
Université Pierre et Marie Curie

DEA Probabilités et Applications

**Chaînes de Markov
et
Processus de Poisson**

J. Lacroix (2001/2002)

Table des Matières

1	Compléments de Probabilités	5
1.1	Intégration	5
1.1.1	Espérances conditionnelles et classes monotones	5
1.1.2	Espérances conditionnelles et lois conditionnelles	6
1.1.3	Équi-intégrabilité	7
1.2	Martingales	8
1.3	Topologies sur les mesures	10
1.4	Transformations de Fourier et Laplace	16
1.4.1	Transformée de Fourier sur $\mathcal{M}_b(\mathbb{R}^d)$	16
1.4.2	Transformée de Fourier-Laplace sur $\mathcal{M}^+(\mathbb{R}^d)$	19
1.5	Convergence en loi	21
1.5.1	T.C.L.	22
1.5.2	T.C.L. local	22
1.6	Grandes déviations	23
2	Chaînes de Markov	29
2.1	Probabilités de transition	29
2.2	Suites markoviennes	31
2.2.1	Définition et exemples	31
2.2.2	Formes équivalentes de la Propriété de Markov	33
2.3	Chaînes canoniques	35
2.3.1	Définition et propriétés	35
2.3.2	Théorie du potentiel	39
2.4	Les chaînes discrètes	42
2.4.1	Rappels	42
2.4.2	Critères de Foster	45
2.5	Marches aléatoires	46
2.5.1	Réurrence, transience	46
2.5.2	Théorème ergodique	48
2.5.3	Critères de Fourier	51
2.6	Chaînes récurrentes	54
2.6.1	Chaînes ergodiques	55
2.6.2	Chaînes récurrentes au sens de Harris	56

2.6.3	Théorie ergodique des chaînes atomiques	60
2.6.4	Chaînes scindées	67
2.6.5	Théorie de Doeblin “locale”	73
2.6.6	Théorie de Doeblin “globale”	74
2.6.7	La méthode de simulation “exacte” de Propp & Wilson	76
2.7	Théorie Spectrale	79
2.7.1	Cas d’un espace d’états fini	80
2.7.2	Le cas d’un espace d’états compact	82
3	Processus de Poisson et de Renouvellement	87
3.1	Processus ponctuels	87
3.2	Processus de Poisson	88
3.3	Processus de Poisson sur \mathbb{R}^+	93
3.3.1	Compteurs	93
3.3.2	Le processus de Poisson homogène sur \mathbb{R}^+	95
3.3.3	Caractérisations du processus de Poisson homogène sur \mathbb{R}^+	97
3.3.4	Processus de Poisson composés	99
3.4	Le Renouvellement sur \mathbb{R}^+	99
3.5	Le Renouvellement sur \mathbb{R}	105

Chapitre 1

Compléments de Probabilités

La preuve des résultats des sections 1.1, 1.2, et 1.5.1 – 2 figure dans les ouvrages suivants:

- Le livre de Neveu [6] pour les espérances conditionnelles et les familles équi-intégrables.
- Le livre de Neveu [7] pour les martingales.
- Le livre de Feller [5] Tome II pour le TCL et ses extensions.

En conséquence les résultats de ces sections seront énoncés sans démonstration.

1.1 Intégration

1.1.1 Espérances conditionnelles et classes monotones

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et Z une variable aléatoire réelle positive (non nécessairement finie) et \mathcal{G} une sous tribu de \mathcal{F} .

Définition 1.1.1. *On appelle espérance conditionnelle de Z quand \mathcal{G} , que l'on note $\mathbb{E}(Z|\mathcal{G})$, la classe d'équivalence (pour l'égalité \mathbb{P} presque sûre) de variables aléatoires U telles que:*

1. U est \mathcal{G} mesurable.
2. $\int_G Z \, d\mathbb{P} = \int_G U \, d\mathbb{P}$ pour tout $G \in \mathcal{G}$.

Si Z est intégrable il en est de même des éléments de $\mathbb{E}(Z|\mathcal{G})$ et l'on peut donc définir par différence l'espérance conditionnelle de toute variable aléatoire intégrable. En fait il y a essentiellement deux façons d'introduire l'espérance conditionnelle:

1. Le théorème de Radon-Nikodym: Soit Z une variable aléatoire réelle positive, on définit la mesure positive $\mu(A) = \mathbb{E}(Z\mathbf{1}_{\{A\}})$. La restriction $\mu_{\mathcal{G}}$ de la mesure μ à la sous-tribu \mathcal{G} est absolument continue par rapport à $\mathbb{P}_{\mathcal{G}}$ et l'espérance conditionnelle est définie comme la densité de $\mu_{\mathcal{G}}$ par rapport à $\mathbb{P}_{\mathcal{G}}$.

2. Le théorème de projection dans les espaces de Hilbert: Si on pose $F = L^2(\Omega, \mathcal{F}, \mathbb{P})$ et $G = L^2(\Omega, \mathcal{G}, \mathbb{P})$ alors l'espérance conditionnelle de $Z \in F$ est définie comme sa projection orthogonale sur G . Cette définition est ensuite étendue aux variables aléatoires positives.

Un résultat très utile pour la caractérisation des espérances conditionnelles est fourni par le théorème des classes monotones:

Théorème 1.1.2. *Soient \mathcal{H} un espace vectoriel de fonctions réelles bornées définies sur Ω et \mathcal{C} un ensemble de parties de Ω stable par intersection finie. On suppose que,*

- (i) $1 \in \mathcal{H}$,
- (ii) si $f_n \in \mathcal{H}$ et si $0 \leq f_n \uparrow f$ bornée, $f \in \mathcal{H}$.
- (iii) pour tout $A \in \mathcal{C}$, $1_A \in \mathcal{H}$.

Alors \mathcal{H} contient toutes les fonctions $\sigma(\mathcal{C})$ -mesurables bornées.

A titre d'application, on peut facilement obtenir le résultat suivant: Soit X une variable aléatoire définie sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans un espace mesurable (E, \mathcal{E}) . On suppose que la tribu \mathcal{G} est engendrée par une famille dénombrable de variables aléatoires Y_n à valeurs dans un espace mesurable (Y, \mathcal{Y}) et soit f une fonction réelle \mathcal{E} mesurable, positive (resp. bornée). Pour tester qu'une variable aléatoire U , \mathcal{G} mesurable positive (resp. bornée) est bien une version de l'espérance conditionnelle de $f(X)$ quand \mathcal{G} il suffit de vérifier que:

$$\mathbb{E}(f(X) \prod_{k=1}^n f_k(Y_k)) = \mathbb{E}(U \prod_{k=1}^n f_k(Y_k))$$

pour tout $n \geq 1$ et tout système f_1, f_2, \dots, f_n de fonctions \mathcal{Y} mesurables, positives (resp. bornées).

1.1.2 Espérances conditionnelles et lois conditionnelles

Soit X une variable aléatoire définie sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans un espace mesurable (E, \mathcal{E}) et \mathcal{G} une sous tribu de \mathcal{F} . On définit alors la loi conditionnelle de X quand \mathcal{G} par la formule:

$$\mathbb{P}_X(B|\mathcal{G}) = \mathbb{E}(1_{\{B\}} \circ X|\mathcal{G}), \quad B \in \mathcal{E}$$

Bien sûr pour B fixé, cette formule ne définit $\mathbb{P}_X(B)$ que comme une classe d'équivalence de variables aléatoires et le théorème de convergence monotone ne fournit qu'une σ additivité \mathbb{P} presque sûre. Il s'agit alors de savoir s'il est possible de trouver pour chaque $B \in \mathcal{E}$ une version de cette classe d'équivalence soit $\mathbb{P}_X(B|\mathcal{G})(\omega)$ de telle sorte que pour chaque ω celle ci soit une loi de probabilité sur (E, \mathcal{E}) . D'après le théorème de Jirina, ceci est toujours possible si la tribu \mathcal{E} est à base dénombrable. On a alors pour toute fonction \mathcal{E} mesurable f , positive ou bornée:

$$\mathbb{E}(f(X)|\mathcal{G}) = \int_E f(x) d\mathbb{P}_X(x|\mathcal{G})$$

Dans le cas où \mathcal{G} est engendrée par une variable aléatoire Y on notera cette loi conditionnelle $\mathbb{P}_X(B|Y)$ et comme cette dernière quantité factorise à travers la variable aléatoire Y on écrit souvent $\mathbb{P}_X(B|Y) = \mathbb{P}_X(B|Y = \bullet) \circ Y$ et donc:

$$\mathbb{E}(f(X)|Y = y) = \int_E f(x) d\mathbb{P}_X(x|Y = y)$$

On a alors la proposition fondamentale suivante qui sera d'un usage constant dans toute la théorie des chaînes de Markov:

Proposition 1.1.3. *Soit U et V deux variables aléatoires et $\varphi(u, v)$ une fonction réelle mesurable positive ou bornée, alors:*

$$\mathbb{E}(\varphi(U, V)|V = v) = \int \varphi(u, v) d\mathbb{P}_U(u|V = v)$$

en particulier si U et V sont indépendantes on a:

$$\mathbb{E}(\varphi(U, V)|V = v) = \mathbb{E}(\varphi(U, v))$$

1.1.3 Équi-intégrabilité

Définition 1.1.4. *Soit X_i , $i \in I$ une famille de variables aléatoires réelles sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$.*

On dit que cette famille est équi-intégrable (en abrégé e.i.) si:

$$\lim_{a \rightarrow \infty} \sup_{i \in I} \int_{|X_i| > a} |X_i| d\mathbb{P} = 0$$

On dit que cette famille est équi-continue (en abrégé e.c.) si:

$$\forall \epsilon > 0, \exists \alpha > 0 \text{ tel que } \mathbb{P}(A) \leq \alpha \Rightarrow \sup_{i \in I} \int_A |X_i| d\mathbb{P} \leq \epsilon$$

Il est très facile de constater que si I est réduit à un élément alors ces propriétés sont satisfaites dès que la variable aléatoire X_i est intégrable.

Proposition 1.1.5. *Soit X_i , $i \in I$ une famille de variables aléatoires réelles sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$.*

Alors les propriétés suivantes sont équivalentes:

1. *La famille X_i est e.i.*
2. *La famille X_i e.c et de plus bornée dans L^1 .*

On déduit facilement de cette caractérisation que la somme de deux familles e.i. est e.i..

Z. Le seul fait qu'une famille soit bornée dans L^1 n'implique pas qu'elle soit équi-intégrable mais la proposition suivante montre que c'est le cas dès que la famille est bornée dans un L^p pour $1 < p \leq \infty$:

Proposition 1.1.6. *S'il existe une application f de \mathbb{R}^+ dans \mathbb{R}^+ avec:*

1. $\lim_{x \rightarrow +\infty} \frac{f(x)}{x} = +\infty$
2. *La famille $f(|X_i|)$ est bornée dans L^1*

alors la famille X_i est e.i..

Le plus souvent on utilise ce critère pour $f(x) = x^p$, $p > 1$ ou bien $f(x) = x \log^+(x)$. En plus de ce critère on a aussi:

Exemple 1.1.7. *Les exemples suivants sont les cas typiques de familles e.i.:*

1. *Une famille X_i majorée en valeur absolue par une variable aléatoire de L^1 .*
2. *Une suite X_n convergente dans L^1 .*
3. *Une famille $X_i = \mathbb{E}(X|\mathcal{F}_i)$ où X est une variable aléatoire intégrable et \mathcal{F}_i est une famille de sous-tribus de \mathcal{F}*
4. *Une famille $X_i = \mathbb{E}(Y_i|\mathcal{B})$ où Y_i est une famille e.i. et \mathcal{B} est une sous-tribu de \mathcal{F}*

Nous terminerons cette section par la généralisation du théorème de Lebesgue suivante:

Proposition 1.1.8. *Soit X_n, X des variables aléatoires et $1 \leq p < \infty$. Les deux assertions suivantes sont équivalentes:*

1. *$X \in L^p$ et X_n converge vers X dans L^p*
2. *$|X_n|^p$ est e.i. et X_n converge en probabilité vers X .*

1.2 Martingales

Les théorèmes de martingales seront souvent utilisés dans ce cours pour établir la convergence de certaines fonctionnelles d'une chaîne de Markov. On se référera principalement aux résultats suivants:

Proposition 1.2.1. *Soit $(\Omega, \mathcal{F}_n, \mathbb{P})$ un espace filtré et $\mathcal{F}_\infty = \bigvee \mathcal{F}_n$ la tribu engendrée par l'algèbre $\bigcup_n \mathcal{F}_n$.*

1. *Soit X_n une surmartingale positive (pouvant prendre la valeur $+\infty$). Elle converge presque sûrement vers une variable aléatoire X_∞ . Si de plus il existe n_0 tel que X_{n_0} soit intégrable il en est de même de X_∞ .*
2. *Soit X_n une martingale bornée dans L^1 . Elle converge presque sûrement vers une variable aléatoire X_∞ (intégrable). L'on n'a pas nécessairement convergence dans L^1 , cette propriété étant équivalente à l'équi-intégrabilité de la suite X_n . Dans ce dernier cas on a $X_n = \mathbb{E}(X_\infty|\mathcal{F}_n)$ et pour tout temps d'arrêt T (fini ou non) $X_T = \mathbb{E}(X_\infty|\mathcal{F}_T)$*

3. Si X_n est une martingale du type $X_n = \mathbb{E}(Y|\mathcal{F}_n)$ avec Y intégrable ou positive (non nécessairement finie) alors $X_\infty = \mathbb{E}(Y|\mathcal{F}_\infty)$.

Z. • On peut d'ores et déjà remarquer que cette proposition ne sera que de peu de secours pour le cas le plus courant de martingale (et de chaîne de Markov!) de la forme $X_n = Y_1 + Y_2 + \dots + Y_n$ où les variables aléatoires Y_n sont i.i.d. et d'espérance nulle. En effet, une telle martingale n'est pas en général bornée dans L^1 car sinon la convergence presque sûre de X_n entraînerait la convergence presque sûre (et donc en loi !) de Y_n vers 0 ce qui n'est possible que si les variables aléatoires Y_n sont presque sûrement nulles!

• Soit X_n une martingale bornée dans L^1 . En général on ne peut pas affirmer que pour un temps d'arrêt fini T on ait X_T intégrable et encore moins que X_T soit l'espérance conditionnelle de X_∞ quand \mathcal{F}_T . Néanmoins on a le résultat suivant:

Proposition 1.2.2. *Soit X_n une martingale de type $X_n = Y_1 + Y_2 + \dots + Y_n$ où les variables aléatoires Y_n sont i.i.d., intégrables et d'espérance nulle. Alors pour tout temps d'arrêt T intégrable on a:*

- La variable aléatoire X_T est intégrable et $X_{n \wedge T} = \mathbb{E}(X_T|\mathcal{F}_n)$.
- La suite $X_{n \wedge T}$ converge donc vers X_T dans L^1 .

Une conséquence importante de ce résultat est la suivante: Si Y_n est une suite i.i.d. intégrable et T un temps d'arrêt intégrable alors $\mathbb{E}(X_T) = \mathbb{E}(Y)\mathbb{E}(T)$ (Lemme de Wald).

On utilisera aussi la théorie des martingales de carré intégrable et les propriétés de la décomposition de Doob:

Proposition 1.2.3. *Soit X_n une martingale de carré intégrable. On peut écrire $X_n^2 = M_n + A_n$ où M_n est une martingale intégrable et A_n un processus croissant prévisible défini par:*

$$A_0 = 0, \text{ et } A_{n+1} - A_n = \mathbb{E}((X_{n+1} - X_n)^2|\mathcal{F}_n)$$

On a alors le résultat de convergence suivant:

1. Si $\mathbb{E}(A_\infty) < \infty$ (ou de façon équivalente, si la suite X_n bornée dans L^2) la suite X_n converge p.s. et dans L^2 .
2. Sur $\{A_\infty < \infty\}$ la suite X_n converge p.s. vers une variable aléatoire finie.
3. Sur $\{A_\infty = \infty\}$ la suite $\frac{X_n}{A_n}$ converge p.s. vers 0.

1.3 Topologies sur les mesures

On se restreint ici au cas de mesures sur \mathbb{R}^d mais en fait, on peut facilement constater que tous les résultats sont valables sur un espace localement compact dont la topologie admet une base dénombrable d'ouverts (espaces L.C.D).

Définition 1.3.1. *On note:*

- $\mathcal{M}(\mathbb{R}^d)$, (resp. $\mathcal{M}_b(\mathbb{R}^d)$) l'ensemble des mesures de Radon sur \mathbb{R}^d , (resp. l'ensemble des mesures de Radon bornées sur \mathbb{R}^d). On rappelle qu'une mesure est de Radon s'il s'agit d'une mesure signée sur la tribu borélienne de \mathbb{R}^d dont la valeur absolue est finie sur les compacts.
- $\mathcal{M}_1^+(\mathbb{R}^d)$ l'ensemble des probabilités sur \mathbb{R}^d .
- $\mathcal{C}_k(\mathbb{R}^d)$, (resp. $\mathcal{C}_b(\mathbb{R}^d)$) l'ensemble des fonctions continues à support compact sur \mathbb{R}^d , (resp. l'ensemble des fonctions continues bornées sur \mathbb{R}^d).

Définition 1.3.2. *On définit la topologie vague sur $\mathcal{M}(\mathbb{R}^d)$ par la famille des écarts associés à $\mathcal{C}_k(\mathbb{R}^d)$ et la topologie étroite sur $\mathcal{M}_b(\mathbb{R}^d)$ par la famille des écarts associés à $\mathcal{C}_b(\mathbb{R}^d)$. Autrement dit on a une base de voisinages de μ en considérant des ensembles de la forme $\{\nu ; |\mu(f_i) - \nu(f_i)| \leq \epsilon, i \in I\}$ où $\epsilon > 0$ et les $f_i, i \in I$ parcourent les sous familles finies de $\mathcal{C}_k(\mathbb{R}^d)$ ou $\mathcal{C}_b(\mathbb{R}^d)$. Il s'en suit que:*

- Une suite $\mu_n \in \mathcal{M}(\mathbb{R}^d)$ converge vaguement vers $\mu \in \mathcal{M}(\mathbb{R}^d)$ si et seulement si $\lim(\mu_n(f)) = \mu(f)$ pour toute fonction $f \in \mathcal{C}_k(\mathbb{R}^d)$.
- Une suite $\mu_n \in \mathcal{M}_b(\mathbb{R}^d)$ converge étroitement vers $\mu \in \mathcal{M}_b(\mathbb{R}^d)$ si et seulement si $\lim(\mu_n(f)) = \mu(f)$ pour toute fonction $f \in \mathcal{C}_b(\mathbb{R}^d)$.

En fait, les restrictions de ces topologies aux cônes positifs $\mathcal{M}^+(\mathbb{R}^d)$ et $\mathcal{M}_b^+(\mathbb{R}^d)$ sont métrisables et l'on peut donc se restreindre à la notion de convergence des suites si l'on ne considère que des mesures positives. De plus $\mathcal{M}_b(\mathbb{R}^d)$ muni de la topologie étroite s'identifie à un sous espace vectoriel du dual de l'espace de Banach $\mathcal{C}_b(\mathbb{R}^d)$ muni de sa topologie affaiblie $\sigma(\mathcal{C}_b'(\mathbb{R}^d), \mathcal{C}_b(\mathbb{R}^d))$.

Z. Les espaces $\mathcal{M}(\mathbb{R}^d)$ et $\mathcal{M}_b(\mathbb{R}^d)$ ne sont pas métrisables ...

Z. Toute forme linéaire positive I sur $\mathcal{C}_b(\mathbb{R}^d)$ n'est pas nécessairement l'extension à $\mathcal{C}_b(\mathbb{R}^d)$ d'une mesure bornée !!! Il faut pour ceci que ce soit un prolongement de Daniell, c'est à dire que pour toute suite de fonctions positives $f_n \in \mathcal{C}_b$ décroissante vers 0 on ait $\lim_n \downarrow I(f_n) = 0$.

La convergence étroite sur $\mathcal{M}_b(\mathbb{R}^d)$ implique bien sûr la convergence vague mais la réciproque est fausse. Cependant cette réciproque est exacte sous certaines conditions:

Proposition 1.3.3. *Soit μ_n et μ des éléments de $\mathcal{M}_b^+(\mathbb{R}^d)$. Si*

- $\mu_n(1) \rightarrow \mu(1)$
- μ_n converge vaguement vers μ

alors μ_n converge étroitement vers μ .

Indications de preuve:

Soit f une fonction continue bornée et φ_ℓ une suite de fonctions continues à support compact qui tend en croissant vers la fonction constante 1. Les fonctions $f\varphi_\ell$ sont continues à support compact et:

$$\lim_n \mu_n(f\varphi_\ell) = \mu(f\varphi_\ell) \quad , \quad \lim_n \mu_n(\varphi_\ell) = \mu(\varphi_\ell) \quad , \quad \lim_l \uparrow \mu(\varphi_\ell) = \mu(1)$$

Par conséquent de l'inégalité

$$|\mu(f) - \mu(f\varphi_\ell)| \leq \|f\|_\infty (\mu(1) - \mu(\varphi_\ell))$$

appliquée aussi à la proba μ_n on déduit:

$$|\mu_n(f) - \mu(f)| \leq |\mu_n(f\varphi_\ell) - \mu(f\varphi_\ell)| + \|f\|_\infty (\mu(1) + \mu_n(1) - \mu(\varphi_\ell) - \mu_n(\varphi_\ell))$$

et on obtient le résultat en faisant tendre successivement n puis ℓ vers l'infini. ■

Proposition 1.3.4. Soit f une fonction s.c.i positive. L'application $\mu \rightarrow \mu(f)$ est elle-même s.c.i sur $\mathcal{M}_b^+(\mathbb{R}^d)$ muni de la topologie étroite, c'est à dire que si μ_n est une suite dans $\mathcal{M}_b^+(\mathbb{R}^d)$ qui converge étroitement vers μ on a:

$$\mu(f) \leq \underline{\lim} (\mu_n(f))$$

De même si g est s.c.s. bornée alors $\mu \rightarrow \mu(g)$ est s.c.s. pour la topologie étroite

Indications de preuve:

Il suffit d'utiliser le fait que f est l'enveloppe supérieure d'une suite de $\mathcal{C}_b^+(\mathbb{R}^d)$. ■

Corollaire 1.3.5. Soit $f \in \mathcal{B}_b^+(\mathbb{R}^d)$ et μ_n est une suite dans $\mathcal{M}_b^+(\mathbb{R}^d)$ qui converge étroitement vers μ . Si l'ensemble des points de discontinuité de f est μ négligeable on a:

$$\lim \mu_n(f) = \mu(f)$$

Indications de preuve:

On note $f_1 = \underline{\lim} (f)$ et $f_2 = \overline{\lim} (f)$ les régularisées s.c.i et s.c.s de f . Il suffit alors d'appliquer le résultat précédent sachant que $f_1 \leq f \leq f_2$ et $\mu(f_1) = \mu(f_2) = \mu(f)$. ■

Corollaire 1.3.6. Soit μ_n une suite de probabilités qui converge étroitement vers μ et B un borélien de \mathbb{R}^d . On a:

$$\mu(\overset{\circ}{B}) \leq \underline{\lim} \mu_n(B) \leq \overline{\lim} \mu_n(B) \leq \mu(\overline{B})$$

En particulier on a convergence des mesures de boréliens dont la frontière est de mesure μ nulle.

Un procédé courant pour prouver qu'une suite est convergente consiste tout d'abord à prouver que cette suite constitue un ensemble relativement compact, puis de prouver que toutes ses valeurs d'adhérence sont égales. Il nous sera donc utile d'avoir des critères de compacité dans les espaces de mesures.

Proposition 1.3.7. *Soit $A \subset \mathcal{M}^+(\mathbb{R}^d)$. Si A vérifie l'une des conditions équivalentes suivantes:*

1. *A est vaguement borné c'est à dire $\sup_{\mu \in A} |\mu(f)| < \infty$ pour tout $f \in \mathcal{C}_k(\mathbb{R}^d)$*
2. *A est fortement borné c'est à dire que pour tout compact K de \mathbb{R}^d il existe une constante finie c_K telle que $\sup_{\mu \in A} |\mu(f)| < c_K \|f\|$ pour tout $f \in \mathcal{C}_k(\mathbb{R}^d)$ à support dans K .*
3. *Pour tout compact K de \mathbb{R}^d on a $\sup_{\mu \in A} \mu(K) < \infty$*

alors A est vaguement relativement compact.

Indications de preuve:

L'équivalence de (2) et (3) ainsi que l'implication (2) \Rightarrow (1) sont faciles. L'implication (1) \Rightarrow (2) est une conséquence du théorème de la borne uniforme. La fermeture vague de A soit \overline{A} possède aussi ces propriétés équivalentes. Soit $\mathcal{C}_{k,n}^+(\mathbb{R}^d)$ le sous ensemble de $\mathcal{C}_k^+(\mathbb{R}^d)$ formé des fonctions à support dans la boule de rayon n et H_n un sous ensemble dense pour la norme uniforme. Il est facile de prouver que la topologie vague sur $\mathcal{M}^+(\mathbb{R}^d)$ peut être définie à partir des écarts associés à $H = \bigcup_n H_n$. Soit alors μ_n une suite dans \overline{A} . Pour $f \in H$ la suite réelle $\mu_n(f)$ est bornée, elle admet donc une sous suite convergente et par un procédé diagonal on obtient une sous suite qui converge pour tout $f \in H$ et donc aussi pour tout $f \in \mathcal{C}_k(\mathbb{R}^d)$ vers un réel noté $I(f)$. Il est facile de constater que I est une forme linéaire positive sur $\mathcal{C}_k(\mathbb{R}^d)$ et c'est donc une mesure de Radon. ■

Proposition 1.3.8. *Soit $A \subset \mathcal{M}_b^+(\mathbb{R}^d)$ vérifiant:*

- *A est borné en norme c'est à dire $\sup_{\mu \in A} |\mu(1)| < \infty$*
- *A est tendu, c'est à dire:*

$$\forall \epsilon > 0 \exists K_\epsilon \text{ compact de } \mathbb{R}^d \text{ tel que } \sup_{\mu \in A} \mu(\mathbb{R}^d \setminus K_\epsilon) \leq \epsilon$$

alors A est étroitement relativement compact.

Indications de preuve:

On commence par remarquer (en utilisant 1.3.4) que l'adhérence étroite de A soit \overline{A} est elle même tendue et bornée en norme. D'après la proposition précédente tout sous ensemble de $\mathcal{M}_b^+(\mathbb{R}^d)$ borné en norme est vaguement compact. La topologie étroite sur $\mathcal{M}_b^+(\mathbb{R}^d)$ étant métrisable, il suffit de considérer des suites. Soit μ_n une suite dans \overline{A} : elle possède une sous suite μ_{n_k} vaguement convergente vers une mesure de Radon positive

μ . D'après 1.3.3 il suffit de montrer que $\mu_{n_k}(1)$ converge vers $\mu(1)$. En considérant une suite φ_ℓ de $\mathcal{C}_k^+(\mathbb{R}^d)$ qui converge en croissant vers 1 et qui vaut 1 sur la boule de rayon ℓ on obtient $\mu(1) \leq \underline{\lim} \mu_{n_k}(1)$. Soit $\epsilon > 0$ et K_ϵ associé. A partir d'un certain rang on a $\varphi_\ell \geq 1$ sur K_ϵ et donc $\mu_{n_k}(\varphi_\ell) \geq \mu_{n_k}(1) - \epsilon$. On en déduit donc que $\mu(1) \geq \overline{\lim} \mu_{n_k}(1) - \epsilon$ d'où la conclusion. ■

En fait la tension est une condition nécessaire de convergence étroite comme le montre la proposition suivante:

Proposition 1.3.9. *Soit μ_n une suite dans $\mathcal{M}_b^+(\mathbb{R}^d)$ qui converge étroitement vers μ . Alors μ_n est tendue.*

Indications de preuve:

Soit $\epsilon > 0$ donné. Il existe alors une boule ouverte B de complémentaire (fermé) F telle que $\mu(F) \leq \epsilon/2$. De l'inégalité $\overline{\lim} \mu_n(F) \leq \mu(F)$ on déduit qu'il existe N à partir duquel on a $\mu_n(F) \leq \epsilon$. Il suffit maintenant de choisir des compacts portant les μ_n à ϵ près pour $n = 1, \dots, N-1$ et de prendre leur réunion avec \overline{B} pour avoir le résultat. ■

Extensions

En fait la notion de convergence étroite peut être définie pour des mesures sur des espaces plus généraux que \mathbb{R}^d et cette extension est nécessaire si l'on veut étudier les convergences en loi de processus. Soit E un espace métrique muni de sa tribu borélienne et $\mathcal{C}_b(E)$ l'espace de Banach des fonctions continues bornées sur E .

Définition 1.3.10. (Voir [1] page 58). *Une mesure de Radon positive μ sur E est définie comme une forme linéaire positive sur $\mathcal{C}_b(E)$ telle que : $\forall \epsilon > 0 \exists K_\epsilon$ compact de E tel que pour $f \in \mathcal{C}_b^+(E)$, $f \leq 1$, f nulle sur K_ϵ on ait $\mu(f) \leq \epsilon$*

On définit toujours la convergence étroite sur $\mathcal{M}_b^+(E)$ comme celle induite par les écarts associés à $\mathcal{C}_b(E)$ et l'on voit que les propositions 1.3.4, 1.3.5, 1.3.6, se généralisent facilement.

Z. Pour étendre la proposition 1.3.9 on doit supposer que les mesures μ_n et μ sont régulières c'est à dire que la mesure d'un borélien est le supremum des mesures des compacts qu'il contient. Cette propriété est toujours vraie si E est polonais c'est à dire métrique séparable et complet.

On peut identifier E à une partie dense d'un compact X tel que tout élément $f \in \mathcal{C}_b(E)$ se prolonge (de façon nécessairement unique) en un élément de $\mathcal{C}(X)$ (compactification de Stone Çech). On peut par exemple considérer l'application i de E dans $\mathcal{C}_b'(E)$ muni de la topologie affaiblie $\sigma(\mathcal{C}_b'(E), \mathcal{C}_b(E))$ définie par $i(x)(f) = f(x)$. On constate facilement que i est injective et bicontinue de E sur $i(E)$, et que de plus on prolonge $f \in \mathcal{C}_b(E)$ à $f \in \mathcal{C}_b(X)$ par $\mu \rightarrow \mu(f)$. Puisque la boule unité de $\mathcal{C}_b'(E)$ est compacte pour la topologie affaiblie on prend pour X l'adhérence de $i(E)$. L'espace $\mathcal{M}_b^+(E)$ muni de la convergence

étroite s'identifie alors à un sous espace de $\mathcal{M}^+(X)$ muni de la convergence vague (ou étroite...).

Z. En général X n'est pas métrisable et E n'est pas borélien dans X .

On peut alors montrer que le critère de compacité étroite établi dans 1.3.8 est toujours valable dans cette situation:

Proposition 1.3.11. *Soit $A \subset \mathcal{M}_b^+(E)$ vérifiant:*

- *A est borné en norme c'est à dire $\sup_{\mu \in A} |\mu(1)| < \infty$*
- *A est tendu, c'est à dire:*

$$\forall \epsilon > 0 \exists K_\epsilon \text{ compact de } E \text{ tel que } \sup_{\mu \in A} \mu(E \setminus K_\epsilon) \leq \epsilon$$

alors A est étroitement relativement compact.

Indications de preuve:

Soit \tilde{A} l'adhérence de A considéré comme sous ensemble de $\mathcal{M}^+(X)$ identifié à $\mathcal{C}_b'(X)$ muni de la topologie affaiblie $\sigma(\mathcal{C}_b'(X), \mathcal{C}_b(X))$. \tilde{A} est donc compact et tout revient à prouver qu'il est en fait contenu dans $\mathcal{M}_b^+(E)$. Tout élément $\lambda \in \tilde{A}$ est une forme linéaire positive sur $\mathcal{C}_b(E)$, il reste à montrer que c'est une mesure de Radon sur E . Soit alors $\epsilon > 0$ et K_ϵ associé par la propriété de tension. Puisque $X \setminus K_\epsilon$ est un ouvert de X , on montre par un raisonnement tout à fait analogue à 1.3.4 que l'ensemble $\{\mu \in \mathcal{M}^+(X) ; \mu(X \setminus K_\epsilon) \leq \epsilon\}$ est fermé dans $\mathcal{M}^+(X)$. Cet ensemble contient A et donc aussi \tilde{A} . Soit alors $f \in \mathcal{C}_b^+(E)$ avec $f \leq 1$ et f nulle sur K_ϵ , on a bien sûr $\lambda(f) \leq \epsilon$. ■

Le cas des processus continus

On considère le cas particulier $E = \mathbf{C}(\mathbb{R}^+, \mathbb{R}^d)$ l'espace des fonctions continues de \mathbb{R}^+ dans \mathbb{R}^d muni de la topologie de la convergence uniforme sur tout compact. C'est un espace métrique séparable complet (polonais) et sa tribu borélienne est engendrée par les applications "coordonnées" $\pi_t(f) = f(t)$, $t \in \mathbb{R}^+$. Pour une probabilité μ on appelle "répartition finie" de μ l'image de μ dans \mathbb{R}^I par l'application $(\pi_{t_i}, i \in I)$ avec I fini. Soit μ_n et μ des éléments de $\mathcal{M}_1^+(E)$. Il est facile de constater que si μ_n converge étroitement vers μ , il en est de même des systèmes de répartitions finies dans (\mathbb{R}^d) . Par contre la réciproque est fausse: il suffit de considérer la suite μ_n des masses de Dirac sur les fonctions f_n telles que $f_n(0) = 0$, $f_n(1/n) = n$, $f_n(x) = 0$ sur $[2/n, \infty[$ et f_n est linéaire sur $[0, 1/n]$ et sur $[1/n, 2/n]$. En effet les répartitions finies convergent vers celles de la masse de Dirac sur la fonction nulle mais μ_n ne converge pas étroitement vers cette masse de Dirac. Par contre on a:

Proposition 1.3.12. *Soit μ_n et μ des éléments de $\mathcal{M}_1^+(E)$. Si les répartitions finies de μ_n convergent étroitement vers celles de μ et si la famille μ_n est tendue, alors μ_n converge étroitement vers μ .*

Indications de preuve:

Soit ν une valeur d'adhérence étroite de la suite μ_n . Alors les répartitions finies de ν sont égales à celles de μ donc $\nu = \mu$. ■

Pour que ce critère soit effectif, il faut caractériser les parties compactes de E , ce qui résultera du théorème d'Ascoli.

Proposition 1.3.13. *Soit $A \subset E$. Si:*

- $\sup_{f \in A} |f(0)| < \infty$
- A est équicontinu en tout point $t \in \mathbb{R}^+$

alors A est relativement compact dans E .

Corollaire 1.3.14. *Soit $a > 0$ et $\alpha = \{\alpha_{N,p}\}$ une suite double strictement positive, définie pour $N, p \geq 1$. Alors le sous ensemble de E*

$$A(a, \alpha) = \{(|f(0)| \leq a) \bigcap_{(N,p) \geq 1} (\sup_{t, t' \leq N; |t-t'| \leq \alpha_{N,p}} |f(t) - f(t')| \leq 1/p)\}$$

est compact dans E .

Nous terminons par une conséquence importante de cette étude:

Théorème 1.3.15. *Soit X_n une suite de processus continus définis sur $(\Omega, \mathcal{F}, \mathbb{P})$, à valeurs réelles. Si:*

- *Les lois conjointes du processus X_n convergent étroitement.*
- *Pour tout $\epsilon > 0$ et $T > 0$ on a:*

$$\lim_{h \rightarrow 0} \left(\sup_n \mathbb{P} \left(\sup_{t, t' \leq T; |t-t'| \leq h} |X_n(t) - X_n(t')| \geq \epsilon \right) \right) = 0$$

alors la loi de X_n converge étroitement sur E vers la loi d'un processus à trajectoires continues.

Indications de preuve:

Soit μ_n la loi de X_n sur E . La convergence étroite de la suite de variables aléatoires $X_n(0)$ implique que leurs lois sont tendues et que donc pour tout $\epsilon > 0$ il existe $a > 0$ tel que $\mu_n\{f ; |f(0)| \geq a\} \leq \epsilon/2$. D'autre part la condition de l'énoncé implique que pour chaque couple (N, p) il existe $\alpha_{N,p} > 0$ avec

$$\sup_n \mathbb{P} \left(\sup_{t, t' \leq N; |t-t'| \leq \alpha_{N,p}} |X_n(t) - X_n(t')| \geq 1/p \right) \leq \frac{\epsilon}{2^{N+p+1}}$$

On en déduit que

$$\begin{aligned}\mu_n(E \setminus A(a, \alpha)) &\leq \epsilon/2 + \sum_{N,p} \frac{\epsilon}{2^{N+p+1}} \\ &\leq \epsilon\end{aligned}$$

et que donc la suite μ_n est tendue. ■

Ce théorème a une application immédiate pour la convergence des marches aléatoires normalisées vers un mouvement brownien (Théorème de Donsker).

1.4 Transformations de Fourier et Laplace

1.4.1 Transformée de Fourier sur $\mathcal{M}_b(\mathbb{R}^d)$

Définition 1.4.1. Soit t et x des éléments de \mathbb{R}^d .

- On pose $\gamma_t(x) = \exp(i \langle t, x \rangle)$ où $\langle t, x \rangle$ désigne le produit scalaire sur \mathbb{R}^d .
- Soit $\mu \in \mathcal{M}_b(\mathbb{R}^d)$, on appelle fonction caractéristique (ou transformée de Fourier) de μ la fonction définie pour $t \in \mathbb{R}^d$ par

$$\hat{\mu}(t) = \int_{\mathbb{R}^d} \gamma_t(x) d\mu(x)$$

Nous commençons par énoncer quelques propriétés immédiates de la fonction caractéristique:

Proposition 1.4.2. La fonction caractéristique de $\mu \in \mathcal{M}_b^+(\mathbb{R}^d)$ vérifie:

1. $\hat{\mu}(0) \geq 0$, $\hat{\mu}(-t) = \overline{\hat{\mu}(t)}$, $|\hat{\mu}(t)| \leq \hat{\mu}(0)$
2. $\hat{\mu}(t)$ est une fonction de type positif c'est à dire que pour tout entier $n \geq 1$ et tout système (t_1, t_1, \dots, t_n) dans \mathbb{R}^d la matrice carrée d'ordre n définie par $M(i, j) = \hat{\mu}(t_i - t_j)$ (qui est hermitienne d'après (1)) est positive c'est à dire que pour tout $z \in \mathbb{C}^n$ alors $\langle z, M\bar{z} \rangle \geq 0$.
3. $t \rightarrow \hat{\mu}(t)$ est une fonction uniformément continue sur \mathbb{R}^d .

Il est bien connu que la fonction caractéristique caractérise la mesure:

Proposition 1.4.3. L'application $\mu \rightarrow \hat{\mu}$ est injective sur \mathcal{M}_b .

Indications de preuve:

Soit \mathcal{E} l'ensemble des transformées de Fourier de fonctions dans $L^1(\mathbb{R}^d, dx)$. D'après le théorème de Stone-Weierstrass \mathcal{E} est dense dans $\mathbf{C}_0(\mathbb{R}^d)$. Soit $f \in L^1(\mathbb{R}^d, dx)$, en utilisant Fubini on obtient $\mu(\hat{f}) = \int f(t) \hat{\mu}(t) dt$. Par conséquent si deux mesures bornées μ et ν ont la même transformée de Fourier elles coïncident sur \mathcal{E} . Or l'application $f \rightarrow \mu(f)$ est continue sur $\mathbf{C}_0(\mathbb{R}^d)$. ■

Le critère suivant est fondamental dans toutes les questions de convergence étroite:

Proposition 1.4.4. Soit μ_n et μ des éléments de $\mathcal{M}_b(\mathbb{R}^d)$. Alors μ_n converge étroitement vers μ si et seulement si $\widehat{\mu_n}(t)$ converge vers $\widehat{\mu}(t)$ pour tout $t \in \mathbb{R}^d$.

Indications de preuve:

En reprenant les notations de la proposition 1.4.3 on voit que $\mu_n(g) \rightarrow \mu(g)$ pour toute fonction $g \in \mathcal{E}$. On prouve sans peine que ceci implique la même convergence pour les fonctions de $\mathbf{C}_0(\mathbb{R}^d)$ d'où la convergence étroite d'après 1.3.3. ■

Z. On se gardera d'affirmer qu'une suite de probabilités μ_n telle que $\widehat{\mu_n}(t)$ converge pour tout t est une suite étroitement convergente! En effet la limite n'est pas nécessairement la transformée de Fourier d'une loi de probabilité.

Le théorème de Paul Lévy ci dessous permet de préciser ce point.

Proposition 1.4.5. (Paul Lévy)

Soit μ_n une suite de probabilités sur \mathbb{R}^d telle que pour tout t on ait $\lim_n \widehat{\mu_n}(t) = \varphi(t)$, la fonction $\varphi(t)$ étant continue à l'origine. Alors $\varphi(t)$ est la transformée de Fourier d'une probabilité μ (et donc la suite μ_n converge étroitement vers μ).

Indications de preuve:

Il suffit de montrer que la suite μ_n est tendue et donc forme un ensemble étroitement relativement compact. La fonction φ sera donc la transformée de Fourier de l'une quelconque de des valeurs d'adhérence étroite de μ_n qui sont donc toutes identiques. On peut facilement généraliser la preuve donnée ci-dessous dans le cas $d = 1$. Soit $u > 0$ on peut écrire:

$$\begin{aligned} \frac{1}{u} \int_{-u}^u (1 - \widehat{\mu_n}(t)) dt &= \int \left(\frac{1}{u} \int_{-u}^u (1 - \gamma_t(x)) dt \right) d\mu_n(x) \\ &= 2 \int \left(1 - \frac{\sin(ux)}{ux} \right) d\mu_n(x) \\ &\geq 2 \int_{\{|ux| \geq 2\}} \left(1 - \frac{1}{|ux|} \right) d\mu_n(x) \\ &\geq \mu_n\{x ; |x| \geq 2/u\} \end{aligned}$$

En utilisant la convergence de $\widehat{\mu_n}(t)$ vers $\varphi(t)$ et la continuité de φ à l'origine on constate facilement que la suite μ_n est tendue. ■

La transformée de Fourier permet de prouver très facilement quelques résultats de convergence:

Proposition 1.4.6. Soit μ_n^i et μ^i , $i \in [1..N]$, des probabilités sur \mathbb{R}^{d_i} telles que pour chaque i la suite μ_n^i converge étroitement vers μ^i . Alors:

1. $\mu_n^1 \otimes \mu_n^2 \dots \otimes \mu_n^N$ converge étroitement vers $\mu^1 \otimes \mu^2 \dots \otimes \mu^N$

2. $\mu_n^1 * \mu_n^2 \dots * \mu_n^N$ converge étroitement vers $\mu^1 * \mu^2 \dots * \mu^N$.

Z. On ne peut affirmer que si μ est une probabilité sur \mathbb{R}^d et ν_n une suite dans $\mathcal{M}^+(\mathbb{R}^d)$ qui converge vaguement vers ν alors $\mu * \nu_n$ converge vaguement vers $\mu * \nu$. On a seulement l'inégalité $\mu * \nu \leq \underline{\lim} \mu * \nu_n$.

Pour être complet nous énonçons le théorème de Bochner dont la preuve nécessite le théorème de Paul Lévy ainsi que quelques propriétés du noyau de Fejer noté K_T . Nous nous retiendrons au cas $d = 1$ dans cette preuve.

Théorème 1.4.7. (Bochner)

Soit φ une fonction de type positif sur \mathbb{R}^d continue à l'origine et telle que $\varphi(0) = 1$. Alors il existe une (unique!) mesure de probabilité sur \mathbb{R}^d soit μ telle que $\varphi(t) = \hat{\mu}(t)$. Si la fonction φ est sommable, alors la mesure μ a une densité f par rapport à la mesure de Lebesgue sur \mathbb{R}^d donnée par:

$$f(t) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}} \varphi(x) \gamma_{-t}(x) dx$$

Indications de preuve:

Pour T réel positif on définit $h_T(t) = (1 - \frac{|t|}{T})^+$ qui est une fonction la fonction positive continue et a support compact. Pour un entier positif N on pose

$$\begin{aligned} g_N(t) &= \frac{1}{2\pi N} \int_0^N \int_0^N \gamma_x(-t) \gamma_y(t) \varphi(x - y) dx dy \\ &= \frac{1}{2\pi} \int h_N(u) \varphi(u) \gamma_{-t}(u) du \end{aligned}$$

La fonction $g_N(t)$ est positive mais on ne peut pas lui appliquer directement Fubini pour calculer son intégrale. Dans ce but on calcule:

$$\int g_N(t) h_T(t) \gamma_u(t) dt = \int h_N(v) \varphi(v) K_T(u - v) dv$$

La fonction $v \rightarrow h_N(v) \varphi(v)$ est continue et intégrable, donc d'après les propriétés du noyau de Féjer on obtient :

$$\lim_{T \rightarrow \infty} \int g_N(t) h_T(t) \gamma_u(t) dt = \lim_{T \rightarrow \infty} \int h_N(v) \varphi(v) K_T(u - v) dv = \varphi(u) h_N(u)$$

En faisant $u = 0$ le théorème de convergence monotone des fonctions positives implique

$$\lim_{T \rightarrow \infty} \int g_N(t) h_T(t) dt = \int g_N(t) dt = \varphi(0) = 1$$

et donc g_N est intégrable. Par conséquent les mesures positives μ_N de densité g_N par rapport à la mesure de Lebesgue sur \mathbb{R} sont des probabilités. La fonction g_N est la

transformée de Fourier inverse de $h_N\varphi$ qui est intégrable et donc $h_N\varphi$ est la transformée de Fourier de μ_N . La suite des transformées de Fourier des mesures μ_N convergent vers la fonction φ continue à l'origine. Le théorème de Paul Levy permet alors d'affirmer qu'il existe une probabilité μ telle que $\varphi(t) = \mu(\gamma_t)$ (et donc les mesures μ_N convergent étroitement vers μ). Si φ est sommable alors les fonctions g_N sont uniformément bornées par $\|\varphi\|_1$ et elles convergent ponctuellement vers f , par conséquent si ψ est une fonction continue à support compact:

$$\mu(\psi) = \lim \mu_T(\psi) = \lim \int g_T(x)\psi(x) dx = \int f(x)\psi(x) dx$$

et la preuve est complète. ■

1.4.2 Transformée de Fourier-Laplace sur $\mathcal{M}^+(\mathbb{R}^d)$

On définit ci dessous la transformée de Laplace complexe, mais elle est le plus souvent utilisée dans le cas d'une variable z réelle.

Définition 1.4.8. Soit μ une mesure de Radon positive sur \mathbb{R}^d et $z \in \mathbb{C}^d$. On définit alors sa transformée de Laplace

$$\mathcal{L}\mu(z) = \int \exp(\langle z, t \rangle) d\mu(t)$$

du moins lorsque l'intégrale est absolument convergente c'est à dire pour les z tels que:

$$\Re(z) \in \Omega_\mu = \{x \in \mathbb{R}^d \mid \int \exp(\langle x, t \rangle) d\mu(t) < \infty\}$$

Si μ possède une densité f par rapport à la mesure de Lebesgue on pose $\mathcal{L}f = \mathcal{L}\mu$.

Théorème 1.4.9. Soit μ une mesure de Radon positive sur \mathbb{R}^d .

1. Ω_μ est un ensemble convexe, la fonction $x \rightarrow \log(\mathcal{L}\mu(x))$ est s.c.i. et convexe sur \mathbb{R}^d et le point 0 appartient à Ω_μ si et seulement si $\mu \in \mathcal{M}_b^+$.
2. La fonction $\mathcal{L}\mu(z)$ est holomorphe dans le domaine $\{z \mid \Re(z) \in \overset{\circ}{\Omega}_\mu\}$. Dans ce domaine, ses dérivées successives sont données par les intégrales absolument convergentes:

$$\frac{\partial^n \mathcal{L}\mu(z)}{\partial^n(z_k)} = \int t_k^n \exp(\langle z, t \rangle) d\mu(t)$$

En particulier si μ est à support compact alors $\Omega_\mu = \mathbb{R}^d$ et $\mathcal{L}\mu(z)$ est holomorphe dans \mathbb{C}^d .

3. Soit μ et ν deux mesures de Radon positives sur \mathbb{R}^d telles que $\Omega_\mu \cap \Omega_\nu$ contienne un ouvert sur lequel $\mathcal{L}\mu$ et $\mathcal{L}\nu$ coïncident. Alors on a $\mu = \nu$.

4. Soit μ et ν deux mesures de Radon positives sur \mathbb{R}^d telles que la convolution $\mu * \nu$ existe dans \mathcal{M}^+ . Alors $\Omega_\mu \cap \Omega_\nu = \Omega_{\mu * \nu}$ et sur cet ensemble on a $\mathcal{L}(\mu * \nu) = (\mathcal{L}\mu) \cdot (\mathcal{L}\nu)$.

Indications de preuve:

Pour alléger les notations, on se restreint au cas $d = 1$.

Preuve de (1): La fonction $x \rightarrow \exp(tx)$ est convexe donc si x_1, x_2 sont dans Ω_μ il en est de même de $\lambda x_1 + (1 - \lambda)x_2$ pour $\lambda \in [0, 1]$ ce qui prouve que Ω_μ est convexe. En utilisant Hölder on voit que le logarithme de la transformée de Laplace est convexe sur Ω_μ et donc sur \mathbb{R} . La propriété de semi-continuité est évidente en utilisant Fatou.

Preuve de (2): Soit z_0 tel que $x_0 = \Re(z_0) \in \overset{\circ}{\Omega}_\mu$. On peut trouver $\epsilon > 0$ tel que $[x_0 - 2\epsilon, x_0 + 2\epsilon] \subset \Omega_\mu$ et alors:

$$\begin{aligned} \frac{\mathcal{L}\mu(z_0 + z) - \mathcal{L}\mu(z_0)}{z} &= \int \exp(tz_0) \frac{(\exp(tz) - 1)}{z} d\mu(t) \\ &= \int \exp(tz_0) \left(\int_0^t \exp(uz) du \right) d\mu(t) \end{aligned}$$

Pour montrer que $\mathcal{L}\mu$ est dérivable en z_0 il suffit d'après Lebesgue de majorer la fonction à intégrer par une fonction μ intégrable fixe sur un voisinage de z_0 . Or pour $|z| \leq \epsilon$:

$$\left| \exp(tz_0) \int_0^t \exp(uz) du \right| \leq |t| \exp(tx_0 + |t|\epsilon)$$

En considérant les valeurs positives ou négatives de t on voit que l'on a le résultat puisque $t \rightarrow t \exp(t(x_0 + \epsilon))$ est μ intégrable sur \mathbb{R}^+ (car $t \rightarrow \exp(t(x_0 + 2\epsilon))$ l'est) et $t \rightarrow t \exp(t(x_0 - \epsilon))$ est μ intégrable sur \mathbb{R}^- (car $t \rightarrow \exp(t(x_0 - 2\epsilon))$ l'est). On obtient alors la formule demandée par récurrence.

Preuve de (3): Soit x_0 un point intérieur à $\Omega_\mu \cap \Omega_\nu$. Les séries de Taylor des deux fonctions holomorphes $\mathcal{L}\mu$ et $\mathcal{L}\nu$ sont donc identiques au point x_0 . On en déduit que les fonctions analytiques réelles $\mathcal{L}\mu(x_0 + iy)$ et $\mathcal{L}\nu(x_0 + iy)$ sont égales. Or ce sont les transformées de Fourier des mesures bornées de densité $\exp(tx_0)$ par rapport à μ et ν . On en déduit donc $\mu = \nu$ puisque l'exponentielle est strictement positive.

Preuve de (4): Application directe de Fubini. ■

Remarque 1.4.10. Ce théorème a des conséquences importantes:

- La "quasi-injectivité" de la transformée de Laplace.
- Si le point 0 est dans l'intérieur de Ω_μ alors la transformée de Fourier de μ est une fonction analytique réelle et μ possède des moments de tous ordres.
- Une mesure à support compact ne peut avoir une transformée de Fourier à support compact.

On utilise beaucoup la transformée de Laplace pour les mesures à support dans \mathbb{R}^+ car cet ensemble est stable par convolution. Dans ce cas on préfère changer de signe dans l'exponentielle pour la définition de la transformée de Laplace de telle sorte que pour une large classe de mesures à support dans \mathbb{R}^+ l'ensemble Ω_μ contienne un intervalle du type $]c, \infty[$ avec $c > 0$. Par exemple si μ est une loi exponentielle de paramètre λ sur \mathbb{R}^+ on vérifie immédiatement que $\sum_{n \geq 1} \mu^{*n}$ est égale à λ fois la mesure de Lebesgue sur \mathbb{R}^+ .

1.5 Convergence en loi

Définition 1.5.1. Soit X_n et X des variables aléatoires à valeurs dans \mathbb{R}^d . On dit que la suite X_n converge en loi vers X (et on note $X_n \xrightarrow{L} X$) si la suite des lois des X_n converge étroitement vers la loi de X .

La propriété de convergence en loi est donc équivalente à:

$$\mathbb{E}(f(X_n)) \rightarrow \mathbb{E}(f(X)) \quad \text{pour tout } f \in \mathcal{C}_b(\mathbb{R}^d)$$

en vertu de 1.3.3 on peut remplacer $\mathcal{C}_b(\mathbb{R}^d)$ par $\mathcal{C}_k(\mathbb{R}^d)$ ou encore par l'espace des fonctions uniformément continues bornées sur \mathbb{R}^d . Contrairement aux autres types de convergence, la convergence en loi n'est pas une notion additive ni de type produit c'est à dire que si X_n et Y_n convergent en loi vers X et Y , on ne peut affirmer que $X_n + Y_n$ converge en loi vers $X + Y$ (sauf si Y est constante....) et de même on ne peut affirmer que le couple (X_n, Y_n) converge en loi vers le couple (X, Y) . Par contre si X_n converge en loi vers X alors pour toute fonction continue g de \mathbb{R}^d dans \mathbb{R}^q la suite $g(X_n)$ converge en loi vers $g(X)$. La convergence en loi est l'une des plus faibles comme le montre la proposition suivante:

Proposition 1.5.2. Si X_n converge en probabilité vers X alors X_n converge en loi vers X .

Indications de preuve:

On pose $B_n^\epsilon = \{|X_n - X| > \epsilon\}$ et soit φ une fonction continue à support compact (donc uniformément continue), $\epsilon > 0$. Il existe $\alpha > 0$ tel que $|x - y| \leq \alpha$ implique $|\varphi(x) - \varphi(y)| \leq \epsilon/2$. En utilisant la majoration $|\varphi| \leq K$ on obtient:

$$\begin{aligned} |\mathbb{E}(\varphi(X_n)) - \mathbb{E}(\varphi(X))| &\leq \mathbb{E}(|\varphi(X_n)) - \mathbb{E}(\varphi(X))| \\ &= \int_{|X - X_n| \leq \alpha} |\varphi(X) - \varphi(X_n)| d\mathbb{P} \\ &\quad + \int_{|X - X_n| > \alpha} |\varphi(X) - \varphi(X_n)| d\mathbb{P} \\ &\leq \epsilon/2 + 2K\mathbb{P}(B_n^\alpha) \end{aligned}$$

Il suffit alors d'utiliser $\mathbb{P}(B_n^\alpha) \rightarrow 0$. ■

On peut remarquer que l'un des seuls cas où la réciproque est exacte est lorsque la suite X_n converge en loi vers une constante.

1.5.1 T.C.L.

Nous allons commencer par rappeler le résultat classique:

Théorème 1.5.3. *Soit X_n une suite i.i.d. à valeurs dans \mathbb{R}^d et de carré sommable. On note m la moyenne des X et Σ leur matrice de covariance. Alors si on pose $S_n = \sum_{k=1}^n X_k$ on a:*

$$W_n = \frac{S_n - nm}{\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}_d(0, \Sigma)$$

où $\mathcal{N}_d(0, \Sigma)$ désigne la loi normale de moyenne nulle et de covariance Σ dans \mathbb{R}^d .

Indications de preuve:

Il suffit décrire que la transformée de Fourier de la loi des $X_n - m$ admet un développement limité à l'origine de la forme: $\hat{\mu}(t) = -t'\Sigma t/2 + ||t||^2\epsilon(t)$ et on utilise 1.4.4 ■

On peut donner différentes extensions de ce théorème sous des hypothèses plus faibles par exemple le théorème de Lindeberg énoncé ici pour $d = 1$.

Soit X_n une suite de variables aléatoires indépendantes à valeurs dans \mathbb{R} , centrées et de carré sommable. On pose $\sigma_k^2 = \mathbb{E}(X_k^2)$, $s_n^2 = \sum_{k=1}^n \sigma_k^2$ et $S_n = \sum_{k=1}^n X_k$.

Définition 1.5.4. *La condition suivante est appelée condition de Lindeberg:*

$$\lim_n \frac{1}{s_n^2} \sum_{k=1}^n \int_{|X_k| > \epsilon s_n} X_k^2 d\mathbb{P} \rightarrow 0$$

pour tout $\epsilon > 0$.

Théorème 1.5.5. *Si la condition de Lindeberg est satisfaite alors $\frac{S_n}{s_n} \xrightarrow{\mathcal{L}} \mathcal{N}_1(0, \sigma^2)$. Réciproquement si l'on a: $s_n \rightarrow \infty$, et $\frac{\sigma_n}{s_n} \rightarrow 0$ alors la convergence en loi ci dessus implique que la condition de Lindeberg est satisfaite.*

1.5.2 T.C.L. local

Il existe plusieurs variantes de l'estimation de la vitesse de convergence dans le T.C.L. appelées théorèmes locaux. Le premier résultat ne nécessite que l'utilisation du théorème de Lebesgue, les autres réclament une étude plus sophistiquée de la fonction caractéristique.

Proposition 1.5.6. *On utilise les notations et hypothèses du théorème 1.5.3 et l'on suppose de plus que la transformée de Fourier de la loi des X_n est sommable (ce qui implique que leur loi a une densité dans $\mathcal{C}_0(\mathbb{R}^d)$). Si on note f_n la densité de W_n alors f_n converge uniformément sur \mathbb{R}^d vers la densité de la loi $\mathcal{N}_d(0, \Sigma)$.*

Une estimation plus précise est donnée par la proposition suivante:

Proposition 1.5.7. (Berry-Essen) Soit X_n une suite i.i.d dans \mathbb{R} , centrée et telle que $\mathbb{E}(X_n^2) = \sigma^2$, $\mathbb{E}(|X_n|^3) = \rho < \infty$. Alors si F_n est la fonction de répartition de la variable aléatoire $\frac{S_n}{\sigma\sqrt{n}}$ et F celle de la loi normale centrée réduite, on a:

$$|F_n(x) - F(x)| \leq \frac{3\rho}{\sigma^3\sqrt{n}}$$

L'énoncé suivant s'apparente à un résultat de dérivation:

Proposition 1.5.8. Soit X_n une suite i.i.d dans \mathbb{R} , centrée avec $\mathbb{E}(X_n^2) = \sigma^2$ et soit η la densité de la loi normale centrée de variance σ^2 . On a alors essentiellement les deux cas suivants:

- Le cas du réseau: la loi des X_n est portée par un ensemble du type $\{b + a\mathbb{Z}\}$. On pose $p_n(x) = \mathbb{P}(S_n/\sqrt{n} = x)$ pour $x \in L_n = \{(nb + a\mathbb{Z})/\sqrt{n}\}$. Alors:

$$\sup_{x \in L_n} \left| \frac{\sqrt{n}}{a} p_n(x) - \eta(x) \right| \rightarrow 0$$

- Le cas "irréductible": le plus petit sous groupe fermé de \mathbb{R} portant la loi des X_n est égal à \mathbb{R} . Si $x_n/\sqrt{n} \rightarrow x$ alors pour $a < b$ on a:

$$\sqrt{n}\mathbb{P}(S_n \in (x_n + a, x_n + b)) \rightarrow (b - a)\eta(x)$$

On peut remarquer que l'irréductibilité est équivalente au fait que $\{t; \hat{\nu}(t) = 1\} = \{0\}$ si ν désigne la loi commune des X_n .

1.6 Grandes déviations

Soit $\mu \in \mathcal{M}_1^+(\mathbb{R}^d)$ possédant un moment d'ordre 1 et $\mathcal{L}\mu$ sa transformée de Laplace. On s'intéresse à la vitesse de convergence dans la loi forte des grands nombres c'est à dire au comportement asymptotique de la loi μ_n de S_n/n où $S_n = \sum_{k=1}^n X_k$, pour une suite X_k de variables aléatoires indépendantes et de même loi μ . Pour $x \in \mathbb{R}^d$ on définit la fonction log-laplace Δ :

$$\Delta(x) = \log\left(\int \exp(\langle t, x \rangle) d\mu(t)\right)$$

D'après 1.4.9 cette fonction est convexe, analytique sur $\overset{\circ}{\Omega}_\mu$ et dans ce domaine, ses dérivées successives sont obtenues par "dérivation sous le signe somme" dans la formule définissant la transformée de Laplace. En particulier si $0 \in \overset{\circ}{\Omega}_\mu$ alors μ possède des moments de tous ordres et S_n/n converge dans L^2 et donc en probabilité vers la constante $m = \mathbb{E}(X)$.

Définition 1.6.1. On définit la transformée de Legendre de $\Delta(x)$ soit $\Delta^*(x)$ par:

$$\Delta^*x = \sup_{y \in \mathbb{R}^d} (< x, y > - \Delta(y))$$

Proposition 1.6.2. La transformée de Legendre est une fonction positive, convexe, s.c.i.

Indications de preuve:

La positivité est évidente puisque $\Delta(0) = 0$. La convexité résulte de celle de Δ et c'est un sup de fonctions continues. ■

Définition 1.6.3. Soit $\mu \in \mathcal{M}_1^+(\mathbb{R}^d)$. On dit que μ satisfait au principe des grandes déviations si:

1. Pour tout fermé $F \subset \mathbb{R}^d$, $\overline{\lim}_n \frac{1}{n} \log \mu_n(F) \leq -\inf_{x \in F} \Delta^*(x)$, (principe de la borne supérieure)
2. Pour tout ouvert $O \subset \mathbb{R}^d$, $\underline{\lim}_n \frac{1}{n} \log \mu_n(O) \geq -\inf_{x \in O} \Delta^*(x)$, (principe de la borne inférieure)

Nous n'allons pas chercher à donner des conditions optimales assurant la validité de ce principe afin de ne pas alourdir l'exposé. Le lecteur intéressé pourra consulter [3] à ce sujet. Nous allons aussi distinguer le cas $d = 1$ du cas $d > 1$ qui nécessite (au moins pour la preuve présentée ici) des hypothèses plus fortes. Afin de clarifier la preuve des théorèmes de Cramér nous allons énoncer quelques résultats intermédiaires.

Lemme 1.6.4. On se place dans le cas d quelconque et $\Omega_\mu = \mathbb{R}^d$. Si $y = \overrightarrow{\nabla} \Delta(\eta)$ pour un certain $\eta \in \mathbb{R}^d$ alors pour tout $\delta > 0$ on a:

$$\underline{\lim}_n \frac{1}{n} \log \mu_n(B(y, \delta)) \geq -\Delta^*(y)$$

où $B(y, \delta)$ est la boule ouverte de centre y et de rayon δ .

Indications de preuve:

On considère la proba ν de densité $x \rightarrow \exp(< x, \eta > - \Delta(\eta))$ par rapport à μ . On constate facilement que ν a pour espérance y et que donc d'après la loi faible des grands nombres on a $\lim_n \nu_n(B(y, \delta)) = 1$. Par conséquent:

$$\begin{aligned} \frac{1}{n} \log \mu_n(B(y, \delta)) &= \Delta(\eta) - < y, \eta > + \frac{1}{n} \log \int_{B(y, \delta)} \exp(n < \eta, \eta - x >) d\nu(x) \\ &\geq \Delta(\eta) - < y, \eta > - |\eta| \delta + \frac{1}{n} \log \nu_n(B(y, \delta)) \end{aligned}$$

On en déduit

$$\underline{\lim}_n \frac{1}{n} \log \mu_n(B(y, \delta)) \geq -\Delta^*(y) - |\eta| \delta$$

D'autre part cette $\underline{\lim}$ est supérieure à $\underline{\lim}_{\delta' \rightarrow 0; 0 < \delta' < \delta} \underline{\lim}_n \frac{1}{n} \log \mu_n(B(y, \delta'))$ et donc à $-\Delta^*(y)$ ■

Ce lemme admet l'extension suivante:

Lemme 1.6.5. *Pour d quelconque et $\Omega_\mu = \mathbb{R}^d$ alors la conclusion du lemme 1.6.4 est valable pour tout $y \in \mathbb{R}^d$.*

Indications de preuve:

On se ramène au lemme précédent en ajoutant une "petite" perturbation gaussienne à la loi de X . Soit V une loi normale centrée réduite dans \mathbb{R}^d . Pour un entier positif M on note $\mu^{(M)}$ la loi de $X^{(M)} = X + V/\sqrt{M}$ et Δ_M le logarithme de sa transformée de Laplace alors $\Delta_M(x) = \Delta(x) + \|x\|^2/(2M)$. Si l'on pose $G_M(x) = \langle x, y \rangle - \Delta_M(x)$ il est facile de constater que la fonction concave et analytique G_M tend vers $-\infty$ à l'infini car $\Delta(x) \geq \langle x, m \rangle$ d'après Jensen. Il existe donc un point η où le gradient de G_M s'annule et donc $y = \overrightarrow{\nabla} \Delta_M(\eta)$. On en déduit donc d'après le lemme précédent que:

$$\lim_n \frac{1}{n} \log \mu_n^{(M)}(B(y, \delta)) \geq -\Delta_M^*(y) \geq -\Delta^*(y)$$

Il suffit bien sûr pour établir le lemme, de considérer le cas d'un point y tel que $\Delta^*(y) < \infty$. Or $S_n^{(M)}/n$ a la même loi que $S_n/n + V/\sqrt{Mn}$ et donc:

$$\mu_n(B(y, 2\delta)) \geq \mu_n^{(M)}(B(y, \delta)) - \mathbb{P}(\|V\| \geq \sqrt{Mn})$$

Or $\lim_n \frac{1}{n} \log(\mathbb{P}(\|V\| \geq \sqrt{Mn})) \leq -\frac{M\delta^2}{2}$. On obtient donc le résultat en passant à la limite sur n puis M . ■

Lemme 1.6.6. *Dans le cas $d = 1$, le lemme précédent est valable sans restrictions sur μ .*

Indications de preuve:

Il faut donc montrer que dans ce cas l'on peut se débarrasser de la restriction $\Omega_\mu = \mathbb{R}$. Il est facile de voir par translation qu'il suffit de se restreindre au cas $y = 0$. Pour cette preuve on va distinguer deux cas:

- Le cas trivial: la loi μ est portée par un demi-axe, par exemple $[0, \infty[$. Dans ce cas Δ est croissante et $-\Delta^*(0) = \log(\mu(\{0\}))$, de plus $\mu_n(B(\delta)) \geq (\mu(\{0\}))^n$ et la preuve est terminée.
- La loi μ vérifie $\mu(]0, \infty[) > 0$ et $\mu(]-\infty, 0]) > 0$. Soit alors M un entier positif et $\mu^{(M)}$ la loi de X conditionnée par $|X| \leq M$ et Δ_M la log-Laplace correspondante. On vérifie facilement que $\mu_n(B(\delta)) \geq \mu_n^{(M)}(B(\delta))(\mu([-M, M]))^n$ et que si l'on pose $H_M(x) = \log \int_{-M}^M \exp(xy) d\mu(y)$ alors $\Delta_M(x) = H_M(x) - \log(\mu([-M, M]))$. La loi $\mu^{(M)}$ étant à support compact on peut lui appliquer le lemme précédent et donc:

$$\lim_n \frac{1}{n} \log \mu_n(B(\delta)) \geq -\Delta_M^*(0) + \log(\mu([-M, M])) = \inf_x (H_M(x)) = I_M$$

La fonction $H_M(\cdot)$ est croissante en M , il en est donc de même de la suite I_M . Il reste donc à prouver qu'en posant $I^* = \lim \uparrow I_M$ alors $I^* \geq \inf_x \Delta(x)$. Puisque $I_M \leq H_M(0) \leq 0$ on a donc $I^* \leq 0$ et pour M assez grand alors $I_M > -\infty$, il en est donc de même pour

I^* et les ensembles $H_M \leq I^*$ sont donc compacts, non vides et décroissants. Il existe donc un point x_0 dans leur intersection et d'après le théorème de limite monotone on a $\Delta(x_0) = \lim H_M(x_0) \leq I^*$ ■

Corollaire 1.6.7. *En dimension $d = 1$ toute loi de probabilité μ satisfait au principe de la borne inférieure dans les grandes déviations. Il en est de même en dimension $d > 1$ si l'on impose $\Omega_\mu = \mathbb{R}^d$.*

Indications de preuve:

Soit O un ouvert. Pour tout $x \in O$ on peut trouver $\delta > 0$ avec $B(x, \delta) \subset O$ et par conséquent:

$$\underline{\lim}_n \frac{1}{n} \log \mu_n(O) \geq \underline{\lim}_n \frac{1}{n} \log \mu_n(B(x, \delta)) \geq -\Delta^*(x)$$

■

Afin d'étudier le principe de la borne supérieure nous utiliserons le résultat suivant:

Lemme 1.6.8. *Dans le cas $d = 1$, si l'on suppose μ intégrable de moyenne m on a:*

1. $\Delta^*(m) = 0$.
2. Si $x \geq m$ alors $\Delta^*(x) = \sup_{y \geq 0} (xy - \Delta(y))$ et Δ^* est croissante.
3. Si $x \leq m$ alors $\Delta^*(x) = \inf_{y \leq 0} (xy - \Delta(y))$ et Δ^* est décroissante.

Indications de preuve:

D'après Jensen on a $\Delta(x) \geq mx$ et donc $\Delta^*(0) = 0$ puisque Δ^* est positive. De plus si $x \geq m$ et $y < 0$ alors $yx - \Delta(y) \leq ym - \Delta(y) \leq \Delta^*(m) = 0$ et donc il suffit de prendre le \sup sur les $y \geq 0$. De plus pour $y \geq 0$ la fonction $xy - \Delta(y)$ est croissante en x . ■

Théorème 1.6.9. *(Théorème de Cramér I) En dimension $d = 1$ toute probabilité μ satisfait au principe des grandes déviations.*

Indications de preuve:

Il ne reste qu'à prouver le principe de la borne supérieure. Pour ce faire nous supposerons d'abord que m existe. En utilisant l'inégalité de Bienaymé-Tchebychev on obtient immédiatement que pour tout $y \geq 0$:

$$\mu_n([x, \infty[) \leq \mathbb{E}(\exp(ny(S_n - x))) = \exp(-n(xy - \Delta(y)))$$

et en appliquant le lemme 1.6.8 pour $x > m$ on a $\mu_n([x, \infty[) \leq \exp(-n\Delta^*(x))$ et de même $\mu_n(]-\infty, x]) \leq \exp(-n\Delta^*(x))$ pour $x < m$. On peut bien sûr supposer que pour le fermé F en question on a $I = \inf_{x \in F} \Delta^*(x) > 0$. Puisque $\Delta^*(m) = 0$ il s'en suit que m appartient au complémentaire de F . On pose $]u, v[= \cup]a, b[$ sur les intervalles $]a, b[$ tels que $a < m < b$ et $]a, b[\cap F = \emptyset$. Alors $u < v$ et l'un au moins de u ou v est fini puisque F est non vide. Si u est fini alors $u \in F$ et donc $\Delta^*(u) \geq I$ de même si v est fini alors $\Delta^*(v) \geq I$. En appliquant alors l'inclusion $F \subset]-\infty, u] \cup [v, +\infty[$ (où l'un des intervalles est absent si ses deux bornes sont identiques) on obtient le résultat avec les majorations

ci dessus.

Si l'on ne fait aucune hypothèse sur Ω_μ la preuve est sensiblement la même en constant que si $\Omega_\mu = \{0\}$ alors Δ^* est nulle et il n'y a rien à prouver. Sinon m existe comme un élément de $\overline{\mathbb{R}}$ (par exemple si $\mathcal{L}\mu(x) < \infty$ pour un point $x > 0$ alors $\int_0^\infty y \, d\mu(y) \leq \mathcal{L}\mu(x)/x < \infty$) et l'on peut recopier la preuve ci-dessus en étudiant séparément les cas triviaux où m est infinie. ■

Théorème 1.6.10. (*Théorème de Cramér II*) Toute probabilité $\mu \in \mathcal{M}_1^+(\mathbb{R}^d)$ telle que $\Omega_\mu = \mathbb{R}^d$ satisfait au principe des grandes déviations.

Indications de preuve:

Il ne reste qu'à prouver le principe de la borne supérieure. Ceci va résulter d'une série de lemmes techniques.

Lemme 1.6.11. Soit $\mu \in \mathcal{M}_1^+(\mathbb{R}^d)$ et $\delta > 0$. On pose $\Delta_\delta^*(x) = \inf(\Delta^*(x) - \delta, 1/\delta)$. Alors le principe de la borne supérieure pour un fermé Δ est équivalent à:

$$\forall \delta > 0, \quad \overline{\lim}_n \frac{1}{n} \log \mu_n(\Delta) \leq \delta - \inf_{x \in \Delta} \Delta_\delta^*(x)$$

Indications de preuve:

Il suffit de remarquer que $\lim_{\delta \rightarrow 0} \inf_{x \in \Delta} \Delta_\delta^*(x) = \inf_{x \in \Delta} \Delta^*(x)$ ■

Lemme 1.6.12. Soit $\mu \in \mathcal{M}_1^+(\mathbb{R}^d)$. Tout compact de \mathbb{R}^d vérifie le principe de la borne supérieure.

Indications de preuve:

Soit $\delta > 0$ et K compact. Alors pour tout $\eta \in K$ on peut trouver $y_\eta \in \mathbb{R}^d$ tel que $\langle y_\eta, \eta \rangle > -\Delta(y_\eta) \geq \Delta_\delta^*(\eta)$ et on choisit $\rho_\eta > 0$ avec $\rho_\eta \|y_\eta\| \leq \delta$. Alors pour tout borélien B de \mathbb{R}^d et tout y on va avoir:

$$\mu_n(B) \leq \mathbb{E} \left\{ \exp[n \langle y, S_n \rangle - \inf_{x \in G} (\langle y, x \rangle)] \right\}$$

et en particulier pour $y = ny_\eta$:

$$\mu_n(B(\eta, \rho_\eta)) \leq \exp n[\Delta(y_\eta) - \inf_{x \in B(\eta, \rho_\eta)} (\langle y_\eta, x \rangle)]$$

et donc:

$$\begin{aligned} \frac{1}{n} \log \mu_n(B(\eta, \rho_\eta)) &\leq \Delta(y_\eta) - \inf_{x \in B(\eta, \rho_\eta)} (\langle y_\eta, x \rangle) \\ &\leq \Delta(y_\eta) + \rho_\eta \|y_\eta\| - \langle y_\eta, \eta \rangle \\ &\leq \Delta(y_\eta) + \delta - \langle y_\eta, \eta \rangle \end{aligned}$$

K étant compact on peut extraire un recouvrement par un nombre fini N de boules de type $B(\eta_i, \rho_{\eta_i})$ et donc

$$\begin{aligned} \frac{1}{n} \log \mu_n(K) &\leq \log(N)/n + \delta - \inf_i (\Delta(y_{\eta_i}) - \langle y_{\eta_i}, \eta_i \rangle) \\ \overline{\lim}_n \frac{1}{n} \log \mu_n(K) &\leq \delta - \inf_i \Delta_\delta^*(\eta_i) \leq \delta - \inf_{x \in K} \Delta_\delta^*(x) \end{aligned}$$

■

Lemme 1.6.13. Soit $\mu \in \mathcal{M}_1^+(\mathbb{R}^d)$ telle que $0 \in \overset{\circ}{\Omega}_\mu$. Alors la suite μ_n est exponentiellement tendue.

Indications de preuve:

Soit μ^i la loi de la i ème composante de X et Δ_i la log-Laplace associée. Si l'on considère le compact $K_\rho = [-\rho, \rho]^d$ on a:

$$\mu_n(\mathbb{R}^d \setminus K_\rho) \leq \sum_{i=1}^d (\mu_n^i([\rho, \infty[) + \mu_n^i(]-\infty, -\rho]))$$

Si l'on reprend la preuve de 1.6.9 on obtient pour $\rho \geq \|m\|$

$$\mu_n^i([\rho, \infty[) \leq \exp(-n\Delta_i^*(\rho)), \quad \text{et} \quad \mu_n^i(]-\infty, -\rho]) \leq \exp(-n\Delta_i^*(-\rho))$$

Or si Ω_μ contient un ouvert $] -\epsilon, \epsilon[$ alors $\Delta^*(x) \rightarrow \infty$ quand $|x| \rightarrow \infty$ car pour tout y , $\Delta^*(x)/|x| \geq y \text{sign}(x) - \Delta(y)/|x|$ et donc $\underline{\lim}_{|x| \rightarrow \infty} \Delta^*(x)/|x| \geq \epsilon$ ■

Corollaire 1.6.14. Soit $\mu \in \mathcal{M}_1^+(\mathbb{R}^d)$ telle que $0 \in \overset{\circ}{\Omega}_\mu$. Alors μ vérifie le principe de la borne supérieure.

Indications de preuve:

Il suffit bien sûr pour établir le lemme, de considérer le cas d'un fermé F tel que $\epsilon = \inf_{y \in F} \Delta^*(y) > 0$. D'après 1.6.13 il existe un compact K_ϵ tel que $\mu_n(\mathbb{R}^d \setminus K_\epsilon) \leq \exp(-n\epsilon)$. En écrivant:

$$\mu_n(F) \leq \mu_n(F \cap K_\epsilon) + \mu_n(\mathbb{R}^d \setminus K_\epsilon)$$

$F \cap K_\epsilon$ est compact on déduit donc de 1.6.12:

$$\begin{aligned} \overline{\lim}_n \frac{1}{n} \log \mu_n(F) &\leq \max\{\overline{\lim}_n \frac{1}{n} \log \mu_n(F \cap K_\epsilon), -\epsilon\} \\ &\leq \max\{-\inf_{x \in F \cap K_\epsilon} \Delta^*(x), -\epsilon\} = -\epsilon \end{aligned}$$

■

En fait la théorie des grandes déviations est essentiellement utilisée pour obtenir un équivalent de $\mu_n(B)$ pour un borélien B qui sera le plus souvent une boule du genre $B(m, \epsilon)$.

Théorème 1.6.15. Soit $\mu \in \mathcal{M}_1^+(\mathbb{R}^d)$ satisfaisant au principe de grandes déviations et B un borélien de \mathbb{R}^d tel que $I(B) = \inf_{x \in B} \overset{\circ}{\Lambda}^*(x) = \inf_{x \in \overline{B}} \Lambda^*(x)$. Alors:

$$\lim_n \frac{1}{n} \log \mu_n(B) = -I(B)$$

Chapitre 2

Chaînes de Markov

On considère un espace filtré $(\Omega, \mathcal{F}_n, \mathbb{P})$ avec $\mathcal{F}_\infty = \bigvee \mathcal{F}_n$ et un espace d'états (E, \mathcal{E}) qui sera le plus souvent métrisable et muni de sa tribu borélienne.

Définition 2.0.16. *On note:*

- $\mathcal{B}^+(E)$ l'ensemble des fonctions mesurables positives sur (E, \mathcal{E}) (pouvant prendre la valeur $+\infty$)
- $\mathcal{B}_b(E)$ l'ensemble des fonctions mesurables bornées sur (E, \mathcal{E})
- $\mathcal{M}^+(E)$ (resp $\mathcal{M}_b(E)$) l'ensemble des mesures positives (resp. bornées) sur (E, \mathcal{E})
- $\mathcal{M}_1^+(E)$ l'ensemble des probabilités sur (E, \mathcal{E}) .

2.1 Probabilités de transition

Soit (E, \mathcal{E}) , (E', \mathcal{E}') deux espaces mesurables.

Définition 2.1.1. *On dit qu'une application \mathbf{Q} de $(E \times \mathcal{E}')$ dans $[0, 1]$ est une probabilité de transition de (E, \mathcal{E}) vers (E', \mathcal{E}') si:*

- Pour tout $x \in E$, $\mathbf{Q}(x, \bullet) \in \mathcal{M}_1^+(E')$.
- Pour tout $B' \in \mathcal{E}'$, $\mathbf{Q}(\bullet, B') \in \mathcal{B}^+(E)$.

Exemple 2.1.2. *Exemples de probabilités de transition:*

1. Si E est dénombrable, on considère une matrice $\mathbf{Q}(i, j)$, avec $(i, j) \in E \times E$ et dont chaque ligne est une répartition de probabilité sur E .
2. Si E est un groupe localement compact métrisable (par exemple \mathbb{R}^d , \mathbb{Z}^d ou un groupe de matrices) et $\mu \in \mathcal{M}_1^+(E)$ on peut définir les deux probabilités de transition $\mathbf{Q}_d(x, \bullet) = \delta_x \star \mu$ et $\mathbf{Q}_g(x, \bullet) = \mu \star \delta_x$ qui correspondront aux marches à "droite" et à "gauche" sur E .

Pour $f \in \mathcal{B}^+(E')$ (resp. $f \in \mathcal{B}_b(E')$) et $\mu \in \mathcal{M}^+(E)$ (resp. $\mu \in \mathcal{M}_b(E)$) on pose:

$$\mathbf{Q}f(x) = \int f(x')\mathbf{Q}(x, dx'), \quad \mu\mathbf{Q}(B') = \int \mathbf{Q}(x, B')d\mu(x)$$

Alors $\mathbf{Q}f \in \mathcal{B}^+(E)$ (resp. $\mathbf{Q}f \in \mathcal{B}_b(E)$), $\mu\mathbf{Q} \in \mathcal{M}^+(E')$ (resp. $\mu\mathbf{Q} \in \mathcal{M}_b(E')$) et l'on a $\mu(\mathbf{Q}f) = (\mu\mathbf{Q})(f)$. Dans le cas d'espaces d'états dénombrables E et E' , on voit que cette définition est compatible avec les produits usuels de matrices à condition d'écrire une mesure comme un vecteur "ligne" et une fonction comme un vecteur "colonne". On peut aussi composer les probabilités de transition : si \mathbf{Q}_1 est une transition de (E, \mathcal{E}) vers (E', \mathcal{E}') et \mathbf{Q}_2 une transition de (E', \mathcal{E}') vers (E'', \mathcal{E}'') alors le produit $\mathbf{Q}_1\mathbf{Q}_2$ défini par

$$\mathbf{Q}_1\mathbf{Q}_2(x, B'') = \int \mathbf{Q}_1(x, dy)\mathbf{Q}_2(y, B'')$$

est une transition de probabilité de (E, \mathcal{E}) vers (E'', \mathcal{E}'') . On constate de nouveau que cette définition coïncide avec la notion de produits de matrices.

Dans la suite, pour $n \geq 0$, on considère l'espace mesurable $\mathfrak{E}_n = (E^{n+1}, \bigotimes_0^n \mathcal{E})$ et l'on se donne une probabilité de transition \mathbf{Q} de (E, \mathcal{E}) dans lui-même. Si f_0, f_1, \dots, f_n , sont des fonctions mesurables sur (E, \mathcal{E}) on définit la fonction mesurable $f_0 \otimes f_1 \dots \otimes f_n$ sur \mathfrak{E}_n par $(f_0 \otimes f_1 \dots \otimes f_n)(x_0, x_1, \dots, x_n) = f_0(x_0)f_1(x_1) \dots f_n(x_n)$ et l'on désigne par $\mathbf{1}$ la fonction qui vaut 1 pour tout x .

Définition 2.1.3. Pour toute probabilité μ sur (E, \mathcal{E}) on peut définir par récurrence la suite de probabilité $\mathbb{P}_\mu^{(n)}$ sur \mathfrak{E}_n par les formules:

- $\mathbb{P}_\mu^{(0)} = \mu$
- $\mathbb{P}_\mu^{(n+1)}(f_0 \otimes \dots \otimes f_{n+1}) = \mathbb{P}_\mu^{(n)}(f_0 \otimes \dots \otimes f_{n-1} \otimes (f_n\mathbf{Q}f_{n+1}))$

et l'on désigne par $\mathbb{P}_x^{(n)}$ la probabilité $\mathbb{P}_\mu^{(n)}$ lorsque μ est la masse de Dirac au point x .

Il est alors facile de prouver:

Proposition 2.1.4. On a les propriétés suivantes:

1. $\mathbb{P}_x^{(n)}$ est une probabilité de transition de \mathfrak{E}_0 vers \mathfrak{E}_n .
2. $\mathbb{P}_\mu^{(n)} = \int \mathbb{P}_x^{(n)} d\mu(x)$
3. Pour $m < n$ la projection de $\mathbb{P}_\mu^{(n)}$ sur \mathfrak{E}_m est égale à $\mathbb{P}_\mu^{(m)}$
4. $\mathbb{P}_\mu^{(n)}(\mathbf{1} \otimes \dots \otimes \mathbf{1} \otimes f) = \mu\mathbf{Q}^n f$
5. $\mathbb{P}_\mu^{(n+m)}(\mathbf{1} \otimes \dots \otimes \mathbf{1} \otimes f_n \otimes \dots \otimes f_{n+m}) = \mathbb{P}_\nu^{(m)}(f_n \otimes \dots \otimes f_{n+m})$ où ν désigne la probabilité $\mu\mathbf{Q}^n$

La propriété 3 de la proposition précédente est très importante car il est alors possible de montrer que pour μ et \mathbf{Q} donnés, il existe une unique probabilité \mathbb{P}_μ sur $(\Omega, \mathcal{F}) = (\prod_0^\infty E, \bigotimes_0^\infty \mathcal{E})$ dont les projections sur chaque \mathfrak{E}_n sont égales à $\mathbb{P}_\mu^{(n)}$ (théorème de Ionescu-Tulcea). Ce résultat est démontré dans [6] page 151.

2.2 Suites markoviennes

2.2.1 Définition et exemples

Soit X_n , $n \geq 0$ une suite de variables aléatoires à valeurs dans (E, \mathcal{E}) , \mathbf{Q} une probabilité de transition sur (E, \mathcal{E}) et μ une probabilité sur (E, \mathcal{E}) .

Définition 2.2.1. . On dit que X_n est une chaîne de Markov de transition \mathbf{Q} et de loi initiale μ si:

1. Pour tout $n \geq 0$, X_n est \mathcal{F}_n mesurable.
2. La variable aléatoire X_0 a pour loi μ .
3. Pour toute fonction $f \in \mathcal{B}_b(E)$ ou bien pour toute fonction $f \in \mathcal{B}^+(E)$ on a:

$$\mathbb{E}(f(X_{n+1})|\mathcal{F}_n) = \mathbf{Q}f(X_n)$$

cette dernière relation est appelée "propriété de Markov".

On peut tout de suite faire quelques remarques sur cette définition:

Remarques

- Si la suite X_n est une chaîne de Markov pour la filtration $\{\mathcal{F}_n\}$ elle l'est aussi pour toute filtration contenue dans celle-ci et rendant le processus X_n adapté. En particulier on peut choisir les tribus "naturelles" du processus X_k engendrées par les variables aléatoires X_k pour $k \leq n$. C'est ce qui est fait implicitement si aucune filtration n'est précisée.
- La définition implique que la loi conditionnelle de X_{n+1} sachant \mathcal{F}_n est identique à la loi conditionnelle de X_{n+1} sachant X_n et que celle-ci possède une version régulière donnée par le noyau \mathbf{Q} . C'est ce qui justifie la phrase courante: une chaîne de Markov est un processus dont l'avenir est uniquement conditionné par l'état présent et non par tout le passé de la trajectoire.
- La propriété de Markov implique que pour $p \geq 1$:

$$\mathbb{E}(f(X_{n+p})|\mathcal{F}_n) = \mathbf{Q}^p f(X_n), \quad \mathbb{E}(f(X_p)) = \mu \mathbf{Q}^p(f)$$

et donc la loi de X_p est donnée par $\mu \mathbf{Q}^p$.

- Dans le cas où les tribus \mathcal{F}_n sont les tribus naturelles d'un processus Y_k à valeurs dans un espace mesurable (E', \mathcal{E}') , on peut vérifier la propriété de Markov en testant que:

$$\mathbb{E}(f(X_{n+1}) \prod_{k=1}^n h_k(Y_k)) = \mathbb{E}(\mathbf{Q}f(X_n) \prod_{k=0}^n h_k(Y_k))$$

pour tout $n \geq 1$ et tout $f \in \mathcal{B}^+(E)$ et tout système h_0, h_1, \dots, h_n de fonctions dans $\mathcal{B}^+(E')$ (ou bien pour tout $f \in \mathcal{B}_b(E)$ et tout système h_0, h_1, \dots, h_n de fonctions dans $\mathcal{B}_b(E')$)

L'exemple le plus courant de chaîne de Markov (et pratiquement toute chaîne peut se ramener à cette situation) est le suivant:

Proposition 2.2.2. Soit Y_n , $n \geq 1$, une suite i.i.d. de loi ν à valeurs dans (E', \mathcal{E}') , f une fonction mesurable de $(E \times E')$ dans E et X_0 une variable aléatoire de loi μ , à valeurs dans E , indépendante de la suite Y_n . Alors la suite X_n définie par récurrence par:

$$X_{n+1} = f(X_n, Y_{n+1}), \quad n \geq 0$$

est une chaîne de Markov de loi initiale μ et de transition définie par

$$\mathbf{Q}(x, A) = \nu\{x' ; f(x, x') \in A\}$$

Indications de preuve:

On voit immédiatement que Y_{n+1} est indépendante de $\mathcal{F}_n = \sigma(X_0, Y_1, \dots, Y_n)$. La preuve est alors une conséquence directe de 1.1.3 dès que l'on a vérifié que \mathbf{Q} est bien une probabilité de transition. ■

Exemples

• Marches aléatoires sur un groupe

Soit E un groupe L.C.D. Les suites $X_n = X_0 Y_1 \dots Y_n$ et $X_n = Y_n \dots Y_1 X_0$ sont des chaînes de Markov de transition $\mathbf{Q}f(x) = \int f(xy) d\nu(y) = f \star \check{\nu}(x)$ et $\mathbf{Q}f(x) = \int f(yx) d\nu(y) = \check{\nu} \star f(x)$ appelées marches aléatoires à droite et à gauche sur E .

• Marches aléatoires sur un espace homogène

Soit G un groupe L.C.D. d'élément neutre e , opérant à droite sur un espace métrique E par $g \rightarrow g \cdot x$. On rappelle que ceci signifie que $(g, x) \rightarrow g \cdot x$ est continue de $G \times E$ dans E , que $e \cdot x = x$ et que $g_1 \cdot (g_2 \cdot x) = (g_1 g_2) \cdot x$. Si Y_n est une suite i.i.d. à valeurs dans G on pose $X_n = Y_n \dots Y_1 \cdot X_0$ dont la fonction de transition est $\mathbf{Q}f(x) = \int f(g \cdot x) d\nu(g)$. On a alors $\mathbf{Q}^n f(x) = \int f(g \cdot x) d\nu^{\star n}(g)$. On remarquera, qu'en général la suite $X_n = Y_1 \dots Y_n \cdot X_0$ n'est pas une chaîne de Markov. Dans la plupart des cas, G est un groupe de matrices opérant sur \mathbb{R}^d ou bien sur un espace projectif de \mathbb{R}^d . Par exemple $G = SL(2, \mathbb{R})$ opère sur l'espace projectif $P(\mathbb{R}^2)$ identifié à $\mathbb{R} \cup \{\infty\}$ par:

$$g \cdot x = \frac{ax + b}{cx + d}, \quad \text{pour } g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

On considère l'équation de récurrence aléatoire (équation de Schrödinger) suivante:

$$u_{n+1} + u_{n-1} = Z_n u_n$$

où Z_n , $n \in \mathbb{Z}$ est une suite de variables aléatoires indépendantes et de même loi θ sur \mathbb{R} . On constate que toute solution s'écrit pour $n \geq 0$:

$$\begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix} = Y_n \dots Y_1 \begin{pmatrix} u_1 \\ u_0 \end{pmatrix}, \quad \text{où } Y_n = \begin{pmatrix} Z_n & -1 \\ 1 & 0 \end{pmatrix}$$

et la suite $X_n = \frac{u_{n+1}}{u_n}$ forme donc une chaîne de Markov sur $P(\mathbb{R}^2)$, partant de $X_0 = \begin{pmatrix} u_1 \\ u_0 \end{pmatrix}$ et de transition $\mathbf{Q}f(x) = \int f(z - 1/x) d\theta(z)$.

• **Files d'attente**

On suppose que Y_n , $n \geq 1$, représente le nombre de clients arrivants dans une file d'attente à l'instant $n - 1/2$ et que si à l'instant n la file est non vide, un client quitte celle-ci. Alors le nombre de clients dans la file à l'instant n , soit X_n , vérifie:

$$X_{n+1} = (X_n - 1)^+ + Y_{n+1}$$

2.2.2 Formes équivalentes de la Propriété de Markov

On reprend ici les notations de la première section, en particulier on considère l'espace mesurable $\mathfrak{E}_n = (E^{n+1}, \bigotimes_0^n \mathcal{E})$ muni de la probabilité $\mathbb{P}_\mu^{(n)}$.

Proposition 2.2.3. *La suite X_n est une chaîne de Markov de transition \mathbf{Q} et de loi initiale μ si et seulement si, pour tout $n \geq 0$, la loi du vecteur (X_0, X_1, \dots, X_n) sur \mathfrak{E}_n est égale à $\mathbb{P}_\mu^{(n)}$.*

Indications de preuve:

Si la loi conjointe des X_i est donnée par $\mathbb{P}_\mu^{(n)}$ alors pour tout $n \geq 1$, tout $f \in \mathcal{B}^+(E)$ et tout système h_0, h_1, \dots, h_n de fonctions dans $\mathcal{B}^+(E)$ on a:

$$\begin{aligned} \mathbb{E}(f(X_{n+1}) \prod_{k=0}^n h_k(X_k)) &= \mathbb{P}_\mu^{(n+1)}(h_0 \otimes h_1 \dots \otimes h_n \otimes f) \\ &= \mathbb{P}_\mu^{(n)}(h_0 \otimes h_1 \dots \otimes h_{n-1} \otimes (h_n \mathbf{Q}(f))) \\ &= \mathbb{E}(\mathbf{Q}f(X_n) \prod_{k=0}^n h_k(X_k)) \end{aligned}$$

d'où la propriété de Markov.

Réciproquement on procède par récurrence sur n , puisque pour $n = 0$ on a $\mathbb{P}_\mu^{(0)} = \mu$. Si la propriété de l'énoncé est vraie à l'ordre n alors pour $f \in \mathcal{B}^+(E)$ et tout système h_0, h_1, \dots, h_n de fonctions dans $\mathcal{B}^+(E)$ on a:

$$\begin{aligned} \mathbb{E}(f(X_{n+1}) \prod_{k=0}^n h_k(X_k)) &= \mathbb{E}\left(\mathbb{E}(f(X_{n+1}) | \mathcal{F}_n) \prod_{k=0}^n h_k(X_k)\right) \\ &= \mathbb{E}(\mathbf{Q}f \prod_{k=0}^n h_k(X_k)) \\ &= \mathbb{P}_\mu^{(n)}(h_0 \otimes h_1 \dots \otimes h_n \otimes (\mathbf{Q}f)) \\ &= \mathbb{P}_\mu^{(n+1)}(h_0 \otimes h_1 \dots \otimes h_n \otimes f) \end{aligned}$$

■

Cette proposition implique que μ et \mathbf{Q} étant donnés, on connaît les lois conjointes du processus et donc la loi du processus. En particulier si E est dénombrable on a :

$$\mathbb{P}(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) = \mu(x_0) \mathbf{Q}(x_0, x_1) \dots \mathbf{Q}(x_{n-1}, x_n)$$

On peut donner la généralisation suivante de la propriété de Markov :

Proposition 2.2.4. *Soit ψ une fonction mesurable positive sur $(\prod_0^\infty E, \otimes_0^\infty \mathcal{E})$. On pose $\Psi_n = \psi(X_n, X_{n+1}, \dots)$ alors :*

$$\mathbb{E}(\Psi_n | \mathcal{F}_n) = \mathbb{E}(\Psi_n | X_n)$$

Indications de preuve :

Ceci revient à prouver que le premier membre de l'égalité ne dépend que de X_n . Il suffit de vérifier cette propriété pour $\psi = h_0 \otimes \dots \otimes h_p$ avec les $h_k \in \mathcal{B}^+(E)$. Pour $p = 0$ la propriété est évidente et pour $p = 1$ c'est la propriété de Markov. On raisonne donc par récurrence sur p en supposant que cette formule est vraie à l'ordre p .

$$\begin{aligned} \mathbb{E}\left(\prod_{k=0}^{p+1} h_k(X_{n+k}) | \mathcal{F}_n\right) &= \mathbb{E}\left(\mathbb{E}\left(\prod_{k=0}^{p+1} h_k(X_{n+k}) | \mathcal{F}_{n+p}\right) | \mathcal{F}_n\right) \\ &= \mathbb{E}(\mathbf{Q}f_{p+1}(X_{n+p}) \prod_{k=0}^p h_k(X_{n+k}) | \mathcal{F}_n) \\ &= \mathbb{E}(\mathbf{Q}f_{p+1}(X_{n+p}) \prod_{k=0}^p h_k(X_{n+k}) | X_n) \end{aligned}$$

■

On peut légèrement généraliser cet énoncé :

Proposition 2.2.5. *Avec les notations de la proposition précédente, si Φ_n est une fonction \mathcal{F}_n mesurable positive, alors :*

$$\mathbb{E}(\Psi_n \Phi_n | X_n) = \mathbb{E}(\Psi_n | X_n) \mathbb{E}(\Phi_n | X_n)$$

Autrement dit, le passé et l'avenir sont conditionnellement indépendants connaissant le présent.

Indications de preuve :

Il suffit, dans le premier membre, de conditionner par \mathcal{F}_n avant X_n . ■

On peut aussi donner une version de la propriété de Markov forte dont la forme sera beaucoup plus lisible avec le formalisme des chaînes canoniques. Soit T un temps d'arrêt (pour la filtration \mathcal{F}_n), on définit le nouvel espace $\widetilde{\Omega} = \Omega \cap (T < \infty)$ muni de la probabilité trace $\widetilde{\mathbb{P}}(A) = \frac{\mathbb{P}(A)}{\mathbb{P}(T < \infty)}$ et de la filtration $\widetilde{\mathcal{F}}_n = \{A \cap (T < \infty) ; A \in \mathcal{F}_{n+T}\}$.

Proposition 2.2.6. (Markov fort) La famille de variables aléatoires $\tilde{X}_n = X_{n+T}$ définie sur $\tilde{\Omega}$ est $(\tilde{\mathcal{F}}_n)$ adaptée et on a :

$$\tilde{\mathbb{E}}(f(\tilde{X}_{n+1})|\tilde{\mathcal{F}}_n) = \mathbf{Q}f(\tilde{X}_n)$$

pour toute fonction f dans $\mathcal{B}^+(E)$. Soit ψ une fonction mesurable positive sur $(\prod_0^\infty E, \otimes_0^\infty \mathcal{E})$. On pose $\Psi_n = \psi(\tilde{X}_n, \tilde{X}_{n+1}, \dots)$ alors :

$$\mathbb{E}(\Psi_n|\tilde{\mathcal{F}}_n) = \mathbb{E}(\Psi_n|\tilde{X}_n)$$

Indications de preuve :

il suffit de découper suivant des ensembles $A \cap (T = k)$ pour $A \in \tilde{\mathcal{F}}_n$ ■

2.3 Chaînes canoniques

2.3.1 Définition et propriétés

Soit μ une probabilité sur E et \mathbf{Q} une probabilité de transition sur E .

Théorème 2.3.1. (Ionescu-Tulcea) Soit \mathbf{Q} une probabilité de transition sur E . Alors pour toute probabilité μ sur E il existe une unique probabilité \mathbb{P}_μ sur l'espace probabilisé $(\Omega, \mathcal{F}) = (\prod_0^\infty E, \otimes_0^\infty \mathcal{E})$ dont les projections sur chaque \mathfrak{E}_n sont égales à $\mathbb{P}_\mu^{(n)}$. De plus si Ψ est mesurable positive ou bornée sur (Ω, \mathcal{F}) alors l'application $x \mapsto \mathbb{P}_x(\Psi)$ est mesurable et $\mathbb{P}_\mu(\Psi) = \int \mathbb{P}_x(\Psi) d\mu(x)$.

Indications de preuve :

La propriété 4. de 2.1.4 implique que l'on peut définir une fonction d'ensemble additive \mathbb{P}_μ sur Ω muni de l'algèbre de Boole des ensembles "cylindriques" c'est à dire mesurables par rapport à la tribu engendrée par les projections d'indice $\leq n$. Le problème est de montrer que cette fonction est en fait σ additive. Comme on l'a déjà indiqué, on peut trouver une preuve complète de ce théorème dans [6] page 151. ■

Une simple application de 2.2.3 fournit le résultat suivant :

Corollaire 2.3.2. Soit μ une probabilité sur E et \mathbf{Q} une probabilité de transition sur E . Alors sur l'espace (Ω, \mathcal{F}) muni de la probabilité \mathbb{P}_μ , la suite des projections X_n définit une chaîne de Markov (pour la filtration naturelle \mathcal{F}_n) de loi initiale μ et de transition \mathbf{Q} . Pour \mathbf{Q} fixé, cette famille de chaînes de Markov est (improprement) appelée "chaîne canonique" et notée $(\Omega, \mathcal{F}_n, X_n, \mathbb{P}_x)$.

On note θ l'opérateur de décalage sur Ω défini par $\theta(x_0, x_1, x_2, \dots) = (x_1, x_2, \dots)$ et θ_n son n ème itéré. On peut alors donner une nouvelle version de la propriété de Markov :

Proposition 2.3.3. (Propriété de Markov) Soit Ψ une fonction mesurable positive ou bornée sur (Ω, \mathcal{F}) . Alors :

$$\mathbb{E}_\mu(\Psi \circ \theta_n | \mathcal{F}_n) = \mathbb{E}_{X_n}(\Psi)$$

Indications de preuve:

Il suffit de tester cette relation sur des fonctions Ψ de la forme $\Psi = \prod_{k=0}^p h_k(X_k)$ avec par exemple les $h_k \in \mathcal{B}^+(E)$. Pour $p = 0$ on a $\mathbb{E}_\mu(f_0(X_n)|\mathcal{F}_n) = f_0(X_n) = \mathbb{E}_{X_n}(f_0(X_0))$ car on a $X_0 = x$, \mathbb{P}_x presque sûrement. On fait ensuite une preuve par récurrence sur p tout à fait analogue à 2.2.4 ■

Cette proposition écrite au temps n peut en fait être écrite à un temps aléatoire T qui est un temps d'arrêt et porte alors le nom de propriété de Markov forte. On commence par énoncer un lemme qui assure que les formules écrites ont un sens.

Lemme 2.3.4. *Soit T un temps d'arrêt et Ψ une variable aléatoire sur (Ω, \mathcal{F}) . Alors $\mathbf{1}_{\{T < \infty\}}(\Psi \circ \theta_T)$ est \mathcal{F} mesurable. Si U_n est un processus \mathcal{F}_n adapté alors $\mathbf{1}_{\{T < \infty\}}U_T$ est \mathcal{F}_T mesurable.*

Proposition 2.3.5. *(Propriété de Markov forte) Soit T un temps d'arrêt et Ψ une variable aléatoire positive ou bornée sur (Ω, \mathcal{F}) . Alors:*

$$\mathbb{E}_\mu\left(\mathbf{1}_{\{T < \infty\}}(\Psi \circ \theta_T)|\mathcal{F}_T\right) = \mathbf{1}_{\{T < \infty\}}\mathbb{E}_{X_T}(\Psi)$$

Indications de preuve:

Il suffit de découper $A \in \mathcal{F}_T$ suivant les valeurs $T = n$ et d'appliquer 2.3.3. ■

Il y a donc deux points de vue sur la notion de chaîne de Markov: celui de la section précédente où la suite des variables aléatoires X_n est donnée et celui de cette section où seule la “dynamique” c'est à dire la transition \mathbf{Q} est donnée et l'on construit une chaîne admettant cette dynamique. Selon les cas, l'on choisit l'un ou l'autre de ces points de vue. En particulier on peut définir une marche aléatoire à droite par la donnée d'une suite Y_n de variables aléatoires indépendantes et de même loi ν à valeurs dans un groupe E ou bien considérer la chaîne canonique construite à partir du noyau $\mathbf{Q}f(x) = \int f(xy) d\nu(y)$. Fort heureusement, la proposition suivante montre que l'on construit le même objet:

Proposition 2.3.6. *Soit $(\Omega, \mathcal{F}_n, X_n, \mathbb{P}_x)$ la chaîne canonique associée à la probabilité de transition $\mathbf{Q}f(x) = \int f(xy) d\nu(y)$ sur un groupe E supposé L.C.D. Pour tout $x \in E$ la suite $Y_n = X_{n-1}^{-1}X_n$, $n \geq 1$ est une suite de variables aléatoires indépendantes et de même loi de loi ν sous \mathbb{P}_x . Par conséquent on a:*

$$X_n = xY_1Y_2 \dots Y_n, \quad \mathbb{P}_x \text{ p.s.}$$

La loi du vecteur (X_0, X_1, \dots, X_n) sous \mathbb{P}_x est donc identique à celle du vecteur translaté $(xy^{-1}X_0, xy^{-1}X_1, \dots, xy^{-1}X_n)$ sous \mathbb{P}_y .

Indications de preuve:

La preuve se fait par récurrence sur n en montrant que pour tout système h_1, h_2, \dots, h_n de fonctions dans $\mathcal{B}^+(E)$ on a

$$\mathbb{E}_x\left(\prod_{k=1}^n h_k(Y_k)\right) = \prod_{k=1}^n \nu(h_k)$$

Pour $n = 1$ on a $\mathbb{E}_x(h(Y_1)) = \mathbb{E}_x(h(x^{-1}X_1)) = \int h_x(xy) d\nu(y) = \nu(h)$ où l'on a posé $h_x(y) = h(x^{-1}y)$. Ensuite pour passer à l'ordre $n + 1$ on écrit:

$$\begin{aligned}\mathbb{E}_x\left(\prod_{k=1}^{n+1} h_k(Y_k)\right) &= \mathbb{E}_x\left(\left(\prod_{k=1}^n h_k(Y_k)\right)\mathbb{E}_x(h_{n+1}(X_n^{-1}X_{n+1})|\mathcal{F}_n)\right) \\ &= \mathbb{E}_x\left(\left(\prod_{k=1}^n h_k(Y_k)\right)\mathbb{E}_{X_n}(h_{n+1}(Y_1))\right) \\ &= \mathbb{E}_x\left(\prod_{k=1}^n h_k(Y_k)\right)\nu(h_{n+1})\end{aligned}$$

Il suffit ensuite d'écrire $X_n = X_0 Y_1 Y_2 \dots Y_n$. ■

Un outil très utile dans la théorie des chaînes de Markov est la notion de chaîne induite sur un ensemble qui correspond à la notion de réduite dans la théorie du potentiel classique.

Définition 2.3.7. Soit $B \in \mathcal{E}$. Dans toute la suite, on adoptera les notations suivantes où, par convention, un inf calculé sur un ensemble vide, est égal à $+\infty$.

- Le temps d'arrêt $T_B = \inf\{n \geq 0 ; X_n \in B\}$ est appelé le temps d'entrée dans B .
- Le temps d'arrêt $S_B = \inf\{n \geq 1 ; X_n \in B\}$ est appelé le temps de retour dans B .

Soit $F \in \mathcal{E}$. On introduit les temps successifs de retour dans F par $S_F^{(0)} = 0$ puis $S_F^{(n+1)} = \inf\{k > S_F^{(n)} ; X_k \in F\}$ (on a donc $S_F = S_F^{(1)}$).

Proposition 2.3.8. On suppose que $\mathbb{P}_x(S_F < \infty) = 1$ pour tout $x \in F$. Alors

1. Pour tout $n \geq 1$ et pour tout $x \in F$, on a $\mathbb{P}_x(S_F^{(n)} < \infty) = 1$.
2. La suite de variables aléatoires $\tilde{X}_n = X_{S_F^{(n)}}$, $n \geq 0$ est, pour tout $x \in F$, une chaîne de Markov sous \mathbb{P}_x , de fonction de transition $\tilde{\mathbf{Q}}f(x) = \mathbb{E}_x(f(X_{S_F})) = \mathbb{E}_x(f(\tilde{X}_1))$
3. Soit $B \subset F$. On pose $S_B = \inf\{k > 0 ; X_k \in B\}$ et $\tilde{S}_B = \inf\{k > 0 ; \tilde{X}_k \in B\}$. Alors:

$$\mathbb{E}_x(S_B) \leq \mathbb{E}_x(\tilde{S}_B) \sup_{y \in F} \mathbb{E}_y(S_F)$$

Indications de preuve:

Preuve de 1.: La preuve se fait par récurrence sur n . Elle est vraie pour $n = 1$ et si on suppose $\mathbb{P}_x(S_F^{(n)} < \infty) = 1$ alors $S_F^{(n+1)} = S_F^{(n)} + S_F \circ \theta_{S_F^{(n)}}$, \mathbb{P}_x p.s. et par application de Markov fort:

$$\mathbb{P}_x(S_F^{(n+1)} < \infty) = \mathbb{E}_x(\mathbf{1}_{\{S_F < \infty\}} \circ \theta_{S_F^{(n)}}) = \mathbb{E}_x(\mathbb{P}_{X_{S_F^{(n)}}}(S_F < \infty)) = 1$$

car $X_{S_F^{(n)}} \in F$ sur $(S_F^{(n)} < \infty)$.

Preuve de 2.: le résultat précédent permet de définir les variables aléatoires \tilde{X}_n , \mathbb{P}_x p.s. pour $x \in F$ et l'on peut remarquer que $\tilde{X}_{n+1} = \tilde{X}_1 \circ \theta_{S_F^{(n)}}$. On pose $\tilde{\mathcal{F}}_n = \mathcal{F}_{S_F^{(n)}}$. On a alors:

$$\begin{aligned} \mathbb{E}_x(f(\tilde{X}_{n+1})|\tilde{\mathcal{F}}_n) &= \mathbb{E}_x(f(\tilde{X}_1 \circ \theta_{S_F^{(n)}})|\tilde{\mathcal{F}}_n) \\ &= \mathbb{E}_{X_{S_F^{(n)}}}(f(\tilde{X}_1)) = \tilde{Q}f(\tilde{X}_n) \end{aligned}$$

Preuve de 3.: On remarque que:

$$S_B = \sum_{0 \leq n < \tilde{S}_B} (S_F^{(n+1)} - S_F^{(n)}) = \sum_{n \geq 0} (S_F^{(n+1)} - S_F^{(n)}) \mathbf{1}_{\{\tilde{S}_B > n\}}$$

et par conséquent:

$$\begin{aligned} \mathbb{E}_x(S_B) &= \sum_n \mathbb{E}_x((S_F \circ \theta_{S_F^{(n)}}) \mathbf{1}_{\{\tilde{S}_B > n\}}) \\ &= \sum_n \mathbb{E}_x(\mathbf{1}_{\{\tilde{S}_B > n\}} \mathbb{E}_{X_{S_F^{(n)}}}(S_F)) \\ &\leq \mathbb{E}_x(\tilde{S}_B) \sup_{y \in F} \mathbb{E}_y(S_F) \end{aligned}$$

■

Nous utiliserons aussi la notion de chaîne tuée avant retour ou après passage dans un ensemble F c'est à dire que l'on ajoute un point Δ à E , on pose $E^* = E \cup \Delta$ et on définit les suites X'_n et X''_n par :

$$X'_n = \begin{cases} X_n & \text{si } n < S \\ \Delta & \text{si } n \geq S \end{cases} \quad X''_n = \begin{cases} X_n & \text{si } n \leq T \\ \Delta & \text{si } n > T \end{cases}$$

où $S = S_F$ et $T = T_F$.

Proposition 2.3.9. *Pour toute probabilité \mathbb{P}_x , $x \in E$, les suites X'_n et X''_n sont des chaînes de Markov à valeurs dans E^* de transition*

$$\begin{aligned} Q'(f^*)(x) &= Q(f \mathbf{1}_{\{F\}})(x) + Q(x, F) f^*(\Delta) \quad \text{si } x \in E, & Q'(f^*)(\Delta) &= f^*(\Delta) \\ Q''(f^*)(x) &= Q(f)(x) \quad \text{si } x \in (E \setminus F), & Q''(f^*)(x) &= f^*(\Delta) \quad \text{si } x \in (F \cup \Delta) \end{aligned}$$

où f^* est une fonction mesurable positive sur E^* et f désigne sa restriction à E .

Indications de preuve:

Pour X'_n , il suffit d'utiliser le fait que $\mathbb{E}_x(f(X_{n+1}) \mathbf{1}_{\{S=n+1\}} | \mathcal{F}_n) = \mathbf{1}_{\{S \geq n+1\}} Qf(X_n)$. Pour X''_n on remarque que $X''_n \in (E \setminus F)$ est équivalent à $T > n$. ■

2.3.2 Théorie du potentiel

Définition 2.3.10. Soit \mathbf{Q} une probabilité de transition sur (E, \mathcal{E}) et $(\Omega, \mathcal{F}_n, X_n, \mathbb{P}_x)$ la chaîne canonique associée.

- On définit la mesure potentielle $\mathbf{G}(x, \bullet)$ par

$$\mathbf{G}(x, B) = \sum_{n \geq 0} \mathbf{Q}^n(x, B)$$

- Soit $f \in \mathcal{B}^+(E)$. On dit que f est sur-harmonique si $\mathbf{Q}f \leq f$, que f est harmonique si $\mathbf{Q}f = f$ et que f est un potentiel s'il existe $u \in \mathcal{B}^+(E)$ telle que $f = \mathbf{G}u$.
- Soit $\nu \in \mathcal{M}^+(E)$. On dit que ν est sur-harmonique si $\nu\mathbf{Q} \leq \nu$, que ν est invariante (ou harmonique) si $\nu\mathbf{Q} = \nu$ et que ν est un potentiel s'il existe $m \in \mathcal{M}^+(E)$ telle que $\nu = m\mathbf{G}$.

Z. La mesure positive $\mathbf{G}(x, \bullet)$ n'est pas nécessairement finie sur les compacts ! Elle peut même être égale à $+\infty$ sur pratiquement tous les ensembles mesurables...

Proposition 2.3.11. Soit $g \in \mathcal{B}^+(E)$. Alors $\mathbf{G}g$ est la plus petite fonction $f \in \mathcal{B}^+(E)$ vérifiant l'équation de Poisson $g + \mathbf{Q}f = f$.

Indications de preuve:

$\mathbf{G}g$ satisfait bien sûr l'équation de Poisson. Soit f une solution, on vérifie facilement que pour tout $n \geq 0$ on a $f = \sum_{k=0}^n \mathbf{Q}^k g + \mathbf{Q}^{n+1} f$ d'où le résultat. ■

Proposition 2.3.12. (Principe du maximum) Soit f et g dans $\mathcal{B}^+(E)$ et $\alpha \geq 0$. Si l'on a $\mathbf{G}f \leq \mathbf{G}g + \alpha$ sur l'ensemble $(f > 0)$ alors cette inégalité est vraie partout. De plus $\mathbf{G}(x, y) = \mathbb{P}_x(T_y < \infty)\mathbf{G}(y, y)$ et pour $B \in \mathcal{E}$ on a $\mathbf{G}(x, B) \leq \sup_{y \in B} \mathbf{G}(y, B)$.

Indications de preuve:

Soit $B = (f > 0)$ et $T = T_B$ le temps d'entrée dans B .

$$\begin{aligned} \mathbf{G}f(x) &= \mathbb{E}_x\left(\sum_n f(X_n)\right) = \mathbb{E}_x\left(\sum_{n \geq T} f(X_n)\mathbf{1}_{\{T < \infty\}}\right) \\ &= \sum_n \mathbb{E}_x\left((f(X_n) \circ \theta_T)\mathbf{1}_{\{T < \infty\}}\right) = \sum_n \mathbb{E}_x\left(\mathbf{1}_{\{T < \infty\}}\mathbb{E}_{X_T}(f(X_n))\right) \\ &= \mathbb{E}_x\left(\mathbf{1}_{\{T < \infty\}}\mathbb{E}_{X_T}\left(\sum_n f(X_n)\right)\right) = \mathbb{E}_x\left(\mathbf{1}_{\{T < \infty\}}\mathbf{G}f(X_T)\right) \\ &\leq \sum_n \mathbb{E}_x\left(\mathbf{1}_{\{T < \infty\}}\mathbb{E}_{X_T}(g(X_n))\right) + \alpha\mathbb{P}_x(T < \infty) \\ &\leq \mathbf{G}g(x) + \alpha \end{aligned}$$

Si l'on considère $f = \mathbf{1}_{\{y\}}$ la même preuve fournit la formule demandée et la dernière inégalité s'obtient pour $f = \mathbf{1}_{\{B\}}$ et $\alpha = 0$. ■

Proposition 2.3.13. Soit $B \subset E$ tel que $\sup_{x \in B} G(x, B) < \infty$ et $A = \{x \in E ; \mathbb{P}_x(T_B < \infty) > 0\}$. Alors A est une réunion dénombrable d'ensembles A_n tels que la fonction $G(x, A_n)$ soit bornée sur E .

Indications de preuve:

On pose $A_{(n,k)} = \{x \in B ; \mathbf{Q}^n(x, B) \geq 1/k\}$. Alors on a:

$$\mathbf{Q}^{m+n}(x, B) \geq \int \mathbf{Q}^m(x, dy) \mathbf{1}_{\{A_{(n,k)}\}}(y) \mathbf{Q}^n(y, B) \geq \mathbf{Q}^m(x, A_{(n,k)})/k$$

et donc $k\mathbf{G}(x, B) \geq \mathbf{G}(x, A_{(n,k)})$ or, d'après le principe du maximum, on a la majoration $\mathbf{G}(x, B) \leq \sup_{x \in B} \mathbf{G}(x, B) < \infty$ ce qui termine la preuve. ■

Lemme 2.3.14. Soit $f \in \mathcal{B}^+(E)$. Si f est sur-harmonique (resp. harmonique) alors la suite $f(X_n)$ est une \mathbb{P}_x surmartingale (resp. martingale) positive (non nécessairement finie) pour tout $x \in E$.

Proposition 2.3.15. Soit $B \in \mathcal{E}$ et $g \in \mathcal{B}^+(E)$. On appelle solution du problème de Dirichlet toute fonction $f \in \mathcal{B}^+(E)$ vérifiant:

$$\mathbf{Q}f(x) = f(x), \text{ si } x \in E \setminus B, \quad f(x) = g(x), \text{ si } x \in B$$

Alors la plus petite solution positive du problème de Dirichlet est donnée par $h(x) = \mathbb{E}_x(\mathbf{1}_{\{T < \infty\}}g(X_T))$ où T est le temps d'entrée dans B .

Indications de preuve:

On va tout d'abord vérifier que h est bien solution. Pour $x \in B$ on a $T = 0$, \mathbb{P}_x p.s. donc $h(x) = g(x)$. Pour $x \notin B$ on a $T = 1 + T \circ \theta$ et $(T < \infty) = (T \circ \theta < \infty)$ \mathbb{P}_x p.s. et donc:

$$\begin{aligned} h(x) &= \mathbb{E}_x(\mathbf{1}_{\{T < \infty\}}g(X_T)) = \mathbb{E}_x(\mathbf{1}_{\{T < \infty\}}(g(X_T)) \circ \theta) \\ &= \mathbb{E}_x\left(\mathbb{E}_x(\mathbf{1}_{\{T < \infty\}}g(X_T)) \circ \theta \mid \mathcal{F}_1\right) = \mathbb{E}_x\left(\mathbb{E}_{X_1}(\mathbf{1}_{\{T < \infty\}}g(X_T))\right) \\ &= \mathbb{E}_x(h(X_1)) = \mathbf{Q}h(x) \end{aligned}$$

Il faut ensuite montrer que si f est solution, alors $f \geq h$. On considère $Z_n = f(X_{T \wedge n})$. On va montrer que Z_n est une \mathbb{P}_x martingale pour tout $x \in E$. Pour ceci on écrit $Z_{n+1} = \sum_{k=0}^n f(X_k) \mathbf{1}_{\{T=k\}} + \mathbf{1}_{\{T > n\}} f(X_{n+1})$ et donc:

$$\begin{aligned} \mathbb{E}_x(Z_{n+1} \mid \mathcal{F}_n) &= \sum_{k=0}^n f(X_k) \mathbf{1}_{\{T=k\}} + \mathbf{1}_{\{T > n\}} \mathbb{E}_x(f(X_{n+1}) \mid \mathcal{F}_n) \\ &= \sum_{k=0}^n f(X_k) \mathbf{1}_{\{T=k\}} + \mathbf{1}_{\{T > n\}} \mathbf{Q}f(X_n) \\ &= Z_n \end{aligned}$$

car sur $(T > n)$ on a $\mathbf{Q}f(X_n) = f(X_n)$. Sur l'ensemble $(T < \infty)$, la suite Z_n converge vers $f(X_T) = g(X_T)$ et par conséquent:

$$h(x) = \mathbb{E}_x(\mathbf{1}_{\{T < \infty\}}f(X_T)) \leq \liminf \mathbb{E}_x(Z_n) = f(x)$$

■

Proposition 2.3.16. (*Décomposition de Riesz*) Soit f une fonction surharmonique finie. Alors f s'écrit $f = h + \mathbf{G}u$ où h est harmonique. De plus h et u sont uniques.

Indications de preuve:

On vérifie immédiatement que $h = \lim \downarrow \mathbf{Q}^n f$ est harmonique (finie). De plus $\sum_{k=0}^N (f - \mathbf{Q}^k f) = f - \mathbf{Q}^{N+1} f$ donc $f - h$ est le potentiel de la fonction positive finie $u = f - \mathbf{Q}f$. Pour établir l'unicité on remarque que $f = h + \mathbf{G}u$ implique que $\mathbf{Q}^n f = h + \sum_{k=n}^{\infty} \mathbf{Q}^k u$, la série définissant $\mathbf{G}u$ étant convergente on en déduit que nécessairement $h = \lim \downarrow \mathbf{Q}^n f$. Il reste à vérifier que si deux potentiels finis $\mathbf{G}u_1$ et $\mathbf{G}u_2$ sont égaux alors il en est de même de u_1 et u_2 ce qui est immédiat en remarquant que $\mathbf{G}u = u + \mathbf{Q}\mathbf{G}u$. ■

On peut obtenir le même genre de décomposition sur les mesures:

Proposition 2.3.17. Soit ν une mesure surharmonique σ -finie. Alors ν s'écrit $\nu = \mu + m\mathbf{G}$ où μ est invariante. De plus μ et m sont uniques.

Indications de preuve:

On vérifie immédiatement que pour $f \in \mathcal{B}^+(E)$ la formule $\mu(f) = \lim \downarrow \nu \mathbf{Q}^n f$ définit une mesure invariante μ et l'on poursuit comme ci-dessus en utilisant la remarque suivante: Soit α et β des mesures positives avec α σ finie et $\beta \leq \alpha$. Si l'on pose $\beta = h\alpha$ alors $0 \leq h \leq 1$ p.s. et la différence $\alpha - \beta$ est définie comme la mesure positive de densité $1 - h$ par rapport à α . Ceci permet de considérer le potentiel $m\mathbf{G}$ de la mesure positive $m = \nu - \nu\mathbf{Q}$ et on utilise le fait que l'égalité $\nu = \mu + m\mathbf{G}$ sur des ensembles B avec $\nu(B) < \infty$ implique l'égalité des mesures. ■

Définition 2.3.18. Soit $B \in \mathcal{E}$. On pose:

$$\pi_B(x) = \mathbb{P}_x(T_B < \infty), \quad N_B = \sum_{n \geq 0} \mathbf{1}_{\{B\}} \circ X_n, \quad h_B(x) = \mathbb{P}_x(N_B = \infty)$$

Proposition 2.3.19. La fonction π_B est sur-harmonique, sa partie harmonique dans la décomposition de Riesz est égale à h_B . De plus pour tout $x \in E$ la \mathbb{P}_x surmartingale $\pi_B(X_n)$ et la \mathbb{P}_x martingale $h_B(X_n)$ convergent \mathbb{P}_x p.s. vers $\mathbf{1}_{\{N_B = \infty\}}$

Indications de preuve:

On pose $T_B^k = \inf\{n \geq k ; X_n \in B\}$ alors $\mathbf{1}_{\{T_B^k < \infty\}} = \mathbf{1}_{\{T_B < \infty\}} \circ \theta_k$ et:

$$\begin{aligned} \mathbf{Q}^k \pi_B(x) &= \mathbb{E}_x(\pi_B \circ X_k) = \mathbb{E}_x\left(\mathbb{P}_{X_k}(T_B < \infty)\right) \\ &= \mathbb{E}_x\left(\mathbb{E}_x(\mathbf{1}_{\{T_B < \infty\}} \circ \theta_k | \mathcal{F}_k)\right) = \mathbb{P}_x(T_B^k < \infty) \\ &\leq \pi_B(x) \end{aligned}$$

Ceci prouve que π_B est sur-harmonique et que sa partie harmonique est

$$\lim_n \downarrow \mathbb{P}_x(T_B^n < \infty) = \mathbb{P}_x(\cap_n (T_B^n < \infty)) = h_B(x)$$

car les T_B^n forment une suite croissante. $\pi_B(X_n)$ est une surmartingale positive et $h_B(X_n)$ est une martingale positive, elles convergent donc \mathbb{P}_x p.s. et ces suites étant bornées, elles convergent dans L^1 . Or on a: $\mathbf{1}_{\{N_B=\infty\}} \circ \theta_n = \mathbf{1}_{\{N_B=\infty\}}$ et donc

$$h_B(X_n) = \mathbb{E}_{X_n}(\mathbf{1}_{\{N_B=\infty\}}) = \mathbb{E}_x(\mathbf{1}_{\{N_B=\infty\}} \circ \theta_n | \mathcal{F}_n) = \mathbb{E}_x(\mathbf{1}_{\{N_B=\infty\}} | \mathcal{F}_n)$$

On en déduit que $h_B(X_n)$ converge \mathbb{P}_x p.s. vers $Y = \mathbf{1}_{\{N_B=\infty\}}$. Si Z est la limite de $\pi_B(X_n)$ on a bien sûr $Z \geq Y$ et le théorème de Lebesgue implique que

$$\mathbb{E}_x(Z) = \lim \mathbb{E}_x(\pi_B(X_n)) = \lim \mathbf{Q}^n \pi_B(x) = h_B(x) = \mathbb{E}_x(Y)$$

et donc $Z = Y$. ■

Comme conséquence de cette proposition on peut donner le résultat suivant:

Corollaire 2.3.20. *Soit Y_n une suite de variables aléatoires indépendantes et de même loi sur \mathbb{R} et dont l'espérance est strictement positive. On pose $S_0 = 0$, $S_n = Y_1 + Y_2 + \dots + Y_n$ pour $n \geq 1$ et $\varphi(x) = \mathbb{P}(\cup_{n \geq 0} (x + S_n \leq 0))$ alors $\lim_{x \rightarrow +\infty} \varphi(x) = 0$.*

Indications de preuve:

Soit X_n la marche canonique associée et T le temps d'entrée de X_n dans $B = \mathbb{R}^-$ alors $\varphi(x) = \mathbb{P}_x(T < \infty) = \pi_B(x)$ La fonction $\varphi(x)$ est décroissante, il suffit donc de trouver une suite $x_n \rightarrow \infty$ avec $\lim \varphi(x_n) = 0$. Or X_n converge \mathbb{P}_x p.s. vers $+\infty$ et $\varphi(X_n)$ converge \mathbb{P}_x p.s. vers $\mathbf{1}_{\{N_B=\infty\}}$ qui est donc \mathbb{P}_x négligeable. ■

2.4 Les chaînes discrètes

On suppose dans cette section que E est dénombrable. Il existe beaucoup d'ouvrages traitant ce cas et l'on se contentera donc de rappeler les définitions et prouver brièvement les résultats essentiels. De plus, la plupart de ces propriétés seront établies plus loin, dans le cadre des "chaînes atomiques". Par contre il est sans doute moins aisé de trouver une preuve complète des critères de Foster et nous donnerons donc les démonstrations de ces critères qui sont très utiles dans l'étude des files d'attente.

2.4.1 Rappels

On considère une chaîne canonique $(\Omega, \mathcal{F}_n, X_n, \mathbb{P}_x)$ sur un espace dénombrable E .

Définition 2.4.1. *Soit $x \in E$, on pose $\rho(x) = \mathbb{P}_x(S_x < \infty)$.*

- *On dit que x est récurrent si $\rho(x) = 1$. Sinon x est dit transitoire.*
- *Une chaîne est dite transiente si tous les états sont transitoires et récurrente si tous les états sont récurrents.*

Proposition 2.4.2. *Si x est transitoire, la variable aléatoire N_x suit une loi géométrique de paramètre $\rho(x)$ sous \mathbb{P}_x c'est à dire que $\mathbb{P}_x(N_x = n) = (\rho(x))^{n-1} (1 - \rho(x))$ pour $n \geq 1$.*

On a donc $\mathbb{E}_x(N_x) = \frac{\rho(x)}{1 - \rho(x)}$, en particulier $\mathbf{G}(x, x) < \infty$.

Indications de preuve:

On a $N_x = 1$ dans le cas $S_x = \infty$ et sinon $N_x = 1 + N_x \circ \theta_{S_x}$. Par conséquent pour $k \geq 1$, on obtient par application de la propriété de Markov forte:

$$\mathbb{P}_x(N_x = k + 1) = \mathbb{P}_x(N_x \circ \theta_{S_x} = k \cap (S_x < \infty)) = \rho(x) \mathbb{P}_x(N_x = k) \blacksquare$$

Proposition 2.4.3. *On a les équivalences:*

1. *Le point x est récurrent.*
2. $\mathbb{P}_x(N_x = \infty) = 1$.
3. $\mathbf{G}(x, x) = \mathbb{E}_x(N_x) = \infty$.

L'équivalence de (1) et (2) résulte de la première assertion de 2.3.8. On a bien sûr $2) \Rightarrow 3)$ et il suffit d'utiliser la proposition précédente pour la réciproque. \blacksquare

Proposition 2.4.4. *Si la chaîne est transiente, $\mathbf{G}(x, y) < \infty$ pour tout couple (x, y) de $E \times E$. La suite X_n tend vers l'infini, c'est à dire que pour toute partie finie $F \subset E$, tout $x \in E$ il existe N tel que \mathbb{P}_x p.s. on ait $X_n \notin F$ pour $n \geq N$.*

Le principe du maximum assure que $\mathbf{G}(x, y) \leq \mathbf{G}(y, y) < \infty$. On a alors $\mathbb{E}_x(N_F) = \sum_{y \in F} \mathbb{E}_x(N_y) < \infty$. La variable aléatoire N_F est donc \mathbb{P}_x presque sûrement finie... \blacksquare

Définition 2.4.5. *On dit que la chaîne est irréductible si $\mathbf{G}(x, y) > 0$ pour tout couple (x, y) . Il revient au même de dire que $\mathbb{P}_x(T_y < \infty) > 0$ pour tout couple (x, y) .*

Proposition 2.4.6. *On suppose que la chaîne est irréductible. S'il existe un point récurrent x alors:*

1. $\mathbb{P}_x(N_y = \infty) = 1$ pour tout y .
2. *Tous les états sont récurrents.*

On en déduit que pour une chaîne irréductible, tous les états sont de même nature. Si la chaîne est récurrente irréductible on a $\mathbf{G}(x, y) = \infty$ pour tout couple (x, y) .

Indications de preuve:

Soit y un point de E . D'après 2.3.19, la \mathbb{P}_x martingale $h(X_n) = \mathbb{P}_{X_n}(N_y = \infty)$ et la \mathbb{P}_x sur-martingale $f(X_n) = \mathbb{P}_{X_n}(T_y < \infty)$ convergent \mathbb{P}_x p.s. vers l'indicatrice de l'ensemble $(N_y = \infty)$. Or \mathbb{P}_x p.s. il existe une infinité de n avec $X_n = x$. On en déduit donc:

$$\mathbb{P}_x(N_y = \infty) = \mathbb{P}_x(T_y < \infty) = \mathbf{1}_{\{N_y = \infty\}}, \quad \mathbb{P}_x \text{ p.s.}$$

Or on sait que $\mathbb{P}_x(T_y < \infty) > 0$, on en déduit donc $\mathbb{P}_x(N_y = \infty) = \mathbb{P}_x(T_y < \infty) = 1$. Pour prouver que y est récurrent on utilise le fait que $\mathbb{P}_x(N_y = \infty) = \mathbb{P}_x(N_y \circ \theta_{T_y} = \infty)$ et on utilise la propriété de Markov forte. \blacksquare

On étudie surtout les chaînes récurrentes irréductibles qui ont les propriétés les plus intéressantes. La proposition et le théorème suivant seront démontrés dans le contexte des chaînes atomiques.

Définition 2.4.7. Soit $\lambda \in \mathcal{M}^+(E)$. On dit que λ est non triviale si ce n'est pas la mesure nulle ni la mesure infinie en tout point.

Proposition 2.4.8. Si \mathbf{Q} définit une chaîne récurrente irréductible alors:

1. Les fonctions sur-harmoniques sont constantes.
2. Il existe une unique mesure invariante non triviale λ à une constante multiplicative près. Si λ est bornée on dit que la récurrence est positive, et sinon que la récurrence est nulle.
3. La mesure λ est strictement positive et finie en tout point. De plus toute mesure surharmonique est multiple de λ .

On a alors le théorème ergodique des chaînes de Markov:

Théorème 2.4.9. Soit $(\Omega, \mathcal{F}_n, X_n, \mathbb{P}_x)$ une chaîne récurrente irréductible.

- Si f et g sont λ intégrables avec $\lambda(g) \neq 0$ on a pour tout $x \in E$:

$$\lim_{n \rightarrow \infty} \frac{\sum_{k=1}^n f(X_k)}{\sum_{k=1}^n g(X_k)} = \frac{\lambda(f)}{\lambda(g)}, \quad \mathbb{P}_x \text{ p.s.}$$

- La récurrence est positive si et seulement si $\mathbb{E}_x(S_x) < \infty$ pour un $x \in E$ (ou pour tout $x \in E$). Alors l'unique probabilité invariante λ vérifie $\lambda(x) = \frac{1}{\mathbb{E}_x(S_x)}$ et pour tout $x \in E$, et toute fonction λ intégrable f on a:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n f(X_k) = \lambda(f), \quad \mathbb{P}_x \text{ p.s.}$$

- Si la récurrence est nulle, pour tout $x \in E$, et toute fonction λ intégrable f on a:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n f(X_k) = 0, \quad \mathbb{P}_x \text{ p.s.}$$

Ce théorème est essentiellement une conséquence de la loi des grands nombres appliquée à la suite de variables aléatoires $Z_n = Z_{n-1} \circ \theta_{S_x}$ avec $Z_1 = \sum_{k=1}^{S_x} f(X_k)$. En effet cette suite forme une famille de variables aléatoires indépendantes et de même loi sous \mathbb{P}_x . On peut citer un théorème ergodique voisin de celui que l'on obtiendra pour les marches aléatoires:

Théorème 2.4.10. Soit \mathbf{Q} une transition définissant une chaîne récurrente irréductible, alors:

$$\lim_{n \rightarrow \infty} \frac{\sum_{k=1}^n \mathbf{Q}^k(x_1, y_1)}{\sum_{k=1}^n \mathbf{Q}^k(x_2, y_2)} = \frac{\lambda(y_1)}{\lambda(y_2)}$$

pour tous points x_1, x_2, y_1, y_2 .

2.4.2 Critères de Foster

Ces critères sont des outils commodes pour montrer qu'une chaîne irréductible est récurrente ou récurrente positive.

Proposition 2.4.11. (*Premier critère de Foster*). Soit X_n une chaîne de Markov et F un sous-ensemble fini de E . S'il existe $f \in \mathcal{B}^+(E)$ partout finie et telle que :

$$\mathbf{Q}f(x) \leq f(x), \quad x \in E \setminus F, \quad \lim_{x \rightarrow \infty} f(x) = +\infty$$

alors il existe un point récurrent.

Indications de preuve :

Soit T le temps d'entrée dans F . En recopiant la preuve de 2.3.15 on montre que $Z_n = f(X_{n \wedge T})$ est une \mathbb{P}_x surmartingale positive pour tout x . Elle converge donc \mathbb{P}_x p.s. vers une variable aléatoire Z qui est égale à $\lim f(X_n)$ sur l'ensemble $(T = \infty)$. Si on suppose la chaîne transiente, alors \mathbb{P}_x p.s. elle tend vers l'infini et donc $Z = \infty$, \mathbb{P}_x p.s. sur $(T = \infty)$. Or $\mathbb{E}_x(Z) \leq \lim_n \mathbb{E}_x(Z_n) \leq \mathbb{E}_x(Z_0) = f(x) < \infty$. On en déduit donc que $\mathbb{P}_x(T < \infty) = 1$ pour tout x . Or si S_F est le temps de retour à F on a $S_F = 1 + T \circ \theta$ et donc

$$\mathbb{P}_x(S_F < \infty) = \mathbb{P}_x(T \circ \theta < \infty) = \mathbb{E}_x(\mathbb{P}_{X_1}(T < \infty)) = 1$$

En utilisant 2.3.8 on peut définir une chaîne induite \tilde{X}_n sur l'ensemble F . Celui-ci étant fini, la chaîne induite possède un point récurrent. Ce point est aussi récurrent pour la chaîne X_n , ce qui infirme notre supposition. ■

Proposition 2.4.12. (*Second critère de Foster*). Soit X_n une chaîne de Markov irréductible et $F \subset E$ un ensemble fini. S'il existe $f \in \mathcal{B}^+(E)$ et $\epsilon > 0$ avec f et $\mathbf{Q}f$ finies tels que :

$$\mathbf{Q}f(x) \leq f(x) - \epsilon, \quad x \in E \setminus F$$

alors la chaîne est récurrente positive.

Indications de preuve :

Soit T le temps d'entrée dans F . On montre à nouveau que $Z_n = f(X_{n \wedge T}) + \epsilon(n \wedge T)$ est une \mathbb{P}_x surmartingale positive pour tout x . On a donc

$$\mathbb{E}_x(n \wedge T) \leq \mathbb{E}_x(Z_n)/\epsilon \leq \mathbb{E}_x(Z_0)/\epsilon = f(x)/\epsilon$$

En prenant la limite lorsque $n \rightarrow \infty$ on obtient $\mathbb{E}_x(T) \leq f(x)/\epsilon$. Si S_F est le temps de retour à F on a $S_F = 1 + T \circ \theta$ et par Markov $\mathbb{E}_x(S_F) \leq (1 + \mathbf{Q}f(x)/\epsilon) < \infty$. En utilisant 2.3.8 on peut définir une chaîne induite \tilde{X}_n sur l'ensemble F . La chaîne induite est donc récurrente irréductible positive puisque F est fini. En reprenant l'alinéa 3. de 2.3.8 pour un choix de B réduit à un seul point $x_0 \in F$ alors $\mathbb{E}_{x_0}(\tilde{S}_{x_0}) < \infty$. Puisque F est fini, on en déduit :

$$\mathbb{E}_{x_0}(S_{x_0}) \leq \mathbb{E}_{x_0}(\tilde{S}_{x_0}) \sup_{y \in F} \mathbb{E}_y(S_F) < \infty$$

■

2.5 Marches aléatoires

Soit E un groupe localement compact métrisable d'élément neutre e . Il existe sur E une mesure de Radon positive non nulle notée m_g , unique à une constante multiplicative près, qui est invariante à gauche c'est à dire $m_g(f) = \int f(yx) dm_g(x)$ pour toute fonction $f \in \mathcal{C}_k(E)$ (mesure de Haar à gauche). De même il existe une mesure de Haar à droite, notée m_d . Ces deux mesures sont équivalentes et on dit que le groupe est unimodulaire si elles sont proportionnelles, c'est le cas des groupes abéliens, semi-simples ou compacts. La topologie de E peut être définie par une distance invariante à gauche et dans la suite on considérera que les boules sont définies par cette distance invariante à gauche c'est à dire vérifiant $d(zx, zy) = d(x, y)$ et on note \mathcal{V}_x la famille des voisinages ouverts de x . Cette famille est la translatée (à droite ou à gauche) de la famille \mathcal{V}_e par l'élément x . On note $(\Omega, \mathcal{F}_n, X_n, \mathbb{P}_x)$ la chaîne canonique associée à la marche aléatoire à droite de loi ν et on pose $\Pi = \sum_{n \geq 0} \nu^{*n}$ (alors $\Pi(B) = \mathbb{E}_e(N_B)$).

2.5.1 Récurrence, transience

On rappelle que si B est un borélien de E , on pose $S_B = \inf\{n \geq 1 ; X_n \in B\}$.

Définition 2.5.1. *On dit que x est un point récurrent si pour tout $V \in \mathcal{V}_x$ on a $\mathbb{P}_x(S_V < \infty) = 1$. Dans le cas contraire il est dit transitoire.*

En vertu de la dernière affirmation de 2.3.6 on constate immédiatement que x est récurrent si et seulement si e est récurrent. Donc soit tous les points de E sont récurrents (on dit alors que la marche est récurrente) soit tous les points sont transitoires (on dit alors que la marche est transitoire).

Proposition 2.5.2. *Les propriétés suivantes sont équivalentes:*

1. *La marche est récurrente.*
2. *Pour tout $V \in \mathcal{V}_x$, on a $\mathbb{P}_x(N_V = \infty) = 1$*
3. *Pour tout $V \in \mathcal{V}_e$, on a $\Pi(V) = +\infty$.*

Indications de preuve:

Dans 2. il suffit de considérer $x = e$ et les implications 2. \implies 1. et 2. \implies 3. sont triviales.

• Preuve de 1. \implies 2.

On note B_r la boule centrée en e et de rayon r . Pour $0 < r' < r$ on a:

$$(X_m \in B'_r) \cap (X_{m+n} \notin Br) \subset (X_m^{-1} X_{m+n} \notin B_{r-r'})$$

et par conséquent:

$$\begin{aligned} \mathbb{P}_e\left((X_m \in B'_r) \cap_{n \geq 1} (X_{m+n} \notin Br)\right) &\leq \mathbb{P}_e\left(\cap_{n \geq 1} \left(\prod_{k=1}^n Y_{m+k} \notin B_{r-r'}\right)\right) \\ &\leq \mathbb{P}_e(S_{B_{r-r'}} = \infty) = 0 \end{aligned}$$

Si on fait tendre $r' \uparrow r$ on obtient $\mathbb{P}_e(X_m \in B_r \cap_{n \geq 1} (X_{m+n} \notin B_r)) = 0$. Or

$$(N_{B_r} < \infty) = \cup_m (X_m \in B_r \cap_{n \geq 1} (X_{m+n} \notin B_r))$$

cette dernière réunion étant disjointe.

• Preuve de 3. \implies 1.

On a l'inclusion:

$$(X_m \in B_r) \cap (X_{m+n} \notin B_r) \supset (X_m^{-1} X_{m+n} \notin B_{2r}) \cap (X_m \in B_r)$$

et par conséquent:

$$\begin{aligned} \mathbb{P}_e \left((X_m \in B_r) \cap_{n \geq 1} (X_{m+n} \notin B_r) \right) &\geq \mathbb{P}_e \left(\cap_{n \geq 1} \left(\prod_{k=1}^n Y_{m+k} \notin B_{2r} \right) \cap (X_m \in B_r) \right) \\ &\geq \mathbb{P}_e(S_{B_{2r}} = \infty) \mathbb{P}_e(X_m \in B_r) \end{aligned}$$

En sommant sur m on obtient $\mathbb{P}_e(N_{B_r} < \infty) \geq \mathbb{P}_e(S_{B_{2r}} = \infty) \mathbb{E}_e(N_{B_r})$ et donc si $\mathbb{E}_e(N_{B_r}) = \infty$ alors $\mathbb{P}_e(S_{B_{2r}} < \infty) = 1$. ■

Proposition 2.5.3. *Si la marche est transiente alors pour tout compact K on a $\Pi(K) < \infty$ (autrement dit Π est une mesure de Radon) et la fonction $x \rightarrow \mathbf{G}(x, K)$ est bornée.*

Indications de preuve:

Il existe $V \in \mathcal{V}_e$ avec $\Pi(V) < \infty$. Soit alors $U \in \mathcal{V}_e$ tel que $U^{-1}U \subset V$. Pour $x \in U$ on a: $\mathbf{G}(x, U) = \Pi(x^{-1}U) \leq \Pi(V)$ et d'après le principe du maximum 2.3.12 on a $\sup_{y \in E} \mathbf{G}(y, U) \leq \Pi(V)$. Si K est compact on peut trouver $(y_1, y_2, \dots, y_N) \in K$ tels que $K \subset (\cup_{i=1}^N y_i U)$ et donc pour x quelconque:

$$\mathbf{G}(x, K) \leq \mathbf{G}(x, \cup_{i=1}^N y_i U) \leq \sum_{i=1}^N \mathbf{G}(x, y_i U) = \sum_{i=1}^N \mathbf{G}(y_i^{-1}x, U) \leq N\Pi(V)$$

■

Remarques

- La récurrence d'une chaîne est la même pour les marches à gauche ou à droite, et pour ν et sa symétrisée $\check{\nu}$.
- Si E est compact, toutes les marches sont récurrentes.
- Si la marche est transiente alors pour tout compact K on a $\mathbb{P}_x(N_K = \infty) = 0$ et donc la marche sort presque sûrement de tout compact (elle tend vers l'infini).

Définition 2.5.4. *Soit $\nu \in \mathcal{M}_1^+(E)$. On désigne par T_ν le plus petit sous semi-groupe fermé de E contenant le support de ν et par G_ν le plus petit sous groupe fermé de E contenant le support de ν . On dit que ν est adaptée (ou que la marche de loi ν est adaptée) si $G_\nu = E$.*

Il est facile de constater que T_ν est le support de la mesure Π puisque le support de ν^{*n} est l'adhérence de la puissance n ème du support de ν et donc $x \in T_\nu$ si et seulement si pour tout $V \in \mathcal{V}_x$ on a $\mathbb{P}_e(T_V < \infty) > 0$.

Proposition 2.5.5. *On suppose que la marche de loi ν est récurrente. Alors:*

1. on a $T_\nu = G_\nu$
2. pour tout couple x, y dans G_ν et pour tout $V \in \mathcal{V}_y$ on a $\mathbb{P}_x(N_V = \infty) = 1$.

Indications de preuve:

On remarque qu'il suffit de prouver 2. pour $x = e$ en remplaçant y par $x^{-1}y$. On pose alors:

$$R = \{y ; \mathbb{P}_e(N_V = \infty) = 1 \text{ pour tout } V \in \mathcal{V}_y\}$$

Par hypothèse $e \in R$ et on constate immédiatement que $R \subset T_\nu$. Si l'on peut prouver que $T_\nu^{-1}R \subset R$ (\spadesuit) on aura:

$$\begin{cases} R \subset T_\nu \\ T_\nu^{-1}R \subset R \end{cases} \quad \text{et puisque } e \in R, \quad T_\nu^{-1} \subset R \subset T_\nu$$

par conséquent T_ν est un groupe, il est donc égal à G_ν et de plus $R = T_\nu = G_\nu$.

Pour prouver l'inclusion (\spadesuit) on considère $x \in T_\nu$, $y \in R$ et $V \in \mathcal{V}_{x^{-1}y}$. On a $xVy^{-1} \in \mathcal{V}_e$ et l'on peut trouver $U \in \mathcal{V}_e$ tel que $U^{-1}U \subset xVy^{-1}$. On a alors:

$$\begin{aligned} (X_m \in Ux) \cap (X_{n+m} \in Uy) &\subset (X_m^{-1}X_{n+m} \in V) \\ (X_m \in Ux) \cap (\overline{\lim}_n (X_n \in Uy)) &\subset \overline{\lim}_n \left(\prod_{k=1}^n Y_{m+k} \in V \right) \end{aligned}$$

Or $\mathbb{P}_e(\overline{\lim}_n (X_n \in Uy)) = \mathbb{P}_e(N_{Uy} = \infty) = 1$ donc:

$$\begin{aligned} \mathbb{P}_e((X_m \in Ux) \cap (\overline{\lim}_n \left(\prod_{k=1}^n Y_{m+k} \in V \right))) &= \mathbb{P}_e(X_m \in Ux) \\ \mathbb{P}_e(X_m \in Ux) \mathbb{P}_e(N_V = \infty) &= \mathbb{P}_e(X_m \in Ux) \end{aligned}$$

Or $\exists m$ avec $\mathbb{P}_e(X_m \in Ux) > 0$ d'où le résultat. ■

Dans le cas d'un groupe dénombrable, une marche récurrente adaptée est l'analogue d'une chaîne récurrente irréductible.

2.5.2 Théorème ergodique

L'importance du théorème ergodique des marches aléatoires ne réside pas tellement dans le résultat lui-même que dans sa conséquence principale: Si un groupe porte une marche adaptée récurrente alors il est unimodulaire! On en déduit donc que toutes les marches adaptées sur des groupes non unimodulaires sont transientes.

On considère ci dessous une marche adaptée récurrente sur E

Proposition 2.5.6. *Les fonctions surharmoniques continues sont constantes.*

Indications de preuve:

Soit f surharmonique continue, $x, y \in E$ et $\epsilon > 0$. Il existe des voisinages U et V de x et y sur lesquels f diffère de $f(x)$ ou $f(y)$ de moins de $\epsilon/2$. Soit $Z = \lim (f(X_n))$, la surmartingale positive $f(X_n)$ converge \mathbb{P}_e p.s. vers Z et puisque $\mathbb{P}_e(N_U = \infty) = 1$ et $\mathbb{P}_e(N_V = \infty) = 1$ on en déduit que \mathbb{P}_e p.s. la suite X_n appartient une infinité de fois à U et une infinité de fois à V et donc que \mathbb{P}_e p.s. on a $|f(x) - Z| \leq \epsilon/2$ et $|f(y) - Z| \leq \epsilon/2$. ■

Proposition 2.5.7. *Soit θ une mesure de Radon positive sur E vérifiant $\nu \star \theta \leq \theta$ alors θ est proportionnelle à la mesure de Haar à gauche m_g (si $\theta \star \nu \leq \theta$ alors θ est proportionnelle à la mesure de Haar à droite m_d)*

Indications de preuve:

Soit $\varphi \in \mathcal{C}_k^+(E)$. On pose $\Psi(x) = \int \varphi(xy) d\theta(y)$ et on obtient par le théorème de Fubini que $\mathbf{Q}\Psi(x) = \int \Psi(xz) d\nu(z) \leq \Psi(x)$. Or Ψ étant continue, elle est donc constante (par exemple égale à sa valeur à l'origine). ■

Théorème 2.5.8. (Théorème ergodique) *Si E porte une marche adaptée récurrente de transition \mathbf{Q} alors:*

1. E est unimodulaire.
2. Soit m une mesure de Haar sur E et $f, g \in \mathcal{C}_k^+(E)$ avec $m(g) \neq 0$.

$$\lim_n \frac{\sum_{k=0}^n \mathbf{Q}^k f(x)}{\sum_{k=0}^n \mathbf{Q}^k g(y)} = \frac{m(f)}{m(g)} \text{ pour tout } x, y \in E$$

Indications de preuve:

En translatant les fonctions f et g on voit qu'il suffit d'établir 2. pour $x = y = e$ et alors $\mathbf{Q}^k f(e) = \nu^{\star k}(f)$. Soit alors $h \in \mathcal{C}_k^+(E)$ fixée non nulle. Il existe un ouvert V et $\alpha > 0$ avec $h \geq \alpha$ sur V . En utilisant 2. de la proposition 2.5.5 on a $\mathbb{P}_x(N_V = \infty) = 1$ et donc pour tout x :

$$\sum_{k=0}^{\infty} \mathbf{Q}^k h(x) = \mathbf{G}h(x) \geq \alpha \mathbf{G}(x, V) = \alpha \mathbb{E}_x(N_V) = \infty$$

On peut donc considérer, pour n assez grand, la suite de mesures de Radon positives

$$\lambda_n = \frac{\sum_{k=0}^n \nu^{\star k}}{\sum_{k=0}^n \nu^{\star k}(h)}$$

Soit $K \subset E$ un compact, il existe n_0 tel que pour $n \geq n_0$ on ait $\sum_{k=0}^n \mathbf{Q}^k h(x) \geq 1$ sur K (en vertu du théorème de Dini) et donc $\nu^{\star r}(K) \leq \sum_{k=r}^{n_0+r} \nu^{\star k}(h)$. On en déduit :

$$\lambda_n(K) \leq \frac{(n_0 + 1) \sum_{k=0}^{n+n_0} \nu^{\star k}(h)}{\sum_{k=0}^n \nu^{\star k}(h)} \leq (n_0 + 1) \left(1 + \frac{n_0 \|h\|}{\sum_{k=0}^n \nu^{\star k}(h)} \right)$$

La suite λ_n est donc vaguement relativement compacte en vertu de la proposition 1.3.7. Soit λ une valeur d'adhérence vague limite vague d'une sous suite λ_{n_k} . On a alors $\lambda \star \nu \leq \lim_k (\lambda_{n_k} \star \nu)$ et $\nu \star \lambda \leq \lim_k (\nu \star \lambda_{n_k})$ Or :

$$\lambda_n \star \nu = \nu \star \lambda_n = \lambda_n - \frac{\partial_e}{\sum_{k=0}^n \nu^{\star k}(h)} + \frac{\nu^{\star(n+1)}}{\sum_{k=0}^n \nu^{\star k}(h)}$$

On en déduit donc que $\lambda \star \nu \leq \lambda$ et $\nu \star \lambda \leq \lambda$. Par conséquent d'après la proposition 2.5.7, λ est à la fois proportionnelle à m_d et m_g . Or λ est non nulle puisque $\lambda(h) = 1$, on en déduit donc que le groupe est unimodulaire et que la suite vaguement relativement compacte λ_n ne possédant qu'une seule valeur d'adhérence est vaguement convergente. Il suffit alors de considérer le rapport $\lambda_n(f)/\lambda_n(g)$ pour conclure. ■

Corollaire 2.5.9. Soit $E = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} ; a > 0, b \in \mathbb{R} \right\}$ le groupe affine de la droite et $\nu \in \mathcal{M}_1^+(E)$ adaptée. Alors la marche de loi ν est transitoire.

Indications de preuve :

On a $m_d = \frac{1}{a} da \otimes db \neq m_g = \frac{1}{a^2} da \otimes db$ ■

On peut utiliser ce corollaire pour des systèmes de production "affines" mis en série, c'est à dire l'étude d'équations de récurrence de la forme $X_{n+1} = A_n X_n + B_n$ où la suite (A_n, B_n) est i.i.d. dans \mathbb{R}^2 . Mais ceci ne règle pas le cas de l'équation de "Schrödinger" citée dans les exemples de chaînes car le groupe $SL(2, \mathbb{R})$ est unimodulaire. Pour ce faire on va utiliser la notion de groupe moyennable.

Définition 2.5.10. On dit que E est moyennable si toute action de E sur un compact métrique U admet une probabilité invariante c'est à dire qu'il existe $\theta \in \mathcal{M}_1^+(U)$ invariante par l'action de tous les éléments de E .

Proposition 2.5.11. S'il existe $\nu \in \mathcal{M}_1^+(E)$ adaptée et récurrente alors E est moyennable.

Indications de preuve :

Soit $\theta \in \mathcal{M}_1^+(U)$, on définit la "pseudo-convolution" $\nu \star \theta$ par $\nu \star \theta(f) = \int_E \int_U f(x \cdot u) d\nu(x) d\theta(u)$ et on considère la suite $\theta_n \in \mathcal{M}_1^+(U)$ définie par $\theta_n = (\sum_{k=0}^{n-1} \nu^{\star k}) \star \theta / n$. U étant compact métrique, il existe une sous suite θ_{n_k} étroitement convergente vers

$\gamma \in \mathcal{M}_1^+(U)$. Il est facile de constater que $\gamma = \nu \star \gamma$. Pour $f \in \mathcal{C}^+(U)$ on définit la fonction continue sur E , $F(x) = \int_U f(x \cdot u) d\gamma(u)$ alors en considérant la marche de loi ν on a:

$$\mathbf{Q}F(x) = \int_E F(xy) d\nu(y) = \int_E \int_U f(xy \cdot u) d\gamma(u) d\nu(y) = F(x)$$

D'après 2.5.6 la fonction F est constante et en la comparant à sa valeur au point e on voit que γ est invariante par l'action de E . ■

Corollaire 2.5.12. *Toute marche adaptée sur $E = SL(2, \mathbb{R})$ est transiente.*

Indications de preuve:

Il suffit de prouver que E n'est pas moyennable. Pour ceci, on considère l'action de E sur l'espace projectif $U = P(\mathbb{R}^2) = \mathbb{R} \cup \infty$ identifié au cercle unité \mathbb{T} par la transformation de Cayley $u \mapsto z = \frac{u+i}{u-i}$. On voit que la seule probabilité γ sur \mathbb{T} laissée invariante par le sous groupe des rotations $K = \left\{ \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \right\}$ est la mesure de Lebesgue puisque l'action d'une matrice de K s'identifie à la translation d'angle 2θ sur \mathbb{T} . La transformation de Cayley inverse s'écrivant $\exp(i\varphi) \mapsto \cot(\varphi/2)$, la seule mesure invariante possible sur U est la loi de Cauchy. Or la loi de Cauchy n'est pas invariante par les matrices $x = \begin{pmatrix} t & 0 \\ 0 & 1/t \end{pmatrix}$ dont l'action projective est donnée par $u \rightarrow t^2 u$. ■

2.5.3 Critères de Fourier

On se place dans le cas d'un groupe E abélien et plus précisément $E = \mathbb{Z}^d$ ou $E = \mathbb{R}^d$. Dans le premier cas il suffit d'utiliser la théorie des séries de Fourier, mais dans le cas continu on doit utiliser les transformées de Fourier qui se manipulent moins aisément...

Le cas de \mathbb{Z}^d

D'après l'étude des séries de Fourier, pour $\nu \in \mathcal{M}_1^+(\mathbb{Z}^d)$ et $t \in \mathbb{T}^d$ on pose

$$\widehat{\nu}(t) = \sum_{n \in \mathbb{Z}^d} \exp(2i\pi \langle n, t \rangle) \nu(n) \text{ et alors } \nu(n) = \int_{\mathbb{T}^d} \exp(-2i\pi \langle n, t \rangle) \widehat{\nu}(t) dt$$

Proposition 2.5.13. *Soit $\nu \in \mathcal{M}_1^+(\mathbb{Z}^d)$. Alors ν définit une marche récurrente si et seulement si:*

$$\lim_{r \uparrow 1} \int_{\mathbb{T}^d} \Re \left(\frac{1}{1 - r\widehat{\nu}(t)} \right) dt = +\infty$$

Indications de preuve:

Il suffit d'étudier la récurrence de 0 c'est à dire la valeur de $\Pi(0)$

$$\Pi(0) = \sum_{n=0}^{\infty} \mathbb{P}_0(X_n = 0) = \sum_{n=0}^{\infty} \int_{\mathbb{T}^d} \widehat{\nu}^n(t) dt$$

Il n'est pas possible d'échanger la série et l'intégrale dans cette dernière expression puisque $\widehat{\nu}(0) = 1$. On utilise alors "l'astuce" suivante: pour $0 < r < 1$

$$\begin{aligned}\Pi(0) &= \lim_{r \uparrow 1} \sum_{n=0}^{\infty} r^n \mathbb{P}_0(X_n = 0) = \lim_{r \uparrow 1} \sum_{n=0}^{\infty} \int_{\mathbb{T}^d} r^n \widehat{\nu}^n(t) dt \\ &= \lim_{r \uparrow 1} \int_{\mathbb{T}^d} \left(\frac{1}{1 - r\widehat{\nu}(t)} \right) dt = \lim_{r \uparrow 1} \int_{\mathbb{T}^d} \Re \left(\frac{1}{1 - r\widehat{\nu}(t)} \right) dt\end{aligned}$$

La dernière formule étant obtenue en prenant la partie réelle des égalités. ■

On déduit immédiatement de ce résultat, la nature des marches symétriques sur \mathbb{Z}^d c'est à dire lorsque ν attribue la masse $1/(2d)$ à chaque vecteur de base et à son opposé. Dans ce cas $\widehat{\nu}(t) = (\sum_{k=1}^d \cos(2\pi t_k))/d$ et $1 - r\widehat{\nu}(t) = (\sum_{k=1}^d (1 - r \cos(2\pi t_k)))/d$.

Corollaire 2.5.14. *Une marche aléatoire symétrique sur \mathbb{Z}^d est récurrente pour $d = 1$ ou $d = 2$ et transiente pour $d \geq 3$.*

Indications de preuve:

Le calcul de la limite de l'intégrale ne pose problème qu'au voisinage de $t = 0$. Sur un tel voisinage, $1 - r\widehat{\nu}(t)$ est positive et croissante en r , il suffit donc d'étudier l'intégrabilité de $\frac{1}{1 - \widehat{\nu}(t)}$. Or au voisinage de 0 on a $(1 - \widehat{\nu}(t)) \sim c\rho^2$ avec $\rho^2 = \sum_{k=1}^d t_k^2$. Il suffit alors de faire un passage en coordonnées polaires pour avoir le résultat puisque $dt = \rho^{d-1} d\sigma(\theta)$ où σ est la mesure de Lebesgue de la boule unité S^{d-1} . ■

Le cas de \mathbb{R}^d

Le lemme suivant résume les quelques résultats sur la transformation de Fourier qui nous seront nécessaires:

Lemme 2.5.15. *Soit $f \in L^1(\mathbb{R}^d, dx)$ et $\nu \in \mathcal{M}_1^+(\mathbb{R}^n)$.*

1. $\int \widehat{\nu}(t) f(t) dt = \int \widehat{f}(t) d\nu(t)$
2. Pour $f(x) = \frac{a - |x|}{a^2} \mathbf{1}_{\{|x| \leq a\}}$ on a $\widehat{f}(t) = \frac{2(1 - \cos(at))}{(at)^2}$
3. Pour $f(x) = \frac{1 - \cos(x/a)}{\pi x^2}$ on a $\widehat{f}(t) = \frac{1 - |at|}{a} \mathbf{1}_{\{|at| \leq 1\}}$

On a bien sûr des résultats équivalents à 2. et 3 sur \mathbb{R}^d en considérant des fonctions de type produit.

Proposition 2.5.16. *Soit $\nu \in \mathcal{M}_1^+(\mathbb{R}^d)$. Alors ν définit une marche récurrente si et seulement si pour tout voisinage V de l'origine on a:*

$$\lim_{r \uparrow 1} \int_V \Re \left(\frac{1}{1 - r\widehat{\nu}(t)} \right) dt = +\infty$$

Indications de preuve:

Soit $a > 0$, on choisit pour V la boule $B(1/a) = \cap_{k=1}^d (|x_k| \leq 1/a)$. Alors:

$$\Pi(V) = \lim_{r \uparrow 1} \sum_n r^n \mathbb{P}_e(X_n \in V) = \lim_{r \uparrow 1} \sum_n r^n \int_V d\nu^{\star n}$$

Or pour $|t| \leq \pi/3$ on a $1 - \cos(t) \geq t^2/4$ et donc:

$$\begin{aligned} \int_V d\nu^{\star n} &\leq 4^d \int_V \prod_{i=1}^d \frac{1 - \cos(ax_i)}{(ax_i)^2} d\nu^{\star n}(x) \\ &\leq 4^d \int \prod_{i=1}^d \frac{1 - \cos(ax_i)}{(ax_i)^2} d\nu^{\star n}(x) \\ &= \Re \left\{ 2^d \int_{B(a)} \widehat{\nu}^n(t) \prod_{i=1}^d \frac{a - |t_i|}{a^2} dt \right\} \end{aligned}$$

d'où:

$$\begin{aligned} \Pi(V) &\leq 2^d \lim_{r \uparrow 1} \int_{B(a)} \left(\prod_{i=1}^d \frac{a - |t_i|}{a^2} \right) \Re \left(\frac{1}{1 - r\widehat{\nu}(t)} \right) dt \\ &\leq \left(\frac{2}{a} \right)^d \lim_{r \uparrow 1} \int_{B(a)} \Re \left(\frac{1}{1 - r\widehat{\nu}(t)} \right) dt \end{aligned}$$

Pour une inégalité dans l'autre sens on peut écrire:

$$\begin{aligned} \int_V d\nu^{\star n} &\geq a^d \int \left(\prod_{i=1}^d \frac{1 - |ax_i|}{a} \mathbf{1}_{\{|ax_i| \leq 1\}} \right) d\nu^{\star n}(x) \\ &= \left(\frac{a}{\pi} \right)^d \int \left(\prod_{i=1}^d \frac{1 - \cos(t_i/a)}{t_i^2} \right) \widehat{\nu}^n(t) dt \\ \Pi(V) &\geq \left(\frac{a}{\pi} \right)^d \lim_{r \uparrow 1} \int \left(\prod_{i=1}^d \frac{1 - \cos(t_i/a)}{t_i^2} \right) \Re \left(\frac{1}{1 - r\widehat{\nu}(t)} \right) dt \\ &\geq \frac{1}{(4\pi a)^d} \lim_{r \uparrow 1} \int_{B(a)} \Re \left(\frac{1}{1 - r\widehat{\nu}(t)} \right) dt \end{aligned}$$

■

On peut montrer qu'en fait le critère est valable en faisant simplement $r = 1$ dans les formules de 2.5.13 et 2.5.16 mais la preuve est relativement compliquée. Nous nous contenterons donc de deux observations qui suffisent dans beaucoup de cas à établir la nature de la marche aléatoire.

Remarque

- On a toujours $(\Re \frac{1}{1 - r\widehat{\nu}(t)}) \geq 0$ donc le lemme de Fatou permet d'obtenir la récurrence dès que $\int_V \Re \left\{ \frac{1}{1 - \widehat{\nu}(t)} \right\} dt = \infty$ (pour \mathbb{Z}^d on remplace V par \mathbb{T}^d).
- On a $(\Re \frac{1}{1 - r\widehat{\nu}(t)}) \leq \frac{1}{1 - \Re(\widehat{\nu}(t))}$ au voisinage de zéro et donc s'il existe V avec $\int_V \frac{1}{1 - \Re(\widehat{\nu}(t))} dt < \infty$ la marche est transiente.

Nous allons terminer par une propriété générale des marches en dimension, ≥ 3 qui généralise (en partie) le corollaire 2.5.14.

Proposition 2.5.17. *Pour $d \geq 3$, toute marche sur \mathbb{R}^d (ou \mathbb{Z}^d) non portée par un hyperplan (ou un sous réseau) de dimension ≤ 2 est transiente.*

Indications de preuve:

On va traiter le cas \mathbb{R}^d celui de \mathbb{Z}^d étant identique. On peut se restreindre au cas où $G_\nu = \mathbb{R}^d$ avec $d \geq 3$. Il existe donc des vecteurs (x_1, x_2, \dots, x_d) linéairement indépendants dans le support de ν . Soit $B(L)$ une boule ouverte contenant tous ces points. Il est alors facile de constater que la forme quadratique définie par

$$\mathbf{Q}_L(y) = \int_{B(L)} \langle x, y \rangle^2 d\nu(x)$$

est définie positive et par conséquent, il existe $\alpha > 0$ avec $\mathbf{Q}_L(y) \geq \alpha \|y\|^2$. D'autre part:

$$\Re(1 - \widehat{\nu}(y)) = 2 \int \sin^2(\langle x, y \rangle / 2) d\nu(x) \geq 2 \int_{B(L)} \sin^2(\langle x, y \rangle / 2) d\nu(x)$$

Or pour $|t| \leq \pi/2$ on a $|\sin(t)| \geq 2|t|/\pi$ et par conséquent pour $\|y\| \leq \pi/L$ on a:

$$\Re(1 - \widehat{\nu}(y)) \geq \frac{2}{\pi^2} \mathbf{Q}_L(y) \geq \frac{2\alpha}{\pi^2} \|y\|^2$$

et il suffit d'utiliser la dernière remarque ci-dessus. ■

Dès que μ possède une espérance non nulle, on obtient bien sûr la transience en toute dimension en vertu de la loi des grands nombres. Il reste à examiner le cas des marches centrées ou sans moment d'ordre 1 en dimension 1 ou 2, ...

2.6 Chaînes récurrentes

On revient au cas d'une chaîne de Markov sur un espace d'états quelconque (E, \mathcal{E}) et de transition \mathbf{Q} . On introduit tout d'abord les notions d'ergodicité puis de récurrence au sens de Harris. Un des buts de cette section est de prouver que si une chaîne possède un ensemble de Doeblin attractif F , alors cette chaîne est Harris récurrente et ergodique. Si de plus $\sup_{x \in F} \mathbb{E}_x(S_F) < \infty$ alors la récurrence est positive et l'on peut alors contrôler

la vitesse de convergence vers la probabilité invariante en fonction des moments de S_F . Pour arriver à ce résultat, on étudie tout d'abord le cas de chaînes atomiques et l'on prouve ensuite que l'on peut se ramener à ce cas pour une chaîne possédant un ensemble de Doeblin attractif par le procédé de "scission". Une grande partie de ce qui suit est adapté du livre de Marie Duflo [2].

2.6.1 Chaînes ergodiques

Soit une chaîne canonique $(\Omega, \mathcal{F}_n, X_n, \mathbb{P}_x)$ de transition \mathbf{Q} . Pour $f \in \mathcal{B}(E)$ on pose $S_n(f) = \sum_{k=1}^n f(X_k)$ et si $f \geq 0$ on note $S_\infty(f) = \lim \uparrow S_n(f)$. Nous commençons par établir un lien entre propriétés ergodiques et mesures invariantes.

Définition 2.6.1. Soit ν une mesure σ -finie non nulle. On dit que la chaîne X_n est ν ergodique si:

1. Soit $B \in \mathcal{E}$ avec $\nu(B) > 0$. Alors pour tout $x \in E$, on a $\mathbb{P}_x(S_B < \infty) = 1$ (et donc $\mathbb{P}_x(N_B = \infty) = 1$).
2. La chaîne satisfait au théorème ergodique ponctuel relativement à ν c'est à dire que pour toutes fonctions f et g dans $\mathcal{B}^+(E)$, ν intégrables avec $\nu(g) \neq 0$ on a:

$$\lim_n \frac{S_n(f)}{S_n(g)} = \frac{\nu(f)}{\nu(g)}, \quad \mathbb{P}_x \text{ p.s. pour tout } x \in E$$

En fait, la condition 1. sert à assurer que les quotients figurant dans condition 2. ci dessus sont bien définis pour n assez grand car:

Lemme 2.6.2. La condition 1. de la définition 2.6.1 implique $S_\infty(g) = \infty$, \mathbb{P}_x p.s. pour toute fonction $g \in \mathcal{B}^+(E)$ avec $\nu(g) > 0$.

Indications de preuve:

Il suffit de vérifier que $\nu(g) > 0$ implique qu'il existe B avec $\nu(B) > 0$ et $\alpha > 0$ tels que $g \geq \alpha \mathbf{1}_{\{B\}}$. ■

Lemme 2.6.3. Dans le théorème ergodique, la limite des quotients $S_n(f)/S_n(g)$ est nulle si $\nu(g) = \infty$ (et $\nu(f) < \infty$). En particulier la suite $S_n(f)/n$ admet une limite nulle si ν est de masse totale infinie (cas nul) et égale à $\nu(f)$ si ν est une probabilité (cas positif).

Indications de preuve:

Soit E_p une suite croissante avec $E = \cup E_p$ et $\nu(E_p) < \infty$. Si on pose $g_p = (g \wedge p) \mathbf{1}_{\{E_p\}}$, la suite g_p est d'intégrale finie et tend en croissant vers la fonction g . Il existe un entier p_0 tel que g_p soit d'intégrale positive et on a alors pour $p \geq p_0$:

$$\overline{\lim}_n S_n(f)/S_n(g) \leq \lim_n S_n(f)/S_n(g_p) = \nu(f)/\nu(g_p)$$

Il suffit maintenant de faire tendre p vers l'infini. Pour obtenir les assertions suivantes, on considère la fonction $g = 1$. ■

Proposition 2.6.4. Soit $f \in \mathcal{B}_b^+(E)$. Alors sur l'ensemble $S_\infty(\mathbf{Q}f) = +\infty$ on a :

$$\lim_n \frac{S_n(f)}{S_n(\mathbf{Q}f)} = 1, \quad \mathbb{P}_\mu \text{ p.s. pour tout } \mu \in \mathcal{M}_1^+(E)$$

Indications de preuve :

On pose $\Sigma_n = \sum_{k=0}^{n-1} \mathbf{Q}f(X_k)$. Alors $\mathbb{E}_\mu(f(X_{n+1})|\mathcal{F}_n) = \mathbf{Q}f(X_n)$ et donc la suite $Z_n = S_n(f) - \Sigma_n$ est une \mathbb{P}_μ martingale de carré intégrable. De plus :

$$\begin{aligned} \mathbb{E}_\mu((Z_{n+1} - Z_n)^2|\mathcal{F}_n) &= \mathbb{E}_\mu((f(X_{n+1}) - \mathbf{Q}f(X_n))^2|\mathcal{F}_n) \\ &= \mathbf{Q}(f^2)(X_n) - (\mathbf{Q}f(X_n))^2 \\ &\leq \|f\|_\infty \mathbf{Q}f(X_n) \end{aligned}$$

Le processus croissant de Z_n est donc majoré par $\|f\|_\infty \Sigma_n$. Le résultat est alors une conséquence directe de la proposition 1.2.3. ■

Proposition 2.6.5. Si la chaîne est ν ergodique relativement à la mesure σ -finie ν , celle ci est invariante.

Indications de preuve :

Soit $p \geq 0$ et B avec $\nu \mathbf{Q}^{p+1}(B) > 0$. On peut appliquer 2.6.4 à la fonction $f = \mathbf{Q}^p \mathbf{1}_{\{B\}}$ en vertu de 2.6.2 et l'on obtient que le quotient $S_n(\mathbf{Q}^p \mathbf{1}_{\{B\}})/S_n(\mathbf{Q}^{p+1} \mathbf{1}_{\{B\}})$ converge vers 1. Si de plus $\nu \mathbf{Q}^p(B) < \infty$ on peut aussi appliquer alors le théorème ergodique (ou son extension 2.6.3) avec $f = \mathbf{Q}^p \mathbf{1}_{\{B\}}$ et $g = \mathbf{Q}^{p+1} \mathbf{1}_{\{B\}}$. On obtient donc :

$$(\nu \mathbf{Q}^{p+1}(B) > 0) \text{ et } (\nu \mathbf{Q}^p(B) < \infty) \implies (\nu \mathbf{Q}^p(B) = \nu \mathbf{Q}^{p+1}(B)) \quad \diamond$$

Puisque ν est σ -finie, il suffit de prouver que $\nu(B) = \nu \mathbf{Q}(B)$ pour B tel que $\nu(B) < \infty$ pour affirmer que ν est invariante. En vertu du résultat \diamond appliqué pour $p = 0$ on obtient $\nu(B) = \nu \mathbf{Q}(B)$ lorsque $\nu \mathbf{Q}(B) > 0$. Il reste à considérer le cas $\nu \mathbf{Q}(B) = 0$. On voit alors facilement, en appliquant \diamond pour $p = 1, 2, \dots$, que dans ce cas, $\nu \mathbf{Q}^p(B) = 0$ pour tout $p \geq 1$ et que donc $\mathbb{P}_x(N_B \leq 1) = 1$ pour ν presque tout x . Ceci est incompatible avec $\nu(B) > 0$ d'après la première condition de la définition de la ν ergodicité. ■

2.6.2 Chaînes récurrentes au sens de Harris

Soit $B \in \mathcal{E}$. On rappelle que T_B est le temps d'entrée dans B , S_B le temps de retour et N_B le nombre de visites à B .

Définition 2.6.6. Soit ν une mesure positive non nulle, σ -finie et invariante sur (E, \mathcal{E}) . On dit qu'une chaîne est ν récurrente au sens de Harris si :

$$(\nu(B) > 0) \implies (\mathbb{P}_x(S_B < \infty) = 1) \text{ pour tout } x \in E \text{ et donc } \mathbb{P}_x(N_B = \infty) = 1$$

On dit que la récurrence est positive si ν est bornée et nulle sinon.

Remarques

- D'après 2.6.5 on obtient qu'une chaîne ν ergodique est une chaîne récurrente au sens de Harris.
- La définition ci-dessus qui impose que ν soit invariante est peu maniable. On définit en général une chaîne de Harris par une certaine forme d'irréductibilité par rapport à une mesure positive σ -finie τ quelconque et le fait que le noyau potentiel ne soit pas "propre", mais ce point de vue ne sera pas utilisé ici (il faudrait alors prouver que cette définition implique que sur un ensemble stochastiquement clos et de complémentaire τ négligeable, il existe une mesure ν conforme à la définition 2.6.6).

Théorème 2.6.7. *Si une chaîne est ν récurrente au sens de Harris, alors:*

1. *Les fonctions surharmoniques sont ν p.s. constantes et les fonctions harmoniques bornées sont constantes.*
2. *Si ρ est une mesure surharmonique non nulle, alors $\rho(B) = 0$ implique que $\mathbb{P}_x(N_B = 0) = 1$ pour ρ presque tout x et donc $\nu(B) = 0$.*
3. *Pour tout $B \in \mathcal{E}$ on a l'alternative:*
 - (a) $\nu(B) > 0$ et alors $\mathbb{P}_x(N_B = \infty) = 1$ pour tout $x \in E$.
 - (b) $\nu(B) = 0$ et alors $\mathbb{P}_x(N_B = \infty) = 0$ pour tout $x \in E$.
4. *Toute mesure sur-harmonique σ -finie est proportionnelle à ν . La récurrence au sens de Harris est donc une propriété intrinsèque de la chaîne et ne dépend pas de la mesure qui sert à définir cette propriété.*

Indications de preuve:

- Preuve de 1.

Si f est une fonction surharmonique non ν p.s. constante, il existe deux réels $a < b$ tels que les ensembles $(f \leq a)$ et $(f \geq b)$ soient de ν mesure strictement positive. La \mathbb{P}_μ surmartingale positive $f(X_n)$ converge \mathbb{P}_μ p.s. vers $Z = \overline{\lim} f(X_n)$ et il existe une infinité d'entiers n avec $f(X_n) \leq a$ et de même avec $f(X_n) \geq b$. Par conséquent on obtient $\mathbb{P}_\mu(Z \leq a) = 1$ et $\mathbb{P}_\mu(Z \geq b) = 1$ ce qui est impossible. Il existe donc une constante c avec $\nu(f \neq c) = 0$ et par le même raisonnement on obtient que $\mathbb{P}_\mu(Z = c) = 1$. Dans le cas d'une fonction harmonique bornée, la martingale bornée $f(X_n)$ converge en moyenne et donc $f(x) = \mathbb{E}_x(f(X_n)) = \mathbb{E}_x(Z) = c$.

- Preuve de 2.

Si $\rho(B) = 0$ alors $0 = \sum_0^\infty \rho \mathbf{Q}^n(B) = \int \mathbb{E}_x(N_B) d\rho(x)$. Par conséquent $\mathbb{E}_x(N_B) = 0$ pour ρ presque tout x . Il existe donc au moins un x avec $\mathbb{P}_x(N_B = \infty) = 0$ ce qui est incompatible avec $\nu(B) > 0$.

- Preuve de 3.

La première alternative résulte de la définition d'une chaîne de Harris. L'ensemble $N_B =$

∞ est invariant par l'opérateur de translation θ et par conséquent la fonction $f(x) = \mathbb{P}_x(N_B = \infty)$ est harmonique bornée et donc constante. D'après 2. (appliqué à ν), on a $\mathbb{P}_x(N_B = \infty) = 0$ pour ν presque tout x . Cette constante est donc nulle.

• Preuve de 4.

D'après 2.3.17 une mesure surharmonique σ -finie ρ s'écrit comme la somme d'une mesure invariante et d'un potentiel de la forme $m\mathbf{G}$. Or m est nécessairement nulle car sinon $(\nu(B) > 0) \implies (m\mathbf{G}(B) = \infty)$ et donc ce potentiel ne peut être σ -fini que si E est une réunion croissante d'ensembles de ν mesure nulle ce qui impliquerait $\nu = 0$. On en déduit donc que toute mesure surharmonique σ -finie est invariante. Supposons pour l'instant que $\rho \leq \nu$. D'après 2., ρ et ν sont alors des mesures équivalentes. En posant $\rho = h\nu$ avec h strictement positive ν p.s. on voit que pour tout $c > 0$ la mesure $\alpha = (h \wedge c)\nu$ est surharmonique donc harmonique et de même pour la mesure $\beta = \rho - \alpha$ de densité $h - h \wedge c$ par rapport à ν . Les mesures positive α et β sont donc soit nulles, soit équivalentes à ν . On en déduit que ν est proportionnelle à ρ . On considère maintenant une mesure surharmonique ρ quelconque. On sait d'après 2. que $\nu = h\rho$ et donc $\rho \wedge \nu = (h \wedge 1)\rho$ est proportionnelle à ν . On en déduit l'alternative: $\rho \geq \nu$ ou bien ρ est proportionnelle à ν . Si l'on ne se trouve pas dans le second cas, on recommence le même raisonnement avec $\rho' = \rho - \nu$ et ainsi de suite. On en déduit donc que si l'on n'a jamais l'alternative ρ proportionnelle à μ , alors $\rho \geq n\nu$ pour tout n ce qui est impossible puisque ρ est σ -finie. ■

Il reste à donner des exemples de chaînes récurrentes au sens de Harris. Le cas le plus simple est celui d'une chaîne récurrente irréductible sur un espace d'états discret. Un exemple plus élaboré est fourni par certaines marches aléatoires. En effet la récurrence ordinaire signifie que cette marche passe une infinité de fois dans tout ouvert non vide. Il s'agit alors de passer des ouverts non vides aux boréliens de mesure de Haar positive.

Pour ne pas alourdir l'exposé on se place sur \mathbb{R} mais on pourrait traiter de la même façon n'importe quel groupe L.C.D. en remplaçant la mesure de Lebesgue ℓ par une mesure de Haar.

Définition 2.6.8. Soit μ une probabilité sur \mathbb{R} . On pose $\mu^{*n} = a_n + b_n$ où a_n est la partie absolument continue de μ^{*n} et b_n sa partie orthogonale par rapport à la mesure de Lebesgue ℓ . On dit que μ est étalée s'il existe $r \geq 1$ avec $a_r \neq 0$.

Lemme 2.6.9. Soit $\mu \in \mathcal{M}_1^+(\mathbb{R})$, on note $\check{\mu}$ la mesure "symétrique" par rapport à l'origine.

1. Si μ est étalée alors le support de $\mu^{*r} \star \check{\mu}^{*r}$ contient un ouvert et donc $G_\mu = \mathbb{R}$.
2. On a $b_{nr} \leq b_r^{*n}$ et par conséquent, si μ est étalée, la masse totale de b_{nr} tend vers 0 lorsque $n \rightarrow \infty$.

Indications de preuve:

• Preuve de 1.

On a $\mu^{*r} \star \check{\mu}^{*r} \geq a_r \star \check{a}_r$, or a_r possède une densité non nulle soit φ_r . Si cette densité est bornée alors la convolution L^1, L^∞ fournit une fonction continue et $\varphi_r \star \check{\varphi}_r(0) = \int \varphi_r^2(y) d(y) > 0$ ce qui prouve que le support de $\mu^{*r} \star \check{\mu}^{*r}$ contient un voisinage de 0. Si la densité φ_r n'est pas bornée on la remplace par $\varphi_r \wedge c$ en choisissant c de telle façon que cette fonction soit d'intégrale strictement positive. Un sous groupe fermé de \mathbb{R} contenant un ouvert est égal à \mathbb{R} et donc $G_\mu = \mathbb{R}$.

• Preuve de 2.

La convolution de deux mesures bornées dont l'une est absolument continue est absolument continue ce qui implique $b_{nr} \leq b_r^{*n}$. Or $b_r(1) < 1$ et donc $b_r^{*n}(1) = (b_r(1))^n$ tend vers 0. ■

Proposition 2.6.10. *On suppose que la marche (X_n) de loi μ est récurrente au sens de Harris. Alors la probabilité μ est étalée et la marche (X_n) est donc récurrente et adaptée.*

Indications de preuve:

Si μ n'est pas étalée, alors pour tout n la probabilité μ^{*n} est orthogonale à ℓ et il existe donc une suite de boréliens B_n avec $\mu^{*n}(B_n) = 1$ et $\ell(B_n) = 0$. En posant $B = \cup_n B_n$ on a $\mu^{*n}(B) = 1$ et $\ell(B) = 0$. Soit S le complémentaire de B , on a alors $\ell(S) > 0$ et $\mathbb{E}_0(N_S) = \sum \mu^{*n}(S) = 0$ ce qui est contradictoire avec la propriété de Harris. ■

Proposition 2.6.11. *Si la probabilité μ est étalée et si la marche (X_n) est récurrente alors (X_n) est récurrente au sens de Harris.*

Indications de preuve:

• Soit U_n une base dénombrable d'ouverts de \mathbb{R} telle que tout ouvert soit une réunion d'éléments de cette famille. On pose $\Omega_n = (N_{U_n} = \infty)$ et $\Omega' = \cap \Omega_n$. Alors $\mathbb{P}_0(\Omega' = 1)$ et si U est un ouvert non vide et $\omega \in \Omega'$ on a $N_U(\omega) = \infty$.

• Soit B un borélien tel que $0 < \ell(B) < +\infty$. On pose $C_n(\omega) = \{x \in \mathbb{R}; N_{B+x}(\omega) \leq n\}$, $C(\omega) = \bigcup_n C_n(\omega)$. Soit $\omega \in \Omega'$, K un compact tel que $K \subset C_n(\omega)$ et $\varphi = \mathbf{1}_{\{K\}} \star \mathbf{1}_{\{B\}}$. La fonction φ est continue et l'on peut écrire:

$$\int_K N_{B+x} d\ell(x) = \sum_n \int \mathbf{1}_{\{K\}}(x) \mathbf{1}_{\{B\}}(X_n - x) d\ell(x) = \sum_n \varphi(X_n)$$

Si φ n'est pas nulle, il existe un ouvert U et $\alpha > 0$ avec $\varphi \geq \alpha \mathbf{1}_{\{U\}}$ et:

$$\infty = \alpha N_U \leq \int_K N_{B+x} d\ell(x) \leq n\ell(K)$$

ce qui est impossible, par conséquent φ est nulle et:

$$0 = \int \varphi(x) d\ell(x) = \ell(B)\ell(K)$$

On en déduit $\ell(K) = 0$ et donc $\ell(C(\omega)) = \ell\{x \in \mathbb{R}; N_{B+x}(\omega) < \infty\} = 0$.

• On pose $W = \{(\omega, x) ; N_{B+x}(\omega) < \infty\}$. C'est un ensemble mesurable pour la tribu produit $\mathcal{F} \otimes \mathcal{B}$ car $N_{B+x}(\omega) = \sum_n \mathbf{1}_{\{B\}}(X_n(\omega) + x)$ et la fonction $(\omega, x) \rightarrow X_n(\omega) + x$

est mesurable en ω et continue en x . Toutes les sections de W à ω fixé sont de mesure de Lebesgue nulle. En utilisant le théorème de Fubini, il existe un borélien Γ de mesure de Lebesgue nulle tel que si $x \notin \Gamma$ alors $\mathbb{P}_0(N_{B+x} = +\infty) = 1 = \mathbb{P}_x(N_B = +\infty)$.

- On veut maintenant éliminer l'ensemble exceptionnel Γ . Pour ceci on écrit:

$$\mathbb{P}_x(X_{nr} \in \Gamma) = \mathbb{P}_0(X_{nr} \in \Gamma - x) = \mu^{\star nr}(\Gamma - x) = b_{nr}(\Gamma - x)$$

car $(\Gamma - x)$ est de mesure de Lebesgue nulle. Et donc $\lim_{n \rightarrow \infty} \mathbb{P}_x(X_{nr} \in \Gamma) = 0$. Par Markov on obtient $\mathbb{P}_x(N_B = +\infty) = \mathbb{E}_x[\mathbb{P}_{X_{nr}}(N_B = +\infty)]$ et l'on peut écrire:

$$\mathbb{P}_x(N_B = \infty) \geq \mathbb{E}_x(\mathbb{P}_{X_{nr}}(N_B = \infty) \mathbf{1}_{\{X_{nr} \notin \Gamma\}}) = \mathbb{P}_x(X_{nr} \notin \Gamma)$$

et donc $\mathbb{P}_x(N_B = \infty) \geq \lim_n \mathbb{P}_x(X_{nr} \notin \Gamma) = 1$. On a donc prouvé que $\mathbb{P}_x(N_B = \infty) = 1$ si $0 < \ell(B) < +\infty$ et c'est à fortiori vrai dans le cas $\ell(B) = \infty$. ■

2.6.3 Théorie ergodique des chaînes atomiques

Définition 2.6.12. Soit F un sous ensemble mesurable de E .

- On dit que F est "attractif" si tout point de E conduit p.s. à F c'est à dire $\mathbb{P}_x(S_F < \infty) = 1$ pour tout $x \in E$.
- On dit que F est un "atome" si $\mathbf{Q}(a, \bullet)$ ne dépend pas du choix de $a \in F$.

Dans la suite, si F est un atome, " a " désignera toujours un point quelconque de F .

Z. Pour a et b dans un atome F on ne peut affirmer que $\mathbb{P}_a = \mathbb{P}_b$. Par contre on a $\mathbb{E}_a(\Psi) = \mathbb{E}_b(\Psi)$ pour une fonction mesurable positive sur l'espace canonique de la forme $\Psi(\omega) = f(X_0)\Phi(\omega)$, lorsque la fonction Φ ne dépend pas de ω_0 et que f est constante sur F . Ceci est le cas de la fonction N_F , par conséquent un atome est dit transient si $G(a, F) < \infty$ et récurrent sinon.

Théorème 2.6.13. Soit X_n une chaîne possédant un atome attractif F . Alors cette chaîne est récurrente au sens de Harris et vérifie le théorème ergodique. Une version de sa mesure invariante est donnée par:

$$\nu(f) = \mathbb{E}_a\left(\sum_{1 \leq n \leq S_F} f(X_n)\right)$$

La récurrence est positive si et seulement si $\mathbb{E}_a(S_F) < \infty$.

Indications de preuve:

Pour alléger les notations on posera $S_F^{(n)} = S^{(n)}$ et l'hypothèse faite implique que tous ces temps d'arrêt sont finis \mathbb{P}_x p.s. pour tout $x \in E$. De même on pose $T = T_F$.

- ν est σ -finie:

Soit X_n'' la chaîne tuée après passage dans F (voir 2.3.9). Pour cette chaîne, l'ensemble

F est un atome transient et tout point de E conduit à F donc d'après 2.3.13 l'ensemble E s'écrit comme une réunion dénombrable d'ensembles E_k par lesquels on a $\mathbf{G}''(x, E_k) = \mathbb{E}_x(\sum_{0 \leq n \leq TS} \mathbf{1}_{\{E_k\}}(X_n)) \leq C_k < \infty$. Or pour un ensemble $A \subset E$ on a :

$$\begin{aligned} \nu(A) &= \mathbb{E}_a(\sum_{1 \leq n \leq S} \mathbf{1}_{\{A\}} \circ X_n) = \mathbb{E}_a(\sum_{1 \leq n \leq 1+T \circ \theta} \mathbf{1}_{\{A\}} \circ X_n) \\ &= \mathbb{E}_a(\{\sum_{0 \leq n \leq T} \mathbf{1}_{\{A\}} \circ X_n\} \circ \theta) = \mathbb{E}_a(\mathbb{E}_{X_1}(\sum_{0 \leq n \leq T} \mathbf{1}_{\{A\}} \circ X_n)) \end{aligned}$$

On en déduit donc que $\nu(E_k) \leq C_k < \infty$.

• Récurrence:

Pour $f \in \mathcal{B}^+(E)$ on pose $Z_1 = \sum_{1 \leq n \leq S_F} f(X_n)$ puis $Z_n = Z_{n-1} \circ \theta_S$. Alors $Z_n = Z_1 \circ \theta_{S(n-1)}$ et on remarque que, sous \mathbb{P}_a , la loi des variables aléatoires Z_k , $k \geq 1$, est indépendante du choix de $a \in F$ et que $\nu(f) = \mathbb{E}_a(Z_1)$. Une simple application de Markov fort montre que pour tout système h_1, \dots, h_n de fonctions boréliennes bornées on a :

$$\begin{aligned} \mathbb{E}_\mu(\prod_{k=1}^n h_k(Z_k)) &= \mathbb{E}_\mu\left(\mathbb{E}_{X_{S(n-1)}}(h_n(Z_1)) \prod_{k=1}^{n-1} h_k(Z_k)\right) \\ &= \mathbb{E}_a(h_n(Z_1)) \mathbb{E}_\mu(\prod_{k=1}^{n-1} h_k(Z_k)) \\ &= \mathbb{E}_\mu(h_1(Z_1)) \prod_{k=2}^n \mathbb{E}_a(h_k(Z_1)) \end{aligned}$$

Par conséquent sous chaque probabilité \mathbb{P}_μ , les variables aléatoires Z_k , $k \geq 2$ sont indépendantes et de même loi donnée par celle de Z_1 sous \mathbb{P}_a . Si on suppose que f est ν intégrable, l'application de la loi forte des grands nombres fournit :

$$\lim_n \frac{1}{n} \sum_{1 \leq k \leq S(n)} f(X_k) = \nu(f), \quad \mathbb{P}_\mu \text{ p.s. pour tout } \mu \in \mathcal{M}_1^+(E)$$

Par conséquent, si $\nu(f) > 0$ on a $S_\infty(f) = \infty$, \mathbb{P}_μ p.s. pour tout $\mu \in \mathcal{M}_1^+(E)$ et la chaîne sera donc récurrente au sens de Harris dès que l'on aura prouvé l'invariance de ν . Or cette chaîne vérifie le théorème ergodique, car pour un entier m avec $S^{(n)} \leq m < S^{(n+1)}$ on a $S_m(\mathbf{1}_{\{F\}}) = n$ et :

$$\frac{1}{n} \sum_{1 \leq k \leq S(n)} f(X_k) \leq \frac{S_m(f)}{S_m(\mathbf{1}_{\{F\}})} \leq \frac{1}{n} \sum_{1 \leq k \leq S^{(n+1)}} f(X_k)$$

et par conséquent on obtient la convergence des quotients $S_n(f)/S_n(g)$ et donc l'invariance de ν d'après la proposition 2.6.5. ■

Remarque

On note \mathcal{G} la sous tribu de \mathcal{E} définie par :

$$\mathcal{G} = \{B \in \mathcal{E} ; B \cap F = \emptyset \text{ ou bien } F \subset B\}$$

alors la restriction de ν à (E, \mathcal{G}) est donnée par:

$$\nu_F(f) = \mathbb{E}_a\left(\sum_{0 \leq n < S_F} f(X_n)\right)$$

En effet soit $f \in \mathcal{B}^+(E)$, constante sur F (c'est à dire \mathcal{G} mesurable), on peut écrire $\mathbb{E}_a(f(X_{S_F})) + \nu_F(f) = \nu(f) + f(a)$ et on a bien sûr $f(X_{S_F}) = f(a)$. Par ailleurs on remarque que \mathbf{Q} est une probabilité de transition sur (E, \mathcal{G}) . L'invariance de ν_F sur (E, \mathcal{G}) peut alors s'obtenir très facilement puisque pour $f \in \mathcal{B}^+(E)$ et constante sur F on a:

$$\begin{aligned} \nu_F \mathbf{Q}(f) &= \mathbb{E}_a\left(\sum_{n \geq 0} \mathbf{Q}f(X_n) \mathbf{1}_{\{S > n\}}\right) \\ &= \mathbb{E}_a\left(\sum_{n \geq 0} \mathbb{E}_a(f(X_{n+1}) \mathbf{1}_{\{S > n\}} | \mathcal{F}_n)\right) \\ &= \mathbb{E}_a\left(\sum_{n \geq 0} f(X_{n+1}) \mathbf{1}_{\{S > n\}}\right) \\ &= \nu_F(f) - f(a) + \mathbb{E}_a(f(X_S)) = \nu_F(f) \end{aligned}$$

On va maintenant passer au résultat essentiel de cette section, à savoir l'estimation de la vitesse de convergence de la loi de X_n vers la probabilité invariante dans le cas récurrent positif. Celui-ci (2.6.18) résultera d'une succession de lemmes "techniques" démontrés ci dessous.

Lemme 2.6.14. *Soit X_n une chaîne de Markov, $F \subset E$ et k un entier ≥ 1 . Alors:*

1. $\left(\mathbb{E}_\mu(S_F^{(n)})^k\right)^{1/k} \leq \left(\mathbb{E}_\mu((S_F)^k)\right)^{1/k} + (n-1) \sup_{a \in F} \left(\mathbb{E}_a((S_F)^k)\right)^{1/k}$
2. $\mathbb{E}_\mu(\mathbb{E}_{X_n}((S_F)^k)) \leq (n+1)^{k-1}(\mathbb{E}_\mu((S_F)^k)) + n \sup_{a \in F} \left(\mathbb{E}_a((S_F)^k)\right)^{1/k}$
3. *On suppose que la chaîne X_n possède un atome attractif F et que sa mesure invariante est de masse totale finie. Soit ν la probabilité invariante, alors:*

$$(\mathbb{E}_\nu((S_F)^k) < \infty) \iff (\mathbb{E}_a((S_F)^{k+1}) < \infty)$$

Indications de preuve:

La première inégalité se prouve par récurrence en écrivant que $S_F^{(n+1)} = S_F^{(n)} + S \circ \theta_{S_F^{(n)}}$ et l'on applique la sous-additivité de la norme L^k . Pour la seconde, on remarque que $S_F \circ \theta_n \leq S_F^{(n+1)}$ et l'on utilise la concavité de $x \rightarrow x^{1/k}$ dans la première relation. En utilisant les notations de 2.6.13 on désigne par C une constante de proportionnalité entre

ν et ν_F et en utilisant le fait que $S_F \circ \theta_n = S_F - n$ sur $(S_F > n)$ on peut écrire:

$$\begin{aligned}\mathbb{E}_\nu((S_F)^k) &= \int \mathbb{E}_x((S_F)^k) d\nu(x) = C \int \mathbb{E}_x((S_F)^k) d\nu_F(x) \\ &= C\mathbb{E}_a\left(\sum_{n=0}^{\infty} \mathbf{1}_{\{S_F > n\}} \mathbb{E}_{X_n}(S_F)^k\right) = C\mathbb{E}_a\left(\sum_{n=0}^{\infty} (\mathbf{1}_{\{S_F > n\}}(S_F - n)^k)\right) \\ &= C\mathbb{E}_a\left(\sum_{1 \leq n \leq S_F} n^k\right) = C\mathbb{E}_a(\mathcal{P}(S_F))\end{aligned}$$

où \mathcal{P} est un polynôme de degré exactement égal à $k + 1$. ■

On va maintenant introduire la notion fondamentale de couplage, qui va permettre d'établir les théorèmes de convergence vers la mesure invariante.

Définition 2.6.15. Soit X_n une chaîne de Markov. On appelle chaîne couplée, deux copies indépendantes X_n et X'_n du processus X_n .

Sur l'espace canonique produit associé au couple (X_n, X'_n) , la mesure de probabilité $\mathbb{P}_{\alpha \otimes \beta}$ que l'on notera aussi $\mathbb{P}_{(\alpha, \beta)}$ est égale au produit $\mathbb{P}_\alpha \otimes \mathbb{P}_\beta$.

Lemme 2.6.16. Soit Y_n et Y'_m des variables aléatoires indépendantes et de même loi μ à valeurs dans \mathbb{Z} et possédant une espérance finie. On pose $U_n = \sum_{k=1}^n Y_k$ et $U'_n = \sum_{k=1}^n Y'_k$. Si l'on suppose que $\mu(1) > 0$ alors pour tout entier p on a:

$$\mathbb{P}(\cup_{(m,n) \geq 1} (U_n - U'_m = p)) = 1$$

Indications de preuve:

On note $Z_n = U_n - U'_n$ la marche de loi $\nu = \mu \star \check{\mu}$ et $H = d\mathbb{Z}$ le sous groupe engendré par le support de ν . La marche Z_n est centrée et intégrable et récurre donc dans H . Si $p \in H$ (ce qui sera toujours le cas lorsque $d = 1$) la preuve est donc terminée. Sinon, on écrit $p = kd + r$ avec $1 \leq r \leq d - 1$. Soit A le support de μ et A_r l'ensemble des sommes de r éléments choisis dans A . Alors pour tout $x \in A_r$, il existe un élément $h_x \in H$ tel que $p = h_x + x$. En effet, pour tout élément $a \in A$ on a $a - 1 \in H$ et donc $a - 1$ est un multiple de d . On en déduit que si $x = a_1 + a_2 + \dots + a_r$ alors $x = r + k'd = p - (k - k')d$. Soit $R(y)$ le temps d'entrée de la marche Z_n dans le point y . Lorsque $y \in H$ on sait que $\mathbb{P}(R(y) < \infty) = 1$. Pour $x \in A_r$, on pose $A_{n,x} = (R(h_x) = n)$ et $B_{n,x} = (Y_{n+1} + \dots + Y_{n+r} = x)$. Ces deux événements sont indépendants et donc

$$\mathbb{P}(\cup_n (A_{n,x} \cap B_{n,x})) = \sum_n \mathbb{P}(A_{n,x}) \mathbb{P}(B_{n,x}) = \mathbb{P}(B_{0,x})$$

On en déduit $\mathbb{P}(\cup_{n,x} (A_{n,x} \cap B_{n,x})) = 1$ et cet ensemble est bien sûr contenu dans $\cup_{(m,n) \geq 1} (U_n - U'_m = p)$. En fait ce lemme est toujours valable si l'on suppose seulement que la loi μ est adaptée sur \mathbb{Z} .

Le lemme suivant récapitule les résultats qui nous seront nécessaires.

Lemme 2.6.17. Soit X_n une chaîne possédant un atome attractif F (elle est donc récurrente au sens de Harris). On suppose que $\mathbb{E}_a(S_F) < \infty$, c'est à dire que la récurrence est positive, et que $\mathbf{Q}(a, F) > 0$. Alors:

1. L'ensemble $(F \times F)$ est un atome attractif pour la chaîne couplée.
2. Soit $R^{(n)}$ la suite des temps de retour du couple dans $(F \times F)$, $S^{(n)}$ la suite des temps de retour de la chaîne initiale dans F et k un entier ≥ 1 . Alors:

(a) Pour α et $\beta \in \mathcal{M}_1^+(E)$, il existe une constante finie C_1 avec:

$$\mathbb{E}_{(\alpha, \beta)}(R^k) \leq C_1(\mathbb{E}_\alpha(S^k))(\mathbb{E}_{(a, \beta)}(R^k) + \mathbb{E}_{(a, a)}(R^k))$$

(b) Si $\mathbb{E}_a(S^k) < \infty$, il existe une constante finie C_2 avec:

$$\mathbb{E}_{(x, a)}(R^k) \leq C_2 \mathbb{E}_x(S^k), \quad \text{pour } a \in F$$

(c) Si $\mathbb{E}_a(S^k) < \infty$, il existe une constante finie C_3 avec:

$$\mathbb{E}_{(\alpha, \beta)}(R^k) \leq C_3(\mathbb{E}_\alpha(S^k))(\mathbb{E}_\beta(S^k))$$

pour tout couple de probabilités α et β .

Indications de preuve:

• Preuve de 1.

On pose $Y_n = S^{(n)} - S^{(n-1)}$, alors pour toute loi de départ $\alpha \otimes \beta$, la suite Y_n est une suite de variables aléatoires indépendantes et, pour $n \geq 2$, elles ont même loi μ qui est la loi de S partant d'un point de F . Cette loi possède une espérance finie et vérifie $\mu(1) > 0$. Si l'on pose $U_n = \sum_{k=2}^n Y_k$ et $U'_n = \sum_{k=2}^n Y'_k$, on a:

$$\cup_{(m, n) \geq 2} (S^{(n)} = S'^{(m)}) = \sum_{k \in \mathbb{Z}} ((Y_1 - Y'_1 = k) \cap (\cup_{(m, n) \geq 2} (U_n - U'_m = -k))) \subset (R < \infty)$$

En utilisant le lemme 2.6.16 on a $\mathbb{P}(\cup_{(m, n) \geq 2} (U_n - U'_m = -k)) = 1$ et on en déduit donc que R est presque sûrement fini pour toute loi de départ. Comme on l'a dit dans ce lemme, la condition $\mathbf{Q}(a, F) > 0$ n'est pas vraiment nécessaire, il suffirait d'éliminer le phénomène de périodicité, c'est à dire le cas où la loi de S , partant d'un point de F , serait portée par un ensemble du type $d\mathbb{N}$ avec $d > 1$.

• Preuve de 2.a.

D'après le lemme 2.6.14 on a:

$$\mathbb{E}_{(\alpha, \beta)} \left(\mathbb{E}_{(X_n, X'_n)}(R^k) \right) \leq (n+1)^{k-1} (\mathbb{E}_{(\alpha, \beta)}(R^k)) + n \mathbb{E}_{(a, a)}(R^k)$$

En utilisant l'inégalité $R \leq S + R \circ \theta_S$ on obtient:

$$\begin{aligned}
\mathbb{E}_{(\alpha,\beta)}(R^k) &\leq 2^{k-1} \left(\mathbb{E}_\alpha(S^k) + \mathbb{E}_{(\alpha,\beta)}(R^k \circ \theta_S) \right) \\
\mathbb{E}_{(\alpha,\beta)}(R^k \circ \theta_S) &= \sum_{n=1}^{\infty} \mathbb{E}_{(\alpha,\beta)} \left(\mathbf{1}_{\{S=n\}} \mathbb{E}_{(a,X'_n)}(R^k) \right) \\
&= \sum_{n=1}^{\infty} \mathbb{P}_\alpha(S=n) \mathbb{E}_{(a,\beta)} \left(\mathbb{E}_{(a,X'_n)}(R^k) \right) \\
&\leq \sum_{n=1}^{\infty} \mathbb{P}_\alpha(S=n) \rho_n^{-1} \mathbb{E}_{(a,\beta)} \left(\mathbb{E}_{(X_n,X'_n)}(R^k) \right) \\
&\leq \sum_{n=1}^{\infty} \mathbb{P}_\alpha(S=n) \rho_n^{-1} \left((n+1)^{k-1} \mathbb{E}_{(a,\beta)}(R^k) + (n+1)^k \mathbb{E}_{(a,a)}(R^k) \right)
\end{aligned}$$

avec $\rho_n = \mathbf{Q}^n(a, F)$. Or $\rho_n > 0$ et cette suite tend vers $\nu(F) > 0$, on a donc le résultat.

• Preuve de 2.b.

Puisque $\mathbb{E}_a(S) < \infty$, la chaîne X_n admet une probabilité invariante ν . La chaîne (X_n, X'_n) admet donc la probabilité invariante $\nu \otimes \nu$ et puisqu'elle récurrente positive, on a $\mathbb{E}_{(a,a)}(R) < \infty$. La majoration souhaitée est donc démontrée dans le cas $k = 1$ en utilisant 2.a avec $\alpha = \delta_x$ et $\beta = \delta_a$. Il reste maintenant à faire une preuve par récurrence sur k . Supposons que l'on ait la majoration à l'ordre k c'est à dire $\mathbb{E}_{(x,a)}(R^k) \leq C(\mathbb{E}_x(S^k))$ et de plus $\mathbb{E}_a(S^{k+1}) < \infty$. D'après 2.6.14 on a $\mathbb{E}_\nu(S^k) < \infty$ et en appliquant 2.a. avec $\alpha = \nu$ et $\beta = \delta_a$ on obtient $\mathbb{E}_{(\nu,a)}(R^k) = \mathbb{E}_{(a,\nu)}(R^k) < \infty$ puisque $\mathbb{E}_{(a,a)}(R^k)$ est fini d'après l'hypothèse de récurrence. En appliquant de nouveau 2.a. avec $\alpha = \beta = \nu$ on obtient que $\mathbb{E}_{(\nu,\nu)}(R^k)$ est fini et donc $\mathbb{E}_{(a,a)}(R^{k+1}) < \infty$. Une nouvelle application de 2.a. à l'ordre $(k+1)$ avec $\alpha = \delta_x$ et $\beta = \delta_a$ fournit la majoration souhaitée à l'ordre $k+1$.

• Preuve de 2.c.

On désigne par C diverses constantes. . . En utilisant 2.a. et 2.b., on montre que $\mathbb{E}_{(\alpha,a)}(R^k) \leq C\mathbb{E}_\alpha(S^k)$, de même $\mathbb{E}_{(a,\beta)}(R^k) = \mathbb{E}_{(\beta,a)}(R^k) \leq C\mathbb{E}_\beta(S^k)$ et enfin la majoration de $\mathbb{E}_{(\alpha,\beta)}(R^k)$. ■

Le théorème suivant résume les principaux résultats sur les vitesses de convergence pour les chaînes atomiques.

Théorème 2.6.18. *Soit X_n une chaîne possédant un atome attractif F (elle est donc récurrente au sens de Harris). On suppose que $\mathbb{E}_a(S_F) < \infty$, c'est à dire que la récurrence est positive, et que $\mathbf{Q}(a, F) > 0$. Alors:*

1. Pour tout couple $\alpha, \beta \in \mathcal{M}_1^+(E)$, on a

$$\|\alpha \mathbf{Q}^n - \beta \mathbf{Q}^n\| \rightarrow 0$$

2. Si ν désigne la probabilité invariante on a

$$\|\mathbf{Q}^n(x, \bullet) - \nu\| \rightarrow 0$$

pour tout $x \in E$.

3. Soit $k \geq 1$. On suppose que $\mathbb{E}_a((S_F)^k) < \infty$ et que pour des probabilités α, β on a $\mathbb{E}_\alpha((S_F)^k) < \infty$ et $\mathbb{E}_\beta((S_F)^k) < \infty$. Alors:

$$n^k \|\alpha \mathbf{Q}^n - \beta \mathbf{Q}^n\| \rightarrow 0, \quad \sum_n n^{k-1} \|\alpha \mathbf{Q}^n - \beta \mathbf{Q}^n\| < \infty$$

4. Soit $k \geq 1$. On suppose que $\mathbb{E}_a((S_F)^{k+1}) < \infty$. Alors il existe une constante $C < \infty$ tel que

$$n^k \|\mathbf{Q}^n(x, \bullet) - \nu\| \leq C \mathbb{E}_x((S_F)^k)$$

pour tout $x \in E$.

Indications de preuve:

On reprend la méthode du “couplage”.

Soit $f \in \mathcal{B}_b(E)$ et $\alpha \in \mathcal{M}_1^+(E)$:

$$\begin{aligned} \alpha \mathbf{Q}^n(f) &= \mathbb{E}_{(\alpha, \beta)} f(X_n) \\ &= \mathbb{E}_{(\alpha, \beta)} (f(X_n) \mathbf{1}_{\{R > n\}}) + \sum_{k=1}^n \mathbb{E}_{(\alpha, \beta)} (f(X_n) \mathbf{1}_{\{R=k\}}) \\ &= \mathbb{E}_{(\alpha, \beta)} (f(X_n) \mathbf{1}_{\{R > n\}}) + \sum_{k=1}^n \mathbb{E}_{(\alpha, \beta)} (\mathbf{1}_{\{R=k\}}) \mathbb{E}_{(a, a)} f(X_{n-k}) \\ &= \mathbb{E}_{(\alpha, \beta)} (f(X_n) \mathbf{1}_{\{R > n\}}) + \sum_{k=1}^n \mathbf{Q}^{n-k} f(a) \mathbb{P}_{(\alpha, \beta)}(R = k) \end{aligned}$$

La troisième ligne est obtenue en écrivant $X_n = X_{n-k} \circ \theta_k$. On en déduit que $|\alpha \mathbf{Q}^n(f) - \beta \mathbf{Q}^n(f)| \leq 2\|f\| \mathbb{P}_{(\alpha, \beta)}(R > n)$ et donc l'inégalité fondamentale:

$$\|\alpha \mathbf{Q}^n - \beta \mathbf{Q}^n\| \leq 2 \mathbb{P}_{(\alpha, \beta)}(R > n)$$

ce qui prouve 1. et 2..

• Preuve de 3.

Le résultat est immédiat car d'après le lemme 2.6.17 alinéa 2.c. la variable aléatoire R^k est $\mathbb{P}_{(\alpha, \beta)}$ intégrable.

• Preuve de 4.

Il suffit de prouver que $\mathbb{E}_{(x, \nu)}(R^k) \leq C \mathbb{E}_x(S^k)$. Or d'après 2.6.14 alinéa 3. on a $\mathbb{E}_\nu(S^k) < \infty$ et il suffit d'appliquer 2.6.17 alinéa 2.c.. ■

Ce théorème s'applique bien sûr aux chaînes récurrentes irréductibles sur un espace d'états dénombrable. En dehors de ce cas, il existe assez peu de chaînes possédant des atomes. La construction des chaînes scindées va permettre d'exhiber une large classe de chaînes atomiques.

2.6.4 Chaînes scindées

On considère deux espaces mesurables (E, \mathcal{E}) et (K, \mathcal{K}) ainsi que les probabilités de transition suivantes:

- $\mathbf{D}(x, du)$ définie sur $E \times \mathcal{K}$.
- $\mathbf{R}(x, u, dy)$ définie sur $E \times K \times \mathcal{E}$.

A partir de ces probabilités de transition on construit:

- $\mathbf{\Pi}(x, dy)$ définie sur $E \times \mathcal{E}$ par la formule:

$$\mathbf{\Pi}f(x) = \int f(y)\mathbf{R}(x, u, dy)\mathbf{D}(x, du)$$

- $\overline{\mathbf{\Pi}}(x, u, dy, dv)$ définie sur $E \times K \times \mathcal{E} \times \mathcal{K}$ par la formule:

$$\overline{\mathbf{\Pi}}f(x, u) = \int f(y, v)\mathbf{R}(x, u, dy)\mathbf{D}(y, dv)$$

- Pour $\mu \in \mathcal{M}_1^+(E)$, on définit la probabilité $\mu \otimes \mathbf{D}$ sur $E \times K$ par la formule:

$$\mu \otimes \mathbf{D}(f) = \int f(x, u)\mathbf{D}(x, du) d\mu(x)$$

On dit que $\mu \otimes \mathbf{D}$ est la mesure “scindée” de μ car $\mu \otimes \mathbf{D}(B, K) = \mu(B)$. On peut alors remarquer que le noyau $\overline{\mathbf{\Pi}}$ résulte de la scission de la probabilité $\mathbf{R}(x, u, \bullet)$.

Définition 2.6.19. On considère les espaces produits $\overline{\Omega} = (E \times K)^{\mathbb{N}}$ et $\Omega = E^{\mathbb{N}}$. On note:

- $\overline{\mathbb{P}}_{\theta}$ la probabilité sur $\overline{\Omega}$ correspondant à la loi initiale $\theta \in \mathcal{M}_1^+(E \times K)$ et la probabilité de transition $\overline{\mathbf{\Pi}}$. On appelle chaîne “scindée” la chaîne canonique (X_n, U_n) sur cet espace.
- \mathbb{P}_{μ} la probabilité sur Ω correspondant à la loi initiale $\mu \in \mathcal{M}_1^+(E)$ et la probabilité de transition $\mathbf{\Pi}$.

Proposition 2.6.20. Soit $\mu \in \mathcal{M}_1^+(E)$. On a alors les propriétés suivantes:

- Sous $\overline{\mathbb{P}}_{\mu \otimes \mathbf{D}}$, la suite X_n est une chaîne de Markov de loi initiale μ et de transition $\mathbf{\Pi}$.
- $\overline{\mathbb{E}}_{\mu \otimes \mathbf{D}}(f(X_n)g(U_n)) = \mathbb{E}_{\mu}(f(X_n)\mathbf{D}g(X_n))$ et donc $(\mu \otimes \mathbf{D})\overline{\mathbf{\Pi}}^n = (\mu\mathbf{\Pi}^n) \otimes \mathbf{D}$.

Indications de preuve:

Soit $(f_0, f_1, \dots, f_n) \in \mathcal{B}^+(E)$ et $g_n \in \mathcal{B}^+(K)$.

$$\begin{aligned}
& \mathbb{P}_{\mu \otimes \mathbf{D}}^{(n)}((\otimes_0^n f_i) \otimes g_n) \\
&= \int_{\substack{u_0, u_1, \dots, u_n \\ x_0, x_1, \dots, x_n}} \left(\left(\prod_0^n f_i(x_i) \right) g_n(u_n) \left(\prod_{k=0}^{n-1} \bar{\Pi}(x_i, u_i; dx_{i+1}, du_{i+1}) \right) \mathbf{D}(x_0, du_0) d\mu(x_0) \right) \\
&= \int_{x_0, x_1, \dots, x_n} \left(\left(\prod_0^n f_i(x_i) \right) \mathbf{D}g_n(x_n) \left(\prod_{k=0}^{n-1} \Pi(x_i; dx_{i+1}) \right) d\mu(x_0) \right) \\
&= \mathbb{P}_{\mu}^{(n)}((\otimes_0^{n-1} f_i) \otimes (f_n \mathbf{D}g_n))
\end{aligned}$$

■

Nous allons maintenant nous restreindre à un cas particulier important de construction de chaîne scindée. On se place dans le cas d'un espace K à deux éléments i.e. $K = \{0, 1\}$ et l'on fixe $F \in \mathcal{E}$.

Définition 2.6.21. Soit \mathbf{P} une probabilité de transition sur $(E \times \mathcal{E})$, m une probabilité sur E portée par F et p un nombre $0 < p < 1$. On effectue alors la construction de la proposition 2.6.20 avec:

$$\mathbf{D}(y, \bullet) = \begin{cases} \delta_0 & \text{si } y \notin F \\ p\delta_1 + q\delta_0 & \text{si } y \in F \end{cases}, \quad \mathbf{R}(x, i, \bullet) = \begin{cases} \mathbf{P}(x, \bullet) & \text{si } i = 0 \\ m & \text{si } i = 1 \end{cases}$$

On a alors:

$$\begin{aligned}
\Pi(x, \bullet) &= \begin{cases} q\mathbf{P}(x, \bullet) + pm & \text{si } x \in F \\ \mathbf{P}(x, \bullet) & \text{si } x \notin F \end{cases} \\
\bar{\Pi}(x, i, \bullet, j) &= \begin{cases} (1 - p\mathbf{1}_{\{F\}})\mathbf{P}(x, \bullet) & \text{pour } (i, j) = (0, 0) \\ p\mathbf{1}_{\{F\}}\mathbf{P}(x, \bullet) & \text{pour } (i, j) = (0, 1) \\ pm & \text{pour } (i, j) = (1, 1) \\ qm & \text{pour } (i, j) = (1, 0) \end{cases}
\end{aligned}$$

On remarque que, partant d'un point $(x, 0)$ et tant que $n \leq S_F$, la chaîne (X_n, U_n) reste dans $E \times 0$ et X_n se comporte comme la chaîne de transition Π sous \mathbb{P}_x .

En fait cette construction peut sembler artificielle mais on verra dans la dernière section de ce chapitre que toute chaîne possédant un ensemble F de Doeblin peut se scinder de telle sorte que le noyau Π ci dessus soit égal à sa probabilité de transition.

Lemme 2.6.22. Soit (X_n, U_n) une chaîne scindée. Alors $(F, 1)$ est un atome pour cette chaîne tel que $\bar{\Pi}(x, 1; (F, 1)) = p > 0$ et:

1. Soit Ψ une fonction mesurable positive sur l'espace canonique, alors

$$\overline{\mathbb{E}}_{(x,1)}(\Psi \circ \theta) = \overline{\mathbb{E}}_{m \otimes \mathbf{D}}(\Psi)$$

2. Si de plus la fonction Ψ ne dépend pas des coordonnées dans K alors

$$\sup_{x \in F, i \in K} \overline{\mathbb{E}}_{(x,i)}(\Psi) \leq \sup_{x \in F} \mathbb{E}_x(\Psi)$$

3. Pour $f \in \mathcal{B}^+(E)$ on a:

$$\overline{\mathbb{E}}_{m \otimes \mathbf{D}}(f(X_n) \mathbf{1}_{\{U_n=1\}}) = p \mathbb{E}_m(f(X_n))$$

En particulier on a donc $\overline{\mathbb{P}}_{(x,1)}(U_{n+1} = 1) = p \mathbb{P}_m(X_n \in F)$ et par conséquent si $\sum_n \mathbb{P}_m(X_n \in F) < \infty$, l'atome $(F, 1)$ est récurrent.

Indications de preuve:

Pour 1., il suffit de remarquer que sous $\overline{\mathbb{P}}_{(x,1)}$ la loi de (X_1, U_1) est égale à $m \otimes \mathbf{D} = pm \otimes \delta_1 + qm \otimes \delta_0$ puis d'appliquer la propriété de Markov. Pour prouver 2., on vérifie que pour $x \in F$ alors $\delta_x \otimes \mathbf{D} = p\delta_{(x,1)} + q\delta_{(x,0)}$ et donc $\mathbb{E}_x(\Psi) = p\overline{\mathbb{E}}_{(x,1)}(\Psi) + q\overline{\mathbb{E}}_{(x,0)}(\Psi)$. Enfin la première relation de 3. est une conséquence directe de la proposition 2.6.20. Pour la dernière formule on remarque que presque sûrement $(U_n = 1) \subset (X_n \in F)$ pour toute loi $\overline{\mathbb{P}}$. ■

Soit \mathcal{H}_n la tribu engendrée par les (X_k, U_k) pour $k \leq n-1$ et X_n . On note $L^{(n)}$ la suite des temps de retour successifs de (X_n, U_n) dans (F, K) .

Z. $L^{(n)}$ est aussi la suite des temps de retour de la composante X_n dans F . Mais il faut distinguer soigneusement L et S_F car ces temps d'arrêt ne sont pas définis sur le même espace de probabilité.

On note $R^{(n)}$ la suite des temps de retour de la chaîne (X_n, U_n) dans $(F, 1)$. La chaîne scindée ne passant pas dans $(E \setminus F, 1)$ après le temps 0, $R^{(n)}$ est aussi la suite des temps de retour de la composante U_n dans le point $\{1\}$, avec les mêmes précautions que ci dessus....

La suite $L^{(n)}$ est une suite de temps d'arrêt pour la filtration \mathcal{H}_n et afin de comparer ces différents temps de retour on commence par établir le lemme suivant:

Lemme 2.6.23. Soit γ une probabilité sur $(E \times K)$.

1. Soit g une fonction sur K alors:

$$\overline{\mathbb{E}}_\gamma(g(U_n) | \mathcal{H}_n) = (pg(1) + qg(0)) \mathbf{1}_{\{F\}}(X_n) + g(0) \mathbf{1}_{\{E \setminus F\}}(X_n)$$

2. Soit g une fonction sur K alors:

$$\overline{\mathbb{E}}_\gamma(g(U_{L^{(k)}}) \mathbf{1}_{\{L^{(k)} < \infty\}} | \mathcal{H}_{L^{(k)}}) = (pg(1) + qg(0)) \mathbf{1}_{\{L^{(k)} < \infty\}}$$

3. On suppose que $\bar{\mathbb{P}}_\gamma(L^{(n)} < \infty) = 1$ pour tout $n \geq 1$. Alors, sous $\bar{\mathbb{P}}_\gamma$, la suite $Y_n = U_{L^{(n)}}$, $n \geq 1$ est une suite de variables aléatoires indépendantes et de même loi de Bernoulli prenant les valeurs 1 et 0 avec les probabilités p et q .

Indications de preuve:

- Preuve de 1.:

Soit $f_k \in \mathcal{B}^+(E)$ et g_k des fonctions sur K pour $k = 0, \dots, n-1$. Alors

$$\begin{aligned} & \bar{\mathbb{E}}_\gamma(f(X_n)g(U_n) \prod_{k=0}^{n-1} f_k(X_k)g_k(U_k)) \\ &= (pg(1) + qg(0))\bar{\mathbb{E}}_\gamma(f\mathbf{1}_{\{F\}}(X_n) \prod_{k=0}^{n-1} f_k(X_k)g_k(U_k)) \\ & \quad + g(0)\bar{\mathbb{E}}_\gamma(f\mathbf{1}_{\{E \setminus F\}}(X_n) \prod_{k=0}^{n-1} f_k(X_k)g_k(U_k)) \end{aligned}$$

En effet, on obtient facilement que

$$\bar{\Pi}(f \otimes g) = (pg(1) + qg(0))\bar{\Pi}(f\mathbf{1}_{\{F\}} \otimes 1) + g(0)\bar{\Pi}(f\mathbf{1}_{\{E \setminus F\}} \otimes 1)$$

Il suffit alors d'appliquer la propriété de Markov dans le calcul de la première espérance.

- Preuve de 2.:

En décomposant l'événement $(L^{(k)} < \infty)$ suivant les valeurs de $L^{(k)} = n$ on est ramené, pour $A \in \mathcal{H}_{L^{(k)}}$ au calcul de:

$$\begin{aligned} \int_{A \cap (L^{(k)}=n)} g(U_n) d\bar{\mathbb{P}}_\gamma &= \int_{A \cap (L^{(k)}=n)} (\bar{\mathbb{E}}_\gamma(g(U_n)|\mathcal{H}_n) d\bar{\mathbb{P}}_\gamma \\ &= \int_{A \cap (L^{(k)}=n)} ((pg(1) + qg(0))\mathbf{1}_{\{F\}}(X_n) + g(0)\mathbf{1}_{\{E \setminus F\}}(X_n)) d\bar{\mathbb{P}}_\gamma \\ &= ((pg(1) + qg(0))\bar{\mathbb{P}}_\gamma(A \cap (L^{(k)} = n))) \end{aligned}$$

ce qui démontre la formule proposée. En fait ce calcul revient à prouver une propriété de Markov forte pour la chaîne $Z_n = (X_n, U_{n-1})$.

- Preuve de 3.:

On applique simplement la propriété ci dessus, en calculant par récurrence une espérance du type $\bar{\mathbb{E}}_\gamma(\prod_{k=1}^n g_k(U_{L^{(k)}}))$, en tenant compte du fait que $U_{L^{(k)}}$ est $\mathcal{H}_{L^{(n)}}$ mesurable pour $k \leq n-1$. ■

Proposition 2.6.24. Soit X_n une chaîne de transition Π correspondant à la définition 2.6.21 et telle que F soit un ensemble attractif. Alors:

1. L'atome $(F, 1)$ est attractif pour la chaîne scindée (et donc la chaîne scindée est récurrente au sens de Harris).

2. On pose $\rho_k = \sup_{x \in F} \mathbb{E}_x((S_F)^k)$. Alors pour $a \in F$ on a

$$\overline{\mathbb{E}}_{(a,i)}(L^k) \leq \rho_k$$

et il existe une constante finie C telle que:

$$\overline{\mathbb{E}}_\gamma(R^k) \leq C(\rho_k + \overline{\mathbb{E}}_\gamma(L^k))$$

pour tout $\gamma \in \mathcal{M}_1^+(E \times K)$.

3. La mesure invariante de la chaîne scindée est du type $\nu \otimes \mathbf{D}$ où ν est une mesure σ -finie sur E . La chaîne initiale est aussi récurrente au sens de Harris avec ν pour mesure invariante. La récurrence est positive si $\sup_{x \in F} \mathbb{E}_x(S_F) < \infty$.

Indications de preuve:

• Preuve de 1.:

On a alors:

$$\sum_{n=1}^{\infty} \mathbf{1}_{\{U_n=1\}} = \sum_{n=1}^{\infty} \mathbf{1}_{\{L^{(n)} < \infty, U_{L^{(n)}}=1\}}$$

En utilisant le lemme 2.6.23, on obtient pour une loi initiale γ sur $(E \times K)$:

$$\overline{\mathbb{P}}_\gamma((U_{L^{(n)}} = 1) \cap (L^{(n)} < \infty) | \mathcal{H}_{L^{(n)}}) = p \mathbf{1}_{\{(L^{(n)} < \infty)\}}$$

Par application du lemme de Borel-Cantelli généralisé, on en déduit que $\overline{\mathbb{P}}_\gamma$ p.s. on a $(N_{(F,1)} = \infty) = (N_{F,K} = \infty)$. Soit $x \notin F$, alors $\delta_x \otimes \mathbf{D} = \delta_{(x,0)}$ donc $\overline{\mathbb{P}}_{(x,0)}(N_{(F,1)} = \infty) = 1$. Pour $x \in F$ on a $\delta_x \otimes \mathbf{D} = p\delta_{(x,1)} + q\delta_{(x,0)}$ et donc $\overline{\mathbb{P}}_{(x,0)}(N_{(F,1)} = \infty) = 1$ et $\overline{\mathbb{P}}_{(x,1)}(N_{(F,1)} = \infty) = 1$. Il ne reste plus que le cas de $(x, 1)$ avec $x \notin F$ mais une simple application de la propriété de Markov montre qu'au "cran" suivant on se retrouve dans un cas déjà traité, puisqu'en utilisant l'invariance de l'événement $(N_{(F,1)} = \infty)$ on a:

$$\overline{\mathbb{P}}_{(x,1)}(N_{(F,1)} = \infty) = \overline{\mathbb{E}}_{(x,1)}(\overline{\mathbb{P}}_{(X_1, U_1)}(N_{(F,1)} = \infty))$$

et $X_1 \in F$.

• Preuve de 2.:

On sait que R est nécessairement égal à l'un des $L^{(n)}$ et l'événement $(R = L^{(n)}) = (U_L = 0, \dots, U_{L^{(n-1)}} = 0, U_{L^{(n)}} = 1)$ a pour probabilité pq^{n-1} . On peut faire le calcul suivant pour l'espérance de R^k :

$$\begin{aligned} \overline{\mathbb{E}}_\gamma((R)^k) &= \sum_{n \geq 1} \overline{\mathbb{E}}_\gamma((R)^k \mathbf{1}_{\{R=L^{(n)}\}}) \\ &= \sum_{n \geq 1} \overline{\mathbb{E}}_\gamma((L^{(n)})^k \mathbf{1}_{\{R=L^{(n)}\}}) \\ &= p \sum_{n \geq 1} \overline{\mathbb{E}}_\gamma((L^{(n)})^k \mathbf{1}_{\{U_L=0, \dots, U_{L^{(n-1)}}=0\}}) \end{aligned}$$

La dernière ligne étant obtenue en conditionnant par rapport à $\mathcal{H}_{L^{(n)}}$. Soit \mathcal{F}_n la tribu naturelle de la chaîne scindée. En utilisant le fait que $\overline{\mathbb{E}}_{(a,i)}(L^k) \leq \rho_k$ et le lemme 2.6.14 on obtient:

$$\begin{aligned} \left(\overline{\mathbb{E}}_{\gamma}((L^{(n)})^k | \mathcal{F}_{L^{(n)}}) \right)^{1/k} &\leq \rho_k^{1/k} + \left(\overline{\mathbb{E}}_{\gamma}((L^{(n-1)})^k) \right)^{1/k} \\ &\leq (n-1)\rho_k^{1/k} + \left(\overline{\mathbb{E}}_{\gamma}(L^k) \right)^{1/k} \\ \overline{\mathbb{E}}_{\gamma}((L^{(n)})^k | \mathcal{F}_{L^{(n)}}) &\leq n^{k-1}((n-1)\rho_k + \overline{\mathbb{E}}_{\gamma}(L^k)) \\ \overline{\mathbb{E}}_{\gamma}((L^{(n)})^k \mathbf{1}_{\{U_L=0, \dots, U_{L^{(n-1)}}=0\}}) &\leq n^{k-1}q^{n-1}((n-1)\rho_k + \overline{\mathbb{E}}_{\gamma}(L^k)) \end{aligned}$$

D'où le résultat par sommation sur n .

• Preuve de 3.:

La chaîne scindée possédant un atome attractif, elle est récurrente au sens de Harris et une version de sa mesure invariante est donnée par

$$\overline{\nu}(f \otimes g) = \overline{\mathbb{E}}_{(a,1)}\left(\sum_{n=1}^R f(X_n)g(U_n)\right)$$

Où $f \in \mathcal{B}_b^+(E)$ et g est une fonction positive sur K . Or

$$Z_n = \sum_{n=1}^R f(X_n)(g(U_n) - \mathbf{D}g(X_n))$$

est une $\overline{\mathbb{P}}_{\gamma}$ martingale centrée, puisque $\overline{\mathbf{\Pi}}(f(g - \mathbf{D}g)) = 0$. On en déduit que pour tout entier N on a:

$$\overline{\mathbb{E}}_{\gamma}\left(\sum_{n=1}^{R \wedge N} f(X_n)g(U_n)\right) = \overline{\mathbb{E}}_{\gamma}\left(\sum_{n=1}^{R \wedge N} f(X_n)\mathbf{D}g(X_n)\right)$$

et en passant à la limite sur N on trouve que $\overline{\nu}(f \otimes g) = \overline{\mathbb{E}}_{(a,1)}(\sum_{n=1}^R f(X_n)\mathbf{D}g(X_n))$. Il est alors très facile de constater que si l'on pose $\nu(f) = \overline{\nu}(f \otimes 1)$ alors $\overline{\nu} = \nu \otimes \mathbf{D}$. On en déduit (en utilisant 2.6.20) que ν est $\mathbf{\Pi}$ invariante. Il reste à prouver la propriété de Harris pour la chaîne X_n mais ceci résulte immédiatement du théorème ergodique pour la chaîne scindée appliqué à des fonctions du type $f \otimes 1$. En reprenant la partie 2. on voit que la récurrence de la chaîne scindée est positive dès que $\rho_1 < \infty$ puisque $\overline{\mathbb{E}}_{(a,1)}(R) \leq 2C\rho_1$.

■

Corollaire 2.6.25. *Soit X_n une chaîne de transition $\mathbf{\Pi}$ correspondant à la définition 2.6.21 telle que F soit un ensemble attractif. Cette chaîne est donc récurrente au sens de Harris et vérifie le théorème ergodique. On suppose que $\sup_{a \in F} \mathbb{E}_a(S_F^{(k)}) < \infty$, alors la récurrence est positive et:*

1. Pour tout couple $\alpha, \beta \in \mathcal{M}_1^+(E)$, on a

$$||\alpha \mathbf{\Pi}^n - \beta \mathbf{\Pi}^n|| \rightarrow 0$$

2. Si ν désigne la probabilité invariante on a

$$\|\Pi^n(x, \bullet) - \nu\| \rightarrow 0$$

pour tout $x \in E$.

3. Soit $k \geq 1$. On suppose que $\sup_{a \in F} \mathbb{E}_a(S_F^{(k)}) < \infty$ et que pour des probabilités α, β on a $\mathbb{E}_\alpha((S_F)^k) < \infty$ et $\mathbb{E}_\beta((S_F)^k) < \infty$. Alors:

$$n^k \|\alpha \Pi^n - \beta \Pi^n\| \rightarrow 0, \quad \sum_n n^{k-1} \|\alpha \Pi^n - \beta \Pi^n\| < \infty$$

4. Soit $k \geq 1$. On suppose que $\sup_{a \in F} \mathbb{E}_a(S_F^{(k+1)}) < \infty$ Alors il existe une constante $C < \infty$ tel que

$$n^k \|\Pi^n(x, \bullet) - \nu\| \leq C \mathbb{E}_x((S_F)^k)$$

pour tout $x \in E$.

Indications de preuve:

D'après la proposition 2.6.24 la chaîne de transition $\overline{\Pi}$ vérifie les hypothèses de 2.6.18 relativement à l'atome $(F \times 1)$. En tenant compte de la majoration $\overline{\mathbb{E}}_{(a,1)}(R^k) \leq 2C \sup_{x \in F} \mathbb{E}_x((S_F)^k)$ et de la relation $(\mu \otimes \mathbf{D})\overline{\Pi}^n = (\mu \Pi^n) \otimes \mathbf{D}$ on obtient immédiatement 1., 2. et 3.. Pour 4. on obtient

$$n^k \|\Pi^n(x, \bullet) - \nu\| \leq C(\sup_{x \in F} \mathbb{E}_x((S_F)^k) + \overline{\mathbb{E}}_{\delta_x \otimes \mathbf{D}}(L^k))$$

On a déjà remarqué que partant de $(x, 0)$ la chaîne scindée se comporte comme la chaîne X_n tant que l'on n'a pas atteint l'ensemble $(F, 0)$ et par conséquent $\overline{\mathbb{E}}_{(x,0)}(L^k) = \mathbb{E}_x(S_F^k)$. D'autre part pour $x \in F$ on a $\overline{\mathbb{E}}_{(x,1)}(L^k) \leq \rho_k$ d'où le résultat. ■

2.6.5 Théorie de Doeblin "locale"

Définition 2.6.26. On dit que $F \in \mathcal{B}$ est un ensemble de Doeblin s'il existe une probabilité m portée par F et $0 < p < 1$ avec: $\mathbf{Q}(x, \bullet) \geq pm$ pour tout $x \in F$.

Il faut remarquer qu'un point b n'est pas nécessairement un ensemble de Doeblin sauf si $\mathbf{Q}(b, b) > 0$. Si l'on considère une chaîne à valeurs dans \mathbb{R}^d du genre $X_{n+1} = f(X_n) + Y_{n+1}$, f étant une fonction mesurable localement bornée, et Y_n une suite de variables aléatoires indépendantes et de même loi possédant une densité continue strictement positive par rapport à la mesure de Lebesgue, alors tout compact de \mathbb{R}^d est un ensemble de Doeblin.

Le fait que F soit un ensemble de Doeblin permet de construire une chaîne scindée "naturelle". En utilisant les notations de 2.6.21, on considère la transition \mathbf{P} définie par:

$$\begin{cases} \mathbf{P}(x, \bullet) = \mathbf{Q}(x, \bullet) & \text{si } x \notin F \\ \mathbf{Q}(x, \bullet) = pm + q\mathbf{P}(x, \bullet) & \text{si } x \in F \end{cases}$$

On voit alors immédiatement que l'on a $\Pi = \mathbf{Q}$ et par conséquent on va pouvoir appliquer les résultats des sections précédentes à la chaîne initiale de transition \mathbf{Q} . Une simple copie de 2.6.24 et 2.6.25 donne alors:

Corollaire 2.6.27. *Si la chaîne de transition \mathbf{Q} admet un ensemble de Doeblin attractif F , alors cette chaîne est récurrente au sens de Harris et l'on peut lui appliquer les résultats sur la convergence du noyau $\mathbf{\Pi}^n$ obtenus dans 2.6.25 en remplaçant le noyau $\mathbf{\Pi}$ par le noyau \mathbf{Q} .*

On peut considérer des ensembles de Doeblin “généralisés”, relatifs à un itéré \mathbf{Q}^r de la transition initiale et il est facile de constater que les résultats obtenus subsistent dans ce cas. On peut montrer que lorsque la tribu \mathcal{E} est à base dénombrable, toute chaîne de Harris apériodique possède un ensemble de Doeblin généralisé qui est chargé par la mesure invariante et donc attractif.

2.6.6 Théorie de Doeblin “globale”

La théorie “globale” est en fait restreinte dans la plupart des applications au cas d'un espace d'états compact (ou fini) car sa définition est beaucoup trop restrictive en dehors de ces cas particuliers. Contrairement à la théorie “locale” on va obtenir très facilement la convergence de $\mathbf{Q}^n(x, \bullet)$ vers la probabilité invariante avec une vitesse exponentielle, qui est indépendante du point de départ x .

On n'envisage ici que l'action de \mathbf{Q} sur l'espace de Banach $\mathcal{B}_b(E)$ muni de la norme uniforme, alors qu'en fait cette technique s'applique à des espaces fonctionnels plus sophistiqués (par exemple des espaces de fonctions höldériennes) sur lesquels cet opérateur aura de “bonnes propriétés spectrales”.

Définition 2.6.28. *On dit qu'une probabilité de transition \mathbf{Q} vérifie la condition de Doeblin globale s'il existe $\nu \in \mathcal{M}_1^+(E)$ et $p > 0$ avec*

$$\mathbf{Q}(x, B) \geq p\nu(B), \text{ pour tout } B \in \mathcal{E}$$

La constante p est nécessairement ≤ 1 et on supposera qu'elle est strictement plus petite que 1 pour ne pas être dans le cas trivial $\mathbf{Q} = \nu$. Il existe alors une probabilité de transition \mathbf{S} telle qu'en posant $q = 1 - p$ on ait:

$$\mathbf{Q}(x, B) = p\nu(B) + q\mathbf{S}(x, B)$$

On considère la topologie forte sur $\mathcal{M}_b(E)$ associée à la norme duale:

$$\|\mu\| = \sup(|\mu(f)| ; f \in \mathcal{B}_b(E), \|f\| \leq 1)$$

L'espace $\mathcal{M}_b(E)$ est alors un espace de Banach réel et une probabilité de transition définit un opérateur positif de norme 1 sur $\mathcal{M}_b(E)$.

Théorème 2.6.29. *Soit \mathbf{Q} une probabilité de transition vérifiant la condition de Doeblin pour $0 < p < 1$. Il existe alors une unique probabilité invariante λ et pour tout $\mu \in \mathcal{M}_1^+(E)$ on a:*

$$\|\mu\mathbf{Q}^n - \lambda\| \leq 2q^n$$

De plus $\lambda = p \sum_{n=0}^{\infty} q^n \nu \mathbf{S}^n$

Indications de preuve:

Une probabilité λ est invariante si et seulement si $\lambda = p\nu + q\lambda\mathbf{S}$ soit encore $\lambda(I - q\mathbf{S}) = p\nu$. Or puisque \mathbf{S} est un opérateur de norme 1, $(I - q\mathbf{S})$ est inversible dans $\mathcal{M}_b(E)$ et d'inverse $\sum_{n=0}^{\infty} q^n \mathbf{S}^n$ ce qui fournit l'unicité de λ et son et la formule $\lambda = p(\sum_{n=0}^{\infty} q^n \nu \mathbf{S}^n)$. Il reste à prouver la convergence exponentielle de $\mu \mathbf{Q}^n$ mais ceci résulte de la formule:

$$\mathbf{Q}^n(x, \bullet) = p\nu \sum_{k=0}^{n-1} q^k \mathbf{S}^k + q^n \mathbf{S}^n(x, \bullet)$$

qui s'obtient facilement par récurrence. Et alors:

$$\lambda - \mu \mathbf{Q}^n = p\nu \left(\sum_{k=n}^{\infty} q^k \mathbf{S}^k \right) - q^n \mu \mathbf{S}^n = q^n (\lambda - \mu) \mathbf{S}^n$$

■

Un exemple typique d'une telle situation est le suivant:

Pour $n \geq 1$ soit (A_n, B_n, C_n) des variables aléatoires toutes indépendantes entre elles et telles que:

- Les variables aléatoires A_n prennent les valeurs 1 et 0 avec les probabilités p et $1 - p$, avec $0 < p < 1$.
- Les variables aléatoires B_n sont à valeurs dans un espace F et de loi μ .
- Les variables aléatoires C_n sont à valeurs dans un espace E et de loi ν .

On se donne de plus une fonction f mesurable de $E \times F$ dans E et on considère la suite X_n de variables aléatoires à valeurs dans E vérifiant:

$$\begin{cases} X_{n+1} = C_{n+1} & \text{si } A_{n+1} = 1 \\ X_{n+1} = f(X_n, B_{n+1}) & \text{si } A_{n+1} = 0 \end{cases}$$

On note \mathbf{S} la fonction de transition de la chaîne Z définie par $Z_{n+1} = f(Z_n, B_{n+1})$. Alors X_n est une chaîne de Markov de transition $\mathbf{Q} = p\nu + q\mathbf{S}$.

Corollaire 2.6.30. *Soit \mathbf{Q} une probabilité de transition et k un entier ≥ 1 tels que \mathbf{Q}^k vérifie la condition de Doeblin. Il existe alors une unique probabilité invariante λ . De plus il existe des constantes $C < \infty$ et $0 < \rho < 1$ tels que:*

$$\|\mu \mathbf{Q}^n - \lambda\| \leq C\rho^n, \quad \text{pour tout } \mu \in \mathcal{M}_1^+(E)$$

Les hypothèses de ce corollaire sont en particulier vérifiées sur un espace d'états fini lorsqu'il existe une puissance de la matrice de transition possédant une colonne strictement positive.

Indications de preuve:

On pose $\mathbf{R} = \mathbf{Q}^k$. Il existe donc une probabilité \mathbf{R} invariante λ avec $\|\mu\mathbf{R}^n - \lambda\| \leq 2q^n$ pour toute probabilité μ . Si on choisit $\mu = \lambda\mathbf{Q}$ alors

$$\mu\mathbf{R}^n = \lambda\mathbf{Q}^{nk+1} = (\lambda\mathbf{R}^n)\mathbf{Q} = \lambda\mathbf{Q}$$

et l'on obtient donc que λ est \mathbf{Q} invariante. Il suffit ensuite d'écrire que si $n = kp + r$ avec $0 \leq r \leq k - 1$ on a:

$$\|\mu\mathbf{Q}^n - \lambda\| = \|\mu\mathbf{Q}^{pk+r} - \lambda\mathbf{Q}^r\| = \|\mu\mathbf{Q}^{pk} - \lambda\| \leq 2q^p \leq C\rho^n$$

avec $\rho = q^{1/k}$ et $C = 2/q$. ■

Remarque

Une chaîne satisfaisant à la conclusion du Corollaire 2.6.30 est nécessairement de Harris. En effet, pour un ensemble mesurable B avec $\lambda(B) = 2\alpha > 0$ il existe N tel que $\mathbf{Q}^N(x, B) \geq \alpha$ et donc $\mathbb{P}_x(T_B < \infty) \geq \alpha$ pour tout $x \in E$. En utilisant la proposition 2.3.19 on obtient que $\mathbf{1}_{\{N_B = \infty\}} \geq \alpha$, \mathbb{P}_x p.s. pour tout $x \in E$ et par conséquent $\mathbb{P}_x(N_B = \infty) = 1$.

2.6.7 La méthode de simulation “exacte” de Propp & Wilson

On considère un ensemble d'états (E, \mathcal{E}) , muni d'une probabilité de transition \mathbf{Q} . Soit T une application aléatoire de E dans lui-même de la forme $T(x) = \Phi(x, Y)$ où Y est une variable aléatoire à valeurs dans un espace auxiliaire, telle que $T(x)$ ait pour loi $\mathbf{Q}(x, \bullet)$ pour tout $x \in E$. Soit Y_n , $n \geq 1$ une suite de variables aléatoires indépendantes et de même loi que Y , et les transformations $T_n = \Phi(\bullet, Y_n)$ associées. On définit les produits gauches et droits par les formules:

$$G_n = T_n \circ T_{n-1} \circ \dots \circ T_1 \quad , \quad D_n = T_1 \circ T_2 \circ \dots \circ T_n$$

Il est facile de constater que les lois de $G_n(x)$ et de $D_n(x)$ sont données par $\mathbf{Q}^n(x, \bullet)$. On en déduit que pour une fonction $f \in \mathcal{B}_b(E)$ et $\mu \in \mathcal{M}_1^+(E)$ on a:

$$\int_E \mathbb{E}(f(D_n(x))) d\mu(x) = \int_E \mathbb{E}(f(G_n(x))) d\mu(x) = \mu\mathbf{Q}^n(f)$$

On pose

$$\tau_G = \inf\{n \geq 1 ; \text{Card}(G_n(E)) = 1\} \quad , \quad \tau_D = \inf\{n \geq 1 ; \text{Card}(D_n(E)) = 1\}.$$

On dit que la transformation T est contractante si $\mathbb{P}\{\text{Card}(T(E)) = 1\} > 0$. Si on suppose T contractante, les événements $A_n = \{\text{Card}(T_n(E)) = 1\}$ sont indépendants et de même probabilité strictement positive. Le lemme de Borel Cantelli permet d'affirmer que, presque sûrement, l'événement A_n est réalisé une infinité de fois, ce qui est plus que suffisant pour affirmer que τ_G et τ_D sont presque sûrement finis. On désigne par Z_G et Z_D les points obtenus lorsque l'ensemble E a été contracté sur un unique point

aux instants τ_G et τ_D et on note ν_G et ν_D leurs lois respectives. Pour $n \geq \tau_D$ on a $D_n(E) = Z_D$ et par conséquent :

$$\lim_n \sum_{x \in E} \mathbb{E}(f(D_n(x))\mu(x) = \nu_D(f)$$

On en déduit qu'il n'existe qu'une seule probabilité invariante et que celle ci est égale à ν_D . Par contre, il n'y a aucune raison pour que ν_G soit invariante car pour $n \geq \tau_G$ l'ensemble $G_n(E)$ est bien réduit à un point, mais celui-ci n'est pas fixe comme dans le cas des produits droits. Pour éviter le calcul de τ_D qui s'avère souvent fastidieux, on préfère souvent utiliser le temps τ de la première occurrence de A_n (ou n'importe quel temps d'arrêt assurant la contraction des produits droits sur un point) car l'inégalité $\tau_D \leq \tau$ implique $Z_D = D_\tau(E)$.

Le cas E fini.

On suppose que \mathbf{Q} satisfait à la propriété de Doeblin. Quitte à modifier la numérotation de E , on peut supposer que la première colonne de \mathbf{Q} est à coefficients strictement positifs. La méthode de simulation de la loi invariante proposée par Propp & Wilson est alors la suivante: Pour chaque $x \in E$, on note $F^{-1}(x, \cdot)$ la fonction de répartition inverse de la ligne d'indice x de la matrice \mathbf{Q} . Soit U une v.a. suivant une loi uniforme sur $[0, 1]$, et T le vecteur aléatoire défini par $T(x) = F^{-1}(x, U)$. La transformation T est alors contractante puisque $\mathbb{P}(T(E) = x_1) = \inf_{x \in E} \mathbf{Q}(x, x_1)$.

Soit $U_n, n \geq 1$ une suite i.i.d. de v.a. suivant une loi uniforme sur $[0, 1]$. La suite aléatoire définie par la relation de récurrence $X_{n+1} = F^{-1}(X_n, U_{n+1})$ est une chaîne de Markov de transition \mathbf{Q} . En posant $T_n(x) = F^{-1}(x, U_n)$ et en choisissant l'état initial $X_0 = x$ on obtient $X_n = G_n(x)$. Le fait que τ_G soit fini est connu sous le nom de phénomène de coalescence car les trajectoires correspondant aux différents points de départ coïncident à partir du temps (aléatoire) τ_G . Par contre pour interpréter les produits droits comme une évolution Markovienne, l'on doit "renverser" le temps, d'où le nom de "coupling from the past" donné à cette méthode. . . . En fait plutôt que de calculer τ_D , on a intérêt à utiliser le temps d'arrêt géométrique :

$$\tau = \inf\{n \geq 1 ; U_n \leq \inf_{x \in E} \mathbf{Q}(x, x_1)\}$$

pour lequel on a $T_\tau(E) = x_1$ et calculer ensuite $D_\tau(E)$ (on remarque que $G_\tau(E) = x_1 \dots$). Tout ce qui vient d'être dit se transpose sans difficulté si l'on suppose seulement qu'il existe un entier r tel que \mathbf{Q}^r soit une matrice de Doeblin. En effet, on peut appliquer le procédé ci dessus à la matrice \mathbf{Q}^r et l'on sait d'après 2.6.30 que toute probabilité \mathbf{Q}^r invariante est en fait \mathbf{Q} invariante. Ceci s'applique en particulier à une chaîne irréductible et apériodique sur E .

Z. Soit X_n une suite Markovienne de la forme $X_{n+1} = f(X_n, Y_{n+1})$ (où Y_n est une suite de variables aléatoires indépendantes et de même loi), admettant la matrice de transition \mathbf{Q} (supposée être de Doeblin). L'application $T(x) = f(x, Y)$ n'a aucune raison d'être contractante (prendre par exemple une marche aléatoire associée à $f(x, y) = x+y$),

ce qui signifie que pour une matrice \mathbf{Q} donnée, il faut soigneusement choisir la fonction f pour que la représentation de la suite Markovienne associée exhibe le phénomène de coalescence!

Le cas général

Le cas d'une chaîne de Doeblin sur un espace d'états quelconque peut se traiter de la même manière: En utilisant les notations de la section précédente, on considère l'application

$$T(x) = \begin{cases} f(x, B) & \text{si } A = 0 \\ C & \text{si } A = 1 \end{cases}$$

et il est bien sûr évident que T est contractante si $\mathbb{P}(A = 1) > 0$. On peut utiliser le temps d'arrêt géométrique $\tau = \inf\{n \geq 1 ; A_n = 1\}$ et D_τ suit donc la loi invariante (on peut d'ailleurs facilement montrer que cette variable aléatoire est de loi $\lambda = p \sum_{n=0}^{\infty} q^n \nu \mathbf{S}^n$).

Algorithme de Hastings Metropolis

Soit E un espace dénombrable. On désire trouver un procédé de simulation d'une variable aléatoire de loi λ que l'on peut toujours supposer strictement positive, cette loi étant connue à une constante multiplicative près, par exemple une mesure de Gibbs de la forme $\lambda_\beta(x) = C \exp(-\beta H(x))$. Pour cela on se donne une probabilité de transition $\mathbf{R}(x, y)$ vérifiant ($\mathbf{R}(x, y) > 0 \implies \mathbf{R}(y, x) > 0$) et l'on construit \mathbf{Q} par la formule:

$$\begin{aligned} \mathbf{Q}(x, y) &= \begin{cases} \mathbf{R}(x, y) \inf\left(\frac{\lambda(y)\mathbf{R}(y, x)}{\lambda(x)\mathbf{R}(x, y)}, 1\right) & \text{si } \mathbf{R}(x, y) \neq 0 \\ 0 & \text{si } \mathbf{R}(x, y) = 0 \end{cases}, \text{ pour } x \neq y, \\ \mathbf{Q}(x, x) &= 1 - \sum_{y \neq x} \mathbf{Q}(x, y) \end{aligned}$$

Il est facile de constater que λ est réversible pour \mathbf{Q} c'est à dire que $\lambda(x)\mathbf{Q}(x, y) = \lambda(y)\mathbf{Q}(y, x)$. Il s'en suit que λ est invariante pour \mathbf{Q} . On doit donc s'assurer que \mathbf{Q} admet une seule probabilité invariante et trouver ensuite un procédé de simulation de celle-ci.

L'algorithme du "recuit simulé", correspond au cas d'une mesure de Gibbs. On se donne un ensemble fini de "voisins" de chaque point $x \in E$, de telle sorte que la mesure $\mathbf{R}(x, \bullet)$ uniformément distribuée sur les "voisins" de x , définisse un noyau symétrique et irréductible. Le noyau \mathbf{Q}_β admet donc l'unique probabilité invariante λ_β . Celle-ci converge vers la loi uniforme sur l'ensemble des minimums de la fonction H lorsque $\beta \rightarrow +\infty$, ce qui donne donc un moyen de déterminer ceux-ci, en évitant les minimums locaux.

Pour construire une variable aléatoire de loi λ , on fait la construction suivante: on se donne une loi de probabilité de base μ que l'on suppose facilement simulable et strictement positive. On choisit $\mathbf{R}(x, y) = \mu(y)$ et l'on pose $\alpha(x, y) = \frac{\mu(x)\lambda(y)}{\lambda(x)\mu(y)}$ qui est une

quantité indépendante de la constante inconnue. La transition de la chaîne de Hastings-Metropolis s'écrit alors:

$$\mathbf{Q}(x, y) = \mu(y) \inf(\alpha(x, y), 1) \text{ , pour } x \neq y \quad \mathbf{Q}(x, x) = 1 - \sum_{y \neq x} \mathbf{Q}(x, y)$$

On se donne deux suites i.i.i.d. de variables aléatoires Y_n et U_n , $n \geq 1$ suivant respectivement la loi μ et la loi uniforme sur $[0, 1]$. On définit alors la suite T_n de transformations de E dans lui-même par:

$$T_n(x) = \begin{cases} Y_n & \text{si } \alpha(x, Y_n) \geq U_n \\ x & \text{sinon} \end{cases}$$

Il est facile de constater que la loi de $T(x)$ est donnée par $\mathbf{Q}(x, \bullet)$ et que T est contractante si la fonction $\varphi(x) = \frac{\mu(x)}{\lambda(x)}$ est bornée inférieurement par un nombre strictement positif, ce que l'on suppose à partir de maintenant. Le procédé général de construction des produits droits fournit donc la réponse à la question posée. Par contre, du point de vue algorithmique, la vérification de la condition d'arrêt, c'est à dire $\text{Card}(D_n(E)) = 1$, est peu réaliste lorsque E est infini. Fort heureusement, si l'on suppose que la fonction $\varphi(x)$ atteint son minimum en un point x_0 , alors la condition d'arrêt devient beaucoup plus simple à tester puisqu'il est facile de constater que dans ce cas on a:

$$\tau_D = \inf\{n \geq 1 \mid \alpha(x_0, Y_n) \geq U_n\}, \quad Z_D = Y_{\tau_D}$$

En fait, on peut alors oublier la notion de chaîne de Markov et montrer directement à partir de la définition ci dessus que la variable $Z_D = Y_{\tau_D}$ suit la loi λ ... On peut donc considérer que dans ce cas, l'algorithme s'apparente à la procédure classique de simulation par la méthode du "rejet".

2.7 Théorie Spectrale

On se place dans le cas général d'un espace d'états (E, \mathcal{E}) . Notre but est de trouver des conditions portant sur les propriétés spectrales de l'opérateur \mathbf{Q} , assurant qu'il existe une et une seule probabilité invariante ν et que la loi de X_n c'est à dire $\mu \mathbf{Q}^n$ converge vers ν quelle que soit la loi initiale μ (stabilisation des chaînes de Markov). Dans le cas d'un espace discret on a vu qu'une chaîne récurrente irréductible était un bon candidat à une telle situation mais il se peut qu'il n'existe pas de mesure finie invariante (cas récurrent nul) et que de plus on ait des phénomènes de périodicité qui font que la suite $\mu \mathbf{Q}^n$ ne soit pas convergente.

Lemme 2.7.1. *Si $\nu \in \mathcal{M}_b(E)$ est invariante, alors ν^+ , ν^- et $|\nu|$ sont aussi invariantes et il existe donc une probabilité invariante. De plus celle ci est unique si et seulement si le sous espace des mesures bornées invariantes est de dimension 1.*

Indications de preuve:

La relation $\nu|_{\mathbf{Q}} = \nu$ implique que $|\nu|_{\mathbf{Q}} \geq |\nu|$. Mais les mesures positives finies $|\nu|$ et $|\nu|_{\mathbf{Q}}$ ont la même masse totale, elles sont donc égales. ■

Sur un espace métrique E , la norme de $\mu \in \mathcal{M}_b(E)$ peut se calculer en considérant μ comme une forme linéaire continue sur $\mathcal{C}_b(E)$ c'est à dire $\|\mu\|_1 = \sup_{\|f\| \leq 1} |\mu(f)|$ ou bien en évaluant sa variation totale c'est à dire $\|\mu\|_2 = \sup \sum_{k=1}^n |\mu(B_k)|$ où la collection (B_1, B_2, \dots, B_n) parcourt les familles finies de boréliens disjoints.

Lemme 2.7.2. *Avec les notations ci-dessus $\|\mu\|_1 = \|\mu\|_2$.*

Indications de preuve:

Si $f \in \mathcal{C}_b(E)$ alors f est une limite uniforme de combinaisons linéaires d'indicatrices de boréliens disjoints et donc $\|\mu\|_1 \leq \|\mu\|_2$. Pour l'inclusion inverse il faut utiliser le fait que μ est régulière c'est à dire que si B est borélien, alors pour tout $\epsilon > 0$ il existe un fermé F contenu dans B et un ouvert O contenant B tels que $|\mu(O \setminus F)| \leq \epsilon$. Il suffit alors de d'utiliser le fait qu'il existe une fonction continue $0 \leq f \leq 1$ qui vaut 0 sur le complémentaire de O et 1 sur F . ■

Le résultat de convexité suivant sera utile dans la suite:

Lemme 2.7.3. *Soit E un espace métrisable, $\mu \in \mathcal{M}_1^+(E)$ et f une fonction continue à valeurs complexes avec $|f(x)| \leq 1$ pour tout $x \in E$. Si $\int_E f(x) d\mu(x) = 1$ alors $f(x) = 1$ sur le support de μ .*

Indications de preuve:

Il suffit d'écrire $\int (1 - \Re(f)) d\mu = 0$ ■

2.7.1 Cas d'un espace d'états fini

On se propose d'établir un certain nombre de résultats dans le cas d'un espace fini E à N éléments. Leur seul intérêt est de servir d'introduction à la section suivante où l'on va remplacer les matrices par des opérateurs compacts, les arguments principaux étant quasiment identiques.

Définition 2.7.4. *Soit M une matrice d'ordre N opérant sur \mathbb{C}^N .*

- *Si λ est une valeur propre de M , on appelle ordre de λ (noté $c(\lambda)$) le plus petit entier k tel que $\ker(I - \lambda M)^k = \ker(I - \lambda M)^{k+1}$. Le sous espace caractéristique $C_{\{\lambda\}} = \ker(I - \lambda M)^{c(\lambda)}$ est stable par M et on note $E_{\{\lambda\}} = \ker(I - \lambda M)$ le sous espace propre.*
- *On dit que la valeur propre λ est semi-simple si $c(\lambda) = 1$ (autrement dit le sous espace caractéristique est égal au sous espace propre) et qu'elle est simple si $\dim C_{\{\lambda\}} = 1$.*

Le lemme suivant résume les quelques propriétés qui nous seront nécessaires.

Lemme 2.7.5. *Soit M une matrice d'ordre N .*

- M et son adjointe M^* admettent les mêmes valeurs propres avec les mêmes dimensions de sous espaces caractéristiques.
- \mathbb{C}^N est la somme directe des sous espaces caractéristiques.
- Lorsque $M^n/n \rightarrow 0$, toute valeur propre de module 1 est semi-simple.

Lemme 2.7.6. *Soit \mathbf{Q} une matrice de transition d'ordre N .*

- Toutes les valeurs propres sont de module ≤ 1 et celles de modules 1 sont semi-simples.
- Le nombre 1 est valeur propre. Il existe donc toujours au moins une probabilité invariante. Celle ci est unique si 1 est valeur propre simple c'est à dire si le sous espace des fonctions invariantes est réduit aux constantes.

Soit \mathbf{Q} une matrice de transition de spectre Σ on a alors:

$$\Sigma = \{1\} \cup \Sigma_\theta \cup \Sigma_\rho, \quad \mathbb{C}^N = E_{\{1\}} \oplus_{\lambda \in \Sigma_\theta} E_{\{\lambda\}} \oplus F$$

où Σ_θ est l'ensemble des valeurs propres de module 1 et différentes de 1, Σ_ρ est l'ensemble des valeurs propres de module strictement plus petit que 1 et F est la somme directe des sous espaces caractéristiques correspondant à Σ_ρ . Soit alors $\Pi_{\{\lambda\}}$ le projecteur spectral sur chaque $E_{\{\lambda\}}$ avec $|\lambda| = 1$ et $\Pi_{\{F\}}$ le projecteur spectral sur F . L'opérateur $R = \mathbf{Q}\Pi_{\{F\}}$ est de rayon spectral inférieur à ρ et l'on a:

$$\mathbf{Q}^n f = \Pi_{\{1\}} f + \left(\sum_{\lambda \in \Sigma_\theta} \lambda^n \right) \Pi_{\{\lambda\}} f + R^n f$$

Proposition 2.7.7. *Soit \mathbf{Q} une matrice de transition et $M_n = \frac{1}{n} \sum_{k=0}^{n-1} \mathbf{Q}^k$.*

1. Si $\Sigma_\theta = \emptyset$ alors il existe $C_1 < \infty$ et $\rho < 1$ tels que $\|\mathbf{Q}^n - \Pi_{\{1\}}\| \leq C_1 \rho^n$.
2. Dans tous les cas, il existe une constante finie C_2 telle que $\|M_n - \Pi_{\{1\}}\| \leq C_2/n$. Pour toute probabilité α , la forme linéaire $\alpha \Pi_{\{1\}}$ est une probabilité et de plus $\|\alpha M_n - \alpha \Pi_{\{1\}}\| \leq C_2/n$.
3. Si 1 est valeur propre simple, $\Pi_{\{1\}}(f) = \nu(f)$ où ν est l'unique probabilité invariante.

Indications de preuve:

Il suffit d'utiliser la décomposition de \mathbf{Q}^n ci dessus. Il existe une constante C_1 telle que $\|R^n\| \leq C_1 \rho^n$ d'où la première affirmation. Pour $\lambda \in \Sigma_\theta$ on a:

$$\left| \sum_{k=0}^{n-1} \lambda^k \right| = \frac{|1 - \lambda^n|}{|1 - \lambda|} \leq \frac{2}{|1 - \lambda|}$$

et donc

$$\|M_n(f) - \Pi_{\{1\}}f\| \leq \frac{2r/d + C_1/(1-\rho)}{n} \|f\|$$

où r est le cardinal de Σ_θ et d la distance de 1 à Σ_θ . Si $\alpha \in \mathcal{M}_1^+(E)$ alors la suite $\alpha_n = \alpha M_n$ est une suite de probabilités et sa limite définit une forme linéaire positive $\alpha \Pi_{\{1\}}$ qui est donc aussi une probabilité. Or la limite de α_n est manifestement invariante et par conséquent égale à ν si on a l'unicité de la probabilité invariante. ■

Proposition 2.7.8. *Soit λ une valeur propre de module 1, alors λ est une racine de l'unité et si la chaîne est irréductible, les fonctions propres associées sont de module constant.*

Indications de preuve:

Soit f une fonction propre de norme 1 associée à λ et $x_0 \in E$ tel que $|f(x_0)| = 1$. On a donc $\mathbf{Q}^n f(x_0) = \sum \mathbf{Q}^n(x_0, y) f(y) = \lambda^n f(x_0)$ et si l'on pose $A_n = \{y ; \mathbf{Q}^n(x_0, y) > 0\}$, le lemme 2.7.3 implique que $f(y) = \lambda^n f(x_0)$ pour $y \in A_n$. Les ensembles A_n sont non vides et donc il ne peut y avoir plus de N d'entre eux qui soient disjoints. Il s'en suit qu'il existe un entier $1 \leq r \leq N$ tel que $\lambda^r = 1$. Si la chaîne est irréductible alors $\cup A_n = E$. ■

On dit qu'un sous ensemble $F \subset E$ est stochastiquement clos pour \mathbf{Q} si $\mathbf{Q}(x, F) = 1$ pour tout $x \in F$. Il est facile de constater qu'un tel ensemble porte toujours une probabilité invariante (s'il est non vide...).

Proposition 2.7.9. *Si pour $1 \leq k \leq N$, la matrice \mathbf{Q}^k admet une unique probabilité invariante alors $\Sigma_\theta = \emptyset$*

Indications de preuve:

Soit λ une valeur propre de module 1. Avec les notations de la proposition 2.7.8, on pose $\Lambda_k = \{y ; f(y) = \lambda^k f(x_0)\}$ pour $k = 0, \dots, r-1$ où r est le plus petit entier p tel que $\lambda^p = 1$. Ces ensembles sont disjoints, non vides car $x_0 \in \Lambda_0$ et $A_k \subset \Lambda_k$ pour $k \geq 1$. De plus ils sont stochastiquement clos pour \mathbf{Q}^r car si $x \in \Lambda_k$ alors

$$\lambda^k f(x_0) = f(x) = \mathbf{Q}^r f(x) = \sum_{y \in E} \mathbf{Q}^r(x, y) f(y)$$

et donc $\mathbf{Q}^r(x, y) > 0 \implies y \in \Lambda_k$ c'est à dire $\mathbf{Q}^r(x, \Lambda_k) = 1$. Par conséquent \mathbf{Q}^r admet une probabilité invariante sur chaque Λ_k ce qui implique $r = 1$ d'après l'unicité et donc $\lambda = 1$. ■

2.7.2 Le cas d'un espace d'états compact

On rappelle qu'un opérateur linéaire sur un Banach B est dit compact s'il envoie la boule unité de B sur un ensemble relativement compact. IL revient au même de dire que son dual T^* envoie toute suite bornée dans le dual B^* et convergente pour la topologie $\sigma(B^*, B)$ sur une suite fortement convergente. Leur importance tient à ce que leur théorie spectrale est très proche de celle des matrices, et leurs propriétés principales sont résumées dans la proposition suivante (voir [4]).

Proposition 2.7.10. *Soit T un opérateur compact dans un espace de Banach complexe B . Alors:*

1. *Le spectre de T est l'adhérence d'un ensemble de valeurs propres dont le seul point d'accumulation éventuel est 0. En particulier si r est son rayon spectral, le cercle $|z| = r$ ne contient qu'un nombre fini de valeurs propres et le reste du spectre est contenu dans un disque $|z| \leq \rho < r$.*
2. *Le sous espace caractéristique de chaque valeur propre est de dimension finie.*
3. *Soit T^* l'opérateur dual de T . Alors T^* est compact et possède les mêmes valeurs propres que T avec la même dimension des sous espaces caractéristiques.*
4. *Si $T^n(x)/n$ converge vers 0 pour tout $x \in B$ alors toute valeur propre de module 1 est semi-simple.*

On se place ici dans le cas d'un espace d'états compact métrisable et l'on suppose que \mathbf{Q} opère sur l'espace de Banach $B = \mathcal{C}(E)$ des fonctions complexes continues sur E muni de la norme uniforme. Son adjoint opère sur l'espace des mesures complexes sur E (c'est à dire l'espace des formes linéaires continues sur $\mathcal{C}(E)$) par $\theta \rightarrow \theta\mathbf{Q}$.

Proposition 2.7.11. *Si \mathbf{Q} définit un opérateur compact sur $\mathcal{C}(E)$ de spectre Σ alors:*

1. *Les valeurs propres de module 1 sont semi-simples et il existe une probabilité invariante unique si et seulement si les fonctions continues invariantes sont constantes.*
2. *On a les décompositions:*

$$\Sigma = \{1\} \cup \Sigma_\theta \cup \Sigma_\rho, \quad \mathcal{C}(E) = E_{\{1\}} \oplus_{\lambda \in \Sigma_\theta} E_{\{\lambda\}} \oplus F$$

où Σ_θ est l'ensemble des valeurs propres de module 1 et différentes de 1, Σ_ρ est un ensemble contenu dans un disque de rayon ρ strictement plus petit que 1 et l'opérateur $R = \mathbf{Q}\Pi_{\{F\}}$ est de rayon spectral inférieur à ρ . On a alors:

$$\mathbf{Q}^n f = \Pi_{\{1\}} f + \left(\sum_{\lambda \in \Sigma_\theta} \lambda^n \right) \Pi_{\{\lambda\}} f + R^n f$$

On peut donc prouver un théorème de façon tout à fait analogue à 2.7.7:

Théorème 2.7.12. *S'il existe une puissance de \mathbf{Q} définissant un opérateur compact sur $\mathcal{C}(E)$ alors:*

1. *Si $\Sigma_\theta = \emptyset$, il existe $C_1 < \infty$ et $\rho < 1$ tels que $\|\mathbf{Q}^n - \Pi_{\{1\}}\| \leq C_1 \rho^n$.*
2. *Dans tous les cas, il existe une constante finie C_2 telle que $\|M_n - \Pi_{\{1\}}\| \leq C_2/n$. Pour toute probabilité α , la forme linéaire $\alpha\Pi_{\{1\}}$ est une probabilité et de plus $\|\alpha M_n - \alpha\Pi_{\{1\}}\| \leq C_2/n$.*

3. Si 1 est valeur propre simple, $\Pi_{\{1\}}(f) = \nu(f)$ où ν est l'unique probabilité invariante.

Indications de preuve:

Le théorème spectral implique que $\Sigma(\mathbf{Q}^n) = (\Sigma(\mathbf{Q}))^n$ et donc la décomposition du spectre de \mathbf{Q} est toujours valable. ■

Proposition 2.7.13. *Si l'une des conditions suivantes est vérifiée:*

1. *Il existe $m \in \mathcal{M}_1^+(E)$ telle que les probabilités $\mathbf{Q}(x, \bullet)$ soient uniformément absolument continues par rapport à m*
2. *Il existe une probabilité invariante m de support E , et l'opérateur \mathbf{Q} envoie $\mathcal{B}_b(E)$ dans $\mathcal{C}(E)$*
3. *Il existe un entier p tel que \mathbf{Q}^p soit un opérateur compact sur $\mathcal{C}(E)$*

alors toute valeur propre de module 1 de \mathbf{Q} est une racine de l'unité.

Indications de preuve:

Soit f une fonction propre (continue) de norme 1 associée à une valeur propre λ de module 1, et x_0 un point tel que $|f(x_0)| = 1$. On pose $\Lambda_n = \{x ; f(x) = \lambda^n f(x_0)\}$. Alors Λ_n est un ensemble fermé tel que:

$$\mathbf{Q}^n(x_0, \Lambda_n) = 1, \quad \mathbf{Q}(x, \Lambda_{n+1}) = 1 \quad \text{si } x \in \Lambda_n$$

Si 1. est vraie alors il existe $\alpha > 0$ avec $m(\Lambda_n) \geq \alpha$. En effet si ce n'est pas le cas on en déduit que pour tout $\epsilon > 0$, il existe n avec $\mathbf{Q}(x, \Lambda_n) \leq \epsilon$ pour tout x , ce qui est impossible. On en déduit donc que les ensembles Λ_n ne peuvent être disjoints puisque m est de masse totale 1.

Si 2. est vraie alors la suite $m(\Lambda_n)$ est croissante:

$$m(\Lambda_{n+1}) = m\mathbf{Q}(\Lambda_{n+1}) = \int \mathbf{Q}(x, \Lambda_{n+1}) dm(x) \geq \int_{\Lambda_n} \mathbf{Q}(x, \Lambda_{n+1}) dm(x) \geq m(\Lambda_n)$$

Or $m(\Lambda_1) = \int \mathbf{Q}(x, \Lambda_1) dm(x)$ et $\mathbf{Q}(x_0, \Lambda_1) = 1$ il existe donc un voisinage de x_0 sur lequel $\mathbf{Q}(x, \Lambda_1) \geq 1/2$ et donc $m(\Lambda_1) > 0$. On en déduit alors que les ensembles Λ_n ne peuvent être disjoints puisque m est de masse 1.

Si 3. est vraie alors d'après 2.7.12 la suite $\theta_n \in \mathcal{M}_1^+(E)$ définie par $\theta_n(f) = \frac{1}{n} \sum_{k=0}^{n-1} \mathbf{Q}^k f(x_0)$ converge en norme (calculée sur $\mathcal{C}(E)$) vers $\theta \in \mathcal{M}_1^+(E)$ et donc aussi en variation totale d'après 2.7.2. On a $\mathbf{Q}^n(x_0, \Lambda_n) = 1$ et par conséquent, si l'on pose $\Lambda = \cup_k \Lambda_k$ alors $\mathbf{Q}^n(x_0, \Lambda) = 1$. Ceci implique que $\theta_n(\Lambda) = 1$ et il en est donc de même pour θ c'est à dire $\theta(\Lambda) = 1$. Si on suppose les Λ_k disjoints alors $\mathbf{Q}^n(x_0, \Lambda_k) = \delta_n^k$, et par conséquent $\theta_n(\Lambda_k) \leq 1/n$. On en déduit $\theta(\Lambda_k) = 0$ puis $\theta(\Lambda) = 0$, ce qui est impossible. ■

Corollaire 2.7.14. *S'il existe un entier p tel que \mathbf{Q}^p soit un opérateur compact sur $\mathcal{C}(E)$ alors il existe un entier N tel que $\Sigma_\theta(\mathbf{Q}^N) = \emptyset$.*

Proposition 2.7.15. *Sous l'une des hypothèses de la proposition 2.7.13, et si \mathbf{Q}^n admet une probabilité invariante unique pour tout $n \geq 1$ alors 1 est la seule valeur propre de module 1.*

Indications de preuve:

Il suffit de recopier la preuve de 2.7.9 en tenant compte du fait qu'un ensemble fermé (donc compact) et stochastiquement clos porte une probabilité invariante. ■

Un exemple

Soit μ une loi de probabilité sur un groupe L.C.D. G opérant sur un espace homogène métrisable compact E et \mathbf{Q} la probabilité de transition associée c'est à dire $\mathbf{Q}f(x) = \int_G f(g \cdot x) d\mu(g)$. L'opérateur \mathbf{Q} envoie $\mathcal{C}(E)$ dans lui-même et il reste à donner une condition suffisante de compacité.

Proposition 2.7.16. *On suppose que la structure uniforme induite par l'action de G sur E définit la topologie de E c'est à dire que pour tout $x \in E$, $\{y ; y = g \cdot x, g \in \mathcal{V}_e\}$ est une base de voisinage de x . Si μ possède une densité par rapport à une mesure de Haar sur G alors \mathbf{Q} est compact sur $\mathcal{C}(E)$.*

Indications de preuve:

Soit φ la densité de μ par rapport à m_d . Alors pour $u \in G$:

$$\begin{aligned} \mathbf{Q}f(u \cdot x) &= \int f(gu \cdot x) \varphi(g) dm_d(g) = \int f(g \cdot x) \varphi(u^{-1}g) dm_d(g) \\ |\mathbf{Q}f(u \cdot x) - \mathbf{Q}f(x)| &\leq \|f\| \int |\varphi(u^{-1}g) - \varphi(g)| dm_d(g) \end{aligned}$$

or la dernière intégrale écrite tend vers 0 lorsque $u \rightarrow e$. Par conséquent \mathbf{Q} envoie la boule unité de $\mathcal{C}(E)$ dans un ensemble équicontinu donc relativement compact. ■

On reprend l'exemple des matrices de Schrödinger en supposant que la loi θ possède une densité sur \mathbb{R} . Alors un simple calcul de Jacobien montre que μ^{*3} possède une densité sur $SL(2, \mathbb{R})$. De plus il suffit de se restreindre aux g de type rotation pour montrer que $\{y ; y = g \cdot x, g \in \mathcal{V}_e\}$ est une base de voisinage de x (en se rappelant que l'action d'une rotation sur le tore se réduit à une translation). On en déduit donc que \mathbf{Q}^3 est un opérateur compact sur $\mathcal{C}(P(\mathbb{R}^2))$.

Pour éliminer les valeurs propres de module 1 distinctes de 1 on peut utiliser:

Proposition 2.7.17. *On suppose que l'action de G sur E est transitive. Si $T_\mu = G$ et si e appartient au support de μ alors toute valeur propre de module 1 de \mathbf{Q} est égale à 1 et les fonctions invariantes sont constantes.*

Indications de preuve:

Soit f une fonction propre (continue) de norme 1 associée à une valeur propre λ de module 1, et x_0 un point tel que $|f(x_0)| = 1$. La relation $\int f(g \cdot x_0) d\mu^{*n}(g) = \lambda^n f(x_0)$

implique par convexité que $f(g \cdot x_0) = \lambda^n f(x_0)$ sur le support de μ^{*n} pour tout $n \geq 1$. Puisque e est un point du support de μ on obtient que $\lambda = 1$. D'autre part, l'action de G étant transitive on a $T_\mu \cdot x_0 = E$ et donc f est constante. On peut remarquer que si l'on n'impose pas que e soit un point du support de μ on obtient seulement que $|f|$ est constante. ■

En fait dans la plupart des applications (et en particulier dans le cas des matrices de Schrödinger) on commence par prouver l'unicité d'une probabilité invariante pour chaque puissance de \mathbf{Q} en utilisant une méthode "directe". On établit ensuite que \mathbf{Q} possède une puissance compacte en utilisant 2.7.16 puis on élimine les éléments de Σ_θ par 2.7.15. On obtient donc ainsi la convergence exponentielle des itérés de \mathbf{Q} vers l'unique probabilité invariante.

Chapitre 3

Processus de Poisson et de Renouvellement

On considère un espace d'états métrisable E , muni de sa tribu borélienne \mathcal{E} , qui sera le plus souvent égal à \mathbb{R}^d . Soit T_n une suite de variables aléatoires définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ et à valeurs dans E . Pour toute fonction $f \in \mathcal{B}^+(E)$ on pose $N(f) = \sum_n f(T_n)$ et l'on adopte la notation $N(B)$ plutôt que $N(\mathbf{1}_{\{B\}})$. On peut remarquer que ces notations sont compatibles avec l'intégrale par rapport à la mesure aléatoire ponctuelle $N = \sum_n \delta_{T_n}$.

3.1 Processus ponctuels

Définition 3.1.1. *On dit que la suite T_n définit un processus ponctuel si $N(B)$ est presque sûrement fini pour tout borélien borné $B \subset E$. L'intensité de ce processus est la mesure positive μ définie par $\mu(B) = \mathbb{E}(N(B))$. Le processus ponctuel est dit intégrable si μ est finie sur les boréliens bornés.*

Une autre façon de présenter ces processus est de considérer une transition $N(\omega, \bullet)$ à valeurs dans les mesures ponctuelles sur E et localement finies, c'est à dire que $N(\bullet, B)$ est fini pour tout borélien borné B . On peut montrer que ces deux points de vue sont équivalents, c'est à dire qu'il est possible de numéroté de "façon mesurable" les points de la mesure ponctuelle $N(\omega, \bullet)$.

On peut aussi considérer des réunions dénombrables de processus ponctuels, pourvu que la mesure ponctuelle somme soit localement finie. L'intensité de la somme est alors la somme des intensités. De même on peut considérer l'image d'un processus ponctuel par une application mesurable, pourvu que celle-ci soit μ propre.

Un outil très utile pour l'étude des processus ponctuels est fourni par la transformée de Laplace définie ci dessous.

Définition 3.1.2. *Soit N un processus ponctuel. Sa transformée de Laplace est définie par $\Psi_N(f) = \mathbb{E}(\exp(-N(f)))$ où f parcourt l'ensemble $\mathcal{B}^+(E)$ des fonctions boréliennes positives sur E .*

Définition 3.1.3. Soit N un processus ponctuel. On appelle famille de lois conjointes de N la donnée de la loi des vecteurs aléatoires $(N(B_1), \dots, N(B_r))$, pour tout système de boréliens (B_1, \dots, B_r) .

On dit que des processus ponctuels N_i , $i \in I$, à valeurs dans des espaces E_i , sont indépendants si leurs familles de lois conjointes sont indépendantes lorsque i parcourt l'ensemble I .

On vérifie facilement que la donnée de la famille des lois conjointes d'un processus est équivalente à celle de sa transformée de Laplace. En effet la loi de $(N(B_1), \dots, N(B_r))$ est caractérisée par la fonction génératrice

$$\mathbb{E}(t_1^{N_{B_1}} t_2^{N_{B_2}} \dots t_r^{N_{B_r}}) = \Psi_N\left(\sum_{k=1}^r -\log(t_k) \mathbf{1}_{\{B_k\}}\right)$$

avec $t_k \in]0, 1]$, $1 \leq k \leq r$.

Soit $N^{(i)}$, $i \in I$ une famille finie de processus ponctuels associés aux points $T_n^{(i)}$. On définit leur somme directe M comme le processus ponctuel sur $\bigoplus_{i \in I} E_i$ associé aux points $(i, T_n^{(i)})$. Soit alors $f^{(i)}$, $i \in I$ des éléments de $\mathcal{B}^+(E_i)$ et f la fonction définie par $f(i, x_i) = f^{(i)}(x_i)$

Proposition 3.1.4. La famille $N^{(i)}$, $i \in I$ est indépendante si et seulement si on a :

$$\Psi_M(f) = \prod_{i \in I} \Psi_{N^{(i)}}(f^{(i)})$$

Indications de preuve :

Il suffit de choisir pour $f^{(i)}$ des fonctions étagées et d'utiliser la caractérisation de l'indépendance de vecteurs aléatoires à partir de leurs fonctions génératrices. Il ne reste plus ensuite qu'à passer à des limites croissantes de telles fonctions $f^{(i)}$ dans les deux membres de l'égalité (qui sont des fonctions décroissantes des $f^{(i)}$). ■

3.2 Processus de Poisson

Définition 3.2.1. On dit qu'un processus ponctuel est un processus de Poisson, s'il est intégrable et si pour tout système de boréliens bornés disjoints (B_1, \dots, B_r) , les variables aléatoires $(N(B_1), \dots, N(B_r))$ sont indépendantes et suivent des lois de Poisson de paramètres $\mu(B_k)$, $k = 1, \dots, r$. La famille des lois conjointes est donc complètement déterminée par la donnée de l'intensité μ .

La proposition suivante permet de construire un processus de Poisson pour toute intensité μ avec $\mu(E) < \infty$.

Proposition 3.2.2. Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes et de même loi μ , à valeurs dans E et S une variable aléatoire indépendante des X_n suivant une loi de Poisson de paramètre λ . La mesure aléatoire ponctuelle :

$$N(\omega, B) = \sum_{k=1}^{S(\omega)} \mathbf{1}_{\{B\}} \circ (X_k(\omega))$$

définit un processus de Poisson d'intensité $\lambda\mu$.

Indications de preuve:

On remarque que pour $t_k \in [0, 1]$, $1 \leq k \leq r$ et des boréliens (B_1, \dots, B_r) on a:

$$t_1^{N_{B_1}} t_2^{N_{B_2}} \dots t_r^{N_{B_r}} = \prod_{k=1}^S v(X_k)$$

où la fonction v est définie par
$$\begin{cases} v(x) = t_k & \text{si } x \in B_k \\ v(x) = 1 & \text{si } x \text{ n'est dans aucun } B_k \end{cases}$$

En prenant l'espérance et en découpant suivant les valeurs de S on obtient que la fonction génératrice du vecteur $(N(B_1), \dots, N(B_r))$ est égale à $\exp(\lambda(\mathbb{E}(v(X)) - 1))$. La proposition résulte alors facilement de l'égalité $\mathbb{E}(v(X)) = 1 + \sum_{k=1}^r \mu(B_k)(t_k - 1)$. ■

On en conclut que toute mesure μ , σ -finie sur E , est l'intensité d'un processus de Poisson. En effet, il suffit d'écrire $\mu = \sum_k \mu_k$ où les μ_k sont des mesures finies et de considérer le processus somme de processus de Poisson indépendants associés à chaque μ_k par le procédé ci dessus.

Proposition 3.2.3. *Soit N un processus de Poisson et f et g des fonctions dans $L^2(\mu)$. Alors $\text{Cov}(N(f), N(g)) = \mu(fg)$.*

Indications de preuve:

Il suffit de prouver cette relation pour des fonctions f et g indicatrices de boréliens bornés F et G . De plus, on peut se restreindre aux deux cas $F = G$ et $F \cap G = \emptyset$. Dans le premier cas on a $\text{Var}(N(F)) = \mu(F)$ et dans le second cas on utilise l'indépendance de $N(F)$ et $N(G)$. ■

On dit qu'un processus est simple si, pour \mathbb{P} presque tout ω , la mesure $N(\omega, \bullet)$ n'a pas de points multiples.

Proposition 3.2.4. *Un processus de Poisson est simple, si et seulement si son intensité est diffuse (c'est à dire $\mu\{x\} = 0$, $\forall x \in E$).*

Indications de preuve:

Soit $N = \sum_n \delta_{T_n}$ un processus de Poisson d'intensité μ . On construit le processus ponctuel $M = \sum_{m \neq n} \delta_{(T_n, T_m)}$, d'intensité ν . En désignant par Δ la diagonale de $E \times E$, on voit que N est simple si et seulement si $M(\Delta) = 0$, \mathbb{P} p.s. soit encore $\nu(\Delta) = 0$. Or pour des boréliens F et G on a $N(F)N(G) = M(F \times G) + N(F \cap G)$ et, en prenant l'espérance, on obtient d'après 3.2.3 que $\nu(F \times G) = \mu(F)\mu(G)$. Le résultat est alors une simple conséquence du théorème de Fubini puisque $\nu(\Delta) = \int \mu\{x\} d\mu(x)$. ■

Une caractérisation importante des processus de Poisson est fournie par la transformée de Laplace.

Proposition 3.2.5. *Le processus ponctuel N d'intensité μ est un processus de Poisson si et seulement si*

$$\Psi_N(f) = \mathbb{E}(\exp(-N(f))) = \exp\left(-\int (1 - \exp(-f(x))) d\mu(x)\right)$$

pour tout $f \in \mathcal{B}^+(E)$.

Indications de preuve:

Si l'on choisit $f = \sum_{k=1}^r \lambda_k \mathbf{1}_{\{B_k\}}$ pour des coefficients positifs λ_k et un système de boréliens bornés disjoints B_k , on voit que cette relation est équivalente à la définition du processus de Poisson puisque

$$\int (1 - \exp(-f(x))) d\mu(x) = \sum_{k=1}^r (1 - \exp(-\lambda_k)) \mu(B_k)$$

et l'on utilise les fonctions génératrices en posant $t_k = \exp(-\lambda_k)$. Il ne reste plus ensuite qu'à passer à des limites croissantes de telles fonctions f dans les deux membres de l'égalité (qui sont des fonctions décroissantes de f). ■

Un exemple d'application de cette formule est donnée par le théorème de Dobrushin dans le cas où E est un espace homogène d'un groupe G . Dans ce cas, si $\theta \in \mathcal{M}_1^+(G)$ et $\mu \in \mathcal{M}^+(E)$ on définit la "convolution" de θ et μ comme la mesure sur E définie par:

$$(\theta * \mu)(f) = \int f(g \cdot x) d\theta(g) d\mu(x), \quad \text{pour } f \in \mathcal{B}^+(E)$$

Il se peut que cette mesure ne soit pas finie sur les boréliens bornés.

Théorème 3.2.6. Dobrushin *Soit E un espace homogène d'un groupe G . On considère une suite T_n définissant les points d'un processus de Poisson, soit N , d'intensité μ sur E , et une suite Y_n de variables aléatoires indépendantes et de même loi θ sur G et indépendantes des T_n . Le processus M dont les points sont associés à la suite $Y_n \cdot T_n$ est alors un processus de Poisson d'intensité $\theta * \mu$ pourvu que cette mesure soit finie sur les boréliens bornés de E .*

Indications de preuve:

Soit $f \in \mathcal{B}^+(E)$. On pose $\exp(-h(x)) = \mathbb{E}(\exp(-f(Y.x))) = \int \exp(-f(g.x)) d\theta(g)$

$$\begin{aligned}
\mathbb{E}(\exp(-M(f))) &= \mathbb{E}(\exp(-\sum_n f(Y_n \cdot T_n))) \\
&= \mathbb{E} \left(\mathbb{E} \left\{ \exp(-\sum_n f(Y_n \cdot T_n)) \mid (T_1 = t_1, \dots, T_n = t_n, \dots) \right\} \right) \\
&= \mathbb{E} \left(\prod_n \mathbb{E} \{ \exp(-f(Y_n \cdot t_n)) \} \right) \\
&= \mathbb{E} \{ \exp(-\sum_n h(T_n)) \} = \mathbb{E}(\exp(-N(h))) \\
&= \exp \left(- \int (1 - \exp(-h(x))) d\mu(x) \right) \\
&= \exp \left(- \int (1 - \exp(-f(x))) d(\theta * \mu)(x) \right)
\end{aligned}$$

■

Ce résultat est surtout utilisé dans le cas $G = E = \mathbb{R}^d$. On peut aussi prouver un résultat voisin dans le cas d'espaces produits.

Théorème 3.2.7. *Soit N un processus ponctuel de Poisson d'intensité μ sur E , associé à la famille de points T_n , et Y_n une famille de variables aléatoires indépendantes et de même loi ν à valeurs dans E' , indépendantes de la famille T_n . Alors le processus ponctuel \tilde{N} défini sur $E \times E'$ par*

$$\tilde{N} = \sum \delta_{(T_n, Y_n)}$$

est un processus de Poisson d'intensité $\mu \otimes \nu$.

Indications de preuve:

La démonstration est tout à fait analogue à celle du théorème de Dobrushin. Pour une fonction mesurable positive f sur $E \times E'$ on introduit la fonction $h(x)$ définie par

$\exp(-h(x)) = \int \exp(-f(x, y)) d\nu(y)$ et on a :

$$\begin{aligned}
\mathbb{E}(\exp(-\tilde{N}(f))) &= \mathbb{E}(\exp(-\sum_n f(T_n, Y_n))) \\
&= \mathbb{E}\left(\mathbb{E}\{\exp(-\sum_n f(T_n, Y_n)) | (T_1 = t_1, \dots, T_n = t_n, \dots)\}\right) \\
&= \mathbb{E}\left(\prod_n \mathbb{E}\{\exp(-f(t_n, Y_n))\}\right) \\
&= \mathbb{E}\{\exp(-\sum_n h(T_n))\} = \mathbb{E}(\exp(-N(h))) \\
&= \exp\left(-\int (1 - \exp(-h(x))) d\mu(x)\right) \\
&= \exp\left(-\int (1 - \exp(-f(x, y))) d\mu \otimes \nu(x, y)\right)
\end{aligned}$$

Nous donnons ci-dessous deux applications importantes de ces théorèmes :

• **Files d'attente.**

On considère un système de service à une infinité de serveurs, les clients arrivant aux instants T_n d'un processus de Poisson sur \mathbb{R}^+ d'intensité μ , le service du client n demandant un temps Y_n . On suppose que la famille Y_n est i.i.d. de loi ν et indépendante du processus des arrivées. Soit X_t le nombre de clients en service à l'instant t .

Proposition 3.2.8. *Avec les notations ci dessus :*

- La variable aléatoire X_t suit une loi de Poisson de paramètre $\lambda_t = \mu \otimes \nu(B_t)$ où $B_t = \{(s, y) ; s \leq t, s + y > t\}$.
- Le processus des sorties associé aux instants $S_n = T_n + Y_n$ est un processus de Poisson d'intensité $\nu \star \mu$.

Il suffit d'utiliser le processus \tilde{N} pour obtenir le premier résultat et le processus M associé aux points $T_n + Y_n$ pour le second. On peut remarquer que dans le cas où $\mu = \lambda \ell^+$ (ℓ^+ étant la mesure de Lebesgue sur \mathbb{R}^+), on a $\lambda_t = \lambda \mathbb{E}(Y \wedge t)$ et donc cette loi de Poisson converge étroitement vers une loi de Poisson de paramètre $\lambda \mathbb{E}(Y)$ lorsque $t \rightarrow \infty$.

• **Effacement de processus.**

Soit N un processus de Poisson d'intensité μ sur E , et Y_n une suite i.i.d. de loi ν , à valeurs dans un espace dénombrable I et indépendante du processus N . On définit alors les processus ponctuels $N^{(i)} = \sum \mathbf{1}_{\{Y_n=i\}} \delta_{T_n}$.

Proposition 3.2.9. *Avec les notations ci dessus, les processus $N^{(i)}$ sont des processus de Poisson indépendants et d'intensité $\nu(i)\mu$.*

Il suffit de remarquer que $N^{(i)}(B) = \tilde{N}(i, B)$.

Z. Le produit direct de deux processus de Poisson n'est pas un processus de Poisson !

Une autre transformation qui conserve trivialement la notion de processus de Poisson est celle d'image par une application mesurable de (E, \mathcal{E}) dans (E', \mathcal{E}') pourvu que celle-ci conserve la notion de mesure ponctuelle localement finie.

3.3 Processus de Poisson sur \mathbb{R}^+

On considère dans cette section, une suite croissante $(0 = T_0 \leq T_1 \leq T_2 \leq \dots \leq T_n \leq \dots)$ de variables aléatoires réelles positives et finies, qui converge presque sûrement vers $+\infty$. Pour $n \geq 1$, on note par Δ_n , la suite des accroissements $\Delta_n = T_n - T_{n-1}$.

3.3.1 Compteurs

On définit les compteurs:

$$N_t = \sum_{n \geq 1} \mathbf{1}_{[0, t]} \circ T_n, \quad \text{et} \quad M_t = \sum_{n \geq 0} \mathbf{1}_{[0, t]} \circ T_n$$

On a bien sûr $M_t \geq 1$ et $M_t = N_t + 1$. Les processus N_t et M_t à valeurs dans \mathbb{N} sont croissants et "cadlag". Si la suite T_n est strictement croissante alors leurs sauts sont d'amplitude 1. La variable aléatoire M_t est un temps d'arrêt pour la filtration engendrée par le processus T_n alors que ce n'est pas le cas pour N_t . De plus, dans le cas où les T_n forment une chaîne de Markov, $\mathbb{E}(M_t)$ représente le potentiel de $[0, t]$ partant de l'origine.

Lemme 3.3.1. *On a $\mathbb{P}(N_t = \infty) = 0$ et donc la suite T_n définit bien un processus ponctuel. De plus $\mathbb{P}(\lim_{t \rightarrow \infty} N_t = \infty) = 1$.*

Indications de preuve:

On a:

$$\mathbb{P}(N_t = \infty) = \lim_n \mathbb{P}(N_t \geq n) = \lim_n \mathbb{P}(T_n \leq t) = \mathbb{P}((\sup T_n) \leq t) = 0$$

De même, en utilisant que $t \rightarrow N_t$ est une fonction croissante:

$$(\lim_{t \rightarrow \infty} N_t < \infty) = \bigcup_n (N_t < n; \forall t) = \bigcup_n (T_n > t; \forall t) = \bigcup_n (T_n = \infty)$$

d'où le résultat. ■

Proposition 3.3.2. *On se place dans le cas où Δ_n est une suite i.i.d. de variables aléatoires positives avec $m = \mathbb{E}(\Delta) > 0$. On a alors $T_n \rightarrow \infty$ p.s. et:*

$$1. \lim_{t \rightarrow \infty} \frac{N(t)}{t} = 1/m, \quad \mathbb{P} \text{ p.s.}$$

$$2. \lim_{t \rightarrow \infty} \frac{\mathbb{E}(N(t))}{t} = 1/m$$

les limites ci-dessus étant nulles si $m = +\infty$.

Indications de preuve:

En écrivant $T_{N_t} \leq t < T_{N_t+1}$, la loi forte des grands nombres (généralisée au cas $m = +\infty$) fournit le premier résultat. On désigne par $Z_t = T_{M_t}$ le plus petit des T_n strictement supérieur à t . On a $Z_t = \sum_{k=1}^{M_t} \Delta_k = \sum_{k=1}^{\infty} \Delta_k \mathbf{1}_{\{(M_t \geq k)\}}$. Or $(M_t \geq k) = (T_{k-1} \leq t)$ et donc $\mathbb{E}(Z_t) = m(\mathbb{E}(M_t))$. En utilisant la relation $Z_t > t$, on obtient $\frac{\mathbb{E}(M_t)}{t} \geq 1/m$ et on en déduit:

$$\lim_{t \rightarrow \infty} \frac{\mathbb{E}(N_t)}{t} \geq 1/m$$

ceci est bien sûr valable pour $m = +\infty$. Si on suppose qu'il existe une constante K avec $\Delta_n \leq K$ on a $Z_t \leq t + K$ et donc $\frac{\mathbb{E}(M_t)}{t} \leq 1/m + K/(mt)$ d'où:

$$\overline{\lim}_{t \rightarrow \infty} \frac{\mathbb{E}(N_t)}{t} \leq 1/m$$

Or il existe un nombre K tel que $Y_n^{(K)} = \Delta_n \wedge K$ soit d'espérance strictement positive et l'inégalité $N_t^{(K)} \geq N_t$ permet de conclure dans tous les cas. ■

Une caractérisation du Processus de Poisson sur \mathbb{R}^+

On reprend les hypothèses du début de cette section, à savoir: $(0 = T_0 \leq T_1 \leq T_2 \leq \dots \leq T_n \leq \dots)$ est une suite croissante de variables aléatoires réelles positives et finies, qui converge presque sûrement vers $+\infty$.

Proposition 3.3.3. *Soit μ l'intensité du processus ponctuel T_n et N_t le compteur associé. Si on suppose que:*

1. *La suite T_n est strictement croissante.*
2. *La mesure μ est une mesure de Radon diffuse.*
3. *Le processus N_t est un P.A.I.*

alors la loi de $N_{t+h} - N_t$ est une loi de Poisson (nécessairement de paramètre $\mu([t, t+h])$).

Indications de preuve:

On ne peut utiliser la Proposition 3.2.4 pour se passer de l'une des deux premières hypothèses puisqu'il s'agit de prouver que N_t est un processus de Poisson. Soit $I =]t, t+h]$, l'uniforme continuité de la fonction de répartition F de la mesure μ permet de construire pour tout $n \geq 1$ une suite

$$t = t_0^{(n)} < t_1^{(n)} < \dots < t_{m(n)}^{(n)} = t + h$$

de telle sorte que:

$$|F(t_{i+1}^{(n)}) - F(t_i^{(n)})| \leq 1/n, \quad \text{et} \quad |t_{i+1}^{(n)} - t_i^{(n)}| \leq h/n$$

On peut de plus imposer que la subdivision à l'ordre $n + 1$ soit une sous-subdivision de celle obtenue à l'ordre n . En notant $\Delta_i^{(n)}$ l'intervalle $]t_{i-1}^{(n)}, t_i^{(n)}]$ on peut écrire:

$$\begin{aligned} N_{t+h} - N_t &= \sum_{i=1}^{m(n)} N_{\Delta_i^{(n)}} \\ &= \sum_{i=1}^{m(n)} N_{\Delta_i^{(n)}} \mathbf{1}_{\{N_{\Delta_i^{(n)}}=1\}} + \sum_{i=1}^{m(n)} N_{\Delta_i^{(n)}} \mathbf{1}_{\{N_{\Delta_i^{(n)}} \geq 2\}} \\ &= W_n + V_n \end{aligned}$$

On vérifie facilement que la suite V_n est décroissante. De plus elle tend presque sûrement vers 0. En effet si $V_n(\omega)$ ne tend pas vers 0, il existe une suite décroissante d'intervalles $A_n(\omega) = \Delta_{i(n)}^{(n)}$ avec $N_{A_n(\omega)} \geq 2$ et dont l'intersection est un point, on en déduit que $\omega \in \cup_{m \neq n} (T_m = T_n)$ ce qui contredit le fait que le processus est simple. La suite V_n converge aussi dans L_1 vers 0 puisqu'elle est majorée par la variable aléatoire intégrable $N_{t+h} - N_t$. Si on revient à la suite W_n qui est une somme de variables aléatoires indépendantes de lois de Bernoulli de paramètres $p_i^{(n)} = \mu(\Delta_i^{(n)})$ on en déduit que :

$$\lim_{n \rightarrow \infty} \sum_{i=1}^{m(n)} p_i^{(n)} = \mu([t, t+h])$$

et que pour $0 \leq s \leq 1$:

$$\begin{aligned} \mathbb{E}(s^{W_n}) &= \prod_{i=1}^{m(n)} (1 - p_i^{(n)}(1 - s)) \\ &= \exp\left(-\sum_{i=1}^{m(n)} \log(1 - p_i^{(n)}(1 - s))\right) \end{aligned}$$

converge vers $\exp(-(1-s)\mu([t, t+h]))$. On utilise pour ceci l'encadrement:

$$-u^2/2 \leq \log(1-u) + u \leq 0 \quad \text{pour } 0 \leq u \leq 1/2$$

■

3.3.2 Le processus de Poisson homogène sur \mathbb{R}^+

On se place dans le cas où Δ_n est une suite i.i.d. de variables aléatoires positives, de loi exponentielle de paramètre $\lambda > 0$ et l'on désigne par ℓ^+ la mesure de Lebesgue sur \mathbb{R}^+ .

Proposition 3.3.4. *Soit φ une fonction borélienne positive sur $(\mathbb{R}^+)^n$. Alors*

$$\mathbb{E}(\varphi(T_1, \dots, T_n) \mathbf{1}_{\{N_t=n\}}) = \lambda^n \exp(-\lambda t) \int_{0 < t_1 < t_2 < \dots < t_n} \varphi(t_1, \dots, t_n) dt_1 \dots dt_n$$

Si de plus φ est invariante par permutation alors cette espérance vaut

$$\frac{\lambda^n \exp(-\lambda t)}{n!} \int_{[0,t]^n} \varphi(t_1, \dots, t_n) dt_1 \dots dt_n$$

et en particulier si f est une fonction borélienne positive sur \mathbb{R} , on a:

$$\mathbb{E} \left(\mathbf{1}_{\{N_t=n\}} \prod_{k=1}^n f(T_k) \right) = \frac{\lambda^n \exp(-\lambda t)}{n!} \left(\int_0^t f(u) du \right)^n$$

Indications de preuve:

La première relation résulte d'un simple changement de variable dans la loi produit du vecteur $(\Delta_1, \Delta_2, \dots, \Delta_{n+1})$ et les autres égalités sont des conséquences évidentes. ■

Définition 3.3.5. On dit qu'un processus X_t est de Lévy si c'est un processus cadlag et qui de plus est à accroissements indépendants et homogènes c'est à dire que:

- Pour toute subdivision $0 \leq t_1 < t_2 < \dots < t_n$, les variables aléatoires $(X_{t_k} - X_{t_{k-1}})$ sont indépendantes pour $k = 2, \dots, n$.
- La loi de $(X_{t+h} - X_t)$ ne dépend pas de t .

Dans la suite on confond le processus ponctuel de Poisson associé à la suite de points T_n avec le processus "ordinaire" fourni par le compteur N_t puisque ces deux données sont équivalentes.

Proposition 3.3.6. Le compteur associé à la marche T_n , $n \geq 1$ est un processus de Poisson d'intensité $\lambda \ell^+$. C'est donc un processus de Lévy.

Indications de preuve:

On reprend les notations de 3.2.2, on considère des boréliens bornés (B_1, B_2, \dots, B_r) et donc contenus dans un intervalle $[0, t]$. On a alors:

$$\begin{aligned} \mathbb{E} \left(t_1^{N_{B_1}} t_2^{N_{B_2}} \dots t_r^{N_{B_r}} \right) &= \mathbb{E} \left(\prod_{k=1}^{\infty} v(T_k) \right) \\ &= \sum_n \mathbb{E} \left(\mathbf{1}_{\{N_t=n\}} \prod_{k=1}^n v(T_k) \right) \\ &= \sum_n \frac{\lambda^n \exp(-\lambda t)}{n!} \left(\int_0^t v(u) du \right)^n \\ &= \exp(-\lambda(t - \int_0^t v(s) ds)) \end{aligned}$$

$$\text{Or } \int_0^t v(s) ds = t - \sum_{k=1}^n (1 - t_k \ell^+(B_k)). \quad \blacksquare$$

On appelle souvent ce processus, "processus de Poisson homogène sur \mathbb{R}^+ d'intensité λ ".

3.3.3 Caractérisations du processus de Poisson homogène sur \mathbb{R}^+

On revient au cas général d'une suite croissante T_n qui converge vers $+\infty$. On note par \mathcal{F}_t la tribu engendrée par les variables aléatoires N_s pour $s \leq t$. La proposition suivante est la traduction de la propriété de Markov forte des processus de Lévy.

Proposition 3.3.7. *On suppose que N_t est un processus de Lévy. Alors pour tout temps d'arrêt S relativement à la filtration \mathcal{F}_t , le processus $N_t^S = N_{S+t} - N_S$ est lui même un processus de Lévy, de même loi que le processus N_t et indépendant de \mathcal{F}_S . En fait, c'est le compteur associé aux points $T_n^S = T_{n+N_S}$ pour $n \geq 1$, à partir de l'origine des temps $T_0^S = S$.*

Indications de preuve:

On considère tout d'abord le cas d'un temps d'arrêt ne prenant qu'un ensemble dénombrable de valeurs $\{s_k\}$, $k \geq 1$. On a alors pour un ensemble $A \in \mathcal{F}_S$, une subdivision $0 \leq t_1 < t_2 < \dots < t_n$ et des entiers m_2, \dots, m_n :

$$\begin{aligned} & \mathbb{P}\left(A \cap \bigcap_{k=2}^n (N_{t_k}^S - N_{t_{k-1}}^S = m_k)\right) \\ &= \sum_r \mathbb{P}\left(A \cap (S = s_r) \cap \bigcap_{k=2}^n (N_{s_r+t_k} - N_{s_r+t_{k-1}} = m_k)\right) \\ &= \sum_r \mathbb{P}(A \cap (S = s_r)) \mathbb{P}\left(\bigcap_{k=2}^n (N_{s_r+t_k} - N_{s_r+t_{k-1}} = m_k)\right) \quad (\text{car } (A \cap (S = s_r)) \in \mathcal{F}_{s_r}) \\ &= \mathbb{P}(A) \prod_{k=2}^n \mathbb{P}(N_{t_k} - N_{t_{k-1}} = m_k) \end{aligned}$$

Pour passer au cas continu, on utilise l'approximation dyadique par excès du temps d'arrêt S , c'est à dire que l'on pose $S^{(n)} = \frac{k}{2^n}$ pour $\frac{k-1}{2^n} < S \leq \frac{k}{2^n}$ et l'approximation dyadique par défaut du réel t , c'est à dire que l'on pose $t^{(n)} = \frac{k}{2^n}$ pour $\frac{k}{2^n} \leq t < \frac{k+1}{2^n}$. La suite $S^{(n)}$ est alors décroissante vers S et l'on remarque que $(S^{(n)} \leq t) = (S \leq t^{(n)})$ ce qui montre que $S^{(n)}$ est un temps d'arrêt pour la filtration \mathcal{F}_t . Il suffit alors d'utiliser la continuité à droite du processus N_t dans le calcul précédent.

Corollaire 3.3.8. *On suppose que la suite T_n est strictement croissante et que N_t est un processus de Lévy. Alors il existe $\lambda > 0$ tel que la suite $\Delta_n = T_n - T_{n-1}$ soit une suite de variables aléatoires i.i.d. de loi exponentielle de paramètre λ .*

Indications de preuve:

En posant $\varphi(t) = \mathbb{P}(N_t = 0)$ on obtient que $\varphi(t+s) = \varphi(t)\varphi(s)$. Or $\varphi(s) = \mathbb{P}(T_1 > s)$ et donc $\lim_{s \rightarrow 0} \varphi(s) = \mathbb{P}(T_1 > 0) = 1 = \varphi(0)$. On en déduit que la fonction φ est continue à droite et par conséquent qu'il existe $\lambda \geq 0$ tel que $\varphi(t) = \exp(-\lambda t)$. On ne peut avoir $\lambda = 0$ car ceci impliquerait $\mathbb{P}(T_1 > t) = 1$ pour tout t ce qui est impossible. Il s'en suit que T_1 est une loi exponentielle de paramètre λ . Il suffit alors de remarquer que la suite T_n est une suite de temps d'arrêt pour la filtration \mathcal{F}_t , en particulier en utilisant le résultat précédent avec $S = T_1$ on obtient que $T_n^S = T_{n+1}$ et que $((T_2 - T_1) > t) = (N_t^S = 0)$ ce dernier événement étant indépendant de \mathcal{F}_S et donc de $T_1 \dots$ ■

Pour un processus de Poisson homogène sur \mathbb{R}^+ , il est facile de constater que $N_t - \lambda t$ est une martingale pour la filtration \mathcal{F}_t . Le résultat suivant établit la réciproque.

Proposition 3.3.9. *On suppose que la suite T_n est strictement croissante et on considère le compteur N_t associé. Si le processus $N_t - \lambda t$ est une martingale pour la filtration \mathcal{F}_t alors N_t est un processus de Poisson homogène de paramètre λ .*

Indications de preuve:

Dans toutes les intégrales ci dessous, les bornes $u \leq v$ signifient que l'intégrale est calculée sur l'intervalle semi-ouvert $]u, v]$. On désigne par N_t^- la limite à gauche de l'application $t \rightarrow N_t$. On vérifie facilement que $N_{T_n}^- = n - 1$ pour $n \geq 1$ et donc pour $0 < u < 1$ on a:

$$1 - u^{N_t} = (1 - u) \int_0^t u^{N_s^-} dN_s$$

Pour $h > 0$ on pose

$$\begin{aligned} A_h &= \frac{1}{h} \int_0^h \left(\int_0^v u^{N_s} ds \right) dN_v \\ B_h &= \frac{1}{h} \int_h^t \left(\int_{v-h}^v u^{N_s} ds \right) dN_v \\ C_h &= \frac{1}{h} \int_t^{t+h} \left(\int_{v-h}^v u^{N_s} ds \right) dN_v \end{aligned}$$

On vérifie alors que $A_h \leq N_h$, que $C_h \leq N_{t+h} - N_t$ et par conséquent que l'intégrale $I_h = \frac{1}{h} \int_0^t u^{N_s} (N_{s+h} - N_s) ds$ qui est égale à $A_h + B_h + C_h$ converge vers $\int_0^t u^{N_s^-} dN_s$ lorsque $h \rightarrow 0$. Or puisque $B_h \leq N_t - N_h$ on a $I_h \leq N_{t+h} \leq N_{t+1}$ pour $h \leq 1$ et le théorème de Lebesgue implique que:

$$\lim_{h \rightarrow 0} \mathbb{E}(I_h) = \frac{1 - \mathbb{E}(u^{N_t})}{1 - u}$$

En utilisant la propriété de martingale on obtient $\mathbb{E}(I_h) = \lambda \int_0^t \mathbb{E}(u^{N_s}) ds$ et donc, en posant $\varphi_t(u) = \mathbb{E}(u^{N_t})$, la relation:

$$1 - \varphi_t(u) = \lambda(1 - u) \int_0^t \varphi_s(u) ds$$

On en déduit que N_t suit une loi de Poisson de paramètre λt . On poursuit la démonstration en remplaçant N_t par $N_t - N_{t'}$, l'intervalle d'intégration $]0, t]$ par $]t', t]$ et l'espérance par l'espérance conditionnelle quand $\mathcal{F}_{t'}$. On prouve de même que, conditionnellement à $\mathcal{F}_{t'}$, la variable aléatoire $N_t - N_{t'}$ suit une loi de Poisson de paramètre $\lambda(t - t')$ et qu'elle est donc indépendante de $\mathcal{F}_{t'}$. Ceci prouve que N_t est un processus de Lévy et on utilise la proposition précédente. ■

3.3.4 Processus de Poisson composés

Soit N un processus de Poisson homogène sur \mathbb{R}^+ d'intensité λ , associé à la famille de points T_n , et Y_n une famille de variables aléatoires indépendantes et de même loi ν à valeurs dans \mathbb{R}^+ , indépendantes de la famille T_n . On définit pour $t \geq 0$ le processus:

$$Z_t = \sum_{n=1}^{N_t} Y_n = \sum_{(n \geq 1; T_n \leq t)} Y_n$$

Pour une mesure ν sur \mathbb{R}^+ on définit sa transformée de Laplace $\mathcal{L}\nu$ par $\mathcal{L}\nu(t) = \int \exp(-tx) d\nu(x)$.

Proposition 3.3.10. *Z_t est un processus de Levy continu à droite et pour des nombres réels positifs t et α on a:*

$$\mathbb{E}(\exp(-\alpha Z_t)) = \exp(-\lambda t(1 - \mathcal{L}\nu(\alpha))), \text{ et donc } \mathbb{E}(Z_t) = \lambda t \mathbb{E}(Y)$$

Indications de preuve:

La continuité à droite est évidente et en reprenant les notations de 3.2.7 on voit que $Z_t = \tilde{N}(\varphi_t)$ pour $\varphi_t(x, y) = y \mathbf{1}_{[0, t]}(x)$. La formule proposée est donc une simple conséquence de l'expression de la transformée de Laplace d'un processus de Poisson d'intensité $\lambda \ell^+ \otimes \nu$ puisque

$$\mathbb{E}(\exp(-\alpha \tilde{N}(\varphi_t))) = \exp\left(-\lambda \int (1 - \exp(-\alpha \varphi_t(x, y))) dx \otimes d\nu\right)$$

Pour une subdivision $0 \leq t_1 < t_2 < \dots < t_n$, on peut faire le même calcul en écrivant que

$$\sum_{k=2}^n \alpha_k (Z_{t_k} - Z_{t_{k-1}}) = \tilde{N}\left(\sum_{k=2}^n \alpha_k (\varphi_{t_k} - \varphi_{t_{k-1}})\right)$$

■

Un exemple classique d'utilisation de tels processus est le suivant: On considère que des accidents se produisent aux instants T_n d'un processus de Poisson homogène sur \mathbb{R}^+ et que le $n^{\text{ème}}$ accident coûte une somme Y_n à la compagnie d'assurance. Si c est le taux des primes par unité de temps, son bilan à l'instant t est donc $X_t = X_0 + ct - Z_t$. Soit W le temps d'entrée de X_t dans \mathbb{R}^- . Le problème est alors de trouver la probabilité de ruine, c'est à dire $\mathbb{P}(W < \infty | X_0 = x)$, et plus précisément son comportement asymptotique lorsque $x \rightarrow +\infty$. La théorie du renouvellement développée dans la section suivante va permettre de répondre à cette question.

3.4 Le Renouvellement sur \mathbb{R}^+

On s'intéresse ici au comportement asymptotique du compteur N_t dans le cas d'une marche aléatoire $T_n = T_0 + \Delta_1 + \dots + \Delta_n$ sur \mathbb{R}^+ où Δ_n est une suite de variables

aléatoires indépendantes et de même loi ν , positives et non identiquement nulles. Dans la suite on supposera $T_0 = 0$ car T_0 étant indépendant de la suite Δ_n il suffira en général “d’intégrer” le résultat obtenu pour $T_0 = 0$ par rapport à la loi de T_0 . On a déjà vu que la courbe $t \rightarrow \mathbb{E}(N_t)/t$ admet la direction asymptotique $1/m$ où m est l’espérance de Y . Il s’agit maintenant de prouver, qu’en général, la pente de la sécante, c’est à dire $\mathbb{E}(N_{t+h} - N_t)/h$, admet la même limite lorsque $t \rightarrow \infty$. Nous commençons par un résultat classique dont la preuve initiale était relativement compliquée, jusqu’à ce que Albert Raugi fournisse un argument très simple que nous reproduisons ici.

Lemme 3.4.1. (Choquet-Deny) *Soit E un groupe abélien L.C.D. et $\nu \in \mathcal{M}_1^+(E)$ adaptée. Alors toute fonction f , continue bornée et harmonique, c’est à dire vérifiant $\overline{\mathbf{Q}f(x)} = \int f(x+y) d\nu(y) = f(x)$ pour tout $x \in E$, est constante.*

Indications de preuve:

Soit On pose

$$H_0(x) = \int (f(x) - f(x+y))^2 d\nu(y) = f^2(x) - \mathbf{Q}(f^2)(x), \quad H_n(x) = \mathbf{Q}^n H_0(x)$$

On a $S_n f(x) = \sum_{k=0}^n H_k(x) = f^2(x) - \mathbf{Q}^{n+1} f^2(x) \leq 2\|f\|$ et par conséquent la suite positive $S_n f(x)$ est bornée. D’autre part on a:

$$\begin{aligned} H_1(x) &= \mathbf{Q}H_0(x) = \int (H_0(x+z)) d\nu(z) \\ &= \int \int (f(x+z) - f(x+y+z))^2 d\nu(y) d\nu(z) \\ &\geq \int \left(\int (f(x+z) - f(x+y+z)) d\nu(z) \right)^2 d\nu(y) \\ &= H_0(x) \end{aligned}$$

La suite $H_n(x)$ est donc croissante et puisque la série de terme général $H_n(x)$ est convergente on a nécessairement $H_0(x) = 0$ et donc $f(x+y) = f(x)$ pour ν presque tout y . La fonction f est continue et ceci implique que $f(x+y) = f(x)$ pour tout y dans le support de ν . Autrement dit, tout point du support de ν est une période pour f . L’ensemble des périodes de f est un sous groupe fermé de E , il est donc égal à E d’après la condition d’adaptation. Il faut noter que la commutativité du groupe E joue un rôle essentiel et que ce résultat est faux dans le cas non abélien. ■

On note $\mathbf{\Pi} = \sum_{n \geq 0} \nu^{*n}$ le noyau potentiel de la marche transiente T_n . D’après 2.5.3 c’est une mesure de Radon et bien sûr $\mathbf{\Pi}([0, t]) = \mathbb{E}(M_t) = \mathbb{E}(N_t) + 1$.

Lemme 3.4.2. *Soit $h_\nu(x)$ la fonction définie pour $x \geq 0$ par $h(x) = \nu([x, \infty[)$ et nulle pour $x < 0$. Alors $\mathbf{\Pi} * h_\nu = \mathbf{1}_{\{\mathbb{R}^+\}}$.*

Indications de preuve:

Pour $t > 0$ on a $\mathcal{L}\mathbf{\Pi}(t) = \frac{1}{1 - \mathcal{L}\nu(t)}$ et une simple application du théorème de Fubini

montre que $\mathcal{L}h(t) = \frac{1 - \mathcal{L}\nu(t)}{t}$. Il ne reste plus qu'à utiliser 1.4.9, puisque la transformée de Laplace de $\mathbf{1}_{\{\mathbb{R}^+\}}$ vaut $1/t$ pour $t > 0$. ■

Pour $t \in \mathbb{R}$ on pose $\mathbf{\Pi}_t = \mathbf{\Pi} * \delta_{-t}$. La proposition suivante résume les résultats techniques nécessaires pour prouver le résultat principal.

Proposition 3.4.3. *Soit K un compact de \mathbb{R} .*

1. *Il existe une constante finie C_K telle que $\mathbf{\Pi}(K + t) \leq C_K$ pour tout $t \in \mathbb{R}$. La famille $\mathbf{\Pi}_t$ est donc vaguement relativement compacte.*
2. *Si $\mathbf{\Pi}_{t_k}$ converge vaguement vers $\gamma \in \mathcal{M}^+(\mathbb{R})$ lorsque $t_k \rightarrow +\infty$, (ou bien $t_k \rightarrow -\infty$) alors, pour $\mu \in \mathcal{M}_1^+(\mathbb{R})$, la famille $\mu * \mathbf{\Pi}_{t_k}$ converge vaguement vers $\mu * \gamma$*
3. *Pour une mesure γ obtenue comme ci-dessus et $\varphi \in \mathcal{C}_k(\mathbb{R})$ la fonction $g(x) = \int (\varphi(x + y) d\gamma(y))$ est continue et bornée.*

Indications de preuve:

• Preuve de 1.

La première affirmation est une conséquence de 2.5.3 puisque $\mathbf{\Pi}(K + t) = \mathbf{G}(-t, K)$. Il suffit ensuite d'appliquer 1.3.7.

• Preuve de 2.

Soit $\varphi \in \mathcal{C}_k(\mathbb{R})$, de support contenu dans le compact K

$$\lim_{t_k \rightarrow \pm\infty} (\mu * \mathbf{\Pi}_{t_k})(\varphi) = \lim_{t_k \rightarrow \pm\infty} \int \left(\int \varphi(x + y) d\mathbf{\Pi}_{t_k}(y) \right) d\mu(x)$$

Or en posant $g_t(x) = \int \varphi(x + y) d\mathbf{\Pi}_t(y)$, on a $\lim_{t_k \rightarrow \pm\infty} g_{t_k}(x) = \int \varphi(x + y) d\gamma(y)$ et de plus:

$$|g_t(x)| \leq \|\varphi\| \mathbf{\Pi}(K - x + t) \leq \|\varphi\| C_K < \infty$$

Il suffit donc d'appliquer le théorème de Lebesgue.

• Preuve de 3.

Avec les notation ci-dessus on a:

$$g(x) = \int \varphi(x + y) d\gamma(y) = \lim_{t_k \rightarrow \pm\infty} \int \varphi(x + y) d\mathbf{\Pi}_{t_k}(y) = \lim_{t_k \rightarrow \pm\infty} g_{t_k}(x)$$

Les majorations obtenues ci dessus montrent que g est bornée et la continuité est évidente. ■

Pour $h > 0$, on pose $I_{n,h} =]nh, (n + 1)h]$. Pour une fonction positive f sur \mathbb{R} on note $M_{k,h}(f) = \sup_{x \in I_{k,h}} f(x)$, $m_{k,h}(f) = \inf_{x \in I_{k,h}} f(x)$. On définit alors la somme de Riemann supérieure $M_h(f) = h \sum_k M_{k,h}(f)$ et la somme de Riemann inférieure $m_h(f) = h \sum_k m_{k,h}(f)$

Définition 3.4.4. *On dit qu'une fonction mesurable positive f sur \mathbb{R} est directement Riemann intégrable (D.R.I. en abrégé) si:*

1. Il existe $h > 0$ tel la somme de Riemann supérieure de soit finie (f est donc intégrable).
2. Les sommes de Riemann inférieures et supérieures convergent vers l'intégrale de f lorsque $h \rightarrow 0$.

Une fonction non nécessairement positive est *D.R.I.* si f^+ et f^- le sont.

Il est facile de constater qu'une fonction monotone positive finie portée par \mathbb{R}^+ et intégrable est *D.R.I.*. De même, une fonction continue telle qu'il existe $h > 0$ pour lequel la somme de Riemann supérieure de $|f|$ est finie est une fonction *D.R.I.*.

Lemme 3.4.5. Soit t_k une suite convergent vers $+\infty$ ou vers $-\infty$ et ℓ la mesure de Lebesgue sur \mathbb{R} . S'il existe $\alpha \in \mathbb{R}$ tel que Π_{t_k} converge vaguement vers $\alpha\ell(I)$, alors on a:

$$\lim_{t_k \rightarrow \pm\infty} \Pi * f(t_k) = \alpha \int f(t) dt$$

pour toute fonction f *D.R.I.* sur \mathbb{R} .

Indications de preuve:

En utilisant le fait qu'un intervalle I est de frontière de mesure nulle pour ℓ , on obtient, d'après 1.3.6, la convergence de $\Pi_{t_k}(-I) = \Pi(-I + t_k) = \Pi * \mathbf{1}_{\{I\}}(t_k)$ vers $\alpha\ell(I)$ pour tout intervalle borné. On considère tout d'abord une fonction f du type $f = \sum_n a_n \mathbf{1}_{\{I_{n,h}\}}$ avec $a_n \geq 0$ et $\sum_n a_n < \infty$. En utilisant 3.4.3 alinéa 1. on remarque que

$$\Pi * \mathbf{1}_{\{I_{n,h}\}}(t) = \Pi([-h, 0] - nh - t) \leq C_h < \infty$$

et par conséquent, si $f_N = \sum_{n=-N}^N a_n \mathbf{1}_{\{I_{n,h}\}}$ et $g_N = f - f_N$ on a:

$$|\Pi * f(x) - \alpha \int f(t) dt| \leq |\Pi * f_N(x) - \alpha \int f_N(t) dt| + (h + C_h) \sum_{|n| > N} a_n$$

ce qui prouve le lemme dans ce cas. On écrit ensuite que toute fonction *D.R.I.* positive est encadrée par deux fonctions de ce type et on fait tendre h vers 0. ■

Théorème 3.4.6. (Renouvellement sur \mathbb{R}^+) Soit ν une mesure adaptée sur \mathbb{R}^+ , d'espérance m (finie ou non). Alors:

$$\lim_{t \rightarrow +\infty} \Pi * f(t) = \frac{1}{m} \int f(x) dx$$

pour toute fonction f qui est *D.R.I.* sur \mathbb{R} . La limite est nulle si $m = +\infty$.

Indications de preuve:

D'après 3.4.3 alinéa 1. la famille Π_t est étroitement relativement compacte. Soit γ une valeur d'adhérence vague lorsque $t \rightarrow +\infty$, limite vague d'une suite Π_{t_k} avec $t_k \rightarrow +\infty$. On déduit de l'équation $\Pi * \nu + \nu = \Pi$ que $\Pi_{t_k} * \nu + \nu * \delta_{-t_k} = \Pi_{t_k}$ et donc, d'après 3.4.3 alinéa 2., l'égalité $\nu * \gamma = \gamma$ car il est facile de constater que $\nu * \delta_{-t_k}$ converge

vaguement vers 0. Il s'en suit que pour $\varphi \in \mathcal{C}_k(\mathbb{R})$ la fonction $g(x) = \int \varphi(x+y) d\gamma(y)$ est une fonction harmonique, continue et bornée d'après 3.4.3 alinéa 3.. Le lemme de Choquet-Deny implique alors que g est constante. La mesure γ est donc proportionnelle à la mesure de Lebesgue ℓ de \mathbb{R} , c'est à dire que $\gamma = \alpha\ell$, la constante α dépendant en principe de la suite t_k choisie. D'après 3.4.5 on obtient donc que:

$$\lim_{t_k \rightarrow +\infty} \Pi * f(t_k) = \alpha \int f(t) dt$$

pour toute fonction f qui est D.R.I. sur \mathbb{R} .

- Identification de α lorsque $m < \infty$.

Soit μ la probabilité sur \mathbb{R}^+ de densité $\frac{1}{m}h_\nu(x)$, d'après 3.4.2, la mesure $m\mu * \Pi_{t_k}$ est égale à la restriction de ℓ à l'intervalle $[-t_k, \infty[$ qui converge vaguement vers ℓ . On en déduit, en utilisant de nouveau 3.4.3 alinéa 2., que $\alpha = 1/m$ et que donc Π_t converge vaguement vers ℓ/m puisque toutes les valeurs d'adhérence sont égales et le théorème est démontré dans ce cas.

- Identification de α lorsque $m = \infty$.

La fonction $h_n(x) = h_\nu(x)\mathbf{1}_{[0,n]}(x)$ est D.R.I et pour tout n fixé:

$$\mathbf{1}_{\{\mathbb{R}^+\}}(t_k) \geq \Pi * h_n(t_k) \rightarrow \alpha \int h_n(t) dt$$

Or $\lim_n \int h_n(t) dt = \int h(t) dt = +\infty$ et donc $\alpha = 0$. ■

La principale application de ce résultat est l'étude de l'équation du renouvellement.

Définition 3.4.7. Soit $\nu \in \mathcal{M}_1^+(\mathbb{R}^+)$, H et h des fonctions positives à support dans \mathbb{R}^+ . On dit que H satisfait à l'équation du renouvellement avec second membre h si:

$$\boxed{H = \nu * H + h}$$

Proposition 3.4.8. Si h possède une transformée de Laplace pour $t > 0$, par exemple si h est bornée ou intégrable, alors la seule solution de l'équation du renouvellement possédant une transformée de Laplace pour $t > 0$ est donnée par $H = \Pi * h$.

Pour trouver le comportement asymptotique de H , il suffit alors d'utiliser le théorème du renouvellement. En fait, cette procédure est toujours valable même lorsque la mesure ν , figurant dans l'équation du renouvellement est une sous probabilité. En effet s'il existe $\alpha > 0$ avec $\int \exp(\alpha x) d\nu(x) = 1$, on pose $d\nu'(x) = \exp(\alpha x)d\nu(x)$, $H'(x) = H(x)\exp(\alpha x)$ et $h'(x) = h(x)\exp(\alpha x)$, et on obtient l'équation $H' = H' * \nu' + h'$ où $\nu' \in \mathcal{M}_1^+(\mathbb{R}^+)$. De plus, si l'on sait que $0 \leq H \leq 1$, l'équation $H = \nu * H$ peut se transformer en une "vraie" équation de renouvellement $G = \nu * G + h$ avec $G = \mathbf{1}_{\{\mathbb{R}^+\}} - H$, $h(x) = \nu([x, +\infty[)$ pour $x \geq 0$ et $h(x) = 0$ pour $x < 0$. On peut remarquer que la fonction h est D.R.I. si $m < \infty$.

On peut maintenant, mener à bien l'étude du problème de ruine évoqué dans les processus de Poisson composés. On considère le processus $X_t = X_0 + ct - \sum_{n=1}^{N_t} Y_n$ où N_t est un

processus de Poisson homogène sur \mathbb{R}^+ d'intensité λ et Y_n est une suite i.i.d. de variables aléatoires positives, indépendantes du processus de Poisson, de loi ν de moyenne finie m . On désigne par W le temps d'entrée de X_t dans $] -\infty, 0[$ et on pose $F(x) = \mathbb{P}(W = \infty | X_0 = x)$.

Lemme 3.4.9. *Sous les hypothèses ci dessus on a:*

- $F(x)$ est croissante et $F(x) = 0$ si $x < 0$.
- $F(x)$ est identiquement nulle si $c \leq \lambda m$.
- Si $c > \lambda m$ alors $\lim_{x \rightarrow +\infty} F(x) = 1$.

Indications de preuve:

La première assertion est triviale. On considère ensuite la marche aléatoire $X_n = X_0 + cT_n - \sum_{k=1}^n Y_k = X_0 + \sum_{k=1}^n U_k$ avec $U_k = c\Delta_k - Y_k$. En notant W' le temps d'entrée de cette marche dans \mathbb{R}^- et $F'(x) = \mathbb{P}_x(W' = \infty)$, on voit que $F(x) = F'(x)$ puisque la ruine ne peut survenir qu'à la suite d'un sinistre.

- Preuve de 2.

La loi de la marche X_n a pour espérance $c - \lambda m$. Par conséquent si $c < \lambda m$ cette marche converge presque sûrement vers $-\infty$. Si $c = \lambda m$ la marche est récurrente et U_k n'étant pas identiquement nulle, cette marche atteint donc presque sûrement des points négatifs.

- Preuve de 3.

D'après 2.3.20 on a $\lim_{x \rightarrow +\infty} \mathbb{P}_x(W' < \infty) = 0$. ■

On pose $\rho = \lambda/c$. La mesure μ de densité $\rho\nu([x, +\infty[)$ a pour masse totale ρm et la fonction $g(t) = \rho \int \exp(tx) d\mu(x)$ est s.c.i. croissante avec $\lim_{t \rightarrow +\infty} g(t) = +\infty$. Si $\rho m < 1$, elle vérifie $g(0) < 1$, mais ceci n'implique pas nécessairement qu'il existe $\alpha > 0$ avec $g(\alpha) = 1$. Cette propriété est cependant vérifiée pour beaucoup de lois ν . . . Dans ce cas, la mesure $d\theta(x) = \exp(\alpha x) d\mu(x)$ est une probabilité et elle est adaptée puisqu'elle possède une densité.

Proposition 3.4.10. *On considère le modèle de ruine avec $m\rho < 1$ et on suppose qu'il existe $\alpha > 0$ avec $\int \exp(\alpha x) d\mu(x) = 1$. Alors*

$$\lim_{x \rightarrow +\infty} ((1 - F(x)) \exp(\alpha x)) = \frac{1 - \rho m}{\alpha \tau}$$

avec $\tau = \int x d\theta(x)$

Indications de preuve:

En conditionnant par l'instant t du premier sinistre et son coût y on voit que

$$\mathbb{P}_x(W = \infty | T_1 = t, Y_1 = y) = F(x + ct - y)$$

et par conséquent:

$$\begin{aligned}
F(x) &= \lambda \int \mathbf{1}_{\{t \geq 0\}} F(x + ct - y) \exp(-\lambda t) dt d\nu(y) \\
&= \rho \exp(\rho x) \int \mathbf{1}_{\{t \geq x\}} F(t - y) \exp(-\rho t) dt d\nu(y) \\
\int_0^u F(x) dx &= \int_0^u F(t - y) dt d\nu(y) + \frac{1}{\rho} (F(u) - F(0))
\end{aligned}$$

Or on montre facilement que $\rho \left(\int_0^u F(x) dx - \int_0^u F(t - y) dt d\nu(y) \right) = (F * \mu)(u)$ et on obtient donc l'équation:

$$F = F(0) + F * \mu$$

On en déduit que $F(x) = F(0) \sum_{n=0}^{\infty} \mu^{*n}([0, x])$ et puisque $\lim_{x \rightarrow +\infty} F(x) = 1$, $\mu(\mathbb{R}^+) = \rho m$ on trouve $F(0) = 1 - \rho m$. En posant $G = \mathbf{1}_{\{\mathbb{R}^+\}} - F$ on obtient l'équation

$$G(x) = G * \mu(x) + \mu(]x, +\infty[)$$

où μ est une sous probabilité. En opérant la multiplication par $\exp(\alpha x)$ on obtient une équation de renouvellement par rapport à la probabilité θ et il suffit de vérifier que $\int_0^{\infty} \exp(\alpha x) \mu(]x, +\infty[) dx = \frac{1 - \rho m}{\alpha}$. ■

Par exemple si ν est une loi exponentielle de paramètre λ' avec $\lambda < c\lambda'$ alors $\alpha = \lambda' - \lambda/c$ et donc la probabilité de ruine avec un capital initial x est équivalente à $\exp(-\alpha x) \frac{c}{\lambda\lambda'}$.

3.5 Le Renouvellement sur \mathbb{R}

La plus grande partie de la section précédente se transpose sans difficulté au cas d'une marche aléatoire adaptée de loi ν sur \mathbb{R} et possédant une espérance m non nulle (sinon la marche est récurrente). C'est le cas de la proposition 3.4.3 et du lemme 3.4.5. Les seuls points non transposables sont ceux faisant appel au calcul de transformée de Laplace. C'est le cas du lemme 3.4.2 qui est utilisé pour l'identification de la constante α dans la preuve du théorème de renouvellement. On le remplace par le résultat suivant:

Lemme 3.5.1. *On pose $H(x) = \mathbf{1}_{\{\mathbb{R}^+\}}(x)$ et on considère la fonction*

$$h(t) = \begin{cases} -\nu(]-\infty, t]) & \text{si } t < 0 \\ \nu(]t, +\infty[) & \text{si } t \geq 0 \end{cases}$$

*Alors $\int h(x) dx = m$, h est D.R.I et $\nu * H + h = H$. De plus $\Pi * h = H$ si $m > 0$ et $\Pi * h = 1 - H$ si $m < 0$.*

Indications de preuve:

La première relation résulte de Fubini et il est évident que h est D.R.I. puisque ses

parties positives et négatives sont monotones et intégrables. D'autre part $\nu * H(x) = \nu([-\infty, x])$ ce qui prouve $\nu * H + h = H$. De plus:

$$\sum_{k=0}^n \nu^{*k} * h = H - \nu^{*(n+1)} * H$$

or $\nu^{*n} * H(x) = \mathbb{P}(\Delta_1 + \Delta_2 + \dots + \Delta_n \leq x)$ et la loi faible des grands nombres implique que cette limite est 1 ou 0 suivant que $m < 0$ ou $m > 0$. ■

Ceci permet d'identifier les valeurs de α_{\pm} pour les limites en $+\infty$ et $-\infty$ du théorème de renouvellement et l'on obtient que:

$$\begin{cases} \alpha_+ = 1/m, \alpha_- = 0 & \text{si } m > 0 \\ \alpha_+ = 0, \alpha_- = -1/m & \text{si } m < 0 \end{cases}$$

Théorème 3.5.2. (Renouvellement sur \mathbb{R}) *Soit ν une mesure adaptée sur \mathbb{R} , possédant une espérance m non nulle (finie ou infinie). Alors:*

$$\lim_{t \rightarrow \pm\infty} \mathbf{\Pi} * f(t) = \alpha_{\pm} \int f(x) dx$$

pour toute fonction f qui est D.R.I. sur \mathbb{R} . La limite est toujours nulle si $|m| = \infty$.

On peut prouver que si m n'existe pas (c'est à dire que ni x^+ ni x^- ne sont ν intégrables) mais en supposant la marche transiente, alors les limites sont nulles.

Il reste à regarder ce que l'on peut dire de l'équation du renouvellement $F = F * \nu + f$ avec F et f positives. On a toujours $F = \sum_{k=0}^n \nu^{*k} * f + \nu^{*(n+1)} * F$ et il faut donc que le dernier terme de cette relation tende vers zéro pour que l'on puisse affirmer que F est de la forme $\mathbf{\Pi} * f$. C'est certainement vrai si on suppose F bornée et de plus soit que F est à support dans $[c, +\infty[$ et $m > 0$ soit que F à support dans $] -\infty, c]$ et $m < 0$. Par exemple dans le premier cas $\nu^{*n} * F(x) \leq \|F\| \mathbb{P}(\Delta_1 + \Delta_2 + \dots + \Delta_n \leq x - c)$ et on conclut avec la loi faible des grands nombres.

Bibliographie

- [1] N. Bourbaki *Intégration Chapitre IX*. Hermann (1969)
- [2] M. Duflo *Méthodes récursives aléatoires* Masson (1990)
- [3] A. Dembo et O. Zeitouni *Large deviations techniques* Jones & Bartlett (1993)
- [4] N. Dunford et J. Schwartz *Linear Operators* Tome I. Interscience Publishers (1957)
- [5] W. Feller *An introduction to probability theory and its application* Tome II. Wiley (1966)
- [6] J. Neveu *Bases Mathématiques du calcul des probabilités*. Masson (1964)
- [7] J. Neveu *Martingales à temps discret*. Masson (1972)