

Année universitaire 2008-2009

UNIVERSITÉ DE NANCY 1

Olivier GARET

Modèles probabilistes sur réseaux
(Master 2ème année – cours spécialisé)



Table des matières

Table des matières	i
1 Théorèmes ergodiques	1
1.1 Définitions	1
1.1.1 mesures invariantes, fonctions invariantes	1
1.1.2 ergodicité, mélange	2
1.1.3 sous-additivité	4
1.2 Théorème de Birkhoff	4
1.2.1 Lemme sous-additif maximal	4
1.2.2 Théorème de Birkhoff	5
1.2.3 Application aux chaînes de Markov	8
1.2.4 Théorème ergodique multidimensionnel	12
1.3 Théorème du retour, système induit	12
1.3.1 Théorème du retour de Poincaré	12
1.3.2 Système induit	12
1.3.3 Théorème de Kac	14
1.4 Théorème ergodique sous-additif	15
1.5 Application à la percolation de premier passage	19
2 Couplage, ordre stochastique	25
2.1 Généralités	25
2.1.1 Ordre stochastique	25
2.1.2 Couplage	26
2.1.3 Quelques propriétés	27
2.1.4 Un contre-exemple	29
2.2 Dynamiques markoviennes monotones	30
2.2.1 Processus de naissance et de mort	31
2.3 Applications à la percolation	32
2.3.1 Transition de percolation	32
2.3.2 Formule de Russo	34
2.4 Théorème de Holley – Inégalités FKG	35

TABLE DES MATIÈRES

2.4.1	Théorème de Holley	35
2.4.2	Inégalité FKG	37
2.4.3	Une inégalité à la Russo	39
3	Quelques outils utiles	41
3.1	L'argument de modification	41
3.2	La technique de renormalisation	42
3.3	Retour sur l'ordre stochastique : théorème de Liggett-Schonmann- Stacey	45

Chapitre 1

Théorèmes ergodiques

1.1 Définitions

1.1.1 mesures invariantes, fonctions invariantes

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 (Ω, \mathcal{F}) 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 (Ω, \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.

On dit qu'un événement A est un invariant du système $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ si $\mathbb{P}(A \Delta \theta^{-1}(A)) = 0$, où, de manière équivalente si $\mathbb{1}_A = \mathbb{1}_A \circ \theta$ \mathbb{P} p.s.. La famille des événements invariants forme une tribu (laissé en exercice), que l'on note souvent \mathcal{I} .

On dit qu'une application mesurable f de (Ω, \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.

1.1. DÉFINITIONS

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 n $f \circ \theta^n = f \circ \theta^{n+1}$, donc presque sûrement $f \circ \theta^n = f$ pour tout n .

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

1.1.2 ergodicité, mélange

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

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

Théorème 1. *Tout système mélangeant est ergodique.*

Démonstration. Soit A un événement invariant. On a vu que pour tout n $\mathbb{1}_A = \mathbb{1}_A \circ \theta^n$ presque sûrement, donc $\mathbb{1}_{A \cap \theta^{-n}(A)} = \mathbb{1}_A \mathbb{1}_{\theta^{-n}(A)} = \mathbb{1}_A^2 = \mathbb{1}_A$ presque sûrement, 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 $\mathbb{P}(A) \in \{0, 1\}$. \square

Théorème 2. *Étant donné un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, pour qu'une transformation $\theta : (\Omega, \mathcal{F}) \rightarrow (\Omega, \mathcal{F})$ conservant la mesure \mathbb{P} soit mélangeante, il suffit qu'il existe une algèbre \mathcal{F}_0 de parties de Ω engendrant la tribu \mathcal{F} et telle que :*

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

Démonstration. Soit $A, B \in \mathcal{F}$. Si $\mathbb{P}(A) = 0$ ou $\mathbb{P}(B) = 0$, alors $\mathbb{P}(A \cap \theta^{-n}(B)) = 0$ pour tout entier $n \geq 1$. 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 un résultat classique de théorie de la mesure¹, on a l'existence de A_ε et B_ε dans \mathcal{F}_0 tels que $\mathbb{P}(A \Delta A_\varepsilon) \leq \varepsilon$ et $\mathbb{P}(B \Delta B_\varepsilon) \leq \varepsilon$.

On a

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

Si x_1, x_2, y_1, y_2 sont de module plus petit que 1, $|x_1 y_1 - x_2 y_2| \leq |x_1 - x_2| + |y_1 - y_2|$. 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

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

Ainsi on a pour tout 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

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

Comme ε est arbitraire, cela donne le résultat voulu. □

Corollaire 1. 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 $x \in \mathbb{Z}^d \setminus \{0\}$, le système $(\Omega, (\mathcal{B}(\mathbb{R}^n))^{\otimes \mathbb{Z}^d}, \nu^{\otimes \mathbb{Z}^d}, \theta_x)$ est ergodique.

¹Si μ est une mesure finie sur \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 $\mu(A \Delta A') \leq \varepsilon$. En effet, il suffit de montrer que les ensembles A qui peuvent s'approcher ainsi forment une tribu, ce qui n'est pas très difficile.

1.2. THÉORÈME DE BIRKHOFF

Démonstration. Il suffit d'appliquer le théorème précédent à l'algèbre 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. \square

1.1.3 sous-additivité

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

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

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

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

est sous-additive (en fait exactement additive).

1.2 Théorème de Birkhoff

Le théorème de Birkhoff-Von Neumann est le résultat de base de la théorie ergodique. La preuve présentée ici est inspirée du texte de Kingman [Kin76]. Comme la plupart des preuves, elle se base sur un lemme sous-additif maximal pour les fonctions sous-additives. La preuve de ce lemme est presque un décalque de la preuve célèbre de Garsia.

1.2.1 Lemme sous-additif maximal

Théorème 3. *Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions sous-additives intégrables. On pose*

$$A = \{\exists n \in \mathbb{N}^* f_n \geq 0\}.$$

Alors

$$\int_A f_1 \geq 0.$$

Démonstration. Posons $M_n = \max(f_1, \dots, f_n)$. On a

$$\begin{aligned}
 M_{n+1} &= \max(f_1, f_2, \dots, f_{n+1}) \\
 &\leq \max(f_1, f_1 + f_1 \circ \theta, \dots, f_1 + f_n \circ \theta) \\
 &\leq f_1 + \max(0, f_1 \circ \theta, \dots, f_n \circ \theta) \\
 &= f_1 + \max(0, \max(f_1, \dots, f_n) \circ \theta) \\
 &= f_1 + \max(0, M_n) \circ \theta
 \end{aligned}$$

Maintenant

$$\begin{aligned}
 \mathbb{E}[M_n \mathbb{1}_{\{M_n \geq 0\}}] &\leq \mathbb{E}[M_{n+1} \mathbb{1}_{\{M_n \geq 0\}}] \\
 &\leq \mathbb{E}[f_1 \mathbb{1}_{\{M_n \geq 0\}}] + \mathbb{E}[\max(0, M_n) \circ \theta \mathbb{1}_{\{M_n \geq 0\}}] \\
 &\leq \mathbb{E}[f_1 \mathbb{1}_{\{M_n \geq 0\}}] + \mathbb{E}[\max(0, M_n) \circ \theta] \\
 &= \mathbb{E}[f_1 \mathbb{1}_{\{M_n \geq 0\}}] + \mathbb{E}[\max(0, M_n)] \\
 &= \mathbb{E}[f_1 \mathbb{1}_{\{M_n \geq 0\}}] + \mathbb{E}[M_n \mathbb{1}_{\{M_n \geq 0\}}]
 \end{aligned}$$

D'où $\mathbb{E}[f_1 \mathbb{1}_{\{M_n \geq 0\}}] \geq 0$.

Cependant, $f_1 \mathbb{1}_{\{M_n \geq 0\}}$ tend presque sûrement vers $f_1 \mathbb{1}_{\mathbb{A}}$, donc le théorème de convergence dominée permet de conclure. \square

1.2.2 Théorème de Birkhoff

Théorème 4. 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 \bar{f} θ -invariante ($f \circ \theta = f$) telle que

$$\frac{S_n}{n} \rightarrow \bar{f},$$

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

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

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

1.2. THÉORÈME DE BIRKHOFF

Démonstration. – Étape 1 : convergence presque sûre

Il est facile de voir que (S_n) est sous-additive (en fait additive). Posons

$$\bar{f} = \overline{\lim}_{n \rightarrow +\infty} \frac{S_n}{n}. \text{ On a également } \bar{f} = \overline{\lim}_{n \rightarrow +\infty} \frac{S_{n+1}}{n}. \text{ Comme } S_{n+1} = f + S_n \circ \theta, \text{ donc on a}$$

$$\bar{f} = \overline{\lim}_{n \rightarrow +\infty} \frac{S_{n+1}}{n} = \overline{\lim}_{n \rightarrow +\infty} S_n \circ \theta = \bar{f} \circ \theta.$$

De même, si l'on pose $\underline{f} = \underline{\lim}_{n \rightarrow +\infty} \frac{S_n}{n}$, on a $\underline{f} = \underline{f} \circ \theta$.

Soit b un réel quelconque tel que $\mathbb{P}(\bar{f} > b) > 0$. Posons $B = \{\bar{f} > b\}$ et définissons \mathbb{P}_B par $\mathbb{P}_B(A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$. Comme \bar{f} est invariante par θ , B est invariant par θ , c'est à dire que $\mathbb{1}_B = \mathbb{1}_B \circ \theta$. On en déduit aisément que $(\Omega, \mathcal{F}, \mathbb{P}_B, \theta)$ est un système dynamique.

Si je pose $g_n = S_n - nb$, il est facile de voir que (g_n) est encore additif, de sorte que d'après le théorème ergodique maximal

$$\int_A g_1 d\mathbb{P}_B \geq 0,$$

avec $A = \{\exists n \geq 0; g_n \geq 0\}$. Cependant $B \subset A$ donc $\mathbb{P}_B(A) = 1$, d'où $\int g_1 d\mathbb{P}_B \geq 0$, soit

$$\int_B f_1 d\mathbb{P} \geq b\mathbb{P}(B).$$

Comme le membre de gauche est borné par $\mathbb{E}|f|$, cela montre que $\mathbb{P}(\bar{f} > t) = O(1/t)$, en particulier $\mathbb{P}(\bar{f} = +\infty) = 0$.

De la même manière, en travaillant avec $h_n = na - S_n$, on peut voir que si $B' = \{\underline{f} < a\}$, on a

$$\int_{B'} -f_1 d\mathbb{P} \geq -a\mathbb{P}(B'),$$

ce qui, en faisant tendre a vers $-\infty$, va entraîner que $\mathbb{P}(\underline{f} = -\infty) = 0$. Soit maintenant deux rationnels a et b , avec $a < b$. On reprend maintenant le même procédé, avec l'événement

$$B'' = \{\underline{f} < a < b < \bar{f}\}.$$

On va montrer que $\mathbb{P}(B'' = 0)$: en effet sinon, avec (g_n) on montre que

$$\int_{B''} f_1 d\mathbb{P} \geq b\mathbb{P}(B'')$$

tandis qu'avec $(-S_n + na)$ on montre

$$\int_{B''} -f_1 d\mathbb{P} \geq -a\mathbb{P}(B'')$$

En faisant la somme, on a $0 \geq (b - a)\mathbb{P}(B'')$, d'où $\mathbb{P}(B'') = 0$. Finalement $\underline{f} = \bar{f}$ et les 2 quantités sont finies presque sûrement, ce qui achève de montrer la convergence presque sûre.

– 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 $S_n(f)$ tend dans L^1 vers \bar{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

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

Mais $\frac{S_n(f)}{n}$ tend dans L^1 vers \bar{f} , donc $\mathbb{E}\left[\frac{S_n(f)}{n} \mathbb{1}_A\right]$ tend vers $\mathbb{E}[\bar{f} \mathbb{1}_A]$, d'où $\mathbb{E}[\bar{f} \mathbb{1}_A] = \mathbb{E}[f \mathbb{1}_A]$. Finalement, on a bien $\bar{f} = \mathbb{E}[f | \mathcal{I}]$.

Passons au cas général : soit f_ε une fonction bornée telle que $\|f - f_\varepsilon\|_1 \leq \varepsilon/2$.

On a

$$S_n(f) - \mathbb{E}[f | \mathcal{I}] = (S_n(f) - S_n(f_\varepsilon)) + (S_n(f_\varepsilon) - \mathbb{E}[f_\varepsilon | \mathcal{I}]) + (\mathbb{E}[f_\varepsilon | \mathcal{I}] - \mathbb{E}[f | \mathcal{I}])$$

Mais

$$\|S_n(f) - S_n(f_\varepsilon)\|_1 \leq \|f - f_\varepsilon\|_1 \leq \varepsilon/2$$

et

$$\|\mathbb{E}[f_\varepsilon | \mathcal{I}] - \mathbb{E}[f | \mathcal{I}]\|_1 \leq \|f - f_\varepsilon\|_1 \leq \varepsilon/2,$$

donc

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

Et en faisant tendre n vers l'infini :

$$\overline{\lim}_{n \rightarrow +\infty} \|S_n(f) - \mathbb{E}[f | \mathcal{I}]\|_1 \leq \varepsilon.$$

Comme ε est arbitraire, cela achève la preuve. □

1.2.3 Application aux chaînes de Markov

Les chaînes de Markov fournissent des exemples de systèmes dynamiques mesurés. Si E est un ensemble dénombrable, $P = (p_{i,j})$ une matrice stochastique et μ une mesure de probabilités sur E , on notera \mathbb{P}^μ 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}^{δ_x} .

On rappelle deux propriétés fondamentales des chaînes de Markov qui seront très utiles dans la suite.

- 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

$$(A \subset \{X_n = x\}) \implies \mathbb{P}^\mu(A \cap \theta^{-n}(B)) = \mathbb{P}^\mu(A) \mathbb{P}^x(B)$$

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

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

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}^μ 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. \square

Théorème 5. *Supposons que μ est invariante, au sens où $\mathbb{P}_{X_1}^\mu = \mu$. Alors $(E^\mathbb{N}, \mathcal{B}(E^\mathbb{N}), \mathbb{P}^\mu, \theta)$ est un système dynamique.*

Démonstration. D'après la propriété de Markov,

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

\square

Théorème 6. *Soit \mathbb{P}^μ 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 $\mathbb{P}^\mu(A, X_n = 0) > 0$. 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

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

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

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

1.2. THÉORÈME DE BIRKHOFF

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

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

Mais

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

donc

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

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, $\mathbb{P}^\mu(A) = \int \mathbb{P}^i(A) d\mu(i) = 1$. \square

Remarque : dans la preuve ci-dessus, on a utilisé le fait qu'une probabilité invariante pour une chaîne de Markov irréductible chargeait tous les points. En effet, prenons $x \in E$ tel que $\mu(x) > 0$ et soit y un autre élément de E . Il existe un entier n et une suite $x = x_0, x_1, \dots, x_n = y$ d'éléments de E avec pour tout i entre 0 et $n-1$, $p_{x_i, x_{i+1}} > 0$. Ainsi

$$\mu(y) = \mathbb{P}^\mu(X_n = y) \geq \mathbb{P}^\mu(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) = \mu(x) \prod_{i=0}^{n-1} p_{x_i, x_{i+1}} > 0.$$

Théorème 7. Soit \mathbb{P}^μ la probabilité sur l'espace canonique associée à une chaîne de Markov irréductible apériodique sur S admettant μ comme probabilité invariante. Alors, le système $(S^{\mathbb{N}}, \mathcal{B}(S^{\mathbb{N}}), \mathbb{P}^\mu)$ 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 $\mathbb{P}^\mu(A, \theta^{-n}(B)) \rightarrow \mathbb{P}^\mu(A) \mathbb{P}^\mu(B)$.

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,

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

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

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

□

Théorème 8. *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*

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

En particulier, pour tout $x \in S$

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

Démonstration. Si l'on note \mathbb{P}^μ 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 $(S^{\mathbb{N}}, \mathcal{B}(S^{\mathbb{N}}), \theta, \mathbb{P}^\mu)$ 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}^μ presque sûrement

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

Notons

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

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 $\mu(x)$. □

1.2.4 Théorème ergodique multidimensionnel

Théorème 9. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé. Soit $(\theta_i)_{i \in \mathbb{Z}^d}$ un groupe de transformations préservant la mesure \mathbb{P} , avec $\theta_{i+j} = \theta_i \circ \theta_j$. Soit \mathcal{I} la tribu des événements invariants par les θ_i .

Soit $(R_n)_{n \in \mathbb{N}}$ une suite de boîtes telle que $R_n \nearrow \mathbb{Z}^d$. Alors,

$$\frac{1}{|R_n|} \sum_{x \in R_n} f \circ \theta_x \rightarrow \mathbb{E}[f | \mathcal{I}] \quad \mathbb{P}\text{-p.s. et dans } L^1(\mathbb{P}) \text{ quand } n \text{ tend vers } +\infty.$$

Une preuve de ce théorème se trouve dans le livre de Hans-Otto Georgii cité en référence.

Le cas le plus fréquent est le cas où $\Omega = E^{\mathbb{Z}^d}$ et où θ_i est l'opérateur de décalage défini par $(\theta_i(\omega))_j = \omega_{i+j}$.

1.3 Théorème du retour, système induit

1.3.1 Théorème du retour de Poincaré

Théorème 10. 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 $\int Y(x) d\mu(x) = \int Z(x) d\mu(x)$. 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 $N = +\infty$. \square

1.3.2 Système induit

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

Pour $x \in \Omega$, on note $n_A(x) = \inf\{n \geq 1; \theta^n x \in A\}$

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 11. $(A, \mathcal{F}_A, \mathbb{P}_A, \theta_A)$ 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

$$\mathbb{P}(A^c, n_A = n, \theta^{-n}(C)) = \mathbb{P}(A, n_A = n+1, \theta^{-(n+1)}(C)) + \mathbb{P}(A^c, n_A = n+1, \theta^{-(n+1)}(C)) \quad (1.1)$$

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 (1.1), 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 12. 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.

1.3. THÉORÈME DU RETOUR, SYSTÈME INDUIT

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

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

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

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

Par définition de C , $\{y \in A; \theta(y) \in C\} = C$, donc

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

Finalement

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

Autrement dit,

$$\mathbb{1}_{\theta_A^{-1}(C)} \leq \mathbb{1}_{\theta_A^{-1}(C)} \circ \theta$$

Comme θ préserve la mesure, on en déduit 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, on a $\mathbb{P}(C) = \mathbb{P}(\theta_A^{-1}(C)) = 0$, d'où $\mathbb{P}_A(C) = 0$.
- Dans le second cas, on a $\mathbb{P}(x \in \Omega : \theta_A(x) \notin C) = 0$, d'où $\mathbb{P}(x \in A : \theta_A(x) \notin C) = 0$, soit $\mathbb{P}_A(\theta_A(x) \notin C) = 0$, et donc, comme θ_A préserve la mesure \mathbb{P}_A , $\mathbb{P}_A(A \setminus C) = \mathbb{P}_A(\theta_A(x) \in A \setminus C) = 0$.

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

1.3.3 Théorème de Kac

Théorème 13. *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

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

où

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

D'autre part, comme θ est ergodique

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

En particulier

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

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

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 $\mu(x) = \frac{1}{\mathbb{E}^x T_x}$.

1.4 Théorème ergodique sous-additif

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

$$\forall n, p \geq 1 \quad u_{n+p} \leq u_n + u_p.$$

Alors $\frac{u_n}{n}$ converge vers $\inf_{n \geq 1} \frac{u_n}{n}$.

Démonstration. Il suffit de montrer que $\overline{\lim}_{n \rightarrow +\infty} \frac{u_n}{n} \leq \inf_{n \geq 1} \frac{u_n}{n}$. On va donc

montrer que pour tout $k \geq 1$ $\overline{\lim}_{n \rightarrow +\infty} \frac{u_n}{n} \leq \frac{u_k}{k}$. Pour tout r entre 0 et $k-1$, on a $u_{kn+r} \leq nu_k + u_r$, d'où

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

Et $\overline{\lim}_{n \rightarrow +\infty} \frac{u_{kn+r}}{nk+r} \leq \frac{u_k}{k}$. Cependant $\overline{\lim}_{n \rightarrow +\infty} \frac{u_n}{n} = \max_{0 \leq r \leq k-1} \overline{\lim}_{n \rightarrow +\infty} \frac{u_{kn+r}}{nk+r}$, d'où le résultat. \square

1.4. THÉORÈME ERGODIQUE SOUS-ADDITIF

Le théorème ergodique sous-additif est dû à Kingman [Kin68]. La preuve présentée ici est celle de Steele [Ste89]. Elle s'inspire des travaux de Kamae [Kam82] et de Katznelson et Weiss [KW82]. Elle présente la particularité de ne pas faire explicitement appel au lemme ergodique maximal, se basant sur des arguments trajectoriels.

Théorème 14. *Soit $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ un système dynamique, $(g_n)_{n \geq 1}$ une famille de fonctions intégrables vérifiant*

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

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

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

Démonstration. – Étape 1 : réduction au cas de fonctions négatives

$$g'_n = g_n - \sum_{i=0}^{n-1} g_1 \circ \theta^i.$$

D'après la propriété de sous-additivité, les fonctions (g'_n) sont négatives. Elles sont également sous-additives. Si on montre le théorème pour les (g'_n) , on l'aura pour les (g_n) car $\frac{1}{n} \sum_{i=0}^{n-1} g_1 \circ \theta^i$ tend vers $\mathbb{E}[g_1 | \mathcal{I}]$. On peut donc dorénavant supposer les g_n négatives.

– Étape 2 : On pose

$$g = \liminf_{n \rightarrow +\infty} \frac{g_n}{n}.$$

$$\frac{g_{n+1}}{n} \leq \frac{g_1}{n} + \frac{g_n}{n} \circ \theta,$$

donc en passant à la limite inférieure $g \leq g \circ \theta$. On pose maintenant

$$G_M = \max(g, -M)$$

On a $G_M \leq G_M \circ \theta$. 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 $g = g \circ \theta$. Définissons maintenant

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

et son complémentaire

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

La clef de la preuve consiste en l'obtention de l'inégalité suivante :

$$\forall N, M \in \mathbb{N}^* \quad \forall n \geq N \quad g_n \leq \left(G_M + \frac{1}{M}\right) \left(\sum_{i=0}^{n-N} \mathbb{1}_{A(N, M)} \circ \theta^i\right). \quad (1.2)$$

La preuve est essentiellement algorithmique. Si $G_M + \frac{1}{M} \geq 0$, il n'y a rien à démontrer car g_n est négative. Sinon, on définit une suite $i_n(\omega)$ par récurrence avec $i_0 = 0$, puis

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

Soit j le plus petit entier tel que $i_n = n$. La sous-additivité donne

$$g_n \leq \sum_{u=0}^{j-1} g_{i_{u+1}-i_u}(\theta^{i_u}(\omega)).$$

Cependant, par construction on a pour tout u :

$$g_{i_{u+1}-i_u} \leq \left(G_M + \frac{1}{M}\right)(i_{u+1} - i_u) \mathbb{1}_{\{i_u \leq n-N\}} \mathbb{1}_{A(N, M)} \circ \theta^{i_u},$$

donc $g_N \leq (G_M + \frac{1}{M})A$, avec

$$A = \sum_{u=0}^{j-1} (i_{u+1} - i_u) \mathbb{1}_{\{i_u \leq n-N\}} (1 - \mathbb{1}_{B(N, M)}) \circ \theta^{i_u}$$

Soit s le plus grand entier tel que $i_s \leq n - N$: on a

$$\begin{aligned} A &= \sum_{u=0}^s (i_{u+1} - i_u) (1 - \mathbb{1}_{B(N, M)}) \circ \theta^{i_u} \\ &= \sum_{u=0}^s (i_{u+1} - i_u) - \sum_{u=0}^s \mathbb{1}_{B(N, M)} \circ \theta^{i_u} \\ &= s_{i_{s+1}} - \sum_{u=0}^s \mathbb{1}_{B(N, M)} \circ \theta^{i_u} \end{aligned}$$

1.4. THÉORÈME ERGODIQUE SOUS-ADDITIF

Or $i_{s+1} \geq n - N + 1$ et $\sum_{u=0}^s \mathbb{1}_{B(N,M)} \circ \theta^{iu} \leq \sum_{i=0}^{n-N} \mathbb{1}_{B(N,M)} \circ \theta^i$, donc

$$A \geq \sum_{i=0}^{n-N} \mathbb{1}_{A(N,M)} \circ \theta^i.$$

Ce qui nous donne bien (1.2). Maintenant, avec le théorème de Birkhoff ponctuel, on a

$$\forall N, M \in \mathbb{N}^* \quad \overline{\lim}_{n \rightarrow +\infty} \frac{g_n}{n} \leq (G_M + \frac{1}{M}) \mathbb{P}(A(N, M) | \mathcal{I}) \text{ p.s.}$$

Mais $\lim_{N \rightarrow +\infty} \mathbb{1}_{A(N,M)} \rightarrow 1$ p.s. D'où, avec le théorème de convergence dominée conditionnel

$$\forall M \in \mathbb{N}^* \quad \overline{\lim}_{n \rightarrow +\infty} \frac{g_n}{n} \leq G_M + \frac{1}{M} \text{ p.s.}$$

On termine la preuve en faisant tendre M vers l'infini. □

Théorème 15. $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ système dynamique $(g_n)_{n \geq 1}$ famille de fonctions intégrables

$$g_{n+p} \leq g_n + g_p \circ \theta^n$$

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

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

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

$$\frac{1}{k} \mathbb{E}[g_k | \mathcal{I}_k] - g \geq 0, \tag{1.3}$$

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 :

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

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

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

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

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

où la dernière inégalité vient de (1.3). Ainsi, on a bien que g_k/k converge dans L^1 vers g . Le dernier point découle alors aisément du Lemme de Fekete. \square

1.5 Application à la percolation de premier passage

La percolation de premier passage a été introduite par Hammersley and 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 passages sont indépendants, distribués suivant une même loi ν admettant un moment d'ordre deux.² 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

²On peut grandement affaiblir cette hypothèse, mais ce n'est pas notre propos ici.

1.5. APPLICATION À LA PERCOLATION DE PREMIER PASSAGE

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 16. *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 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 [Kes86] 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

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

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 1 D'après l'inégalité ci-dessus, on a

$$0 \leq \mu(x) \leq m\|x\|_1. \tag{1.4}$$

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 $+\infty$

$$\forall x, y \in \mathbb{Z}^d \quad \mu(x+y) \leq \mu(x) + \mu(y) \tag{1.5}$$

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 l'infini $\mu(x) = \mu(-x)$. Soit $p \geq 1$. La suite $d(0, (np)x)/np$ est une suite extraite de $d(0, nx)/n$ donc sa limite est $\mu(x)$. Mais $d(0, n(px))/n$ tend vers $\mu(px)$. Donc $\mu(px) = p\mu(x)$. Finalement

$$\forall x \in \mathbb{Z}^d \quad \forall n \in \mathbb{Z} \quad \mu(nx) = |n|\mu(x). \quad (1.6)$$

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

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

On vérifie sans difficulté que l'inégalité triangulaire (1.5), l'homogénéité (1.6), et la "continuité" (1.4) 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 1.4

$$|\mu(x) - \mu(y)| \leq m\|x - y\|_1. \quad (1.7)$$

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 2.

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

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

1.5. APPLICATION À LA PERCOLATION DE PREMIER PASSAGE

Soit $\delta > 0$ quelconque. Son choix précis sera fait ultérieurement. Soit p_0 donné par le lemme précédent 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 η_1, \dots, η_n iid 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

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

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

$$\mathbb{P}(B_t \not\subset \Lambda_{\frac{2}{\alpha}t}) \leq C \exp(-\beta t) \tag{1.8}$$

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

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

D'après le lemme de Borel-Cantelli, $d(0, nu) \geq \frac{\alpha}{2}n\|u\|_1$ pour n assez grand : on en déduit que $\mu(u) \geq \frac{\alpha}{2}\|u\|_1$. Comme μ est homogène et continue, l'inégalité s'étend d'abord à \mathbb{Q}^d , puis à \mathbb{R}^d . Et comme $\|\cdot\|_1$ 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 17. *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 $t \geq t_0$*

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

où l'on a noté

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

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,

$$\overline{\lim}_{|x| \rightarrow +\infty} \frac{|d(0, x) - \mu(x)|}{|x|} = 0.$$

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 $z_1, \dots, z_n \in \mathbb{Z}^d$ tel 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

1.5. APPLICATION À LA PERCOLATION DE PREMIER PASSAGE

$x \in \mathbb{Z}^d$ avec $\varepsilon|x| \geq M$: il existe i tel que $|\frac{z_i}{|z_i|} - \frac{x}{|x|}| \leq \varepsilon|x|$, soit $|z_i - \frac{|z_i|}{|x|}x| \leq \varepsilon|x|$ et, en notant n l'entier le plus proche de $\frac{|z_i|}{|x|}$, $|x - nz| \leq 2\varepsilon|x|$. On a

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

d'où

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

D'où

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

avec $m = \mathbb{E}[X_1]$. On va montrer que $\overline{\lim}_{|x| \rightarrow +\infty} \sup_{|x-y| \leq 2\varepsilon|x|} \frac{|d(x, y)|}{|x|} \leq 4m\varepsilon$, ce qui achèvera la preuve. Pour cela, avec le lemme de Borel-Cantelli, il suffit de montrer que

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

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. □

Chapitre 2

Couplage, ordre stochastique

2.1 Généralités

2.1.1 Ordre stochastique

Si Ω est un ensemble muni d'un ordre partiel \preceq , on dit qu'une fonction f de Ω dans \mathbb{R} est croissante si

$$\forall (x, y) \in \Omega^2 \quad (x \preceq y) \implies f(x) \leq f(y).$$

Dans la pratique, on considèrera le plus souvent le cas où $\Omega = E^S$, où E est une partie de \mathbb{R} et S un ensemble fini ou dénombrable, l'ordre partiel considéré étant l'ordre classique

$$\forall (x, y) \in \Omega^2 \quad (x \preceq y) \iff (\forall i \in S \quad x_i \leq y_i).$$

Soient μ et ν deux mesures sur (Ω, \mathcal{F}) . On dit que μ est stochastiquement dominée par ν , et l'on note $\mu \preceq \nu$ si pour toute fonction croissante mesurable bornée, on a

$$\int_{\Omega} f \, d\mu \leq \int_{\Omega} f \, d\nu.$$

On dit qu'un ensemble (un événement) est croissant si son indicatrice est une fonction croissante.

Théorème 18. *Soient μ et ν deux mesures sur \mathbb{R} . Les deux conditions suivantes sont équivalentes :*

- $\mu \preceq \nu$
-

$$\forall t \in \mathbb{R} \quad Q_{\mu}(t) \leq Q_{\nu}(t),$$

avec $Q_{\nu}(t) = \nu(]t, +\infty[)$.

2.1. GÉNÉRALITÉS

Démonstration. Preuve de (ii) \implies (i)

Posons, pour $\gamma \in \{\mu, \nu\}$

$$\forall u \in [0, 1] \quad Q_\gamma^*(u) = \inf\{x; Q_\gamma(x) \leq u\}.$$

Par construction $Q_\mu^* \leq Q_\nu^*$. Soit maintenant U une variable aléatoire suivant la loi uniforme sur $[0, 1]$. $\{Q_\gamma^*(U) > x\} = \{Q_\gamma(x) > U\}$, donc

$$\forall x \in \mathbb{R} \quad \mathbb{P}(Q_\gamma^*(U) > x) = \mathbb{P}(Q_\gamma(x) > U) = Q_\gamma(x),$$

ce qui signifie que γ est la loi de $Q_\gamma^*(U)$.

Soit maintenant f une fonction croissante mesurable bornée : on a $Q_\mu^* \leq Q_\nu^*$, donc $Q_\mu^*(U) \leq Q_\nu^*(U)$

$$f(Q_\mu^*(U)) \leq f(Q_\nu^*(U)),$$

d'où

$$\int_{\mathbb{R}} f \, d\mu = \mathbb{E}f(Q_\mu^*(U)) \leq \mathbb{E}f(Q_\nu^*(U)) = \int_{\mathbb{R}} f \, d\nu,$$

□

Exemple : pour tous $p, p' \in [0, 1]$, $p \leq p'$ entraîne $\text{Ber}(p) \preceq \text{Ber}(p')$.

Au cours de la preuve, on a montré qu'on pouvait fabriquer des variables aléatoires X et Y telles que $X \leq Y$, avec μ pour loi de X et ν pour loi de Y : c'est ce qu'on appelle un couplage.

2.1.2 Couplage

Définition : Soit μ et ν deux lois sur Ω : on appelle couplage de μ et ν la construction d'un espace probabilisé $(\Omega', \mathcal{F}', \mathbb{P})$ et de deux variables aléatoires X et Y sur $(\Omega', \mathcal{F}', \mathbb{P})$ telles que $\mathbb{P}_X = \mu$ et $\mathbb{P}_Y = \nu$.

Remarque : un couplage existe toujours ! Il suffit en effet de prendre $\Omega' = \Omega^2$, $\mathcal{F}' = \mathcal{F} \otimes \mathcal{F}$ et $\mathbb{P} = \mu \otimes \nu$. Généralement, on recherche des couplages qui vérifient des conditions supplémentaires sur la loi du couple (X, Y) . Dans notre exemple, la condition supplémentaire était $\mathbb{P}_{(X,Y)}(M) = 1$, avec

$$M = \{(x, y) \in \mathbb{R}^2; x \leq y\}.$$

Théorème 19 (Strassen). *Soit (Ω, \mathcal{F}) un espace mesuré polonais (métrique séparable complet) muni d'un ordre partiel. On suppose que*

$$M = \{(x, y) \in \Omega^2; x \preceq y\}$$

est fermé pour la topologie produit. Alors, quelque soient les lois μ et ν sur (Ω, \mathcal{F}) , il y a équivalence entre

- $\mu \preceq \nu$
- Pour toute fonction f croissante continue bornée

$$\int_{\Omega} f d\mu \leq \int_{\Omega} f d\nu.$$

- Il existe un couplage de μ et ν sur Ω^2 avec $\mathbb{P}_{(X,Y)}(M) = 1$, $\mathbb{P}_X = \mu$ et $\mathbb{P}_Y = \nu$.

Démonstration. La preuve est trop longue et complexe pour être exposée ici. On peut trouver les grandes lignes de la preuve dans Liggett [Lig85b]. \square

2.1.3 Quelques propriétés

Corollaire 2. Si Ω est muni d'une addition compatible avec l'ordre, alors la convolution associée préserve l'ordre stochastique.

Démonstration. Il suffit de réaliser un couplage. \square

Exemple : pour tout $n \in \mathbb{N}^*$ et tous $p, p' \in [0, 1]$, $p \leq p'$ entraîne $\mathcal{B}(n, p) \preceq \mathcal{B}(n, p')$.

Corollaire 3. La convergence en loi préserve l'ordre stochastique.

Démonstration. Cela vient de la caractérisation avec les fonctions croissantes continues bornées. \square

Exemple : pour tout $\lambda, \lambda' > 0$, $\lambda \leq \lambda'$ entraîne $\mathcal{P}(\lambda) \preceq \mathcal{P}(\lambda')$.

Ainsi, si $(u_n)_{n \geq 0}$ est une suite croissante, la fonction

$$x \mapsto e^{-x} \sum_{n=1}^{+\infty} \frac{u_n}{n!} x^n$$

est croissante sur \mathbb{R}_+ .

Définition : on dit qu'une fonction définie sur E^S est locale si elle ne dépend que d'un nombre fini de coordonnées.

Corollaire 4. On prend $\Omega = E^S$, où E est une partie de \mathbb{R} et S un ensemble dénombrable, l'ordre partiel considéré étant l'ordre classique.

$$\forall (x, y) \in \Omega^2 \quad (x \preceq y) \iff (\forall i \in S \quad x_i \leq y_i).$$

2.1. GÉNÉRALITÉS

Soient μ et ν deux mesures sur (Ω, \mathcal{F}) . Pour que μ soit stochastiquement dominée par ν , il suffit d'avoir pour toute fonction croissante locale continue bornée :

$$\int_{\Omega} f d\mu \leq \int_{\Omega} f d\nu.$$

Démonstration. Soit f une fonction croissante continue bornée. Soit $y \in \Omega$ quelconque, Λ une partie finie de S . On pose $f_{\Lambda}(x) = f(x_{\Lambda}y_{\Lambda^c})$. f_{Λ} est locale croissante continue, donc

$$\int_{\Omega} f_{\Lambda} d\mu \leq \int_{\Omega} f_{\Lambda} d\nu.$$

Soit Λ_n une suite croissante de parties finies de S dont la réunion est S . Quand n tend vers l'infini $x_{\Lambda_n}y_{\Lambda_n^c}$ tend vers x (pour la topologie produit), donc, comme f est continue $f_{\Lambda_n}(x) = f(x_{\Lambda_n}y_{\Lambda_n^c})$ tend vers $f(x)$. Par convergence dominée, $\int_{\Omega} f_{\Lambda_n} d\mu$ tend donc vers $\int_{\Omega} f d\mu$, idem pour ν , d'où

$$\int_{\Omega} f d\mu \leq \int_{\Omega} f d\nu.$$

□

Remarque : si S est fini, une mesure de probabilité sur $\{0, 1\}^S$ est complètement déterminée par la donnée des mesures des événements croissants. En effet si $\omega \in S$, on a $\{\omega\} = A^{\omega} \setminus (A^{\omega} \cap B^{\omega})$ où

$$A^{\omega} = \bigcap_{e:\omega_e=1} \{x \in \Omega; x_e = 1\}$$

et

$$B^{\omega} = \bigcup_{e:\omega_e=0} \{x \in \Omega; x_e = 1\}.$$

Si S est dénombrable, comme les marginales fini-dimensionnelles caractérisent une mesure, une mesure de probabilité sur $\{0, 1\}^S$ est complètement déterminée par la donnée des mesures des événements locaux croissants.

Théorème 20. *Soit E un ensemble compact, S un ensemble dénombrable. Pour que la suite μ_n de mesures de probabilités sur E^S converge vers μ , il suffit que pour toute fonction continue locale (bornée), $\int f d\mu_n$ converge vers $\int f d\mu$.*

Démonstration. Soit f une fonction continue. Comme E^S est compact, f est uniformément continue. Quand n tend vers l'infini $x_{\Lambda_n}y_{\Lambda_n^c}$ tend uniformément vers x (pour la topologie produit), donc, comme f est uniformément continue $f_{\Lambda_n}(x) = f(x_{\Lambda_n}y_{\Lambda_n^c})$ tend uniformément vers $f(x)$. Ainsi, on peut uniformément approcher toute fonction continue par une fonction locale. La suite est classique. □

Théorème 21. *Soit S un ensemble fini ou dénombrable, et $\Omega = \{0, 1\}^S$. Soit μ_n une suite de mesures de probabilités sur Ω . Si μ_n est monotone pour l'ordre stochastique, alors μ_n converge vers une mesure de probabilité μ_∞ .*

Démonstration. Traitons d'abord le cas où S est fini. Soit A un événement croissant. La suite $\mu_n(A)$ est croissante (ou décroissante), majorée par 1, minorée par 0 donc converge. Soit $\omega \in \Omega$. $(\mu_n(\omega))_{n \geq 1} = (\mu_n(A^\omega) - \mu_n((A^\omega \cap B^\omega)))_{n \geq 1}$ est une suite à valeurs positives. La suite converge (comme différence de deux suites convergentes) et la limite est positive (comme limite de quantités positives). Notons $\mu_\infty(\omega)$ la limite. Pour tout n , on a $\sum_{\omega \in \Omega} \mu_n(\omega) = 1$; comme la somme est finie, à la limite on a $\sum_{\omega \in \Omega} \mu_\infty(\omega) = 1$, donc μ_∞ est bien une mesure de probabilités. Il est maintenant évident que μ_n converge vers μ .

Passons au cas général : pour tout k , les marginales $\mu_n^{\Lambda^k}$ convergent vers une limite $\mu_\infty^{\Lambda^k}$. Comme le système $(\mu_n^{\Lambda^k})_{k \geq 0}$ est compatible, le système $(\mu_\infty^{\Lambda^k})_{k \geq 0}$ l'est aussi, ce qui définit donc une mesure sur $\{0, 1\}^S$ d'après le théorème de prolongement de Kolmogorov. D'après ce qui précède, il y a convergence de $\int f d\mu_n$ vers $\int f d\mu$ pour les fonctions locales, ce qui suffit à caractériser la convergence en loi d'après le théorème précédent. \square

2.1.4 Un contre-exemple

On a vu que dans \mathbb{R} , la comparaison des fonctions de répartition permet de caractériser l'ordre stochastique. Il est légitime de se demander si c'est encore vrai pour l'ordre partiel classique sur l'espace produit \mathbb{R}^S . On va voir tout de suite que c'est faux.

Théorème 22. *On peut construire des vecteurs aléatoires (X_1, X_2) et (Y_1, Y_2) avec*

$$\forall (t_1, t_2) \in \mathbb{R}^2 \quad \mathbb{P}(X_1 \geq t_1, X_2 \geq t_2) \leq \mathbb{P}(Y_1 \geq t_1, Y_2 \geq t_2),$$

mais tels qu'on n'a pas $\mathbb{P}_{(X_1, X_2)} \preceq \mathbb{P}_{(Y_1, Y_2)}$.

Démonstration. On va prendre des variables de Bernoulli. Ainsi, il suffira de montrer que $\mathbb{P}(X_1 = 1) \leq \mathbb{P}(Y_1 = 1)$, $\mathbb{P}(X_2 = 1) \leq \mathbb{P}(Y_2 = 1)$ et $\mathbb{P}(X_1 = X_2 = 1) \leq \mathbb{P}(Y_1 = Y_2 = 1)$.

Prenons X_1 et X_2 indépendantes suivant la loi Bernoulli de paramètre $p \in (0, 1)$. Posons $Y_1 = Y_2 = X_1$.

$$\mathbb{P}(X_1 = 1) = \mathbb{P}(Y_1 = 1) = p \text{ et } \mathbb{P}(X_2 = 1) = p = \mathbb{P}(Y_2 = 1) = p.$$

$$\mathbb{P}(X_1 = X_2 = 1) = p^2 \text{ tandis que } \mathbb{P}(Y_1 = Y_2 = 1) = \mathbb{P}(X_1 = 1) = p > p^2.$$

Cependant

$$\mathbb{E}[\max(X_1, X_2)] = 1 - (1 - p)^2 = p(2 - p)$$

tandis que

$$\mathbb{E}[\max(Y_1, Y_2)] = \mathbb{E}X_1 = p < p(2 - p).$$

□

2.2 Dynamiques markoviennes monotones

Si E est un ensemble dénombrable muni d'un ordre partiel \preceq , on dit qu'une dynamique markovienne donnée par la matrice $p_{i,j}$ est monotone si, lorsqu'on note μ_i la probabilité définie par $\mu_i(j) = p_{i,j}$, on a

$$(i \preceq j) \implies (\mu_i \preceq \mu_j)$$

Avec l'aide du théorème de Strassen, on peut démontrer le théorème suivant :

Théorème 23. *Soient μ, ν deux mesures sur E avec $\mu \preceq \nu$; on suppose qu'une dynamique markovienne monotone $p_{i,j}$ sur E est donnée.*

Alors, on peut construire sur le même espace deux chaînes de Markov $(X_n)_{n \geq 0}, (Y_n)_{n \geq 0}$ suivant la dynamique monotone donnée par $(p_{i,j})$ telles que $X_0 \sim \mu, Y_0 \sim \nu$ et pour tout $n, X_n \preceq Y_n$.

Démonstration. On prend d'abord (X_0, Y_0) avec $X_0 \preceq Y_0$, X_0 de loi μ et Y_0 de loi ν (Strassen). Posons ensuite

$$M = \{(x, y) \in E^2; x \preceq y\}.$$

Pour $(i, j) \in M$, notons $\mathbb{P}_{i,j}$ telle que $\mathbb{P}_{i,j}(M) = 1$ ayant μ_i comme première marginale et ν_j comme deuxième marginale (l'existence est donnée par le théorème de Strassen). Ensuite on prend des variables $V_{(i,j,n)}$ indépendantes avec $V_{(i,j,n)} \sim \mathbb{P}_{i,j}$, ces variables étant elles mêmes indépendantes de (X_0, Y_0) . Ensuite, on définit par récurrence $(X_n, Y_n)_{n \geq 1}$ par $(X_{n+1}, Y_{n+1}) = V_{X_n, Y_n, n+1}$ - par construction, on a pour tout $n (X_n, Y_n) \in M$ ce qui légitime la définition de $(X_{n+1}, Y_{n+1}) \in M$. Ainsi, pour tout n , on a $X_n \preceq Y_n$. Posons $\mathcal{G}_n = \sigma(X_i, Y_i, i \leq n)$. La suite $(X_n, Y_n)_{n \geq 0}$ est clairement markovienne, avec

$$\mathbb{P}(X_{n+1} = x, Y_{n+1} = y | \mathcal{G}_n) = \mathbb{P}_{(X_n, Y_n)}(\{(x, y)\}),$$

d'où

$$\mathbb{P}(X_{n+1} = x | \mathcal{G}_n) = \sum_{y \in E} \mathbb{P}_{(X_n, Y_n)}(\{(x, y)\}) = \mu_{X_n}(x) = p_{X_n, x},$$

On en déduit que si l'on pose $\mathcal{F}_n = \sigma(X_i, i \leq n)$, on a

$$\mathbb{P}(X_{n+1} = x | \mathcal{F}_n) = \mu_{X_n}(x) = p_{X_n, x},$$

ce qui montre que $(X_n)_{n \geq 0}$ suit bien la dynamique voulue. On procède de même pour $(Y_n)_{n \geq 0}$. \square

On va maintenant étudier un cas simple où tous les calculs peuvent se faire sur des espaces "concrets". On pourrait utiliser le théorème que l'on vient de démontrer, mais il est aussi simple, et plus instructif, de procéder de manière directe.

2.2.1 Processus de naissance et de mort

On appelle processus (à temps discret) de naissance et de mort une chaîne de Markov $(X_n)_{n \in \mathbb{N}}$ sur \mathbb{Z} (ou sur une partie de \mathbb{Z}) qui vérifie $X_{n+1} - X_n \in \{-1; 0; +1\}$. Ainsi, il existe p_i, q_i, r_i avec $\mathbb{P}(X_{n+1} = i - 1 | X_n = i) = q_i$, $\mathbb{P}(X_{n+1} = i | X_n = i) = r_i$, $\mathbb{P}(X_{n+1} = i + 1 | X_n = i) = p_i$ et $p_i + q_i + r_i = 1$ pour tout i .

Théorème 24. Soient $x, y \in \mathbb{Z}$ avec $x \leq y$; $(p_i)_{i \in \mathbb{Z}}, (q_i)_{i \in \mathbb{Z}}, (r_i)_{i \in \mathbb{Z}}$ trois suites de réels positifs indexées par \mathbb{Z} avec $p_i + q_i + r_i = 1$ pour tout i .

On suppose que pour tout $i \in \mathbb{Z}$, on a $q_{i+1} \leq 1 - p_i$.

Alors, on peut construire sur le même espace deux chaînes de naissance et de mort $(X_n)_{n \geq 0}, (Y_n)_{n \geq 0}$ de dynamique donnée par $(p_i)_{i \in \mathbb{Z}}, (q_i)_{i \in \mathbb{Z}}, (r_i)_{i \in \mathbb{Z}}$ telles que $X_0 = x, Y_0 = y$ et pour tout n , $X_n \leq Y_n$.

Démonstration. Posons $\Psi(a, b, u) = -\mathbb{1}_{\{u \leq a\}} + \mathbb{1}_{\{u > 1-b\}}$. Soit $(U_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes suivant la loi uniforme sur $[0, 1]$ On définit X_n par $X_0 = x$ et $X_{n+1} = X_n + \Psi(q_{X_n}, p_{X_n}, U_n)$ et Y_n par $Y_0 = y$ et $Y_{n+1} = Y_n + \Psi(q_{Y_n}, p_{Y_n}, U_n)$. Il est facile de voir que $(X_n)_{n \geq 0}$ et $(Y_n)_{n \geq 0}$ suivent les dynamiques voulues. Reste à voir que $X_n \leq Y_n$ pour tout n . Il suffit de démontrer que pour tous x, y dans \mathbb{Z} et tout $u \in \mathbb{R}$,

$$(x \leq y) \implies (x + \Psi(q_x, p_x, u) \leq y + \Psi(q_y, p_y, u))$$

Si $x = y$, il n'y a rien à démontrer. Si $y \leq x + 2$, on a $x + \Psi(q_x, p_x, u) \leq x + 1 \leq y - 1 \leq y + \Psi(q_y, p_y, u)$. Reste donc le cas où $y = x + 1$: il faut montrer que

$$x + \Psi(q_x, p_x, u) \leq x + 1 + \Psi(q_{x+1}, p_{x+1}, u),$$

soit

$$\Psi(q_x, p_x, u) \leq 1 + \Psi(q_{x+1}, p_{x+1}, u),$$

ou encore

$$\Psi(q_x, p_x, u) - \Psi(q_{x+1}, p_{x+1}, u) < 2.$$

L'inégalité large est toujours vérifiée. Supposons qu'on a égalité : on aurait $\Psi(q_x, p_x, u) = 1$ et $\Psi(q_{x+1}, p_{x+1}, u) = -1$, soit $u > 1 - p_x$ et $u \leq q_{x+1}$: on en déduit $1 - p_x < q_{x+1}$: contradiction. \square

Exemple : processus de naissance et de mort avec le taux de mort 1 et le taux de naissance λ : $(q_0, r_0, p_0) = (0, 1, 0)$ et pour $i > 0$ $(q_i, r_i, p_i) = (\frac{1}{\lambda+1}, 0, \frac{\lambda}{\lambda+1})$.

Remarque : en réalité, avec la même construction, on peut faire mieux : on peut faire vivre sur le même espace des chaînes $(X_n(i))_{n \geq 0}$ où i décrit \mathbb{Z} tout entier, avec $X_0(i) = i$ et telles que $i \leq j$ entraîne $X_n(i) \leq X_n(j)$ pour tout n . C'est un résultat plus fort que celui énoncé dans le théorème général sur les dynamiques monotones. Cette particularité est liée au fait que \mathbb{Z} est totalement ordonné.

2.3 Applications à la percolation

2.3.1 Transition de percolation

Soit μ une mesure de probabilités sur $\Omega = \{0, 1\}^{\mathbb{Z}^d}$. On note $(X_i)_{i \in \mathbb{Z}^d}$ la projection canonique sur la i -ième composante. On dit que le site i est ouvert dans la configuration ω si $X_i(\omega) = 1$.

On appelle chemin toute suite de points de \mathbb{Z}^d telle que chacun ne diffère du précédent que d'un élément de norme 1.

Notons

$$I = \{ \text{Il existe un chemin infini de sites ouverts} \}.$$

Si I est vérifié, on dit qu'il y a percolation.

Il est facile de voir que I est invariant pour l'action de \mathbb{Z}^d .

Ainsi, si μ est ergodique $\mu(I)$ ne peut valoir que 0 ou 1.

Le cas le plus fréquemment étudié est celui de la percolation indépendante : on a $\mu = \mathbb{P}_p = \text{Ber}(p)^{\otimes \mathbb{Z}^d}$.

Théorème 25. *Il existe $p_c \in (0, 1)$ tel que pour $p < p_c$ $\mathbb{P}_p(I) = 0$ tandis que pour $p > p_c$, $\mathbb{P}_p(I) = 1$.*

Démonstration. Soit $(U_i)_{i \in \mathbb{Z}^d}$ une collection de variables aléatoires indépendantes suivant la loi uniforme sur $[0, 1]$. On pose $X_i(p) = \mathbb{1}_{\{U_i \leq p\}}$. À p fixé, la loi des

$(X_i(p))_{i \in \mathbb{Z}^d}$ est \mathbb{P}_p . Cependant par construction on a $(X_i(p)) \preceq (X_i(q))$ dès que $p \leq q$. Ainsi pour $p \leq q$, on a

$$\begin{aligned} \mathbb{P}_p(I) &= \mathbb{P}(\text{les } i \text{ avec } X_i(p) = 1 \text{ percolent}) \\ &\leq \mathbb{P}(\text{les } i \text{ avec } X_i(q) = 1 \text{ percolent}) = \mathbb{P}_q(I). \end{aligned}$$

Comme les \mathbb{P}_p sont ergodiques, on sait maintenant que $p \mapsto \mathbb{P}_p(I)$ est une fonction croissante de $[0, 1]$ dans $\{0, 1\}$, d'où l'existence d'une valeur critique $p_c \in [0, 1]$.

Il reste à voir que $p_c \in]0, 1[$. Pour cela, il suffit d'exhiber un $p > 0$ tel que $\mathbb{P}_p(I) = 0$ et un autre $p < 1$ tel que que $\mathbb{P}_p(I) > 0$.

Notons $A_n(x)$ l'événement : "le point x est le point de départ d'un chemin ouvert sans recoupement de longueur n " : on voit facilement qu'il existe moins de $2d(2d-1)^n$ tels chemins, donc on a $\mathbb{P}_p(A_n(x)) \leq 2d(2d-1)^n p^n$, de sorte que si $p < \frac{1}{2d-1}$, la probabilité de l'événement $A(x)$: "le point x est le point de départ d'un chemin ouvert infini" est nulle. Comme $I = \bigcup_{x \in \mathbb{Z}^d} A(x)$, on a $\mathbb{P}_p(I) = 0$ pour $p < \frac{1}{2d-1}$. Ainsi $p_c \geq \frac{1}{2d-1} > 0$.

On suppose ici que $d = 2$. Si la composante connexe de 0 est finie, sa frontière extérieure $A = \partial C(0)$ forme un contour *-connexe fini qui coupe l'axe des abscisses en un point $(x, 0)$ avec $1 \leq x \leq |A|$, et les points de A sont tous fermés. Ainsi

$$\begin{aligned} \mathbb{P}_p(|C(0)| < +\infty) &\leq \sum_{n=1}^{+\infty} \sum_{1 \leq x \leq n} \sum_{B \in R(x,n)} \mathbb{P}_p(\forall y \in B \quad X_y = 0) \\ &\leq \sum_{n=1}^{+\infty} \sum_{1 \leq x \leq n} \sum_{B \in R(x,n)} (1-p)^n \end{aligned}$$

où $R(x, n)$ est l'ensemble des contours *-connexes passant par $(x, 0)$ et de longueur n . Cependant, $|R(x, n)| \leq 8 \cdot 7^{n-1}$, donc

$$\begin{aligned} \mathbb{P}_p(|C(0)| < +\infty) &\leq \sum_{n=1}^{+\infty} n 8 \cdot 7^{n-1} (1-p)^n \\ &\leq \frac{8(1-p)}{(1-7(1-p))^2} \end{aligned}$$

En prenant p assez grand (par exemple $p = 0.7$), on trouve $\mathbb{P}_p(|C(0)| < +\infty) < 1$, ce qui achève la preuve en dimension 2.

Cependant, si il y a percolation en dimension $d = 2$, il y a percolation en dimension plus grande car le même raisonnement assure l'existence d'un chemin infini dans un sous-espace de dimension 2 de \mathbb{Z}^d .

□

2.3.2 Formule de Russo

Théorème 26. Soit E un ensemble fini, et $\Omega = \{0, 1\}^E$. Soit A un événement croissant, au sens où $\mathbb{1}_A$ est une fonction croissante. On dit que $e \in E$ est pivotale pour A dans la configuration ω si $\mathbb{1}_A(0_e \omega_{e^c}) \neq \mathbb{1}_A(1_e \omega_{e^c})$. Soit $N(A)$ la variable aléatoire représentant le nombre de points pivotaux. Alors si \mathbb{P}_p désigne la mesure produit de Bernoulli de paramètre p sur Ω , on a

$$\frac{d}{dp} \mathbb{P}_p(A) = \mathbb{E}_p[N(A)].$$

Si $p > 0$, on a aussi

$$\frac{d}{dp} \mathbb{P}_p(A) = \mathbb{E}_p[N(A)] = \frac{1}{p} \mathbb{E}_p[N(A) \mathbb{1}_A] = \frac{1}{p} \mathbb{E}_p[N(A) | A] \mathbb{P}_p(A).$$

Ainsi, si $0 \leq p < q \leq 1$, on a

$$\mathbb{P}_q(A) = \mathbb{P}_p(A) \exp \left(\int_p^q \frac{1}{x} \mathbb{E}_x[N(A) | A] dx \right).$$

Démonstration. Soit $(\Omega', \mathcal{F}, \mathbb{P}) = ([0, 1]^E, \mathcal{B}(\Omega), U[0, 1]^{\otimes E})$.

On pose $X_e^{p'}(\omega') = \mathbb{1}_{\{\omega'_e \leq p\}}$. Pour tout $p \in [0, 1]$, le champ $(X_e^{p'})_{e \in E}$ prend ses valeurs dans Ω et suit \mathbb{P}_p sous la loi \mathbb{P} . Soit $G : \mathbb{R}^E \rightarrow \mathbb{R}$ défini par

$$G((p_e)_{e \in E}) = \mathbb{E}[\mathbb{1}_A((X_e^{p_e})_{e \in E})].$$

G est un polynôme en les composantes de p , donc G est C^1 . Or, si on pose $\bar{p} = (p, \dots, p)$

$$\forall p \in [0, 1] \quad G(\bar{p}) = \mathbb{P}_p(A),$$

donc d'après la formule de dérivation composée

$$\frac{d}{dp} \mathbb{P}_p(A) = \sum_{f \in E} \frac{\partial G}{\partial f}(p, \dots, p)$$

Notons $(u_f)_{f \in E}$ les vecteurs de base : on a

$$\begin{aligned} G(\bar{p} + hu_f) - G(\bar{p}) &= \mathbb{E}[\mathbb{1}_A((X_e^{p_e + h\delta_{e,f}})_{e \in E})] - \mathbb{E}[\mathbb{1}_A((X_e^{p_e})_{e \in E})] \\ &= \mathbb{E}[\mathbb{1}_{\{f \text{ pivotale pour } A\}}((X_e^{p_e})_{e \in E}) \mathbb{1}_{\{p < \omega'_f \leq p+h\}}] \\ &= \mathbb{E}[\mathbb{1}_{\{f \text{ pivotale pour } A\}}((X_e^{p_e})_{e \in E})] \mathbb{E}[\mathbb{1}_{\{p < \omega'_f \leq p+h\}}] \\ &= \mathbb{E}_p[\mathbb{1}_{\{f \text{ pivotale pour } A\}}] h \end{aligned}$$

Ici, on a utilisé que l'événement $\{f \text{ pivotal pour } A\}$ est indépendant de ω'_f . Ainsi,

$$\frac{\partial G}{\partial f}(\bar{p}) = \mathbb{E}_p[\mathbb{1}_{\{f \text{ pivotal pour } A\}}],$$

d'où en faisant la somme

$$\begin{aligned} \frac{d}{dp} \mathbb{P}_p(A) &= \sum_{f \in E} \mathbb{E}_p[\mathbb{1}_{\{f \text{ pivotal pour } A\}}] \\ &= \mathbb{E}_p\left[\sum_{f \in E} \mathbb{1}_{\{f \text{ pivotal pour } A\}}\right] \\ &= \mathbb{E}_p[N(A)] \end{aligned}$$

Cependant, on a également

$$\begin{aligned} \mathbb{E}_p[\mathbb{1}_{\{f \text{ pivotal pour } A\}}] &= \frac{1}{p} \mathbb{E}_p[\mathbb{1}_{\{f \text{ pivotal pour } A\}} \mathbb{1}_{\{f \text{ ouverte}\}}] \\ &= \frac{1}{p} \mathbb{E}_p[\mathbb{1}_{\{f \text{ pivotal pour } A\}} \mathbb{1}_A] \end{aligned}$$

d'où en faisant la somme la deuxième formule voulue. □

2.4 Théorème de Holley – Inégalités FKG

2.4.1 Théorème de Holley

Lemme 3 (Echantillonneur de Gibbs). *Soient (X, Y) deux vecteurs aléatoires à valeurs respectivement dans \mathbb{R}^n et \mathbb{R}^p . On suppose que la loi ν du couple (X, Y) admet la décomposition*

$$\int_{\mathbb{R}^n \times \mathbb{R}^p} f(x, y) d\nu(x, y) = \int_{\mathbb{R}^n \times \mathbb{R}^p} \left(\int_{\mathbb{R}^n} f(x, y) d\nu_x \right) d\nu(x, y)$$

On suppose que sur l'espace $(\Omega, \mathcal{F}, \mathbb{P})$ vivent des variables aléatoires $(Y_x)_{x \in \mathbb{R}^n}$, avec ν_x comme loi de Y_x , et une variable aléatoire X' indépendante des $(Y_x)_{x \in \mathbb{R}^n}$ ayant la même loi que X . Alors, si l'on pose $Y' = Y_{X'}$, le couple (Y', X') a la loi ν .

Démonstration. Soit g une fonction mesurable bornée.

$$\mathbb{E}[g(X', Y') | X'] = \mathbb{E}[(g(X', Y_{X'}) | X') = \psi(X'),$$

2.4. THÉORÈME DE HOLLEY – INÉGALITÉS FKG

avec $\psi(x) = \mathbb{E}[g(x, Y_x)] = \int_{\mathbb{R}^p} g(x, y) d\nu_x(y)$. Ainsi

$$\mathbb{E}[g(X', Y')] = \mathbb{E}[\psi(X')] = \int_{\mathbb{R}^n \times \mathbb{R}^p} \psi(x) d\nu(x, y) = \int_{\mathbb{R}^n \times \mathbb{R}^p} g d\nu.$$

□

Théorème 27 (Holley). *Soit E une partie finie de \mathbb{R} et S un ensemble fini. Soient μ et ν des mesures de probabilités sur E^S . On suppose que ν charge tous les éléments de E^S . On suppose de plus que si $a \in E$ et $s \in S$ alors pour tous $\zeta, \eta \in E^{S \setminus \{s\}}$ avec $\zeta \preceq \eta$, on a*

$$C_\mu(x, \zeta, a) \leq C_\nu(x, \eta, a),$$

avec

$$C_\mu(x, \zeta, a) = \mu(X_x \geq a | X_{S \setminus \{x\}} = \zeta)$$

et

$$C_\nu(x, \eta, a) = \nu(X_x \geq a | X_{S \setminus \{x\}} = \eta).$$

Alors $\mu \preceq \nu$.

Démonstration. On peut supposer sans perte de généralité que E s'écrit $E = \{0, 1, \dots, N\}$. On va utiliser la technique de l'échantillonneur de Gibbs. Soit (X_x^0) de loi μ . On pose $(Y_x^0) = (X_x^0)$. Soit $(U_n)_{n \geq 0}$ une suite de variables aléatoires suivant la loi uniforme sur $[0, 1]$, $(V_n)_{n \geq 0}$ une suite de variables aléatoires suivant la loi uniforme sur S , ces suites étant indépendantes entre elles et indépendantes de (X_x^0) . On va construire une chaîne de Markov par récurrence comme suit : Pour $x \neq V_n$, $X_x^{n+1} = X_x^n$ et $Y_x^{n+1} = Y_x^n$. En revanche, on définit $X_{V_n}^{n+1}$ comme le plus petit entier v tel que $C_\mu(V_n, X^n, v) \leq U_n$ et $Y_{V_n}^{n+1}$ comme le plus petit entier v tel que $C_\mu(V_n, Y^n, v) \leq U_n$. Par construction, on montre par récurrence que pour tout n , on a $(X_x^n)_{x \in S} \preceq (Y_x^n)_{x \in S}$.

La chaîne $(X^n)_{n \geq 0}$ est une chaîne de Markov stationnaire de loi μ . La chaîne $(Y^n)_{n \geq 0}$ est une chaîne de Markov irréductible apériodique admettant comme loi invariante la loi μ .

Soit f une fonction croissante : pour tout n , on a $f(X^n) \leq f(Y^n)$, d'où

$$\int f d\mathbb{P}_{X^n} \leq \int f d\mathbb{P}_{Y^n},$$

soit

$$\int f d\mu \leq \int f d\mathbb{P}_{Y^n}.$$

En faisant tendre n vers $+\infty$, on obtient

$$\int f d\mu \leq \int f d\nu.$$

□

2.4.2 Inégalité FKG

On dit qu'une probabilité μ sur E^S est monotone si pour tout $x \in S$ et tout $\eta, \zeta \in E^{S \setminus \{x\}}$ et $a \in E$, on a

$$(\zeta \leq \eta) \implies C_\mu(x, \zeta, a) \leq C_\mu(x, \eta, a),$$

avec

$$C_\mu(x, \zeta, a) = \mu(X_x \geq a | X_{S \setminus \{x\}} = \zeta).$$

Remarque : une mesure produit est monotone.

Théorème 28. *Soit E une partie finie de \mathbb{R} et S un ensemble fini. Soit μ une mesure sur E^S qui charge tous les points de E^S . Si μ est monotone, alors μ est positivement corrélée, c'est à dire que quelles que soient les fonctions croissantes f et g , on a*

$$\mathbb{E}_\mu[fg] \geq \mathbb{E}_\mu[f]\mathbb{E}_\mu[g].$$

Démonstration. Soit α une constante telle que $g + \alpha > 0$ sur Ω . Il est équivalent de montrer que

$$\mathbb{E}_\mu[fh] \geq \mathbb{E}_\mu[f],$$

où l'on a posé $h = \frac{g+\alpha}{\mathbb{E}_\mu[g+\alpha]}$. Autrement dit, il suffit de montrer que $\mu \preceq \nu$ où ν est la mesure dont la densité par rapport à μ est h . Comme h est strictement positive, μ et ν ont même support, donc on peut espérer appliquer le théorème de Holley.

$$C_\nu(x, \eta, a) = \frac{\sum_{s \geq a} \nu(s_x \eta_{x^c})}{\sum_{s \in S} \nu(s_x \eta_{x^c})}$$

D'où, en utilisant la croissance de h

$$\begin{aligned} \frac{C_\nu(x, \eta, a)}{1 - C_\nu(x, \eta, a)} &= \frac{\sum_{s \geq a} \mu(s_x \eta_{x^c}) h(s_x \eta_{x^c})}{\sum_{s < a} \mu(s_x \eta_{x^c}) h(s_x \eta_{x^c})} \\ &\geq \frac{\sum_{s \geq a} \mu(s_x \eta_{x^c}) h(a_x \eta_{x^c})}{\sum_{s < a} \mu(s_x \eta_{x^c}) h(a_x \eta_{x^c})} = \frac{C_\mu(x, \eta, a)}{1 - C_\mu(x, \eta, a)} \end{aligned}$$

2.4. THÉORÈME DE HOLLEY – INÉGALITÉS FKG

Ainsi, si $\zeta \leq \eta$, on a bien $C_\mu(x, \zeta, a) \leq C_\mu(x, \eta, a) \leq C_\nu(x, \eta, a)$ et on peut appliquer le théorème de Holley. \square

Corollaire 5 (FKG pour la percolation). *Soit $\Omega = \{0, 1\}^{\mathbb{Z}^d}$. Pour tout $p \in [0, 1]$, on note \mathbb{E}_p l'espérance sous \mathbb{P}_p .*

Si f et g deux fonctions croissantes dans $L^2(\mathbb{P}_p)$, alors

$$\mathbb{E}_p[fg] \geq \mathbb{E}_p[f]\mathbb{E}_p[g].$$

Démonstration. Si $p = 0$ ou $p = 1$, le résultat est évident. On supposera donc $0 < p < 1$. Dans le cas où f et g ne dépendent que des X_i pour i dans la boîte finie $\Lambda_n = [-n, n]^d$, il suffit d'appliquer l'inégalité FKG avec $\mu = \text{Ber}(p)^{\otimes \Lambda_n}$.

Dans le cas général, on va appliquer l'inégalité FKG à f_n et g_n avec $f_n = \mathbb{E}[f|\mathcal{F}_{\Lambda_n}]$ et $g_n = \mathbb{E}[g|\mathcal{F}_{\Lambda_n}]$. f_n et g_n sont des fonctions croissantes (ce n'est pas si évident, il faut utiliser l'indépendance). $f_n, g_n, f_n + g_n$ tendent respectivement dans L^2 vers $f, g, f + g$. La convergence dans L^2 impliquant la convergence dans L^1 , on a $\mathbb{E}_p[(f_n + g_n)^2] \rightarrow \mathbb{E}_p[(f + g)^2]$, $\mathbb{E}_p[f_n^2] \rightarrow \mathbb{E}_p[f^2]$, $\mathbb{E}_p[g_n^2] \rightarrow \mathbb{E}_p[g^2]$, $\mathbb{E}_p[f_n] \rightarrow \mathbb{E}_p[f]$, $\mathbb{E}_p[g_n] \rightarrow \mathbb{E}_p[g]$. Mais d'après la formule de polarisation $xy = \frac{1}{2}((x + y)^2 - x^2 - y^2)$, on a également $\mathbb{E}_p[f_n g_n] \rightarrow \mathbb{E}_p[fg]$. Ainsi l'inégalité

$$\mathbb{E}_p[f_n g_n] \geq \mathbb{E}_p[f_n]\mathbb{E}_p[g_n]$$

donne à la limite

$$\mathbb{E}_p[fg] \geq \mathbb{E}_p[f]\mathbb{E}_p[g].$$

\square

Corollaire 6 (FKG continue). *Soit $N \geq 1$ et ν une mesure quelconque sur \mathbb{R} . Soient f et g deux fonctions croissantes continues bornées sur \mathbb{R}^N . Alors*

$$\int_{\mathbb{R}^N} fg \, d\mu^{\otimes N} \geq \int_{\mathbb{R}^N} f \, d\mu^{\otimes N} \int_{\mathbb{R}^N} g \, d\mu^{\otimes N}.$$

Démonstration. Posons

$$F_n(x) = \begin{cases} 0 & \text{si } x < -n \\ F(\text{Ent}(nx)/n) & \text{si } -n \leq x < n \\ 1 & \text{si } x \geq n \end{cases}$$

Il est facile de voir que F_n est la fonction de répartition d'une loi μ_n dont le support E_n est fini. Par ailleurs, si F est continue au point x , alors $F_n(x)$ tend vers $F(x)$: ainsi μ_n converge en loi vers μ , ce qui entraîne que $\mu_n^{\otimes N}$

converge en loi vers $\mu^{\otimes N}$. $\mu_n^{\otimes N}$ est une mesure sur E_n^N qui est monotone et charge tous les points de E_n^N : on a donc

$$\int_{\mathbb{R}^N} fg \, d\mu_n^{\otimes N} \geq \int_{\mathbb{R}^N} f \, d\mu_n^{\otimes N} \int_{\mathbb{R}^N} g \, d\mu_n^{\otimes N}.$$

Comme f, g et fg sont continues bornées, la convergence en loi de $\mu_n^{\otimes N}$ vers $\mu^{\otimes N}$ entraîne

$$\int_{\mathbb{R}^N} fg \, d\mu^{\otimes N} \geq \int_{\mathbb{R}^N} f \, d\mu^{\otimes N} \int_{\mathbb{R}^N} g \, d\mu^{\otimes N}.$$

□

2.4.3 Une inégalité à la Russo

Théorème 29. Soit $(X_n)_{n \geq 0}$ une suite de variables aléatoires à valeurs dans $\{0; 1\}$ telle que pour tout n $\mathbb{E}[X_{n+1} | \mathcal{F}_n] \geq p$, où $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$. Alors

$$\mathbb{P}_X \succeq \text{Ber}(p)^{\otimes N}$$

Démonstration. On note f_0 la fonction constante égale à $\mathbb{E}X_0$, puis, pour $n \geq 1$, on prend pour f_n une fonction de \mathbb{R}^n dans \mathbb{R} dans telle que $\mathbb{E}[X_n | \mathcal{F}_{n-1}] = f_n(X_0, \dots, X_{n-1})$. D'après l'hypothèse qui est faite, on peut supposer que f_n est à valeurs dans $[p, 1]$. Soit $(U_n)_{n \geq 0}$ une suite de variables aléatoires à valeurs dans $[0, 1]$. On pose $X_0 = \mathbb{1}_{\{U_0 \leq f_0\}}$, $Y_0 = \mathbb{1}_{\{U_0 \leq p\}}$, puis pour $n \geq 1$: $X_n = \mathbb{1}_{\{U_n \leq f_n(X_0, \dots, X_{n-1})\}}$ et $Y_n = \mathbb{1}_{\{U_n \leq p\}}$. Par construction, on a $X_n \geq Y_n$ pour tout n .

On va pouvoir montrer que $(X_n)_{n \geq 0}$ a même loi que Z_n . Pour cela, on montre par récurrence sur n que pour tout n , (X_0, \dots, X_n) et (Z_0, \dots, Z_n) ont même loi. Pour $n = 0$, c'est évident. Pour $n \geq 1$, on applique l'hypothèse de récurrence et le lemme de l'échantillonneur de Gibbs, avec $X = (Z_0, \dots, Z_{n-1})$, $Y = Z_n$, $X' = (X_0, \dots, X_{n-1})$ et $Y_{x_0, \dots, x_n} = \mathbb{1}_{\{U_n \leq f_n(x_0, \dots, x_{n-1})\}}$.

Ainsi, on a $\text{Ber}(p)^{\otimes N} = \mathbb{P}_{Y'} \preceq \mathbb{P}_{X'} = \mathbb{P}_X$.

□

2.4. THÉORÈME DE HOLLEY – INÉGALITÉS FKG

Chapitre 3

Quelques outils utiles

3.1 L'argument de modification

Soit μ une mesure sur E^S , avec E fini et S dénombrable. On dit que μ a la propriété d'énergie finie si pour tout Λ partie finie de S et $B \in \mathcal{F}_\Lambda$ tel que $\mu(B) > 0$, on a

$$\mu(B|\mathcal{F}_{\Lambda^c}) > 0 \quad \mu \text{ p.s.}$$

Théorème 30. *Soit \mathbb{P} une mesure ayant la propriété d'énergie finie. Soit $A \in \mathcal{F}_\Lambda^c$ tel que $\mathbb{P}(A) > 0$ et Φ une application \mathcal{F}_Λ^c -mesurable à valeurs dans E^Λ . Alors*

$$\mathbb{P}(\{X_\Lambda = \Phi\} \cap A) > 0$$

Démonstration. Comme

$$\mathbb{P}(A) = \sum_{\eta \in E^\Lambda} \mathbb{P}(\Phi = \eta, A),$$

il existe $\eta \in E^\Lambda$ tel que $\mathbb{P}(\Phi = \eta, A) > 0$. Posons donc $A' = \{\Phi = \eta\} \cap A$ et $B = \{X_\Lambda = \eta\}$. Comme A' est \mathcal{F}_{Λ^c} -mesurable, on a $\mathbb{P}(A' \cap B|\mathcal{F}_{\Lambda^c}) = \mathbb{1}_{A'} \mathbb{E}(B|\mathcal{F}_{\Lambda^c})$. Ainsi

$$\mathbb{P}(A' \cap B) = \int \mathbb{1}_{A'} \mathbb{P}(B|\mathcal{F}_{\Lambda^c}) d\mathbb{P}.$$

Donc si on avait $\mathbb{P}(A' \cap B) = 0$, on aurait $\mathbb{1}_{A'} \mathbb{P}(B|\mathcal{F}_{\Lambda^c}) = 0$ p.s., donc $\mathbb{1}_{A'} = 0$ p.s. puisque $\mathbb{P}(\mathbb{P}(B|\mathcal{F}_{\Lambda^c}) = 0) = 0$. Ceci contredirait que $\mathbb{P}(A') > 0$. Ainsi $\mathbb{P}(A' \cap B) > 0$. Comme $A' \cap B \subset \{X_\Lambda = \Phi\} \cap A$, cela donne le résultat. \square

Théorème 31 (Newman et Schulman). *Soit \mathbb{P} une mesure sur $\{0, 1\}^{\mathbb{Z}^d}$. On suppose que \mathbb{P} à la propriété d'énergie finie et que \mathbb{P} est ergodique pour la translation de vecteur e_1 . On note $N(\omega)$ le nombre de composantes connexes infinies dans le graphe aléatoire associé à ω . Alors*

$$\mathbb{P}(N \in \{0, 1, +\infty\}) = 1.$$

Démonstration. N est invariant par θ_{e_1} , donc d'après l'hypothèse d'ergodicité, N est presque sûrement constant. Supposons qu'il existe $k \geq 2$ avec $\mathbb{P}(N = k) = 1$. Soit A_n l'événement "les amas infinis rencontrent tous la boîte Λ_n ".

$$\mathbb{P}(\cup_{n \geq 1} A_n) = 1$$

(ici on utilise le fait que $k < +\infty$), donc il existe n tel que $\mathbb{P}(A_n) > 0$. Cependant, si A_n est réalisé, tous les amas infinis rencontrent le bord extérieur de Λ_n . On a donc $\mathbb{P}(A'_n) > 0$ où A'_n est l'événement "tous les amas infinis rencontrent le bord extérieur de Λ_n ." Un instant de réflexion montre que A'_n est $\mathcal{F}_{\Lambda_n^c}$ mesurable. Soit $B = \{\forall i \in \Lambda_n \quad X_i = 1\}$. D'après le théorème de modification

$$\mathbb{P}(A'_n \cap B) > 0.$$

Si A'_n est réalisé, tous les amas infinis touchent le bord extérieur de Λ_n . Mais si en plus B est réalisé, alors il n'y a qu'un seul amas infini, donc $\mathbb{P}(N = 1) > 0$. Contradiction. \square

3.2 La technique de renormalisation

Schématiquement, la technique de renormalisation (ou arguments de blocs) est utile quand on sait que des événements $A_n(\varepsilon)$ vérifient $\mathbb{P}(A_n(\varepsilon)) \rightarrow 1$ et que l'on veut montrer que la convergence se fait à vitesse exponentielle.

La technique est la suivante :

- on considère un événement $B_N(\varepsilon)$ qui entraîne $A_N(\varepsilon)$ mais qui en plus est local.
- On choisit N tel que $\mathbb{P}(B_N(\varepsilon)) \geq p$, avec p grand.
- Ensuite, on regarde une famille de translatés des événements $B_N(\varepsilon)$: $B_x^N(\varepsilon) = \theta_{Nx}^{-1}(B_N(\varepsilon))$, où x décrit \mathbb{Z}^d ou une partie de \mathbb{Z}^d .
- Enfin, on décompose n en $n = Nk + r$ (division euclidienne) et on essaie de démontrer que
 - si de nombreux événements $B_x^N(\varepsilon)$ sont réalisés, alors $A_{Nk+r}(\eta)$ est réalisé.
 - Si p est assez grand, la probabilité de "de nombreux événements $B_x^N(\varepsilon)$ sont réalisés" dépasse $1 - K \exp(-Bk)$, où les constantes K et B ne doivent pas dépendre de N .

Pour comprendre la technique, on va la mettre en oeuvre sur un modèle simple : la percolation de premier passage avec des temps de passage bornés.

Avant cela, on va donner un petit lemme pour expliquer l'intérêt d'avoir de remplacer des événements par des événements locaux. En effet, souvent, quand on a des événements locaux, les choses ne sont pas très loin de l'indépendance.

Rappelons qu'une famille d'événements de variables aléatoire X_1, \dots, X_n est dite m -dépendante si pour tout I et J inclus dans $\{1, \dots, n\}$ les vecteurs X_I et X_J sont indépendants dès que la distance de I à J est au moins égale à m . En particulier, si X_1, \dots, X_n est m -dépendante et que I est tel que deux éléments distincts de I sont à une distance au moins m , les variables $(X_i)_{i \in I}$ sont indépendantes.

On va voir qu'on peut énoncer pour les variables m -dépendantes un lemme qui est le décalque exact du lemme binomiale pour les événements indépendants.

Lemme 4. *Soit m un entier naturel non nul. Pour tout $\varepsilon > 0$, pour tout $\delta > 0$, il existe $p_0 < 1$ et $K > 0$ tel que pour toute suite A_1, \dots, A_n d'événements m -dépendants de même probabilité $p \geq p_0$, on ait*

$$\forall n \geq 1 \quad \mathbb{P}\left(\sum_{i=1}^n \mathbb{1}_{A_i} \geq [n(1 - \varepsilon)]\right) \geq 1 - K\delta^n.$$

Démonstration. Pour tout $j \in \{0, \dots, m - 1\}$, on note E_j l'ensemble des entiers entre 1 et n qui sont congrus à j modulo m .

$$1 - \mathbb{P}\left(\sum_{i=1}^n \mathbb{1}_{A_i} \geq n(1 - \varepsilon)\right) = \mathbb{P}\left(\sum_{i=1}^n \mathbb{1}_{A_i^c} > n\varepsilon\right)$$

Mais

$$\left\{\sum_{i=1}^n \mathbb{1}_{A_i^c} > n\varepsilon\right\} \subset \bigcup_{i=0}^{m-1} \left\{\sum_{i \in E_i} \mathbb{1}_{A_i^c} > |E_i|\varepsilon\right\}.$$

Ainsi

$$1 - \mathbb{P}\left(\sum_{i=1}^n \mathbb{1}_{A_i} \geq n(1 - \varepsilon)\right) \leq \sum_{i=0}^{m-1} \mathcal{B}(|E_i|, p)([|E_i|(1 - \varepsilon), +\infty))$$

3.2. LA TECHNIQUE DE RENORMALISATION

Pour tout $\delta' > 0$ on peut choisir p_0 tel que pour $p \geq p_0$, on ait $\mathcal{B}(|E_i|, p)([|E_i|(1-\varepsilon), +\infty) \leq \delta'^{|E_i|}$, et donc

$$\begin{aligned} 1 - \mathbb{P}\left(\sum_{i=1}^n \mathbb{1}_{A_i} \geq n(1-\varepsilon)\right) &\leq \sum_{i=0}^{m-1} \delta'^{|E_i|} \\ &\leq m\delta'^k \leq m\delta'^{n/m-1} \end{aligned}$$

Ainsi, si on a choisi $\delta' = \delta^m$, on a le résultat voulu avec $K = \frac{m}{\delta'}$. \square

Théorème 32. *On considère le modèle de percolation de premier passage iid avec une loi ν des temps de passage à support dans $[0, M]$, avec $\nu(0) = 0$. On note μ la norme associée. Alors pour tout x dans \mathbb{Z}^d et tout $\varepsilon > 0$, il existe des constantes A et B strictement positives telles que*

$$\mathbb{P}(d(0, nx) > (1 + \varepsilon)n\mu(x)) \leq A \exp(-Bn).$$

Démonstration. On pose $A_n(\varepsilon) = \{d(0, nx) \leq (1 + \varepsilon)\mu(x)\}$ et $B_n = A_n \cap \{B_{(1+2\varepsilon)\mu(x)n} \subset \Lambda_{Kn}\}$. L'événement B_n est mesurable par rapport à la tribu engendrée par les variables des arêtes qui touchent au moins un point de Λ_{Kn} . On a

$$\mathbb{P}(B_n^c(\varepsilon)) \leq \mathbb{P}(A_n^c(\varepsilon)) + \mathbb{P}(B_{(1+2\varepsilon)\mu(x)n} \not\subset \Lambda_{Kn})$$

Lorsqu'on a montré que μ était une norme, on a établi l'inégalité (1.8). Ainsi, il existe $\alpha > 0$ tel que si on choisit $K = \frac{(1+2\varepsilon)\mu(x)}{\alpha}$, la probabilité de $\{B_{(1+2\varepsilon)\mu(x)n} \not\subset \Lambda_{Kn}\}$ décroît exponentiellement vite avec n . Comme $\frac{d(0, nx)}{n\mu(x)}$ tend presque sûrement vers 1, il y a aussi convergence en probabilité : ainsi $\mathbb{P}(A_n^c(\varepsilon))$ tend vers 0 et finalement $\mathbb{P}(B_n^c(\varepsilon))$ tend vers 0.

Ainsi, pour tout $p < 1$ on peut trouver un N tel que que $\mathbb{P}(B_N(\varepsilon)) \geq p$.

Soit n un entier. On effectue la division euclidienne de n par N : $n = kN + r$. On a

$$\begin{aligned} d(0, nx) &\leq \left(\sum_{i=0}^{k-1} d(iNx, (i+1)Nx)\right) + d(kNx, (kN+r)x) \\ &\leq \left(\sum_{i=0}^{k-1} N\mu(x)(1+\varepsilon)\mathbb{1}_{A_N} \circ \theta_{iNx} + MN\mathbb{1}_{A_N^c} \circ \theta_{iNx}\right) + MN \\ &\leq n\mu(x)(1+\varepsilon) + MN \sum_{i=0}^{k-1} \mathbb{1}_{A_N^c} \circ \theta_{iNx} + MN \\ &\leq n\mu(x)(1+\varepsilon) + MN \sum_{i=0}^{k-1} \mathbb{1}_{B_N^c} \circ \theta_{iNx} + MN \\ &\leq n\mu(x)(1+\varepsilon) + MnY_k^N + MN \end{aligned}$$

où on a posé

$$Y_k^N = \frac{1}{k} \sum_{i=0}^{k-1} \mathbb{1}_{B_N^c} \circ \theta_{iNx}$$

Ainsi, dès que n est assez grand, on a

$$d(0, nx) \leq n\mu(x)(1 + 2\varepsilon + \frac{M}{\mu(x)} Y_k^N)$$

Si $\|iNx - jNx\|_1 > (2K + 1)N$, il n'y a aucune arête qui touche à la fois les boîtes $iNx + \Lambda_{KN}$ et $jNx + \Lambda_{KN}$, donc les variables $\mathbb{1}_{B_N^c} \circ \theta_{iNx}$ et $\mathbb{1}_{B_N^c} \circ \theta_{jNx}$ sont indépendantes : le champ $(\mathbb{1}_{B_N^c} \circ \theta_{iNx})_{i \geq 0}$ est donc $\frac{2K+1}{\|x\|_1}$ -dépendant. Ainsi, d'après le lemme (4), on peut trouver p et K' , tel que

$$\forall N \geq 1 \quad \mathbb{P}(B_N) \geq p \implies \forall k \geq 1 \quad \mathbb{P}\left(\frac{M}{\mu(x)} Y_k^N > \varepsilon\right) \leq KK' \cdot \exp(-k).$$

Faisons ce choix pour p : cela induit le choix de N et on a maintenant pour $n = Nk + r$ suffisamment grand

$$\mathbb{P}(d(0, nx) > n\mu(x)(1+3\varepsilon)) \leq K' \exp(-k) \leq K' \exp(-n/N+1) = K'' \exp(-\frac{1}{N}n).$$

□

On aura remarqué que le lemme (4) laisse entendre qu'un champ de variables de Bernoulli m -dépendantes qui prend souvent des grandes valeurs se comporte en gros comme un champ de Bernoulli m indépendantes de grand paramètre.

Pendant longtemps, de nombreuses preuves en théorie de la percolation se sont appuyés sur cette idée, avec à chaque fois un raisonnement ad hoc du genre du découpage que nous avons fait. Cependant, en 1997, un résultat de Liggett Schonman et Stacey [LSS97], basé sur des techniques dues à Pisztora (voir par exemple [Pis96] ou [AP96], a permis de donner un principe général qui peut s'appliquer dans tous les cas

3.3 Retour sur l'ordre stochastique : théorème de Liggett-Schonmann-Stacey

Théorème 33 (Liggett-Schonmann-Stacey). *Soit S un ensemble dénombrable et Δ un entier naturel. On suppose qu'à chaque x on associe un voisinage épointé $V(x) \subset S \setminus \{x\}$ avec $|V(x)| \leq \Delta - 1$.*

3.3. RETOUR SUR L'ORDRE STOCHASTIQUE : THÉORÈME DE LIGGETT-SCHONMANN-STACEY

On suppose que $(X_x)_{x \in S}$ est un champ à valeurs dans $\{0, 1\}$ et, $(Y_x)_{x \in S}$ un champ de Bernoulli de paramètre r indépendant de $(X_x)_{x \in S}$. On suppose enfin que

$$\forall x \in S \quad \mathbb{E}[X_x | \mathcal{F}_{V(x)^c}] \geq p, \quad (3.1)$$

où $p \in [0, 1]$ est tel que

$$(1 - \alpha) \min(\alpha, 1 - r)^{\Delta-1} \geq p = 1 - p.$$

Pour tout $x \in S$, on pose $Z_x = X_x Y_x$.

On a alors

$$\mathbb{P}_X \succeq \mathbb{P}_Z \succeq \text{Ber}(\alpha r)^{\otimes S}.$$

En particulier, pour $P \geq 1/4$, on a $\mathbb{P}_X \succeq \text{Ber}(P)^{\otimes S}$ si on a

$$\forall x \in S \quad \mathbb{E}[X_x | \mathcal{F}_{V(x)^c}] \geq 1 - (1 - \sqrt{P})^\Delta \quad (3.2)$$

Démonstration. On va montrer d'abord pour I fini avec $x \notin I$, on a

$$\mathbb{E}[X_x | Z_I] \geq \alpha \text{ et } \mathbb{E}[Z_x | Z_I] \geq \alpha r.$$

Notons (H_n) l'hypothèse de récurrence : pour tout $I \subset S$ avec $|I| \leq n$, tout $x \in S \setminus I$ on a

$$\mathbb{E}[X_x | Z_I] \geq \alpha.$$

L'hypothèse est évidemment vérifiée pour $n = 0$.

Avant d'entamer la preuve, remarquons que si (H_n) est réalisé, alors pour tous J et K disjoints avec $|J \cup K| \leq n + 1$, on a pour tout A $\sigma(Z_K)$ mesurable

$$\mathbb{E}[\mathbb{1}_A \prod_{j \in J} X_j] \geq \alpha^{|J|} \mathbb{P}(A). \quad (3.3)$$

La preuve se fait par récurrence sur $|J|$: pour $|J| = 0$, il suffit d'intégrer l'inégalité (3.1) et d'utiliser que $p \geq \alpha$. Sinon on écrit $|J \cup K| = \{x\} \cup J'$ avec $|J'| = |J \cup K| - 1 \leq n$.

$$\mathbb{E}[\mathbb{1}_A \prod_{j \in J} X_j | \mathcal{F}_{J'}] = \mathbb{1}_A \left(\prod_{j \in J \setminus \{x\}} X_j \right) \mathbb{E}[X_x | \mathcal{F}_{J'}] \geq \alpha \mathbb{1}_A \left(\prod_{j \in J \setminus \{x\}} X_j \right)$$

d'après (H_n) . En intégrant, on obtient $\mathbb{E}[\mathbb{1}_A \prod_{j \in J} X_j] \geq \alpha \mathbb{E}(\mathbb{1}_A \prod_{j \in J \setminus \{x\}} X_j)$, d'où

le résultat par récurrence sur $|J|$.

On va maintenant montrer que la propriété (H_n) est héréditaire : supposons (H_{n-1}) vérifié et montrons (H_n) . Soit $|I|$ de cardinal inférieur ou égal

à n et $x \in S \setminus I$. Soit $z \in \{0, 1\}^I$. On va déterminer une fonction φ telle que $\mathbb{E}[X_x | Z_I = z] \geq \varphi(z)$. Pour simplifier les écritures, on pose $R = I \setminus V(x)$. On note aussi $N_1(z) = \sum_{i \in I \cap V(x)} z_i$ et $N_0(z) = |I \cap V(x)| - N_1(z)$.

On a

$$\mathbb{E}[X_x | Z_I = z] = \frac{\mathbb{E}[X_x \mathbb{1}_{\{Z_I=z\}}]}{\mathbb{P}(Z_I = z)} = 1 - \frac{\mathbb{E}[(1 - X_x) \mathbb{1}_{\{Z_I=z\}}]}{\mathbb{P}(Z_I = z)}$$

$$\begin{aligned} \frac{\mathbb{E}[(1 - X_x) \mathbb{1}_{\{Z_I=z\}}]}{\mathbb{P}(Z_I = z)} &\leq \frac{\mathbb{E}[(1 - X_x) \mathbb{1}_{\{Z_R=z_R\}} \prod_{\substack{i \in I \cap V(x) \\ z_i=1}} \mathbb{1}_{\{Y_i=1\}}]}{\mathbb{P}(Z_I = z)} \\ &= r^{N_1(z)} \frac{\mathbb{E}[(1 - X_x) \mathbb{1}_{\{Z_R=z_R\}}]}{\mathbb{P}(Z_R = z_R)} \frac{\mathbb{P}(Z_R = z_R)}{\mathbb{P}(Z_I = z)} \end{aligned}$$

En utilisant l'indépendance, puis (3.1) on a

$$\mathbb{E}[X_x | \sigma(X_R, Y_R)] = \mathbb{E}[X_x | X_R] \geq p,$$

d'où, comme $\sigma(Z_R)$ est une sous-tribu de $\sigma(X_R, Y_R)$:

$$\mathbb{E}[X_x | Z_r] \geq p \text{ d'où } \frac{\mathbb{E}[(1 - X_x) \mathbb{1}_{\{Z_R=z_R\}}]}{\mathbb{P}(Z_R = z_R)} \leq q.$$

Maintenant

$$\mathbb{P}(Z_I = z) \geq \mathbb{E} \left[\mathbb{1}_{\{Z_R=z_R\}} \left(\prod_{\substack{i \in I \cap V(x) \\ z_i=1}} \mathbb{1}_{\{X_i=1\}} \right) \left(\prod_{\substack{i \in I \cap V(x) \\ z_i=0}} \mathbb{1}_{\{Y_i=0\}} \right) \left(\prod_{\substack{i \in I \cap V(x) \\ z_i=1}} \mathbb{1}_{\{Y_i=1\}} \right) \right]$$

En utilisant l'indépendance, on a

$$\begin{aligned} \mathbb{P}(Z_I = z) &\geq r^{N_1(z)} (1 - r)^{N_0(z)} \mathbb{E} \left[\mathbb{1}_{\{Z_R=z_R\}} \left(\prod_{\substack{i \in I \cap V(x) \\ z_i=1}} X_i \right) \right] \\ &\geq r^{N_1(z)} (1 - r)^{N_0(z)} \alpha^{N_1(z)} \mathbb{P}(Z_R = z_R) \text{ d'après (3.3)} \\ &\geq r^{N_1(z)} \min(1 - r, \alpha)^{|I \cap V(x)|} \mathbb{P}(Z_R = z_R) \\ &\geq r^{N_1(z)} \min(1 - r, \alpha)^{\Delta-1} \mathbb{P}(Z_R = z_R) \end{aligned}$$

3.3. RETOUR SUR L'ORDRE STOCHASTIQUE : THÉORÈME DE LIGGETT-SCHONMANN-STACEY

Finalement

$$\forall z \in \{0, 1\}^I \quad \mathbb{E}[X_x | Z_I = z] \geq 1 - \frac{q}{\min(1-r, \alpha)^{\Delta-1}} \geq \alpha,$$

soit $\mathbb{E}[X_x | Z_I] \geq \alpha$.

Ceci achève la preuve de la récurrence.

Finalement, par indépendance,

$$\mathbb{E}[X_x Y_x | \sigma(X, Y_I)] = X_x \mathbb{E}[Y_x | \sigma(X, Y_I)] = X_x \mathbb{E}[Y_x] = r X_x,$$

d'où $\mathbb{E}[Z_x | Z_I] = r \mathbb{E}[X_x | Z_I] \geq \alpha r$.

Soit φ une bijection de \mathbb{N} dans S . On pose $Z'_n = Z_{\varphi(n)}$ et $I_n = (\varphi(\{0, \dots, n-1\}))$. On a pour tout n $\mathbb{E}[Z'_n | \sigma(Z'_0, \dots, Z'_{n-1})] = \mathbb{E}[Z_{\varphi(n)} | Z_{\varphi(I_n)}] \geq \alpha r$. Ainsi d'après le lemme 29, on a $\mathbb{P}_{Z'_n} \succeq \text{Ber}(\alpha r)^{\otimes \mathbb{N}}$, soit $\mathbb{P}_Z \succeq \text{Ber}(\alpha r)^{\otimes S}$. Comme $Z_x = X_x Y_x \leq X_x$ pour tout $x \in S$, on a évidemment $\mathbb{P}_Y \succeq \mathbb{P}_Z$.

Si $P \geq 1/4$, et que l'on pose $\alpha = r = \sqrt{P}$. On a

$$1 - r = 1 - \sqrt{P} \leq 1/2 \leq \sqrt{P} = \alpha,$$

donc on a $(1 - \alpha) \min(\alpha, 1 - r)^{\Delta-1} = (1 - \alpha)(1 - r)^{\Delta-1} = (1 - \sqrt{P})^\Delta$, ce qui donne bien la condition voulue. \square

Corollaire 7. *Soit M un entier naturel non nul. Il existe une fonction $f_M : [1/4, 1] \rightarrow [0, 1]$ avec $\lim_{p \rightarrow 1} f_M(p) = 1$ tel que tout champ $(X_x)_{x \in \mathbb{Z}^d}$ M -dépendant à valeurs dans $\{0, 1\}$ de loi μ vérifie*

$$(\forall x \in \mathbb{Z}^d \quad \mathbb{E}X_x \geq f_M(p)) \implies (X_x)_{x \in \mathbb{Z}^d} \succeq \text{Ber}(p).$$

Démonstration. Il suffit d'appliquer le lemme précédent avec $V(x) = (x + \{-M, \dots, M\}^d) \setminus \{x\}$ et $\Delta = |x + \{-M, \dots, M\}^d| = (2M+1)^d$, ce qui donne l'expression explicite $f_M(p) = 1 - (1 - \sqrt{p})^{(2M+1)^d}$. \square

Bibliographie

- [AP96] Peter Antal and Agoston Pisztora. On the chemical distance for supercritical Bernoulli percolation. *Ann. Probab.*, 24(2) :1036–1048, 1996.
- [CD81] J. Theodore Cox and Richard Durrett. Some limit theorems for percolation processes with necessary and sufficient conditions. *Ann. Probab.*, 9(4) :583–603, 1981.
- [CPS99] Jennifer Tour Chayes, Amber L. Puha, and Ted Sweet. Independent and dependent percolation. In *Probability theory and applications (Princeton, NJ, 1996)*, volume 6 of *IAS/Park City Math. Ser.*, pages 49–166. Amer. Math. Soc., Providence, RI, 1999.
- [Geo88] Hans-Otto Georgii. *Gibbs measures and phase transitions*, volume 9 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1988.
- [GHM01] Hans-Otto Georgii, Olle Häggström, and Christian Maes. The random geometry of equilibrium phases. In *Phase transitions and critical phenomena, Vol. 18*, volume 18 of *Phase Transit. Crit. Phenom.*, pages 1–142. Academic Press, San Diego, CA, 2001.
- [GK84] Geoffrey Grimmett and Harry Kesten. First-passage percolation, network flows and electrical resistances. *Z. Wahrsch. Verw. Gebiete*, 66(3) :335–366, 1984.
- [Kam82] Teturo Kamae. A simple proof of the ergodic theorem using non-standard analysis. *Israel J. Math.*, 42(4) :284–290, 1982.
- [Kes86] Harry Kesten. Aspects of first passage percolation. In *École d’été de probabilités de Saint-Flour, XIV—1984*, volume 1180 of *Lecture Notes in Math.*, pages 125–264. Springer, Berlin, 1986.
- [Kin68] J. F. C. Kingman. The ergodic theory of subadditive stochastic processes. *J. Roy. Statist. Soc. Ser. B*, 30 :499–510, 1968.
- [Kin76] J. F. C. Kingman. Subadditive processes. In *École d’Été de Probabilités de Saint-Flour, V—1975*, pages 167–223. *Lecture Notes in Math.*, Vol. 539. Springer, Berlin, 1976.

BIBLIOGRAPHIE

- [Kre85] Ulrich Krengel. *Ergodic theorems*, volume 6 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1985. With a supplement by Antoine Brunel.
- [KW82] Yitzhak Katznelson and Benjamin Weiss. A simple proof of some ergodic theorems. *Israel J. Math.*, 42(4) :291–296, 1982.
- [Lig85a] Thomas M. Liggett. An improved subadditive ergodic theorem. *Ann. Probab.*, 13(4) :1279–1285, 1985.
- [Lig85b] Thomas M. Liggett. *Interacting particle systems*, volume 276 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1985.
- [Lin02] Torgny Lindvall. *Lectures on the coupling method*. Dover Publications Inc., Mineola, NY, 2002. Corrected reprint of the 1992 original.
- [LSS97] T.M. Liggett, R.H. Schonmann, and A.M. Stacey. Domination by product measures. *Ann. Probab.*, 25 :71–95, 1997.
- [Pis96] Agoston Pisztora. Surface order large deviations for Ising, Potts and percolation models. *Probab. Theory Related Fields*, 104(4) :427–466, 1996.
- [Ste89] J. Michael Steele. Kingman’s subadditive ergodic theorem. *Ann. Inst. H. Poincaré Probab. Statist.*, 25(1) :93–98, 1989.
- [Str93] Daniel W. Stroock. *Probability theory, an analytic view*. Cambridge University Press, Cambridge, 1993.