ et $\pi^{2m+1} = \pi$ pour tout $m \in \mathbb{N}$. On n'a donc pas convergence de $\pi^n(x,y)$ lorsque $n \to +\infty$, pour aucun $x, y \in \{0, 1\}$.*

**Définition 6.11** *On dit qu'un état $x \in E$ est apériodique si $\pi^n(x,x) > 0$ pour tout $n$ suffisamment grand.*

**Exercice 6.12** *En utilisant la relation $\pi^{n+m}(x,x) \geq \pi^n(x,x)\pi^m(x,x)$ (que l'on démontrera), démontrer que $x$ est apériodique si et seulement si l'ensemble $\{n \in \mathbb{N} : \pi^n(x,x) > 0\}$ n'a pas de diviseur commun autre que 1.*

**Proposition 6.13** *Si la chaîne est irréductible et a un état apériodique $x \in E$, alors pour tous états $y, z \in E$, $\pi^n(y, z) > 0$ pour $n$ suffisamment grand. En particulier, tous les états sont apériodiques. On dit alors que la chaîne (ou la probabilité de transition $\pi$ associée) est apériodique.*

**Exercice 6.14** *En vous inspirant de l'exercice 6.12, prouver la proposition 6.13.*

**Théorème 6.15** *Supposons que la probabilité de transition $\pi = (\pi(x, y), x, y \in E)$ est irréductible et apériodique et qu'elle admet un probabilité invariante $\lambda$ (c'est-à-dire qu'elle est positive récurrente). Soit $\mu$ une probabilité quelconque sur $E$ et $(X_n, n \geq 0)$ une chaîne de Markov de loi initiale $\mu$ et de probabilité de transition $\pi$. Alors*

$$\mathbb{P}(X_n = y) \to \lambda(y) \quad \text{quand } n \to +\infty \text{ pour tout } y \in E.$$

*En particulier,*

$$\pi^n(x, y) \to \lambda(y) \quad \text{quand } n \to +\infty \text{ pour tout } x, y \in E.$$

**Preuve.** La preuve repose sur un argument de couplage. Soit $(Y_n, n \geq 0)$ une chaîne de Markov de loi initiale $\lambda$ et de probabilité de transition $\pi$ indépendante de $(X_n, n \geq 0)$. En particulier, la loi de $Y_n$ est $\lambda$ pour tout $n \geq 0$. On fixe $y \in E$ quelconque et on pose

$$T = \inf\{n \geq 0 : X_n = Y_n = y\}.$$

**1)** Montrons que $\mathbb{P}(T < \infty) = 1$. Le processus $W_n = (X_n, Y_n)$ et une chaîne de Markov sur $E \times E$ de probabilités de transition $\tilde{\pi}((x_1, y_1), (x_2, y_2)) = \pi(x_1, x_2)\pi(y_1, y_2)$ et de loi initiale $\mu(x, y) = \mu(x)\lambda(y)$ (voir l'exercice 6.16).

Puisque $\pi$ est apériodique, pour tout $x_1, x_2, y_1, y_2 \in E$, on a $\tilde{\pi}^n((x_1, y_1), (x_2, y_2)) = \pi^n(x_1, x_2)\pi^n(y_1, y_2) > 0$ pour tout $n$ suffisamment grand. Donc $\tilde{\pi}$ est irréductible. De plus $\tilde{\pi}$ a pour probabilité invariante $\tilde{\lambda}(x, y) = \lambda(x)\lambda(y)$. Donc, d'après la proposition 6.9, $\tilde{\pi}$ est récurrente positive. Or $T$ est le premier temps d'atteinte de $(y, y)$ par la chaîne de Markov $(W_n, n \geq 0)$, donc $\mathbb{P}(T < \infty) = 1$ d'après le théorème 4.2.

**2)** Pour tout $n \geq 0$, on définit

$$Z_n = \begin{cases} X_n & \text{si } n < T \\ Y_n & \text{si } n \geq T. \end{cases}$$

Montrons que $Z_n$ est une chaîne de Markov de probabilité de transition $\pi$ et de loi initiale $\mu$. La loi initiale est évidente puisque $Z_0 = X_0$. D'après la propriété de Markov forte (proposition 3.8 6.) à $(W_n, n \geq 0)$ au temps $T$, le processus $(X_{T+n}, Y_{T_n}, n \geq 0)$ est une chaîne de Markov de loi initial $(y, y)$ et de probabilité de transition $\tilde{\pi}$, indépendante de $(X_0, Y_0), (X_1, Y_1), \ldots, (X_T, Y_t)$. Par symmétrie, on peut remplacer le processus $(X_{T+n}, Y_{T_n}, n \geq 0)$ par $(Y_{T+n}, X_{T_n}, n \geq 0)$ dans la phrase précédente. En particulier, $(Z_{T+n}, n \geq 0) = (Y_{T+n}, n \geq 0)$ est une chaîne de Markov issue de $y$ et de transition $\pi$, indépendante de $Z_0 = X_0, Z_1 = X_1, \ldots, Z_T = X_T$. D'après la réciproque de la proposition 3.8 5., on en déduit que $(Z_n, n \geq 0)$ est un chaîne de Markov de probabilité de transition $\pi$.

**3)** On a

$$\begin{aligned}\mathbb{P}(X_n = x) = \mathbb{P}(Z_n = x) &= \mathbb{P}(X_n = x,\ n < T) + \mathbb{P}(Y_n = x,\ n \geq T) \\ &= \mathbb{P}(Y_n = x) + \mathbb{P}(X_n = x,\ n < T) - \mathbb{P}(Y_n = x,\ n < T) \\ &= \lambda(x) + \mathbb{P}(X_n = x,\ n < T) - \mathbb{P}(Y_n = x,\ n < T).\end{aligned}$$

Donc

$$|\mathbb{P}(X_n = x) - \lambda(x)| = |\mathbb{P}(X_n = x,\ n < T) - \mathbb{P}(Y_n = x,\ n < T)| \leq \mathbb{P}(n < T)$$

et $\mathbb{P}(n < T) \to 0$ quand $n \to +\infty$. ■

**Exercice 6.16** *Démontrer que le processus $W_n = (X_n, Y_n)$ et une chaîne de Markov sur $E \times E$ de probabilités de transition $\tilde{\pi}((x_1, y_1), (x_2, y_2)) = \pi(x_1, x_2)\pi(y_1, y_2)$ et de loi initiale $\mu(x, y) = \mu(x)\lambda(y)$ (avec les notations de la preuve précédente).*

**Exercice 6.17** *On lance un dé plusieurs fois de suite. Soit $X_n$ la somme des $n$ premiers résultats. Calculer*

$$\lim_{n \to +\infty} \mathbb{P}(X_n \text{ est un multiple de } 13).$$

**Exercice 6.18** *Chaque matin un étudiant prend un des trois livres qu'il possède sur son étagère. La probabilité qu'il choisisse le livre $i$ est $\alpha_i$, où $0 < \alpha_1 < 1$ pour $i = 1, 2, 3$, et les choix sont indépendants d'un jour à l'autre. Le soir il remet son livre à gauche de l'étagère. On note $p_n$ la probabilité que, le matin du jour $n$, les livres soient dans l'ordre 1,2,3 de gauche à droite sur l'étagère. Montrer que $p_n$ converge lorsque $n \to +\infty$ vers une limite indépendante de la position initiale des livres sur l'étagère et calculer cette limite.*

**Exercice 6.19** ***(Processus de renouvellement)*** *Soit $Y_1, Y_2, \ldots$ des v.a. i.i.d. à valeurs dans $\mathbb{N}^*$. On suppose que l'ensembe d'entiers*

$$\{n \geq 1 : \mathbb{P}(Y_1 = n) > 0\}$$

*a pour plus grand diviseur commun 1. Soit $m = \mathbb{E}(Y_1)$.*

1. *Montrer que le processus suivant est une chaîne de Markov :*

$$X_n = \inf\{k \geq n : k = Y_1 + \ldots + Y_l \text{ pour un certain } l \geq 0\} - n.$$

2. *Déterminer la limite de $\mathbb{P}(X_n = 0)$ lorsque $n \to +\infty$.*
3. *Montrer que, lorsque $n \to +\infty$,*

$$\mathbb{P}(n = Y_1 + \ldots + Y_k \text{ pour un certain } k \geq 0) \to \frac{1}{m}.$$

## 6.4 Théorème ergodique

Nous terminons ce chapitre avec un autre résultat de convergence, appelé théorème ergodique, qui ne nécessite pas de supposer que la chaîne est récurrente positive, mais seulement récurrente.

**Théorème 6.20** *(**Théorème ergodique**)*
*Soient $(X_n)$ une chaîne de Markov récurrente et irréductible, $\lambda$ une mesure invariante strictement positive, et $x_0$ un point quelconque de $E$. Alors :*

1. *Pour tout $y$ de $E$,*

$$\lim_{n\to\infty} \frac{\sum_{0\le k<n} \mathbb{1}_{\{X_k=y\}}}{\sum_{0\le k<n} \mathbb{1}_{\{X_k=x_0\}}} = \frac{\lambda(y)}{\lambda(x_0)}, \quad \mathbb{P}_{x_0} \; p.s..$$

2. *Plus généralement, pour toute fonction $f : E \to \mathbb{R}$ positive ou $\lambda$-intégrable (i.e. $\sum_{z\in E} |f(z)|\lambda(z) < \infty$), on a :*

$$\lim_{n\to\infty} \frac{\sum_{0\le k<n} f(X_k)}{\sum_{0\le k<n} \mathbb{1}_{\{X_k=x_0\}}} = \frac{1}{\lambda(x_0)} \sum_{z\in E} f(z)\lambda(z), \quad \mathbb{P}_{x_0} \; p.s.$$

**Remarque 6.21**

1. *La chaîne étant récurrente et irréductible, $\lambda$ est une unique à une constante multiplicative près. Il est clair que les deux rapports $\frac{\lambda(y)}{\lambda(x_0)}$ et $\frac{1}{\lambda(x_0)} \sum_{z\in E} f(z)\lambda(z)$ ne dépendent pas du choix de cette constante.*
2. *Soit $y \in E$. La suite $\sum_{0\le k<n} \mathbb{1}_{\{X_k=y\}}$ converge, lorsque $n$ tend vers l'infini, vers $N_y$. Puisque la chaîne est récurrente, cette v.a. est infinie.*

**Preuve du Théorème 6.20.**
Voir le livre de P. Vallois, *Modélisations et simultations stochatiques.* Ellipses 2007, Théorème 6.56, Section 6.6. ■

## 7 Exemples de chaînes de Markov

### 7.1 Un modèle de diffusion gazeux

On considère un récipient contenant un nombre $N$ de particules. Ce récipient est séparé en deux parties $A$ et $B$ par une cloison poreuse qui laisse passer les particules et leur permet ainsi de "diffuser". On suppose qu'à l'instant initial il y a $X_0 = a$ particules dans le compartiment $A$ et donc $N-a$ particules dans $B$. Le phénomène de diffusion entre $A$ et $B$ a lieu à des instants fixés $1, 2, \cdots$ et une seule particule à la fois peut passer d'un compartiment à l'autre. La dynamique de diffusion est la suivante : supposons qu'à l'instant $n$, il y a $X_n = i$ particules dans $A$, à l'instant $n+1$, une particule, prise au hasard, change de compartiment. Si $1 \leq i \leq N-1$, alors $X_{n+1} = i - 1$ (resp. $X_{n+1} = i + 1$) si une particule de $A$ migre vers $B$ (resp. une particule de $B$ migre vers $A$) ; ce qui se produit avec probabilité $\dfrac{i}{N}$ (resp. $\dfrac{N-i}{N}$). Lorsque $X_n = 0$ (resp. $X_n = N$) alors $X_{n+1} = 1$ (resp. $X_{n+1} = N - 1$).
La pression dans un compartiment étant proportionnelle au nombre de particules, les particules situées dans un compartiment haute pression ont tendance à rejoindre le compartiment basse pression. Il parait intuitif qu'au bout d'un certain temps le système va vers un **état d'équilibre** ; ce qui se traduit mathématiquement par l'existence d'une probabilité invariante et la convergence (en loi) de $X_n$, lorsque $n \to \infty$, vers la probabilité invariante.
Le modèle est appelé modèle de **diffusion d' Ehrenfest** . On peut aussi présenter ce modèle en remplaçant le récipient par une urne possédant deux compartiments, et les particules par des boules.
**1)** Il est clair que $(X_n)$ est une chaîne de Markov, à valeurs dans $E = \{0, 1, \cdots, N\}$ et de probabilité de transition :

$$\pi(i,j) = \begin{cases} \dfrac{i}{N} & \text{lorsque } j = i-1, j \in E, \\ \dfrac{N-i}{N} & \text{lorsque } j = i+1, j \in E, \\ 0 & \text{sinon.} \end{cases}$$

pour tout $0 \leq i \leq N$.
Puisque $\pi(i, i-1) > 0$ et $\pi(i, i+1) > 0$ alors $i$ communique avec $i + 1$, lorsque $0 < i < N$. De plus $0 \rightleftarrows 1$ et $N \rightleftarrows N - 1$, on en déduit que tous les états communiquent. La chaîne est donc irréductible. L'ensemble $E$ étant fini, $(X_n)$ est récurrente positive et irréductible.
**2)** Soit $\nu$ sa probabilité invariante. Par définition $\nu$ vérifie les équations de récurrence :

$$\begin{aligned} \nu(j) &= \sum_{i=0}^{N} \nu(i)\pi(i,j) = \nu(j-1)\pi(j-1,j) + \nu(j+1)\pi(j+1,j) \\ &= \nu(j-1)\frac{N-j+1}{N} + \nu(j+1)\frac{j+1}{N}, \end{aligned}$$

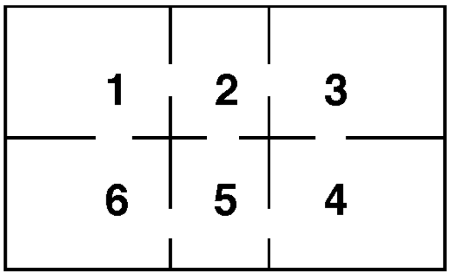
pour tout $0 \leq j \leq N$ avec la convention $\nu(-1) = \nu(N+1) = 0$.
Posons $\nu(j) = C_N^j\, 2^{-N} \nu'(j)$, pour tout $0 \leq j \leq N$. On exprime $\nu(j)$ à l'aide de $\nu'(j)$, on déduit facilement de la relation précédente :

$$N\nu'(j) = j\nu'(j-1) + (N-j)\nu'(j+1), \quad 0 \leq j \leq N.$$

En particulier si $j = 0$, on a : $\nu'(0) = \nu'(1)$. On montre ensuite par récurrence sur $j$ que $\nu'(j) = \nu'(0)$. Puisque $\nu(0) + \cdots + \nu(N) = 1$, alors $\nu'(0) = \cdots = \nu'(N) = 1$. Par conséquent la probabilité invariante est $\mathcal{B}(N, \frac{1}{2})$.
L'interprétation de ce résultat est le suivant. On décide d'affecter une particule au compartiment $A$ (resp. $B$) avec probabilité 1/2. On répète $N$ fois cette opération. Ainsi, la distribution initiale de particules dans $A$ est $\mathcal{B}(N, \frac{1}{2})$. On fait ensuite agir la dynamique de diffusion. La loi de $X_1$ est inchangée, ce qui parait intuitivement exact.
On déduit de la proposition 6.9 que, partant de $i$, la moyenne du temps de retour est $\frac{1}{\nu(i)} = \frac{2^N}{C_N^i}$.

## 7.2 Déplacement dans un labyrinthe

On considère une souris se déplaçant dans un espace divisé en 6 compartiments. Il existe des portes de communication entre certains compartiments comme le montre la figure suivante :




Lorsque la souris est dans un compartiment à un instant donné, elle choisit au hasard une des portes donnant sur un compartiment contigu et s'y déplace à l'instant suivant. Par exemple, si la souris est dans le compartiment 1, elle choisit de migrer vers le compartiment 2 ou 6, avec probabilité 1/2. Il est clair que le déplacement de la souris est Markovien. Soit $E = \{1, 2, \cdots, 6\}$. La matrice de transition est

$$\pi = \begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 \end{pmatrix}$$

Il est clair que la chaîne de Markov est récurrente et irréductible. On désigne par $\nu = (\nu(1), \cdots, \nu(6))$ sa probabilité invariante. Il est facile de vérifier que $\nu(i)$ est proportionnel au nombre de portes de sortie du compartiment $i$. Ainsi :

$$\nu(1) = 2c, \quad \nu(2) = 3c, \quad \nu(3) = 2c, \quad \nu(4) = 2c, \quad \nu(5) = 3c, \quad \nu(6) = 2c.$$

Puisque $\nu(1) + \cdots + \nu(6) = 1$, on en déduit que $c = 1/14$. Par conséquent la probabilité invariante est $\nu = (\frac{2}{14}, \frac{3}{14}, \frac{2}{14}, \frac{2}{14}, \frac{3}{14}, \frac{2}{14})$.

## 7.3 Chaine de Wright

On suppose que l'on a une population de $2N$ gènes de types "$A$" ou "$a$". On note $X_n$ le nombre de gènes "$A$" à la $n$-ième génération.

**1)** On commence par définir le cas de la dérive pure. Le passage d'une génération à la suivante s'effectue suivant la règle suivante : si la population mère comporte $j$ gènes de type "$A$", la génération suivante est obtenue en effectuant $2N$ épreuves de Bernoulli indépendantes, avec deux résultats possibles "$A$" ou "$a$", et où la probabilité d'avoir le type "$A$" est $p_j := \dfrac{j}{2N}$. La suite est rédigée sous la forme d'un exercice.

**Exercice 7.1** *(Chaine de Wright avec dérive pure)*

1. *Vérifier que* $(X_n)$ *est une chaine de Markov.*
2. *Que dire des états* $0$ *et* $2N$ *?*
3. *La chaine est-elle irréductible ?*
4. *Déterminer la nature des états.*
5. *Montrer que* $(X_n)$ *est une martingale. Calculer la probabilité de fixation en* $0$ *ou* $2N$*.*

**2)** On considère maintenant le cas où il y a un avantage sélectif en faveur de "$A$". La dynamique est la même que précédemment mais cette fois :

$$p_j = \frac{1-q^j}{1-q^{2N}} \quad \textit{avec} \quad q = e^{-s/N}, s > 0$$

**Exercice 7.2** *(Chaine de Wright avec avantage sélectif en faveur de "A")*

1. *Reprendre les questions de 1 à 4 de l'excercice précédent.*
2. *Expliquer comment on peut retrouver, du moins formellement, le cas de la "dérive pure".*
3. *Montrer que* $(q^{2NX_n}, n \geq 0)$ *est une martingale. Calculer la probabilité de fixation en* $0$ *ou* $2N$*.*

## 8 Références bibliographiques

On peut donner comme bonnes références sur les chaînes de Markov les ouvrages suivant :

— D. Foata et A. Fuchs. *Calcul des probabilités.* Masson, 1996 (chapitres 5 et 6).

— E. Çinlar. *Introduction to stochastic processes*, Prentice-Hall-Inc, 1975 (chapitres 5, 6 et 7).

— R. Durrett, *Probability : theory and examples*, Duxbury Press, Belmont, CA, second edition, 1996 (chapitres 5 et 6).

On trouvera dans :

— W. Feller, *An introduction to probability theory and its applications. Vol. I*, John Wiley & Sons Inc., 1968 (chapitre XVI)

et

— S. Karlin, *Initiation aux processus aléatoires*, Dunod, 1969 (chapitre 4)

des développements algébriques très intéressants prolongeant la présentation.

Des exemples de chaînes de Markov intervenant dans la modélisation de certains phénomènes biologiques figurent dans :

— S. Karlin, *Initiation aux processus aléatoires*, Dunod, 1969 (chapitre 13)

et

— R. Durrett, *Probability models for DNA sequence evolution*, Springer, 2002.