

# THÉORÈMES LIMITES

ADRIEN RICHOU

*Master MSS Bordeaux*

On considère  $X_n$  une chaîne de Markov homogène sur un espace d'états  $E$  dénombrable et de matrice de transition  $P$ . Pour  $x \in E$ , si  $X_0 = x$ , on note  $\tau_x = \inf\{p \geq 1 : X_p = x\}$  le temps de retour en  $x$ . On s'intéresse dans ce chapitre au comportement asymptotique de  $(X_n)_{n \in \mathbb{N}}$ .

*Remarque 0.1.* Notons  $\mu_n$  la loi de  $X_n$  (i.e. un vecteur ligne) pour tous  $n \in \mathbb{N}$ . Si  $X_n$  converge en loi, alors  $\mu_n \rightarrow \mu_\infty$  avec  $\mu_\infty$  une mesure de probabilité. Vu que l'on a  $\mu_{n+1} = \mu_n P$ , en passant à la limite on trouve que  $\mu_\infty$  vérifie

$$\mu_\infty = \mu_\infty P.$$

## 1. MESURES INVARIANTES

**Définition 1.1.** Une mesure  $\pi$  sur  $E$  est dite invariante pour  $P$  si  $\pi P = \pi$ . Si  $\pi$  est de plus une mesure de probabilité, on dit que  $\pi$  est une mesure de probabilité invariante (ou stationnaire).

Exercice: Commençons par un exemple. Soit

$$P = \begin{pmatrix} 1/2 & 0 & 0 & 0 & 1/2 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1/4 & 1/4 & 1/4 & 1/4 \\ 1/2 & 0 & 0 & 0 & 1/2 \end{pmatrix}$$

Les classes sont:  $\{1, 5\}$  (récurrente),  $\{3\}$  (récurrente) et  $\{2, 4\}$  (transiente).

L'ensemble des mesures de probabilités invariantes est:  $\lambda(1/2, 0, 0, 0, 1/2) + (1 - \lambda)(0, 0, 1, 0, 0)$ , pour  $0 \leq \lambda \leq 1$ .

Les mesures invariantes ne chargent pas les points transients. Et elles sont des combinaisons convexes de mesures invariantes sur chacune des 2 classes récurrentes.

On ne considérera donc dans la suite que le cas irréductible (et récurrent).

*Remarque 1.2.* Si  $P$  est une matrice stochastique et si  $C \subset E$  est une classe fermée alors  $P|_C$  est encore une matrice stochastique.

### 1.1. Existence d'une mesure invariante.

Cas d'un espace d'états fini : il existe toujours une mesure de probabilité invariante.

*Proof.* On note  $d$  le cardinal de  $E$ . Soit  $\mu_0$  une probabilité sur  $E$ , c'est-à-dire un vecteur de  $\mathbb{R}^d$  à entrées positives et de somme 1. L'ensemble des probabilités sur  $E$  forme donc un ensemble compact pour la topologie usuelle sur  $\mathbb{R}^d$ .

On considère la suite:  $\nu_n = \frac{\mu_0 + \mu_0 P + \dots + \mu_0 P^n}{n+1}$ . Il s'agit de la moyenne de Césaro de la loi de la chaîne de Markov. C'est encore une probabilité sur  $E$ . Par compacité, on peut extraire une sous suite  $\nu_{n_k}$  convergeant vers une probabilité  $\nu_\infty$ .

On a  $\nu_{n_k} P = \nu_{n_k} + \frac{\mu_0 P^{n_k+1} - \mu_0}{n_k+1}$ . Puisque  $\mu_0 P^{n_k+1}$  reste borné;  $n_k \rightarrow +\infty$  donne :  $\nu_\infty P = \nu_\infty$ . □

Cas d'un espace d'états dénombrable : il n'existe pas nécessairement de mesure de probabilité invariante à la marche aléatoire simple sur  $\mathbb{Z}$ , mais une mesure uniforme sur  $\mathbb{Z}$  est invariante.

## 1.2. Unicité de la mesure invariante.

**Théorème 1.3.** *Soit  $X_n$  une chaîne de Markov irréductible et récurrente. Alors il existe une mesure strictement positive invariante. De plus, cette mesure est unique à multiplication près par une constante.*

*Proof.* Pour chaque  $x \in E$ , on pose:

$$(1.1) \quad \lambda^x(y) = \mathbb{E}_x \left( \sum_{k=0}^{\tau_x-1} 1_{\{X_k=y\}} \right).$$

Cette quantité désigne le nombre moyen de passage en  $y$  entre 2 passages en  $x$ . On a clairement:

$$(1.2) \quad \lambda^x(x) = 1$$

flight simulator

La chaîne étant récurrente, on a que  $\mathbb{P}_x(\{\tau_x < +\infty\}) = 1$  et que  $X_{\tau_x} = x$ , on peut donc remarquer que

$$(1.3) \quad \lambda^x(y) = \mathbb{E}_x \left( \sum_{k=1}^{\tau_x} 1_{\{X_k=y\}} \right).$$

Maintenant,

$$\begin{aligned}
\lambda^x(y) &= \mathbb{E}_x \left( \sum_{k=1}^{\tau_x} 1_{\{X_k=y\}} \right) \\
&= \mathbb{E}_x \left( \sum_{k=1}^{+\infty} 1_{\{\tau_x \geq k\}} 1_{\{X_k=y\}} \right) \\
&= \sum_{k=1}^{+\infty} \mathbb{P}_x (\{\tau_x \geq k\}, \{X_k = y\}) \\
&= \sum_{z \in E} \sum_{k=1}^{+\infty} \mathbb{P}_x (\{\tau_x \geq k\}, \{X_k = y\}, \{X_{k-1} = z\})
\end{aligned}$$

(1.4)

Or  $\{\tau_x \geq k\} = \{\tau_x \leq k-1\}^c \in \sigma(X_0, X_1, \dots, X_{k-1})$ , donc est indépendant de l'événement  $\{X_k = y\}$  conditionnellement à  $\{X_{k-1} = z\}$  par la propriété de Markov faible d'où:

$$\begin{aligned}
&\mathbb{P}_x (\{\tau_x \geq k\}, \{X_k = y\}, \{X_{k-1} = z\}) \\
&= \mathbb{P} (\{X_k = y\} | \{X_{k-1} = z\}) \mathbb{P}_x (\{\tau_x \geq k\} | \{X_{k-1} = z\}) \mathbb{P}_x (\{X_{k-1} = z\}) \\
&= P_{z,y} \mathbb{P}_x (\{\tau_x \geq k\}, \{X_{k-1} = z\}).
\end{aligned}$$

D'où

$$\begin{aligned}
\lambda^x(y) &= \sum_{z \in E} \sum_{k=1}^{+\infty} P_{z,y} \mathbb{P}_x (\{\tau_x \geq k\}, \{X_{k-1} = z\}) \\
&= \sum_{z \in E} P_{z,y} \sum_{k=1}^{+\infty} \mathbb{P}_x (\{\tau_x \geq k\}, \{X_{k-1} = z\}) \\
&= \sum_{z \in E} P_{z,y} \sum_{l=0}^{+\infty} \mathbb{P}_x (\{\tau_x \geq l+1\}, \{X_l = z\}) \\
&= \sum_{z \in E} P_{z,y} \mathbb{E}_x \left( \sum_{l=0}^{+\infty} 1_{\{\tau_x \geq l+1\}} 1_{\{X_l=z\}} \right) \\
&= \sum_{z \in E} P_{z,y} \mathbb{E}_x \left( \sum_{l=0}^{\tau_x-1} 1_{\{X_l=z\}} \right) \\
&= \sum_{z \in E} \lambda^x(z) P_{z,y} \\
&= (\lambda^x P)(y).
\end{aligned}$$

Donc  $\lambda^x$  est une mesure positive invariante par  $P$  !

Maintenant, montrons que pour tout  $x, y \in E$ ,

$$0 < \lambda^x(y) < +\infty.$$

En effet, d'une part,  $P$  étant irréductible, il existe  $n \geq 1$  tel que  $P_{x,y}^n > 0$  et

$$\lambda^x(y) = (\lambda^x P^n)(y) = \sum_{z \in E} \lambda^x(z) P_{z,y}^n \geq \lambda^x(x) P_{x,y}^n = P_{x,y}^n > 0.$$

D'autre part, il existe  $m \geq 1$  tel que  $P_{y,x}^m > 0$  et

$$1 = \lambda^x(x) = (\lambda^x P^m)(x) = \sum_{z \in E} \lambda^x(z) P_{z,x}^m \geq \lambda^x(y) P_{y,x}^m,$$

donc

$$\lambda^x(y) \leq \frac{1}{P_{y,x}^m} < +\infty.$$

Il reste à prouver l'unicité de cette mesure positive invariante. Soit donc  $\nu$  une autre mesure positive (non identiquement nulle) invariante pour  $P$ . Soit  $i \in E$  tel que  $\nu_i \neq 0$ . Montrons d'abord que:

$$(1.5) \quad \nu_j \geq \nu_i \lambda^i(j).$$

Cette partie n'utilise que l'irréductibilité.

Soit  $j \neq i \in E$ , on a

$$\begin{aligned} \nu_j = (\nu P)_j &= \sum_{i_1 \in E} \nu_{i_1} P_{i_1,j} \\ &= \nu_i P_{i,j} + \sum_{i_1 \in E, i_1 \neq i} \nu_{i_1} P_{i_1,j}. \end{aligned}$$

Et de même  $\nu_{i_1} = \sum_{i_2 \in E} \nu_{i_2} P_{i_2,i_1}$ , donc,

$$\begin{aligned} \nu_j &= \nu_i P_{i,j} + \sum_{i_1, i_1 \neq i} \sum_{i_2 \in E} \nu_{i_2} P_{i_2,i_1} P_{i_1,j} \\ &= \nu_i P_{i,j} + \sum_{i_1, i_1 \neq i} \nu_{i_1} P_{i_1,i_1} P_{i_1,j} + \sum_{i_1, i_1 \neq i} \sum_{i_2, i_2 \neq i} \nu_{i_2} P_{i_2,i_1} P_{i_1,j} \end{aligned}$$

Ainsi, en poursuivant le même raisonnement, on a pour  $n \geq 1$ ,

$$\begin{aligned} \nu_j &= \nu_i P_{i,j} \\ &+ \sum_{i_1 \in E, i_1 \neq i} \nu_{i_1} P_{i,i_1} P_{i_1,j} \\ &+ \sum_{i_1 \in E, i_1 \neq i} \sum_{i_2, i_2 \neq i} \nu_{i_2} P_{i,i_2} P_{i_2,i_1} P_{i_1,j} \\ &+ \dots \\ &+ \sum_{i_1 \in E, i_1 \neq i} \sum_{i_2 \in E, i_2 \neq i} \dots \sum_{i_{n-1} \in E, i_{n-1} \neq i} \nu_{i_{n-1}} P_{i,i_{n-1}} P_{i_{n-1},i_{n-2}} P_{i_{n-2},i_{n-3}} \dots P_{i_2,i_1} P_{i_1,j} \\ &+ \sum_{i_1 \in E, i_1 \neq i} \sum_{i_2 \in E, i_2 \neq i} \dots \sum_{i_n \in E, i_n \neq i} \nu_{i_n} P_{i,i_n} P_{i_n,i_{n-1}} P_{i_{n-1},i_{n-2}} \dots P_{i_2,i_1} P_{i_1,j}. \end{aligned}$$

Donc,

$$(1.6) \quad \nu_j \geq \nu_i \left( P_{i,j} + \sum_{i_1 \in E, i_1 \neq i} P_{i,i_1} P_{i_1,j} + \cdots + \sum_{i_1 \in E, i_1 \neq i} \sum_{i_2 \in E, i_2 \neq i_1} \cdots \sum_{i_{n-1} \in E, i_{n-1} \neq i_{n-2}} P_{i,i_{n-1}} P_{i_{n-1},i_{n-2}} \cdots P_{i_2,i_1} P_{i_1,j} \right)$$

Et on a par exemple,

$$\sum_{i_1 \in E, i_1 \neq i} P_{i,i_1} P_{i_1,j} = \sum_{i_1 \in E, i_1 \neq i} \mathbb{P}_i(X_2 = j, X_1 = i_1) = \mathbb{P}_i(X_2 = j, X_1 \neq i) = \mathbb{P}_i(X_2 = j, \tau_i \geq 2)$$

Ainsi, on déduit de (1.6), que

$$\begin{aligned} \nu_j &\geq \nu_i (\mathbb{P}_i(X_1 = j) + \mathbb{P}_i(X_2 = j, \tau_i \geq 2) + \cdots + \mathbb{P}_i(X_n = j, \tau_i \geq n)) \\ &= \nu_i (\mathbb{E}_i(1_{\{X_1=j\}}) + \mathbb{E}_i(1_{\{X_2=j\}} 1_{\{\tau_i \geq 2\}}) + \cdots + \mathbb{E}_i(1_{\{X_n=j\}} 1_{\{\tau_i \geq n\}})) \\ &= \nu_i \mathbb{E}_i \left( \sum_{k=1}^n 1_{\{X_k=j\}} 1_{\{\tau_i \geq k\}} \right) \end{aligned}$$

En faisant tendre  $n$  vers l'infini dans l'inégalité précédente, on obtient,

$$\nu_j \geq \nu_i \mathbb{E}_i \left( \sum_{k=1}^{+\infty} 1_{\{X_k=j\}} 1_{\{\tau_i \geq k\}} \right) = \nu_i \mathbb{E}_i \left( \sum_{k=1}^{\tau_i-1} 1_{\{X_k=j\}} \right) = \nu_i \lambda^i(j),$$

c'est-à-dire, pour tout  $(i, j) \in E^2$ ,

$$(1.7) \quad \nu_j - \nu_i \lambda^i(j) \geq 0.$$

Maintenant montrons que

$$(1.8) \quad \nu_j = \nu_i \lambda^i(j).$$

Or  $\nu$  est une mesure invariante et  $P$  étant irréductible et récurrente,  $\lambda^i$  est aussi une mesure invariante, donc  $\nu - \nu_i \lambda^i$  est une mesure (positive) invariante. Par irréductibilité de  $P$ , pour tout  $j \in E$ , il existe  $n \geq 1$  tel que  $P_{j,i}^n > 0$ , et,

$$\begin{aligned} 0 = \nu_j - \nu_i \lambda^i(j) &= ((\nu - \nu_i \lambda^i) P^n)_j \\ &= \sum_{k \in E} (\nu_k - \nu_i \lambda^i(k)) P_{k,i}^n \\ &\geq (\nu_j - \nu_i \lambda^i(j)) P_{j,i}^n \end{aligned}$$

Or, par (1.7) et  $P_{j,i}^n > 0$ , il s'ensuit que pour tout  $j \in E$ ,  $\nu_j - \nu_i \lambda^i(j) = 0$ , donc la mesure  $\nu$  est proportionnelle à la mesure  $\lambda^i$ . Ceci finit la preuve du théorème 1.3 □

### 1.3. Récurrence positive.

*Remarque 1.4.* on peut calculer la masse totale de la mesure  $\lambda^x$ , elle est donnée par la formule:

$$\begin{aligned}
\lambda^x(E) &= \sum_{y \in E} \lambda^x(y) \\
&= \sum_{y \in E} \mathbb{E}_x \left( \sum_{k=0}^{\tau_x-1} 1_{\{X_k=y\}} \right) \\
&= \mathbb{E}_x \left( \sum_{k=0}^{\tau_x-1} \sum_{y \in E} 1_{\{X_k=y\}} \right) \\
&= \mathbb{E}_x \left( \sum_{k=0}^{\tau_x-1} 1 \right) \\
&= \mathbb{E}_x(\tau_x)
\end{aligned}$$

**Définition 1.5.** Un point  $i$  de  $E$  est dit **récurrent positif** pour la chaîne de Markov  $(X_n)_{n \in \mathbb{N}}$  si  $\mathbb{E}_i(\tau_i) < \infty$ . Un point récurrent de  $E$  qui n'est pas récurrent positif est dit **récurrent nul**.

**Théorème 1.6.** *La récurrence positive et la récurrence nulle sont des propriétés de classe.*

*Proof.* Résultat admis. □

**Proposition 1.7.** *Si  $E$  est fini, tous les états récurrents sont récurrents positifs.*

*Proof.* Résultat admis. □

Précisons les résultats précédents dans les cas finis et dénombrables:

Cas d'un espace d'états fini :

**Proposition 1.8.** *Si  $P$  est irréductible. Elle est récurrente positive.  $P$  admet une unique mesure de probabilité invariante  $\pi$ . De plus pour tout  $i \in E$ , on a*

$$\pi(i) = 1/\mathbb{E}_i(\tau_i).$$

Cas d'un espace d'états dénombrable :

**Proposition 1.9.** *Supposons  $P$  irréductible. Les assertions suivantes sont équivalentes :*

- (1) *Il existe une (unique) mesure de probabilité invariante  $\pi$ .*
- (2) *La mesure  $\pi$  définie pour tout  $i \in E$  par  $\pi(i) = 1/\mathbb{E}_i(\tau_i)$  est la mesure de probabilité invariante.*
- (3) *Tous les états sont récurrents positifs.*
- (4) *Il existe un état récurrent positif.*

Exemple: On considère la marche aléatoire asymétrique sur  $\mathbb{Z}$ .

Si  $p = q = 1/2$ , la marche est récurrente nulle et il existe une unique mesure positive invariante: la mesure uniforme qui n'est pas de masse finie.

Si  $p \neq 1/2$ , la marche est transiente et il existe une famille de mesure invariante:  $\lambda + \nu(\frac{p}{q})^n$ , pour  $\lambda, \nu \geq 0$ .

**1.4. Mesure réversible.** Une mesure invariante vérifie  $\pi P = P$ . cela peut se réécrire sous la forme:

$$\sum_{y \neq x} \pi(y)P(y, x) = \pi(x)(1 - P(x, x))$$

Cela signifie que sous  $\pi$ , la masse qui arrive (globalement) en  $x$  est la même que celle qui part de  $x$ .

**Définition 1.10.** On dit que  $\pi$  est réversible si pour tout  $x, y \in E$

$$\pi(x)P(x, y) = \pi(y)P(y, x).$$

Cela signifie que sous  $\pi$ , pour chaque arête  $(x, y)$ , la masse allant de  $x$  à  $y$  est la même que celle allant de  $y$  à  $x$ .

**Proposition 1.11.** Si  $\pi$  est une mesure réversible alors elle est invariante.

*Proof.* Soit  $x \in E$ ,

$$\begin{aligned} (\pi P)(x) &= \sum_{y \in E} \pi(y)P(y, x) \\ &= \sum_{y \in E} \pi(x)P(x, y) \\ &= \pi(x) \sum_{y \in E} P(x, y) \\ &= \pi(x). \end{aligned}$$

□

*Remarque 1.12.* Une chaîne de Markov n'admet pas forcément de mesure réversible mais si c'est le cas, il est facile de trouver les mesures réversible pour une chaîne de Markov.

## 2. THÉORÈME ERGODIQUE

Comme pour les suites de variables i.i.d., il existe pour les chaînes de Markov une "loi forte des grands nombres" appelée théorème ergodique.

Commençons par rappeler la loi forte des grands nombres usuelle.

**Théorème 2.1.** Soit  $(Y_n)_n$  une suite de variables aléatoires indépendantes et de même loi. On suppose qu'elles sont intégrables i.e.  $E[|Y_1|] < +\infty$ . Notons alors  $m = \mathbb{E}[Y_1]$ . Alors

$$\frac{Y_1 + \dots + Y_n}{n} \xrightarrow[n \rightarrow +\infty]{p.s.} m.$$

Pour  $x \in E$ , notons  $V_x(n) = \sum_{i=0}^{n-1} \mathbf{1}_{\{X_k=x\}}$  le nombre de visite de  $x$  avant  $n$ . La quantité  $\frac{V_x(n)}{n}$  correspond à la proportion de temps passé dans l'état  $x$  avant  $n$ .

**Théorème 2.2** (Théorème ergodique). *Supposons  $P$  irréductible et récurrente. Alors, pour toute mesure initiale  $\mu$ , et pour tout  $x \in E$ ,*

$$\frac{V_x(n)}{n} = \frac{1}{n} \sum_{k=0}^{n-1} \mathbf{1}_{\{X_k=x\}} \xrightarrow[n \rightarrow +\infty]{p.s.} \frac{1}{\mathbb{E}_x(\tau_x)}.$$

Dans le cas récurrent positif, La quantité  $\frac{1}{\mathbb{E}_x(\tau_x)}$  est nulle dans le cas récurrent nul. Elle est positive et vaut  $\pi(x)$  dans le cas récurrent positif où  $\pi$  désigne l'unique probabilité invariante.

**Théorème 2.3.** *Supposons que la chaîne est récurrente positive. Alors pour toute mesure initiale  $\mu$ , et pour toute fonction  $f$  de  $E$  dans  $\mathbb{R}$  bornée,*

$$\frac{1}{n} \sum_{k=0}^{n-1} f(X_k) \xrightarrow[n \rightarrow +\infty]{p.s.} \int_E f d\pi := \sum_{x \in E} \pi(x) f(x)$$

où  $\pi(x) = \frac{1}{\mathbb{E}_x(\tau_x)}$  désigne l'unique mesure de probabilité invariante.

Pour la preuve, on utilisera le lemme suivant.

**Lemme 2.4.** *Soit  $P$  une matrice stochastique irréductible et récurrente. Soit  $x \in E$  et soit  $\mu$  une mesure de probabilité sur  $E$ . Alors:*

$$\mathbb{P}_\mu(T_x < +\infty) = 1$$

avec  $T_x := \inf\{n \geq 0, X_n = x\}$  le temps d'atteinte de  $x$ .

Ce lemme dit juste que partant de n'importe quelle mesure initiale, on finira par toucher le point  $x$  dans le cas irréductible et récurrent.

Venons en à la démonstration du théorème ergodique.

*Proof.* Soit  $x \in E$  fixé. Puisque  $\mathbb{P}_\mu(T_x < +\infty) = 1$  et que  $(X_{T_x+n})_n$  est une chaîne de Markov partant de  $x$  et est indépendante de  $(X_0, \dots, X_{T_x})$ . Puisque  $T_x$  est fini, le comportement en temps long de la proportion de temps passé en  $x$  est la même pour  $(X_{T_x+n})_n$  et pour  $(X_n)_n$ . On peut donc considérer que la chaîne part de  $x$ .

Maintenant considérons les temps successifs de passage en  $x$

$$\begin{aligned} \tau_x^1 &= \inf\{p \geq 1, X_p = x\}, \\ \tau_x^n &= \inf\{p \geq \tau_x^{n-1}, X_p = x\}, \quad n \geq 2 \end{aligned}$$

ainsi que les durées entre 2 visites de  $x$ ,

$$S_x^n = \tau_x^n - \tau_x^{n-1}, \quad n \geq 1$$

(avec  $\tau_x^0 = 0$ ). On a montré (chapitre sur la récurrence et la transience) que les  $(S_x^n)_n$  sont des variables aléatoires indépendantes et de même loi. On a de plus clairement que

$$\mathbb{E}[S_x^n] = \mathbb{E}_x[\tau_x].$$

Maintenant  $V_x(n)$  désigne le nombre de visite en  $x$  avant  $n$ . On part au temps 0 de  $x$ , donc  $\tau_x^n$  correspond à l'instant où la chaîne est en  $x$  pour la  $n + 1$  ème fois. Donc:

$$\tau_x^{V_x(n)-1} < n \leq \tau_x^{V_x(n)}.$$



Or  $\tau_x^k = S_x^1 + \dots + S_x^k$ , donc en divisant par  $V_x(n)$ ,

$$\frac{S_x^1 + \dots + S_x^{V_x(n)-1}}{V_x(n)} < \frac{n}{V_x(n)} \leq \frac{S_x^1 + \dots + S_x^{V_x(n)}}{V_x(n)}.$$

Maintenant d'après la loi forte des grands nombres usuelle:

$$\frac{S_x^1 + \dots + S_x^n}{n} \xrightarrow[n \rightarrow +\infty]{p.s.} \mathbb{E}_x[\tau_x].$$

(Ce résultat est en fait valable même si  $\mathbb{E}_x[\tau_x] = +\infty$  car les variables aléatoires sont positives.) De plus, comme on est dans le cas récurrent:

$$V_x(n) \xrightarrow[n \rightarrow +\infty]{p.s.} +\infty$$

Donc pour  $\omega \in \Omega$  appartenant aux 2 ensembles précédents de probabilité 1:

$$\frac{S_x^1(\omega) + \dots + S_x(\omega)^{V_x(n,\omega)-1}}{V_x(n,\omega)} \xrightarrow[n \rightarrow +\infty]{} \mathbb{E}_x[\tau_x].$$

d'où par encadrement, on déduit que

$$\frac{n}{V_x(n)} \xrightarrow[n \rightarrow +\infty]{p.s.} \mathbb{E}_x[\tau_x]$$

et enfin

$$\frac{V_x(n)}{n} \xrightarrow[n \rightarrow +\infty]{p.s.} \frac{1}{\mathbb{E}_x[\tau_x]}.$$

□

*Remarque 2.5.* Dans le cas transient, on a que  $V_x(\infty) = \sum_{k \geq 0} \mathbf{1}_{\{X_k=x\}}$  est fini presque sûrement. Donc avec probabilité 1,

$$\frac{V_x(n)}{n} \leq \frac{V_x(\infty)}{n} \xrightarrow[n \rightarrow +\infty]{p.s.} 0.$$

Venons en à la preuve de la conséquence.

*Proof.* Pour simplifier, faisons la juste dans le cas où  $E$  est fini. Soit  $P$  irréductible récurrente positive. Soit  $f$  bornée par  $M$  sur  $E$ . Le point clé est d'écrire, avec les notations précédentes:

$$\frac{1}{n} \sum_{k=0}^{n-1} f(X_k) = \sum_{x \in E} \frac{V_x(n)}{n} f(x).$$

Maintenant:

$$\begin{aligned} \left| \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) - \sum_{x \in E} \pi_x f(x) \right| &= \left| \sum_{x \in E} \left( \frac{V_x(n)}{n} - \pi_x \right) f(x) \right| \\ &\leq M \sum_{x \in E} \left| \frac{V_x(n)}{n} - \pi_x \right|. \end{aligned}$$

Par hypothèse la somme est finie et d'après le résultat précédent, chaque terme tend (p.s.) vers 0, donc (p.s.):

$$\frac{1}{n} \sum_{k=0}^{n-1} f(X_k) \rightarrow \sum_{x \in E} \pi_x f(x).$$

□

### 3. CONVERGENCE EN LOI

Revenons à la question initiale sur la convergence en loi de la chaîne de Markov. On se place dans un cadre irréductible récurrent positif. Il existe donc une unique mesure de probabilité invariante  $\pi$ . As-t-on convergence en loi vers  $\pi$ ? Ce n'est pas forcément le cas comme le montre l'exemple très simple suivant.

Contre exemple : considérons  $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  sur  $E = \{0, 1\}$  Partant de  $X_0 = 0$  on a  $X_n = 0$  si  $n$  est pair et on a  $X_n = 1$  si  $n$  est impair (c'est-à-dire:  $P_{0,0}^n = 1$  si  $n$  est pair et on a  $P_{0,0}^n = 0$  si  $n$  est impair).

**3.1. Périodicité.** La non convergence dans l'exemple précédent vient du fait que la chaîne est périodique. On va définir maintenant la notion de période.

**Définition 3.1.** Soit  $i \in E$  un état. On définit la **période**  $d(i)$  de  $i$  par

$$d(i) := \text{pgcd}\{n \geq 1, P_{i,i}^n > 0\}$$

avec la convention  $\text{pgcd}(\emptyset) = \infty$ . Si  $d(i) = 1$  on dit que l'état  $i$  est **apériodique**.

*Exercice 3.2.* Montrer sur un exemple que l'on peut avoir:  $P_{i,i}^{d(i)} = 0$ .

*Remarque 3.3.* Si  $P_{i,i} > 0$  alors  $d(i) = 1$  et  $i$  est apériodique.

**Exemple 3.4.** Marche aléatoire sur  $\mathbb{Z}$

Soit  $P = (P_{i,j})_{i,j \in \mathbb{Z}}$  donnée, pour tout  $i \in \mathbb{Z}$  par  $P_{i,i-1} = P_{i,i+1} = 1/2$ . Déterminer la période de tous les états de cette matrice.

**Théorème 3.5.** *La période est une propriété de classe. C'est-à-dire, si les états  $i$  et  $j$  sont dans la même classe communicante, alors  $d(i) = d(j)$ .*

*Proof.* Soient  $i, j \in E$ ,  $i \neq j$ , appartenant à la même classe communicante, c'est-à-dire  $i \longleftrightarrow j$ . Montrons que  $d(i) = d(j)$ . Puisque  $i \longleftrightarrow j$ , il existe deux entiers  $n$  et  $m \geq 1$  tels que  $P_{i,j}^n > 0$  et  $P_{j,i}^m > 0$ . On a donc  $P_{i,i}^{n+m} > 0$  et donc  $d(i) | n + m$ .

On remarque également que l'ensemble  $\{k \geq 1, P_{j,j}^k > 0\}$  n'est pas vide (car  $P_{j,j}^{n+m} > 0$ ).

Soit  $k \geq 1$  tel que  $P_{j,j}^k > 0$ , on a  $P_{i,i}^{n+m+k} \geq P_{i,j}^n P_{j,j}^k P_{j,i}^m > 0$  et donc  $d(i) | n + m + k$ . Or  $d(i) | n + m$  d'où  $d(i) | k$ . On a donc obtenu:

$$\text{si } k \geq 1 \text{ satisfait } P_{j,j}^k > 0, \text{ alors } d(i) | k.$$

On en déduit que  $d(i) | \text{pgcd}\{k \geq 1, P_{j,j}^k > 0\} := d(j)$ . Par symétrie des rôles de  $i$  et de  $j$ , on a également  $d(j) | d(i)$  et donc  $d(i) = d(j)$ . □

**Proposition 3.6.**

$i$  est apériodique  $\iff \exists n_0 \geq 1, \forall n \geq n_0, P_{i,i}^n > 0$ .

*Proof.*

- Tout d'abord, si  $P_{i,i}^{n_0} > 0$  et  $P_{i,i}^{n_0+1} > 0$ , le pgcd de  $n_0$  et  $n_0 + 1$  étant 1, on obtient immédiatement que la période de  $i$  est 1.
- Supposons maintenant que  $i$  est apériodique. Soit  $N = \{n > 0, P_{i,i}^n > 0\}$ ,  $i$  est apériodique donc  $\text{pgcd}(N) = 1$ . Montrons qu'il existe  $m, n \in N$  tels que  $n = m + 1$ . L'ensemble  $N$  est clairement stable par addition. Posons  $G = \{n - m, n \in N, m \in N\}$ .  $G$  est donc un sous-groupe de  $\mathbb{Z}$ , donc  $G$  s'écrit  $G = k\mathbb{Z}$ . De plus, par Bezout,  $k = \text{pgcd}(N) = 1$ . Donc  $G = \mathbb{Z}$  et en particulier il existe  $n, m \in N$  tel que  $n = 1 + m$ .

Soit maintenant  $n_0 := m^2$ . Soit  $k \geq n_0$ , la division euclidienne de  $k - m^2$  par  $m$  donne

$$\begin{aligned} k &= m^2 + mq + r \quad \text{avec } q \geq 0 \text{ et } 0 \leq r < m \\ &= m^2 + mq + mr + r - mr \\ &= m(m - r + q) + r(m + 1) \\ &= m(m - r + q) + rn \end{aligned}$$

Puisque  $m - r + q \geq 0$ , la stabilité de  $N$  par addition donne que  $k \in N$ ; c'est à dire  $P_{i,i}^k > 0$ .  $\square$

**Proposition 3.7.** *Si  $E$  est fini alors  $P$  est irréductible et apériodique si et seulement si  $\exists n \geq 1$  tel que  $\forall i, j \in E, P_{i,j}^n > 0$ .*

*Proof.*

- Supposons que  $P$  est irréductible et apériodique. Alors

$$\begin{cases} \forall (i, j) \in E^2, \exists n_{i,j} \in \mathbb{N}, P_{i,j}^{n_{i,j}} > 0, \\ \forall i \in E, \exists m_i \geq 1, \forall m \geq m_i, P_{i,i}^{m_i} > 0. \end{cases}$$

Soit  $N = \max_{(i,j) \in E^2} (n_{i,j})$  et  $M = \max_{i \in E} (m_i)$ , posons  $N_0 = N + M$ . Soit  $l, k \in E$  on a

$$\begin{aligned} P_{l,k}^{N_0} &= \left( P_{l,j}^{n_{l,j}} P_{j,k}^{M+N-n_{l,j}} \right)_{l,k,j} \\ &= \sum_{p \in E} P_{l,p}^{n_{l,p}} P_{p,j}^{M+N-n_{l,p}} \geq P_{l,k}^{n_{l,k}} P_{k,k}^{M+N-n_{l,k}} > 0 \end{aligned}$$

- Réciproquement, supposons qu'il existe  $n \in \mathbb{N}$ , tel que pour tout  $i, j \in E$ ,  $P_{i,j}^n > 0$ . On a donc  $i \rightarrow j$ ,  $j \rightarrow i$  et donc  $i \leftrightarrow j$ . Tous les états communiquent. On a une seule classe communicante: la matrice  $P$  est irréductible. Montrons maintenant que  $P$  est apériodique, soit  $i \in E$ , on a

$$P_{i,i}^{n+1} = \sum_{j \in E} P_{i,j}^n P_{j,i}$$

La matrice  $P$  étant irréductible, il existe  $j_0 \in E$  tel que  $P_{i,j_0} > 0$ . Par hypothèse  $P_{i,j_0}^n > 0$ . Par conséquent,  $P_{i,i}^{n+1} > 0$ . Or par hypothèse on a

également  $P_{i,i}^n > 0$ . Le pgcd de  $n$  et  $n + 1$  est 1 on obtient que  $i$  et donc  $P$  est apériodique. □

### 3.2. Résultat de convergence.

En rajoutant une hypothèse de plus de mélange (l'apériodicité) on a le résultat de convergence en loi suivant.

**Théorème 3.8.** *Soit  $P$  irréductible, récurrente positive et apériodique. On note  $\pi$  la mesure de probabilité invariante. Alors, pour toute mesure initiale  $\mu$ ,*

$$P_\mu(X_n = j) \xrightarrow{n \rightarrow +\infty} \pi(j).$$

*Remarque 3.9.* En particulier si  $P$  est irréductible, récurrente positive et apériodique de mesure invariante  $\pi$ . On déduit que pour tout  $i, j \in E$ ,

$$P_{i,j}^{(n)} \xrightarrow{n \rightarrow +\infty} \pi(j).$$

et quand  $n$  tend vers l'infini,  $P^n$  converge vers une matrice dont les lignes sont identiques et égales au vecteur ligne  $\pi$ .

*Proof.* Nous allons ici juste donner l'idée de la preuve. Elle utilise un argument de couplage de 2 chaînes de Markov.

On considère  $(X_n)$  et  $(Y_n)$  2 chaînes de Markov indépendantes de même matrice de transition  $P$ . On note  $\mu$  la loi initiale de  $(X_n)$  et on suppose que la loi initiale de  $(Y_n)$  est la mesure de probabilité invariante  $\pi$ .

**Fait 1:**  $(X_n, Y_n)$  est une chaîne de Markov de matrice de transition  $Q$  définie par

$$Q_{(i,j),(k,l)} = P_{i,k} \cdot P_{j,l}.$$

**Fait 2:** Pour  $N \geq 1$  les puissances de  $Q$  sont données

$$(Q^N)_{(i,j),(k,l)} = (P^N)_{i,k} \cdot (P^N)_{j,l}.$$

**Fait 3:**  $P$  est apériodique donc  $Q$  est irréductible.

En effet montrons que  $(i, j)$  conduit à  $(k, l)$ . Par apériodicité de  $P$ , il existe  $n_1$  tel que si  $n \geq n_1$ ,  $p_{i,k}^n > 0$  et il existe  $n_2$  tel que si  $n \geq n_2$ ,  $p_{j,l}^n > 0$ . Donc pour  $N = \max(n_1, n_2)$ ,  $(Q^N)_{(i,j),(k,l)} > 0$ .

**Fait 4:**  $Q$  admet une probabilité invariante  $\nu$  donnée par  $\nu_{(i,j)} = \pi_i \pi_j$ .

**Fait 5:**  $Q$  est irréductible et admet une probabilité invariante. Donc  $Q$  est récurrente (positive).

**Fait 6:** Soit  $a \in E$ . Posons  $T = T_{(a,a)}$  le temps d'atteinte de  $(a, a)$  pour  $(x_n, Y_n)$ :

$$T = \inf\{n \geq 0, X_n = Y_n = a\}.$$

$Q$  est irréductible et récurrente, donc pour toute mesure initiale (en particulier pour  $\mu \times \pi$ )

$$\mathbb{P}(T < +\infty) = 1.$$

**Fait 7:** Par la propriété forte de Markov,  $(X_{T+n}, Y_{T+n})_{n \geq 0}$  est une chaîne de Markov de matrice de transition  $Q$  et de loi initiale  $\delta_{(a,a)}$ , qui est indépendante de  $(X_0, Y_0), \dots, (X_T, Y_T)$ .

**Fait 8:** On remarque également que  $(Y_{T+n}, X_{T+n})_{n \geq 0}$  est aussi une chaîne de Markov de matrice de transition  $Q$  et de loi initiale  $\delta_{(a,a)}$  (qui est indépendante de  $(X_0, Y_0), \dots, (X_T, Y_T)$ ). On en déduit que

$$(A_n, B_n) = \begin{cases} (X_n, Y_n) & \text{si } n \geq T \\ (Y_n, X_n) & \text{si } n \leq T \end{cases}$$

est encore une chaîne de Markov de matrice de transition  $Q$ . Sa loi initiale est la même que celle de  $(X_n, Y_n)$  ( $\mu \times \pi$ ).

On en déduit que pour tout  $n$  la loi de  $(A_n, B_n)$  est la même que celle de  $(X_n, Y_n)$ . En particulier, pour tout  $n$  la loi de  $B_n$  est  $\pi$ .

**Fait 9:** Soit  $i \in E$ , on a

$$|\mathbb{P}(X_n = j) - \pi_j| \leq \mathbb{P}(T > n).$$

En effet,

$$\begin{aligned} \mathbb{P}(X_n = j) &= \mathbb{P}(A_n = j) \\ &= \mathbb{P}(A_n = j, T > n) + \mathbb{P}(A_n = j, T \leq n) \\ &= \mathbb{P}(X_n = j, T > n) + \mathbb{P}(Y_n = j, T \leq n) \end{aligned}$$

et

$$\begin{aligned} \pi_j &= \mathbb{P}(Y_n = j) \\ &= \mathbb{P}(Y_n = j, T > n) + \mathbb{P}(Y_n = j, T \leq n); \end{aligned}$$

d'où

$$|\mathbb{P}(X_n = j) - \pi_j| = |\mathbb{P}(X_n = j, T > n) - \mathbb{P}(Y_n = j, T > n)| \leq \mathbb{P}(T > n).$$

**Fait 10:**  $Q$  est récurrente donc  $T$  est fini p.s. et

$$\mathbb{P}(T > n) \rightarrow 0 \text{ quand } n \rightarrow +\infty.$$

□

*Remarque 3.10.* Si  $P$  est irréductible, apériodique mais non nécessairement récurrente positive, on a pour toute mesure initiale  $\mu$ , et pour tout  $i \in E$ ,

$$\lim_{n \rightarrow \infty} \mathbb{P}_\mu(X_n = i) = \frac{1}{\mathbb{E}_i(\tau_i)}.$$

*Proof.* Admis.

□