Chapitre 9

Théorèmes ergodiques

9.1 Définitions

9.1.1 Mesures invariantes, fonctions invariantes

Définition: On appelle système dynamique un quadruplet $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ où $(\Omega, \mathcal{F}, \mathbb{P})$ est un espace probabilisé et θ une application mesurable de $(\Omega, \mathcal{F})$ dans lui-même qui préserve la mesure $\mathbb{P}$, c'est à dire que

$$\forall A \in \mathcal{F} \quad \mathbb{P}(\theta^{-1}(A)) = \mathbb{P}(A).$$

On dira aussi que la mesure $\mathbb{P}$ est invariante par θ.

Exemple fondamental : si $(X_n)_{n\in T}$ est un processus réel stationnaire indexé par $T = \mathbb{Z}$ ou $\mathbb{N}$, alors $(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T), \mathbb{P}_X, \theta)$ est un système dynamique, où θ est l'opérateur de décalage usuel : $(\theta x)_n = x_{n+1}$. On rappelle que si Π_n est l'opérateur de projection canonique de projection de $\mathbb{R}^T$ sur $\mathbb{R}$, on a $\Pi_0 \circ \theta^n = \Pi_n$ et la loi du processus canonique $(\Pi_n)_{n\in T}$ sous $\mathbb{P}_X$ est exactement la loi du processus $(X_n)_{n\in T}$ sous $\mathbb{P}$.

Inversement, si $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ est un système dynamique et f une fonction mesurable quelconque de $(\Omega, \mathcal{F})$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, le processus

$$f(\theta^k \omega)_{k\in\mathbb{N}}$$

est stationnaire. Si θ est inversible, alors le processus $f(\theta^k \omega)_{k\in\mathbb{Z}}$ est encore un processus stationnaire.

Définition: On dit qu'un événement A est un invariant du système $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ si il vérifie

$$\mathbb{P}(A \Delta \theta^{-1}(A)) = 0,$$

où, de manière équivalente si

$$\mathbb{1}_A = \mathbb{1}_A \circ \theta \quad \mathbb{P}-\text{p.s.}.$$

La famille des événements invariants forme une tribu (laissé en exercice, voir l'exercice 113), que l'on note souvent $\mathcal{I}$.

On dit qu'une application mesurable f de $(\Omega, \mathcal{F})$ est invariante si $f = f \circ \theta$ $\mathbb{P}$-presque sûrement. Les fonctions invariantes par θ sont exactement les fonctions $\mathcal{I}$-mesurables.

Soit f une fonction invariante. Posons

$$Y = \sum_{n=0}^{+\infty} |f - f \circ \theta| \circ \theta^n.$$

On a

$$\begin{aligned} \mathbb{E}(Y) &= \sum_{n=0}^{+\infty} \mathbb{E}|f - f \circ \theta| \circ \theta^n = \sum_{n=0}^{+\infty} \mathbb{E}|f - f \circ \theta| \\ &= \sum_{n=0}^{+\infty} 0 = 0. \end{aligned}$$

Ainsi, si f est une fonction invariante, on a presque sûrement pour tout entier naturel n, $f \circ \theta^n = f \circ \theta^{n+1}$, donc presque sûrement $f \circ \theta^n = f$ pour tout n.

En particulier, si f est invariante pour $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$, alors pour tout n, elle est invariante pour $(\Omega, \mathcal{F}, \mathbb{P}, \theta^n)$.

9.1.2 Ergodicité, mélange

Définition: On dit que le système $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ est ergodique si la tribu $\mathcal{I}$ est triviale sous $\mathbb{P}$, c'est à dire que pour tout $A \in \mathcal{I}$, $\mathbb{P}(A) \in \{0, 1\}$.

On dit que le système $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ est mélangeant si quels que soient les événements A, B dans $\mathcal{F}$, on a

$$\mathbb{P}(A \cap \theta^{-n}(B)) \to \mathbb{P}(A)\mathbb{P}(B),$$

Théorème 92. *Tout système mélangeant est ergodique.*

Démonstration. Soit A un événement invariant. On a vu que pour tout n, on a $\mathbb{1}_A = \mathbb{1}_A \circ \theta^n$ presque sûrement, donc

$$\mathbb{1}_{A \cap \theta^{-n}(A)} = \mathbb{1}_A \mathbb{1}_{\theta^{-n}(A)} = \mathbb{1}_A^2 = \mathbb{1}_A \ \mathbb{P} - \text{p.s.}$$

d'où $\mathbb{P}(A \cap \theta^{-n}(A)) = \mathbb{P}(A)$. En faisant tendre n vers $+\infty$, on obtient $\mathbb{P}(A) = \mathbb{P}(A)^2$, d'où $\mathbb{P}(A) \in \{0, 1\}$. ✓

Il est rare que l'on démontre le mélange directement sur des ensembles A et B quelconques de la tribu $\mathcal{F}$. Le plus souvent, on se base sur un lemme classique de la théorie des probabilités.

Théorème 93. *Si μ est une mesure finie sur la tribu $\mathcal{F}$ et $\mathcal{A}$ une algèbre engendrant $\mathcal{F}$, alors pour tout $A \in \mathcal{F}$ et tout $\varepsilon > 0$, il existe $A' \in \mathcal{A}$ tel que $\mu(A \Delta A') \le \varepsilon$, où $A \Delta A' = (A \cup A') \setminus (A \cap A')$.*

Démonstration. Il suffit de montrer que les ensembles A qui peuvent s'approcher ainsi forment une tribu. Pour les détails, on pourra se reporter à Garet–Kurtzmann [17], page 33. ✓

Théorème 94. *Soit $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ un système dynamique. Pour que le système soit mélangeant, il suffit qu'il existe une algèbre $\mathcal{F}_0$ de parties de Ω engendrant la tribu $\mathcal{F}$ et telle que :*

$$\forall A, B \in \mathcal{F}_0, \mathbb{P}(A \cap \theta^{-n}(B)) \longrightarrow_{n \to \infty} \mathbb{P}(A)\mathbb{P}(B).$$

Démonstration. Soit $A, B \in \mathcal{F}$. Si $\mathbb{P}(A) = 0$ ou $\mathbb{P}(B) = 0$, on a évidemment $\mathbb{P}(A \cap \theta^{-n}(B)) = 0$ pour tout entier $n \ge 1$. La limite est donc claire dans ce cas.

On peut donc supposer que $\mathbb{P}(A) > 0$ et $\mathbb{P}(B) > 0$. Soit $\varepsilon > 0$. D'après le théorème 93, on a l'existence de A_ε et B_ε dans $\mathcal{F}_0$ tels que

$\mathbb{P}(A\triangle A_\varepsilon) \leq \varepsilon$ et $\mathbb{P}(B\triangle B_\varepsilon) \leq \varepsilon$.

On a

$$\begin{aligned}\mathbb{P}(A\cap\theta^{-n}(B)) - \mathbb{P}(A)\mathbb{P}(B) &= \mathbb{P}(A\cap\theta^{-n}(B)) - \mathbb{P}(A_\varepsilon\cap\theta^{-n}(B_\varepsilon))\\ &+ \mathbb{P}(A_\varepsilon\cap\theta^{-n}(B_\varepsilon)) - \mathbb{P}(A_\varepsilon)\mathbb{P}(B_\varepsilon)\\ &+ \mathbb{P}(A_\varepsilon)\mathbb{P}(B_\varepsilon) - \mathbb{P}(A)\mathbb{P}(B).\end{aligned}$$

Si x_1, x_2, y_1, y_2 sont des nombres réels ou complexes de module plus petit que 1, on a

$$|x_1y_1 - x_2y_2| \leq |x_1 - x_2| + |y_1 - y_2|.$$

Ainsi, $|\mathbb{P}(A_\varepsilon)\mathbb{P}(B_\varepsilon) - \mathbb{P}(A)\mathbb{P}(B)| \leq 2\varepsilon$ et

$$|\mathbb{1}_A\mathbb{1}_{\theta^{-n}(B)} - \mathbb{1}_{A_\varepsilon}\mathbb{1}_{\theta^{-n}(B_\varepsilon)}| \leq |\mathbb{1}_A - \mathbb{1}_{A_\varepsilon}| + |\mathbb{1}_{\theta^{-n}(B)} - \mathbb{1}_{\theta^{-n}(B_\varepsilon)}|,$$

d'où en intégrant

$$\begin{aligned}\mathbb{P}((A\cap\theta^{-n}(B))\triangle(A_\varepsilon\cap\theta^{-n}(B_\varepsilon))) &\leq \mathbb{P}(A\triangle A_\varepsilon) + \mathbb{P}(\theta^{-n}(B\Delta B_\varepsilon))\\ &\leq \mathbb{P}(A\triangle A_\varepsilon) + \mathbb{P}(B\Delta B_\varepsilon)\\ &\leq 2\varepsilon.\end{aligned}$$

On a donc pour tout entier naturel n,

$$|\mathbb{P}(A\cap\theta^{-n}(B) - \mathbb{P}(A)\mathbb{P}(B)| \leq 4\varepsilon + \mathbb{P}(A_\varepsilon\cap\theta^{-n}(B_\varepsilon)) - \mathbb{P}(A_\varepsilon)\mathbb{P}(B_\varepsilon),$$

d'où

$$\overline{\lim_{n\to+\infty}}\ |\mathbb{P}(A\cap\theta^{-n}(B) - \mathbb{P}(A)\mathbb{P}(B)| \leq 4\varepsilon.$$

Comme ε est arbitraire, cela donne le résultat voulu.

✓

Corollaire 34. *Soit $n \geq 1$ et $\Omega = (\mathbb{R}^n)^{\mathbb{Z}^d}$. Pour tout $i \in \mathbb{Z}^d$, on note θ_i l'opérateur de décalage défini par $(\theta_i(\omega))_j = \omega_{i+j}$. Soit ν une mesure quelconque sur $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$.*

Pour tout $x \in \mathbb{Z}^d\backslash\{0\}$, le système $(\Omega, (\mathcal{B}(\mathbb{R}^n))^{\otimes\mathbb{Z}^d}, \nu^{\otimes\mathbb{Z}^d}, \theta_x)$ est ergodique.

Démonstration. Il suffit d'appliquer le théorème précédent à l'algèbre $\mathcal{A}$ formée par les événements locaux : $\mathcal{A}$ est la réunion des tribus $\sigma(\Pi_\Lambda)$, où Π_Λ est la projection canonique de Ω sur $(\mathbb{R}^d)^\Lambda$ et où Λ décrit l'ensemble des parties finies de $\mathbb{Z}^d$. Pour A et B dans cette algèbre, A et $\theta^{-n}(B)$ sont indépendants dès que n est assez grand, ce qui donne la convergence voulue. ✓

9.1.3 Sous-additivité

On dit qu'une suite de fonctions à valeurs réelles $(f_n)_{n\in\mathbb{N}^*}$ est sous-additive pour le système $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ si on a

$$\forall (n,p) \in \mathbb{N}^* \times \mathbb{N}^* \quad f_{n+p} \le f_p + f_n \circ \theta^p.$$

Exemple: si f est une fonction quelconque, la suite $(f_n)_{n\ge 1}$ définie par

$$f_n = \sum_{k=0}^{n-1} f \circ \theta^k$$

est sous-additive (en fait exactement additive).

Lemme 18 (Fekete)**.** *Soit $(u_n)_{n\ge 1}$ une suite vérifiant*

$$\forall n, p \ge 1 \quad u_{n+p} \le u_n + u_p.$$

Alors $\frac{u_n}{n}$ converge vers $\inf\limits_{n\ge 1} \frac{u_n}{n}$.

Démonstration. Il suffit de montrer que $\overline{\lim\limits_{n\to+\infty}} \frac{u_n}{n} \le \inf_{n\ge 1} \frac{u_n}{n}$. On va donc montrer que pour tout $k \ge 1$ $\overline{\lim\limits_{n\to+\infty}} \frac{u_n}{n} \le \frac{u_k}{k}$. Pour tout r entre 0 et $k-1$, on a $u_{kn+r} \le nu_k + u_r$, d'où

$$\frac{u_{kn+r}}{nk+r} \le \frac{nu_k}{nk+r} + \frac{u_r}{nk+r}.$$

Et $\overline{\lim\limits_{n\to+\infty}} \frac{u_{kn+r}}{nk+r} \le \frac{u_k}{k}$. Cependant $\overline{\lim\limits_{n\to+\infty}} \frac{u_n}{n} = \max\limits_{0\le r\le k-1} \overline{\lim\limits_{n\to+\infty}} \frac{u_{kn+r}}{nk+r}$, d'où le résultat. ✓

9.2 Le théorème ergodique et le théorème ergodique sous-additif

Le but de cette section est de présenter deux théorèmes importants et assez liés :

— Le théorème de Birkhoff (ou Birkhoff–Von Neumann) (1931), qui traite du comportement asymptotique des suites additives,

— Le théorème de Kingman (1968), qui traite du comportement asymptotique des suites additives.

Les preuves les plus répandues du théorème de Birkhoff passent par une version ou l'autre du lemme sous-additif maximal. Cependant, Avila et Bochi ont récemment remarqué qu'une des preuves du théorème ergodique sous-additif pouvait permettre d'obtenir en même temps le théorème ergodique de Birkhoff – alors que les preuves classiques du théorème de Kingmann utilisaient le théorème ergodique de Birkhoff. C'est le chemin que nous empruntons ici.

9.2.1 Un lemme

Le lemme qui suit est dérivé d'une note d'Avila et Bochi [2], inspirée par la preuve de Steele [35] du théorème ergodique sous-additif. La preuve de Steele s'inspirait elle-même des travaux de Kamae [21] et de Katznelson et Weiss [23].

On notera que la preuve est basée sur un algorithme de majoration, trajectoire par trajectoire, des fonctions sous-additives, en découpant l'intervalle $\{0,\dots,n\}$ en un certain nombre de pas, d'une manière qui n'est pas sans évoquer la preuve du lemme de Fekete, à la différence notable que les longueurs des pas sont aléatoires.

Lemme 19. *Soit $(\Omega,\mathcal{F},\mathbb{P},\theta)$ un système dynamique, $(g_n)_{n\ge 1}$ une famille de fonctions intégrables vérifiant*

$$\forall n,p\ge 1 \quad g_{n+p}\le g_n+g_p\circ\theta^n.$$

On pose

$$g=\underline{\lim}_{n\to+\infty}\frac{g_n}{n} \text{ et } L=\inf_{n\ge 1}\mathbb{E}[g_n]/n.$$

g est mesurable par rapport à la tribu des invariants $\mathcal{I}$. Soient M, N des entiers naturels quelconques. On pose $G_M = \max(g, -M)$, puis

$$A(N,M) = \bigcup_{1 \le \ell \le N} \left\{ \frac{g_\ell}{\ell} \le G_M + \frac{1}{M} \right\} \text{ (les gentils)}$$

et son complémentaire

$$B(N,M) = \bigcap_{1 \le \ell \le N} \left\{ \frac{g_\ell}{\ell} > G_M + \frac{1}{M} \right\} \text{ (les méchants).}$$

On pose $\psi_{N,M} = (G_M + \frac{1}{M})\mathbb{1}_{A_{N,M}} + g_1 \mathbb{1}_{B(N,M)}$. On a les inégalités

$$\begin{aligned} &\forall N, M \in \mathbb{N}^* \quad \forall n \ge N \\ &g_n \le \sum_{i=0}^{n-N} \psi_{N,M} \circ \theta^i + \sum_{i=n-N+1}^{n-1} \max(g_1, \psi_{N,M}) \circ \theta^i \end{aligned} \tag{9.1}$$

et $L \le \mathbb{E}[g]$.

Démonstration. — Étape 1 : D'après le lemme de Fekete, la suite $\mathbb{E}[g_n]/n$ converge vers $L = \inf_{n \ge 1} \mathbb{E}[g_n]/n$.

— Étape 2 : On a
$$\frac{g_{n+1}}{n} \le \frac{g_1}{n} + \frac{g_n}{n} \circ \theta,$$
donc en passant à la limite inférieure $g \le g \circ \theta$.

On a $g^+ = \underline{\lim}_{n \to +\infty} \frac{g_n^+}{n}$, donc avec le lemme de Fatou,
$$\mathbb{E}[g^+] \le \underline{\lim}_{n \to +\infty} \mathbb{E}[\frac{g_n^+}{n}] \le \mathbb{E}[g_1^+],$$
où la dernière inégalité vient de la sous-additivité. On en déduit que $G_M = \max(g, -M)$ est intégrable. On a $G_M \le G_M \circ \theta$. G_M est intégrable, de même $G_M \circ \theta - G_M$ est intégrable et positive, mais son intégrale est nulle, donc $G_M = G_M \circ \theta$ presque sûrement. En prenant l'intersection pour $M \in \mathbb{N}$, on obtient $g = g \circ \theta$.

On peut remarquer que $\psi_{N,M} \ge G_M + \frac{1}{M}$.

La preuve de l'inégalité (9.1) est essentiellement algorithmique. On définit une suite $(i_j(\omega))_{j \ge 0}$ par récurrence avec $i_0 = 0$, puis

— si $i_j > n - N$, $i_{j+1} = n$
— sinon, si $\theta^{i_j}(\omega) \in B(N, M)$, $i_{j+1} = i_j + 1$
— sinon, $i_{j+1} = i_j + k$, où k est un entier entre 1 et N tel que $g_k(\theta^{i_j}\omega) \le (G_M(\theta^{i_j}\omega) + \frac{1}{M})k = (G_M + \frac{1}{M})k$.

Soit j le plus petit entier tel que $i_j = n$. Soit s le plus grand entier tel que que $i_s \le n - N$: La sous-additivité donne

$$g_n \le \sum_{u=0}^{s} g_{i_{u+1}-i_u}(\theta^{i_u}(\omega)) + \sum_{i=i_{s+1}}^{n-1} g_1 \circ \theta^i.$$

Cependant, par construction on a pour tout u :

$$g_{i_{u+1}-i_u} \le \sum_{i=i_u}^{i_{u+1}-1} \psi_{N,M} \circ \theta^i.$$

En effet, si $\mathbb{1}_{B(N,M)} \circ \theta^{i_u} = 1$, on a $i_{u+1} - i_u = 1$ et la somme ne comprend qu'un terme ; dans le cas inverse, on a

$$\begin{aligned} g_{i_{u+1}-i_u} &\le (i_{u+1} - i_u)(G_M + 1/M) \circ \theta^{i_u} \\ &= \sum_{i=i_u}^{i_{u+1}-1} (G_M + 1/M) \circ \theta^i \\ &\le \sum_{i=i_u}^{i_{u+1}-1} \psi_{N,M} \circ \theta^i. \end{aligned}$$

On a utilisé successivement la définition de la suite, l'invariance de G_M par translation, et l'inégalité $G_M + 1/M \le \psi_{N,M}$. On a donc

$$g_n \le \sum_{i=0}^{i_{s+1}-1} \psi_{N,M} \circ \theta^i + \sum_{i=i_{s+1}}^{n-1} g_1 \circ \theta^i.$$

Comme $n - N \le i_{s+1} - 1$, on a

$$g_n \le \sum_{i=0}^{n-N} \psi_{N,M} \circ \theta^i + \sum_{n-N+1}^{n-1} \max(g_1, \psi_{N,M}) \circ \theta^i.$$

Notons que G_M et g_1 sont intégrables, on peut prendre l'espérance et on obtient

$$\mathbb{E}[g_n] \leq (n-N+1)\mathbb{E}[\psi_{N,M}] + (N-1)\mathbb{E}[\max(g_1, \psi_{N,M})].$$

En divisant par n et en faisant tendre n vers l'infini, on a

$$L \leq \mathbb{E}[\psi_{N,M}].$$

En faisant tendre N vers l'infini, le théorème de convergence dominée donne $L \leq \mathbb{E}[G_M] + 1/M$, soit

$$L \leq \mathbb{E}[g^+] - \mathbb{E}[\min(g^-, M)] + 1/M,$$

d'où $L \leq \mathbb{E}[g]$ en faisant tendre M vers l'infini. (Notons qu'il n'est pas exclus que $\mathbb{E}[g] = -\infty$.)

✓

9.2.2 Théorème de Birkhoff

Le théorème de Birkhoff–Von Neumann est le résultat de base de la théorie ergodique. Il en existe de nombreuses preuves. Nous allons le démontrer simplement à partir du lemme 19

Théorème 95. *Soit $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ un système dynamique et f une fonction intégrable. On pose*

$$S_n = S_n(f) = \sum_{k=0}^{n-1} f \circ \theta^k.$$

Alors, il existe une fonction $\overline{f}$ θ-invariante ($\overline{f} \circ \theta = \overline{f}$) telle que

$$\frac{S_n}{n} \to \overline{f},$$

presque sûrement. La convergence a lieu également dans L^1 et on a

$$\overline{f} = \mathbb{E}[f|\mathcal{I}],$$

où $\mathcal{I}$ est la tribu des événements invariants par θ.

Démonstration. — Étape 1 : convergence presque sûre
Il est facile de voir que (S_n) est sous-additive (en fait additive).

Avec le lemme 19, cela nous donne l'inégalité $\mathbb{E}[f] \leq \mathbb{E}[\varliminf_{n\to+\infty} \frac{S_n}{n}]$.

De même, $-\mathbb{E}[f] \leq \mathbb{E}[\varliminf_{n\to+\infty} \frac{-S_n}{n}]$. Ainsi

$$\mathbb{E}[\varlimsup_{n\to+\infty} \frac{S_n}{n}] \leq \mathbb{E}[f] \leq \mathbb{E}[\varliminf_{n\to+\infty} \frac{S_n}{n}],$$

ce qui entraîne la convergence presque sûre de $\frac{S_n}{n}$ vers une fonction $\overline{f}$. On a déjà montré qu'elle était invariante par θ.

— Étape 2 : convergence L^1 et identification de la limite.
Pour la suite, on va d'abord traiter le cas où f est bornée : dans ce cas $S_n(f)/n$ tend dans L^1 vers $\overline{f}$ grâce au théorème de convergence dominée.
Soit $A \in \mathcal{I}$:

$$\frac{S_n(f)}{n}\mathbb{1}_A = \frac{1}{n}\sum_{i=0}^{n-1}(f\mathbb{1}_A)\circ\theta^i,$$

donc

$$\mathbb{E}[\frac{S_n(f)}{n}\mathbb{1}_A] = \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}[(f\mathbb{1}_A)\circ\theta^i] = \mathbb{E}[f\mathbb{1}_A].$$

Comme $\frac{S_n(f)}{n}$ tend dans L^1 vers $\overline{f}$ quand n vers l'infini, on a aussi la convergence de $\mathbb{E}[\frac{S_n(f)}{n}\mathbb{1}_A]$ vers $\mathbb{E}[\overline{f}\mathbb{1}_A]$, ce qui permet l'identification $\mathbb{E}[\overline{f}\mathbb{1}_A] = \mathbb{E}[f\mathbb{1}_A]$. Finalement, on a bien $\overline{f} = \mathbb{E}[f|\mathcal{I}]$.
Passons au cas général : soit f_ε une fonction bornée telle que $\|f - f_\varepsilon\|_1 \leq \varepsilon/2$.
On a

$$\begin{aligned}
&S_n(f)/n - \mathbb{E}[f|\mathcal{I}] \\
&= (S_n(f)/n - S_n(f_\varepsilon)/n) + (S_n(f_\varepsilon)/n - \mathbb{E}[f_\varepsilon|\mathcal{I}]) \\
&\quad + (\mathbb{E}[f_\varepsilon|\mathcal{I}] - \mathbb{E}[f|\mathcal{I}])\,.
\end{aligned}$$

Or,

$$\|S_n(f)/n - S_n(f_\varepsilon)/n\|_1 \le \|f - f_\varepsilon\|_1 \le \varepsilon/2$$
$$\text{et } \|\mathbb{E}[f_\varepsilon|\mathcal{I}] - \mathbb{E}[f|\mathcal{I}]\|_1 \le \|f - f_\varepsilon\|_1 \le \varepsilon/2,$$

donc

$$\forall n \ge 1 \quad \|S_n(f)/n - \mathbb{E}[f|\mathcal{I}]\|_1 \le \varepsilon + \|S_n(f_\varepsilon)/n - \mathbb{E}[f_\varepsilon|\mathcal{I}]\|_1.$$

Et en faisant tendre n vers l'infini :

$$\overline{\lim_{n\to+\infty}} \ \|S_n(f)/n - \mathbb{E}[f|\mathcal{I}]\|_1 \le \varepsilon.$$

Comme ε est arbitraire, cela achève la preuve.

✓

9.2.3 Théorème de Kingman

Le théorème ergodique sous-additif est dû à Kingman [25]. La preuve présentée ici permet de déduire le théorème de Kingman du lemme 19 et du théorème de Birkhoff. Elle ressemble beaucoup à la preuve de Steele [35] [1].

Théorème 96. *Soit* $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ *un système dynamique,* $(g_n)_{n\ge 1}$ *une famille de fonctions intégrables vérifiant*

$$\forall n, p \ge 1 \quad g_{n+p} \le g_n + g_p \circ \theta^n.$$

Alors il existe une fonction g *mesurable par rapport à la tribu des invariants* $\mathcal{I}$ *telle que*

$$g = \lim_{n\to+\infty} \frac{g_n}{n} \text{ p.s.}$$

1. Il est possible de déduire directement le théorème de Kingman du lemme 19, sans utiliser le théorème de Birkhoff : c'est ce qu'ont fait Avila et Bochi [2]. Toutefois, passer par Birkhoff nous a paru plus naturel, on a donc simplement « factorisé » les deux preuves.

Démonstration. On a montré dans le lemme l'inégalité

$$\forall N, M \in \mathbb{N}^* \quad \forall n \geq N \quad g_n \leq \sum_{i=0}^{n-N} \psi_{N,M} \circ \theta^i + \sum_{i=n-N+1}^{n-1} \max(g_1, \psi_{N,M}) \circ \theta^i. \tag{9.2}$$

Notons que si G est une fonction intégrable $\frac{G \circ \theta^n}{n}$ tend presque sûrement vers 0 : c'est une conséquence du lemme de Borel–Cantelli. On en déduit que

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{i=n-N+1}^{n-1} \max(g_1, \psi_{N,M}) \circ \theta^i = 0 \text{ p.s.}$$

Avec le théorème de Birkhoff, il vient alors

$$\forall N, M \in \mathbb{N}^* \qquad \overline{\lim_{n \to +\infty}} \frac{g_n}{n} \leq \mathbb{E}[\psi_{N,M} | \mathcal{I}],$$

puis, avec le théorème de convergence dominée conditionnel, en faisant tendre N vers l'infini

$$\forall M \in \mathbb{N}^* \qquad \overline{\lim_{n \to +\infty}} \frac{g_n}{n} \leq G_M + \frac{1}{M},$$

d'où finalement $\overline{\lim\limits_{n \to +\infty}} \frac{g_n}{n} \leq g = \underline{\lim\limits_{n \to +\infty}} \frac{g_n}{n}$, ce qui montre la convergence presque sûre. ✓

Théorème 97. *Soit* $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ *un système dynamique,* $(g_n)_{n \geq 1}$ *une famille de fonctions intégrables avec*

$$\forall n, p \geq 1 g_{n+p} \leq g_n + g_p \circ \theta^n.$$

On suppose qu'il existe $c > 0$ *telle que pour tout* n $\mathbb{E} g_n \geq -nc$. *Alors il existe une fonction* g *mesurable par rapport à la tribu des invariants* $\mathcal{I}$ *telle que*

$$g = \lim_{n \to +\infty} \frac{g_n}{n} \textit{ p.s. et dans } L^1.$$

Enfin,

$$\mathbb{E} g = \inf_{n \geq 1} \frac{\mathbb{E} g_n}{n}.$$

Démonstration. On a déjà obtenu la convergence presque sûre par le théorème précédent, reste à voir la convergence dans L^1. La première étape consiste à établir l'intégrabilité de g.

La sous-additivité nous donne

$$\forall n,k \geq 1 \quad \frac{1}{nk}\sum_{j=0}^{n-1} g_k \circ \theta^{kj} - \frac{g_{nk}}{nk} \geq 0.$$

En faisant tendre n vers l'infini, le théorème de Birkhoff entraîne

$$\frac{1}{k}\mathbb{E}[g_k|\mathcal{I}_k] - g \geq 0, \tag{9.3}$$

où $\mathcal{I}_k$ est la tribu des invariants de θ^k.
En particulier, la fonction $\mathbb{E}[g_1|\mathcal{I}] - g$ est positive. On peut majorer son intégrale à l'aide du lemme de Fatou :

$$\begin{aligned}\mathbb{E}(\mathbb{E}[g_1|\mathcal{I}] - g) &= \mathbb{E} \lim_{n\to+\infty} [\mathbb{E}[g_1|\mathcal{I}] - \frac{g_n}{n}] \\ &\leq \underline{\lim}_{n\to+\infty} (\mathbb{E}g_1 - \frac{\mathbb{E}g_n}{n}) \leq \mathbb{E}g_1 + c.\end{aligned}$$

Ainsi, $\mathbb{E}[g_1|\mathcal{I}] - g$ est intégrable. g est donc également intégrable.
Maintenant, on a

$$\begin{aligned}(\frac{g_k}{k} - g)^+ &\leq (\frac{g_k}{k})^+ + |g| \\ &\leq \frac{1}{k}\sum_{j=0}^{k-1} g_1^+ \circ \theta^j + |g|.\end{aligned}$$

Mais, d'après le théorème ergodique L^1, la suite $(\frac{1}{k}\sum_{j=0}^{k-1} g_1^+ \circ \theta^j)_k$ est équi-intégrable, donc la suite $((\frac{g_k}{k} - g)^+)_k$ est équiintégrable et la convergence L^1 de $(g_k/k - g)^+$ vers 0 s'ensuit. Cependant

$$\mathbb{E}|\frac{g_k}{k} - g| = 2\mathbb{E}(\frac{g_k}{k} - g)^+ - (\frac{\mathbb{E}g_k}{k} - \mathbb{E}g) \leq 2\mathbb{E}(\frac{g_k}{k} - g)^+,$$

où la dernière inégalité vient de (9.3). Ainsi, on a bien que g_k/k converge dans L^1 vers g. Le dernier point découle alors aisément du Lemme de Fekete. ✓

9.3 Application aux chaînes de Markov

Les chaînes de Markov fournissent des exemples de systèmes dynamiques mesurés. Le but de section est de revisiter un certain nombre de résultats de convergence des chaînes de Markov, en particulier tout ce qui tourne autour du théorème ergodique des chaînes de Markov, à la lumière de ce que nous savons désormais de la théorie ergodique, et notamment du théorème de Birkhoff.

On va en particulier donner une version un peu plus forte du théorème ergodique des chaînes de Markov. Dans ce qui suit, on n'utilisera pas le théorème ergodique ponctuel des chaînes de Markov, en revanche on utilisera le théorème de la probabilité stationnaire (le théorème de convergence en loi).

Si E est un ensemble dénombrable, $P = (p_{i,j})$ une matrice stochastique et μ une mesure de probabilité sur E , on notera $\mathbb{P}^\mu$ la loi sur $(\Omega, \mathcal{F}) = (E^{\mathbb{N}}, \mathcal{B}(E^{\mathbb{N}})$ d'une chaîne de Markov de dynamique $P = (p_{i,j})$ et de loi initiale μ ; c'est à dire que

$$\mathbb{P}^\mu(X_0 = x_0, \dots, X_n = x_n) = \mu(x_0) \prod_{i=0}^{n-1} p_{x_i, x_{i+1}}.$$

On note comme d'habitude $\mathbb{P}^x$ pour $\mathbb{P}^{\delta_x}$.

On rappelle deux propriétés fondamentales des chaînes de Markov qui seront très utiles dans la suite [2] :

— Expression d'une loi Markovienne partant d'une probabilité quelconque comme un mélange de lois issues d'une Dirac : pour tout $A \in \mathcal{F}$
$$\mathbb{P}^\mu(A) = \int \mathbb{P}^x(A)\, d\mu(x).$$

— Propriété de Markov. Si $A \in \mathcal{F}_n = \sigma(X_0, \dots, X_n)$ et $B \in \mathcal{F}$, on a
$$\mathbb{P}^\mu(A \cap \theta^{-n}(B)) = \int \mathbb{1}_A \varphi_B(X_n)\, d\mathbb{P}^\mu, \text{ avec } \varphi_B(x) = \mathbb{P}^x(B).$$
En particulier
$$(A \subset \{X_n = x\}) \Longrightarrow \mathbb{P}^\mu(A \cap \theta^{-n}(B)) = \mathbb{P}^\mu(A)\mathbb{P}^x(B)$$

2. Ce qui suit ne surprendra pas les courageux qui ont cherché l'exercice 57.

et si on prend $A = \Omega$, on a

$$\mathbb{P}^\mu(\theta^{-n}(B)) = \int \mathbb{P}^x(B)\, d\mathbb{P}^\mu_{X_n}(x).$$

Démonstration. Posons $A = \{X_0 = x_0, \dots, X_n = x_n\}$ et $B = \{X_0 = y_0, \dots X_p, = y_p\}$. Dans ce cas, les vérifications des identités proposées ne sont ni très difficiles, ni très amusantes. On prendra soin de distinguer selon que $x_n = y_0$ ou non. Tout élément de $\mathcal{F}_n$ est réunion dénombrable disjointe d'éléments de type $A = \{X_0 = x_0, \dots, X_n = x_n\}$, idem pour B et $\mathcal{F}_p$. Ainsi, par sommation dénombrable, on obtient les formules pour A dans $\mathcal{F}_n$ quelconque et B dans $\mathcal{F}_p$ quelconque. La première formule est donc vérifiée dans la réunion des $\mathcal{F}_n$, qui est un Π-système qui engendre $\mathcal{F}$, or $\mathbb{P}_\mu$ et $A \mapsto \int \mathbb{P}^x(A)\, d\mu(x)$ sont deux probabilités sur $\mathcal{F}$; coïncidant sur un Π-système qui engendre $\mathcal{F}$, elles coïncident sur $\mathcal{F}$ tout entier. Passons à la propriété de Markov. Fixons $A \in \mathcal{F}_n$. Si $\mathbb{P}^\mu(A) = 0$, l'identité est évidente. Sinon,

$$B \mapsto \varphi(B) = \frac{\mathbb{P}^\mu(A \cap \theta^{-n}(B))}{\mathbb{P}^\mu(A)}$$

et

$$B \mapsto \psi(B) = \frac{\int \mathbb{1}_A \varphi_B(X_n)\, d\mathbb{P}^\mu}{\mathbb{P}^\mu(A)}$$

sont encore deux probabilités qui coïncident sur un Π-système qui engendre $\mathcal{F}$, donc qui coïncident sur $\mathcal{F}$ tout entier. ✓

Théorème 98. *Supposons que μ est invariante, au sens où $\mathbb{P}^\mu_{X_1} = \mu$. Alors $(E^{\mathbb{N}}, \mathcal{B}(E^{\mathbb{N}}), \mathbb{P}^\mu, \theta)$ est un système dynamique.*

Démonstration. D'après la propriété de Markov,

$$\begin{aligned}\mathbb{P}^\mu(\theta^{-1}(A)) &= \int \mathbb{P}^x(A)\, d\mathbb{P}^\mu_{X_1} \\ &= \int \mathbb{P}^x(A)\, d\mu(x) \\ &= \mathbb{P}^\mu(A).\end{aligned}$$

✓

Théorème 99. *Soit $\mathbb{P}^\mu$ la probabilité sur l'espace canonique associée à une chaîne de Markov irréductible sur S admettant μ comme probabilité invariante. Alors, le système $(S^{\mathbb{N}}, \mathcal{B}(S^{\mathbb{N}}), \mathbb{P}^\mu)$ est ergodique.*

Démonstration. Soit A un événement invariant tel que $\mathbb{P}^\mu(A) > 0$. On veut montrer que $\mathbb{P}^\mu(A) = 1$. Soit i tel que $\mathbb{P}^\mu(A, X_n = i) > 0$. Comme on l'a déjà vu, il existe $n \geq 1$ et A_ε tel que $A_\varepsilon \in \sigma(X_0, \dots, X_n)$ et $\mathbb{P}^\mu(A \Delta A_\varepsilon) \leq \varepsilon$. On en déduit que

$$\mathbb{P}^\mu(\theta^{-n}(A) \Delta \theta^{-n}(A_\varepsilon)) = \mathbb{P}^\mu(\theta^{-n}(A \Delta A_\varepsilon)) \leq \varepsilon.$$

On a

$$\begin{array}{rcl} \mathbb{P}^\mu(A, X_n = i) & = & \mathbb{P}^\mu(A, X_n = i, \theta^{-n}(A)) \\ & \leq & \mathbb{P}^\mu(A_\varepsilon, X_n = i, \theta^{-n}(A_\varepsilon)) + 2\varepsilon \\ & = & \mathbb{P}^\mu(A_\varepsilon, X_n = i)\mathbb{P}^i(A_\varepsilon) + 2\varepsilon \\ & \leq & \mathbb{P}^\mu(A, X_n = i)\mathbb{P}^i(A) + 4\varepsilon + \varepsilon^2. \end{array}$$

En faisant tendre ε vers 0 dans la dernière inégalité, on obtient

$$\mathbb{P}^\mu(A, X_n = i) \leq \mathbb{P}^\mu(A, X_n = i)\mathbb{P}^i(A),$$

d'où $\mathbb{P}^i(A) = 1$. On a donc montré

$$\forall i \in E \quad (\mathbb{P}^\mu(X_n = i, A) > 0) \Longrightarrow (\mathbb{P}^i(A) = 1).$$

Mais

$$\mathbb{P}^\mu(X_n = i, A) = \mathbb{P}^\mu(X_n = i, \theta^{-n}(A)) = \mathbb{P}^\mu(X_n = i)\mathbb{P}^i(A) = \mu(i)\mathbb{P}^i(A).$$

D'après le théorème 79, on en déduit

$$\forall i \in E \quad (\mathbb{P}^i(A) > 0) \Longrightarrow (\mathbb{P}^i(A) = 1).$$

Il existe i tel que $\mathbb{P}^\mu(X_n = i, A) > 0$ (sinon on aurait $\mathbb{P}^\mu(A) = 0$). Soit j un entier quelconque. Il existe n tel que $\mathbb{P}^j(X_n = i) > 0$.

$$\mathbb{P}^j(A) = \mathbb{P}^j(\theta^{-n}(A)) \geq \mathbb{P}^j(X_n = i, \theta^{-n}(A)) = \mathbb{P}^j(X_n = i)\mathbb{P}^i(A) > 0,$$

donc $\mathbb{P}^j(A) = 1$. Finalement, $\mathbb{P}^\mu(A) = \int \mathbb{P}^i(A)\, d\mu(i) = 1$. ✓

Théorème 100. *Soit $\mathbb{P}^\mu$ la probabilité sur l'espace canonique associée à une chaîne de Markov irréductible apériodique sur S admettant μ comme probabilité invariante. Alors, le système $(S^{\mathbb{N}}, \mathcal{B}(S^{\mathbb{N}}), \mathbb{P}^\mu, \theta)$ est mélangeant.*

Démonstration. On va utiliser le critère de mélange et le théorème de la probabilité stationnaire.

En effet, prenons $A \in \mathcal{F}_p$ et $B \in \mathcal{F}$ un événement quelconque. On va montrer que $\mathbb{P}^\mu(A, \theta^{-n}(B)) \to \mathbb{P}^\mu(A)\mathbb{P}^\mu(B)$.

Soit $n \geq p$. On peut écrire $\theta^{-n}(B) = \theta^{-p}(\theta^{-(n-p)}(B))$. D'après la propriété de Markov, on a ainsi

$$\mathbb{P}^\mu(A, \theta^{-n}(B)) = \int \mathbb{1}_A \varphi_{\theta^{-(n-p)}(B)}(X_p)\ d\mathbb{P}^\mu.$$

Il suffit donc de montrer que pour tout x, $\mathbb{P}^x(\theta^{-n}B)$ tend vers $\mathbb{P}^\mu(B)$ et on pourra conclure avec le théorème de convergence dominée. Or, d'après la propriété de Markov,

$$\mathbb{P}^x(\theta^{-n}(B)) = \int \mathbb{P}^y(B)\ d\mathbb{P}^x_{X_n}(y).$$

D'après le théorème de la probabilité stationnaire, $\mathbb{P}^x_{X_n}$ converge en loi vers μ, donc

$$\mathbb{P}^x(\theta^{-n}(B)) \to \int \mathbb{P}^y(B)\ d\mu = \mathbb{P}^\mu(B).$$

✓

Théorème 101. *Soit $(X_n)_{n \geq 0}$ une chaîne de Markov irréductible à valeurs dans l'ensemble dénombrable S et admettant la probabilité invariante μ. Pour toute fonction g μ-intégrable, on a presque sûrement*

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n-1} g(X_k) \to \int\ g\ d\mu.$$

En particulier, pour tout $x \in S$

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{1}_{\{X_k = x\}} \to \mu(x).$$

Démonstration. Si l'on note $\mathbb{P}^\mu$ la loi sur $S^{\mathbb{N}}$ d'une chaîne de Markov avec la même dynamique que (X_n) et démarrant avec la loi initiale μ, le quadruplet $(S^N, \mathcal{B}(S^N), \theta, \mathbb{P}^\mu)$ est un système dynamique. Ce système est ergodique d'après le théorème précédent. D'après le théorème ergodique de Birkhoff avec $f = g \circ X_0$, on a $\mathbb{P}^\mu$ presque sûrement

$$\frac{1}{n}\sum_{k=0}^{n-1} g(X_k) = \frac{1}{n}\sum_{k=0}^{n-1}(g \circ X_0) \circ \theta^k \to \int f \, d\mathbb{P}^\mu = \int g \, d\mu.$$

Notons

$$A = \left\{ \frac{1}{n}\sum_{k=0}^{n-1} g(X_k) \to \int g \, d\mu \right\}.$$

On a

$$0 = \mathbb{P}^\mu(A^c) = \int \mathbb{P}^x(A^c) \, d\mu(x),$$

donc $\mathbb{P}^x(A^c) = 0$ pour μ-presque tout x. Mais μ charge tous les points de S, donc pour tout x, $\mathbb{P}^x(A^c) = 0$. Finalement, pour tout ν, on a

$$\mathbb{P}^\nu(A^c) = \int \mathbb{P}^x(A^c) \, d\nu(x) = \int 0 \, d\nu(x) = 0,$$

d'où $\mathbb{P}^\nu(A) = 1$: ainsi, quelque soit la probabilité de départ, même si elle n'est pas invariante, on a bien $\frac{1}{n}\sum_{k=0}^{n-1} g(X_k)$ qui tend presque sûrement vers $\mu(x)$. ✓

9.4 Facteur d'un système dynamique

Théorème 102. *Soit* $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ *un système dynamique mesuré. Soit* $(\Omega', \mathcal{F}')$ *un espace mesuré et* T *une application de* Ω' *dans lui-même. On suppose qu'on a en outre une application mesurable* π *de* $(\Omega, \mathcal{F})$ *dans* $(\Omega', \mathcal{F}')$ *telle que*

$$\pi \circ \theta = T \circ \pi.$$

Alors $(\Omega', \mathcal{F}', \mathbb{P}_\pi, T)$ *est un système dynamique mesuré.*

On a le diagramme commutatif suivant :

$$\begin{tikzcd} \Omega \arrow[r, "\theta"] \arrow[d, "\pi"'] & \Omega \arrow[d, "\pi"] \\ \Omega' \arrow[r, "T"] & \Omega' \end{tikzcd}$$

De plus, si $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ *est ergodique, alors* $(\Omega', \mathcal{F}', \mathbb{P}_\pi, T)$ *est ergodique.*

Définition: On dit que $(\Omega', \mathcal{F}', \mathbb{P}_\pi, T)$ est un facteur du système dynamique $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$.

Démonstration. Soit $A \in \mathcal{F}'$. On a $\mathbb{P}_\pi(A) = \mathbb{P}(\pi^{-1}(A))$.
Comme $\pi^{-1}(A) \in \mathcal{F}$ et que $\mathbb{P}$ est invariante par θ, on a

$$\begin{aligned} \mathbb{P}_\pi(A) &= \mathbb{P}(\pi^{-1}(A)) = \mathbb{P}_\theta(\pi^{-1}(A)) \\ &= \mathbb{P}(\theta^{-1}(\pi^{-1}(A))) \\ &= \mathbb{P}((\pi \circ \theta)^{-1}(A)) = \mathbb{P}((T \circ \pi)^{-1}(A)) \\ &= \mathbb{P}(\pi^{-1}(T^{-1}(A)) = \mathbb{P}_\pi(T^{-1}(A)). \end{aligned}$$

Si A est invariant par T et que $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ est ergodique, alors on a successivement $\mathbb{1}_A \circ T = \mathbb{1}_A$, $\mathbb{1}_A \circ T \circ \pi = \mathbb{1}_A \circ \pi$, $\mathbb{1}_A \circ \pi \circ \theta = \mathbb{1}_A \circ \pi$, soit $\mathbb{1}_{\pi^{-1}(A)} \circ \theta = \mathbb{1}_{\pi^{-1}(A)}$. L'hypothèse d'ergodicité nous donne alors $\mathbb{P}(\pi^{-1}(A)) \in \{0, 1\}$, soit $\mathbb{P}_\pi(A) \in \{0, 1\}$. ✓

Exemple fondamental : On considère $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ un système dynamique mesuré quelconque, et on suppose qu'on a $f : \Omega \to \mathbb{R}$. On considère alors $\Omega' = \mathbb{R}^{\mathbb{N}}$, $\pi(\omega) = (f(\theta^n \omega))_{n \geq 0}$, avec T l'opérateur de translation sur $\mathbb{R}^{\mathbb{N}}$: $T(x) = (x_{n+1})_{n \geq 0}$. On a bien $\pi \circ \theta = T \circ \pi$ et le système $(\Omega', \mathcal{B}(\mathbb{R}^{\mathbb{N}}), \mathbb{P}_\pi, T)$ est un facteur du système dynamique $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$, qui est ergodique si le système initial l'est.
Exemple : Le doublement modulo 1. On prend $\Omega = \{0, 1\}^{\mathbb{N}}$, on le munit de sa tribu borélienne $\mathcal{F} = \mathcal{B}(\Omega)$ et de la mesure $\mathbb{P} = \text{Ber}(1/2)^{\otimes \mathbb{N}}$.

Pour l'opérateur θ de décalage, le système $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ est ergodique (on a même vu qu'il était mélangeant). Si l'on pose

$$\pi(\omega) = \sum_{n=0}^{+\infty} \frac{\omega_n}{2^{n+1}} \in [0,1[\text{ et } T(x) = \{2x\},$$

il n'est pas difficile de voir que $\pi \circ \theta = T \circ \pi$.

Ainsi, le système $([0,1[, \mathcal{B}([0,1[), \mathbb{P}_\pi, T)$ est un système dynamique. Comme les ω_i sont des variables aléatoires indépendantes suivant la loi de Bernoulli de paramètre $1/2$, on peut décrire la loi P_π : c'est la loi uniforme sur $[0,1[$, ou encore la restriction de la mesure de Lebesgue à $[0,1[$.

9.5 Théorème du retour, système induit

9.5.1 Théorème du retour de Poincaré

Théorème 103. *Soit $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ un système dynamique mesuré. Pour tout $A \in \mathcal{F}$, la suite $\theta^n x$ passe une infinité de fois dans A pour presque tout point x de A.*

Démonstration. On pose

$$N(x) = \sum_{n=0}^{+\infty} \mathbb{1}_A(\theta^n(x)).$$

ainsi que $Y(x) = \exp(-N(x))$, avec la convention $\exp(-\infty) = 0$. On pose enfin $Z(x) = Y(\theta x)$. Comme θ préserve la mesure, on a

$$\mathbb{E}(Z) = \int Y \circ \theta \; d\mathbb{P} = \int Y \; d\mathbb{P}_\theta = \int Y \; d\mathbb{P}_\theta = \mathbb{E}(Y).$$

Cependant, on vérifie facilement que $Y(x) = e^{-\mathbb{1}_A(x)} Z(x)$.

Comme $Y \leq Z$, Y et Z sont donc presque sûrement égales, ce qui veut dire que pour presque tout $x \in A$, $Z(x) = 0$, soit $N = +\infty$. ✓

9.5.2 Système induit

Soit $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ un système dynamique mesuré, A un événement de la tribu $\mathcal{F}$ avec $\mathbb{P}(A) > 0$.

Pour $x \in \Omega$, on note $n_A(x) = \inf\{n \geq 1; \theta^n x \in A\}$: c'est le temps de retour du système dans A.

Notons θ_A l'opérateur de Ω dans lui même défini par

$$\theta_A(x) = \begin{cases} \theta^{n_A(x)}(x) & \text{si } n_A(x) < +\infty \\ x & \text{sinon.} \end{cases}$$

On note $\mathbb{P}_A$ la mesure de probabilité définie sur la trace $\mathcal{F}_A$ de $\mathcal{F}$ sur A par $\mathbb{P}_A(C) = \frac{\mathbb{P}(C)}{\mathbb{P}(A)}$.

Théorème 104. $(A, \mathcal{F}_A, \mathbb{P}_A, \theta_A)$ *est un système dynamique.*

Démonstration. Soit $C \in \mathcal{F}$ avec $C \subset A$. Pour tout $n \geq 1$, on a

$$\mathbb{1}_{\{n_A=n+1\}} = (\mathbb{1}_{A^c}\mathbb{1}_{\{n_A=n\}}) \circ \theta.$$

D'où, en multipliant par $\mathbb{1}_C \circ \theta^{n+1}$:

$$\mathbb{1}_{\{n_A=n+1\}}(\mathbb{1}_C \circ \theta^{n+1}) = (\mathbb{1}_{A^c}(\mathbb{1}_{\{n_A=n\}}\mathbb{1}_C \circ \theta^n)) \circ \theta.$$

Puis, en prenant l'espérance et en utilisant l'invariance de $\mathbb{P}$ par θ

$$\mathbb{P}(n_A = n+1, \theta^{-(n+1)}(C)) = \mathbb{P}(A^c, n_A = n, \theta^{-n}(C)),$$

d'où la décomposition

$$\begin{aligned} \mathbb{P}(A^c, n_A = n, \theta^{-n}(C)) = \mathbb{P}(A, n_A = n+1, \theta^{-(n+1)}(C)) \\ + \mathbb{P}(A^c, n_A = n+1, \theta^{-(n+1)}(C)). \end{aligned} \tag{9.4}$$

D'autre part, on a

$$\begin{aligned} \mathbb{P}(C) &= \mathbb{P}(\theta^{-1}(C)) \\ &= \mathbb{P}(n_A = 1, \theta^{-1}(C)) \\ &= \mathbb{P}(n_A = 1, A, \theta^{-1}(C)) + \mathbb{P}(n_A = 1, A^c, \theta^{-1}(C)). \end{aligned}$$

Ainsi, avec (9.4), on montre par récurrence sur n que pour tout $n \geq 1$, on a

$$\mathbb{P}(C) = \left(\sum_{k=1}^{n} \mathbb{P}(n_A = k, A, \theta^{-k}(C))\right) + \mathbb{P}(n_A = n, A^c, \theta^{-1}(C)).$$

La série de terme général $\mathbb{P}(n_A = n)$ converge, donc le dernier terme de la somme tend vers 0 quand n tend vers l'infini et on a

$$\mathbb{P}(C) = \sum_{k=1}^{+\infty} \mathbb{P}(n_A = k, A, \theta^{-k}(C)),$$

d'où $\mathbb{P}(C) = \mathbb{P}(\theta_A^{-1}(C))$. ✓

Théorème 105. *Si* $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ *est un système dynamique ergodique, alors* $(A, \mathcal{F}_A, \mathbb{P}_A, \theta_A)$ *est un système dynamique ergodique.*

Démonstration. Soit $C \in \mathcal{F}_A$ un événement invariant pour θ_A. On va d'abord montrer que $\theta_A^{-1}(C)$ est un invariant pour θ, ce qui en fera un événement trivial sous $\mathbb{P}$. On va commencer par établir l'inclusion

$$\{\theta_A(x) \in C\} \subset \{\theta_A(\theta(x)) \in C\} \cup \{\theta(x) \in C \backslash \theta_A^{-1}(C)\}.$$

Si $\theta(x) \notin A$, alors $\theta_A(x) = \theta_A(\theta(x))$, donc

$$\{\theta_A(x) \in C, \theta(x) \notin A\} \subset \{\theta_A(\theta(x)) \in C\}.$$

Si, au contraire, $\theta(x) \in A$, on a $\theta_A(x) = \theta(x)$, d'où l'on déduit

$$\begin{aligned}\{\theta_A(x) \in C, \theta(x) \in A\} &\subset \{\theta(x) \in C\} \\ &\subset \{\theta_A(\theta(x)) \in C\} \cup \{\theta(x) \in C \backslash \theta_A^{-1}(C)\}.\end{aligned}$$

On a donc bien

$$\{\theta_A(x) \in C\} \subset \{\theta_A(\theta(x)) \in C\} \cup \{\theta(x) \in C \backslash \theta_A^{-1}(C)\}.$$

Cependant, l'invariance de $\mathbb{P}$ par θ donne

$$\mathbb{P}(\theta(x) \in C \backslash \theta_A^{-1}(C)) = \mathbb{P}(C \backslash \theta_A^{-1}(C)) = 0,$$

puisque C est invariant pour θ_A. Autrement dit,

$$\mathbb{1}_{\theta_A^{-1}(C)} \le \mathbb{1}_{\theta_A^{-1}(C)} \circ \theta \quad \mathbb{P}-\text{p.s.}$$

Comme θ préserve la mesure, on en déduit, en prenant l'espérance, que

$$\mathbb{1}_{\theta_A^{-1}(C)} = \mathbb{1}_{\theta_A^{-1}(C)} \circ \theta \quad \mathbb{P}-\text{p.s.}$$

Ainsi, $\theta_A^{-1}(C)$ est un invariant pour θ : comme θ est ergodique, on a $\mathbb{P}(\theta_A^{-1}(C)) = 0$ ou $\mathbb{P}(\theta_A^{-1}(C)) = 1$.

— Dans le premier cas, $\mathbb{P}(C) = \mathbb{P}(\theta_A^{-1}(C)) = 0$, d'où $\mathbb{P}_A(C) = 0$.
— Dans le second cas, on a $\mathbb{P}(x \in \Omega : \theta_A(x) \notin C) = 0$, a fortiori $\mathbb{P}(x \in A : \theta_A(x) \notin C) = 0$, soit $\mathbb{P}_A(\theta_A(x) \notin C) = 0$; comme θ_A préserve la mesure $\mathbb{P}_A$, $\mathbb{P}_A(A \backslash C) = \mathbb{P}_A(\theta_A(x) \in A \backslash C) = 0$.

Dans les deux cas $\mathbb{P}_A(C) \in \{0, 1\}$ et θ_A est ergodique. ✓

9.5.3 Théorème de Kac

Théorème 106. *Soit* $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ *est un système dynamique ergodique et* $A \in \mathcal{F}$ *avec* $\mathbb{P}(A) > 0$. *Alors,*

$$\int_A n_A \, d\mathbb{P}_A = \frac{1}{\mathbb{P}(A)}.$$

Démonstration. Comme θ_A est ergodique, d'après le théorème ergodique, pour $\mathbb{P}_A$-presque tout ω

$$\int_A n_A \, d\mathbb{P}_A = \lim_{n \to +\infty} \frac{1}{n} S_n(\omega),$$

où

$$S_n = \sum_{k=0}^{n-1} n_A \circ \theta_A^k.$$

D'autre part, comme θ est ergodique

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_A \circ \theta^k(\omega) = \mathbb{P}(A) \quad \mathbb{P}-\text{p.s.}$$

En particulier

$$\lim_{n\to+\infty} \frac{1}{S_n} \sum_{k=1}^{S_n} \mathbb{1}_A \circ \theta^k = \mathbb{P}(A) \quad \mathbb{P}-\text{p.s.}$$

Mais par construction,

$$\sum_{k=1}^{S_n} \mathbb{1}_A \circ \theta^k = n.$$

Ainsi $\frac{n}{S_n}$ tend presque sûrement vers $\mathbb{P}(A)$, d'où $\int_A n_A \, d\mathbb{P}_A = \frac{1}{\mathbb{P}(A)}$. ✓

Remarque: dans le cas d'une chaîne de Markov sur l'espace canonique, en appliquant la formule de Kac à l'événement $\{X_0 = x\}$, on retrouve la formule classique $\mu(x) = \frac{1}{\mathbb{E}^x T_x}$.

9.6 Application à la percolation de premier passage

La percolation de premier passage a été introduite par Hammersley et Welsh comme un modèle de propagation d'un fluide dans un média poreux. À chaque arête du graphe $\mathbb{Z}^d$, on associe une variable aléatoire positive qui représente le temps nécessaire pour traverser l'arête. Dans le cas le plus classique, on suppose que les temps de passage sont indépendants, distribués suivant une même loi ν admettant un moment d'ordre deux. [3] On peut alors définir le temps nécessaire pour suivre un chemin : c'est la somme des temps des arêtes qui le composent. On peut alors parler de temps minimal pour aller d'un point x à un point y : c'est la borne inférieure des temps des chemins qui vont de x à y.

Ainsi, cette famille de temps de passage induit une distance (aléatoire) $d(.,.)$ sur $\mathbb{Z}^d$

Théorème 107. *On suppose que les temps de passage des arêtes sont des variables aléatoires indépendantes identiquement distribuées suivant une*

3. On peut grandement affaiblir cette hypothèse, mais ce n'est pas notre propos ici.

loi ν avec un moment d'ordre 1 et telle que $\nu(\mathbb{R}^-) = 0$. Alors, il existe une norme μ sur $\mathbb{R}^d$ telle que pour tout $x \in \mathbb{Z}^d$, $d(0, nx)/n$ converge presque sûrement et dans L^1 vers $\mu(x)$.

Avant de faire la preuve, il convient de noter que les hypothèses présentées sont loin d'être minimales. On pourra par exemple se reporter au cours de Saint-Flour de Kesten [24] pour plus de renseignements.

Démonstration. Quelques remarques simples : évidemment $d(0, x) \geq 0$ pour tout x. Ensuite, pour tout $x \in \mathbb{Z}^d$, on peut construire un chemin déterministe de longueur $\|x\|_1$ de 0 à x : soit $\eta_1, \dots, \eta_{\|x\|_1}$ les nombres inscrits sur ces arêtes : on a

$$d(0, x) \leq \eta_1 + \cdots + \eta_{\|x\|_1},$$

d'où $0 \leq \mathbb{E}d(0, x) \leq m\|x\|_1$ avec $m = \mathbb{E}\eta_1$.

D'après l'inégalité triangulaire

$$d(0, (n+p)x) \leq d(0, nx) + d(nx, (n+p)x) = d(0, nx) + d(0, px) \circ \theta_x^n).$$

Posons $f_n = d(0, nx)$: les f_n sont dans L^1 et sont positives. On peut donc appliquer le théorème ergodique sous additif et il existe $\mu(x) \geq 0$ tel que $d(0, nx)/n$ converge presque sûrement et dans L^1 vers $\mu(x)$ – noter qu'on utilise ici le corollaire 34. D'après l'inégalité ci-dessus, on a

$$0 \leq \mu(x) \leq m\|x\|_1. \tag{9.5}$$

Soient x, y dans $\mathbb{Z}^d$. On a

$$d(0, n(x+y)) \leq d(0, nx) + d(nx, nx + ny) = d(0, 0, n) + d(0, ny) \circ \theta_{nx}$$

d'où en prenant l'espérance et en divisant par n :

$$\frac{\mathbb{E}d(0, n(x+y))}{n} \leq \frac{\mathbb{E}d(0, nx)}{n} + \frac{\mathbb{E}d(0, ny)}{n}.$$

Et, en faisant tendre n vers $+\infty$,

$$\forall x, y \in \mathbb{Z}^d \quad \mu(x+y) \leq \mu(x) + \mu(y). \tag{9.6}$$

En utilisant l'invariance par translation $\mathbb{E}d(0,nx) = \mathbb{E}d(-nx,0)$ mais $d(-nx,0) = d(0,nx)$, donc en divisant par n et en faisant tendre vers vers l'infini $\mu(x) = \mu(-x)$. Soit $p \geq 1$. La suite $(d(0,(np)x)/np)_{n\geq 1}$ est une suite extraite de $(d(0,nx)/n)_{n\geq 1}$ donc sa limite est $\mu(x)$. Mais $d(0,n(px))/n$ tend vers $\mu(px)$. Donc $\mu(px) = p\mu(x)$. Finalement

$$\forall x \in \mathbb{Z}^d \quad \forall n \in \mathbb{Z} \quad \mu(nx) = |n|\mu(x). \tag{9.7}$$

Cela permet de prolonger μ à $\mathbb{Q}^d$ par homogénéité en posant

$$\mu(\frac{p_1}{N},\dots,\frac{p_d}{N}) = \frac{\mu(p_1,\dots,p_d)}{N}.$$

On vérifie sans difficulté que l'inégalité triangulaire (9.6), l'homogénéité (9.7), et la "continuité" (9.5) sont encore vérifiées pour x,y dans $\mathbb{Q}^d$ et n dans $\mathbb{Q}$.

Grâce à l'inégalité triangulaire, on a pour tout

$$\forall x,y \in \mathbb{Q}^d \quad -\mu(y-x) \leq \mu(x)-\mu(y) \leq \mu(x-y),$$

d'où, grâce à 9.5

$$|\mu(x)-\mu(y)| \leq m\|x-y\|_1. \tag{9.8}$$

Ainsi, l'application $x \mapsto \mu(x)$ est une application uniformément continue sur $\mathbb{Q}^d$. Comme $\mathbb{Q}^d$ est une partie dense de $\mathbb{R}^d$, μ admet un prolongement unique continu sur $\mathbb{R}^d$.

Il est maintenant facile de voir que μ est une semi-norme.

Reste à voir que μ est une norme.

On utilise le lemme suivant :

Lemme 20.

$$\forall \varepsilon,\delta > 0\ \exists p_0 < 1\ \forall p \geq p_0\ \forall n \geq 1 \quad \mathcal{B}(n,p)([n(1-\varepsilon),+\infty[) \geq 1-\delta^n.$$

Démonstration. Soit $X \sim \mathcal{B}(n,p)$ et $\theta > 0$. On pose $Y = n - X$ et $q = 1-p$.

$$\begin{aligned}\mathbb{P}(X < n(1-\varepsilon)) &= \mathbb{P}(Y > n\varepsilon)\\ &= \mathbb{P}(e^{\theta Y} > e^{n\theta\varepsilon})\\ &\leq (\mathbb{E}e^{\theta Y})e^{-n\theta\varepsilon}\\ &\leq (((1-q)+qe^{\theta})e^{-\theta\varepsilon})^n.\end{aligned}$$

On choisit θ tel que $e^{-\theta\varepsilon} < \delta$. Alors, pour q suffisamment proche de 0 (c'est à dire p suffisamment proche de 1), on a

$$((1-q)+qe^{\theta})e^{-\theta\varepsilon} < \delta.$$

✓

Soit $\delta > 0$ quelconque. Son choix précis sera fait ultérieurement. Soit p_0 donné par le lemme 20 avec $\varepsilon = 1/2$. Comme $\mu(\mathbb{R}^-) = 0$, on peut trouver $\alpha > 0$ tel que $\mu([\alpha, +\infty[) \geq p_0$. Si je prends des variables aléatoires $\eta_1, \dots, \eta_n$ i.i.d. de loi μ, j'ai

$$\begin{aligned}\mathbb{P}(\eta_1 + \dots + \eta_n \leq \frac{\alpha}{2}n) &\leq \mathbb{P}(\mathbb{1}_{\{\eta_1 \geq \alpha\}} + \dots + \mathbb{1}_{\{\eta_n \geq \alpha\}} \leq n/2) \\ &\leq \delta^n.\end{aligned}$$

Notons

$$B_t = \{x \in \mathbb{Z}^d; d(0,x) < t\} \text{ et } \Lambda_t = \{x \in \mathbb{Z}^d; \|x\|_1 < t\}$$

$$\{B_t \not\subset \Lambda_{\frac{2}{\alpha}t}\} \subset \bigcup_{n \geq \frac{2}{\alpha}t} \bigcup_{\gamma \in C_n} \left\{\sum_{e \in \gamma} \eta_e \leq t\right\},$$

où C_n est la famille des parties de $\mathbb{E}^d$ qui sont le support d'un chemin sans recoupement partant de 0 et de longueur n.

D'où

$$\begin{aligned}\{B_t \not\subset \Lambda_{\frac{2}{\alpha}t}\} &\subset \bigcup_{n \geq \frac{2}{\alpha}t} \bigcup_{\gamma \in C_n} \left\{\sum_{e \in \gamma} \alpha \mathbb{1}_{\{\eta_e \geq \alpha\}} \leq t\right\} \\ &\subset \bigcup_{n \geq \frac{2}{\alpha}t} \bigcup_{\gamma \in C_n} \left\{\sum_{e \in \gamma} \mathbb{1}_{\{\eta_e \geq \alpha\}} \leq \frac{n}{2}\right\}.\end{aligned}$$

Ainsi

$$\begin{aligned}
\mathbb{P}(B_t \not\subset \Lambda_{\frac{2}{\alpha}t}) &\le \sum_{n\ge\frac{2}{\alpha}t} \sum_{\gamma\in C_n} \mathbb{P}\left(\sum_{e\in\gamma} \mathbb{1}_{\{\eta_e\ge\alpha\}} \le \frac{n}{2}\right) \\
&\le \sum_{n\ge\frac{2}{\alpha}t} |C_n|\delta^n \\
&\le \sum_{n\ge\frac{2}{\alpha}t} 2d(2d-1)^{n-1}\delta^n \\
&\le \frac{2d}{(2d-1)(1-(2d-1)\delta)}((2d-1)\delta)^{\frac{2}{\alpha}t-1}.
\end{aligned}$$

Ainsi, il existe des constantes C, β strictement positives telles que

$$\mathbb{P}(B_t \not\subset \Lambda_{\frac{2}{\alpha}t}) \le C\exp(-\beta t) \tag{9.9}$$

Soit $u \in \mathbb{Z}^d$. On a

$$\begin{aligned}
\mathbb{P}(d(0,nu) < \frac{\alpha}{2}n\|u\|_1) &\le \mathbb{P}(B_{\frac{\alpha}{2}n\|u\|_1} \not\subset \Lambda_{n\|u\|_1}) \\
&\le C\exp(-\beta\frac{\alpha}{2}n\|u\|_1).
\end{aligned}$$

D'après le lemme de Borel-Cantelli, $d(0,nu) \ge \frac{\alpha}{2}n\|u\|_1$ pour n assez grand : on en déduit que $\mu(u) \ge \frac{\alpha}{2}\|u\|_1$. Comme μ est homogène et continue, l'inégalité s'étend d'abord à $\mathbb{Q}^d$, puis à $\mathbb{R}^d$. Et comme $\|.\|_1$ est une norme, il en est de même pour μ.

✓

En utilisant les ingrédients de la preuve ci-dessus, on peut même démontrer le théorème de forme asymptotique :

Théorème 108. *On suppose que les temps de passage des arêtes sont des variables aléatoires indépendantes identiquement distribuées suivant une loi ν avec un moment d'ordre 2 et telle que $\nu(\mathbb{R}^-) = 0$. Alors, il existe une norme μ sur $\mathbb{R}^d$ telle que, pour tout $\varepsilon > 0$, il existe presque sûrement un $t_0(\omega)$ tel que pour $t \ge t_0$*

$$B_\mu((1-\varepsilon)t) \subset B_t \subset B_\mu((1+\varepsilon)t),$$

où l'on a noté

$$B_t = \{x \in \mathbb{Z}^d; d(0,x) < t\} \textit{ et } B_\mu(t) = \{x \in \mathbb{Z}^d; \mu(x) < t\}.$$

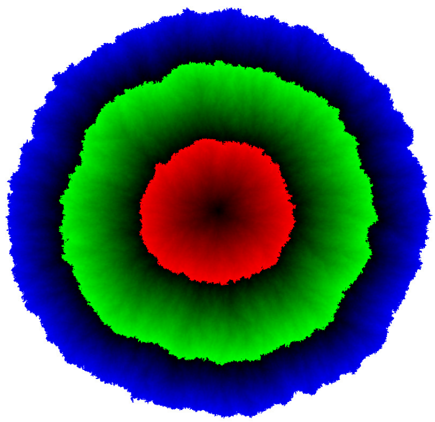

FIGURE 9.1 – Un ensemble B_t. Les points de même couleur sont atteints en même temps.

On va donner la preuve dans le cas plus simple où les temps de passage ont un moment exponentiel. Cela couvre en particulier le cas où les temps de passages sont des variables exponentielles (modèle de Richardson).

Démonstration. On va montrer que, presque sûrement,

$$\overline{\lim_{|x|\to+\infty}} \frac{|d(0,x) - \mu(x)|}{|x|} = 0.$$

C'est la forme analytique du théorème ; il n'est pas très difficile de voir qu'elle est équivalente à la forme géométrique.

Soit $\varepsilon > 0$. La famille $\mathbb{Q}\mathbb{Z}^d$ est dense dans $\mathbb{R}^d$, donc son image par $z \mapsto \frac{z}{|z|}$ est dense dans le "cercle" unité : $\{z : |z| = 1\}$: on peut trouver une famille finie $z_1, \dots z_n$ d'éléments de $\mathbb{Z}^d$ telle que pour tout z avec $|z| = 1$, il existe i avec $|\frac{z_i}{|z_i|} - z| \leq \varepsilon$. On pose $M = \max|z_i|$. Avec probabilité 1, on a pour tout i entre 1 et n, $\lim d(0,kz_i)/k = \mu(z_i)$. Soit $x \in \mathbb{Z}^d$ avec $\varepsilon|x| \geq M$: il existe $i = i(x)$ tel que $|\frac{z_i}{|z_i|} - \frac{x}{|x|}| \leq \varepsilon$, soit $|z_i - \frac{|z_i|}{|x|}x| \leq \varepsilon$ et, en notant n l'entier le plus proche de $\frac{|x|}{|z_i|}$, $|x - nz| \leq 2\varepsilon\|x\|$. Notant $z = z_{i(x)}$, on a

$$d(0,x) - \mu(x) = d(0,x) - d(0,nz) + d(0,nz) - \mu(nz) + \mu(nz) - \mu(x),$$

d'où

$$|d(0,x) - \mu(x)| \leq d(x,nz) + |d(0,nz) - \mu(nz)| + \mu(nz - x),$$

puis

$$\overline{\lim_{|x|\to+\infty}} \frac{|d(0,x) - \mu(x)|}{|x|} \leq \overline{\lim_{|x|\to+\infty}} \sup_{|x-y|\leq 2\varepsilon\|x\|} \frac{|d(x,y)|}{|x|} + 4m\varepsilon,$$

avec $m = \mathbb{E}[X_1]$.

On va montrer que $\overline{\lim\limits_{|x|\to+\infty}} \sup_{|x-y|\leq 2\varepsilon\|x\|} \frac{|d(x,y)|}{|x|} \leq 4m\varepsilon$, ce qui achèvera la preuve. Pour cela, avec le lemme de Borel-Cantelli, il suffit de montrer que

$$\sum_{x\in\mathbb{Z}^d\backslash\{0\}} \mathbb{P}\left(\sup_{|x-y|\leq 2\varepsilon\|x\|} \frac{|d(x,y)|}{|x|} > 4m\varepsilon|x|\right) < +\infty.$$

Mais, pour $\alpha > 0$, on a

$$\begin{aligned}
\mathbb{P}\left(\sup_{|x-y|\leq 2\varepsilon\|x\|} \frac{|d(x,y)|}{|x|} > 4m\varepsilon\right) &\leq \sum_{|x-y|\leq 2\varepsilon\|x\|} \mathbb{P}(e^{\alpha d(x,y)} > e^{\alpha 4m\varepsilon|x|}) \\
&\leq (1 + 4\varepsilon|x|)^d \mathbb{P}(e^{\alpha S} > e^{\alpha 4m\varepsilon|x|}),
\end{aligned}$$

où S est la somme de $p = \lceil 2\varepsilon\|x\|\rceil$ variables aléatoires indépendantes ayant la loi du temps de passage des arêtes. Ainsi,

$$\begin{aligned}
\mathbb{P}\left(\sup_{|x-y|\leq 2\varepsilon\|x\|} \frac{|d(x,y)|}{|x|} > 4m\varepsilon\right) &\leq (1 + 4\varepsilon|x|)^d \mathbb{E}[e^{\alpha X_1}]^p e^{-(p-1)2\alpha m} \\
&\leq (1 + 4\varepsilon|x|)^d e^{2\alpha m} (\mathbb{E}[e^{\alpha X_1}] e^{-2\alpha m})^p,
\end{aligned}$$

ce qui donne le résultat voulu en prenant α suffisamment petit.

✓

9.7 Une application à la théorie des nombres

9.7.1 Transformation de Gauss

Pour $\omega \in]0,1[$, on pose $T(\omega) = \{\frac{1}{\omega}\}$.

Théorème 109. *La mesure $\mathbb{P}$ sur $]0,1[$ de densité $f(x) = \frac{\mathbb{1}_{[0,1]}(x)}{(1+x)\log 2}$ est invariante par T.*

Démonstration. Posons $X(\omega) = \omega$. On compare les fonctions de répartition de X et T sur leur support commun $[0,1]$. Pour $t \in [0,1]$, on a aisément $F_X(t) = \frac{\log(1+t)}{\log 2}$, tandis que

$$\begin{aligned} F_T(t) &= \sum_{k=1}^{+\infty} \mathbb{P}(T \le t, \lfloor \frac{1}{X} \rfloor = k) = \sum_{k=1}^{+\infty} \mathbb{P}(k \le \frac{1}{X} \le k+t) \\ &= \sum_{k=1}^{+\infty} \mathbb{P}(X \in [\frac{1}{k+t}, \frac{1}{k}]) = \sum_{k=1}^{+\infty} \frac{1}{\log 2}\Big(\log(1+\frac{1}{k}) - \log(1+\frac{1}{k+t})\Big) \\ &= \frac{1}{\log 2}\sum_{k=1}^{+\infty} \log(1+k) - \log(k) + \log(k+t) - \log(k+t+1) \\ &= \sum_{k=1}^{+\infty} g_t(k) - g_t(k+1), \end{aligned}$$

avec $g_t(k) = \frac{\log(k+t)-\log(k)}{\log 2}$.

On obtient finalement $F_T(t) = g_1(t) = F_X(t)$. ✓

Dans la suite, Ω est l'ensemble des irrationnels de $[0,1]$. On note encore T la restriction de T à Ω. T est maintenant une application de Ω dans Ω, que l'on peut donc itérer.

La suite de la section est consacrée à la preuve des résultats énoncés dans le théorème suivant :

Théorème 110. — *Si on pose, pour tout $n \geq 1$, $X_n = [\frac{1}{T^{n-1}\omega}]$, on a pour tout $\omega \in \Omega$*

$$\omega = \lim_{n \to +\infty} f_n \text{ avec } f_n = \cfrac{1}{X_1 + \cfrac{1}{X_2 + \cfrac{1}{\cdots + \cfrac{1}{X_n}}}}.$$

La suite $(X_n(\omega))_{n \geq 1}$ est appelée développement de ω en fraction continue.

— *Le système $(\Omega, \mathcal{B}(\Omega), \mathbb{P}, T)$ est ergodique.*

— *La suite $(X_1.X_2 \dots .X_n)^{1/n}$ converge presque sûrement vers une constante K strictement positive, appelée constante de Khintchine. La constante K est donnée par*

$$K = \prod_{k=1}^{+\infty} \left(\frac{(k+1)^2}{k(k+2)} \right)^{\log_2 k}. \tag{9.10}$$

— $-\frac{1}{n} \log |\omega - f_n(\omega)|$ *converge presque sûrement vers* $\frac{\pi^2}{6 \log 2}$.

Remarque: Le « presque sûr » dans l'équation (9.10) est établi par rapport à la mesure de Gauss $\mathbb{P}$, mais c'est vrai également relativement à la restricttion à $[0, 1]$ de la mesure de Lebesgue ; les deux mesures étant à densité l'une par rapport à l'autre, elles ont les mêmes ensembles négligeables.

9.7.2 Ergodicité de la transformation de Gauss

En appliquant au point $\frac{1}{T^{n-1}(\omega)}$ l'identité $x = \lfloor x \rfloor + \{x\}$, on a

$$\frac{1}{T^{n-1}(\omega)} = X_n + \{\frac{1}{T^{n-1}(\omega)}\} = X_n + T^n(\omega),$$

soit

$$T^{n-1}\omega = \frac{1}{X_n + T^n(\omega)}.$$

Pour tout $\omega \in \Omega$, il existe un unique couple $(n, \omega') \in \mathbb{N}^* \times \Omega$ tel que

$$\frac{1}{\omega} = n + \omega'.$$

Par suite, si l'on pose $f_n(\omega') = \frac{1}{n+\omega'}$, les ensembles $(f_n(\Omega))_{n\geq 0}$ forment une partition de Ω.

Un raisonnement classique montre alors que pour tout $n \geq 1$, les ensembles $f_{a_1} \circ f_{a_2} \circ f_{a_n}(\Omega)$, où $(a_1, \dots, a_n) \in (\mathbb{N}^*)^n$ forment encore une partition de Ω.

Comme $T^{n-1}(\omega) = f_{X_n}(T^n(\omega))$, on a

$$\omega = f_{X_1}(T(\omega)) = f_{X_1}(f_{X_2}(T^2(\omega))) = \dots (f_{X_1} \circ \dots \circ f_{X_n})(T^n(\omega))$$

On en déduit que quel que soit $(a_1, \dots, a_n) \in (\mathbb{N}^*)^n$ et $I \subset \Omega$.

$$\{X_1 = a_1, \dots, X_n = a_n\} = (f_{a_1} \circ \dots \circ f_{a_n})(\Omega).$$

et

$$\{X_1 = a_1, \dots, X_n = a_n, T^n\omega \in I\} = (f_{a_1} \circ \dots \circ f_{a_n})(I).$$

Étudions plus précisément ces itérées. Rappelons qu'il y a un morphisme entre le groupe des matrices et celui des homographies : si l'on note pour

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \mathcal{H}_M(z) = \frac{az+b}{cz+d},$$

on a $\mathcal{H}_{AB} = \mathcal{H}_A \circ \mathcal{H}_B$. En particulier, si on pose

$$A_n = \begin{pmatrix} 0 & 1 \\ 1 & X_n \end{pmatrix}, \text{ on a } f_{X_1} \circ \dots \circ f_{X_n} = \mathcal{H}_{A_1.A_2 \times .A_n}.$$

Précisons la valeur de $M_n = A_1.A_2 \times .A_n$.

D'abord $\det M_n = \prod_{i=1}^n \det A_i = (-1)^n$, ensuite si je note

$$M_{n-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

j'ai

$$M_n = M_{n-1}A_n = \begin{pmatrix} a & b \\ c & f \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & a_n \end{pmatrix} = \begin{pmatrix} b & a + bX_n \\ d & c + dX_{n-1} \end{pmatrix}.$$

On remarque que la $1^{\text{ère}}$ colonne de M_n est la deuxième colonne de M_{n-1}. On peut donc simplifier l'écriture de la récurrence et écrire

$$\begin{aligned} M_n = \begin{pmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{pmatrix} &= \begin{pmatrix} p_{n-2} & p_{n-1} \\ q_{n-2} & q_{n-1} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & X_n \end{pmatrix} \\ &= \begin{pmatrix} p_{n-1} & p_{n-2} + X_n p_{n-1} \\ q_{n-1} & q_{n-2} + X_n q_{n-1} \end{pmatrix}. \end{aligned}$$

Comme $\frac{az+b}{cz+d} - \frac{b}{d} = \frac{(ad-bc)z}{d(cz+d)}$ et que $\det M_n = (-1)^n$,

$$\omega - \frac{p_n}{q_n} = \mathcal{H}_{M_n}(T^n(\omega)) - \frac{p_n}{q_n} = \frac{(-1)^n}{q_n(q_n/T^n(\omega) + q_{n-1})}. \tag{9.11}$$

Vue l'identité (9.11), on a

$$|\omega - \frac{p_n(\omega)}{q_n(\omega)}| = \frac{1}{q_n(\omega)(q_n(\omega)/T^n(\omega) + q_{n-1}(\omega))}.$$

Comme $X_{n+1}(\omega) = X_1(T^n(\omega)) = \lfloor 1/T^n(\omega) \rfloor$, on a

$$X_{n+1}(\omega) \leq \frac{1}{T^n(\omega)} \leq X_{n+1} + 1,$$

d'où

$$X_{n+1} q_n(\omega) + q_{n-1}(\omega) \leq \frac{q_n(\omega)}{T^n(\omega)} + q_{n-1}(\omega) \leq q_n(\omega)(X_{n+1} + 1) + q_{n-1}(\omega),$$

soit, avec les relations de récurrence,

$$q_{n+1}(\omega) \leq q_n(\omega)/T^n(\omega) + q_{n-1}(\omega) \leq q_{n+1}(\omega) + q_n(\omega)$$

et finalement

$$\frac{1}{q_n(q_n + q_{n+1})} \leq |\omega - \frac{p_n}{q_n}| \leq \frac{1}{q_n q_{n+1}}. \tag{9.12}$$

Comme $X_n \geq 1$, on a $p_n \geq p_{n-2} + p_{n-1}$, $q_n \geq q_{n-2} + q_{n-1}$.

Si l'on pose $F_0 = 0$, $F_1 = 1$, puis $F_{n+1} = F_n + F_{n-1}$, il est facile de voir par récurrence que pour tout $n \geq 0$, $p_n \geq F_n$ et $q_n \geq F_{n+1}$. Comme $F_n \sim C\varphi^n$ avec $\varphi = \frac{1+\sqrt{5}}{2}$, p_n/q_n converge vers ω.

Munissons maintenant Ω de la mesure $\mathbb{P}$. On veut montrer qu'il existe une constante C telle que

$$\forall A \in \mathcal{B}(\Omega) \quad \frac{1}{C}\mathbb{P}(A) \leq \mathbb{P}(T^n(\omega) \in A|\mathcal{F}_n) \leq C\mathbb{P}(A). \tag{9.13}$$

Grâce au théorème 93, il suffit de le montrer lorsque A est la trace d'un intervalle.

Comme la densité de $\mathbb{P}$ par rapport à la mesure de Lebesgue est bornée inférieurement et supérieurement, on a c tel que

$$c^{-1} \leq \frac{\mathbb{P}(T^n(\omega) \in A|X_1 = a_1, \dots X_n = a_n)}{\lambda(T^n(\omega) \in A|X_1 = a_1, \dots X_n = a_n)} \leq c.$$

D'après la formule

$$\lambda(T^n(\omega) \in A|\mathcal{F}_n) = \frac{\lambda(\mathcal{H}_{M_n}(A))}{\lambda(\mathcal{H}_{M_n}(\Omega))}.$$

$\mathcal{H}_{M_n}$ étant monotone, on a pour $A = [x, y]$

$$\frac{\lambda(\mathcal{H}_{M_n}(A))}{\lambda(\mathcal{H}_{M_n}(\Omega))} = \left|\frac{\mathcal{H}_{M_n}(y) - \mathcal{H}_{M_n}(x)}{\mathcal{H}_{M_n}(1) - \mathcal{H}_{M_n}(0)}\right|.$$

Un calcul simple montre que si $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, on a

$$\frac{\mathcal{H}_M(y) - \mathcal{H}_M(x)}{\mathcal{H}_M(1) - \mathcal{H}_M(0)} = (y - x)\frac{d(c+d)}{(cx+d)(cy+d)}.$$

Si $0 \leq c \leq d$ et $0 \leq x, y \leq 1$, alors

$$\frac{1}{2} \leq \frac{d}{c+d} = \frac{d(c+d)}{(c+d)^2} \leq \frac{d(c+d)}{(cx+d)(cy+d)} \leq \frac{d(c+d)}{d^2} = \frac{c+d}{d} \leq 2.$$

Ici, on a $c = q_{n-1}$ et $d = q_n$.

Comme $q_n - q_{n-1} = (q_{n-2} + X_n q_{n-1}) - q_n = q_{n-2} + (X_n - 1)q_{n-1} \geq 0$, on obtient l'inégalité

Ainsi

$$\frac{1}{2}|x-y| \leq \left|\frac{\mathcal{H}_{M_n}(y) - \mathcal{H}_{M_n}(x)}{\mathcal{H}_{M_n}(1) - \mathcal{H}_{M_n}(0)}\right| \leq 2|x-y|,$$

soit

$$\frac{1}{2}\lambda(A) \leq \lambda(T^n(\omega) \in A|\mathcal{F}_n) \leq 2\lambda(A),$$

ce qui nous donne 9.13 avec $C = 2c$. Comme $\omega = \lim_{n\to+\infty} \frac{p_n}{q_n}$ et que p_n et q_n sont $\mathcal{F}_n$ mesurable, on a $\vee_{n\geq 1}\mathcal{F}_n = \mathcal{B}(\Omega)$.

Soit A un événement invariant. Par le théorème de convergence des martingales, $\mathbb{P}(T^n(\omega) \in A|\mathcal{F}_n) = \mathbb{P}(A|\mathcal{F}_n)$ converge presque sûrement vers $\mathbb{1}_A$ et $\mathbb{P}(A|\mathcal{F}_n)\mathbb{1}_{A^c}$ converge presque sûrement vers 0.
Par convergence dominée $\mathbb{E}[\mathbb{P}(A|\mathcal{F}_n)\mathbb{1}_{A^c}]$ tend vers 0 Comme l'inégalité (9.13) entraîne

$$\mathbb{E}[\mathbb{P}(A|\mathcal{F}_n)\mathbb{1}_{A^c}] \geq \frac{1}{C}\mathbb{P}(A)\mathbb{E}[\mathbb{1}_{A^c}] = \frac{1}{C}\mathbb{P}(A)(1-\mathbb{P}(A),$$

on a $\mathbb{P}(A)(1-\mathbb{P}(A)) = 0$, donc $\mathbb{P}(A) \in \{0;1\}$.

9.7.3 La constante de Khintchine

Le théorème ergodique appliqué à $\log \circ X_1$ nous dit que

$$\frac{1}{n}\sum_{k=0}^{n-1} \log \circ X_1 \circ T^k = \frac{1}{n}\sum_{k=0}^{n-1} \log(X_n)$$

converge presque sûrement vers $\mathbb{E}_{\mathbb{P}} \log(X_1)$. Il suffit de vérifier que $\mathbb{E}_{\mathbb{P}} \log(X_1) < +\infty$. On a

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}(\log X_1) &= \frac{1}{\log 2} \int_0^1 \frac{\log\lfloor 1/x \rfloor}{1+x}\, dx. \\
&= \frac{1}{\log 2} \sum_{k=1}^{+\infty} \int_{1/(k+1)}^{1/k} \frac{\log\lfloor 1/x \rfloor}{1+x}\, dx \\
&= \frac{1}{\log 2} \sum_{k=1}^{+\infty} \int_{1/(k+1)}^{1/k} \frac{\log k}{1+x}\, dx \\
&= \frac{1}{\log 2} \sum_{k=1}^{+\infty} \int_0^1 \frac{\log k}{(k+\theta)(k+\theta+1)}\, d\theta \\
&\leq \frac{1}{\log 2} \sum_{k=1}^{+\infty} \frac{\log k}{k^2} < +\infty.
\end{aligned}$$

Précisons la somme :

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}(\log X_1) &= \frac{1}{\log 2} \sum_{k=1}^{+\infty} \log k \int_0^1 \frac{1}{k+\theta} - \frac{1}{k+\theta+1}\, d\theta \\
&= \frac{1}{\log 2} \sum_{k=1}^{+\infty} \log k \left(\log(\frac{k+1}{k}) - \log(\frac{k+2}{k+1} \right) \\
&= \sum_{k=1}^{+\infty} \log \left(\frac{(k+1)^2}{k(k+2)} \right)^{\log_2 k},
\end{aligned}$$

ce qui nous donne la valeur de la constante de Khintchine.

9.7.4 Un théorème de Paul Lévy

Passons à la preuve du dernier point. On va commencer par montrer, à l'aide du théorème ergodique, que $-\frac{1}{n} \log q_n(\omega)$ converge presque sûrement vers une constante (qui se trouve être $\frac{\pi^2}{12 \log 2}$) [4].

4. Selon Lévy, la convergence presque sûre de $-\frac{1}{n} \log q_n(\omega)$ vers une constante non précisée lui a été notifiée par Khintchine ; ses travaux [28] ont permis l'identification de la constante.

L'équation $\omega = \mathcal{H}_{M_n}(T^n(\omega))$ entraîne $T^n(\omega) = \mathcal{H}_{M_n^{-1}}(\omega)$. Notons que deux matrices proportionnelles induisent la même homographie, donc $\mathcal{H}_{M_n^{-1}} = \mathcal{H}_{(\mathrm{com}\ M_n)^*}$.

Or $(\mathrm{com}\ M_n)^* = \begin{pmatrix} q_n & -p_n \\ -q_{n-1} & p_{n-1} \end{pmatrix}$, donc

$$T^n(\omega) = \frac{q_n\omega - p_n}{-q_{n-1}\omega + p_{n-1}} = -\frac{q_n\omega - p_n}{q_{n-1}\omega - p_{n-1}},$$

ce qui nous donne

$$\begin{aligned} q_n(\omega)\omega - p_n(\omega) &= (q_0(\omega)\omega - p_0(\omega)) \prod_{i=1}^{n} \frac{q_i(\omega)\omega - p_i(\omega)}{q_{i-1}(\omega)\omega - p_{i-1}(\omega)} \\ &= \omega \prod_{i=1}^{n} (-T^i(\omega)) = (-1)^n \prod_{i=0}^{n} T^i(\omega). \end{aligned}$$

On en déduit $|q_n(\omega)\omega - p_n(\omega)| = \prod_{i=1}^n T^i(\omega)|$, puis

$$\frac{1}{n} \log(|q_n(\omega)\omega - p_n(\omega)|) = -\frac{1}{n} \sum_{i=1}^{n} - \log T^i(\omega).$$

Sur $[0,1]$, $-\log\omega$ est positive. Calculons l'intégrale par rapport à $\mathbb{P}$. On a

$$\begin{aligned} &\int \log(\omega)\ d\mathbb{P}(\omega) \\ &= \frac{1}{\log 2}\int_0^1 \frac{\log x}{1+x}\ dx = \frac{1}{\log 2}\int_0^1 -\frac{\log(1+x)}{x}\ dx \qquad (I.P.P.) \\ &= \frac{1}{\log 2}\int_0^{+\infty} \log(1+e^{-u})\ du = \frac{1}{2\log 2}\int_0^{+\infty} -\log(1-e^{-u})\ du \qquad (*) \\ &= \frac{1}{2\log 2}\int_0^{+\infty} \sum_{k=1}^{+\infty} \frac{e^{-ku}}{k}\ du = \frac{1}{2\log 2}\sum_{k=1}^{+\infty}\frac{1}{k^2} = \frac{\pi^2}{12\log 2}. \end{aligned}$$

Pour (*), on utilise $\log(1+e^{-u}) = \log(1-e^{-2u}) - \log(1-e^{-u})$.

Ainsi le théorème ergodique entraîne que $\frac{1}{n}\log(|q_n(\omega)\omega - p_n(\omega)|)$ converge $\mathbb{P}$-presque sûrement vers $-\frac{\pi^2}{12\log 2}$.

En multipliant l'inégalité (9.12) par q_{n+1}, on obtient

$$\frac{1}{2} = \frac{q_{n+1}}{q_{n+1}(\omega) + q_{n+1}(\omega)} \leq \frac{q_{n+1}}{q_n(\omega) + q_{n+1}(\omega)}$$
$$\leq q_{n+1}(\omega)|q_n(\omega)\omega - p_n(\omega)| \leq 1,$$

d'où

$$-\log 2 \leq \log q_{n+1}(\omega) + \log|q_n(\omega)\omega - p_n(\omega)| \leq 0,$$

et donc

$$\frac{\log q_{n+1}}{n+1} = -\frac{\log|q_n(\omega)\omega - p_n(\omega)|}{n+1} + o(1),$$

ce qui, vu la question précédente, entraîne que $\frac{\log q_{n+1}}{n+1}$ tend presque sûrement vers $\frac{\pi^2}{12\log 2}$, et par décalage d'indice, $\frac{\log q_n}{n}$ aussi. Par ailleurs,

$$\frac{1}{q_n q_{n+2}} \leq \frac{1}{q_n(\omega)(q_n(\omega) + q_{n+1}(\omega))} \leq |\omega - \frac{p_n(\omega)}{q_n(\omega)}|$$
$$\leq \frac{1}{q_n(\omega)q_{n+1}(\omega)} \leq \frac{1}{q_n(\omega)^2},$$

donc

$$-\frac{(\log q_n)}{n} - \frac{\log q_{n+2}}{n} \leq \frac{\log|\omega - \frac{p_n(\omega)}{q_n(\omega)}|}{n} \leq -2\frac{\log q_n}{n},$$

et par encadrement,

$$\frac{\log|\omega - \frac{p_n(\omega)}{q_n(\omega)}|}{n} \to -2 \times \frac{\pi^2}{12\log 2} = -\frac{\pi^2}{6\log 2} \quad \mathbb{P} - \text{p.s.}$$

On a fini de démontrer les résultats annoncés.

Faisons une remarque, afin de conclure cet ouvrage en entrouvrant une dernière porte : on a

$$\log \mathbb{P}(X_1 = X_1(\omega), \ldots, X_n = X_n(\omega))$$
$$= \log \lambda(X_1 = X_1(\omega), \ldots, X_n = X_n(\omega)) + O(1).$$

Or

$$\begin{aligned}
&\lambda(X_1 = X_1(\omega), \dots, X_n = X_n(\omega)) \\
&= \lambda(\mathcal{H}_{M_n}([0,1]) \\
&= |\mathcal{H}_{M_n}(1) - \mathcal{H}_{M_n}(0)| = \frac{1}{q_n(q_n + q_{n-1})}.
\end{aligned}$$

Comme $\frac{1}{q_n q_{n+1}} \leq \frac{1}{q_n(q_n+q_{n-1})} \leq \frac{1}{q_n^2}$, on a

$$-\log q_n - \log q_{n+1} \leq \log \lambda(\mathcal{H}_{M_n}([0,1]) \leq -2\log q_n.$$

On en déduit que

$$\begin{aligned}
\lim_{n\to+\infty} -\frac{1}{n}\log \mathbb{P}(X_1 = X_1(\omega), \dots, X_n = X_n(\omega)) &= \lim_{n\to+\infty} \frac{2}{n}(\log q_n) \\
&= \frac{\pi^2}{6\log 2}.
\end{aligned}$$

La quantité

$$\lim_{n\to+\infty} -\frac{1}{n}\log \mathbb{P}(X_1 = X_1(\omega), \dots, X_n = X_n(\omega))$$

est très étudiée en théorie ergodique. C'est l'entropie du système dynamique. Elle joue un rôle important dans la classification des systèmes dynamiques. Pour en savoir plus, on pourra, dans la littérature francophone, consulter le livre de Coudène [11] ou les notes de cours de de La Rue [12]. La littérature anglophone contient de nombreux livres consacrés à la théorie ergodique, l'ouvrage le plus souvent conseillé aux débutants est le cours de Walters [39]. Pour une lecture plus orientée vers la théorie des nombres, on pourra se tourner vers Einsiedler et Ward [16].

Ainsi, nous venons de démontrer [5] que l'entropie de la transformation T vaut $H(T) = \frac{\pi^2}{6\log 2}$.

5. Enfin, cela dépend de la définition de l'entropie. En toute rigueur, il faudrait encore utiliser le théorème de Kolmogorov-Sinaï et le théorème de Shannon-Mc Millan, mais ceci est une autre histoire...

9.8 Exercices sur les théorèmes ergodiques

9.8.1 Exercices corrigés

Exercice 113. Soit $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ un système dynamique. Montrer que l'ensemble des événements invariants du système forme une tribu.

Exercice 114. Soit A une application affine de $\mathbb{R}^d$ dans lui-même, avec $|\det A| = 1$. Pour $x \in \mathbb{R}^d$, on pose $\pi(x) = (\{x_1\}, \dots, \{x_d\})$, puis

$$\tilde{A} = \pi \circ A.$$

On pose $C = [0, 1[^d$. On note $\mathbb{P}$ la loi uniforme sur C. Montrer que $\tilde{A}$ préserve $\mathbb{P}$.

Lorsque $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$, $\tilde{A}$ est connue sous le nom de transformation du chat d'Arnold.

Exercice 115. Soit μ la mesure invariante d'une chaîne de Markov irréductible sur un ensemble S dénombrable. On suppose que le système $(S^{\mathbb{N}}, \mathcal{B}(S^{\mathbb{N}}), \mathbb{P}^\mu, \theta)$ est mélangeant. Montrer que la chaîne est irréductible.

Exercice 116. *Théorème de Kesten–Spitzer–Whitman.*
Soit G un groupe d'élément neutre e. Son produit est noté multiplicativement. On suppose que $\mathbb{P}$ est une mesure de probabilité sur $G^{\mathbb{Z}}$ telle que $(G^{\mathbb{Z}}, \mathcal{P}(G^{\mathbb{Z}}), \mathbb{P}, \theta)$ est un système dynamique ergodique. On note X_k la projection sur la k-ième composante, de sorte que les variables $(X_n)_{n\in\mathbb{Z}}$ sont reliées par la relation $X_{n+j} = X_j \circ \theta^n$.

On considère la marche aléatoire définie $(S_n)_{n\geq 0}$ définie par $S_0 = e$, puis, pour $n \geq 0$, par la récurrence $S_{n+1} = S_n X_{n+1}$. De même, la récurrence $S_n = S_{n+1} X_{n+1}^{-1}$ permet de définir S_n pour $n < 0$.

On s'intéresse au nombre de points vérifiés entre les temps 0 et n :

$$R_n(\omega) = |\{S_k(\omega); 0 \leq k \leq n-1\}.$$

1. On pose $N_k = \prod_{i=0}^{k-1} \mathbb{1}_{\{S_i \neq S_k\}}$. Montrer que

$$\lim_{k\to+\infty} \mathbb{E}(N_k) = \mathbb{P}(\forall n < 0; S_n \neq e).$$

On note désormais $\alpha = \mathbb{P}(\forall n < 0; S_n \neq e)$.

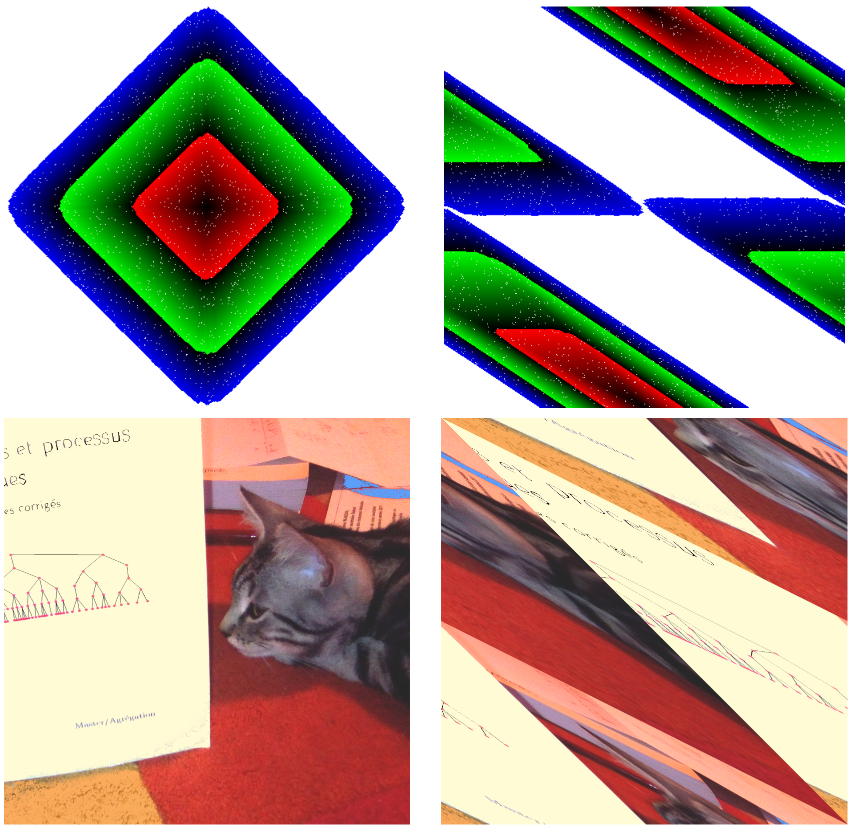

FIGURE 9.2 – Effet de la transformation du chat

2. Montrer que $\lim_{n\to+\infty} \frac{\mathbb{E}[R_n]}{n} = \alpha$.
3. Montrer soigneusement que pour tous $n, p \geq 0$, on a

$$R_{n+p} \leq R_n + R_p \circ \theta^n.$$

   En déduire que $\frac{R_n}{n}$ converge presque sûrement et dans L^1 vers α.
4. On dit que le système est réversible si $\mathbb{P}$ est laissée invariante par $x \mapsto (x_{-n})_{n\in\mathbb{Z}}$. Montrer que dans ce cas $\alpha = \mathbb{P}(\forall n > 0; S_n \neq e)$. Que trouve-t-on si $\mathbb{P}$ est de la forme $\mathbb{P} = \mu^{\otimes\mathbb{Z}}$?

9.8.2 Exercices non corrigés

Exercice 117. Soit $(X_n)_{n\geq 0}$ un processus stationnaire défini sur $(\Omega, \mathcal{F}, \mathbb{P})$. On dit que $(X_n)_{n\geq 0}$ est ergodique si $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}), \mathbb{P}_X, \theta)$ est ergodique. Montrer qu'une suite de variables aléatoires iid est ergodique.

Exercice 118. Soit $(X_n)_{n\geq 0}$ un processus stationnaire ergodique ; avec $\mathbb{E}[|X_0|] < +\infty$. Que peut-on dire du comportement asymptotique de la suite $\frac{1}{n}\sum_{k=0}^{n-1} X_k$?

Exercice 119. Soit $(X_n)_{n\geq 0}$ un processus stationnaire ergodique ; f une application mesurable de $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}))$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. On pose $Y_n = f(X_n, X_{n+1}, \dots,)$. Montrer que la suite $(Y_n)_{n\geq 0}$ forme un processus stationnaire ergodique.

Exercice 120. Soit $(X_n)_{n\geq 0}$ une suite de variables aléatoires de Bernoulli de paramètre $p > 0$. On note $Y_n = \inf\{k \geq 0; X_{n+k} = 1\}$. Montrer que $\frac{1}{n}\sum_{k=0}^{n-1} Y_k$ converge vers une limite que l'on déterminera.

Exercice 121. Soit $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ un système ergodique. On dit que $B \in \mathcal{F}$ est un événement invariant strict si $\theta^{-1}(A) = A$. Montrer que pour tout événement invariant A, il existe un événement invariant strict B tel que $\mathbb{P}(A\Delta B) = 0$. En déduire que le système est ergodique si et seulement si pour tout événement invariant strict B, $\mathbb{P}(B) \in \{0, 1\}$.

Exercice 122. Soit $(X_n)_{n\geq 1}$ un processus gaussien stationnaire centré. On suppose que pour tout entier $n \geq 1$, la loi de $(X_1, \dots X_n)$ n'est pas dégénérée (elle admet donc une densité par rapport à la mesure de Lebesgue sur $\mathbb{R}^n$). On note ρ la fonction de covariance. Montrer que $(X_n)_{n\geq 1}$ est mélangeant si et seulement si $\lim_{n\to+\infty} \rho(n) = 0$.

Exercice 123. *Décomposition en composantes ergodiques.*
Soit Ω un espace polonais. On note $\mathcal{F}$ sa tribu borélienne.

1. Montrer qu'il existe une algèbre $\mathcal{A}$ dénombrable qui engendre la tribu $\mathcal{F}$ (on pourra fabriquer cette algèbre à l'aide d'un π-système formé d'ouverts).

2. Soit μ une mesure de probabilité sur $(\Omega, \mathcal{F})$, et T une transformation mesurable de Ω dans lui-même. On suppose que pour tout $A \in \mathcal{A}$, $\mu(T^{-1}(A)) = \mu(A)$ et

$$\frac{1}{n}\sum_{k=0}^{n-1} \mathbb{1}_A \circ T^k \to \mu(A) \quad \mu - \text{p.s.}$$

   Montrer que $(\Omega, \mathcal{F}, \mu, T)$ est un système dynamique ergodique.

3. On suppose maintenant que $(\Omega, \mathcal{F}, \mathbb{P}, T)$ est un système dynamique. On note $\omega \mapsto \mathbb{P}_\omega^{\mathcal{I}}$ la loi conditionnelle de $\mathbb{P}$ sachant la tribu des invariants.

   (a) Soit $A \in \mathcal{F}$. Montrer que pour $\mathbb{P}$-presque tout ω,

$$\mathbb{P}_\omega^{\mathcal{I}}(A) = \mathbb{P}_\omega^{\mathcal{I}}(T^{-1}(A)),$$

   puis que pour $\mathbb{P}$-presque tout ω, $(\Omega, \mathcal{F}, \mathbb{P}_\omega^{\mathcal{I}}, T)$ est un système dynamique.

   (b) Soit $A \in \mathcal{A}$. On pose

$$\ell_A = \varliminf_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} \mathbb{1}_A \circ T^k \text{ et } L_A = \varlimsup_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} \mathbb{1}_A \circ T^k.$$

   Montrer que pour $\mathbb{P}$-presque tout ω,

$$\mathbb{P}_\omega^{\mathcal{I}}(\ell_A = L_A) = 1, \text{ et } L_A = \mathbb{E}_\omega^{\mathcal{I}}(L_A) \quad \mathbb{P}_\omega^{\mathcal{I}} - \text{p.s.}$$

   (c) Montrer enfin que pour $\mathbb{P}$-presque tout ω, $(\Omega, \mathcal{F}, \mathbb{P}_\omega^{\mathcal{I}}, T)$ est un système dynamique ergodique.